

# Stochastik für die Informatik

Prof. Dr. Noemi Kurt

FB 12, Institut für Mathematik, Goethe-Universität Frankfurt

Wintersemester 2022/23

# Organisatorisches: Vorlesung

## Dozentin

- ▶ Prof. Dr. Noemi Kurt
- ▶ E-Mail: [kurt@math.uni-frankfurt.de](mailto:kurt@math.uni-frankfurt.de)

## Vorlesungsmaterial

- ▶ **Buch**: Link zum E-Book auf der Moodle-Seite
- ▶ **Folien**: Jeweils kurz vor der Vorlesung auf der Moodle-Seite
- ▶ **Übungszettel**: Wöchentlich auf der Moodle-Seite

# Organisatorisches: Tutorien

**Klausurtermine** 14. Februar und 31. März (Änderungen vorbehalten). Zeit und Raum werden rechtzeitig mitgeteilt.

## Dauer, Voraussetzungen und Hilfsmittel

- ▶ Dauer: 90 Minuten
- ▶ Zulassungsvoraussetzungen: 72 der 160 Punkte aus den 10 Hausaufgabenblättern.
- ▶ **Hilfsmittel:** Ein (beidseitig) handschriftlich beschriebenes A4-Blatt, sowie ein nichtprogrammierbarer Taschenrechner. Smartphones, Tablets, etc. dürfen nicht als Taschenrechnerersatz genutzt werden, internetfähige Geräte müssen während der Klausur ausgeschaltet bleiben.

# Literaturvorschläge

## Lehrbücher

- ▶ L. Dümbgen: Stochastik für Informatiker (Springer)
- ▶ N. Henze: Stochastik für Einsteiger (Vieweg)
- ▶ G. Kersting und A. Wakolbinger: Elementare Stochastik (Birkhäuser)
- ▶ M. Mitzenmacher und E. Upfal: Probability and Computing (Cambridge University Press)
- ▶ R. Motwani und P. Raghavan: Randomized Algorithms (Cambridge University Press)
- ▶ H.-O. Georgii: Stochastik (De Gruyter)
- ▶ L. Papula: Mathematik für Ingenieure und Naturwissenschaftler, Band 3, (Vieweg)

## Software

- ▶ R Software-Paket mit vielen Statistik-Anwendungen.  
<http://www.r-project.org>, Download z.B. hier  
<http://ftp5.gwdg.de/pub/misc/cran/>

# Vorlesung 1

## Inhalt

- ▶ Was ist Stochastik?
- ▶ Wo wird Stochastik angewendet?
- ▶ Grundkonzepte der mathematischen Stochastik

## Lernziele

- ▶ Wissen was der Begriff Stochastik bedeutet
- ▶ Anwendungsgebiete der Stochastik kennen
- ▶ Erste Grundkonzepte der Wahrscheinlichkeitstheorie kennen

## Benötigte Vorkenntnisse

- ▶ Begriffe und Notation für Mengen und Funktionen, z.B. aus AnNuMa, Vorkurs Mathematik

# Was ist Stochastik?

Stochastik =  
Lehre von der Gesetzmäßigkeit des Zufalls

Später dazu mehr.

Warum könnte mich das interessieren?

- um einen Zufallszahlengenerator programmieren zu können

```
int getRandomNumber()  
{  
    return 4; // chosen by fair dice roll.  
              // guaranteed to be random.  
}
```

Bild: [www.xkcd.com](http://www.xkcd.com)

- um PageRank (Google-Algorithmus) zu verstehen

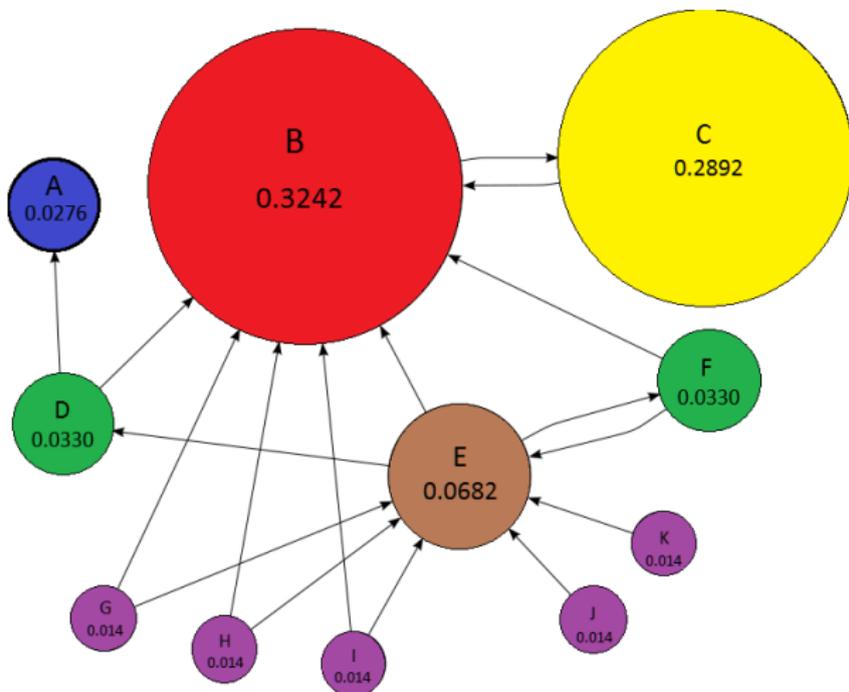


Bild: Wikipedia, „PageRank-Beispiel“ von Zetkin. Lizenziert unter CC BY-SA 3.0 über Wikimedia Commons

- Um Telekommunikationsnetzwerke zu optimieren

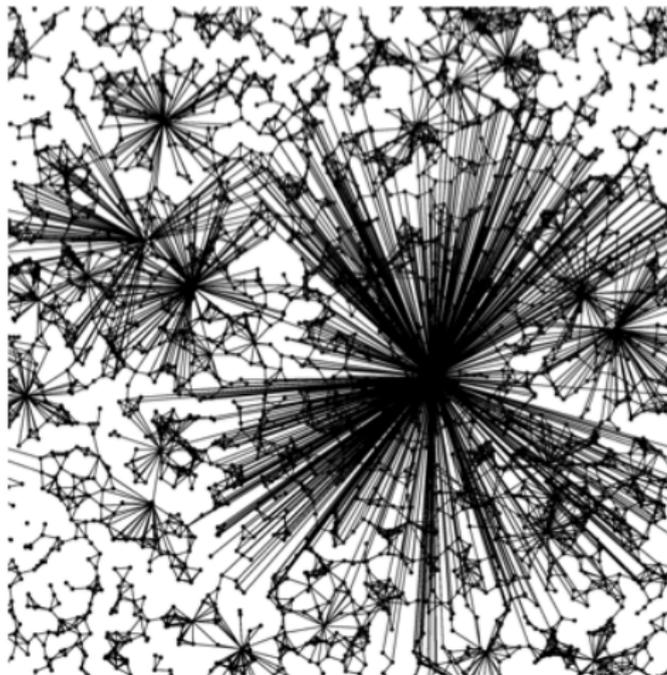


Bild: Christian Hirsch, U. Mannheim

- Um Wartesysteme zu modellieren und zu analysieren

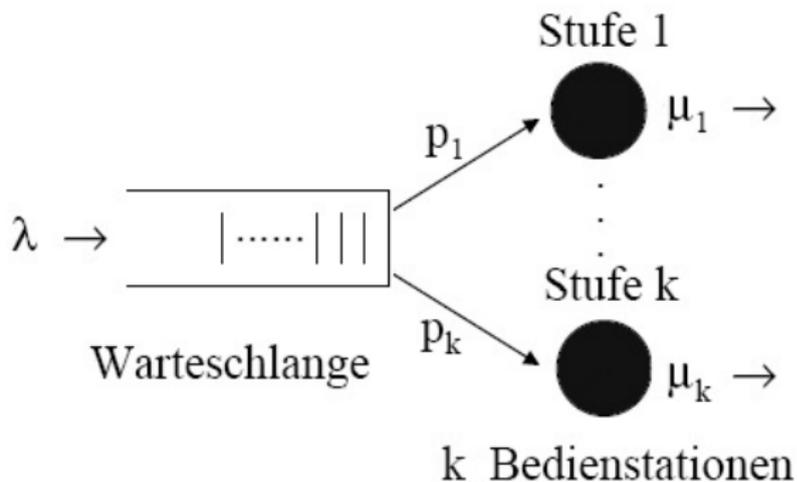


Bild: Wikipedia, frei verfügbar

## - Um Wahlergebnisse interpretieren zu können

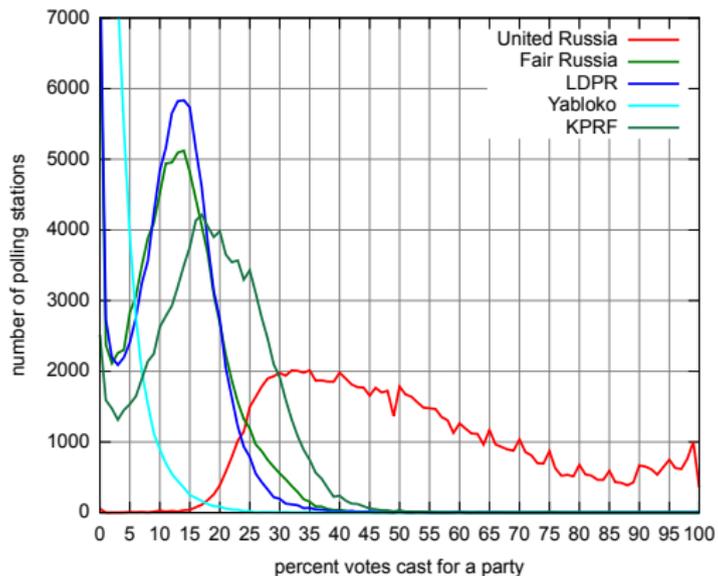


Bild: Wikipedia, „Russian legislative election 2011“ von Gritzko. Lizenziert unter CC BY-SA 3.0 über Wikimedia Commons

**Ziel dieser Veranstaltung:** Beherrschen der mathematischen Grundlagen, welche diesen und anderen Anwendungen zugrunde liegen.

**Inhalt und Aufbau der Vorlesung:**

- ▶ Mathematische Grundkonzepte: Wahrscheinlichkeiten, Zufallsvariablen
- ▶ Kenngrößen von Zufallsvariablen und ihre Berechnung
- ▶ Wichtige Klassen von (mathematischen) Beispielen und ihr Auftreten
- ▶ Modellierung zufällig ablaufender Prozesse
- ▶ Grundlagen und Methoden der Statistik
- ▶ Markov-Ketten, Stochastische Algorithmen

Stochastik = Lehre von der Gesetzmäßigkeit des Zufalls.

Stochastik = Überbegriff für:

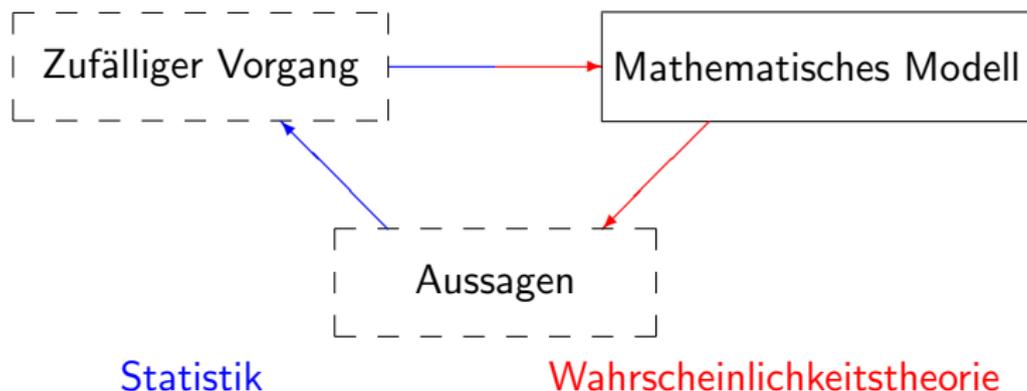
- ▶ **Wahrscheinlichkeitstheorie:** Mathematische Modellierung und Formalisierung, theoretische Grundlage
- ▶ **Statistik:** Analyse und Interpretation von Messwerten und Daten aus zufälligen Vorgängen mittels wahrscheinlichkeitstheoretischen Methoden

**Zufall:** “Echter” Zufall, oder Mangel an Information, oder Effekt vieler kleiner Ereignisse, welche einzeln nicht genau kontrolliert werden können, und deshalb als zufällig angenommen werden.

Beispiele für zufällig ablaufende Prozesse:

- ▶ Würfeln, Lottoziehung
- ▶ Warteschlangenbildung, z.B. in der Verkehrsmodellierung
- ▶ Bildung von Netzwerken
- ▶ Stochastische Algorithmen, z.B. Googles PageRank

# Stochastik



## Wahrscheinlichkeitstheorie:

Idealisierung und Formalisierung  $\Rightarrow$  **Universelle Anwendbarkeit**

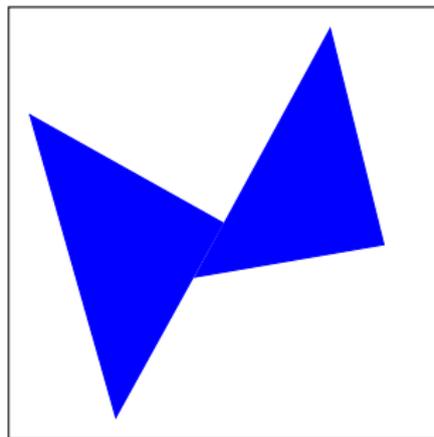
Eine Theorie, viele Anwendungsbereiche

## Statistik:

Viele Methoden, Auswahl einsprechend der konkreten Anwendung

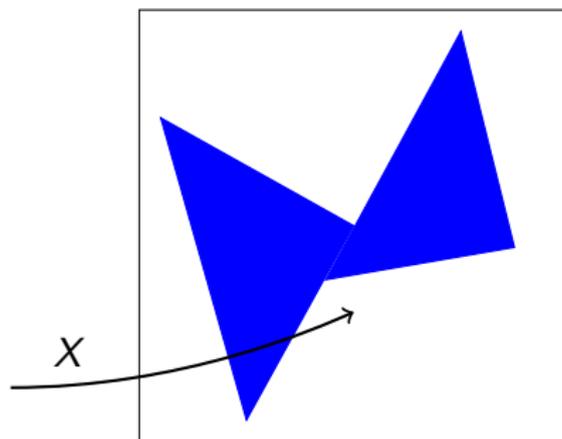
# Grundbegriffe der Wahrscheinlichkeitstheorie

Ein erstes Beispiel



Wie groß ist der Anteil der blauen Fläche an der Fläche des Quadrats?

## Ein erstes Beispiel



Wie wahrscheinlich ist es, dass eine **rein zufällige Auswahl** eines Punktes aus dem Quadrat auf die blaue Fläche führt?

(Angenommen wir haben ein Werkzeug mit dem wir **rein zufällig** einen Punkt auswählen können, z.B. einen Dartspeil)

## Ein erstes Beispiel

Sei  $Q$  das Quadrat. “ $X$  ist rein zufälliger Punkt von  $Q$ ” bedeutet: für jede “messbare” Teilmenge  $A \subseteq Q$  gilt

$$\mathbb{P}(X \in A) = \frac{\text{Fläche von } A}{\text{Fläche von } Q}$$

in Worten: Die **Wahrscheinlichkeit** dass ein zufälliger Punkt  $X$  in der Menge  $A$  liegt, ist gegeben durch Anteil der Fläche von  $A$  an der Gesamtfläche.

$\mathbb{P}$  steht für “probabilitas” oder “probability”, Wahrscheinlichkeit.

Die Notation  $\mathbb{P}(X \in A)$  wird bald noch einmal formal eingeführt werden.

## Ein erstes Beispiel

Angenommen wir können einen rein zufälligen Punkt  $X$  auswählen, und zwar nicht nur einmal, sondern beliebig oft:

Wir erhalten also eine zufällige Folge  $X_1, X_2, X_3, \dots$  von Punkten aus  $Q$ .

Für vorgegebenes  $A \subseteq Q$  betrachten wir eine daraus abgeleitete zufällige Folge  $Z_1, Z_2, \dots$  wobei

$$Z_i = \begin{cases} 1 & \text{falls } X_i \in A, \\ 0 & \text{falls } X_i \notin A. \end{cases}$$

Die zufällige Zahl

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

ist ein **Schätzer** für die gesuchte Wahrscheinlichkeit, d.h.

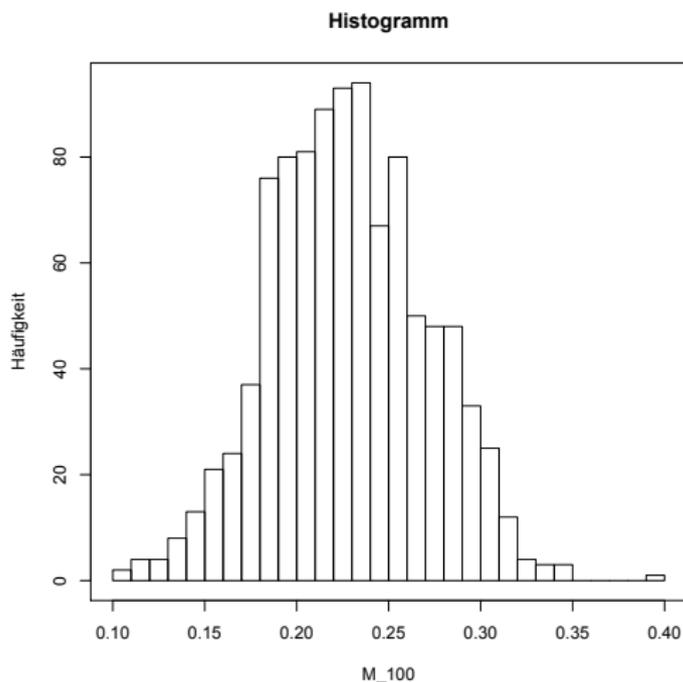
$$\mathbb{P}(X \in A) \approx M_{100}.$$

Welche Werte kann die zufällige Zahl  $M_{100}$  annehmen?

$$M_{100} \in \left\{0, \frac{1}{100}, \frac{2}{100}, \dots, \frac{99}{100}, 1\right\}$$

Jedoch ist nicht jeder dieser Werte gleich wahrscheinlich.

Wir erzeugen (am Rechner) 1000 “unabhängige Kopien” von  $M_{100}$ , und zeichnen das entsprechende **Histogramm**:



Der echte Wert  $\mathbb{P}(X \in A)$  liegt dort, wo die **Verteilung** von  $M_{100}$  **konzentriert** ist.

Im Laufe dieser Vorlesung werden die in diesem Beispiel verwendeten Begriffe, wie **Wahrscheinlichkeit**,  $\mathbb{P}(X \in A)$ , **Verteilung**, **zufällige Folge**, **Schätzer...** präzise definiert und angewandt werden. Sie gehören zu den Grundbegriffen der Wahrscheinlichkeitstheorie und Statistik.

# Teil 1: Grundlagen der Wahrscheinlichkeitstheorie

## 1. Zufallsexperimente und Wahrscheinlichkeitsräume

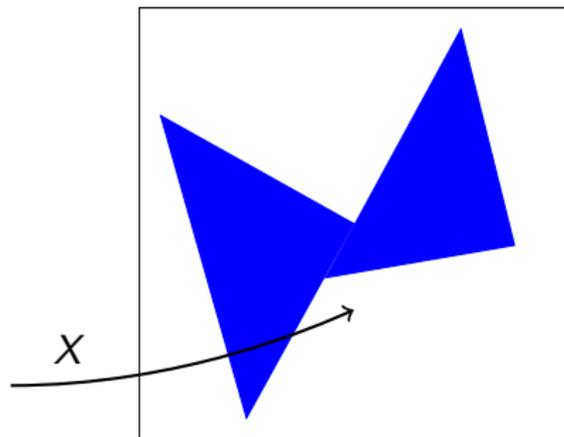
### 1.1. Wahrscheinlichkeiten und Ereignisse

(Def. 1.1) Ein **Zufallsexperiment** ist ein Vorgang mit einem zufälligen oder (a priori) unbestimmten Ausgang.

#### Beispiele

- ▶ Zufällige Auswahl eines Punktes aus dem Quadrat
- ▶ Münzwurf
- ▶ Lottoziehung
- ▶ Messung der Lebensdauer eines Bauteils (zufällig ausgewählt aus einer Menge gleichartiger Teile)
- ▶ Messung der Anzahl Zugriffe auf eine Website an einem zufällig ausgewählten Tag
- ▶ ...

# Zufallsexperiment: Cartoon



# Ergebnisse und Ereignisse

(Def. 1.2) Ein möglicher Ausgang eines Zufallsexperiments ist ein **Ergebnis**. Die Menge aller Ergebnisse ist die **Ergebnismenge**  $\Omega$ . Ein **Ereignis** ist eine Teilmenge  $A \subseteq \Omega$  (ggf. mit gewissen günstigen Eigenschaften).

- ▶ Notation:  $\omega \in \Omega$  ist ein *Ergebnis*
- ▶ Ergebnisse sind Elemente, Ereignisse Teilmengen der Ergebnismenge  $\Omega$ . Ein Ereignis besteht aus einem oder mehreren Ergebnissen.
- ▶ Im Beispiel mit dem Quadrat: Ein einzelner Punkt ist ein Ergebnis, eine Teilmenge ist ein Ereignis.

(Beispiel 1.1: Würfel: siehe Vorlesung)

(Bemerkung: Ereignisse und Aussagen, Tabelle)

# Endliche Wahrscheinlichkeitsräume

(Def. 1.3) Ein **endlicher Wahrscheinlichkeitsraum** ist ein Paar  $(\Omega, \mathbb{P})$ , bestehend aus

- ▶ Einer endlichen Ergebnismenge  $\Omega$
- ▶ Einer Funktion  $\mathbb{P}$  welche jedem Ereignis  $A \subseteq \Omega$  eine Wahrscheinlichkeit  $\mathbb{P}(A)$  zuordnet.

Dabei soll die Funktion  $\mathbb{P}$  folgende Eigenschaften haben:

(P0)  $0 \leq \mathbb{P}(A) \leq 1$  für alle  $A \subseteq \Omega$ .

(P1)  $\mathbb{P}(\Omega) = 1$

(P2) Für  $A, B \subseteq \Omega$  mit  $A \cap B = \emptyset$  gilt  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Eine solche Funktion  $\mathbb{P}$  heißt **Wahrscheinlichkeitsmaß** auf  $\Omega$ .

# Wahrscheinlichkeitsraum, Ereignisse

- ▶ Notation:  $\mathbb{P}(\omega) := \mathbb{P}(\{\omega\})$ , für  $\omega \in \Omega$
- ▶ Was hat diese Definition mit unserer intuitiven Vorstellung von Wahrscheinlichkeiten zu tun? (Skizze siehe VL)
- ▶ (Beispiel 1.3: Berechnung von Wahrscheinlichkeiten, siehe VL)

# Wahrscheinlichkeitsmaß

(Satz 1.1) Sei  $\mathbb{P}$  ein Wahrscheinlichkeitsmaß auf einer Ergebnismenge  $\Omega$ . Sei  $A \subseteq \Omega$  ein Ereignis. Dann gelten

$$(P3) \quad \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

$$(P4) \quad \mathbb{P}(\emptyset) = 0$$

(P5) Für  $A \subseteq B \subseteq \Omega$  gilt  $\mathbb{P}(A) \leq \mathbb{P}(B)$ .

(P6) Für  $A, B \subseteq \Omega$  gilt  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

(Herleitung und Interpretation siehe VL)