Programm

Wiederholung

Gleichverteilung
- Diskrete Gleichverteilung
- Stetige Gleichverteilung

Binomialverteilung

Hypergeometrische Verteilung

Wiederholung

- verschiedene Mittelwerte für verschiedene Skalenniveaus
- robuste und nicht robuste Mittelwerte
- Erwartungswert und Varianz

Erwartungswert

Definition

Der Erwartungswert einer Zufallsvariablen X ist derjenige Wert der Zufallsvariablen, dessen Eintreffen vor der Durchführung des Zufallsexperimentes im Mittel zu erwarten ist.

- diskrete ZV:
$$E(X) = \sum_i x_i \cdot f(x_i)$$
- stetige ZV:
$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

Varianz

Definition

Die Varianz ist die mittlere quadratische Abweichung vom arithmetischen Mittel.

- diskrete ZV:

$$\mathrm{Var}(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

- stetige ZV:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty}(x-\mu)^2 \cdot f(x)dx$$

3 Wichtige Verteilungsmodelle

3.1 Gleichverteilung

Diskrete Gleichverteilung

Eine diskrete Zufallsvariable X mit den endlich vielen Realisationen x_i, $(i = 1, \ldots, n)$ heißt diskret gleichverteilt, wenn jeder Wert von X die gleiche Wahrscheinlichkeit der Realisierung hat.

Wahrscheinlichkeitsfunktion

$$f(x_i) = \begin{cases} \frac{1}{n} & \text{für } i = 1, \ldots, n \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion

$$F(x) = \begin{cases} 0 & \text{für } x < x_1 \\ \frac{i}{n} & \text{für } x_i \leq x < x_{i+1};\ i = 1, \ldots, n-1 \\ 1 & \text{für } x_n \leq x \end{cases}$$

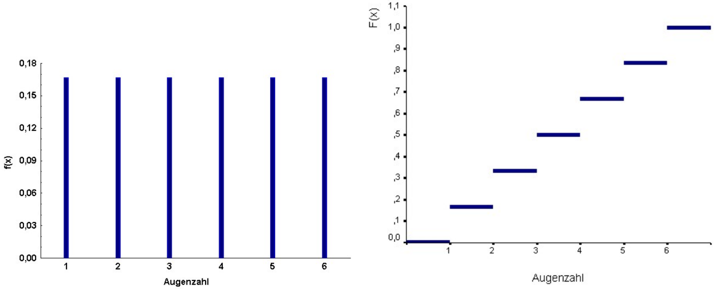
Beispiel 3.1:

Einmaliges Werfen eines idealen Würfels X = “Augenzahl”;

Wertebereich: 1, 2, 3, 4, 5, 6

klassische Definition der Wahrscheinlichkeit:

$f(x_i) = 1/6$ für $i = 1, \ldots, 6$.



Erwartungswert und Varianz

$$
\begin{aligned}
E(X) &= \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \\
Var(X) &= \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Diskrete Gleichverteilung hängt vom Parameter n ab

Stetige Gleichverteilung

Eine stetige Zufallsvariable X, die nur Werte im Intervall $[a, b]$ annehmen kann und deren Dichtefunktion in diesem Intervall positiv und konstant und sonst Null ist, heißt stetig gleichverteilt.

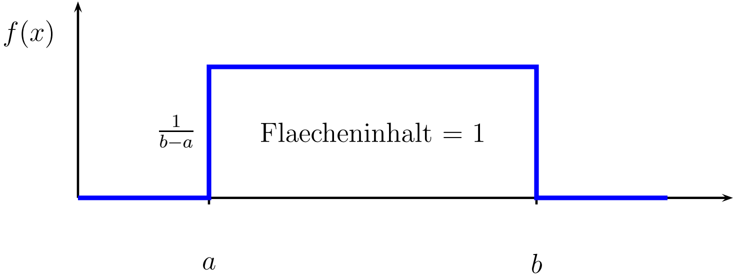
Wahrscheinlichkeitsdichte

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

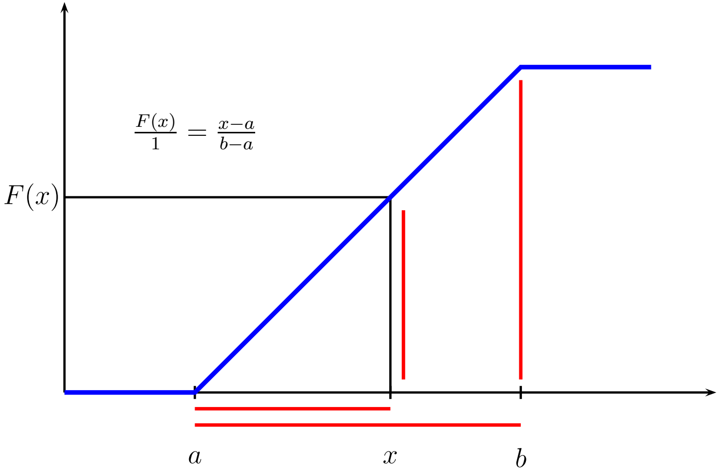
Verteilungsfunktion

$$F(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x < b \\ 1 & \text{für } b \leq x \end{cases}$$

Herleitung Dichtefunktion



Herleitung Verteilungsfunktion



Erwartungswert und Varianz

$$
\begin{aligned}
E(X) &= \frac{b+a}{2} \\
Var(X) &= \frac{(b-a)^2}{12}
\end{aligned}
$$

Stetige Gleichverteilung hängt von den Parametern a und b ab

Herleitung Erwartungswert und Erwartungswert

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x \cdot \frac{1}{b-a} = \int_{a}^{b} x \cdot \frac{1}{b-a} \\
&= \left| \frac{x^2}{2(b-a)} \right|_a^b \\
&= \frac{1}{2}\frac{1}{(b-a)} \cdot b^2 - \frac{1}{2}\frac{1}{b-a} a^2 \\
&= \frac{1}{2}\frac{1}{(b-a)} \cdot (b^2 - a^2) \\
&= \frac{1}{2}\frac{1}{(b-a)} \cdot (b-a)(b+a) \\
&= \underline{\underline{\frac{a+b}{2}}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \int_{-\infty}^{+\infty} (x - \mathrm{E}(X))^2 f(x)dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{(b-a)} dx \\
&= \left| \frac{1}{(b-a)} \cdot \frac{1}{3} \cdot \left(x - \frac{a+b}{2}\right)^3 \right|_a^b \\
&= \frac{1}{3(b-a)} \cdot \left(\left(b - \frac{a+b}{2}\right)^3 - \left(a - \frac{a+b}{2}\right)^3 \right) \\
&= \frac{1}{3(b-a)} \cdot \left(\left(\frac{2b-a-b}{2}\right)^3 - \left(\frac{2a-a-b}{2}\right)^3 \right) \\
&= \frac{1}{3(b-a)} \cdot \left(\left(\frac{b-a}{2}\right)^3 + \left(\frac{b-a}{2}\right)^3 \right) \\
&= \frac{1}{3(b-a)} \cdot \left(\left(\frac{b-a}{2}\right)\left(\frac{b-a}{2}\right)^2 + \left(\frac{b-a}{2}\right)\left(\frac{b-a}{2}\right)^2 \right) \\
&= \frac{1}{3} \cdot \left[\frac{1}{2}\left(\frac{b-a}{2}\right)^2 + \frac{1}{2}\left(\frac{b-a}{2}\right)^2 \right] \\
&= \frac{1}{3} \cdot \left(\frac{b-a}{2}\right)^2 = \underline{\underline{\frac{(b-a)^2}{12}}}
\end{aligned}
$$

Beispiel 3.2

ZV X 'Wartezeit auf die nächste S-Bahn in Minuten'
Wertebereich $[a, b] = [0, 20]$

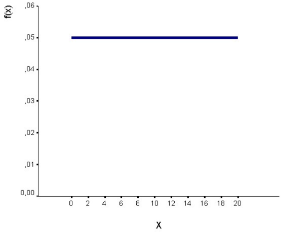
$$P(0 \leq X \leq 20) = 1$$

$$P(a \leq X \leq b) = 1 = k \cdot (b - a) = \int_a^b k \quad dx$$

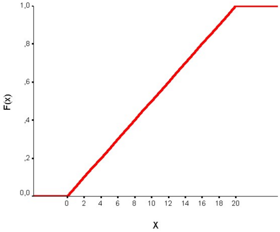
$$\Rightarrow k = \frac{1}{b - a} = \frac{1}{20} = 0.05$$

$$f(x) = \begin{cases} \frac{1}{20} & 0 < x \leq 20 \\ 0 & \text{sonst} \end{cases} \qquad F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{20} \cdot x & 0 \leq x < 20 \\ 1 & \text{sonst} \end{cases}$$

Dichtefunktion



Verteilungsfunktion



$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x f(x)\, dx = \int_{0}^{20} x \frac{1}{20}\, dx \\
&= \frac{1}{20}\left[\frac{1}{2}x^2\right]_0^{20} \\
&= \frac{1}{20}\left[\frac{1}{2}20^2 - \frac{1}{2}0^2\right] = 10
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx \\
&= \int_{0}^{20} (x-10)^2 \cdot \frac{1}{20}\,dx \\
&= \frac{1}{20} \int_{0}^{20} (x^2 - 20x + 100)\,dx \\
&= \frac{1}{20} \left[\frac{1}{3}x^3 - \frac{1}{2}20x^2 + 100x \right] \\
&= \frac{1}{20} \left[\frac{1}{3}20^3 - \frac{1}{2}20^3 + 100 \cdot 20 \right] = 33,33
\end{aligned}
$$

Bernoulli-Verteilung

- Zufallsvariable X mit zwei Ausprägungen, 1 und 0
- zugrundeliegendes Zufallsexperiment mit A (Erfolg) und $\overline{A}$ (Misserfolg)
- A tritt mit W. p ein, $\overline{A}$ mit Gegenwahrscheinlichkeit $1 - p$
- A tritt ein, ZV X erhält den Wert 1, bei $\overline{A}$ den Wert 0

$$P(X = 1) = P(A) = p$$

$$P(X = 0) = P(\overline{A}) = 1 - p$$

Bernoulli-Verteilung

Wahrscheinlichkeitsfunktion

$$f_{\text{Be}}(x,p) = \begin{cases} 1-p & \text{für x=0} \\ p & \text{für x=1} \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion

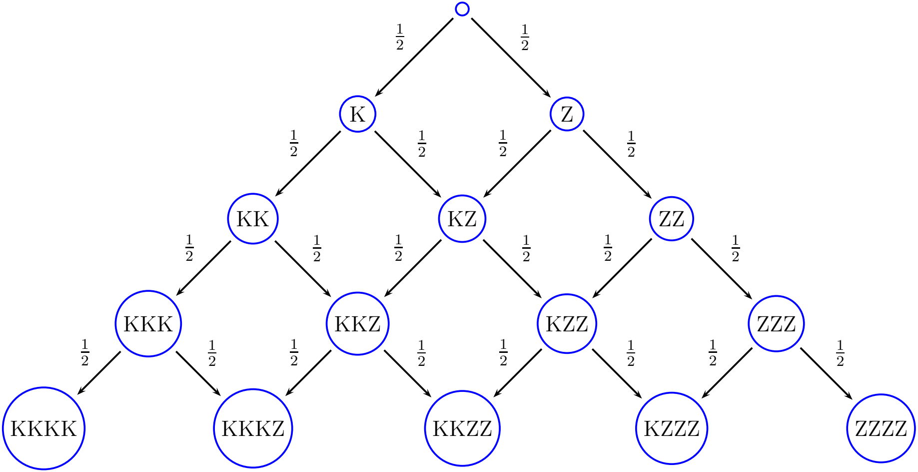
$$F_{\text{Be}}(x,p) = \begin{cases} 0 & \text{für } x < 0 \\ 1-p & \text{für } 0 \le x < 1 \\ 1 & \text{für } 1 \le x \end{cases}$$

$$E(X) = 0 \cdot (1-p) + 1 \cdot p = p$$

$$\text{Var}(X) = (0-p)^2(1-p) + (1-p)^2 p = p(1-p)$$

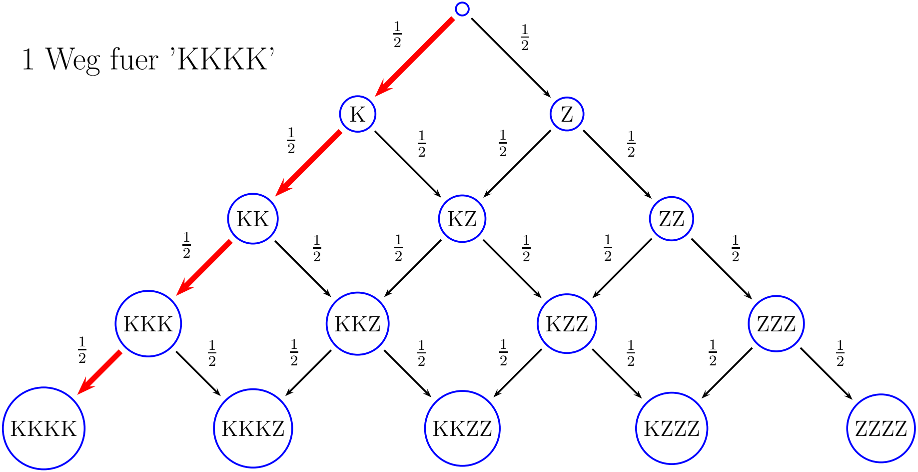
Binomialverteilung - Einführendes Beispiel

- 4-maliger Münzwurf
- P('Kopf') = P('Zahl') = $\frac{1}{2}$
- Würfe sind unabhängig
- Zufallsvariable X 'Anzahl von Kopf bei n=4 Würfen'

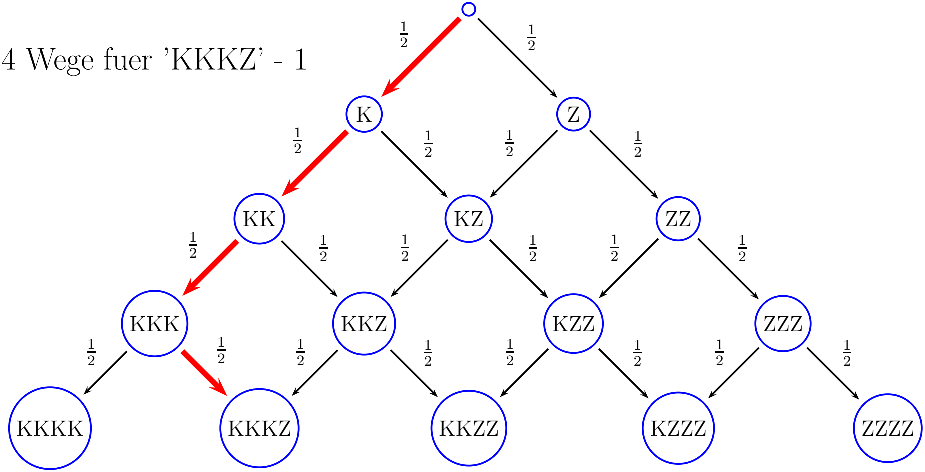


K
Z
KK
KZ
ZZ
KKK
KKZ
KZZ
ZZZ
KKKK
KKKZ
KKZZ
KZZZ
ZZZZ
$\frac{1}{2}$

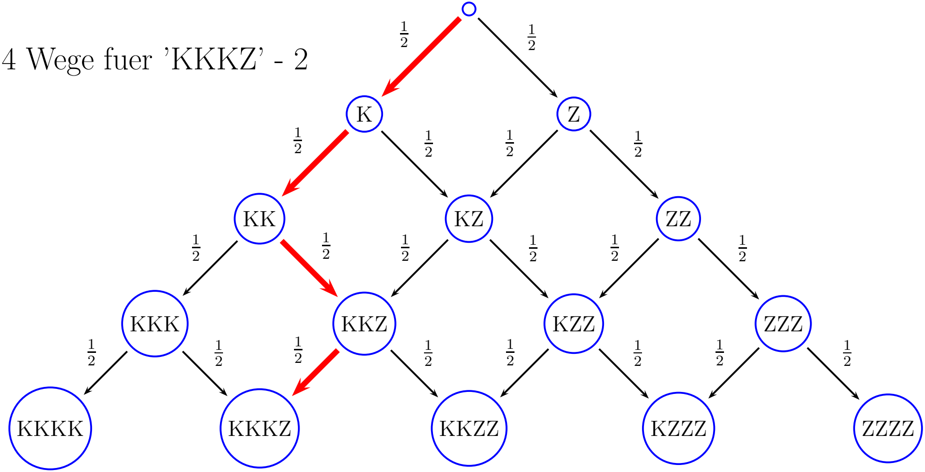
1 Weg fuer 'KKKK'



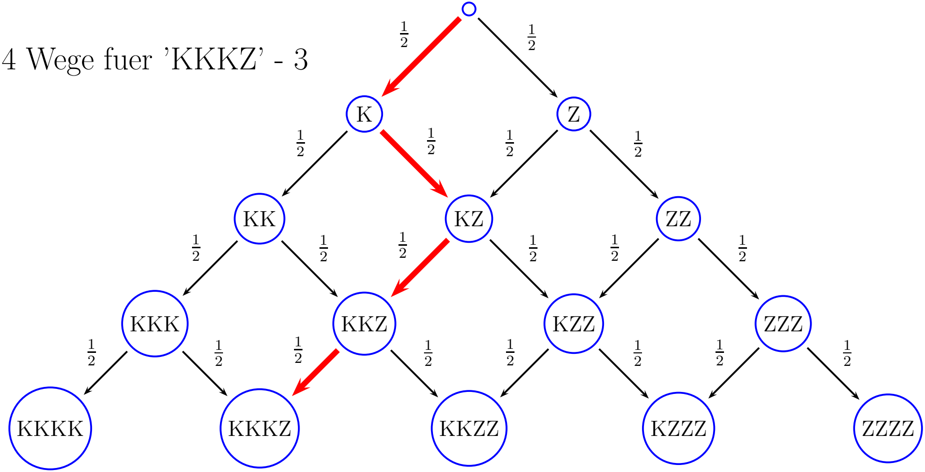
4 Wege fuer 'KKKZ' - 1



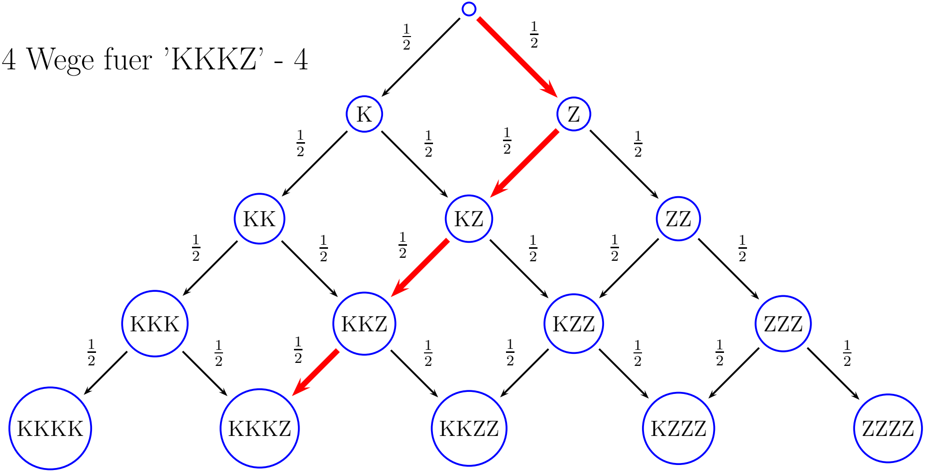
4 Wege fuer 'KKKZ' - 2



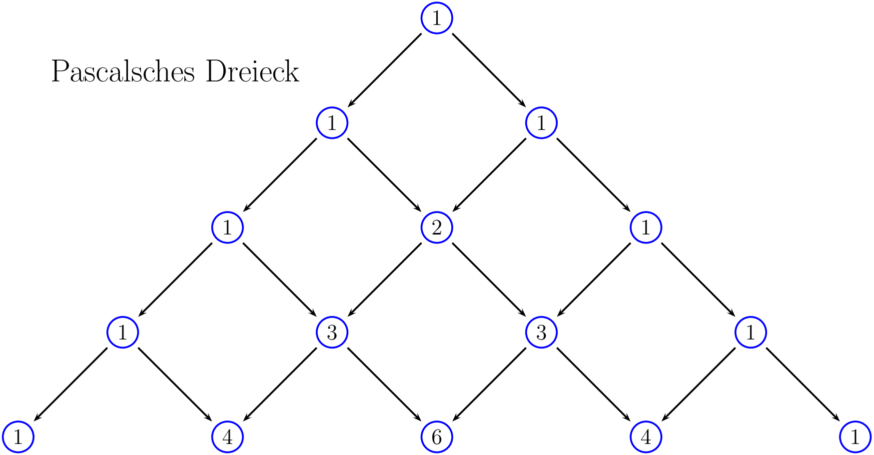
4 Wege fuer 'KKKZ' - 3



4 Wege fuer 'KKKZ' - 4



Pascalsches Dreieck



Binomialkoeffizient

Binomialkoeffizient

$$\text{Anzahl der Kombinationen} = \left(\begin{array}{c} n \\ x \end{array} \right) = \frac{n!}{x! \cdot (n-x)!}$$

Kombinationen

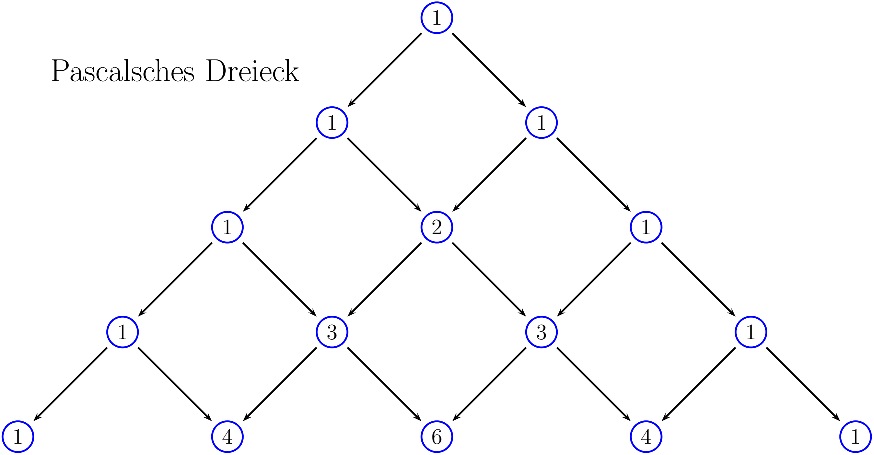
Eine Auswahl von x Objekten aus einer Menge von n Objekten ohne Berücksichtigung der Reihenfolge nennt man Kombination.

$(n+1)$-Zeile, $(x+1)$-Spalte

$$\left(\begin{array}{c} 4 \\ 2 \end{array} \right) = \frac{4!}{2! \cdot (4-2)!} = 6$$

$(4+1)$-Zeile, $(2+1)$-Spalte $\Rightarrow 6$

Pascalsches Dreieck



$X = \{$Anzahl des Auftretens von A bei n Versuchen$\}$ Wertebereich: $x = 0, 1, 2, \ldots, n$
$\Rightarrow X$ – diskrete Zufallsvariable

$$X = \sum_{i=1}^{n} X_i$$

$$P(X = x) = f(x) = ?$$

Eintreten der Realisation $X = x \rightarrow$ z.B. wenn Ereignisfolge

$$\underbrace{A_1 \cap A_2 \cap \ldots \cap A_x}_{x \text{ mal } A} \cap \underbrace{\overline{A}_{x+1} \cap \overline{A}_{x+2} \cap \ldots \cap \overline{A}_n}_{(n-x) \text{ mal } \overline{A}}$$

eintritt.

Reihenfolge der Anordnung spielt keine Rolle

$$\overline{A}_{x+1}, A_1, A_2, \ldots, A_x, \overline{A}_{x+2}, \ldots, \overline{A}_n$$

Wahrscheinlichkeit, diese Ereignisfolge zu erhalten:

$$\begin{aligned}
&P(A_1 \cap A_2 \cap \ldots \cap A_x \cap \overline{A}_{x+1} \cap \overline{A}_{x+2} \cap \ldots \cap \overline{A}_n) \\
&\quad = \quad P(A_1) \cdot \ldots \cdot P(A_x) \cdot P(\overline{A}_{x+1}) \cdot \ldots \cdot P(\overline{A}_n) \\
&\quad = \quad \underbrace{p \cdot \ldots \cdot p}_{x} \cdot \underbrace{(1-p) \cdot \ldots \cdot (1-p)}_{n-x} \\
&\quad = \quad p^x \cdot (1-p)^{n-x}
\end{aligned}$$

Anzahl der verschiedenen Ereignisfolgen, x-mal das Ereignis A bei n Versuchen zu erhalten $=$ Anzahl der Kombinationen ohne Wiederholung

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

$$\Rightarrow \quad P(X = x) = f(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$

Wahrscheinlichkeitsfunktion

$$f_B(x;n,p) = \begin{cases} \binom{n}{x} \, p^x \, (1-p)^{n-x} & \text{für } x = 0, 1, \ldots, n \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion

$$F_B(x;n,p) = \begin{cases} \sum_{k=0}^{x} \binom{n}{k} \, p^k \, (1-p)^{n-k} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

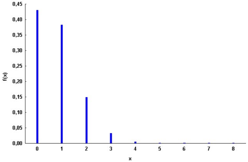
$$E(X) = n \cdot p$$

$$Var(X) = n \cdot p \cdot (1-p)$$

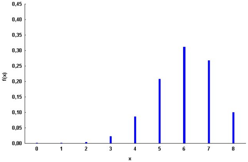
Die Binomialverteilung hängt von den Parametern n und p ab

Kurzschreibweise:

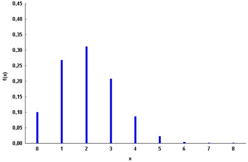
$$X \sim B(n;p)$$



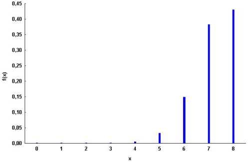
B(8; 0,10)



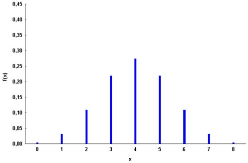
B(8; 0,75)



B(8; 0,25)



B(8; 0,90)



B(8; 0,50)

Beispiel 3.3:

Urne mit 10 Kugeln, 3 weiße, 7 rote Kugeln

$A = \{$Ziehen einer weißen Kugel$\}$
$\bar{A} = \{$Ziehen einer roten Kugel$\}$

Ziehen mit Zurücklegen $\Rightarrow$ Versuche unabhängig
$P(A) = p = 0,3 \quad P(\bar{A}) = 1 - p = 0,7 \quad \rightarrow$ konstant

$n = 5$ Ziehungen $X_i = \{$Anzahl des Auftretens einer weißen Kugel bei der i-ten Ziehung$\}$, $i = 1, \ldots, 5$

X_i	$P(X_i = x_i) = f(x_i)$
1	0,3
0	0,7

$X = \{$Anzahl des Auftretens weißer Kugeln bei $n = 5$ Ziehungen mit Zurücklegen$\}$

$X = \sum_i X_i$

$X \sim B(n;p) = B(5;0,3)$

$X = 2$ (weiße Kugeln)

$$\begin{aligned} P(X=2) &= f_B(2;5;0,3) \\ &= \binom{5}{2} \cdot 0,3^2 \cdot 0,7^3 = 0,3087 \end{aligned}$$

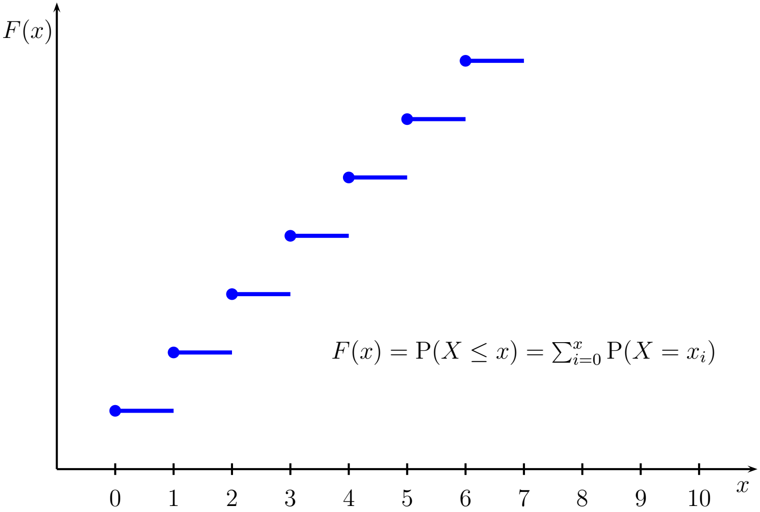
Wahrscheinlichkeits- und Verteilungsfunktion der Binomialverteilung $B(5; 0,3)$

x	$f_B(x; n, p)$	$F_B(x; n.p)$
0	0,1681	0,1681
1	0,3601	0,5282
2	0,3087	0,8369
3	0,1323	0,9692
4	0,0284	0,9976
5	0,0024	1,0000

Berechnung mit Verteilungsfunktion:

$$
\begin{aligned}
f_B(2;5;0,3) &= F_B(2;5;0,3) - F_B(1;5;0,3) \\
&= 0,8369 - 0,5282 \\
&= 0,3087
\end{aligned}
$$

Verteilungsfunktion Binomialverteilung



$Y_i = \{$Anzahl des Auftretens einer roten Kugel bei der i-ten Ziehung$\}$, $i = 1, \ldots, 5$

Y_i	$P(Y_i = y_i) = f(y_i)$
1	0,7
0	0,3

$Y = \{$Anzahl des Auftretens roter Kugeln bei $n = 5$ Ziehungen mit Zurücklegen$\}$

$Y = \Sigma_i Y_i$

$Y \sim B(n; p) = B(5; 0,7)$

$Y = 2$ (rote Kugeln)

$$\begin{aligned} P(Y = 2) &= f_B(2; 5; 0,7) \\ &= \binom{5}{2} \cdot 0,7^2 \cdot 0,3^3 = 0,1323 \end{aligned}$$

Alternative Berechnung

$$P(\text{'2 rote Kugeln'}) = P(\text{'3 weisse Kugeln'})$$

$$f_B(2, 5, 0.7) = f_B(3, 5, 0.3)$$

$$f_B(y, n, p) = f_B(n - y, n, 1 - p)$$

$$f_B(3, 5, 0.3) = 0.1323$$

$$f_B(3, 5, 0.3) = F_B(3, 5, 0.3) - F_B(2, 5, 0.3)$$

$$f_B(3, 5, 0.3) = 0.9692 - 0.8369 = 0.1323$$

Wahrscheinlichkeits- und Verteilungsfunktion der Binomialverteilung $B(5; 0,3)$

x	$f_B(x; n, p)$	$F_B(x; n.p)$
0	0,1681	0,1681
1	0,3601	0,5282
2	0,3087	0,8369
3	0,1323	0,9692
4	0,0284	0,9976
5	0,0024	1,0000

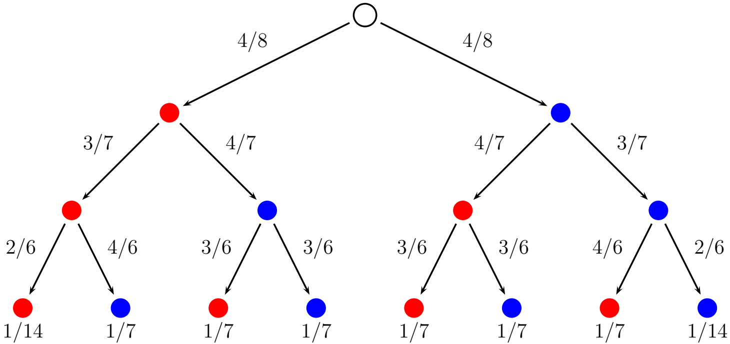
3.3 Hypergeometrische Verteilung

Zufallsexperiment:

- Gesamtheit: endliche Anzahl N von Objekten
  - $\rightarrow$ M Objekte mit Eigenschaft A
  - $\rightarrow$ $N - M$ Objekte ohne Eigenschaft A
- zwei Ereignisse A und $\bar{A}$
- zufällige Auswahl von n Objekten
- Zufallsauswahlmodell ohne Zurücklegen
  - $\rightarrow$ keine Unabhängigkeit der Ziehungen
  - $\rightarrow$ $P(A)$, $P(\bar{A})$ nicht konstant

Die Hypergeometrische Verteilung

N=8 Kugeln, M=4 davon rot, n=3-mal ziehen OHNE Zuruecklegen



P(1 rote Kugel)=$\frac{1}{7} \cdot 3$ P(3 rote Kugeln)=$\frac{1}{14}$

P(2 rote Kugeln)=$\frac{1}{7} \cdot 3$ P(0 rote Kugeln)=$\frac{1}{14}$