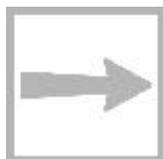


3 Grundlagen zur Wahrscheinlichkeitsrechnung

Durch die Simulation haben wir eine gute Vorstellung von der Bedeutung und der Grösse von \hat{p} resp. \hat{q} erhalten. In diesem Kapitel wollen wir die Grundlagen erarbeiten, um \hat{p} und \hat{q} im nächsten Kapitel formelmässig *berechnen* zu können. Das kann sich als sehr hilfreich erweisen: Weiss man beispielsweise, wie \hat{p} von N , S und n abhängt, so kann man evtl. ausrechnen, wie gross der Stichprobenumfang n gewählt werden muss, damit \hat{p} einen gewissen gewünschten Wert nicht überschreitet.

Wir setzen hier voraus, dass Sie die wichtigsten Grundbegriffe aus der Kombinatorik und der Wahrscheinlichkeitsrechnung mitbringen. Trotzdem ist das Kapitel sehr ausführlich gehalten.

Wir können \hat{p} und \hat{q} berechnen, sobald wir die folgende Grundaufgabe gelöst haben.



Die Grundaufgabe

Eine Urne enthält $N=100'000$ Kugeln, davon sind $S=4000$ schwarz. Man entnimmt der Urne mit irgendeinem Zufallsmechanismus – denken Sie an die Auswahlmaschinen beim Zahlenlotto – $n=200$ Kugeln (Fig. 3.1).

Gesucht ist die Wahrscheinlichkeit dafür, dass die Stichprobe s schwarze Kugeln enthält. (Dabei ist s eine bestimmte Zahl mit $0 \leq s \leq n$.)

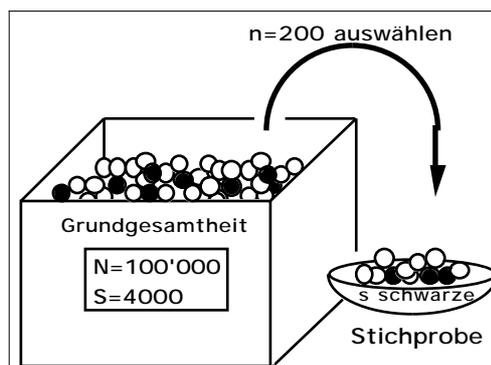


Fig. 3.1: Urnenmodell für die Ziehung von 200 Kugeln

Die Kugelnummern

Wie schon im letzten Kapitel nehmen wir an, die 100'000 Kugeln seien beginnend mit 0 durchnummeriert, wobei die Nummern 0, 1, ..., 3999 für die schwarzen Kugeln reserviert sind. Mit Hilfe dieses Tricks erreichen wir, dass die Kugeln künstlich unterscheidbar werden, was die nachfolgenden Erklärungen hoffentlich vereinfacht.

Stichproben und Teilmengen

Wenn wir zur schwarzen Kugel mit der Nr. 0 weitere 199 weisse Kugeln aussuchen, entsteht eine „Stichprobe mit 1 schwarzen Kugel“. Wenn wir die gleichen 199 weissen Kugeln mit der schwarzen Kugel Nr. 1 kombinieren, haben wir wieder eine „Stichprobe mit genau einer schwarzen Kugel“. Das Ereignis „Stichprobe mit 1 schwarzen Kugel“ kann somit allein hinsichtlich der schwarzen Kugel auf 4000 verschiedene Arten zustande kommen.

Etwas mathematischer können wir sagen: Es gibt viele verschiedene *Teilmengen* mit 200 Kugeln, die auf das Ereignis „Stichprobe mit 1 schwarzen Kugel“ führen. Wir bezeichnen solche Teilmengen als *günstig* für das Ereignis. Für die *Anzahl* der *günstigen* Teilmengen verwenden wir das Symbol g_1 .

Alles Gesagte gilt analog für Stichproben mit s schwarzen Kugeln. Insbesondere verstehen wir unter g_s die Anzahl günstiger Teilmengen für eine "Stichprobe mit s schwarzen Kugeln".

Unterschied zwischen Stichprobe und Teilmenge



Vielleicht wird der Unterschied zwischen Stichprobe und Teilmenge noch durch die folgende Bemerkung verdeutlicht: Zwei *Teilmengen* mit je 200 Kugeln betrachten wir dann als verschieden, wenn es in der einen Teilmenge wenigstens eine Kugelnummer gibt, die in der anderen Teilmenge nicht vorkommt. Dagegen betrachten wir zwei *Stichproben* als verschieden, wenn sie sich in der *Anzahl der schwarzen Kugeln* unterscheiden; es gibt mit anderen Worten in unserer Betrachtungsweise nur 201 verschiedene Stichproben, nämlich solche mit 0, 1, 2, ..., 200 schwarzen Kugeln.



Aufgabe 3.1

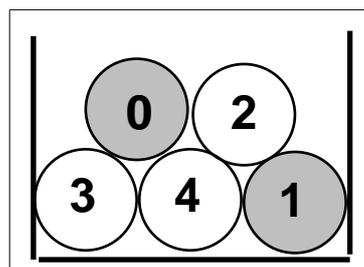
Eine Urne enthält 2 schwarze und 3 weiße (numerierte) Kugeln.

a

Bilden Sie alle möglichen Teilmengen mit 3 Kugeln.

b

Welche bzw. wie viele der Teilmengen führen auf Stichproben mit 0, 1, 2, 3 schwarzen Kugeln?

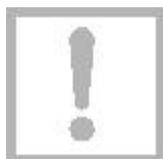


Alle Teilmengen mit 200 Elementen sind gleich wahrscheinlich

Für die weiteren Berechnungen ist nun entscheidend, dass wir postulieren: Alle möglichen Teilmengen mit 200 Kugeln haben die gleiche Wahrscheinlichkeit, ausgewählt zu werden.

Beachten Sie, dass dies für *Stichproben* im oben definierten Sinne nicht zutrifft: Eine Stichprobe mit 200 schwarzen Kugeln hat praktisch die Wahrscheinlichkeit 0, Stichproben mit ungefähr 10 schwarzen Kugeln sind jedoch recht häufig.

Die Wahrscheinlichkeit einer Stichprobe ist um so grösser, je grösser die zugehörige Anzahl der günstigen Teilmengen ist. In der Wahrscheinlichkeitsrechnung haben Sie vermutlich die folgende Formel kennen gelernt:



$$P(\text{„s schwarze Kugeln“}) = \frac{\text{Anzahl günstiger Teilmengen}}{\text{Anzahl möglicher Teilmengen}} = \frac{g_s}{m}$$

Hierbei bedeuten:

- g_s = Anzahl der (günstigen) Teilmengen mit $n=200$ Kugeln, von denen s schwarz sind (s ist eine bestimmte Zahl mit $0 \leq s \leq 200$).
- m ist die Anzahl der (möglichen) Teilmengen mit $n=200$ Kugeln.



Aufgabe 3.2

Eine Urne enthält 2 schwarze und 3 weiße Kugeln (siehe Aufgabe 3.1). Es wird eine Stichprobe vom Umfang 3 gezogen. Mit welchen Wahrscheinlichkeiten enthält sie 0, 1, 2 oder 3 schwarze Kugeln?

m = Anzahl der möglichen Teilmengen

Wiederholen wir das Ergebnis der Aufgabe 3.1. Mit 5 Kugeln, die mit 0, 1, 2, 3, 4 numeriert sind, kann man 10 verschiedene Teilmengen mit 3 Kugeln bilden, nämlich:

$$\begin{array}{ccccc} \{0, 1, 2\} & \{0, 1, 3\} & \{0, 1, 4\} & \{0, 2, 3\} & \{0, 2, 4\} \\ \{0, 3, 4\} & \{1, 2, 3\} & \{1, 2, 4\} & \{1, 3, 4\} & \{2, 3, 4\}. \end{array}$$

In der Kombinatorik wird gezeigt, dass diese Anzahl gleich $\binom{5}{3}$ ist. Wie Sie sich wohl erinnern, berechnet man diesen sog. *Binomialkoeffizienten* mit einem der folgenden beiden Brüche:

$$\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} = 10$$

Konsultieren Sie evtl. Ihre Formelsammlung! (Berechnungsbeispiele folgen!)

Analog gilt: Mit $N=100'000$ (unterscheidbaren) Kugeln kann man

$$m = \binom{N}{n} = \frac{100'000}{200}$$

verschiedene Teilmengen mit $n=200$ Kugeln bilden.

Binomialkoeffizienten auf dem Taschenrechner

Die nachstehende Fig. 3.2 zeigt, wie man diesen Binomialkoeffizienten auf einem modernen Taschenrechner berechnet (nCr bedeutet hier "number of combinations", deutsch Binomialkoeffizient. Um die Eingabe von Grossbuchstaben zu umgehen, steht nn anstelle von N).

100000→n

200→n

nCr(nn,n)

1.03903E625

Fig. 3.2 : Die Berechnung eines Binomialkoeffizienten, wie sie auf einem modernen Taschen-computer aussehen mag.

Das Ergebnis $m=1.04 \cdot 10^{625}$ ist eine unvorstellbar grosse Zahl. Vergleiche: Die Anzahl der Elektronen im ganzen Weltall ist deutlich kleiner als 10^{80} .



Aufgabe 3.3

Eine Urne enthält 5 (numerierte) Kugeln. Wie viele verschiedene Teilmengen mit 3 Kugeln kann man aufgrund der eben angegebenen Formel bilden? Vergleichen Sie Ihr Ergebnis mit demjenigen in Aufgabe 3.1.

g_s = Anzahl der günstigen Teilmengen

Zu den günstigen Teilmengen gehören solche mit genau s schwarzen und $(n-s)$ weissen Kugeln.

Ihre Anzahl ergibt sich aufgrund der folgenden *Idee*: Wir kombinieren alle möglichen Mengen mit s schwarzen Kugeln mit allen möglichen Mengen mit $(n-s)$ weissen Kugeln.

Gemäss den vorherigen Überlegungen können wir aus den $S=4000$ schwarzen Kugeln $\binom{S}{s}$ verschiedene Teilmengen mit s Kugeln wählen.

Jede dieser „schwarzen“ Mengen können wir nun mit den $\binom{N-S}{n-s}$ verschiedenen „weissen“ Mengen zu einer Menge mit $n=200$ Kugeln ergänzen (Fig. 3. 3).

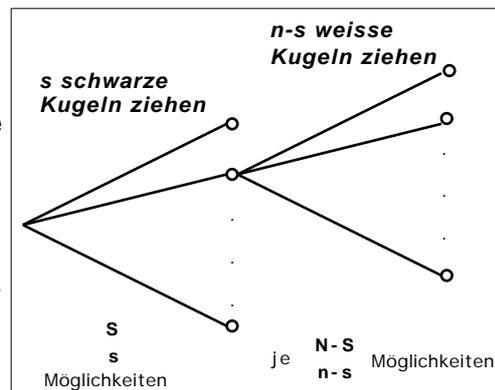
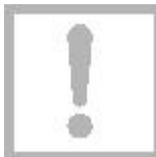


Fig. 3.3: Baumdiagramm für die Erzeugung von Stichproben mit s schwarzen Kugeln.

Es gibt also total

$$g_s = \binom{S}{s} \binom{N-S}{n-s}$$

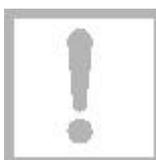
verschiedene Teilmengen mit $n=200$ Kugeln, von welchen s schwarz sind.



Aufgabe 3.4

Eine Urne enthält 2 schwarze und 3 weisse (numerierte) Kugeln. Wie viele verschiedene Teilmengen mit genau 1 schwarzen Kugel gibt es?

Wenden Sie die eben hergeleitete Formel an und vergleichen Sie das Ergebnis mit demjenigen in Aufgabe 3.1.



Die Lösung der Grundaufgabe

$$P(\text{"s schwarze Kugeln"}) = \frac{g_s}{m} = \frac{\binom{S}{s} \binom{N-S}{n-s}}{\binom{N}{n}}$$

- N = Anzahl Kugeln in der Urne (z. B. $N=100'000$)
- S = Anzahl schwarzer Kugeln in der Urne (z. B. $S = 4000$)
- n = Umfang der Stichprobe (z. B. $n = 200$)
- s = Anzahl schwarzer Kugeln in der Stichprobe (z. B. $s = 13$)

Auch diese Formel finden Sie in Ihrer Formelsammlung. Suchen Sie die Formel unter dem Stichwort „Ziehen ohne Zurücklegen“ oder unter „Hypergeometrische Verteilung“.

Rechenbeispiel

Berechnen wir die Wahrscheinlichkeit, dass die Stichprobe genau die Grundgesamtheit wiedergibt, d.h. genau 8 schwarze Kugeln enthält. Gemäss Fig. 3.4 erhalten wir auf dem Taschencomputer:

$$P(\text{„8 schwarze Kugeln“}) = 14.3\%.$$

$$\begin{array}{l} 100000 \rightarrow n \\ 4000 \rightarrow ss \\ 200 \rightarrow n \\ 8 \rightarrow s \\ \frac{nCr(ss,s) \cdot nCr(nn-ss,n-s)}{nCr(nn,n)} \quad .142605 \end{array}$$

Fig. 3.4: So oder ähnlich könnte sich die Berechnung der Wahrscheinlichkeit

$P(\text{„8 schwarze“})$

auf einem modernen Taschencomputer präsentieren.



Aufgabe 3.5

Eine Urne enthält 2 schwarze und 3 weisse Kugeln. Es wird eine Stichprobe vom Umfang 3 gezogen. Mit welchen Wahrscheinlichkeiten enthält sie 0, 1, 2 oder 3 schwarze Kugeln?

Rechnen Sie mit der eben hergeleiteten Formel und zwar mit und ohne Taschenrechner und vergleichen Sie die Ergebnisse mit denjenigen in Aufgabe 3.2.

Aufgabe 3.6

Einmal tippen beim Schweizerischen Zahlenlotto heisst, die 6 Zahlen zu erraten, die bei der nächsten Ziehung aus den Zahlen 1, 2, ..., 45 ausgelost werden.

Wie gross ist die Wahrscheinlichkeit, dass Sie mit einem Tip 4 richtige Zahlen erraten?

Hinweis für diese nicht ganz einfache Aufgabe: Betrachten Sie eine Urne mit 45 nummerierten Kugeln, wovon einige (welche?) schwarz sind.

In diesem kurzen Kapitel ging es darum, die Wahrscheinlichkeiten für gewisse Stichproben zu berechnen. Prüfen Sie nun anhand der folgenden Lernkontrolle, ob Sie in der Lage sind, die hergeleitete Formel in einem praktischen Beispiel anzuwenden.

Wenn Sie ohne Hilfe des Lösungsteils zum richtigen Ergebnis gekommen sind, werden Sie auch den anschliessenden Kapitteltest ohne Schwierigkeiten bestehen.