# Stochastik

**Barth, Friedrich**

**München, [20]03**

4. 2. Eigenschaften der relativen Häufigkeit

urn:nbn:de:hbz:466:1-83580

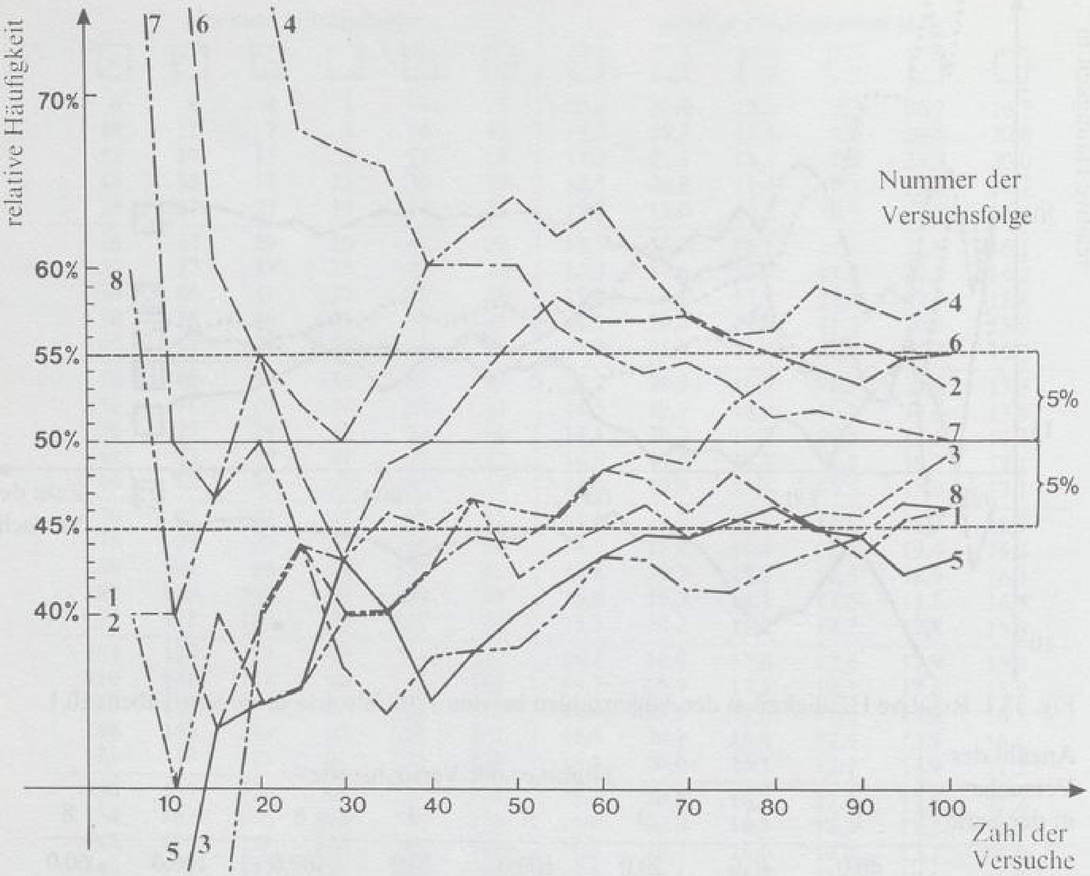


Fig. 34.1 Relative Häufigkeit von »Adler« bei je 100 Münzenwürfen in Abhängigkeit von der Versuchsfolge

## 4.2. Eigenschaften der relativen Häufigkeit

Wir betrachten die erste Zeile von Tabelle 10.1. Sie stellt die Ergebnisse einer Folge von 30 Versuchen (Würfelwurf) dar. Die relativen Häufigkeiten der Ereignisse $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$ und $\{6\}$ sind in folgender Tabelle zusammengestellt:

| Ereignis | $\{1\}$ | $\{2\}$ | $\{3\}$ | $\{4\}$ | $\{5\}$ | $\{6\}$ |
|---|---|---|---|---|---|---|
| relative Häufigkeit | $\frac{6}{30}$ | $\frac{6}{30}$ | $\frac{4}{30}$ | $\frac{1}{30}$ | $\frac{8}{30}$ | $\frac{5}{30}$ |

Diese relativen Häufigkeiten sind positive rationale Zahlen unter 1.
Allgemein kann man sagen: Tritt das Ereignis $A$ bei $n$ Versuchen $k$-mal ein, so gilt offenbar $0 \leqq k \leqq n$ und damit $0 \leqq \frac{k}{n} \leqq 1$. Also:

Die relative Häufigkeit eines Ereignisses $A$ in einer Versuchsfolge der Länge $n$ ist eine rationale Zahl aus dem Intervall $[0; 1]$, d.h.

(1) $\boxed{0 \leqq h_n(A) \leqq 1}$

Daß die Grenzfälle 0 und 1 auch wirklich auftreten, erkennt man sofort, wenn man den trivialen Fall $n = 1$ betrachtet. Der erste Wurf hatte das Ergebnis 4, die relative Häufigkeit $h_1(\{4\})$ hat somit den Wert 1, die relative Häufigkeit aller anderen Ereignisse jedoch den Wert 0. Die ersten 3 Versuche aus Zeile 10 zeigen, daß diese Werte auch für $n > 1$ auftreten können.

Wir wollen uns nun überlegen, wie man die relative Häufigkeit eines Ereignisses $A$ berechnen kann, wenn man die relativen Häufigkeiten der Elementarereignisse bei dieser Versuchsfolge kennt. Dazu betrachten wir als Beispiel das Ereignis $A :=$ »Augenzahl ist gerade« bei der oben betrachteten Versuchsfolge der Länge 30. Wir müssen in der ersten Zeile von Tabelle 10.1 zählen, wie oft eines der Ergebnisse 2, 4 oder 6 sich eingestellt hat. Wir erhalten $h_{30}(A) = \frac{12}{30} = 40\%$. Diese Zahl hätten wir aber auch aus der obigen Aufstellung der relativen Häufigkeiten der Elementarereignisse erhalten können. Wir müssen nämlich nur die relativen Häufigkeiten derjenigen Elementarereignisse addieren, deren Vereinigung das Ereignis $A$ ergibt, also $h_{30}(A) = h_{30}(\{2\}) + h_{30}(\{4\}) + h_{30}(\{6\}) = \frac{6}{30} + \frac{1}{30} + \frac{5}{30} = \frac{12}{30} = 40\%$.

Die eben durchgeführten Überlegungen lassen sich leicht verallgemeinern. Man erhält dann in endlichen Ergebnisräumen für relative Häufigkeiten bei einer festen Versuchsfolge von $n$ Versuchen:

Die relative Häufigkeit eines Ereignisses $A\,(\neq \emptyset)$ ist gleich der Summe der relativen Häufigkeiten derjenigen Elementarereignisse, deren Vereinigung $A$ ist; in Zeichen

$$(2) \qquad \boxed{h_n(A) = \sum_{\omega \in A} h_n(\{\omega\})}$$

Wir wenden uns nun den Ereignissen $\emptyset$ und $\Omega$ zu.

Das unmögliche Ereignis $\emptyset$ tritt nie ein; in Definition 30.1 ist also $k = 0$, woraus folgt

$$(3) \qquad \boxed{h_n(\emptyset) = 0}$$

Für das sichere Ereignis $\Omega$ gilt andererseits $k = n$, weil es bei jedem Versuch eintritt. Somit gilt

$$(4) \qquad \boxed{h_n(\Omega) = 1}$$

$A$ und $B$ seien nun 2 Ereignisse bei derselben Versuchsfolge, deren relative Häufigkeiten $h_n(A)$ und $h_n(B)$ bekannt sind. Kann man damit die relative Häufigkeit des Ereignisses »$A$ oder $B$« $= A \cup B$ bei dieser Versuchsfolge berechnen? Zur Beantwortung dieser Frage betrachten wir bei den obigen 30 Würfelwürfen die Ereignisse $A :=$ »Augenzahl ist gerade« und $B :=$ »Augenzahl ist von 1 und 6 verschieden«. $h_{30}(A)$ war 40%. Aufgrund von Eigenschaft (2) errechnen wir für

$$\begin{aligned} h_{30}(B) &= h_{30}(\{2\}) + h_{30}(\{3\}) + h_{30}(\{4\}) + h_{30}(\{5\}) = \\ &= \tfrac{6}{30} + \tfrac{4}{30} + \tfrac{1}{30} + \tfrac{8}{30} = \\ &= \tfrac{19}{30} = 63\tfrac{1}{3}\%. \end{aligned}$$

Der naive Vorschlag, die relative Häufigkeit von $A \cup B$ als Summe der relativen Häufigkeiten von $A$ bzw. $B$ zu berechnen, schlägt fehl, da sich hier für die Summe der Wert $\frac{31}{30} = 103\frac{1}{3}\%$ ergibt. Man sieht aber auch sofort, woran das liegt: Die Ergebnisse 2 und 4 treten sowohl in $A$ als auch in $B$ auf, die relativen Häufigkeiten der zugehörigen Elementarereignisse $\{2\}$ bzw. $\{4\}$ wurden also bei der Summenbildung doppelt gezählt. Um diesen Fehler zu korrigieren, müssen wir diese relativen Häufigkeiten vom Summenwert $\frac{31}{30}$ subtrahieren; wir erhalten also $h_{30}(A \cup B) = \frac{31}{30} - (\frac{6}{30} + \frac{1}{30}) = \frac{24}{30} = 80\%$. Dieser Wert ist richtig, wie wir durch direkte Berechnung von $h_{30}(A \cup B)$ überprüfen können:

$$h_{30}(A \cup B) = h_{30}(\{2, 3, 4, 5, 6\}) = \tfrac{6}{30} + \tfrac{4}{30} + \tfrac{1}{30} + \tfrac{8}{30} + \tfrac{5}{30} = \tfrac{24}{30} = 80\%.$$

Was wir am Beispiel gesehen haben, gilt aber sogar allgemein:
In einer festen Versuchsfolge ist die relative Häufigkeit des Ereignisses »$A$ oder $B$« gleich der Summe der relativen Häufigkeiten der beiden Ereignisse abzüglich der relativen Häufigkeit des Ereignisses »$A$ und $B$«, kurz

$$(5) \qquad \boxed{h_n(A \cup B) = h_n(A) + h_n(B) - h_n(A \cap B)}$$

**Beweis:** Bezeichnen wir mit $k(A \cup B)$ die absolute Häufigkeit des Ereignisses $A \cup B$ in der Serie von $n$ Versuchen, so erkennt man an Hand von Figur 36.1 leicht, daß für die absoluten Häufigkeiten $k(A \cup B)$, $k(A)$, $k(B)$ und $k(A \cap \mathrm{B})$ gilt:

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

Dividiert man diese Gleichung durch $n$, so erhält man die Behauptung.

| | $B$ | $\bar{B}$ | |
|---|---|---|---|
| $A$ | $k(A \cap B)$ | $k(A \cap \bar{B})$ | $k(A)$ |
| $\bar{A}$ | $k(\bar{A} \cap B)$ | $k(\bar{A} \cap \bar{B})$ | $k(\bar{A})$ |
| | $k(B)$ | $k(\bar{B})$ | |

Fig. 36.1 Mehrfeldertafel der absoluten Häufigkeiten. $k(M)$ bedeutet die absolute Häufigkeit des Ereignisses $M$ in der Versuchsserie.

Sind $A$ und $B$ unvereinbare Ereignisse, d.h., ist $A \cap B = \emptyset$, so wird aus (5)

$$(6) \qquad \boxed{A \cap B = \emptyset \;\Rightarrow\; h_n(A \cup B) = h_n(A) + h_n(B)}$$

Ist im besonderen $B = \bar{A}$, $B$ also das Gegenereignis zu $A$, so ergibt (6) wegen $A \cap \bar{A} = \emptyset$:

$h_n(A \cup \bar{A}) = h_n(A) + h_n(\bar{A})$. Andererseits ist nach (4)
$h_n(A \cup \bar{A}) = h_n(\Omega) = 1$. Somit gilt

$$(7) \qquad \boxed{h_n(\bar{A}) = 1 - h_n(A)}$$

Die relative Häufigkeit eines Ereignisses und die seines Gegenereignisses ergeben in einer festen Versuchsfolge stets 100%.
An einem **Beispiel** wollen wir zeigen, wie man mit Hilfe der Eigenschaften der relativen Häufigkeiten eine Mehrfeldertafel erstellen und damit zusammenhängende Aufgaben lösen kann.

Am 31. 12. 1973 hatte die Bundesrepublik Deutschland $n = 62\,101\,400$ Einwohner. Davon waren 29713800 männlich und davon wieder 20002000 volljährig. Insgesamt waren 43151600 Einwohner volljährig.
Aus diesen Daten können wir die relativen Häufigkeiten für die Ereignisse
$♂ :=$ »Ein beliebig herausgegriffener Einwohner ist männlich« und
$V :=$ »Ein beliebig herausgegriffener Einwohner ist volljährig« berechnen:

$$h_n(♂) = \tfrac{29\,713\,800}{62\,101\,400} = 47{,}8\%, \qquad h_n(V) = \tfrac{43\,151\,600}{62\,101\,400} = 69{,}5\%.$$

Für das Ereignis, daß ein beliebig herausgegriffener Einwohner männlich und volljährig ist, erhalten wir $h_n(♂ \cap V) = \tfrac{20\,002\,000}{62\,101\,400} = 32{,}2\%$.

Diese gegebenen Zahlen sind in der Mehrfeldertafel (Figur 37.1) schwarz eingetragen. Die restlichen Zahlen berechnen wir unter Verwendung der Eigenschaften der relativen Häufigkeit.

$$h_n(\bar{V}) = 1 - h_n(V) = \\ = 1 - 0{,}695 = 30{,}5\%$$

$$h_n(♀) = 1 - h_n(♂) = \\ = 1 - 0{,}478 = 52{,}2\%$$

| | $V$ | $\bar{V}$ | |
|---|---|---|---|
| ♂ | 32,2% | 15,6% | 47,8% |
| ♀ | 37,3% | 14,9% | 52,2% |
| | 69,5% | 30,5% | |

Fig. 37.1 Mehrfeldertafel der relativen Häufigkeiten

Weil $V$ die Vereinigung der unvereinbaren Ereignisse $V \cap ♂$ und $V \cap ♀$ ist, gilt nach (6):

$h_n(V) = h_n(V \cap ♂) + h_n(V \cap ♀)$; also ist

$$h_n(V \cap ♀) = h_n(V) - h_n(V \cap ♂) = 69{,}5\% - 32{,}2\% = 37{,}3\%.$$

Analog erhalten wir

$$h_n(\bar{V} \cap ♂) = h_n(♂) - h_n(V \cap ♂) = 47{,}8\% - 32{,}2\% = 15{,}6\%$$

und

$$h_n(\bar{V} \cap ♀) = h_n(♀) - h_n(V \cap ♀) = 52{,}2\% - 37{,}3\% = 14{,}9\%.$$

Damit ist die Vierfeldertafel gefüllt.
Jede weitere einschlägige Fragestellung läßt sich nun direkt aus der Vierfeldertafel beantworten; zum Beispiel:

$$h_n(V \cup ♀) = h_n(V) + h_n(\bar{V} \cap ♀) = 69{,}5\% + 14{,}9\% = 84{,}4\%$$

oder auch

$$h_n(V \cup ♀) = h_n(♀) + h_n(V \cap ♂) = 52{,}2\% + 32{,}2\% = 84{,}4\%$$

oder umständlicher mit Eigenschaft (5)

$$h_n(V \cup ♀) = h_n(V) + h_n(♀) - h_n(V \cap ♀) = 69{,}5\% + 52{,}2\% - 37{,}3\% = 84{,}4\%.$$