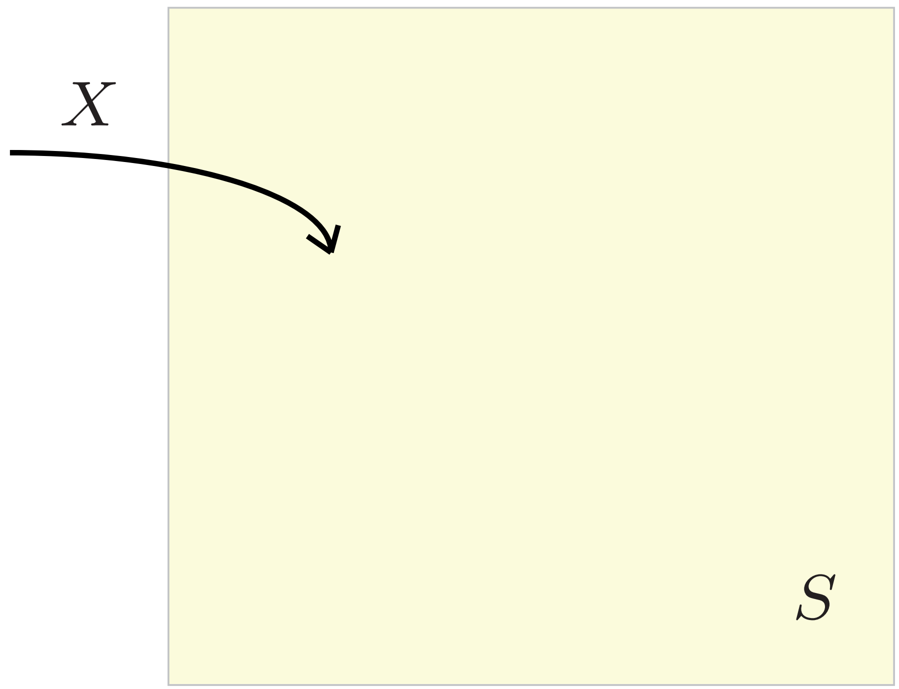
Vorlesung 1a

Zufallsvariable und Ereignisse, Wahrscheinlichkeiten und Verteilungen

Ein erster Blick

1. Rein zufällige Wahl aus einer endlichen Menge

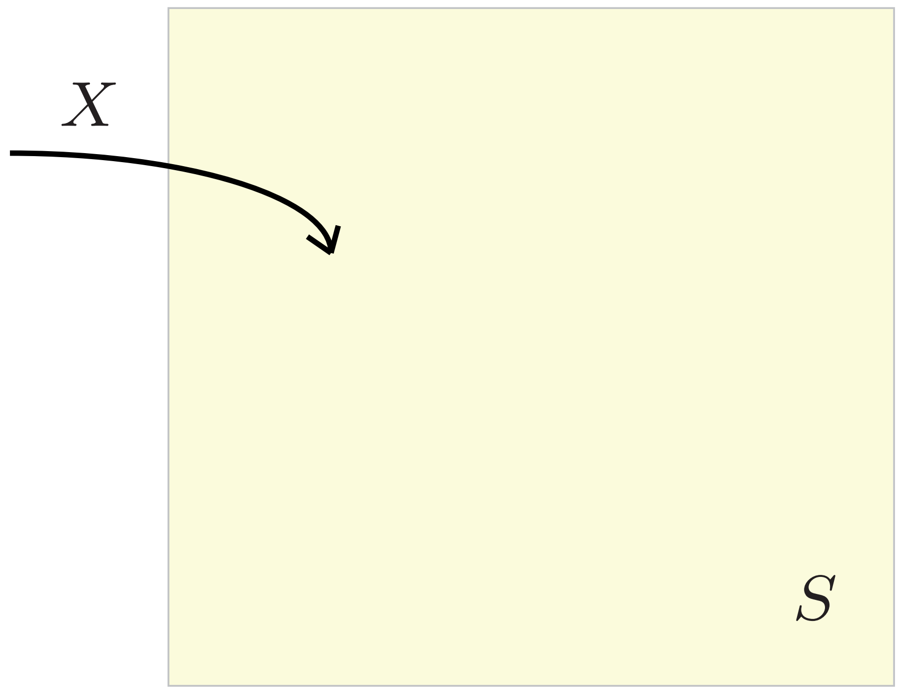
Prototypisches Beispiel einer Zufallsvariablen:



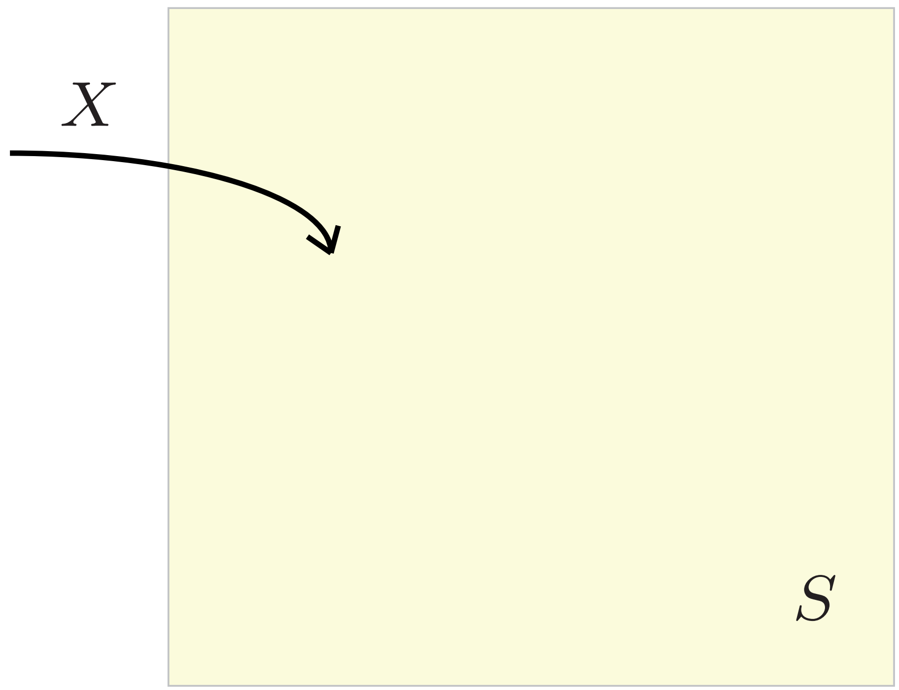
rein zufällige Wahl eines Punktes aus einer Fläche

(z. B. aus dem beigen Quadrat S)

Prototypisches Beispiel einer Zufallsvariablen:



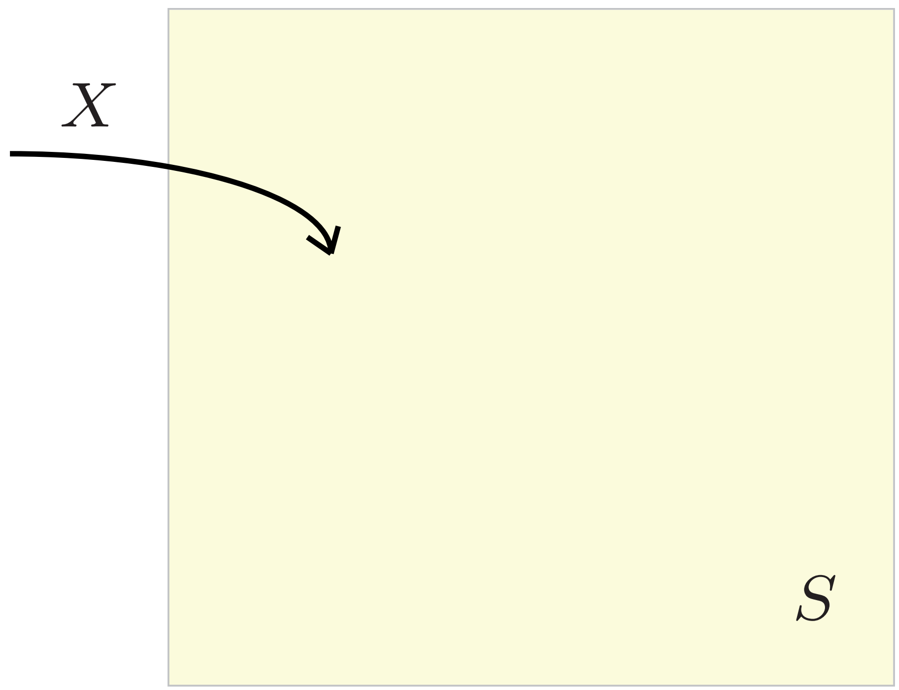
Was heißt “rein zufällige Wahl” eines Punktes aus S?



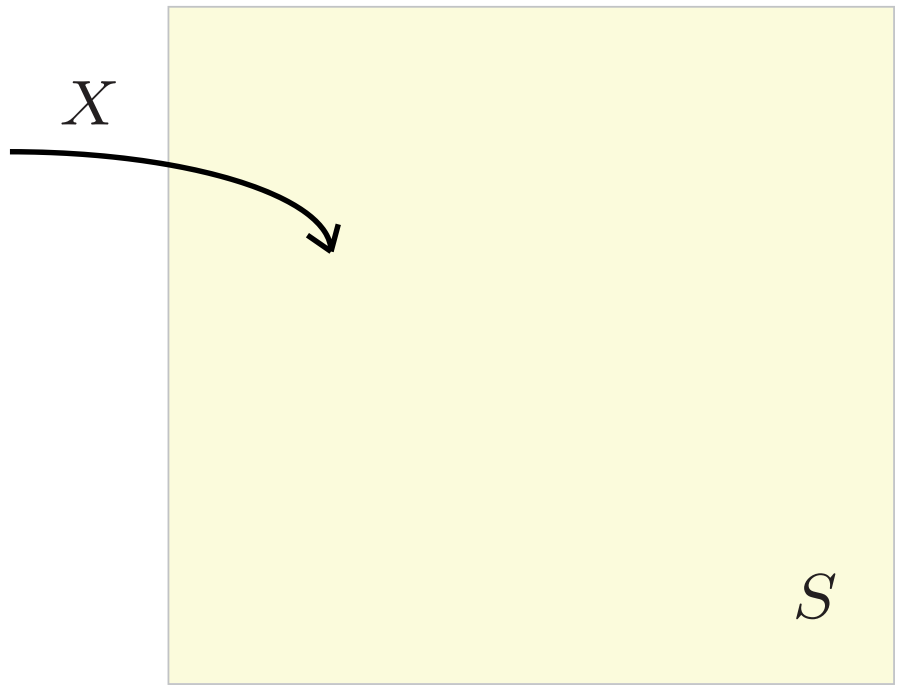
Fürs Erste sind endlich viele Pixel leichter vorstellbar

als ein Kontinuum aus unendlich vielen Punkten.

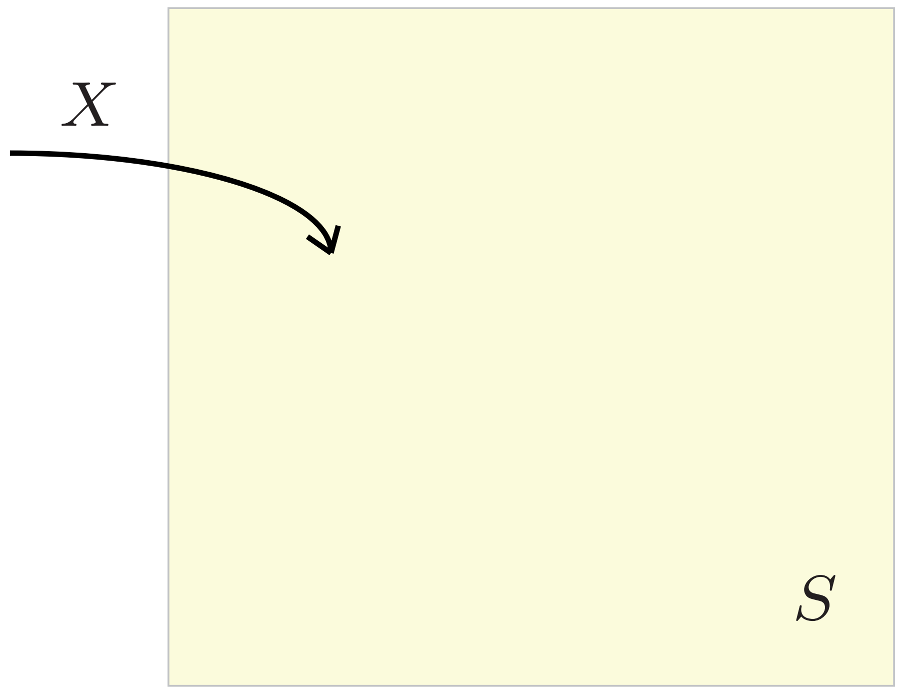
Stellen wir uns vor, S besteht aus 1000×1000 Pixeln



Stellen wir uns vor, S besteht aus g Pixeln, mit $g \in \mathbb{N}$

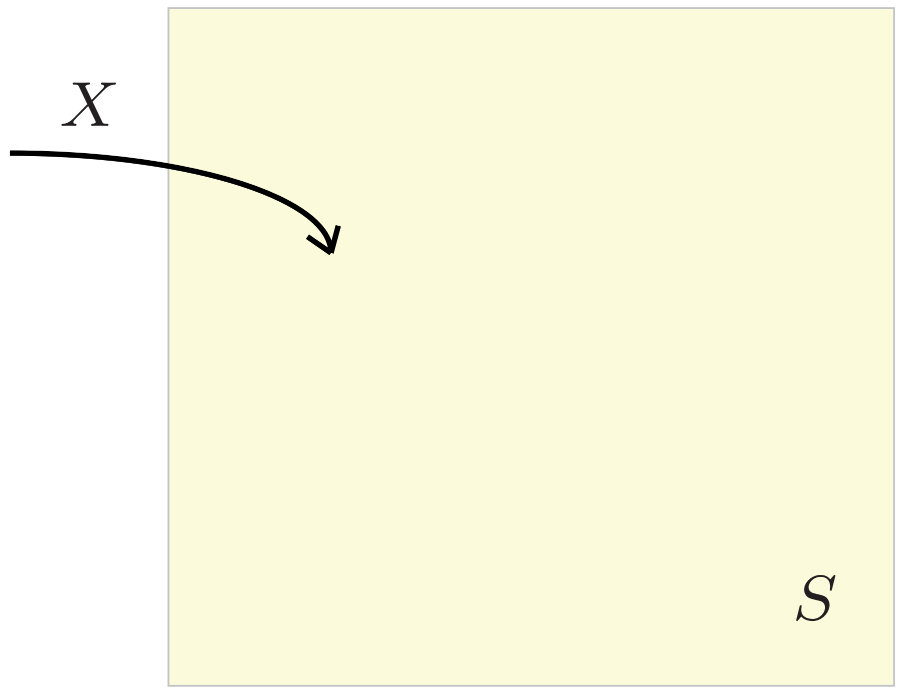


“Rein zufällige Wahl aus S” soll heißen:



alle Pixel in S haben dieselbe Chance,

zum Zug zu kommen.



Man spricht dann von einer

uniform auf S verteilten Zufallsvariablen X

Analogie zum **fairen Würfeln**:

Die Menge der möglichen Ausgänge ist hier

$$S = \{1, 2, 3, 4, 5, 6\}.$$

Auch hier wird der zufällige Ausgang beschrieben durch eine uniform auf S verteilte Zufallsvariable X.

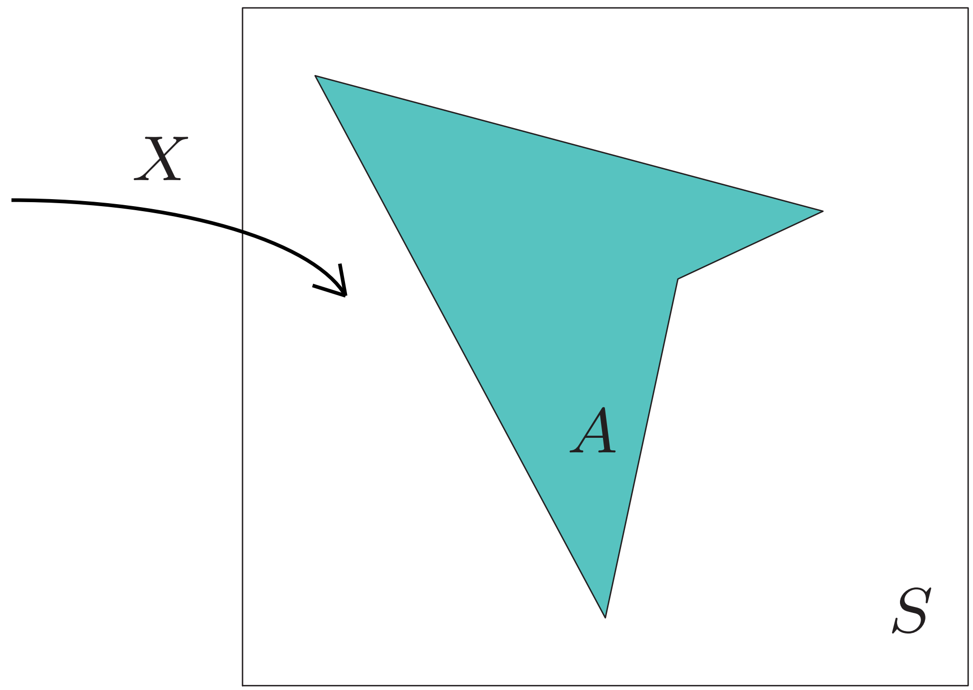
Wir halten fest:

Zufallsvariable mit Werten in einer endlichen Menge S
entsprechen einer zufälligen Wahl aus S.

Auf S **uniform** verteilte Zufallsvariable
entsprechen einer **rein** zufälligen Wahl aus S,
bei der alle Elemente von S die gleiche Chance haben,
zum Zug zu kommen.

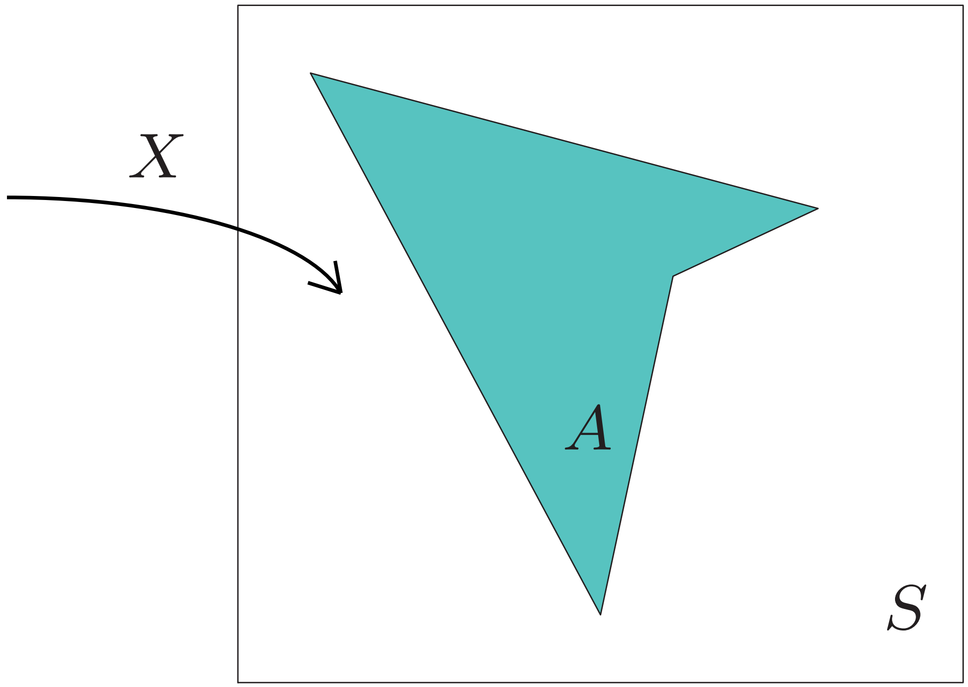
2. Ereignisse

Betrachten wir wieder unser Quadrat S



diesmal zusammen mit einer bestimmten Teilmenge A von S

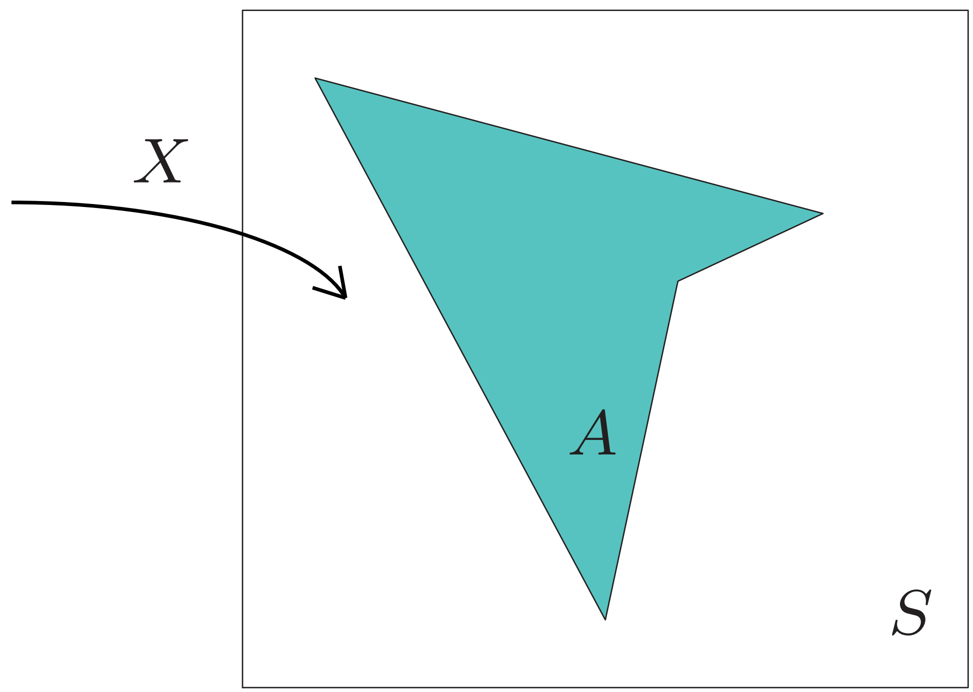
Bei der Wahl eines Pixels aus S



kann die Wahl auf A fallen

– oder auch nicht.

Bei der Wahl eines Pixels aus S



kann das Ereignis "X fällt in A" eintreten

– oder auch nicht.

Das Ereignis "X fällt in A"

notiert man als

$\{X \in A\}$.

Die Menge aller Ereignisse $\{X \in A\}, \quad A \subset S$,

nennt man auch die

"von X erzeugte Kollektion von Ereignissen".

Ereignisse kann man (aussagen-)logisch verknüpfen, z.B. gilt:

$$\{X \in A\} \text{ und } \{X \in B\} = \{X \in A \cap B\}.$$

Mehr dazu später!

3. Wahrscheinlichkeiten

Wie wahrscheinlich ist es,

dass bei einer rein zufälligen Wahl eines Pixels aus S

die Wahl auf A fällt?

Wie groß ist die Wahrscheinlichkeit

des Ereignisses $\{X \in A\}$?

.... eine Zahl zwischen 0 ($= 0\%$) und 1 ($= 100\%$)

Merke: Das Ereignis $\{X \in S\}$ hat Wahrscheinlichkeit 1.

Für die Wahrscheinlichkeit des Ereignisses $\{X \in A\}$

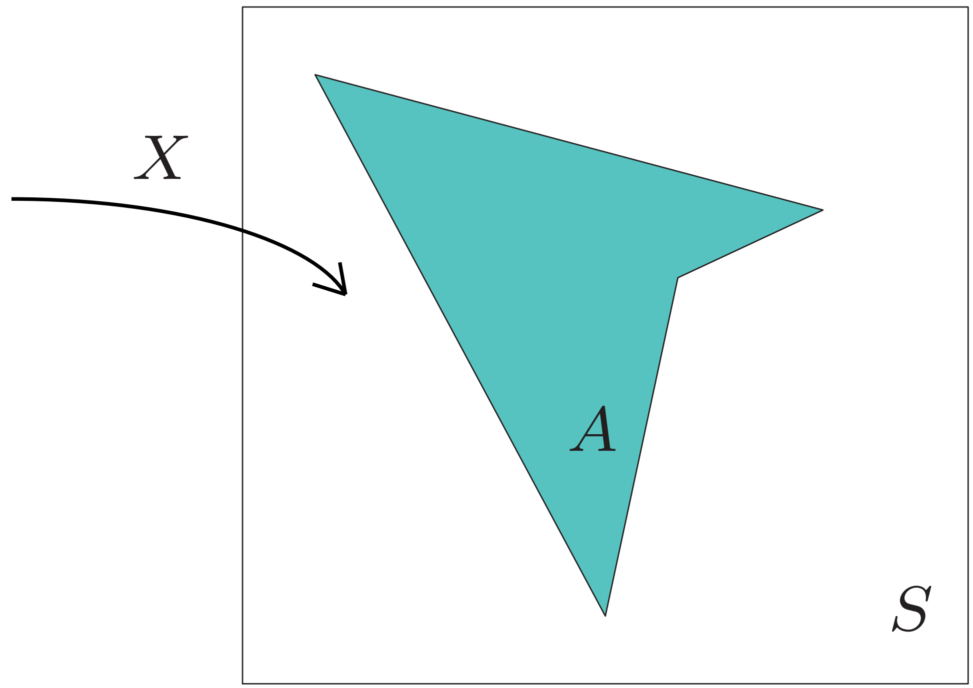
schreiben wir

$$\mathbf{P}(\{X \in A\})$$

oder kurz

$$\mathbf{P}(X \in A).$$

P steht für **probabilitas** (= Wahrscheinlichkeit)



Bei der rein zufälligen Wahl eines Pixels aus S,

beschrieben durch die Zufallsvariable X, ist

die Wahrscheinlichkeit des Ereignisses $\{X \in A\}$

proportional zur Anzahl der Pixel in A.

Zusammengefasst:

X ist *rein zufälliger Pixel aus dem Quadrat* S

bedeutet:

Für jede Teilmenge A von S ist

$$\mathbf{P}(\{X \in A\}) = \frac{\text{Anzahl der Pixel in } A}{\text{Anzahl der Pixel in } S}.$$

Lies und merke:

die Wahrscheinlichkeit, dass X in A fällt

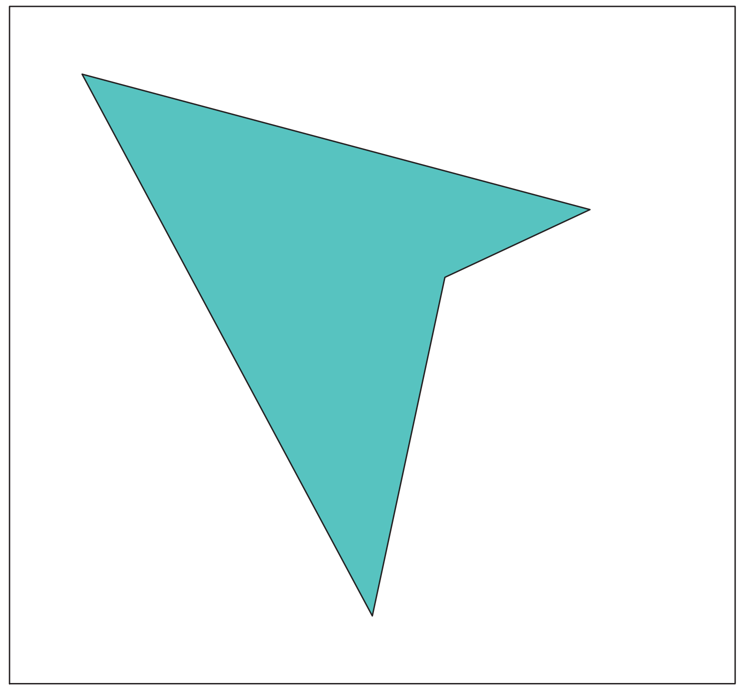
ist der Anteil der Menge A an der Menge S.

4. Schätzung eines Flächenanteils

Eine Anwendung der “rein zufälligen Wahl”:

Monte-Carlo Schätzung eines Flächenanteils.

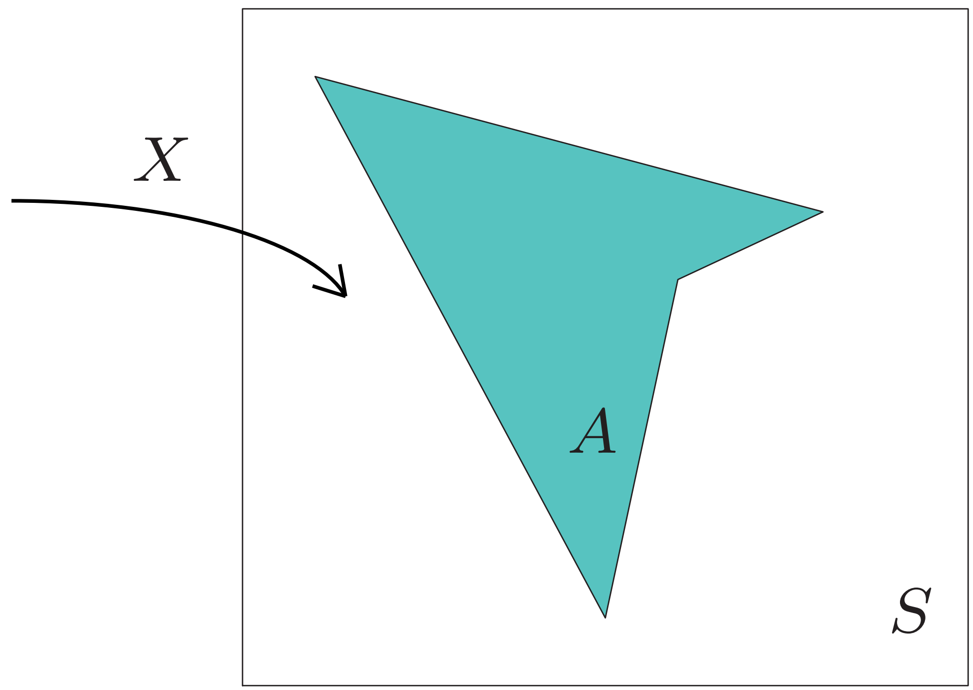
Wir fragen:



Wie groß ist der Anteil der blauen Fläche

an der Fläche des Quadrats?

und übersetzen in die Sprache der Stochastik:



Wie wahrscheinlich ist es, dass die

rein zufällige Wahl eines Pixels aus dem Quadrat S

in die blaue Fläche trifft?

Wie wir bald sehen werden

(und wie auch intuitiv klar ist)

gibt es einen engen Zusammenhang zwischen

Wahrscheinlichkeiten und Trefferquoten.

Angenommen wir haben ein Werkzeug, mit dem man einen **rein zufälligen Pixel** aus dem Quadrat wählen kann

– und das nicht nur einmal, sondern “immer wieder neu”.

Wir bekommen dann mit unserem Werkzeug
nicht nur *einen einzigen* rein zufälligen Pixel,

sondern sogar beliebig viele,

genauer:

eine *rein zufällige Folge* $(X_1, X_2, ...)$

von Pixeln in S.

Sei $A \subset S$.

$$Z_i := \mathbf{1}_A(X_i), \quad i = 1, 2, \ldots$$

zählt, ob X_i in A fällt.

Dabei ist

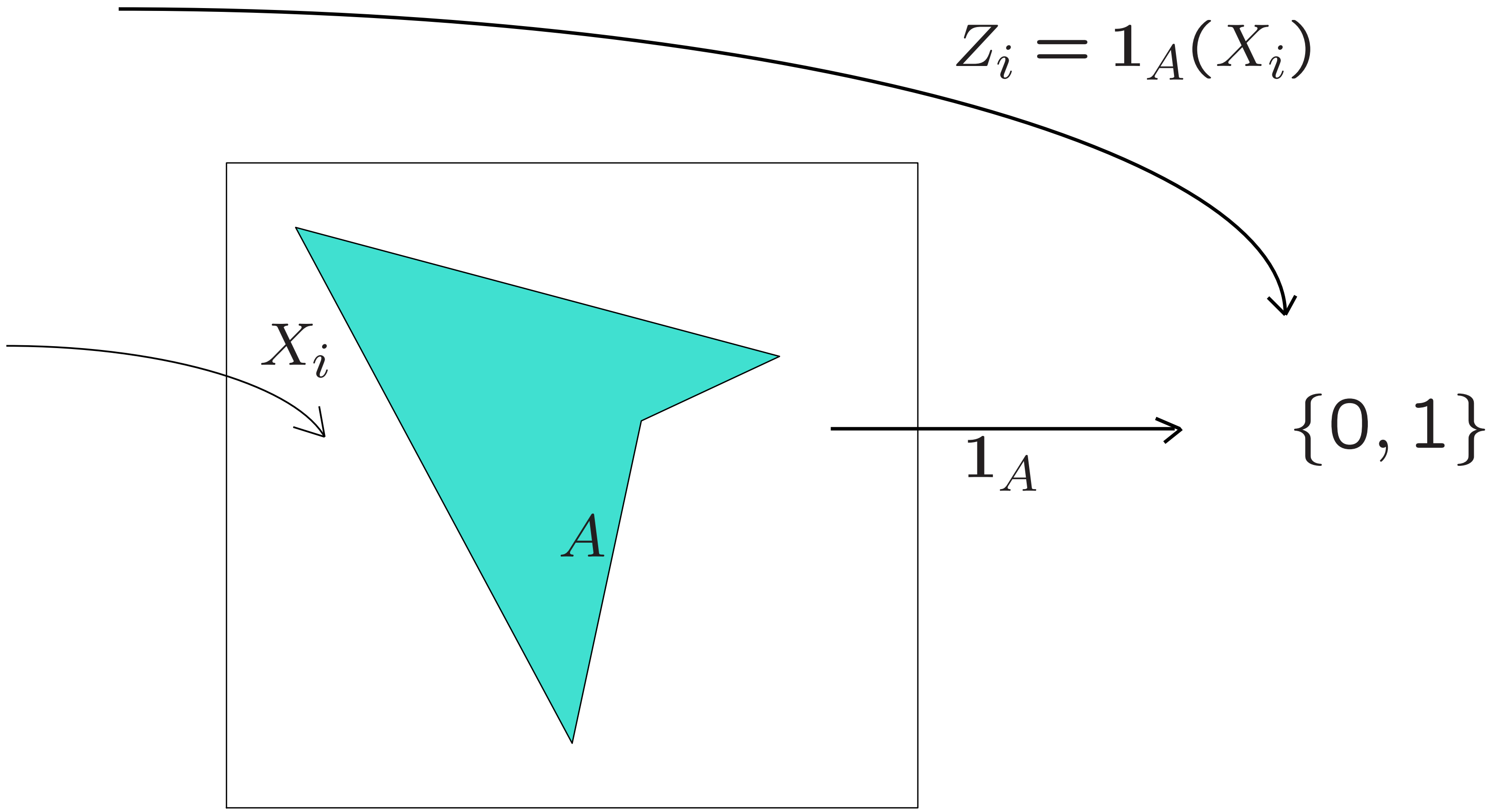
$$\mathbf{1}_A(x) = \begin{cases} 1 \text{ für } x \in A \\ 0 \text{ für } x \in S \setminus A \end{cases}$$

die Indikatorfunktion der Menge A.

Mit dieser Abbildung von S nach $\{0, 1\}$

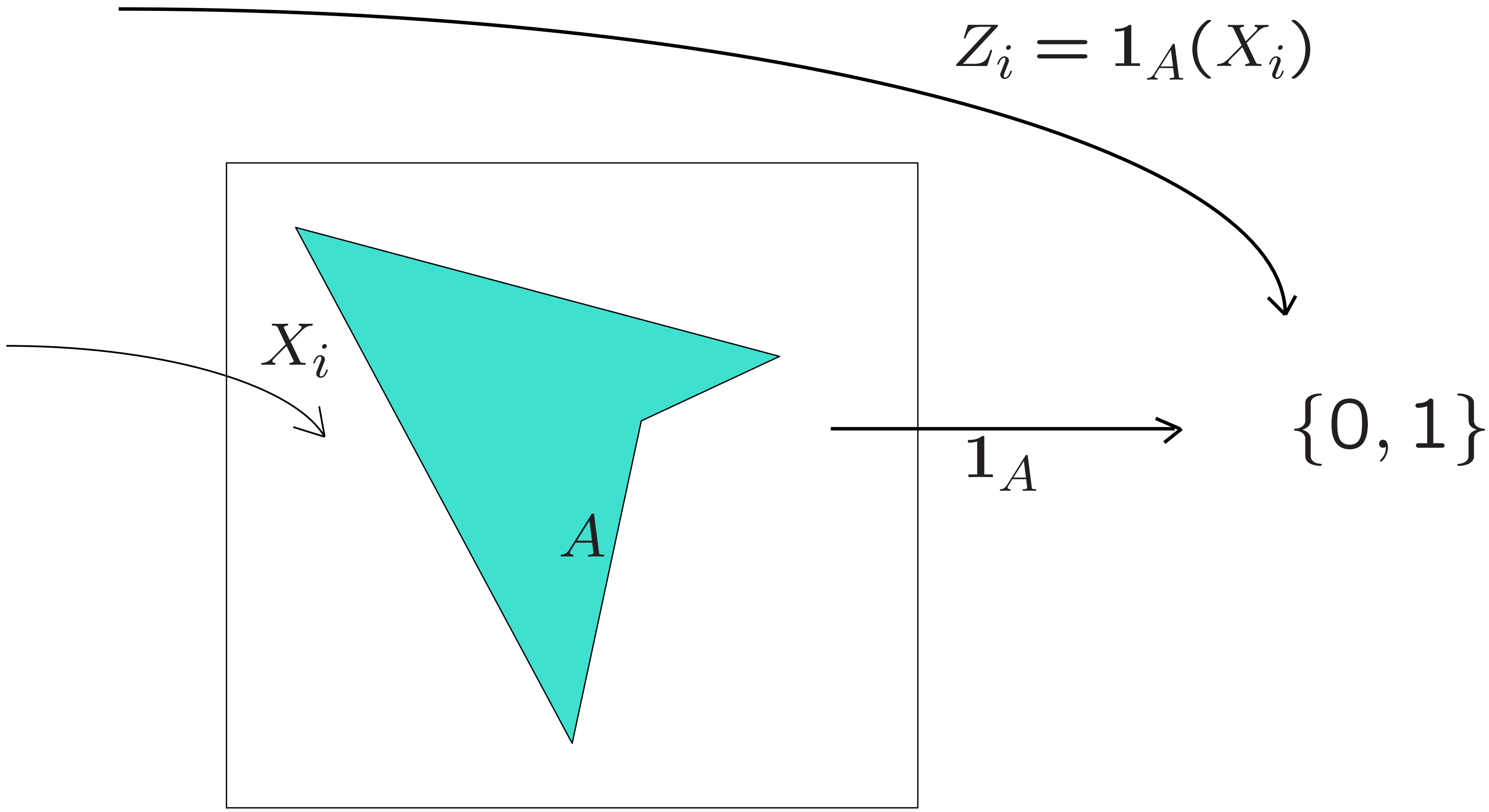
wird die S-wertige Zufallsvariable X_i

zur $\{0, 1\}$-wertigen Zufallsvariablen Z_i “verarbeitet”.



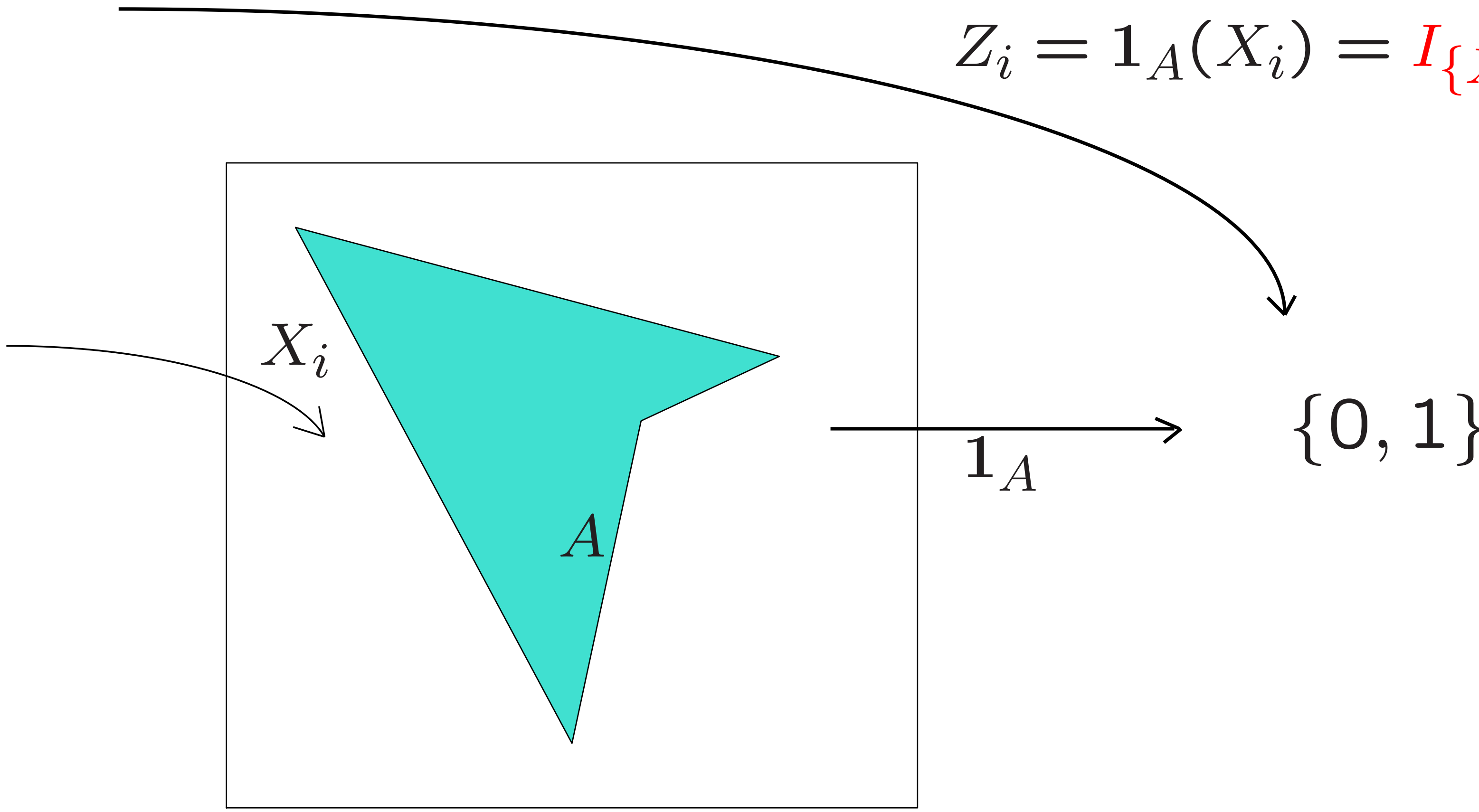
$Z_i = \mathbf{1}_A(X_i)$ ist der *der Indikator* (*“der Zähler”*)

des Ereignisses “X_i fällt in A”.

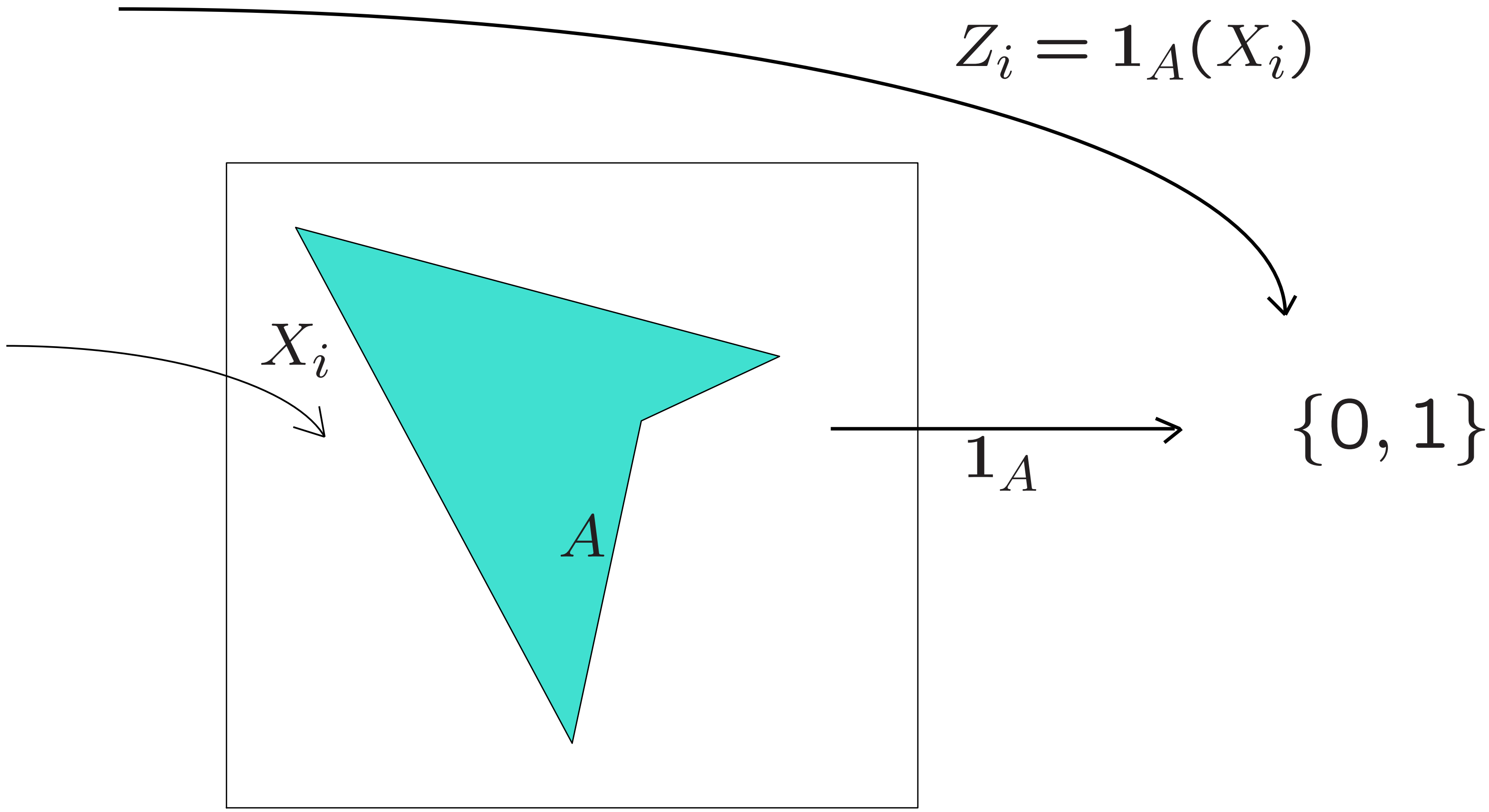


Eine alternative Schreibweise für $\mathbf{1}_A(X_i)$ ist $I_{\{X_i \in A\}}$

$$Z_i = \mathbf{1}_A(X_i) = I_{\{X_i \in A\}}$$



Eine alternative Schreibweise für $\mathbf{1}_A(X_i)$ ist $I_{\{X_i \in A\}}$



$\{X_i \in A\}$ und $\{Z_i = 1\}$ sind zwei Schreibweisen für

dasselbe Ereignis "X_i fällt in A"

5. Die Verteilung der zufälligen Trefferquote

Illustriert durch ein R-Programm

werden $n = 100$ Punkte

“rein zufällig” aus dem Quadrat gewählt.

$$Z_i = \mathbf{1}_A(X_i), \;\; i = 1, 2, \ldots$$

Die zufällige Zahl

$$M := \frac{1}{100}(Z_1 + \cdots + Z_{100})$$

(die “Trefferquote”)

ist ein *Schätzer* für die Wahrscheinlichkeit

$$p := \mathbf{P}(X \in A)$$

(und damit für den gefragten Flächenanteil).

Ein Ergebnis (“eine Realisierung”) von $(X_1, \ldots, X_{100})$

liefert eine Realisierung von $(Z_1, \ldots, Z_{100})$

und damit eine Realisierung von M

(einen Schätzwert für p).

Wie “zuverlässig” ist dieser Schätzwert?

$$M := \tfrac{1}{100}(Z_1 + \cdots + Z_{100})$$

Es sei an dieser Stelle verraten:

Der Anteil der blauen Fläche am Quadrat

(den man in der Realität ja nicht kennt) ist in unserem Beispiel

$$p = 0.195$$

Damit hat M gar keine Chance, exakt auf p zu fallen, denn

der *Wertebereich* von M

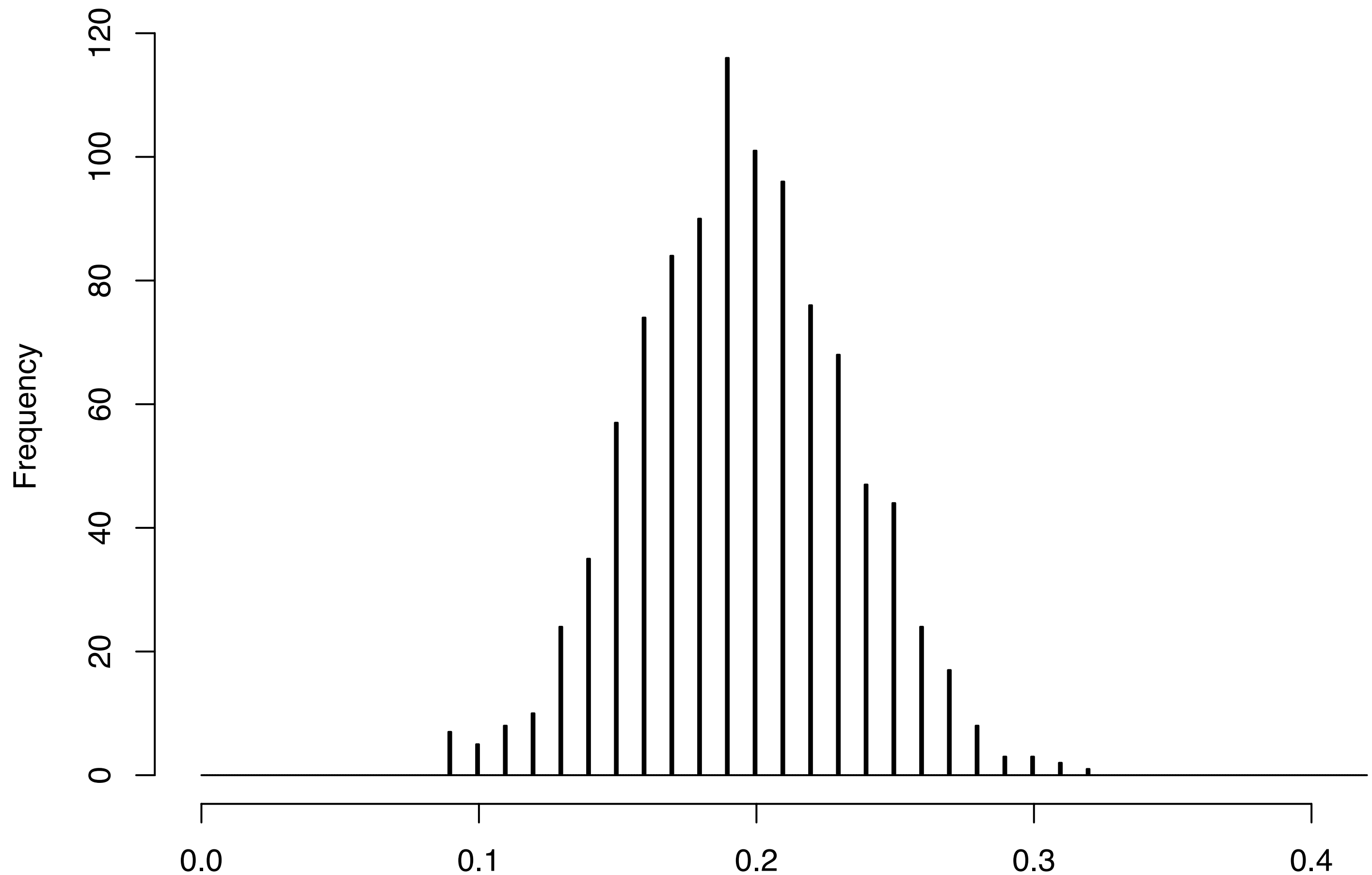
(d.h. die Menge der möglichen Ausgänge) ist

$$S' := \left\{ \frac{0}{100}, \frac{1}{100}, \ldots, \frac{99}{100}, \frac{100}{100} \right\}$$

Wie zuverlässig ist M als Schätzer für p?

Davon machen wir uns ein Bild, indem wir viele (z.B. 1000) “unabhängige Kopien” von M erzeugten und in einem *Histogramm* darstellen, wie oft welche Ausgänge realisert wurden.

Verteilung von M (Näherung aus 1000 Wiederholungen):



Wie zuverlässig ist M als Schätzer für p?

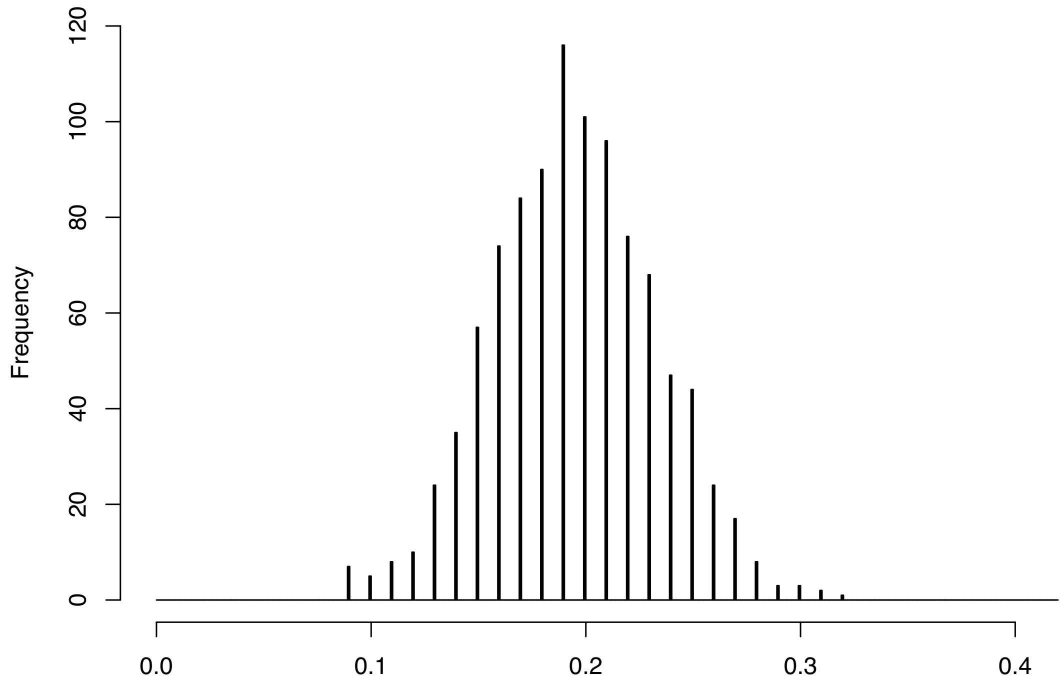
Davon machen wir uns ein Bild, indem wir viele (z.B. 1000) “unabhängige Kopien” von M erzeugten und in einem *Histogramm* darstellen, wie oft welche Ausgänge realisert wurden.

So bekommen wir eine näherungsweise Darstellung der Verteilung von M.

Die Verteilung von M ist bestimmt durch ihre Gewichte

$$\rho(b) := \mathbf{P}(M = b), \quad b \in S'.$$

Verteilung von M (Näherung aus 1000 Wiederholungen):



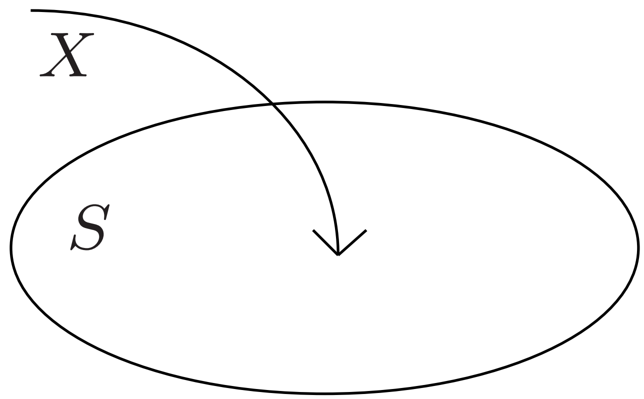
Die 101 möglichen Ausgänge von M

sind (bei weitem) nicht gleich wahrscheinlich:

die Verteilung von M “ist um p konzentriert”.

6. Zusammenfassung der wichtigsten Begriffe der ersten Stunde

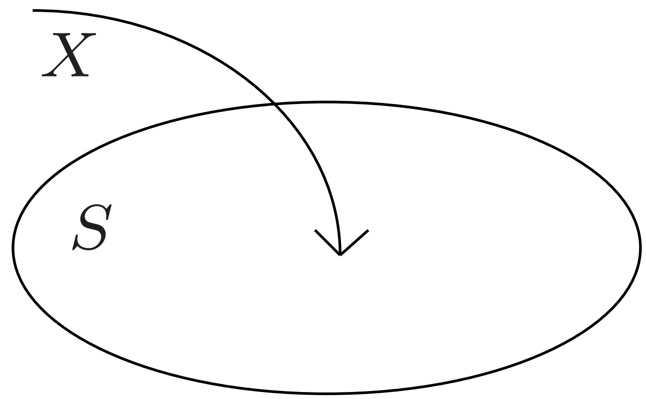
Ein Logo der Elementaren Stochastik:



X …zufällige Wahl eines Elements aus S

S …Menge von möglichen Ausgängen

Ein Logo der Elementaren Stochastik:

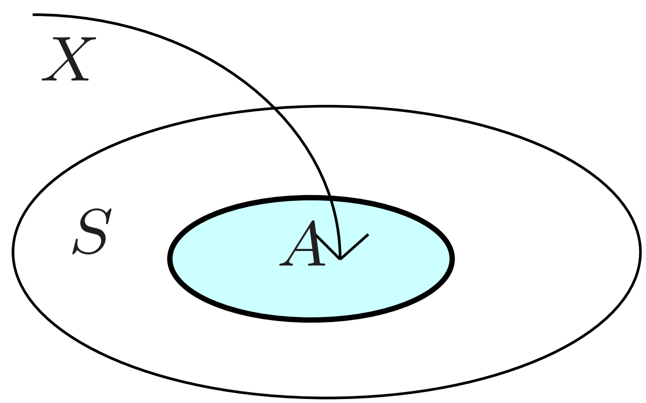


X ... **Zufallsvariable**

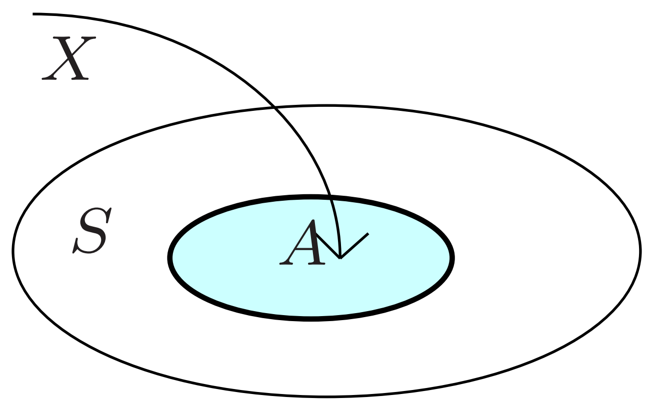
mit *Zielbereich (Wertebereich)* S

Wir interessieren uns für die *Wahrscheinlichkeit*

des *Ereignisses* “X fällt in A”



Dabei ist A eine bestimmte Teilmenge von S.



Ereignisse werden (wie Mengen)

in geschweiften Klammern notiert:

$$\{X \in A\}$$

Lies:

"X fällt in A".

“X ist *rein zufällig*”

heißt im Fall einer *endlichen* Menge S:

alle Elemente von S haben die gleiche W’keit

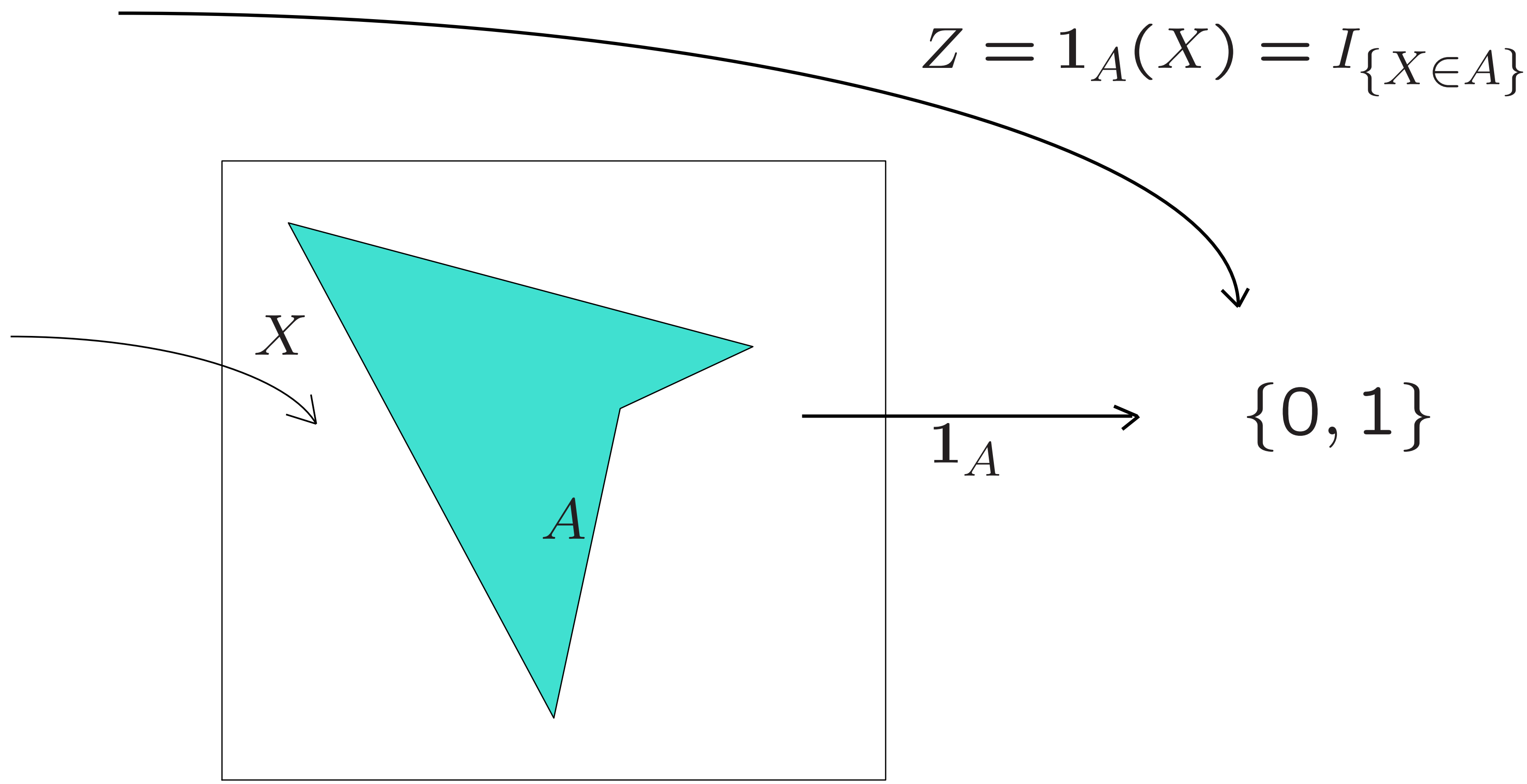
gewählt zu werden.

Dann ist die Wahrscheinlichkeit des Ereignisses “X fällt in A”:

$$\mathbf{P}(\{X \in A\}) = \frac{\#A}{\#S}.$$

Statt $\mathbf{P}(\{X \in A\})$ scheiben wir kurz:

$$\mathbf{P}(X \in A).$$



Die Ereignisse $\{X \in A\}$ und $\{Z = 1\}$ stimmen überein!

Ein Ausblick ins Kontinuum:

“X ist *rein zufällige Wahl aus S*”

heißt im Fall einer **kontinuierlichen** Menge $S \subset \mathbb{R}^d$:

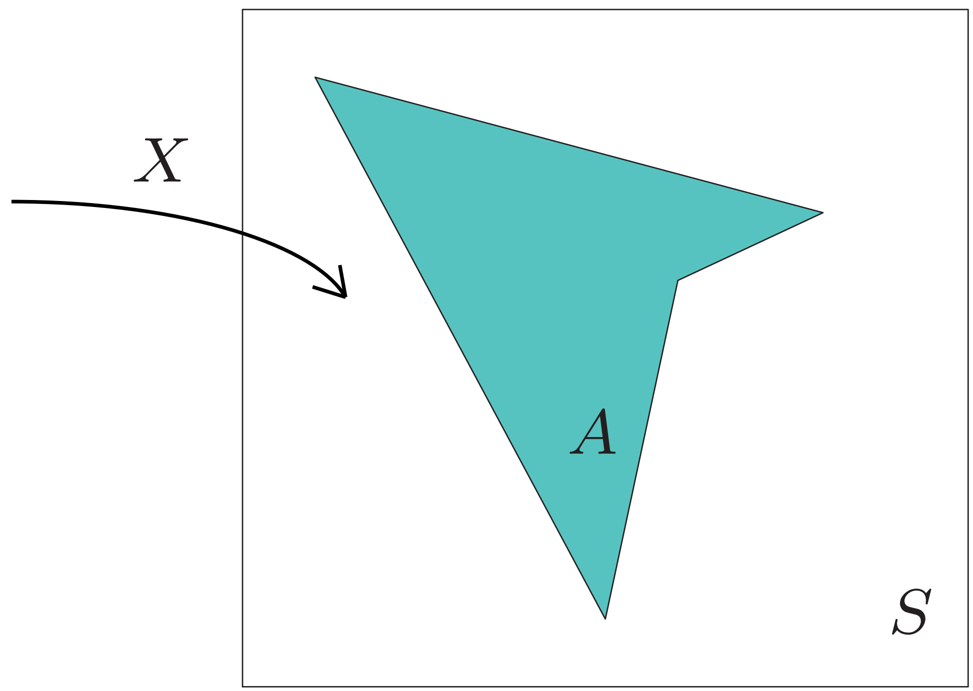
jede (messbare) Teilmenge A von S

kommt mit einer W’kt zum Zug,

die dem relativen Anteil ihres Volumens

am Volumen von S entspricht:

$$\mathbf{P}(X \in A) = \frac{\text{Volumen von } A}{\text{Volumen von } S}$$


X
A
S

Stochastik für die Informatik

Wintersemester 2020/21

Anton Wakolbinger

StofI-Webseite:

https://www.math.uni-frankfurt.de/

∼ismi/wakolbinger/teaching/StofI2021/

Oder:

google: “Wakolbinger”

$\rightarrow$ Goethe-Universität – Homepage A. Wakolbinger

$\rightarrow$ Veranstaltungen WS 2020/21

Dort: Link zum

Moodle-Kurs “Stochastik für die Informatik”

Anmeldung zu einer **Übungsgruppe:**

elektronisch ($\leftarrow$ Stofl - Webseite)

bis Donnerstag 05.11.2020, 24 Uhr

nach dem first come - first serve Prinzip

Gruppenzeiten:

Di 14-16, Di 16-18,

Mi 12-14, 12-14

Do 12-14 , Do 12-14, Do 14-16

Fr 8-10, Fr 10-12, Fr 14-16

Seit gestern (02.11.) sind alle 10 Gruppen (à 40) voll.

Wenn Sie sich noch zu den Übungen anmelden wollen, dann schreiben Sie bitte eine E-Mail an den Übungskoordinator: Herrn Florin Boenkost, boenkost@math.uni-frankfurt.de mit der Angabe von Ihnen möglichen Terminen.

Und bitte halten Sie Ihre Anmeldung nur dann aufrecht, wenn Sie wirklich an den Übungen teilnehmen wollen.

jeweils spätestens Dienstags:

Neues Übungsblatt im Moodle-Kurs

Tipps zu den Übungsaufgaben:

in den Tutorien

(und zu Blatt 1 auch schon in der Vorlesung diesen Freitag)

Termin für die elektronische Abgabe

der schriftlichen Lösungen der "S-Aufgaben":

Freitags (10 Tage nach Ausgabe des Blattes)

Leitfaden zur elektronischen Abgabe: siehe Moodle-Kurs.

In der Woche nach der Abgabe

werden die Lösungen in den Tutorien besprochen.

Abschlussklausur: Dienstag, 16 Februar 2021

Zweitklausur: Donnerstag 08. April 2021.

Beide Termine sind derzeit noch unter Vorbehalt.

In der Klausur können 100 Klausurpunkte erreicht werden.

Die Note errechnet sich aus

der Summe der Anzahl der erreichten Klausurpunkte

plus der Anzahl der im Tutorium erreichten Bonuspunkte.

Abschlussklausur: Dienstag, 16 Februar 2021

Zweitklausur: Donnerstag 08. April 2021.

Beide Termine sind derzeit noch unter Vorbehalt.

In der Klausur können 100 Klausurpunkte erreicht werden.

Die Note errechnet sich aus

der Summe der Anzahl der erreichten Klausurpunkte

plus der Anzahl der im Tutorium erreichten Bonuspunkte.

Beträgt diese Summe mindestens 45,
gilt die Abschlussprüfung über die Veranstaltung
als bestanden.

Bonuspunkte (maximal 12):

durch aktive Beteiligung in den Tutorien.

Bonuspunkte bekommt man nur, wenn man mindestes zweimal im Semester Lösungen von Übungsaufgaben (oder Teile davon) im Tutorium vorstellt,

und grundsätzlich nur für die Aufgaben, bei deren Lösungsbesprechung man im Tutorium anwesend ist.

Lehrbuch:

Götz Kersting, Anton Wakolbinger

Elementare Stochastik, Birkhäuser, 2. Aufl. 2010,

Preis: 19,99 EUR

Semesterausleihe möglich aus der

Bibliothek des Mathematischen Seminars,

Robert-Mayer-Str. 8, 4. Stock

in der UB als E-Book vorhanden

Vorlesung 1b

Wahrscheinlichkeit von Kollisionen

bei der wiederholten rein zufälligen Wahl

aus endlich vielen möglichen Ausgängen

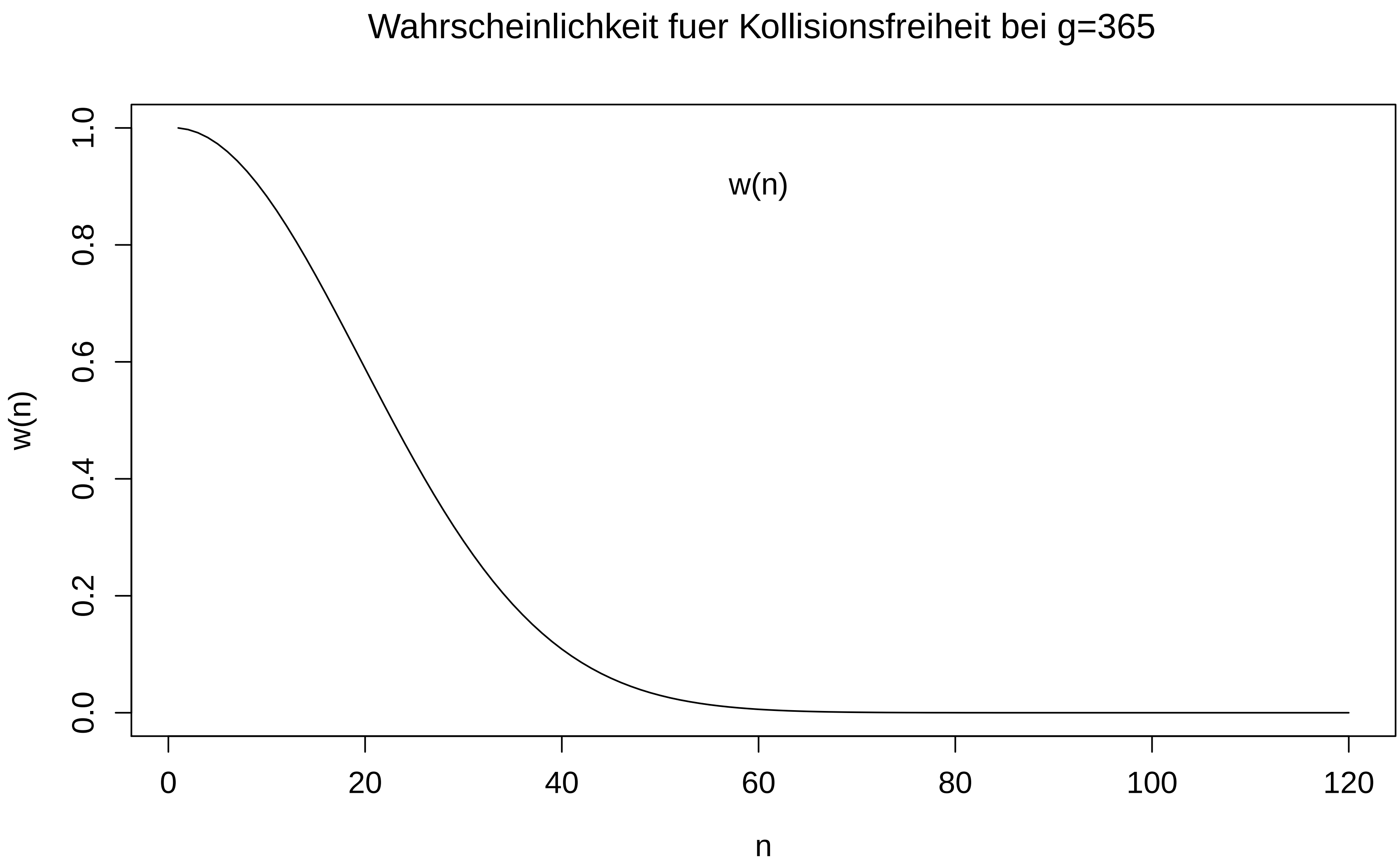
(Buch S. 1-5)

Teil Zwei:

Approximationen

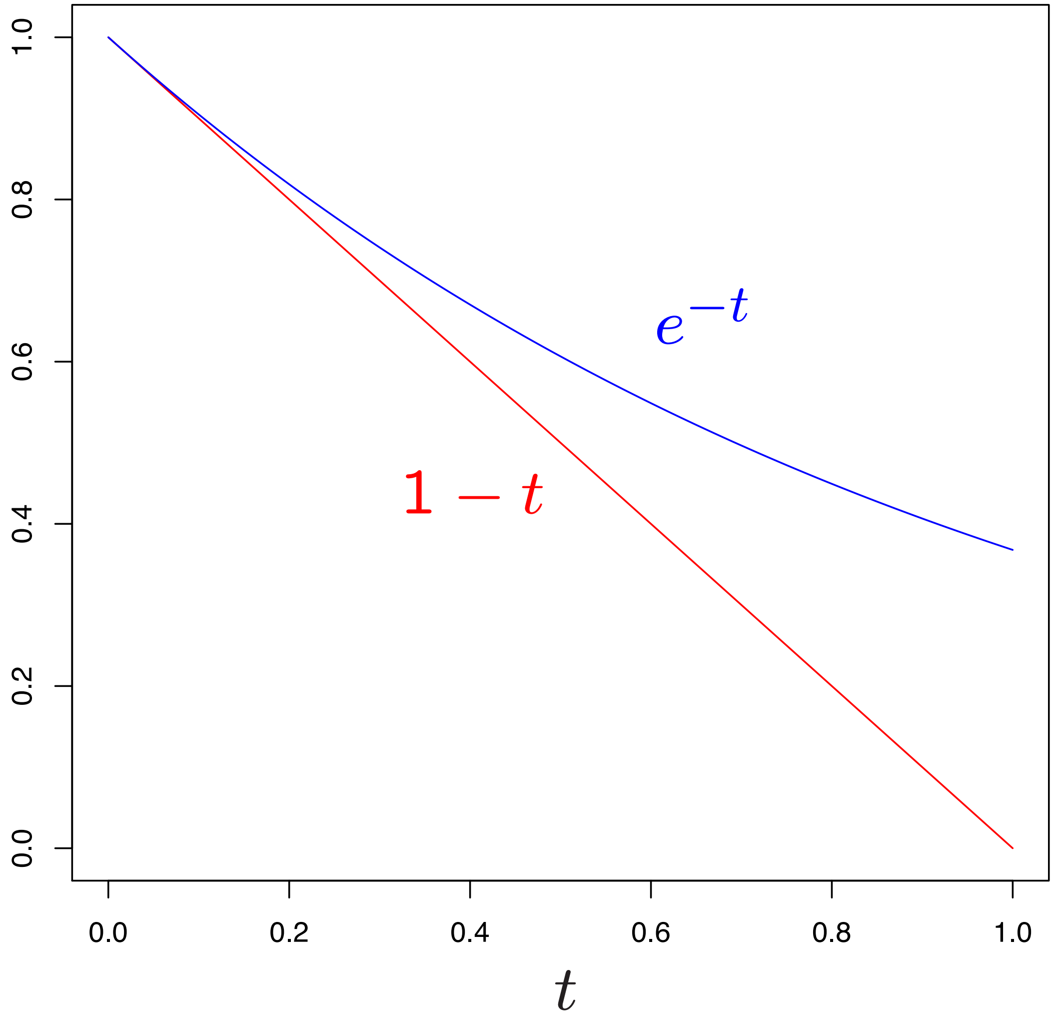
der Wahrscheinlichkeit für Kollisionsfreiheit

$$w(n, g) = \prod_{i=1}^{n-1} \left(1 - \frac{i}{g}\right)$$

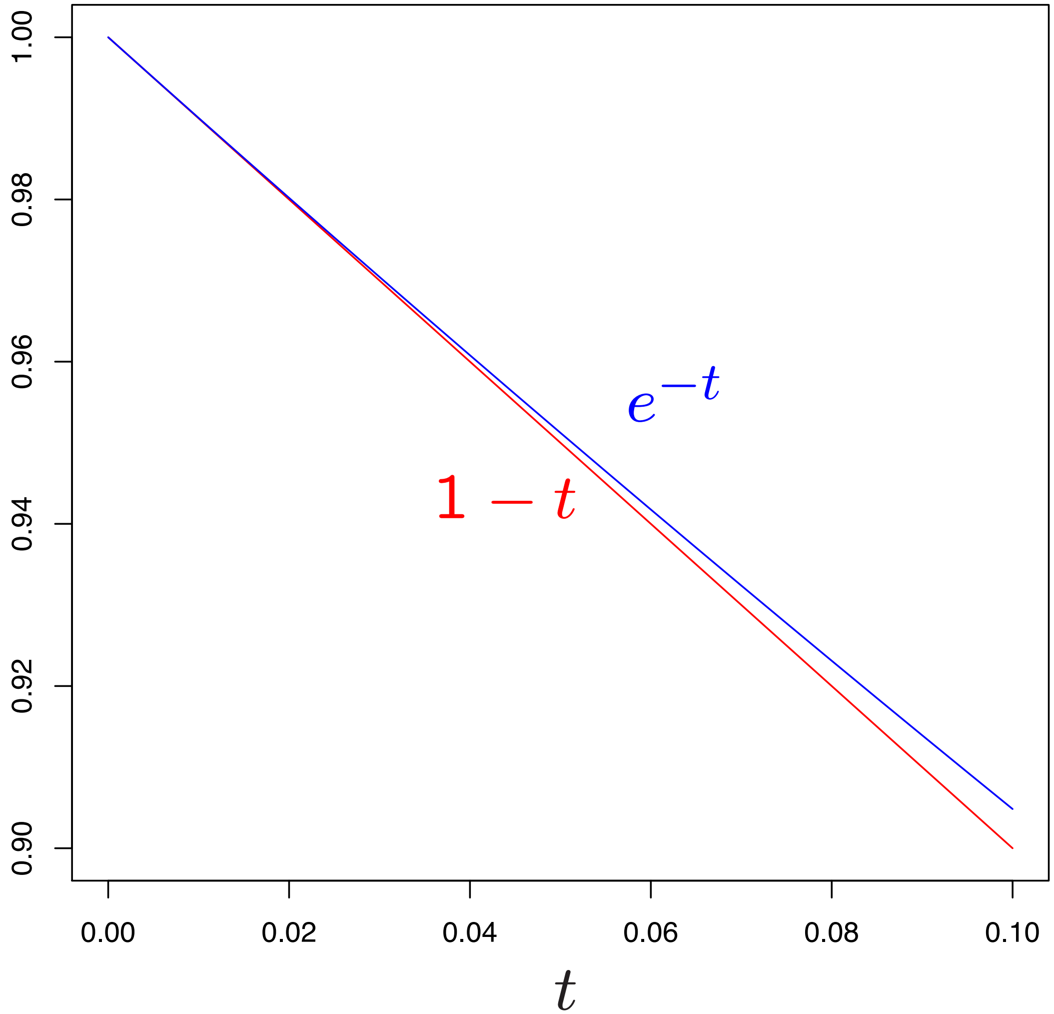

Wahrscheinlichkeit fuer Kollisionsfreiheit bei g=365
w(n)
w(n)
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
120
n

1. Approximation über die Linearisierung von exp

für $0 \le t \le 1$:



für $0 \le t \le 0.1$:



Lokale lineare Approximation von $t \mapsto e^{-t}$ bei $t = 0$:

$$e^{-t} = 1 - t + o(t) \qquad \text{für } t \to 0.$$

Dabei ist der Term $o(t)$ *von kleinerer Ordnung als* t, d.h. $\lim_{t \to 0} \frac{o(t)}{t} = 0$.

Salopp geschrieben:

$$e^{-t} \approx 1 - t \qquad \text{für kleine } t.$$

Damit bekommen wir für kleines $\frac{n}{g}$ die Approximation

$$w(n,g) = \prod_{i=1}^{n-1} \left(1 - \frac{i}{g}\right) \approx \prod_{i=1}^{n-1} \exp\left(-\frac{i}{g}\right)$$

$$= \exp\left(-\sum_{i=1}^{n-1} \frac{i}{g}\right) = \exp\left(-\frac{(n-1)n}{2g}\right).$$

Also für kleines $\frac{n}{g}$, d.h. für $n \ll g$:

$$w(n,g) \approx \exp\left(-\frac{(n-1)n}{2g}\right) \; .$$

Gibt es auch eine Näherungsformel,

die brauchbar ist für alle $n < g$?

In der Tat!

2. Die Stirling-Approximation

(Buch S. 4-5)

$$w(n,g) = \frac{g(g-1)\cdots(g-(n-1))}{g^n}$$

$$= \frac{g(g-1)\cdots(g-(n-1))}{g^n}\frac{(g-n)(g-n-1)\cdots 1}{(g-n)(g-n-1)\cdots 1}$$

$$w(n,g) = \frac{g!}{g^n(g-n)!}$$

mit $k! := 1 \cdot 2 \cdots k$,   lies: k-Fakultät

Die Stirling-Formel:

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^k$$

$$k! \approx \sqrt{2\pi\, k}\ \left(\frac{k}{e}\right)^{k}$$

$$\frac{g!}{g^n (g-n)!} \approx \frac{1}{g^n}\sqrt{\frac{g}{g-n}}\ \frac{g^g}{(g-n)^{g-n}}\ \frac{e^{g-n}}{e^g}$$

$$= \sqrt{\frac{g}{g-n}}\, e^{-n} \left(\frac{g}{g-n}\right)^{g-n}$$

$$= \frac{1}{\sqrt{1-\frac{n}{g}}}\, e^{-n} \left(1 - \frac{n}{g}\right)^{n-g}$$

$$\frac{g!}{g^n(g-n)!} \approx \frac{1}{\sqrt{1-\frac{n}{g}}}\, e^{-n} \left(1-\frac{n}{g}\right)^{n-g}$$

Wir formen jetzt den Ausdruck

$$e^{-n} \left(1-\frac{n}{g}\right)^{n-g}$$

weiter um. Ein Trick dabei ist es,

$$\left(1-\frac{n}{g}\right)^{n-g} \quad \text{als} \quad \exp\left((n-g)\ln\left(1-\frac{n}{g}\right)\right)$$

zu schreiben

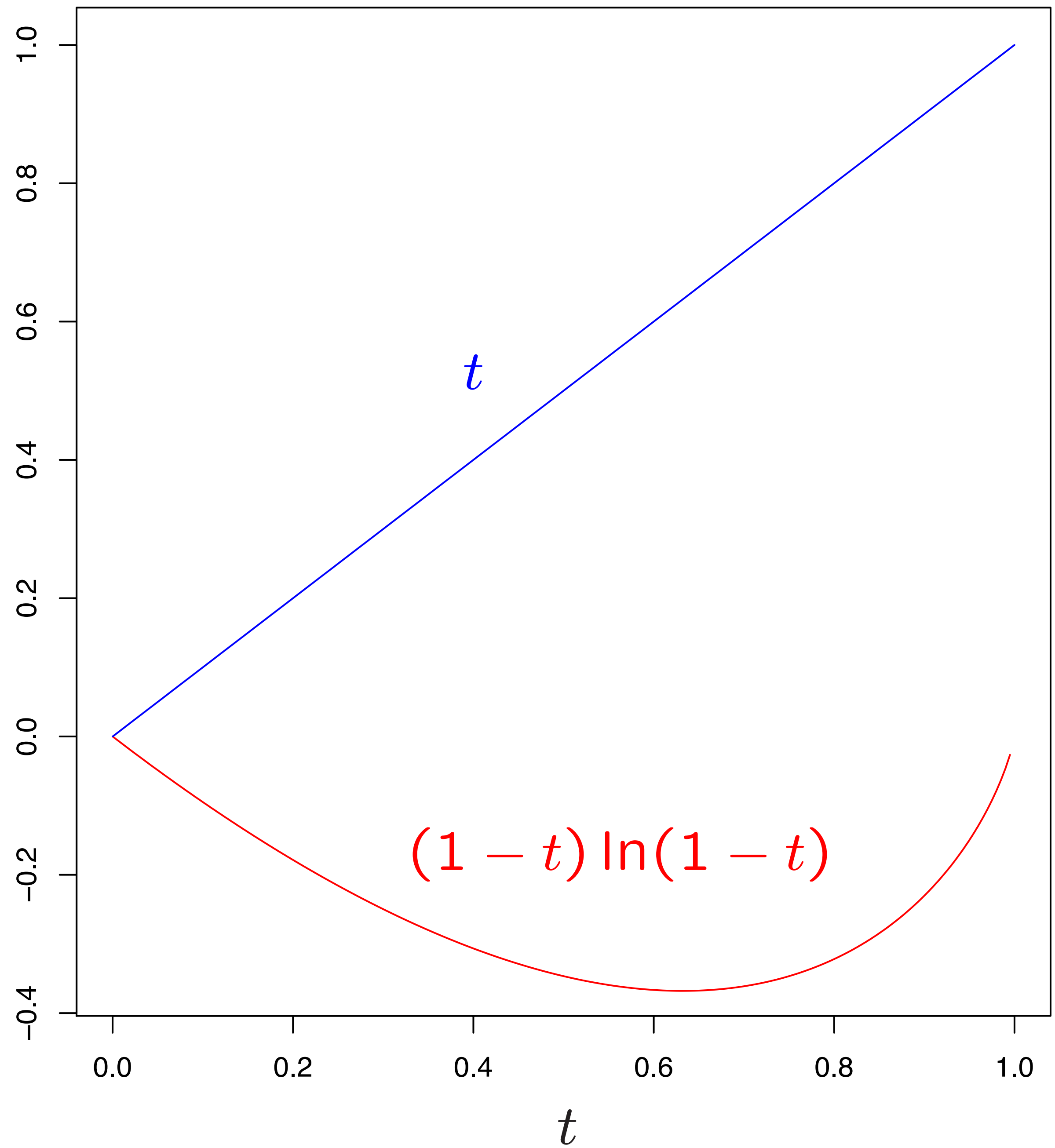
und dann die Beziehung $e^a e^b = e^{a+b}$ zu verwenden:

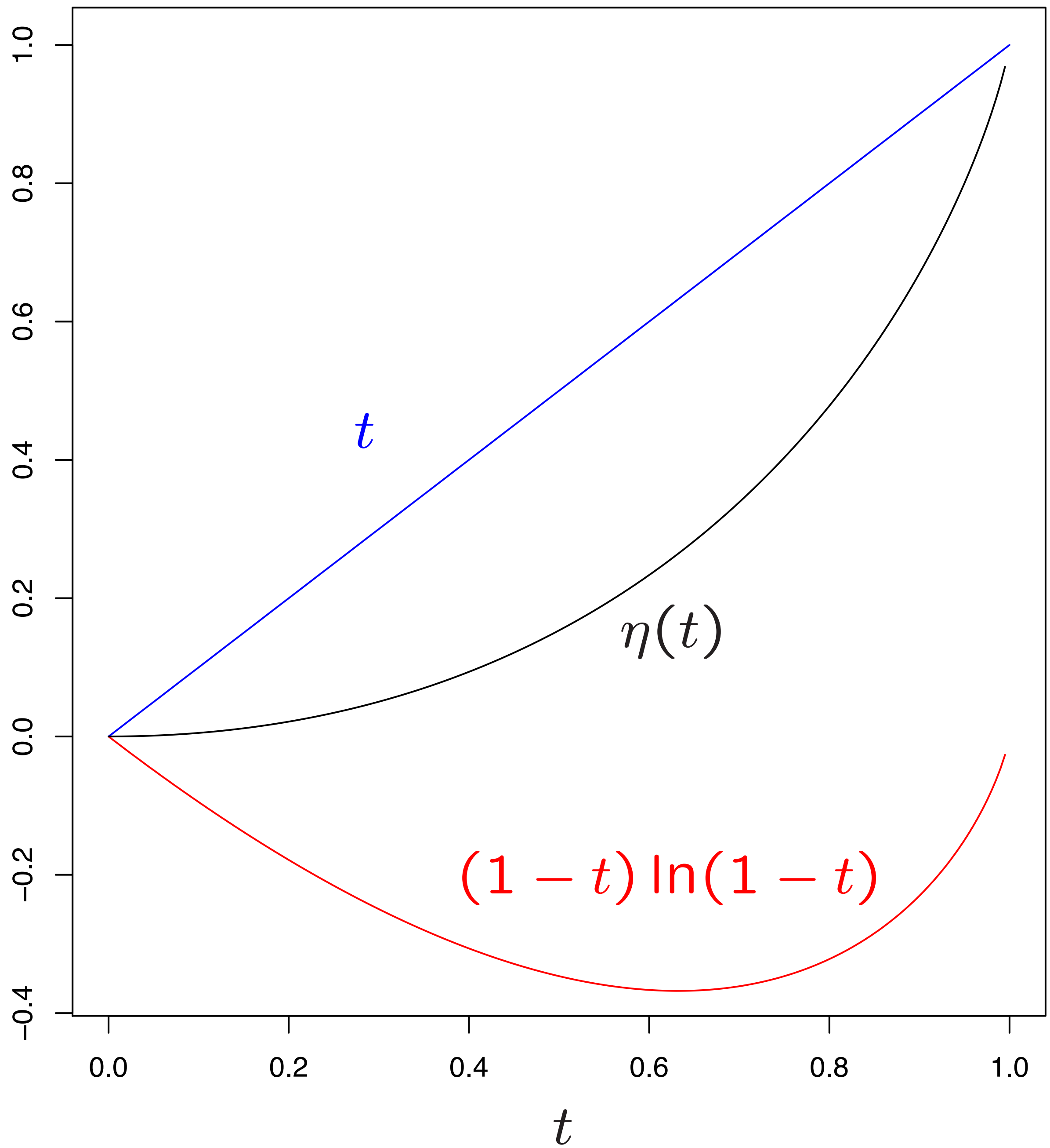
$$e^{-n}\left(1-\frac{n}{g}\right)^{n-g} = \exp\left(-n + (n-g)\ln\left(1-\frac{n}{g}\right)\right)$$

$$= \exp\left(-g\left(\frac{n}{g} + \left(1-\frac{n}{g}\right)\ln\left(1-\frac{n}{g}\right)\right)\right)$$

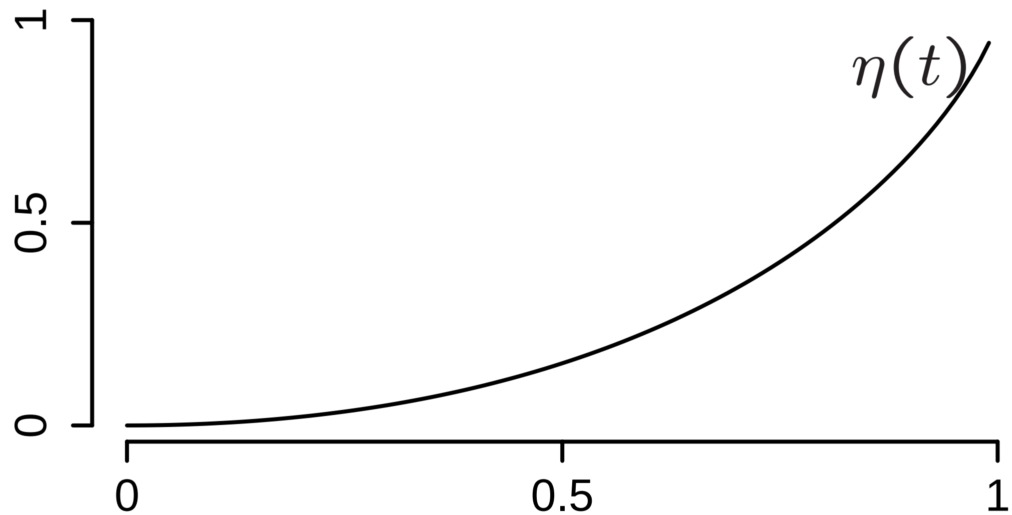
$$= \exp\left(-g\,\eta\left(\frac{n}{g}\right)\right)$$

mit $\quad \eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0
t
t
$(1-t)\ln(1-t)$


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0
t
$\eta(t)$
$(1-t)\ln(1-t)$
t

$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$

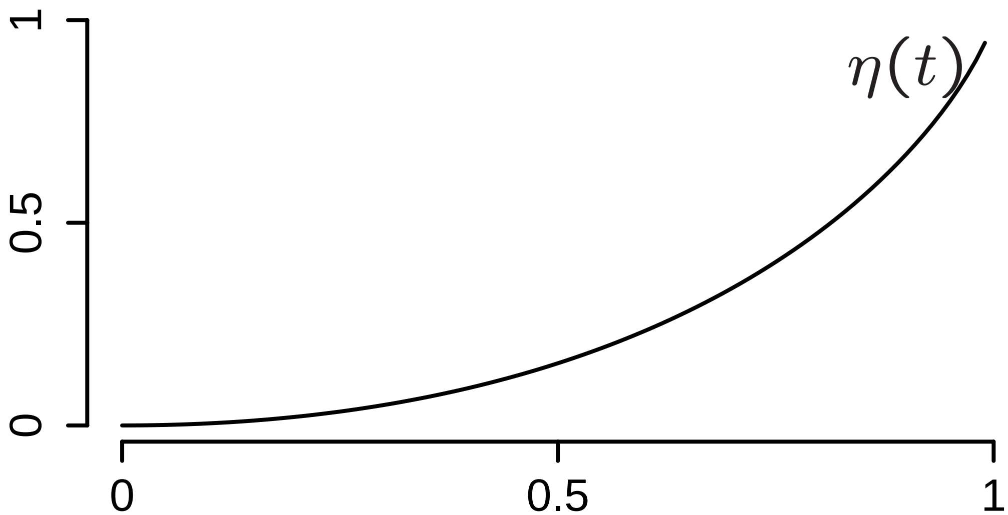


$$\mathbf{P}(X \in A) \approx \frac{1}{\sqrt{1-\frac{n}{g}}}\exp\left(-g\,\eta\left(\frac{n}{g}\right)\right)$$

Stirling-Approximation

der Wahrscheinlichkeit für Kollisionsfreiheit

$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$



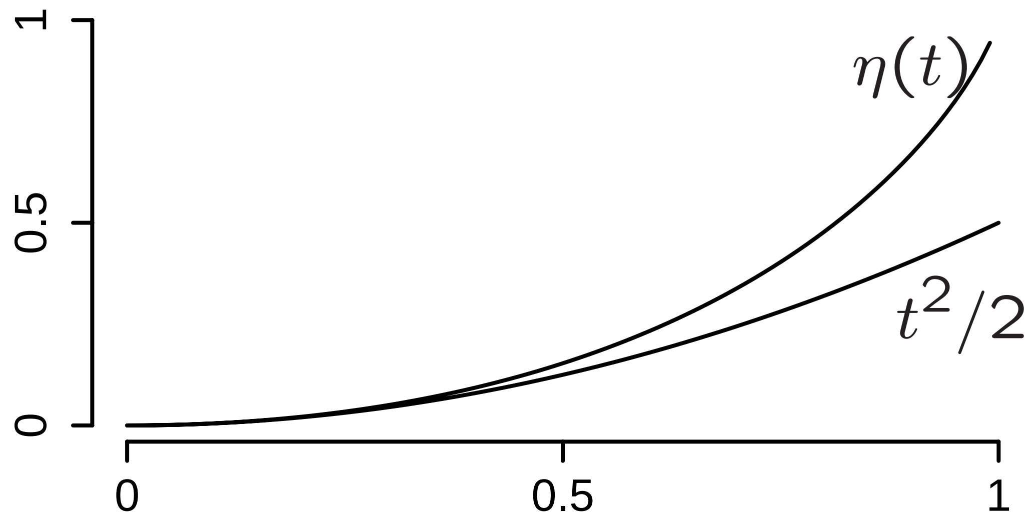
$$w(n,g) \approx \frac{1}{\sqrt{1-\frac{n}{g}}} \exp\left(-g\,\eta\left(\frac{n}{g}\right)\right)$$

Die Folien 19 bis 23 wurden in der Vorlesung nicht besprochen. Ihre Lektüre ist für das Weitere nicht notwendig, kann aber (auch in der Remeinszenz an Mathe 1 und 2) genüsslich sein.

Für $n \ll g$, also $t := \frac{n}{g} \ll 1$,
können wir

(i) $\eta(t)$ quadratisch um $t = 0$ approximieren
und
(ii) den Vorfaktor $\frac{1}{\sqrt{1-\frac{n}{g}}}$ mittels $e^{-t} \approx 1 - t$ approximieren:

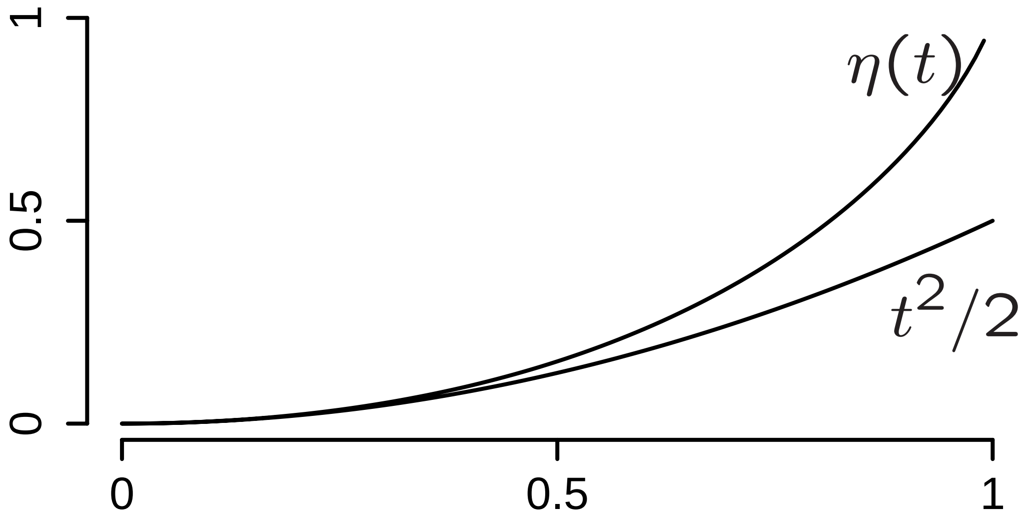
$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$



$$\ln(1-t) = -t - \frac{t^2}{2} - \frac{t^3}{3} - \cdots$$

$$(1-t)\ln(1-t) = -t + \frac{t^2}{2} + o(t^2) \text{ für } t \to 0$$

$$\eta(t) = \frac{t^2}{2} + o(t^2) \text{ für } t \to 0$$



$\frac{1}{2}t^2$ ist die quadratische Approximation von $\eta(t)$ um $t = 0$.

Für $n \ll g$ (wie z. B. für $n = 25,\ g = 365$) ist also

$$\eta\left(\frac{n}{g}\right) \approx \frac{1}{2} \cdot \left(\frac{n}{g}\right)^2.$$

Jetzt zum Vorfaktor $\frac{1}{\sqrt{1-\frac{n}{g}}}$:

$$\frac{1}{\sqrt{1-t}} = (1-t)^{-1/2} \approx (e^{-t})^{-1/2} = e^{t/2}.$$

Insgesamt ist also für kleine t

$$\frac{1}{\sqrt{1-t}} e^{-g\eta(t)} \approx e^{t/2} e^{-gt^2/2}$$

Für $t := \frac{n}{g} \ll 1$ ergibt sich die Stirling+Taylor-Approximation:

$$w(n,g) \approx \exp\left(\frac{n}{2g}\right) \exp\left(-\frac{n^2}{2g}\right) = \exp\left(-\frac{n(n-1)}{2g}\right)$$

Man beachte:

Die Stirling+Taylor Approximation

$$\mathbf{P}(X \in A) \approx \exp\left(-\frac{n(n-1)}{2g}\right)$$

ist identisch mit der

Approximation über die Linearisierung von exp.

Im Buch haben wir hier den Faktor $\sqrt{1 - n/g}$
in der Stirling-Approximation vernachlässigt, d.h. gleich 1 gesetzt.
Dadurch erhielten wir im Buch die etwas weniger genaue Näherung

$$w(n, g) \approx \exp\left(-\frac{n^2}{2g}\right).$$

Fazit:

Für $n < g$ ist $\quad w(n,g) \approx \frac{1}{\sqrt{1-\frac{n}{g}}} \exp\left(-g\,\eta\left(\frac{n}{g}\right)\right)$

mit $\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$

(Stirling-Approximation).

Für $n \ll g$ ist $\quad w(n,g) \approx \exp\left(-\frac{n(n-1)}{2g}\right)$

(Approximation über die Linearisierung von exp)

Für $n^2 \ll g$ ist $\quad w(n,g) \approx 1.$

Beispiel: $n = 25, g = 365$:

$$\mathbf{P}(X \in A) = \frac{g(g-1)\cdots(g-n+1)}{g^n} = 0.431300$$

$$\frac{1}{\sqrt{1-\frac{n}{g}}} \exp\left(-g\,\eta\left(\frac{n}{g}\right)\right) = 0.431308$$

$$\exp\left(-\frac{n(n-1)}{2g}\right) = 0.4396$$

Vorlesung 1b

Wahrscheinlichkeit von Kollisionen

bei der wiederholten rein zufälligen Wahl

aus endlich vielen möglichen Ausgängen

(Buch S. 1-5)

Teil Eins:

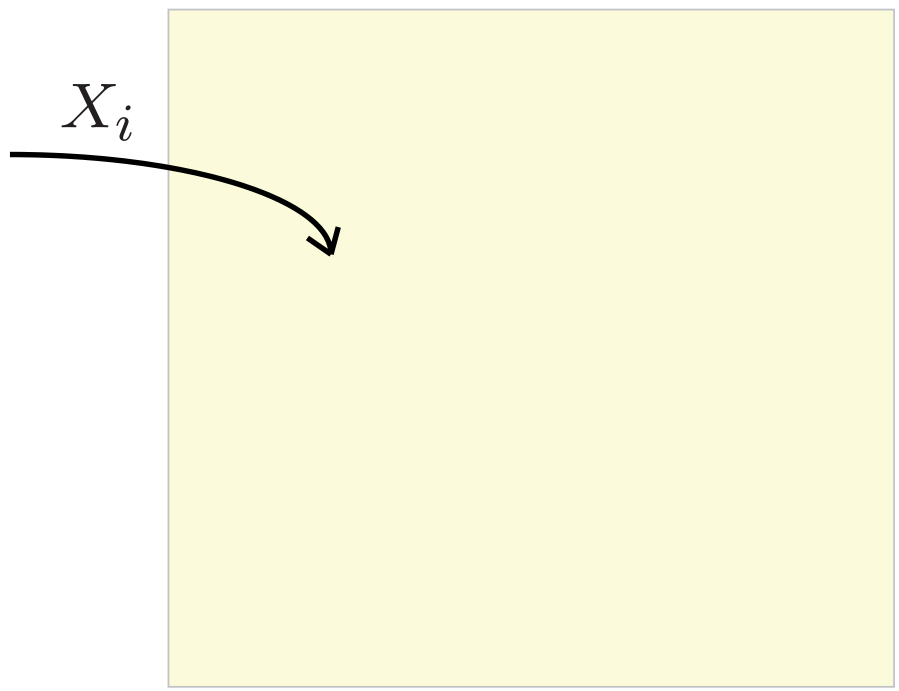
Vom Modell zur Formel

1. Die Fragestellung

– in verschiedenen Verpackungen

n-mal wiederholt wird rein zufällig

ein Pixel aus $g = 10^6$ Pixeln gewählt.



$$i = 1, \ldots, n$$

n-mal wiederholt wird rein zufällig

ein Pixel aus $g = 10^6$ Pixeln gewählt.

Wie wahrscheinlich ist es, dass dabei

lauter verschiedene Pixel gewählt werden?

Ziemlich unwahrscheinlich

oder ziemlich hoch wahrscheinlich,

oder so “mittendrin”, für

$n = 100$?

$n = 1000$?

$n = 10000$?

Was schätzen Sie?

Bei der “wiederholten rein zufälligen Pixelwahl”

handelt es sich um ein

Ziehen mit Zurücklegen.

Man kann auch denken an eine Urne

mit g Kugeln, nummeriert mit $1, 2, \ldots, g$.

Es wird n-mal gezogen.

Nach jedem Zug wird die gezogene Kugel zurückgelegt,

vor jedem Zug wird “perfekt durchmischt”.

Tema con variazioni:

n Objekte,

g Plätze.

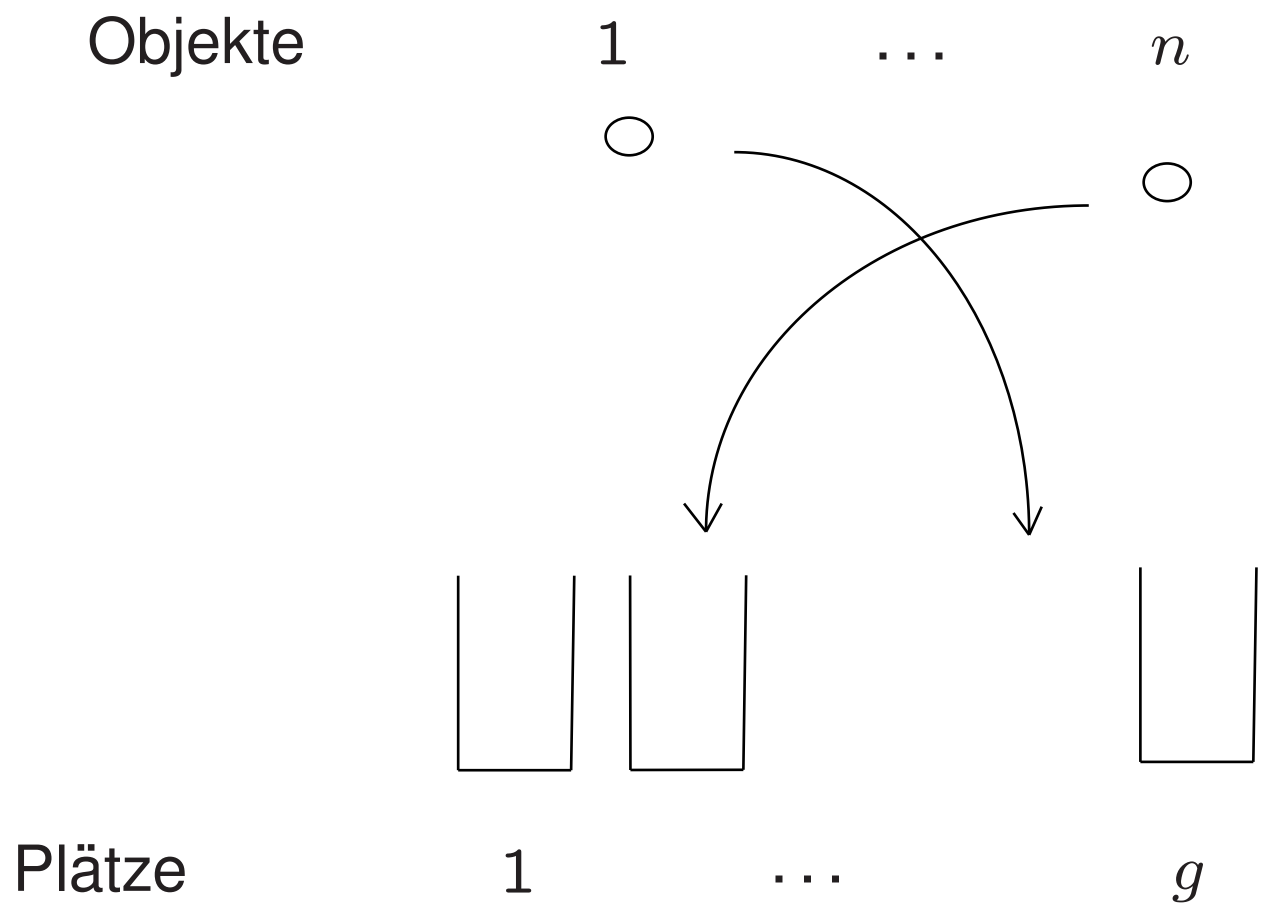
Jedes Objekt wird auf einen

rein zufällig ausgewählten Platz gesetzt.

(Mehrfachbelegungen sind erlaubt!)

Wie wahrscheinlich ist es, dass dabei

keine Mehrfachbelegung ("Kollision") auftritt?



Anders als im Buch wird hier die Anzahl der Kennzeichen mit g (und nicht mit r) bezeichnet - wegen des besseren typographischen Unterschieds von g zu n und der Assoziation von g mit **G**esamtzahl der Plätze, **G**eburtstage,...

Populäre Version:

$n = 25$ Personen auf einer Party

Platz $\longleftrightarrow$ Geburtstag $(\in \{1, 2, \ldots, 365\})$

Wie wahrscheinlich ist es,
dass keine zwei Personen am selben Tag Geburtstag haben?

2. Beschreibung durch eine Zufallsvariable und ein Ereignis.

(Buch S. 2-3)

Die Objekte denken wir uns mit 1 bis n

und die Plätze mit 1 bis g nummeriert.

Ein Ausgang der Platzwahl (ein "Wahlprotokoll")

lässt sich beschreiben durch das n-tupel

$$a = (a_1, \ldots, a_n) ,$$

wobei a_i den für das i-te Objekt gewählte Platz bezeichnet

$(1 \leq a_i \leq g)$.

Die Menge der möglichen Wahlprotokolle ist

$$S := \{1, \dots, g\}^n \,,$$

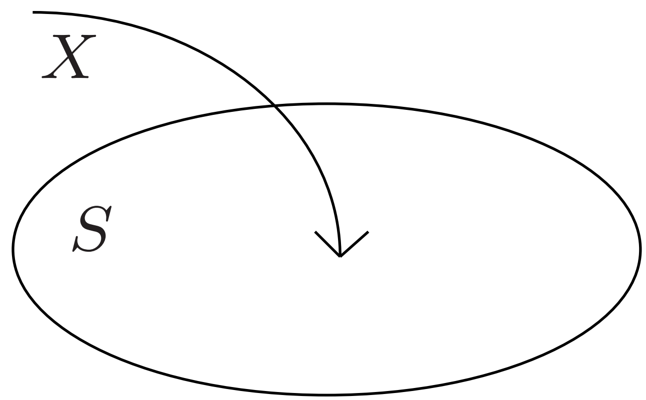
die Menge aller n-tupel $(a_1, \dots, a_n)$

mit Einträgen (Komponenten)

$$a_i \in \{1, \dots, g\}, \quad i = 1, \dots, n.$$

Die zufällige Platzwahl

beschreiben wir durch eine S*-wertige Zufallsvariable* X.



X kommt durch zufällige Wahl

eines Elementes aus S zustande.

Die Menge S heißt *Zielbereich* (oder *Wertebereich*)

der Zufallsvariable X.

Wie jedes Element $(a_1, \ldots, a_n)$ unserer Menge S

besteht auch die Zufallsvariable X aus n Komponenten:

$$X = (X_1, \ldots, X_n) \; .$$

Wir interessieren uns für das *Ereignis*,

dass keine zwei Komponenten von X gleich sind.

Dieses Ereignis schreiben wir als

$$\{X_i \neq X_j \text{ für alle } i \neq j\}$$

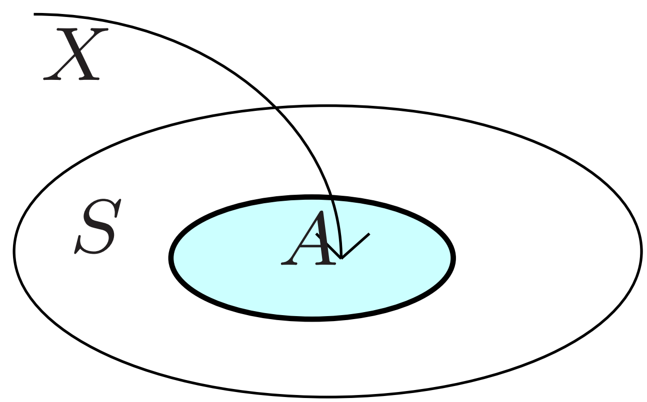
oder auch als

$$\{X \in A\}$$

mit der Teilmenge

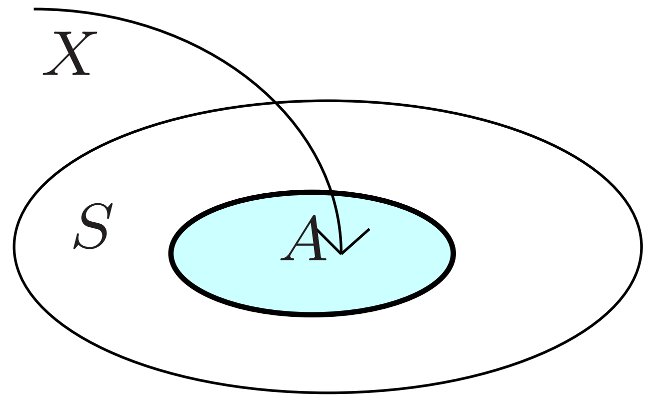
$$A := \{a \in S \; : \; a_i \neq a_j \text{ für alle } i \neq j\} \; .$$

Ein Logo für das Ereignis $\{X \in A\}$:



3. Die Wahrscheinlichkeit für Kollisionsfreiheit

(Buch S. 3)



Wie kommt man zur **Wahrscheinlichkeit**

des Ereignisses $\{X \in A\}$?

Zum Merken:

Wahrscheinlichkeiten gehören zu Ereignissen.

Die Wahrscheinlichkeit eines Ereignisses

misst dessen Chance einzutreten

mit einer Zahl zwischen 0 und 1:

$$\mathbf{P}(\{X \in A\})$$

Statt $\mathbf{P}(\{X \in A\})$

schreiben wir kurz

$$\mathbf{P}(X \in A)$$

Statt $\mathbf{P}(X \in \{a\})$

schreiben wir kurz

$$\mathbf{P}(X = a)$$

Eine einleuchtende Regel

für das Rechnen mit Wahrscheinlichkeiten ist die

Additivität:

Für disjunkte $A,\ A' \subset S$ gilt

$$\mathbf{P}(X \in A \cup A') = \mathbf{P}(X \in A) + \mathbf{P}(X \in A').$$

Daraus folgt im Fall endlich vieler möglicher Ausgänge,

d.h. für $\#S < \infty$:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a)$$

Um die Wahrscheinlichkeit $\mathbf{P}(X \in A)$ berechnen zu können,

muss man eine Modellannahme treffen.

Eine prominente Modellannahme ist die einer

rein zufälligen Wahl.

Damit ist gemeint, dass für je zwei $a, a' \in S$

$$\mathbf{P}(X = a) = \mathbf{P}(X = a').$$

Das heißt: kein Ausgang ist bevorzugt.

Gemäß der Additivität gilt:

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

Und wir hatten schon vereinbart: $\mathbf{P}(X \in S) = 1$

Für eine rein zufällige Wahl folgt daraus sofort:

$$P(X = a) = \frac{1}{\#S}\,, \quad a \in S\,.$$

Und

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}\,.$$

Wir haben nun die Aufgabe des Abzählens der zwei Mengen

$$S := \{1, \ldots, g\}^n = \{(a_1, \ldots, a_n) : 1 \leq a_i \leq g\}$$

und

$$A := \{a \in S \,:\, a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#S = g^n$$

$$S = \{(a_1, \ldots, a_n) : 1 \leq a_i \leq g\}$$

$$A := \{a \in S \ : \ a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#A = ?$$

Für a_1 gibt es g mögliche Werte, für a_2 dann noch $g - 1$, usw. Also:

$$\#A = g(g-1)\cdots(g-(n-1))$$

$$\mathbf{P}(X \in A) = \frac{g(g-1)\cdots(g-(n-1))}{g^n}$$

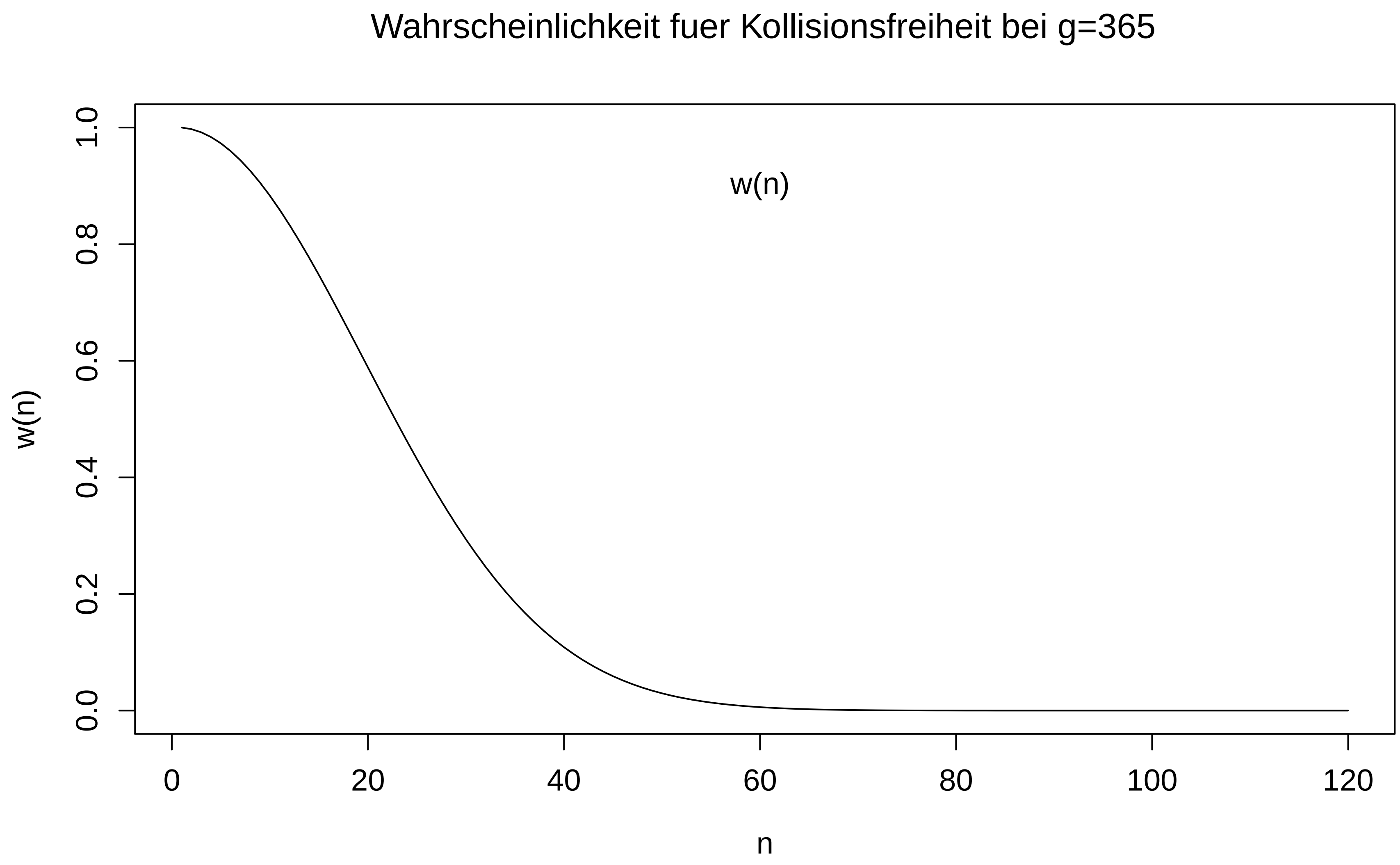
$$= \frac{g-1}{g}\,\frac{g-2}{g}\cdots\frac{g-(n-1)}{g}$$

$$\mathbf{P}(X \in A) = \prod_{i=1}^{n-1}\left(1 - \frac{i}{g}\right) =: w(n,g)$$

ist die Wahrscheinlichkeit für Kollisionsfreiheit

bei der n-maligen wiederholten rein zufälligen Platzwahl

aus g Plätzen



Wahrscheinlichkeit fuer Kollisionsfreiheit bei g=365
w(n)
1.0
0.8
0.6
0.4
0.2
0.0
w(n)
0
20
40
60
80
100
120
n

4. Der Zeitpunkt der ersten Kollision und seine Verteilung

Ein dynamisches Bild:

Die Anzahl g der Plätze ist fest.

Jetzt wird ein Objekt nach dem anderen ($i = 1, 2, \ldots, g + 1$)
auf einen (immer wieder neu)
rein zufällig gewählten Platz gesetzt.

Es sei X_i der von Objekt i gewählte Platz.

Eines ist sicher:

Spätestens bis $i = g + 1$muss eine Kollision kommen

(warum?)

Das “zufällige Wahlprotokoll”

$$(X_1, X_2, \ldots, X_{g+1})$$

ist somit eine Zufallsvariable mit Wertebereich

$$S :=$$

$$\{(a_1, \ldots, a_{g+1}\}) : a_i \in \{1, \ldots, g\}; \exists i \neq j \text{ mit } a_i = a_j\}$$

$$T := \min\{n \le g+1 : \exists i < n : X_i = X_n\}$$

ist der *Zeitpunkt der ersten Kollision*.

T ist eine aus $(X_1, \ldots, X_{g+1})$ ablesbare Zufallvariable.

Für das Ereignis

E_n := "keine Kollision bis (einschließlich) n" gilt:

$$E_n = \{T > n\}$$

also insbesondere auch

$$\mathbf{P}(E_n) = \mathbf{P}(T > n).$$

$$\mathbf{P}(E_n) = \mathbf{P}(T > n)$$

In Abschnitt 3 hatten wir berechnet:

$$\mathbf{P}(E_n) = \prod_{i=1}^{n-1}\left(1 - \frac{i}{g}\right) =: w(n,g) =: w(n)$$

$$w(n) = \mathbf{P}(T > n)$$

Wie bekommen wir daraus die Verteilungsgewichte

$$\mathbf{P}(T = n),\ n = 1, 2, \ldots$$

der $\mathbb{N}$-wertigen Zufallsvariablen T?

$$w(n) = \mathbf{P}(T > n)$$

Aus der Additivität der Wahrscheinlichkeit erhalten wir

$$\mathbf{P}(T > n-1) = \mathbf{P}(T = n) + \mathbf{P}(T > n).$$

Also:

$$\mathbf{P}(T = n) = w(n-1) - w(n).$$

Die $\mathbb{N}$-wertige Zufallsvariable T

(den Zeitpunkt der ersten Kollision) haben wir bekommen

als "Verarbeitung" des zufälligen Wahlprotokolls X:

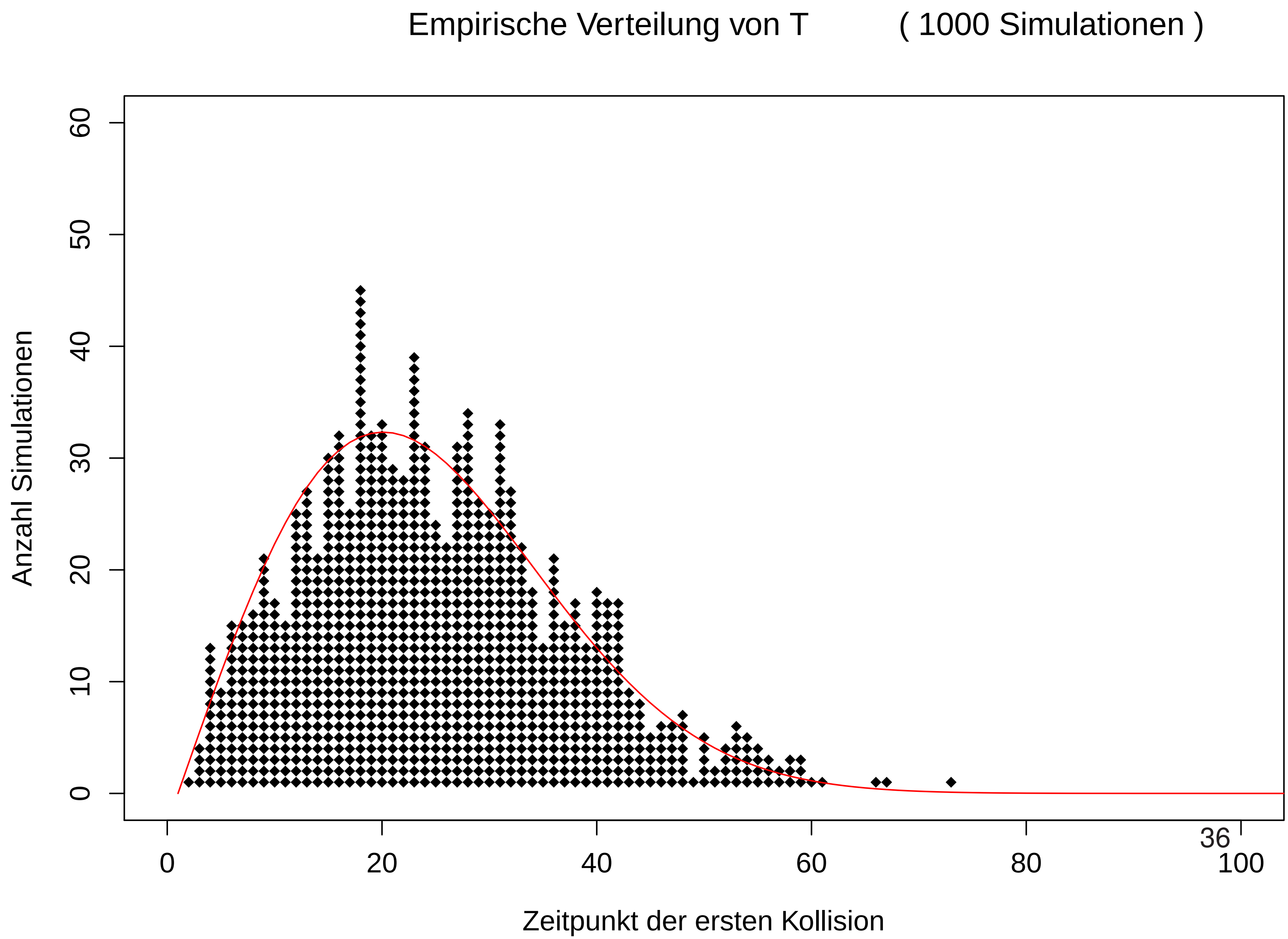
$T = h(X)$ mit passendem $h : S \to \mathbb{N}$.

Aus der Gleichheit der Ereignisse

$$\{T = n\} = \{h(X) = n\} = \{X \in h^{-1}(n))\}$$

ergibt sich die Gleichheit

$$\mathbf{P}(T = n) = \mathbf{P}(X \in h^{-1}(n)), \quad n \in \mathbb{N}.$$


Empirische Verteilung von T (1000 Simulationen)
Anzahl Simulationen
0
10
20
30
40
50
60
0
20
40
60
80
100
Zeitpunkt der ersten Kollision

Vorlesung 2a

Diskret uniform verteilte Zufallsvariable

Teil 3: Besetzungszahlen

(vgl. Buch S. 10-11)

1. Begriffsbildung

Sei $a = (a_1, \ldots, a_n)$ eine $1, \ldots, g$ - Folge der Länge n.

Vorstellung: Objekt Nr. i kommt auf Platz a_i.

Wie oft wird laut Protokoll a der Platz j besetzt?

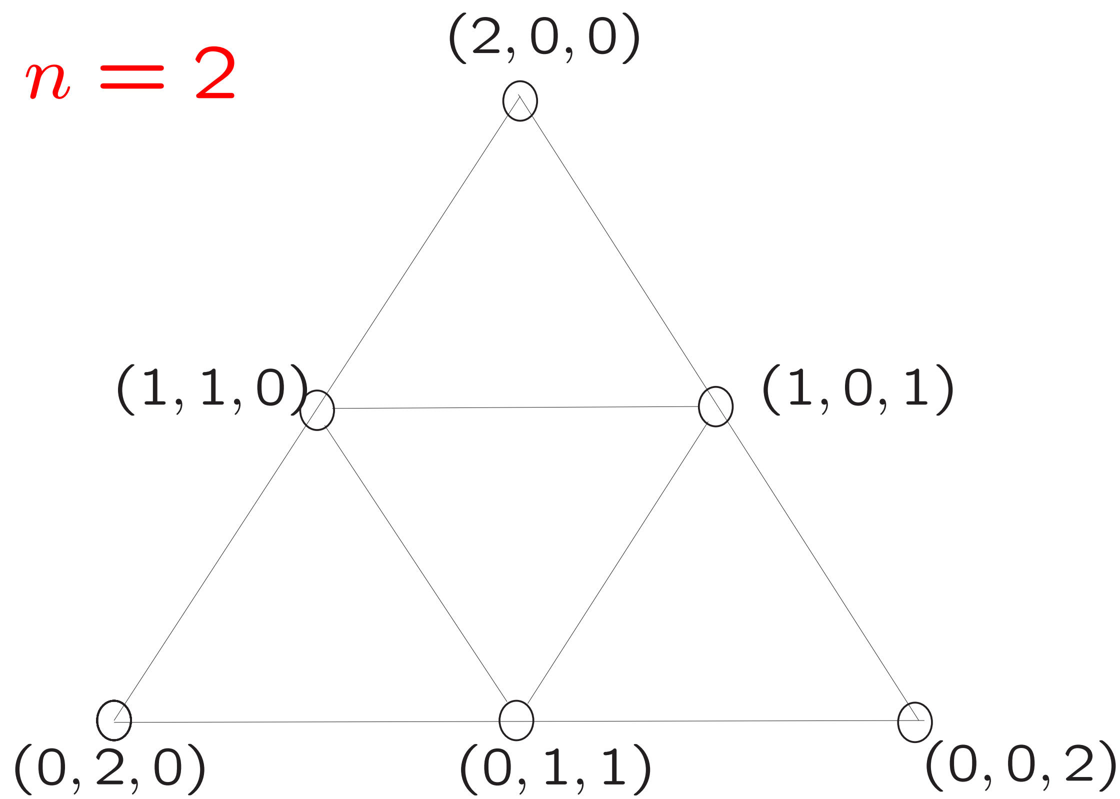
Anders gesagt: Für wieviele i ist $a_i = j$?

$$b_j(a) := \#\{i : \, a_i = j, 1 \leq i \leq n\}$$

Das g-tupel der Besetzungszahlen $b_j(a)$ nennen wir kurz
“die durch a induzierte **Besetzung**”.

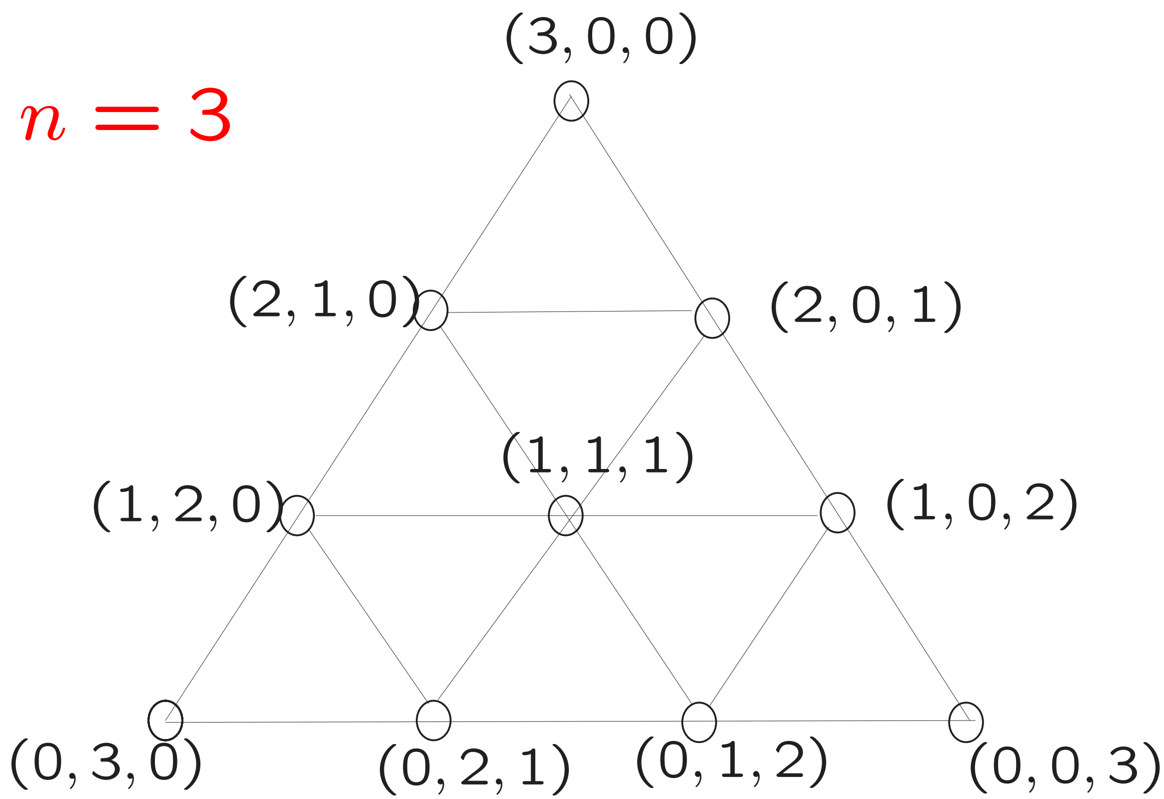
In der Vorstellung des
Setzens von n Objekten auf g mögliche Plätze
gibt sie an, wieviele Objekte auf welchem Platz landen
(und unterscheidet nicht, welche Objekte das sind).

Im Fall $g = 3$ gibt es eine nette Darstellung der Menge der Besetzungen mittels des sogenannten de Finetti-Dreiecks:



Im Fall $g = 3$ gibt es eine nette Darstellung der Menge der Besetzungen mittels des sogenannten de Finetti-Dreiecks:

$n = 3$



2. Besetzungszahlen bei wiederholter rein zufälliger Platzwahl

Machen wir uns ein Bild von der zufälligen Besetzung,

die aus einem auf $\{1, \dots, g\}^n$ uniform verteilten X entsteht.

Das war das Szenario aus Vorlesung 1b:

n-malige rein zufällige Wahl

aus g möglichen Plätzen.

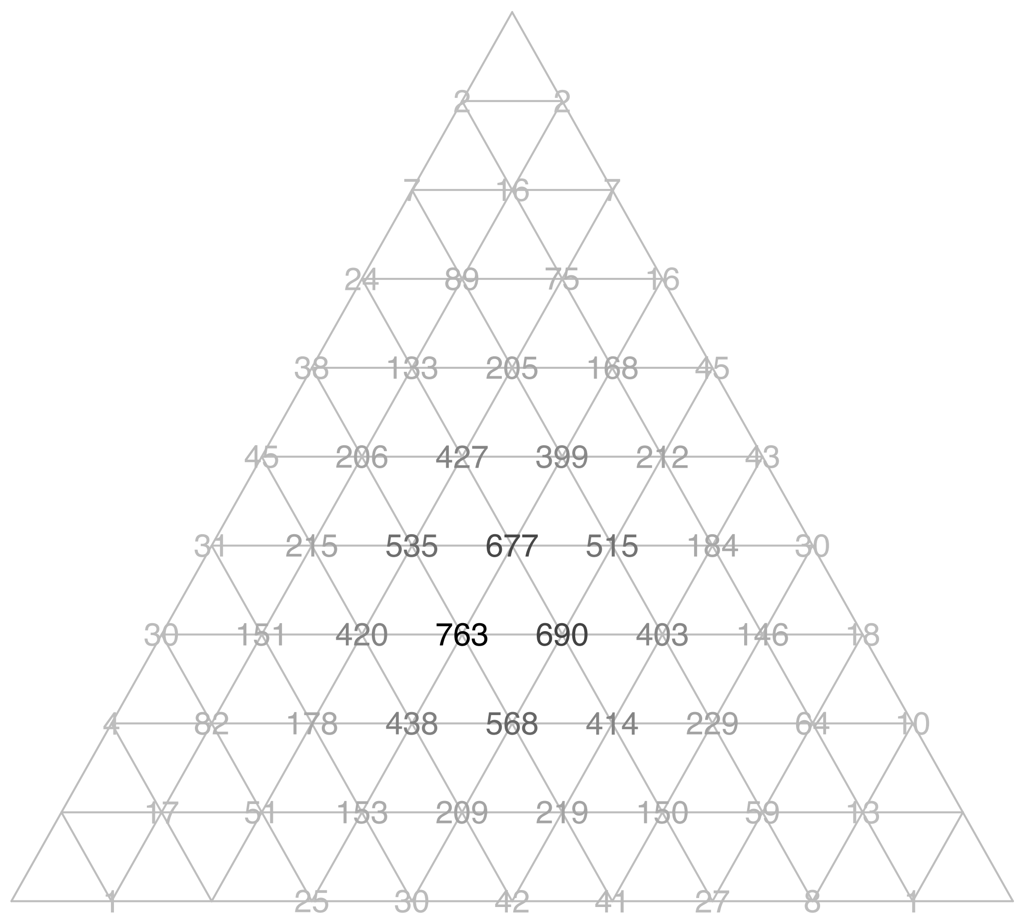
Es folgt das Ergebis einer Simulation für $n = 10$ und $g = 3$.

Die Ecken des *de Finettii-Dreiecks* auf der nächsten Folie

entsprechen den Besetzungen

$(10, 0, 0)$ (oben), $(0, 10, 0)$ (links) und $(0, 0, 10)$ (rechts).

Häufigkeiten der Besetzungen bei 10000 Wiederholungen



Wir sehen aus der Simulation:

Die Verteilung dieser zufälligen Besetzung ist
(bei weitem) nicht uniform.

Wir kommen auf diese Verteilung
in der nächsten Vorlesung zurück.

Zum Kontrast betrachten wir jetzt die

3. Uniform verteilte Besetzung

von g Plätzen mit n Objekten:

Der Zielbereich ist

$$S_{n,g} := \{b = (b_1, \ldots, b_g) : b_j \in \mathbb{N}_0,\ b_1 + \cdots + b_g = n\}$$

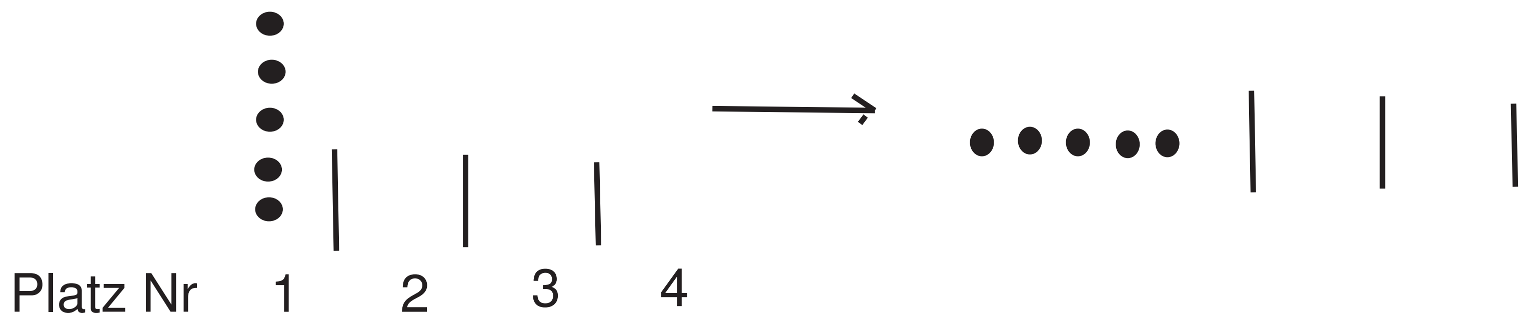
b ist ein g-tupel von Besetzungszahlen, kurz: eine Besetzung.

$$S_{n,g} := \{b = (b_1, \ldots, b_g) : b_j \in \mathbb{N}_0,\ b_1 + \cdots + b_g = n\}$$

$$\# S_{n,g} = ?$$

Hier hilft ein hübscher Trick, $S_{n,g}$ anders darzustellen.

Beispiele: $g = 4$ Plätze, $n = 5$ Objekte



$$(5, 0, 0, 0) \longrightarrow (1, 1, 1, 1, 1, 0, 0, 0)$$

$$(1,0,2,2) \longrightarrow (1,0,0,1,1,0,1,1)$$

$$(2,0,3,0) \longrightarrow (1,1,0,0,1,1,1,0)$$

Fakt:

$$h(b_1, b_2, \ldots, b_g) := \underbrace{1 \ldots 1}_{b_1\text{-mal}} 0 \underbrace{1 \ldots 1}_{b_2\text{-mal}} 0 \ldots 0 \underbrace{1 \ldots 1}_{b_g\text{-mal}}$$

ist eine bijektive Abbildung von $S_{n,g}$ nach

$S :=$ Menge der 01-Folgen der Länge $n + g - 1$

mit genau n Einsen

$$h(b_1, b_2, \ldots, b_g) := \underbrace{1\ldots1}_{b_1\text{-mal}}0\underbrace{1\ldots1}_{b_2\text{-mal}}0\ldots0\underbrace{1\ldots1}_{b_g\text{-mal}}$$

Voriges Beispiel: $n = 5$, $g = 4$:

$$(b_1, \ldots, b_4) = (2, 0, 3, 0)$$

$$h(2, 0, 3, 0) = 11001110$$

Der zweite und der vierte Block aus Einsen sind hier leer.

$$h(b_1, b_2, \ldots, b_g) := \underbrace{1\ldots1}_{b_1\text{-mal}}0\underbrace{1\ldots1}_{b_2\text{-mal}}0\ldots0\underbrace{1\ldots1}_{b_g\text{-mal}}$$

Die Länge des j-ten Blocks aus Einsen steht für die Anzahl der Objekte auf Platz j.

Die Blöcke aus Einsen sind durch Nullen getrennt. Die Nullen fungieren als “Trennwände” zwischen den g Plätzen, insgesamt gibt es $g - 1$ solche Trennwände.

$S :=$ Menge der 01-Folgen der Länge $n + g - 1$

mit genau n Einsen

$$\#S = ?$$

$$\#S = \binom{n + g - 1}{n}$$

Also (wegen der Bijektion mittles h) auch:

$$\#S_{n,g} = \binom{n + g - 1}{n}$$

Eine Möglichkeit zum Erzeugen

einer uniform verteilten Besetzung:

Ziehe aus einer Urne mit n weißen und $g - 1$ schwarzen

Kugeln sukzessive ohne Zurücklegen.

Notiere 0 beim Zug einer schwarzen

und 1 beim Zug einer weißen Kugel.

Erzeuge so ein rein zufälliges Element aus S.

Übersetze dieses (mit der Umkehrung von h)

in eine rein zufällige Besetzung.

Hier ist noch ein weiteres Zufallsexperiment, das auf eine
uniform verteilte Besetzung $(Z_1, \ldots, Z_g)$ führt –
dass das tatsächlich so ist, werden wir im Kapitel über mehrstufige Zufallsexperimente unter der Überschrift *Pólya-Urne* sehen.

Das Experiment (unter dem Stichwort *Wo Tauben sind, fliegen Tauben zu*) läuft wie folgt:
Anfangs sind g Personen auf der Bühne, nummeriert mit $1, \ldots, g$.
Sukzessive kommen Ankömmlinge dazu.
Ankömmling 1 wählt rein zufällig eine der g Personen und stellt sich zu ihr, damit steigt die Zahl der Personen auf der Bühne auf $g + 1$.
Ankömmling 2 wählt rein zufällig eine der $g + 1$ Personen und stellt sich zu ihr, u.s.w.
Betrachtet wird dann (bei insgesamt n Ankömmlingen) das g-Tupel $(Z_1, \ldots, Z_g)$,
wobei Z_j die Anzahl der Ankömmlinge ist, die bei der Person Nr. j stehen.

Vorlesung 2a

Diskret uniform verteilte Zufallsvariable

(Buch S. 6-11)

Erinnerung und Auftakt

Sei S eine endliche Menge.

Eine Zufallsvariable X heißt *uniform verteilt auf S*,

wenn

$$\mathbf{P}(X = a) = \frac{1}{\#S} \quad \text{für alle } a \in S.$$

Damit beschreibt X eine rein zufällige Wahl aus S.

Beispiel aus Vorlesung 1b:

$$S = \{1, 2, \ldots, g\}^n$$

X := rein zufällige $1 \ldots g$ - Folge der Länge n.

Eine auf einem endlichen Wertebereich uniform verteilte Zufallsvariable nennt man auch diskret uniform verteilt.

Heute lernen wir drei weitere Beispiele

von diskret uniform verteilten Zufallsvariablen kennen:

1. Rein zufällige Permutation
2. Rein zufällige k-elementige Teilmenge
3. Uniform verteilte Besetzung

Bei der Gelegenheit erarbeiten wir auch ein paar

Hilfen fürs Abzählen.

Teil 1

Rein zufällge Permutation

(Buch S. 6-8)

1. Elementares

Eine *Permutation* von $1, \ldots, n$

ist eine bijektive Abbildung der Menge $\{1, \ldots, n\}$ auf sich.

Z. B. mit $n = 7$

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Wie wahrscheinlich ist es,

dass eine rein zufällige Permutation

genau **so** ausfällt?

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

mal $(n-1)$ Möglichkeiten für das Bild von 2

mal $(n-2)$ Möglichkeiten für das Bild von 3

$$\ldots$$

$$= n(n-1)(n-2)\cdots 2 \cdot 1 = n!$$

Sei X eine rein zufällige Permutation von $1, \dots, n$,

d.h. eine Zufallsvariable, deren Zielbereich

$S :=$ die Menge aller Permutationen von $1, \dots, n$

ist, und die auf S uniform verteilt ist.

Für alle Elemente $a \in S$ gilt also:

$$\mathbf{P}(X = a) = \frac{1}{n!}$$

2. Zufällige Permutation und zufälliges Ziehen

Wie kann man sich eine rein zufällige Permutation entstanden denken?

Zum Beispiel: als Folge der gezogenen Nummern beim n-maligen rein zufälligen Ziehen *ohne Zurücklegen* aus $\{1, 2, \ldots, n\}$.

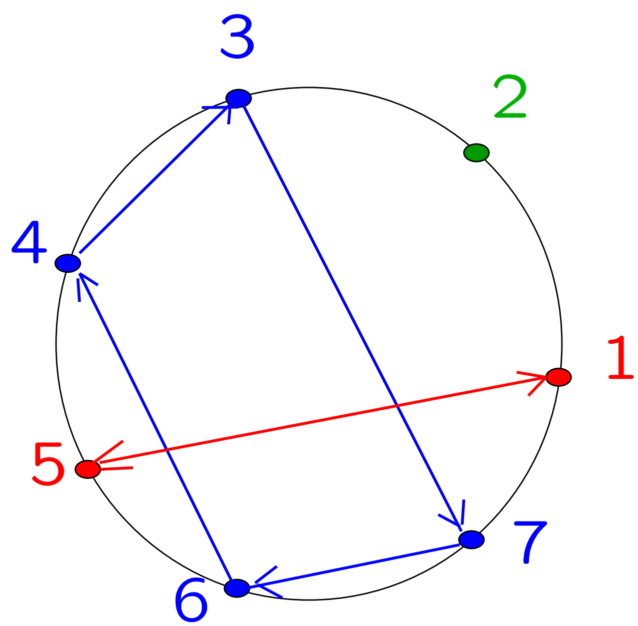
Szenario: eine stets ideal durchmischte Urne mit anfangs n Kugeln, beschriftet mit den Nummern $1, \ldots, n$. Ziehe sukzessive **ohne Zurücklegen** alle n Kugeln und notiere die gezogenen Nummern in Reihenfolge.

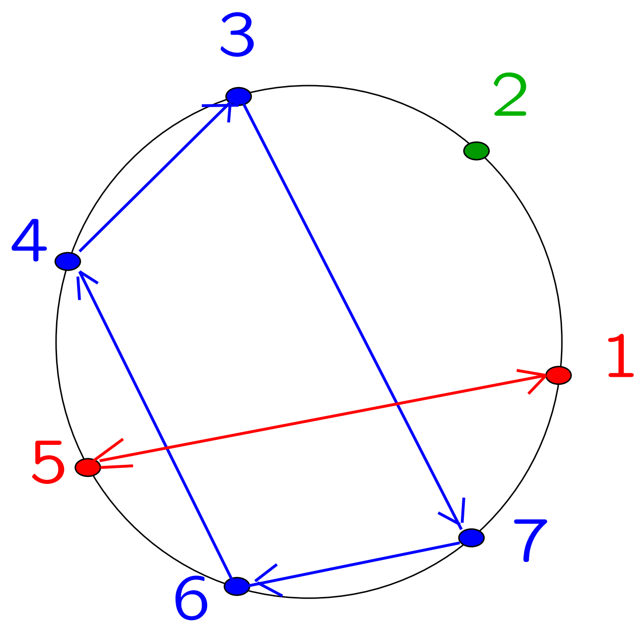
3. Zyklendarstellung einer Permutation

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$





Die Länge des Zyklus, der die Eins enthält, ist hier

zwei.

Sei X eine rein zufällige Permutation von $1, \dots, 7$,

Wie wahrscheinlich ist es, dass der die 1 enthaltende Zykel genau die Länge 3 hat?

Wieviele Permutationen von $1, \dots, 7$ gibt es, bei denen der die 1 enthaltende Zykel genau die Länge 3 hat?

Es gibt davon $6 \cdot 5 \cdot 1 \cdot 4!$ Stück (warum?)

Also ist die gefragte W'keit: $\dfrac{6!}{7!} = \dfrac{1}{7}$.

Jetzt allgemein:

Für eine Permutation $a \in S$ bezeichne

$$h(a)$$

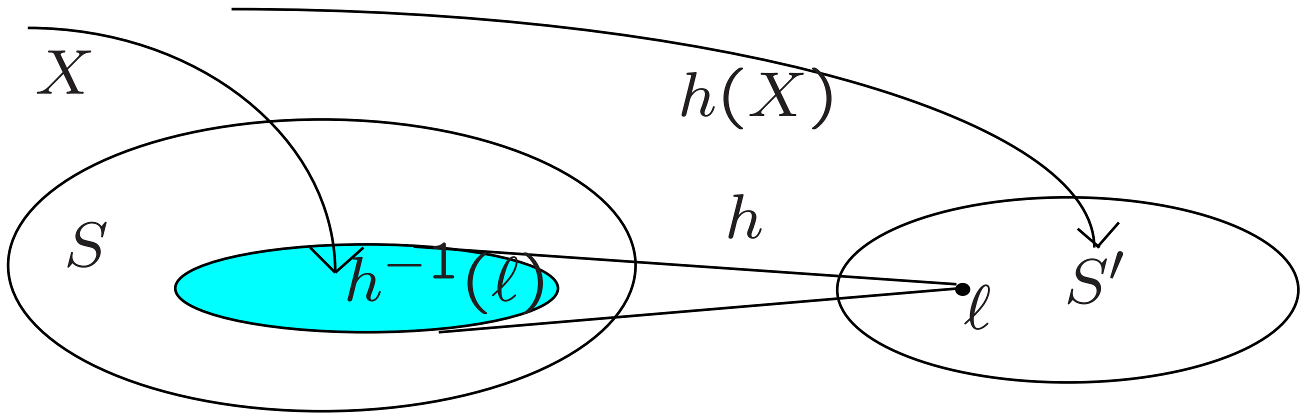
die Länge des Zyklus von a, der die Eins enthält.

Sei X eine rein zufällige Permutation von $\{1, \ldots, n\}$,

also eine rein zufällige Wahl aus S,

und sei $\ell \in \{1, 2, \ldots, n\}$.

$$\mathbf{P}(h(X) = \ell) = ?$$



$$S' = \{1, \ldots, n\}$$

Wieviele Permutationen $a \in S$ gibt es mit $h(a) = \ell$?

$$A := \{a \in S : h(a) = \ell\}$$

$$\#A = ?$$

$$A = \{a \in S : a(1) \neq 1, a^2(1) \neq 1, \ldots, a^{\ell-1}(1) \neq 1, a^{\ell}(1) = 1\}$$

$$\#A = (n-1)(n-2)\cdots(n-(\ell-1)) \cdot 1 \cdot (n-\ell)\cdots 1$$

$$= (n-1)!$$

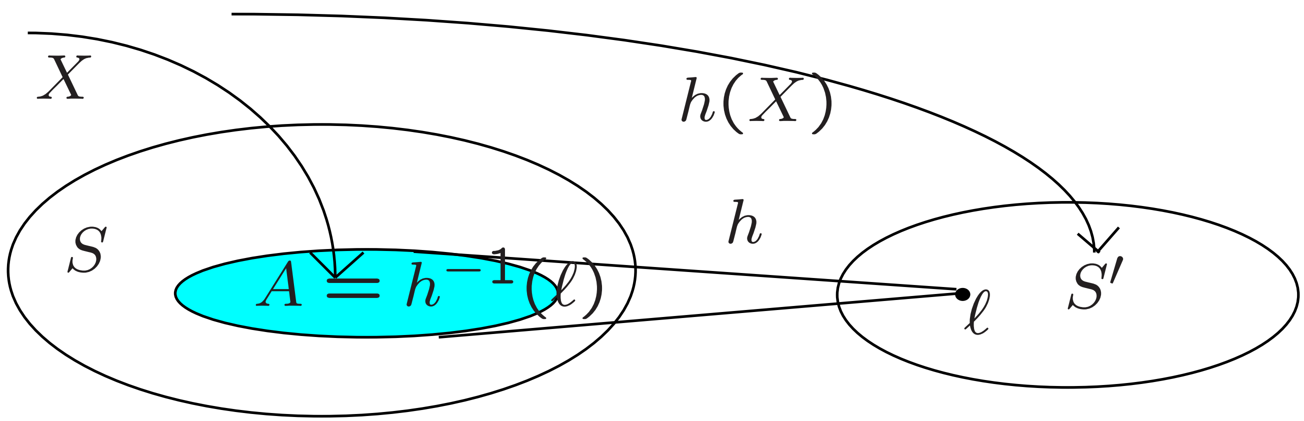
$$\#A = (n-1)!, \qquad \#S = n!$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}$$

$$= \frac{(n-1)!}{n!}$$

$$= \frac{1}{n}$$

$$\mathbf{P}(X \in A) = \frac{1}{n}$$



$$S' = \{1, \ldots, n\}$$

$$A = \{a \in S : h(a) = \ell\}$$

$$\{X \in A\} = \{h(X) = \ell\}$$

$$\mathbf{P}(h(X) = \ell) = \frac{1}{n}$$

$$\mathbf{P}(h(X) = \ell) = \frac{1}{n}, \quad \ell = 1, \ldots n.$$

Fazit:

Die Länge desjenigen Zyklus

einer rein zufälligen Permutation von $1, \ldots, n$,

der die Eins enthält,

ist uniform verteilt auf $\{1, \ldots, n\}$.

Vorlesung 2a

Diskret uniform verteilte Zufallsvariable

Teil 2: Rein zufällige Teilmenge einer festen Größe

(Buch S. 9-10)

1. Die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n\}$: der Binomialkoeffizient.

Sei $k \leq n$.

Jetzt sei S

die Menge aller k-elementigen Teilmengen von $\{1, \ldots, n\}$.

Wieviele k-elementige Teilmengen von $\{1, \ldots, n\}$ gibt es?

Wieviele Möglichkeiten gibt es, aus n Personen

ein k-köpfiges Komitee ohne Reihung zu bilden?

Wird “nach der Reihe” ausgewählt, dann gibt es

$n(n-1)\cdots(n-(k-1))$ mögliche Wahlprotokolle.

Auf die Reihenfolge kommt es am Ende nicht an,

somit führen jeweils $k!$ dieser Wahlprotokolle

auf dieselbe k-elementige Teilmenge.

Also:

$$\#S = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

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

Binomialkoeffizient „ n über k “

die Anzahl der Möglichkeiten für „ k aus n “

Beispiel: Binomischer Lehrsatz:

$$(x+y)^n = \underbrace{(x+y)(x+y)\cdots(x+y)}_{n-\text{mal}} =?$$

Multipliziert man aus, dann ergeben sich 2^n Summanden

(entsprechend den 2^n $x\,y$-Folgen der Länge n).

Jeder dieser Summanden ist von der Form $x^k y^{n-k}$.

Für festes k gibt es $\binom{n}{k}$ Möglichkeiten,

k aus den n Faktoren $(x+y)$ auszuwählen,

bei denen x zum Zug kommt.

Also $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

$\binom{n}{k}$ ... die Anzahl der Möglichkeiten für „ k aus n “

Rekursion: $$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Rekursion: $$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Interpretation: Anzahl der Möglichkeiten,

aus n Hessen und einem Bayern

ein $k+1$ köpfiges Komitee auszuwählen.

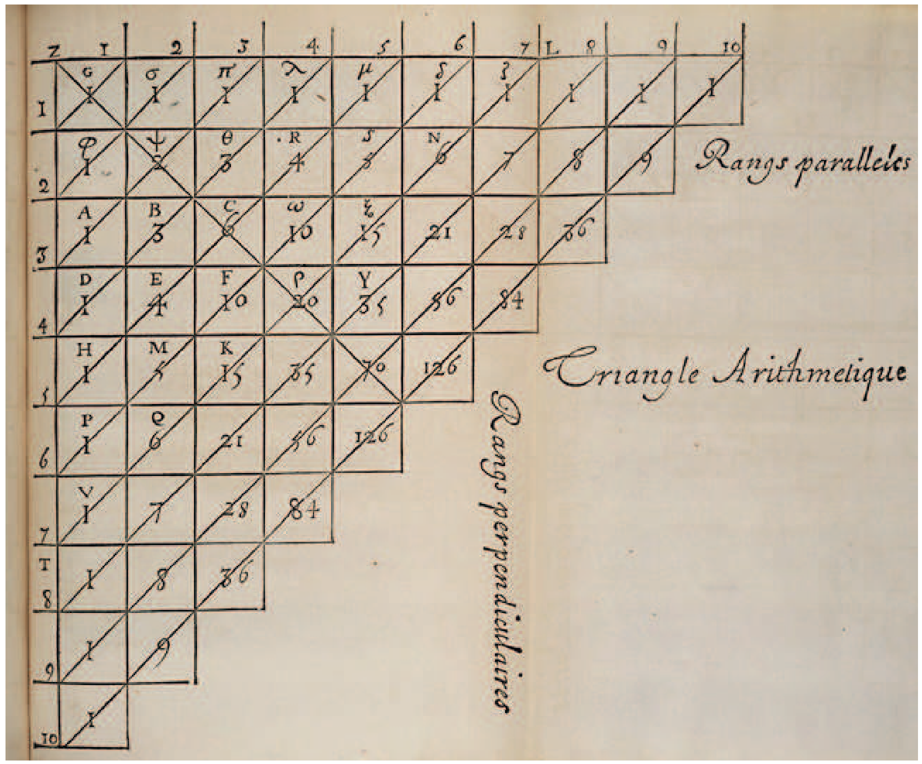
Entweder der Bayer ist im Komitee...

oder er ist nicht im Komitee...

Pascal'sches Dreieck

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & & & & \vdots & & & & &
\end{array}
$$

Rekursion: $$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$



"La nombre de chaque cellule est égal a celui de la cellule
qui la précède dans son rang perpendiculaire,
plus à celui de la cellule qui la précède dans son rang parallèle."
(Pascal 1654)

2. Rein zufällige Teilmengen

Sei $0 \leq k \leq n$

und sei Y eine rein zufällige k-elementige Teilmenge von $\{1, \ldots, n\}$.

Wie wahrscheinlich ist das Ereignis $\{Y = \{1, \ldots, k\}\}$?

Der Zielbereich von Y ist

$$S := \{t : t \subset \{1, \dots, n\},\ \#t = k\}\ ,$$

die Menge der k-elementigen Teilmengen von $\{1, \dots, n\}$.

Wir haben gesehen:

$$\#S = \binom{n}{k}$$

Fazit:

$$\mathbf{P}(Y = \{1, \dots, k\}) = \frac{1}{\binom{n}{k}}\ .$$

3. Ein Zusammenhang mit rein zufälligen Permutationen

Wie bekommt man eine rein zufällige k-elementige Teilmenge

aus einer rein zufälligen Permutation?

Fakt (für $k \leq n$):

Ist $X = (X_1, \ldots, X_n)$

eine rein zufälligen Permutation von $1, \ldots, n$,

dann ist $\{X_1, \ldots, X_k\}$

eine rein zufällige k-elementige Teilmenge von $\{1, \ldots, n\}$.

Anders gesagt:

Beim Ziehen ohne Zurücklegen

führen die ersten k Züge auf eine

rein zufällige Teilmenge des anfänglichen Reservoirs.

Noch eine Möglichkeit zum Erzeugen

einer rein zufälligen k-elementigen Teilmenge:

Ziehe sukzessive ohne Zurücklegen aus einer Urne

mit k roten und $n - k$ blauen Kugeln.

Notiere die Nummern $X_1, \ldots, X_k$ der Züge,

bei denen eine rote Kugel gezogen wird.

Dann ist $\{X_1, \ldots, X_k\}$

eine rein zufällige k-elementige Teilmenge von $\{1, \ldots, n\}$.

Vorlesung 2b

Diskrete Zufallsvariable

und ihre Verteilungen

Teil 3:

Zufällige Besetzungen:

Vom Würfeln zur Multinomialverteilung

1. Das $(p_1, \ldots, p_g)$-Würfeln als Verallgemeinerung des p-Münzwurfs

Oder: Was 2 recht ist, soll g billig sein!

(vgl. Buch S. 28)

Definition ("n-faches $(p_1, \ldots, p_g)$-Würfeln"):

Seien $g \in \mathbb{N}$ und $p_1, \ldots, p_g \geq 0$ mit $p_1 + \ldots + p_g = 1$.

Wir definieren Verteilungsgewichte auf

$$S := \{a = (a_1, \ldots, a_n) \,:\, a_i \in \{1, \ldots, g\}\}$$

durch

$$\rho(a_1, \ldots, a_n) := p_{a_1} \cdot p_{a_2} \cdots p_{a_n}.$$

Eine Zufallsvariable X mit diesem Zielbereich S und diesen Verteilungsgewichten ρ nennen wir n-faches $(p_1, \ldots, p_g)$-Würfeln.

Für *jedes* $a \in S$ *mit*

b_1 Komponenten gleich 1,

b_2 Komponenten gleich 2,

...

b_g Komponenten gleich g

ist dann

$$\mathbf{P}(X = a) = p_1^{b_1} p_2^{b_2} \cdots p_g^{b_g}$$

2. Vom $(p_1, \ldots, p_g)$-Würfeln zur Multinomialverteilung

$X = (X_1, \ldots, X_n)$ sei ein n-faches $(p_1, \ldots, p_g)$-Würfeln

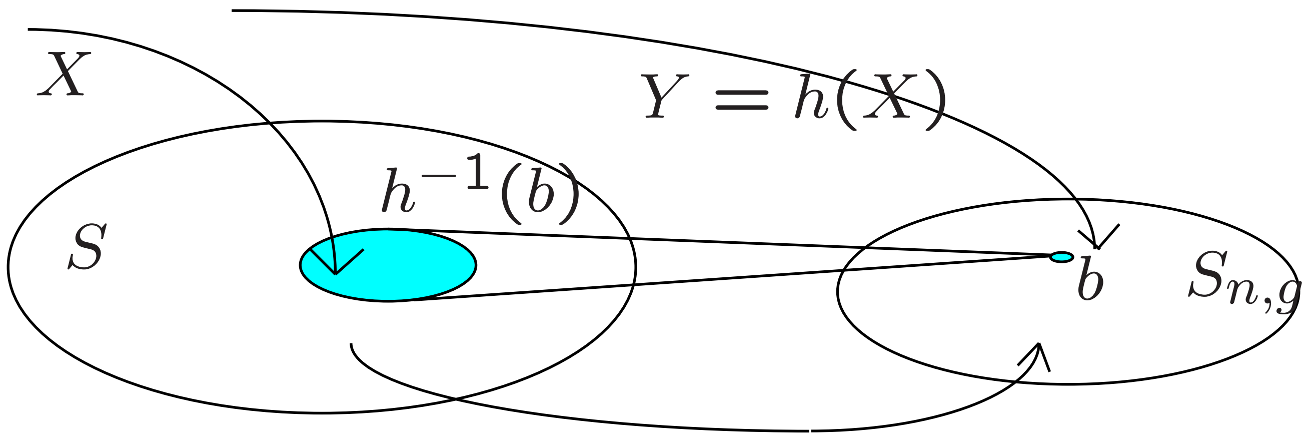
$$Y_j := \#\{i \; : \; X_i = j\}$$

(die Anzahl der Würfe mit Ergebnis j).

$Y := (Y_1, \ldots, Y_g)$ hat dann den Zielbereich

$$S_{n,g} = \{(b_1, \ldots, b_g) \; : b_j \in \mathbb{N}_0, b_1 + \ldots + b_g = n\}.$$

Verteilung von Y = ?



$$h(a_1, \ldots, a_n) = (b_1, \ldots, b_g) =: b$$

mit $b_j := \#\{i : a_i = j\}), \ j = 1, \ldots, g$

Jedes $a \in S$ mit $h(a) = (b_1, \ldots, b_g)$

hat Gewicht $p_1^{b_1} \ldots p_g^{b_g}$

Wieviele solche a gibt es?

Dazu überlegen wir:

Auf wieviele Arten kann man

n Objekte so auf g Fächer verteilen,

dass das j-te Fach jeweils genau b_j Objekte enthält?

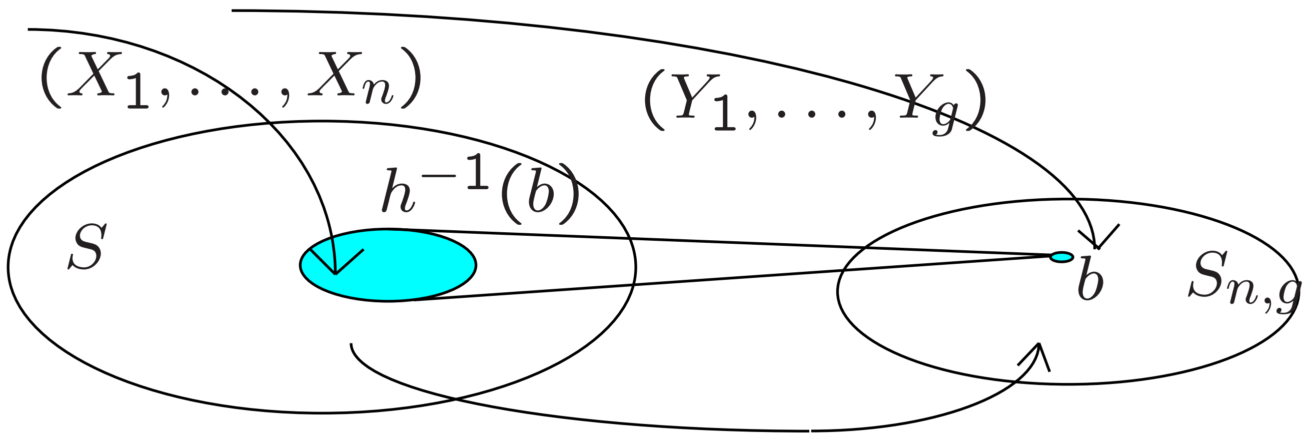
Dabei ist $b_1 + \cdots + b_g = n$.

Die Antwort ist:

$$\binom{n}{b_1} \cdot \binom{n-b_1}{b_2} \cdots \binom{n-b_1-\cdots-b_{g-1}}{b_g}$$

$$= \frac{n!}{b_1!\, b_2! \cdots b_g!} =: \binom{n}{b_1, \ldots, b_g}$$

Multinomialkoeffizient, lies: n über $b_1, \ldots, b_g$



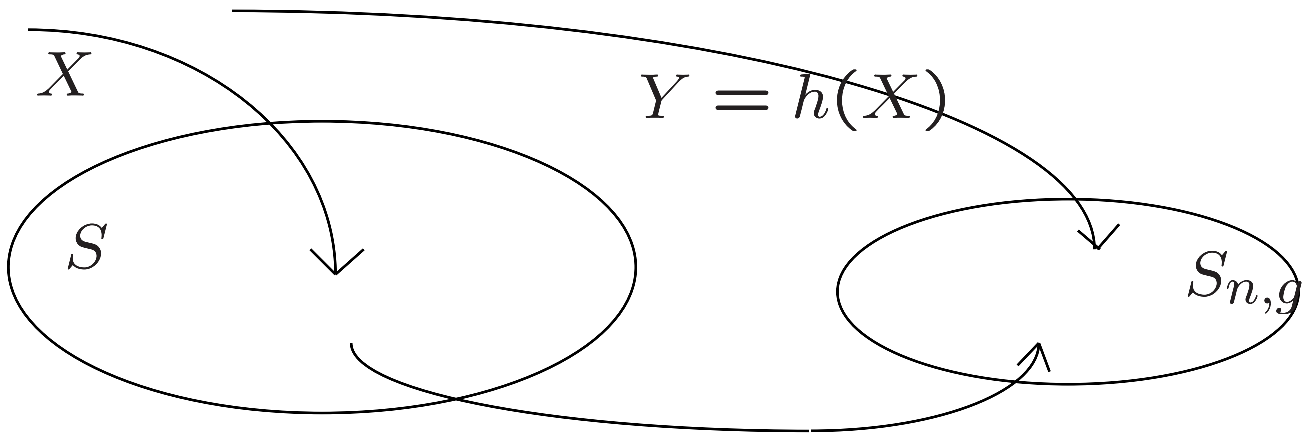
$$h(a_1, \ldots, a_n) = (b_1, \ldots, b_g)$$

Jedes $a \in S$ mit $h(a) = (b_1, \ldots, b_g)$

hat Gewicht $p_1^{b_1} \ldots p_g^{b_g}$

Es gibt $\binom{n}{b_1, \ldots, b_g}$ solche a.

$$\mathbf{P}(Y_1 = b_1, \ldots, Y_g = b_g) = \binom{n}{b_1, \ldots, b_g} p_1^{k_1} \ldots p_g^{b_g}$$



$$h(a_1, \ldots, a_n) = (b_1, \ldots, b_g) =: k$$

mit $b_j := \#\{i : a_i = j\},\ j = 1, \ldots, g$

$$\mathbf{P}(Y_1 = b_1, \ldots, Y_g = b_g) = \binom{n}{b_1, \ldots, b_g} p_1^{b_1} \ldots p_g^{b_g}$$

Definition:

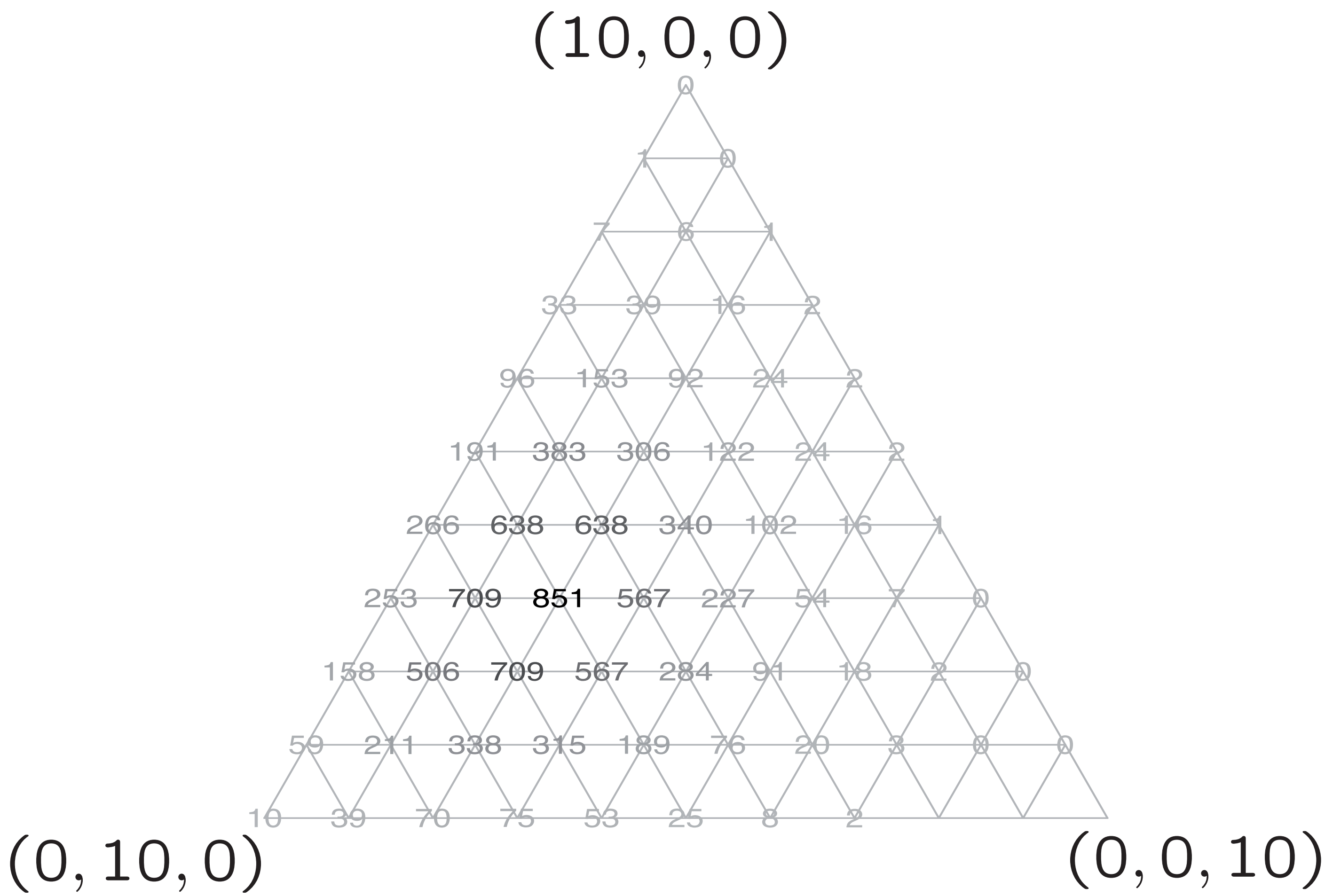
Eine Zufallsvariable X mit Zielbereich $S_{n,g}$

heißt *multinomialverteilt* mit Parametern $n; p_1, \ldots, p_g$,

wenn

$$\mathbf{P}(X = (b_1, \ldots, b_g)) = \binom{n}{b_1, \ldots, b_g} p_1^{b_1} \ldots p_g^{b_g},$$

$$(b_1, \ldots, b_g) \in S_{n,g} \,.$$



Gewichte der Multinomialverteilung, notiert in Vielfachen von $\frac{1}{10000}$,
für $n = 10$, $g = 3$, $p_1 = 0.3$, $p_2 = 0.5$, $p_3 = 0.2$

Resumé

$g \in \mathbb{N}$	$g = 2$
Würfeln	Münzwurf
Besetzung der Ausgänge	Anzahl der Erfolge
Multinomialverteilung	Binomialverteilung

Vorlesung 2b

Diskrete Zufallsvariable

und ihre Verteilungen

Teil 2:

Die Anzahl der Erfolge:

Vom Münzwurf zur Binomialverteilung

(vgl. Buch S. 22)

1. Vom n-fachen fairen Münzwurf

zur Binomial$(n, \frac{1}{2})$-Verteilung:

$$S := \{0, 1\}^n$$

die Menge der 01-Folgen der Länge n

X sei uniform verteilt auf S,

jeder Ausgang hat somit das Gewicht

$$\frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2}.$$

Man sagt dann auch:

X ist ein n-facher "fairer Münzwurf".

$Y :=$ die Anzahl der Einsen in X.

Was ergibt sich für die Verteilungsgewichte von Y?

Jede einzelne 01-Folge a der Länge n mit genau k Einsen hat Gewicht $\frac{1}{2^n}$

Es gibt $\binom{n}{k}$ derartiger Folgen a.

Also:

$$\mathbf{P}(Y = k) = \binom{n}{k}\frac{1}{2^n}, \quad k = 0, \dots, n.$$

Das sind die Gewichte der Binomial$(n, \frac{1}{2})$-Verteilung.

2. Die Anzahl der Sechsen beim fairen Würfeln:

Vom n-fachen Würfeln zur Binomial$(n, \frac{1}{6})$-Verteilung:

(vgl. Buch S. 28)

$X = (X_1, \dots, X_n)$ uniform verteilt auf

$S := \{1, \dots, 6\}^n$.

$Z := (Z_1, \dots, Z_n)$, mit

$Z_i := \mathbf{1}_{\{6\}}(X_i)$

Z ist also eine zufällige 01-Folge, mit

$Z_i = 1$ falls der i-te Wurf eine Sechs ergibt

und $Z_i = 0$ sonst.

Wie ist Z verteilt?

$$\mathbf{P}(Z_1 = 1, \ldots, Z_k = 1,\ Z_{k+1} = 0, \ldots, Z_n = 0)$$

$$= \mathbf{P}(X_1 = 6, \ldots, X_k = 6,\ X_{k+1} \neq 6, \ldots, X_n \neq 6)$$

$$= \frac{1^k \cdot 5^{n-k}}{6^n}$$

$$= p^k q^{n-k},$$

mit $p := \frac{1}{6}$ und $q := \frac{5}{6}$.

Auch für jede andere Platzierung von genau k “Sechsen”

in den n Würfen ergibt sich diese W’keit.

Es gibt Es gibt $\binom{n}{k}$ derartige Platzierungen.

Allso hat die Verteilung von

$$Y := Z_1 + \cdots + Z_n$$

die Gewichte

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

Das sind die Gewichte der Binomial$(n, \frac{1}{6})$-Verteilung.

3. Vom p-Münzwurf zur Binomialverteilung

(“Was $1/6$ recht ist, soll p billig sein...”)

(Buch S. 22)

Definition (p-Münzwurf):

Sei $p \in [0, 1]$, $q := 1 - p$.

Eine Zufallsvariable Z mit Zielbereich

$$S = \{0, 1\}^n = \{a = (a_1, \ldots, a_n) \ : a_i \in \{0, 1\}\}$$

heißt n-facher p-Münzwurf,

wenn für alle $a \in S$ mit k Einsen und $n - k$ Nullen gilt:

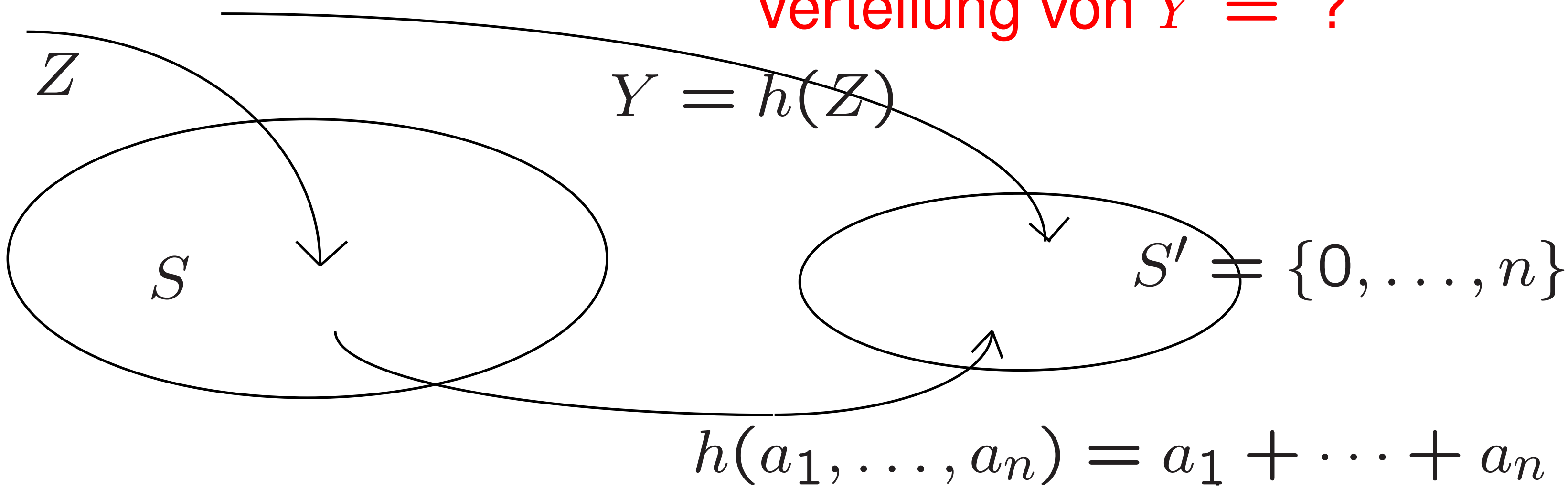
$$\mathbf{P}(Z = a) = p^k q^{n-k}.$$

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

und $Y = Z_1 + \cdots + Z_n =: h(Z)$ die *Anzahl der Erfolge*

(die Anzahl der Einsen in der zufälligen 0-1 Folge Z)

Verteilung von $Y =$?



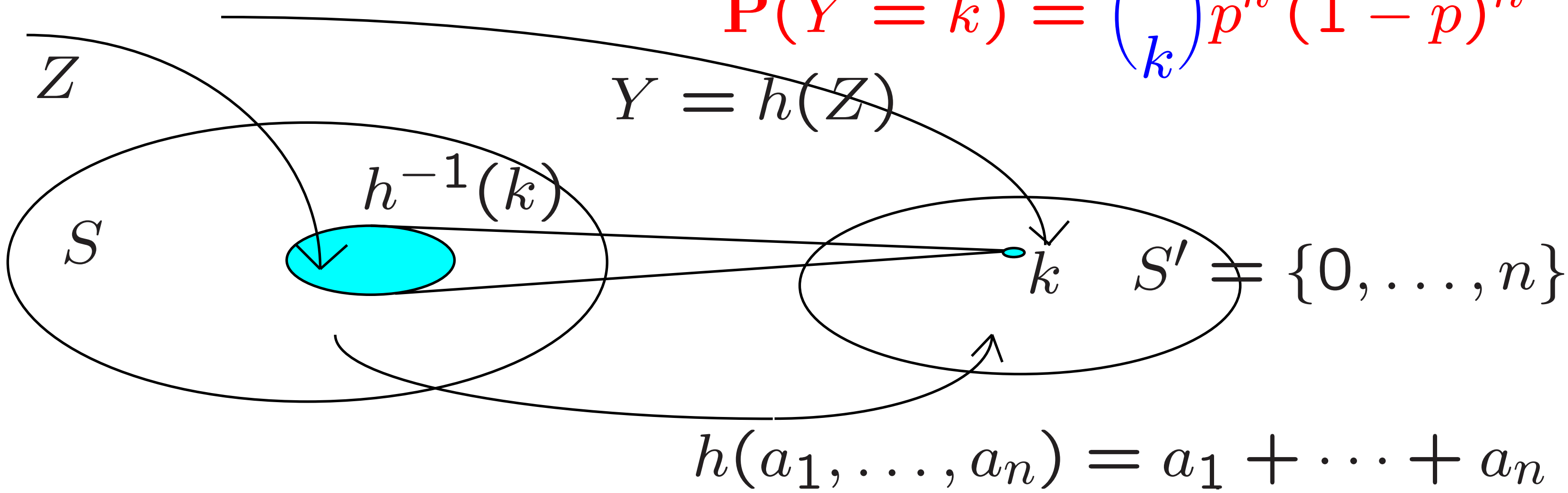
Jedes $a \in S$ mit $h(a) = k$

(d.h. mit k Einsen und $n - k$ Nullen)

hat Gewicht $p^k (1-p)^{n-k}$.

Es gibt $\binom{n}{k}$ solche a.

$$\mathbf{P}(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$



Definition:

Eine Zufallsvariable Y mit Zielbereich $\{0, 1, \ldots, n\}$

heißt *binomialverteilt* mit Parametern n und p,

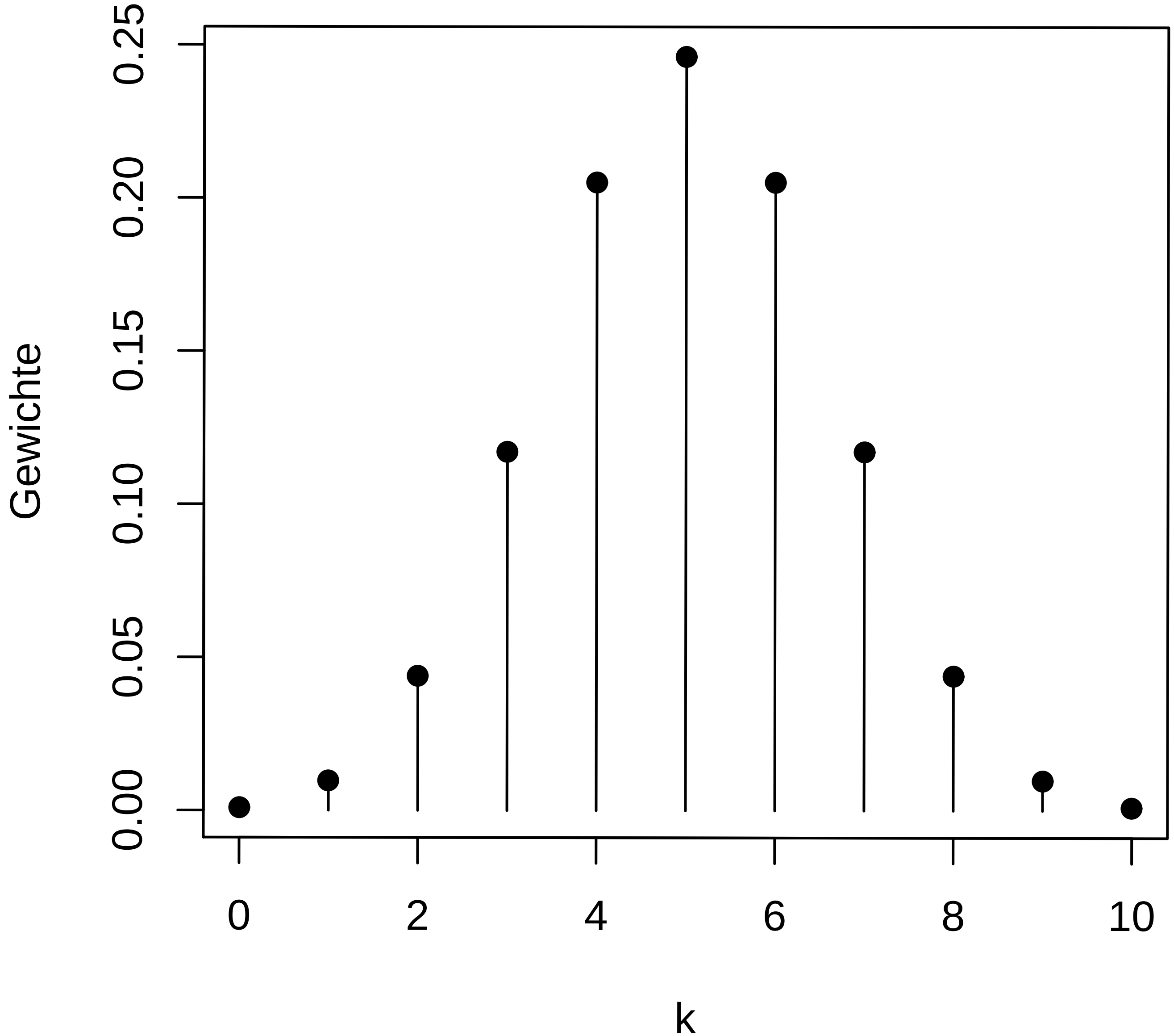
kurz

Bin(n, p)-verteilt,

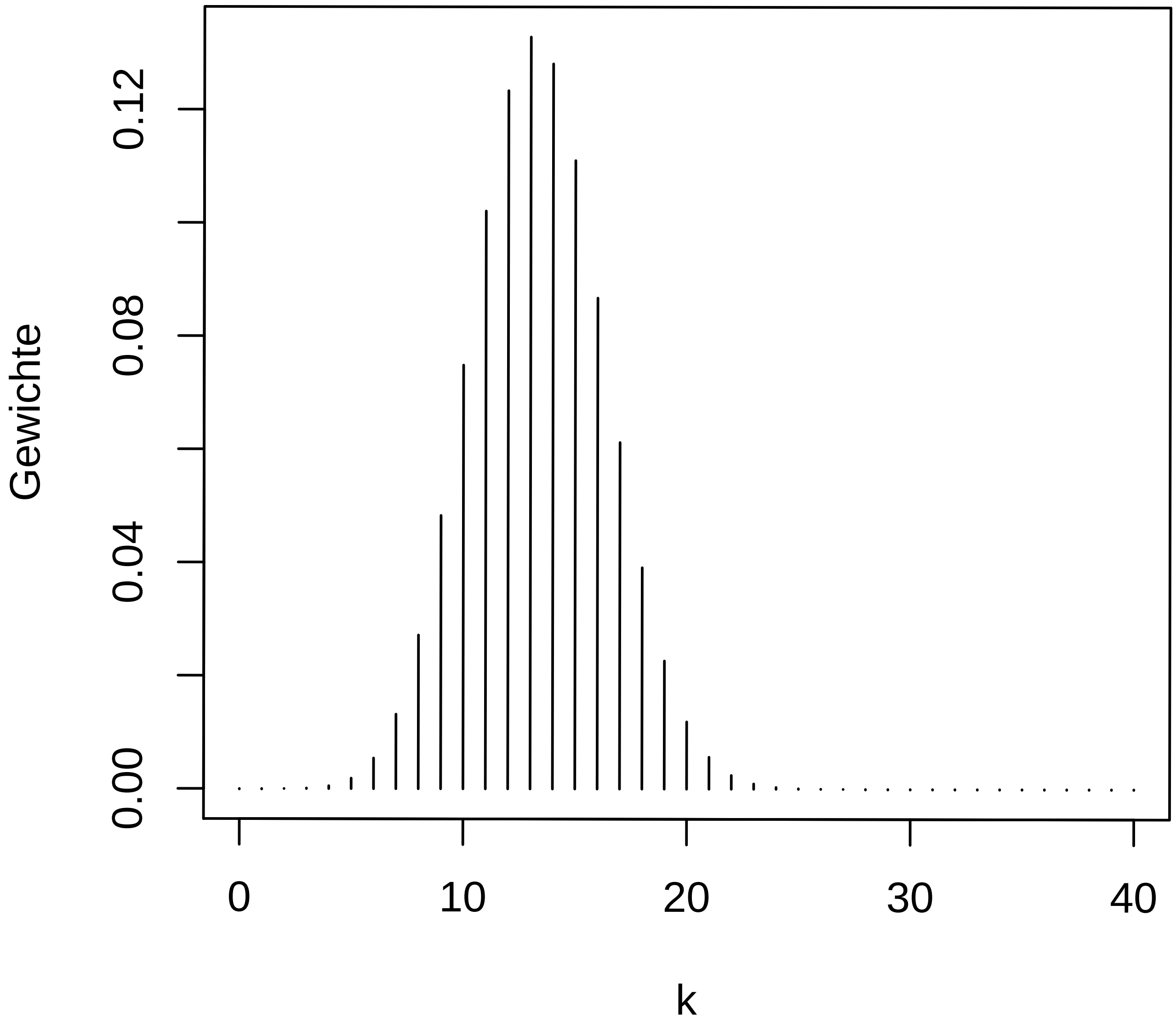
wenn

$$\mathbf{P}(Y = k) = \binom{n}{k} p^k q^{n-k} \ , \quad k = 0, 1, \ldots, n \ ,$$

mit $q = 1 - p$.



Gewichte der Bin(10, 1/2) Verteilung



Gewichte der Bin(40, 1/3) Verteilung

4. Vom Ziehen mit Zurücklegen zum p-Münzwurf

(vgl. Buch S. 33)

n-maliges Ziehen *mit Zurücklegen*

aus einer ideal durchmischten Urne.

Ein Anteil p der Kugeln ist rot,

der restliche Anteil $q = 1 - p$ ist blau.

Zufällige 0-1 Folge $Z = (Z_1, \dots, Z_n)$:

$Z_i = 1$ wenn beim i-ten Zug eine rote Kugel kommt,

und $Z_i = 0$ wenn beim i-ten Zug eine blaue Kugel kommt.

Sei a eine vorgegebene 0-1 Folge der Länge n mit k Einsen,

z. B.: $\quad a := (\underbrace{1, \ldots, 1}_{k\text{-mal}}, \underbrace{0, \ldots, 0}_{(n-k)\text{-mal}})$

$$\mathbf{P}(Z = a) = \,?$$

Sei g die Gesamtanzahl der Kugeln in der Urne.

Die Anzahl der roten Kogeln ist pg,

die der blauen Kugeln ist qg. Für obiges a gilt:

$$\mathbf{P}(Z = a) = \frac{(pg)^k (qg)^{n-k}}{g^n} = p^k q^{n-k}$$

Das ist so für jede 0-1 Folge a mit k Einsen und $n - k$ Nullen.

Zur Erinnerung:

Definition (p-Münzwurf):

Sei $p \in [0, 1]$, $q := 1 - p$.

Eine Zufallsvariable Z mit Zielbereich

$$S = \{0, 1\}^n = \{a = (a_1, \dots, a_n) \ : a_i \in \{0, 1\}\}$$

heißt n-facher p-Münzwurf,

wenn für alle $a \in S$ mit k Einsen und $n - k$ Nullen gilt:

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

Vorlesung 2b

Diskrete Zufallsvariable

und ihre Verteilungen

mit den Beispielen

Anzahl der Erfolge beim n-fachen p-**Münzwurf**

und

Besetzungen beim n-fachen $(p_1, \ldots, p_g)$-**Würfeln**

Teil 1: Grundbegriffe

(Buch S. 20-21)

Abzählbare Additivität von Wahrscheinlichkeiten

Verteilung und Verteilungsgewichte

Zufällige Paare: gemeinsame Verteilung und Projektionen

Weiterverarbeitung von Zufallsvariablen und

Transport von Verteilungen

Bisher hatten wir uns mit Zufallsvariablen beschäftigt, deren Wertebereich S endlich war.

Die (schon in Vorlesung 1b formulierten)

zwei Grundregeln

für Wahrscheinlichkeiten lauteten für diesen Fall:

Normiertheit auf Eins:

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

(Folgerung aus der) Additivität:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a), \qquad A \subset S$$

Diese beiden Regeln behalten ihren Sinn, wenn der Wertebereich nicht endlich, sondern abzählbar unendlich ist.

Beispiel: $S = \mathbb{N}$

$$\mathbf{P}(X = 1) = \frac{1}{2}, \quad \mathbf{P}(X = 2) = \frac{1}{4}, \quad \mathbf{P}(X = 3) = \frac{1}{8}, \ldots$$

$$\mathbf{P}(X = a) = 1/2^a, \quad a \in \mathbb{N}.$$

Auch wenn der Wertebereich von X

eine überabzählbare Menge ist

(wie z.B. die Menge der reellen Zahlen $\mathbb{R}$

oder das “Einheitsintervall” $[0, 1]$

oder das “Einheitsquadrat” $[0, 1] \times [0, 1]$),

behalten beide Regeln ihren Sinn, wenn man fordert,

dass der Wertebereich

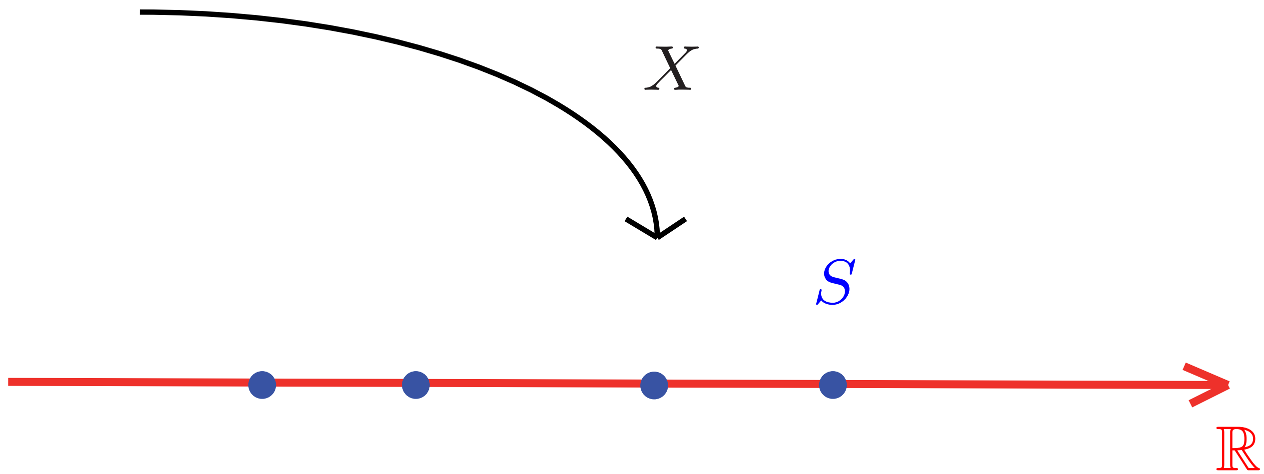
eine endliche oder abzählbar unendliche Menge S enthält

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

Eine Zufallsvariable X mit dieser Eigenschaft nennen wir

diskret.

Beispiel: Wertebereich $\mathbb{R}$



$S \subset \mathbb{R}$ endlich oder abzählbar unendlich mit $\mathbf{P}(X \in S) = 1$

(z.B. $S = \{1, 2, \ldots, g\}$, oder $S = \mathbb{N}$)

Warum ist das interessant?

Wie wir sehen werden (und wie jetzt schon intuitiv klar ist),

kann man mit reellwertigen Zufallsvariablen rechnen.

Man kann z.B. eine reellwertige Zufallsvariable X

durch 2 teilen,

und wenn die Zufallsvariable X diskret ist,

ist auch die Zufallvariable $X/2$ diskret.

Nochmal die Definition:

Eine Zufallsvariable X heißt **diskret**,

falls ihr Wertebereich (nennen wir ihn W)

eine diskrete (d.h. endliche oder abzählbar unendliche)

Menge S enthält mit

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

Für diskretes X mit Wertebereich W

und diskretem S mit $\mathbf{P}(X \in S) = 1$ gilt:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a), \qquad A \subset S$$

(Wahrscheinlichkeiten sind abzählbar additiv)

$$\mathbf{P}(X \in W \setminus S) = \mathbf{P}(X \in W) - \mathbf{P}(X \in S) = 1 - 1 = 0$$

(X fällt mit Wahrscheinlichkeit 0 außerhalb von S.)

Die Abbildung $A \mapsto \rho(A) := \mathbf{P}(X \in A), \qquad A \subset S,$

heißt die **Verteilung** von X.

Die Zahlen $\rho(a) := \mathbf{P}(X = a), \qquad a \in S$,

sind die (atomaren) **Gewichte** der Verteilung ρ.

Wir nennen sie auch die **Verteilungsgewichte** von X.

Zufällige Paare und ihre Komponenten

(Buch S. 21)

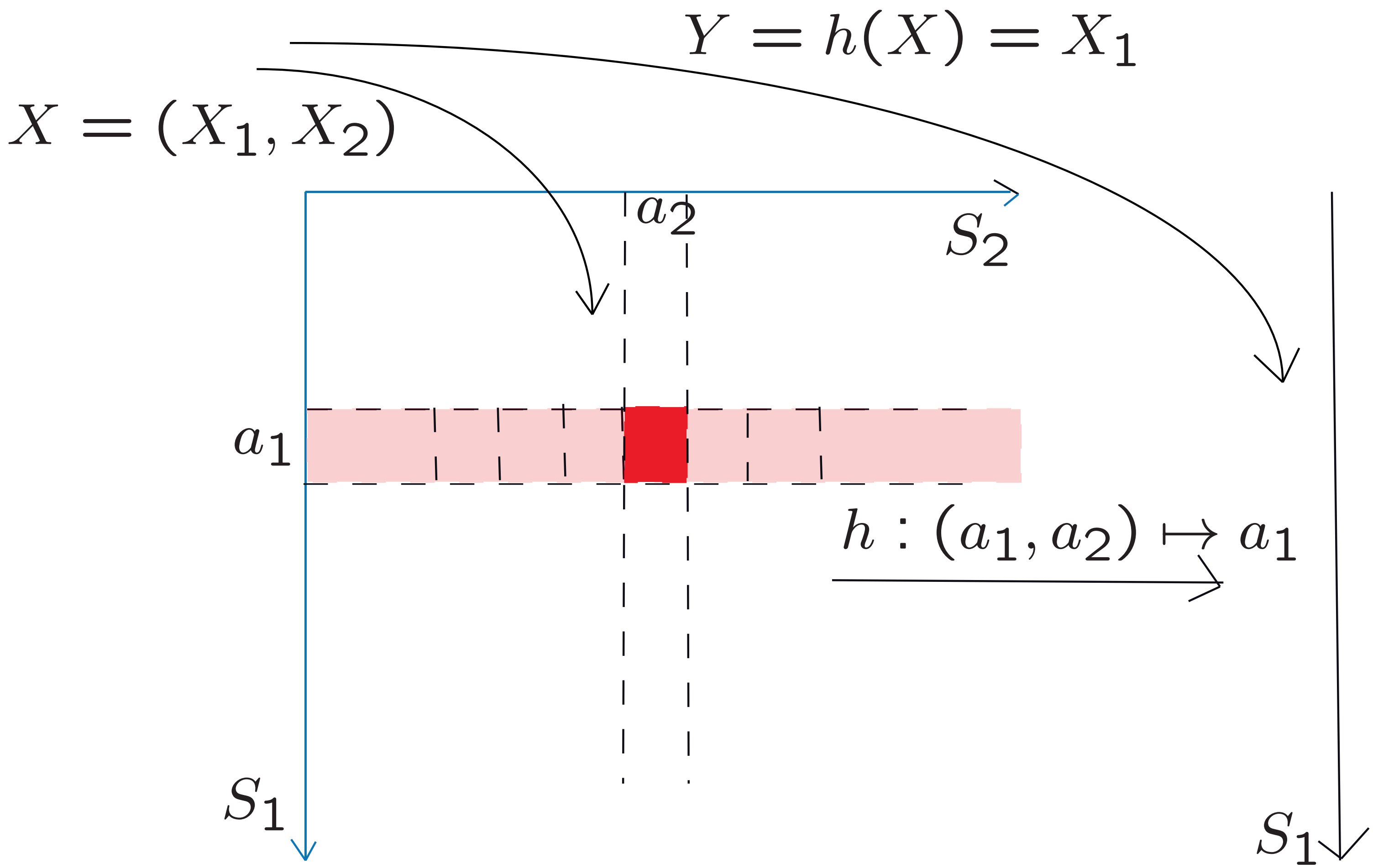
$X = (X_1, X_2)$ sei eine Zufallsvariable

(ein “zufälliges Paar”) mit diskreten Komponenten X_1, X_2.

Dann ist auch X diskret, denn die Diskretheit von S_1 und S_2 vererbt sich auf $S_1 \times S_2$, und $\mathbf{P}(X_i \in S_i) = 1$, $i = 1, 2$, impliziert $\mathbf{P}(X \in S_1 \times S_2) = 1$.

Die Ereignisse $\{(X_1, X_2) = (a_1, a_2)\}$ notieren wir auch als

$$\{X_1 = a_1, X_2 = a_2\}.$$


$Y = h(X) = X_1$
$X = (X_1, X_2)$
a_2
S_2
a_1
$h : (a_1, a_2) \mapsto a_1$
S_1
S_1

Die Verteilungsgewichte von X (man sagt auch:
die gemeinsamen Verteilungsgewichte von X_1 und X_2)
schreiben wir als

$$\begin{aligned}\rho(a_1, a_2) &= \mathbf{P}\big((X_1, X_2) = (a_1, a_2)\big)\\ &= \mathbf{P}(X_1 = a_1, X_2 = a_2)\ ,\end{aligned}$$

Die Verteilungsgewichte von X_1 erhält man
durch Summation über die a_2:

$$\rho_1(a_1) = \sum_{a_2 \in S_2} \rho(a_1, a_2)\ .$$

3. Weiterverarbeitung von Zufallsvariablen

und

Transport von Verteilungen

(Buch S. 21-22)

Der Übergang von $X = (X_1, X_2)$

zur Komponente X_1

ist ein Beispiel einer

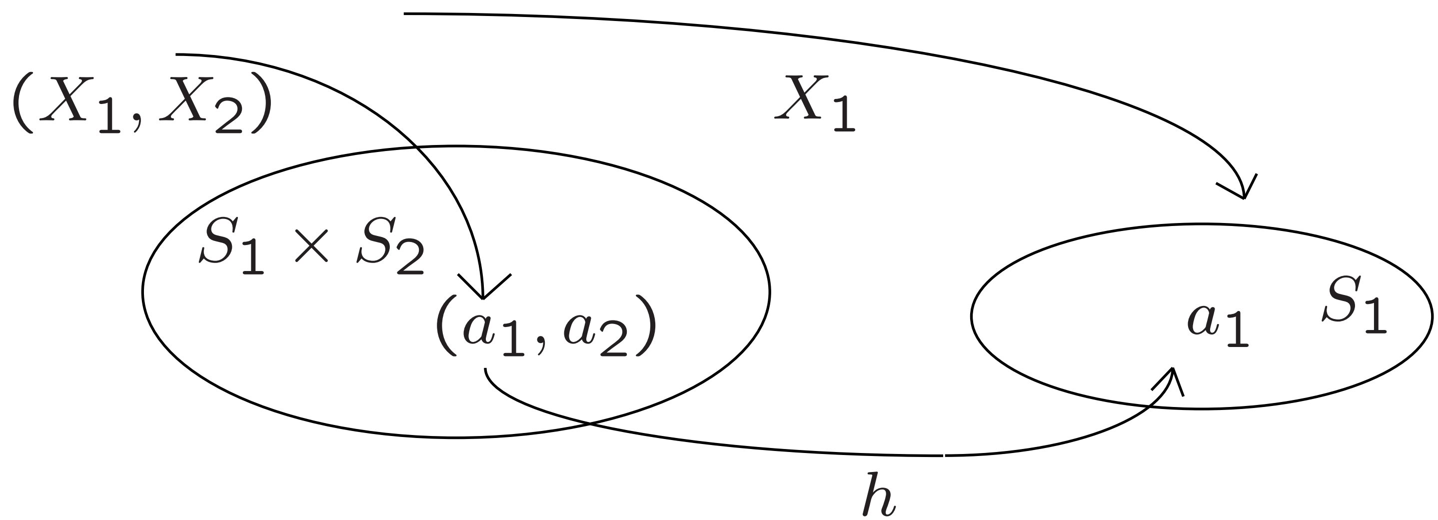
Weiterverarbeitung einer Zufallsvariablen:

$$X_1 = h(X)$$

mit $h((a_1, a_2)) := a_1$,

also die Projektion des Paares (a_1, a_2)

auf seine erste Komponente.


(X_1, X_2)
X_1
$S_1 \times S_2$
(a_1, a_2)
a_1
S_1
h

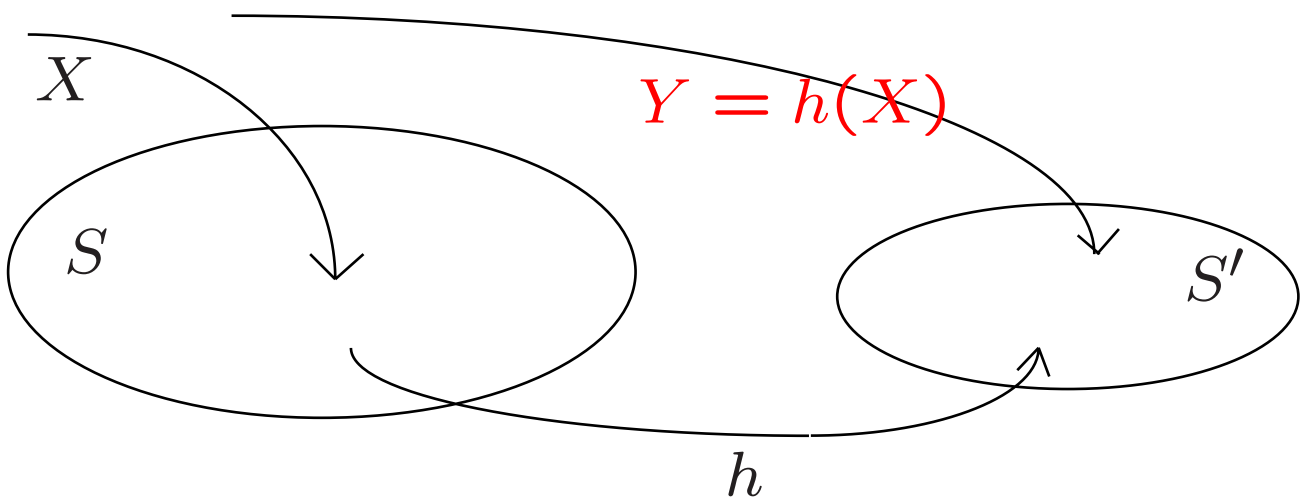
Sind S und S' zwei Mengen,

X eine Zufallsvariable mit Zielbereich S,

h eine Abbildung von S nach S',

und nimmt man X als zufällige Eingabe von h,

dann bekommt man eine Zufallsvariable Y mit Zielbereich S':

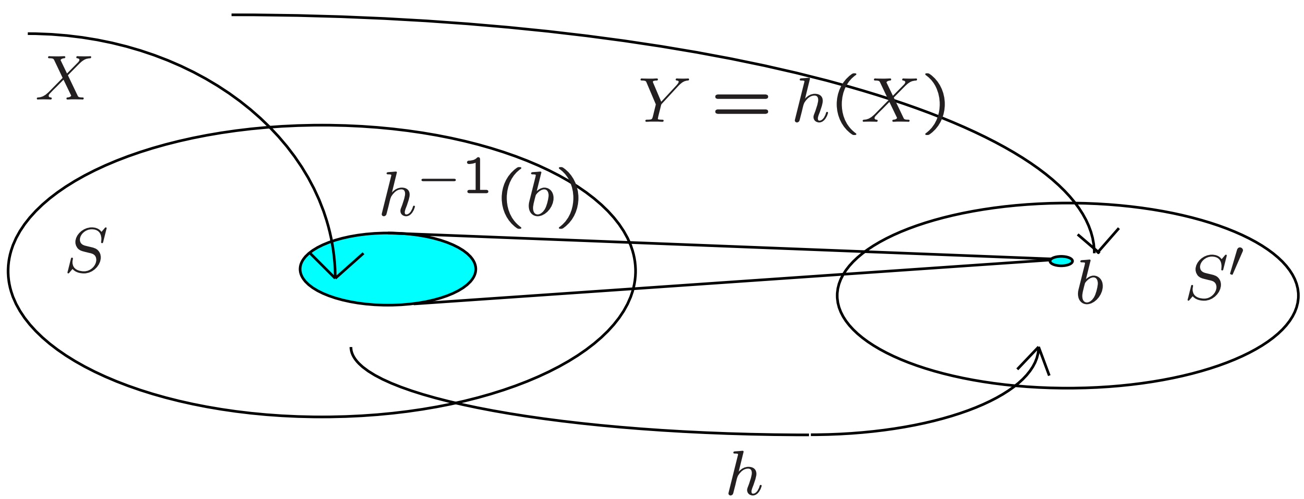


Für jedes $b \in S'$ gilt:

$$\{h(X) = b\} = \{X \in h^{-1}(b)\}.$$

Für die Verteilungsgewichte von $Y = h(X)$ ergibt sich:

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

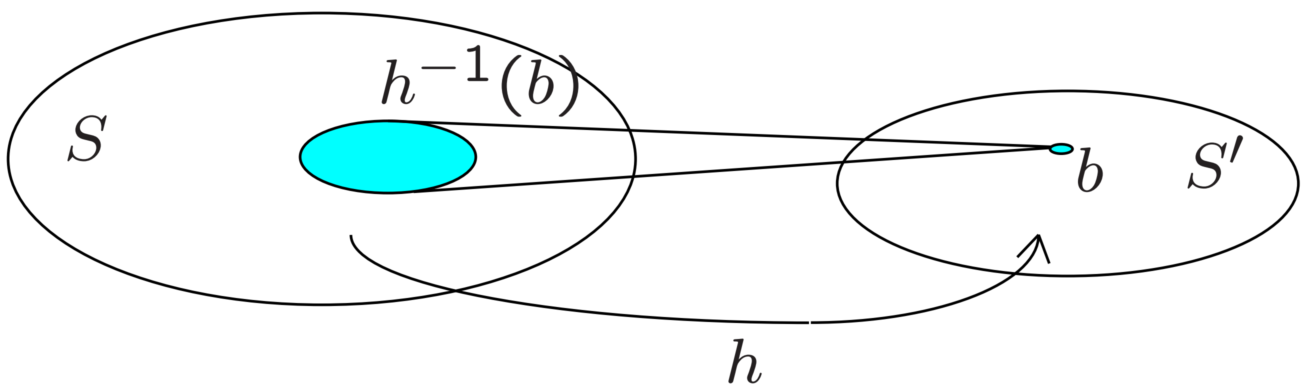


Bezeichnet ρ die Verteilung von X und ρ' die von Y,

dann ist

$$\rho'(b) = \sum_{a \in h^{-1}(b)} \rho(a).$$

Man sagt: Die Verteilung ρ wird durch die Abbildung h in die Verteilung ρ' transportiert.



Diese Situation haben wir schon mehrmals angetroffen:

in Vorlesung 1b:

$X :=$ rein zufällige $1, \ldots, g$-Folge der Länge $g + 1$

$T = h(X)$:= Zeitpunkt der ersten Kollision.

in Vorlesung 2a:

$X :=$ rein zufällige Permutation von $1, \ldots, n$

$h(X) :=$ Länge des Zyklus von X, der die Eins enthält.

Heutiges Programm:

Zusätzliche Beispiele für

“Weiterverarbeitungen von zufälligen Folgen”

Daraus bekommen wir wichtige Beispiele

diskreter Zufallsvariabler und diskreter Verteilungen.

Vorlesung 5a

Zufallsvariable mit Dichten:

Transformationen, Exponentialverteilung,

Normalverteilung

Teil 2

Exponentialverteilung

Definition: Die reellwertige Zufallsvariable X heißt standard-exponentialverteilt, falls

$$\mathbf{P}(X > b) = e^{-b}, \;\; b \geq 0.$$

Äquivalent dazu ist:

X ist Zufallsvariable mit Verteilungsfunktion

$$b \mapsto \begin{cases} 0 & \text{für } b < 0 \\ 1 - e^{-b} & \text{für } b \geq 0. \end{cases}$$

Äquivalent dazu ist:

X ist Zufallsvariable mit Dichte $e^{-a}\,da, \quad a \geq 0.$

Ist X standard-exponentialverteilt,

dann hat $\frac{X}{\alpha}$ die Dichte

$$\alpha e^{-\alpha b} db, \quad b \geq 0.$$

Eine Zufallsvariable Y mit dieser Dichte heit

exponentialverteilt mit Parameter α, kurz Exp(α)-verteilt.

Merke: Ist Y Exp(α)-verteilt, dann ist αY Exp(1)-verteilt.

Von der geometrischen zur Exponentialverteilung

Betrachten wir eine Geom(p)-verteilte Zufallsvariable Y mit

$$p = \frac{1}{100}.$$

Es gilt: $\mathbf{E}[Y] = 100$, also hat $\frac{Y}{100}$ Erwartungswert 1.

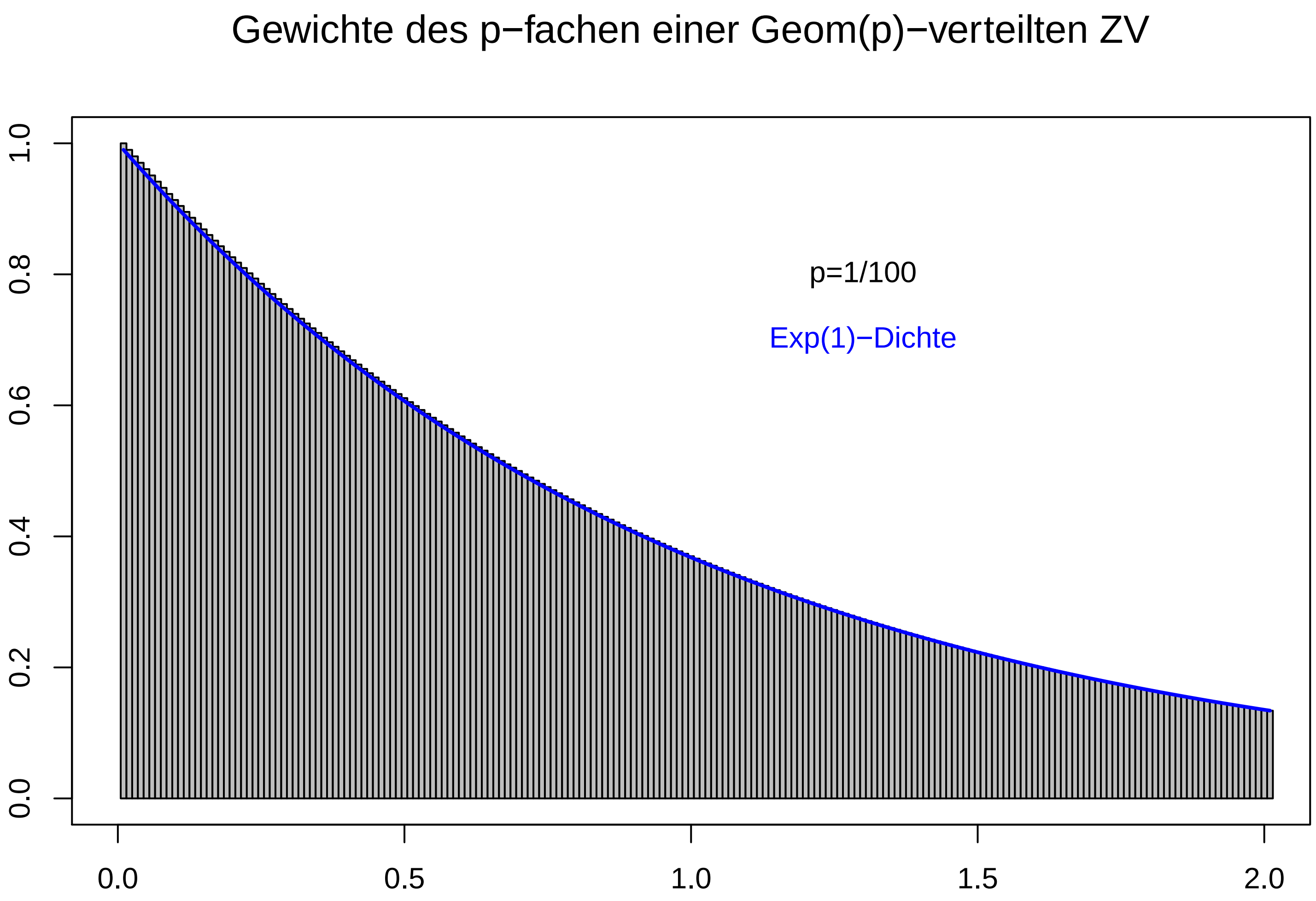
$$\mathbf{P}\left(\frac{Y}{100} = \frac{k}{100}\right) = \mathbf{P}(Y = k)$$

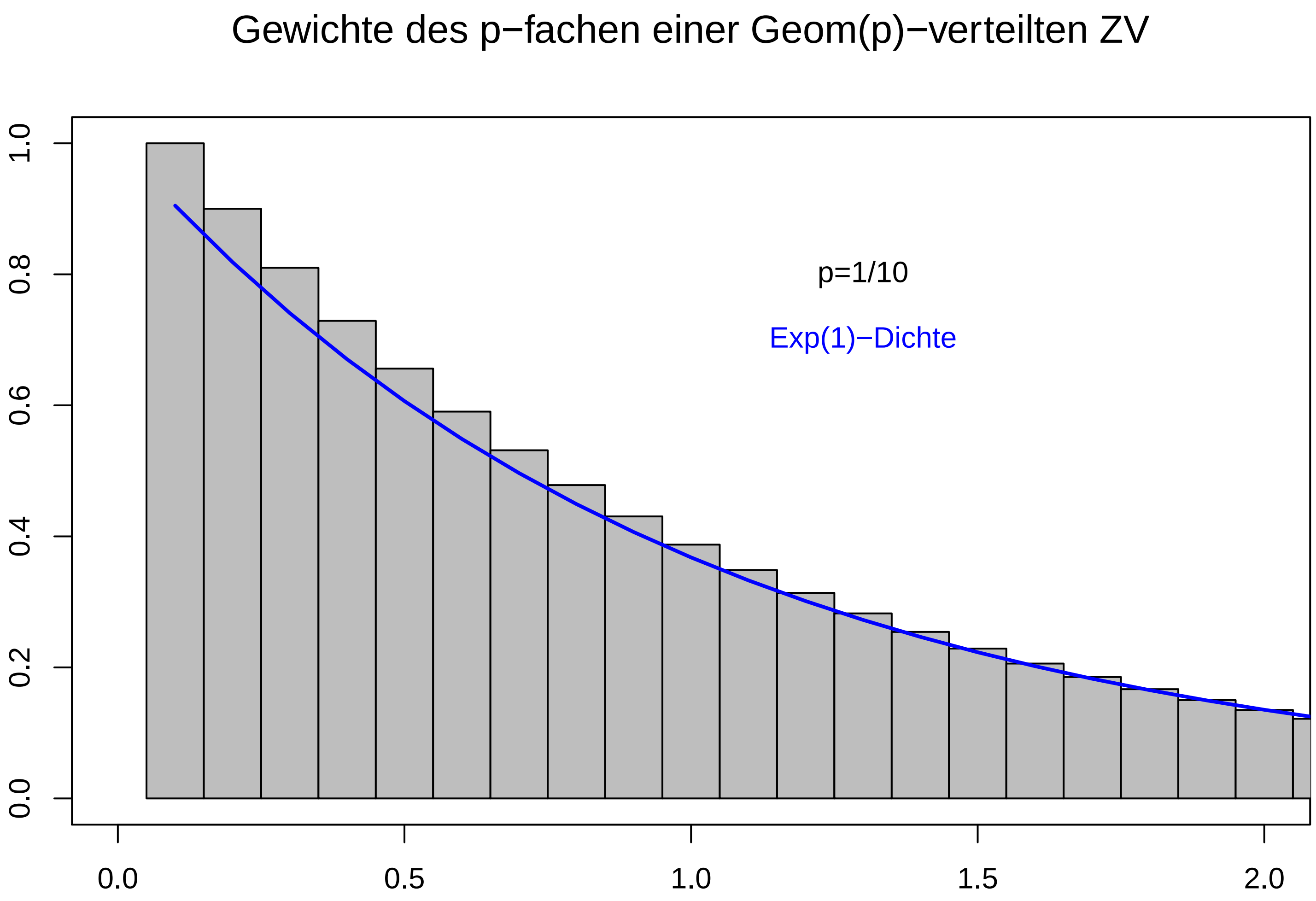
$$= \frac{1}{100}\left(1 - \frac{1}{100}\right)^{k-1} = \frac{1}{100}\left(\left(1 - \frac{1}{100}\right)^{100}\right)^{(k-1)/100}$$

$$\approx \frac{1}{100}\, e^{-(k-1)/100} \approx \mathbf{P}\left(X \in \left[\frac{k-1}{100}, \frac{k}{100}\right]\right)$$

für ein standard-exponentialverteiltes X.

Das wird durch das folgende Bild veranschaulicht:


Gewichte des p−fachen einer Geom(p)−verteilten ZV
p=1/100
Exp(1)−Dichte
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0


Gewichte des p−fachen einer Geom(p)−verteilten ZV
p=1/10
Exp(1)−Dichte
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0

Salopp gesprochen:

Für kleines p hat eine Geom(p)-verteilte Zufallsvariable Y einen großen Erwartungswert, nämlich $\frac{1}{p}$.

Man holt die Verteilung von Y zurück ins Schaubild, indem man pY betrachtet.

Wir erinnern an die “Exponentialapproximation”:

(Vorlesung 4a und Buch Seite 42):

Ist für jedes $n \in \mathbb{N}$ die Zufallsvariable X_n geometrisch verteilt mit $\mathbf{E}[X_n] \to \infty$ für $n \to \infty$, dann gilt:

$$\mathbf{P}\Big(\frac{X_n}{\mathbf{E}[X_n]} > t\Big) \to e^{-t} \quad \text{für alle } t \in \mathbb{R}_+.$$

$$\mathbf{P}\Big(\frac{X_n}{\mathbf{E}[X_n]} > t\Big) \to e^{-t} \quad \text{für alle } t \in \mathbb{R}_+.$$

Also:

$$\mathbf{P}\Big(\frac{X_n}{\mathbf{E}[X_n]} > t\Big) \to \mathbf{P}(X > t), \quad t \in \mathbb{R}_+,$$

wobei X eine standard-exponentialverteilte Zufallsvariable ist.

Man sagt dafür auch:

Die Folge der Zufallsvariablen $X_n/\mathbf{E}[X_n]$

konvergiert in Verteilung

gegen die Zufallsvariable X.

Vorlesung 5a

Zufallsvariable mit Dichten:

Transformationen, Exponentialverteilung, Normalverteilung

Teil 4

Die Standardnormalverteilung auf $\mathbb{R}$

Die folgende Definition beinhaltet

die wichtigste Verteilung der Stochastik:

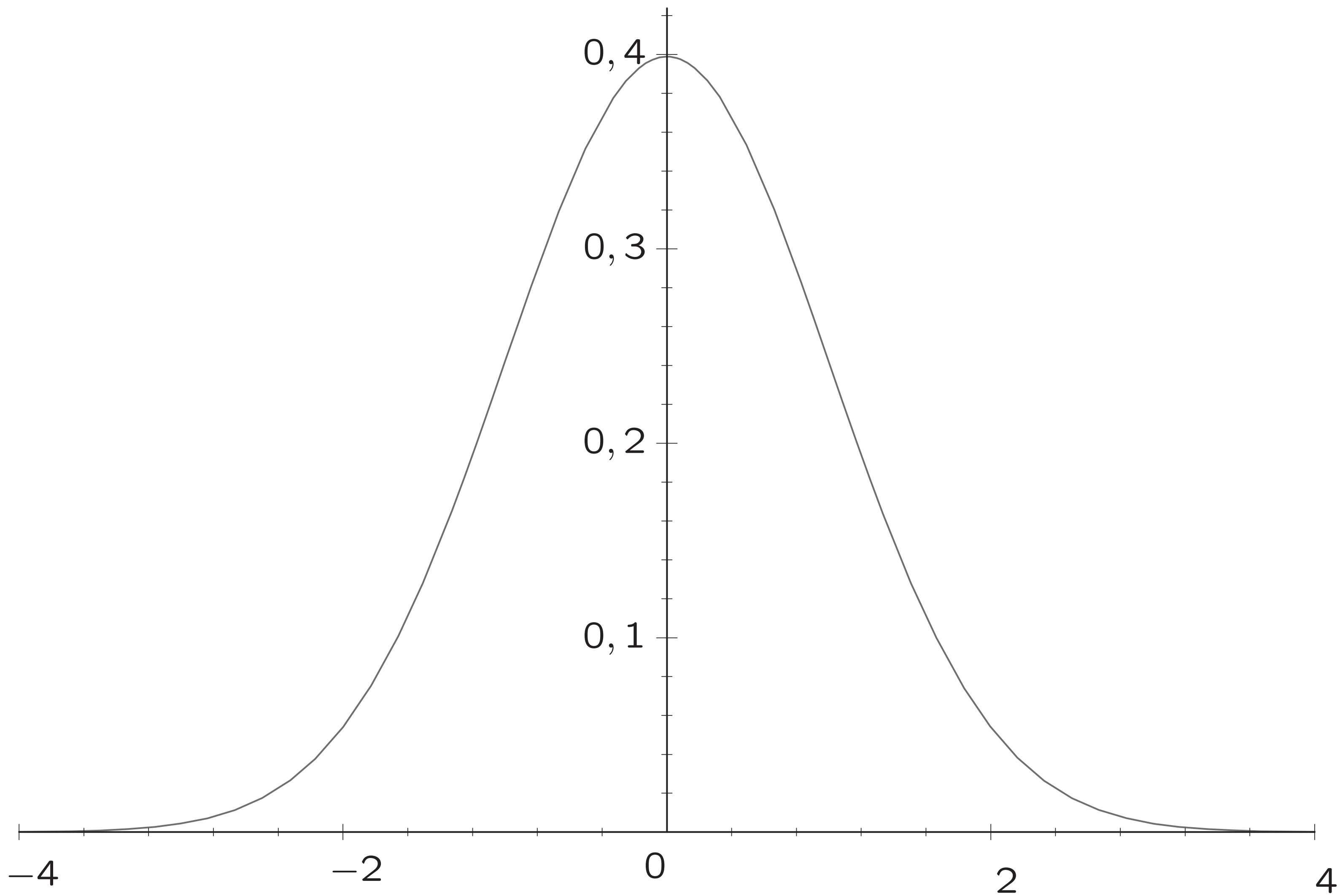
Eine $\mathbb{R}$-wertige Zufallsvariable Z mit Dichte

$$\boxed{\varphi(a)\,da := \frac{1}{\sqrt{2\pi}} e^{-a^2/2}\,da}$$

heißt standard-normalverteilt.

Wir werden bald auf geometrischem Wege sehen, dass gilt
(vgl. die erste Aufgabe auf S. 72 im Buch):

$$\int_{-\infty}^{+\infty} e^{-a^2/2} = \sqrt{2\pi}, \quad \text{also} \int_{-\infty}^{+\infty} \varphi(a)\,da = 1.$$


0, 4
0, 3
0, 2
0, 1
−4
−2
0
2
4

Für ein standard-normalverteiltes Z ist

$$\mathbf{E}[Z] = 0, \quad \mathbf{E}[Z^2] = 1.$$

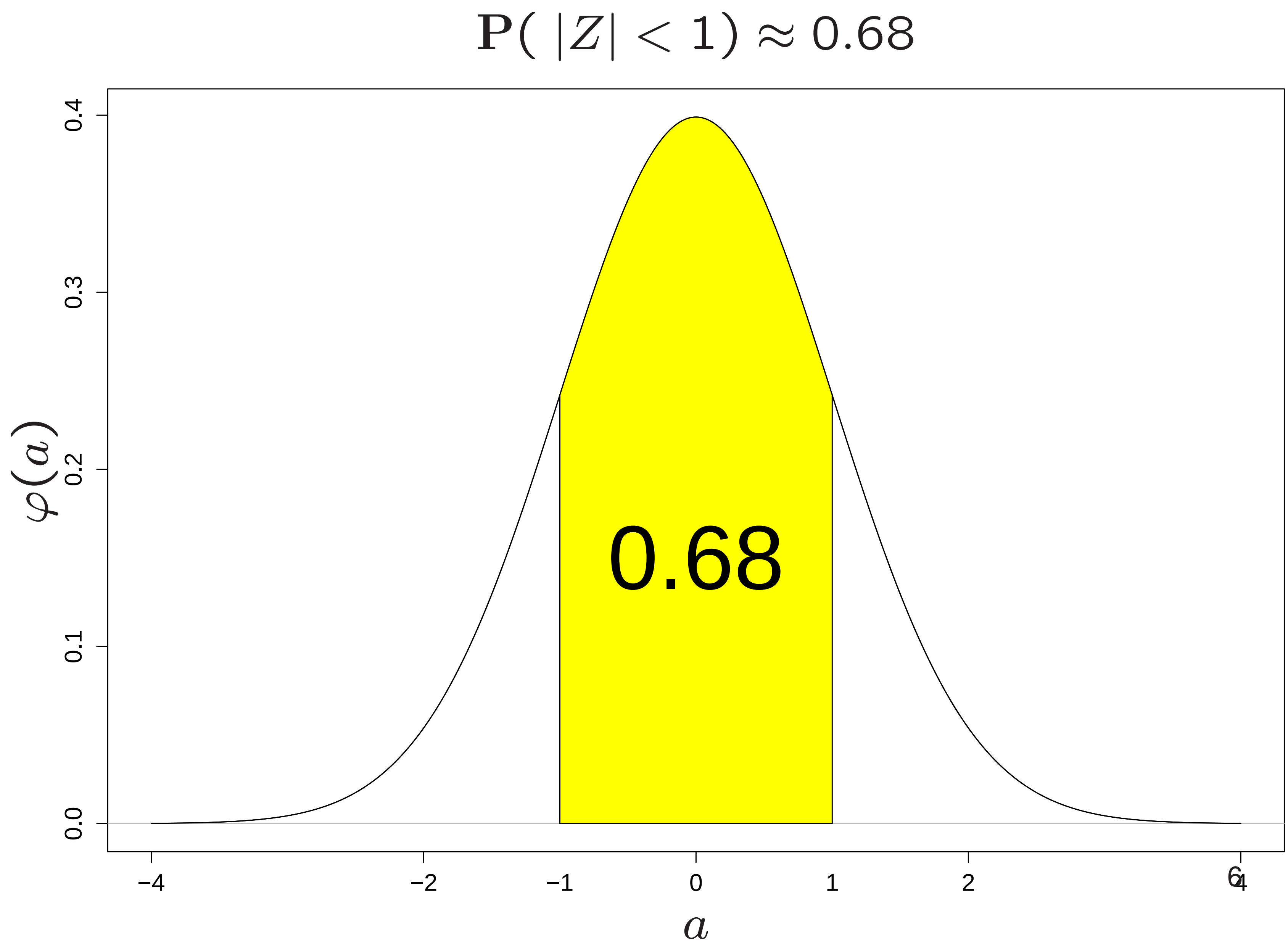
Denn aus Symmetriegründen ist $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a\, e^{-a^2/2}\, da = 0$,

und mit partieller Integration bekommt man

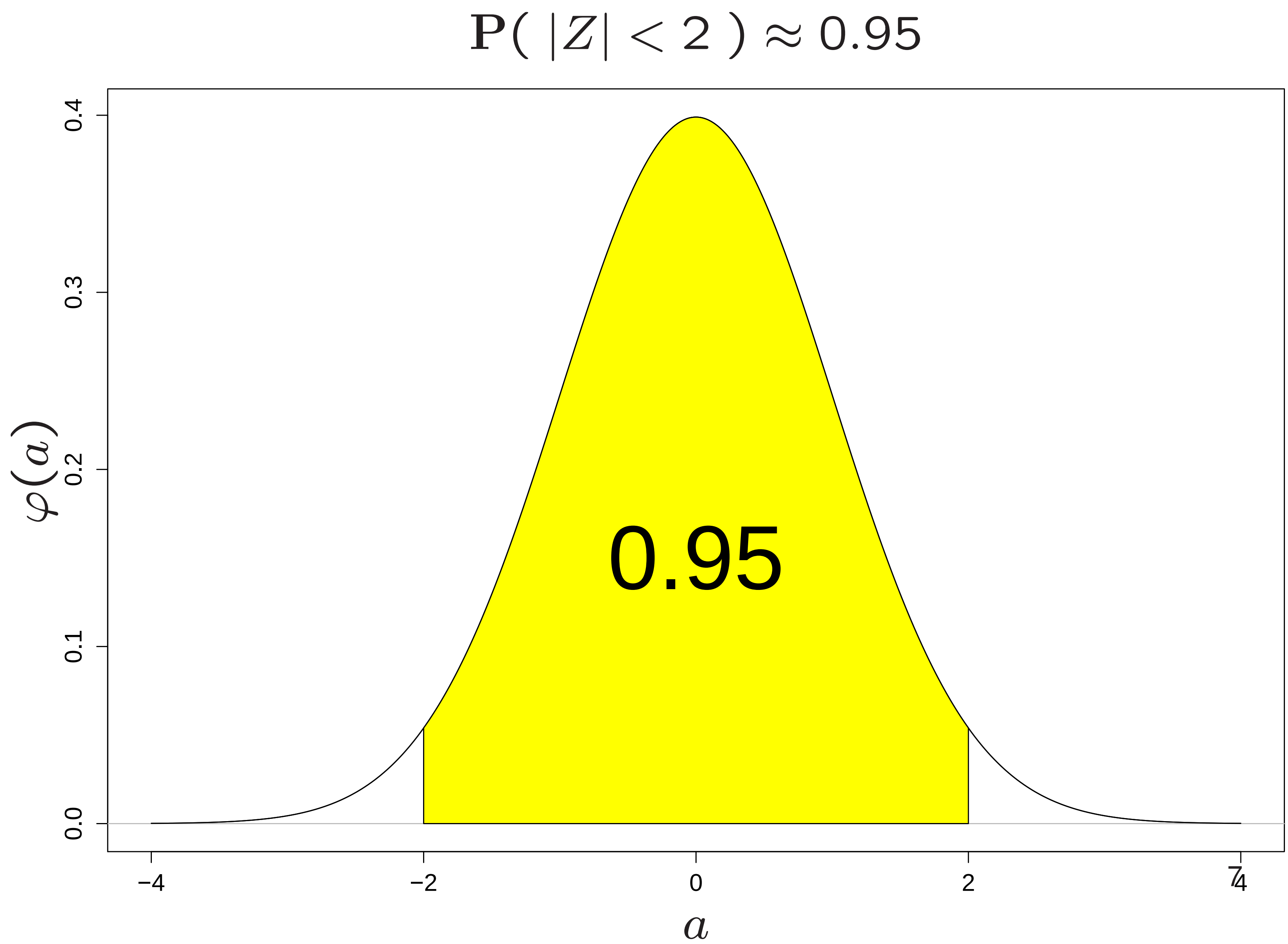
$$\int_{-\infty}^{\infty} a^2\, e^{-a^2/2}\, da = \int_{-\infty}^{\infty} a\, a e^{-a^2/2}\, da$$

$$= -a\, e^{-a^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-a^2/2}\, da = 0 + \sqrt{2\pi}.$$

Zwei für die Praxis wichtige Zahlen:



$\mathbf{P}(\,|Z| < 1) \approx 0.68$
0.68
$\varphi(a)$
0.0
0.1
0.2
0.3
0.4
−4
−2
−1
0
1
2
4
6
a


$\mathbf{P}(\ |Z| < 2\ ) \approx 0.95$
0.95
$\varphi(a)$
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
a

$$\Phi(b) := \int_{-\infty}^{b} \varphi(a)\, da$$

ist die Standard-Normalverteilungsfunktion.

Es gibt für sie keinen expliziten analytischen Ausdruck

(der ohne die Formulierung als Integral bzw. Stammfunktion auskommt).

Der R-Befehl dafür ist `pnorm(b)` .

Affin lineare Transformation einer standard-normalverteilten ZV'e

Sei Z standard-normalverteilt, $\mu \in \mathbb{R}$, $\sigma > 0$. Dann hat

$$X := \sigma Z + \mu \ :$$

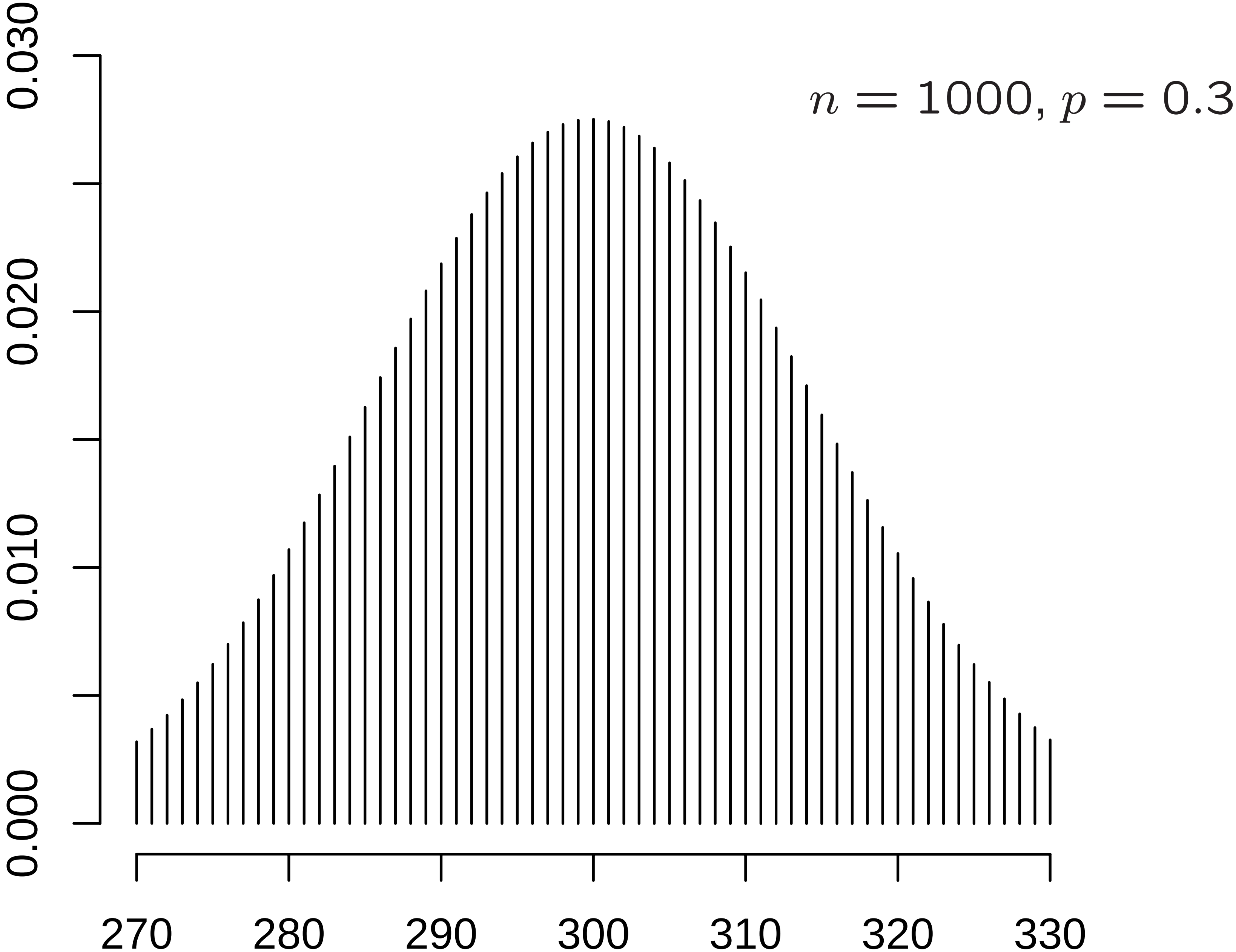
die Dichte

$$\frac{1}{\sigma}\varphi\left(\frac{a-\mu}{\sigma}\right)\,da = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(a-\mu)^2}{2\sigma^2}}\,da\,.$$

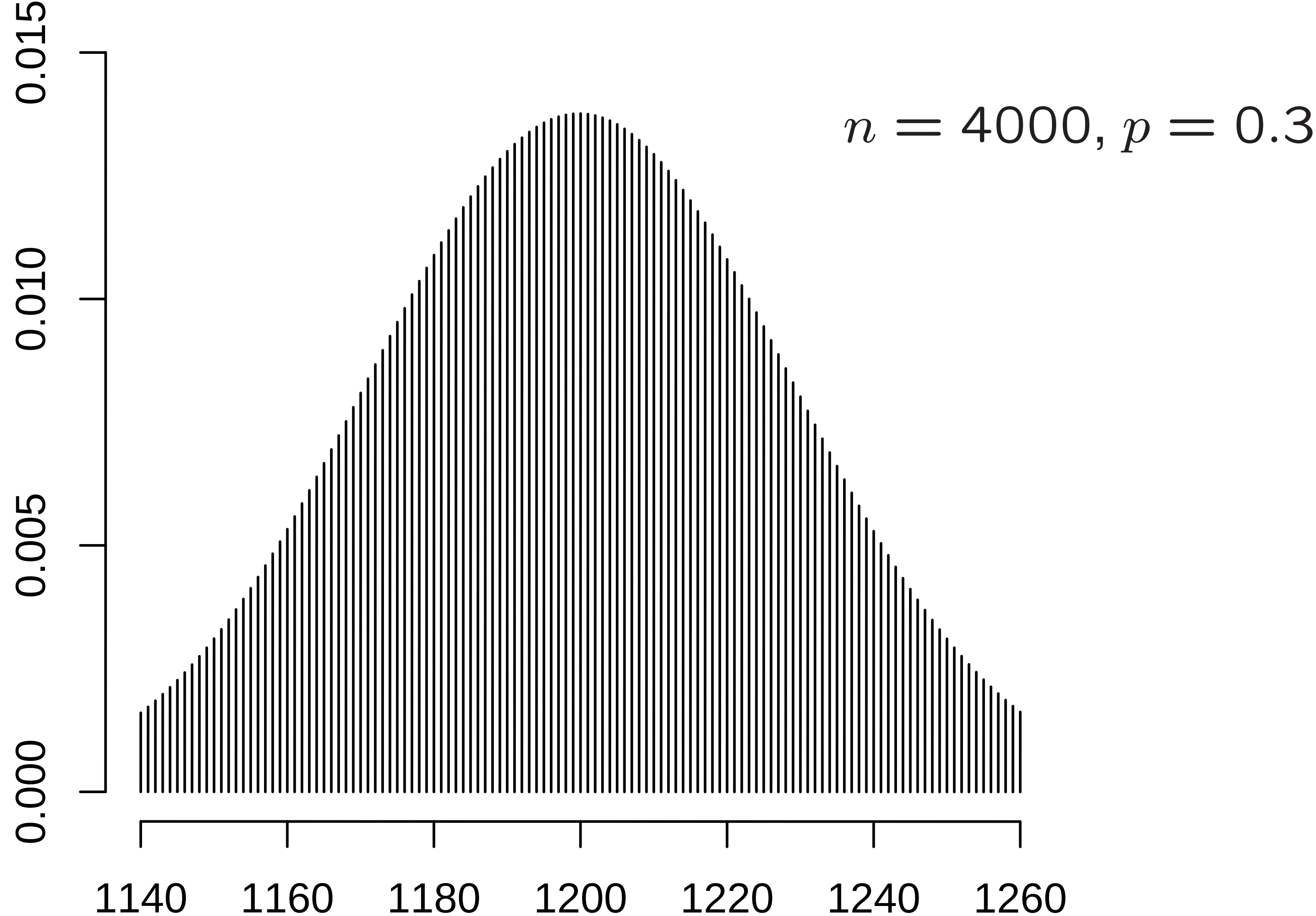
Die Gaußsche Glockenkurve sieht man auch entstehen

aus den Binomialgewichten mit großem n und großem npq,

wenn man sie “geeignet ins Bild bringt”:



$n = 1000, p = 0.3$
0.000
0.010
0.020
0.030
270
280
290
300
310
320
330


0.015
0.010
0.005
0.000
$n = 4000, p = 0.3$
1140
1160
1180
1200
1220
1240
1260

Mehr dazu bald in der Vorlesung!

Vorlesung 5a

Zufallsvariable mit Dichten:

Transformationen, Exponentialverteilung, Normalverteilung

Teil 3

Erwartungswert und Transformationsformel

Für diskrete reellwertige Zufallsvariable hatten wir

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

Das hat sein Analogon im Fall reellwertiger ZV'er mit Dichten:

Den Verteilungsgewichten $\rho(a)$ entspricht die Dichte $f(a)\, da$.

Und aus der Summe wird ein Integral:

$$\mu = \mathbf{E}[X] = \int_l^r a\, f(a)\, da$$

Im Diskreten hatten wir für $h : S \to \mathbb{R}$

die Transformationsformel

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

Analog gilt im Fall mit Dichten:

$$\mathbf{E}[h(X)] = \int_l^r h(a)\,f(a)\,da$$

Der Erwartungswert einer standard-exponentialverteilten Zufallsvariablen X:

$$\mathbf{E}[X] = \int_0^\infty x\, e^{-x}\, dx$$

Mit partieller Integration

$$\int u\, v' = uv - \int u' v$$

ergibt sich

$$\int_0^\infty x\, e^{-x}\, dx = -xe^{-x}\Big|_0^\infty + \int_0^\infty 1 \cdot e^{-x} dx = 1$$

Also:

$$\boxed{\mathbf{E}[X] = 1.}$$

Der Erwartungswert des Quadrates einer standard-exponentialverteilten Zufallsvariablen X:

$$\mathbf{E}\left[X^2\right] = \int_0^\infty x^2\, e^{-x}\, dx$$

Wieder mit partieller Integration:

$$\int_0^\infty x^2\, e^{-x}\, dx = -x^2 e^{-x}\Big|_0^\infty + \int_0^\infty 2x \cdot e^{-x} dx = 2$$

Also:

$$\boxed{\mathbf{E}\left[X^2\right] = 2.}$$

Der Erwartungswert einer Exp(α)-verteilten Zufallsvariablen Y:

Wir wissen schon: Ist Y Exp(α)-verteilt,

dann ist αY Exp(1)-verteilt.

Also ist

$$\mathbf{E}[Y] = \frac{1}{\alpha}\mathbf{E}[\alpha Y] = \frac{1}{\alpha} \cdot 1.$$

$$\boxed{\mathbf{E}[Y] = \frac{1}{\alpha}.}$$

Vorlesung 5a

Zufallsvariable mit Dichten:

Transformationen, Exponentialverteilung, Normalverteilung

Teil 1

Transformationen

Drei Beispiele:

A. Sei U uniform verteilt auf $[0, 1]$. Gefragt ist nach Verteilungsfunktion und Dichte von $X := U^2$.

$$F_X(b) = \mathbf{P}(X \le b) = \mathbf{P}(U \le \sqrt{b}) = \sqrt{b}, \quad 0 \le b \le 1.$$

$$\sqrt{b} \overset{!}{=} \int_0^b f(a)\, da, \quad 0 \le b \le 1.$$

$$f(b) = \begin{cases} \frac{1}{2\sqrt{b}}, & 0 < b \le 1 \\ 0 & \text{sonst} \end{cases}$$ ist Dichtefunktion von X.

B. Sei U uniform verteilt auf $[0,2]$. Gefragt ist nach Verteilungsfunktion und Dichte von $X := U^2$.

X hat Wertebereich $[0,4]$.

$$F_X(b) = \mathbf{P}(X \le b) = \mathbf{P}(U \le \sqrt{b}) = \frac{1}{2}\sqrt{b}, \quad 0 < b \le 4.$$

$$\frac{1}{2}\sqrt{b} \overset{!}{=} \int_0^b f(a)\,da, \quad 0 < b \le 4.$$

$$f(b) = \begin{cases} \frac{1}{4}\frac{1}{\sqrt{b}}, & 0 \le b \le 4 \\ 0 & \text{sonst} \end{cases} \quad \text{ist Dichtefunktion von} X.$$

C. Sei U uniform verteilt auf $[0,1]$.

Gefragt ist nach der Dichte von $X := -\ln U$.

$$\begin{aligned}\mathbf{P}(X \le b) &= \mathbf{P}(-\ln U \le b) = \mathbf{P}(\ln U \ge -b)\\ &= \mathbf{P}(U \ge e^{-b}) = \mathbf{P}(U \in [e^{-b}, 1])\\ &= 1 - e^{-b} \overset{!}{=} \int_0^b f(a)\, da, \quad b \ge 0.\end{aligned}$$

$$f(b) = \begin{cases} e^{-b}, & b \ge 0 \\ 0 & \text{sonst} \end{cases}$$ ist Dichtefunktion von X.

Zufallsvariable X mit der Eigenschaft

$$\mathbf{P}(X > b) = e^{-b}, \;\; b \geq 0,$$

sind uns schon (implizit) begegnet

bei der Approximation der Verteilung von pT;

dabei war T Geom(p)-verteilt mit kleinem p.

Wir sprachen damals von der

Exponentialapproximation der geometrischen Verteilung,

siehe V4a.

Affin lineare Transformation:

X habe Verteilungsfunktion F_X.

Was ist dann die Verteilungsfunktion von $Y := \beta X + \gamma$?

Dabei sei $\beta > 0$.

$$
\begin{aligned}
&F_Y(b)\\
&= \mathbf{P}(Y \le b) = \mathbf{P}(\beta X + \gamma \le b)\\
&= \mathbf{P}(\beta X \le b - \gamma) = \mathbf{P}\left(X \le \frac{b-\gamma}{\beta}\right)\\
&= F_X\left(\frac{b-\gamma}{\beta}\right)
\end{aligned}
$$

X habe Dichte $f_X(a)da$.

Was ist dann die Dichte von $Y := \beta X + \gamma$?

Der Einfachheit halber nehmen wir an: Die Dichtefunktion f_X ist stückweise stetig. Dann ist in allen Stetigkeitspunkten

$$f_Y(b) = F_Y'(b) = F_X'\left(\frac{b-\gamma}{\beta}\right)\frac{1}{\beta}$$

Die Dichte von Y ist somit $f_Y(b)\,db = f_X'\left(\frac{b-\gamma}{\beta}\right)\frac{db}{\beta}$

Zum Merken: Schlag nach beim Urbild,

und vergiss den Streckungsfaktor nicht!

Vorlesung 4b

Zufallsvariable mit Dichten

Teil 1

Kontinuierlich uniform verteilte Zufallsvariable

Kontinuierlich anstelle von diskret:

Bisher hatten wir im Fokus der Vorlesung:

Diskrete Zufallsvariable.

Sie fallen mit W'keit 1 in eine diskrete

(d.h. endliche oder abzählbar unendliche)

Menge S.

Jetzt wenden wir uns Zufallsvariablen zu,

die kontinuierlich verteilt sind.

Dann ist der Wertebereich *überabzählbar*.

Ein prominentes Beispiel ist der (faire) Münzwurf.

Man kann ihn auffassen als

rein zufällige Wahl eines Elementes aus $\{0,1\}^{\mathbb{N}}$.

Ein weiteres einprägsames Beispiel ist die rein zufällige Wahl

eines Punktes aus dem Einheitsintervall

oder aus dem Einheitsquadrat.

Idee bei der rein zufälligen Wahl aus einem Kontinuum:

$\mathbf{P}(X \in A)$ ist gegeben durch den Anteil von A an S

Uniforme Verteilung auf dem Einheitsintervall

Eine Zufallsvariable X mit Wertebereich $S = [0, 1]$ heißt

uniform verteilt auf S,

wenn für alle $A \subset S$ mit wohldefiniertem Längenmaß $\lambda(A)$

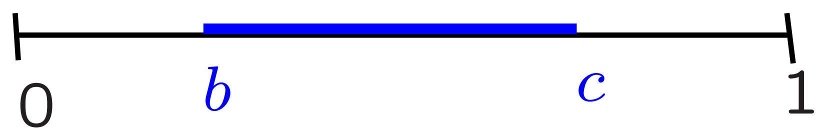
gilt:

$$\mathbf{P}(X \in A) = \lambda(A)$$

(denn hier ist ja $\lambda(S) = 1$).

Beispiel 1:

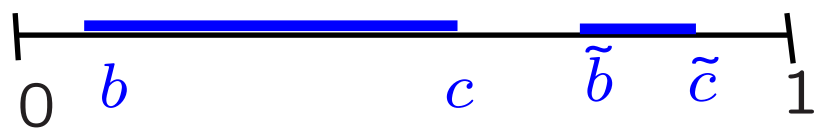
$A := [b, c]$ mit $0 \leq b \leq c \leq 1$



$$\mathbf{P}(X \in A) = c - b.$$

Beispiel 2:

$A := [b, c] \cup [\tilde{b}, \tilde{c}]$ mit $0 \leq b \leq c < \tilde{b} \leq \tilde{c} \leq 1$



$$\mathbf{P}(X \in A) = (c - b) + (\tilde{c} - \tilde{b}).$$

Beispiel 3:

$$A := \{a\} \quad \text{mit } 0 \le a \le 1$$

$$\mathbf{P}(X = a) = a - a = 0.$$

Uniforme Verteilung auf einem Rechteck in $\mathbb{R}^2$:

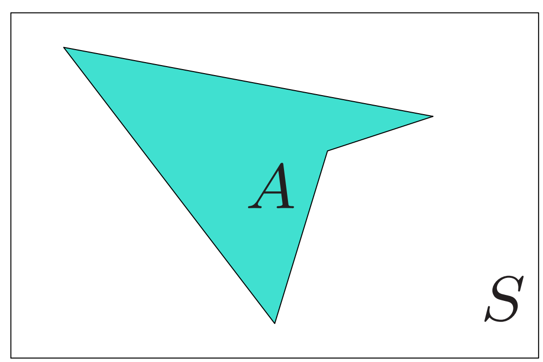
Eine Zufallsvariable X mit Zielbereich

$$S := [0, \ell] \times [0, b] \subset \mathbb{R}^2$$

heißt *uniform verteilt auf* S, wenn

für alle $A \subset S$ mit wohldefiniertem Flächenmaß $\lambda^2(A)$ gilt:

$$\mathbf{P}(X \in A) = \frac{\lambda^2(A)}{\lambda^2(S)} = \frac{\lambda^2(A)}{\ell \cdot b}.$$



Uniforme Verteilung auf einer Teilmenge des $\mathbb{R}^d$

Definition (Buch S. 12)

Sei S eine Teilmenge des $\mathbb{R}^d$ mit endlichem Inhalt $\lambda^d(S) > 0$.

Eine Zufallsvariable X mit Wertebereich S heißt

uniform verteilt auf S,

wenn für alle $A \subset S$ mit wohldefiniertem (Volums-)Inhalt

$v(A) := \lambda^d(A)$ gilt:

$$\mathbf{P}(X \in A) = \frac{v(A)}{v(S)}.$$

$$\mathbf{P}(X \in A) = \frac{v(A)}{v(S)}.$$

Man beachte die Analogie zu

“Anzahl günstige durch Anzahl mögliche Fälle”:

Der zahlenmäßige Anteil von A

an einem endlichen Wertebereich S

wird jetzt ersetzt durch den volumsmäßigen Anteil von A

am (überabzählbar) unendlichen Wertebereich S.

Wie im Diskreten werden wir uns nicht nur

mit *rein* zufälliger Wahl begnügen.

Vorlesung 4b

Zufallsvariable mit Dichten

Teil 3

Verteilungsfunktionen

Wieder sei S ein Intervall in $\mathbb{R}$, und X eine (diskrete oder kontinuierliche) Zufallsvariable mit Wertebereich S.

Die Funktion $F(b) := F_X(b) := \mathbf{P}(X \le b), \quad b \in \mathbb{R}$,

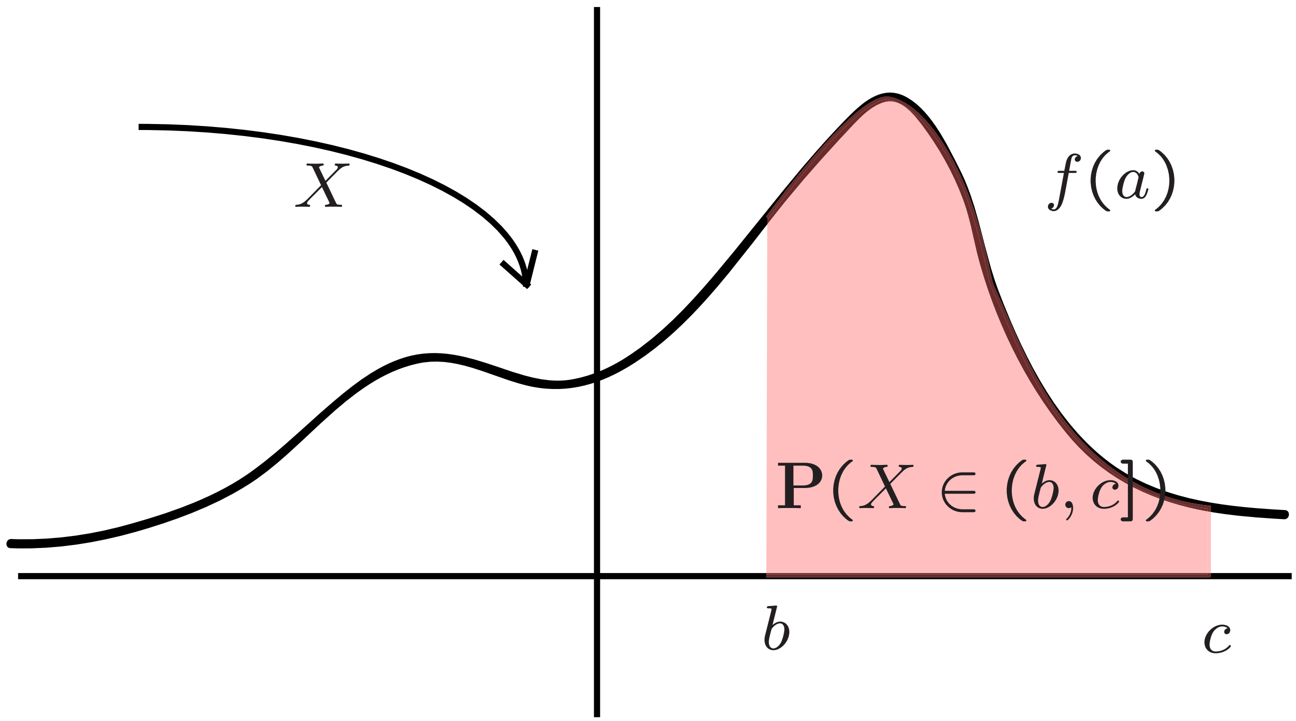
heißt *Verteilungsfunktion* von X.

Hat X die Dichte $f(a)\,da$, so gilt

$$F(b) = \int_{-\infty}^{b} f(a)\,da\ , \quad b \in \mathbb{R}$$

(mit $f(a) := 0$ für $a \notin S$)

Ist f stetig in a, dann ist $f(a) = F'(a)$.



$$\mathbf{P}(X \leq c) - \mathbf{P}(X \leq b) = \mathbf{P}(b < X \leq c)$$

$$F(c) - F(b) = \int_b^c f(a)da$$

Man findet den Hauptsatz der Differential-und Integralrechnung wieder!

Ein hinreichendes Kriterium für die Existenz einer Dichte.

Sei F die Verteilungsfunktion einer reellwertigen Zufallsvariablen X. Hat F keine Sprünge *und* ist F *stückweise stetig differenzierbar**, dann besitzt X eine Dichte.

Denn für jeden Randpunkt a eines der Intervalle gilt: $\mathbf{P}(X = a) = 0$ (ansonsten hätte F in a einen Sprung).
Und innerhalb eines jeden Intervalls gilt nach Voraussetzung der Hauptsatz der Differential-und Integralrechnung, siehe die vorige Folie.

*d.h. es gibt endlich viele disjunkte Intervalle, deren Vereinigung $\mathbb{R}$ ist, so dass F eingeschränkt auf jedes dieser Intervalle eine stetige Ableitung hat.

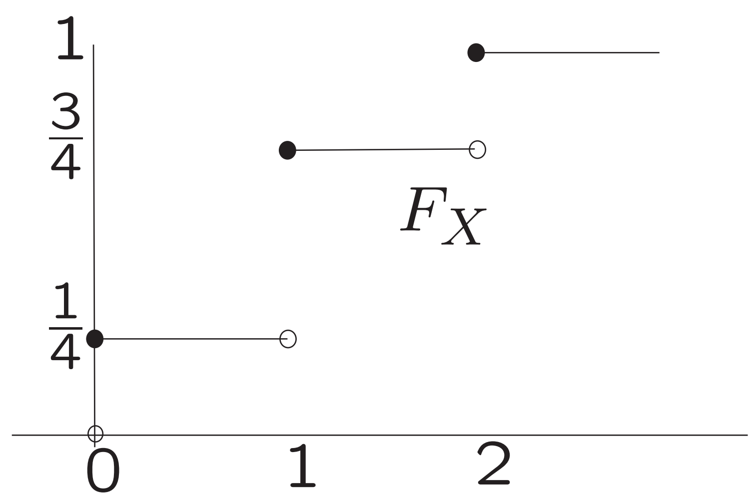
Für *diskrete* reellwertige Zufallsvariable X

ist F_X stückweise konstant, mit Sprüngen der Höhe

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

Beispiel:

X Binomial$(2, 1/2)$-verteilt



Vorlesung 4b

Zufallsvariable mit Dichten

Teil 2

Dichten auf $\mathbb{R}$ bzgl des Längenmaßes

Beginnen wir mit einem **Beispiel**:

X sei eine $\mathbb{R}_+$-wertige Zufallsvariable mit

$$\mathbf{P}(X > b) = e^{-b} \quad \text{für alle } b \geq 0.$$

$$\mathbf{P}(X \geq b) = ?$$

$$\mathbf{P}(X = b) = ?$$

Für $b = 0$ ist

$$\mathbf{P}(X \geq 0) = 1 = e^{-0} = \mathbf{P}(X > 0).$$

Für $b = 0$ ist

$$\mathbf{P}(X \geq 0) = 1 = e^{-0} = \mathbf{P}(X > 0).$$

Für $b > 0$ und $0 < \varepsilon < b$ ist

$$e^{-b} = \mathbf{P}(X > 0) \leq \mathbf{P}(b \leq X)$$

$$\leq \mathbf{P}(b - \varepsilon < X) = e^{-b+\varepsilon}$$

$$\Longrightarrow \mathbf{P}(X \geq b) = e^{-b} = \mathbf{P}(X > b).$$

$$\mathbf{P}(X = b) = \mathbf{P}(X \geq b) - \mathbf{P}(X > b) = 0.$$

Für alle $0 \le b \le c$ folgt:

$$\mathbf{P}(b \le X \le c) = \mathbf{P}(b \le X) - \mathbf{P}(c < X)$$

$$= e^{-b} - e^{-c} = \int_b^c e^{-a}\, da.$$

Dieses Beispiel wird im Folgenden verallgemeinert:

Sei S ein (offenes, abgeschlossenes oder halboffenes)

Intervall mit linker Grenze l und rechter Grenze r,

und X eine $\overline{\mathbb{R}}$- wertige Zufallsvariable,

für die $\{X \in S\}$ das sichere Ereignis ist.

(Dabei ist $l = -\infty$ und/oder $r = +\infty$ erlaubt.)

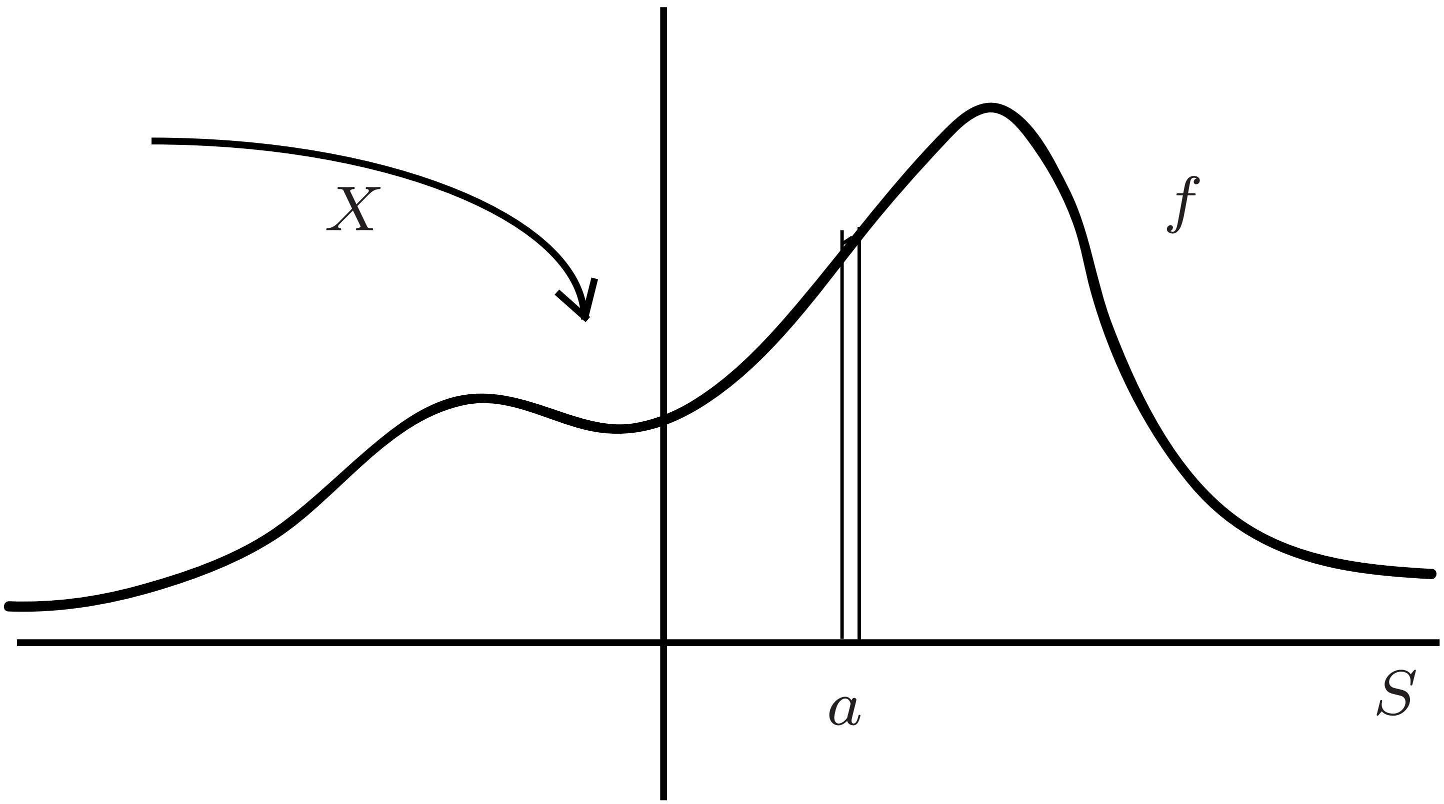
Das Analogon zu den Verteilungsgewichten $\rho(a)$ ist jetzt

gegeben durch "infinitesimale Gewichte" $f(a)\,da$, wobei

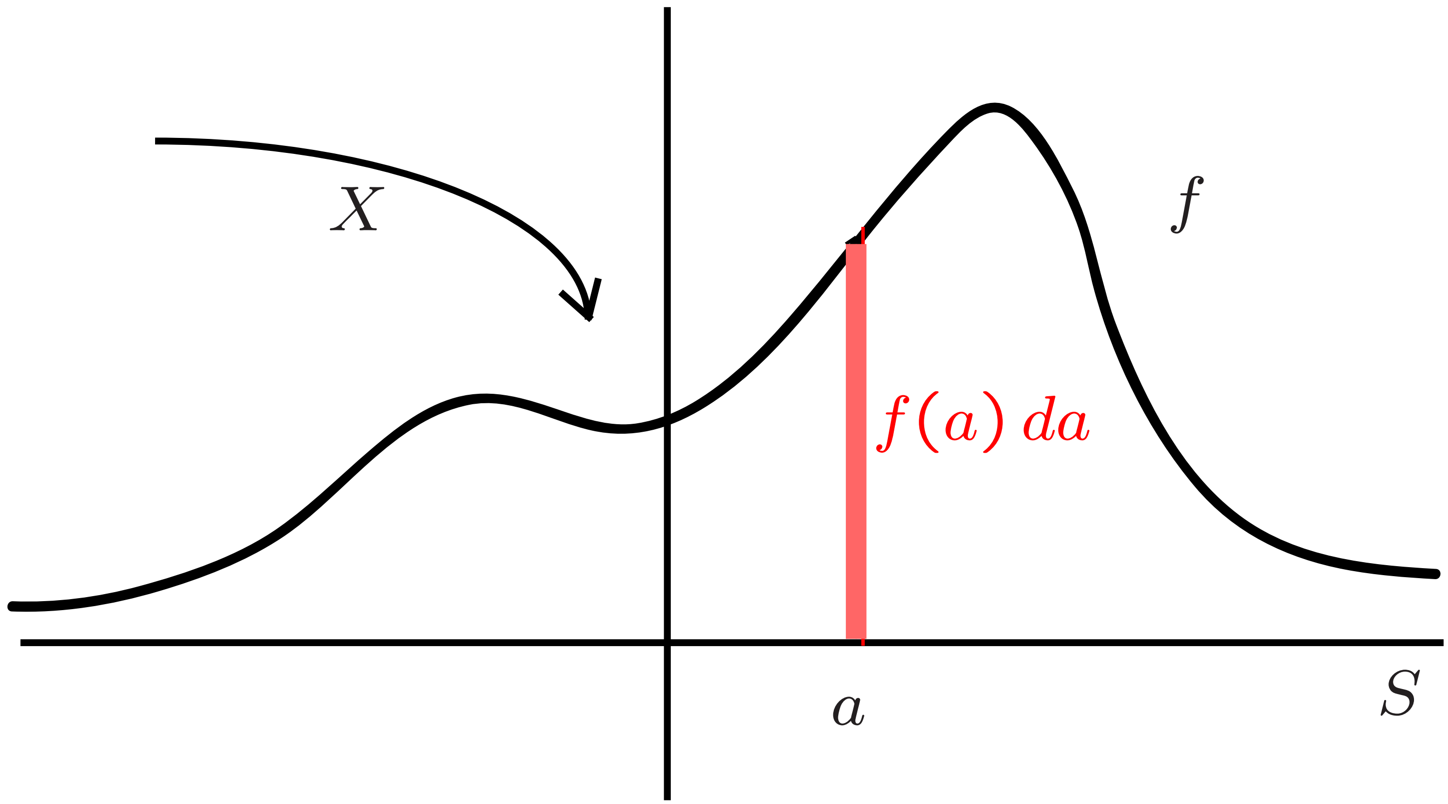
$f : S \to \mathbb{R}_+$ eine "integrierbare"

(z.B. stückweise stetige) Funktion ist mit

$$\int_l^r f(a)\,da = 1.$$



X
f
a
S



X
f
$f(a)\,da$
a
S

Sei X eine Zufallsvariable mit Wertebereich S.

Gilt für alle Intervalle $[b, c] \subset S$ die Gleichung

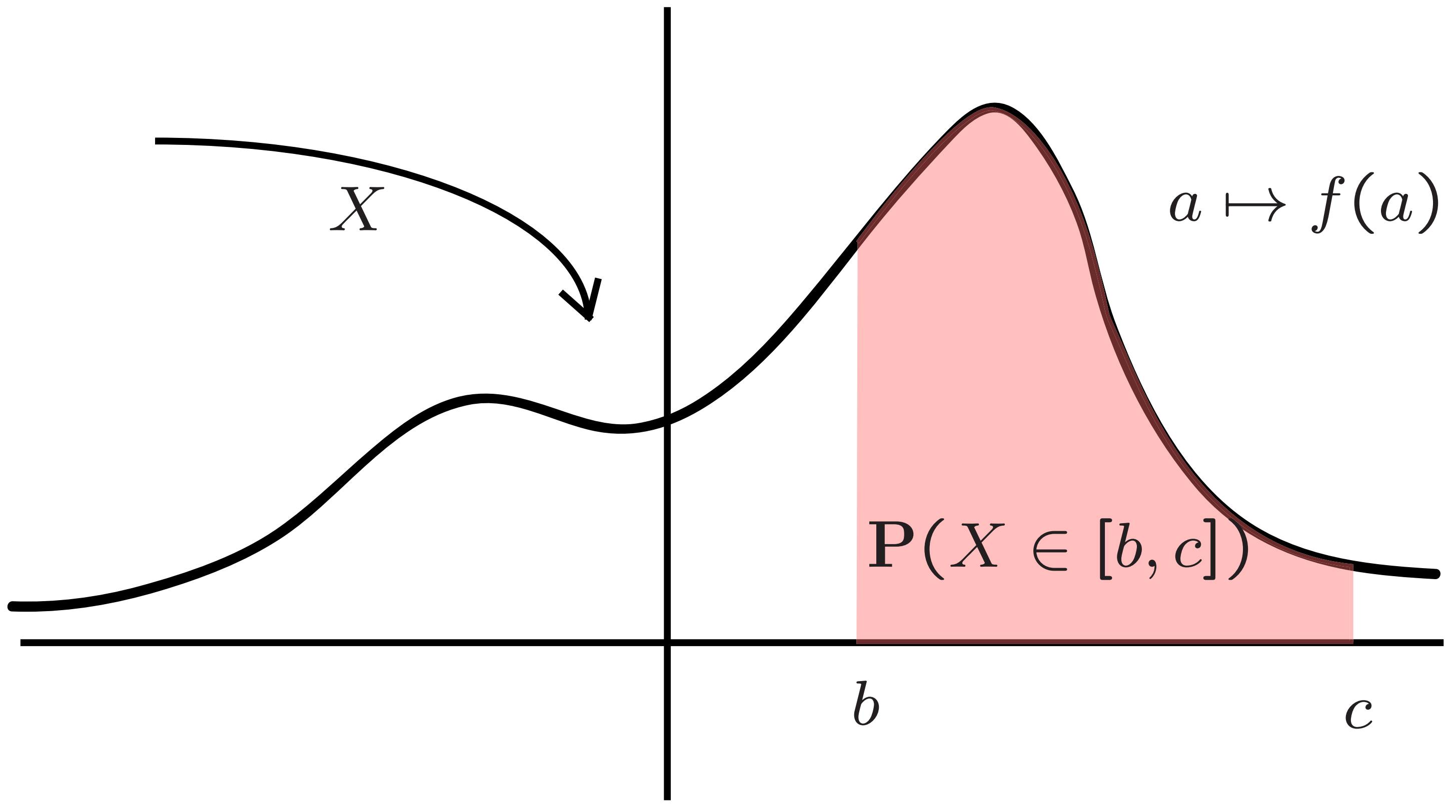
$$\mathbf{P}\big(X \in [b, c]\big) = \int_b^c f(a)\, da \ ,$$

so sagten wir, dass

X die *Dichte* $f(a)\, da$ besitzt*,

und nennen f *Dichtefunktion* (der Verteilung) von X.

*Genauer spricht man von der Dichte bzgl des natürlichen Längenmaßes (des Lebesguemaßes) auf $\mathbb{R}$


X
$a \mapsto f(a)$
$\mathbf{P}(X \in [b, c])$
b
c

Wir schreiben dann kurz

$$\mathbf{P}(X \in da) = f(a)\, da\,, \quad a \in S\,,$$

$$\mathbf{P}(X \in da) = f(a)\, da.$$

Der Ausdruck da taucht hier in zwei Bedeutungen auf:

links als infinitesimales Raumstück da (um den Punkt a)

und rechts als dessen (infinitesimale) Länge da.

$$\mathbf{P}(X \in da) = f(a)\,da.$$

Der Ausdruck da taucht hier in zwei Bedeutungen auf:

links als infinitesimales Raumstück da (um den Punkt a)

und rechts als dessen (infinitesimale) Länge da.

Diese Gleichung bekommt ihre exakte Bedeutung

“unter dem Integral”:

$$\mathbf{P}(X \in [c, d]) = \int_c^d f(a)\,da$$

Beispiele:

Eine auf dem Intervall $[0, 2]$

uniform verteilte Zufallsvariable

hat die Dichte $\frac{1}{2}\, da, \;\; 0 \le a \le 2.$

Eine auf einem endlichen Intervall $S = [l, r]$

uniform verteilte Zufallsvariable

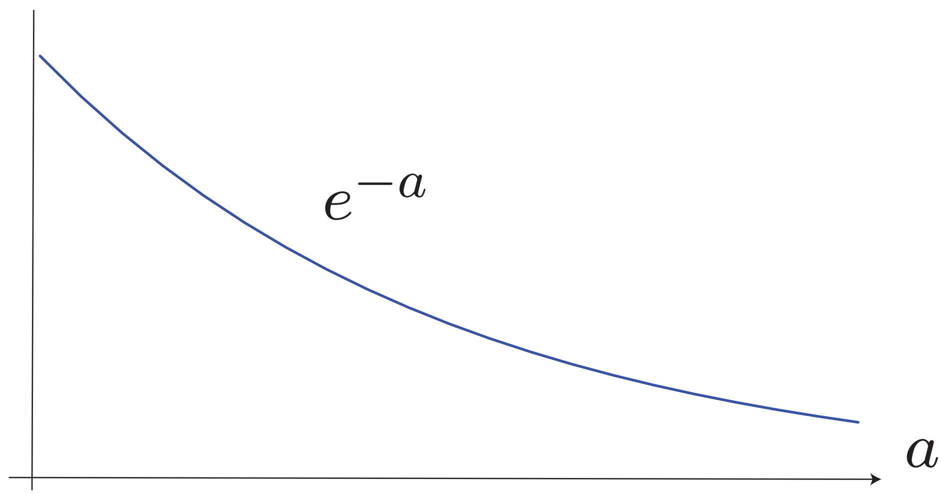
hat die Dichte $$\frac{1}{r-l}\, da\,, \quad a \in S.$$

Die Bedingung $\int_S f(a)\, da = 1$ kann auch erfüllt sein,

wenn S unendlichen Inhalt hat.

Man denke an das Eingangsbeispiel

$$S = [0, \infty); \;\; f(a) = e^{-a}, \quad a \geq 0.$$



Merke:

Für eine Zufallsvariable X mit Dichte $f(a)\,da$

ist für jedes $b \in \mathbb{R}$

$$\mathbf{P}(X = b) = \int_b^b f(a)\,da = 0.$$

Also gilt (mit naheliegender Schreibweise)

für $b \leq c \in \mathbb{R}$:

$$\int_{(b,c]} f(a)\,da = \int_{[b,c]} f(a)\,da = \int_b^c f(a)\,da.$$

Hat die Zufallsvariable X eine Dichte, so gilt

für jede endliche oder abzählbar unendliche Menge A:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) = 0.$$

Insbesondere ist X dann nicht diskret.

Umgekehrt gilt also:

Eine diskrete reellwertige Zufallsvariable besitzt keine Dichte

(bzg. des natürlichen Längenmaßes auf $\mathbb{R}$).

Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Teil 2

Der Zeitpunkt des ersten Erfolgs und die geometrische Verteilung

(Buch S. 34-35)

$(Z_1, Z_2, \ldots)$ sei ein fortgesetzter p-Münzwurf.

$$T := \mathrm{min}\{i \ : \ i \in \mathbb{N}, Z_i = 1\}$$

ist der *Zeitpunkt des ersten Erfolges*.

Wie sieht die Verteilung von T aus?

$$\mathbf{P}(T = n) = ?$$

$$\{T = n\} = \{Z_1 = 0, \ldots Z_{n-1} = 0,\ Z_n = 1\}$$

Also:

$$\mathbf{P}(T = n) = \mathbf{P}(Z_1 = 0, \ldots Z_{n-1} = 0,\ Z_n = 1)$$

$$= q^{n-1}\, p.$$

Alternativ:

$$\{T > n\} = \{Z_1 = 0, \ldots, Z_n = 0\}$$

Also

$$\mathbf{P}(T > n) = q^n.$$

$$\mathbf{P}(T = n) = q^{n-1}\, p$$

$$\mathbf{P}(T > n) = q^{n}$$

Das passt zusammen.

Denn das Ereignis $\{T > n - 1\}$ ist die disjunkte Vereinigung

der beiden Ereignisse $\{T = n\}$ und $\{T > n\}$.

Dementsprechend:

$$q^{n-1} = q^{n-1}p + q^{n}.$$

Definition

Sei $p \in (0,1)$. Eine Zufallsvariable T mit Zielbereich $\mathbb{N}$ heißt

geometrisch verteilt mit Parameter p,

kurz Geom(p)-verteilt,

wenn

$$\mathbf{P}(T > a) = q^a, \quad a = 0, 1, 2 \ldots\ ,$$

mit $q := 1 - p$.

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

Anschaulich ist klar:

Beim gewöhnlichen Würfeln kommt im Mittel

jedes 6-te Mal eine Sechs.

Beim Münzwurf mit Erfolgswahrscheinlichkeit p

kommt im Mittel jedes $(1/p)$-te Mal ein Erfolg.

Also wird gelten:

$$\mathbf{E}[T] = \frac{1}{p}.$$

Das beweist man auch schnell

mit dem folgenden

Lemma zum Erwartungswert $\mathbb{N}$-wertiger ZV'er:

(Buch S. 34)

Ist X eine Zufallsvariable mit Zielbereich $\mathbb{N}$ oder $\mathbb{N}_0$, dann ist

$$\mathbf{E}[X] = \sum_{i \geq 0} \mathbf{P}(X > i)$$

Folgerung aus dem Lemma:

Für eine Geom(p)-verteilte Zufallsvariable T ist

$$\mathbf{E}[T] = \sum_{i \geq 0} \mathbf{P}(T > i) = \sum_{i \geq 0} q^i = \frac{1}{1-q} = \frac{1}{p}.$$

$$\boxed{\mathbf{E}[T] = \frac{1}{p}}$$

Lemma

Ist X eine Zufallsvariable mit Zielbereich $\mathbb{N}$ oder $\mathbb{N}_0$, dann ist

$$\mathbf{E}[X] = \sum_{i \geq 0} \mathbf{P}(X > i)$$

Beweis.

$\rho(j)$ seien die Verteilungsgewichte von X.

$$\mathbf{E}[X] = \sum_{j \geq 1} j\rho(j)$$

$$\sum_{i \geq 0} \mathbf{P}(X > i) = \sum_{i \geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$

Warum sind die beiden rechten Seiten gleich?

Beweis.

$\rho(j)$ seien die Verteilungsgewichte von X.

$$\mathbf{E}[X] = \sum_{j \geq 1} j\rho(j)$$

$$\sum_{i \geq 0} \mathbf{P}(X > i) = \sum_{i \geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$

Warum sind die beiden rechten Seiten gleich?

Beweis.

$\rho(j)$ seien die Verteilungsgewichte von X.

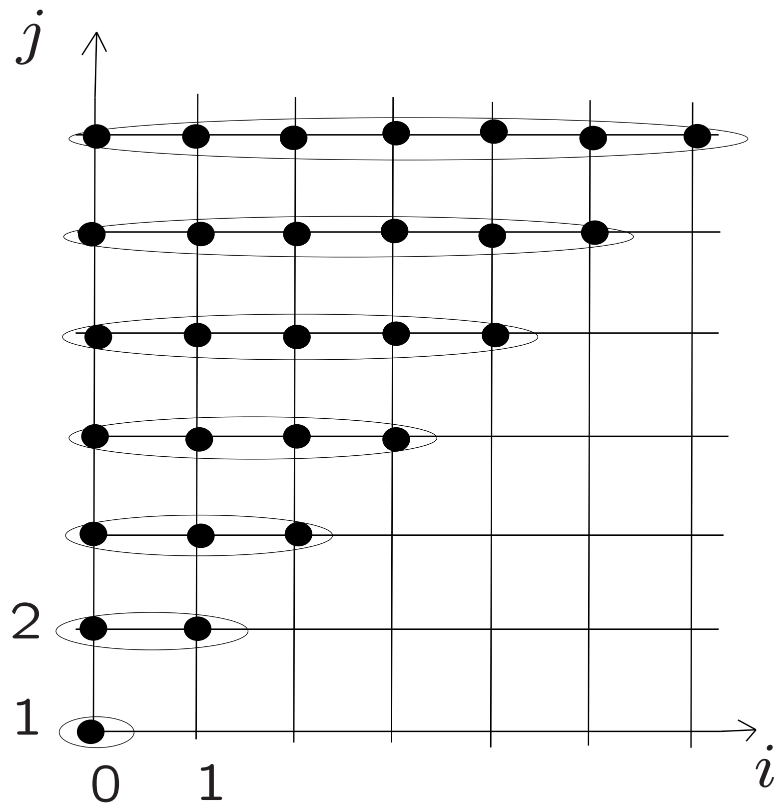
$$\mathbf{E}[X] = \sum_{j\geq 1} j\rho(j) = \sum_{j\geq 1}\sum_{i=0}^{j-1} \rho(j)$$

$$\sum_{i\geq 0} \mathbf{P}(X > i) = \sum_{i\geq 0}\sum_{j=i+1}^{\infty} \rho(j)$$

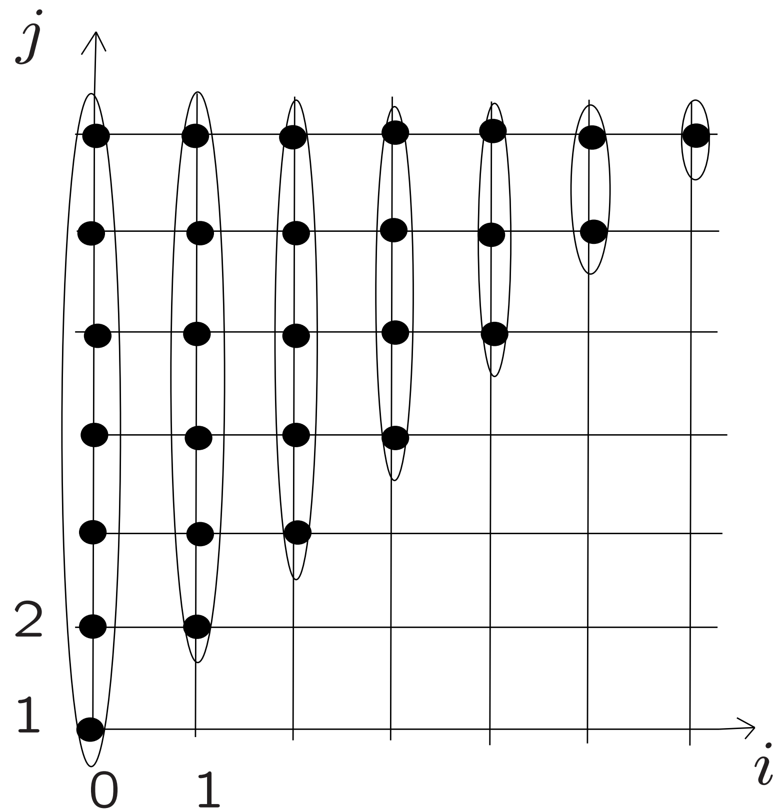
Warum sind die beiden rechten Seiten gleich?

Wie sieht man die Gleichheit

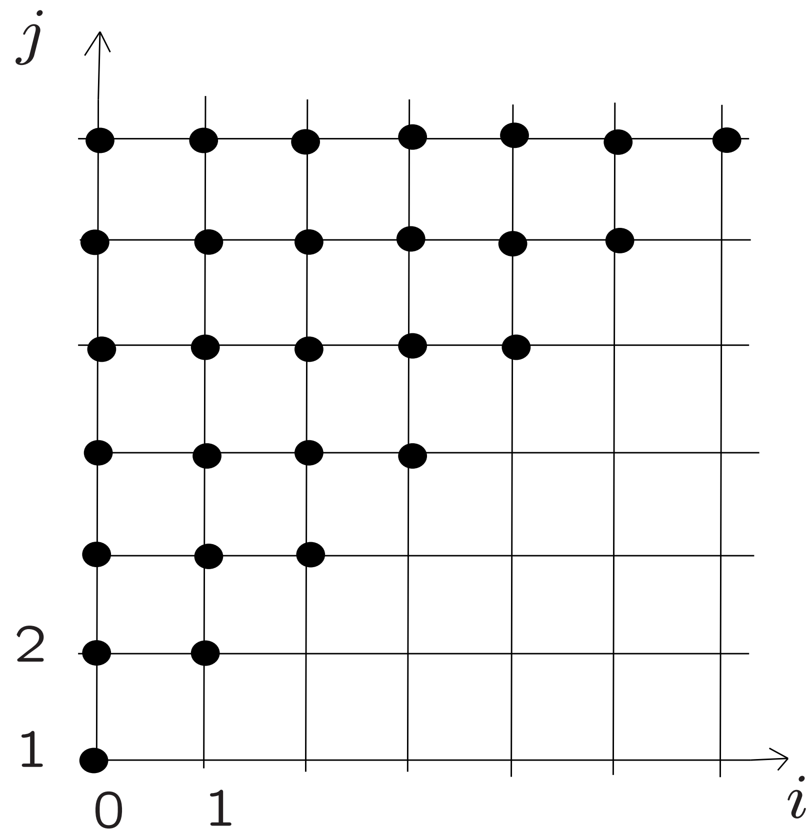
$$\sum_{j\geq 1}\sum_{i=0}^{j-1}\rho(j)=\sum_{i\geq 0}\sum_{j=i+1}^{\infty}\rho(j) \qquad ?$$



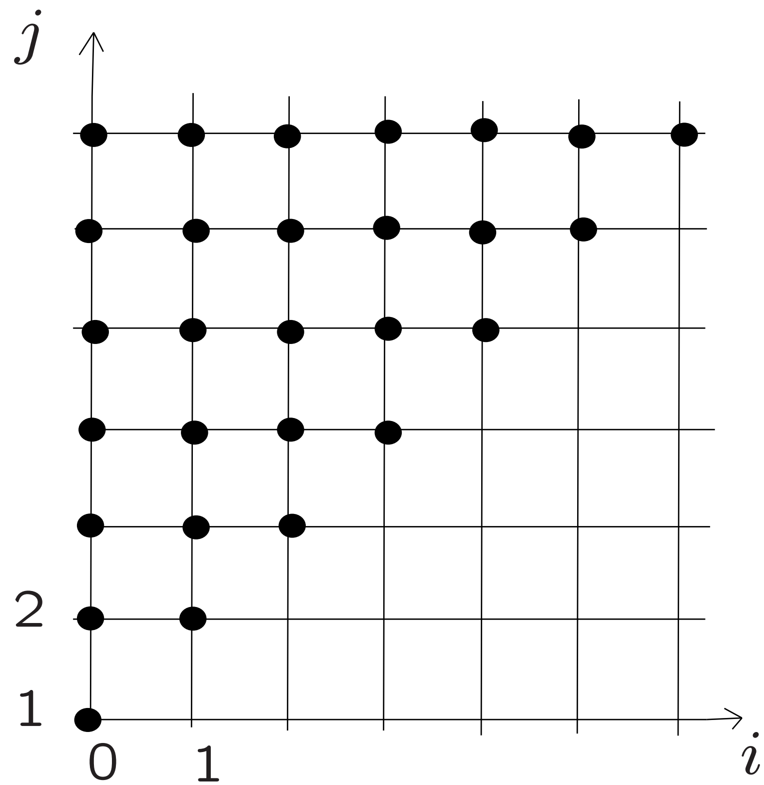
$$\sum_{j \geq 1} \sum_{i=0}^{j-1} \rho(j)$$



$$\sum_{i\geq 0}\ \sum_{j=i+1}^{\infty} \rho(j)$$



Es kommt bei nichtnegativen Summanden
nicht auf die Reihenfolge der Summation an!



$$\sum_{j\geq 1}\sum_{i=0}^{j-1}\rho(j)=\sum_{i\geq 0}\sum_{j=i+1}^{\infty}\rho(j) \qquad \square$$

Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Teil 4

Die Poissonapproximation.

Oder:

Münzwurf mit kleiner Erfolgswahrscheinlichkeit:

Wie ist die Anzahl der Erfolge verteilt

bei einer großen Zahl von Versuchen?

(Buch S. 29-30)

p klein, $\quad n$ groß

$$X := Z_1 + Z_2 + \ldots + Z_n$$

$$\mathbf{P}(X = k) \approx ?$$

Beispiel:

$$p = \frac{1}{1000}, \quad n = 3000$$

$$\mathbf{P}(X = 0) = q^n = \left(1 - \frac{1}{1000}\right)^{3000}$$
$$\approx e^{-3}$$

$$\mathbf{P}(X = 1) = npq^{n-1} \approx 3e^{-3}$$

$$\mathbf{P}(X = 2) = \binom{n}{2} p^2 q^{n-2} \approx \frac{1}{2}(np)^2 q^n \approx \frac{1}{2} 3^2 e^{-3}$$

Clou:

p klein, n groß:

$$q^n = (1-p)^n \approx e^{-np}$$

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{k!} n^k p^k q^n \approx \frac{1}{k!} (np)^k e^{-np}$$

Fazit

Sei p eine kleine positive Zahl,

n eine große natürliche Zahl

und X eine Bin(n, p)-verteilte Zufallsvariable.

Man kann dann die Verteilungsgewichte von X

approximativ als Funktion von $\mathbf{E}[X] = n\,p$ ausdrücken.

Rigoros fasst man diese Behauptung im folgenden

Grenzwertsatz:

Satz (Poissons Gesetz der seltenen Ereignisse)

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n, $n = 1, 2, \ldots$,

eine Folge von Bin(n, p_n)-verteilten Zufallsvariablen,

so dass für $n \to \infty$

$$\mathbf{E}[X_n] \to \lambda\,, \quad \text{d. h. } p_n \sim \frac{\lambda}{n}\,.$$

Dann gilt für jedes $k = 0, 1, 2, \ldots$

$$\boxed{\mathbf{P}(X_n = k) \to \frac{\lambda^k}{k!} e^{-\lambda}\,.}$$

Beweis:

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} \; =$$

$$\frac{1}{k!} \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\to 1} \underbrace{(np_n)^k}_{\to \lambda^k} \underbrace{\left(1-\frac{np_n}{n}\right)^n}_{\to e^{-\lambda}} \underbrace{(1-p_n)^{-k}}_{\to 1}$$

$$\to \frac{1}{k!}\lambda^k e^{-\lambda} \, . \qquad \square$$

Definition (Poissonverteilung)

(Buch S. 29)

Sei $\lambda \in \mathbb{R}_+$.

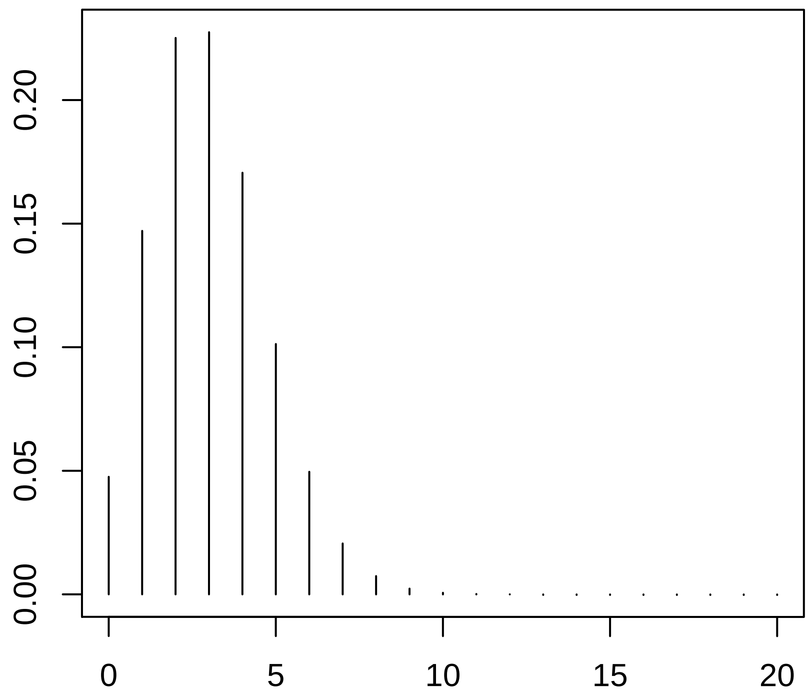
Eine Zufallsvariable X mit Zielbereich $\mathbb{N}_0$ heißt

Poissonverteilt mit Parameter λ,

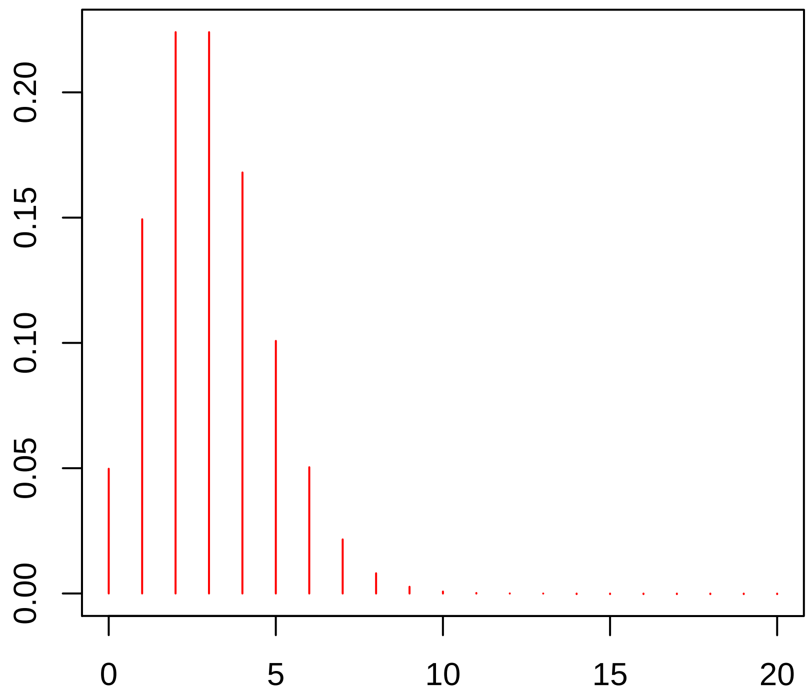
kurz Pois(λ)-verteilt,

wenn

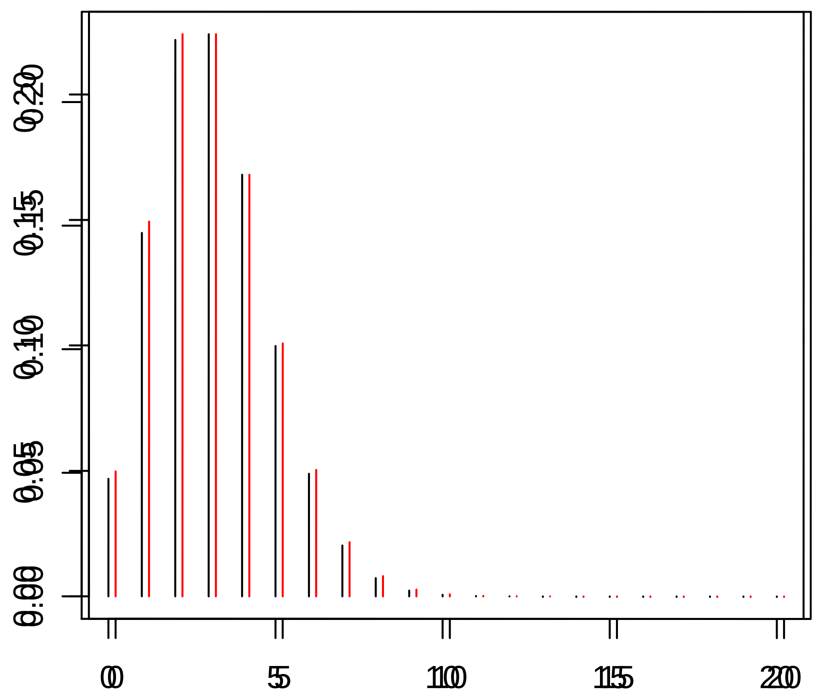
$$\boxed{\mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}\,, \quad k = 0, 1, 2, \ldots\ .}$$



Binomialgewichte zu $n = 100$ und $p = 0.03$



Poissongewichte zum Parameter λ =3



0.00
0.05
0.10
0.15
0.20
0
5
10
15
20

Satz.

Der Erwartungswert

einer Pois(λ)-verteilten Zufallsvariablen X ist

$$\mathbf{E}[X] = \lambda.$$

Beweis:

$$\mathbf{E}[X] = \sum_{k=0}^{\infty} k \, \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \cdot 1 \qquad \square$$

Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Teil 3

Die Exponentialapproximation.

Oder:

Münzwurf mit kleiner Erfolgswahrscheinlichkeit:

Wie lange dauert es bis zum ersten Erfolg?

(Buch S. 42)

Wieder sei

$$T$$

der zufällige Zeitpunkt des ersten Erfolgs

in einem fortgesetzten p-Münzwurf.

Beispiel:

$$p = \frac{1}{1000}$$

$$\mathbf{P}(T > 2000) = q^{2000} = \left(1 - \tfrac{1}{1000}\right)^{2000} \approx e^{-2}$$

$$\mathbf{P}(T > 2 \cdot \mathbf{E}[T]) \qquad \approx \qquad e^{-2}$$

$$\boxed{\mathbf{P}\left(\frac{T}{\mathbf{E}[T]} > 2\right) \approx e^{-2}}$$

Betrachten wir T *auf der Skala seines Erwartungswertes*:

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für $t \in \mathbb{R}_+$ ist

$$\mathbf{P}\{\widetilde{T} > t\} = \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right)$$

$$= \Big(1 - p\Big)^{\left\lfloor \frac{t}{p} \right\rfloor}$$

$$= \Big(1 - p\Big)^{\frac{1}{p}\, p \left\lfloor \frac{t}{p} \right\rfloor}$$

Für $p \to 0$ konvergiert dies gegen

$$(e^{-1})^t = e^{-t}.$$

Diese Tatsache formulieren wir als einen *Grenzwertsatz*:

(vgl. Buch S. 42)

Satz Sei $T_1, T_2, \ldots$ eine Folge von geometrisch verteilten Zufallsvariablen mit der Eigenschaft

$$\mathbf{E}[T_m] \underset{m\to\infty}{\longrightarrow} \infty.$$

Dann gilt für jedes $t \geq 0$:

$$\boxed{\mathbf{P}\left(\frac{T_m}{\mathbf{E}[T_m]} > t\right) \underset{m\to\infty}{\longrightarrow} e^{-t}}$$

Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Die Welt des p-Münzwurfs -

von Bernoulli zu Poisson

Teil 1

Der fortgesetze p-Münzwurf

(Buch S. 19-20)

Zur Erinnerung:

Der n-fache p-Münzwurf ...

$\ldots$ ist eine $\{0,1\}^n$-wertige ZV'e $(Z_1,\ldots,Z_n)$

mit

$$(*) \qquad \mathbf{P}(Z_1 = a_1,\ldots,Z_n = a_n) = p^k(1-p)^{n-k}$$

falls $a_1 + \cdots + a_n = k$.

Ereignisse kann man oft auf verschiedene Weise darstellen.

Für einen $(n+1)$-fachen p-Münzwurf gilt z.B.

$$\begin{aligned}
&\{Z_1 = a_1, \ldots, Z_n = a_n\} \\
&\qquad = \{Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 1\} \\
&\qquad\qquad \cup \{Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 0\}
\end{aligned}$$

Weil rechts zwei disjunkte Ereignisse stehen, muss gelten:

$$\begin{aligned}
&\mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n) \\
&\qquad = \mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 1) \\
&\qquad\qquad + \mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 0)
\end{aligned}$$

Die Definition $(*)$ auf der vorigen Folie ist damit verträglich:

$$\mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n)$$
$$= \mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 1)$$
$$+ \mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n, Z_{n+1} = 0)$$

$$p^k (1-p)^{n-k}$$
$$= p^{k+1}(1-p)^{n-k}$$
$$+ p^k (1-p)^{n-k+1}$$

Dies zeigt, dass auch die folgende Definition

konsistent über n ist.

Definition: Sei $p \in (0, 1)$, $\quad q := 1 - p$.

Eine *Bernoulli-Folge zum Parameter p*

(man sagt manchmal auch: ein *fortgesetzter p-Münzwurf*)

ist eine zufällige 01-Folge $(Z_1, Z_2, \ldots)$, deren Verteilung die folgende Eigenschaft hat:

Für jedes $n \in \mathbb{N}$ und jede endliche 01-Folge $(a_1, \ldots, a_n)$ mit k Einsen und $n - k$ Nullen ist

$$\mathbf{P}(Z_1 = a_1, \ldots, Z_n = a_n) = p^k q^{n-k}.$$

(d.h. für jedes n ist $(Z_1, \ldots, Z_n)$ ein n-facher p-Münzwurf)

Wir wissen schon:

Für jedes n ist dann

die Anzahl der Einsen in $(Z_1, \ldots, Z_n)$

(die “Anzahl der Erfolge in n Versuchen”)

binomial(n, p)-verteilt:

$$\mathbf{P}(Z_1 + \cdots + Z_n = k) = \binom{n}{k} p^k q^{n-k}.$$

Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

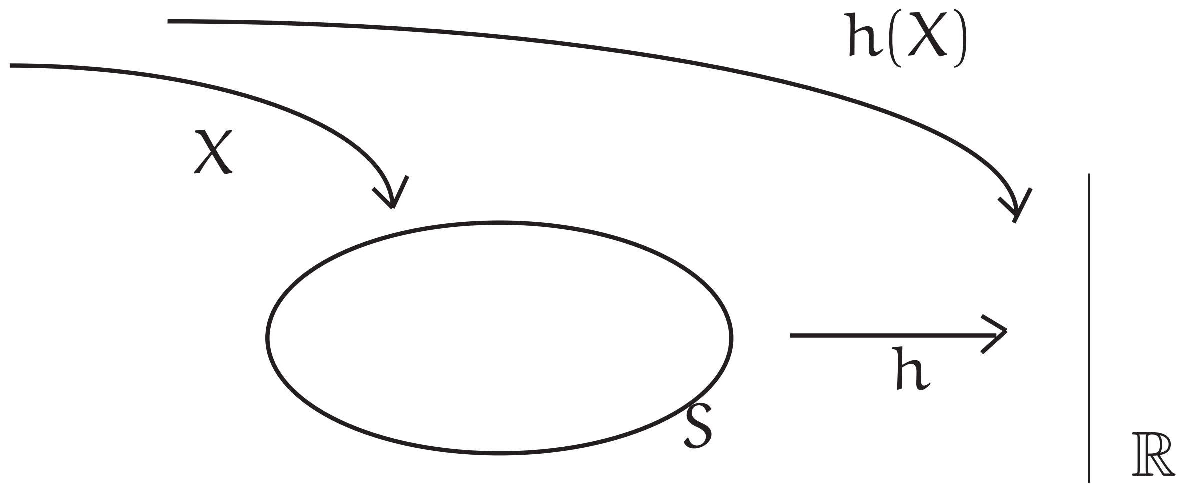
Teil 2

Die Transformationsformel für Erwartungswerte

(vgl. Buch S. 23)

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 $\mathbb{R}$

mit $\sum\limits_{a \in S} |h(a)| \, \mathbf{P}(X = a) < \infty$.

Dann ist

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

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 nächsten Teil helfen,

die Linearität des Erwartungswertes herzuleiten.

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

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 dann mit den Gewichten der Werte a.

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

Denn:

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

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

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

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

Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

Teil 4

Beispiele

1. Der Erwartungswert der Binomialverteilung

(als Erwartungswert der Anzahl der Erfolge

beim n-fachen p-Münzwurf)

(Buch S. 49)

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 (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.

2. Der Erwartungswert der hypergeometrischen Verteilung

(als Erwartungswert der Anzahl der “Erfolge” beim n-fachen Ziehen ohne Zurücklegen)

(Buch S. 50 und S. 28)

BEISPIEL

Ziehen ohne Zurücklegen

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

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

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

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

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

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.

$\{Z_i = 1\}$ ist das Ereignis “die i-te gezogene Kugel ist rot”.

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

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

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

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

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

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

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

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

BEISPIEL

Ziehen ohne Zurücklegen

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

Verteilung von R?

Verteilungsgewichte von R ?

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

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

$$(k = 0, \ldots, n)$$

Eine ZV mit diesen Verteilungsgewichten

heißt

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

(vgl. 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 wie wir eben gesehen haben,

(über die Darstellung von R als Summe von Zählern)

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

3. 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

$00000000 \qquad R = 1$

$11100011 \qquad R = 3$

$10101010 \qquad R = 8$

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

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

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

$$Y_1 \equiv 1$$

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

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

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

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

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

Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

Teil 1

Der Erwartungswert als gewichtetes Mittel

(Buch S. 23)

Unter einer diskreten reellwertigen Zufallsvariablen

verstehen wir eine ZV'e, deren Wertebereich

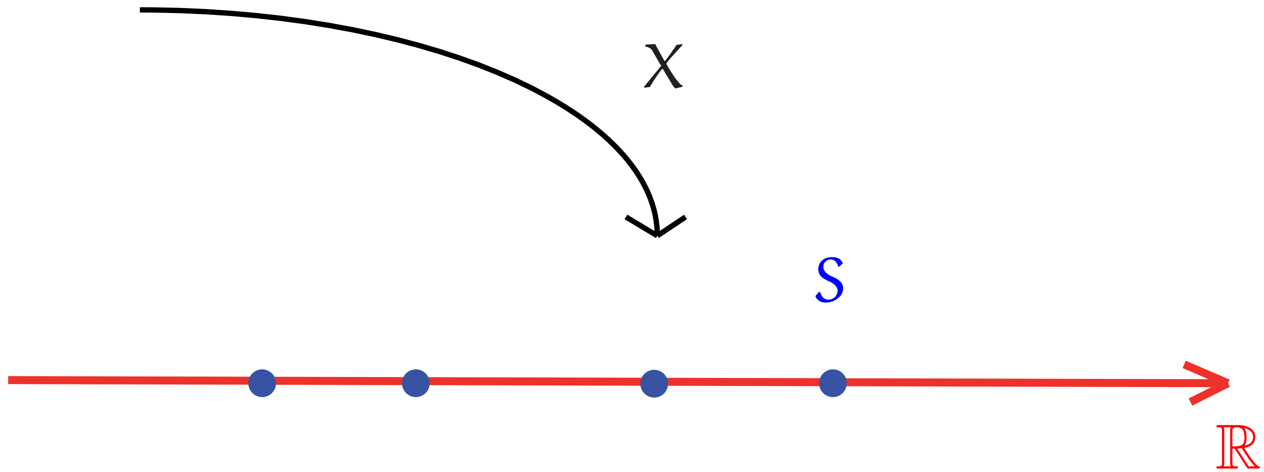
die Menge $\mathbb{R}$ der reellen Zahlen

(oder einer Teilmenge davon) ist,

und für die eine

endliche oder abzählbar unendliche Menge S ($\subset \mathbb{R}$) existiert

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


X
S
ℝ

Eine einprägsame Kenngröße

für die *Lage* der Verteilung von X

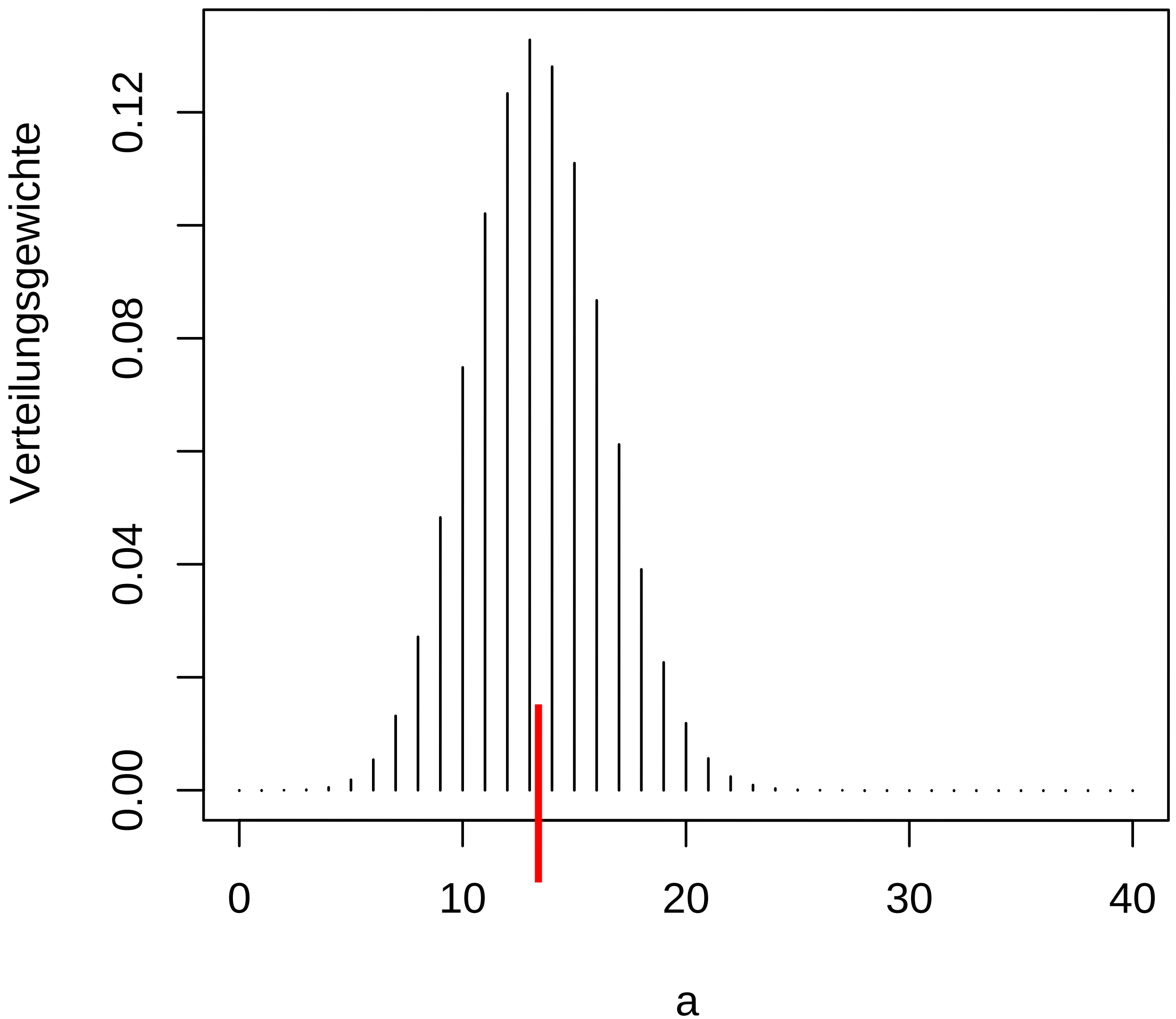
ist das mit den Wahrscheinlichkeiten gewichtete Mittel

der möglichen Werte von X:

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

Man spricht vom *Erwartungswert* von X.

(Wir bezeichnen ihn auch mit μ oder μ_X.)



Verteilungsgewichte
0.00
0.04
0.08
0.12
0
10
20
30
40
a

Das elementarste Beispiel:

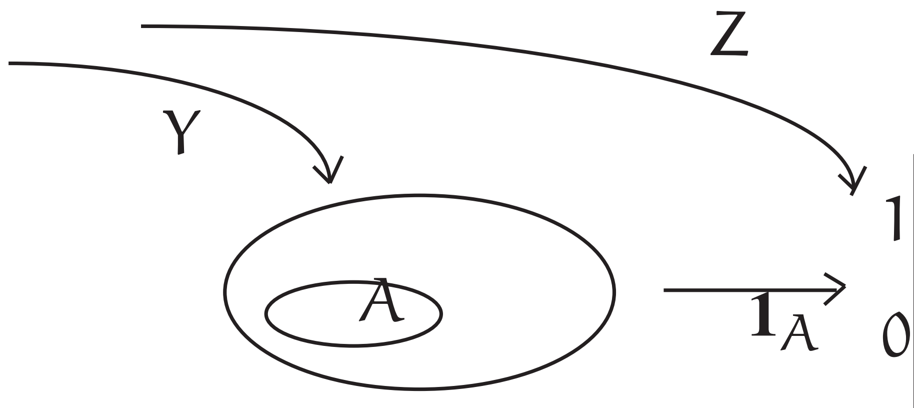
$$Z = \begin{cases} 1 \text{ mit Wahrscheinlichkeit } p \\ 0 \text{ mit Wahrscheinlichkeit } q = 1 - p \end{cases}$$

Z beschreibt also einen einfachen p-Münzwurf.

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

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

Indikatorvariable von Ereignissen



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

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

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

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

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

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

Die Wahrscheinlichkeit eines Ereignisses ist gleich dem Erwartungswert der Indikatorvariable des Ereignisses.

Für eine diskrete reellwertige ZV'e 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 Zufallsvariablen X

hängt nur von ihrer Verteilung ρ ab.

Synonym sprechen wir daher auch manchmal vom

Erwartungswert der Verteilung ρ .

Gilt $S \subset \mathbb{R}_+$ oder $\sum\limits_{a \in S} |a| \, \rho(a) < \infty$,

dann hängt der Summenwert $\sum\limits_{a \in S} a \, \rho(a)$

nicht von der Reihenfolge der Summanden ab.

Man sagt dann:

Der Erwartungswert der Zufallsvariablen X mit Verteilung ρ

ist wohldefiniert

oder kurz: Der Erwartungswert von X *existiert*.

Merke:

X

ist eine Zufallsgröße;

$\mathbf{E}[X]$

ist eine Zahl.

Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

Teil 3

Linearität des Erwartungswertes

(vgl. (Buch S. 52)

Satz 1. (Homogenität des EW)

Der Erwartungswert des c-fachen

einer diskreten reellwertigen Zufallsvariablen X

ist das c-fache des Erwartungswertes:

$$\mathbf{E}[cX] = c\mathbf{E}[X]$$

Satz 1. (Homogenität des EW)

Der Erwartungswert des c-fachen

einer diskreten reellwertigen Zufallsvariablen X

ist das c-fache des Erwartungswertes:

$$h(a) := ca, \quad Y := h(X) = cX.$$

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

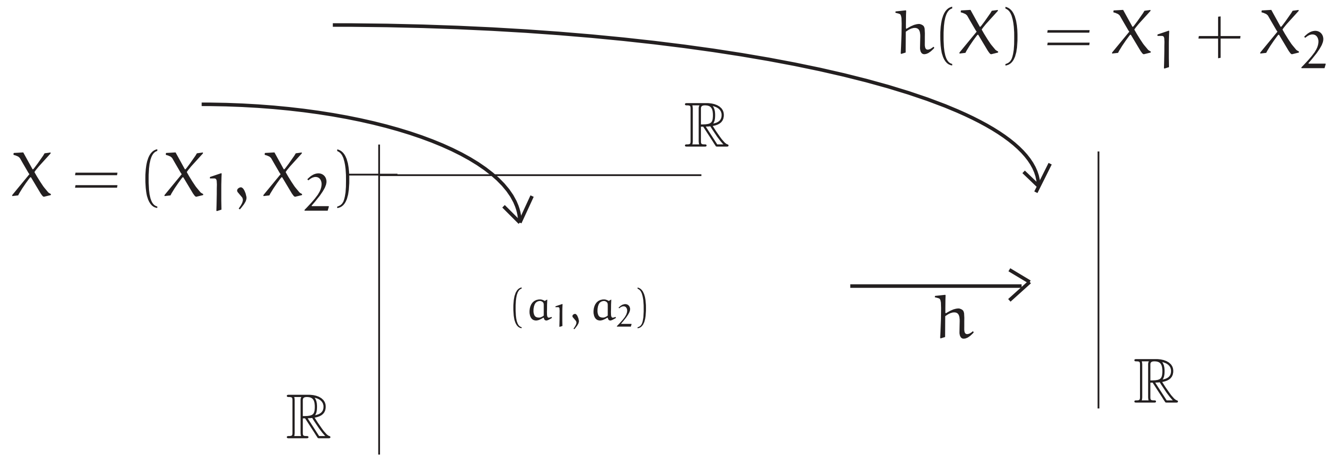
$$= \sum_{a \in S} ca\ \mathbf{P}(X = a) = c \sum_{a \in S} a\,\mathbf{P}(X = a) = c\mathbf{E}[X].$$

Satz 2. (Additivität des Erwartungswertes)

Sei (X_1, X_2) ein zufälliges Paar reellwertiger ZV'er, für die $\mathbf{E}[X_1]$ und $\mathbf{E}[X_2]$ existieren und endlich sind.

Dann gilt

$$\mathbf{E}[X_1 + X_2] = \mathbf{E}[X_1] + \mathbf{E}[X_2]$$



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

Beweis von Satz 2

(hier nur für **diskrete** Zufallsvariable):

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

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

$$\mathbf{E}[X_1 + X_2] = \sum_{(a_1,a_2)\in S_1 \times S_2} (a_1 + a_2)\,\mathbf{P}(X_1 = a_1, X_2 = a_2)$$

In der nächsten Umformung werden wir

die in V2b1 besprochene Beziehung

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

verwenden.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\square$$

Korollar aus Satz 1 und Satz 2

(**Linearität des Erwartungswertes**)

(Buch S. 52)

Für reellwertige Zufallsvariable X_1, X_2

für die $\mathbf{E}[X_1]$ und $\mathbf{E}[X_2]$ existieren und endlich sind,

gilt

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

Vorlesung 3a

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

Teil 5

Wie erlebt man den Erwartungswert?

$$X$$

eine Zufallsgröße;

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

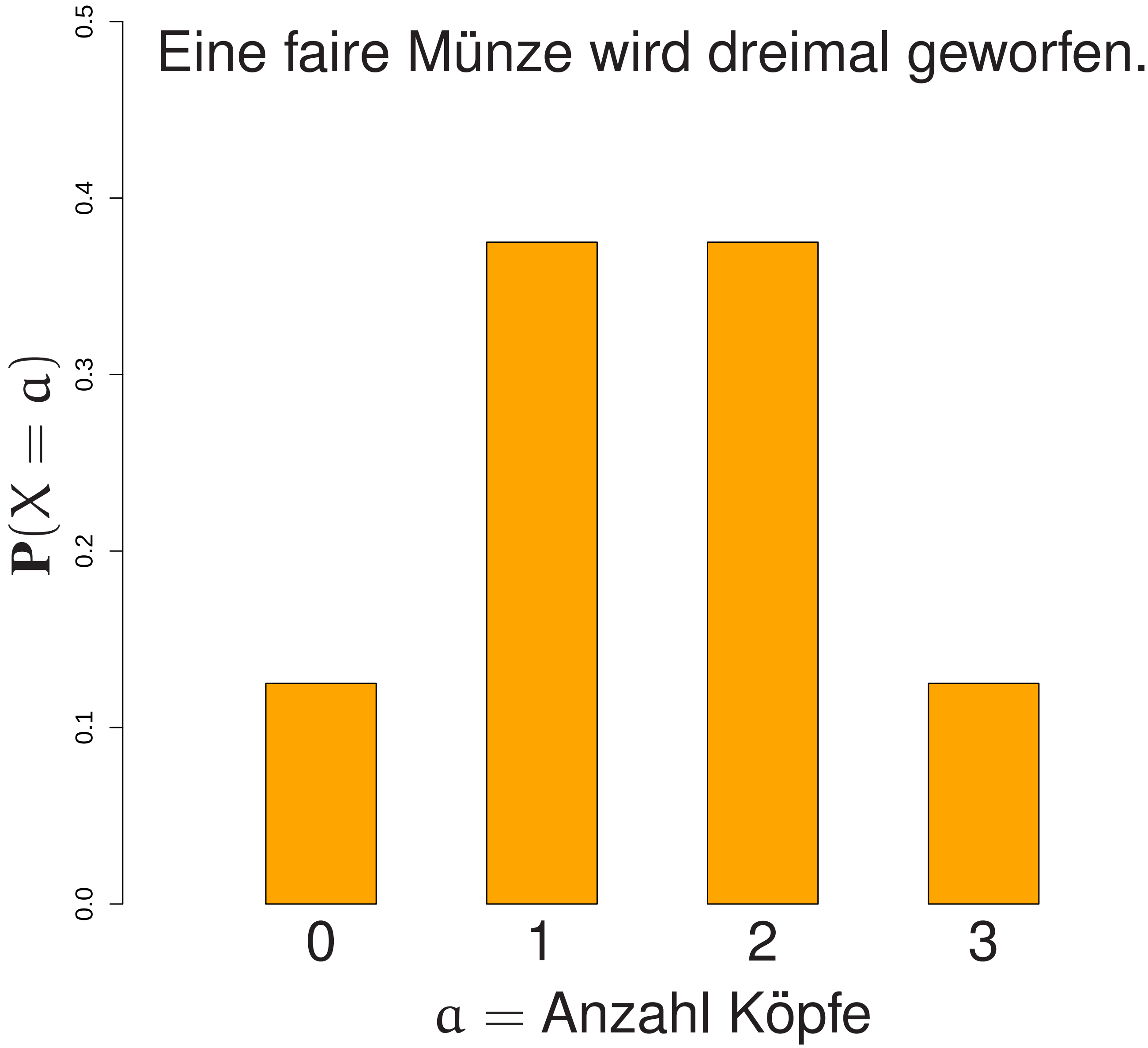
eine Zahl.

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

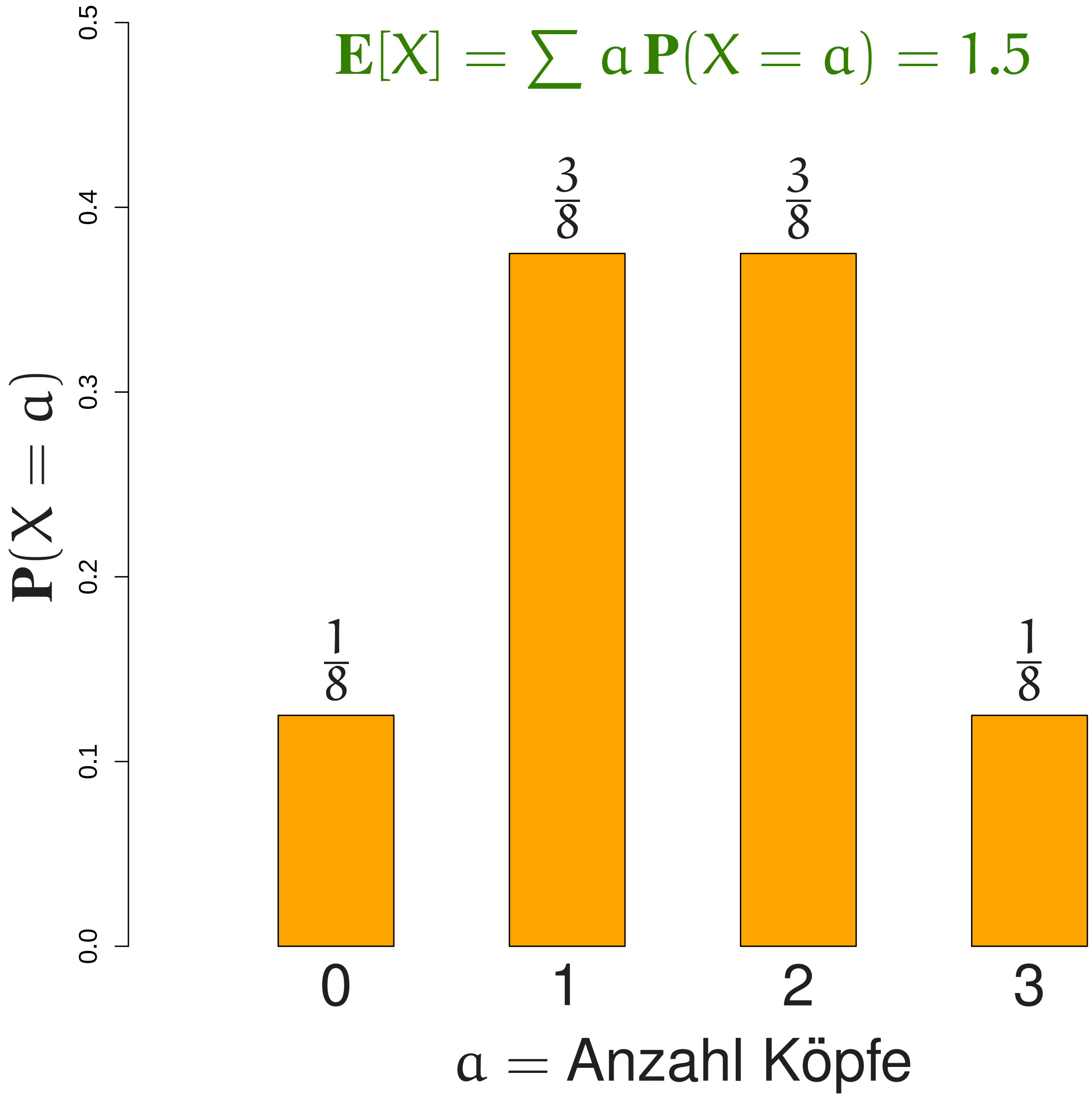
Eine faire Münze wird dreimal geworfen.

$X :=$ Anzahl der geworfenen Köpfe.

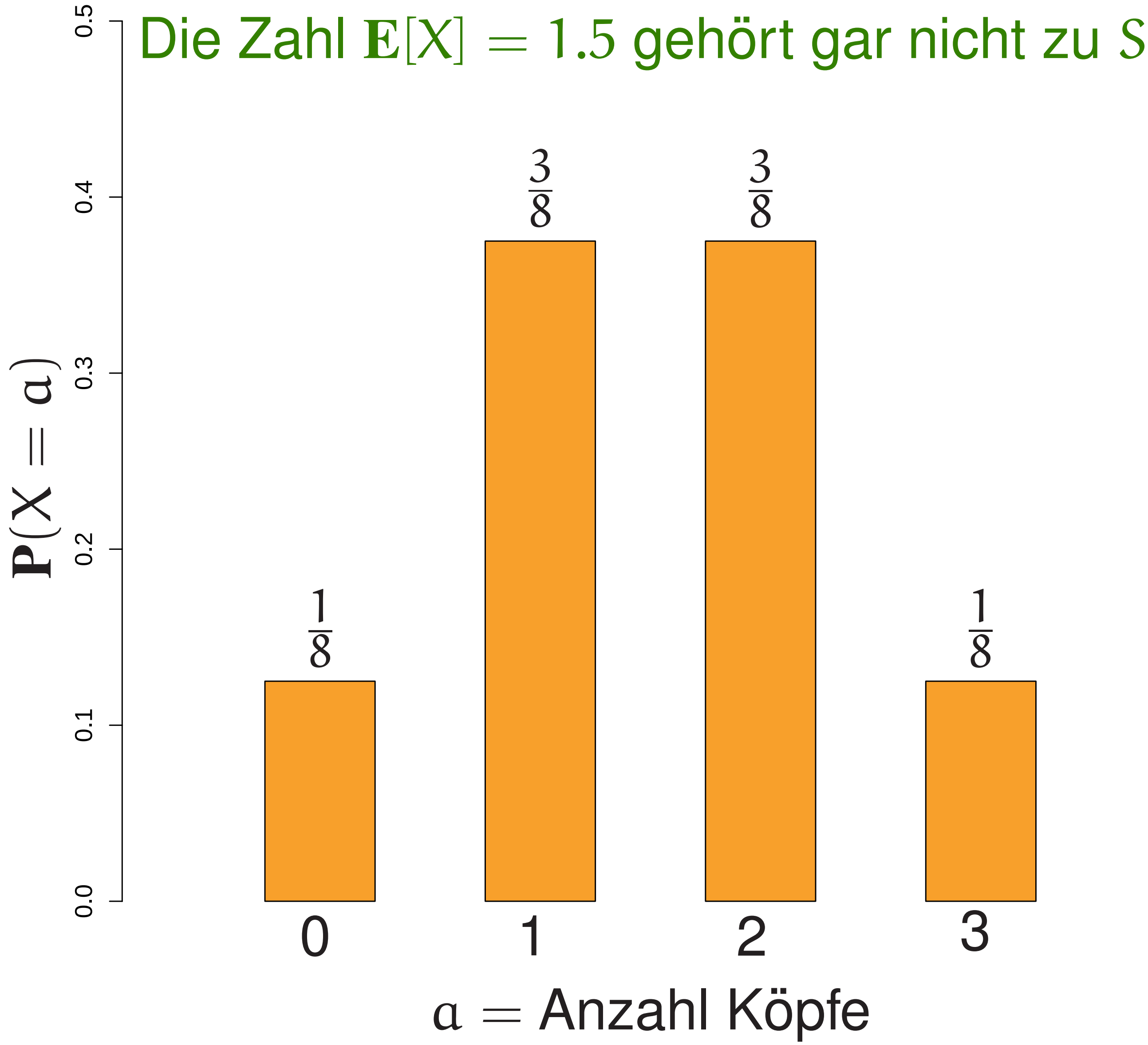
$S := \{0, 1, 2, 3\}.$



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 Köpfe

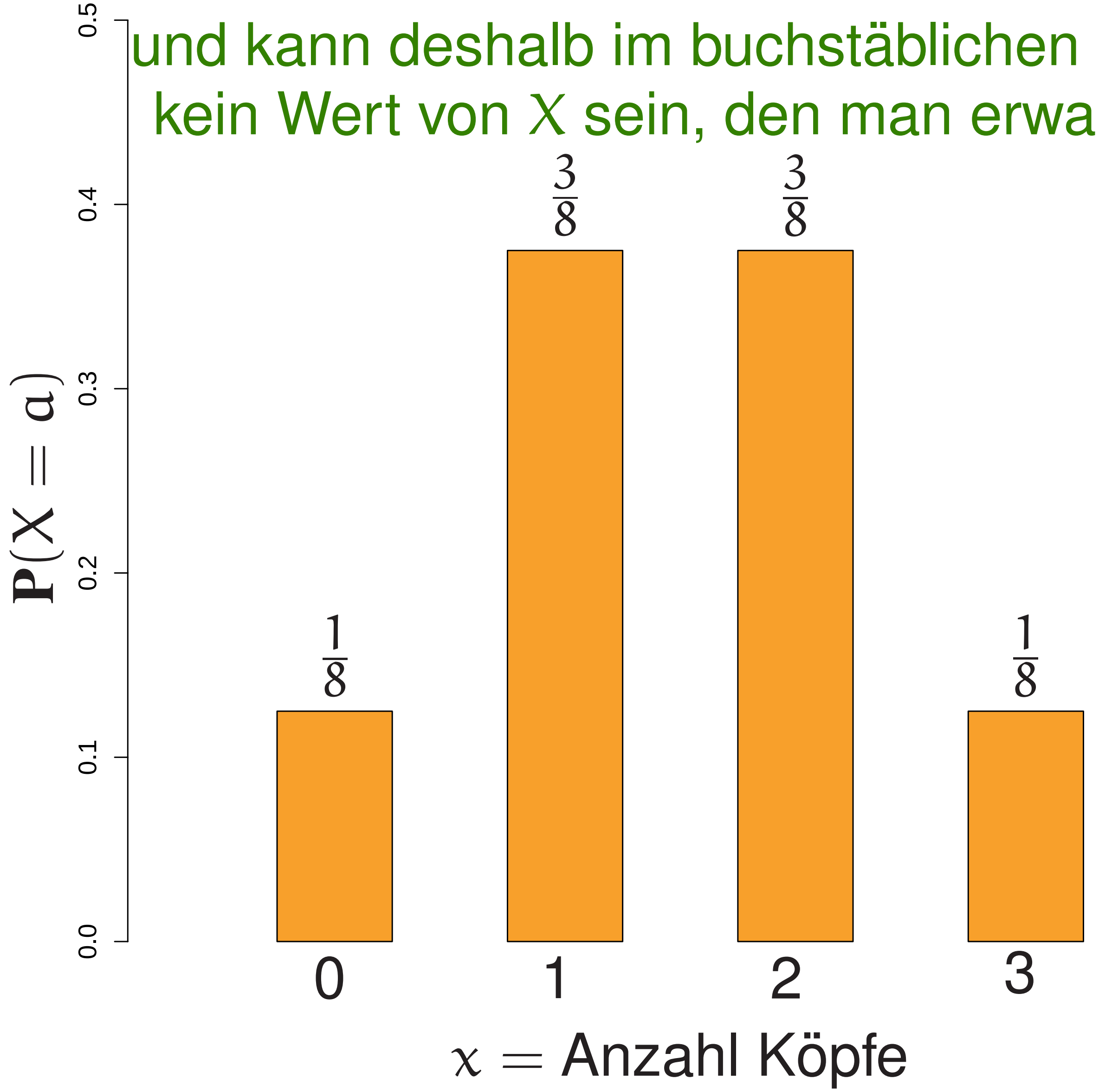


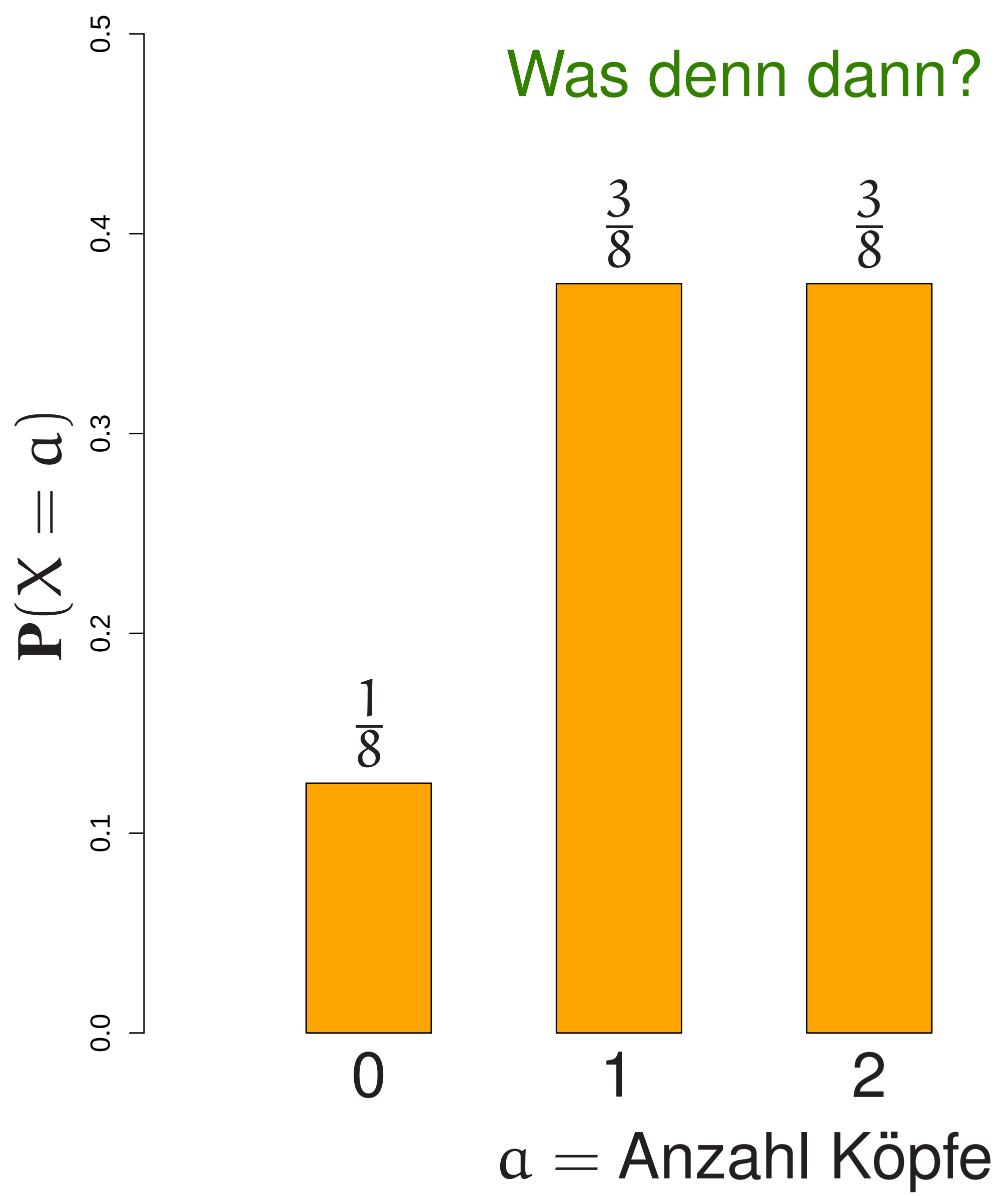
$\mathbf{E}[X] = \sum a\,\mathbf{P}(X = a) = 1.5$
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe

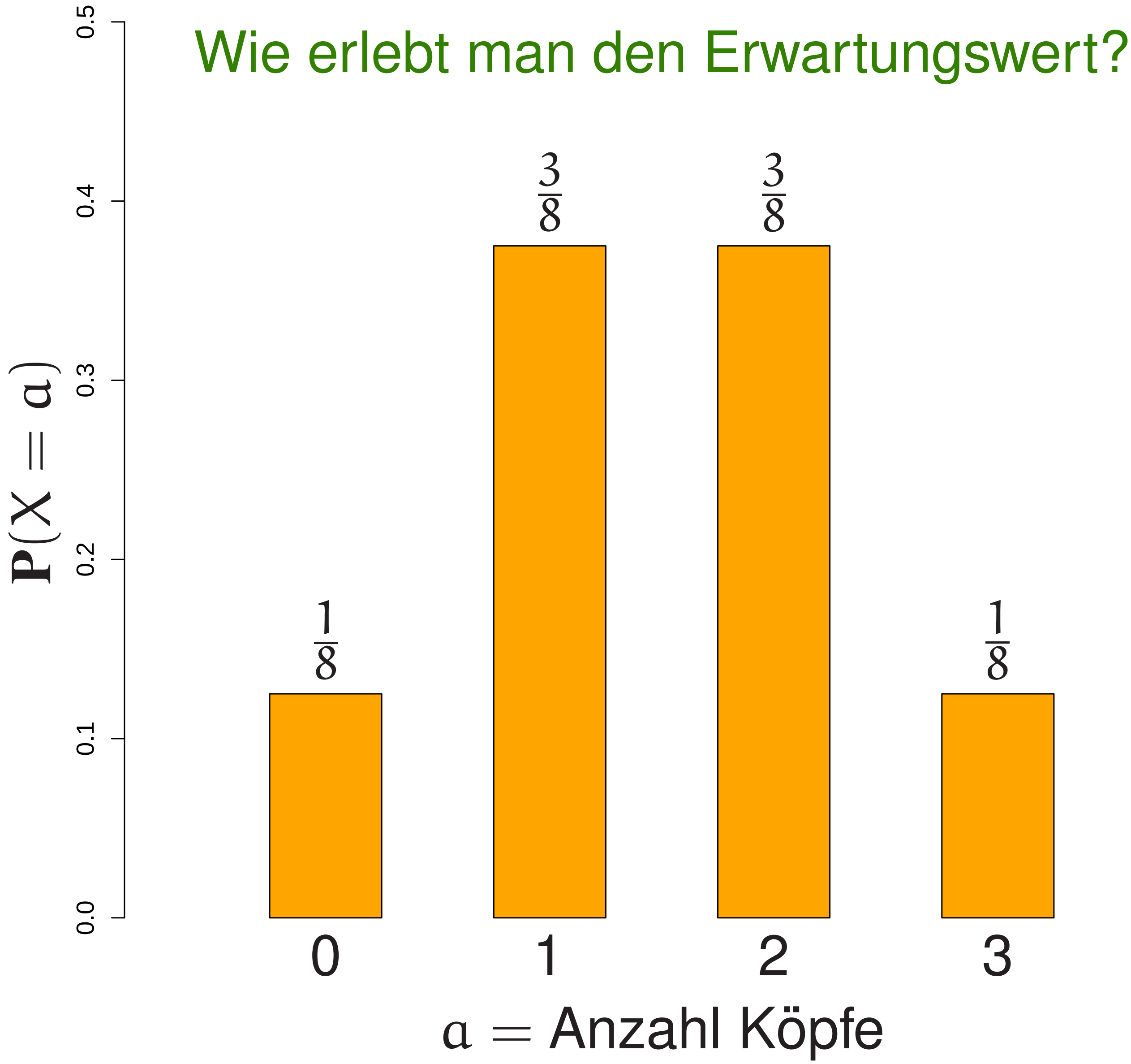


Die Zahl $\mathbf{E}[X] = 1.5$ gehört gar nicht zu S
$\mathbf{P}(X = a)$
0.0
0.1
0.2
0.3
0.4
0.5
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0
1
2
3
a = Anzahl Köpfe

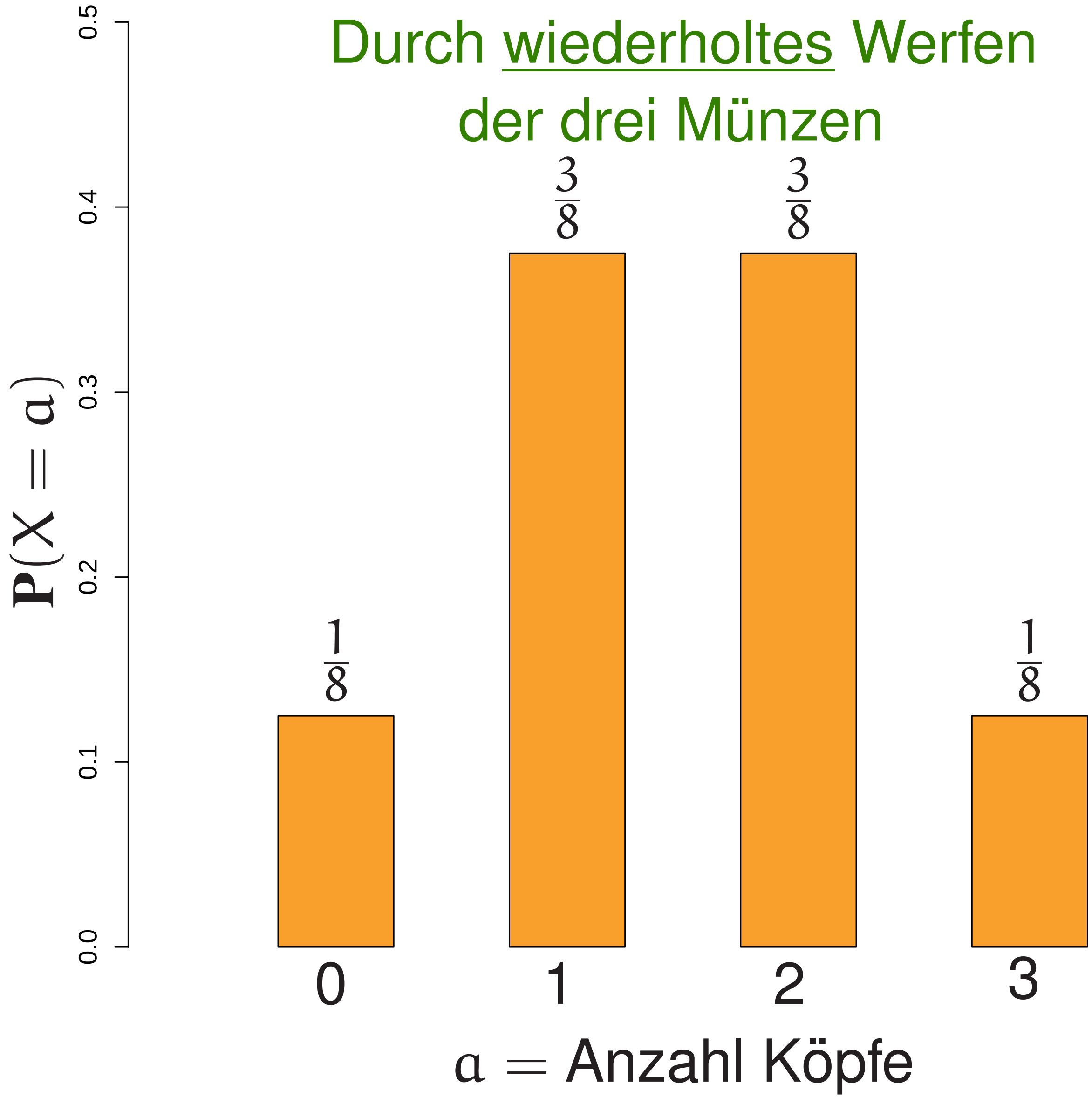
und kann deshalb im buchstäblichen Sinn
kein Wert von X sein, den man erwartet.




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



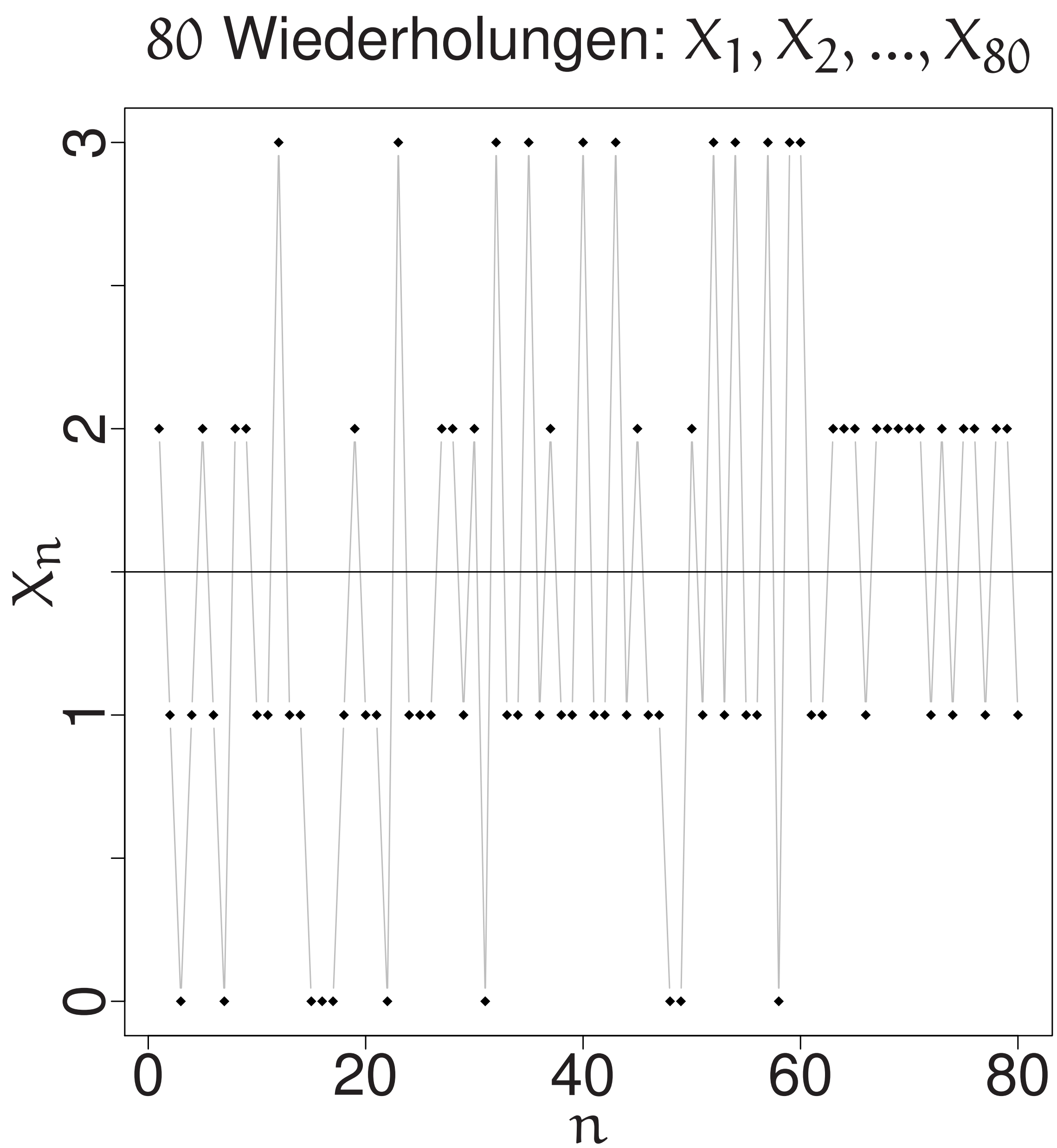
Wie erlebt man den Erwartungswert?
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe

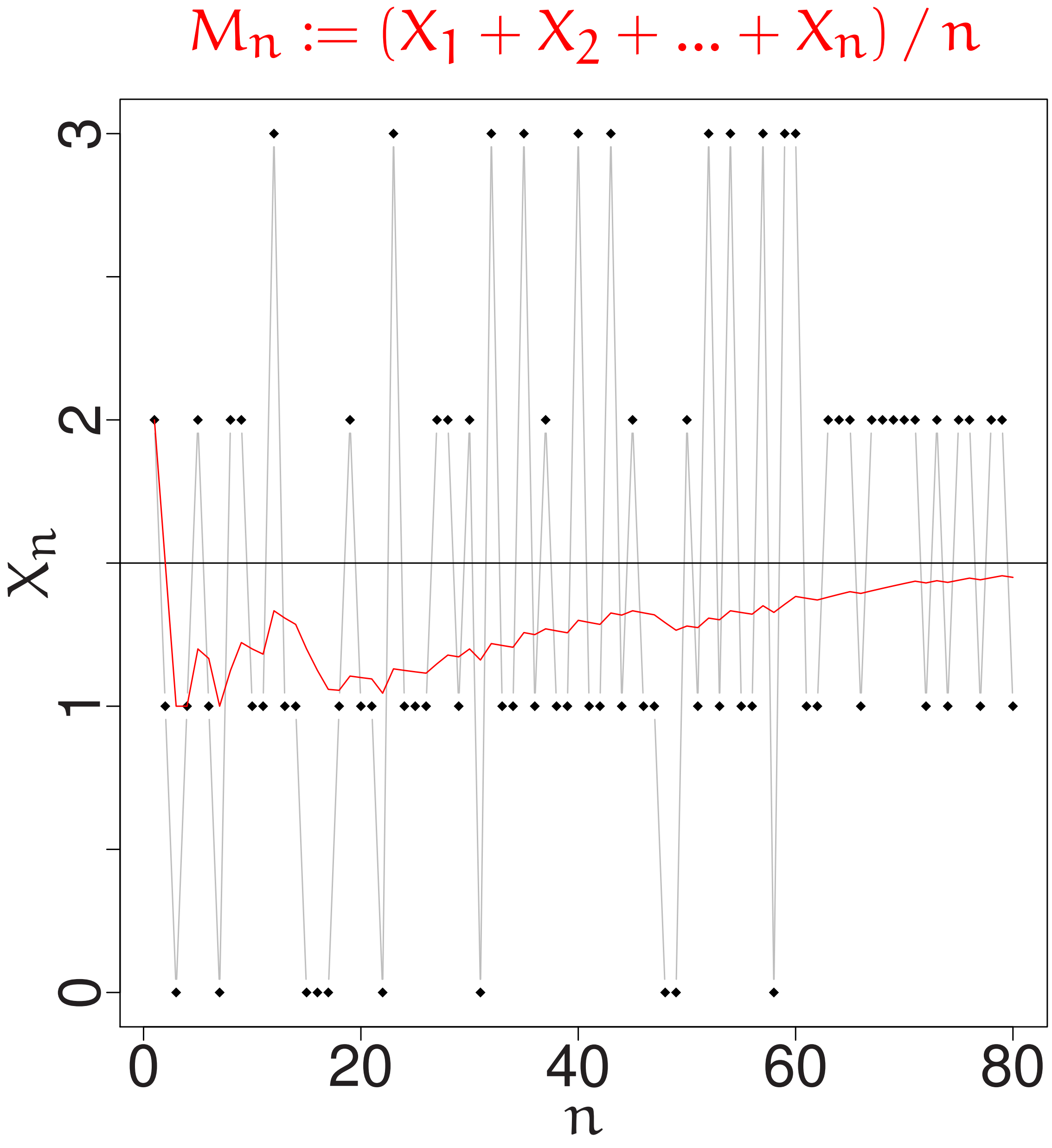

Durch wiederholtes Werfen
der drei Münzen
$\frac{1}{8}$
$\frac{3}{8}$
$\frac{3}{8}$
$\frac{1}{8}$
0.0
0.1
0.2
0.3
0.4
0.5
$\mathbf{P}(X = a)$
0
1
2
3
a = Anzahl Köpfe

Der Erwartungswert als Langzeitmittel

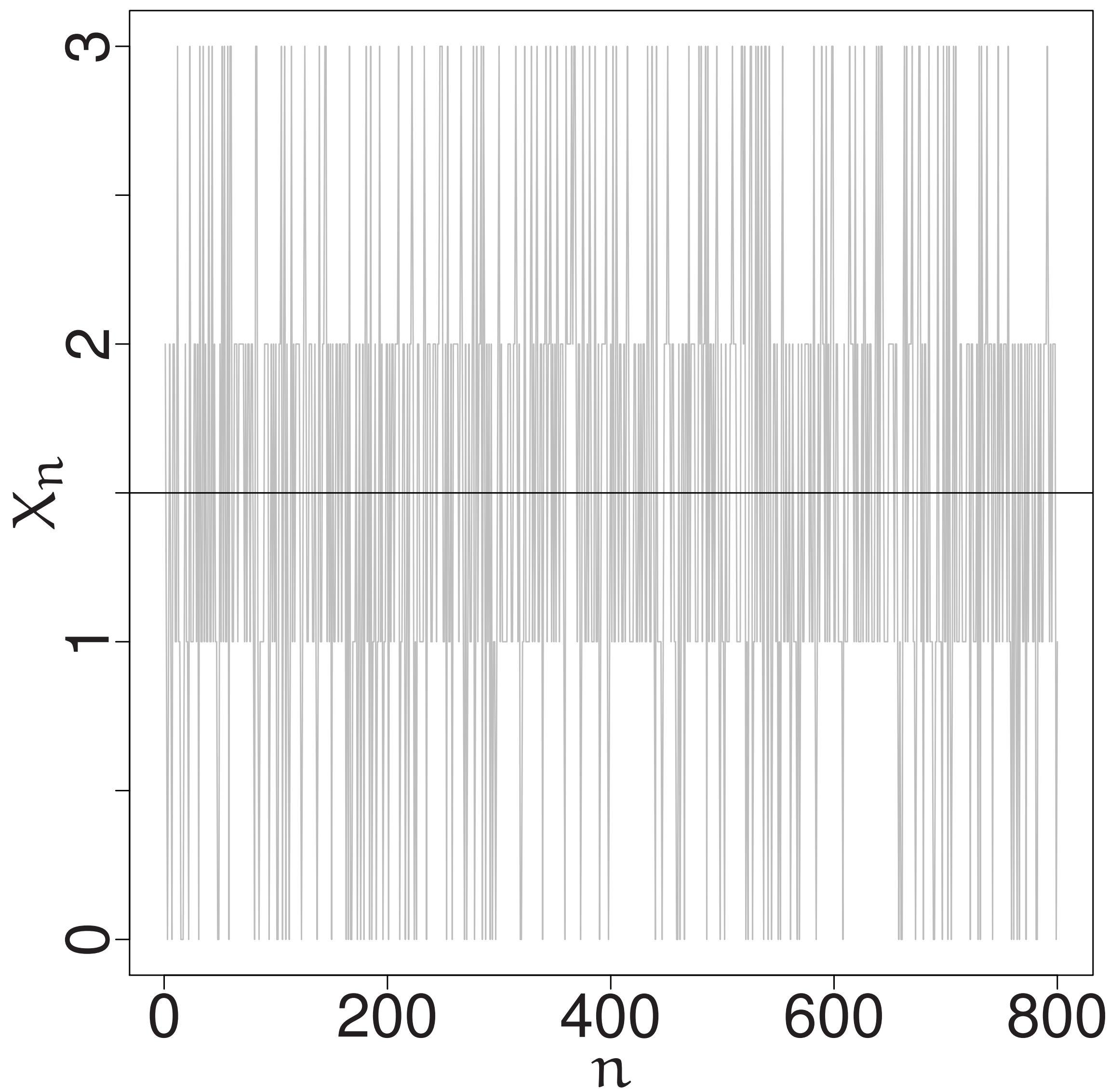
Beispiel:

X . . . Anzahl Köpfe beim dreimaligen fairen Münzwurf

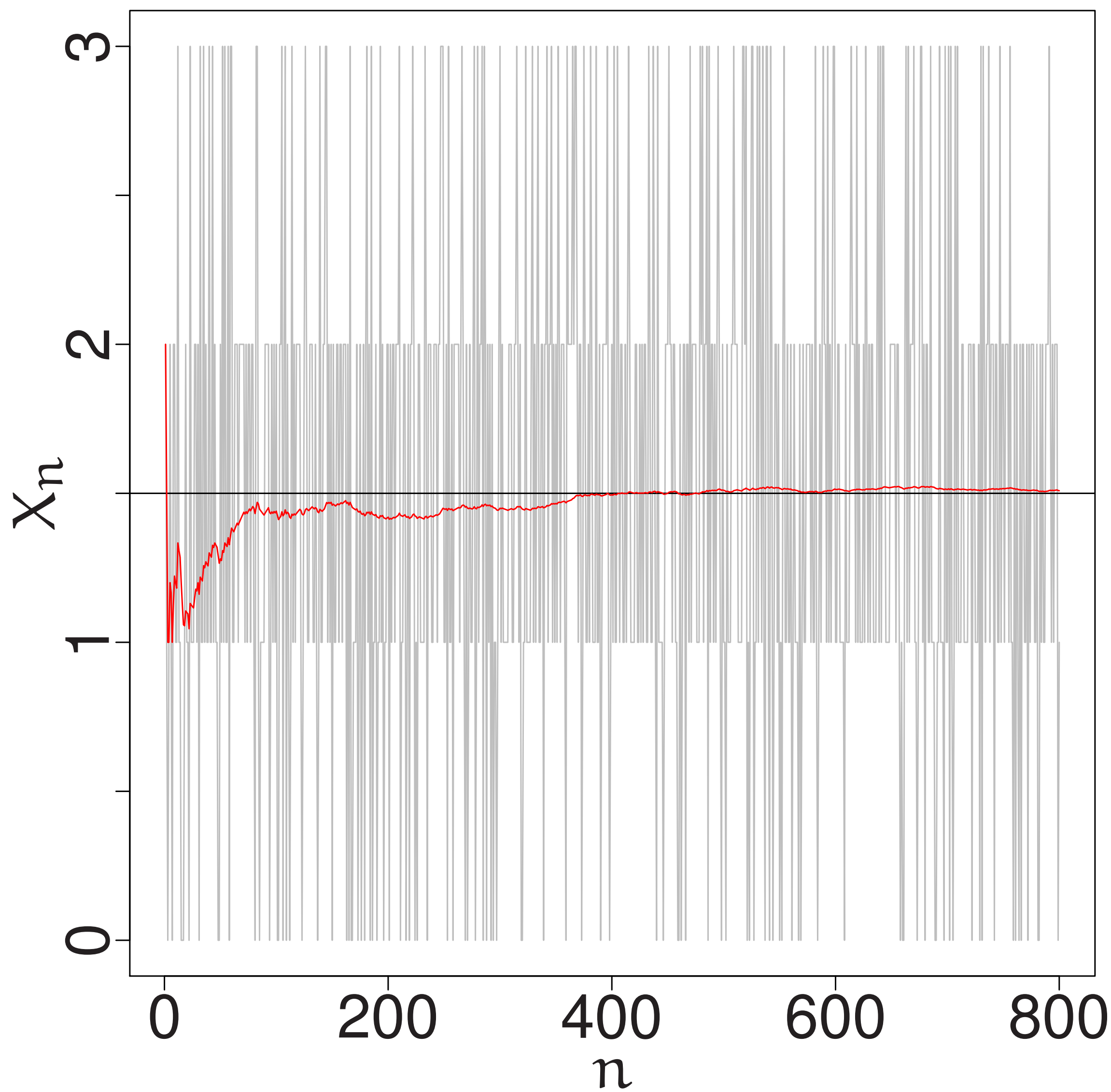

80 Wiederholungen: $X_1, X_2, ..., X_{80}$
X_n
0
1
2
3
0
20
40
60
80
n

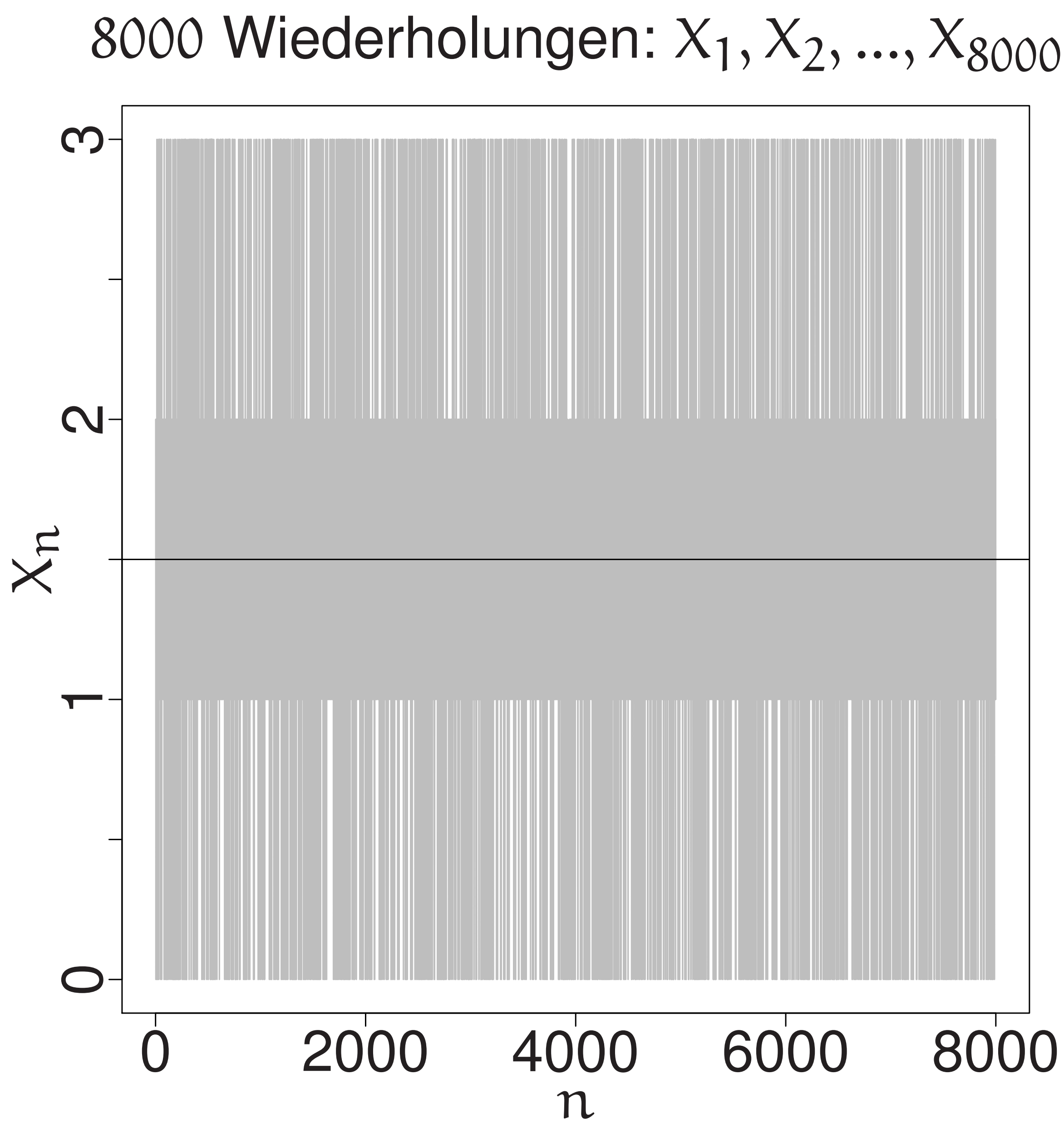

$M_n := (X_1 + X_2 + ... + X_n) / n$
X_n
0
1
2
3
0
20
40
60
80
n

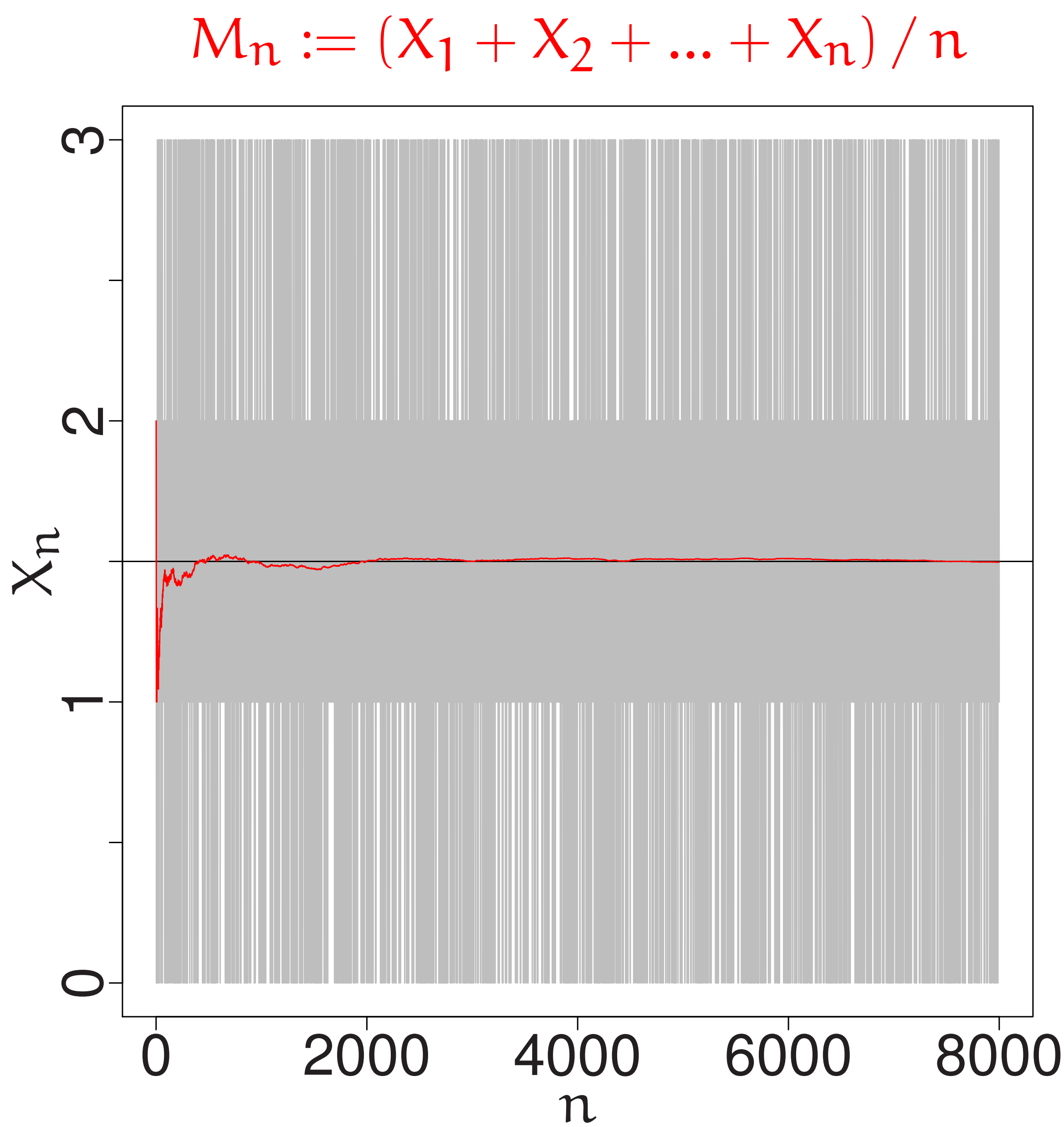
800 Wiederholungen: $X_1, X_2, ..., X_{800}$

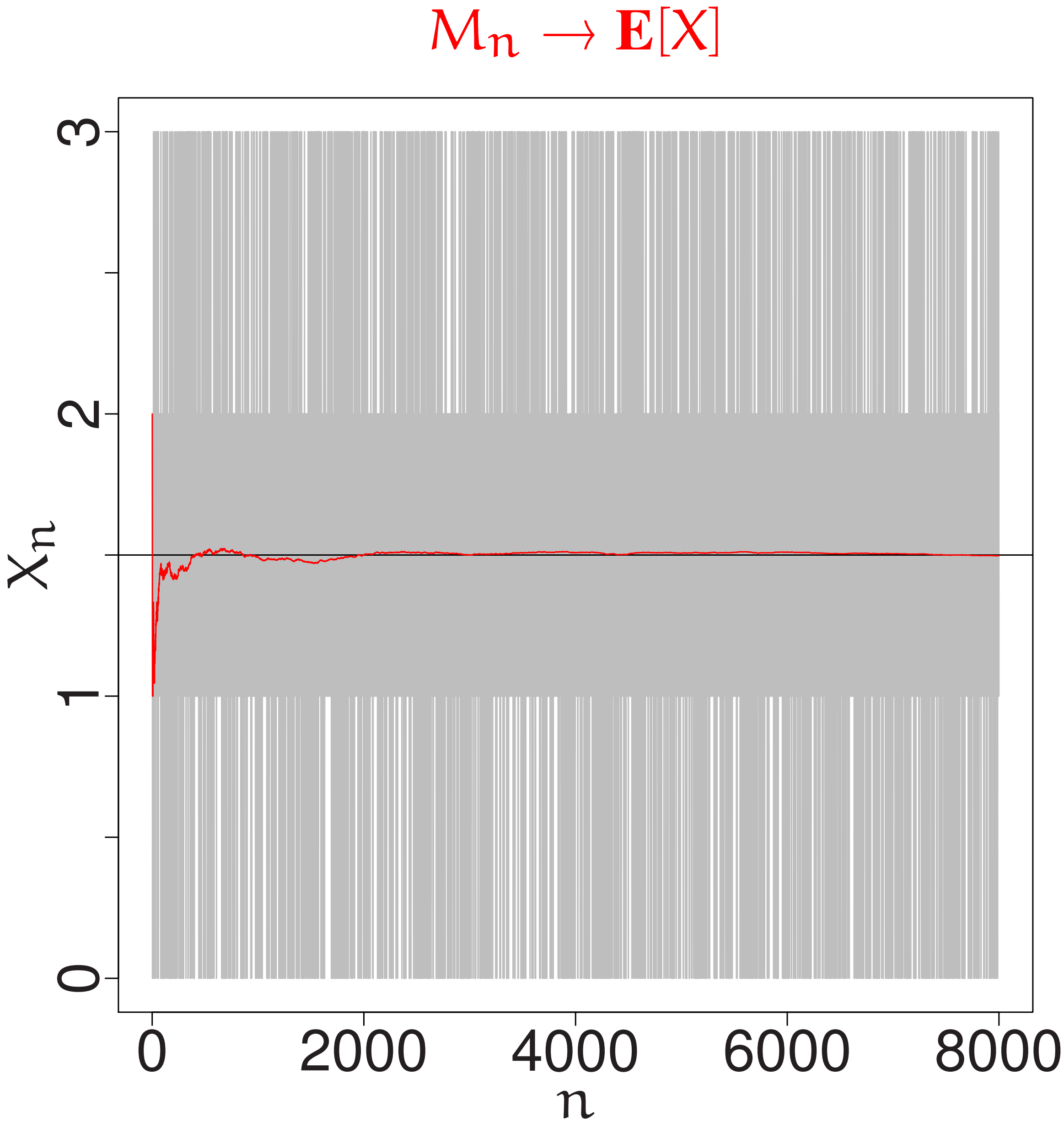


$$M_n := (X_1 + X_2 + \ldots + X_n)\,/\,n$$

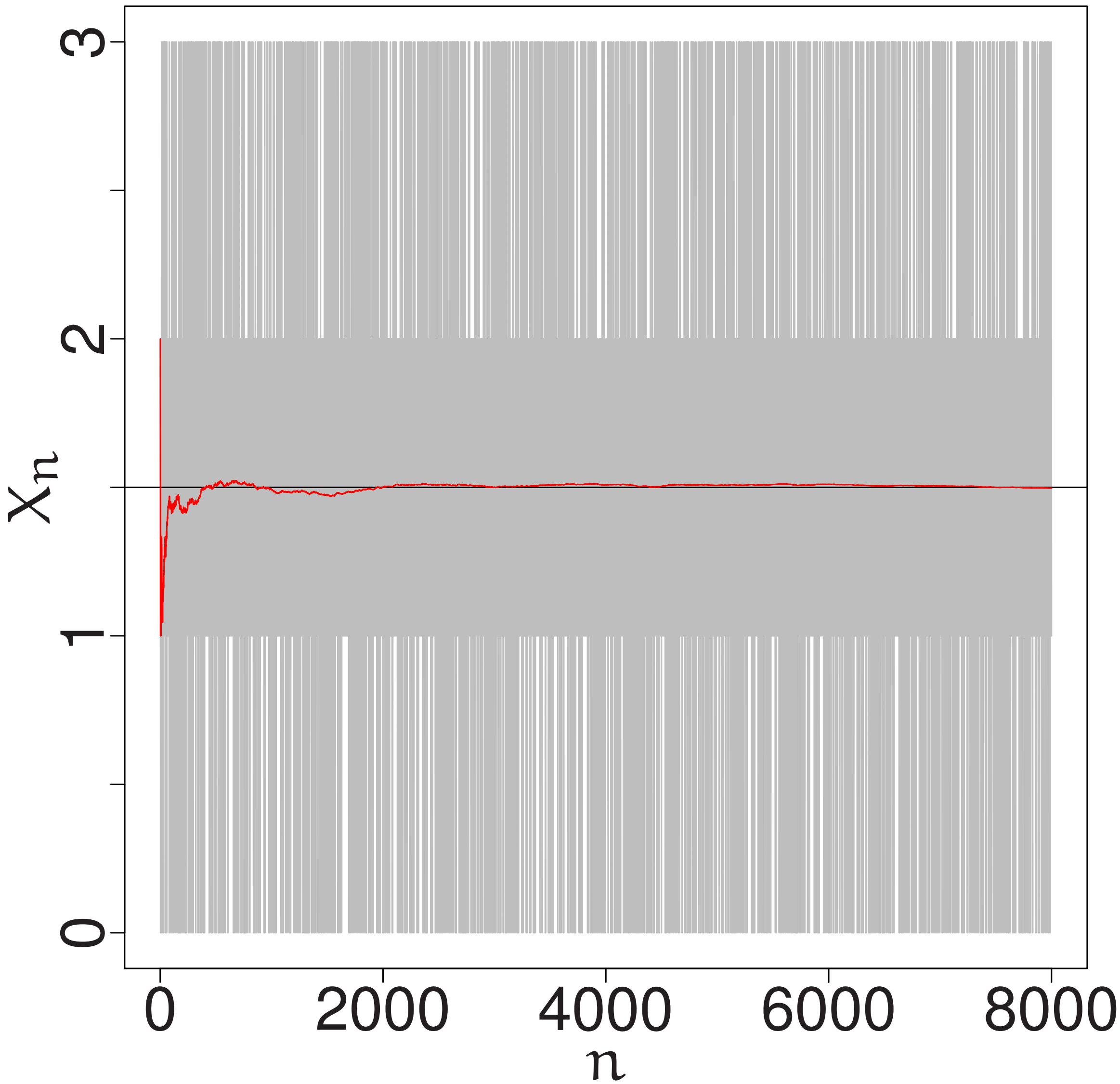



8000 Wiederholungen: $X_1, X_2, ..., X_{8000}$
3
2
1
0
X_n
0
2000
4000
6000
8000
n

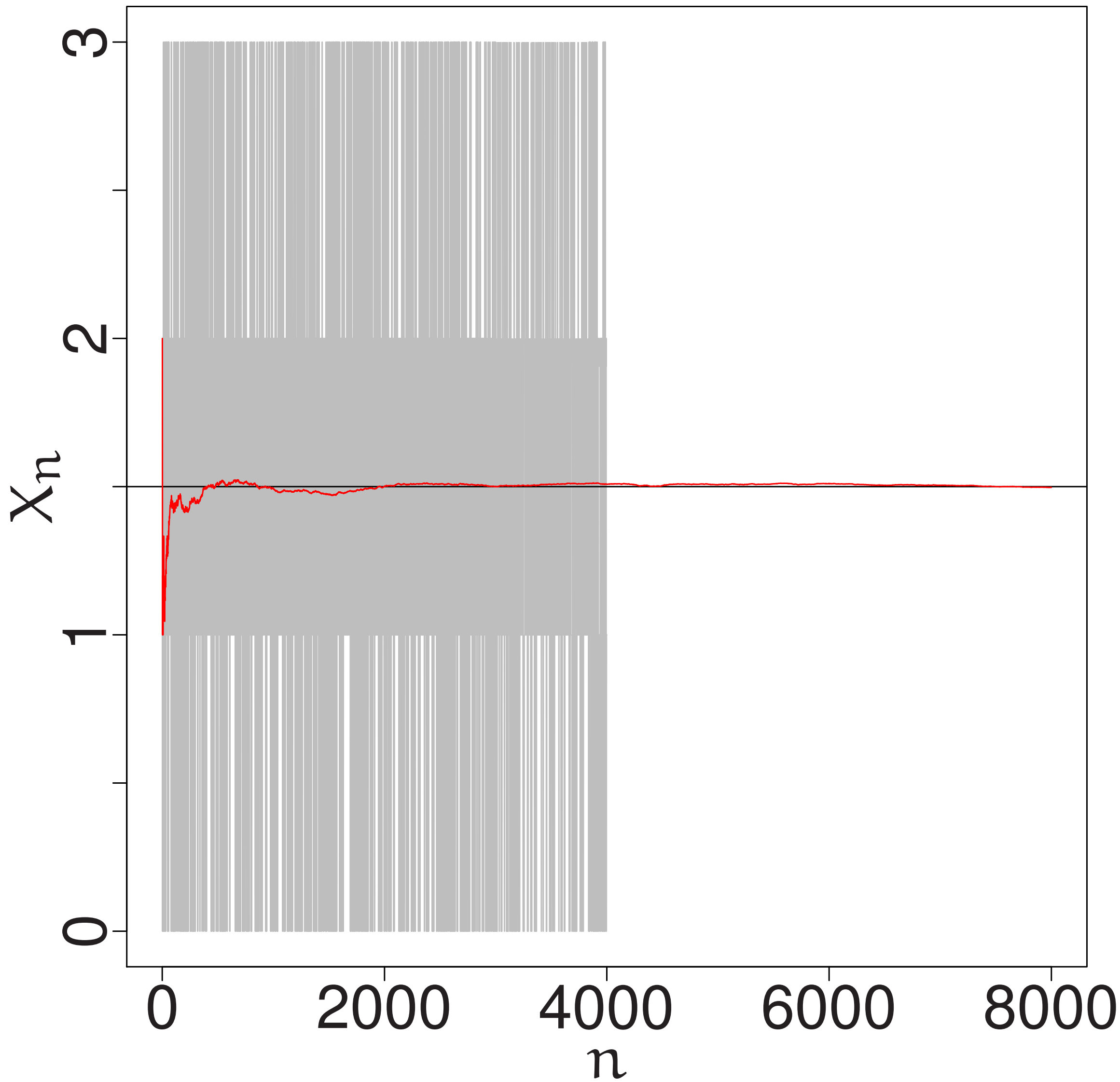

$M_n := (X_1 + X_2 + ... + X_n) / n$
3
2
1
0
X_n
0
2000
4000
6000
8000
n


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

Warum?



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



Die Summe von n Zahlen $x_1, \ldots, x_n \in \{0, 1, 2, 3\}$

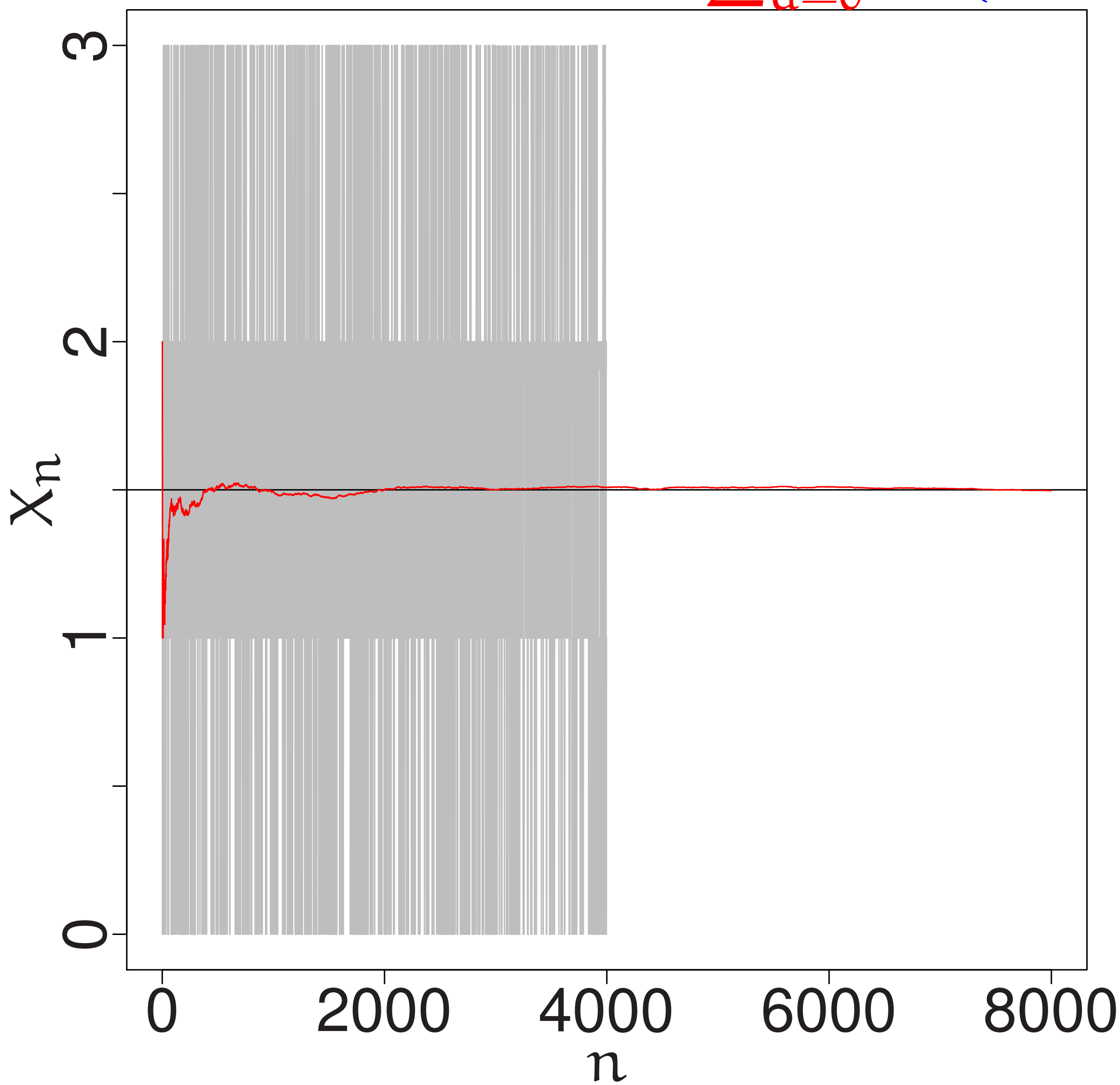
kann man auch so berechnen:

Man zählt für $a = 1, 2, 3$, wieviele der x_i gleich a sind

und bekommt

$$x_1 + \ldots + x_n = \sum_{a=0}^{3} a \, \#\{i \leq n : x_i = a\}.$$

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



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 „ $\rightarrow$ “?

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:

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

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:

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

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

Vorlesung 3b

Indikatorvariable

Rechnen mit Ereignissen und Wahrscheinlichkeiten.

Teil 3

Rechnen mit Wahrscheinlichkeiten.

(Buch S. 57)

Für Ereignisse und ihre Indikatorvariable

gelten die Beziehungen

$$E = \{I_E = 1\},$$

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

Aus dem Rechnen mit Indikatorvariablen

und aus der Linearität des Erwartungswertes

ergeben sich die Regeln

für das Rechnen mit Wahrscheinlichkeiten:

$$\text{(i)}\ \ \mathbf{P}(E_{\mathsf{s}}) = 1, \qquad \mathbf{P}(E_{\mathsf{u}}) = 0,$$

denn

eine $\{0, 1\}$-wertige ZV'e, die stets den Wert 1 annimmt,

hat Erwartungswert 1.

Das passt auch zu unserer Vereinbarung der ersten Stunde:

Für eine S-wertige Zufallsvariable X ist

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

$$\text{(ii)}\ \mathbf{P}(E_1) + \mathbf{P}(E_2) = \mathbf{P}(E_1 \cup E_2) + \mathbf{P}(E_1 \cap E_2)$$

Das sieht man aus der Identität

$$b_1 + b_2 = \max(b_1, b_2) + \min(b_1, b_2), \quad b_1,\, b_2 \in \{0, 1\}$$

mit ihrer Folgerung

$$I_{E_1} + I_{E_2} = I_{E_1 \cup E_2} + I_{E_1 \cap E_2}$$

zusammen mit der Linearität des Erwartungswertes.

Aus (ii) folgt sofort die Subadditivität:

$$\mathbf{P}(E_1 \cup E_2) \le \mathbf{P}(E_1) + \mathbf{P}(E_2)$$

Eine weitere unmittelbare Konsequenz aus (ii) ist

(iii) $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2)$

falls E_1 und E_2 disjunkt.

“(endliche) Additivität”

In der Wahrscheinlichkeitstheorie arbeitet man mit einer stärkeren Form dieser Eigenschaft, der σ-Additivität. Sie besagt:
Für unendliche Folgen $E_1, E_2, \ldots$ von paarweise disjunkten Ereignissen gilt $\mathbf{P}(E_1 \cup E_2 \cup \cdots) = \mathbf{P}(E_1) + \mathbf{P}(E_2) + \cdots$.
Diese Formel entspricht zusammen mit der Eigenschaft (i) den Axiomen von Kolmogorov für Wahrscheinlichkeitsmaße.

Eine unmittelbare Konsequenz aus

der Additivität des Erwartungswertes ist

(iv) $\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$

(Regel von der Gegenwahrscheinlichkeit)

denn: $I_{E^c} = 1 - I_E$.

Für je zwei Ereignisse E_1, E_2 gilt die folgende "Zerlegung von E_2":

$$E_2 = (E_1 \cap E_2) \cup (E_1^c \cap E_2),$$

dabei sind $E_1 \cap E_2$ und $E_1^c \cap E_2$ disjunkt.

Also folgt aus der Additivität:

$$\mathbf{P}(E_1 \cap E_2) \leq \mathbf{P}(E_2).$$

Daraus ergibt sich

(v) $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$

falls das Ereignis E_1 das Ereignis E_2 nach sich zieht

(d.h. falls gilt: $E_1 = E_1 \cap E_2$)

Wir fassen zusammen:

(i) $\mathbf{P}(E_{\mathsf{s}}) = 1, \qquad \mathbf{P}(E_{\mathsf{u}}) = 0.$

(ii) $\mathbf{P}(E_1) + \mathbf{P}(E_2) = \mathbf{P}(E_1 \cup E_2) + \mathbf{P}(E_1 \cap E_2)$,
insbesondere $\mathbf{P}(E_1 \cup E_2) \leq \mathbf{P}(E_1) + \mathbf{P}(E_2)$.

(iii) $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2)$
falls E_1 und E_2 disjunkt.

(iv) $\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$.

(v) $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$, falls $E_1 \subset E_2$.

Vorlesung 3b

Indikatorvariable

Rechnen mit Ereignissen und

Wahrscheinlichkeiten.

Teil 4

Die Einschluss-Ausschlussformel

(Buch S. 57-58)

Die im vorigen Teil diskutierte Eigenschaft

(ii) $\mathbf{P}(E_1) + \mathbf{P}(E_2) = \mathbf{P}(E_1 \cup E_2) + \mathbf{P}(E_1 \cap E_2)$

bzw. äquivalent dazu

$$\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2) - \mathbf{P}(E_1 \cap E_2)$$

wird verallgemeinert durch die

Einschluss-Ausschluss-Formel:

$$\mathbf{P}(E_1 \cup \cdots \cup E_n)$$
$$= \sum_i \mathbf{P}(E_i) - \sum_{i<j} \mathbf{P}(E_i \cap E_j) + \cdots - \cdots \pm \mathbf{P}(E_1 \cap \cdots \cap E_n) .$$

Beweis der Einschluss-Ausschluss-Formel:

$$1 - I_{E_1 \cup \cdots \cup E_n}$$

fällt genau dann auf 0,

wenn mindestens eines der I_{E_i} auf 1 fällt,

ist also gleich dem Produkt

$$(1 - I_{E_1}) \cdots (1 - I_{E_n})$$

Ausmultiplizieren ergibt

$$1 - \sum_i I_{E_i} + \sum_{i<j} I_{E_i \cap E_j} - \cdots .$$

Gehe dann über zum Erwartungswert. $\square$

Beispiel (vgl Buch S. 58). $X = (X_1, \dots, X_n)$ sei eine

rein zufällige Permutation von $1, \dots, n$.

Was ist die W'keit, dass X mindestens einen Fixpunkt hat?

Sei $E_i := \{X_i = i\}$. Wir arbeiten mit der E-A-Formel.

Offenbar gilt für $i_1 < \cdots < i_k$:

$$\mathbf{P}(E_{i_1} \cap \cdots\cdots \cap E_{i_k}) = \frac{(n-k)!}{n!}$$

(denn es gibt $(n-k)!$ Permutationen von $1, \dots, n$

mit Fixpunkten bei $i_1, \dots, i_k$)

Für jedes k gibt es $\binom{n}{k}$ Mögl'k'ten, $i_1 < \cdots < i_k$ zu wählen.

Also:

$$\sum_{i_1<i_2<\cdots<i_k} \mathbf{P}(E_{i_1} \cap \cdots\cdots \cap E_{i_k}) = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

Mit der E-A-Formel folgt für die gefragte W'keit

$$\mathbf{P}(E_1 \cup \cdots \cup E_n) = 1 - \tfrac{1}{2!} + \cdots \pm \tfrac{1}{n!}. \quad \square$$

(Für $n \to \infty$ konvergiert das übrigens gegen $1 - e^{-1}$,

denn aus der Mathe 1 ist bekannt:

$$e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!} = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots,$$
$$e^{-1} = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots$$

Vorlesung 3b

Indikatorvariable

Rechnen mit Ereignissen und Wahrscheinlichkeiten.

Teil 5

Positivität und Monotonie des Erwartungswertes

(Buch S. 55)

Wir beweisen jetzt

(hier nur für *diskrete* rellwertige Zufallsvariable)

zwei weitere fundamentale

Eigenschaften des Erwartungswerts:

die Positivität und die Monotonie.

Als Vorbereitung dazu ist hier ein

Nachtrag zu Teil 1 der heutigen Vorlesung.

Die Aussage "$X \geq 0$" und das Ereignis $\{X \geq 0\}$:

X sei eine reellwertige Zufallsvariable.

Die Aussage "$X \geq 0$"
definieren wir als gleichbedeutend damit,
dass $\{X \geq 0\}$ das sichere Ereignis ist.

Ist $\{X \geq 0\} = E_{\mathsf{S}}$, dann können wir wahlweise $[0, \infty)$
oder $\mathbb{R}$ (oder jede andere Obermenge von $[0, \infty)$)
als Wertebereich von X verwenden.

Die Aussage $X \leq Y$ und das Ereignis $\{X \leq Y\}$:

Es sei $H := \{(x, y) \in \mathbb{R}^2 : x \leq y\}$,

der Halbraum über (und einschließlich) der Diagonalen.

Für reellwertige Zufallsvariable X, Y setzen wir

$$\{X \leq Y\} := \{(X, Y) \in H\}.$$

Die Aussage "$X \leq Y$"

definieren wir als gleichbedeutend damit,

dass $\{X \leq Y\}$ das sichere Ereignis ist.

Positivität des Erwartungswertes

Für die reellwertige Zufallsvariable X gelte $X \geq 0$. Dann gilt

(i) $\mathbf{E}[X] \geq 0$,

(ii) $\mathbf{E}[X] = 0$ genau dann, wenn $\mathbf{P}(X = 0) = 1$.

Monotonie des Erwartungswertes

Für reellwertige Zufallsvariable $X_1 \leq X_2$

mit wohldefinierten Erwartungswerten gilt

$$\mathbf{E}[X_1] \leq \mathbf{E}[X_2].$$

Positivität

Für die reellwertige Zufallsvariable X gelte $X \geq 0$. Dann gilt

(i) $\mathbf{E}[X] \geq 0$,

(ii) $\mathbf{E}[X] = 0$ genau dann, wenn $\mathbf{P}(X = 0) = 1$.

Wir geben hier einen Beweis nur im diskreten Fall:

Nach Voraussetzung (siehe die “Vorbereitung” am Beginn dieses Teils) können wir $[0, \infty)$ als Wertebereich ansehen.

Weil X als diskret vorausgesetzt war, existiert dann eine abzählbare Teilmenge $\tilde{S} \subset [0, \infty)$ mit $\mathbf{P}(X \in \tilde{S}) = 1$.

Dann gilt auch $\mathbf{P}(X \in S) = 1$ mit $S := \tilde{S} \cup \{0\}$.

Nach der Definition des Erwartungswerts aus V3b1 gilt:

$$\mathbf{E}[X] = \sum_{a \in S} a\mathbf{P}(X = a) = \sum_{a \in S: a > 0} a\mathbf{P}(X = a) \qquad (1)$$

Aus $\mathbf{P}(X = S) = 1$ folgt

$$1 = \mathbf{P}(X = 0) + \sum_{a \in S: a > 0} \mathbf{P}(X = a) \qquad (2)$$

Aus (1) folgt sofort: $\mathbf{E}[X] \geq 0$ also die Aussage (i). Weiter gilt:

$$\mathbf{P}(X = 0) = 1$$

$$\overset{(2)}{\Longleftrightarrow}$$

$$\mathbf{P}(X = a) = 0 \text{ für alle strikt positiven } a \in S$$

$$\overset{(1)}{\Longleftrightarrow}$$

$$\mathbf{E}[X] = 0 \quad \square.$$

Monotonie

Für reellwertige Zufallsvariable $X_1 \leq X_2$

mit wohldefiniertem Erwartungswert gilt

$$\mathbf{E}[X_1] \leq \mathbf{E}[X_2].$$

Beweis:

$X_1 \leq X_2$ ist gleichbedeutend mit $X_2 - X_1 \geq 0$.

Aus der Positivität und der Linearität des Erwartungswertes

folgt $\mathbf{E}[X_2] - \mathbf{E}[X_1] \geq 0.$ $\square$

Die Ungleichung von Markov

X reellwertige Zufallsvariable mit $X \geq 0$, $c > 0$. Dann gilt

$$\boxed{\mathbf{P}(X \geq c) \leq \frac{1}{c}\mathbf{E}[X]}$$

Beweis:

Wegen $\quad c\mathbf{1}_{[c,\infty)}(a) \leq a, \quad a \geq 0,$

gilt

$$cI_{\{X \geq c\}} \leq X.$$

Aus Linearität und Monotonie des Erwartungswertes folgt:

$$c\mathbf{E}[I_{\{X \geq c\}}] \leq \mathbf{E}[X]. \quad \square$$

Vorlesung 3b

Indikatorvariable

Rechnen mit Ereignissen und

Wahrscheinlichkeiten.

Teil 2

Rechnen mit Ereignissen

(Buch S. 37-38)

Das *sichere Ereignis* E_{s} hatten wir dadurch charakterisiert, dass seine Indikatorvariable stets den Wert 1 annimmt:

$$I_{E_{\mathrm{s}}} = 1$$

Das *unmögliche Ereignis* E_{u} ist dadurch charakterisiert, dass seine Indikatorvariable stets auf den Wert 0 fällt:

$$I_{E_{\mathrm{u}}} = 0.$$

Für zwei Ereignisse E_1, E_2

hat deren "Oder-Ereignis" die Indikatorvariable

$$I_{E_1 \cup E_2} := \max(I_{E_1}, I_{E_2})$$

und deren "Und-Ereignis" die Indikatorvariable

$$I_{E_1 \cap E_2} := \min(I_{E_1}, I_{E_2}).$$

Man nennt $E_1 \cup E_2$ bzw. $E_1 \cap E_2$ auch

die "Vereinigung" bzw. den "Durchschnitt"

der Ereignisse E_1 und E_2.

Für $b_1, b_2 \in \{0, 1\}$ gilt die Identität:

$$b_1 + b_2 = \max(b_1, b_2) + \min(b_1, b_2).$$

Dies überträgt sich auf eine Gleichheit von Zufallsvariablen:

$$I_{E_1} + I_{E_2} = I_{E_1 \cup E_2} + I_{E_1 \cap E_2}$$

Falls

$$E_1 \cap E_2 = E_{\mathrm{u}} \; ,$$

so heißen E_1 und E_2

disjunkte oder *sich ausschließende Ereignisse*.

Gilt $E_1 = E_1 \cap E_2$, so schreiben wir

$$E_1 \subset E_2 \, .$$

und sagen:

“Mit E_1 tritt sicher auch E_2 ein”

oder auch

“Das Ereignis E_1 zieht das Ereignis E_2 nach sich.”

Für jedes Ereignis E ist sein *Komplementärereignis*

$$E^c$$

definiert durch

$$I_{E^c} := 1 - I_E \quad \text{bzw.} \quad E^c := \{I_E = 0\}\ .$$

Wegen $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$ gilt

$$\{X \in A\}^c = \{X \in A^c\}\ .$$

Vorlesung 3b

Indikatorvariable

Rechnen mit Ereignissen und Wahrscheinlichkeiten.

Teil 1

Ereignisse und ihre Indikatorvariablen

(Buch S. 36-37)

Wir wissen schon:

Ein- und dasselbe Ereigniss kann man

auf verschiedene Weisen darstellen:

Beispiel:

Sei $X = (X_1, X_2)$ das Paar der Augenzahlen

beim zweimaligen (gewöhnlichen) Würfeln. Dann gilt:

$$\{X_1 = 3\} = \big\{X_1 = 3,\ X_2 \in \{1, 2, 3, 4, 5, 6\}\big\}$$

Allgemeiner:

Sei $X = (X_1, X_2)$ ein zufälliges Paar

mit Wertebereich $S_1 \times S_2$.

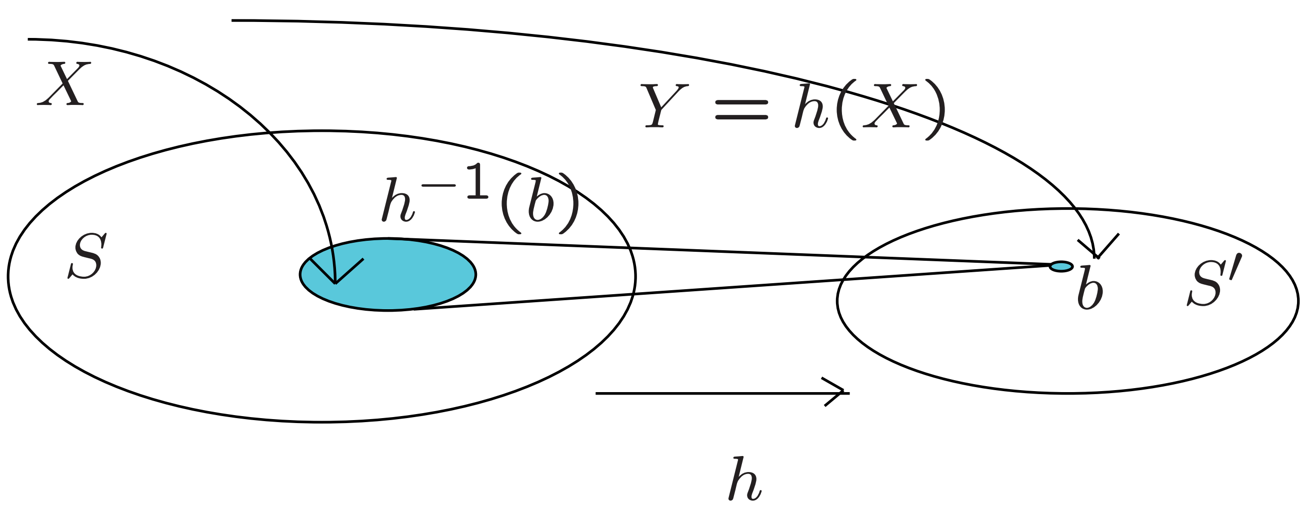
Dann ist für $a_1 \in S_1$

$\{X_1 = a_1\} = \{X_1 \in a_1, X_2 \in S_2\}$.

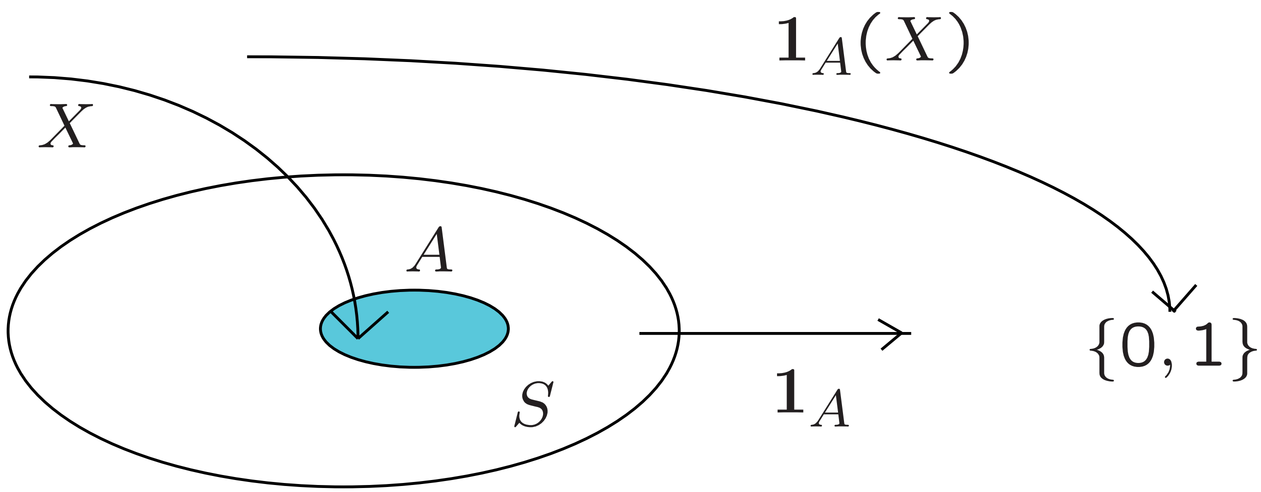
Noch allgemeiner gilt

für die “Verarbeitung” $Y = h(X)$ einer Zufallsvariablen X:

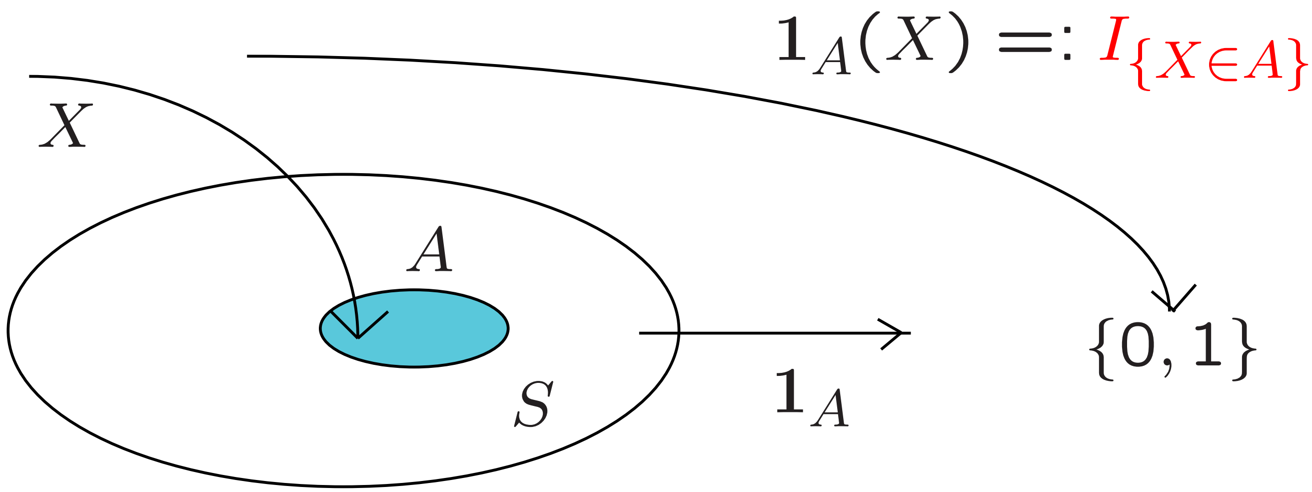
$$\boxed{\{Y = b\} = \{X \in h^{-1}(b)\}.}$$



Ein wichtiger Spezialfall hiervon ist die Verarbeitung von X mittels der Indikatorfunktion $\mathbf{1}_A$ einer Teilmenge A von S:



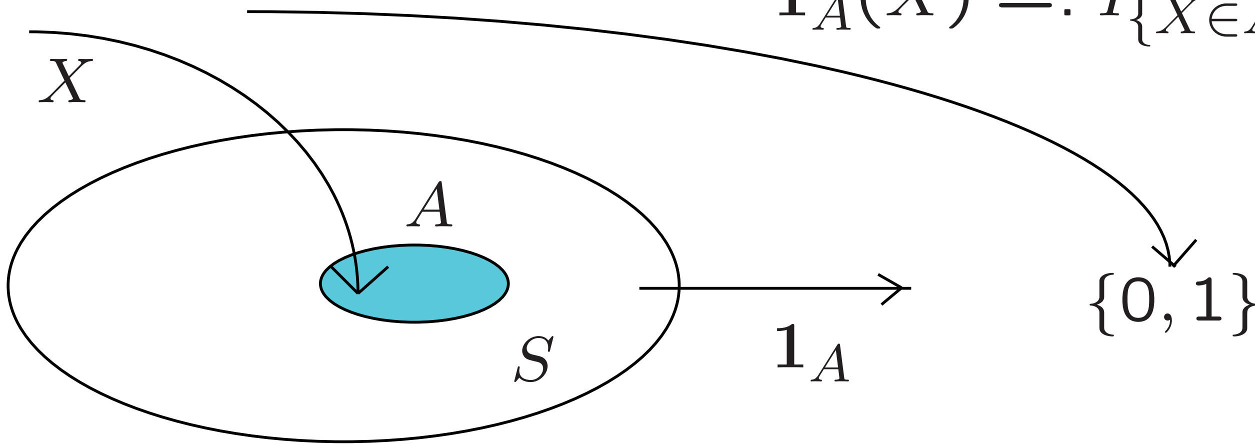
$$\{X \in A\} = \{\mathbf{1}_A(X) = 1\}$$



$$\{X \in A\} = \{\mathbf{1}_A(X) = 1\}$$

$I_{\{X\in A\}}$ heißt Indikatorvariable des Ereignisses $\{X \in A\}$

$$\mathbf{1}_A(X) =: I_{\{X \in A\}} = I_E$$

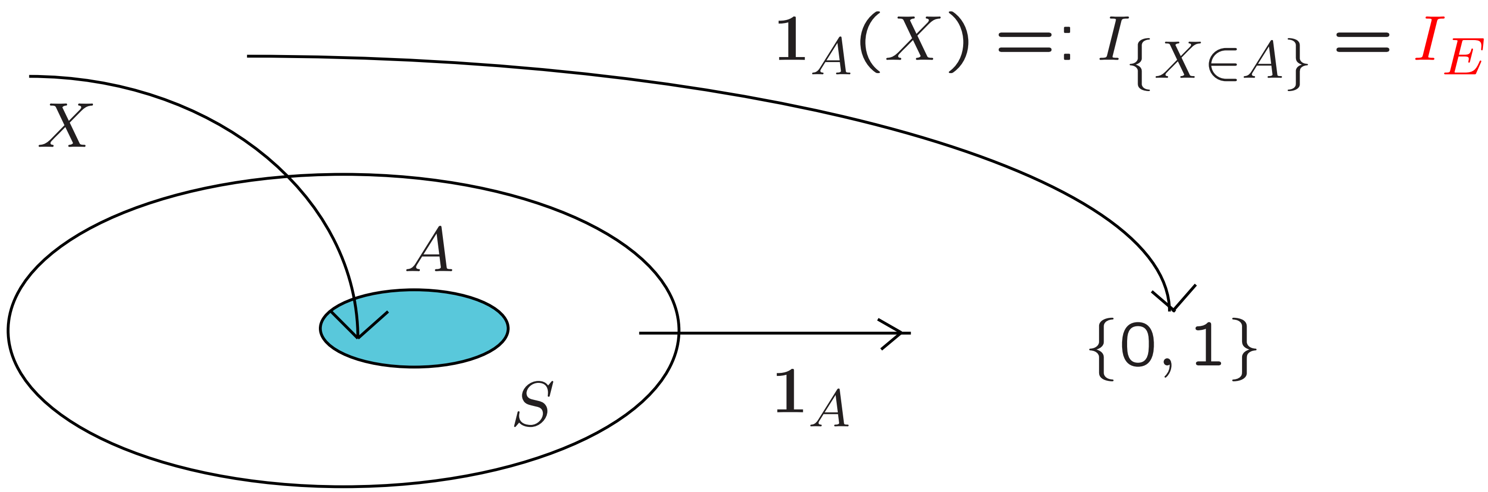


$$E := \{X \in A\} = \{\mathbf{1}_A(X) = 1\}$$

$$E = \{I_E = 1\}$$

I_E fällt auf den Wert 1 genau dann,

wenn das Ereignis E eintritt.



$$E := \{X \in A\} = \{\mathbf{1}_A(X) = 1\}$$

$$E = \{I_E = 1\}$$

I_E fällt auf den Wert 0 genau dann,

wenn das Ereignis E nicht eintritt.

$$E = \{I_E = 1\}.$$

Ereignisse sind gleich,

wenn ihre Indikatorvariablen gleich sind

(in dem Sinn, dass beide stets gemeinsam auf 0

oder gemeinsam auf 1 fallen)

Wann “sind zwei Zufallsvariablen gleich”?

Intuitiv dann, wenn stets

die eine auf denselben Wert fällt wie die andere.

Das präzisieren wir im Folgenden.

Das sichere Ereignis

Für jede Zufallsvariable X mit Wertebereich S gilt:

$$I_{\{X \in S\}} = \mathbf{1}_S(X)$$

ist eine Zufallsvariable, die “sicher” auf den Ausgang 1 fällt.

Das *sichere Ereignis* E_S ist dadurch charakterisiert,

dass seine Indikatorvariable stets auf den Wert 1 fällt:

$$I_{E_\mathsf{S}} = 1.$$

Die Aussage $X = Y$ und das Ereignis $\{X = Y\}$

Seien X, Y Zufallsvariable mit demselben Wertebereich S.

$D := \{(x, y) \in S^2 : x = y\}$, die „Diagonale“ in S^2.

Wir definieren das Ereignis “X und Y fallen gleich aus”

als

$$\{X = Y\} := \{(X, Y) \in D\}$$

Ist dieses Ereignis gleich dem sicheren Ereignis,

so schreiben wir dafür kurz:

$$X = Y$$