Stochastik für die Informatik

Wintersemester 2016/17

Anton Wakolbinger

StofI-Webseite:

http://ismi.math.uni-

frankfurt.de/wakolbinger/teaching/StofI1617/

Oder:

http://www.uni-frankfurt.de/53215133

Oder:

google $\rightarrow$ "Wakolbinger"

Übungsgruppen zur Stochastik für die Informatik, WS 2016/17

Gruppe	Zeit	Ort		Tutorin/Tutor
1	Mi 8-10	Raum 711 gr	RM10	Jasmin Straub
2	Mi 10-12	Raum 902	RM10	Hans-Joachim Kunz
3	Do 10-12	Raum 903	RM10	Theresa Kumpitsch
4	Do 12-14	Raum 902	RM10	Thomas Fischer
5	Fr 8-10	Raum 711 gr	RM10	Anna Kremer
6	Fr 14-16	Raum 711 gr	RM10	Hung The Tran

Anmeldung zu einer Übungsgruppe:

elektronisch über OLAT ($\leftarrow$ StofI - Webseite)

bis Sonntag 23. Oktober 2016, 24 Uhr

nach dem first come - first serve Prinzip

Freitags: Ausgabe des Übungsblatts.

Tipps zu den Übungsaufgaben:

in den Tutorien (ab nächster Woche)

Termin für die Abgabe der schriftlichen Lösungen der “S-Aufgaben”:

am Dienstag (11 Tage nach Ausgabe des Blattes)

In der Woche der Abgabe

werden die Lösungen in den Tutorien besprochen.

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.

Abschlussklausur:

Dienstag, 23. Februar 2016, 10:15-11:45 Uhr.

Dabei können 100 Klausurpunkte erreicht werden.

Die Note errechnet sich aus

der Summe der Anzahl der erreichten Klausurpunkte

plus der Anzahl der erreichten Bonuspunkte.

Beträgt diese Summe mindestens 50,

gilt die Abschlussprüfung über die Veranstaltung

als bestanden.

Lehrbuch:

Götz Kersting, Anton Wakolbinger

Elementare Stochastik, Birkhäuser, 2. Aufl. 2010,

Preis: 18,90 EUR

Semesterausleihe möglich aus der

Bibliothek des Mathematischen Seminars,

Robert-Mayer-Str. 8, 4. Stock

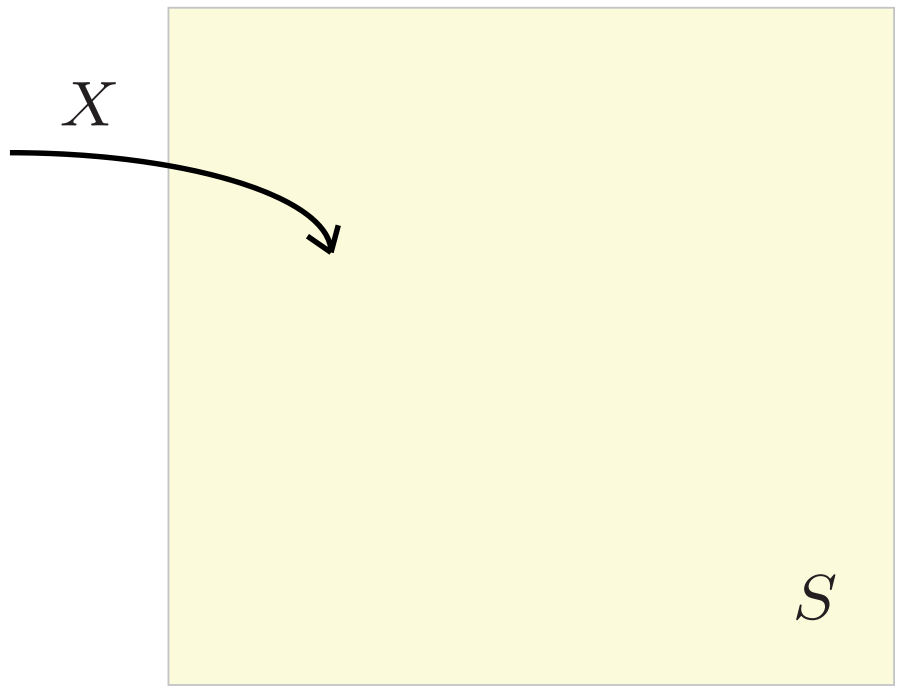
in der UB als E-Book vorhanden

Vorlesung 1a

Zufallsvariable und Ereignisse, Wahrscheinlichkeiten und Verteilungen

vorgestellt an einem Beispiel

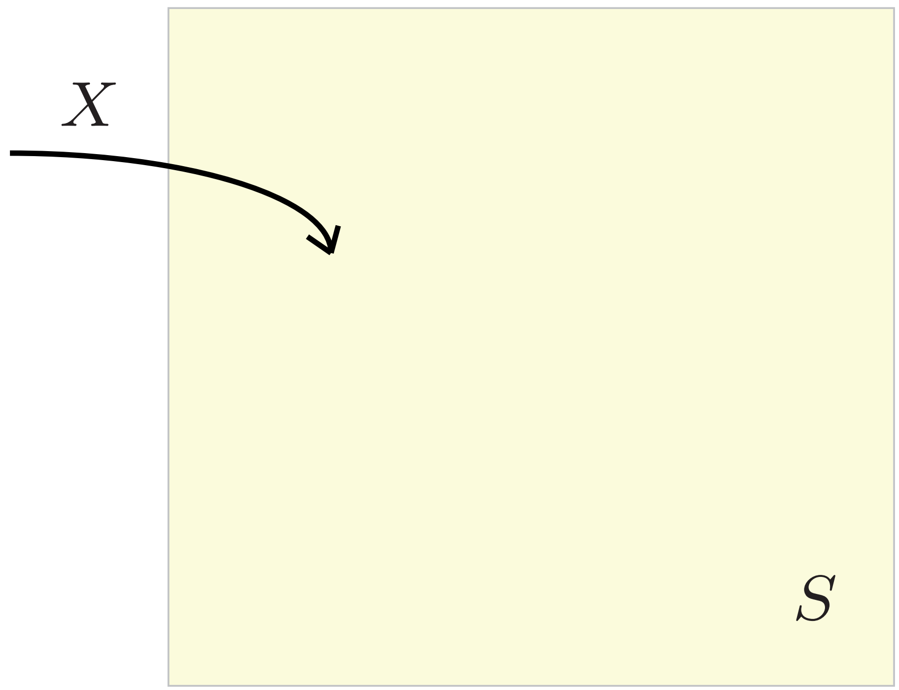
Prototypisches Beispiel einer Zufallsvariablen:



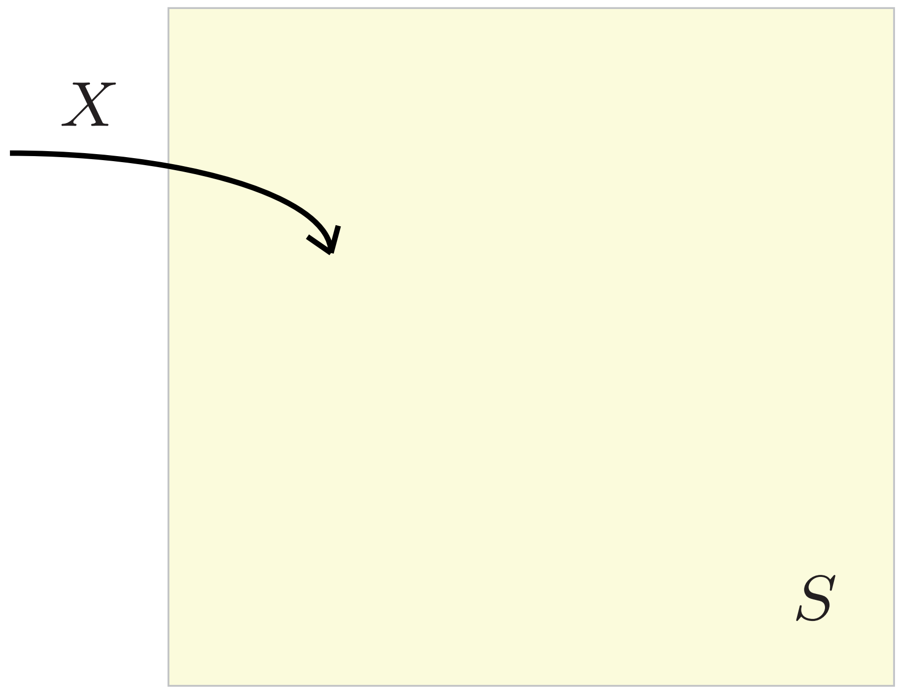
rein zufällige Wahl eines Punktes aus einer Fläche

(z. B. aus dem angegebenen Quadrat S)

Prototypisches Beispiel einer Zufallsvariablen:



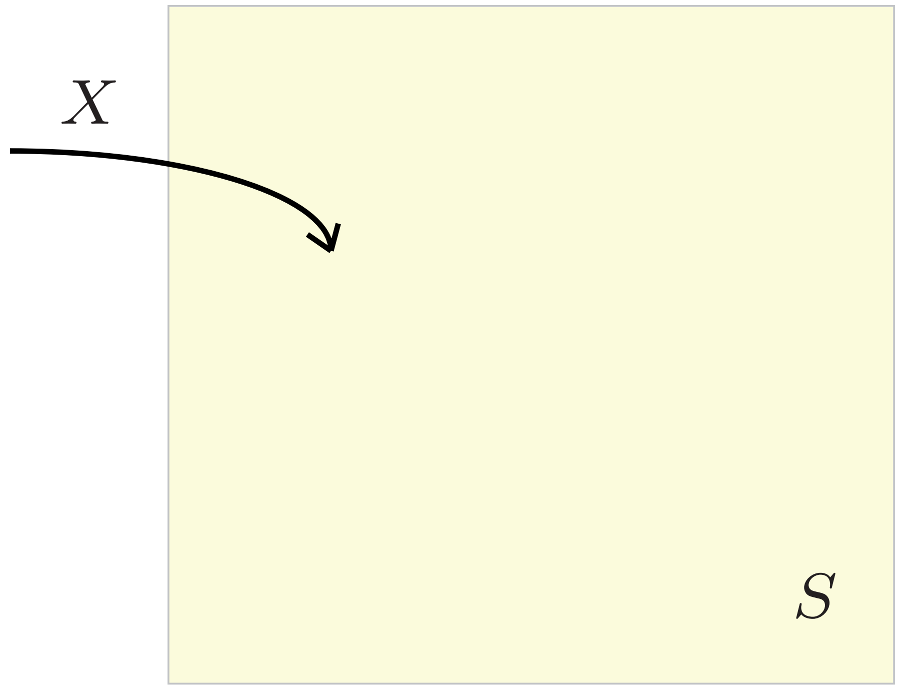
Was heißt “rein zufällige Wahl” eines Punktes?



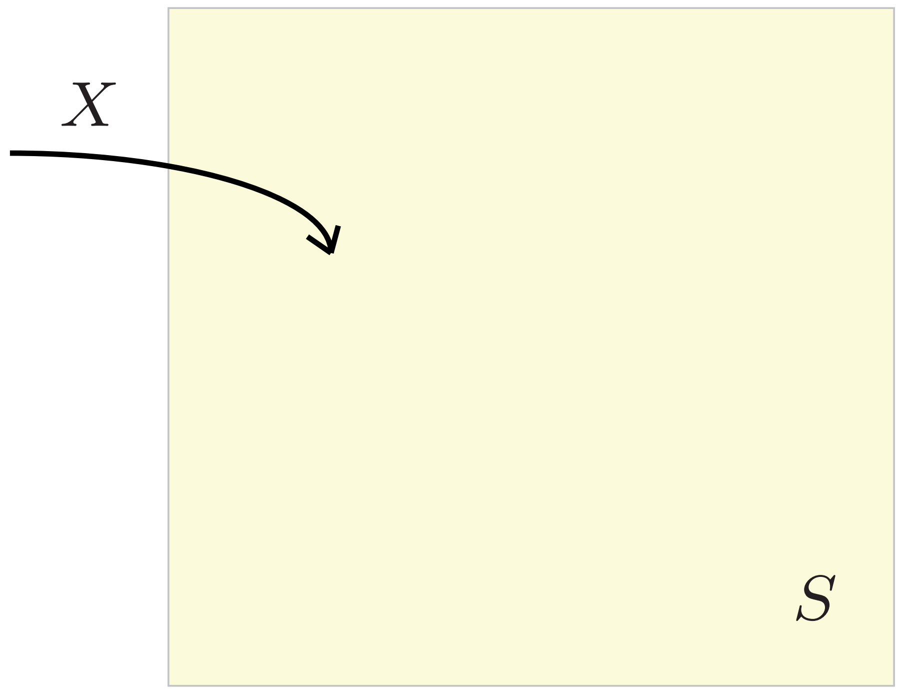
Fürs Erste sind (endlich viele) Pixel leichter vorstellbar

als (kontinuierlich viele) Punkte.

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

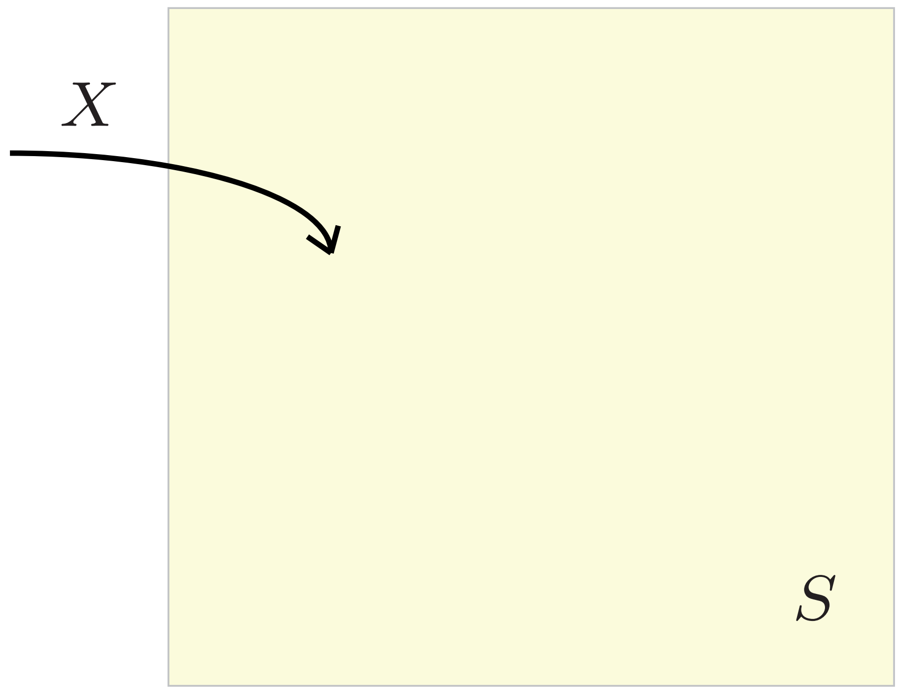


"Rein zufällige Wahl aus S" soll heißen:



alle Pixel in S

haben dieselbe Chance, zum Zug zu kommen



Man spricht dann von einer

uniform auf S verteilten Zufallsvariablen X

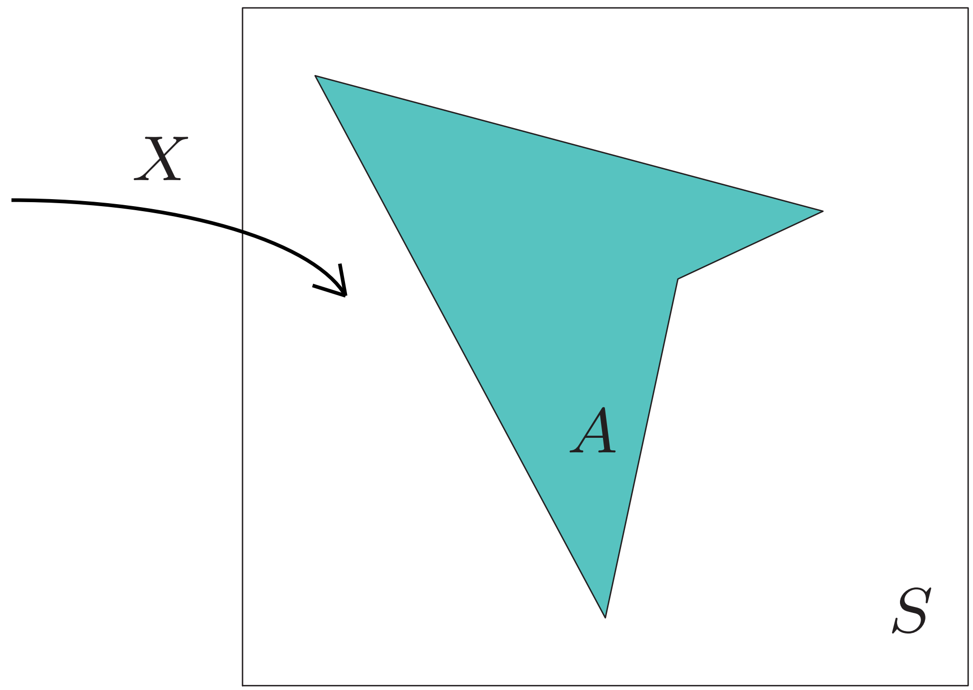
Analogie zum fairen Würfeln:

Die Menge der möglichen Ausgänge ist hier

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

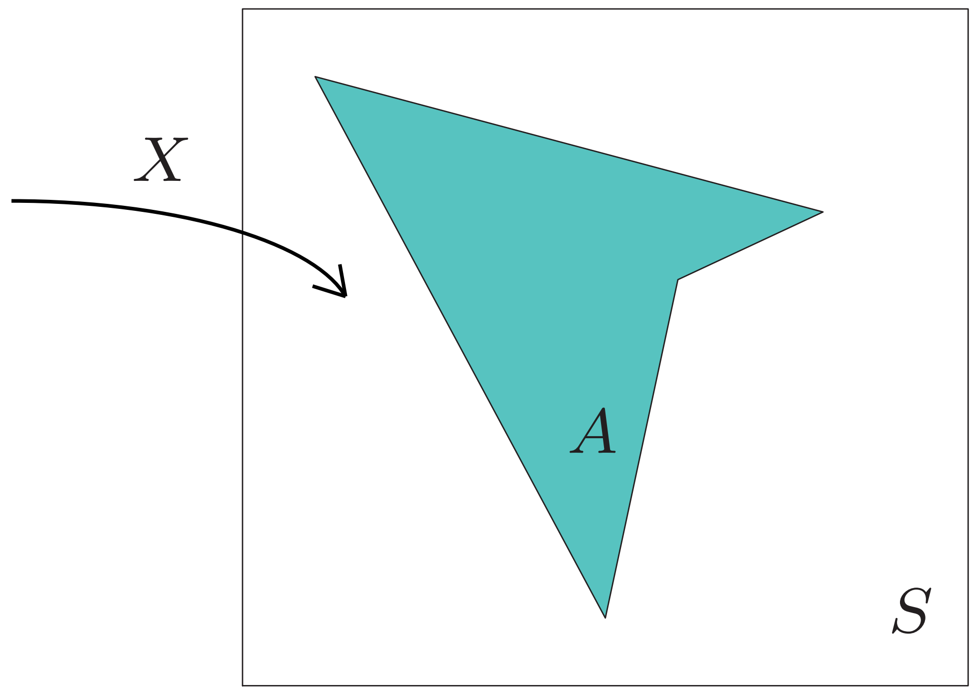
Der zufällige Ausgang wird hier beschrieben durch eine uniform auf S verteilte Zufallsvariable X.

Betrachten wir wieder unser Quadrat S



diesmal zusammen mit einer bestimmten Teilmenge A von S

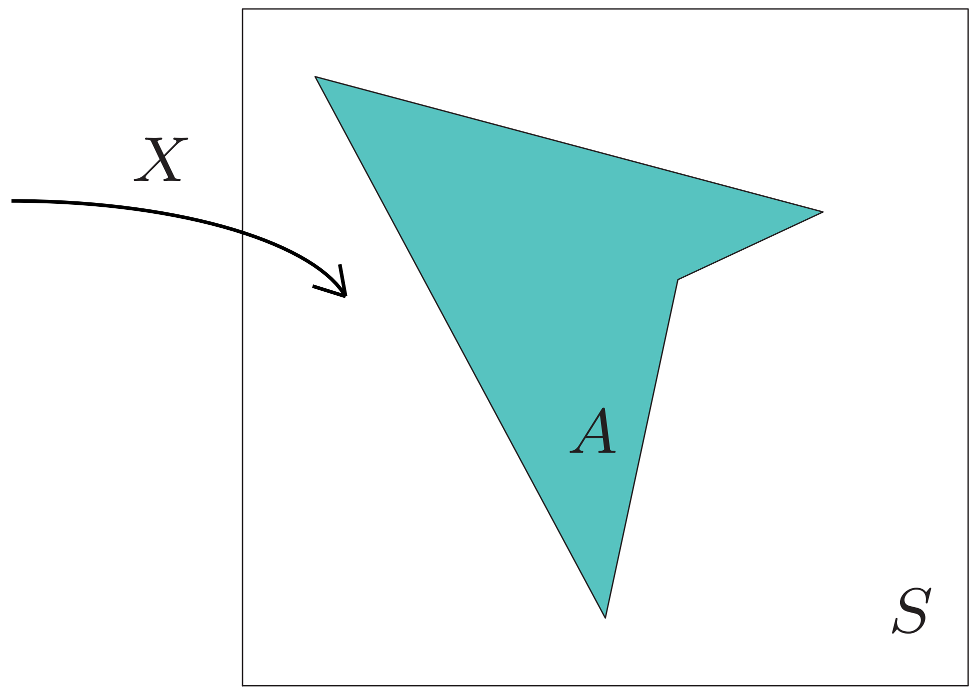
Bei der rein zufälligen Wahl eines Pixels aus S



kann die Wahl auf A fallen

– oder auch nicht.

Bei der rein zufälligen 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 $\{X \in A\}$, $A \subset S$,

nennt man auch die

“von X erzeugte Kollektion von Ereignissen”.

Ereignisse kann man aussagenlogisch verknüpfen, z.B.:

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

Mehr dazu später!

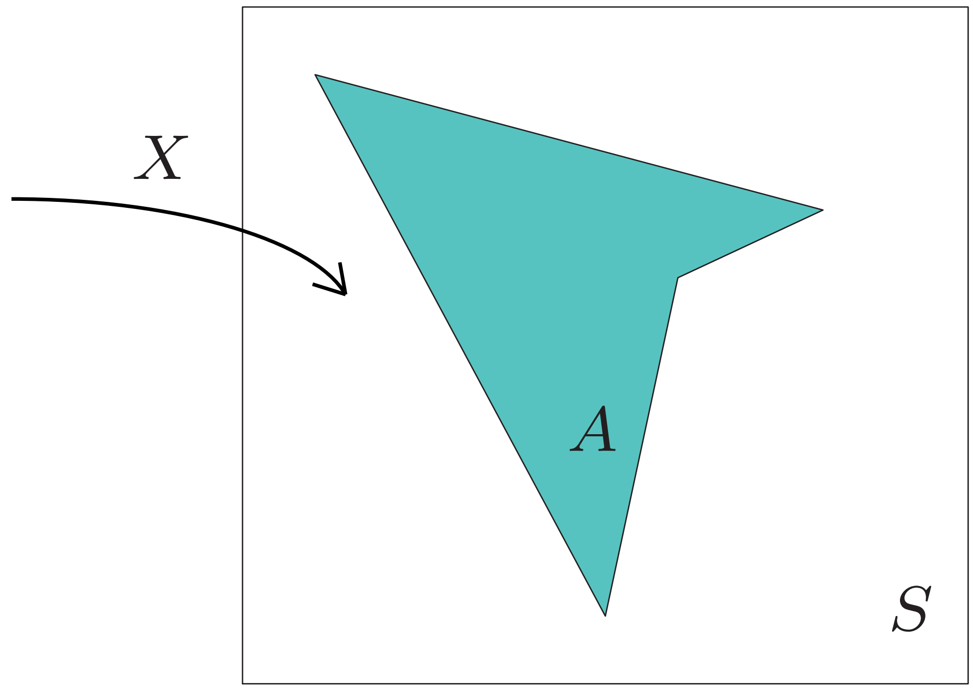
Wie wahrscheinlich ist es,

dass bei einer rein zufälligen Wahl eines Pixels aus S

die Wahl auf A fällt?

Wie groß ist die Wahrscheinlichkeit

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



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.

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)

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:

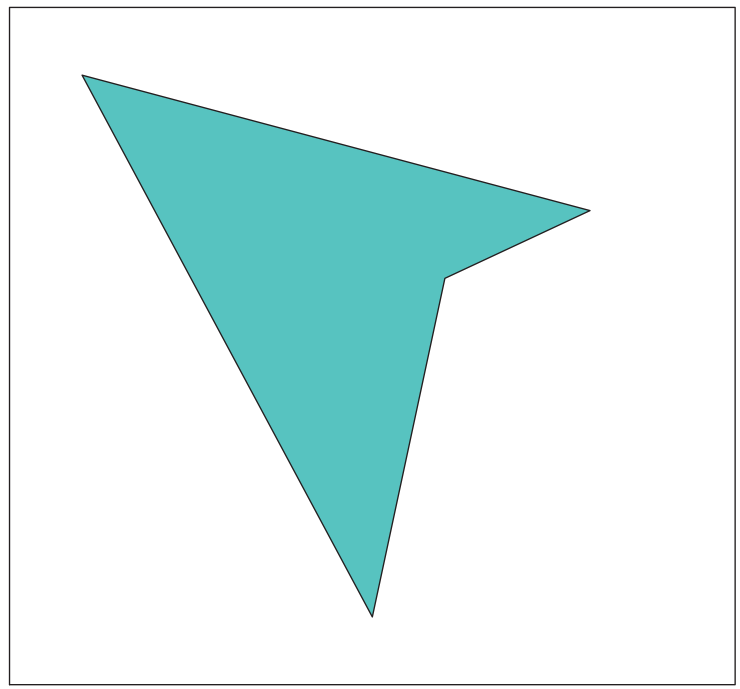
die Wahrscheinlichkeit, dass X in A fällt

ist der Anteil der Menge A an der Menge S.

Eine Anwendung der “rein zufälligen Wahl”:

Monte-Carlo Schätzung eines Flächenanteils.

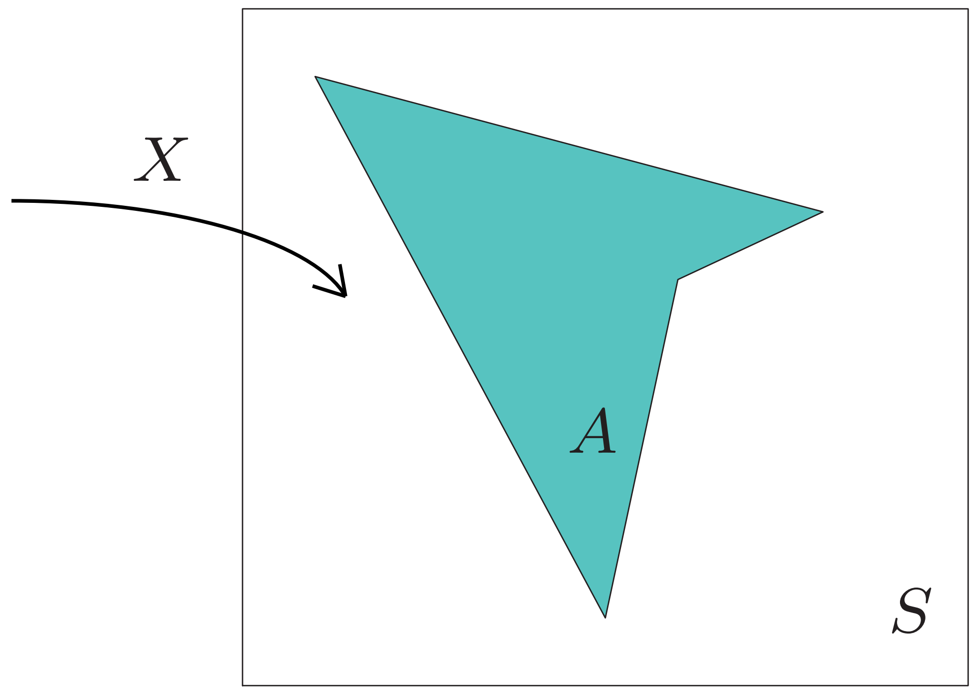
Wir fragen:



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
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

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

Z_i ist der *der Indikator* (*"der Zähler"*)

des Ereignisses "X_i fällt in A".

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

Die zufällige Zahl (die “Trefferquote”)

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

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?

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

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

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

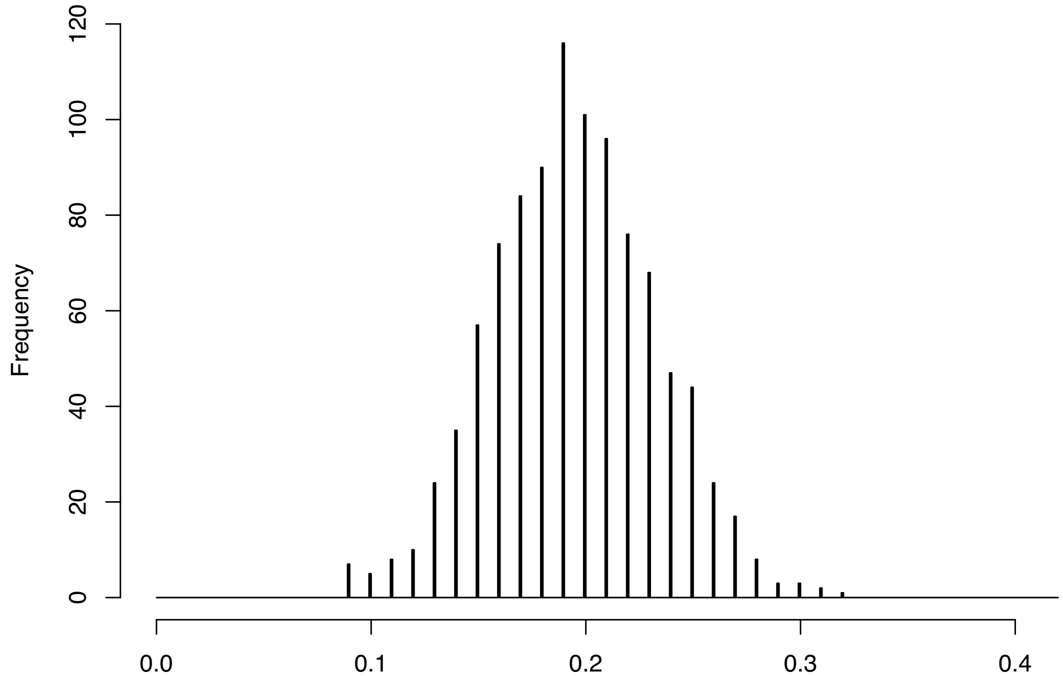
Wie zuverlässig ist M als Schätzer für p?
Davon machen wir uns ein Bild, indem wir viele (z.B. 1000) “unabhängige Kopien” von M erzeugten und in einem *Histogramm* darstellen, wie oft welche Ausgänge realisert wurden.

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:

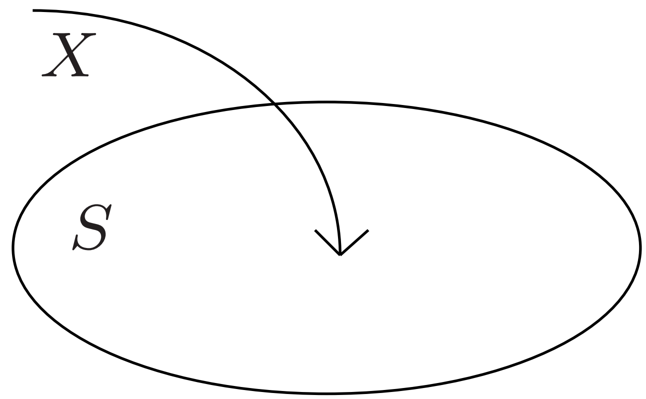
die Verteilung von M “ist um p konzentriert”.

Zusammenfassung

der wichtigsten Begriffe

der ersten Stunde:

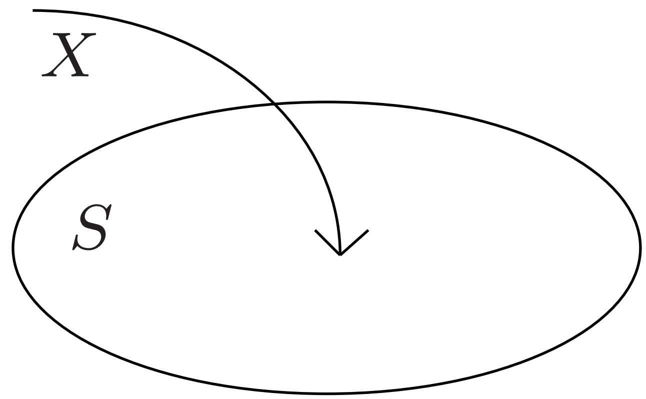
Ein Logo der Elementaren Stochastik:



X ... zufällige Wahl eines Elements aus S

S ... Menge von möglichen Ausgängen

Ein Logo der Elementaren Stochastik:

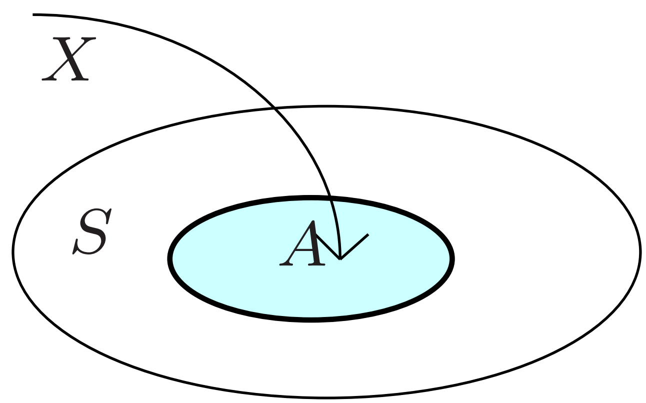


X . . . **Zufallsvariable**

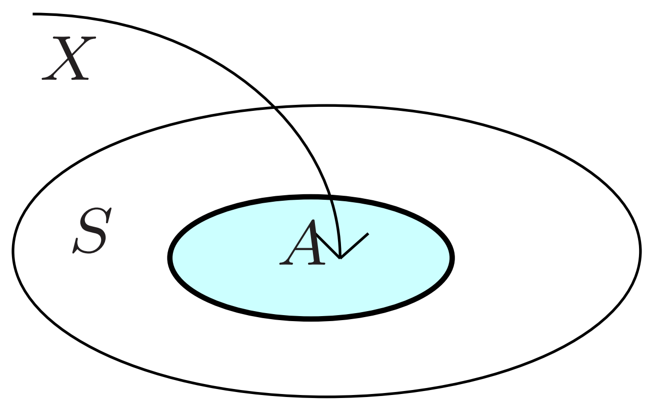
mit *Zielbereich (Wertebereich)* S

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:

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

“X ist *rein zufällig*”

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

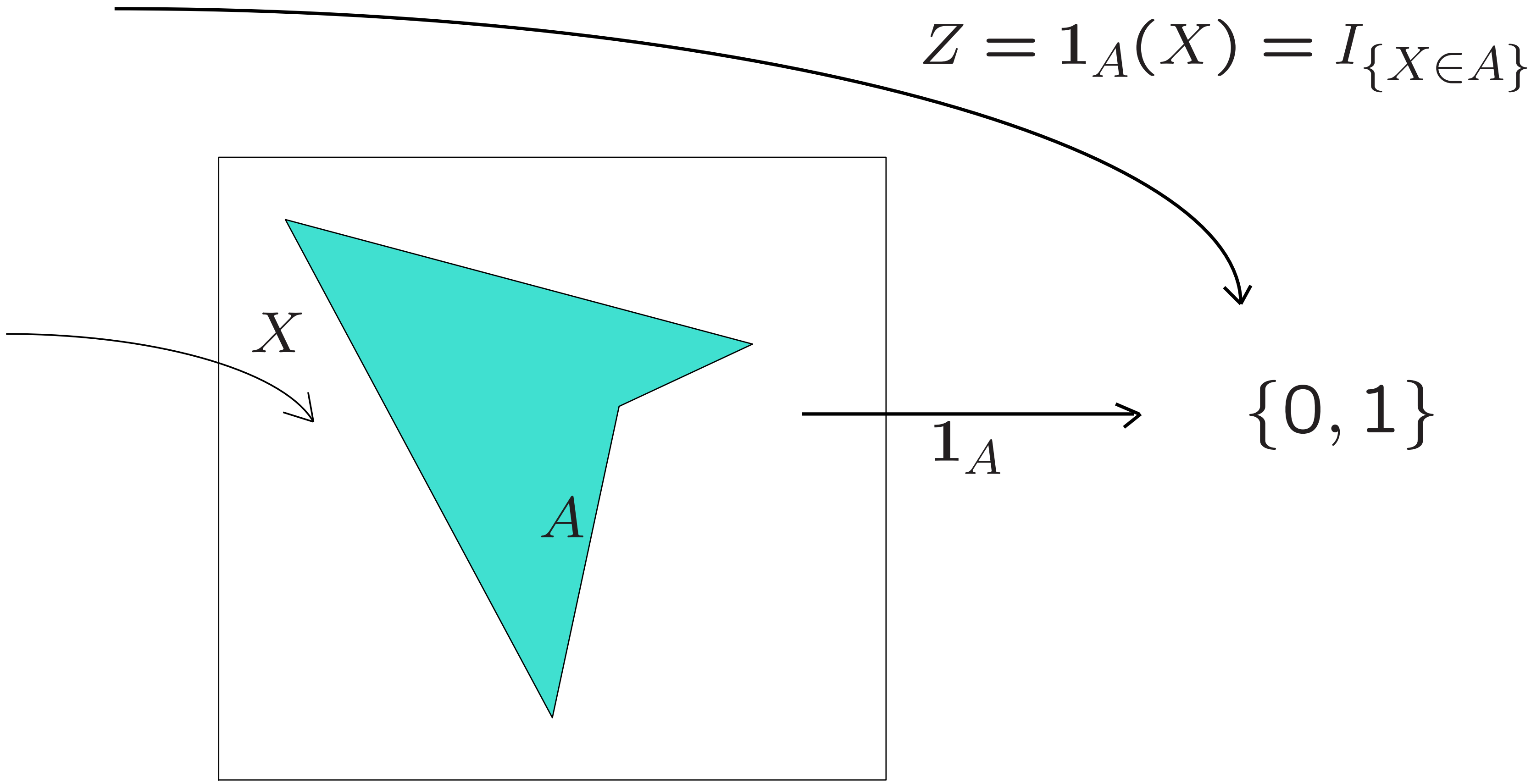
jede (messbare) Teilmenge A von S

kommt mit einer W’kt zum Zug,

die dem relativen Anteil ihres Volumens

am Volumen von S entspricht:

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



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