Karlsruher Institut für Technologie

# Rechnernutzung in der Physik

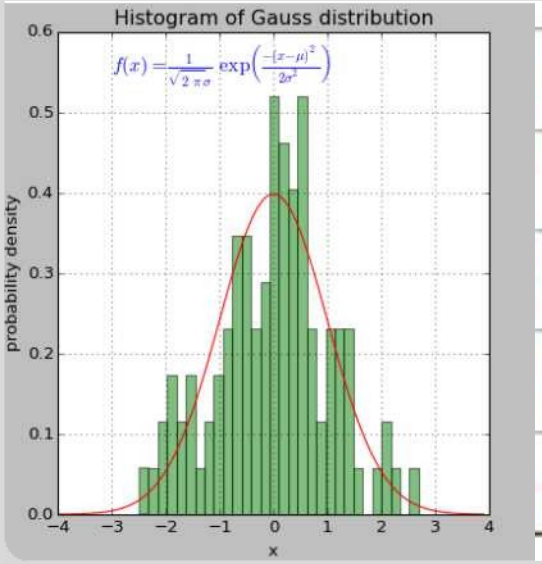
## Zusammenfassung der Vorlesung

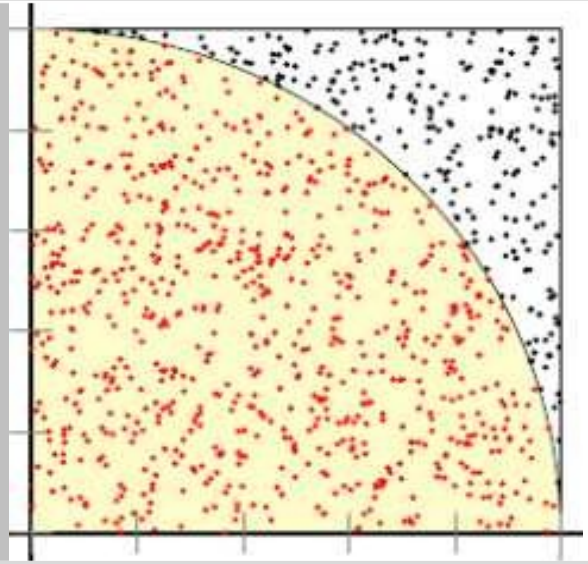
# Computergestützte Datenauswertung

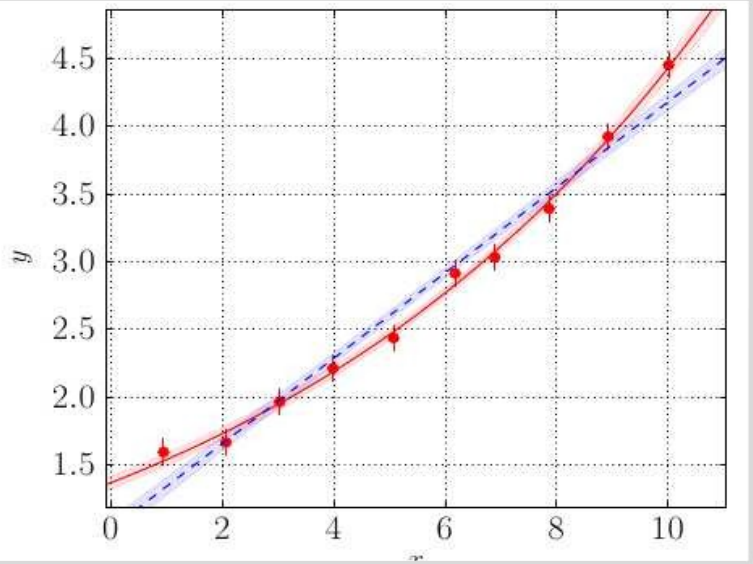
**Günter Quast**

Fakultät für Physik
Institut für Experimentelle Teilchenphysik

**WS 2023/24**





# Organisatorisches

**Die Konten zur Nutzung von bwUniCluster2 sind seit Mo 30.10. nachmittags freigeschaltet.**

**Hinweise s. ILIAS-Link ILIAS-Link**

Auszug aus E-Mail vom Rechenzentrum (Kurzanleitung):

„Für das Anmelden am bwUniCluster ist eine einmalige Registrierung auf der Seite https://bwidm.scc.kit.edu erforderlich.
Danach können Sie sich auf einem Login-Knoten des Clusters mit Ihrem KIT-Account einloggen (uc2.scc.kit.edu).

Informationen zum Systemzugang finden Sie auf der SCC-Webseite:
https://wiki.bwhpc.de/e/BwUniCluster_2.0_User_Access
http://www.scc.kit.edu/dienste/hpc.php
http://www.bwhpc-c5.de/wiki/index.php/BwHPC_Best_Practice_Repository

Falls Sie Fragen zum KIT-Kennwort haben, wenden Sie sich bitte an den SCC-Servicedesk (http://www.scc.kit.edu/hotline/index.php).“

**Bitte ausprobieren !**

**Bei Problemen:** Beratungstutorium am Di, 07.11.2023 15:00-16:30

# Block 1: Statistik und Datenanalyse

**Themen:**

31.10. Grundlagen der Statistik (Wiederholung CgDA);
python & friends, jupyter-Tutorials

07.11. Monte Carlo-Methode als numerisches Hilfsmittel
MC-Tools in PhyPraKit

14.11. Parameterschätzung mit der Maximum-Likelihood-Methode
Hilfsfunktionen in PhyPraKit, kafe2

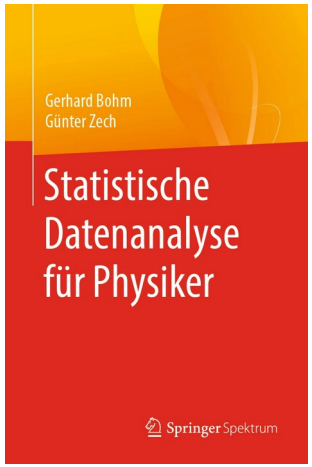
21.11. Maximum Likelihood (2) und numerische Optimierung
Ensemble-Tests

28.11. Hypothesentests

# Literatur

- **Eigene Skripte und Jupyter-Tutorials**

  http://www.ekp.kit.edu/~quast




- **Einführende Literatur, z.B.**
  - G. Bohm, G. Zech, Statistische Datenanalyse für Physiker (2020), über KIT-Bibliothek als Download
  - G. Cowan, *Statistical data analysis*, Oxford (1997).
  - R. J. Barlow, *Statistics: A Guide to the use of statistical methods in the physical sciences*, Wiley (1989).
  - V. Blobel, E. Lormann, *Statistische und numerische Methoden der Datenanalyse*, DESY (2012) (Webseite).

- **weitere Literatur**
  - G. Bohm, G. Zech, *Einführung in Statistik und Messwertanalyse für Physiker*, DESY eBook, http://www-library.desy.de/preparch/books/vstatmp.pdf
    also in English: http://www-library.desy.de/preparch/books/vstatmp_engl.pdf
  - W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes, Cambridge Univ. Press (2007) (Webseite).

# Überblick

über ausgewählte Inhalte der Vorlesung

# Computergestützte Datenauswertung

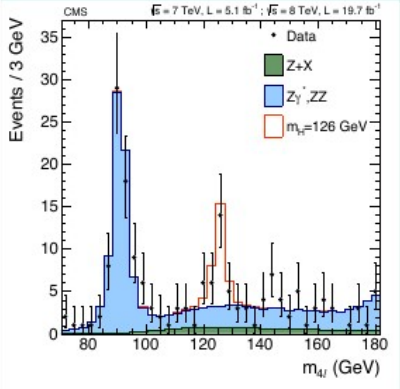
# Physik: Experiment und Theorie

*Die Aufgabe des Wissenschaftlers:*

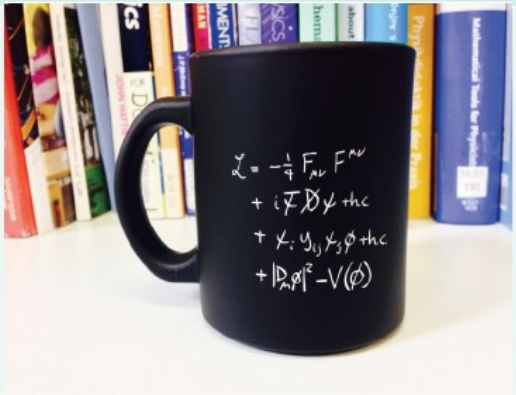
## Vergleich von Modellen mit der Wirklichkeit

**Frage:**
Passt das (theoretische) Modell zu den Beobachtungen (Messungen) ?

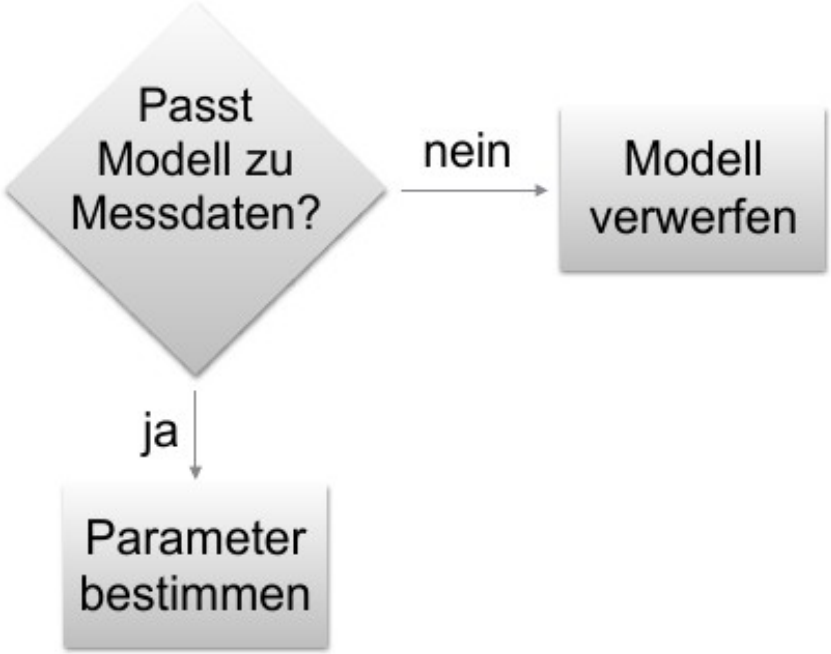
| Experiment | Theorie |
|---|---|
| Messdaten mit **Unsicherheiten**, z. B. als Häufigkeits-verteilung | Mathematisches **Modell** der Daten, oft mit freien **Parametern** |




PRD 89 (2014) 092007



Phys. Educ. 52 (2017) 034001

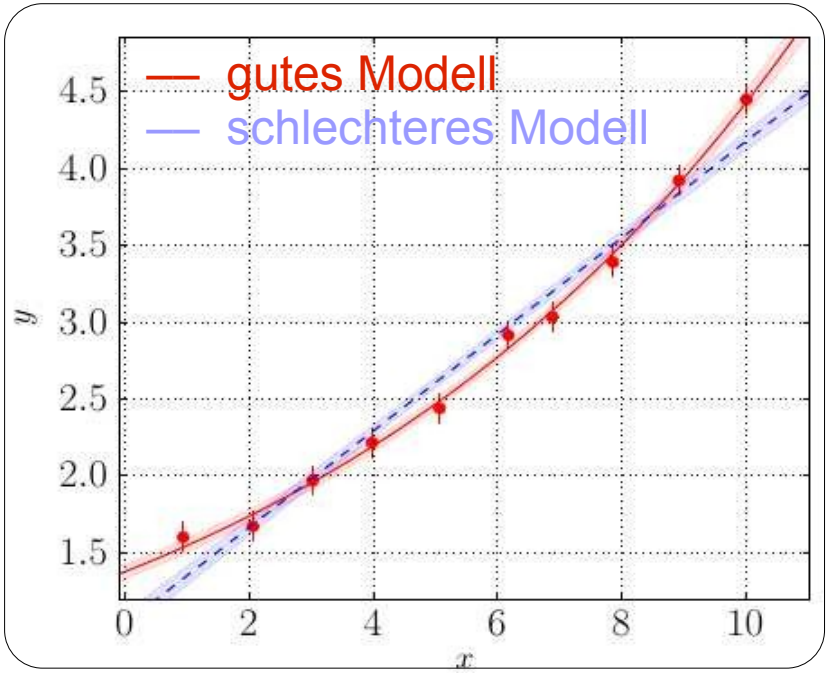


# Wozu Statistische Datenanalyse ?

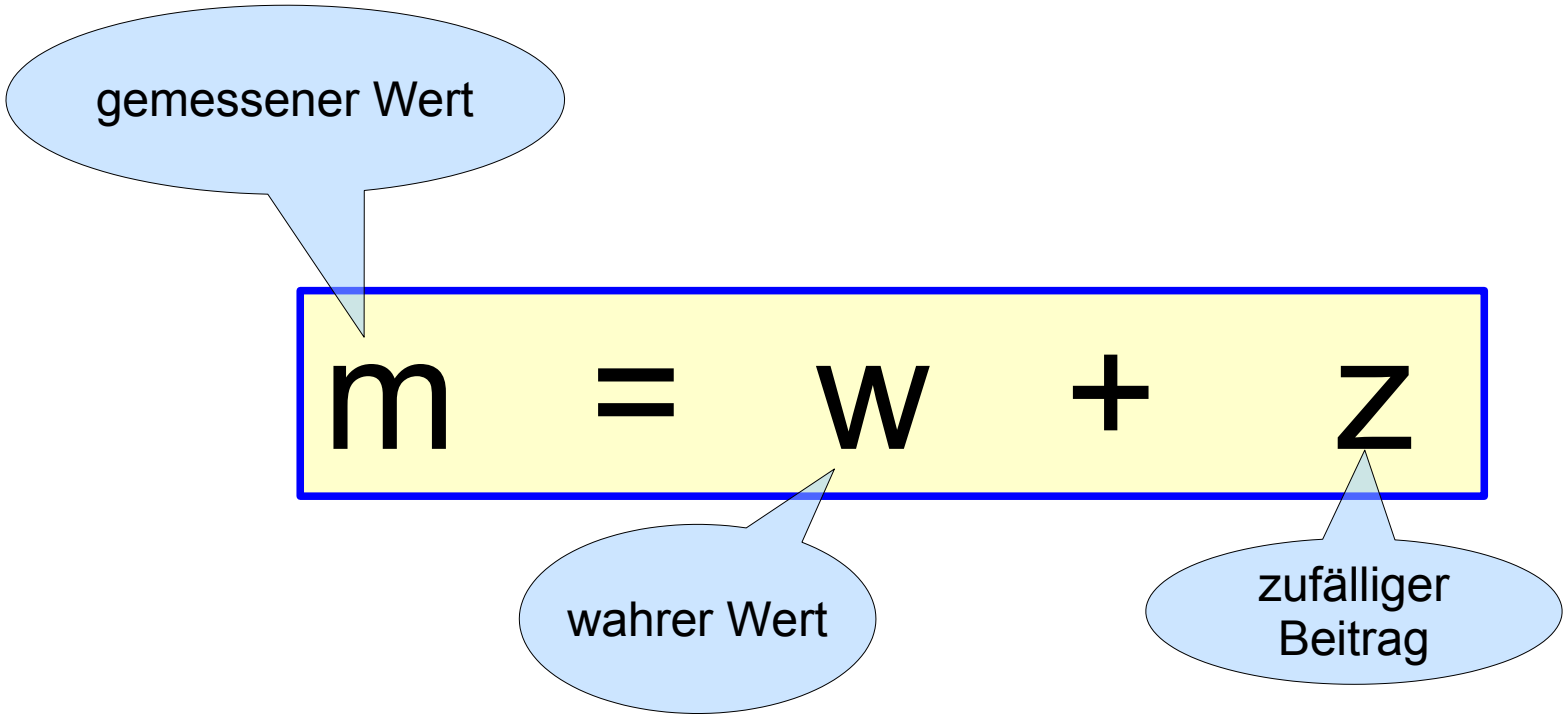
- **Messungen sind immer mit Ungenauigkeiten behaftet**
  (Rauschen, Ablese- oder Digitalisierungs-genauigkeit, Eichung, ...)
- **in der Quantenphysik sind die relevanten Modellvorhersagen und -Größen selbst Parameter von Zufallsverteilungen**
  (mittlere Lebensdauer eines Zustands, Erwartungswert eines Operators, ...)

Im einfachsten Fall:
Modell =
Funktion einer Variablen $x$
mit Modellparametern ***p***

$$y = f(x; \mathbf{p})$$



# Modell einer Messung

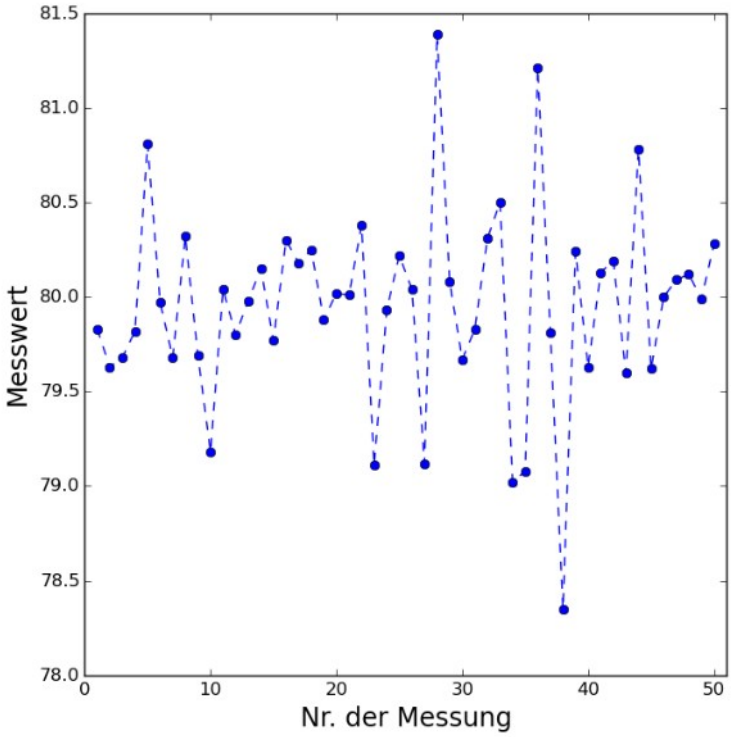



***z* kann viele Ursachen haben:**

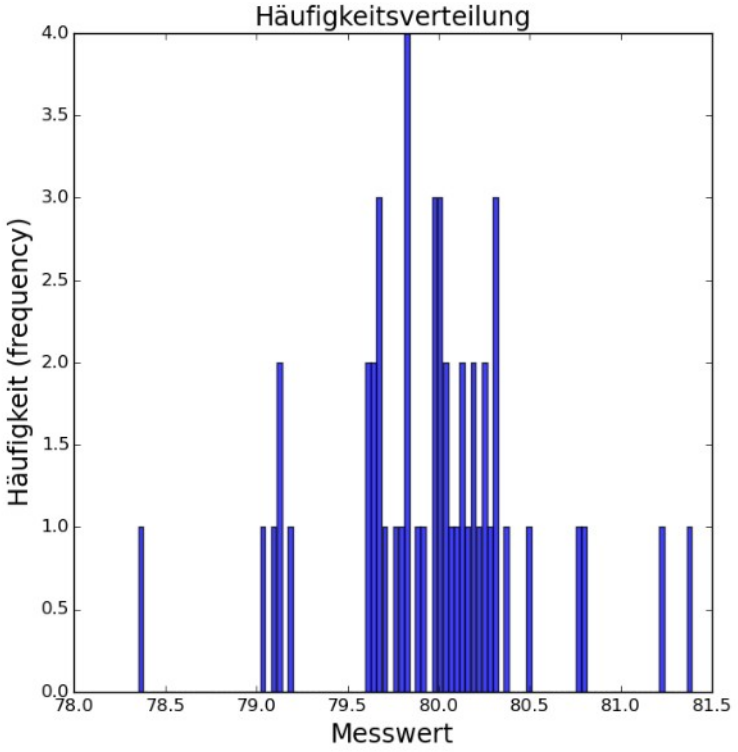
- zufälliger Beitrag zum Messwert („Rauschen“) → „statistische Unsicherheit“
- Genauigkeit des verwendeten Messinstruments → „systematische Unsicherheit“
- mitunter gibt es auch eine Unsicherheit auf den „wahren“ Wert, den man oft z zuschlägt → „theoretische Unsicherheit“
- Fehler im Messprozess – sollten nicht passieren !

$$\Rightarrow \quad z = z_{stat} + z_{syst} + z_{theo}$$

# **Beispiel Längenmessung** (aus Kurs CgDA)




Visualisierung der Messreihe
(typisch wie z.B. mit MS Excel)




Darstellung als
„**Häufigkeitsverteilung**“
oder „**Histogramm**“
(*engl.* frequency distribution“, „histogram“)

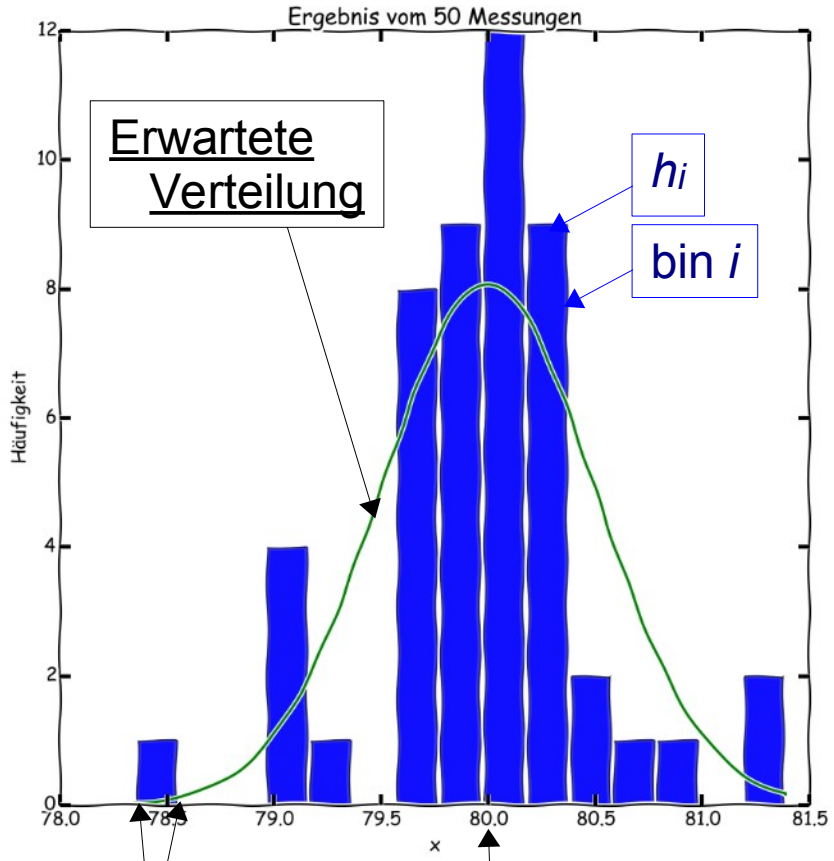
Statistische Information: →

Mittelwert: 79.954
Standardabweichung: 0.522
Unsicherheit auf Mittelwert: 0.074

Anm.: Standardabweichung beschreibt „Breite der Verteilung“

# Häufigkeitsverteilung und Wahrscheinlichkeitsdichte

*Beispiel* **Messen**:
Häufigkeit von N Messergebnissen in bestimmten Intervallen („Bins“)




Erwartete Verteilung

$h_i$

bin $i$

Für eine große Anzahl von Messungen nähert sich die Häufigkeitsverteilung der erwarteten Verteilung immer mehr an

Intervallgrenzen („bin edges“)

Erwartungswert der Verteilung

Häufigkeitsdefinition der Wahrscheinlichkeit:

$$p_i = h_i/N, \;\; N \to \infty$$

# Wahrscheinlichkeit

# Wahrscheinlichkeit

**P(A): Wahrscheinlichkeit für Ereignis A**

unterschiedliche Definitionen (alle erfüllen Kolmogorov-Axiome), z. B.

- **Frequentistisch:** relative Häufigkeit, A in n Messungen k mal zu beobachten (im Grenzfall unendlich vieler Messungen)

$$P(A) = \lim_{n \to \infty} \frac{k}{n}$$

- Nach **Bayes**: Grad der persönlichen Überzeugung („subjektiv“)

(Details siehe weiter unten)

# Einschub: Wahrscheinlichkeit (1)

## 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 $\omega_i$ heißen Ergebnisse des Zufallsexperiments.

- **Beispiele:**
  - Münzwurf $\Omega = \{\text{Kopf}, \text{Zahl}\}$.
  - Würfelwurf $\Omega = \{1, 2, 3, 4, 5, 6\}$.
  - Bestimmung der Fallbeschleunigung (→ wie lautet der Ergebnisraum?).
  - Bestimmung der Anzahl radioaktiver Atome nach einer Zeitspanne $\Delta t$ (→ 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 $\Omega$), $\mathfrak{P}(\Omega)$ heißt Ereignisraum.

# Einschub: Wahrscheinlichkeit (2)

Wahrscheinlichkeit nach Kolmogorov,

nur 3 Axiome:

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 $\Omega$, 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).

# Einschub: Wahrscheinlichkeit

Folgerungen (ohne Beweis, aber einfach zu zeigen):

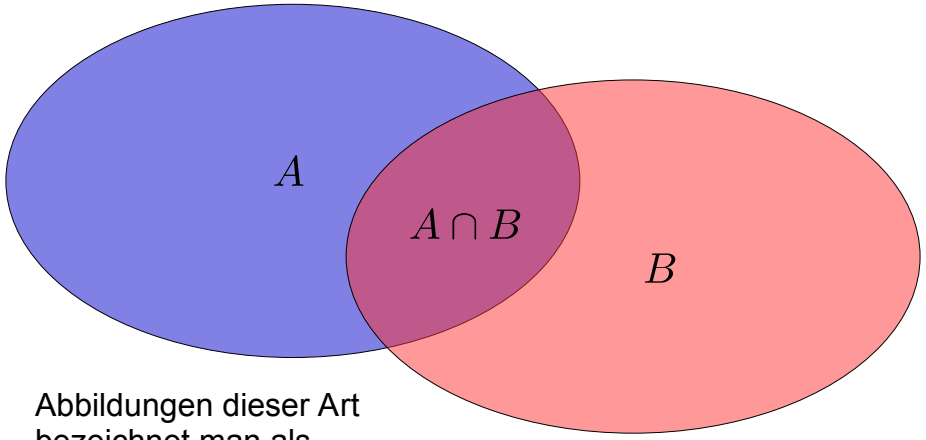
$P(A) \in [0;1]$

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

$P(\emptyset) = 0$ (Unmögliches Ereignis)

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

$A \subset B$ impliziert $P(A) \leq P(B)$




Abbildungen dieser Art bezeichnet man als Venn-Diagramm.

# Einschub: Wahrscheinlichkeit (3)

## Bedingte Wahrscheinlichkeit

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

# Einschub: Wahrscheinlichkeit (4)

## Satz von Bayes

Seien $P$ eine Wahrscheinlichkeitsverteilung über dem Ergebniraum $\Omega$ 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

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

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

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

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

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

Thomas Bayes (1702 – 1761)

wichtig für statistische Interpretation:

Wahrscheinlichkeit, die Daten zu beobachten (A) wenn das Modell (B) richtig ist

→ **Wahrscheinlichkeit für die Richtigkeit des Modells**

Dieser Übergang benötigt die **Bayes'sche Prior-Wahrscheinlichkeit *P*(B)**

# Einschub: Wahrscheinlichkeit (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) Unabhängigkeit 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.

# Zufallsvariable

**Zufallsvariable:** reelle Variable x aus Wahrscheinlichkeitsraum (hier S genannt), die abhängig vom Zufall verschiedene Werte annehmen kann

- Diskret: $P(x_i)$: **Wahrscheinlichkeitsfunktion** (engl.: probability mass function, PMF)

  $\longrightarrow$ Wahrscheinlichkeit, dass x diskreten Wert $x_i$ annimmt

- Kontinuierlich: f(x0) dx = Wahrscheinlichkeit, x in infinitesimalem Intervall $[x_0, x_0 + dx]$ zu finden $\rightarrow$ f(x): **Wahrscheinlichkeitsdichtefunktion** (engl.: probability density function, PDF)

Normierung:

$$\sum_{x_i \in S} P(x_i) = 1 \qquad \int_S f(x)\,dx = 1$$

$F(x_0)$: **Verteilungsfunktion** (engl.: cumulative distribution function, CDF)

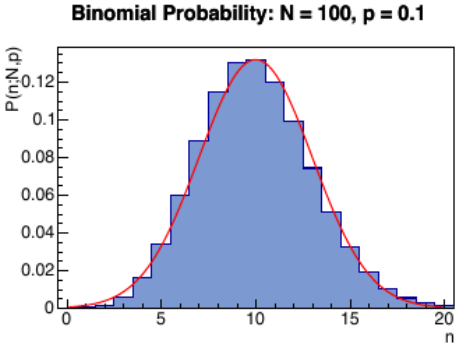
$$F(x_0) = \sum_{x_i \leq x_0} P(x_i) \qquad F(x_0) = \int_{-\infty}^{x_0} f(x)\,dx$$

in python: z = np.random.rand()

(diskrete) Verteilungen
und Verteilungsdichten,

mehrdimensionale Verteilungsdichten
und Kovarianz

# Gauß ↔ Poisson ↔ Binomial - Verteilungen




**Binomialverteilung:**

$$P(n;N,p) = \binom{N}{n} p^n (1-p)^{N-n}$$

$N$ „groß“

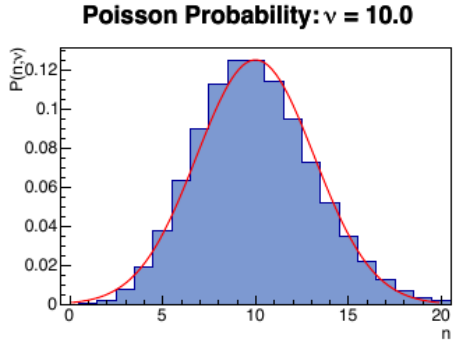
$Np = \mu,\ \sigma^2 = Np(1-p)$

**Gaußverteilung:**

$$G(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

$N \to \infty,\ p \to 0$

$Np = \nu$

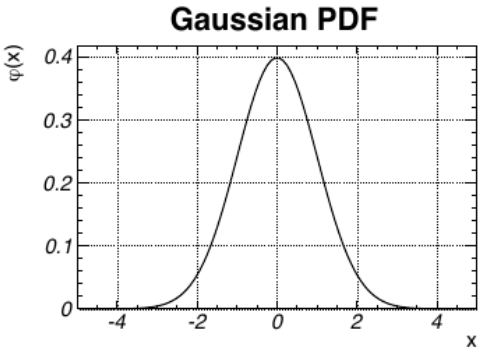



**Poissonverteilung:**

$$P(n;\nu) = \frac{\nu^n}{n!} e^{-\nu}$$

$\nu$ „groß“

$\mu = \nu,\ \sigma^2 = \nu$




für diskrete Verteilungen „diskretisierte Gaußverteilung“ mit $\sigma^2 = \nu$ :

$$P(n;\nu) = \frac{1}{\sqrt{2\pi}\,\nu} \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} \exp\left[-\frac{(n-\nu)^2}{2\nu}\right]$$

# Charakterisierung v. Verteilungen: Formeln

| | Stichprobe | diskrete Verteilung | kontinuierliche Vert. |
|---|---|---|---|
| **Erwartungswert** $\mathrm{E}[x] \equiv \langle x \rangle = \mu$ | $= \frac{1}{N} \sum_{i=1}^{N} x_i$ | $= \sum_{i=1}^{N} x_i p(x_i)$ | $= \int_{-\infty}^{\infty} x f(x)\, dx$ |
| **Varianz** $V \equiv \sigma^2 = \mathrm{E}\left[(x-\mu)^2\right] = \mathrm{E}\left[x^2\right] - \mu^2$ | $= \frac{1}{(N-1)^*} \sum_{i=1}^{N} {x_i}^2 - \mu^2$ | $= \sum_{i=1}^{N} {x_i}^2 p(x_i) - \mu^2$ | $= \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2$ |
| **Schiefe** $\gamma_1 = \mathrm{E}\left[(x-\mu)^3\right] / \sigma^3$ | | | |
| **Kurtosis** $\gamma_2 = \beta_2 - 3$ mit $\beta_2 = \mathrm{E}\left[(x-\mu)^4\right] / \sigma^4$ | | | |
| $\gamma_2 = 0$ für Gauß-Vert. | | | |
| **+ höhere ...** | | | |

* „Bessel-Korrektur“: N → N-1 vermeidet Verzerrung
Plausibilitätsargument: für N=1 ist V nicht definiert !

$$\mathrm{E}(x^n) = \sum_{i=1}^{N} {x_i}^n p(x_i) \text{ bzw. } \int_{-\infty}^{\infty} x^n f(x)\, dx$$

nennt man das n-te Moment der Verteilung

Eine Verteilung ist über ihre Momente vollständig charakterisiert

# **Zentraler Grenzwertsatz** oder warum sind Messfehler Gauß-verteilt ?

Im Grenzfall von großen N ist die Summe von N unabhängigen Zufallszahlen eine Zufallszahl, die einer Gauß-Verteilung folgt.

$$x = \lim_{N\to\infty} \sum_{i=1}^{N} x_i$$

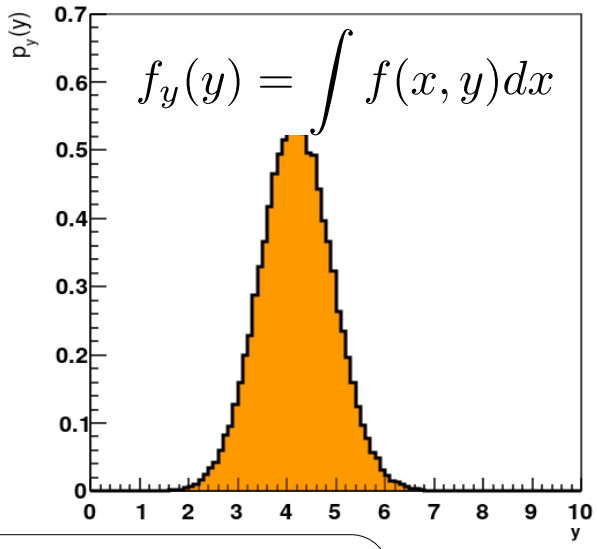
$x_i$ aus beliebiger Verteilung mit Mittelwert $\mu_i$ und **endlicher Varianz** $\sigma_i$

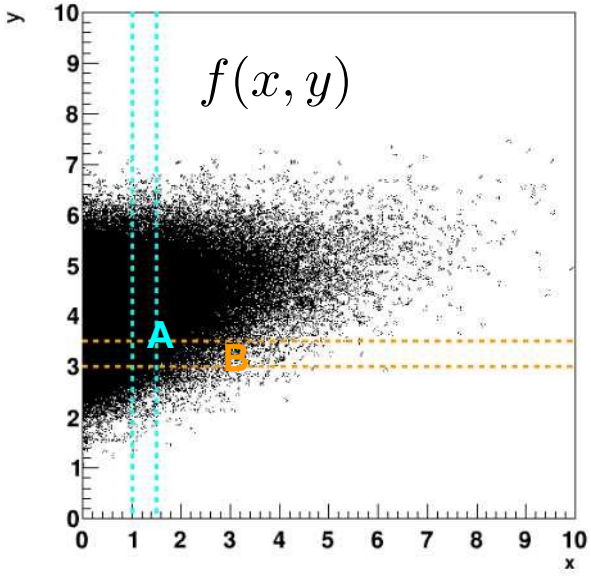
**unter recht schwachen Annahmen** (Lyapunov-Bedingung) **gilt:**

$x$ **ist gaußverteilt mit Erwartungswert** $\mu = \sum_{i=1}^{N} \mu_i$ **und Varianz** $\sigma^2 = \sum_{i=1}^{N} {\sigma_i}^2$

# Mehrdimensionale Verteilungen

***f(x, y)*: Multivariate PDF**
in zwei Dimensionen




Projektion auf x-Achse

Projektion auf y-Achse

**Normierung**

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\mathrm{d}x\mathrm{d}y = 1$$

**kumulativ**

$$\mathcal{P}(x' \le x, y' \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(x',y')\mathrm{d}x'\mathrm{d}y'$$

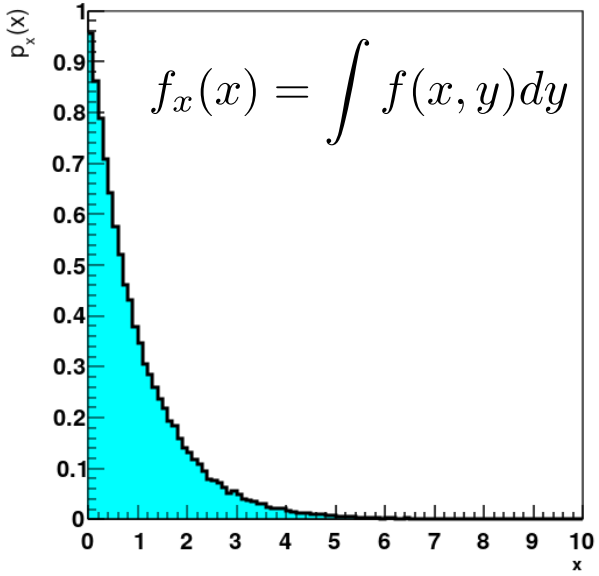
$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y$$

$$\mathrm{f}_y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x$$

**Randverteilungen**
*engl.:* ***marginal distribution***

$$\mathcal{P}_x(x' \le x) = \int_{-\infty}^{x} f_x(x')\mathrm{d}x'$$

kumulativ




- Sind die zwei Zufallsvariablen $x$ und $y$ **unabhängig**, dann gilt für die Wahrscheinlichkeitsdichte: $f(x,y) = f_x(x) \cdot f_y(y)$

# Kovarianz und Korrelation

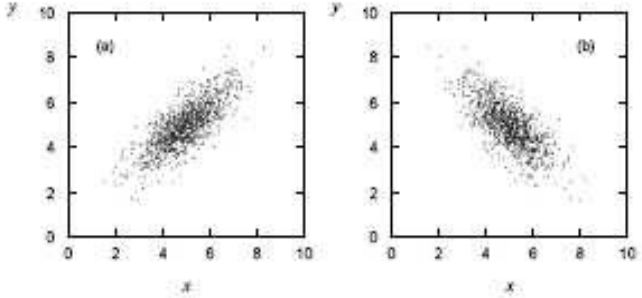
**Definition Kovarianz:** $\mathrm{cov}_{i,j} = E[(x_i - \mu_i)(x_j - \mu_j)] = \ldots = E[x_i x_j] - \mu_i \mu_j$

Normiere so, dass Diagonalelemente = 1

⇒ **Korrelationskoeffizenten :** $\rho_{xy} = \dfrac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}\ , \quad -1 \le \rho_{xy} \le 1$

**Streudiagramme von zwei korrelierten Zufallsvariablen**

$\rho = 0.75$




$\rho = -0.75$

$\rho = 0.95$

$\rho = 0.25$

Korrelationskoeffizienten $\rho_{ij}$ sind anschaulicher als Kovarianzen !

x, y
- unkorreliert ≠ unabhängig
- korreliert ≠ kausal verknüpft
- Korrelation kann zufällig sein,
  * x ⇒ y möglich
  * y ⇒ x möglich
  * z ⇒ (x, y) möglich

# Kovarianzmatrix von korrelierten Messungen

Mehrere ($n$) Messwerte $m_i$ , jede Messung hat
- eine individuelle, zufällige Unsicherheit $z_i$
- sowie eine gemeinsame Unsicherheit $z_g$

$\rightarrow$ $m_i = w + z_i + z_g$

ein gemeinsamer wahrer Wert und $n$+1 Zufallszahlen
*(eine allen Messungen gemeinsame sowie n individuelle)*

Für $n$ Messwerte $m_i$ erhält man die Kovarianz-Matrix, indem man die Varianz der gemeinsamen Unsicherheit $z_g$, $(\sigma_g)^2$, in die Nebendiagonale setzt. Die Diagonalelemente sind die Varianzen der Gesamtunsicherheiten, $(\sigma_i)^2 + (\sigma_g)^2$.

Die vollständige Kovarianzmatrix sieht so aus:

$$V = \begin{pmatrix} {\sigma_1}^2 + {\sigma_g}^2 & {\sigma_g}^2 & \dots & {\sigma_g}^2 \\ {\sigma_g}^2 & {\sigma_2}^2 + {\sigma_g}^2 & \dots & {\sigma_g}^2 \\ & \dots & & \\ {\sigma_g}^2 & {\sigma_g}^2 & \dots & {\sigma_n}^2 + {\sigma_g}^2 \end{pmatrix}$$

# Praktisches Beispiel: Konstruktion d. Kovarianzmatrix

6 Personen in 3 Gruppen messen mit jeweils eigenem Messgerät pro Gruppe, es gibt zusätzlich eine von allen angewandte „Theorie-Korrektur“ mit Unsicherheit.

**Fehlerbeiträge:**

- Systematischer Fehler eines Messgeräts: $\Delta_s$ (korreliert innerhalb einer Gruppe, d.h. Studierende 1-2, 3-4 und 5-6, unabhängig zwischen den Gruppen)
- Theoriefehler: $\Delta_t$ (korreliert für allen Messungen)
- einzelne unabhängige Messunsicherheiten: $\Delta_{f1}, \ldots, \Delta_{f6}$

**Jede Messung hat die Gesamtunsicherheit**

$$\Delta g_i = \sqrt{{\Delta_{f_i}}^2 + {\Delta_s}^2 + {\Delta_t}^2}$$

**Konstruktionsprinzip:**

- Quadrate der *Gesamt-unsicherheiten* stehen in der Diagonalen
- *gemeinsame Unsicherheiten* werden zu den betreffenden Nebendiagonalelementen quadratisch addiert

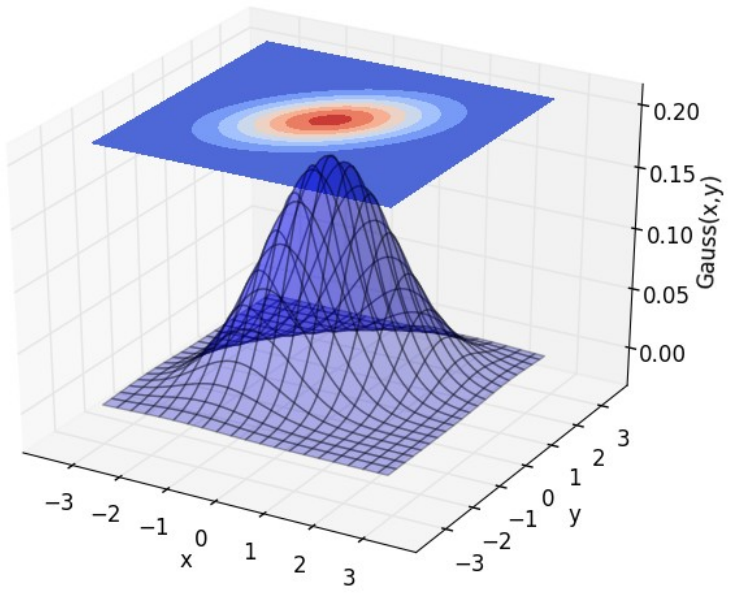
Anm:: unabhängige Unsicher-heiten tauchen nur in der Diagonalen auf

$$V = \begin{pmatrix} {\Delta g_1}^2 & \Delta_s^2 + \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 \\ \Delta_s^2 + \Delta_t^2 & {\Delta g_2}^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 \\ \Delta_t^2 & \Delta_t^2 & {\Delta g_3}^2 & \Delta_s^2 + t^2 & \Delta_t^2 & \Delta_t^2 \\ \Delta_t^2 & \Delta_t^2 & \Delta_s^2 + \Delta_t^2 & {\Delta g_4}^2 & \Delta_t^2 & \Delta_t^2 \\ \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & {\Delta g_5}^2 & \Delta_s^2 + \Delta_t^2 \\ \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_t^2 & \Delta_s^2 + \Delta_t^2 & {\Delta g_6}^2 \end{pmatrix}$$

# Multidimensionale Gaußverteilung

$$p(\vec{x}|\vec{\mu}, V) = \frac{1}{(2\pi)^{\frac{n}{2}}|V|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T V^{-1}(\vec{x}-\vec{\mu})\right]$$

**2d Gauß** $\mu_i=0,\ \sigma_i=1,\ \rho=0.5$




Script: `plot3dGauss.py`

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \cdot & \cdot & \cdot & \sigma_1\sigma_n\rho_{1n} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \cdot & \cdot & \cdot & \sigma_2\sigma_n\rho_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_n\sigma_1\rho_{n1} & \cdot & \cdot & \cdot & \sigma_n\sigma_{n-1}\rho_{n-1} & \sigma_n^2 \end{pmatrix}$$

allg. Kovarianzmatrix
mit Korrelationskoeffizienten $\rho_{ij}$

2-dimensional

$$g(x_1, x_2 \,|\, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\cdot \exp\left(-\frac{1}{2(1-\rho^2)}\left[u_1^2 + u_2^2 - 2\rho u_1 u_2)\right]\right)$$

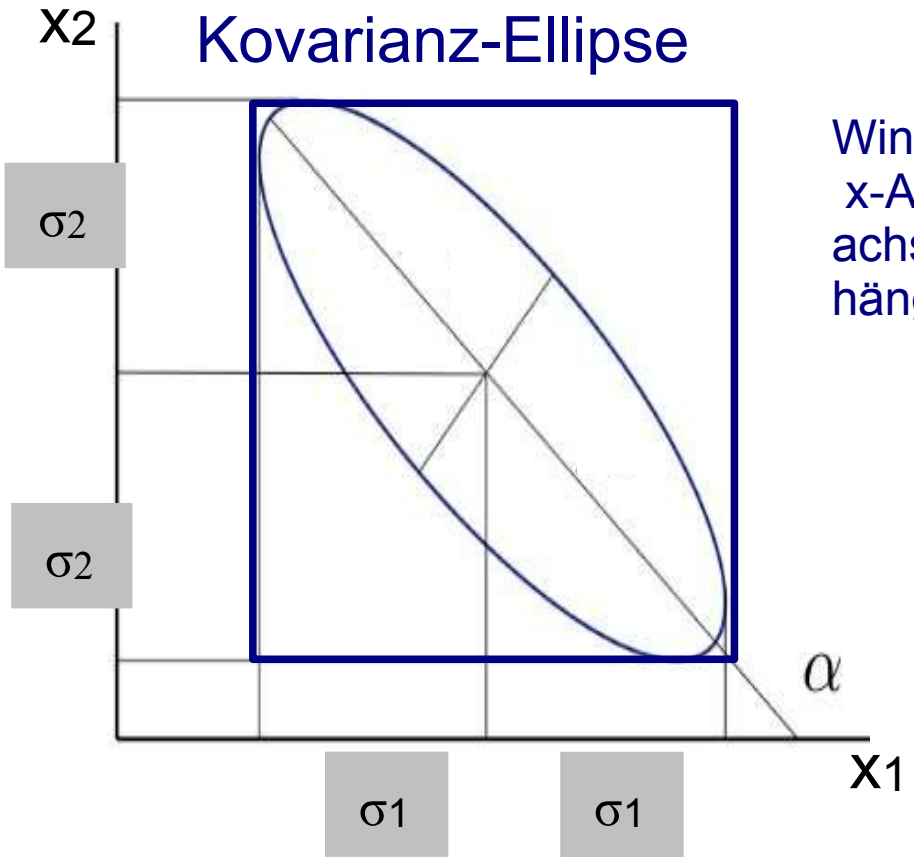
$$\text{mit } u_i = \frac{x_i - \mu_i}{\sigma_i}$$

# Kovarianz-Ellipse

Kontur konstanter Wahrscheinlichkeitsdichte („Höhenlinie“) bei 2d-Gaußverteilung

$$\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_1-\mu_2}{\sigma_2}\right)=1-\rho_{12}^2$$

ist eine **Ellipsengleichung**




Winkel zwischen x-Achse und Haupt-achse der Ellipse hängt von $\rho_{12}$ ab:

$$\tan 2\alpha = \frac{2\rho_{12}\sigma_1\sigma_2}{\sigma_1^2-\sigma_2^2}$$

$\alpha=0°$ für $\rho_{12}=0$

Fläche innerhalb der 1-σ Ellipse entspricht ~39% Wahrscheinlichkeit

gebräuchliche Darstellung bei Problemen mit vielen Variablen:
**Kovarianz-Ellipsen von Variablen-Paaren**

# Kovarianzmatrix u. Fehlerfortpflanzung

Für nicht verschwindende Kovarianzen $V_{ij} = \mathrm{cov}(x_i, x_j)$ lautet das **linearisierte Fehlerfortpflanzungsgesetz**

$$y(x_1, \ldots, x_n) \Rightarrow \sigma_y^2 \simeq \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \frac{\partial y}{\partial x_i} \right) V_{i,j} \left( \frac{\partial y}{\partial x_j} \right)$$

Spezialfälle:

$$y = x_1 \pm x_2$$

$$\Rightarrow {\sigma_y}^2 = {\sigma_1}^2 + {\sigma_2}^2 \pm 2\,\mathrm{cov}(x, y)$$

$$y = x_1 \cdot x_2 \text{ oder } y = \frac{x_1}{x_2}$$

$$\Rightarrow \left( \frac{\sigma_y}{y} \right)^2 \simeq \left( \frac{\sigma_1}{x_1} \right)^2 + \left( \frac{\sigma_2}{x_2} \right)^2 \pm 2 \frac{\mathrm{cov}(x, y)}{xy}$$

Erinnerung CgDA

Monte Carlo (später)

# Monte Carlo-Methode zur Simulation

Die **Monte-Carlo-Methode** (auch „MC-Simulation“)

- abgeleitet vom Namen der durch ihr Spielkasino berühmten Stadt Monte Carlo -

ist eine auf **(Pseudo-)Zufallszahlen basierende numerische Methode** zur

- Integration in hochdimensionalen Räumen
- Bestimmung der Eigenschaften von Verteilungen von Zufallszahlen
  (z.B. von Messgrößen in Experimenten)
- Nachbildung von komplexen Prozessen
  (z.B. in Thermodynamik, Wirtschaft oder von experimentellen Apparaturen, u.v.a.)

**Grundsätzliche Vorgehensweise:**

I. Erzeuge Folge von Zufallszahlen $r_1, r_2, \ldots, r_m$, gleichverteilt in ]0,1].

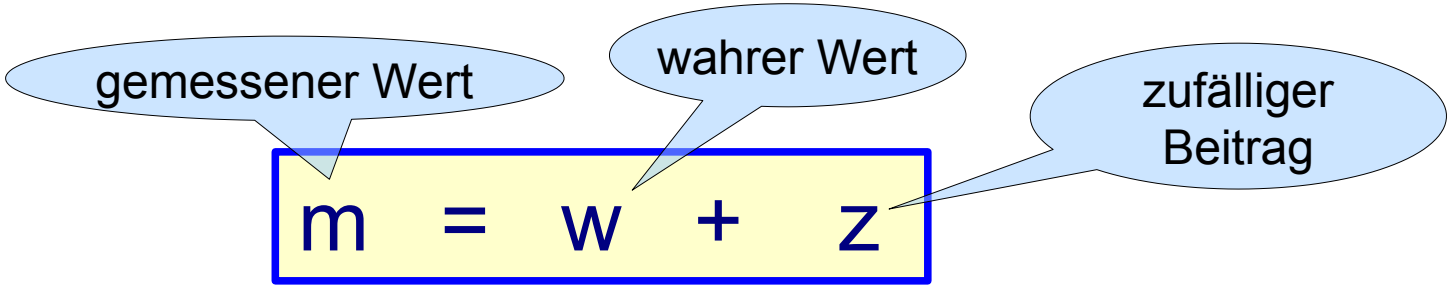
II. Transformiere diese zur Erzeugung einer anderen Sequenz $x_1, x_2, \ldots, x_n$, die einer Verteilungsdichte $f(x)$ folgen (Anm.: $x_i$ können auch Vektoren sein !)

III. Aus den $x_i$ werden dann Eigenschaften von $f(x)$ bestimmt

**Vorteil:** *einfache Operationen mit den $x_i$ ersetzen sehr viel komplexere Operationen auf den Verteilungsdichten*

# Simulierte Daten

Zufallskomponente *z* bei der Gewinnung empirischer Daten ebenfalls mit Zufallszahlen modellierbar:




Funktion (=Modell) des wahren Wertes, ***y = f(x),*** und Modellierung der Messwerte mit MC:

$$m_y = f( w_x + z_x ) + z_y$$

$z_x$ und $z_y$ entsprechen den Unsicherheiten der Messgrößen *x* und *y*

# Erinnerung CgDA

Parameterschätzung –

Methode der kleinsten Quadrate

( → mehr später)

# Parameterschätzung mit der Methode der kleinsten Quadrate

## („$\chi^2$-Methode")

**Minimiere**

„Summe der Residuen-Quadrate" S

bzgl. der $k$ Parameter $\mathbf{p}$

$$S(\mathbf{p}) = \sum_{i=1}^{N} \frac{(y_i - f(x_i, \mathbf{p}))^2}{{\sigma_i}^2}$$

$\sigma_i^2$ sind Varianzen der $N$ Messungen $y_i$

Falls die Fehler korreliert sind, ersetze

$1/\sigma_i^2 \rightarrow V^{-1}$ (Inverse der Kovarianzmatrix)

$S_{\min} = S(\hat{\mathbf{p}})$ folgt

bei gaußverteilten Fehlern $\sigma_i$

einer **$\chi^2$-Verteilung** mit

$n_f = N - k$ Freiheitsgraden

**Erwartungswert:** $\langle S_{\min} \rangle = n_f$

**bzw.** $\langle S_{\min} / n_f \rangle = 1$

**$\chi^2$-Wahrscheinlichkeit:**

$$\chi^2_{\text{prob}}(S_{\min}) := \int_{S_{\min}}^{\infty} \chi^2(s, n_f) ds = 1. - \int_{0}^{S_{\min}} \chi^2(s, n_f) ds$$

dient zur Quantifizierung der Qualität einer Anpassung

**Aussage, mit welcher Wahrscheinlichkeit ein größerer Wert von S am Minimum als der tatsächlich beobachtete zu erwarten wäre.**

# Minimierung von S

- analytisch: $\frac{S}{\partial p_j} = 0\,,\ j = 1, \ldots, k$ als notwendige Bedingung für ein Minimum

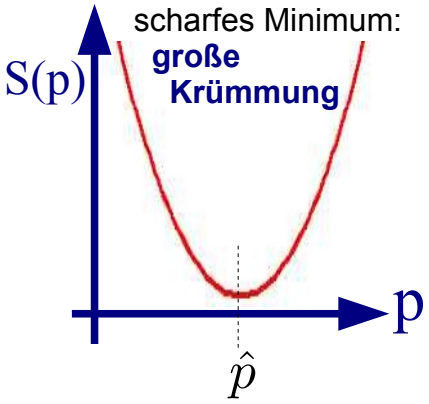
  lösbar in Spezialfällen, z. B. für lineare Probleme $f(x_i, \mathbf{p}) = \sum_{j=0}^{k} p_j\, f_j(x_i)$

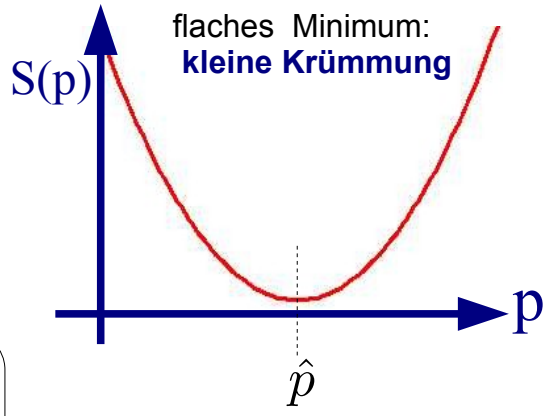
- i. a. numerisch

  „numerische Optimierung: Algorithmen zur Suche nach dem (einem?) Minimum einer Skalaren Funktion im *k*-dimensionalen Parameterraum

z. B. **kafe2:** https://www.etp.kit.edu/~quast/kafe2/htmldoc oder https://philfitters.github.io/kafe2/

## Parameterunsicherheiten



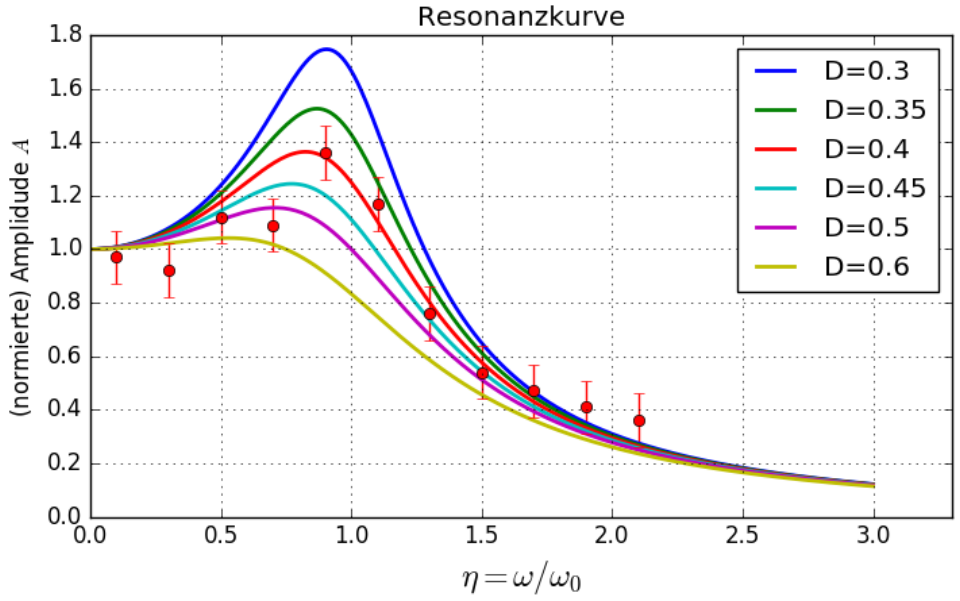

Je schärfer das Minimum von $\chi^2(\boldsymbol{a})$, desto kleiner die Parameterfehler:

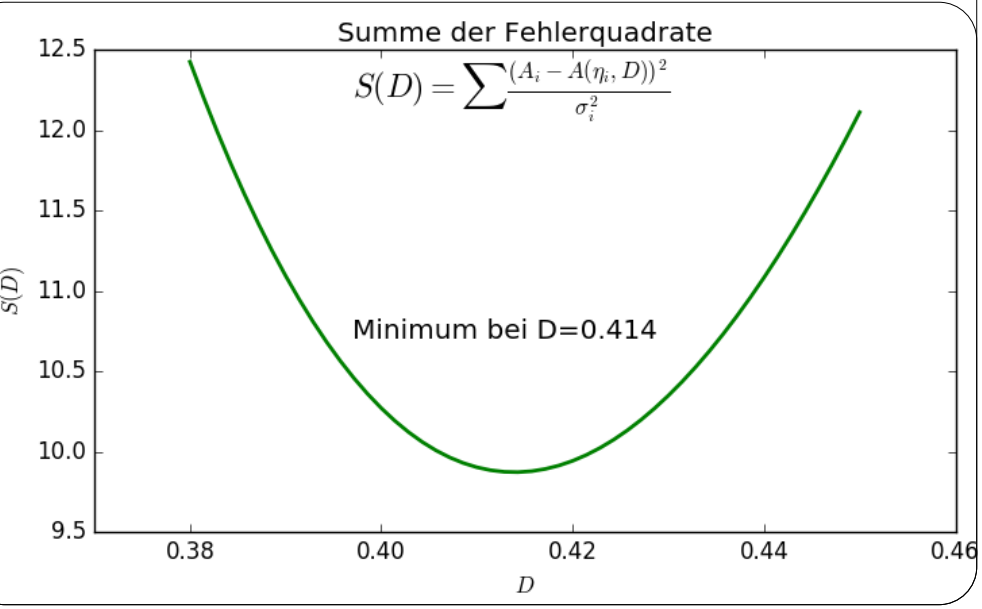
$$\left(V_{\hat{p}}^{-1}\right)_{ij} = \frac{1}{2}\,\frac{\partial^2 S(\mathbf{p})}{\partial p_i \partial p_j}\bigg|_{\hat{p_i}, \hat{p_j}}$$

# Numerische Bestimmung des Minimums

- 11 Messpunkte $A_i$
- Modell „Resonanzkurve“, freier Parameter Dämpfung $D$
- Abstandsmaß

$$S(D) = \sum_{i=1}^{11} \frac{(A_i - A(\eta; D))^2}{\sigma^2}$$




Die Funktion *S(D)* ist (nahezu) parabelförmig mit klarem Minimum bei *D* = 0.414

# Unsicherheiten mit Toy-MC

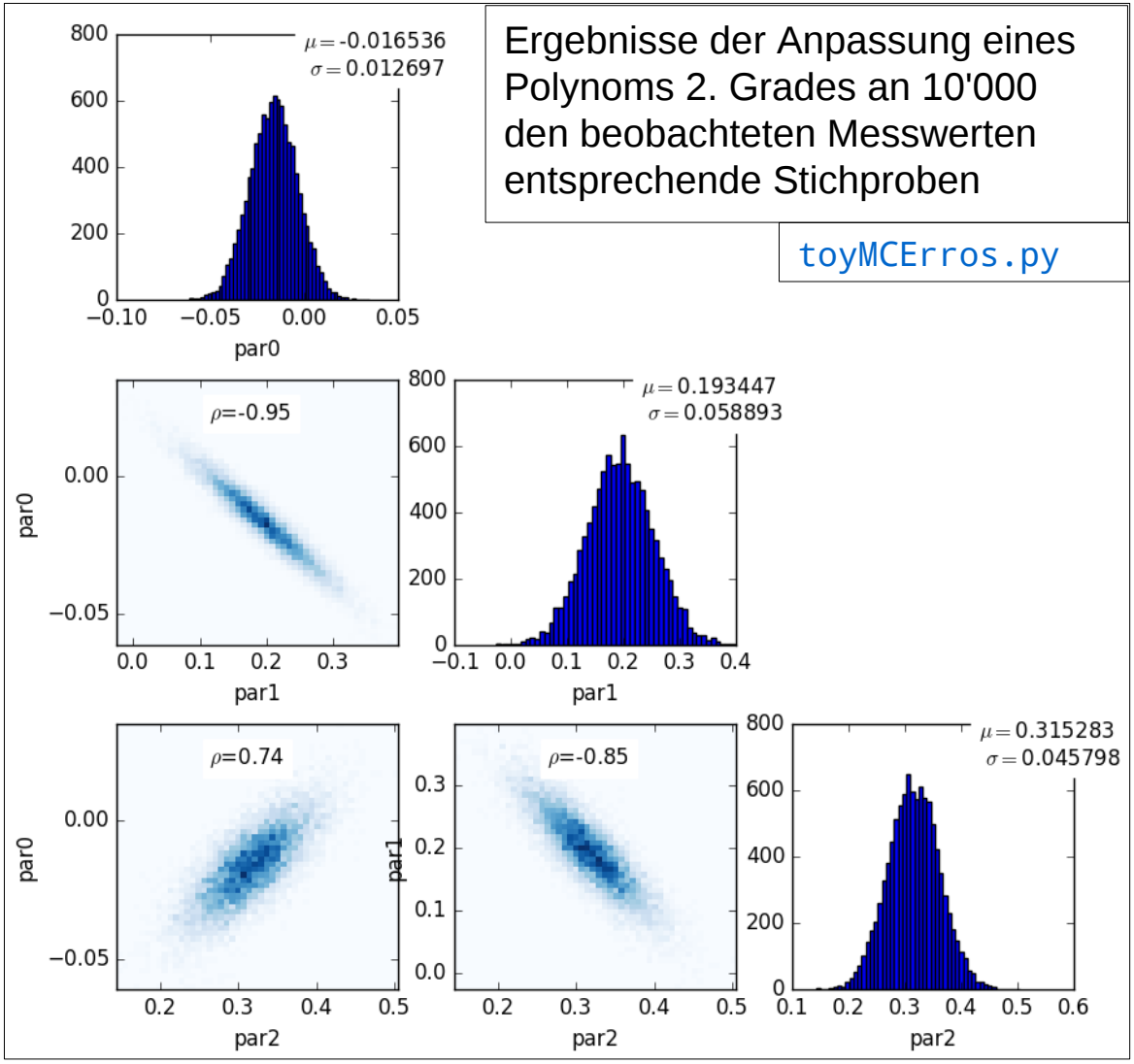
Die Unsicherheiten können ganz allgemein bestimmt werden, indem man die Anpassung für viele Stichproben wiederholt und die Verteilungen der Parameterwerte analysiert.

Auf diese Weise lassen sich
- **Parameterunsicherheiten**
- **Korrelationen der Parameter**

bestimmen.

Das Verfahren ist
**aufwändig**
**aber**
**exakt**

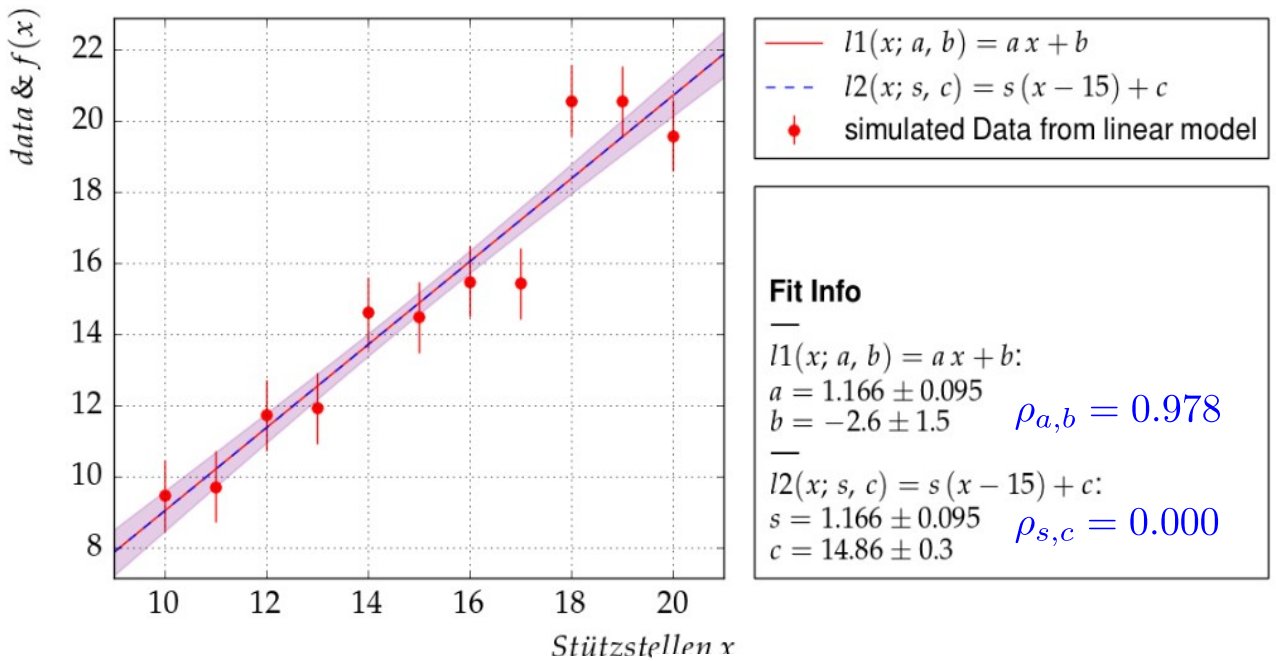
Im Zweifelsfall wird die Exaktheit statistischer Verfahren mit solchen **„Toy – Monte Carlos“** untersucht.




Ergebnisse der Anpassung eines Polynoms 2. Grades an 10'000 den beobachteten Messwerten entsprechende Stichproben

`toyMCErros.py`

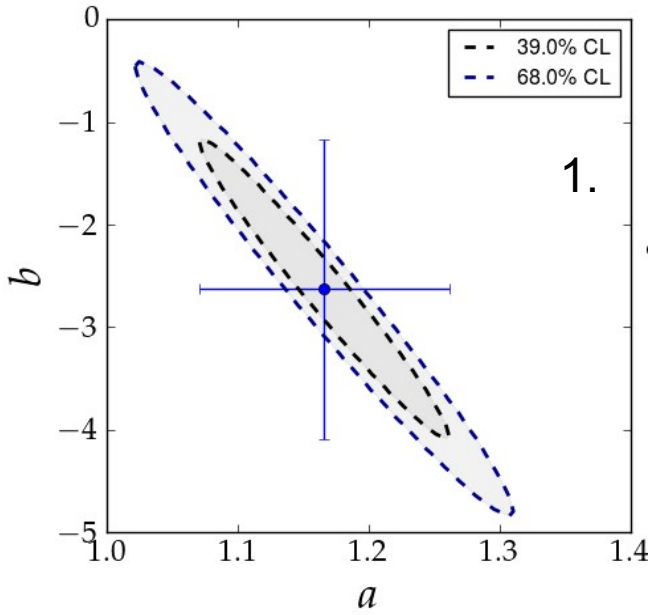
# Geradenanpassung: verschiedene Parametrisierungen

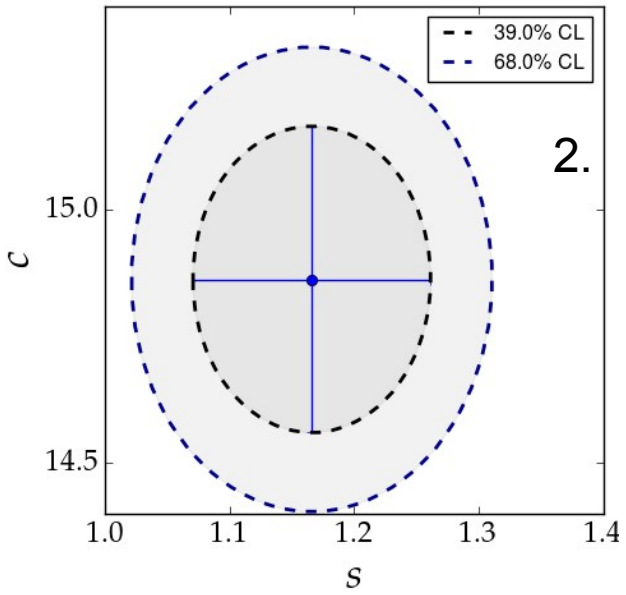



## Parametrisierungen:

1. $y(x) = ax + b$

2. $y(x) = s(x - \bar{x}) + c$




Ergebnisse für das angepasste Modell sind identisch, aber im 1. Fall sind die Parameter sehr stark korreliert.

# Nächste Vorlesung

## Monte Carlo - Methode

- Material:
  V03a_MonteCarlo.pfd