

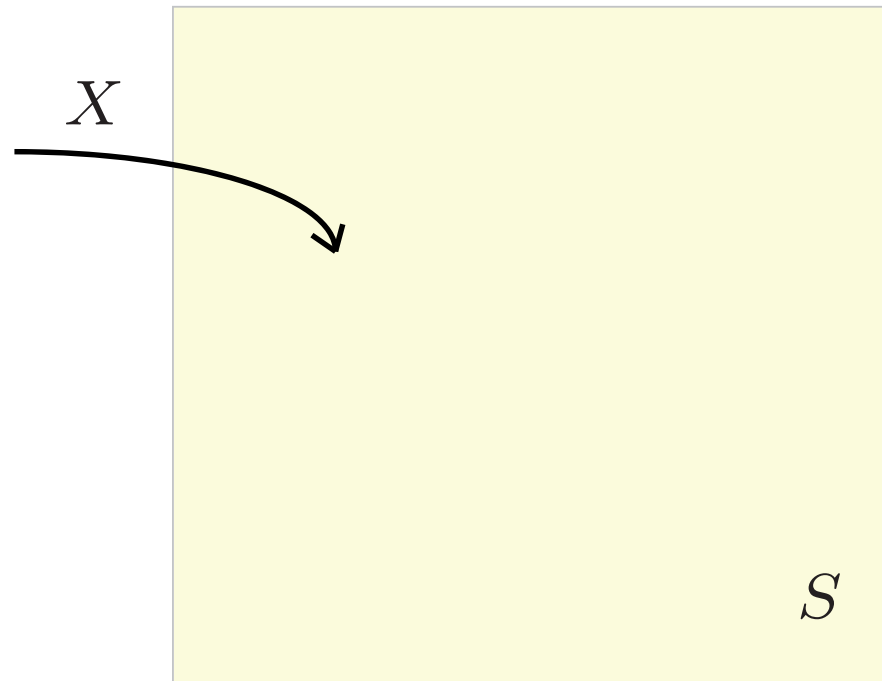
Vorlesung 1a

Zufallsvariable und Ereignisse, Wahrscheinlichkeiten und Verteilungen

Ein erster Blick

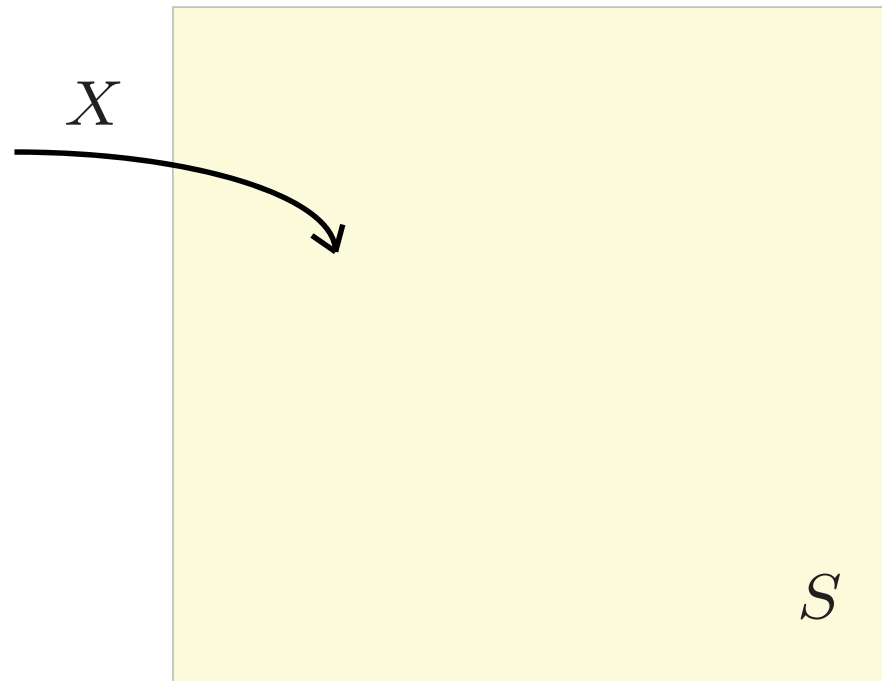
1. Rein zufällige Wahl aus einer endlichen Menge

Prototypisches Beispiel einer **Zufallsvariablen**:

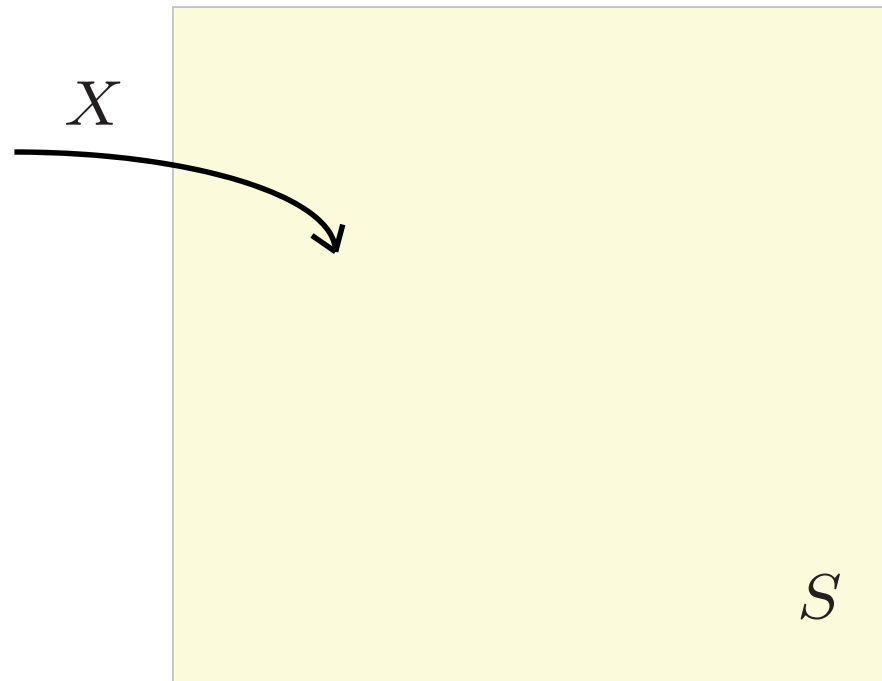


rein zufällige Wahl eines Punktes aus einer Fläche
(z. B. aus dem beigeen Quadrat S)

Prototypisches Beispiel einer **Zufallsvariablen**:

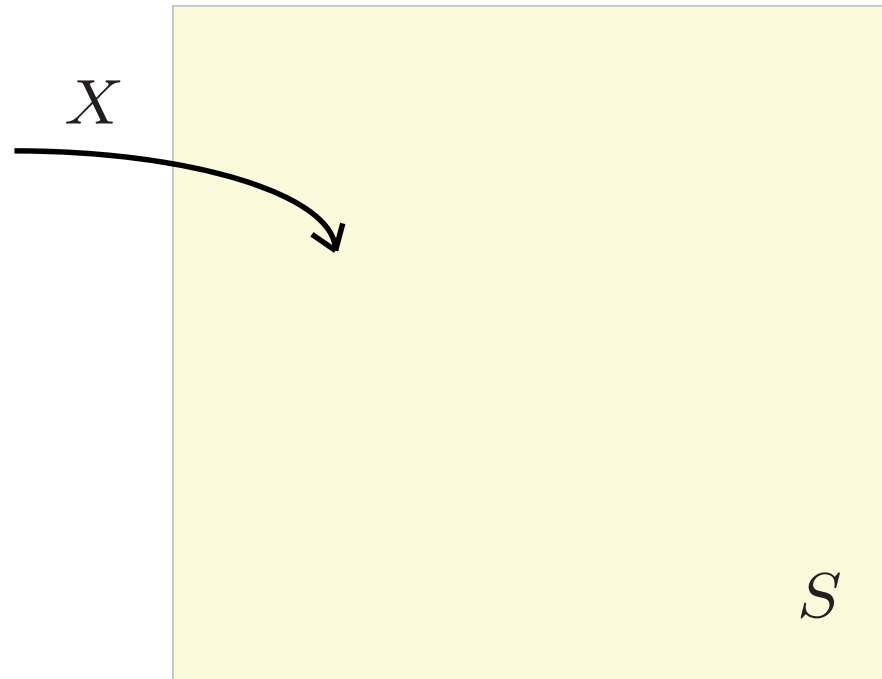


Was heißt “rein zufällige Wahl” eines Punktes aus S ?

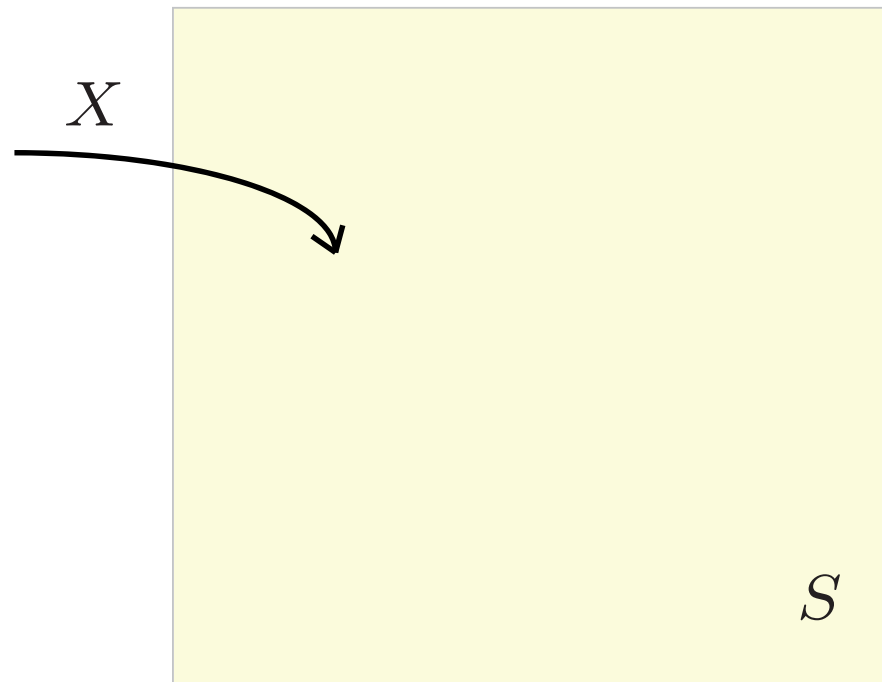


Fürs Erste sind endlich viele Pixel leichter vorstellbar
als ein Kontinuum aus unendlich vielen Punkten.

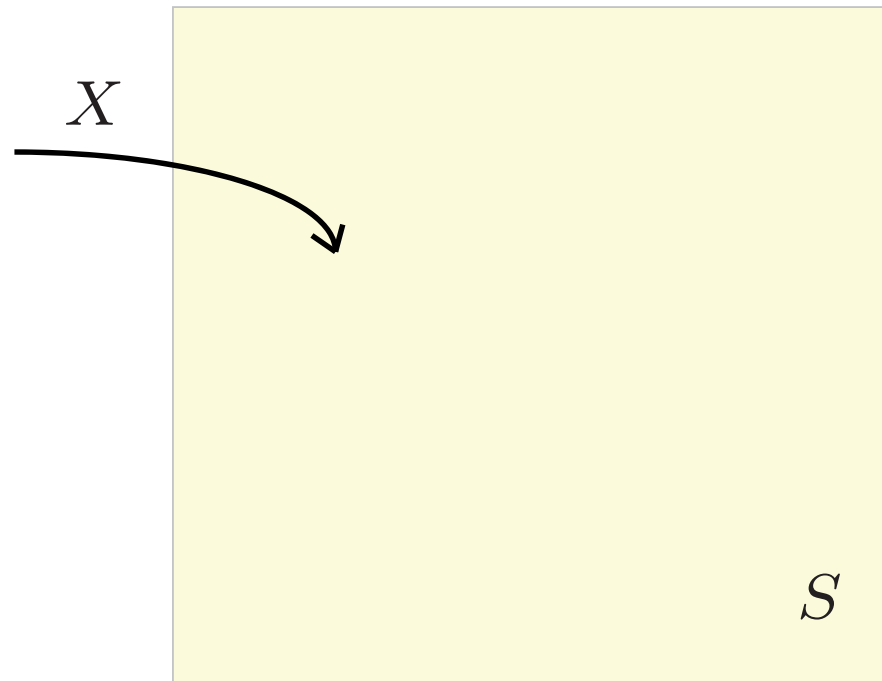
Stellen wir uns vor, S besteht aus 1000×1000 Pixeln



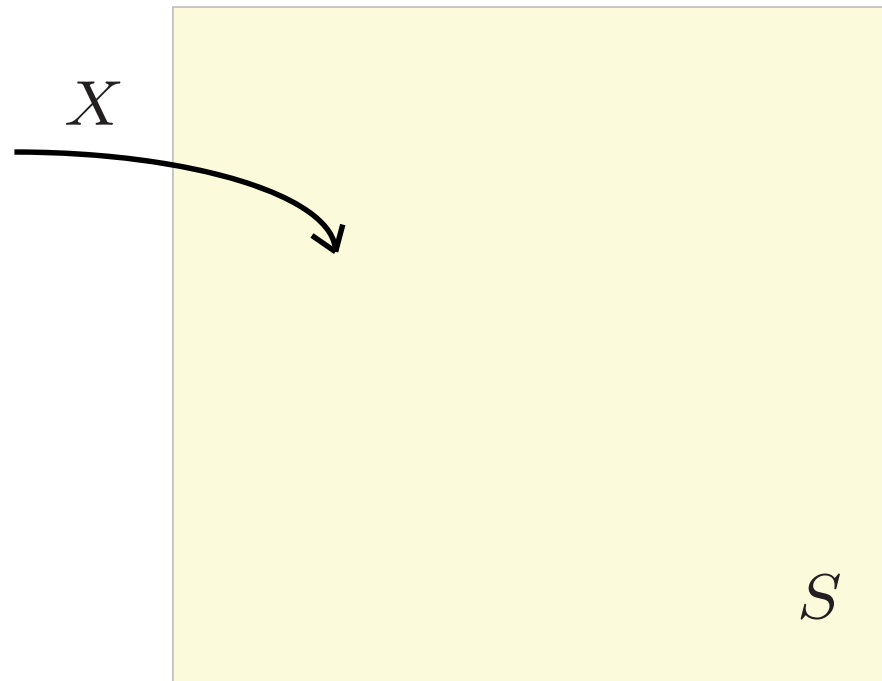
Stellen wir uns vor, S besteht aus g Pixeln, mit $g \in \mathbb{N}$



“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** .

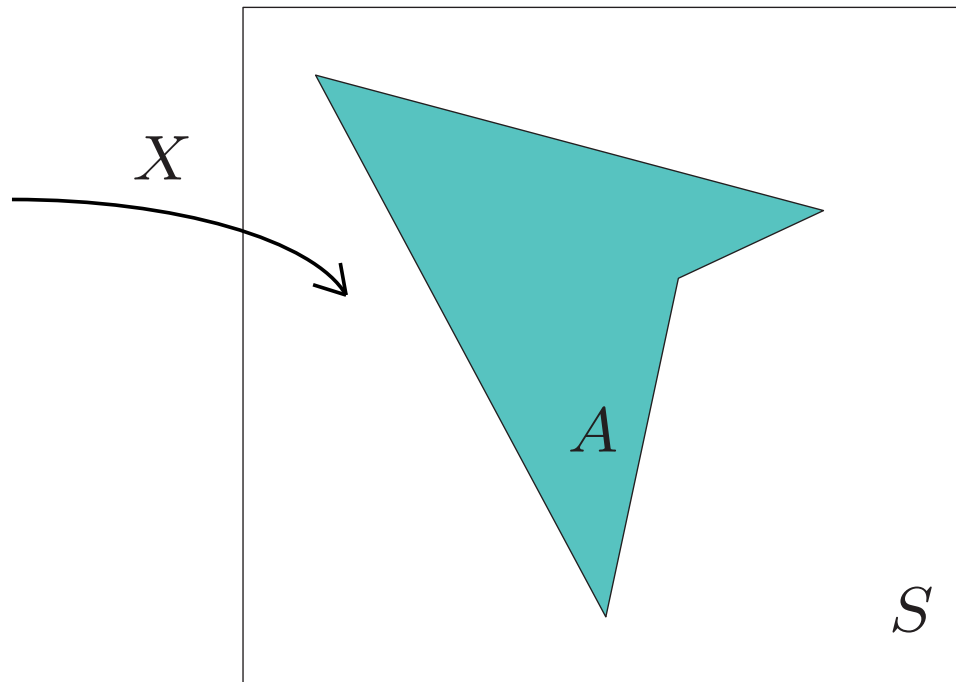
Wir halten fest:

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.

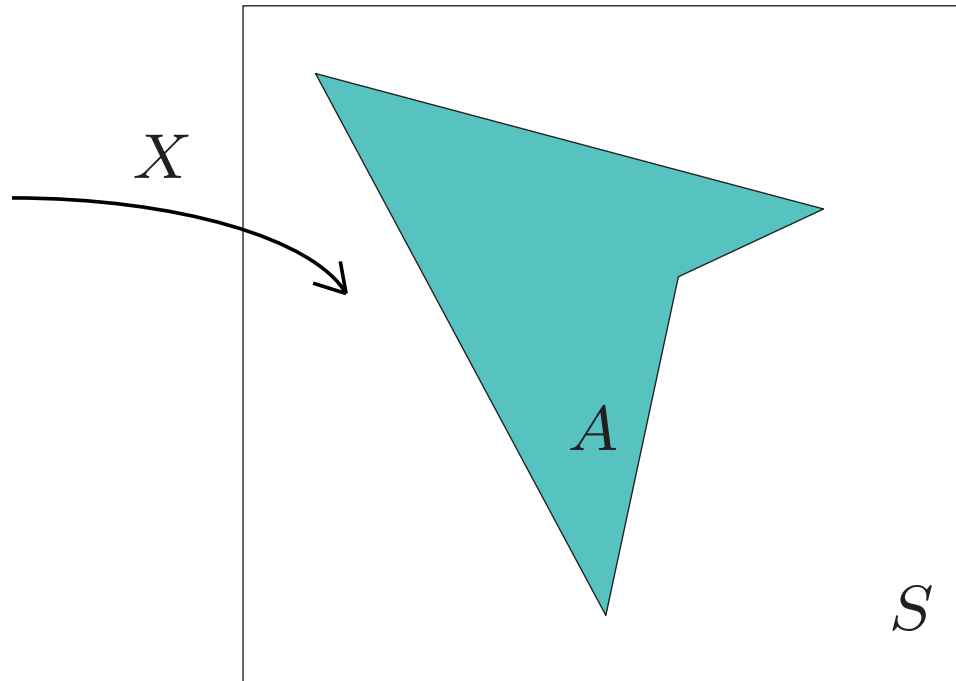
2. Ereignisse

Betrachten wir wieder unser Quadrat S



diesmal zusammen mit einer bestimmten **Teilmenge** A von S

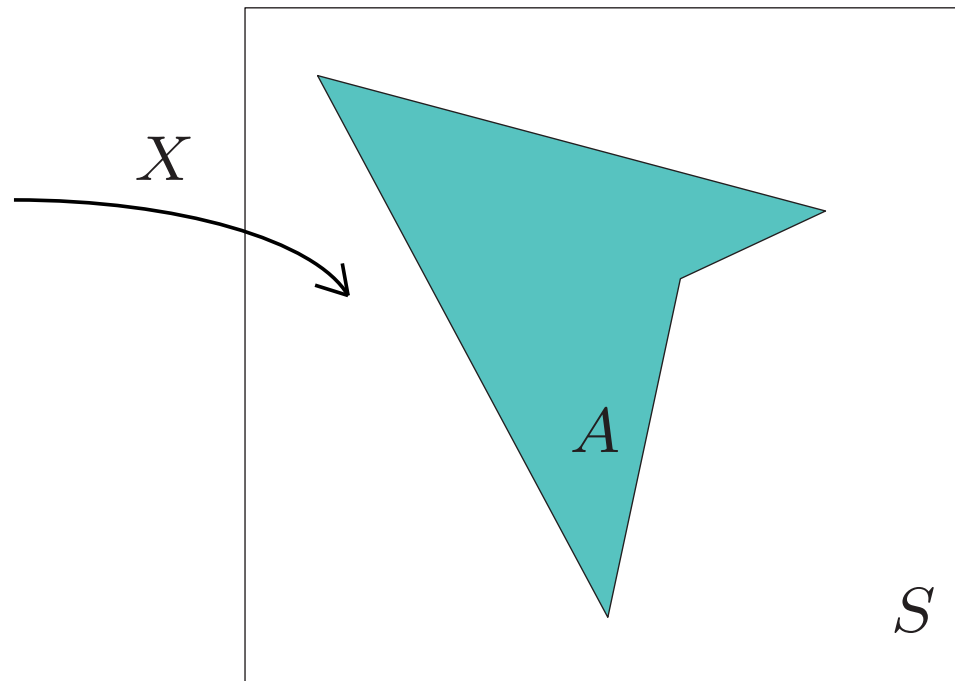
Bei der Wahl eines Pixels aus S



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 $\{X \in A\}$, $A \subset S$,

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.

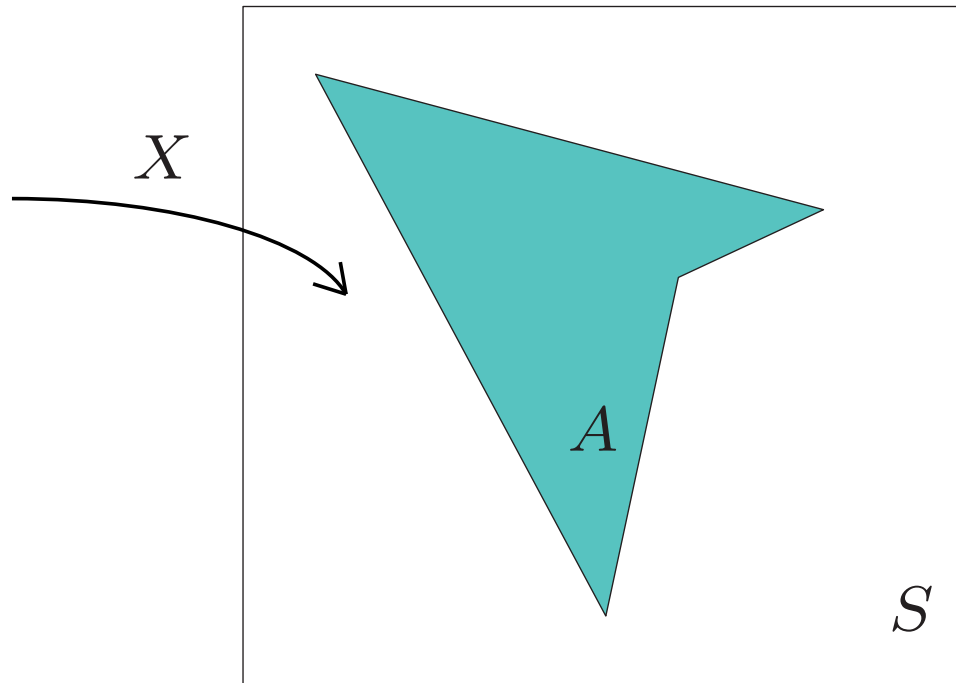
Für die Wahrscheinlichkeit des Ereignisses $\{X \in A\}$
schreiben wir

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

oder kurz

$$P(X \in A).$$

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

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

Lies und merke:

die Wahrscheinlichkeit, dass X in A fällt

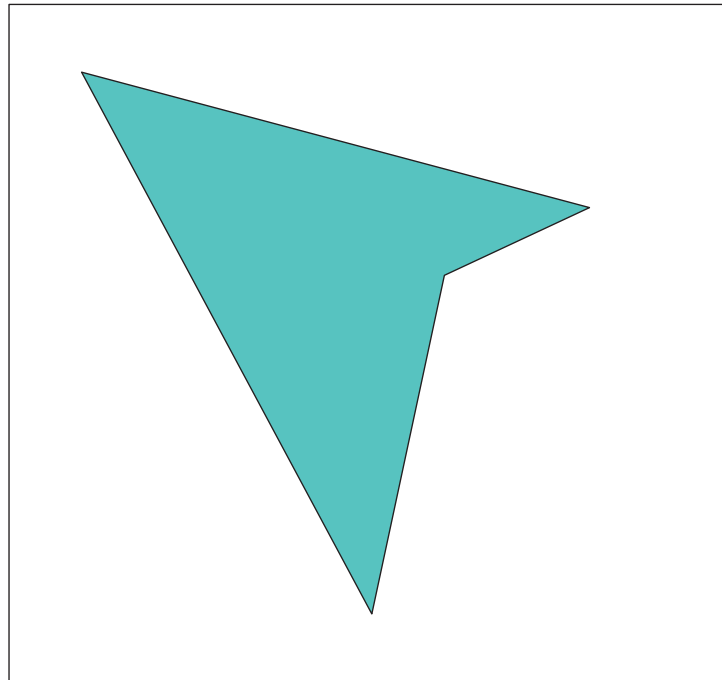
ist der Anteil der Menge A an der Menge S .

4. Schätzung eines Flächenanteils

Eine Anwendung der “rein zufälligen Wahl”:

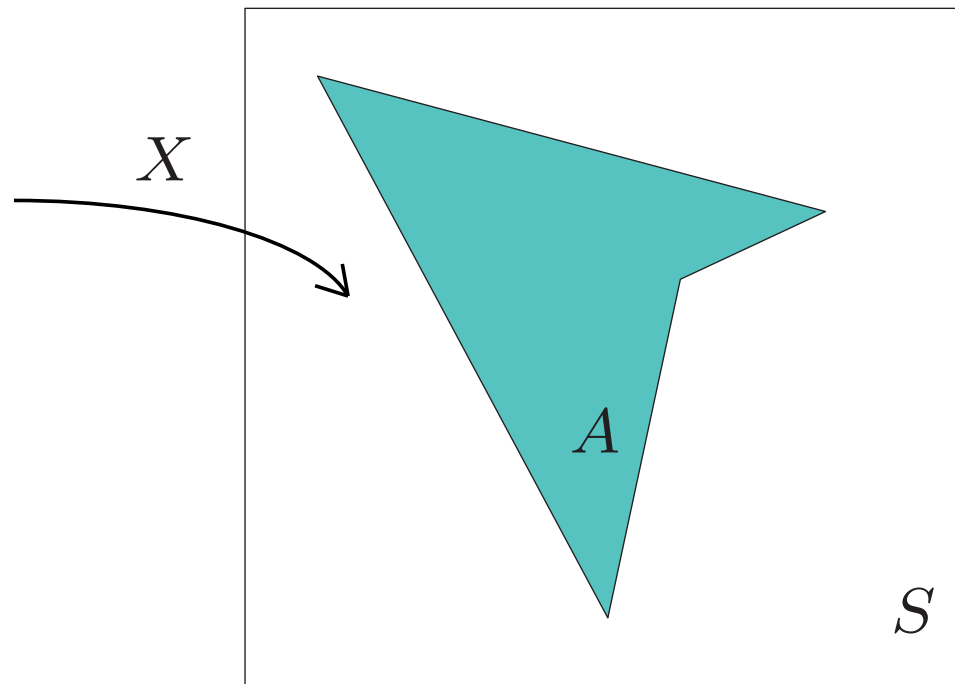
Monte-Carlo Schätzung eines Flächenanteils.

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, \dots)

von Pixeln in S .

Sei $A \subset S$.

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

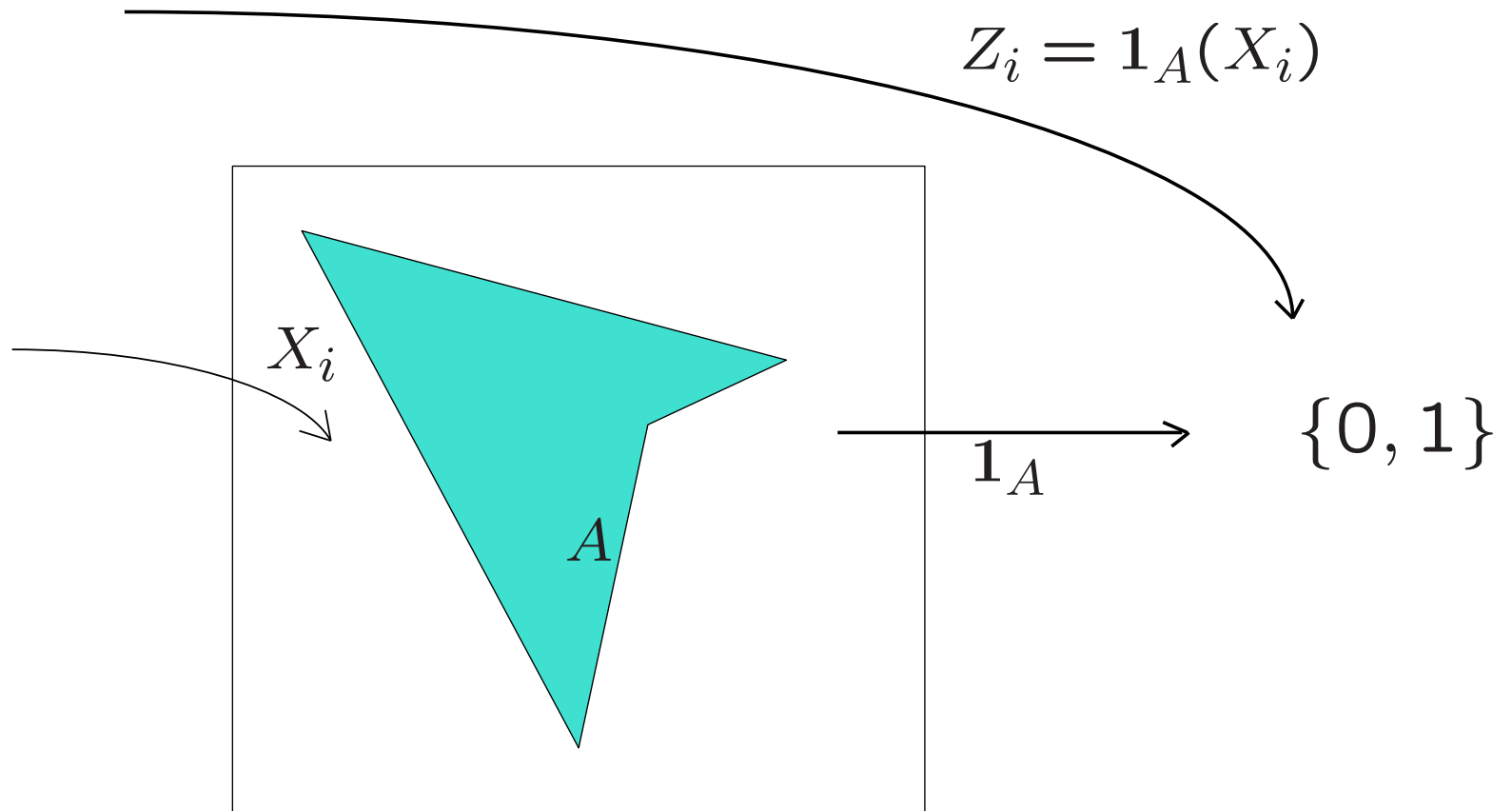
zählt, ob X_i in A fällt.

Dabei ist

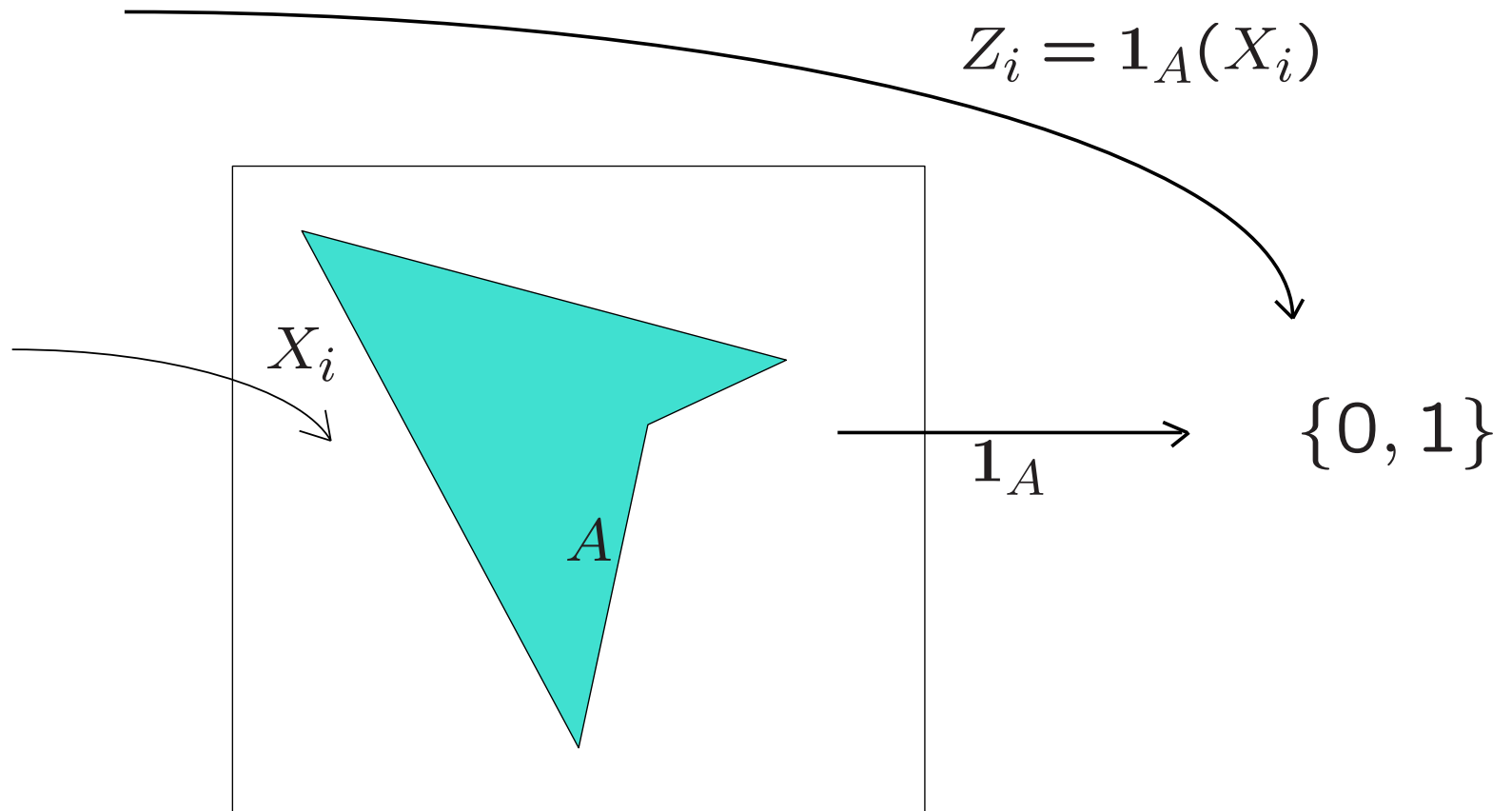
$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{für } x \in A \\ 0 & \text{für } x \in S \setminus A \end{cases}$$

die Indikatorfunktion der Menge A .

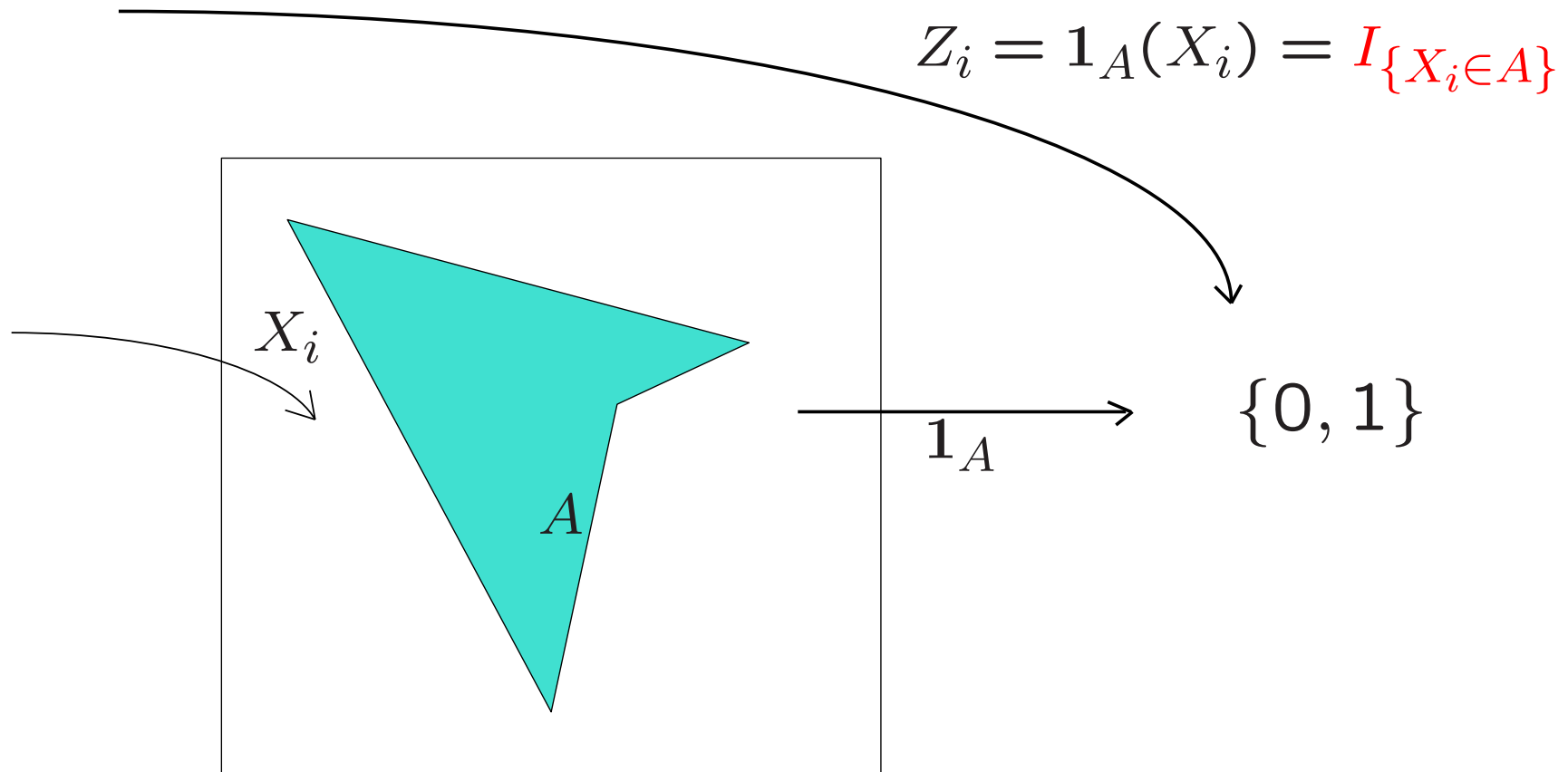
Mit dieser Abbildung von S nach $\{0, 1\}$
wird die S -wertige Zufallsvariable X_i
zur $\{0, 1\}$ -wertigen Zufallsvariablen Z_i “verarbeitet”.



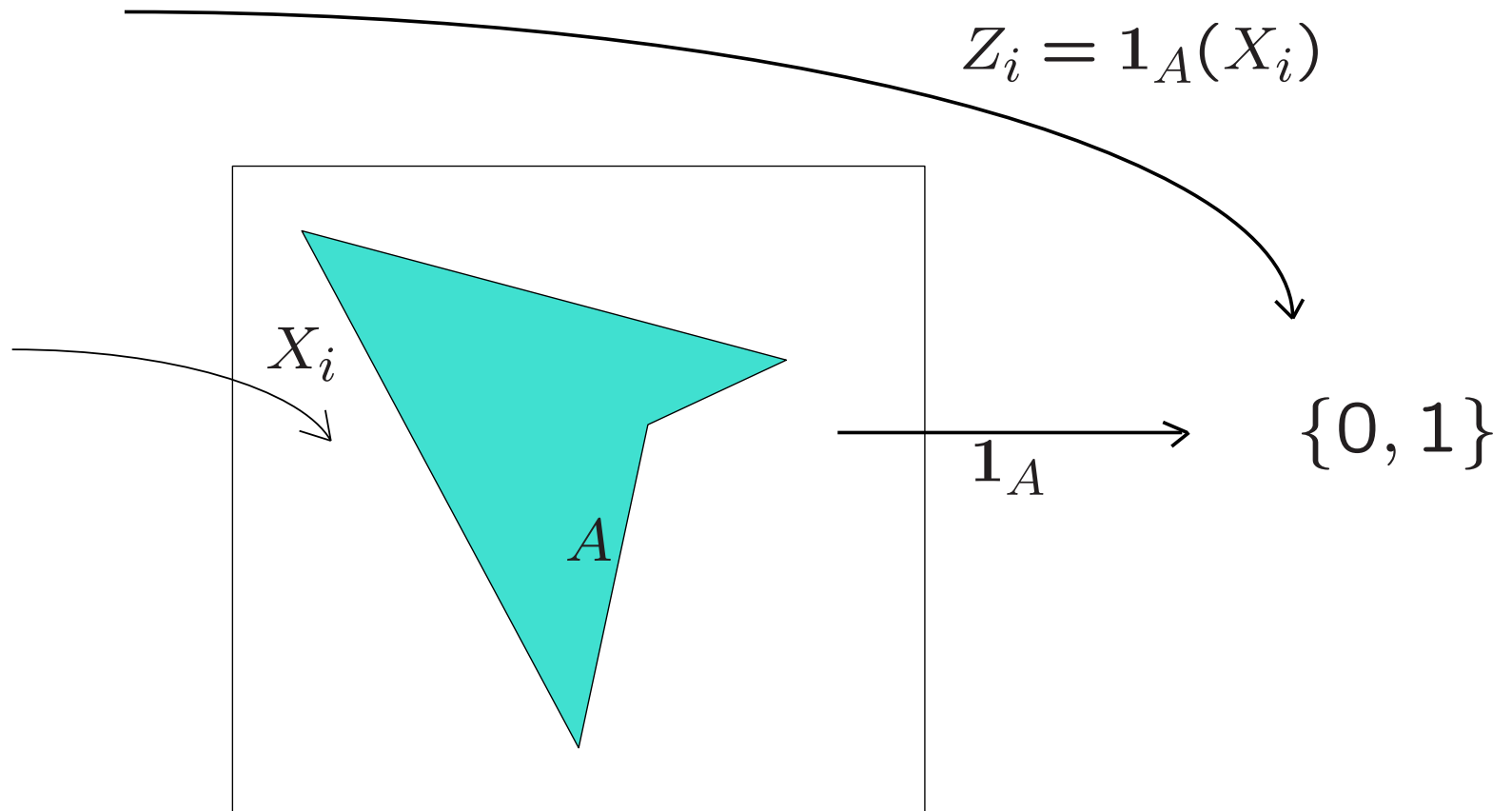
$Z_i = \mathbf{1}_A(X_i)$ ist der *der Indikator* (“der Zähler”) des Ereignisses “ X_i fällt in A ”.



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



Eine alternative Schreibweise für $1_A(X_i)$ 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.

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

Die zufällige Zahl

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

(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, \dots, X_{100})
liefert eine Realisierung von (Z_1, \dots, 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

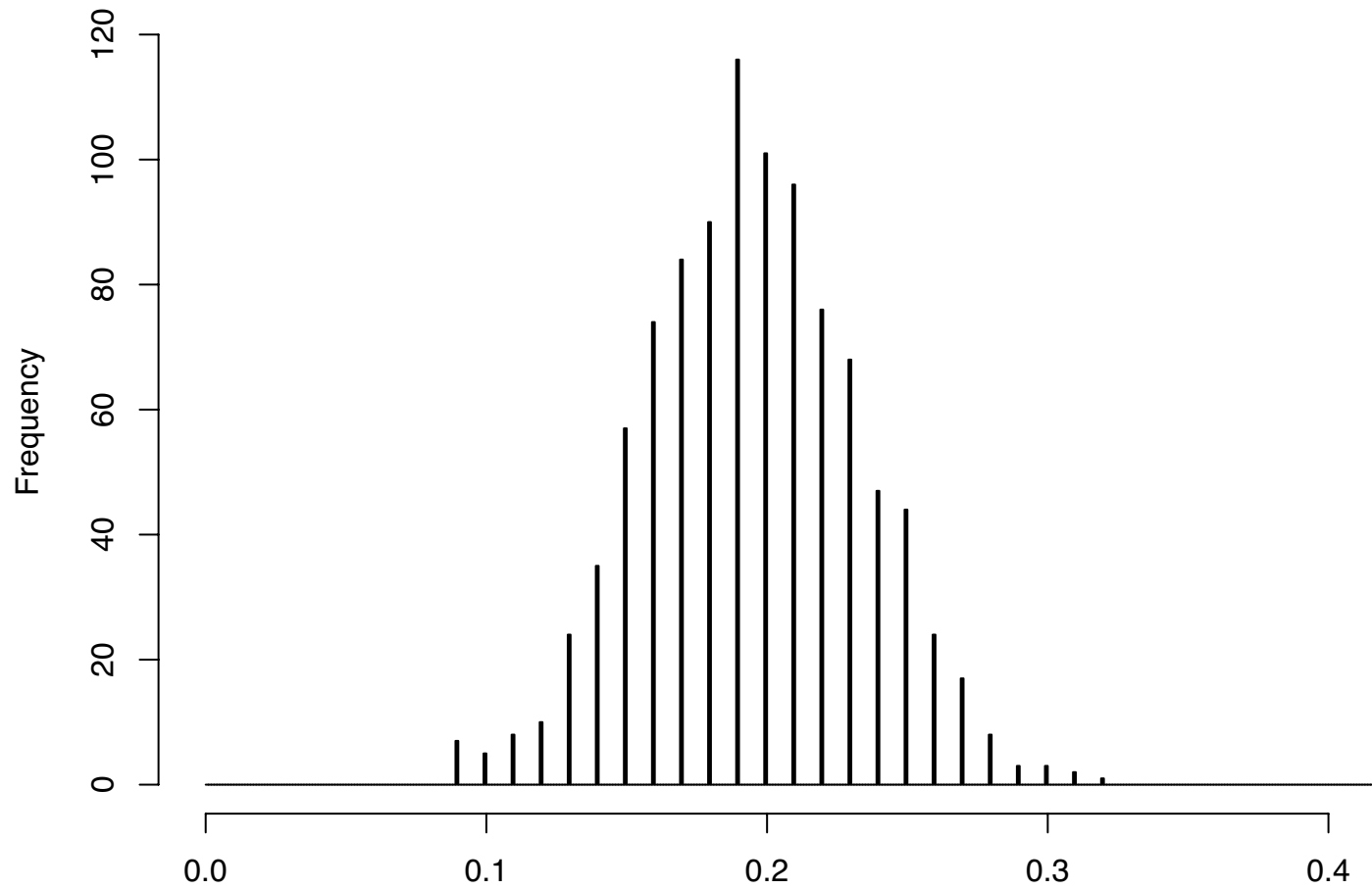
(d.h. die Menge der möglichen Ausgänge) ist

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

Wie zuverlässig ist M als Schätzer für p ?

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

Verteilung von M (Näherung aus 1000 Wiederholungen):



Wie zuverlässig ist M als Schätzer für p ?

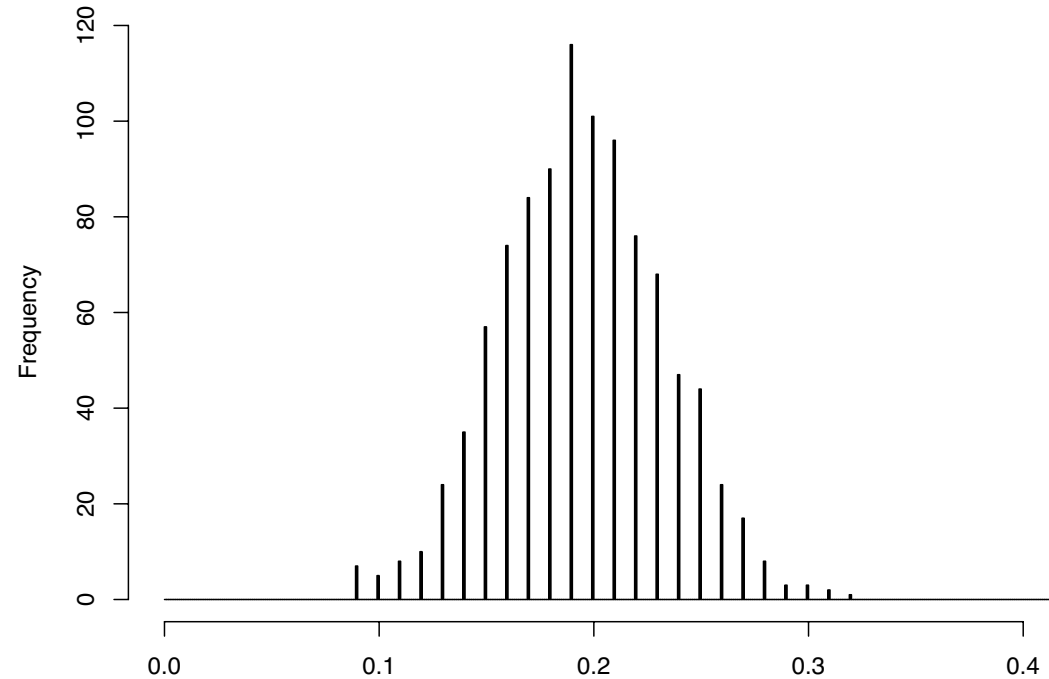
Davon machen wir uns ein Bild, indem wir viele (z.B. 1000) “unabhängige Kopien” von M erzeugen und in einem *Histogramm* darstellen, wie oft welche Ausgänge realisiert wurden.

So bekommen wir eine näherungsweise Darstellung der **Verteilung** von M .

Die **Verteilung** von M ist bestimmt durch ihre **Gewichte**

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

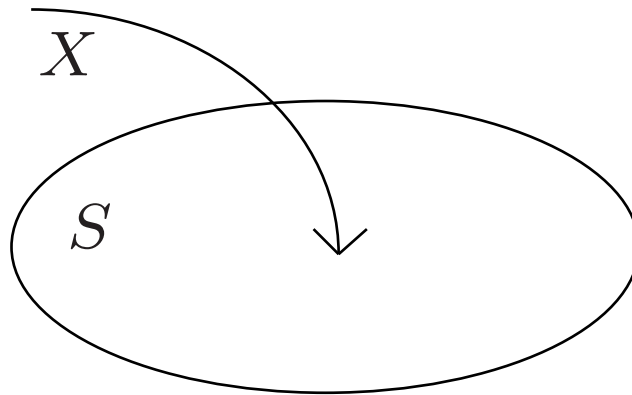
Verteilung von M (Näherung aus 1000 Wiederholungen):



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

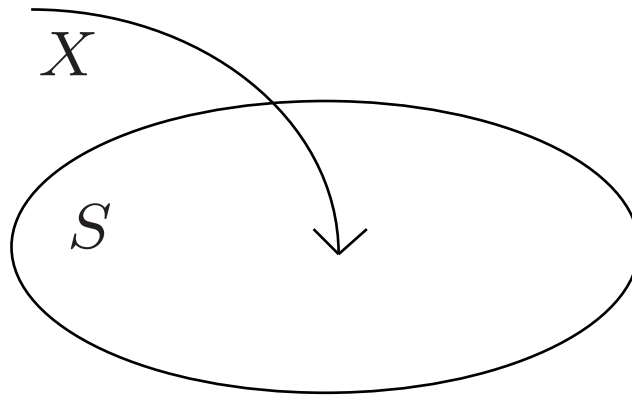
Ein Logo der Elementaren Stochastik:



X ... zufällige Wahl eines Elements aus S

S ... Menge von möglichen Ausgängen

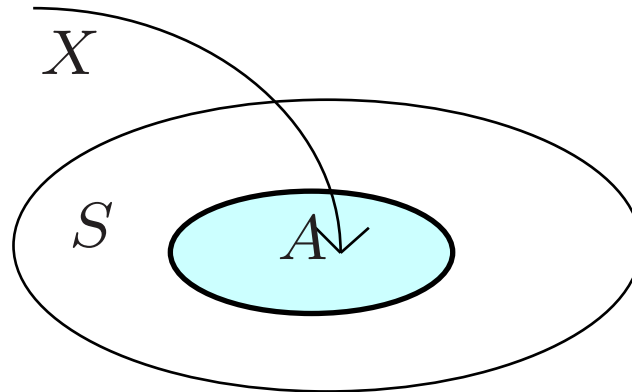
Ein Logo der Elementaren Stochastik:



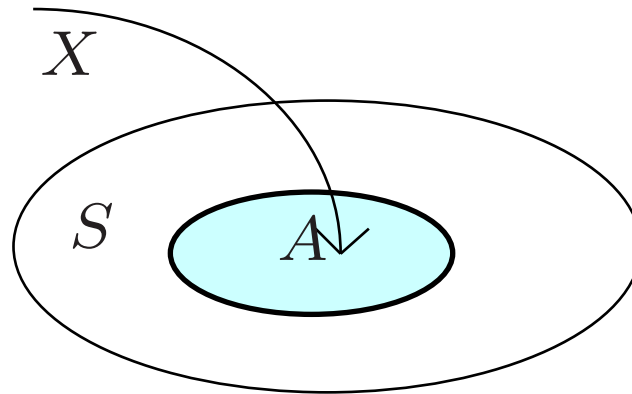
X ... **Zufallsvariable**

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 :

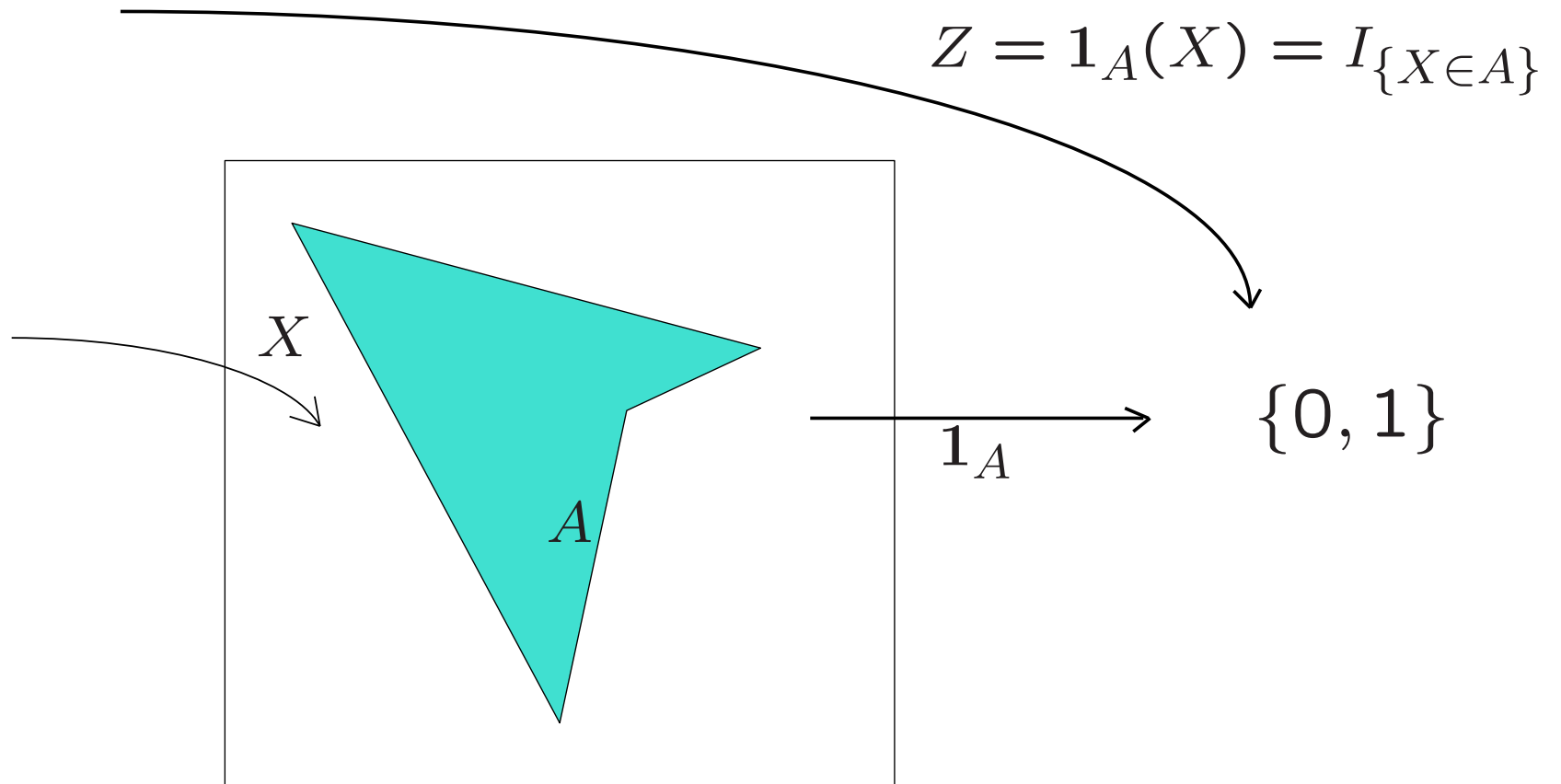
alle Elemente von S haben die gleiche *W*'keit
gewählt zu werden.

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

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

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

$$P(X \in A).$$



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:

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

