Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

0. Diskrete reellwertige Zufallsvariable

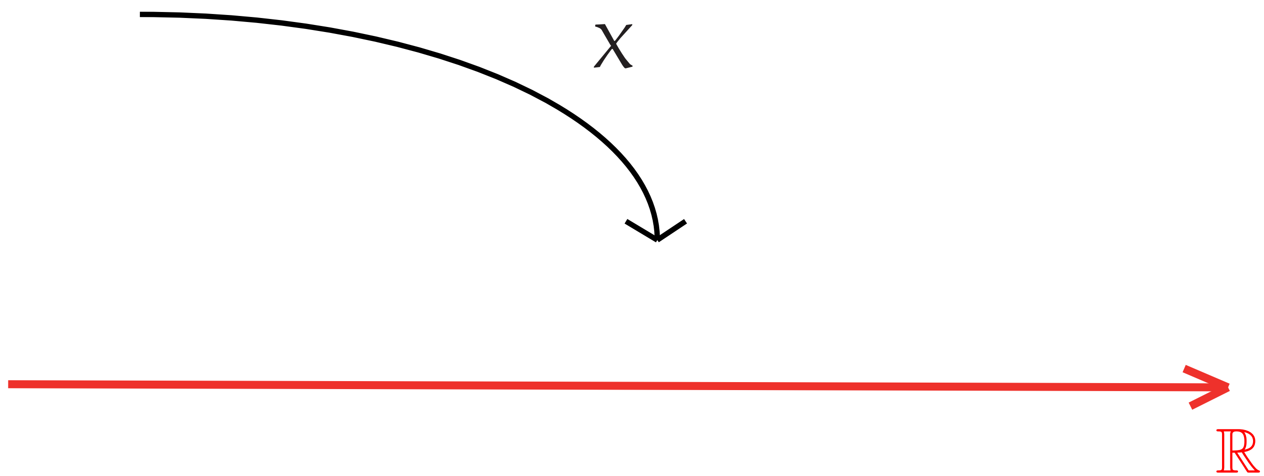
(Zur Erinnerung an VI 2b)

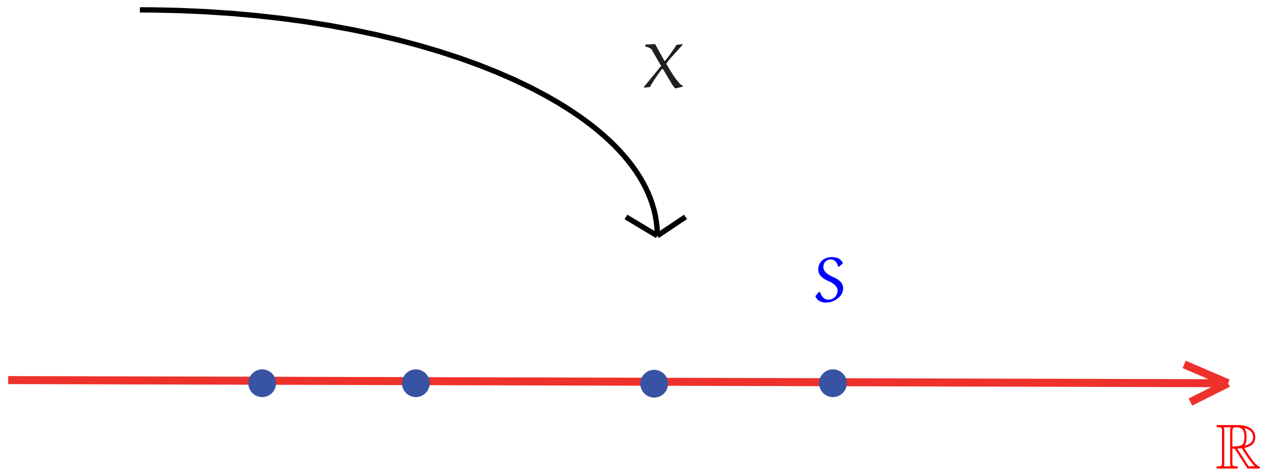
X sei eine Zufallsvariable, deren Zielbereich

$\mathbb{R}$ (die Menge der reellen Zahlen)

oder eine Teilmenge von $\mathbb{R}$

ist.

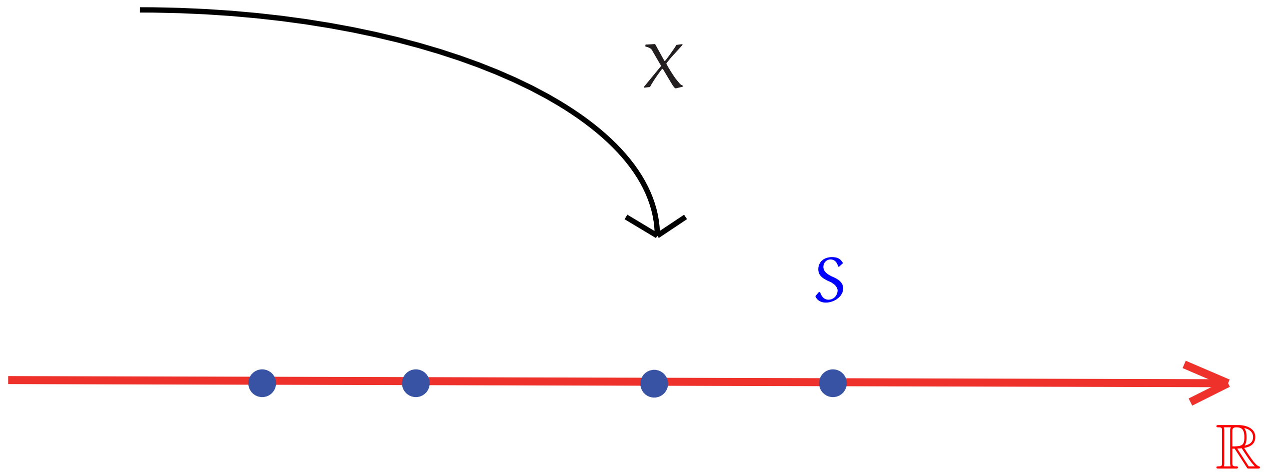




Außerdem existiere eine abzählbare* Menge $S \subset \mathbb{R}$ mit

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

*d.h. endliche oder abzählbar unendliche



Wir sagen dann:

X ist eine diskrete reellwertige Zufallsvariable

1. Der Erwartungswert als gewichtetes Mittel

Eine einprägsame Kenngröße

für die *Lage* der Verteilung von X

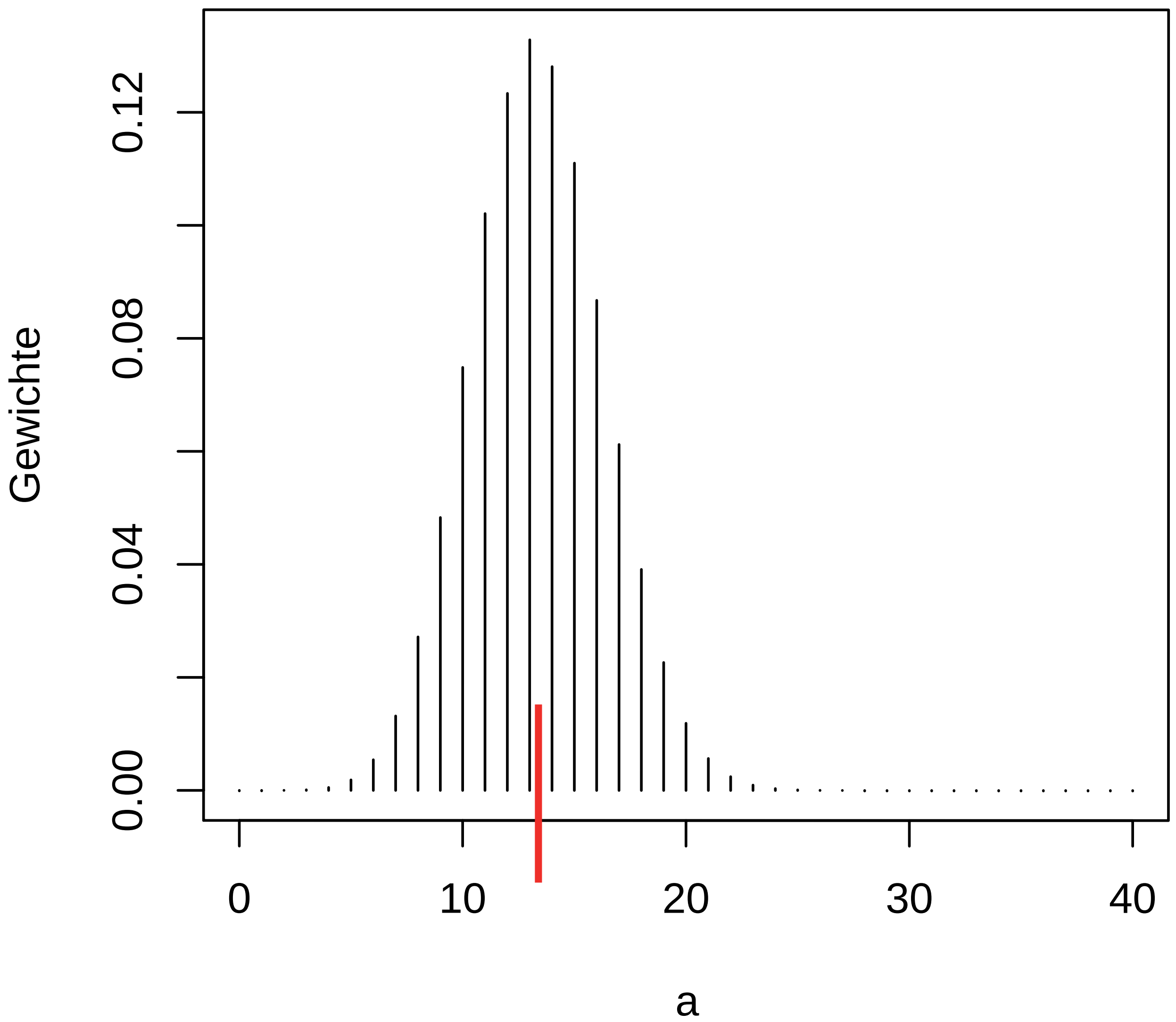
ist das mit den Wahrscheinlichkeiten gewichtete Mittel

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:

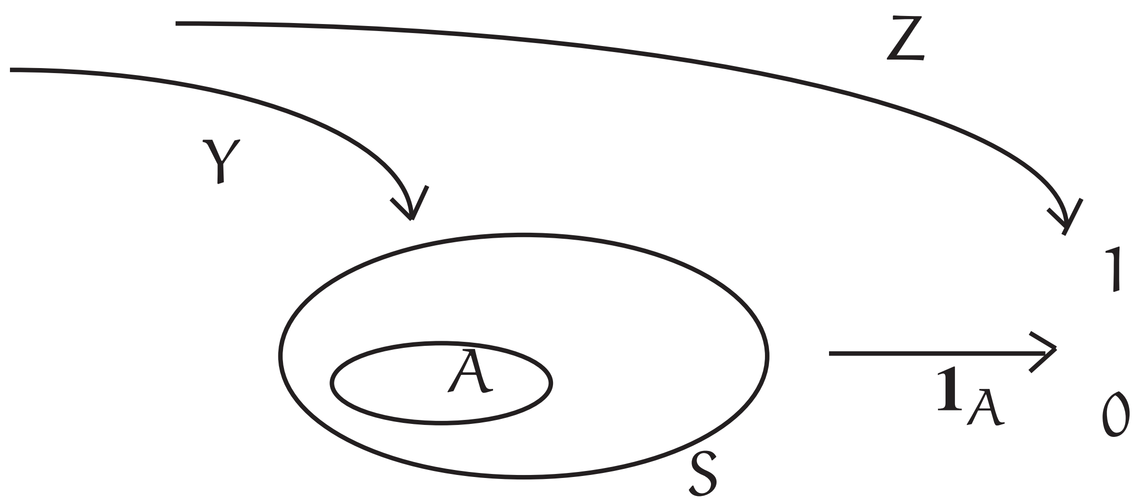
$$Z = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } p \\ 0 & \text{mit Wahrscheinlichkeit } q = 1 - p \end{cases}$$

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



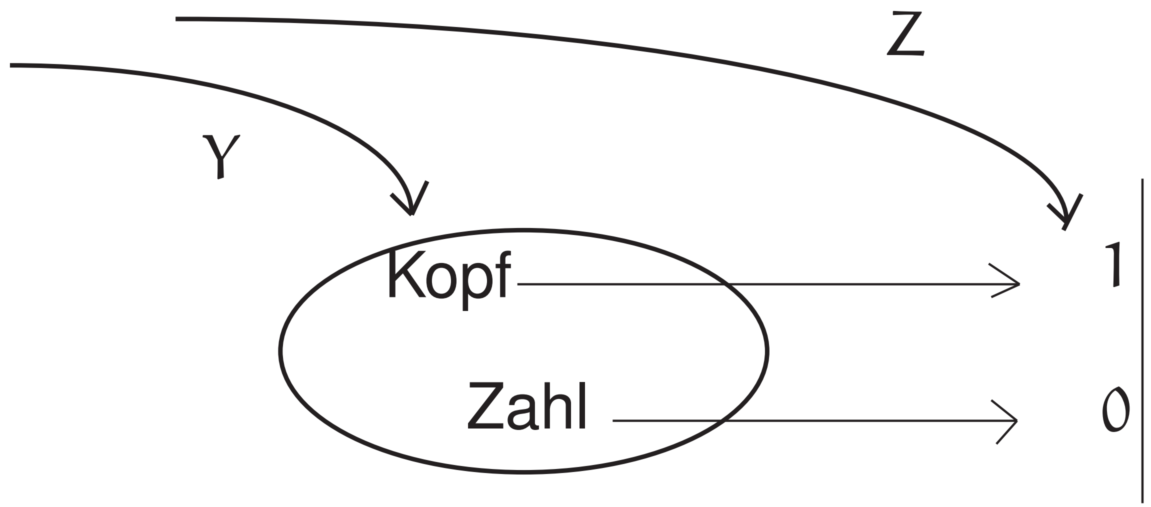
$$\{Y \in A\} = \{Z = 1\}$$

Man sagt auch:

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

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

... und entspricht dem Szenario des einfachen Münzwurfs:



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

Z ist die *Indikatorvariable* des Ereignisses $\{Y = \text{Kopf}\}$

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

Für allgemeines diskretes, reellwertiges X hatten wir

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

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

mit $\rho(a)$:= Verteilungsgewichte von X

Wohlgemerkt:

Der Erwartungswert der Zufallvariablen X

hängt nur von deren Verteilung ρ ab.

Synonym sprechen wir daher auch manchmal vom

Erwartungswert der Verteilung ρ .

X

eine Zufallsgröße;

$\mathbf{E}[X]$

eine Zahl.

2. 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: $\mathbf{P}(X = (-2)^n) := 2^{-n},\ n = 1, 2, \dots$

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

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

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

Definition: Wir sagen

Die diskrete reellwertige Zufallsvariable X

hat einen wohldefinierten Erwartungswert

wenn die beiden Summen

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

nicht beide zugleich den Wert $+\infty$ annehmen.

Anders gesagt:

Damit die Summe $\sum_{a \in S} a\,\mathbf{P}(X = a)$

in $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ exisitiert, muss gelten:

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

oder

$$\sum_{a \in S,\, a < 0} a\,\mathbf{P}(X = a) > -\infty$$

∞ ist als Summenwert erlaubt, $-\infty$ auch.

Aber $\infty - \infty$ gibt keinen Sinn.

Beispiele:

1. $\mathbf{P}(X = 2^j) = 2^{-j}, \qquad j = 1, 2, \ldots$:

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

2. $\mathbf{P}\left(X = (-2)^j\right) = 2^{-j}, \qquad j = 1, 2, \ldots$:

$\mathbf{E}[X]$ exisitiert nicht.

3. Transformationsformel für den Erwartungswert

Satz: Sei X diskrete Zufallsvariable mit $\mathbf{P}(X \in S) = 1$ und h eine Abbildung von S nach $\mathbb{R}$ (wobei der Erwartungswert der Zufallsvariablen $h(X)$ wohldefiniert sei). Dann ist

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

Die Idee ist einfach: anstatt mit den Gewichten der Werte $b = h(a)$, $a \in S$ zu mitteln, zerlegt man nach den Elementen $a \in h^{-1}(b)$ und mittelt mit deren Gewichten.

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

Beweis:

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

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

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

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

4. Die Linearität des Erwartungswertes

Wir betrachten

zwei diskrete reellwertige Zufallsvariable X_1, X_2,

zusammengefasst zu einem zufälligen Paar (X_1, X_2).

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

$$c_1 X_1 + c_2 X_2$$

eine diskrete reellwertige Zufallvariable.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\square$$

Warum ist die Additivität des Erwartungwerts so wichtig?

Oft lassen sich Zufallsvariable als Summen von einfacheren Bausteinen (z.B. Zählvariablen) darstellen.

Mittels der Additivität wird die Berechnung des Erwartungswertes dann einfach.

Dies illustrieren wir an ein paar Beispielen.

5. Der Erwartungswert der Binomialverteilung

X sei Bin(n, p) verteilt.

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

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

Es GEHT so (vgl Buch Seite 23-24)

Aber es geht auch einfacher:

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

Eine Urne enthält r rote und b blaue Kugeln.

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

Aus der Urne werden **ohne Zurücklegen** n Kugeln gezogen.

ooooooooo $\qquad n = 9$

$R :=$ Anzahl der gezogenen roten Kugeln

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

Verteilung von R ?

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

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

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

heißt

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

(vg. Buch Seite 28)

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

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

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

Es GEHT so (vgl. Buch Seite 32)

Aber es geht auch einfacher.

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

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

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

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

$$\mathbf{P}(Z_i = 1) = ?$$

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

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

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?

$$R = Z_1 + Z_2 + ... + Z_n$$

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

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

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

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

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

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

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

7. Der Erwartungswert einer Anzahl von Runs

beim fairen Münzwurf

$Z := (Z_1, Z_2, \ldots, Z_n)$ 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

$00000000 \qquad R = 1$

$11100011 \qquad R = 3$

$10101010 \qquad R = 8$

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

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

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

$$Y_1 \equiv 1$$

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

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

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

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

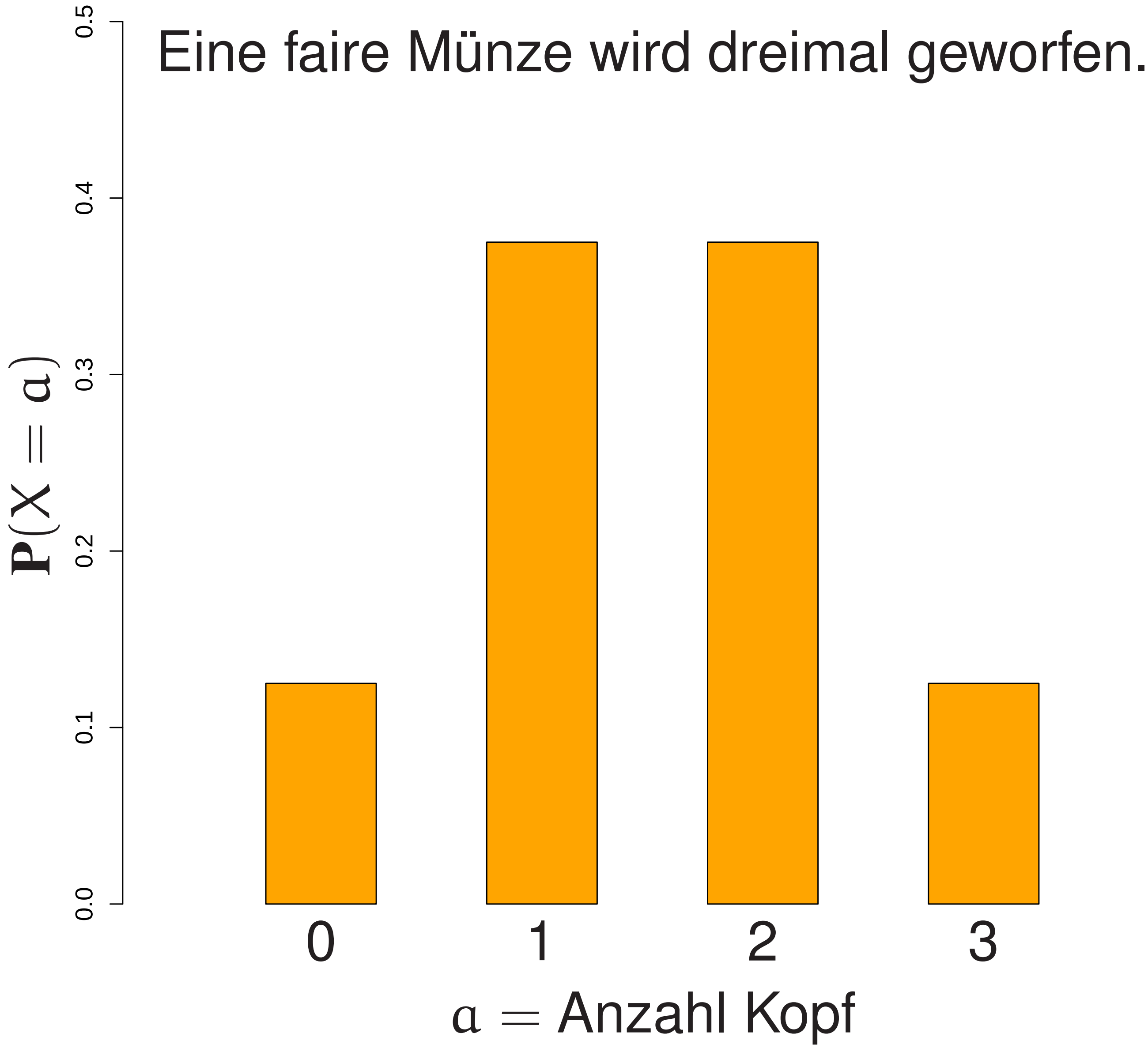
$$\boxed{\mathbf{E}[R] = 1 + \frac{1}{2}(n-1)}$$

8. Wie erlebt man den Erwartungswert?

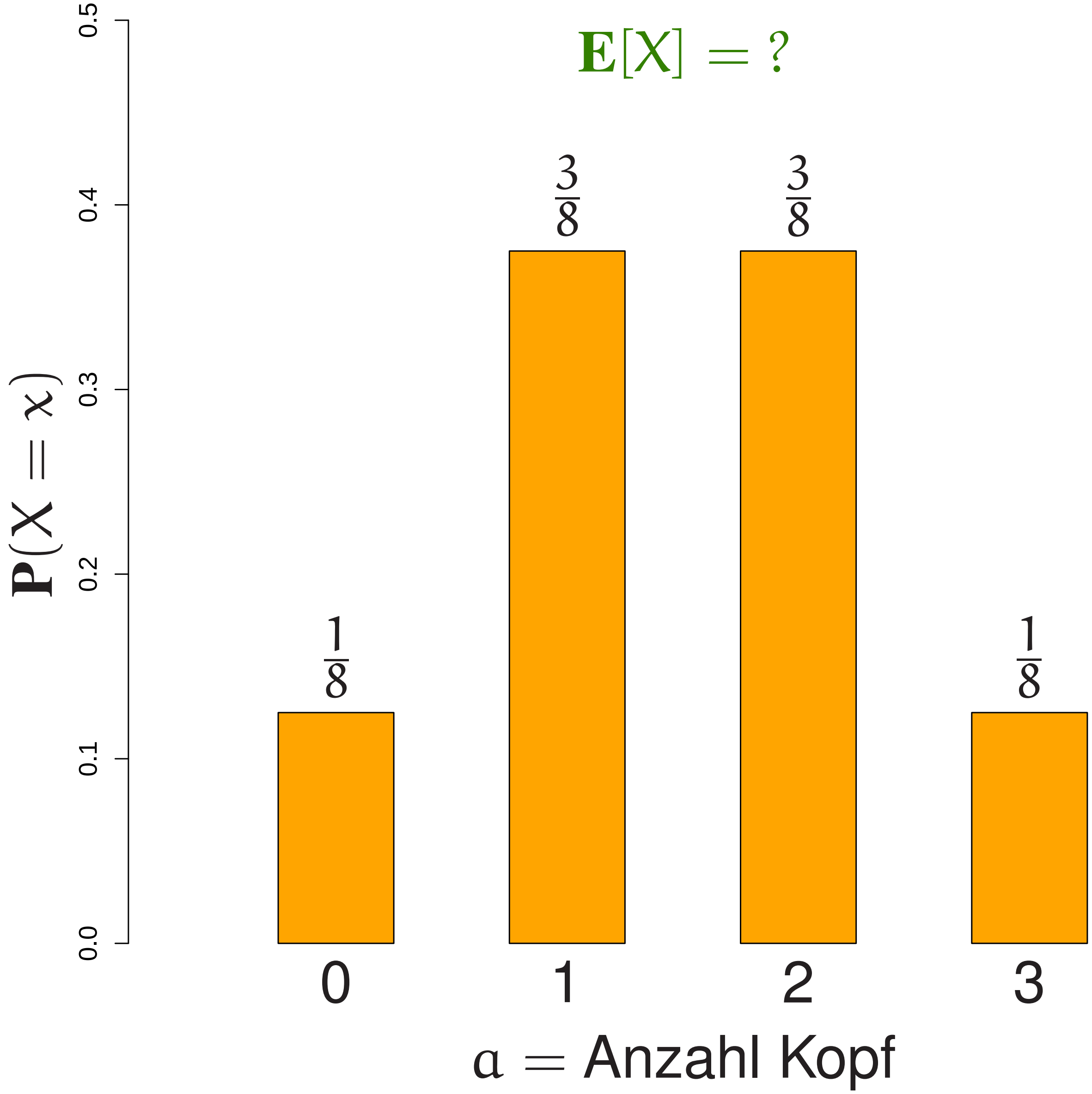
Beispiel:

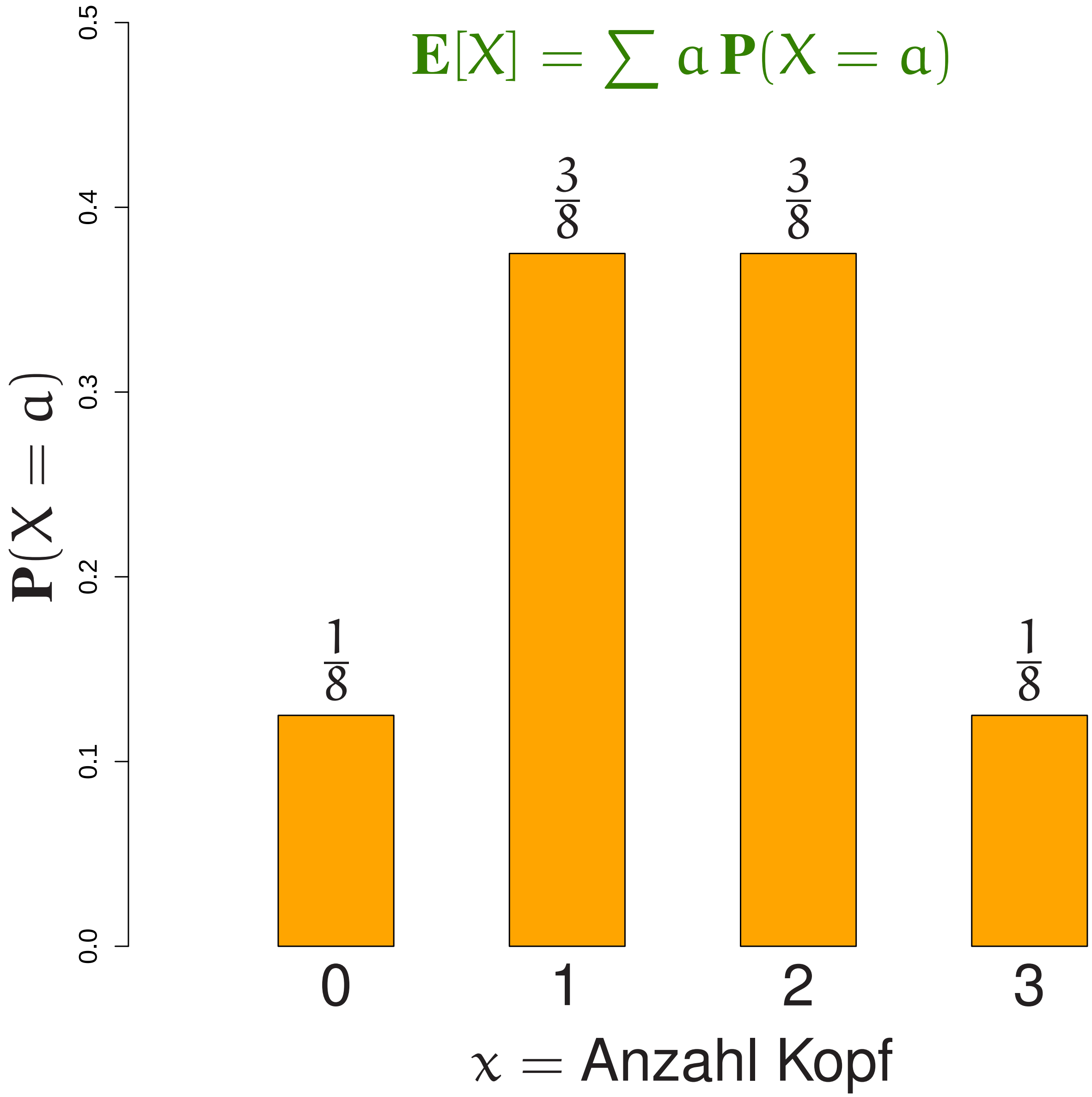
Eine faire Münze wird dreimal geworfen.

$X :=$ Anzahl der geworfenen Köpfe.

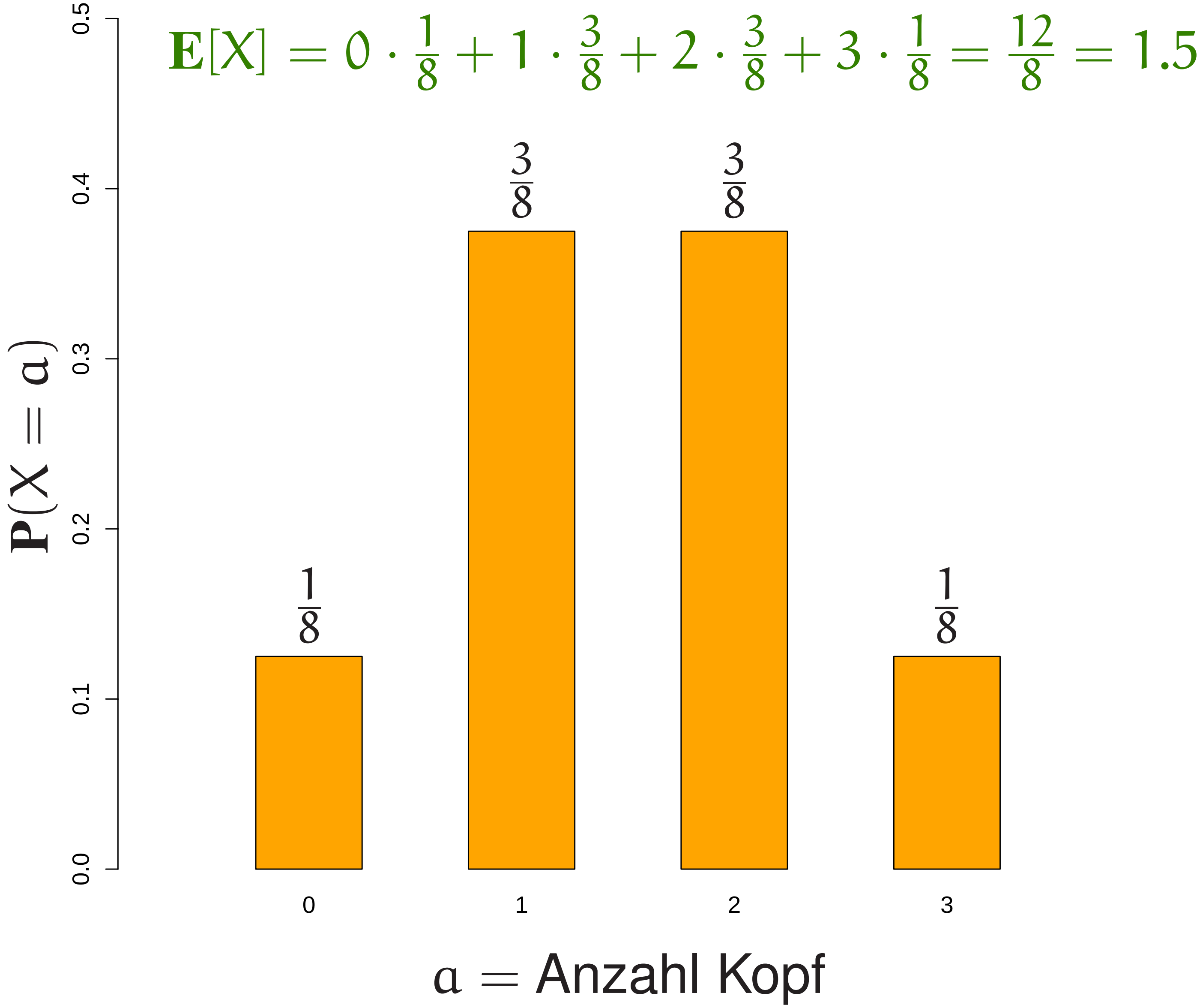


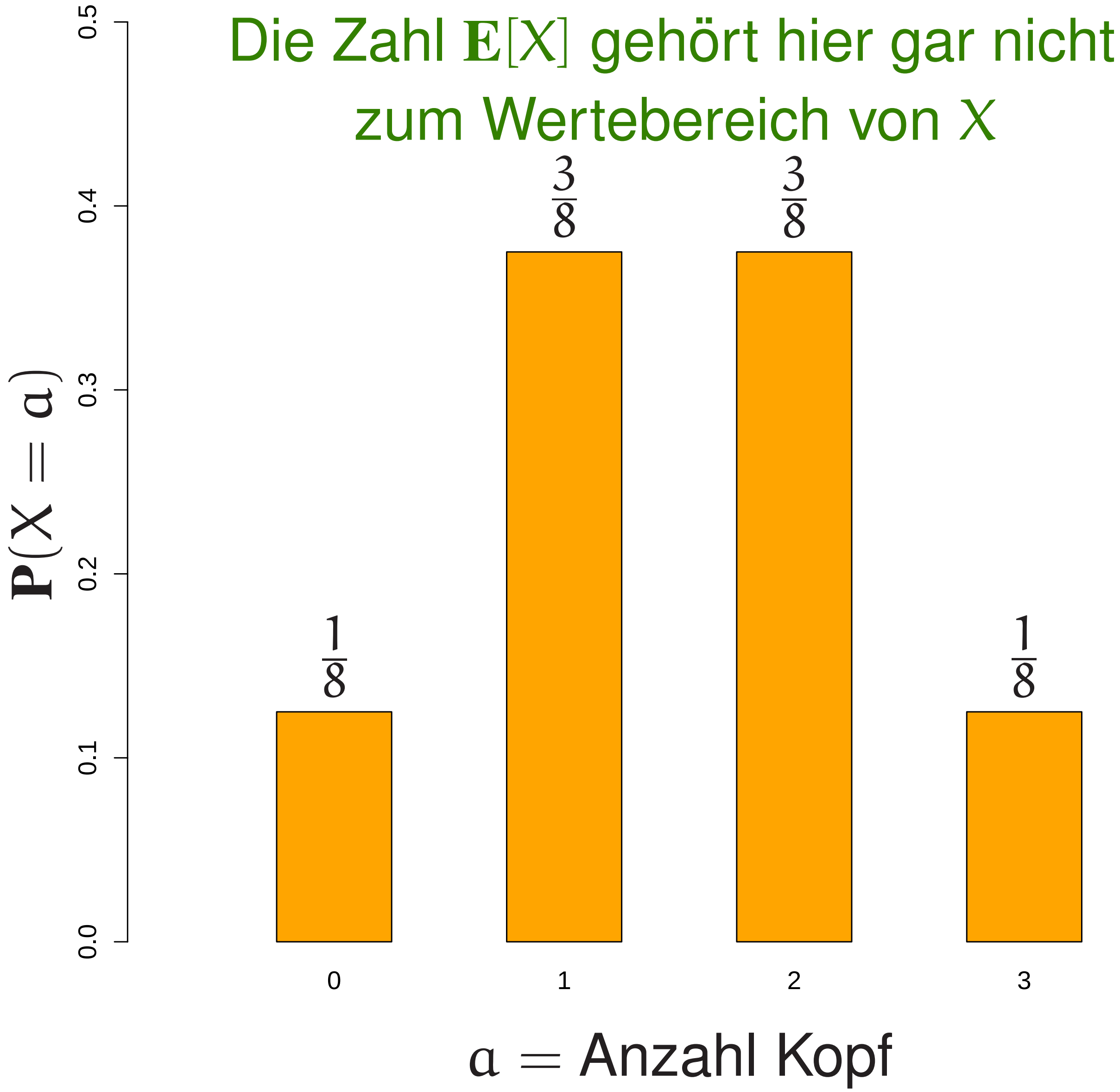
Eine faire Münze wird dreimal geworfen.
0.0
0.1
0.2
0.3
0.4
0.5
P(X = a)
0
1
2
3
a = Anzahl Kopf

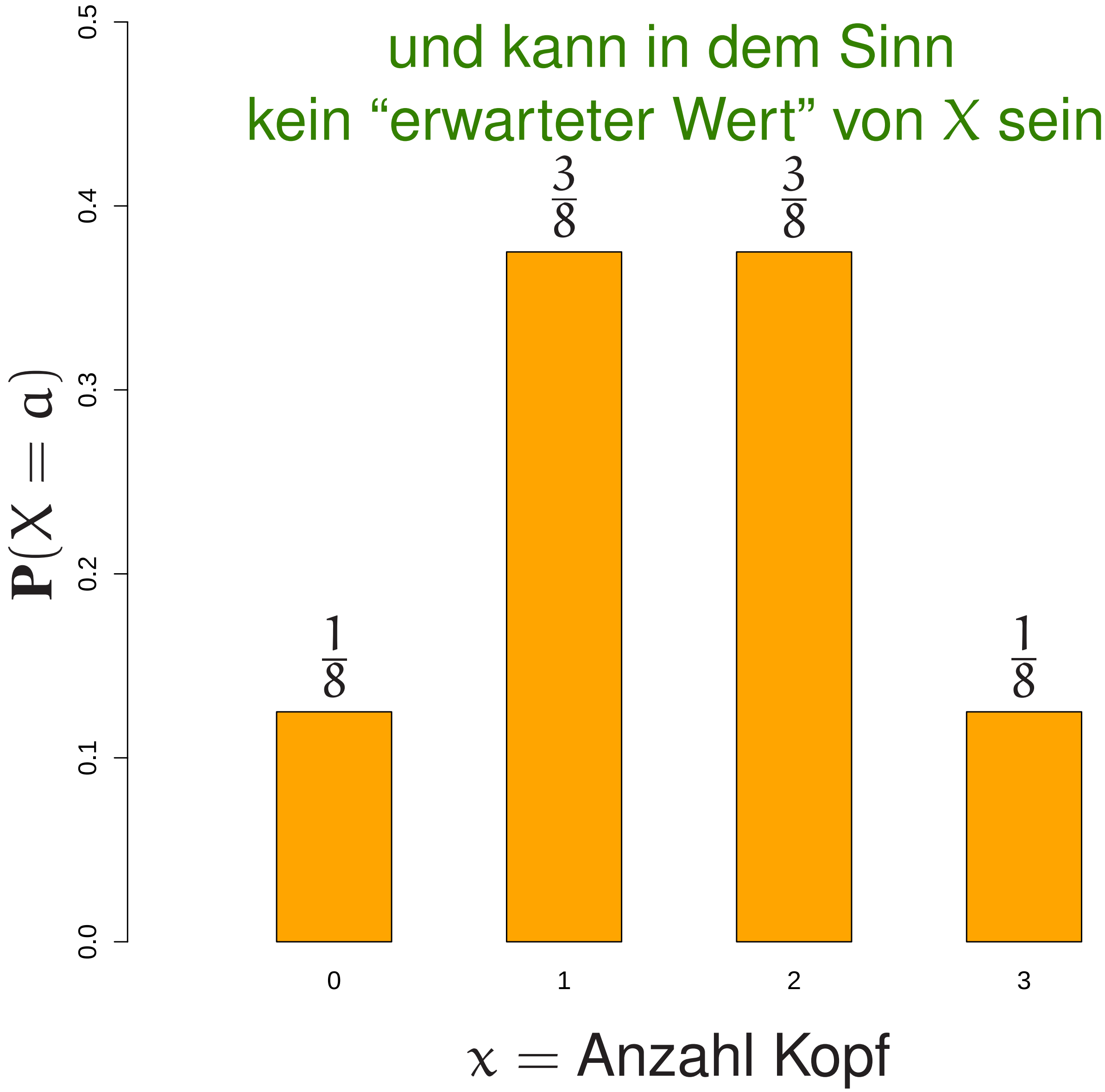

$\mathbf{E}[X] = ?$
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = x)$
0
1
2
3
x = Anzahl Kopf

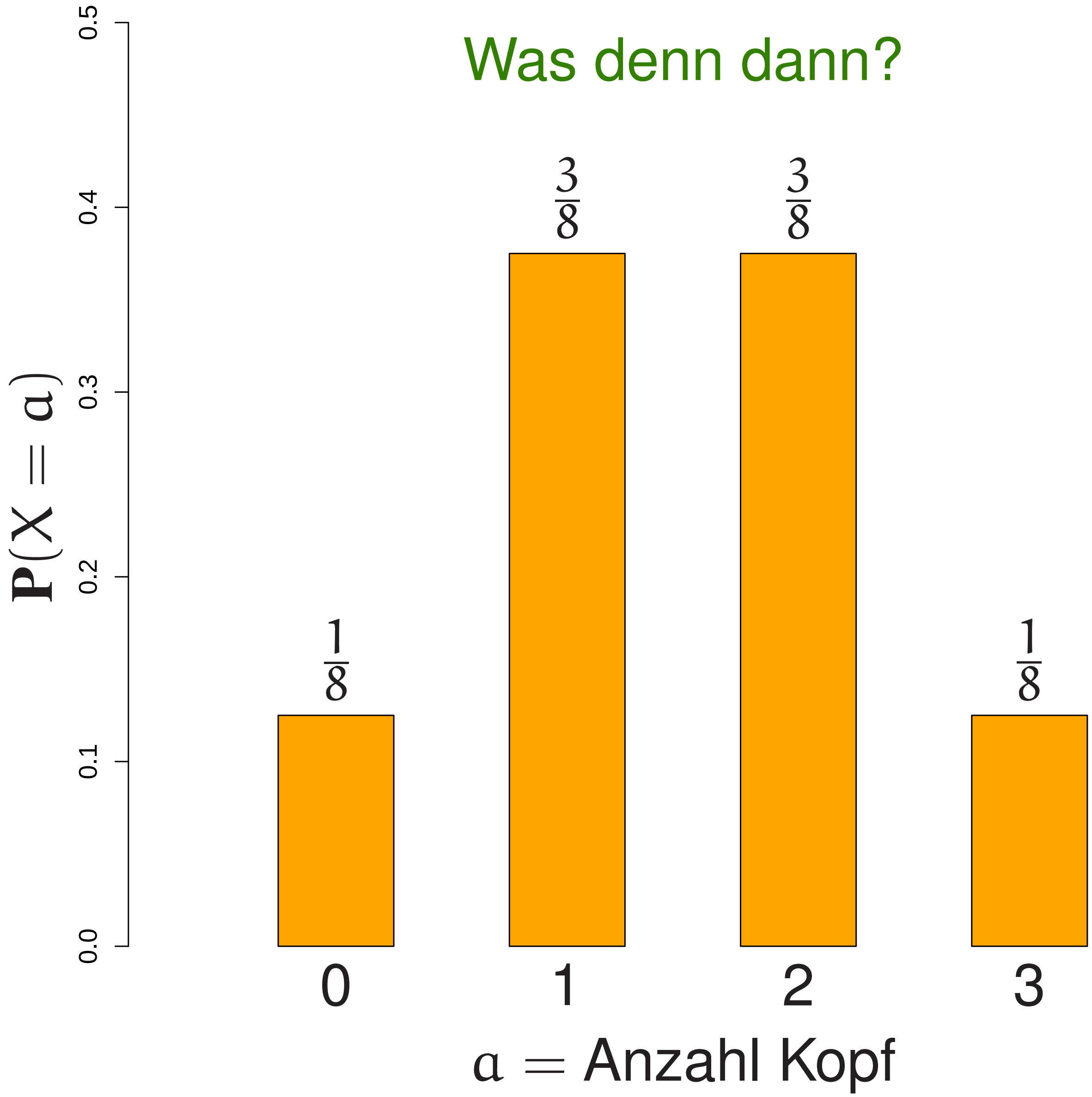


$\mathbf{E}[X] = \sum a\,\mathbf{P}(X = a)$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
x = Anzahl Kopf

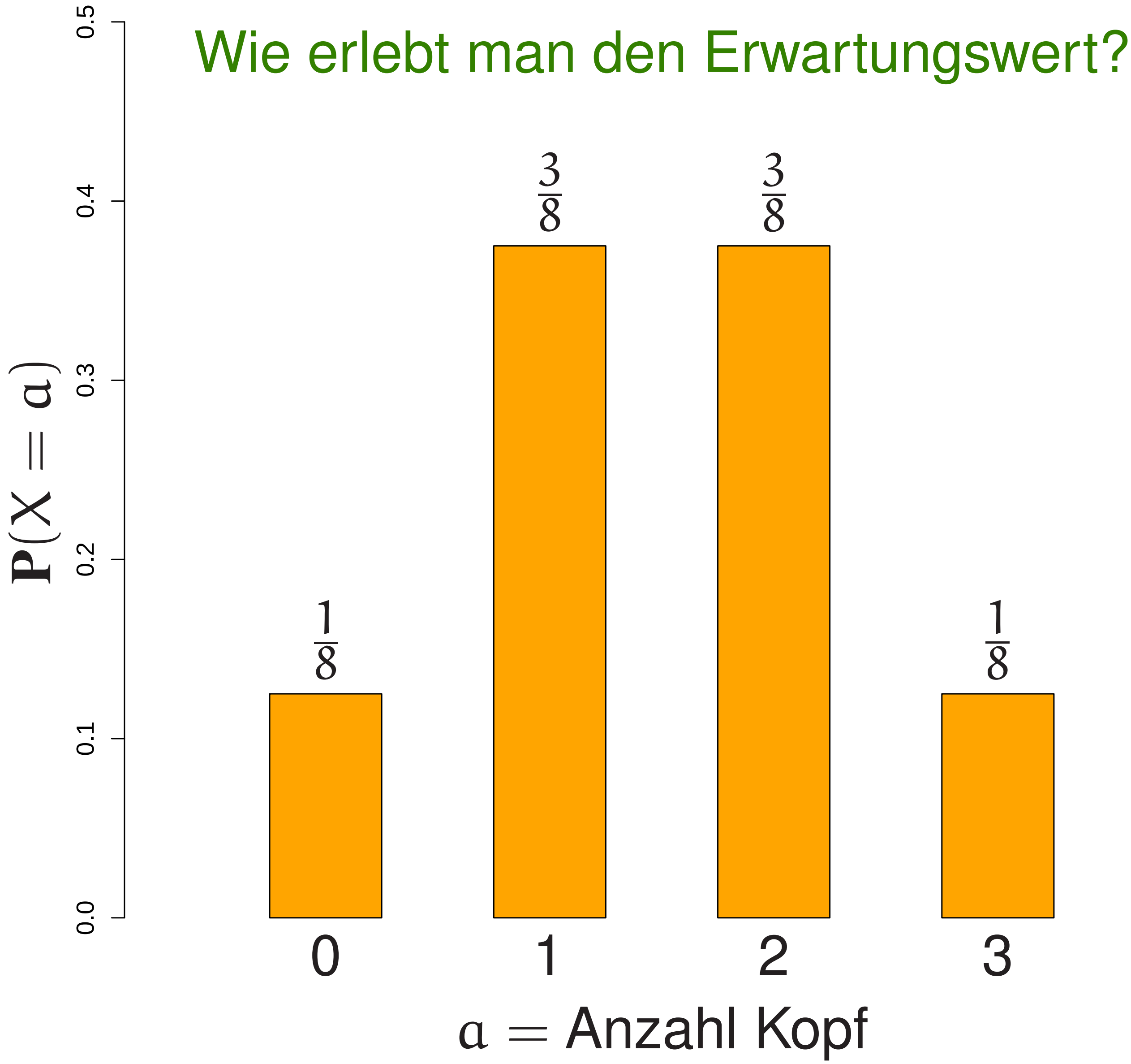

$\mathbf{E}[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Kopf

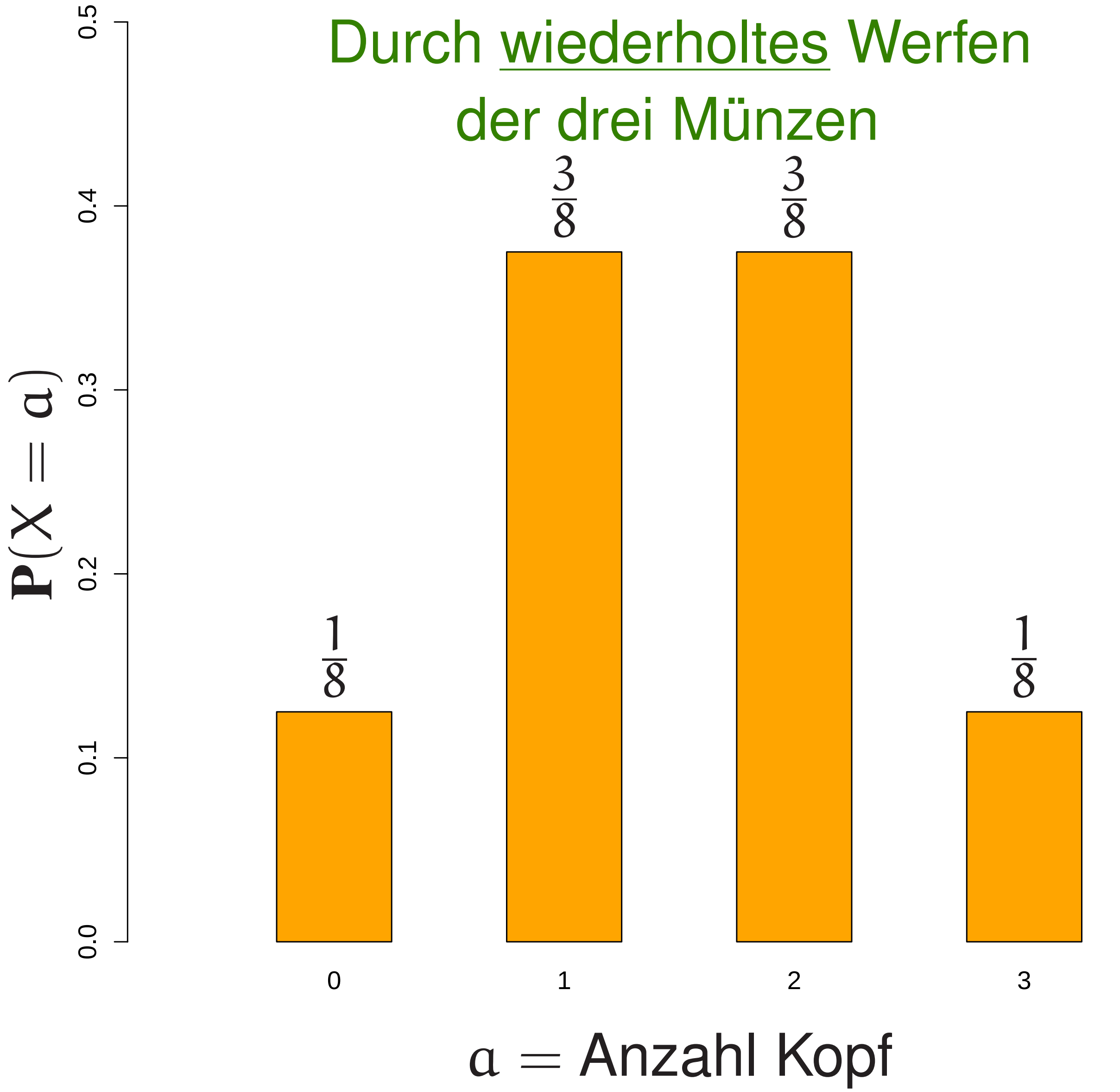

Die Zahl $\mathbf{E}[X]$ gehört hier gar nicht
zum Wertebereich von X
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
a = Anzahl Kopf


und kann in dem Sinn
kein “erwarteter Wert” von X sein
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
x = Anzahl Kopf



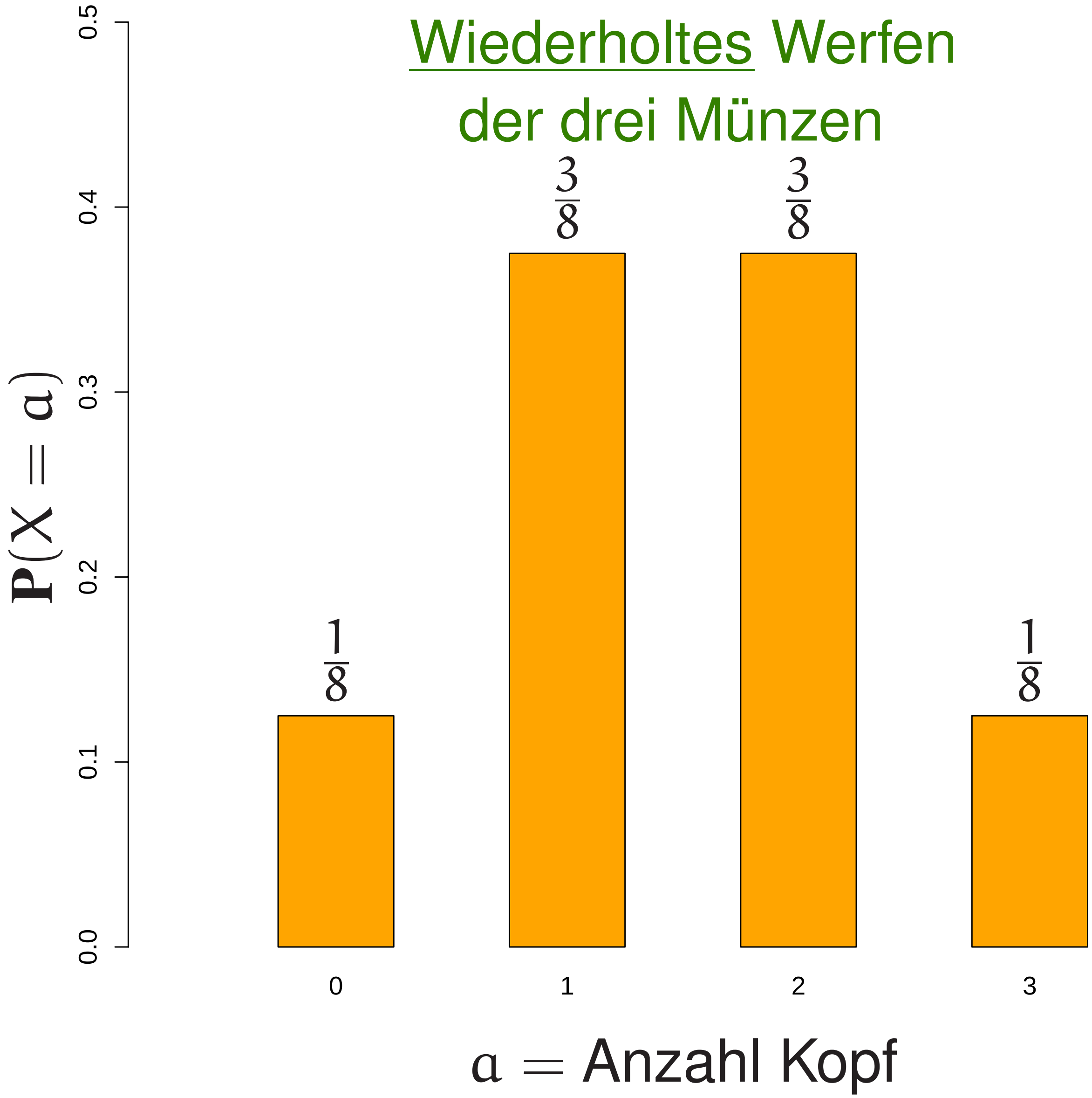
Was denn dann?
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Kopf


Wie erlebt man den Erwartungswert?
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 Kopf

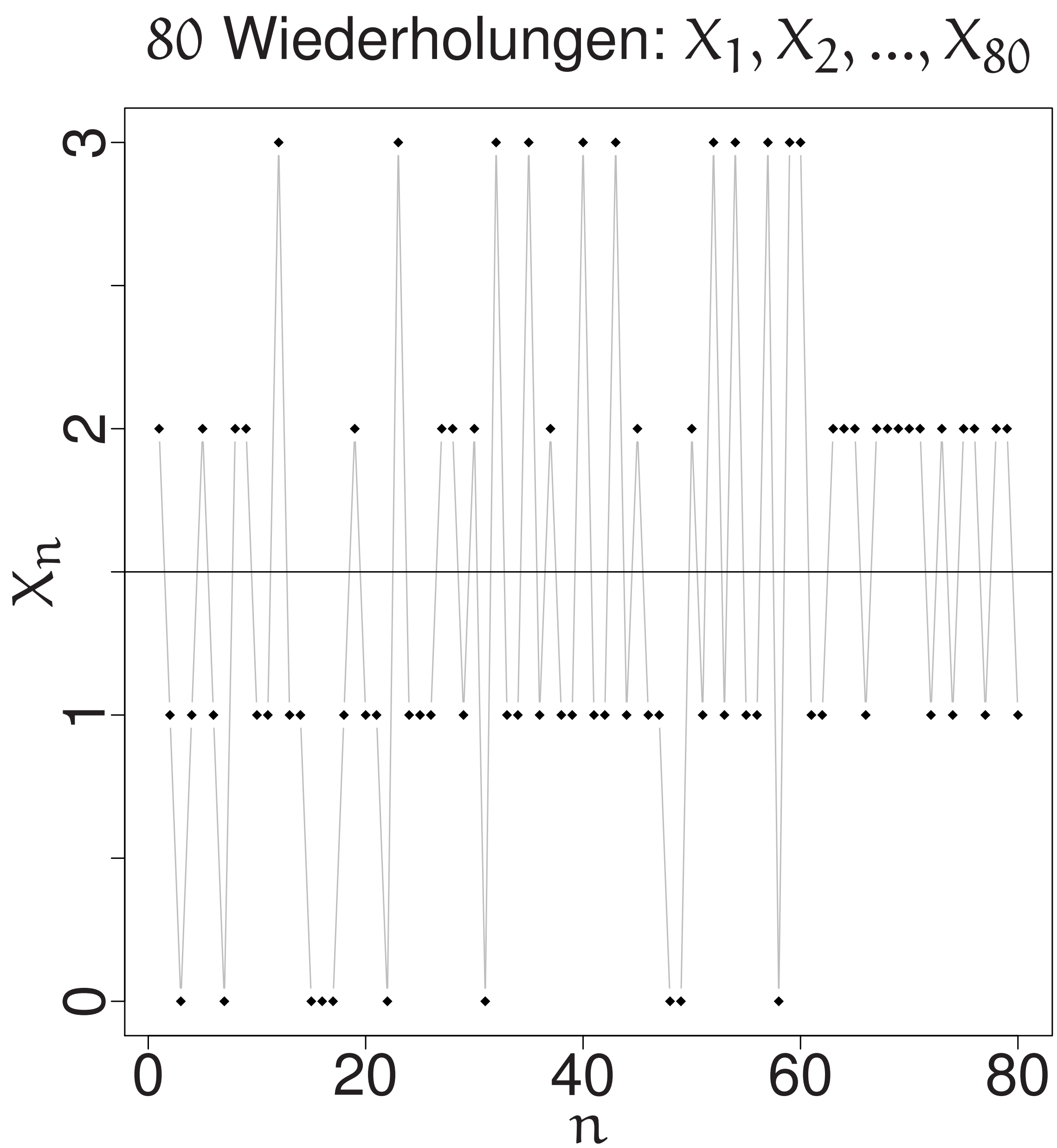


Durch wiederholtes Werfen
der drei Münzen
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
$a =$ Anzahl Kopf

Der Erwartungswert als Langzeitmittel

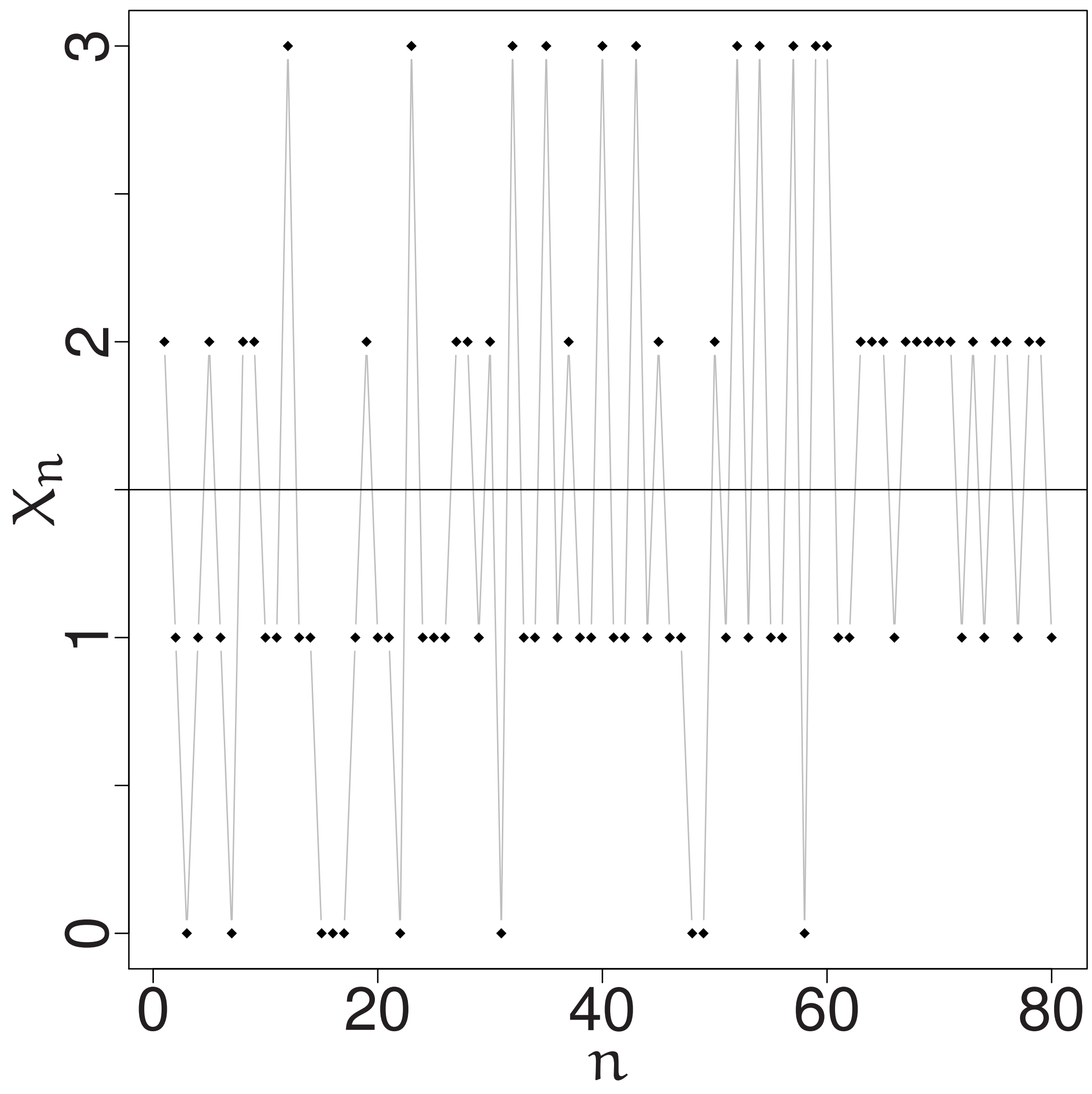


Wiederholtes Werfen der drei Münzen
$P(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Kopf

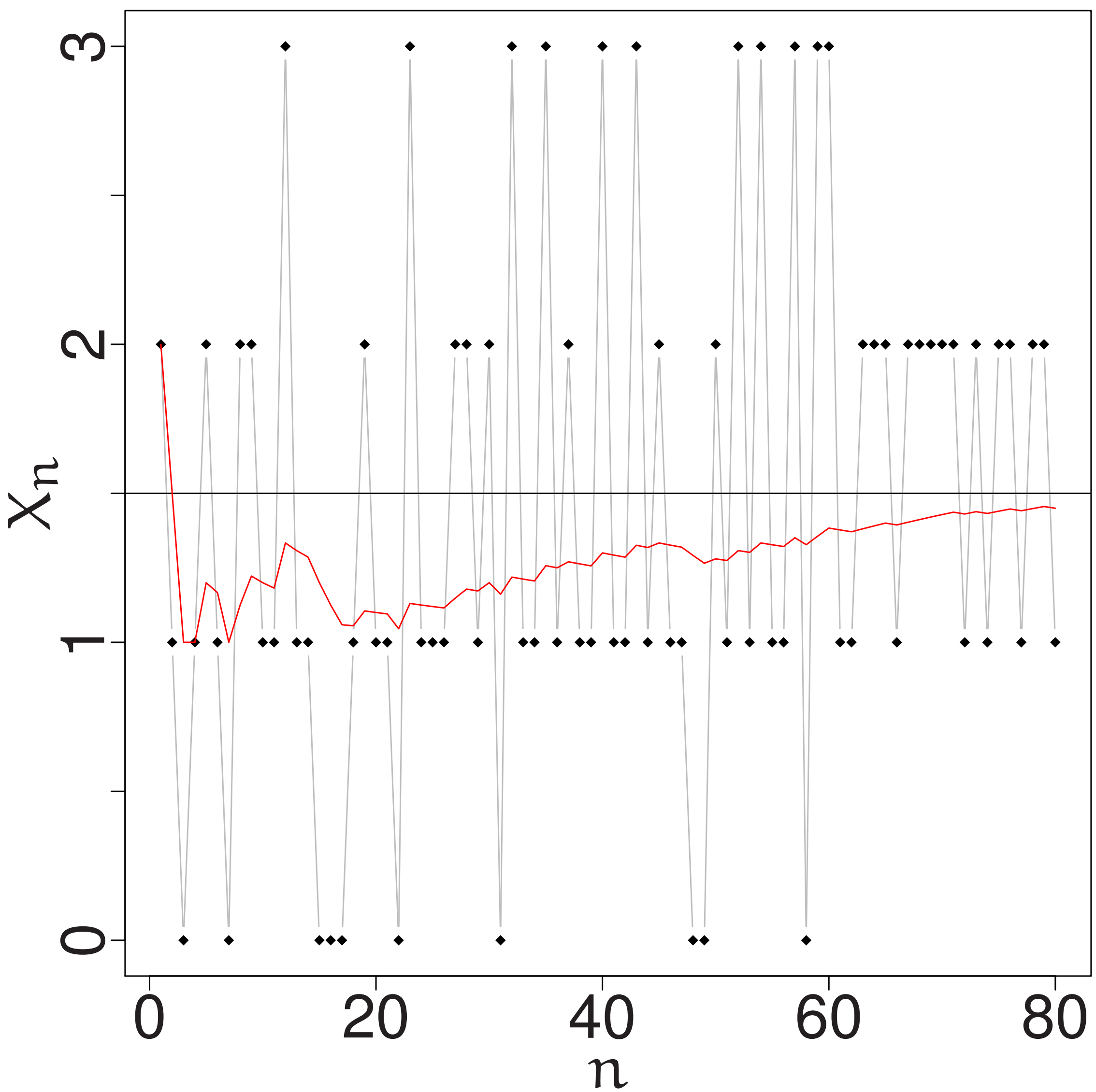


80 Wiederholungen: $X_1, X_2, ..., X_{80}$
X_n
n
0
1
2
3
0
20
40
60
80

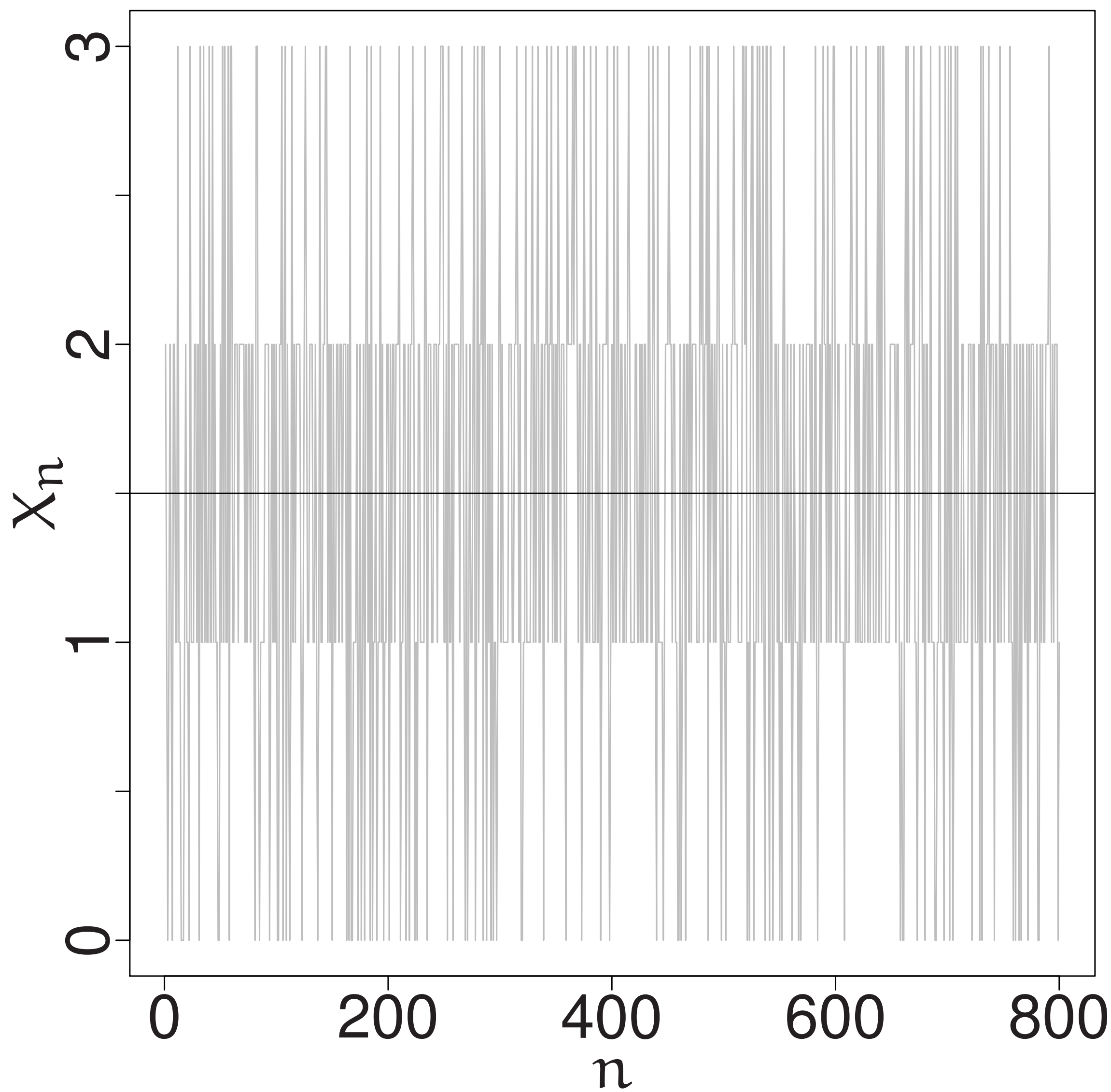
$$M_n := (X_1 + X_2 + \ldots + X_n) \,/\, n$$



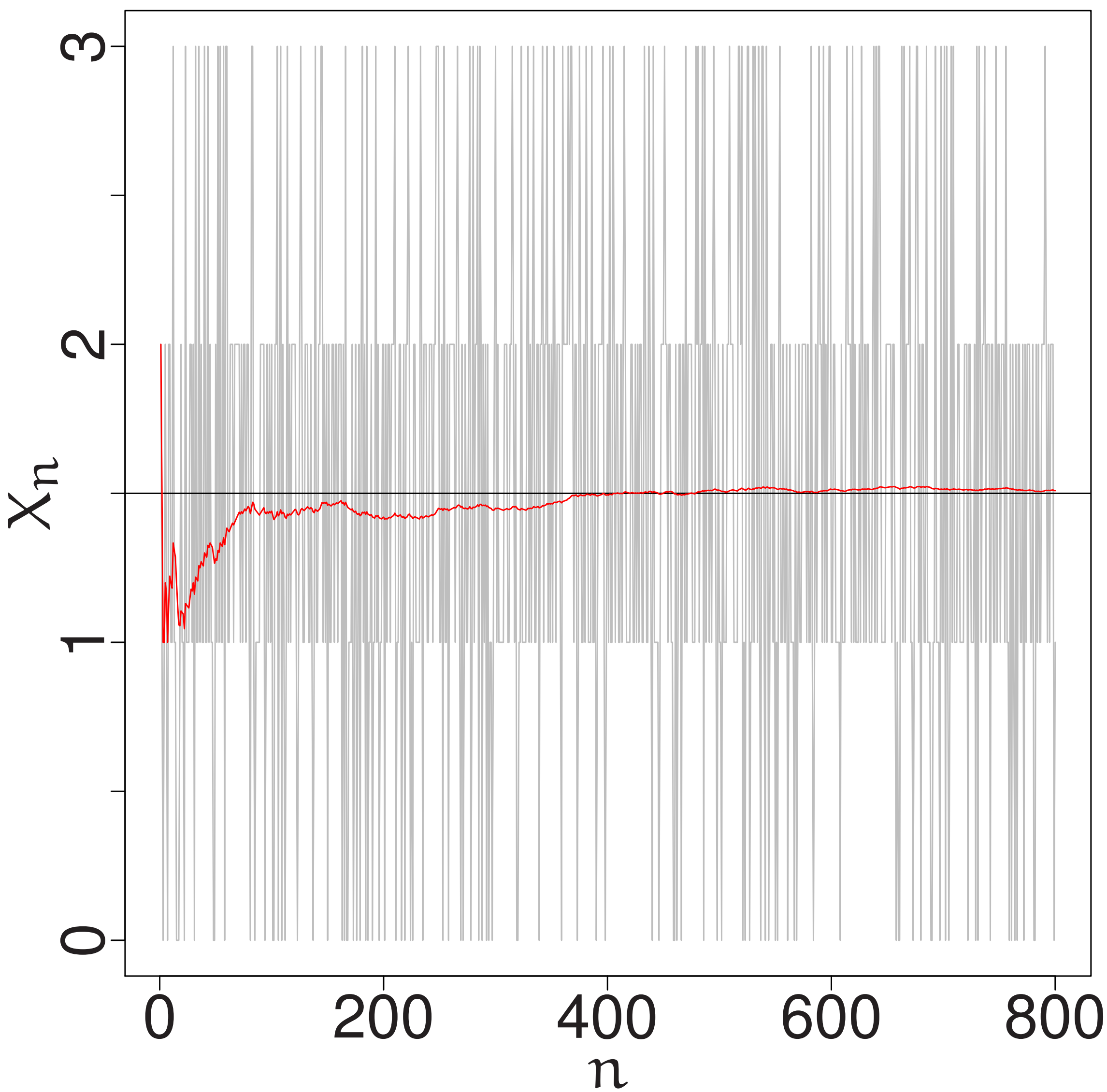
$$M_n := (X_1 + X_2 + \ldots + X_n)\,/\,n$$

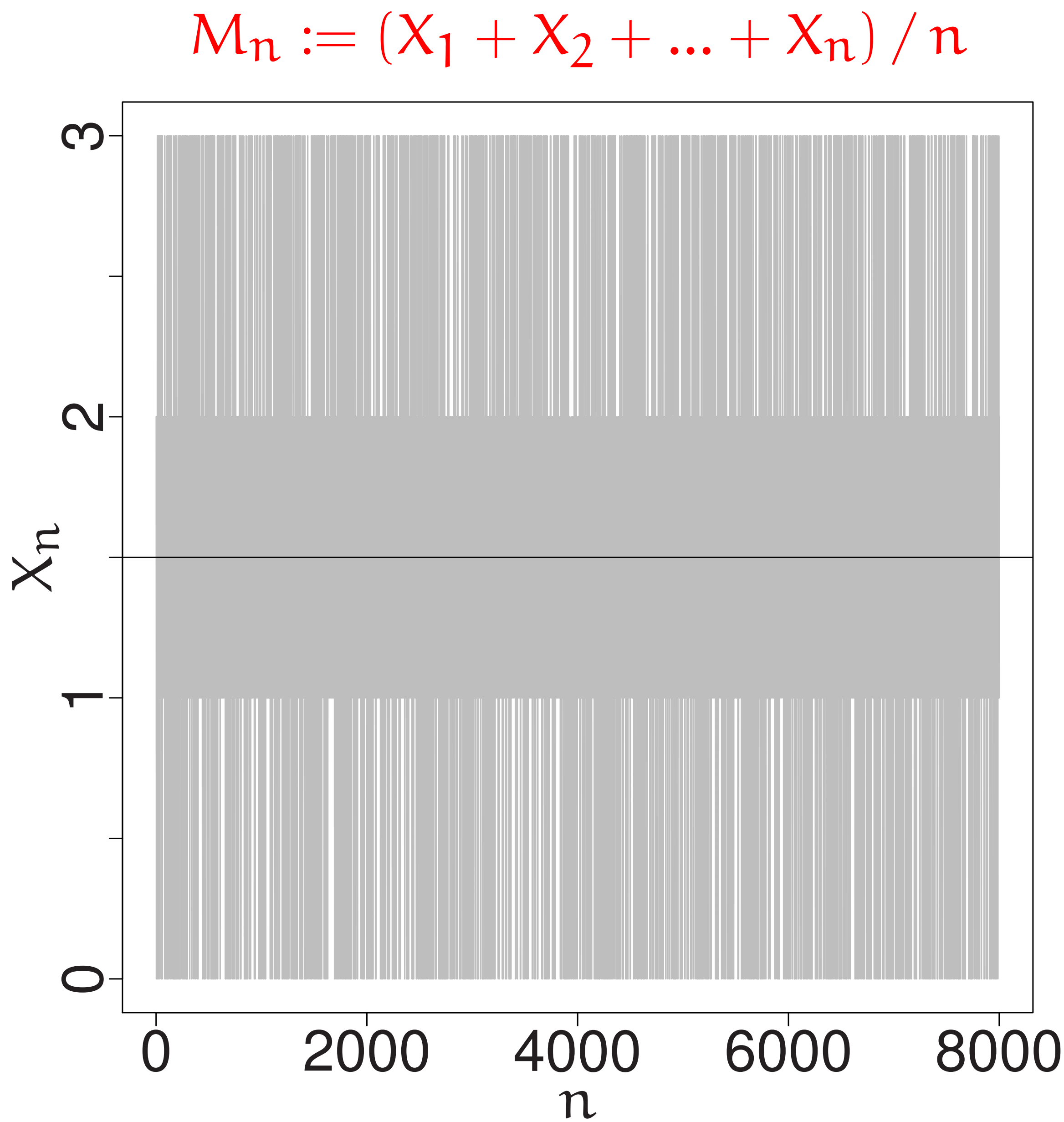


$M_n := (X_1 + X_2 + \ldots + X_n) \,/\, n$

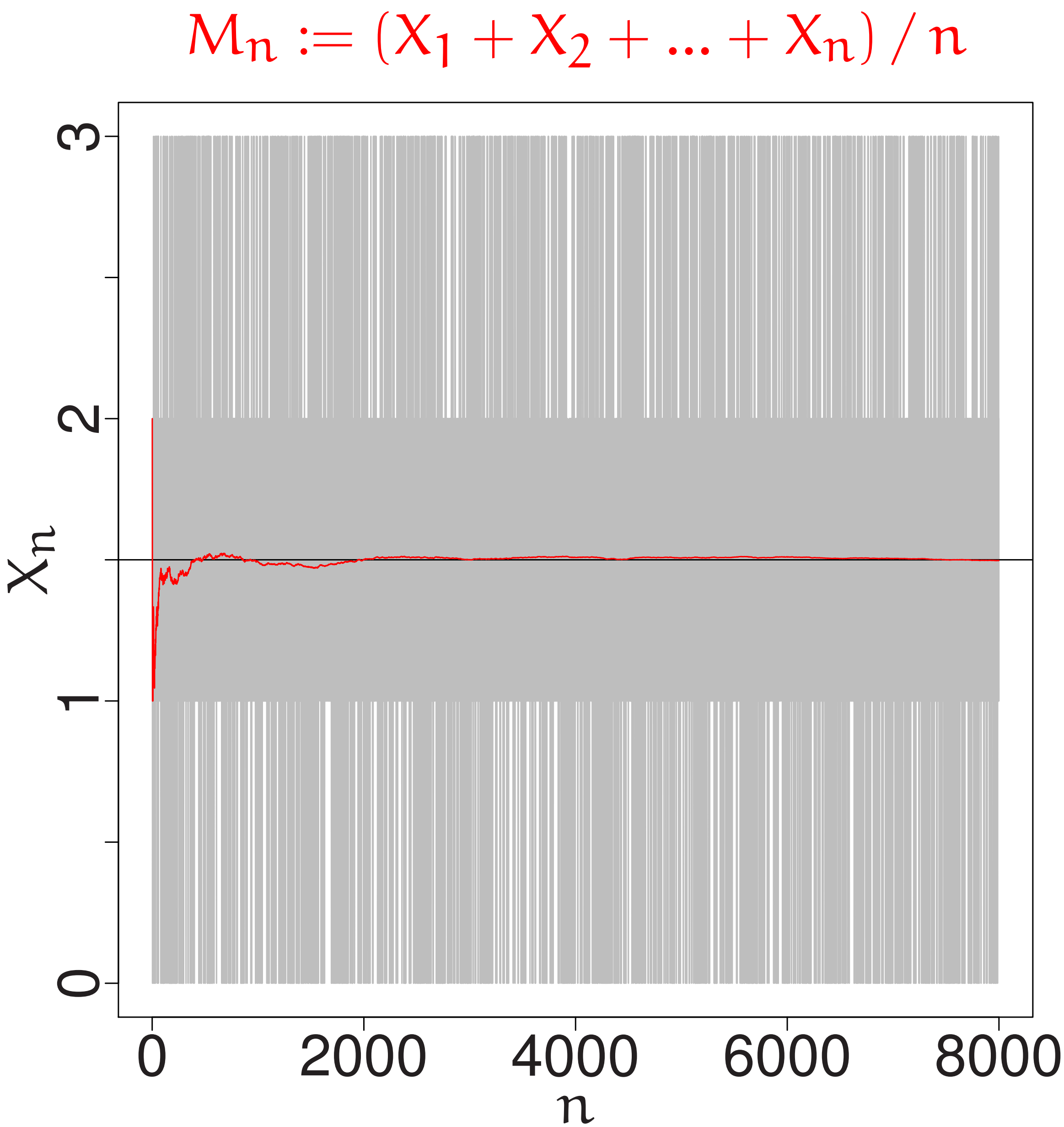


$$M_n := (X_1 + X_2 + \ldots + X_n) \,/\, n$$

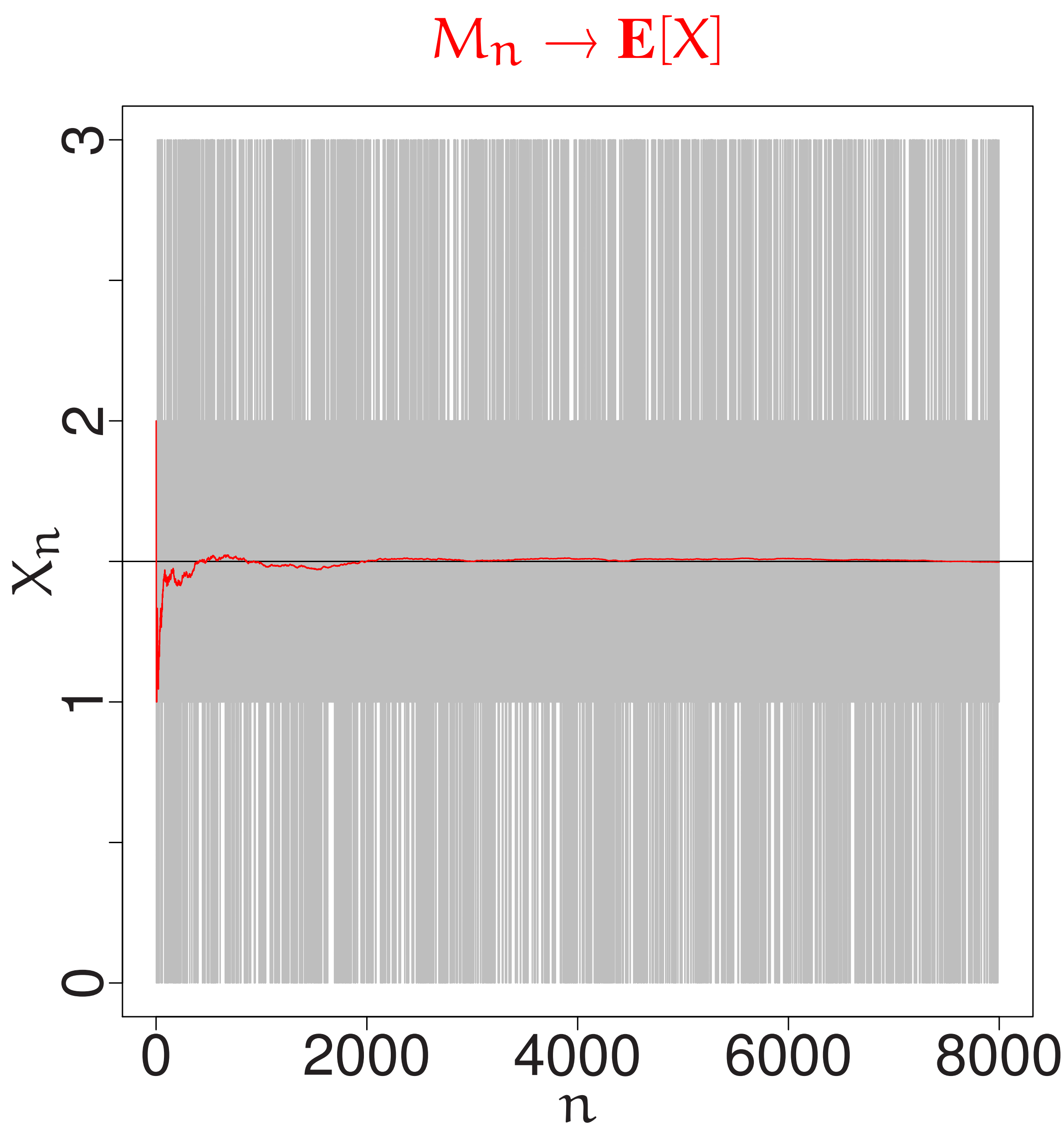




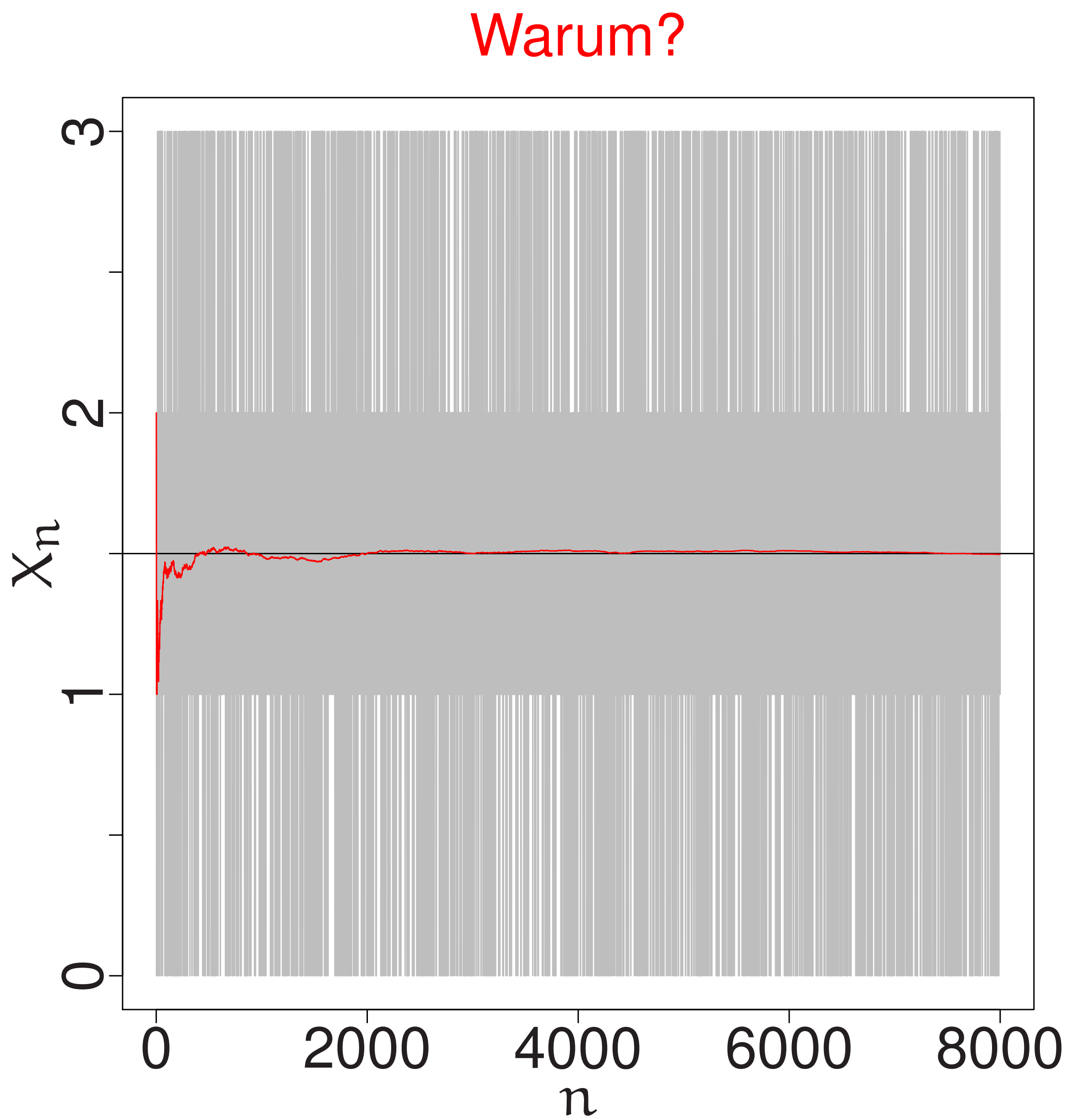
$M_n := (X_1 + X_2 + ... + X_n) / n$
X_n
0
1
2
3
0
2000
4000
6000
8000
n



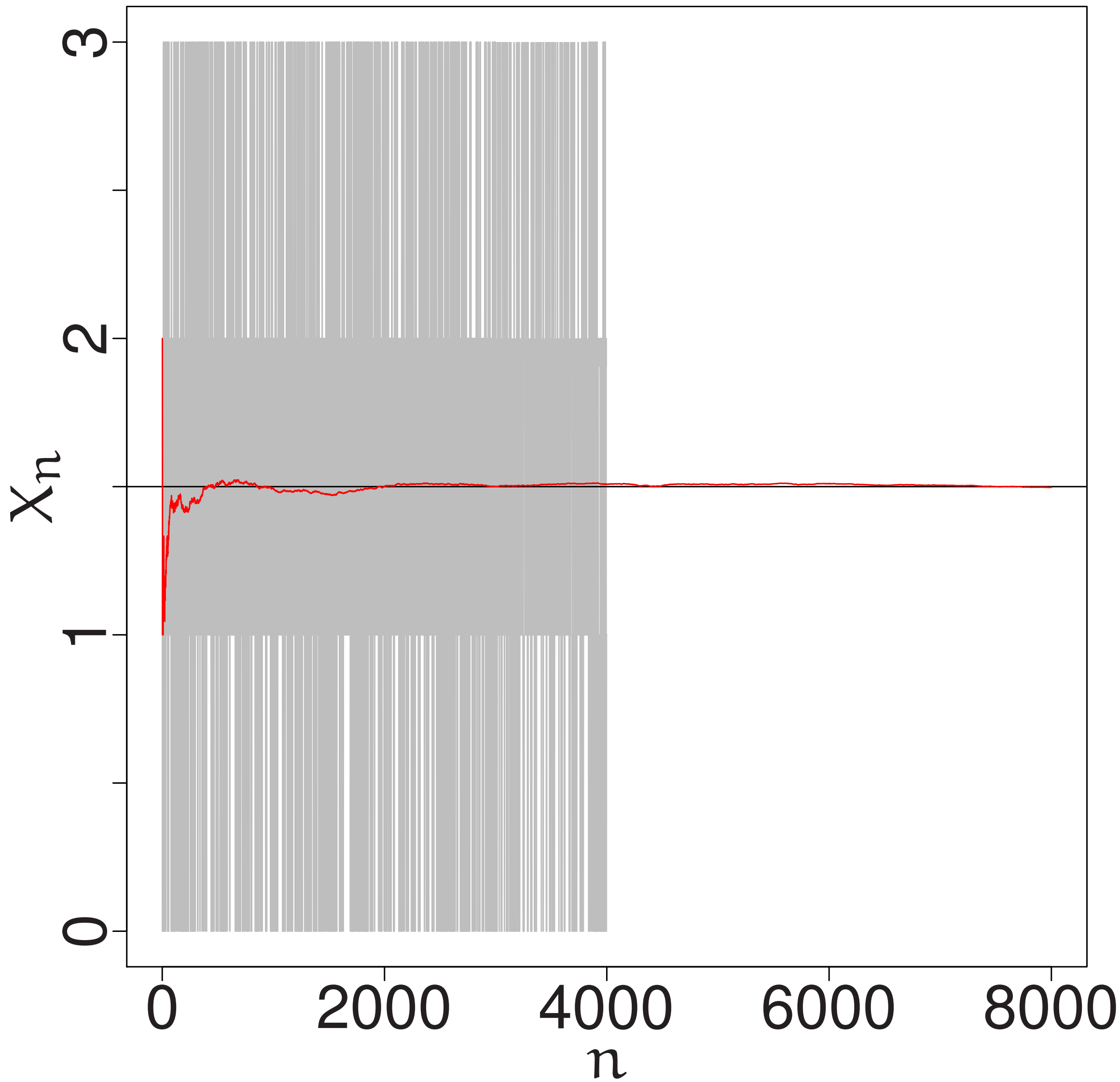
$M_n := (X_1 + X_2 + ... + X_n) / n$
3
2
1
0
X_n
0
2000
4000
6000
8000
n



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


Warum?
0
1
2
3
X_n
0
2000
4000
6000
8000
n

$$M_n = \sum a \ \#\{\text{Würfe mit Ergebis } a\}/n \to \sum a \ \mathbf{P}(X = a)$$



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. <u>Gewichtetes Mittel</u>

der möglichen Werte:

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

2. <u>Langzeitmittelwert</u>

bei unabhängigen Wiederholungen:

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