Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

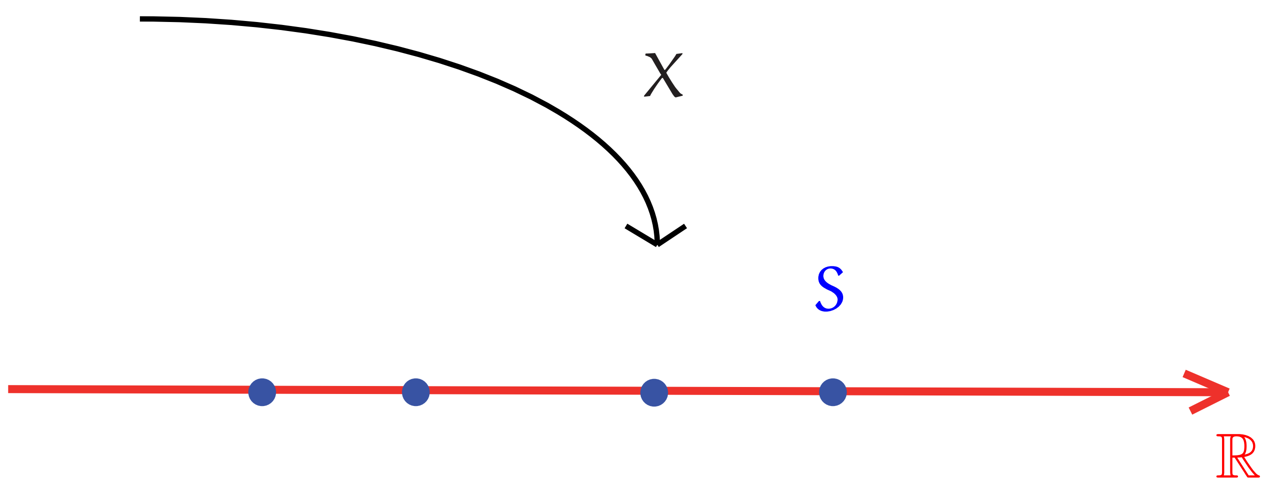
X sei eine diskrete reellwertige Zufallsvariable, d.h.

eine ZV'e mit Wertebereich $\mathbb{R}$ (oder einer Teilmenge davon),

sodass eine

endliche oder abzählbar unendliche Menge S existiert mit

$$\mathbf{P}(X \in S) = 1.$$



1. Der Erwartungswert als gewichtetes Mittel

Eine einprägsame Kenngröße

für die *Lage* der Verteilung von X

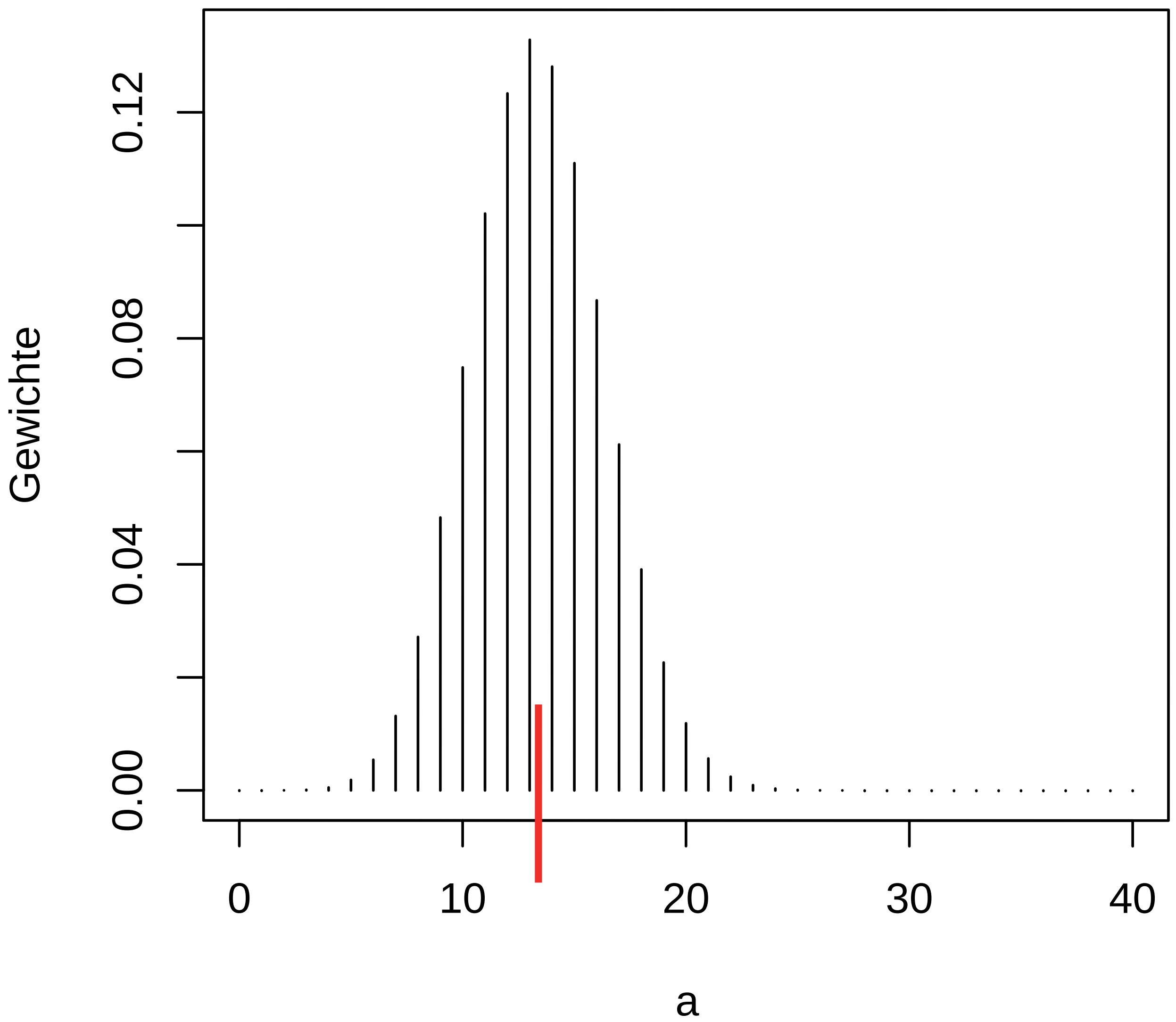
ist das mit den Wahrscheinlichkeiten gewichtete Mittel

der möglichen Werte von X:

$$\mathbf{E}[X] := \sum_{a \in S} a \, \mathbf{P}(X = a) \ .$$

Man spricht vom *Erwartungswert* von X.

(Wir bezeichnen ihn auch mit μ oder μ_X.)



Gewichte
0.00
0.04
0.08
0.12
0
10
20
30
40
a

Das elementarste Beispiel:

$$Z = \begin{cases} 1 \text{ mit Wahrscheinlichkeit } p \\ 0 \text{ mit Wahrscheinlichkeit } q = 1 - p \end{cases}$$

Man erinnere sich an die Situation der ersten Stunde:

Rein zufällige Wahl eines Pixels aus einem Quadrat,

die Teilmenge A hatte den Pixelanteil p;

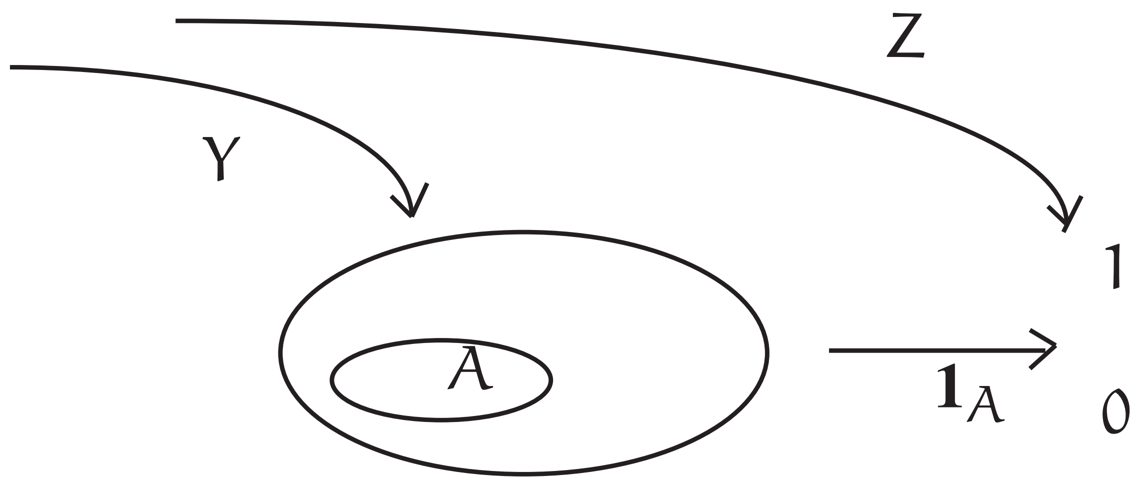
gezählt wird, wenn der Pixel in A fällt.

Das elementarste Beispiel:

$$Z = \begin{cases} 1 \text{ mit Wahrscheinlichkeit } p \\ 0 \text{ mit Wahrscheinlichkeit } q = 1 - p \end{cases}$$

$$\mathbf{E}[Z] = p \cdot 1 + q \cdot 0 = p$$

Das passt gut zu unserem Logo der ersten Stunde



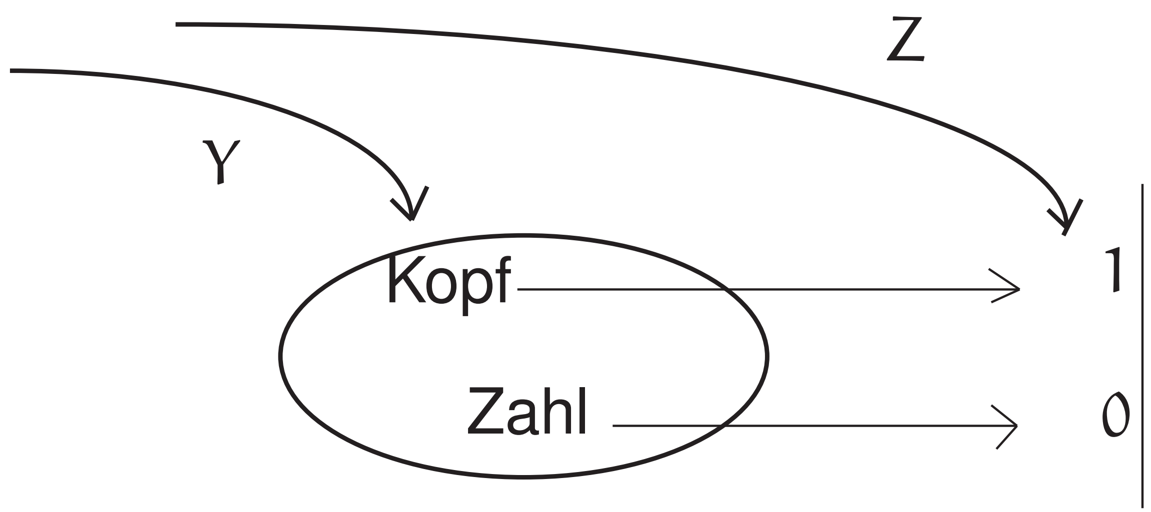
$$Z = \mathbf{1}_A(Y) =: I_{\{Y \in A\}}$$

.... die *Indikatorvariable* des Ereignisses $\{Y \in A\}$

$$\{Z = 1\} = \{Y \in A\}$$

$$\mathbf{E}[Z] = \mathbf{P}(Y \in A).$$

... und entspricht dem Szenario des einfachen Münzwurfs:



$$Z = I_{\{Y=\text{Kopf}\}}$$

$$\{Z = 1\} = \{Y = \text{Kopf}\}$$

$$\mathbf{E}[Z] = \mathbf{P}(Y = \text{Kopf}).$$

Für allgemeines diskretes, reellwertiges X hatten wir

$$\mathbf{E}[X] = \sum_{a \in S} a \, \mathbf{P}\{X = a\}$$

$$= \sum_{a \in S} a \, \rho(a)$$

Dabei sind die Zahlen $\rho(a)$ die Verteilungsgewichte von X.

Merke:

Der Erwartungswert der Zufallvariablen X

hängt nur von ihrer Verteilung ρ ab.

Synonym sprechen wir daher auch manchmal vom

Erwartungswert der Verteilung ρ .

$$X$$

eine Zufallsgröße;

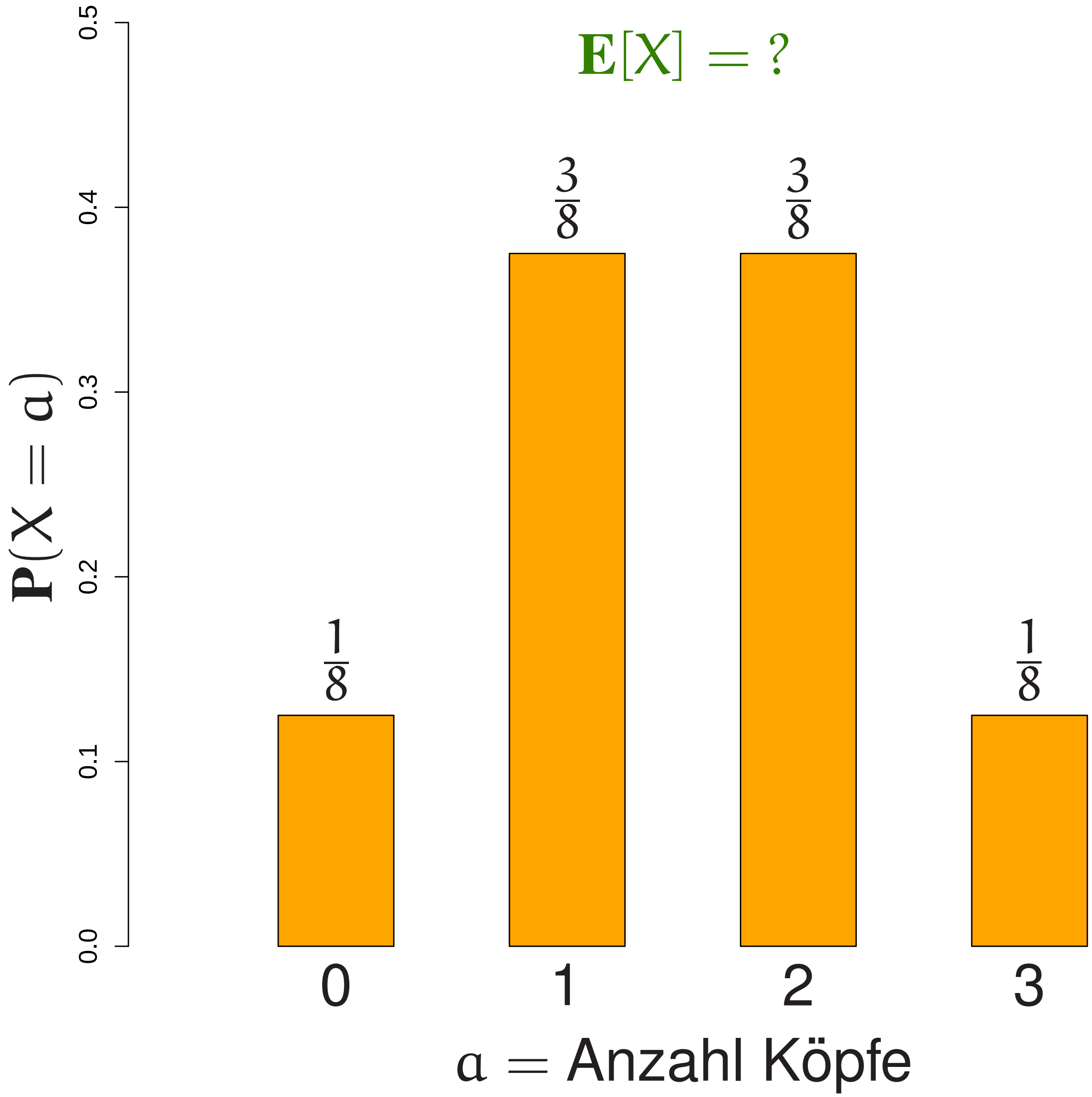
$$\mathbf{E}[X]$$

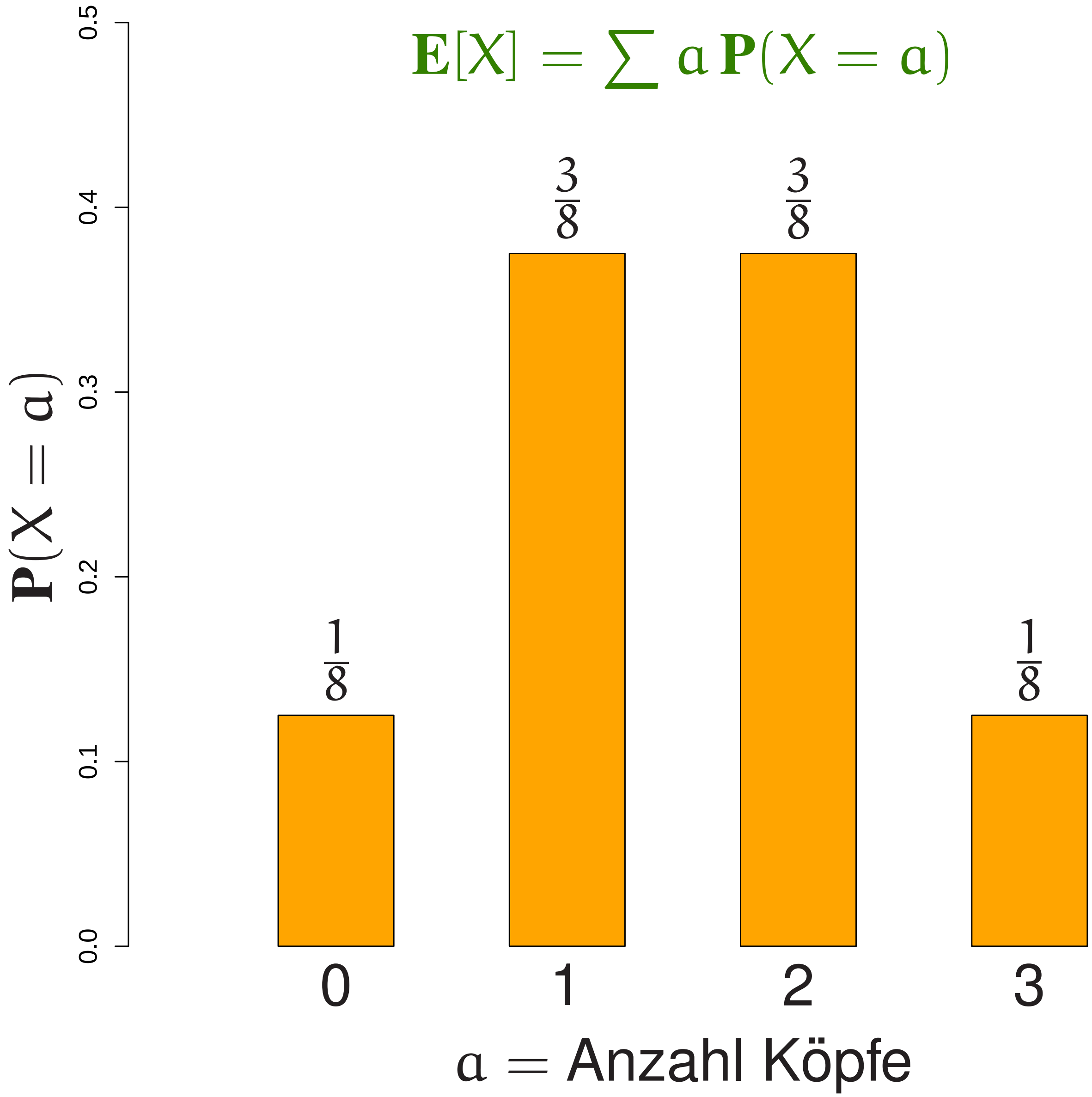
eine Zahl.

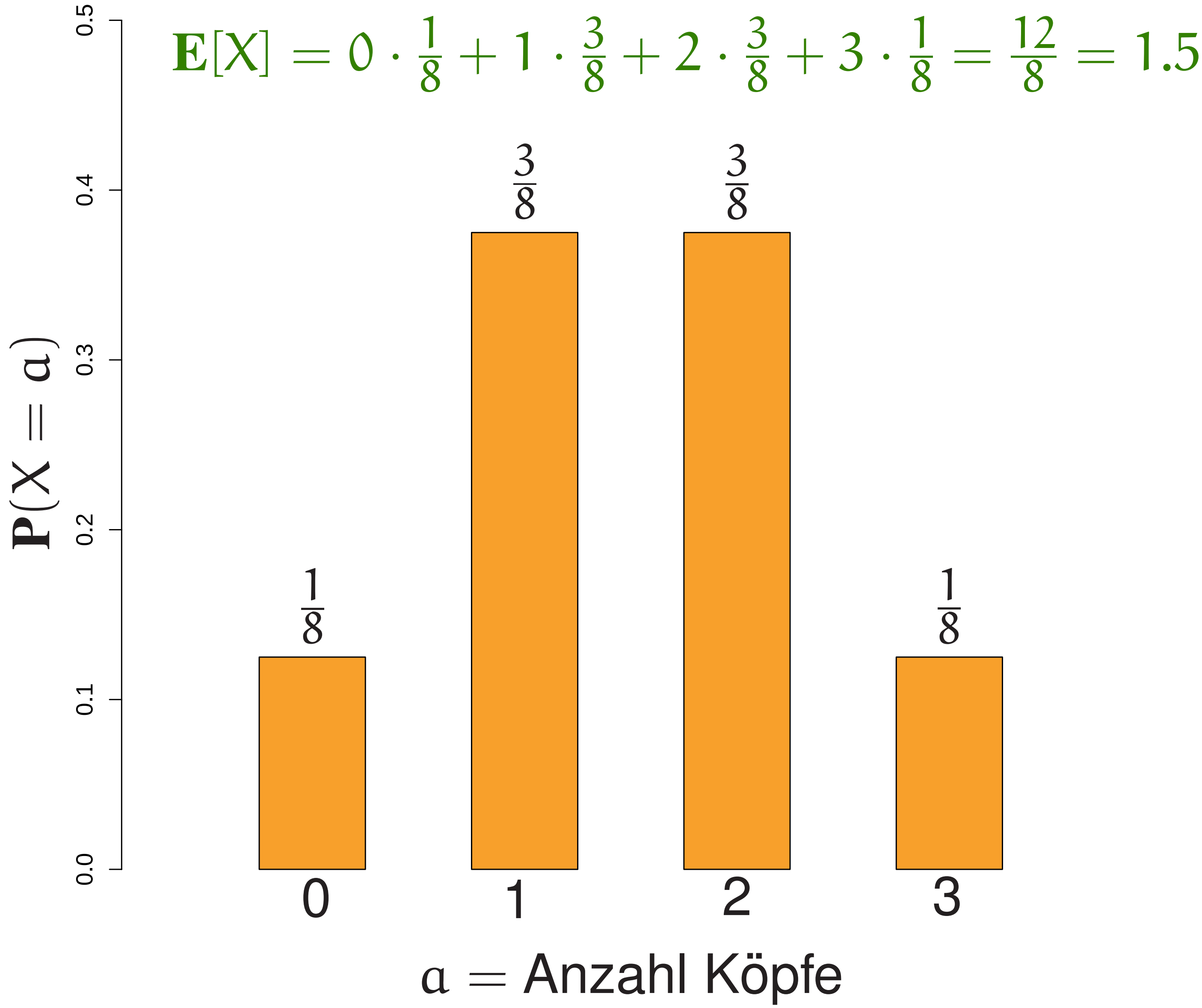
2. Ein Beispiel: Der Erwartungswert der Anzahl der Erfolge beim dreifachen Münzwurf

Eine faire Münze wird dreimal geworfen.

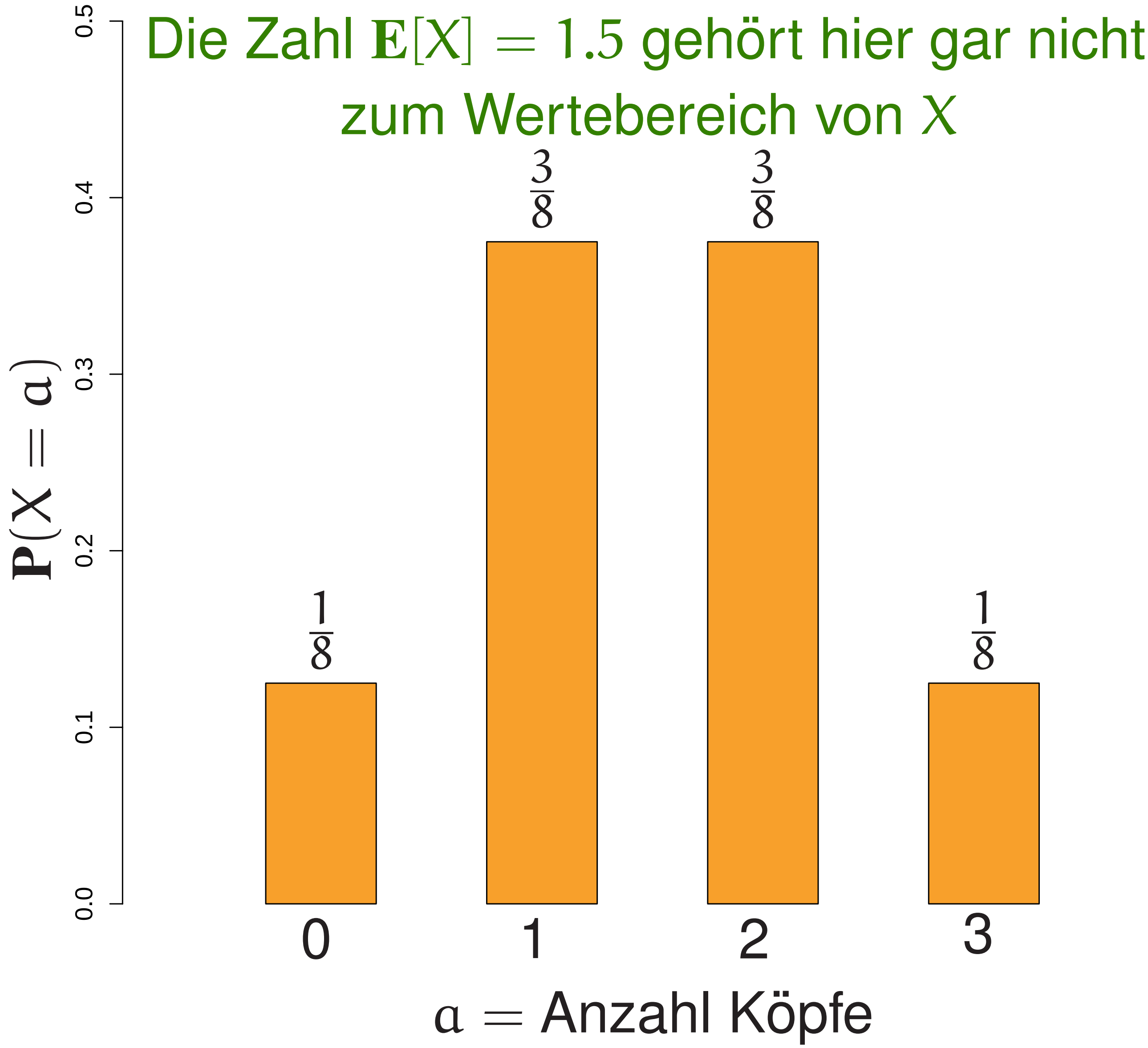
$X :=$ Anzahl der geworfenen Köpfe.

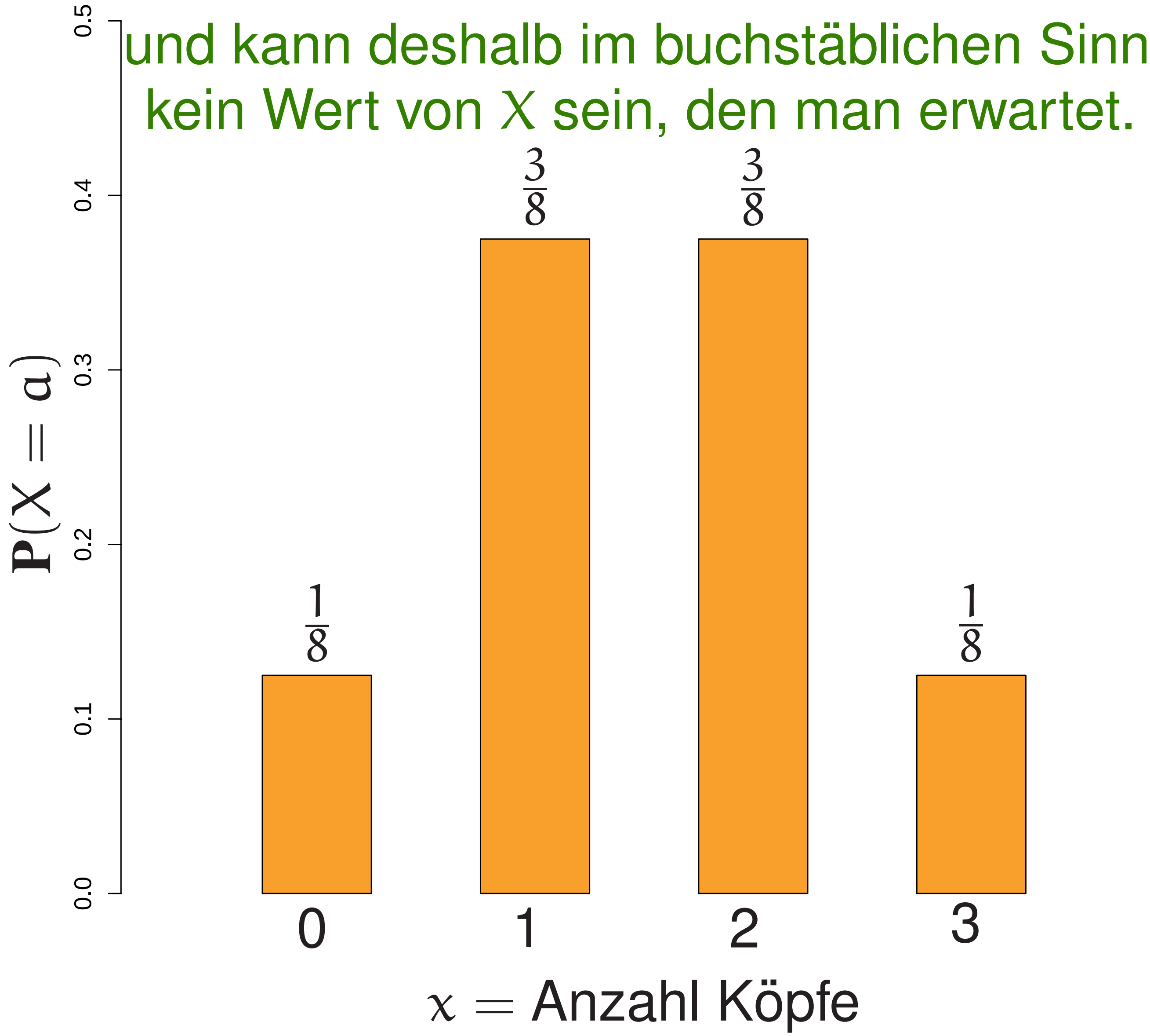

$\mathbf{E}[X] = ?$
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe

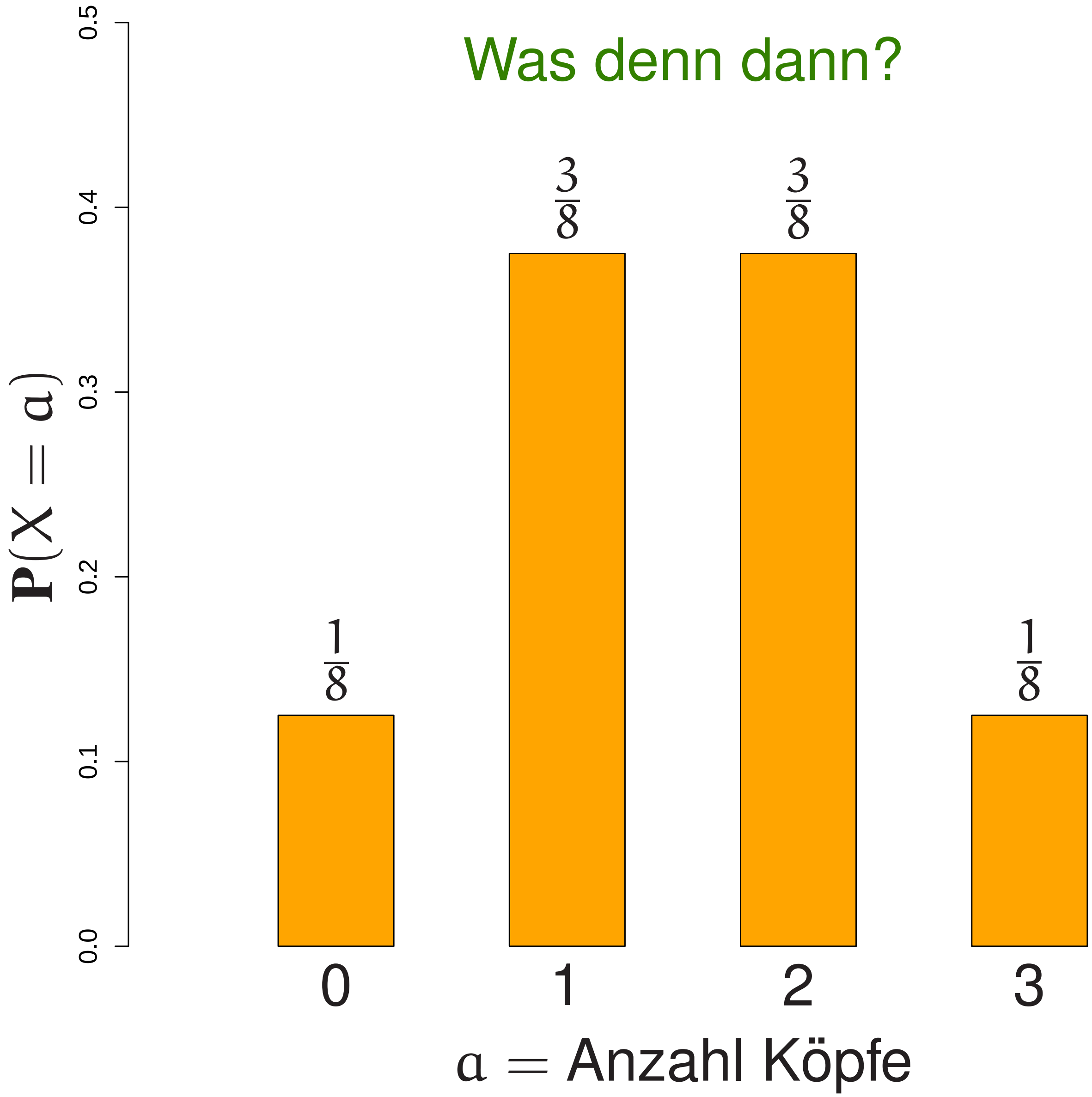

$\mathbf{E}[X] = \sum a \, \mathbf{P}(X = a)$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Köpfe

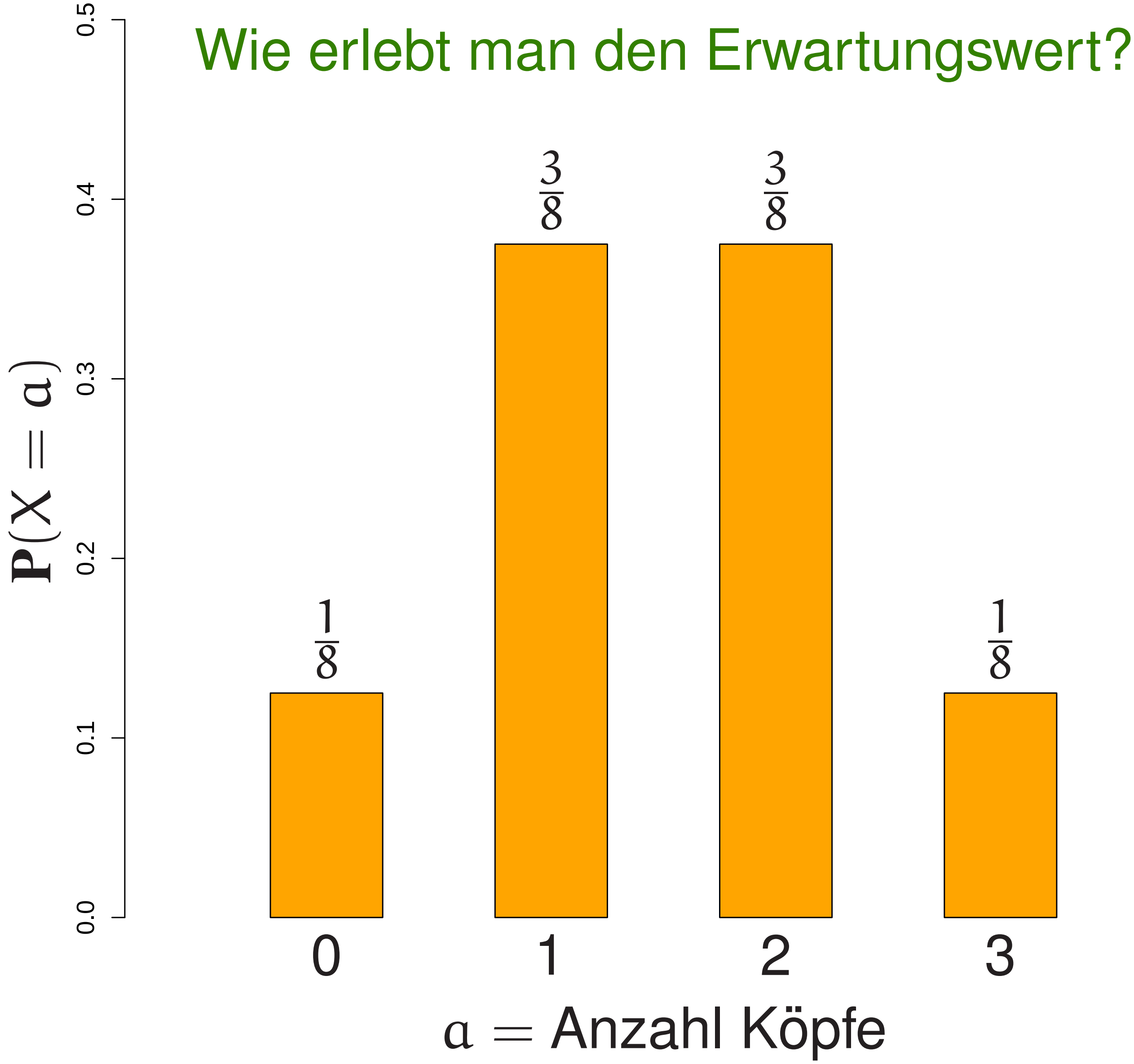


$\mathbf{E}[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Köpfe

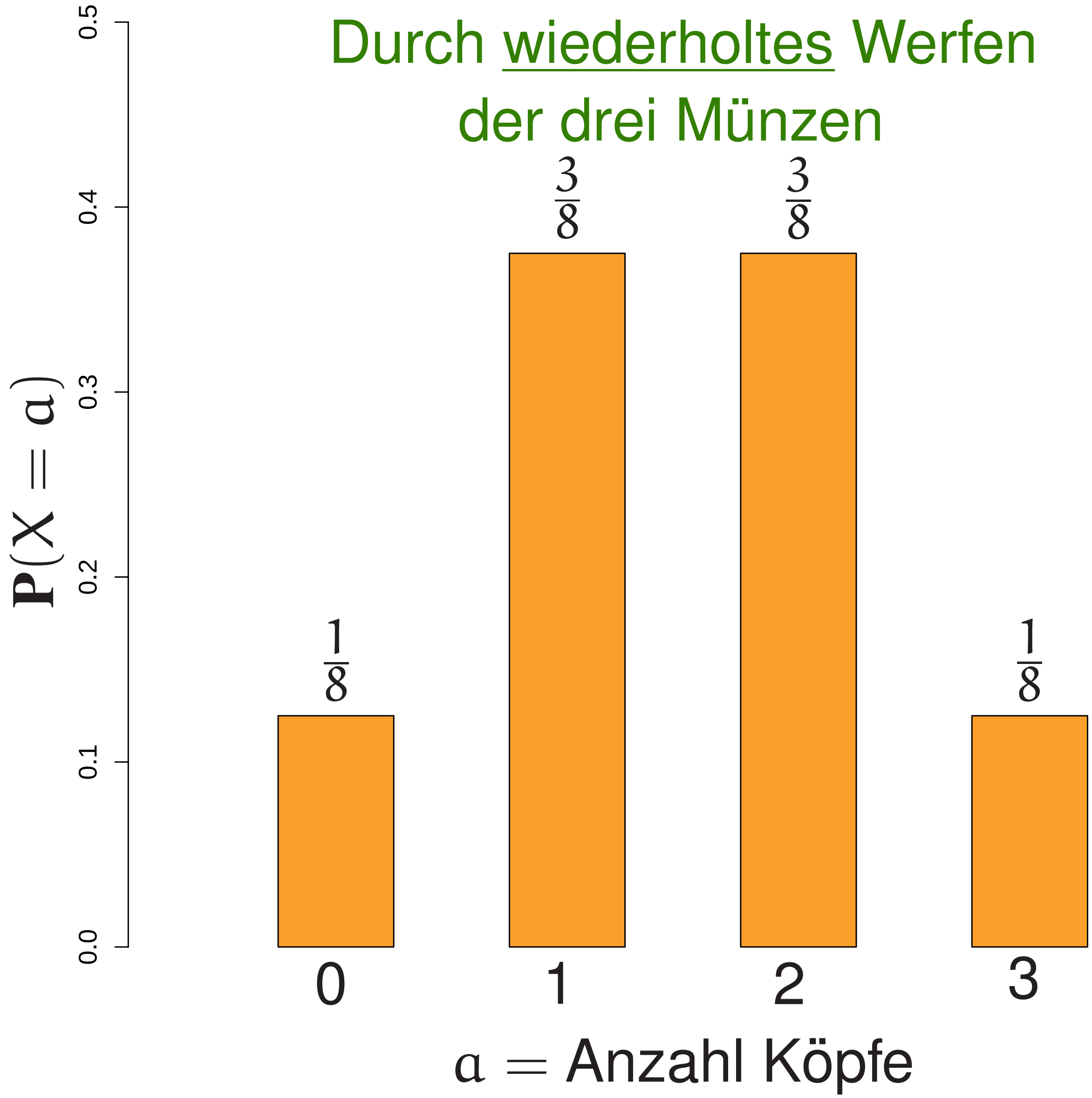

Die Zahl $\mathbf{E}[X] = 1.5$ gehört hier gar nicht
zum Wertebereich von X
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
$a =$ Anzahl Köpfe


und kann deshalb im buchstäblichen Sinn
kein Wert von X sein, den man erwartet.
0.5
0.4
0.3
0.2
0.1
0.0
P(X = a)
3/8
3/8
1/8
1/8
0
1
2
3
x = Anzahl Köpfe


Was denn dann?
0.0
0.1
0.2
0.3
0.4
0.5
P(X = a)
1/8
3/8
3/8
1/8
0
1
2
3
a = Anzahl Köpfe



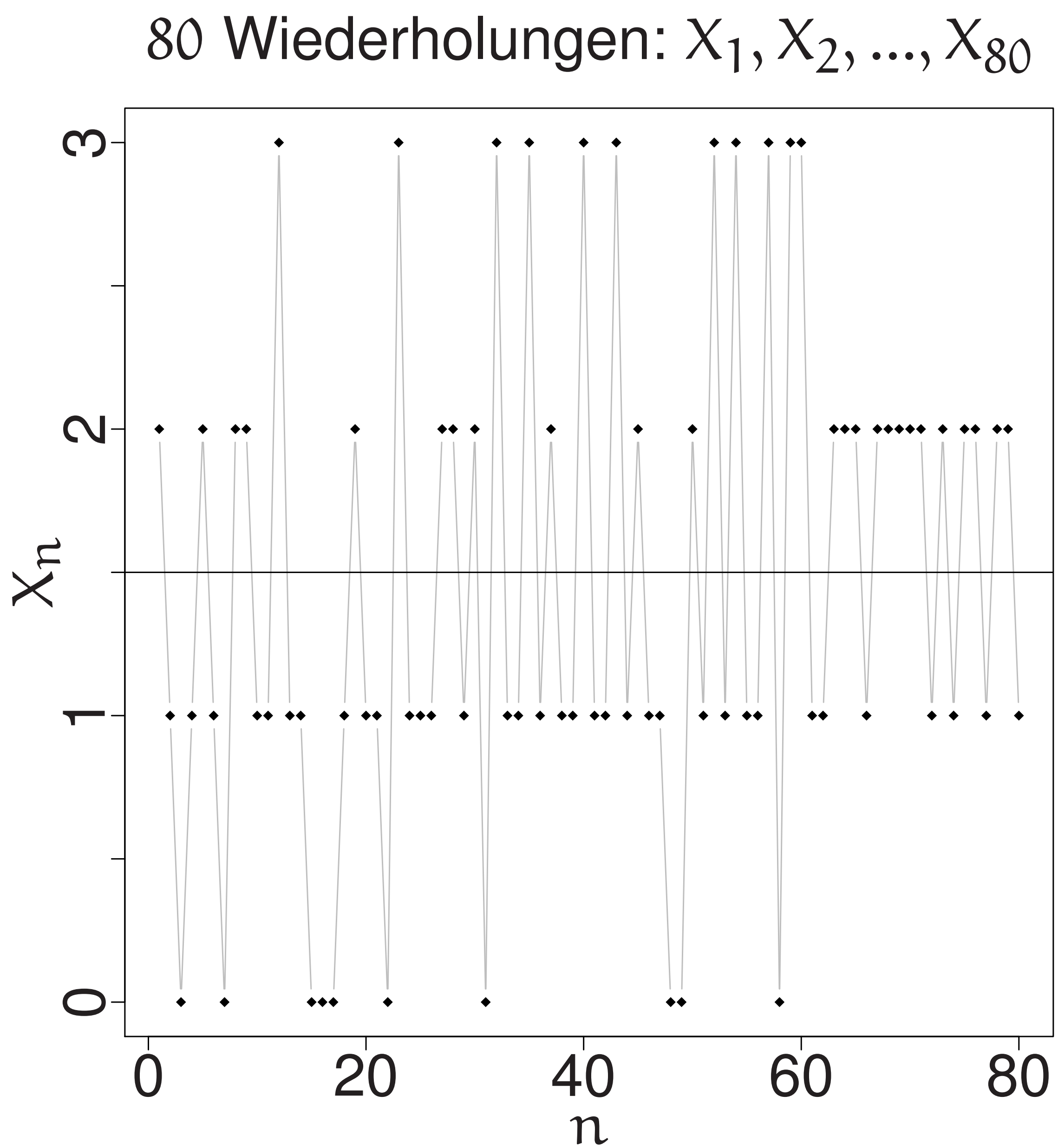
Wie erlebt man den Erwartungswert?
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe


Durch wiederholtes Werfen
der drei Münzen
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe

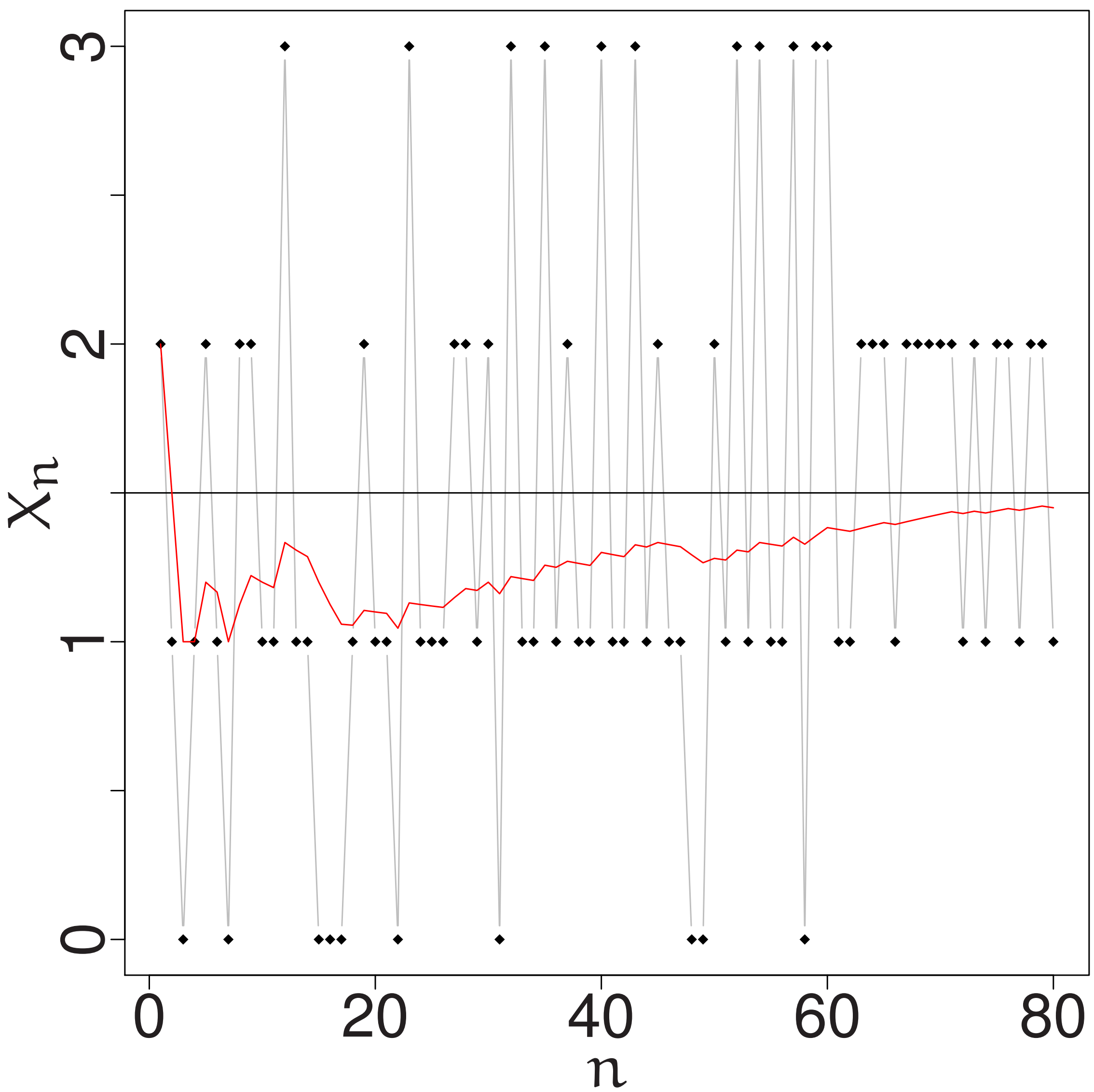
3. Der Erwartungswert als Langzeitmittel

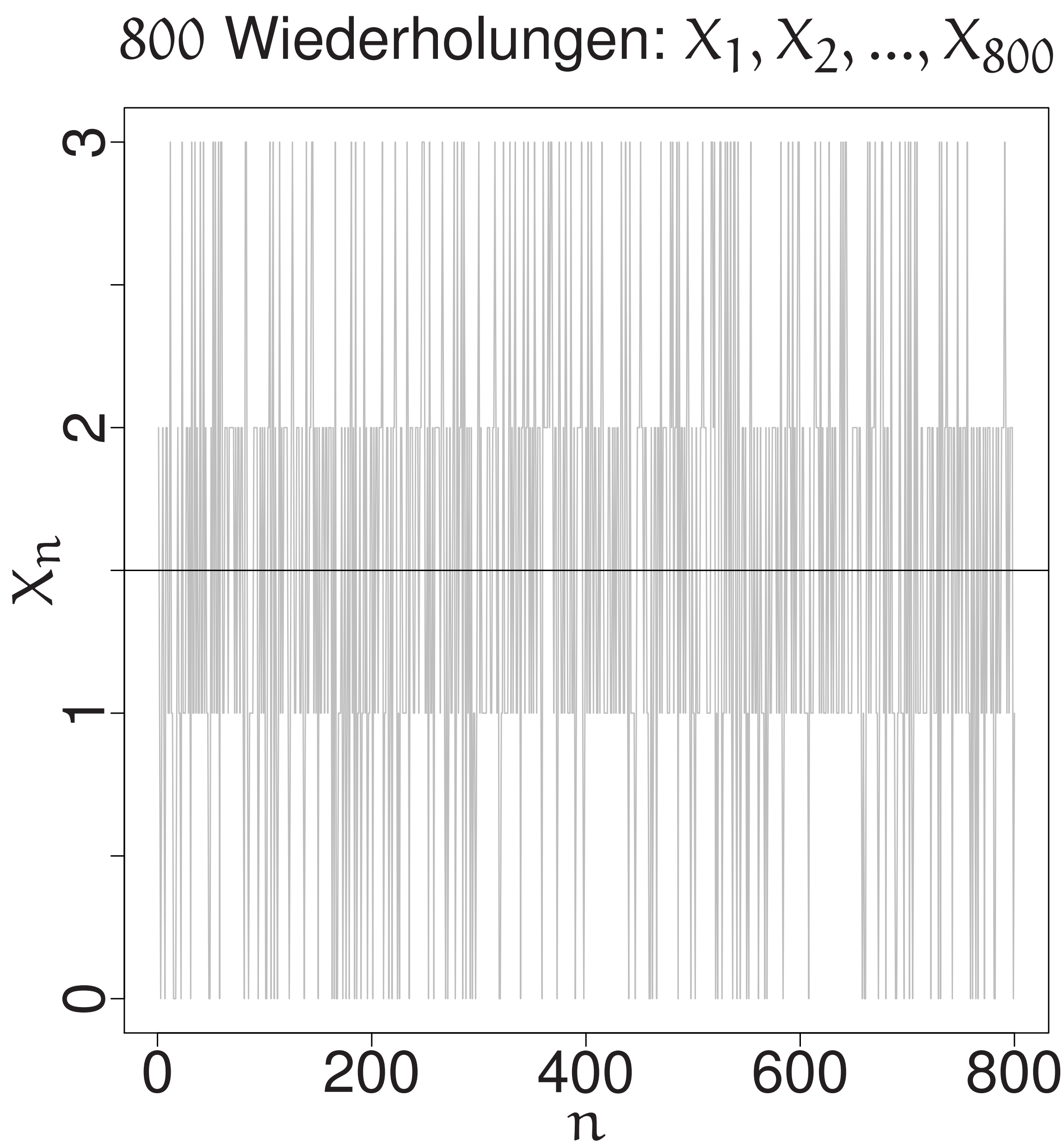
Beispiel:

X . . . Anzahl Köpfe beim dreimaligen fairen Münzwurf


80 Wiederholungen: $X_1, X_2, ..., X_{80}$
X_n
0
1
2
3
0
20
40
60
80
n

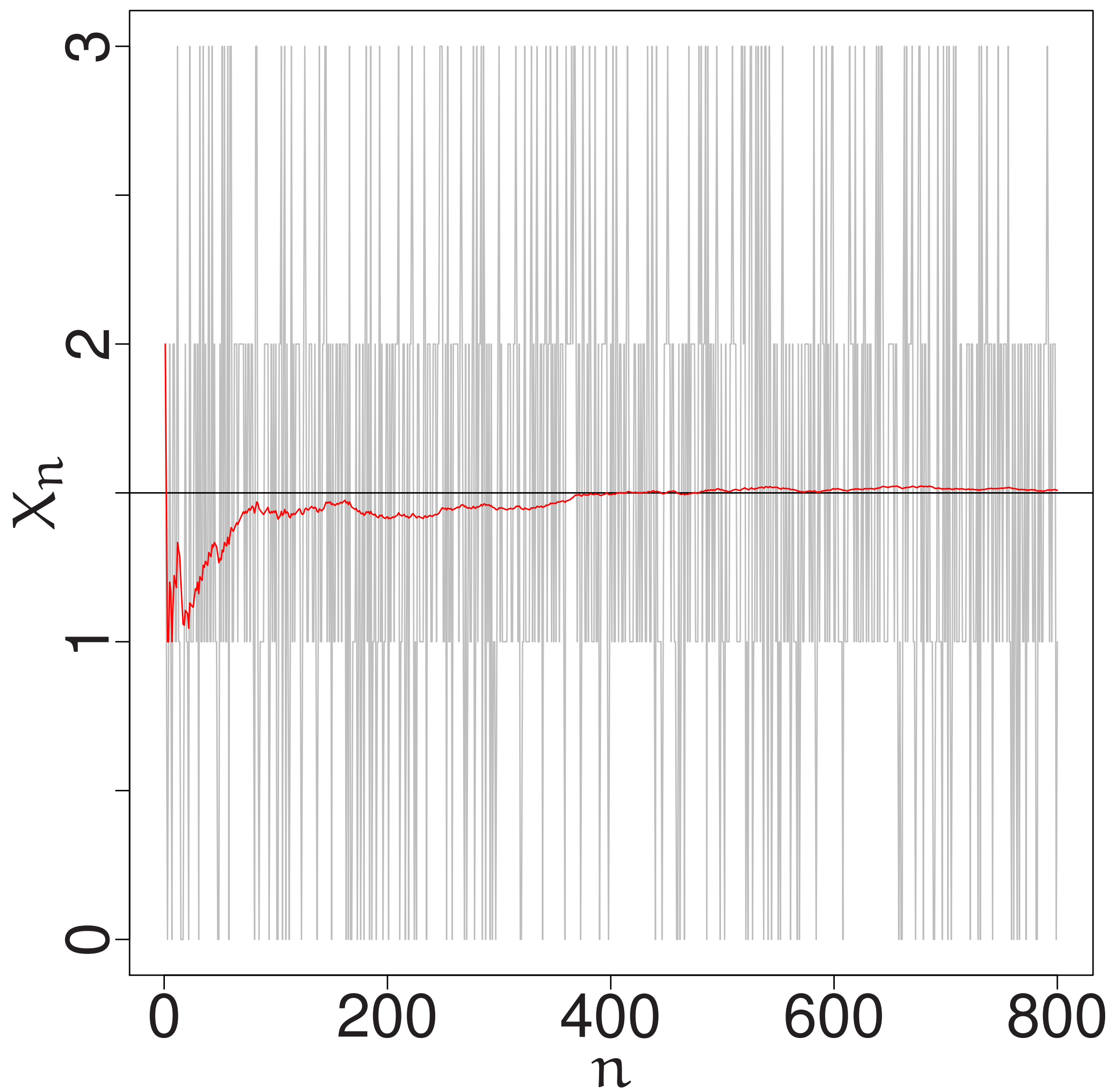
$M_n := (X_1 + X_2 + \ldots + X_n) / n$



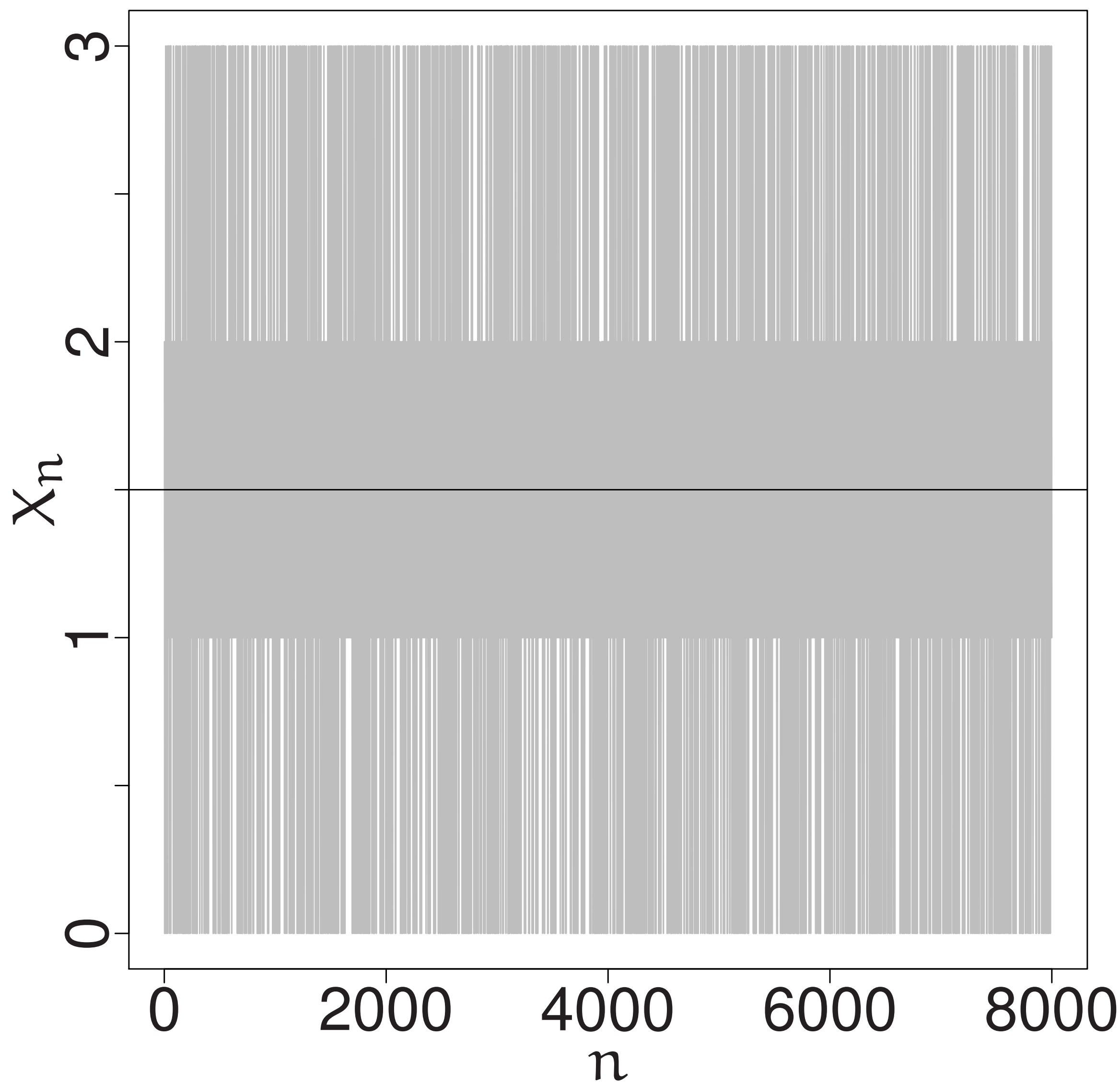


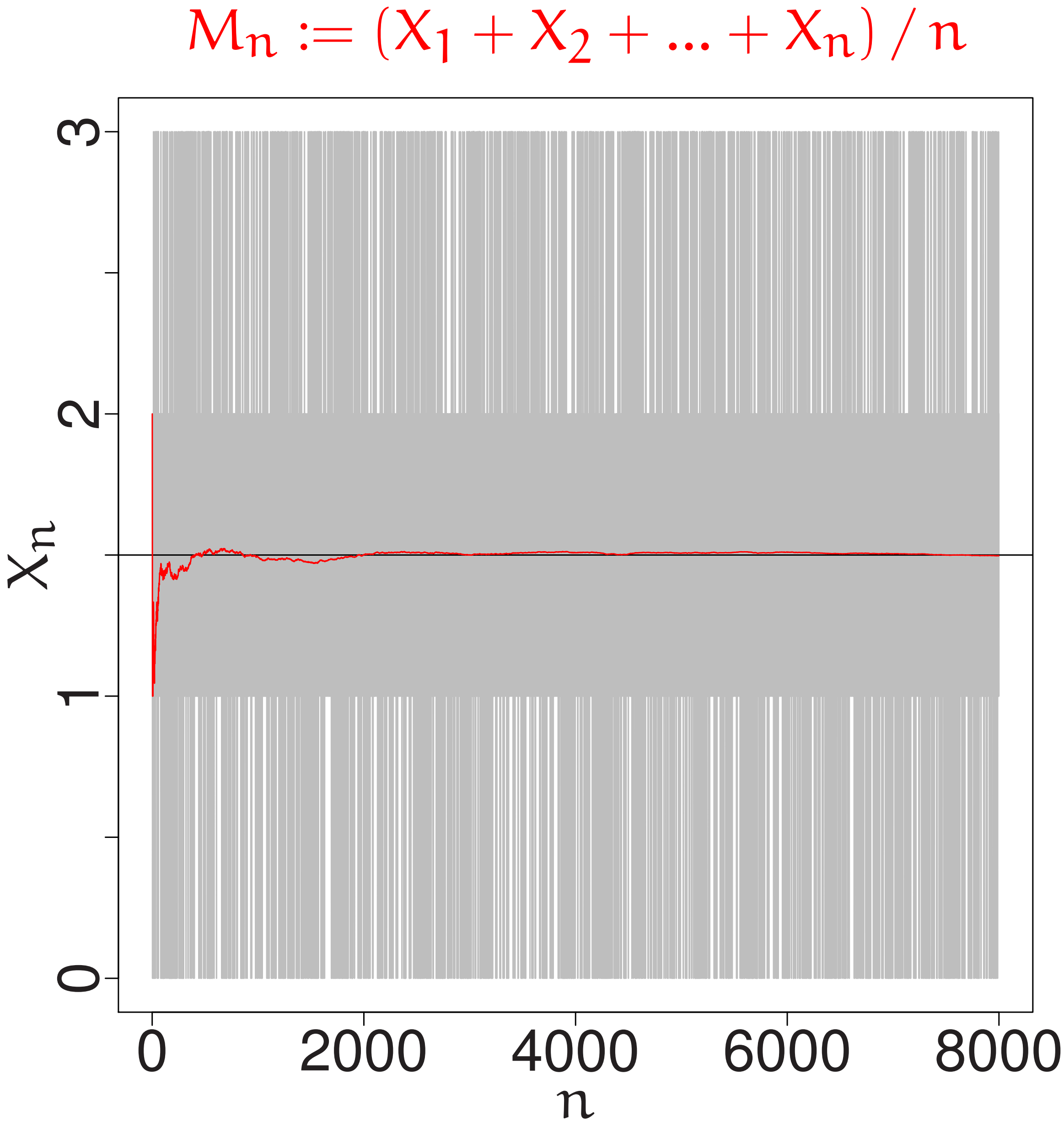
800 Wiederholungen: $X_1, X_2, ..., X_{800}$
3
2
1
0
X_n
0
200
400
600
800
n

$$M_n := (X_1 + X_2 + \ldots + X_n)\ /\ n$$

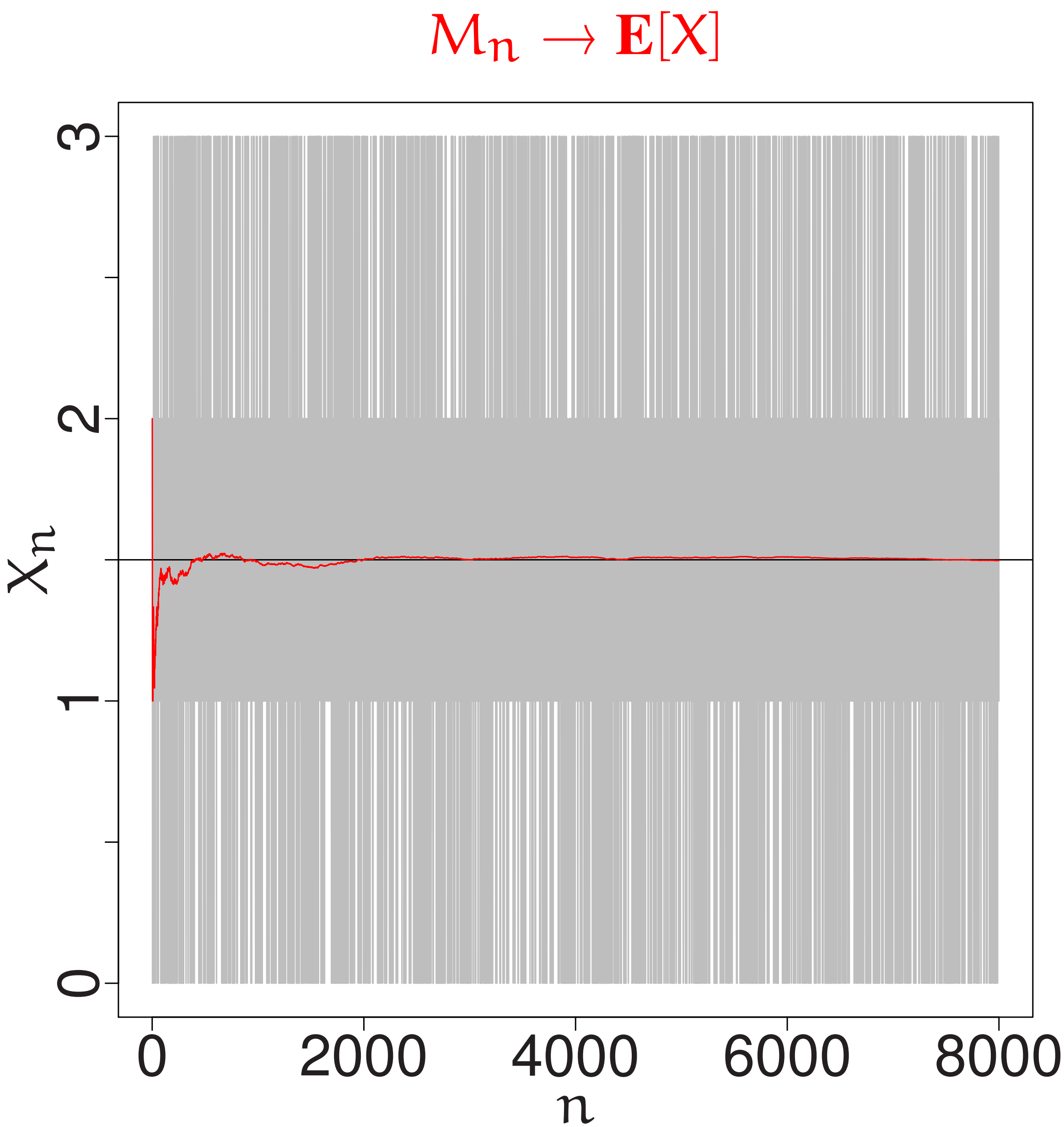


8000 Wiederholungen: $X_1, X_2, ..., X_{8000}$



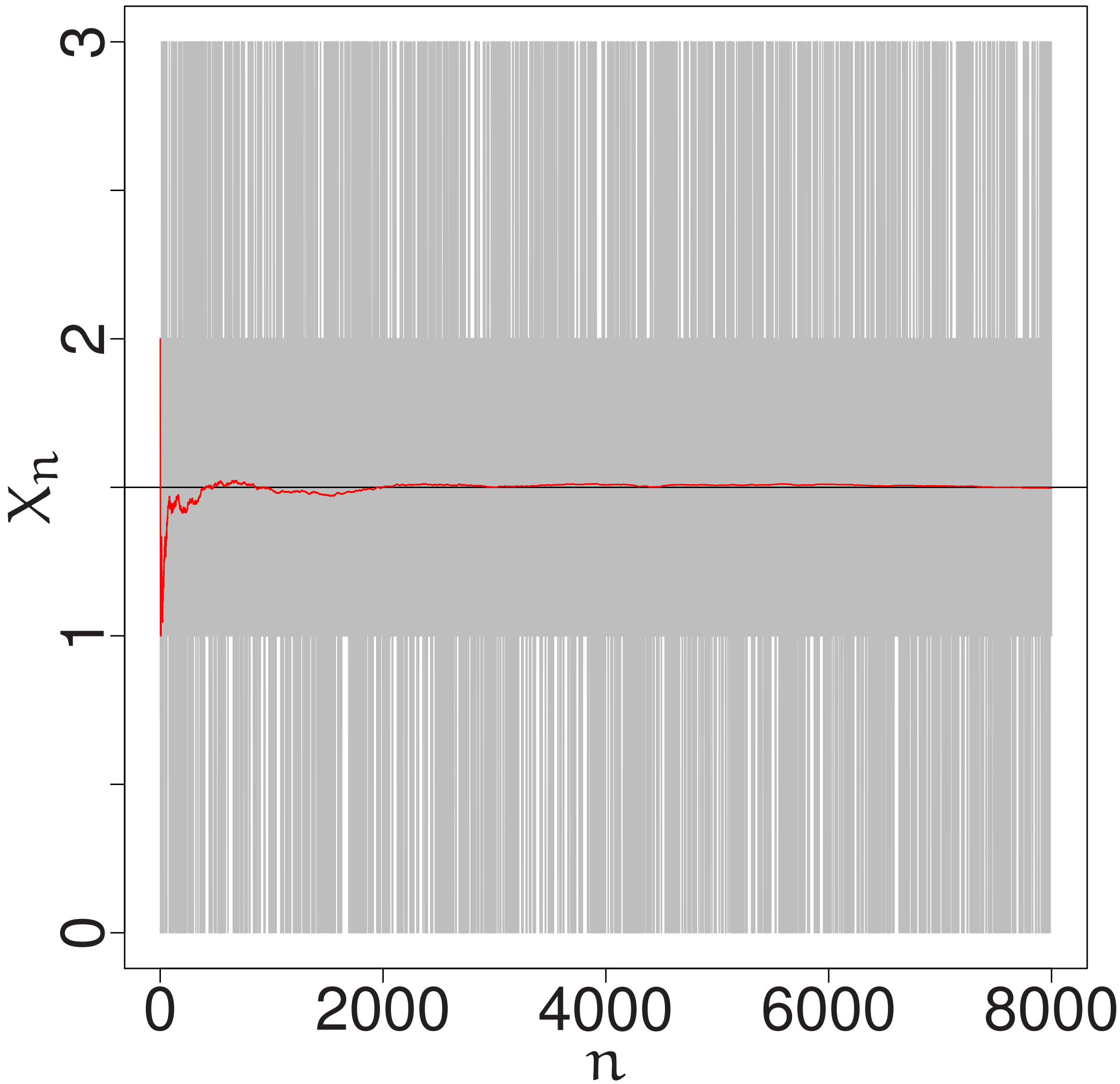


$M_n := (X_1 + X_2 + ... + X_n) / n$
X_n
0
1
2
3
0
2000
4000
6000
8000
n

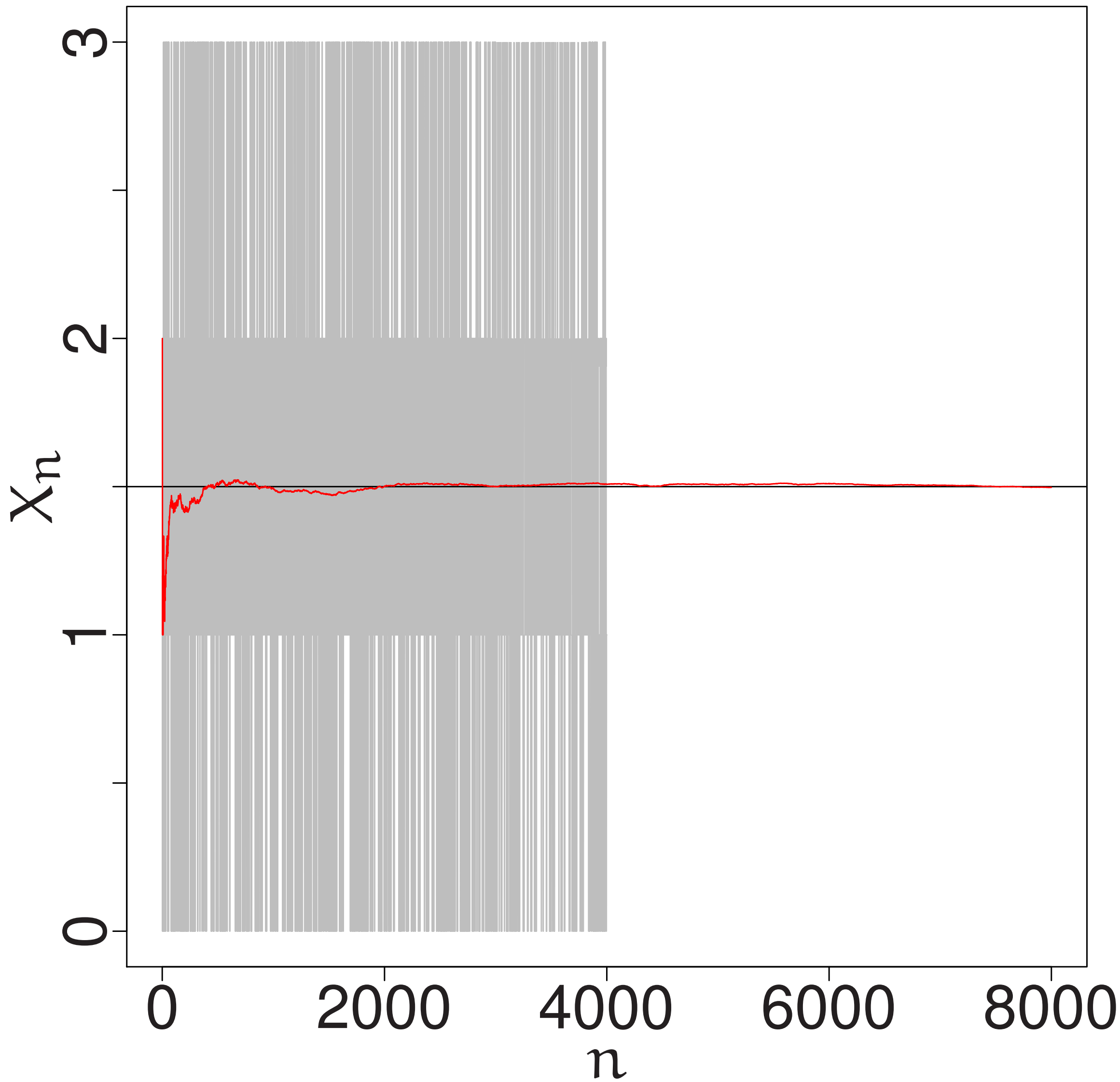


$M_n \to \mathbf{E}[X]$
3
2
1
0
X_n
0
2000
4000
6000
8000
n

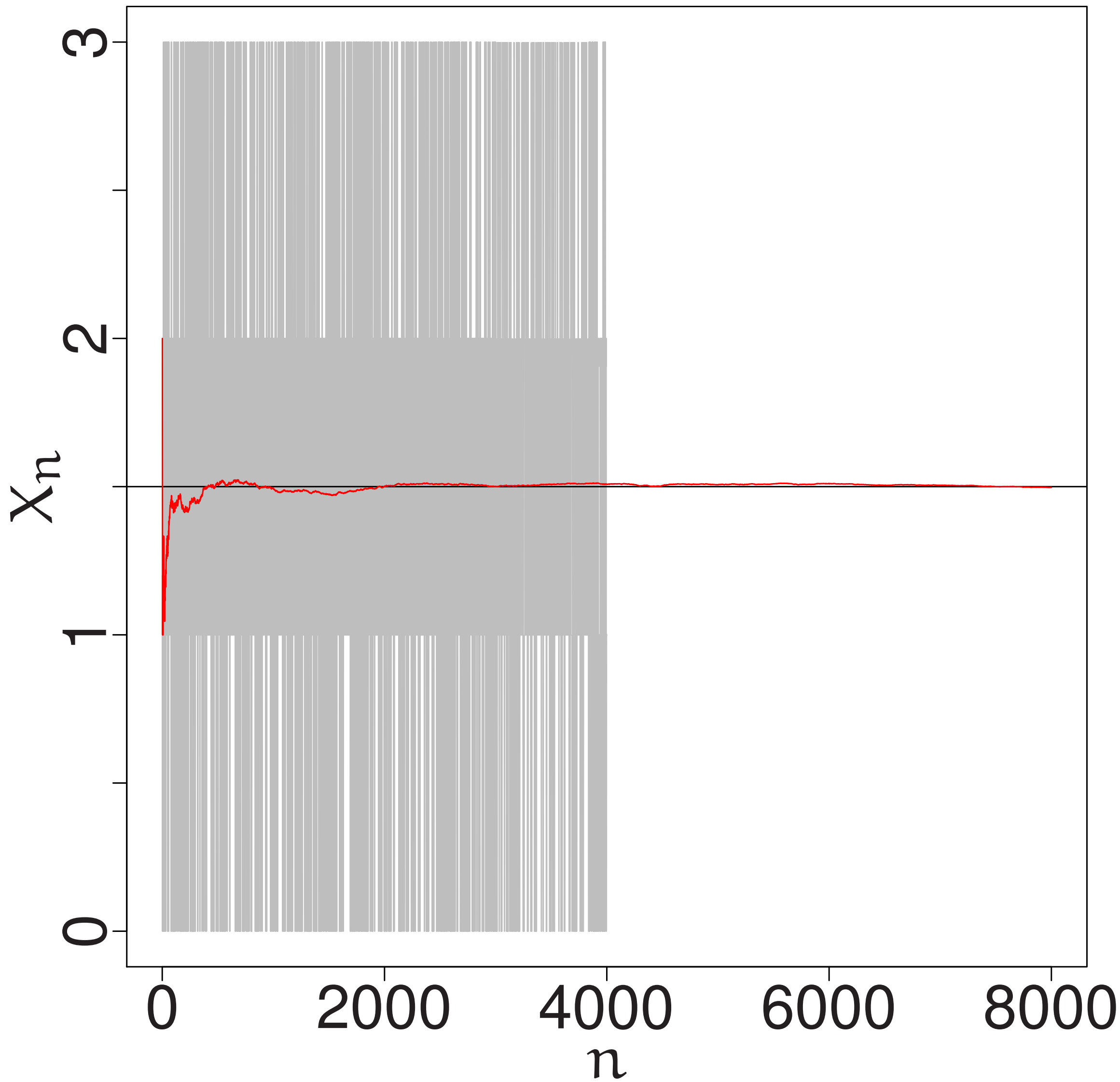
Warum?



$$M_n = (X_1 + X_2 + \ldots + X_n) \,/\, n$$



$$M_n = \sum_{a=0}^{3} a \, \#\{i \leq n : X_i = a\}/n$$
$$\to \sum a \, \mathbf{P}(X = a)$$



Dazu später mehr.

Für den Moment nur als kurzer Ausblick:

Das Gesetz der grossen Zahlen

Sei X eine Zufallsgröße mit Erwartungswert $\mathbf{E}[X]$.

Seien $X_1, X_2, \ldots$ unabhängige Kopien von X.

Dann gilt

$$\frac{X_1 + \ldots + X_n}{n} \to \mathbf{E}[X]$$

Zu klären

1. Was heißt „ unabhängig “?
2. Was heißt „ $\to$ “?

Diese Klärung wird in der Vorlesung

in wenigen Wochen erfolgen.

Jetzt halten wir erst einmal fest:

Zwei Vorstellungen von $\mathbf{E}[X]$

1. Gewichtetes Mittel

der möglichen Werte:

$$\mathbf{E}[X] := \sum a\,\mathbf{P}(X = a)$$

2. Langzeitmittelwert

bei unabhängigen Wiederholungen:

$$\frac{X_1 + \ldots + X_n}{n} \to \mathbf{E}[X]$$

4. Die Additivität des Erwartungswertes

- anschaulich

und als Werkzeug

Die wichtigste Eigenschaft des Erwartungswerts ist die

Additivität

Die wichtigste Eigenschaft des Erwartungswerts ist die

<u>Additivität</u>

$$\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$$

Die Additivität des Erwartungswerts wird

intuitiv sofort klar aus der Vorstellung als Langzeitmittelwert

bei “unabhängigen Wiederholungen”:

$$\frac{1}{n}((X_1 + Y_1) + ... + (X_n + Y_n))$$

$$= \frac{1}{n}(X_1 + ... + X_n) + \frac{1}{n}(Y_1 + ... + Y_n)$$

$$\to \mathbf{E}[X] + \mathbf{E}[Y]$$

Ein prominenter Fall ist

$$X = Z_1 + \cdots + Z_n,$$

wobei die $Z_1, \ldots, Z_n$ nur die Werte 0 oder 1 annehmen.

Dann gilt

$$\mathbf{E}[Z_i] = \mathbf{P}(Z_i = 1)$$

und somit

$$\mathbf{E}[X] = \mathbf{P}(Z_1 = 1) + \cdots + \mathbf{P}(Z_n = 1)\ .$$

5. Der Erwartungswert der Binomialverteilung

(als Erwartungswert der Anzahl der Erfolge

beim n-fachen p-Münzwurf)

X sei Bin(n, p) verteilt.

$$\mathbf{E}[X] = ?$$

$$\sum_{k=0}^{n} k\,\mathbf{P}(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \ ...$$

Es GEHT so (vgl Buch Seite 23-24)

Aber es geht auch einfacher (vgl. Buch S. 49):

Sei $Z = (Z_1, \ldots, Z_n)$ ein n-facher p-Münzwurf.

Dann ist $(Z_1 + \cdots + Z_n)$ Bin(n, p)-verteilt.

$$\mathbf{E}[Z_1 + \cdots + Z_n] = \mathbf{E}[Z_1] + \cdots + \mathbf{E}[Z_n]$$

$$\mathbf{E}[Z_i] = 1 \cdot p + 0 \cdot q = p$$

Fazit:

Der Erwartungswert einer Bin(n, p) verteilten ZV ist

np.

6. Der Erwartungswert der hypergeometrischen Verteilung

(als Erwartungswert der Anzahl der “Erfolge”

beim n-fachen Ziehen ohne Zurücklegen)

BEISPIEL

Ziehen ohne Zurücklegen

ause Eine Urne enthält r rote und b blaue Kugeln.

ooooooooooooo $\quad r = 8 \quad b = 5$

Aus der Urne werden ohne Zurücklegen n Kugeln gezogen.

ooooooooo $\quad n = 9$

$R :=$ Anzahl der gezogenen roten Kugeln

$$\mathbf{E}[R] = ?$$

$$R = Z_1 + Z_2 + \ldots + Z_n$$

$Z_i = 1$ falls i-te gezogogene Kugel rot

$Z_i = 0$ falls i-te gezogene Kugel blau

OOOOOOOOOOOOO $\qquad r = 8 \quad b = 5$

$$\mathbf{P}(Z_i = 1) = \frac{r}{r+b}$$

Man stelle sich vor, die Nummern der Züge werden

als rein zufällige Permutation an die $r + b$ Kugeln vergeben.

Wie wahrscheinlich ist es,

dass Nummer i auf eine rote Kugel fällt?

$$R = Z_1 + Z_2 + \ldots + Z_n$$

$Z_i = 1$ falls i-te gezogogene Kugel rot

$Z_i = 0$ falls i-te gezogene Kugel blau

OOOOOOOOOOOOO $\qquad r = 8 \quad b = 5$

$$\mathbf{P}(Z_i = 1) = \frac{r}{r+b}$$

$$\mathbf{E}[Z_i] = \tfrac{r}{r+b}$$

$$\mathbf{E}[R] = \mathbf{E}[Z_1] + \mathbf{E}[Z_2] + \ldots + \mathbf{E}[Z_n]$$

$$\boxed{\mathbf{E}[R] = n\frac{r}{r+b}}$$

BEISPIEL

Ziehen ohne Zurücklegen

vspace0.6cm Eine Urne enthält r rote und b blaue Kugeln.

$\color{red}{oooooooo}\color{blue}{ooooo}$ $\qquad$ $\color{red}{r = 8}$ $\quad$ $\color{blue}{b = 5}$

Aus der Urne werden ohne Zurücklegen n Kugeln gezogen.

$\color{red}{oo}\color{blue}{o}\color{red}{oooo}\color{blue}{o}\color{red}{o}$ $\qquad$ $n = 9$

$R :=$ Anzahl der gezogenen roten Kugeln

Verteilung von R?

Verteilung von R ?

$$\mathbf{P}(R = k) = ?$$

vspace1cm $\mathbf{P}(R = k) = \binom{r}{k} \binom{b}{n-k} / \binom{r+b}{n}$

Eine ZV mit diesen Verteilungsgewichten ($k = 0, \ldots, n$)

heißt

hypergeometrisch verteilt zu den Parametern $(n,\ r+b,\ r)$.

(vg. Buch Seite 28)

$$\mathbf{P}(R = k) = \binom{r}{k}\binom{b}{n-k} \Big/ \binom{r+b}{n}$$

$$\mathbf{E}(R) = ?$$

$$\mathbf{E}[R] = \sum_{k=0}^{n} k \binom{r}{k}\binom{b}{n-k} \Big/ \binom{r+b}{n} = ...$$

Es GEHT so (vgl. Buch Seite 32)

Aber wie wir eben gesehen haben,

(über die Darstellung von R als Summe von Zählern)

geht's auch einfacher (vgl. Buch S. 50/51).

7. Der Erwartungswert einer Anzahl von Runs

Runs beim fairen Münzwurf

$Z := (Z_1, Z_2, ..., Z_n)$ n-facher fairer Münzwurf

$$P\{Z_i = 1\} = \frac{1}{2} \qquad P\{Z_i = 0\} = \frac{1}{2}$$

Run: ein Block von Nullen (Einsen),

der nicht echt in einem größeren Block enthalten ist

$R :=$ Anzahl Runs in Z

$$00000000 \qquad R = 1$$

$$11100011 \qquad R = 3$$

$$10101010 \qquad R = 8$$

$$\mathbf{E}[R] = ?$$

Dazu schreiben wir R als Summe von Zählern.

Bei jedem Wurf zählen wir eins dazu,

wenn bei diesem Wurf ein Run beginnt:

$Y_i := 1$ falls bei i ein Run beginnt, $\quad Y_i := 0$ sonst

$$R = Y_1 + Y_2 + \ldots + Y_n$$

$$Y_1 \equiv 1$$

$$\{Y_i = 1\} = \{(Z_{i-1}, Z_i) = (0,1) \text{ oder } (1,0)\} \qquad (i > 1)$$

$$P(Y_i = 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \qquad (i > 1)$$

$$\mathbf{E}[Y_i] = \frac{1}{2} \qquad (i > 1)$$

$$\mathbf{E}[R] = \mathbf{E}[Y_1] + \mathbf{E}[Y_2] + \mathbf{E}[Y_3] + \ldots + \mathbf{E}[Y_n]$$

$$\boxed{\mathbf{E}[R] = 1 + \frac{1}{2}(n-1)}$$

8. Zur Wohldefiniertheit des Erwartungswertes

Wie kann es sein, dass für eine

diskrete reellwertige Zufallsvariable X

mit $\mathbf{P}(X \in S)$, S abzählbar,

die Summe $\sum_{a \in S} a\mathbf{P}(X = a)$ nicht wohldefiniert ist?

Ein Beispiel: $\mathbf{P}(X = (-2)^n) := 2^{-n}$, $n = 1, 2, \ldots$

Dann ist $$\sum_{n\in\{1,3,\ldots\}} -2^n\mathbf{P}(X = -2^n) = -\infty$$

und $$\sum_{n\in\{2,4,\ldots\}} 2^n\mathbf{P}(X = 2^n) = +\infty.$$

Aber die Summe von $-\infty$ und $+\infty$ gibt keinen Sinn!

Wenn wir sagen

Die diskrete reellwertige Zufallsvariable X

hat einen wohldefinierten Erwartungswert

meinen wir, dass nicht zugleich

$$\sum_{a \in S, a > 0} a \mathbf{P}(X = a) \text{ und } \sum_{a \in S, a < 0} |a| \, \mathbf{P}(X = a)$$

Unendlich sein dürfen.

Wenden wir uns nun

der Herleitung der Linearitätseigenschaft

aus der Definition des Erwartungswertes zu.

9. Transformationsformel für den Erwartungswert

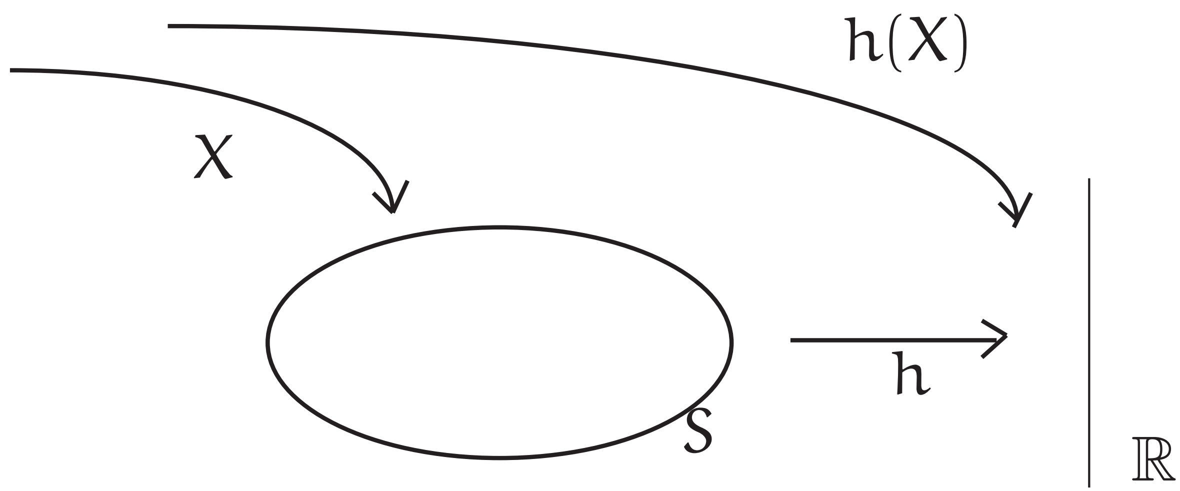
Diese Formel ist oft hilfreich
bei der Berechnung von Erwartungswerten.

Sie erinnert an die Einsetzungsregel(Substitutionsregel)
zum Berechnen von Summen und Integralen,

und wird uns im Abschnitt 10 helfen,
die Linearität des EW aus seiner Definition herzuleiten.

Sei X diskrete Zufallsvariable mit $\mathbf{P}(X \in S) = 1$

und h eine Abbildung von S nach $\mathbb{R}$



Sei X diskrete Zufallsvariable mit $\mathbf{P}(X \in S) = 1$

und h eine Abbildung von S nach $\mathbb{R}$

so, dass der Erwartungswert der Zufallsvariablen $h(X)$

wohldefiniert ist. Dann ist

$$\mathbf{E}[h(X)] = \sum_{a \in S} h(a)\, \mathbf{P}(X = a)\ .$$

Die Idee ist einfach: anstatt die Werte $b = h(a)$, $a \in S$,

mit deren Gewichten zu mitteln,

“zerlegt man nach dem Urbild”

und mittelt mit den Gewichten der Werte a.

$$\mathbf{E}[h(X)] = \sum_{a \in S} h(a)\, \mathbf{P}(X = a)$$

Denn: $\displaystyle\sum_{b \in h(S)} b\, \mathbf{P}(h(X) = b)$

$$= \sum_{b \in h(S)} b \sum_{a \in h^{-1}(b)} \mathbf{P}(X = a)$$

$$= \sum_{b \in h(S)} \sum_{a \in h^{-1}(b)} h(a)\, \mathbf{P}(X = a)$$

$$= \sum_{a \in S} h(a)\, \mathbf{P}(X = a)\,. \quad \square$$

10. Die Linearität des Erwartungswertes

- Beweis

Wir betrachten

zwei diskrete reellwertige Zufallsvariable X_1, X_2,

die gemeinsam in einem Zufallsexperiment auftreten

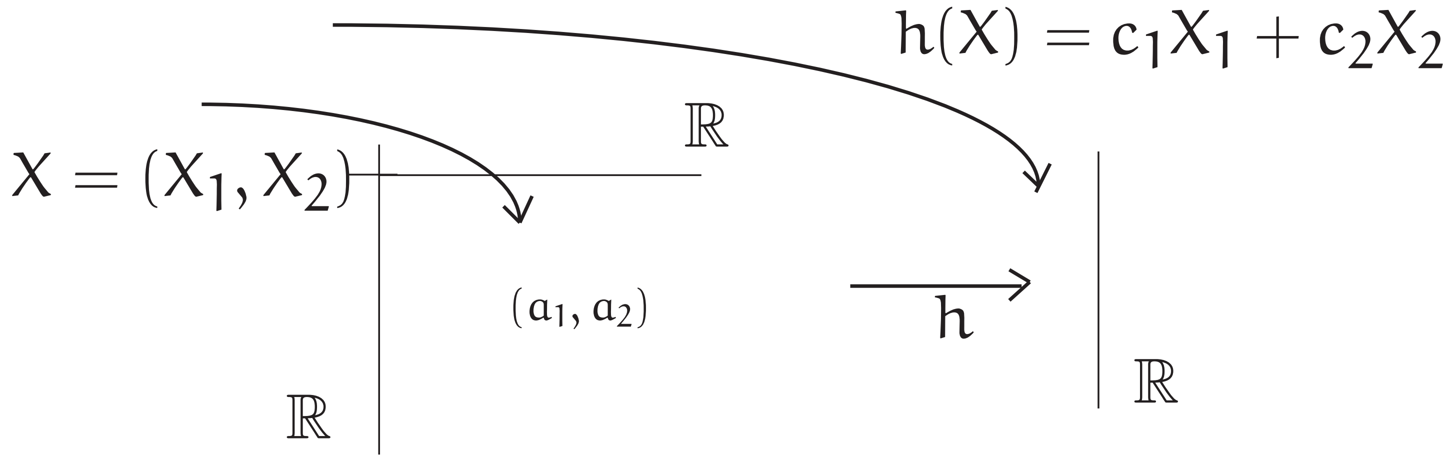
und sich damit zu einem zufälligen Paar (X_1, X_2)

zusammenfassen lassen.

Für $c_1,\ c_2 \in \mathbb{R}$ ist dann auch

$$c_1 X_1 + c_2 X_2$$

eine diskrete reellwertige Zufallvariable.



$$h(a_1, a_2) = c_1a_1 + c_2a_2$$

Satz [**Linearität des Erwartungswertes**]

(Buch S. 52) Für reellwertige Zufallsvariable X_1, X_2

mit wohldefiniertem Erwartungswert gilt

$$\mathbf{E}[c_1X_1 + c_2X_2] = c_1\mathbf{E}[X_1] + c_2\mathbf{E}[X_2]\ , \quad c_1, c_2 \in \mathbb{R}\ .$$

Den Beweis führen wir hier nur für *diskrete* Zufallsvariable,
und zwar über die Transformationsformel mit $h(a_1, a_2) := c_1a_1 + c_2a_2$.

Beweis.

Seien $S_1, S_2 \subset \mathbb{R}$ abzählbar mit

$\mathbf{P}(X_1 \in S_1) = \mathbf{P}(X_2 \in S_2) = 1$.

Aus der Transformationsformel folgt mit

$h(a_1, a_2) := c_1 a_1 + c_2 a_2$:

$$\mathbf{E}[c_1X_1 + c_2X_2]$$

$$= \sum_{a_1\in S_1} \sum_{a_2\in S_2} (c_1a_1 + c_2a_2)\,\mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= c_1 \sum_{a_1\in S_1} a_1 \sum_{a_2\in S_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$+ c_2 \sum_{a_2\in S_2} a_2 \sum_{a_1\in S_1} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$\sum_{a_1 \in S_1} a_1 \sum_{a_2 \in S_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} a_1 \mathbf{P}(X_1 = a_1)$$

$$= \mathbf{E}[X_1]$$

$$\mathbf{E}[c_1X_1 + c_2X_2]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} (c_1a_1 + c_2a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= c_1 \sum_{a_1 \in S_1} a_1 \sum_{a_2 \in S_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$+ c_2 \sum_{a_2 \in S_2} a_2 \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$\mathbf{E}[c_1 X_1 + c_2 X_2]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} (c_1 a_1 + c_2 a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= c_1 \mathbf{E}[X_1]$$

$$+ c_2 \sum_{a_2 \in S_2} a_2 \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$\mathbf{E}[c_1 X_1 + c_2 X_2]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} (c_1 a_1 + c_2 a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= c_1 \mathbf{E}[X_1]$$

$$+ c_2 \mathbf{E}[X_2]$$

$$\square$$

Zusammenfassung

des Wichtigsten

A.

Was ist der Erwartungswert?

$$\mathbf{E}[X] = \sum a \, \mathbf{P}(X = a)$$

<u>und</u>

$$\mathbf{E}[X] = \lim \frac{X_1 + ... + X_n}{n}$$

für "unabhängige Wiederholungen" $X_1, X_2, \ldots$

B.

Was ist die wichtigste Eigenschaft des Erwartungswertes?

Die Linearität:

$$\mathbf{E}[\alpha X + \beta Y] = \alpha \mathbf{E}[X] + \beta \mathbf{E}[Y]$$

C.

Wie berechnet man $\mathbf{E}[X]$ am besten?

Oft dadurch,

dass man X als Summe schreibt:

$$X = Z_1 + ... + Z_n$$

$$\mathbf{E}[X] = \mathbf{E}[Z_1] + ... + \mathbf{E}[Z_n]$$