Karlsruher Institut für Technologie

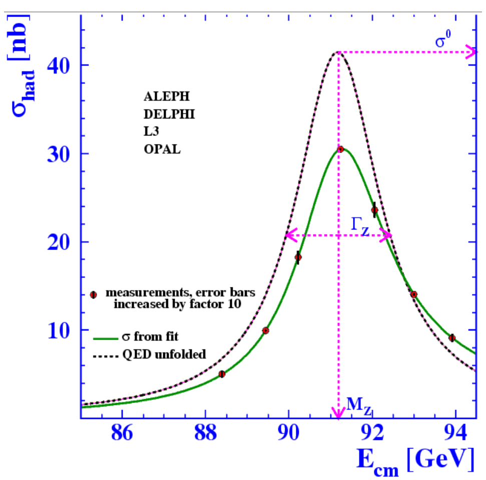
Moderne Methoden der Datenanalyse

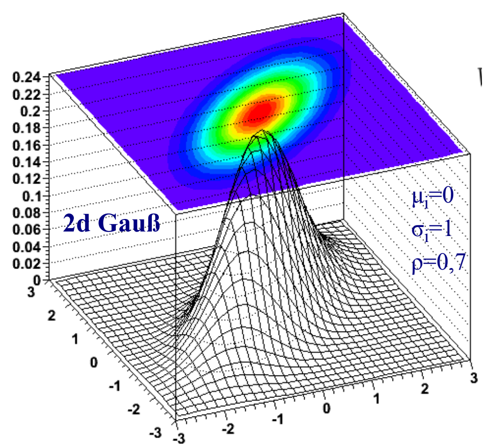
Zusammenfassung Einführung (2)

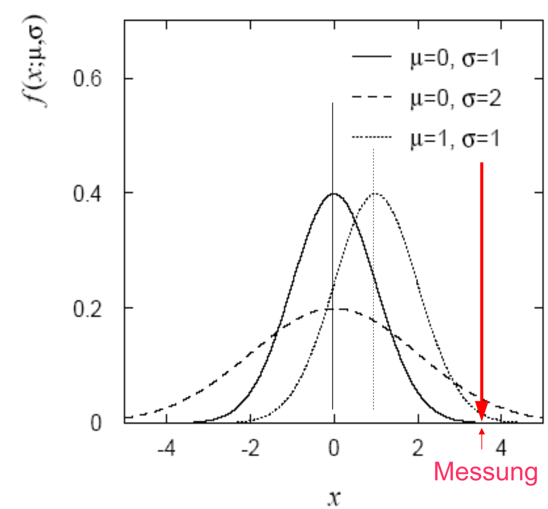
Andreas B.Meyer

2. Mai 2017

Fakultät für Physik
Institut für Experimentelle Kernphysik - IEKP







Korrelation in zwei Dimensionen

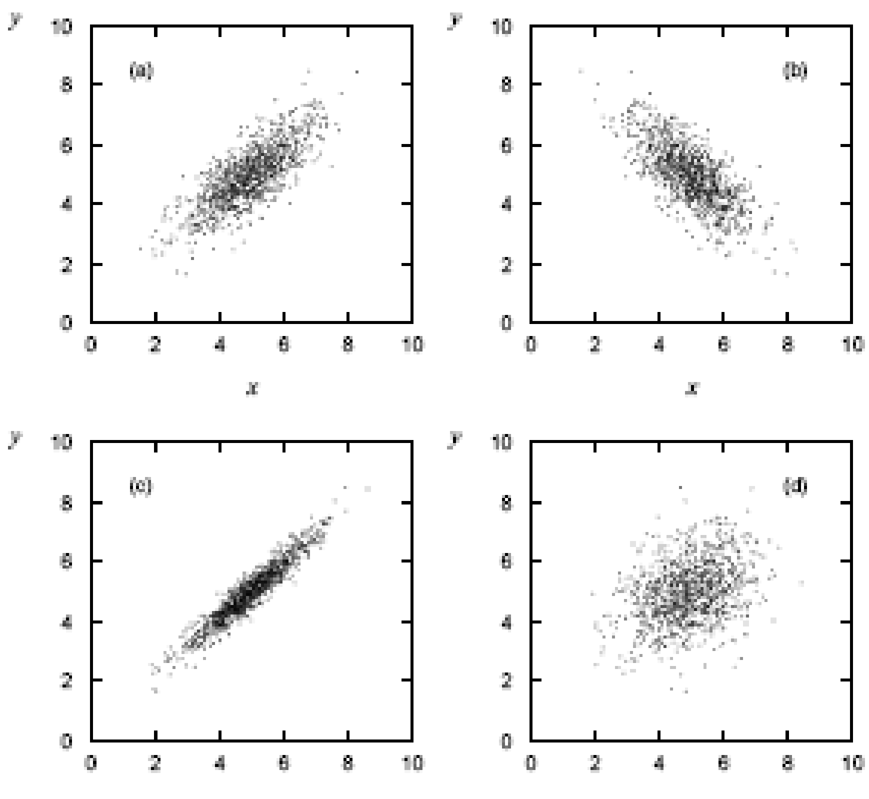
- Kovarianzmatrix V

$$V = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix}$$

- Korrelationsmatrix: Transformiere V so, dass Diagonalelemente alle 1. Dann sind ρ_{xy} die Korrelationskoeffizienten.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \qquad -1 \leq \rho_{xy} \leq 1$$

$\rho = 0.75$ $\rho = -0.75$



$\rho = 0.95$ $\rho = 0.25$

Blobel/Lohrmann Abschnitt 4.8

Beispiel: Gauß-Verteilung in n Dimensionen

- Wahrscheinlichkeitsdichte der Gauß-Verteilung:

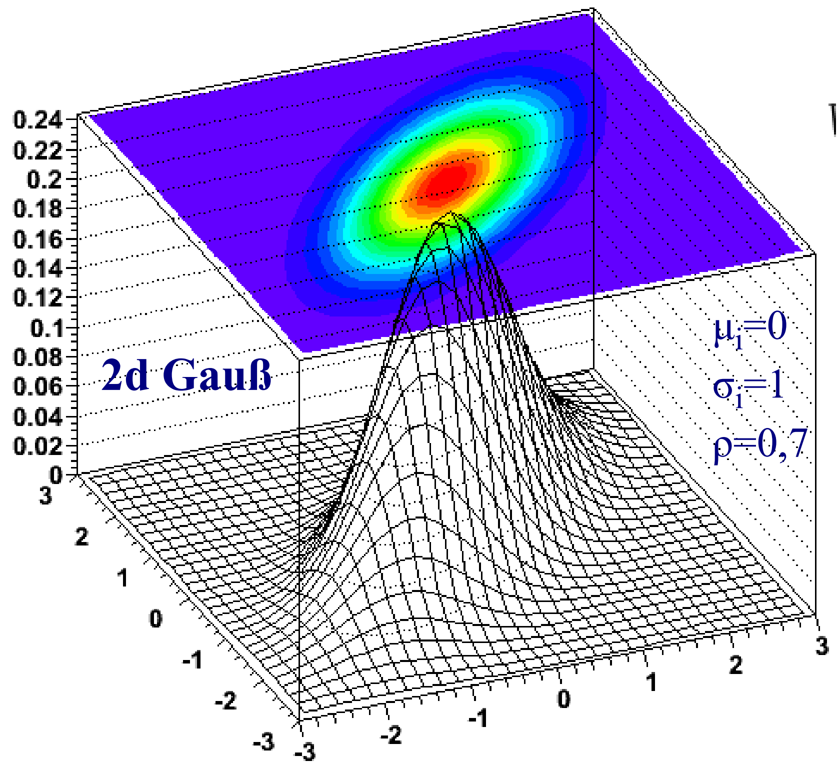
$$P(\vec{x},\vec{\mu},V) = \frac{1}{(2\pi)^{n/2}\sqrt{|V|}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T V^{-1}(\vec{x}-\vec{\mu})\right)$$

- Nicht-Diagonalelemente von V erzeugen Korrelation
- In 2 Dimensionen:

$$V = \begin{pmatrix} \sigma_1^2 & \rho\,\sigma_1\sigma_2 \\ \rho\,\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$|V| = \det(V) = \sigma_1^2\sigma_2^2 - (\rho\sigma_1\sigma_2)^2$$

$$V^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \times \begin{pmatrix} \sigma_2^2 & -\rho\ \sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$



Fehlerfortpflanzung mit mehreren Variablen

Karlsruher Institut für Technologie

- Messung besteht häufig aus Kombination mehrerer Größen in eine einzelne $y = y(x_1, \ldots x_n)$. B ist dann eine 1 x n Matrix
- Beispiel für m=1 und n=2:

$$
\begin{aligned}
V[y] &= B \cdot V[\vec{x}] \cdot B^T \\
&= \begin{pmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} \end{pmatrix} \cdot \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{pmatrix} \\
&= \left(\frac{\partial y}{\partial x_1} \right)^2 \sigma_1^2 + 2 \frac{\partial y}{\partial x_1} \frac{\partial y}{\partial x_2} \sigma_{12} + \left(\frac{\partial y}{\partial x_2} \right)^2 \sigma_2^2
\end{aligned}
$$

- Falls die $x_1 \ldots x_n$ unkorreliert, dann sind die $\sigma_{i \neq j} = 0$, d.h. V ist diagonal

$$
V[y_j] = \sigma_{y_j}^2 = \sum_{i=1}^{n} \left(\frac{\partial y_j}{\partial x_i} \right)^2 \sigma_{x_i}^2
$$

Fehlerfortpflanzungsgesetz für <u>unkorrelierte</u> Variablen x_i

"Goodness of Fit" (am Beispiel: χ^2 - Verteilung)

Karlsruher Institut für Technologie

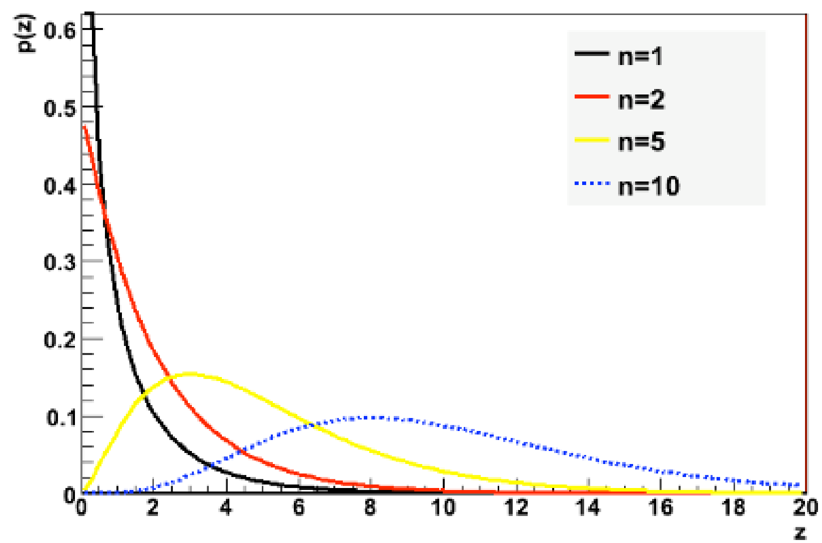
- Gegeben eine Vorhersage: Wie gut stimmt sie mit den Daten überein ?
- Beispiel: Gewichtete Summe der Abstands- (auch: Residuen-) quadrate zwischen Daten (mit Werten x_i und Unsicherheiten σ_i) und Funktionswerten $f(x_i)$

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

N Messwerte y_i
k Parameter p

- Bei Gauß-förmig verteilten y_i mit Unsicherheiten σ_i folgt S einer χ^2-Verteilung mit $n = N - k$ Freiheitsgraden

$$f_n(\chi^2) = \frac{\frac{1}{2}\left(\frac{\chi^2}{2}\right)^{\frac{n}{2}-1} e^{-\frac{\chi^2}{2}}}{\Gamma\left(\frac{n}{2}\right)}$$



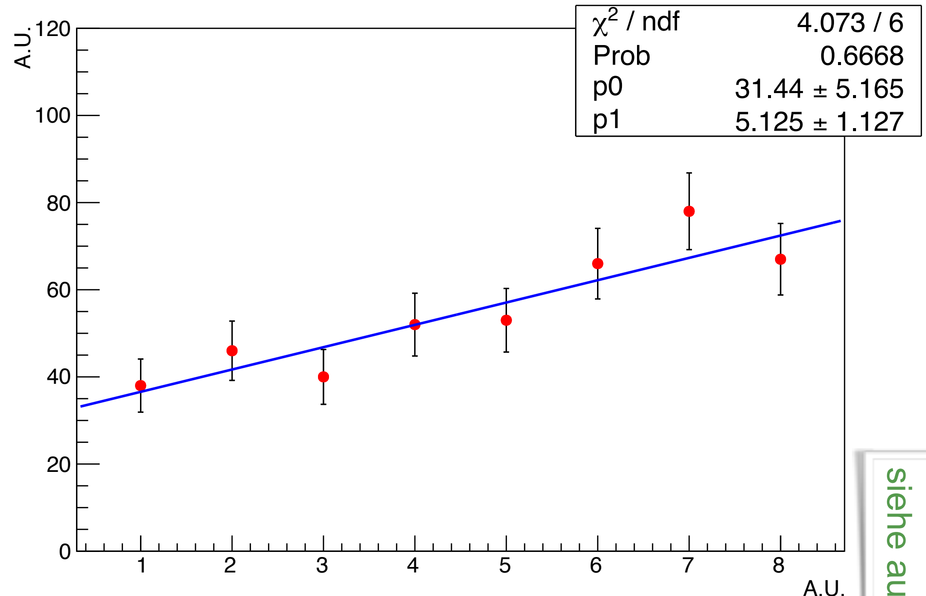
- Erw.wert: $\boxed{\langle \chi^2 \rangle = n}$ $\Rightarrow \langle \chi^2/n \rangle = 1$
- Varianz: $\boxed{V[\chi^2] = 2n}$

Parameterschätzung

- Beispiel: Geraden-Anpassung (2 Parameter)

Beispiel: fitf1.C

- Ansatz: $f(x) = p_0 + p_1\, x$



- Ergebnis:
  - χ^2/ndof: 4.1/6, p-Wert: 66%
  - Kovarianz: $\begin{pmatrix} 26.7 & -5.07 \\ -5.07 & 1.27 \end{pmatrix}$
  - Korrelation: $\begin{pmatrix} 1 & -0.86 \\ -0.86 & 1 \end{pmatrix}$

- Bemerke: Parameter p_0 und p_1 sind stark (anti-) korreliert!

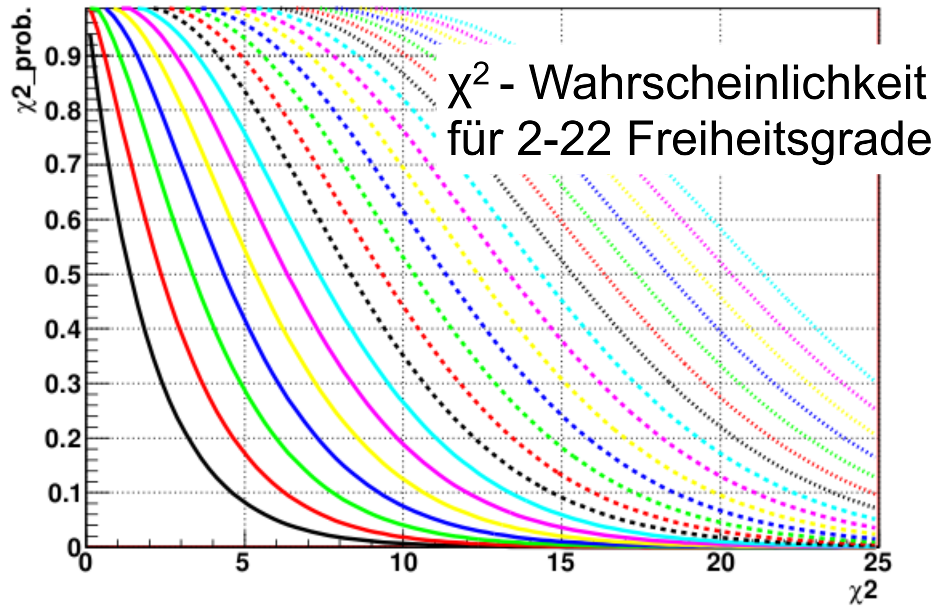
(Anschaulich: Größeres p_0 verlangt kleineres p_1, so dass $f(x)$ die Daten noch beschreibt)

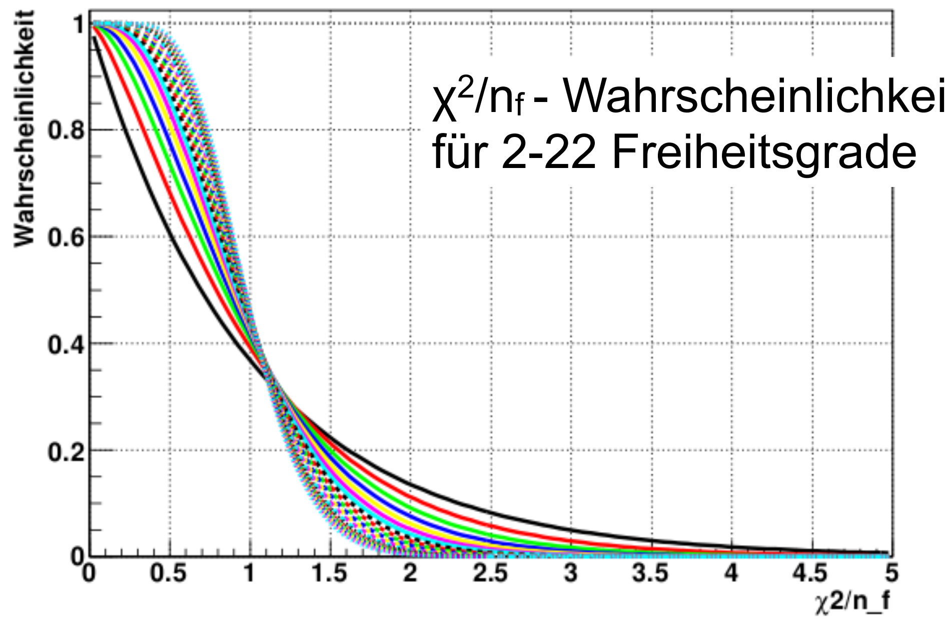
siehe auch Cowan Abschnitt 7.3

χ^2 - Wahrscheinlichkeit

$$\chi^2_{\text{Prob}} = \int_{\chi^2}^{\infty} f_n(v)dv$$
$$= 1 - \int_0^{\chi^2} f_n(v)dv$$

- dient zur Quantifizierung der Übereinstimmung
- Aussage, mit welcher Wahrscheinlichkeit ein größerer Wert für χ^2 als der tatsächlich gemessene zu beobachten wäre
- Wahrscheinlichkeit für χ^2/n einen Wert > 1 zu beobachten ist ~40% (ziemlich unabhängig von n)



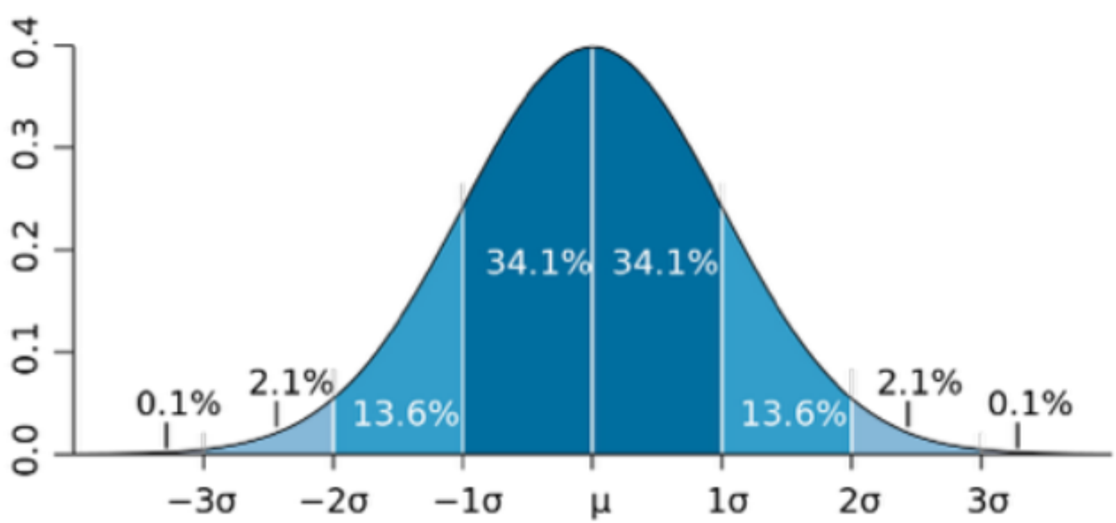


In Root: TMath:chisquared_cdf()

Quantile der Gauß-Verteilung

- Angabe von Konfidenz- und Signifikanz-Niveaus durch Quantile der Gauß-Verteilung: Statt p-Wert wird häufig “sigma” verwendet
- Äquivalente Wahrscheinlichkeit einer Normalverteilung ausgedrückt in Einheiten der Standardabweichung σ

$$\alpha = 1 - \int_{-\delta}^{\delta} \mathrm{Gauss}(x; \mu = 0, \sigma = 1)\, dx = 1 - \mathrm{erf}\left(\frac{\delta}{\sqrt{2}}\right)$$



$1-\alpha$	α	δ
$\boxed{0.683}$	$\boxed{0.317}$	$\boxed{1\sigma}$
0.90	0.1	1.65σ
0.95	0.05	1.96σ
0.9545	0.0455	2σ
0.9973	0.0027	3σ
	5.7×10^{-7}	5σ

Maximum Likelihood

- Likelihood-Funktion:
  - Die Werte ***x*** = *(x_i, … x_n)* seien *n* unabhängige Messpunkte
  - Das Produkt der Werte der Wahrscheinlichkeitsdichten *P(**x**|**a**)*

$$\boxed{\mathcal{L}(\mathbf{a}) = P(x_1|\mathbf{a}) \cdot P(x_2|\mathbf{a}) \cdot \ldots \cdot P(x_n|\mathbf{a}) = \prod_{i=1}^{n} P(x_i|\mathbf{a})}$$

hängt nur von den Parametern ***a*** ab.

- Maximum-Likelihood Prinzip:
  - Der beste Schätzwert ***â*** für den Parametervektor ***a*** ist der, der die Likelihood-Funktion $\mathcal{L}$(***a***) maximiert
  - In der Praxis: minimiere den negativen Logarithmus:

$$\boxed{F(\mathbf{a}) = -\ln \mathcal{L}(\mathbf{a}) = -\sum_{i=1}^{n} \ln\left(P(x_i|\mathbf{a})\right)}$$

Vergleich: ML und LS

- Fehlerbestimmung: $\Delta(-\ln L)$ und $\Delta\chi^2$: in Quantilen der Gauß-Verteilung

$$\boxed{\Delta(-\ln L) = \frac{1}{2}\Delta\chi^2}$$

	$\Delta(-\ln L)$	$\Delta\chi^2$
1σ	0.5	1
2σ	2.0	4
3σ	4.5	9
$n\sigma$	$n^2/2$	n^2

	Maximum Likelihood	Kleinste Quadrate
Methode	Höhe der PDF	Abweichung vom Mittelwert
Voraussetzungen	PDF ist bekannt	Mittelwert und Varianz
Effizienz	maximal	maximal bei linearen Problemen
Komplexität	aufwändig	oft exakt analytisch lösbar
Güte der Anpassung	nein	ja:χ^2-Wahrscheinlichkeit
Robustheit	nein	nein