Rechnernutzung in der Physik

Karlsruher Institut für Technologie
Wintersemester 2020/21

Teil – 1 “Computerbasierte Methoden der Statistik”

Roger Wolf (roger.wolf@cern.ch), Ralf Ulrich (ralf.ulrich@kit.edu)

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Inhalt

Roger Wolf

Ralf Ulrich

Literaturempfehlungen

- **Einführende Literatur zu Statistik und Numerik:**
  - G. Cowan, *Statistical data analysis*, Oxford (1997) (KIT-Bibliothek).
  - G. Bohm, G. Zech, *Einführung in Statistik und Messwertanalyse für Physiker*, DESY (2006) (eBook deutsch, eBook english).
  - V. Blobel, E. Lormann, *Statistische und numerische Methoden der Datenanalyse*, DESY (2012) (Webseite).
  - R. J. Barlow, *Statistics: A Guide to the use of statistical methods in the physical sciences*, Wiley (1989) (KIT-Bibliothek).
  - W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical recipes*, Cambridge Univ. Press (2007) (Webseite).

- Skriptensammlung von Prof. G. Quast (Link):

1 Einführung und Grundlagen

1.1 Physik vs. Statistik

In der Physik studieren wir Naturgesetze und damit Gewissheiten. Was hat Physik also mit der Anwendung von Statistik zu tun?

Messung und Reproduzierbarkeit

- In den Naturwissenschaften werden Erkenntnisse durch **Experimente** gewonnen und daraus gezogene Schlussfolgerungen durch Experimente überprüft. Das Ergebnis eines Experiments ist eine **Beobachtung**.

- Eine auf Reproduzierbarkeit abzielende, objektive, wissenschaftliche Beobachtung bezeichnen wir als **Messung**.

- Jede (glaubwürdige) Messung besteht aus einem nachvollziehbaren Wert (mit Einheit) und einer Messunsicherheit als Aussage über die (abgeschätzte) **Reproduzierbarkeit** des Wertes.

Messunsicherheit

- Messunsicherheiten können verschiedene **Quellen** haben (Beispiel: Messung der Masse des Higgs Bosons):

$$\boxed{m_{\mathrm{H}} = (125.38 \pm 0.18\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.}))\;\mathrm{GeV}}$$

PLB 805 (2929) 135425

Ungenauigkeit der Messapperatur/-technik.

Unvollständige Kenntnis der Randbedingungen.

Immanent sind nur Wahrscheinlichkeitsaussagen möglich (z.B. QM).

Messunsicherheit

- Messunsicherheiten können verschiedene **Quellen** haben (Beispiel: Messung der Masse des Higgs Bosons):

$$m_{\mathrm{H}} = (125.38 \pm 0.18\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.}))\;\mathrm{GeV}$$

PLB 805 (2929) 135425

Ungenauigkeit der Messapperatur/-technik.

Eigenschaften der Messung

Unvollständige Kenntnis der Randbedingungen.

Immanent sind nur Wahr-scheinlichkeitsaussagen möglich (z.B. QM).

Messunsicherheit

- Messunsicherheiten können verschiedene **Quellen** haben (Beispiel: Messung der Masse des Higgs Bosons):

$$\boxed{m_{\mathrm{H}} = (125.38 \pm 0.18\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.}))\;\mathrm{GeV}}$$

PLB 805 (2929) 135425

Ungenauigkeit der Messapperatur/-technik.

Eigenschaften der Messung

Unvollständige Kenntnis der Randbedingungen.

Immanent sind nur Wahrscheinlichkeitsaussagen möglich (z.B. QM).

Eigenschaften des zu messenden Systems

Messunsicherheit

- Messunsicherheiten können verschiedene **Quellen** haben (Beispiel: Messung der Masse des Higgs Bosons):

$$\boxed{m_{\mathrm{H}} = (125.38 \pm 0.18\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.}))\,\mathrm{GeV}}$$

PLB 805 (2929) 135425

Ungenauigkeit der Messapperatur/-technik.

Eigenschaften der Messung

Unvollständige Kenntnis der Randbedingungen.

Immanent sind nur Wahr-scheinlichkeitsaussagen möglich (z.B. QM).

Eigenschaften des zu messenden Systems

- Unsicherheiten müssen so objektiv und realistisch wie möglich abschätzt werden.

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Messunsicherheit

- Messunsicherheiten können verschiedene **Quellen** haben (Beispiel: Messung der Masse des Higgs Bosons):

$$\boxed{m_{\mathrm{H}} = \left(125.38 \pm 0.18\,(\mathrm{stat.}) \pm 0.08\,(\mathrm{syst.})\right)\,\mathrm{GeV}}$$

PLB 805 (2929) 135425

Ungenauigkeit der Messapperatur/-technik.

Eigenschaften der Messung

Unvollständige Kenntnis der Randbedingungen.

Immanent sind nur Wahrscheinlichkeitsaussagen möglich (z.B. QM).

Eigenschaften des zu messenden Systems

- Unsicherheiten müssen so objektiv und realistisch wie möglich abschätzt werden.

- Manchmal aufgrund mathematischer Gesetzmäßigkeiten gut möglich (z.B. für „stat.“). Manchmal Bestandteil der Kunst, Erfahrung und **Glaubwürdigkeit** des Messenden (t.w. für „syst.“).

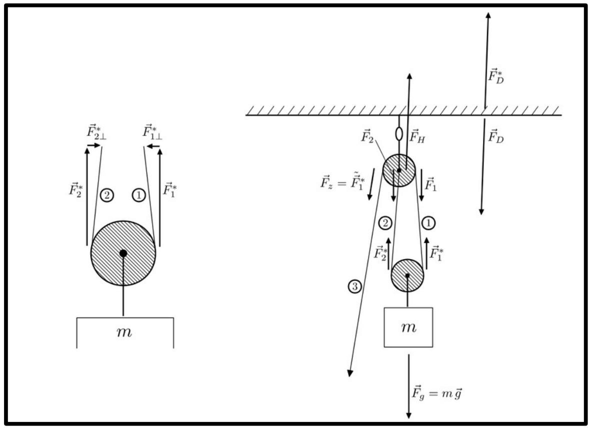
Zufallsverteilte Größen und Zufallsexperimente

- Eine Eigenschaft eines zu vermessenden Systems, die nicht mit exakter Genauigkeit vorhergesagt werden kann, bezeichnen wir als **zufallsverteilt**, die Größe als **Zufallsgröße**, die Messung selbst ist ein **Zufallsexperiment**.

- Wenn wir von Einzelbeobachtungen absehen, haben wir fast ausschließlich mit zufallsverteilten Größen zu tun.

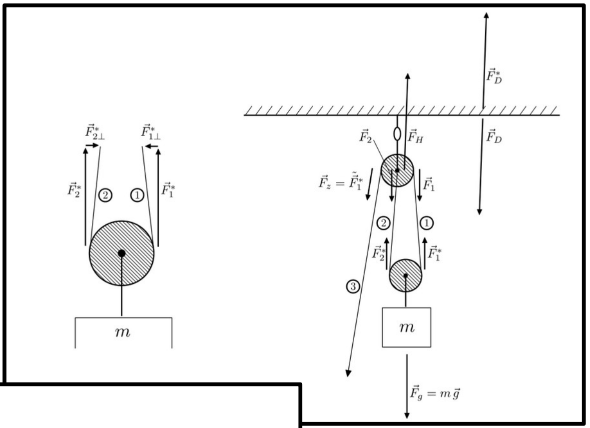
Vorhersagbarkeit und Zufall in der Physik

- **Klassische Physik:**
  - Dynamik idealisierter Körper grundsätzlich exakt vorhersagbar (→ Randwertprobleme).
  - Realität: (i) Vorhersagen i.a. nicht analytisch möglich (→ numerische Methoden); (ii) Randwerte nicht beliebig genau bestimmbar.

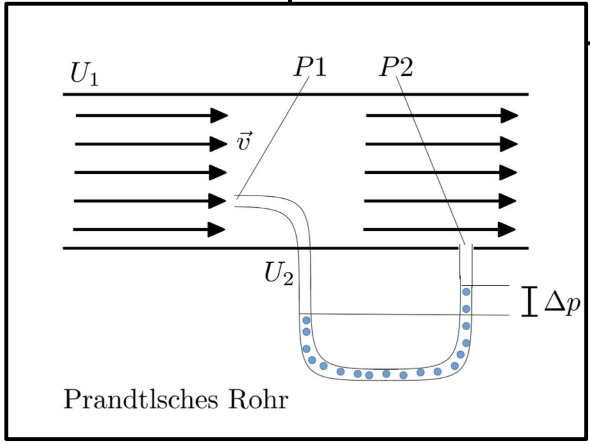


Vorhersagbarkeit und Zufall in der Physik

- **Klassische Physik:**
  - Dynamik idealisierter Körper grundsätzlich exakt vorhersagbar (→ Randwertprobleme).
  - Realität: (i) Vorhersagen i.a. nicht analytisch möglich (→ numerische Methoden); (ii) Randwerte nicht beliebig genau bestimmbar.

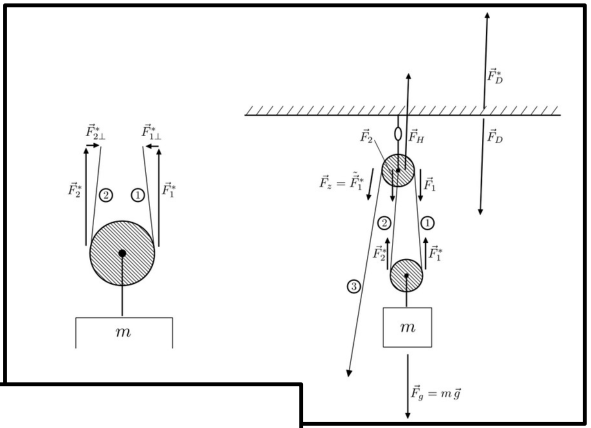


- **Thermodynamik/Statistische Physik:**
  - Eigenschaften ausgedehnter Systeme → statistische Kenngrößen großer Stichpro-ben/Ensembles.

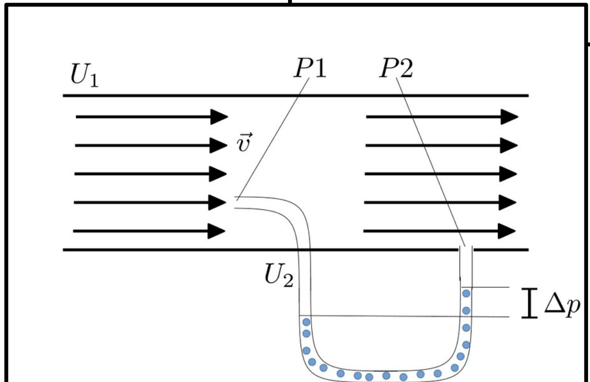


Vorhersagbarkeit und Zufall in der Physik

- **Klassische Physik:**
  - Dynamik idealisierter Körper grundsätzlich exakt vorhersagbar ($\rightarrow$ Randwertprobleme).
  - Realität: (i) Vorhersagen i.a. nicht analytisch möglich ($\rightarrow$ numerische Methoden); (ii) Randwerte nicht beliebig genau bestimmbar.

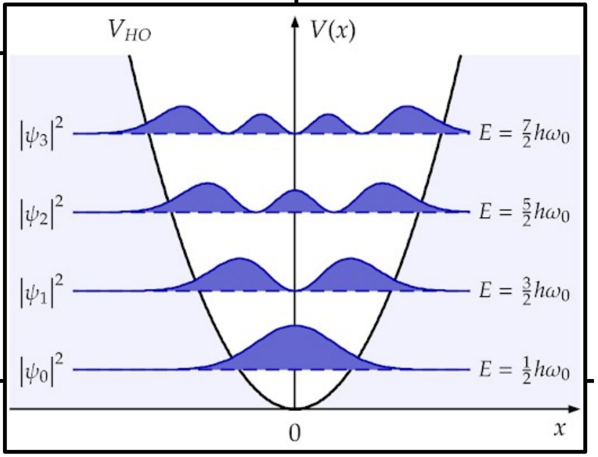


- **Thermodynamik/Statistische Physik:**
  - Eigenschaften ausgedehnter Systeme $\rightarrow$ statistische Kenngrößen großer Stichproben/Ensembles.



Prandtlsches Rohr

- **Quantenphysik:**
  - Nur Wahrscheinlichkeitsaussagen auf Grundlage der Wellenfunktion möglich.



Statistische Methoden in der Teilchenphysik

- **Im Experiment:**
  - Alle Gesetzmäßigkeiten beruhen auf expliziten Häufigkeitsmessungen.
  - Am LHC zeichnen wir Billionen unabhängiger Teilchenkollisionen auf.

- **In der Theorie:**
  - Explizite Interpretation der Wellenfunktion in der Quantenmechanik als Wahrscheinlichkeitsdichte.
  - Matrixelement → Wahrscheinlichkeit Endzustand f aus Anfangszustand i vorzufinden.
  - Prozesse setzen sich aus statistisch unabhängigen Unterprozessen zusammen (Partondichtefunktion → Matrixelement → Hadronisierung → erlaubt *event-by-event* Simulation).

<u>Perfektes Anwendungsgebiet</u> statistischer Methoden!

1 Einführung und Grundlagen

1.2 Statistik und Wahrscheinlichkeitsrechnung

Die mathematische Disziplin der Wahrscheinlichkeitsrechnung hat ihren Ursprung in der Formalisierung des Glücksspiels zur Zeit der Emanzipation des Geistes in der beginnenden Renaissance.

Quantifizierbarkeit des Zufalls

- Das **Maß an Zufall** läßt sich durch das Konzept der Wahrscheinlichkeit quantifizieren.
- Die Anfänge der Wahrscheinlichkeitstheorie (Stochastik) gehen bis in die Renaissance zurück.(1)
- Der moderne Wahrscheinlichkeitsbegriff begründet sich auf die mengentheoretische Formulierung von Kolmogorov (1933).(2)

(1)
Luca Pacioli: *Summa de Arithmetica Geometrica et Proportionality* (1491).

(2)
Andrei Nikolajewitsch Kolmogorov (1903 – 1987).

Ergebnis- & Ereignisraum

Eine Menge $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ heißt Ergebnisraum eines Zufallsexperiments, wenn jedem Ausgang des Experiments höchstens ein Element $\omega_i \in \Omega$ zugeordnet ist. Die ω_i heißen Ergebnisse des Zufallsexperiments.

- **Beispiele:**
  - Münzwurf $\Omega = \{\mathrm{Kopf}, \mathrm{Zahl}\}$.
  - Würfelwurf $\Omega = \{1, 2, 3, 4, 5, 6\}$.
  - Bestimmung der Fallbeschleunigung ($\rightarrow$ wie lautet der Ergebnisraum?).
  - Bestimmung der Anzahl radioaktiver Atome nach einer Zeitspanne Δt ($\rightarrow$ wie lautet der Ergebnisraum?).

Jede Teilmenge $A \subset \Omega$ heißt Ereignis. A tritt genau dann ein, wenn sich ein Ergebnis $\omega_i \in A$ einstellt, das in A enthalten ist. Die Menge aller Ereignisse (d.h. die Potenzmenge über Ω), $\mathfrak{P}(\Omega)$ heißt Ereignisraum.

Wahrscheinlichkeitverteilung (nach Kolmogorov)

Eine auf dem Ereignisraum $\mathfrak{P}(\Omega)$ definierte Funktion

$$P\,:\,\mathfrak{P}(\Omega) \to \mathbb{R}\,; \qquad A \to P(A),$$

heißt Wahrscheinlichkeitsverteilung über dem Ergebnisraum Ω, wenn sie die folgenden Eigenschaften erfüllt:

- Für jedes Ereignis $A \in \mathfrak{P}(\Omega)$ gilt $P(A) \geq 0$ (Nichtnegativität);
- Für die Wahrscheinlichkeit zweier disjunkter Ereignisse ($A \cap B = \emptyset$) gilt $P(A \cup B) = P(A) + P(B)$ (Linearität);
- Die Wahrscheinlichkeit für das Eintreffen irgendeines Ereignisses ist 1 (Normierungsbedingung).

Wahrscheinlichkeitverteilung (nach Kolmogorov)

- Folgerungen (ohne Beweis, aber einfach zu beweisen):

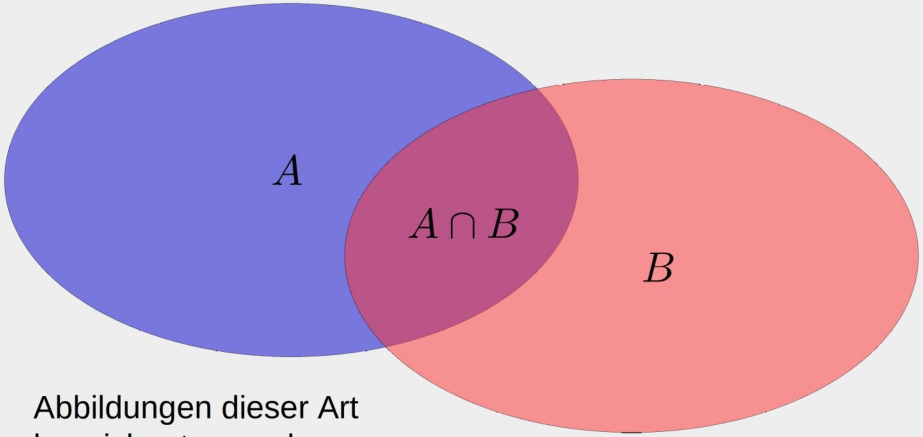
$$P(\bar{A}) = 1 - P(A) \text{ (Komplementäres Ereignis)}$$

$$P(\emptyset) = 0 \text{ (Unmögliches Ereignis)}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$A \subset B \text{ impliziert } P(A) \leq P(B)$$

$$P(A) \in [0;1]$$



Abbildungen dieser Art bezeichnet man als Venn-Diagramm.

Bedingte Wahrscheinlichkeit

Sei P eine Wahrscheinlichkeitsverteilung über dem Ergebniraum Ω, B ein Ereignis mit $P(B) > 0$ und A ein beliebiges Ereignis, dann heißt

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}$$

die Wahrscheinlichkeit von A unter der Bedingung B.

- $P(A|B)$ ist selbst eine Wahrscheinlichkeitsverteilung über Ω (die den Axiomen von Kolmogorov genügt).
- $P(A)$ wiederum kann als bedingte Wahrscheinlichkeit $P(A|S)$ betrachtet werden mit $S = \mathfrak{P}(\Omega)$ und $P(S) = 1$.
- $P(A \cap B) = P(A|B)\,P(B)$

Satz von Bayes

Seien P eine Wahrscheinlichkeitsverteilung über dem Ergebniraum Ω und A und B Ereignisse mit $P(A) > 0$ und $P(B) > 0$, dann erhält man $P(B|A)$ aus $P(A|B)$ durch die Umformung

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)}\,.$$

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \qquad \boxed{P(B|A) = \frac{P(B \cap A)}{P(A)}}$$

$$P(A \cap B) = P(B \cap A)$$

$$P(A|B)\,P(B) = P(B|A)\,P(A)$$

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)}$$

Thomas Bayes (1702 – 1761)

Satz von der totalen Wahrscheinlichkeit

Bilden die Ergeinisse $A_1,\ A_2,\ \ldots,\ A_n$ mit $P(A_i) > 0\ \forall i$ eine Zerlegung des Ergebniraums Ω (mit den Eigenschaften $\cup_i A_i = \Omega$ und $A_i \cap A_j = \emptyset, \forall i \neq j$), dann gilt für die Wahrscheinlichkeit eines beliebigen Ereignisses B

$$P(B) = \sum_{i=1}^{n} P(B|A_i)\,P(A_i)\,.$$

- Oft sieht man den Satz von der totalen Wahrscheinlichkeit auch in einer Formulierung in Kombination mit dem Satz von Bayes:

$$P(A_j|B) = \frac{P(B|A_j)\,P(A_j)}{\sum\limits_i P(B|A_i)\,P(A_i)}$$

- Spezialfall $A_1 = A$ und $A_2 = \bar{A}$:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|\bar{A})\,P(\bar{A})}$$

Zurück zu Folie 6

Unabhängigkeit

Zwei Ereignisse A und B heißen stochastisch unabhängig, wenn gilt

$$P(A \cap B) = P(A)\,P(B)$$

und stochastisch abhängig sonst.

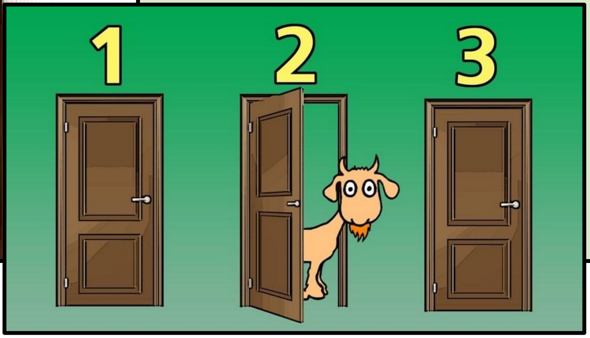
- Im Rahmen bedingter Wahrscheinlichkeiten bekommt (stochastische) Unabhängikeit eine anschauliche Bedeutung:

  $$P(B|A) = P(B)\,; \quad P(A|B) = P(A)$$

  d.h. die Wahrscheinlichkeit für das Eintreten von Ereignis $B\,(A)$ hängt nicht vom Eintreten des Ereignisses $A\,(B)$ ab.

Geh auf's Ganze (a.k.a. Ziegenproblem)

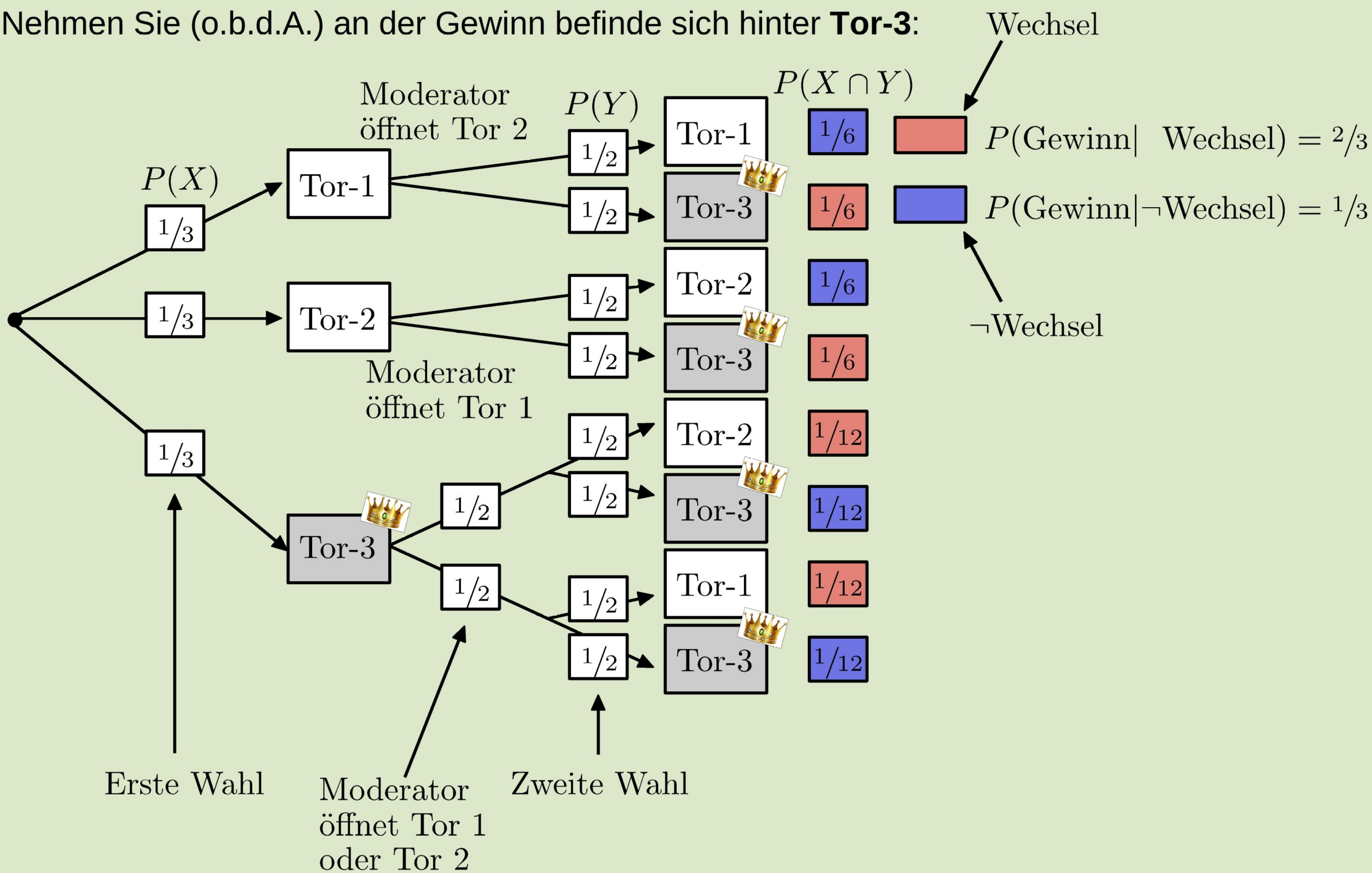
- Beispiel aus einer fürchterlichen Fernsehshow der 80er Jahre:



- **Spielidee:**
  - Drei verschlossene Tore. Hinter einem befindet sich der Gewinn.
  - Nachdem der Kandidat ein Tor gewählt hat öffnet der Moderator eines der <u>nicht</u> gewählten Tore.
  - Wecheln oder nicht wechseln?

Geh auf‘s Ganze (a.k.a. Ziegenproblem)

- Nehmen Sie (o.b.d.A.) an der Gewinn befinde sich hinter **Tor-3**:



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Bedingte Wahrscheinlichkeit und COVID

- Zahl an COVID-19 Erkrankungen in Deutschland (Stand 09.11.2020):

$$P(\mathrm{COVID}) = \frac{682\,624}{83\,020\,000} = 0,82\% \qquad P(\neg\mathrm{COVID}) = 1 - P(\mathrm{COVID}) = 99,18\%$$

- Zuverlässigkeit eines COVID-19 Tests (Stand 09.11.2020):

$$P(+|\mathrm{COVID}) = 95\% \qquad P(+|\neg\mathrm{COVID}) = 10\%$$

Sie unterziehen sich einem Test. Dieser <u>Test fällt positiv aus</u>. Sind Sie krank?

Bedingte Wahrscheinlichkeit und COVID

- Zahl an COVID-19 Erkrankungen in Deutschland (Stand 09.11.2020):

$$P(\mathrm{COVID}) = \frac{682\,624}{83\,020\,000} = 0,82\% \qquad P(\neg\mathrm{COVID}) = 1 - P(\mathrm{COVID}) = 99,18\%$$

- Zuverlässigkeit eines COVID-19 Tests (Stand 09.11.2020):

$$P(+|\mathrm{COVID}) = 95\% \qquad P(+|\neg\mathrm{COVID}) = 10\%$$

Sie unterziehen sich einem Test. Dieser <u>Test fällt positiv aus</u>. Sind Sie krank?

$$\begin{aligned} P(\mathrm{COVID}|+) &= \frac{P(+|\mathrm{COVID})\,P(\mathrm{COVID})}{P(+|\mathrm{COVID})\,P(\mathrm{COVID}) + P(+|\neg\mathrm{COVID})\,P(\neg\mathrm{COVID})} \\ &= \frac{0.95 \cdot 0.0082}{0.95 \cdot 0.0082 + 0.10 \cdot 0.9912} = 0.073 \end{aligned}$$

Bedingte Wahrscheinlichkeit und COVID

- Zahl an COVID-19 Erkrankungen in Deutschland (Stand 09.11.2020):

$$P(\mathrm{COVID}) = \frac{682\,624}{83\,020\,000} = 0{,}82\% \qquad P(\neg\mathrm{COVID}) = 1 - P(\mathrm{COVID}) = 99{,}18\%$$

- Zuverlässigkeit eines COVID-19 Tests (Stand 09.11.2020):

$$P(+|\mathrm{COVID}) = 95\% \qquad P(+|\neg\mathrm{COVID}) = 10\%$$

Sie unterziehen sich einem Test. Dieser <u>Test fällt positiv aus</u>. Sind Sie krank?

Sie unterziehen sich einem weiteren Test. Dieser <u>Test fällt wieder positiv aus</u>. Sind Sie krank?

Bedingte Wahrscheinlichkeit und COVID

- Zahl an COVID-19 Erkrankungen in Deutschland (Stand 09.11.2020):

$$P(\mathrm{COVID}) = \frac{682\,624}{83\,020\,000} = 0,82\% \qquad P(\neg\mathrm{COVID}) = 1 - P(\mathrm{COVID}) = 99,18\%$$

- Zuverlässigkeit eines COVID-19 Tests (Stand 09.11.2020):

$$P(+|\mathrm{COVID}) = 95\% \qquad P(+|\neg\mathrm{COVID}) = 10\%$$

Sie unterziehen sich einem Test. Dieser <u>Test fällt positiv aus</u>. Sind Sie krank?

Sie unterziehen sich einem weiteren Test. Dieser <u>Test fällt wieder positiv aus</u>. Sind Sie krank?

$$P(\mathrm{COVID}|+) = \frac{P(+|\mathrm{COVID})\,P(\mathrm{COVID})}{P(+|\mathrm{COVID})\,P(\mathrm{COVID}) + P(+|\neg\mathrm{COVID})\,P(\neg\mathrm{COVID})}$$

$$= \frac{0.95 \cdot 0.073}{0.95 \cdot 0.073 + 0.10 \cdot 0.927} = 0.428$$

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Interpretation von Wahrscheinlichkeiten

- Viele Funktionen genügen den Axiomen von Kolmogorov. Es obliegt jedoch dem Anwender Ω und P geeignet zu **interpretieren**.

- **Zwei Hauptschulen:**

Frequentistische (klassische)
Wahrscheinlichkeitsinterpretation

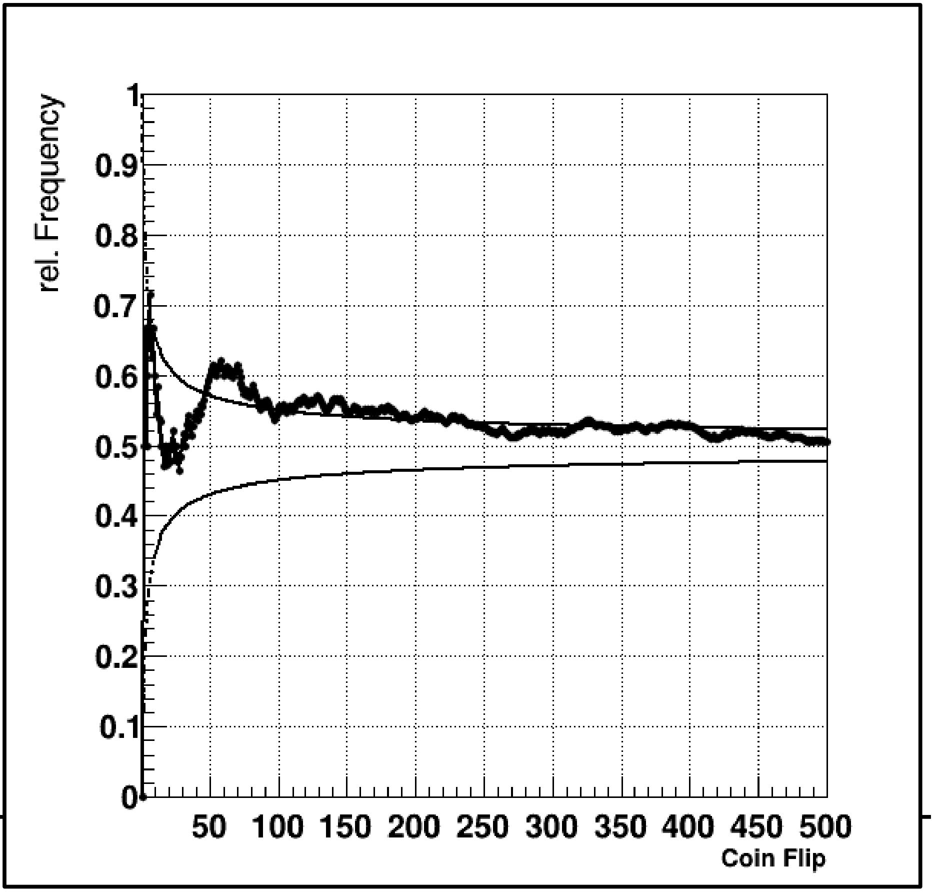
Bayesianische (subjektive)
Wahrscheinlichkeitsinterpretation

Zurück zu Folie 6

Frequentistische Wahrscheinlichkeitsinterpretation

- **Annahme:** Ein wahrer Wert existiert, der im Grenzwert einer erschöpfenden Stichprobe ermittelt werden kann: $$P(A) = \lim_{N\to|\mathfrak{P}(\Omega)|} \frac{A}{N}$$

- **Beispiel:** Relative Häufigkeit für „Kopf“ bei 500-fachem Münzwurf. Annahme: die wahre Wahrscheinlichkeit „Kopf“ zu erzielen ist 0.5 (erschöpfend wäre hier $n \to \infty$).



Zurück zu Folie 6

Frequentistische Wahrscheinlichkeitsinterpretation

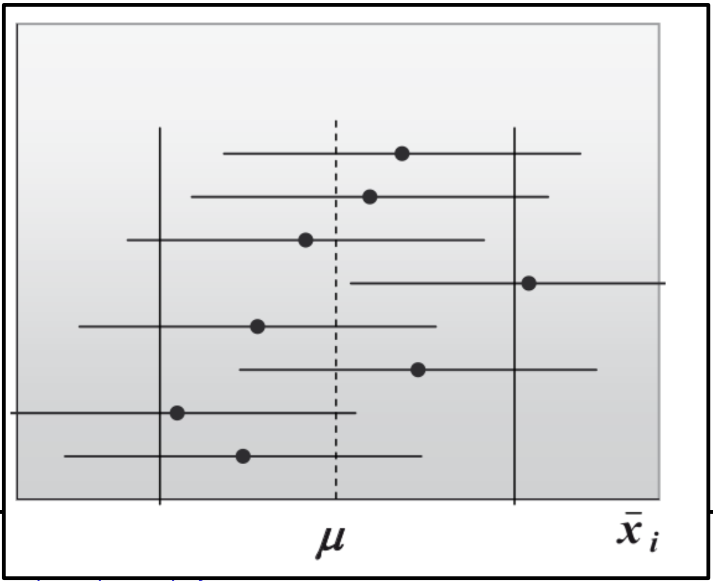
- **Annahme:** Ein wahrer Wert existiert, der im Grenzwert einer erschöpfenden Stichprobe ermittelt werden kann:

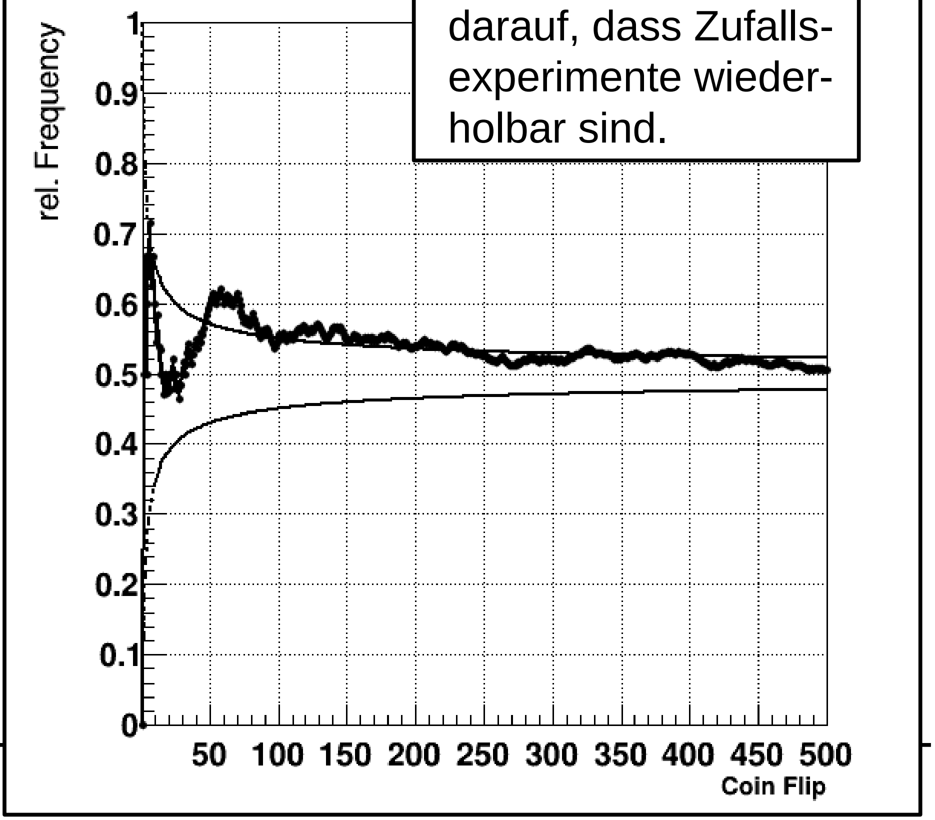
$$P(A) = \lim_{N \to |\mathfrak{P}(\Omega)|} \frac{A}{N}$$

- **Beispiel:** Relative Häufigkeit für „Kopf“ bei 500-fachem Münzwurf. Annahme: die wahre Wahrscheinlichkeit „Kopf“ zu erzielen ist 0.5 (erschöpfend wäre hier $n \to \infty$).

Problem:
Interpretation beruht darauf, dass Zufallsexperimente wiederholbar sind.

- Wahrer Wert und/oder Messung ansich haben keine Unsicherheit ($\rightarrow$ Gewissheiten).
- Angegebene Fehlerbalken der Messung überdecken wahren Wert (bei vielfacher Wiederholung) im Anteil α aller Fälle ($\rightarrow$ Repräsentationsschluss, i.a. $\alpha = 68\%$).





Bayesianische Wahrscheinlichkeitsinterpretation

- In diesem Fall werden die Elemente des Ergebnisraums als Hypothesen bezeichnet. Der Ereignisraum $\mathfrak{P}(\Omega)$ heißt Hypothesenraum.

- $P(A)$ ist ein Maß für das **subjektive Fürwahrhalten** der Hypothese A .

- Zentrale Gesetzmäßigkeit in diesem Paradigma $\rightarrow$ Satz von Bayes:

Hypothese wird mit Messung abgeglichen:

$$P(\mathrm{Hypothese}|\mathrm{Messung}) \propto P(\mathrm{Messung}|\mathrm{Hypothese})\, P(\mathrm{Hypothese})$$

Posteriori Wahrschein- lichkeit (**Posterior**)

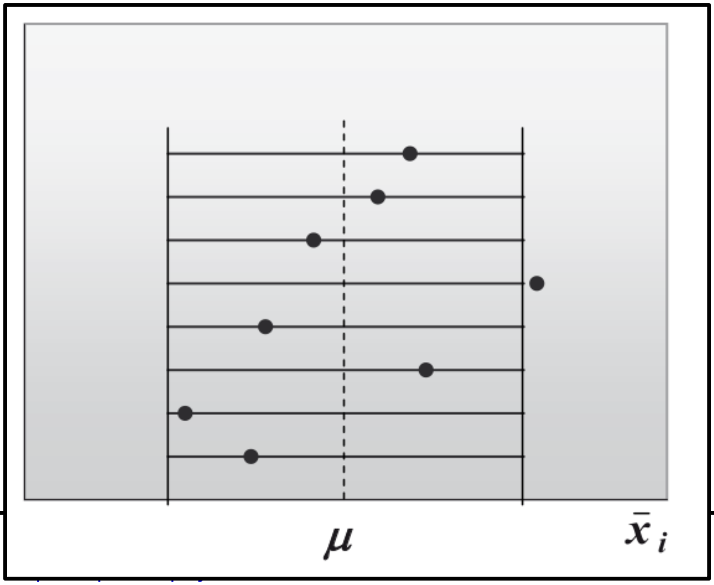
Likelihood

A priori Wahr- scheinlichkeit (**Prior**)

Bayesianische Wahrscheinlichkeitsinterpretation

- In diesem Fall werden die Elemente des Ergebnisraums als Hypothesen bezeichnet. Der Ereignisraum $\mathfrak{P}(\Omega)$ heißt Hypothesenraum.
- $P(A)$ ist ein Maß für das **subjektive Fürwahrhalten** der Hypothese A .

- Messung selbst ist die (neue) „Wahrheit“, mit Unsicherheit behaftet ($\rightarrow$ Maß des Fürwahrhaltens).
- Anteil α aller weiteren Messungen würde in dem durch den Fehlerbalken angegebenen Intervall liegen ($\rightarrow$ Inklusionsschluss).



Problem:

- Konstruktion des Hypothesenraums, so dass er Wahrscheinlichkeitsdefinition erfüllt.
- Keine Aussage über Prior.
- **NB:** bei vorgegebenem Prior Aussage über Prior im Lichte der Messung.

Ein Beispiel für einen aktualisierten Prior finden Sie im Backup zu dieser Vorlesung.

Erkenntnisgewinn in der Moderne

- Seit der Frühmoderne geleitet von mathematischen Modellen:

Konstruktion des Modells

Erklärungskraft

Experiment:

Eintreten eines <u>realen</u> Ereignisses A

Modell:

Modellraum, Modell-ereignis A'

Überprüfung

Vorhersage

Eintreten des realen Er-eignisses B

Mathematisches Resul-tat B'

Übertragung in die Realität

Vorhersagekraft

- Einfache Erklärung & Vorhersage reproduzierbarer Naturereignisse.

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

1 Einführung und Grundlagen

1.3 Wahrscheinlichkeitsverteilung/-dichte

Beim Übergang von Ereignisräumen mit endlicher Mächtigkeit zu Ereignisräumen mit unendlicher Mächtigkeit vollziehen wir den Übergunang von der Wahrscheinlichkeitsverteilung zur Wahrscheinlichkeitsdichte.

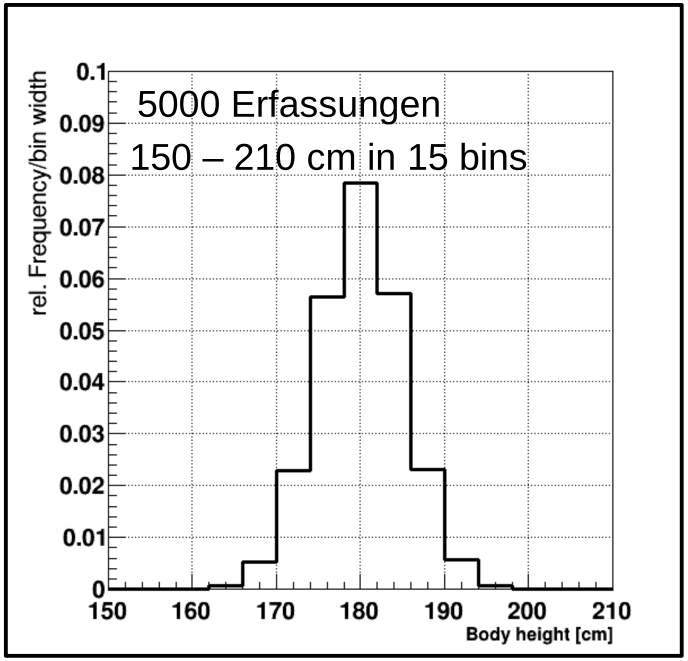
Achtung:
Wir werden beide Begriffe im folgenden äquivalent benutzen. Was gemeint ist sollte sich aus dem Zusammenhang erschließen.

Übergang zur Wahrscheinlichkeitsdichte

- **Beispiel:** Erfassung der Körpergröße x von 5000 Einwohnern über 18 Jahre in Karlsruhe.
  - Die Menge aller Einwohner in Karlsruhe würden wir als **Grundgesamtheit** bezeichnen.
  - Die Messung mit 5000 Probanden heißt **Stichprobe** (engl. *sample*). Das Ziehen von Stichproben bezeichnet man auch als *sampling*.
  - Die Körpergröße x ist eine (kontinuierlich verteilte) **Zufallsvariable**, deren gemessener Wert jeweils das Ergebnis eines **Zufallsexperiments** ist.

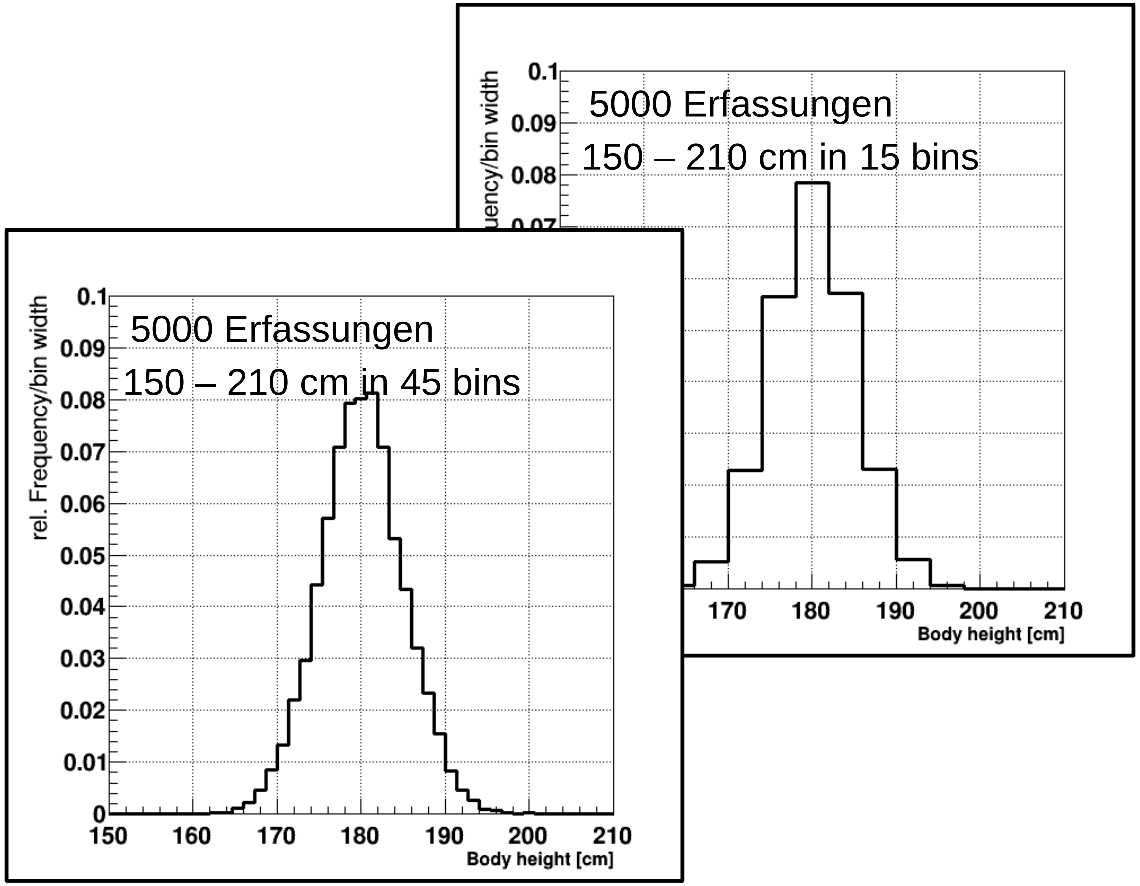
Übergang zur Wahrscheinlichkeitsdichte

- **Darstellung** durch Histogramm, dividiert durch Stichprobenlänge (=5000) und Binbreite.



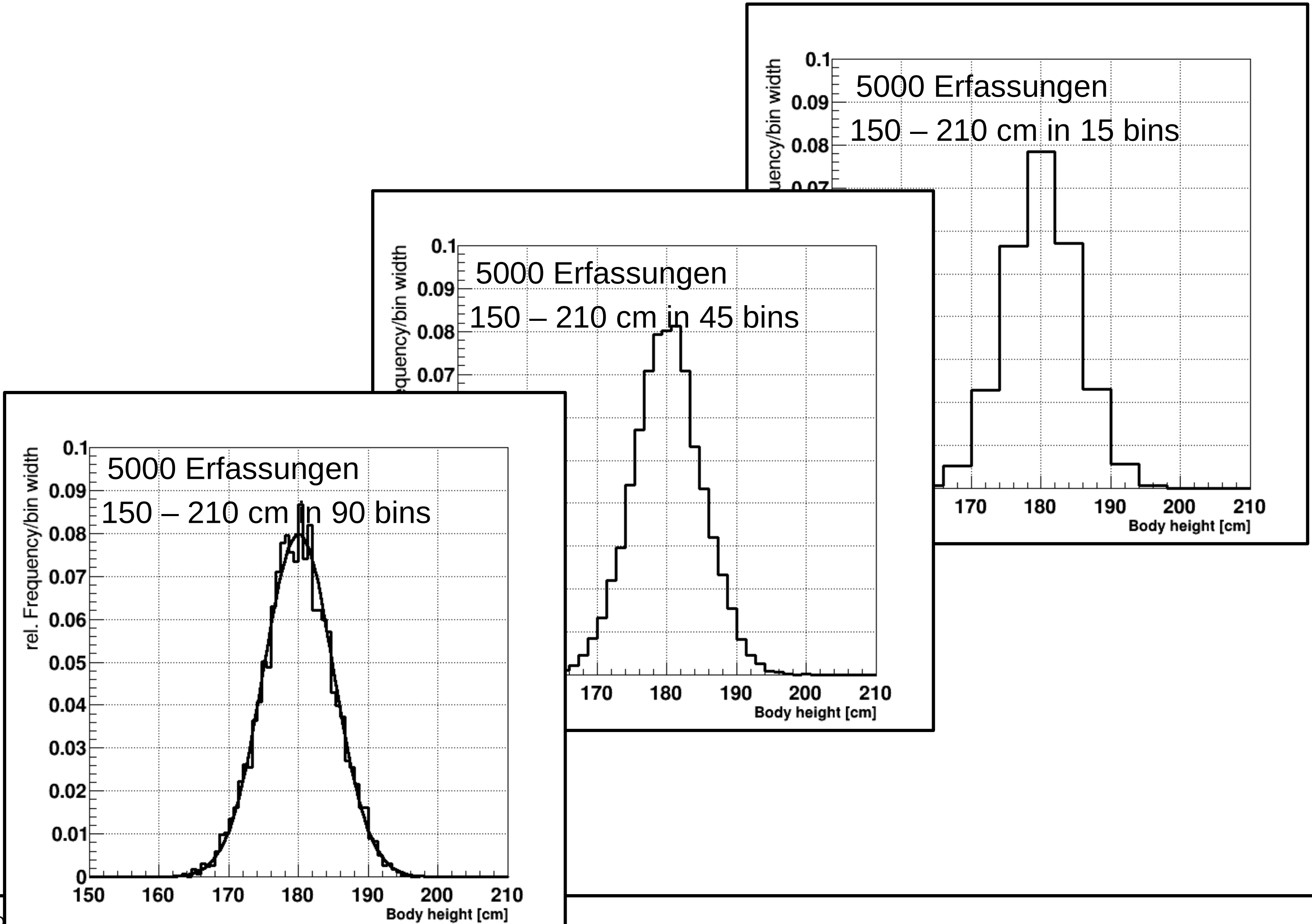
Übergang zur Wahrscheinlichkeitsdichte

- **Darstellung** durch Histogramm, dividiert durch Stichprobenlänge (=5000) und Binbreite.



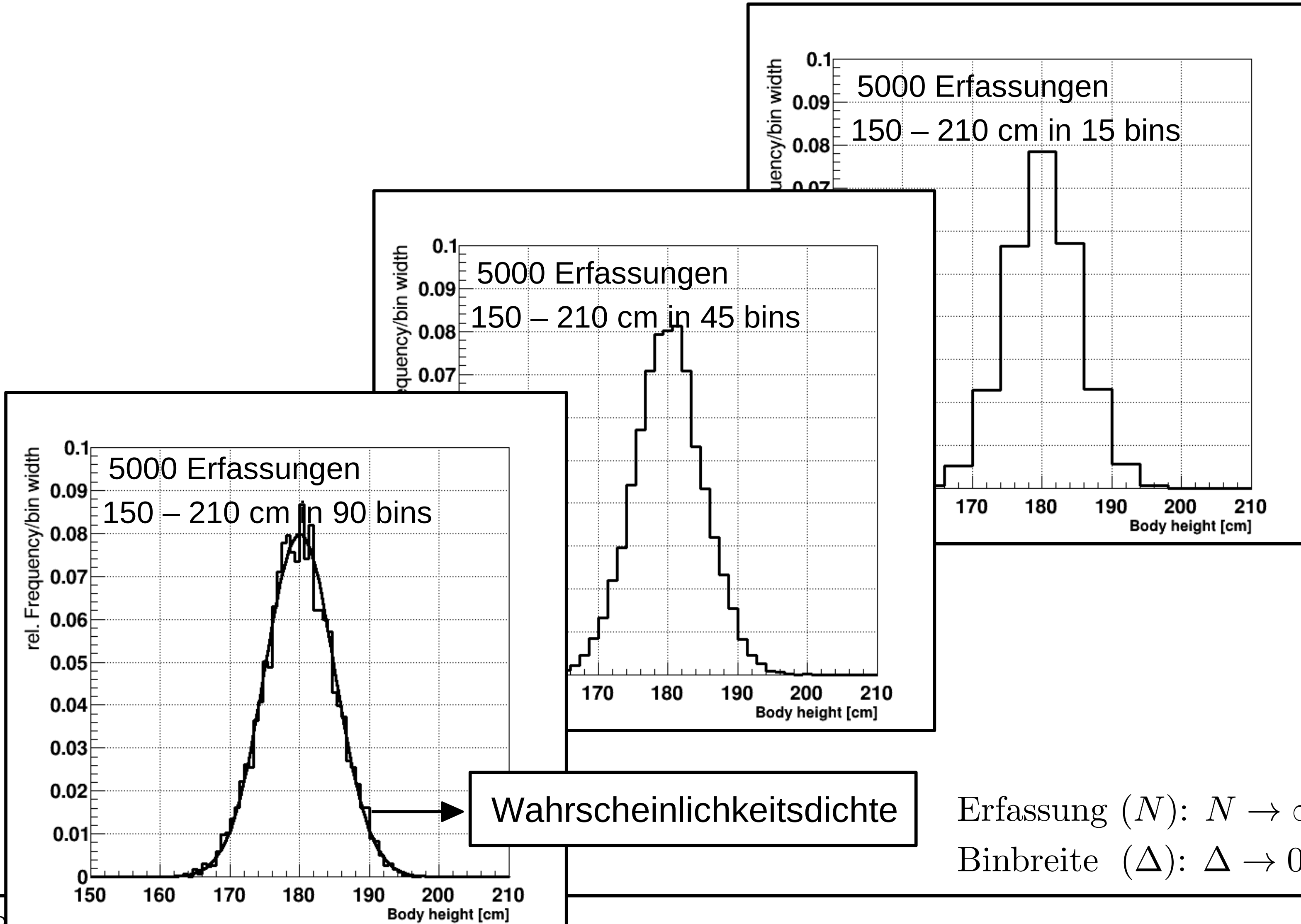
Übergang zur Wahrscheinlichkeitsdichte

- **Darstellung** durch Histogramm, dividiert durch Stichprobenlänge (=5000) und Binbreite.



Übergang zur Wahrscheinlichkeitsdichte

- **Darstellung** durch Histogramm, dividiert durch Stichprobenlänge (=5000) und Binbreite.



Erfassung (N): $N \to \infty$

Binbreite (Δ): $\Delta \to 0$

Zurück zu Folie 40

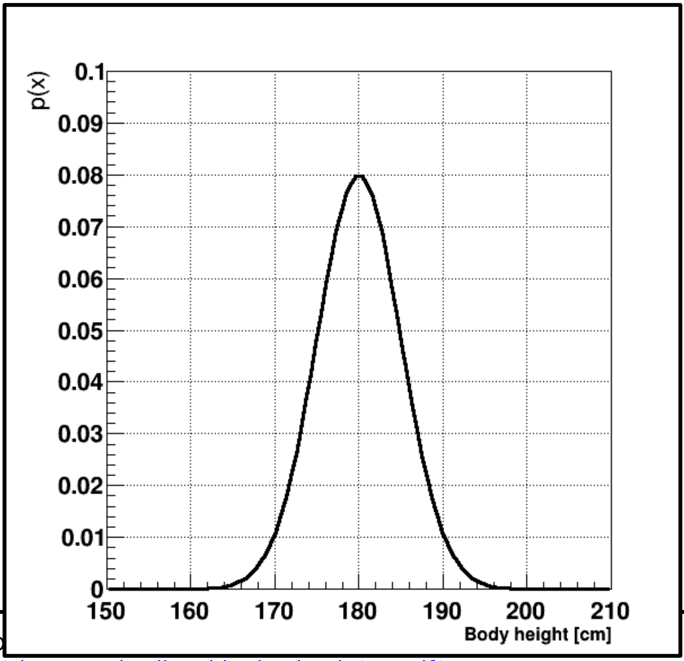
Wahrscheinlichkeitsdichte

Ist x eine kontinulierlich verteilte Zufallsvariable und die Wahrscheinlichkeitsverteilung $P(x)$ über dem Ergebnisraum Ω stetig differenzierbar, dann bezeichnen wir

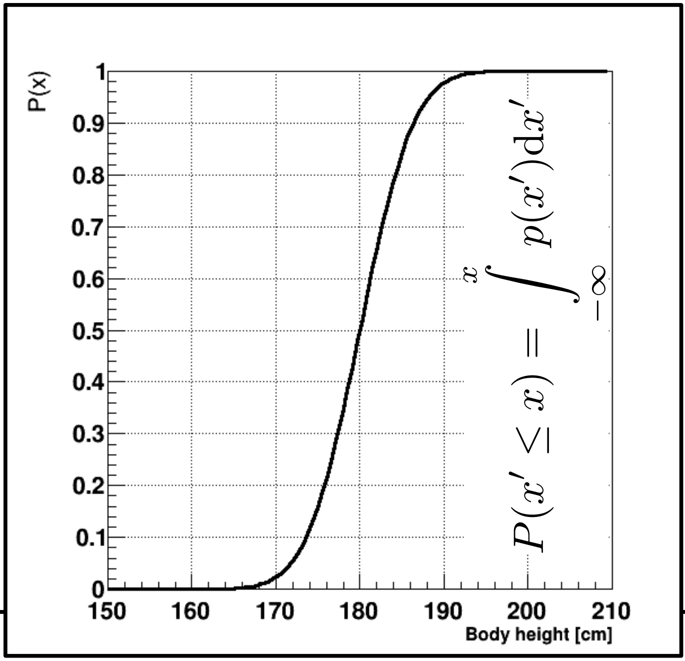
$$p(x) = \frac{\mathrm{d}P}{\mathrm{d}x}(x)$$

als Wahrscheinlichkeitsdichtefunktion von x.

Wahrscheinlichkeitsdichte (engl. *probability density function*, PDF)



Kumulative Wahrscheinlichkeitsfunktion (engl. *cumulative distribution function*, CDF)



Wahrscheinlichkeitsdichte – Wahrscheinlichkeitsverteilung

- Die Wahrscheinlichkeit für das Ereignis $\mathrm{Körpergröße} = (180 \pm 1)\mathrm{cm}$ erhalten Sie nicht aus $p(x)|_{x=180}$ sondern aus dem Integral:

$$\int\limits_{179}^{181} p(x')\mathrm{d}x' = P(x' \leq 181) - P(x' \leq 179)$$

- **NB:** Üblicherweise ist die Wahrscheinlichkeitsdichte einer Grundgesamtheit ($\rightarrow$ Einwohner in Karlsruhe) nicht bekannt! Sie muss angenommen oder aus einer Stichprobe, wie in unserem Beispiel, erschlossen werden (siehe auch Folie 47ff).

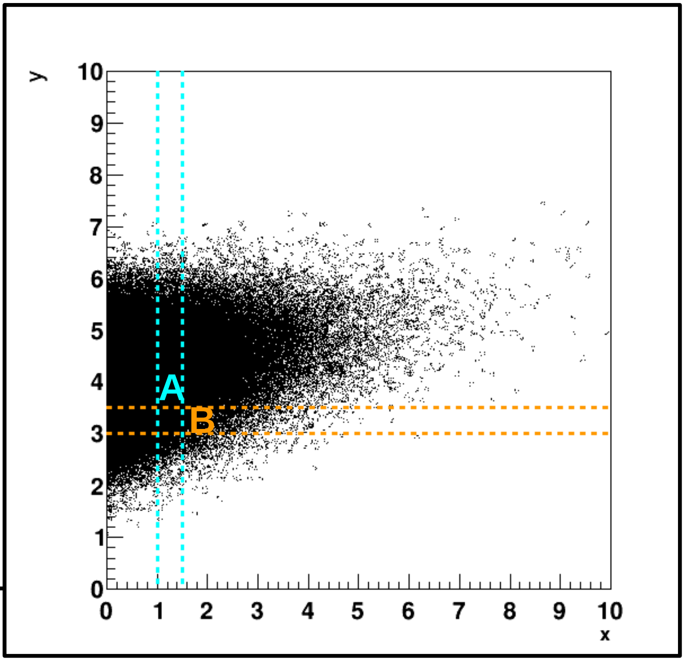
Mehrdimensionale Wahrscheinlichkeitsdichten

- Der Ausgang einer Messung kann durch mehrere Zufallsgrößen charakterisiert sein (z.B. Körpergröße und Alter). In diesem Fall ist auch die Wahrscheinlichkeitsverteilung mehrdimensional.

$$A : x \in [x_0, x_0 + \mathrm{d}x]$$
$$B : y \in [y_0, y_0 + \mathrm{d}y]$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x\mathrm{d}y = 1 \quad \text{(Normierung)} \qquad P(x' \leq x, y' \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(x',y')\mathrm{d}x'\mathrm{d}y' \quad \text{(CDF)}$$

$$\begin{aligned} P(A\cap B) = & \int_{x_0}^{x}\int_{y_0}^{y} p(x,y)\mathrm{d}x\mathrm{d}y \\ = & \quad P(x' \leq (x_0 + dx), y' \leq (y_0 + dy)) \\ & - P(x' \leq x_0, y' \leq (y_0 + \mathrm{d}y)) \\ & - P(x' \leq (x_0 + \mathrm{d}x), y' \leq y_0) \\ & + P(x' \leq x_0, y' \leq y_0) \end{aligned}$$

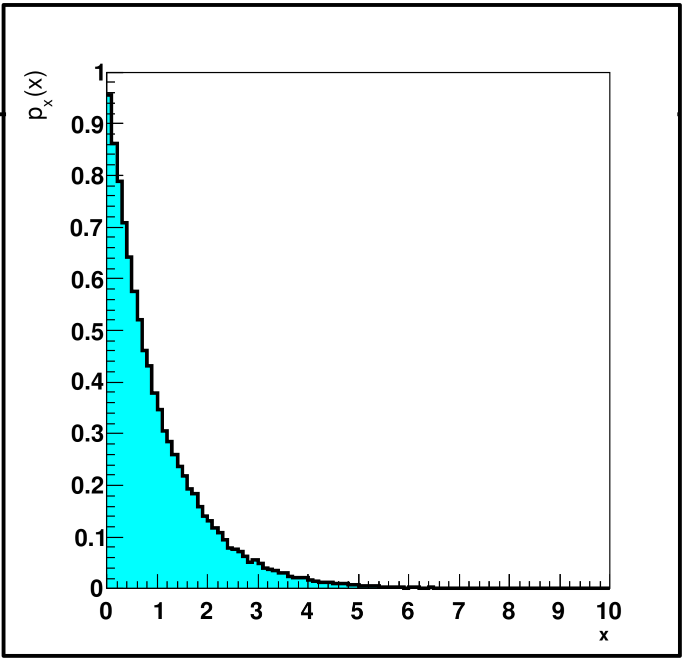


Randverteilungen

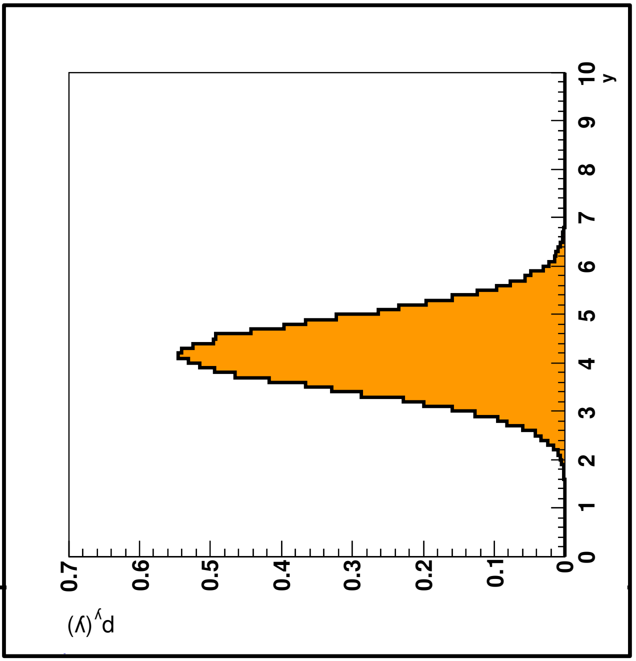
- Randverteilungen (engl. *marginal distributions*) erhält man durch Ausintegration einer oder mehrerer Zufallsvariablen (→ Marginalisieren).

$$p_x(x) = \int_{-\infty}^{+\infty} p(x,y)\,\mathrm{d}y \qquad P_x(x' \leq x) = \int_{-\infty}^{x} p_x(x')\,\mathrm{d}x'$$

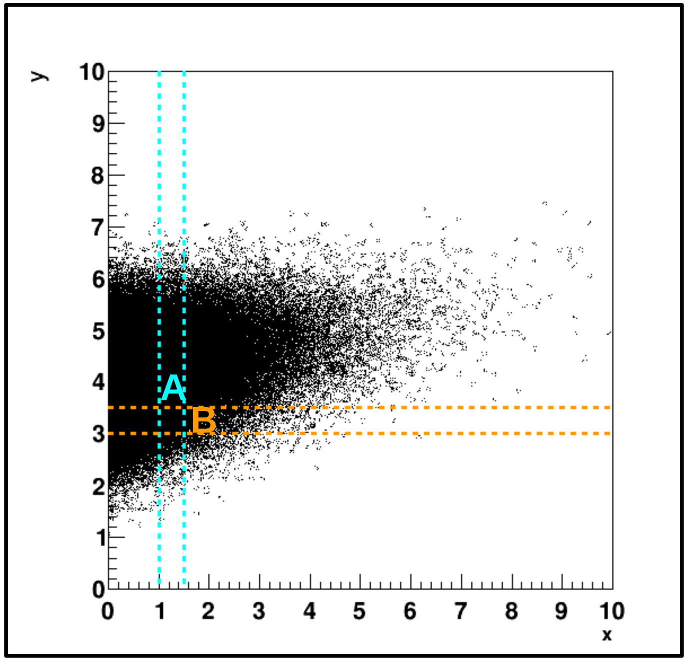
- Achtung: In der Physik sind Wahrscheinlichkeitsdichten üblicherweise (sehr) hochdimensional!



Projektion auf x-Achse



Projektion auf y-Achse



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de

Gängige Wahrscheinlichkeitsdichten

- Bei Stichprobenuntersuchungen ergeben sich einige Wahrscheinlichkeitsdichten aus fundamentalen Annahmen oder statistischen Gesetzmäßigkeiten. Sie sind daher von fundamentaler und universeller Natur. Die wichtigsten Beispiele sind:

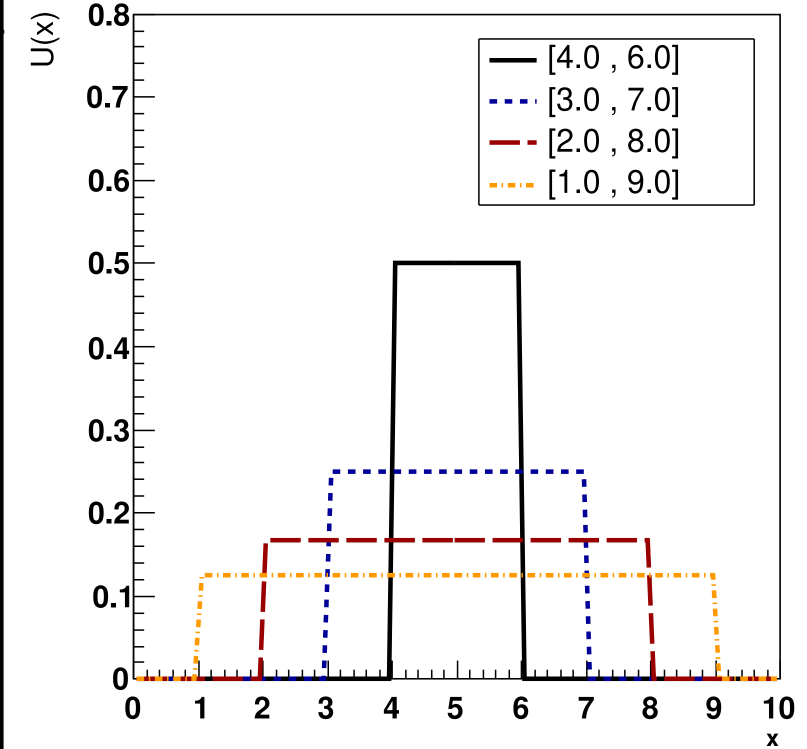
  - **Uniforme Verteilung**: → Alle Ergebnisse gleichwahrscheinlich!
  - **Exponentialverteilung**: → Wahrscheinlichkeit für das Eintreten eines Ergebnisses über Beobachtungszeit gleichverteilt.
  - **Binomialverteilung**: → Ein Ergebnis tritt mit fester Wahrscheinlichkeit ein.
  - **Poissonverteilung**: → Näherung für die Binomialverteilung falls diese Wahrscheinlichkeit (sehr) klein ist.
  - **Normalverteilung**: → Näherung für die Binomialverteilung für (sehr) große Stichproben (→ Zentraler Grenzwertsatz).

Uniforme Verteilung

$$U_{[a,b]}(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{sonst} \end{cases}$$

$$E[x] = {}^1\!/\!_2(a+b) \qquad \text{(Erwartungswert)}$$

$$\text{var}[x] = {}^1\!/\!_{12}(b-a)^2 \qquad \text{(Varianz)}$$



- **NB:**
  Jede beliebig verteilte Zufallsvariable x läßt sich auf eine uniform verteilte Zufallsvariable $a(x)$ mit $q(a) \equiv 1$ trans-formieren. Die Transformation ist:

$$a(x) = P(x' \le x)$$

$$\frac{\mathrm{d}a}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} \int_{-\infty}^{x} p(x')\mathrm{d}x' = p(x)$$

$$q(a) = p(x) \left| \frac{\mathrm{d}x}{\mathrm{d}a} \right| = p(x) \left| \frac{\mathrm{d}a}{\mathrm{d}x} \right|^{-1} \equiv 1 \quad (\text{für } 0 \le a \le 1)$$

- Jeder Beginn einer Mote Carlo Integration.
- Allgemeinster *prior* einer Bayesianischen *likelihood* Abschätzung.

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

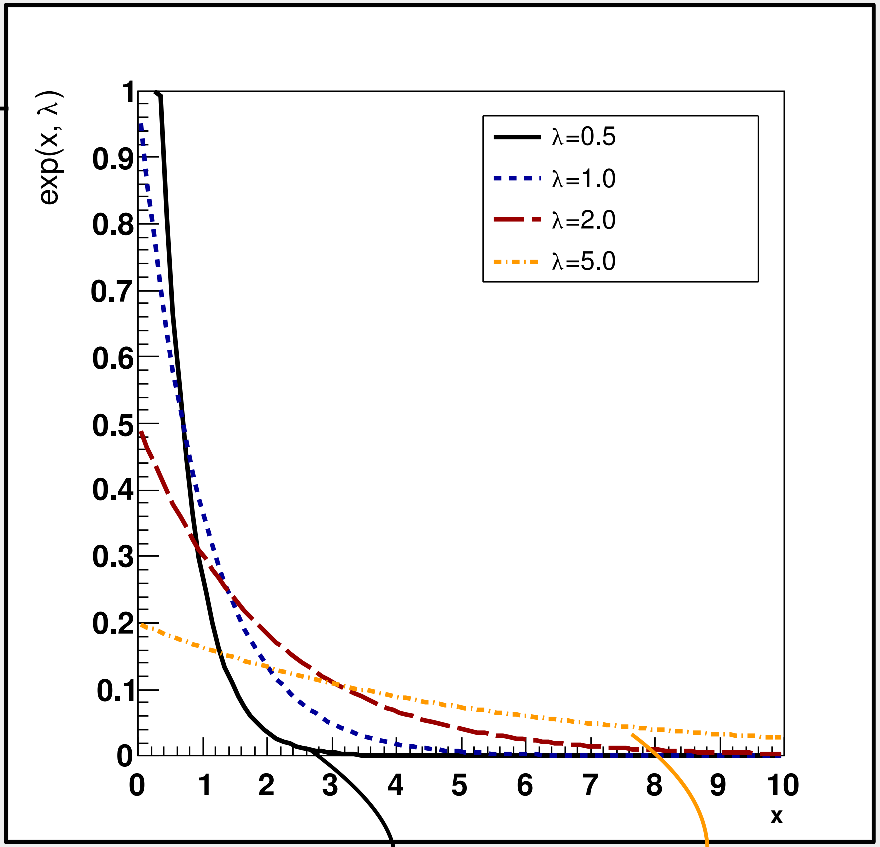
Exonentialverteilung

$$\exp(x,\lambda) = \frac{1}{\lambda}e^{-x/\lambda}$$

$$E[x] \quad = \lambda \qquad\qquad\qquad\qquad \text{(Erwartungswert)}$$

$$\text{var}[x] = \lambda^2 \qquad\qquad\qquad\qquad \text{(Varianz)}$$

- **NB:**
  Klassisches Beispiel: radioaktiver Zerfall.
  In diesem Bild entspricht λ der Lebens-
  dauer des Präparats.



λ klein

λ groß

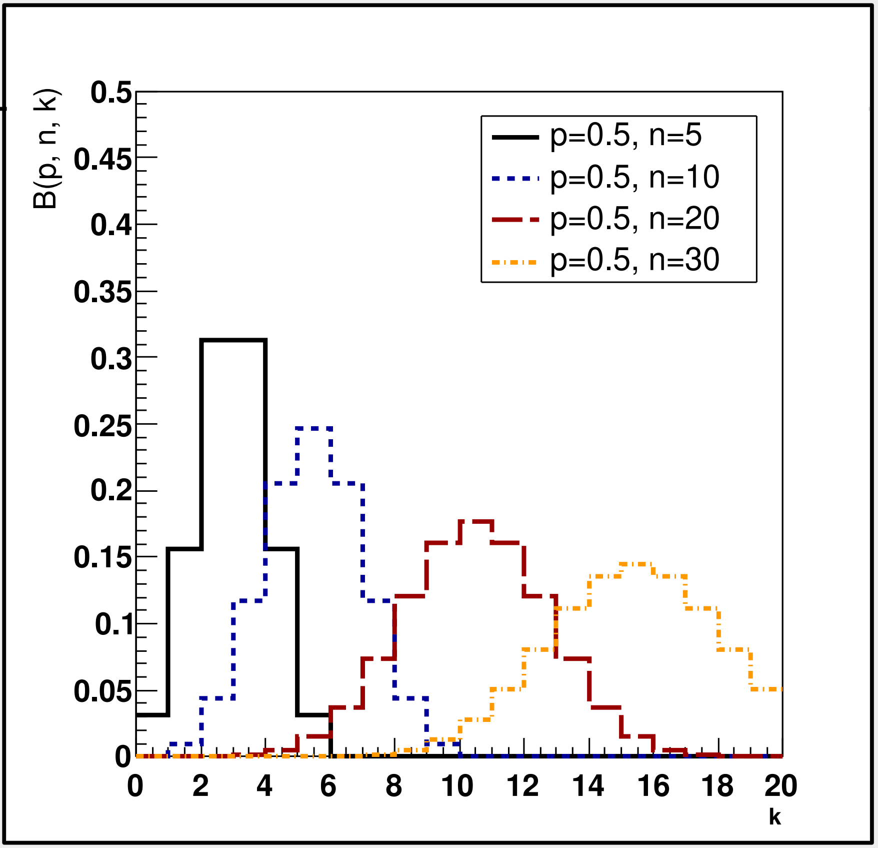
λ und damit die Zerfallswahr-
scheinlicheikt ist zu jedem
Zeitpunkt gleich.

Binomialverteilung

$$B(p,n,k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$$E[x] \quad = np \qquad\qquad (\text{Erwartungswert})$$

$$\text{var}[x] = np(1-p) \qquad (\text{Varianz})$$



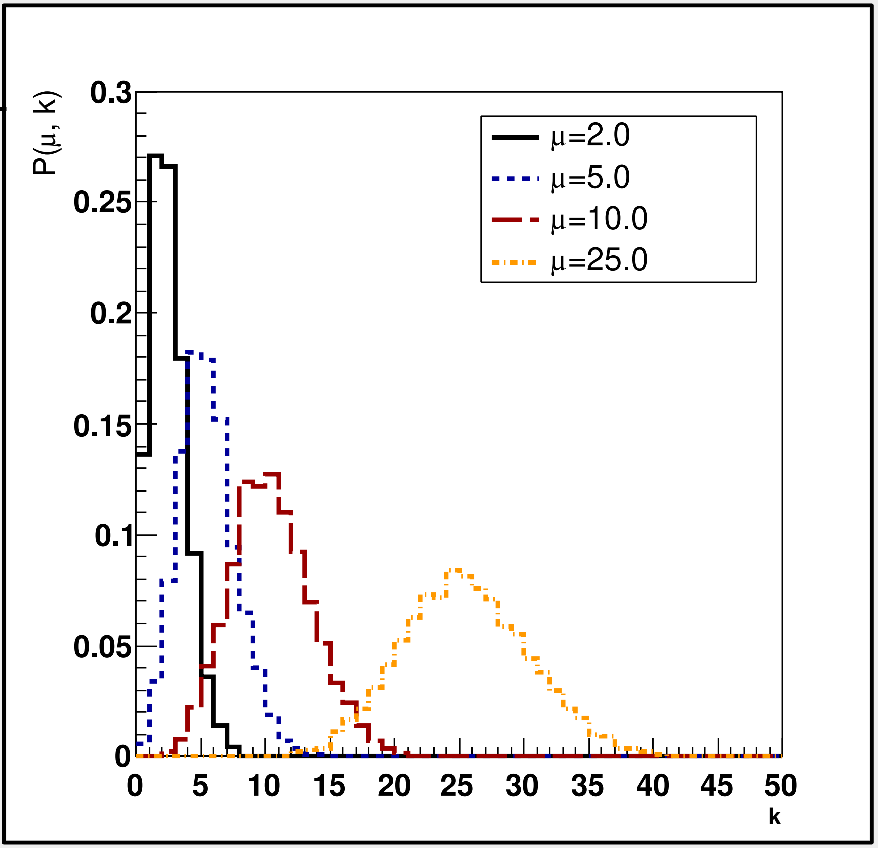
- **NB:**
  Verteilung von k nicht mehr gleich sondern bestimmt durch Wahrscheinlichkeit p.

Poissonverteilung

$$P(\mu, k) = \frac{\mu^k}{k!} e^{-\mu}$$

$$E[x] \ = np = \mu \qquad\qquad \text{(Erwartungswert)}$$

$$\mathrm{var}[x] = np = \mu \qquad\qquad \text{(Varianz)}$$



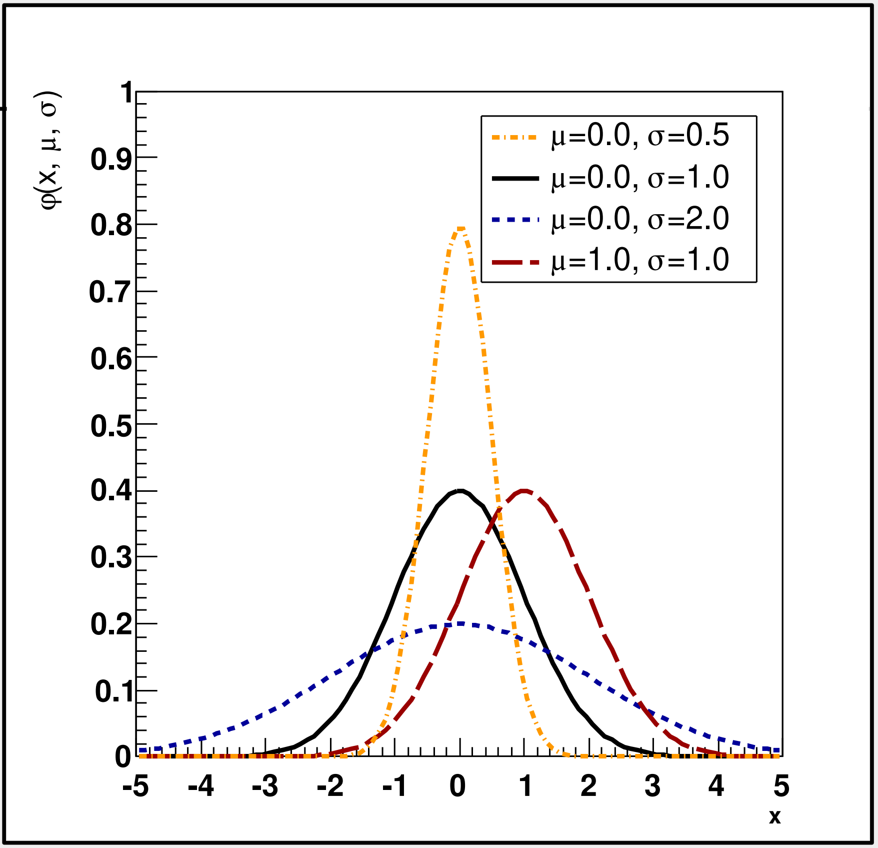
- **NB:** Binomialverteilung im Grenzwert $p \to 0$, $n \to \infty$.

Binomial- vs. Poissonverteilung

$$
\begin{aligned}
B(p,n,k) &= \begin{pmatrix} n \\ k \end{pmatrix} p^k \cdot (1-p)^{n-k} \\
&= \frac{n(n-1)(n-2)\cdot\ldots\cdot(n-k+1)}{k!} \cdot \frac{\mu^k}{n^k} \cdot \frac{\left(1-\frac{\mu}{n}\right)^n}{\left(1-\frac{\mu}{n}\right)^k} \\
&= \frac{1\cdot(1-\frac{1}{n})(1-\frac{2}{n})\cdot\ldots\cdot(1-\frac{k-1}{n})}{\left(1-\frac{\mu}{n}\right)^k} \cdot \frac{\mu^k}{k!} \cdot \left(1-\frac{\mu}{n}\right)^n \\
&= \underbrace{\frac{1}{\left(1-\frac{\mu}{n}\right)} \cdot \frac{(1-\frac{2}{n})}{\left(1-\frac{\mu}{n}\right)} \cdot \frac{(1-\frac{2}{n})}{\left(1-\frac{\mu}{n}\right)} \cdot\ldots\cdot \frac{(1-\frac{k-1}{n})}{\left(1-\frac{\mu}{n}\right)}}_{\longrightarrow 1} \cdot \frac{\mu^k}{k!} \cdot \underbrace{\left(1-\frac{\mu}{n}\right)^n}_{\longrightarrow e^{-\mu}} = \frac{\mu^k}{k!} e^{-\mu}
\end{aligned}
$$

für $\mu = const, n \to \infty$

Normalverteilung

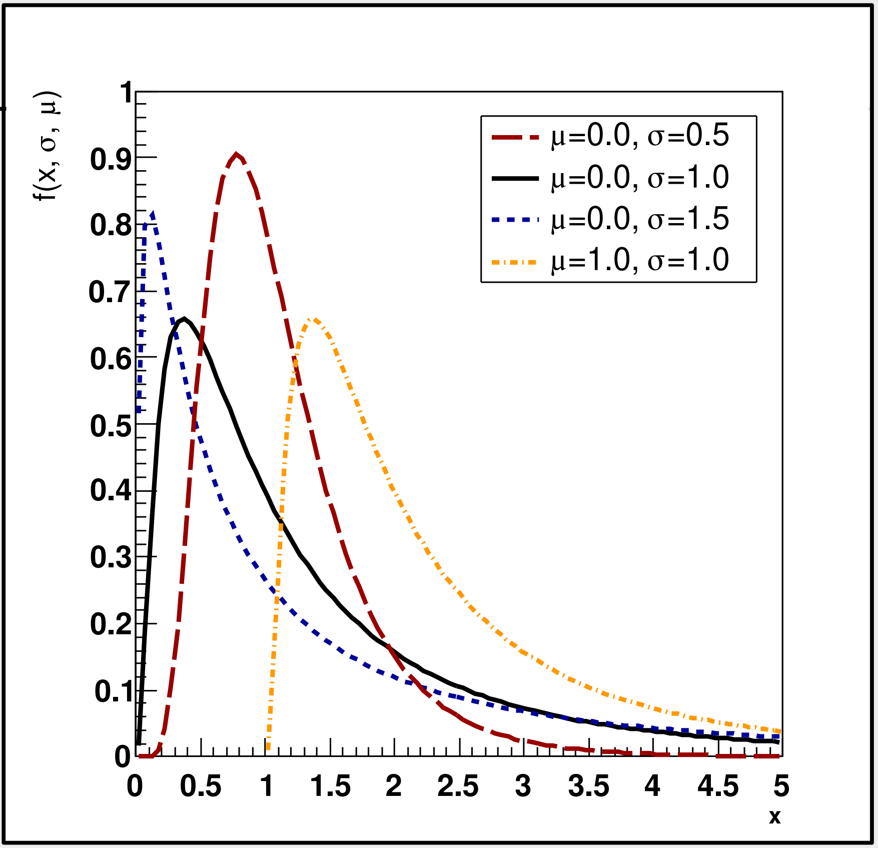


$$\varphi(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

$$E[x] \quad = \mu \qquad\qquad\qquad \text{(Erwartungswert)}$$

$$\mathrm{var}[x] = \sigma^2 \qquad\qquad\qquad \text{(Varianz)}$$

Log-Normalverteilung



$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\frac{1}{x}e^{-(\ln(x)-\mu)^2/2\sigma^2}$$

$$E[x] \;= e^{\mu+1/2\sigma^2} \qquad \text{(Erwartungswert)}$$

$$\mathrm{var}[x] = e^{2\mu+\sigma^2}\cdot\left(e^{\sigma^2}-1\right) \qquad \text{(Varianz)}$$

- **NB:**
  Die Variablensubstitution

  $$x' = \ln(x)\,; \qquad \mathrm{d}x' = \mathrm{d}\ln(x) = \frac{\mathrm{d}x}{x}$$

  führt die Log-Normalverteilung in die Normalverteilung über:

  $$f(x,\mu,\sigma)\mathrm{d}x = \varphi(x',\mu,\sigma)\mathrm{d}x'|_{x'=\ln(x)}$$

- Die multiplikative Überlagerung viele kleiner unabhängiger Zufallseffekte folgt also (näher-ungsweise) einer Log-Normalverteilung

χ^2-Verteilung

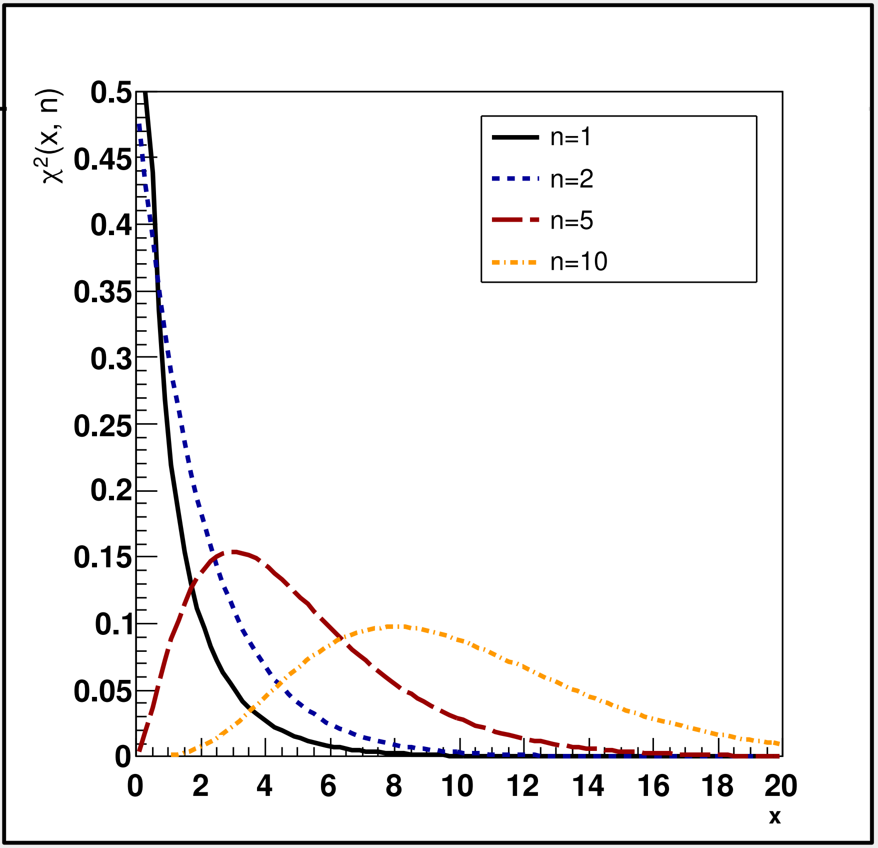
$$\chi^2(x,n) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}$$

mit:

$$\Gamma(x) = \int e^{-t} t^{x-1} \mathrm{d}t$$

$$E[x] \ = n \qquad \text{(Erwartungswert)}$$

$$\mathrm{var}[x] = 2n \qquad \text{(Varianz)}$$



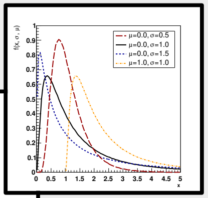
- **NB:**
  Die Summe der Quadrate von n normal-verteilten Zufallsgrößen x_i

$$z = \sum_{i=0}^{N} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$

ist χ^2-verteilt. Die Zahl n heißt Freiheitsgrad. Sie entspricht der Anzahl unabhängiger Normalverteilungen.

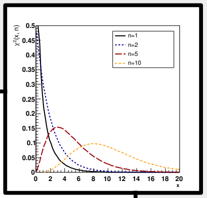
Stammbaum der Normalverteilung

Log-Normalverteilung



Zufallsvariable:
Produkt von Einzelmessungen

χ^2-Verteilung

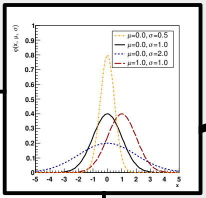


Log Likelihood für
normalverteilte Zufallsgrößen

Logarithmus über
Argument der Nor-
malverteilung

Logarithmus über
Argument Normal-
verteilung selbst

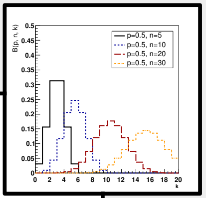
Normalverteilung



Zufallsvariable:
Summe von Einzelmessungen

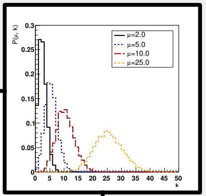
$n \to \infty, p = const$

Binomialverteilung



$n \to \infty, np = \mu = const$

Poissonverteilung

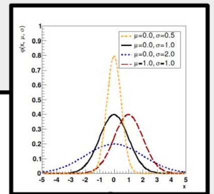


Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Binomialverteilung – Nomalverteilung – Poissonverteilung

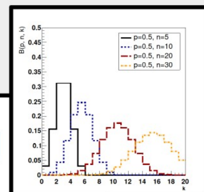
Normalverteilung

Zufallsvariable:
Summe von Einzelmessungen



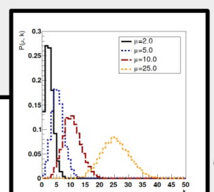
$n \to \infty, p = const$

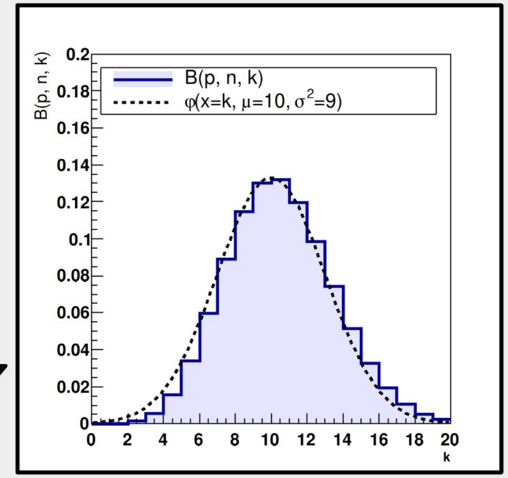
Binomialverteilung



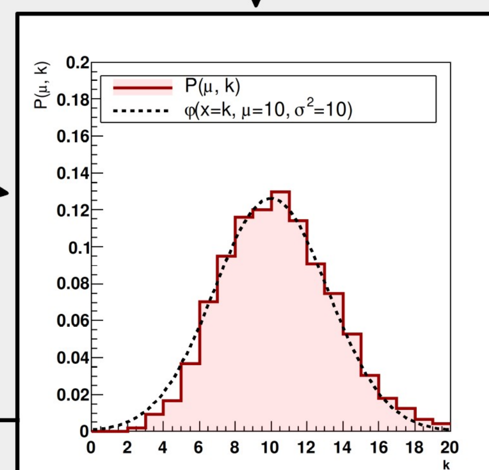
$n \to \infty, np = \mu = const$

Poissonverteilung





$n \to \infty$
$p \to 0$
$np = \mu = const$



1.4 Kenngrößen von Wahrscheinlichkeitsverteilungen

Ziel der Verwendung statistischer Kenngrößen (z.B. in der deskriptiven Statistik) ist die Reduktion der Information der vollständigen Wahrscheinlichkeitsverteilung auf ein handhabares Maß.

Kennzahlen von Wahrscheinlichkeitsdichten

- Oftmals ist es im wissenschaftlichen Diskurs nicht praktikabel/möglich, die ganze Wahrscheinlichkeitsdichte anzugeben.

- In diesem Fall reduziert man die Wahrscheinlichkeitsdichte auf einige wenige **charakteristische Kennzahlen**.

- Die folgenden (Gruppen von) Kennzahlen sollten Ihnen geläufig sein:
  - Quantilen
  - Lagemaße (→ Erwartungswert)
  - Algebraische Momente (→ Varianz, Kovarianz)

- In manchen Fällen ist die Wahrscheinlichkeitsdichte bei Kenntnis dieser Kennzahlen **vollständig bestimmt** (→ z.B. bei der Normalverteilung).

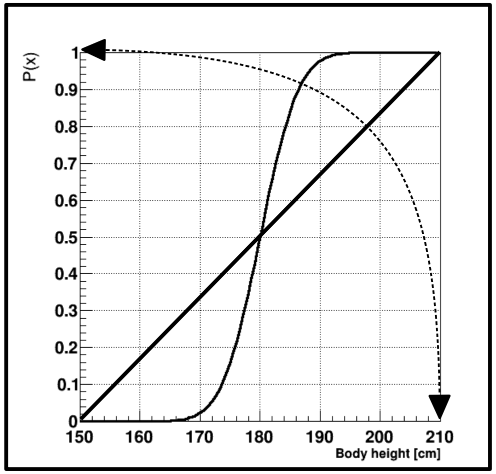
Quantilen

Die Umkehrfunktion der kumulativen Wahrscheinlichkeitsfunktion

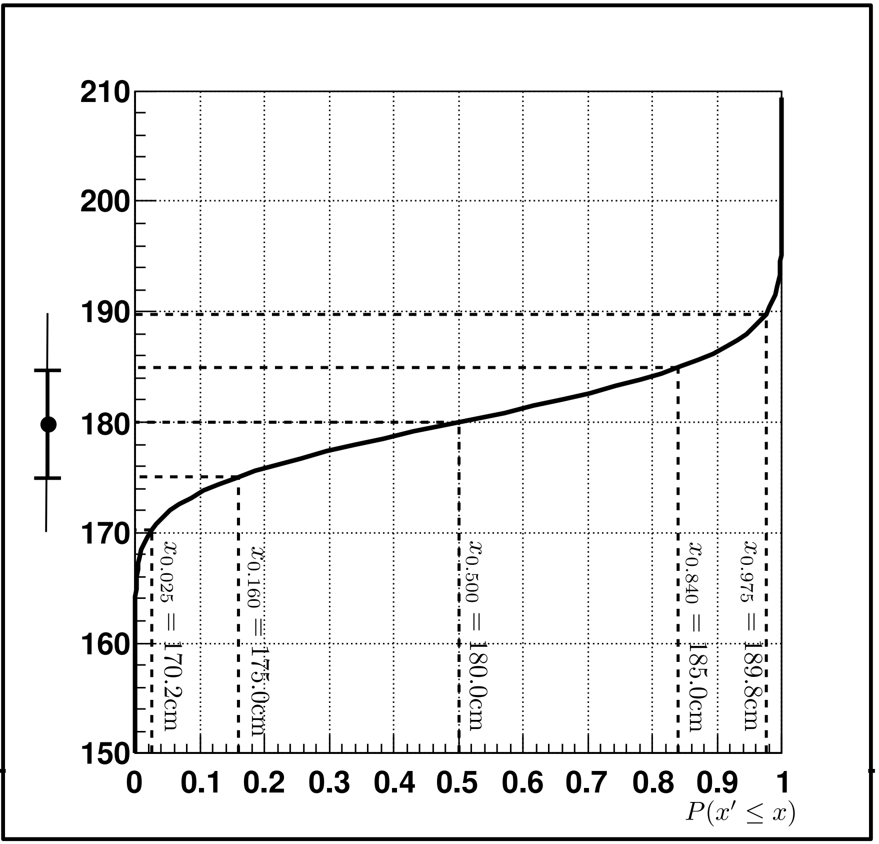
$$x_\alpha = P^{-1}(x' \leq x)\big|_{x=\alpha}$$

liefert zu jeder Wahrscheinlichkeit $P(x' \leq x) = \alpha$ einen Wert der Zufallsgröße x_α. Diesen Wert bezeichnet man als Quantil der Ordnung α oder α-Punkt.

- Umkehrfunktion → Spiegelung an Hauptdiagonale



(Vgl. mit Folie 23)



Lagemaße

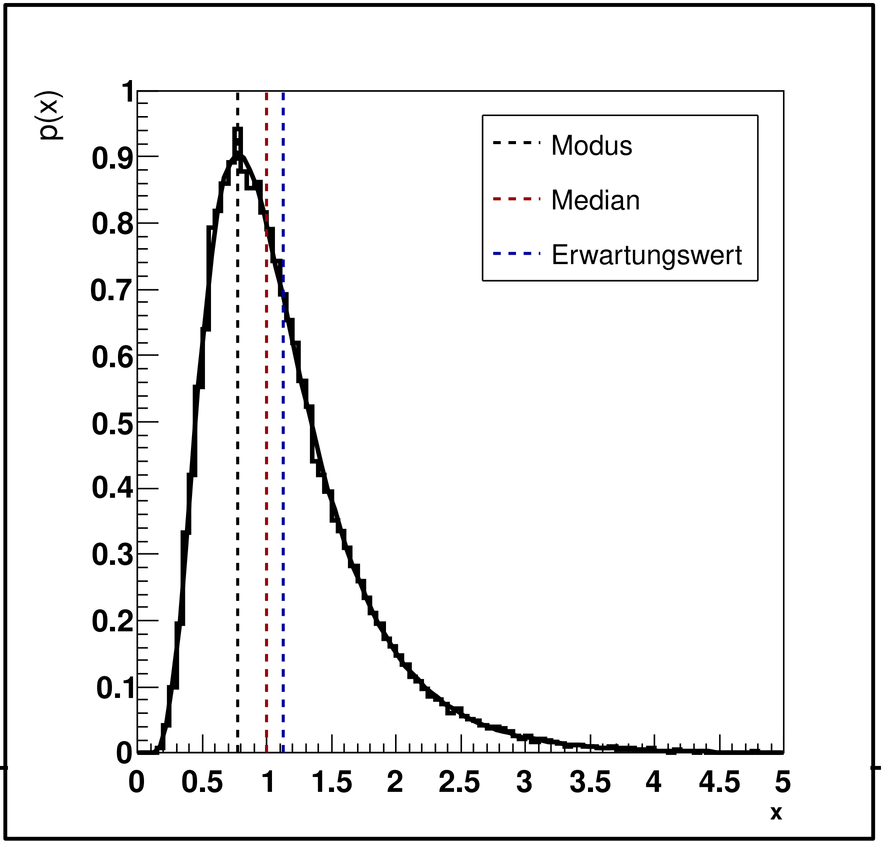
- Gängige Lagemaße für Wahrscheinlichkeitsverteilungen:

- Modus:
  Maximum der Verteilung ($\rightarrow$ wahrscheinlichster Wert).

- Median ($x_{1/2}$):
  Hälfte der Werte ober-/unterhalb.

- Erwartungswert ($E[x]$):
  Mit Wahrscheinlichkeit gewichtetes Mittel.

- Das Beispiel rechts zeigt, dass diese drei Maße nicht gleich sein müssen ($\rightarrow$ wann wäre das der Fall?).



Erwartungswert

Ist x eine kontinuierlich (diskret) verteilte Zufallsvariable und $p(x)$ $(P(x))$ die zugehörige Wahrscheinlichkeitsdichte(-verteilung) über Ω. Dann bezeichnet man die Größe

$$E[x] = \int\limits_{\Omega} x\, p(x) \mathrm{d}x \qquad \text{(kontinuierlich)}$$

$$E[x] = \sum_{\Omega} x_i\, P(x_i) \qquad \text{(diskret)}$$

als den Erwartungswert für x über Ω.

- Für eine vorgegebene Wahrscheinlichkeitsdichte(-verteilung) ist $E[x]$ eine Zahl und keine Funktion von x (andere gänige Bezeichnungen: μ, $\langle x\rangle$).

- Der Erwartungswert ist linear in x:

$$E\left[\alpha\, x + \beta\, y\right] = \alpha\, E[x] + \beta\, E[y]$$

Algebraische Momente und Varianz

Man bezeichnet $E\left[(x-x_0)^n\right]$ als das n-te algebraische Moment um x_0. In der Statistik sind die folgenden Spezialfälle für $x_0 = E[x]$ von Relevanz:

$$\text{0-tes Moment:}\qquad E[(x-E[x])^0] = \int\limits_\Omega p(x)\,\mathrm{d}x = 1$$

$$\text{1-tes Moment:}\qquad E[(x-E[x])^1] = \int\limits_\Omega (x-E[x])\,p(x)\,\mathrm{d}x = 0$$

$$\begin{aligned}\text{2-tes Moment:}\qquad E[(x-E[x])^2] &= \int\limits_\Omega (x-E[x])^2\,p(x)\,\mathrm{d}x \\ &= \underbrace{\int\limits_\Omega x^2\,p(x)\,\mathrm{d}x}_{E[x^2]} - \underbrace{2E[x]\int\limits_\Omega x\,p(x)\,\mathrm{d}x}_{E[x]^2} + \underbrace{E[x]^2\int\limits_\Omega p(x)\,\mathrm{d}x}_{E[x]^2}\end{aligned}$$

Das 2-te algebraische Moment um $E[x]$

$$\mathrm{var}[x] = E\left[(x-E[x])^2\right] = E[x^2] - E[x]^2$$

heißt Varianz von x über Ω, $\sigma_x = \sqrt{\mathrm{var}[x]}$ heißt Standardabweichung.

Kovarianz und Korrelationskoeffizient

Analog zur Varianz definiert man die Kovarianz

$$\mathrm{cov}[x,y] = E\left[(x - E[x])\,(y - E[y])\right] = E[x\,y] - E[x]E[y]$$

zur Beschreibung der Beziehung zweier Zufallsvariablen x und y zueinander. Die Matrix

$$V_{xy} = \begin{pmatrix} \sigma_x^2 & \mathrm{cov}[x,y] \\ \mathrm{cov}[x,y] & \sigma_y^2 \end{pmatrix}$$

bezeichnet man als Kovarianzmatrix. Die Größe

$$\rho_{xy} \equiv \frac{\mathrm{cov}[x,y]}{\sigma_x \sigma_y}$$

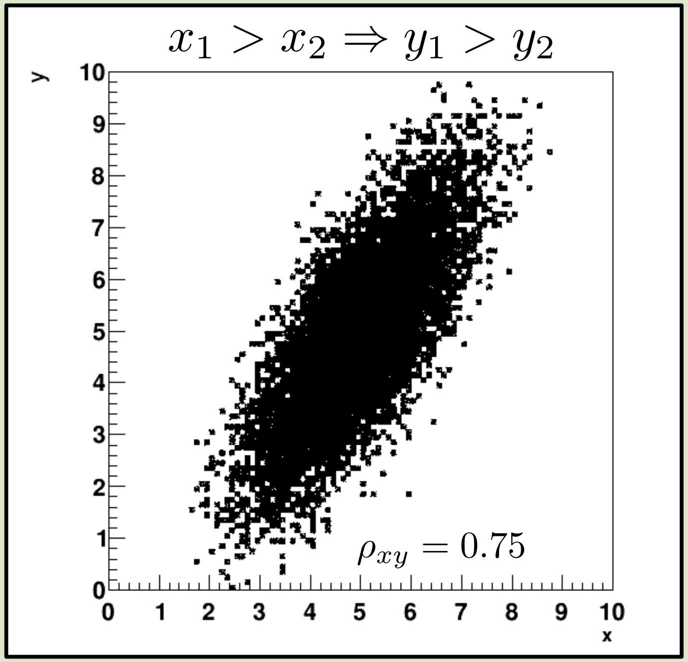
als (linearen) Korrelationskoeffizienten.

- V_{xy} is symmetrisch, d.h. es gibt immer eine Hauptachsentransformation und reelle Eigenwerte.
- $\rho_{xy} \in [-1,+1]$. Für unabhängige Zufallsvariablen gilt:

$$E[xy] = E[x]E[y] \quad \text{d.h.} \quad \rho_{xy} = 0$$

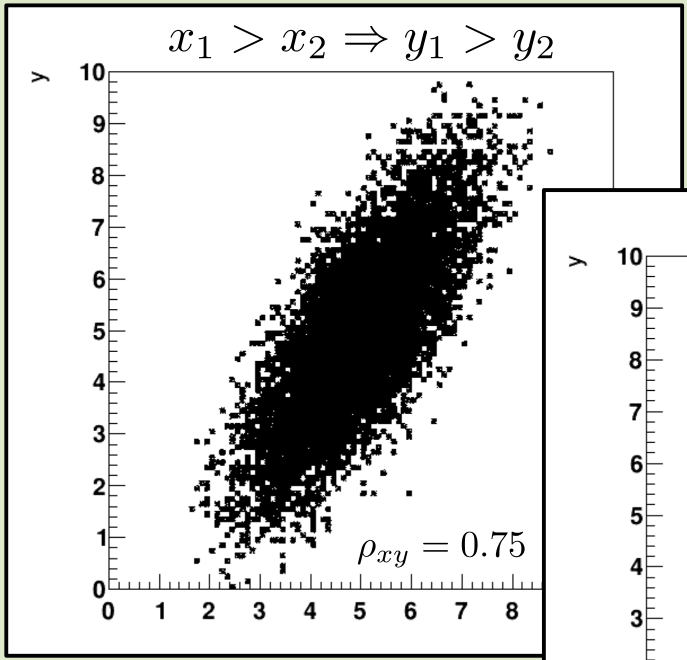
Korrelation anschaulich

Positive Korrelation:

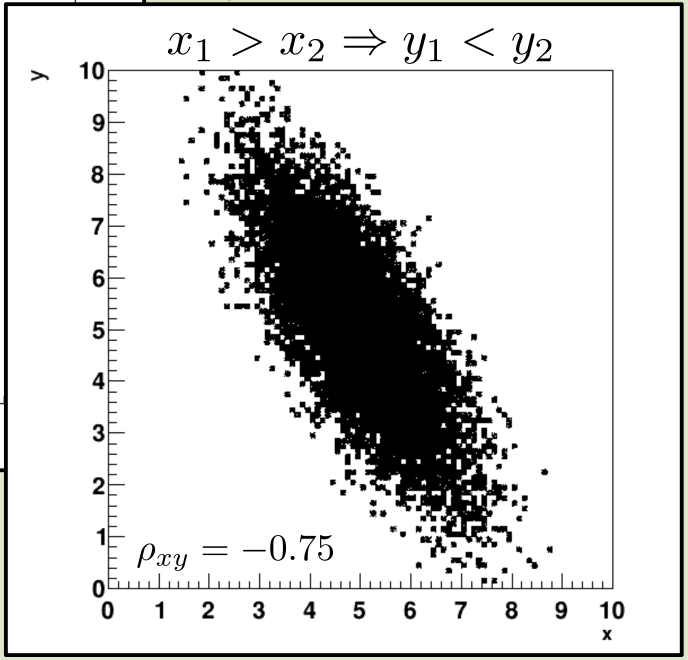


Korrelation anschaulich

Positive Korrelation:

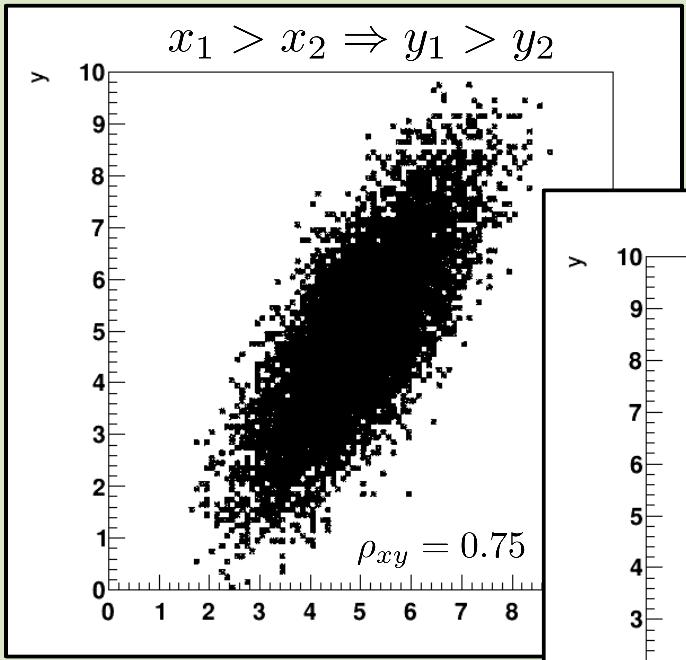


Negative Korrelation:

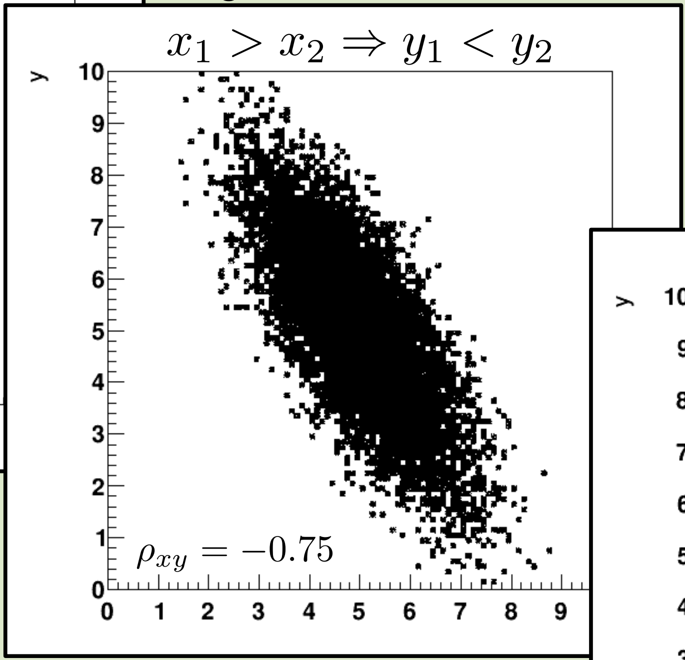


Korrelation anschaulich

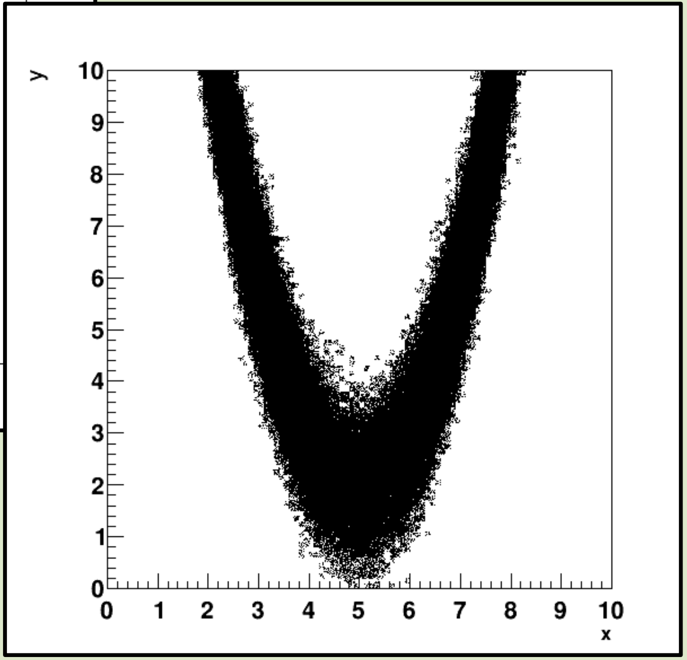
Positive Korrelation:



Negative Korrelation:

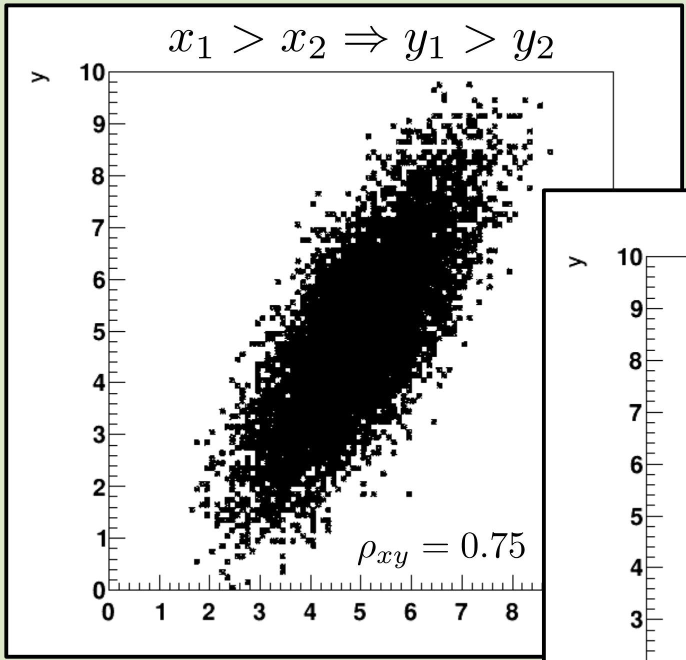


Welche Korrelation?

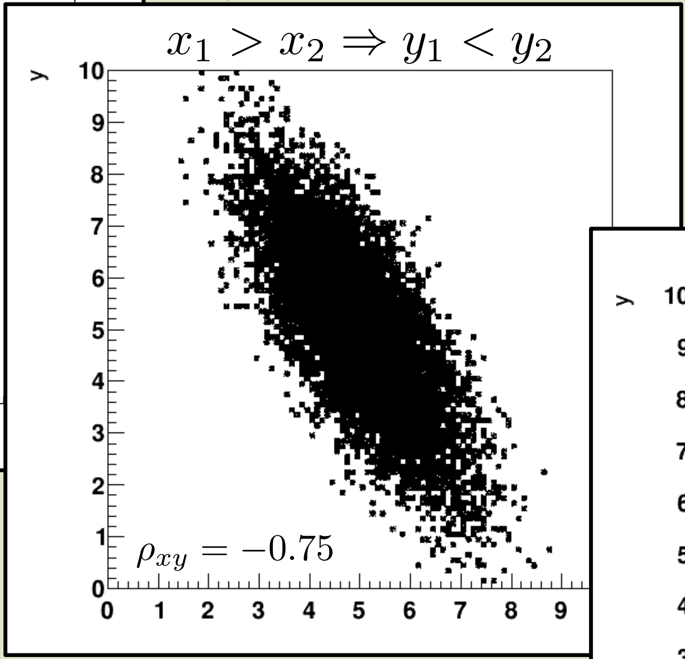


Korrelation anschaulich

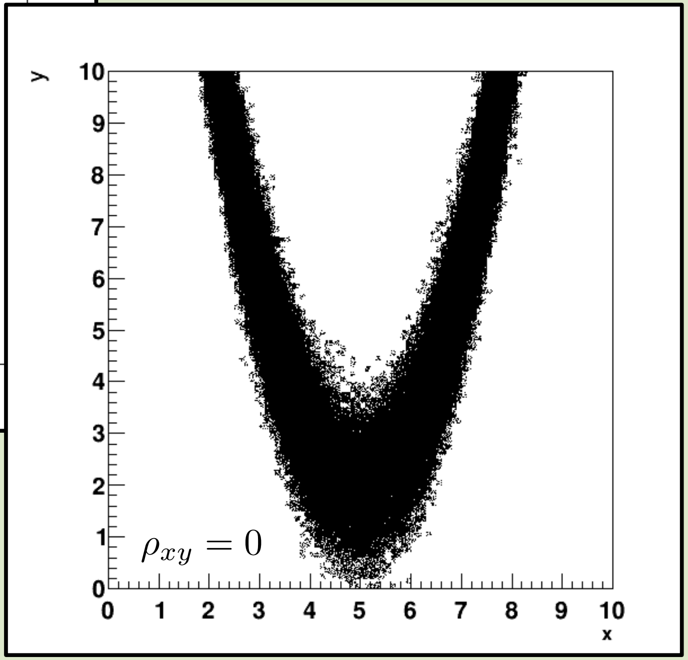
Positive Korrelation:



Negative Korrelation:

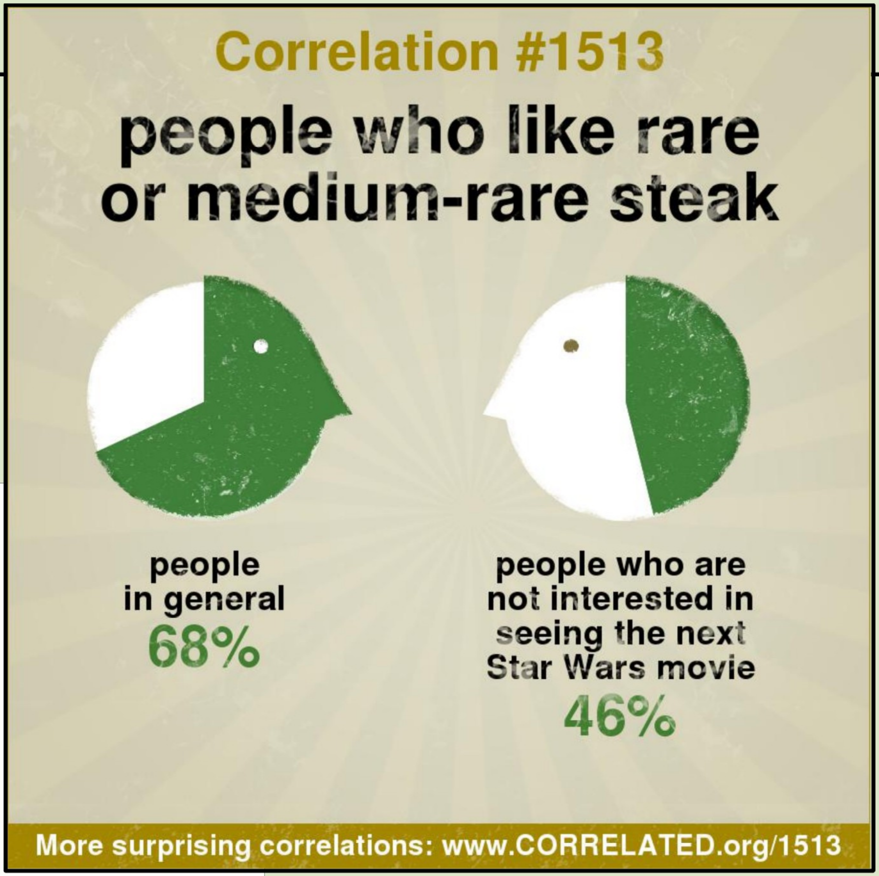


Welche Korrelation?



Korrelation vs. Kausalität

- Verwechseln Sie nicht – wie 99% der Bevölkerung – die Begriffe Korrelation und Kausalität.



- Unkorreliert $\cancel{\Rightarrow}$ unabhängig!
- Korreliert $\cancel{\Leftrightarrow}$ kausal verknüpft!
  - Korrelation kann zufällig sein.
  - $x \Rightarrow y$.
  - $y \Rightarrow x$.
  - $z \Rightarrow (x, y)$.

Zentraler Grenzwertsatz (Lindeberg-Lévy)

Betrachte einen Ereignisraum mit n unabhängigen Zufallsvariablen x_i, z.B. Ergebnissen einer Stichprobenziehung. Die x_i können dabei einer *beliebigen* Wahrscheinlichkeitsverteilung mit endlichem Erwartungswert (μ) und endlicher Varianz (σ^2) auf der Grundgesamtheit folgen. Dann gilt für die Größe:

$$z_n = \frac{\sum\limits_{i=1}^{n}(x_i - \mu)}{\sigma\sqrt{n}}$$

$$\lim_{n\to\infty} z_n = \varphi(x, \mu = 0, \sigma = 1)\,,$$

d.h. für den Grenzübergang $n \to \infty$ strebt die Verteilung von z_n einer Standardnormalverteilung (mit $\mu = 0$ und $\sigma = 1$) entgegen.

- Beachten Sie: μ und σ^2 sind Erwartungswert und Varianz der Wahrscheinlichkeits-dichte(-verteilung) auf der Grundgesamtheit.

Zentraler Grenzwertsatz – Verallgemeinerung –

- Der zentrale Grenzwertsatz von Lindeberg-Lévy läßt sich so verallgemeinern, dass die Gleichheit der Wahrscheinlichkeitsverteilung keine notwendige Voraussetzung mehr ist ($\rightarrow$ siehe z.B. zentraler Grenzwertsatz von Ljapunow).

- Die Folge der Zufallsvariablen $\{x_i\}$ muss in diesem Fall qualitativ den folgenden Bedingungen genügen:

  - Stochastische Unabhängigkeit.

  - Endliche Varianz $0 < \mathrm{var}[x_i] < \infty\,; \quad \forall i$.

  - Für die Summe der Varianzen gilt:

$$S_n \equiv \sum_{i=1}^{n} \mathrm{var}[x_i]$$

$$\exists\,\delta > 0 \quad \text{mit:} \quad \lim_{n\to\infty} \frac{1}{S_n^{1+\delta/2}} \sum_{i=1}^{n} E\left[|x_i - E[x_i]|^{2+\delta}\right] = 0 \quad \text{(Ljapunow-Bedingung)}$$

- Der Satz von Ljapunow erklärt, warum die Summe vieler kleiner unabhängiger Zufallsvariablen normalverteilt ist.

Zurück zu Folie 24
Zurück zu Folie 52
Zurück zu Folie 55

Mittelwert, Varianz, Kovarianz von Stichproben

- Bisher haben wir Erwartungswert ($E[x]$), Varianz ($\mathrm{var}[x]$) und Kovarianz (ρ_{xy}) der Wahrscheinlichkeitsdichte(-verteilung) von Einzelmessungen aus der Grundgesamtheit diskutiert.

- Deren Bestimmung setzt die Kenntnis der zugrunde liegenden Wahrscheinlichkeitsdichte(-verteilung) voraus.

- Erwartungswert, Varianz und Kovarianz der Wahrscheinlichkeitsdichte(-verteilung) der Grundgesamtheit lassen sich <u>selbst ohne Kenntnis der Wahrscheinlichkeitsdichte (-verteilung) der Grundgesamtheit</u> aus Mittelwert ($\overline{x}$), Varianz (s^2), und Kovarianz (r) von Stichproben ermitteln!

- Dies ist unabhängig von der Wahrscheinlichkeitsdichte(-verteilung) der Fall.

- Dies ist eine direkte Konsequenz des zentralen Grenzwertsatzes.

Mittelwert der Stichprobe

Die Größe

$$\overline{x} = \frac{1}{n}\sum_{i\leq n} x_i$$

heißt Mittelwert der Stichprode. Ihr Erwartungswert ($E[\overline{x}]$) und ihre Varianz ($\mathrm{var}[\overline{x}]$) sind:

$$E[\overline{x}] = E\left[\frac{1}{n}\sum_{i\leq n} x_i\right] = \frac{1}{n}\sum_{i\leq n}\underbrace{E[x_i]}_{\equiv\,\mu} = \frac{1}{n}\sum_{i\leq n}\mu = \mu$$

$$\mathrm{var}[\overline{x}] = E[\overline{x}^2] - E[\overline{x}]^2 = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)\right] - \mu^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} E[x_i x_j] - \mu^2$$

$$= \frac{1}{n^2}\Big[(n^2-n)\mu^2 + n(\mu^2+\sigma^2)\Big] - \mu^2 = \sigma^2/n\,.$$

Dabei sind μ und σ^2 der Erwartungswert und die Varianz der Einzelmessung und n die Länge der Stichprobe.

Mittelwert der Stichprobe

Die Größe

$$\overline{x} = \frac{1}{n}\sum_{i\leq n} x_i$$

heißt Mittelwert der Stichprode. Ihr Erwartungswert ($E[\overline{x}]$) und ihre Varianz ($\mathrm{var}[\overline{x}]$) sind:

$$E[\overline{x}] = E\left[\frac{1}{n}\sum_{i\leq n} x_i\right] = \frac{1}{n}\sum_{i\leq n}\underbrace{E[x_i]}_{\equiv\,\mu} = \frac{1}{n}\sum_{i\leq n}\mu = \mu$$

$$\begin{aligned}\mathrm{var}[\overline{x}] &= E[\overline{x}^2] - E[\overline{x}]^2 = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)\right] - \mu^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} E[x_i x_j] - \mu^2 \\ &= \frac{1}{n^2}\left[\underbrace{(n^2-n)\mu^2}+\underbrace{n(\mu^2+\sigma^2)}\right] - \mu^2 = \sigma^2/n\,.\end{aligned}$$

Dabei sind μ und σ^2 der Erwartungswert und die Varianz der Einzelmessung und n die Länge der Stichprobe.

Mittelwert der Stichprobe

Die Größe

$$\overline{x} = \frac{1}{n} \sum_{i \leq n} x_i$$

heißt Mittelwert der Stichprode. Ihr Erwartungswert ($E[\overline{x}]$) und ihre Varianz ($\mathrm{var}[\overline{x}]$) sind:

$$E[\overline{x}] = E\left[\frac{1}{n} \sum_{i \leq n} x_i\right] = \frac{1}{n} \sum_{i \leq n} \underbrace{E[x_i]}_{\equiv \mu} = \frac{1}{n} \sum_{i \leq n} \mu = \mu$$

$$\mathrm{var}[\overline{x}] = E[\overline{x}^2] - E[\overline{x}]^2 = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)\right] - \mu^2 = \frac{1}{n^2} \sum_{i,j=1}^{n} E[x_i x_j] - \mu^2$$

$$= \frac{1}{n^2}\Big[\underbrace{(n^2 - n)\mu^2}_{} + \underbrace{n(\mu^2 + \sigma^2)}_{}\Big] - \mu^2 = \sigma^2/n\,.$$

Dabei sind μ und σ^2 der Erwartungswert und die Varianz der Einzelmessung und n die Länge der Stichprobe.

$n(n-1)$ "off-diagonale" Elemente $(i \neq j)$ mit:
$E\left[x_i x_j\right] = E\left[x_i\right] E\left[x_j\right] = \mu^2$

Mittelwert der Stichprobe

Die Größe

$$\overline{x} = \frac{1}{n} \sum_{i \le n} x_i$$

heißt Mittelwert der Stichprode. Ihr Erwartungswert ($E[\overline{x}]$) und ihre Varianz ($\mathrm{var}[\overline{x}]$) sind:

$$E[\overline{x}] = E\left[\frac{1}{n}\sum_{i\le n} x_i\right] = \frac{1}{n}\sum_{i\le n} \underbrace{E[x_i]}_{\equiv \mu} = \frac{1}{n}\sum_{i\le n} \mu = \mu$$

$$\mathrm{var}[\overline{x}] = E[\overline{x}^2] - E[\overline{x}]^2 = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)\right] - \mu^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} E[x_i x_j] - \mu^2$$

$$= \frac{1}{n^2}\left[\underbrace{(n^2-n)\mu^2}+\underbrace{n(\mu^2+\sigma^2)}\right] - \mu^2 = \sigma^2/n\,.$$

Dabei sind μ und σ^2 der Erwartungswert und die Varianz der Einzelmessung und n die Länge der Stichprobe.

$n(n-1)$ "off-diagonale" Elemente
($i \neq j$) mit:
$E\left[x_i x_j\right] = E\left[x_i\right] E\left[x_j\right] = \mu^2$

n "diagonale" Elemente
($i = j$) mit:
$E\left[x_i x_i\right] = \mu^2 + \sigma^2$

Varianz der Stichprobe

Die Größe

$$s^2 = \frac{1}{n-1}\sum_{i\leq n}(x_i - \overline{x})^2 = \frac{n}{n-1}\left(\overline{x^2} - \overline{x}^2\right)$$

heißt Varianz der Stichprobe. Ihr Erwartungswert ($E[s^2]$) und ihre Varianz sind:

$$E[s^2] = \sigma^2$$

$$\mathrm{var}[s^2] = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-4}\mu_2^2\right),$$

wobei μ_k das k-te zentrale Moment um μ, und μ und σ^2 der Erwartungswert und die Varianz der Einzelmessung sind.

- Die μ_k können durch $m_k = \frac{1}{n-1}\sum_{i\leq n}(x_i - \overline{x})^k$ abgeschätzt werden.
- Beachten Sie bei der Berechnung der Varianz der Stichprobe die Normierung auf $n-1$ statt n . Den Faktor $n/(n-1)$ bezeichnet man als **Besselkorrektur**.

Korrelationskoeffizient der Stichprobe

Die Größe

$$r = \frac{\hat{V}_{xy}}{s_x s_y} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_j - \overline{x})^2 \sum(y_k - \overline{y})^2} = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\sqrt{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)}}$$

ist eine Schätzfunktion für den Korrelationskoeffizienten zweier einzelner Zufallsvariablen x und y. Sie hat den Erwartungswert und die Varianz:

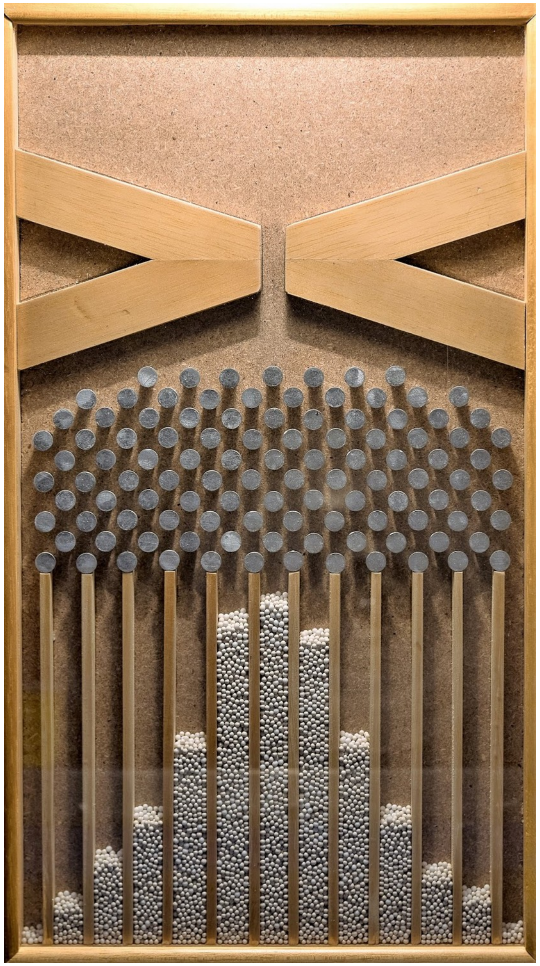
$$E[r] = \rho - \frac{\rho(1-\rho^2)}{2n} + \mathcal{O}({}^1\!/_{n^2})$$

$$\mathrm{var}[r] = \frac{1}{n}(1-\rho^2)^2 + \mathcal{O}({}^1\!/_{n^2})$$

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

1.5 Softwarewerkzeuge zur statistischen Datenanalyse

NB: Das hier ist ein Galtonbrett. Können Sie erklären welche Verteilung durch die Kugeln abgebildet wird und v.a. warum?



Moderne Anwendungen der Statistik basieren heutzutage auf vielfacher Simulation. Hierzu sind Sie zwingend auf compultergestützte Datenerfassung (und Erzeugung) angewiesen.

Computer in der statistischen Datenanalyse

- Gerade aus den Folien 47-50 dieser Vorlesung können Sie erahnen, wie groß die Bedeutung von **Stichproben** in der statistischen Datenanalyse ist.

- Im (ernsthaften) wissenschaftlichen Alltag werden Sie zwingend darauf angewiesen sein große Datenmengen mit Hilfe von Computern zu verarbeiten, z.B.:

  - Zur automatisierten Erfassung der Daten (→ Nachweis der Zerfälle eines radioaktiven Präparats).

  - Zur Simulation komplexer Abläufe (→ „Welchen Ausgang des Experiments erwarte ich überhaupt?“ – **siehe nächste Vorlesung!**).

  - Zur Evaluation von Rechnungen mit Hilfe numerischer Methoden (→ Auffinden eines Minimums).

- Zur statistischen Datenanalyse existieren mehrere Computerwerkzeuge.

- Für diesen Kurs empfehlen wir dringend die Nutzung der Programmiersprache **python** mit entsprechenden Bibliotheken.

Python – Linksammlung –

- Python (Tutorial, Compact Course(login: python2019 passwd: PythonIntro), Jupyter-Notebook):
  - … ist eine C basierte Skript und Programmiersprache.
  - **Einfach**, intuitiv erlernbar, **plattformunabhängig**!
- SciPy (Project Website mit Tutorial, Lecture notes):
  - … ist eine python Bibliothek für wissenschaftliches Rechnen (→ für numerische Integration, Interpolation, Optimierung, lineare Algebra, Statistik, …).
- NumPy (Documentation Website, Guide to NumPy):
  - … ist eine python Bibliothek zur Verwendung mehrdimensionaler Arrays (→ Vektoren und Matrizen).
- Matplotlib (Project Website mit weiteren Links):
  - … ist eine python Bibliothek zur mathematischen Darstellung im Stile von MATLAB.
- Die auf dieser Folie angegebenen Links statten Sie mit genügend Information aus, um mit Ihrem Laptop wissenschaftlich arbeiten zu können.

Übungen zur Vorlesung

- Unsere Übungen werden auf python und den auf Folie 53 angegebenen Bibliotheken basieren.

- Wir werden Sie auf den ersten Übungsblättern **Schritt für Schritt** in die Teile der Sprache und der entsprechenden Bibliotheken einführen, die Sie zur Bearbeitung der Aufgaben im weiteren Verlauf dieses Vorlesungsteils benötigen.

- Darüber hinaus hoffen wir, Sie **inspirieren** zu können, sich weiter mit den entsprechenden Funktionen der angegebenen Bibliotheken vertraut zu machen.

- „Science is a kid‘s game!“ – **Spielen** Sie mit dem, was sie im Laufe dieses Kurses lernen! Probieren Sie! (Programmier-)Sprachen lernen Sie ausschließlich, indem Sie sie anwenden und ausprobieren. Das gilt explizit auch für die Verwendung statistischer Methoden ansich!

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“
- Außenstehende empfinden die Methoden der Statistik oft als undurchdringliche Ansammlung von **Kochrezepten**, deren Anwendbarkeit nicht (immer) 100% klar ist.

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“
- Außenstehende empfinden die Methoden der Statistik oft als undurchdringliche Ansammlung von **Kochrezepten**, deren Anwendbarkeit nicht (immer) 100% klar ist.
- Hierfür bestehen u.a. die folgenden Gründe:

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“
- Außenstehende empfinden die Methoden der Statistik oft als undurchdringliche Ansammlung von **Kochrezepten**, deren Anwendbarkeit nicht (immer) 100% klar ist.
- Hierfür bestehen u.a. die folgenden Gründe:
  - Statistische Probleme folgen i.a. <u>nicht</u> Ihrer alltäglichen Intuition.

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“
- Außenstehende empfinden die Methoden der Statistik oft als undurchdringliche Ansammlung von **Kochrezepten**, deren Anwendbarkeit nicht (immer) 100% klar ist.
- Hierfür bestehen u.a. die folgenden Gründe:
  - Statistische Probleme folgen i.a. <u>nicht</u> Ihrer alltäglichen Intuition.
  - Oft sind die Ergebnisse einer Analyse <u>mathematisch richtig</u>, die **Interpretation** ist jedoch eine andere, als Sie beabsichtigen (→ statt einer Wahrscheinlichkeit haben Sie z.B. eine bedingte Wahrscheinlichkeit bestimmt).

Mensch und Statistik

- „Glaube keiner Statistik, die du nicht selbst gefälscht hast…“

- Außenstehende empfinden die Methoden der Statistik oft als undurchdringliche Ansammlung von **Kochrezepten**, deren Anwendbarkeit nicht (immer) 100% klar ist.

- Hierfür bestehen u.a. die folgenden Gründe:

  - Statistische Probleme folgen i.a. nicht Ihrer alltäglichen Intuition.

  - Oft sind die Ergebnisse einer Analyse mathematisch richtig, die **Interpretation** ist jedoch eine andere, als Sie beabsichtigen (→ statt einer Wahrscheinlichkeit haben Sie z.B. eine bedingte Wahrscheinlichkeit bestimmt).

  - Wir haben Ihnen auf den Folien 47-50 einige sehr mächtige statistische Größen gezeigt, die die Kenntnis der Wahrscheinlichkeitsdichte(-verteilung) der Grundgesamtheit nicht voraussetzen. Dies sind Ausnahmen! Einige bekannte Methoden und Tests machen Annahmen über die Wahrscheinlichkeitsdichte(-verteilung) der Grundgesamtheit, die nicht auf Ihr Problem zutreffen müssen (→ z.B. Normalverteilung der Zufallsgröße).

Mensch und Statistik und diese Vorlesung

- Im Rahmen dieses Kursteils werden wir zwei Ziele verfolgen:
  - Wir werden versuchen, Ihnen die zugrundeliegenden **systematischen Strukturen** der Methoden aufzuzeigen, die wir (in der kurzen uns verbleibenden Zeit) behandeln (können).
  - Wir werden versuchen, Ihnen die Werkzeuge und Einstellungen zur Hand zu geben, um sich eine **eigene Intuition** im Feld der statistischen Datenanalyse zu „erobern".
- Lassen Sie uns gerne am Ende des Kurses wissen, inwieweit uns dies gelungen sein wird und was wir noch verbessern könnten. Ihr Feedback interessiert uns!
- Unabhängig davon beachten Sie: „erobern" ist ein aktiver Vorgang. Sie sollten diese Vorlesung nicht „konsumieren", Sie sollten sie „antizipieren", um den maximalen Gewinn daraus zu beziehen.

Zusammenfassung

- Bedeutung des Zufalls in der Physik.
- Klärung der Grundlagen und Begriffe der Statistik.
- Wahrscheinlichkeitverteilung und Wahrscheinlichkeitsdichte.
- Kennzahlen von Wahrscheinlicheitsverteilungen.
- Werkzeuge der statistischen Datenanalyse.

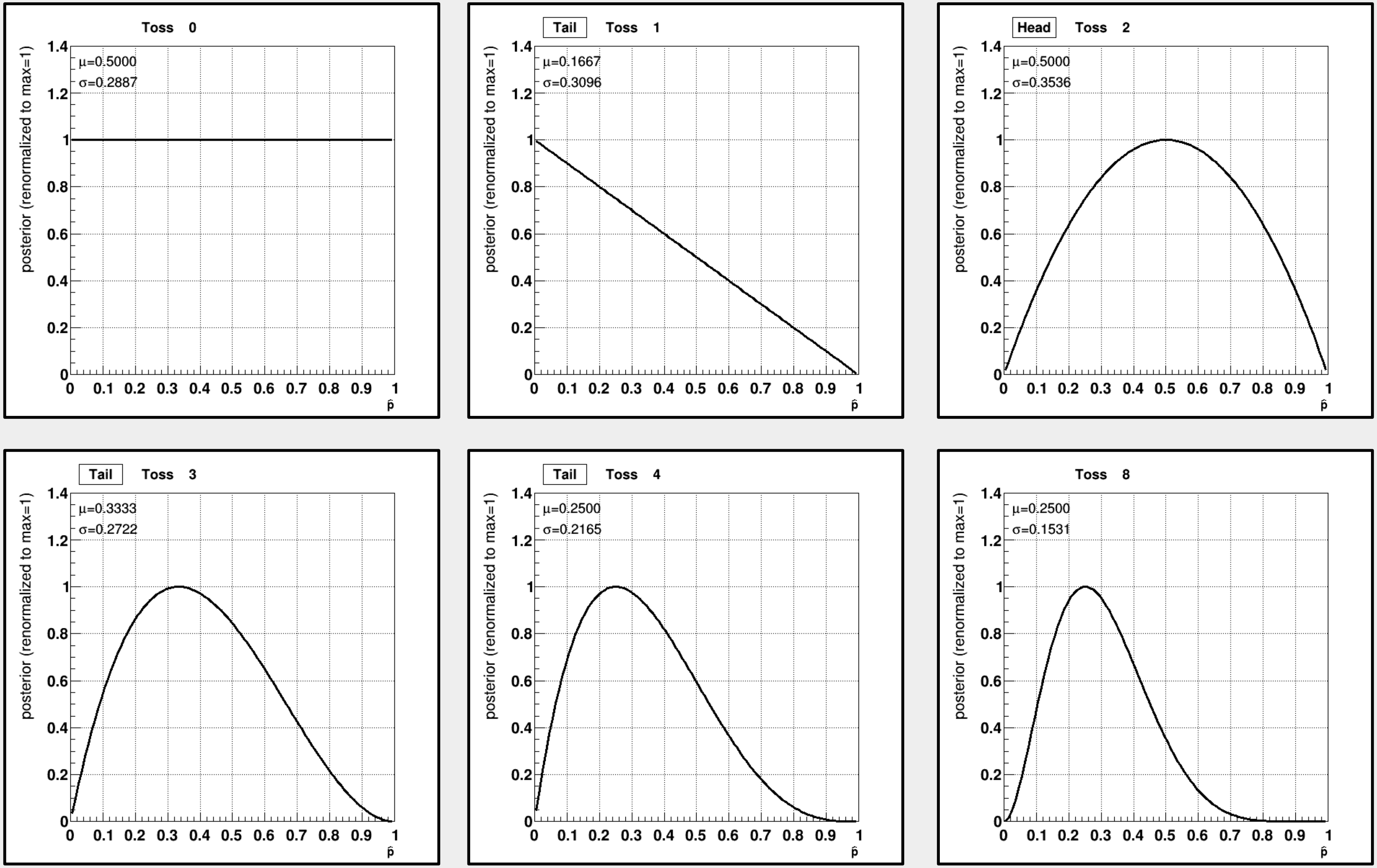
Backup

- Bayesian statistics@work

Bayesianische Statistik vs. Likelihood

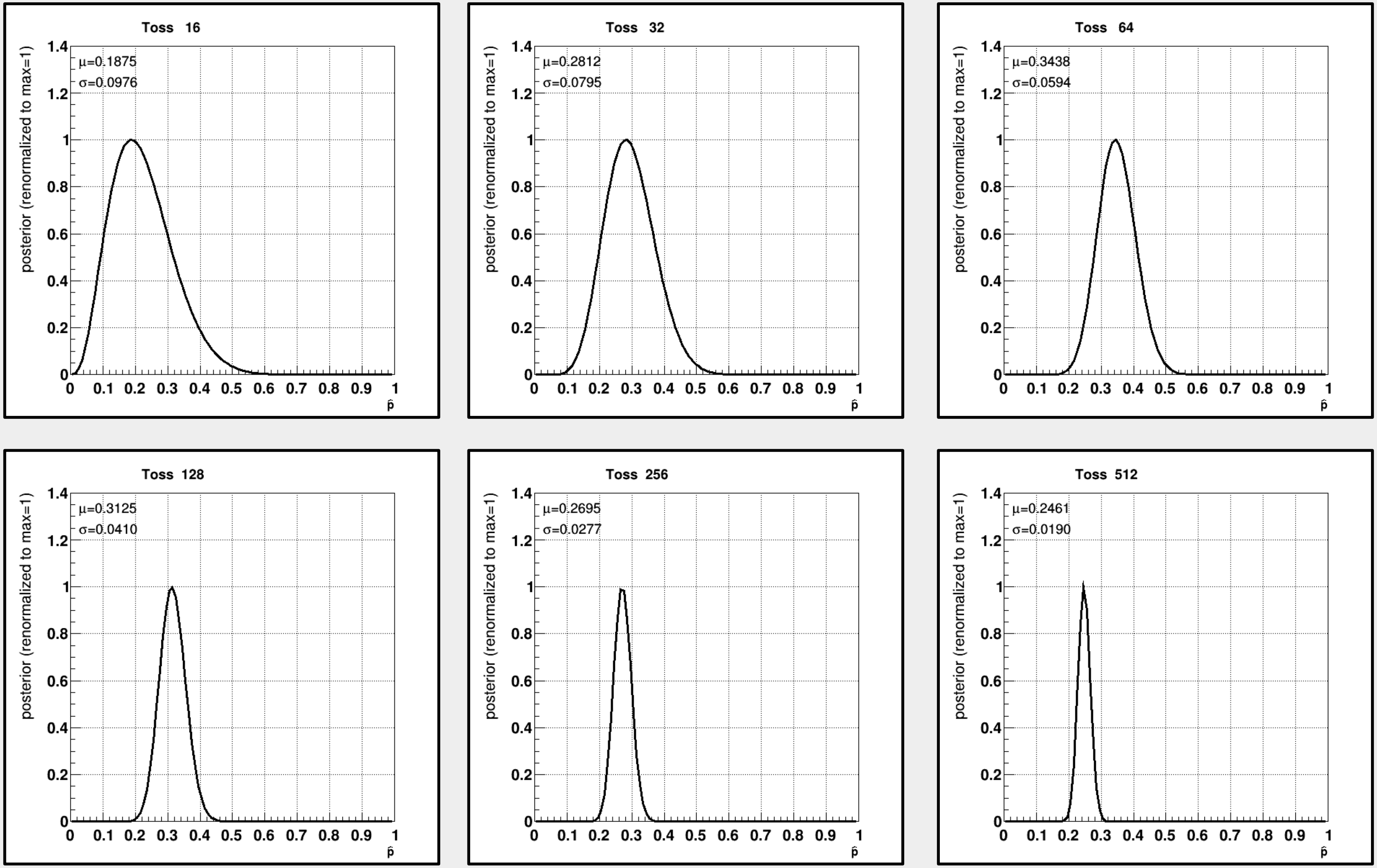
- Es folgt Beispiel wie man im Bayesianischen Paradigma den Posterior nach Auswertung der Daten sukzessive aktualisieren kann. Wir wählen hierzu die folgende Fragestellung:
  - n-facher Münzwurf.
  - Wie hoch ist die Wahrscheinlichkeit $P(\text{"Kopf"})$ bei der Münze „Kopf" zu erhalten?
  - Das zugrunde liegende Modell ist die Binomialverteilung $B(p, n, k)$ nach n-fachem Wurf.
  - Vor dem ersten Wurf ist $P(\text{"Kopf"})$ gleichverteilt (Bayesianischer Prior).
  - Der Prior wird nach jedem Münzwurf aktualisiert, $\hat{P}(\text{"Kopf"}) = \dfrac{k}{n} \equiv \hat{p}$. Die Daten tragen so zur sukzessiven Verfeinerung des Modells bei.

Bayesianischer Posterior bei n-fachem Münzwurf



Ein lauffähiges RooT macro finden Sie hier.

Bayesianischer Posterior bei n-fachem Münzwurf



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Ein lauffähiges RooT macro finden Sie hier.

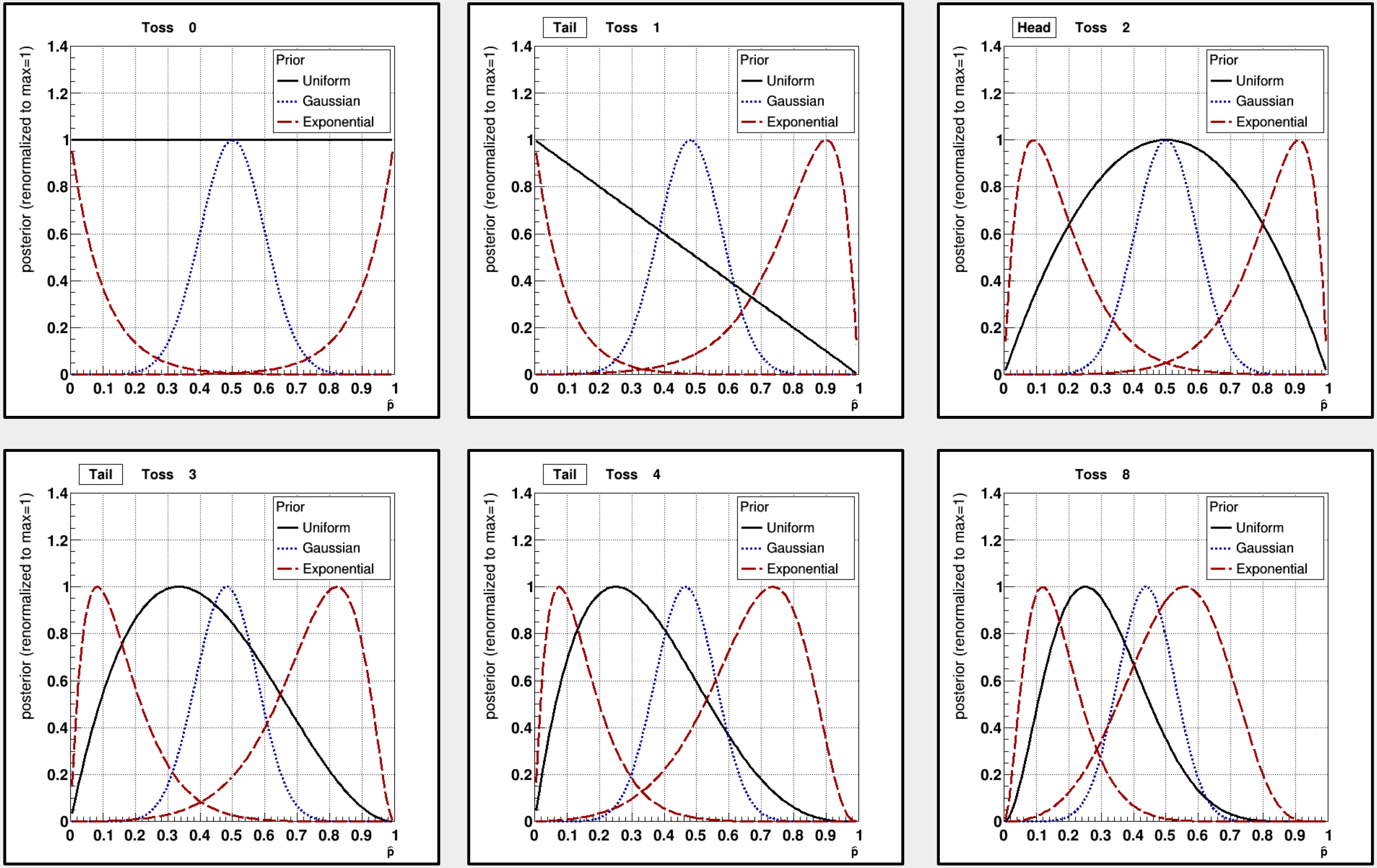
Bayesianischer Posterior bei n-fachem Münzwurf

- Im folgenden zeige ich das gleiche Beispiel nochmal, aber ich starte mit einem anderen Prior!

- Sie sehen das folgende:
  - Gerade zu Beginn, wo die Messung nicht mächtig genug ist, um $P(\text{"Kopf"})$ zu bestimmen, hat der Prior signifikanten Einfluss auf den Posterior.
  - Das ist kein „bug“, das ist ein „feature“. Es wohnt dieser Wahrscheinlichkeitsinterpretation inne.
  - Ihr Prior wird jedoch mit jeder weiteren Messung aktualisiert. Nach eine Messreihe der Länge 512 ist $P(\text{"Kopf"})$ eindeutig durch die Messung festgelegt. Dann spielt der ursprüngliche Prior keine Rolle mehr.

  - **NB:** Die Stichprobe mit uniformem Prior ist zu den Folien B2/B3 identisch.

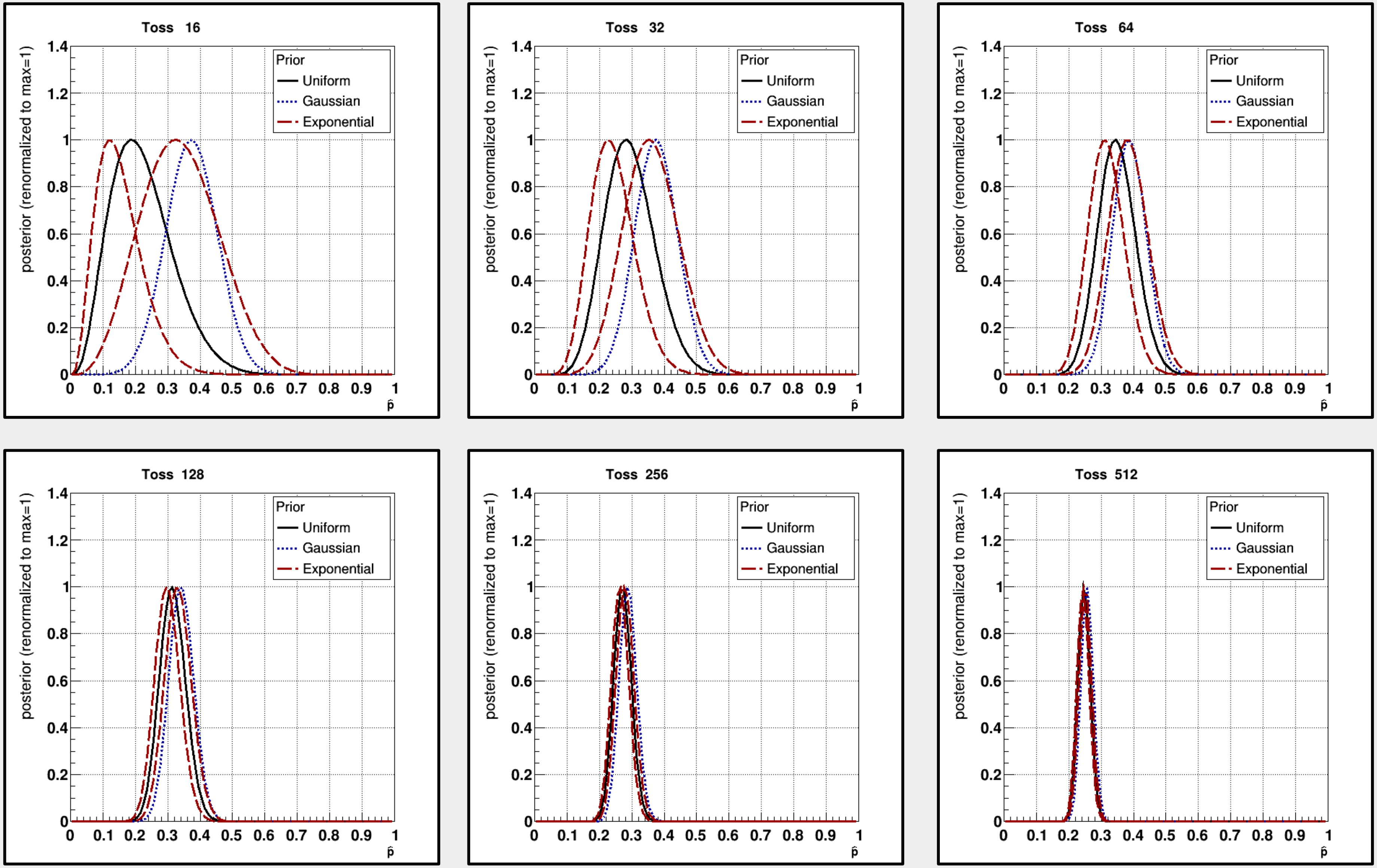
Bayesianischer Posterior bei n-fachem Münzwurf



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Ein lauffähiges RooT macro finden Sie hier.

Bayesianischer Posterior bei n-fachem Münzwurf



Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

Ein lauffähiges RooT macro finden Sie hier.