Stochastik

Barth, Friedrich

München, [20]03

14. 1. Einführung

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

14. Die Binomialverteilung

14.1. Einführung

Abraham de Moivre (1667–1754) veröffentlichte im Jahre 1711 die Abhandlung *De Mensura Sortis, seu, de Probabilitate Eventuum in Ludis a Casu Fortuito Pendentibus* (Bild 75.1), in der er 26 Probleme abhandelte. Problem I lautet:

PROB. I.

A & B *una teſſera ludunt, ea conditione, ut ſi* A *bis vel pluries, octo jactibus teſſeræ monada jecerit, ipſe* A *vincat; ſin ſemel tantum, vel non omnino,* B *vincat; quænam erit ratio ſortium?*

»A und B spielen mit einem Würfel so, daß A gewinnen soll, wenn er bei 8 Würfen zweimal oder öfters ein As [d.h. eine Eins] wirft; fällt das As nur einmal oder gar nicht, so gewinne B. Wie groß ist das Verhältnis der Chancen?«

Wir wollen diese Aufgabe mit unseren Hilfsmitteln lösen. Versuchen wir, zunächst die Wahrscheinlichkeit für genau 2 Asse bei diesen 8 Würfen zu ermitteln. Das zugrundeliegende Zufallsexperiment kann als *Bernoulli*-Kette der Länge 8 mit dem Parameter $\frac{1}{6}$ gedeutet werden, falls man als Treffer an der Stelle i das Erscheinen eines Asses beim i-ten Wurf nimmt. Diese Annahme ist zulässig, weil man davon ausgehen darf, daß die Ereignisse $A_i :=$ »As beim i-ten Wurf« $(i = 1, 2, \ldots, 8)$ stochastisch unabhängig sind, da sich die Würfe gegenseitig nicht beeinflussen. Der Ergebnisraum Ω besteht aus den 2^8 Oktupeln, die aus den Ziffern 0 und 1 gebildet werden können. Bezeichnet man mit Z die Zufallsgröße »Anzahl der Treffer«, in unserem Fall also die Anzahl der gefallenen Asse, so besteht unsere Aufgabe darin, die Wahrscheinlichkeit des Ereignisses »$Z = 2$« zu berechnen. Dieses Ereignis besteht aus denjenigen 8-Tupeln aus Ω, die aus 2 Einsen und 6 Nullen gebildet werden können. Beispiele hierfür sind die 8-Tupel 11000000, 00100010, 00010100 usw. Für das Ereignis »$Z = 2$« spielt es dabei keine Rolle, an welchen Stellen die beiden Einsen stehen, d.h., bei welchen der 8 Würfe die beiden Asse fallen werden. Da man die 2 Einsen auf die 8 Stellen des 8-Tupels auf $\binom{8}{2}$ Arten verteilen kann, gibt es $\binom{8}{2}$ Oktupel, die für das Ereignis »$Z = 2$« günstig sind. Jedes dieser 8-Tupel hat als Elementarereignis gemäß Definition 221.1 die Wahrscheinlichkeit $\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^6$. Damit erhalten wir für die Wahrscheinlichkeit des Ereignisses »$Z = 2$« den Wert $\binom{8}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^6$; der erste Teil unserer Aufgabe ist somit gelöst.

Analog gewinnen wir nun die Wahrscheinlichkeit für genau k Treffer, also für das Ereignis »$Z = k$«, indem wir in den obigen Überlegungen die Zahl 2 durch k ersetzen. Also ist

$$P(Z=k)=\binom{8}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{8-k}.$$

Damit ergibt sich für die Gewinnchance von A der Wert

$$P(Z \geqq 2)=\sum_{k=2}^{8}\binom{8}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{8-k}$$

Die numerische Berechnung dieser Wahrscheinlichkeit ist etwas mühsam. Leichter erhalten wir ihren Wert über das Gegenereignis »$Z \leqq 1$«, d.h. über die Gewinnchance von B:

$$\begin{aligned}
P(Z \geqq 2) &= 1-P(Z \leqq 1) = \\
&= 1-\sum_{k=0}^{1}\binom{8}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{8-k} = \\
&= 1-\binom{8}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^8-\binom{8}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^7 = \\
&= \frac{6^8-5^8-8\cdot 5^7}{6^8} = \\
&= \frac{1679616-390625-625000}{1679616} = \\
&= \frac{1679616-1015625}{1679616} = \\
&= \frac{663991}{1679616} \approx \\
&\approx 39{,}5\%.
\end{aligned}$$

Die Chancen von A und B verhalten sich also wie $663991 : 1015625 \approx 2:3$.

Das Typische an der Aufgabe von *de Moivre* ist, daß man sich nicht mehr für die Nummer des Versuchs interessiert, bei dem der Treffer eintritt, sondern daß man nach der Anzahl der Treffer fragt, die sich bei einer Serie von Versuchen ergeben kann. Man betrachtet im stochastischen Modell also die Zufallsgröße $Z :=$ »Anzahl der Treffer bei einer *Bernoulli*-Kette der Länge n mit dem Parameter p«.

Für ihre Wahrscheinlichkeitsverteilung gilt nach dem Obigen die von *Jakob Bernoulli* (1655–1705) in der *Ars Conjectandi* (Seite 40) hergeleitete Formel:

$$P(Z=k)=\binom{n}{k}p^k(1-p)^{n-k}=\binom{n}{k}p^k q^{n-k}.$$

In dieser Verteilung spielen die Binomialkoeffizienten eine wichtige Rolle. Man sagt daher, Z sei binomial verteilt. Allgemein definiert man:

Definition 231.1: Eine Zufallsgröße X heißt **binomial nach $B(n; p)$ verteilt**, wenn
1. die Wertemenge von X die Menge $\{0, 1, 2, \dots, n\}$ ist, und
2. für die Wahrscheinlichkeitsverteilung von X gilt:

$$B(n; p): \quad x \mapsto B(n; p; x) := \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{für} \quad x \in \{0, 1, \dots, n\}, \\ 0 & \text{sonst.} \end{cases}$$

Bemerkungen:

1) Interessant sind eigentlich nur die Werte $B(n; p; x)$ für $x \in \{0, 1, 2, \dots, n\}$. Für ein derartiges x schreibt man gerne k, um anzudeuten, daß es sich um eine ganze Zahl handelt.
2) Mit $q := 1 - p$ erhält man den kürzeren Ausdruck $B(n; p; k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$.
3) Jede Wahrscheinlichkeitsverteilung $B(n; p)$ heißt **Binomialverteilung**. Der Name rührt davon her, daß $B(n; p; k)$ gerade der k-te Summand in der Entwicklung der n-ten Potenz des Binoms $p + q$ ist; es gilt nämlich

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot q^{n-k}.$$

4) Die obige Definition 231.1 ist nur sinnvoll für den nicht-trivialen Fall $0 < p < 1$. Ist $p = 0$, so liefert jeder Versuch eine Niete; das führt zur Verteilung

$$B(n; 0; x) := \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{sonst.} \end{cases}$$

Ist hingegen $p = 1$, so liefert jeder Versuch einen Treffer; das führt zur Verteilung

$$B(n; 1; x) := \begin{cases} 1 & \text{für } x = n, \\ 0 & \text{sonst.} \end{cases}$$

Für die **kumulative Verteilungsfunktion** einer nach $B(n; p)$ verteilten Zufallsgröße hat sich die Bezeichnung F_p^n bewährt. Es gilt also nach Satz 176.1:

$$\boxed{F_p^n(x) := \sum_{i \leq x} B(n; p; i)}$$

Ist insbesondere x eine der interessierenden Zahlen aus $\{0, 1, 2, \dots, n\}$, so schreibt man an Stelle von x wieder gerne k und erhält damit

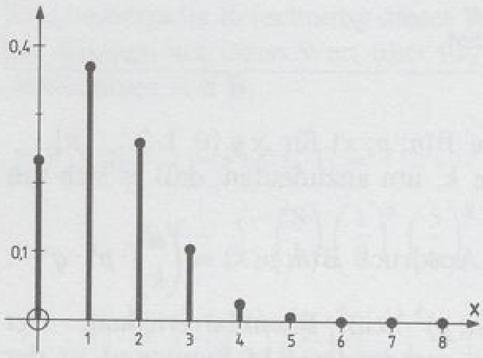
$$\boxed{F_p^n(k) = \sum_{i=0}^{k} B(n; p; i)}$$

Wenn keine Verwechslung möglich ist, lassen wir die Indizes bei F_p^n weg. Unter Verwendung dieses Symbols lautet die Lösung des Problems von *de Moivre*

$$\frac{P(Z \geqq 2)}{P(Z \leqq 1)} = \frac{1 - F_{1/6}^8(1)}{F_{1/6}^8(1)}.$$

Wir veranschaulichen die Binomialverteilung $B(8; \frac{1}{6})$ sowohl durch ein Stabdiagramm (Figur 232.1) als auch durch ein Histogramm (Figur 232.2).
Den Graphen der zugehörigen kumulativen Verteilungsfunktion $F^8_{1/6}$ zeigt Figur 232.3.

Fig. 232.1 Stabdiagramm von $B(8; \frac{1}{6})$



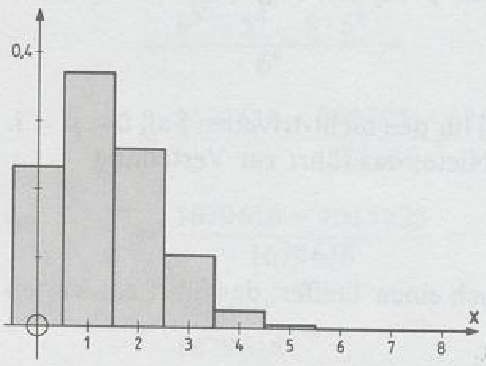


Fig. 232.2 Histogramm von $B(8; \frac{1}{6})$

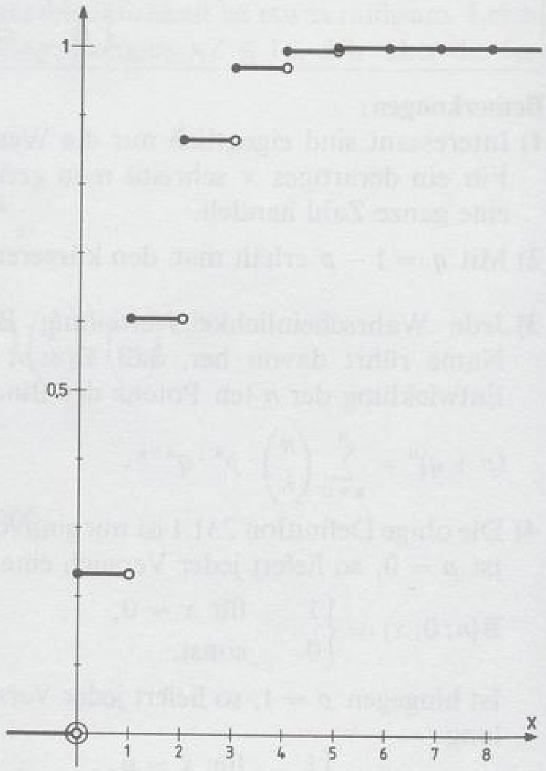


Fig. 232.3 Graph von $F^8_{\frac{1}{6}}$.

14.2. Ziehen mit bzw. ohne Zurücklegen

Die Formel von Definition 231.1 für die Binomialverteilung kennen wir schon lange. Beim Ziehen mit Zurücklegen aus einer Urne erhielten wir in Satz 107.1 für die Wahrscheinlichkeit, genau s schwarze Kugeln zu ziehen, den Wert $\binom{n}{s}p^s q^{n-s}$, also gerade $B(n; p; s)$. Die Zufallsgröße »Anzahl der Treffer« beim Ziehen mit Zurücklegen ist demnach binomial verteilt. Weil man viele Experimente auf das *Ziehen mit Zurücklegen* reduzieren kann, ist diese Zufallsgröße gewissermaßen der Prototyp einer binomial verteilten Zufallsgröße.
Andererseits lassen sich viele Zufallsexperimente durch das Urnenexperiment *Ziehen ohne Zurücklegen* simulieren. In diesem Fall liegt keine Bernoulli-Kette vor, wie in Aufgabe 223/2 gezeigt wurde. Die Zufallsgröße »Anzahl der Treffer«