# Iudex non calculat (?)

## Kulturgut Mathematik und die Justiz

Oliver Ernst

Professur Numerische Mathematik, TU Chemnitz

Vortrag in der Reihe »Kulturgut Mathematik«
25. April 2024
Gerichtssaal 036, Landgericht Chemnitz

Mathematik! TU Chemnitz

TECHNISCHE UNIVERSITÄT IN DER KULTURHAUPTSTADT EUROPAS CHEMNITZ

Evangelisches FORUM CHEMNITZ

# Kulturgut

Bundesamt
für Bevölkerungsschutz
und Katastrophenhilfe

**Mögliche Kriterien zur Identifizierung von Kulturgut:** (i.S.d. Haager Konvention zum Schutz von Kulturgut bei bewaffneten Konflikten, 1954)

- es zeigt Entwicklungen der Menschheit auf nationaler, regionaler und globaler Ebene in einem oder mehreren Zeitabschnitten
- es ist ein Meisterwerk menschlicher Kreativität
- es bezeugt auf bedeutsame Weise eine kulturelle Tradition oder eine noch existierende oder bereits verschwundene Zivilisation
- es belegt bedeutende wissenschaftliche und kulturelle Errungenschaften des Menschen in einem bestimmten Zeitraum
- es hat eine große Bedeutung für die kulturelle Identität der betreffenden Gesellschaften

# Inhalt

# Inhalt

MS 3047 (Schøyen-Sammlung)
Multiplikationstabelle, Flächenberechnung
Sumer, 27. Jh. v. Chr.


Codex Ur-Nammu (Archäologisches Museum Istanbul)
Ältestes bekanntes Gesetzeswerk
Sumer, ca. 2100 v. Chr.

Seit der Spätantike: die Sieben Freien Künste


| **Trivium** (Kommunikation) | **Quadrivium** (Die Welt) |
|---|---|
| Grammatik | Arithmetik |
| Rhetorik | Geometrie |
| Dialektik | Musik |
| | Astronomie |

Vorbereitung auf die Fakultäten

Theologie, Jurisprudenz und Medizin

# Inhalt

# Wahrheitsbegriff

| **Mathematik** | **Justiz** |
| --- | --- |
| Rechnen | Urteilen, Entscheiden |

# Wahrheitsbegriff

**Mathematik**

~~Rechnen~~ Denken

**Justiz**

Urteilen, Entscheiden

# Wahrheitsbegriff

**Mathematik**

~~Rechnen~~ Denken

**Justiz**

Urteilen, Entscheiden

**Wahrheit**

# Wahrheitsbegriff

**Mathematik**

~~Rechnen~~ Denken

**Justiz**

Urteilen, Entscheiden

**Wahrheit**

**Erkenntnistheorie** (Epistemologie)

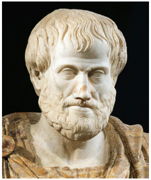
- Induktives Vorgehen (empirisch): Verallgemeinerung einzelner Beobachtungen (David Hume, Karl Popper, Paul Feyerabend)
  Natur-, Ingenieur-, Human- und Sozialwissenscchaften
- Deduktives Vorgehen (logisches Schließen): Ableiten neuer Erkenntnisse aus bereits gesicherten (Aristoteles)
  Mathematik, Philosophie

Klassische logische Denkregeln: $A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder

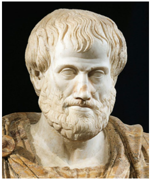


Klassische logische Denkregeln: $A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder



Satz vom ausgeschlossenen Dritten (tertium non datur):

| $A$ | $\neg A$ | $A \vee (\neg A)$ |
|---|---|---|
| $w$ | $f$ | $w$ |
| $f$ | $w$ | $w$ |

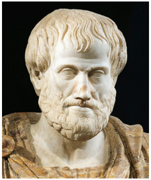
„Wenn der Hahn kräht auf dem Mist, ändert sich das Wetter oder es bleibt, wie es ist.“

Klassische logische Denkregeln: $A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder



Satz vom (ausgeschlossenen) Widerspruch (principium contradictionis):

| $A$ | $\neg A$ | $A \wedge (\neg A)$ |
|---|---|---|
| $w$ | $f$ | $f$ |
| $f$ | $w$ | $f$ |

„You can't have your cake and eat it, too."

„Man kann nicht alles haben."

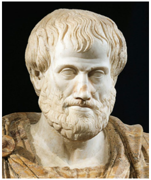
Klassische logische Denkregeln:

$A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder



Schlussregel: Ausschluss einer Disjunktion

| $A$ | $B$ | $A \vee B$ | $\neg A$ |
|---|---|---|---|
| $w$ | $w$ | $w$ | $f$ |
| $w$ | $f$ | $w$ | $f$ |
| $f$ | $w$ | $w$ | $w$ |
| $f$ | $f$ | $f$ | $w$ |

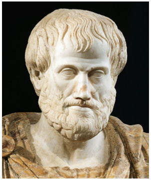
| | | |
|---|---|---|
| | $A \vee B$ | „Diese Idee ist blöd oder saublöd." |
| | $\neg A$ | „Diese Idee ist nicht blöd." |
| *ergo* | $B$ | „Diese Idee ist saublöd." |

Klassische logische Denkregeln: $A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder



Schlussregel: Bestätigung der Prämise (modus ponendo ponens)

| $A$ | $B$ | $A \Rightarrow B$ | $A$ |
|---|---|---|---|
| $w$ | $w$ | $w$ | $w$ |
| $w$ | $f$ | $f$ | $w$ |
| $f$ | $w$ | $w$ | $f$ |
| $f$ | $f$ | $w$ | $f$ |

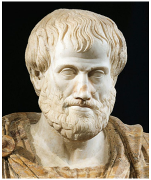
| | | |
|---|---|---|
| | $A \Rightarrow B$ | „Wenn es regnet, wird die Straße nass.“ |
| | $A$ | „Es regnet.“ |
| *ergo* | $B$ | „Die Straße wird nass.“ |

Klassische logische Denkregeln: $A, B$ : Aussagen (wahr oder falsch)

$\neg$ Verneinung

$\wedge$ und

$\vee$ oder

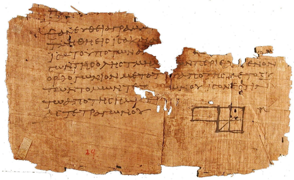


Schlussregel: Verneinung der Folgerung (modus tollendo tollens)

| $A$ | $B$ | $A \Rightarrow B$ | $\neg B$ |
|---|---|---|---|
| $w$ | $w$ | $w$ | $f$ |
| $w$ | $f$ | $f$ | $w$ |
| $f$ | $w$ | $w$ | $f$ |
| $f$ | $f$ | $w$ | $w$ |

| | | |
|---|---|---|
| | $A \Rightarrow B$ | „Wenn es geregnet hat, ist die Straße nass.“ |
| | $\neg B$ | „Die Straße ist nicht nass.“ |
| *ergo* | $\neg A$ | „Es hat nicht geregnet.“ |

**Euklid von Alexandria** (3. Jh. v. Chr.)



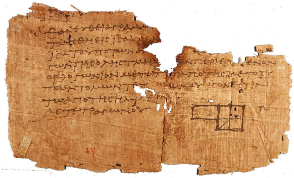
Elemente: systematische Abhandlung in 13 Büchern zu Geometrie, Arithmetik;

bis in 2. Hälfte des 19. Jh. das nach der Bibel meistverbreitete Werk der Weltliteratur

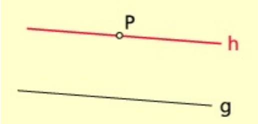
**Euklid von Alexandria** (3. Jh. v. Chr.)

Elemente: systematische Abhandlung in 13 Büchern zu Geometrie, Arithmetik;

bis in 2. Hälfte des 19. Jh. das nach der Bibel meistverbreitete Werk der Weltliteratur



- Aussagen (Sätze, Theoreme) aus begrenztem Vorrat von Definitionen und Axiomen durch Beweise abgeleitet.
- Vorbild für axiomatischen Aufbau der Mathematik
- Parallelenaxiom lange für überflussig gehalten.




- Erst im 19. Jh. (Gauss, Lobatschewski, Bolyai) als unabhängig von den übrigen euklidischen Axiomen bewiesen.
- Hilberts Programm (1920er Jahre), gesamte Mathematik axiomatisch aufzubauen, durch den Gödelschen Unvollständigkeitssatz (1931) als undurchführbar erwiesen.

Mathematischer **Beweis:** rhetorisches Konstrukt, um andere von der Richtigkeit einer Behauptung zu überzeugen.

- Folge von Aussagen verbunden durch logische Schlussweisen (z.B. modus ponendo ponens)
- Zahlreiche Techniken: Beweis durch vollständige Induktion, durch Widerspruch, durch Aufzählung etc.
- Einmal bewiesen, gilt ein Satz für immer (dagegen Literatur in z.B. Medizin, Informatik nach wenigen Jahren überholt).
- Stilfragen unterliegen Schwankungen (Bourbaki-Gruppe).


- Auch werden zunehmend Computer eingesetzt, um Beweise zu finden.

- Gutachtliche Fallösungstechnik, spielt zentrale Rolle in deutscher Juristenausbildung bei Übungen, Hausarbeiten, Examina.
- Zentraler Bestandteil: systematische Herantragen von Normhypothesen an den Sachverhalt, die mittels Subsumtion im Wege des „Justizsyllogismus" als anwendbar erkannt oder verworfen werden.

| | | |
|---|---|---|
| Grundform: | $\forall x : Tx \Rightarrow Rx$ | „Wenn $x$ den Tatbestand $T$ erfüllt, dann gilt für $x$ die Rechtsfolge $R$." |
| | $Ta$ | „$a$ erfüllt den Tatbestand." |
| *ergo* | $Ra$ | „Für $a$ gilt die Rechtsfolge $R$." |

- strenge Disziplin der Gedankenführung; komplette rechtliche Würdigung, Sachverhalt wird unter allen vernünftigerweise in Betracht kommenden Gesichtspunkten geprüft.
- Stuckenberg[1]: Ähnlichkeit mit Descartes „Methode zum rechten Vernunftgebrauch und zur Wahrheitssuche in den Wissenschaften"
- USA: IRAC-Schema für juristische Analysen (Issue, Rule, Application, Conclusion)

[1] C.-F. Stuckenberg. "Der juristische Gutachtenstil als cartesische Methode". In: *Zeitschrift für Didaktik der Rechtswissenschaft* 6.4 (2019), S. 323–341. DOI: 10.5771/2196-7261-2019-4-323.

> *In the course of my law-reading I constantly came upon the word demonstrate. I thought, at first, that I understood its meaning, but soon became satisfied that I did not. [...] At last I said, 'L[incoln], you can never make a lawyer if you do not understand what demonstrate means;' and I left my situation in Springfield, went home to my father's house, and staid there till I could give any propositions in the six books of Euclid at sight. I then found out what 'demonstrate' means, and went back to my law studies.*
>
> *Abraham Lincoln (1864)*[2]

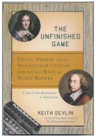
Übersetzung:

> *Im Verlauf meines Jurastudiums stieß ich ständig auf das Wort „demonstrieren“. Zuerst dachte ich, dass ich seine Bedeutung verstand, erkannte aber bald, dass ich es nicht tat. [...] Schließlich sagte ich: „L[incoln], du kannst niemals Anwalt werden, wenn du nicht verstehst, was 'demonstrieren' bedeutet“; und ich verließ meine Stelle in Springfield, ging nach Hause zu meinem Vater und blieb dort, bis ich jede Proposition in den sechs Büchern von Euklid auf Anhieb geben konnte. So fand ich heraus, was „demonstrieren“ bedeutet, und kehrte zu meinem Jurastudium zurück.*

[2] M. Termini. "Proving the Point: Connections Between Legal and Mathematical Reasoning". In: *Suffolk University Law Review* 52.1 (2019), S. 5–36.

# Wahrheitsbegriff

## Einzug der Wahrscheinlichkeiten




- 24. August 1654: Briefwechsel zwischen Blaise Pascal und Pierre de Fermat gilt als Geburtsstunde der Wahrscheinlichkeitsrechnung.
- Naturwissenschaft: rationaler Umgang mit ungewissen Ausgängen und Ursachen („probability of causes“); mathematisches Kalkül für induktives Schließen.
- In Verbindung mit Statistik Anwendung auch auf (frühe) Human-, Sozial-, Wirtschafts- und Rechtswissenschaften („Moral Sciences“)
- Bemerkenswert viele Pioniere der Wahrscheinlichkeitstheorie wandten diese auf Gerichtsentscheidungen an.
- Leitlinie: Laplaces Losung, Wahrscheinlichkeitstheorie sei lediglich in Kalkül übersetzter gesunder Menschenverstand.
- Historischer Hintergrund: gesellschaftliches Verlangen nach Justizreform im vorrevolutionären Frankreich, auch angesichts spektakulärer Justizirrtümer
  - Familie Calas
  - Chevalier de la Barre
  - Elisabeth Canning
- Bekannteste Vertreter:
  Condorcet (1743–1794), Laplace (1749–1827), Poisson (1781–1840)

Wie groß darf die Wahrscheinlichkeit eines Fehlurteils sein?
Condercets Antwort:

$$\frac{1}{144\,768} \approx 1,9 \cdot 10^{-6}$$

(Wahrscheinlichkeit eines gesunden Mannes in den nächsten 24 Stunden zu sterben (aus Sterbetafeln statistisch ermittelt).

Wie groß darf die Wahrscheinlichkeit eines Fehlurteils sein?
Condercets Antwort:

$$\frac{1}{144\,768} \approx 1,9 \cdot 10^{-6}$$

(Wahrscheinlichkeit eines gesunden Mannes in den nächsten 24 Stunden zu sterben (aus Sterbetafeln statistisch ermittelt).

Condorcets Betrachtung:

- Justiz als Gesellschaftsvertrag: Risikoausgleich zwischen sozialer Ordnung und individueller Freiheit.
- Für lange Serie von Urteilen genügt es, wenn durchschnittlich

$$\frac{\mathbf{P}(\text{verurteilter Angeklagter schuldig})}{\mathbf{P}(\text{freigesprochener Angeklagter unschuldig})} \approx \frac{\text{Nachteile Verurteilung Unschuldiger}}{\text{Nachteile Freisprechung Schuldiger}}$$

- Für den Einzelnen genügt diese gesamtgesellschaftliche Betrachtung nicht, viel geringeres Risiko erforderlich.
- O.g. Risiko gehen Menschen jeden Tag ein, wird als akzeptabel betrachtet („moralische Gewissheit“).

Optimale Gestaltung einer Gruppenentscheidung (ja/nein) in Bezug auf

$n$ : Anzahl Personen

$p$ : Einzelwahrscheinlichkeit für richtige Entscheidung ($0 \leqslant p \leqslant 1$)

$m$ : erforderliche Mehrheit

Optimale Gestaltung einer Gruppenentscheidung (ja/nein) in Bezug auf

$$\begin{aligned} n &: \text{Anzahl Personen} \\ p &: \text{Einzelwahrscheinlichkeit für richtige Entscheidung } (0 \leqslant p \leqslant 1) \\ m &: \text{erforderliche Mehrheit} \end{aligned}$$

Ist $X \in \{0, 1, 2, \ldots, n\}$ die Anzahl Personen, die richtig entschieden haben, so gilt

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 1, 2, \ldots, n. \qquad \text{(Binomialverteilung)}$$

Optimale Gestaltung einer Gruppenentscheidung (ja/nein) in Bezug auf

$$\begin{aligned} n &: \text{Anzahl Personen} \\ p &: \text{Einzelwahrscheinlichkeit für richtige Entscheidung } (0 \leqslant p \leqslant 1) \\ m &: \text{erforderliche Mehrheit} \end{aligned}$$

Ist $X \in \{0, 1, 2, \dots, n\}$ die Anzahl Personen, die richtig entschieden haben, so gilt

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 1, 2, \dots, n. \qquad \text{(Binomialverteilung)}$$

Die Wahrscheinlichkeit, dass die Mehrheit richtig entschieden hat, ergibt sich dann zu

$$\mathbf{P}\left(X > \frac{n}{2}\right) = \sum_{k=\lceil n/2 \rceil}^{n} \mathbf{P}(X = k).$$

Optimale Gestaltung einer Gruppenentscheidung (ja/nein) in Bezug auf

$$\begin{aligned} n &: \text{Anzahl Personen} \\ p &: \text{Einzelwahrscheinlichkeit für richtige Entscheidung } (0 \leqslant p \leqslant 1) \\ m &: \text{erforderliche Mehrheit} \end{aligned}$$

Ist $X \in \{0, 1, 2, \dots, n\}$ die Anzahl Personen, die richtig entschieden haben, so gilt

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 1, 2, \dots, n. \qquad \text{(Binomialverteilung)}$$

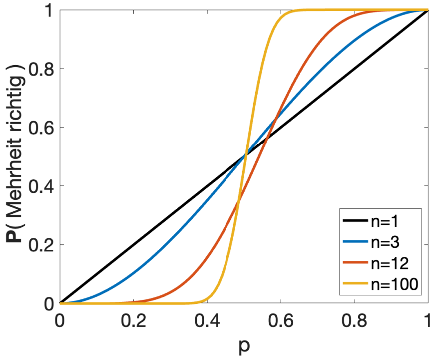
Die Wahrscheinlichkeit, dass die Mehrheit richtig entschieden hat, ergibt sich dann zu

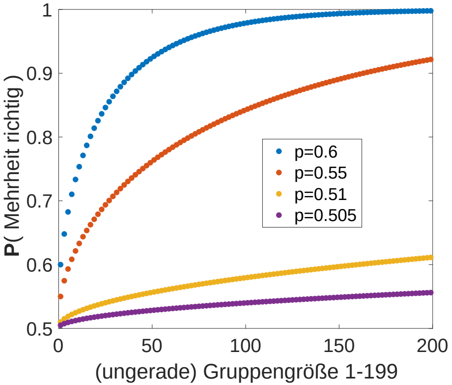
$$\mathbf{P}\left(X > \frac{n}{2}\right) = \sum_{k=\lceil n/2 \rceil}^{n} \mathbf{P}(X = k).$$

**Theorem (Condorcet 1785)**

Für $p > \frac{1}{2}$, $n > 1$ und $p$ ungerade gilt

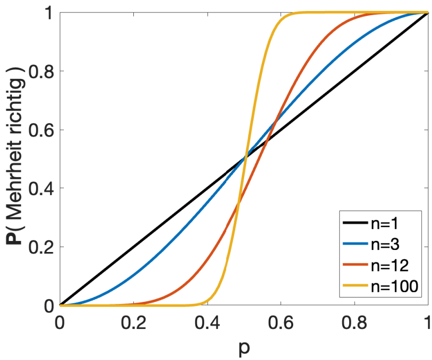
$$\mathbf{P}\left(X > \frac{n}{2}\right) > p \qquad \text{und} \qquad \mathbf{P}\left(X > \frac{n}{2}\right) \nearrow 1 \quad \text{für } n \to \infty.$$

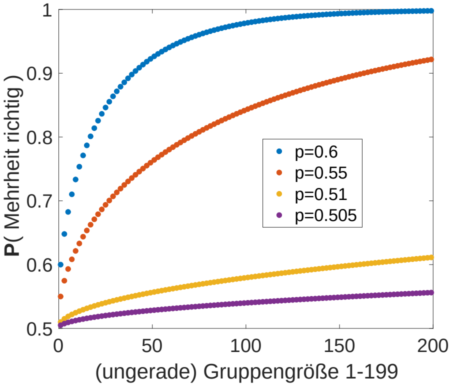

P( Mehrheit richtig )
0
0.2
0.4
0.6
0.8
1
0
0.2
0.4
0.6
0.8
1
p
n=1
n=3
n=12
n=100


P( Mehrheit richtig )
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
(ungerade) Gruppengröße 1-199
p=0.6
p=0.55
p=0.51
p=0.505

# Probabilité des Jugements

## Condorcets Jury-Theorem (1785)




- Bei $n = 12$ erfordert $\mathbf{P}\,(X \geqslant 8) > 1 - \frac{1}{144768}$ eine Einzelwahrscheinlichkeit von $p = 0.98$.
- Condorcets Schlussfolgerung: große demokratische Versammlungen nur in primitiven Gesellschaften mit eher ignoranten Einzelpersonen sinnvoll; in höher entwickelten Gesellschaften genügt eine aufgeklärte Oligarchie.

# Inhalt

In der Forensik stellen assoziative Beweismittel Verbindungen her zwischen Angeklagten und Straftat anhand von Merkmalen wie

- Blutgruppe
- Fingerabdrücke
- Textilfasern
- DNA-Proben

Oft wird als statistische Information angegeben, wie häufig ein Merkmal bei zwei zufällig ausgewählten Personen übereinstimmt (Inzidenz).

Bei der Beweiswürdigung solcher Inzidenzstatistik kommt es leicht (oft) zu Fehlschlüssen: Trugschluss des Anklägers (prosecutor's fallacy)[3,4].

Typisches Argumentationsmuster: die vorliegende Übereinstimmung (z.B. DNA) wäre extrem selten bei zufällig ausgewählten Personen, ergo ist die Wahscheinlichkeit, dass der Angeklagte unschuldig ist, extrem klein.

---

[3] W. C. Thompson und E. L. Schumann. "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy". In: *Law and Human Behavior* 11.3 (1987), S. 167–187. DOI: 10.1007/BF01044641.

[4] N. Sesardić. "Guilt by Statistical Association: Revisiting the Prosecutor's Fallacy and the Interrogator's Fallacy". In: *Journal of Philosophy* 105.6 (2008), S. 320–332.

Hintergrund: bedingte Wahrscheinlichkeit, die Änderung der Wahrscheinlichkeit $\mathbf{P}(A)$ eines Ereignisses $A$ infolge Kenntnis eines anderen bereits eingetretenen Ereignisses $B$.

Hintergrund: bedingte Wahrscheinlichkeit, die Änderung der Wahrscheinlichkeit $\mathbf{P}(A)$ eines Ereignisses $A$ infolge Kenntnis eines anderen bereits eingetretenen Ereignisses $B$.

**Beispiel:** $A =$ „Beide Kinder unserer neuen Nachbarn sind Mädchen."

Hintergrund: bedingte Wahrscheinlichkeit, die Änderung der Wahrscheinlichkeit $\mathbf{P}(A)$ eines Ereignisses $A$ infolge Kenntnis eines anderen bereits eingetretenen Ereignisses $B$.

**Beispiel:** $A =$ „Beide Kinder unserer neuen Nachbarn sind Mädchen."

$$\mathbf{P}(A) = \frac{|\{MM\}|}{|\{JJ, JM, MJ, MM\}|} = \frac{1}{4}.$$

Hintergrund: bedingte Wahrscheinlichkeit, die Änderung der Wahrscheinlichkeit $\mathbf{P}(A)$ eines Ereignisses $A$ infolge Kenntnis eines anderen bereits eingetretenen Ereignisses $B$.

**Beispiel:** $A =$ „Beide Kinder unserer neuen Nachbarn sind Mädchen."

$$\mathbf{P}(A) = \frac{|\{MM\}|}{|\{JJ, JM, MJ, MM\}|} = \frac{1}{4}.$$

Zusatzinformation: $B =$ „Eines der Kinder ist ein Mädchen".

$$\mathbf{P}(B) = \frac{|\{MM, JM, MJ\}|}{|\{JJ, JM, MJ, MM\}|} = \frac{3}{4}.$$

Hintergrund: bedingte Wahrscheinlichkeit, die Änderung der Wahrscheinlichkeit $\mathbf{P}(A)$ eines Ereignisses $A$ infolge Kenntnis eines anderen bereits eingetretenen Ereignisses $B$.

**Beispiel:** $A =$ „Beide Kinder unserer neuen Nachbarn sind Mädchen."

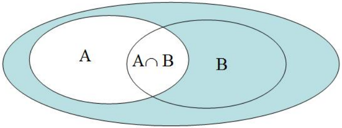
$$\mathbf{P}(A) = \frac{|\{MM\}|}{|\{JJ, JM, MJ, MM\}|} = \frac{1}{4}.$$

Zusatzinformation: $B =$ „Eines der Kinder ist ein Mädchen".

$$\mathbf{P}(B) = \frac{|\{MM, JM, MJ\}|}{|\{JJ, JM, MJ, MM\}|} = \frac{3}{4}.$$

Bedingte Wahrscheinlichkeit:

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A)}{\mathbf{P}(B)} = \frac{1/4}{3/4} = \frac{1}{3}.$$



Sei nun

$$A = \text{„DNA-Probe zweier beliebiger Personen stimmen überein“},$$
$$\mathbf{P}(A) = 10^{-5} = 0,00001 =: \varepsilon.$$
$$S = \text{„Angeklagter schuldig“}, \qquad \neg S = \text{„Angeklagter nicht schuldig“}$$
$$M = \text{„DNA von Angeklagtem und Täter stimmen überein“ (Match)}$$

Sei nun

$$
\begin{aligned}
A &= \text{„DNA-Probe zweier beliebiger Personen stimmen überein“},\\
&\quad \mathbf{P}(A) = 10^{-5} = 0,00001 =: \varepsilon.\\
S &= \text{„Angeklagter schuldig“}, \qquad \neg S = \text{„Angeklagter nicht schuldig“}\\
M &= \text{„DNA von Angeklagtem und Täter stimmen überein“ (Match)}
\end{aligned}
$$

Dann ist

$$\mathbf{P}(M|\neg S) = \mathbf{P}(A) = \varepsilon.$$

Mit anderen Worten: eine unschuldige Erklärung der vorliegenden Übereinstimmung ist höchst unwahrscheinlich, ergo ist die Schuld des Angeklagten höchst wahrscheinlich.

Sei nun

$$A = \text{„DNA-Probe zweier beliebiger Personen stimmen überein“},$$
$$\mathbf{P}(A) = 10^{-5} = 0,00001 =: \varepsilon.$$
$$S = \text{„Angeklagter schuldig“}, \qquad \neg S = \text{„Angeklagter nicht schuldig“}$$
$$M = \text{„DNA von Angeklagtem und Täter stimmen überein“ (Match)}$$

Dann ist

$$\mathbf{P}(M|\neg S) = \mathbf{P}(A) = \varepsilon.$$

Mit anderen Worten: eine unschuldige Erklärung der vorliegenden Übereinstimmung ist höchst unwahrscheinlich, ergo ist die Schuld des Angeklagten höchst wahrscheinlich.

Von Interesse ist hier aber $\mathbf{P}(\neg S|M)$.

Sei nun

$$A = \text{„DNA-Probe zweier beliebiger Personen stimmen überein“},$$
$$\mathbf{P}(A) = 10^{-5} = 0,00001 =: \varepsilon.$$
$$S = \text{„Angeklagter schuldig“}, \qquad \neg S = \text{„Angeklagter nicht schuldig“}$$
$$M = \text{„DNA von Angeklagtem und Täter stimmen überein“ (Match)}$$

Dann ist

$$\mathbf{P}(M|\neg S) = \mathbf{P}(A) = \varepsilon.$$

Mit anderen Worten: eine unschuldige Erklärung der vorliegenden Übereinstimmung ist höchst unwahrscheinlich, ergo ist die Schuld des Angeklagten höchst wahrscheinlich.

Von Interesse ist hier aber $\mathbf{P}(\neg S|M)$.

Formel von Bayes:

$$\mathbf{P}(\neg S|M) = \frac{\mathbf{P}(M|\neg S) \cdot \mathbf{P}(\neg S)}{\mathbf{P}(M)}$$

Sei nun

$$A = \text{„DNA-Probe zweier beliebiger Personen stimmen überein“},$$

$$\mathbf{P}(A) = 10^{-5} = 0,00001 =: \varepsilon.$$

$$S = \text{„Angeklagter schuldig“}, \qquad \neg S = \text{„Angeklagter nicht schuldig“}$$

$$M = \text{„DNA von Angeklagtem und Täter stimmen überein“ (Match)}$$

Dann ist

$$\mathbf{P}(M|\neg S) = \mathbf{P}(A) = \varepsilon.$$

Mit anderen Worten: eine unschuldige Erklärung der vorliegenden Übereinstimmung ist höchst unwahrscheinlich, ergo ist die Schuld des Angeklagten höchst wahrscheinlich.

Von Interesse ist hier aber $\mathbf{P}(\neg S|M)$.

Formel von Bayes:

$$\mathbf{P}(\neg S|M) = \frac{\mathbf{P}(M|\neg S) \cdot \mathbf{P}(\neg S)}{\mathbf{P}(M)}$$

Die beiden stimmen also genau dann überein, wenn $\mathbf{P}(\neg S) = \mathbf{P}(M)$.

Nun ist (totale Wahrscheinlichkeit)

$$\mathbf{P}(M) = \underbrace{\mathbf{P}(M|\neg S)}_{=\varepsilon} \cdot \mathbf{P}(\neg S) + \underbrace{\mathbf{P}(M|S)}_{\approx 1} \cdot \mathbf{P}(S)$$

Somit

$$\mathbf{P}(\neg S|M) = \frac{\varepsilon \cdot \mathbf{P}(\neg S)}{\varepsilon \mathbf{P}(\neg S) + 1 \cdot [1 - \mathbf{P}(\neg S)]} = \frac{\varepsilon \cdot \mathbf{P}(\neg S)}{1 - (1-\varepsilon)\mathbf{P}(\neg S)}$$

- Entscheidend: $\mathbf{P}(\neg S)$ unabhängig von DNA Beweismittel.
- Wäre etwa infolge weiterer Beweise $\mathbf{P}(\neg S) = 0.5$, so ergäbe sich
$$\mathbf{P}(\neg S|M) = \frac{\varepsilon \cdot 0.5}{1 - (1-\varepsilon)0.5} \approx \epsilon.$$
- Wäre der Angeklagte abgesehen vom DNA-Match einer von 10 000 in Frage kommenden Tätern, also $\mathbf{P}(\neg S) = 0.9999$, so hieße dies
$$\mathbf{P}(\neg S|M) = \frac{\varepsilon \cdot 0.9999}{1 - (1-\varepsilon)0.9999} \approx 10\%.$$
- **Fazit:** Trugschluss des Anklägers droht weniger bei sehr kleinem $\varepsilon$ oder wenn weitere Beweismittel $\mathbf{P}(\neg S)$ stark erhöhen.

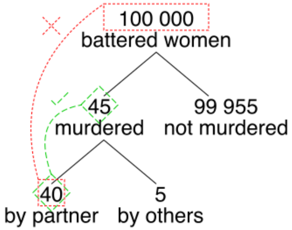
**Prominenter Fall:** Mordprozess gegen O. J. Simpson, Los Angeles 1995

- Blutspuren am Tatort stimmten mit Simpsons Blutgruppe überein (1:400)
- Verteidigung: Personen in LA mit dieser Übereinstimmung würden ein Fußballstadion füllen, daher unbedeutend (s.o.).
- Weiteres Indiz: Simpson hatte Mordopfer (Ehefrau) im Vorfeld der Tat misshandelt.
- Verteidigung: Nur in $\frac{1}{2500}$ von Misshandlungsfällen durch Partner kommt es zu einem Mord.

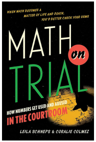
**Prominenter Fall:** Mordprozess gegen O. J. Simpson, Los Angeles 1995

- Blutspuren am Tatort stimmten mit Simpsons Blutgruppe überein (1:400)
- Verteidigung: Personen in LA mit dieser Übereinstimmung würden ein Fußballstadion füllen, daher unbedeutend (s.o.).
- Weiteres Indiz: Simpson hatte Mordopfer (Ehefrau) im Vorfeld der Tat misshandelt.
- Verteidigung: Nur in $\frac{1}{2500}$ von Misshandlungsfällen durch Partner kommt es zu einem Mord.

- Außer Acht gelassen: Ehefrau gehört zu den Misshandlungsfällen, die mit Mord enden.
- Weitere Statistik: von allen Misshandlungsfällen, die in Mord enden, war der Partner in 90% der Fälle der Täter.



10 mathematische Fehlschlüsse und ihr Vorkommen in Gerichtsverfahren:
Leila Schneps und Coralie Colmez. Math on Trial: How Numbers Get Used and Abused in the Courtroom. Basic Books (2013)




1. Unzulässige Multiplikation von Wahrscheinlichkeiten (Sally Clark, SIDS, UK 1999)
2. Unfundierte Schätzungen (Janet Collins, Raubüberfall, USA 1964)
3. Mathematische Einschüchterung bei dürftiger Beweislage (Joe Sneed, Mord, USA 1964)
4. Unterschätzung der Effektivität doppelter DNA-Tests (Amanda Knox, Mord, Italien 2007)
5. Geburtstagsparadoxon (Diana Sylvester, Mord, USA 2003)
6. Simpsons Paradoxon (Geschlechterdiskriminierung bei Zulassung, UC Berkeley, 1988)
7. Seltene Ereignisse (Lucia de Berk, Patientenmorde, Niederlande, 2001)
8. Exponentielles Wachstum, Schneeballsysteme (Ponzi/Madoff, Betrug, USA 1920/2009)
9. Falsche Modellierung (Hetty Green, Testamentsanfechtung, USA 1865)
10. Überzeugender mathematischer Irrsinn (Dreyfus-Affäre, Landesverrat, Frankreich 1894)

# Inhalt

Grundmodell im statistischen Lernen (überwachtes Lernen):

datengenerierendes Modell: $Y = f(X) + \varepsilon$, Approximation: $\hat{f} \approx f$.

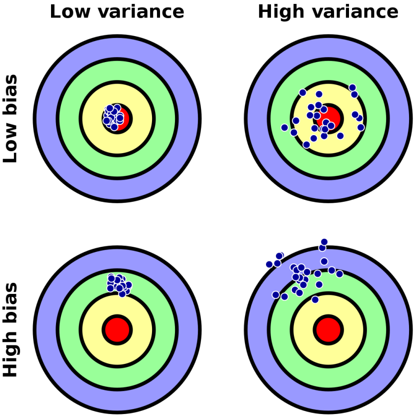
Gütemaß für Approximation: Quadratmittelfehler (MSE), Bias-Varianz-Zerlegung

$$\mathrm{MSE}(\hat{f}) := \mathbf{E}_{X,\varepsilon}\left[(Y - \hat{f}(X))^2\right] = \mathrm{Bias}(\hat{f})^2 + \mathbf{Var}\,\hat{f} + \mathbf{Var}\,\varepsilon$$

$$\mathrm{Bias}(\hat{f}) = f - \mathbf{E}\left[\hat{f}\right]$$

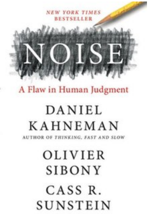
Grundmodell im statistischen Lernen (überwachtes Lernen):

datengenerierendes Modell: $Y = f(X) + \varepsilon$, Approximation: $\hat{f} \approx f$.



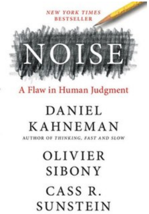
**Daniel Kahneman: Noise in der Justiz**

Im öffentlichen Diskurs geht es im Zusammenhang mit Diskriminierungsvorwürfen an Polizei und Justiz meist um systematische Ungleichbehandlung anhand von Merkmalen wie Ethnizität, sozioökonomischer Hintergrund, Geschlecht, sexuelle Orientierung etc.




Kahneman et al. zeigen, dass oft Noise die größere Rolle spielt.

**Daniel Kahneman: Noise in der Justiz**

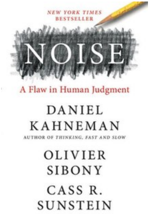



Im öffentlichen Diskurs geht es im Zusammenhang mit Diskriminierungsvorwürfen an Polizei und Justiz meist um systematische Ungleichbehandlung anhand von Merkmalen wie Ethnizität, sozioökonomischer Hintergrund, Geschlecht, sexuelle Orientierung etc.

Kahneman et al. zeigen, dass oft Noise die größere Rolle spielt.

- 1981 Studie mit 208 Bundesrichtern, Vorlage von 16 hypothetischen Fallbeispielen
  - In nur 3 von 16 Fällen einstimmige Empfehlung einer Haftstrafe
  - In den Fällen mit vorwiegender Haftempfehlung erhebliche Schwankungen bei Haftdauer
  - In einem Betrugsdelikt: mittlere Haftdauer 8,5 Jahre, maximale lebenslänglich
  - Anderer Fall: mittlere Haftstrafe 1,1 Jahre, maximale 15 Jahre.
  - Variabilität eher unterschätzt, da in Praxis viel mehr Details eines Falles bekannt

**Daniel Kahneman: Noise in der Justiz**

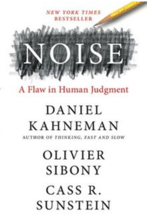



Im öffentlichen Diskurs geht es im Zusammenhang mit Diskriminierungsvorwürfen an Polizei und Justiz meist um systematische Ungleichbehandlung anhand von Merkmalen wie Ethnizität, sozioökonomischer Hintergrund, Geschlecht, sexuelle Orientierung etc.

Kahneman et al. zeigen, dass oft Noise die größere Rolle spielt.

- Strafzumessung und Football:
  - Studie von über 1000 Jugendstrafverfahren: im Schnitt härtere Strafzumessung am Montag, wenn am Wochenende zuvor die lokale Football-Mannschaft gewonnen hatte.
  - Weitere Studie, 1,5 Mio Urteile über 3 Jahrzehnte: Richter erteilen höhere Strafzumessung an Tagen nach Niederlage der lokalen Football-Mannschaft.

**Daniel Kahneman: Noise in der Justiz**

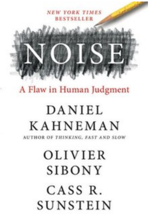



Im öffentlichen Diskurs geht es im Zusammenhang mit Diskriminierungsvorwürfen an Polizei und Justiz meist um systematische Ungleichbehandlung anhand von Merkmalen wie Ethnizität, sozioökonomischer Hintergrund, Geschlecht, sexuelle Orientierung etc.

Kahneman et al. zeigen, dass oft Noise die größere Rolle spielt.

- Strafmaß und Geburtstag:
  - Studie von 6 Mio Urteile französischer Richter über 12 Jahre: Angeklagte erhalten mildere Urteile an ihrem Geburtstag.
- Asylentscheidungen und Außentemperatur
  - Auswertung von 207 000 Entscheidungen eines Einwanderungsgerichts über 4 Jahre (USA): negative Asylentscheidungen korrelieren mit höherer Außentemperatur.

**Daniel Kahneman: Noise in der Justiz**



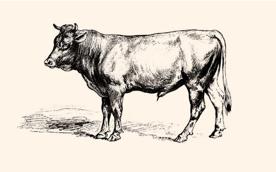

Im öffentlichen Diskurs geht es im Zusammenhang mit Diskriminierungsvorwürfen an Polizei und Justiz meist um systematische Ungleichbehandlung anhand von Merkmalen wie Ethnizität, sozioökonomischer Hintergrund, Geschlecht, sexuelle Orientierung etc.

Kahneman et al. zeigen, dass oft Noise die größere Rolle spielt.

- 1984: US-Kongress verabschiedet *Sentencing Reform Act*, Schaffung einer Kommission zur Festlegung vepflichtender Richtlinien für Strafmaß zur Reduktion von Variabilität in Strafzumessung
- In den Folgejahren tatsächlich deutlich geringere Variabilität.
- In den Folgejahren ebenfalls starker Protest von Richterseite
- 2005 annulliert der Supreme Court die Richtlinien
- In der Folge wieder Rückkehr zu starker Variabilität.

## Klassische Maßnahme gegen Noise: Mittelung

Schätzexperiment Sir Francis Galton (1906)



Nutztiermesse in Plymouth (England), Wettbewerb zur korrekten Schätzung des Gewichts eines Ochsen

Von Galton zunächst als Demonstration für Unzulänglichkeit von Mehrheitsentscheidungen vorgesehen

- 787 gültige Schätzungen abgegeben
- Tatsächliches Gewicht: 1198 lb.
- Median der Schätzungen: 1207 lb. (Fehler $< 0.8$ %)
- Extrema der Schätzfehler: -124 lb. (-10%) bis +95 lb. (+8%)

*“This result is, I think, more creditable to the trustworthiness of a democratic judgment than might have been expected.“*

*Francis Galton, Vox Populi. Nature 75, p. 450–451 (1907)*

## Klassische Maßnahme gegen Noise: Mittelung

Statistik: Schätzvarianz des Mittelwerts unabhängiger Stichproben

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}, \qquad \mathbf{E}\left[X_i\right] = 0, \; \mathbf{Var}\, X_i = \sigma^2.$$

### Klassische Maßnahme gegen Noise: Mittelung

Statistik: Schätzvarianz des Mittelwerts unabhängiger Stichproben

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}, \qquad \mathbf{E}\left[X_i\right] = 0,\ \mathbf{Var}\, X_i = \sigma^2.$$

Dann ist

$$\mathbf{E}\left[\ \overline{X}\ \right] = 0,$$
$$\mathbf{Var}\,\overline{X} = \frac{\sigma^2}{n^2} + \frac{\sigma^2}{n^2} + \cdots + \frac{\sigma^2}{n^2} = n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Mit anderen Worten: bei doppelt sovielen Stimmen halbiert sich die Variabilität.

## Klassische Maßnahme gegen Noise: Mittelung

Statistik: Schätzvarianz des Mittelwerts unabhängiger Stichproben

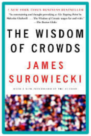
$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}, \qquad \mathbf{E}\left[X_i\right] = 0,\ \mathbf{Var}\, X_i = \sigma^2.$$

Dann ist

$$\begin{aligned} \mathbf{E}\left[\ \overline{X}\ \right] &= 0, \\ \mathbf{Var}\,\overline{X} &= \frac{\sigma^2}{n^2} + \frac{\sigma^2}{n^2} + \cdots + \frac{\sigma^2}{n^2} = n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}. \end{aligned}$$

Mit anderen Worten: bei doppelt sovielen Stimmen halbiert sich die Variabilität.

Schwarmintelligenz (wisdom of crowds) stellt sich nur unter bestimmten Voraussetzungen ein, insbesondere müssen die Stimmen unabhängig sein.



- Level of Service Inventory-Revised (LSI-R): standardisiertes Verfahren zur Untersuchung von Rechtsbrechern mit dem Ziel der Risikobewertung z.B. bei vorzeitiger Entlassung. 1995 in USA entwickelt, 2012 an deutsche Verhältnisse angepasst.
- Oberlandesgerichtsassistent (OLGA), OLG Stuttgart. Manuelle Sachverhaltsprüfung in Massenverfahren (Dieselklagen), Zuordnung von Verfahren zu Fallgruppen. Erkennung von Tenor, Tatbestand, Anträge; Analyse von erstinstanzlichem Urteil, Berufungsbegründung sowie Berufungserwiderung; keine Entscheidungen, eher Suchmaschine.
- FRAnkfurter Urteils-Konfigurator Elektronisch (FRAUKE), AG Frankfurt a. M.; anhand abgeschlossener Fälle in den Fluggastrechteverfahren erprobt; extrahiert aus Prozessvorbringen der Parteien wesentliche Kernelemente. Aufbereiten der Informationen, Heranziehung von Entscheidungen zu vergleichbaren Sachverhalten, Formulierungsvorschläge für richterliche Entscheidung. (Machbarkeitsstudie)
- **Fazit:** Angesichts der rasanten Entwicklung (LLM, Gesichtserkennung, prädiktive Polizeiarbeit) sind bisherige Initiativen nur der Anfang.

# Inhalt

**Ursprung:** Römische Digestes[5], vermutlich Aemilius Macer, *De Appellationibus* (3. Jh.)

Ein Urteil bleibt gültig auch wenn ihm Rechenfehler zugrundeliegen.

Verwandt:

*Error calculi non nocet. (Ein Rechenfehler schadet nicht.)*

[5] Im Auftrag des oströmischen Kaisers Justinian zusammengestellte spätantike Kompilation der Jurisprudenz der Rechtsgelehrten der klassisch-römischen Kaiserzeit.

**Ursprung:** Römische Digestes[5], vermutlich Aemilius Macer, *De Appellationibus* (3. Jh.)

Ein Urteil bleibt gültig auch wenn ihm Rechenfehler zugrundeliegen.

Verwandt:

*Error calculi non nocet. (Ein Rechenfehler schadet nicht.)*

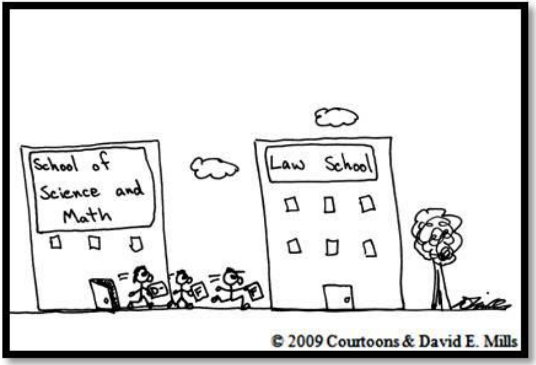
**§ 319 ZPO** *(1) Schreibfehler, Rechnungsfehler und ähnliche offenbare Unrichtigkeiten, die in dem Urteil vorkommen, sind jederzeit von dem Gericht auch von Amts wegen zu berichtigen.*

Rechenfehler im Urteil haben also keine Auswirkung auf die Rechtskraft des Urteiles an sich.

[5] Im Auftrag des oströmischen Kaisers Justinian zusammengestellte spätantike Kompilation der Jurisprudenz der Rechtsgelehrten der klassisch-römischen Kaiserzeit.

*I think there are a lot of people who go to law school because they are not good at math and can't think of anything else to do.*

*Chief Justice John Roberts (Oct. 26, 2012)*




Where lawyers come from.

**Problem der Mathematik vor Gericht:** In einem folgenreichen Aufsatz[6] argumentiert Prof. Laurence Tribe (Harvard Law School) gegen die Verwendung von Mathematik im Strafprozess.

- Gefahr, die von Irrtümern ausgeht, ist zu groß.
- Logischer/quantitativer Ansatz im mathematischen Denken unterscheidet sich so stark vom intuitiven Ansatz, dem Geschworenr/Richter bei der Beweiswürdigung folgen müssen, dass es zu oft nicht gelingt, diese beiden Ansätze richtig zu kombinieren.
- Eine Situation ist selten so einfach, dass ein mathematisches Modell alle lebensweltlichen Nuancen berücksichtigt.
- Psychologische Wirkung der Mathematik kann Geschworene/Richter überwältigen; Gefahr, "dass das Rechtssystem noch fremdartiger und unmenschlicher wirkt, als es ohnehin bei beunruhigend vielen Menschen der Fall ist."
- Aufsatz drängte Einsatz von Statistik in US-Gerichten viele Jahre stark zurück.

[6] L. H. Tribe. "Trial by Mathematics: Precision and Ritual in the Legal Process". In: *Harvard Law Review*, 84.6 (1971), S. 1329–1393.

**45 Jahre später:** Forschungssemester “Probability and Statistics in Forensic Science” am Isaac Newton Institute for Mathematical Sciences, Cambridge (UK) Juli-Dezember, 2016.

Verabschiedung von 12 Richtlinien[7] für Mathematiker, Forensiker und Juristenzur verantwortungsvollen Würdigung quantitativen Beweismaterials in Strafprozessen.

- Prinzip 1: Wahrscheinlichkeit ist grundlegend bei der Beweiswürdigung.
- Prinzip 11: Bayes’sche Analyse ist eine Standardmethode zur Aktualisierung von Wahrscheinlichkeiten nach Beobachtung mehrerer Beweismittel und daher hervorragend geeignet für die Evidenzsynthese.
- Prinzip 12: Bei der Kommunikation statistischer Analysen sowie Synthesen von Beweismitteln für Richter und Geschworenen ist au,erordentliche Sorgfalt erforderlich.

[7] N. Balding D. u. a. *Twelve Guiding Principles and Recommendations for Dealing with Quantitative Evidence in Criminal Law: For the Use of Statisticians, Forensic Scientists and Legal Professionals*. Techn. Ber. Cambridge, UK: Isaac Newton Institute for Mathematical Sciences, 2017.

# Fazit

## Iudex non calculat?

Persönliche Schlussfolgerungen nach Beschäftigung mit der Thematik:

- Mathematische wie juristische Praxis erfordert neben Vermittlung von Wissen auch eine vorbereitende Denkschule.
- Mathematik und Justiz haben unterschiedliche Wahrheitsbegriffe.
- Mathematik bietet formale Methoden der korrekten Beweiswürdigung in vielen Situationen. Diese sind oft der einzige rettende Anker beim Entscheiden unter Unsicherheit.
- Diskussion über angemessene Rolle von Mathematik in Justiz mehrere Jahrhunderte alt. Künstliche Intelligenz stellt ähnliche Fragen erneut.
- Es wäre gut, wenn Juristen für ihr Berufsleben mit mehr als Abiturmathematik ausgestattet würden.

Persönliche Schlussfolgerungen nach Beschäftigung mit der Thematik:

- Mathematische wie juristische Praxis erfordert neben Vermittlung von Wissen auch eine vorbereitende Denkschule.
- Mathematik und Justiz haben unterschiedliche Wahrheitsbegriffe.
- Mathematik bietet formale Methoden der korrekten Beweiswürdigung in vielen Situationen. Diese sind oft der einzige rettende Anker beim Entscheiden unter Unsicherheit.
- Diskussion über angemessene Rolle von Mathematik in Justiz mehrere Jahrhunderte alt. Künstliche Intelligenz stellt ähnliche Fragen erneut.
- Es wäre gut, wenn Juristen für ihr Berufsleben mit mehr als Abiturmathematik ausgestattet würden.

**Iudex non calculat? Er/sie kann nicht anders!** ☺

# Weitere Informationen

## Statistik im Alltag

Unstatistik des Monats (Harding-Zentrum, Berlin)

https://www.hardingcenter.de/de/transfer-und-nutzen/unstatistik-des-monats

Podcast „More or Less“ (BBC)

Klarstellung irreführender statistischer Behauptungen in den Medien.

**Vielen Dank**

**für’s Zuhören !**