Stochastik

Barth, Friedrich

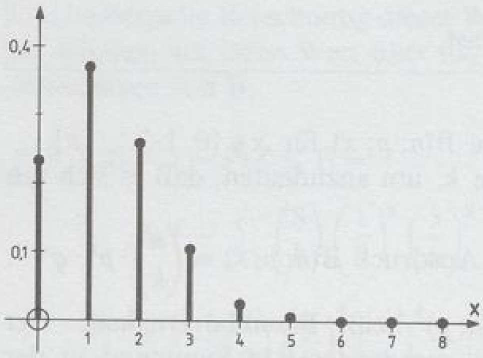
München, [20]03

14. 2. Ziehen mit bzw. ohne Zurücklegen

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

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 $B(8; \frac{1}{6})$



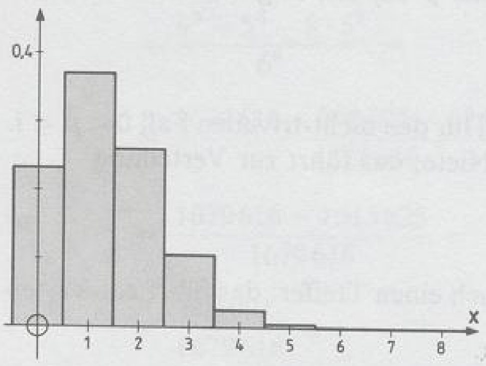


Fig. 232.2 Histogramm von $B(8; \frac{1}{6})$

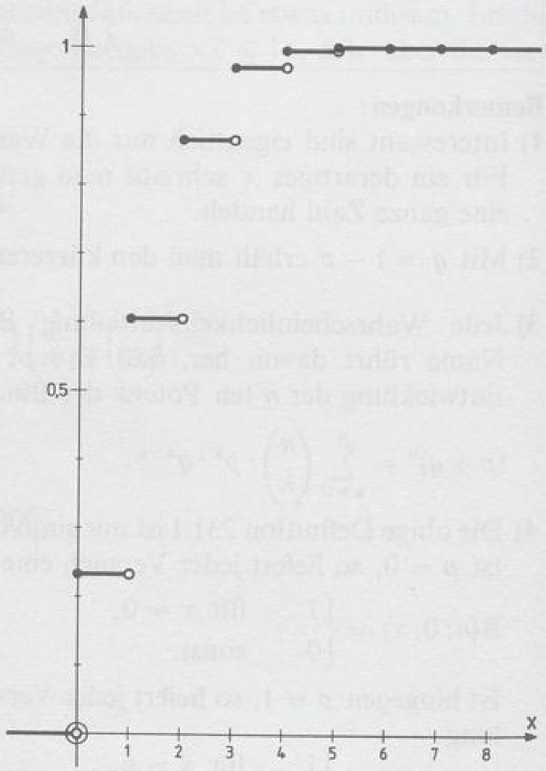


Fig. 232.3 Graph von $F^8_{\frac{1}{6}}$.

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«

ist dann auch nicht binomial verteilt. Für ihre Verteilung erhielten wir in Satz 106.1

$$P(Z = s) = \frac{\binom{S}{s}\binom{N-S}{n-s}}{\binom{N}{n}}.$$

Allgemein definieren wir:

Definition 233.1: Eine Zufallsgröße X heißt für $K \leqq N$ und $n \leqq N$ **hypergeometrisch nach $\mathbf{H}(N; K; n)$ verteilt**, wenn gilt:
1. die Wertemenge von X ist eine Teilmenge von $\{0, 1, 2, \ldots, n\}$, und
2. die Wahrscheinlichkeitsverteilung von X lautet

$$\mathrm{H}(N; K; n)\colon x \mapsto \mathrm{H}(N; K; n; x) := \begin{cases} \dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} & \text{für } x \in \{0, 1, \ldots, n\}, \\ 0 & \text{sonst.} \end{cases}$$

Auch hier schreibt man gerne für $x \in \{0, 1, \ldots, n\}$ den Buchstaben k.
In der Praxis spielt die hypergeometrische Verteilung eine große Rolle. Der Prototyp einer hypergeometrisch verteilten Zufallsgröße ist die »Anzahl der Treffer« beim Ziehen ohne Zurücklegen aus einer Urne. So sind z.B. die Zufallsgrößen »Anzahl der defekten Stücke« bei einer Qualitätskontrolle und »Anzahl der Ja-Antworten« bei einer Umfrage hypergeometrisch verteilt.
Die hypergeometrische Verteilung erfordert wegen der drei Binomialkoeffizienten einen sehr hohen rechnerischen Aufwand. Rechnerisch leichter zugänglich ist die Binomialverteilung. Glücklicherweise läßt sich die hypergeometrische Verteilung für $n \ll \min\{N, K, N-K\}$ recht gut durch die Binomialverteilung $\mathrm{B}\left(n; \frac{K}{N}\right)$

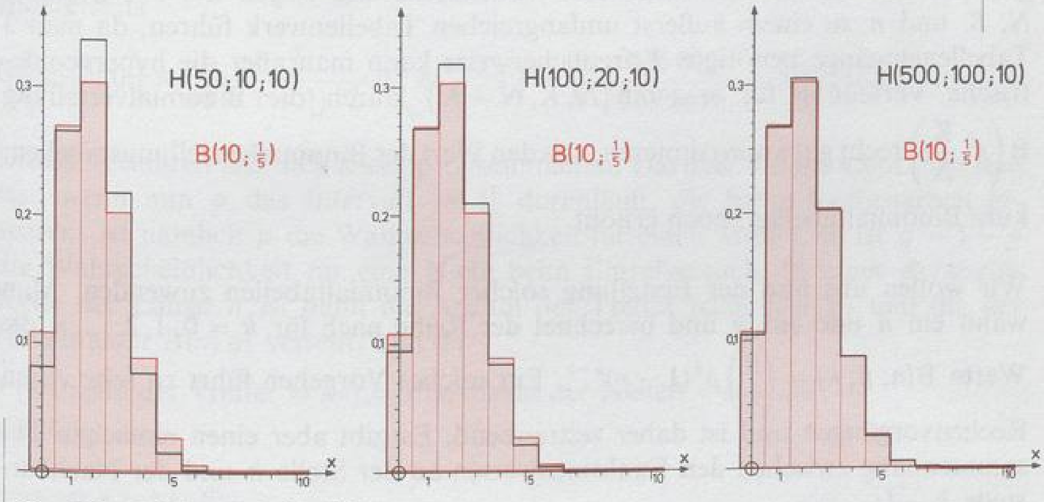


Fig. 233.1. Bild zu Tabelle 234.1.

approximieren. (Vergleiche dazu Aufgabe 264/**27**.) Dies ist gar nicht so erstaunlich, weil ja bei großen Kugelzahlen die Entnahme einiger weniger Kugeln keine wesentliche Änderung der Anteile in der Urne bewirkt. Man kann dann also das Ziehen mit Zurücklegen als gute Näherung für das Ziehen ohne Zurücklegen nehmen. Eine Veranschaulichung geben Tabelle 234.1 und Figur 233.1.

k	$B(10; \frac{1}{5}; k)$	$H(N; K; 10; k)$							
		N	50	100	500	1000	100000	1000000	1000000000
		K	10	20	100	200	20000	200000	200000000
0	0,107374		0,082519	0,095116	0,104951	0,106164	0,107362	0,107373	0,107374
1	0,268435		0,266192	0,267933	0,268417	0,268431	0,268435	0,268435	0,268435
2	0,301990		0,336898	0,318170	0,305050	0,303510	0,302005	0,301991	0,301990
3	0,201327		0,217792	0,209208	0,202849	0,202085	0,201334	0,201327	0,201327
4	0,088080		0,078469	0,084107	0,087395	0,087744	0,088077	0,088080	0,088080
5	0,026424		0,016142	0,021531	0,025488	0,025959	0,026419	0,026424	0,026424
6	0,005505		0,001868	0,003541	0,005096	0,005299	0,005503	0,005505	0,005505
7	0,000786		0,000115	0,000368	0,000689	0,000737	0,000786	0,000786	0,000786
8	0,000074		0,000003	0,000023	0,000060	0,000067	0,000074	0,000074	0,000074
9	0,000004		$4 \cdot 10^{-8}$	0,000001	0,000003	0,000004	0,000004	0,000004	0,000004
10	$1 \cdot 10^{-7}$		$1 \cdot 10^{-10}$	$1 \cdot 10^{-8}$	$7 \cdot 10^{-8}$	$9 \cdot 10^{-8}$	$1 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$1 \cdot 10^{-7}$

Tab. 234.1 Vergleich einer Binomialverteilung mit verschiedenen hypergeometrischen Verteilungen mit gleichem $p = \frac{K}{N}$

14.3. Tabellen der Binomialverteilung

Die Berechnung von Werten einer Binomialverteilung ist rechnerisch meist sehr aufwendig. Da die Binomialverteilung aber eine sehr häufig auftretende Wahrscheinlichkeitsverteilung ist, hat man sie für oft vorkommende Werte der Parameter n und p tabellarisiert. Für die ebenfalls sehr häufig auftretende hypergeometrische Verteilung würde eine Tabellarisierung wegen der 3 Parameter N, K und n zu einem äußerst umfangreichen Tabellenwerk führen, da man 3 Tabelleneingänge benötigte. Erfreulicherweise kann man aber die hypergeometrische Verteilung für $n \ll \min\{N, K, N-K\}$ durch die Binomialverteilung $B\left(n; \frac{K}{N}\right)$ recht gut approximieren, was den Wert der Binomialverteilungstabellen, kurz Binomialtabellen, noch erhöht.

Wir wollen uns nun der Erstellung solcher Binomialtabellen zuwenden. Man wählt ein n und ein p und berechnet der Reihe nach für $k = 0, 1, 2, \ldots, n$ die Werte $B(n; p; k) = \binom{n}{k} p^k (1-p)^{n-k}$. Ein solches Vorgehen führt zu sehr vielen Rechenvorgängen und ist daher zeitraubend. Es gibt aber einen einfachen Zusammenhang zwischen den Funktionswerten an der Stelle k und der Nachbarstelle $k-1$: