

Datenanalyse in der Physik

Vorlesung 8

Methode der kleinsten Quadrate (χ^2 -Methode)

Prof. Dr. J. Mnich

DESY und Universität Hamburg



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 1

Methode der kleinsten Quadrate

Die Methode der kleinsten Quadrate oder χ^2 -Methode kann man ableiten als Spezialfall der Maximum-Likelihood-Methode für gaußverteilte Zufallsvariablen

Nebenbemerkung: Die Methode ist auch für andere Verteilungen anwendbar unter Beibehaltung bestimmter, guter Eigenschaften der Parameterschätzung

- Wir haben n Messwerte x_i vorliegen, die gaußverteilt sein sollen

$$f(x_i; a) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i - a)^2}{2\sigma_i^2}}$$

Die Fehler der einzelnen Messungen σ_i können dabei unterschiedlich sein
Der zu schätzende Parameter ist hier zunächst einfach der Mittelwert $a = \langle x \rangle$

- Bilden der Likelihood-Funktion

$$L(a) = \prod_{i=1}^n f(x_i; a) \quad \text{bzw.} \quad \ell(a) = -\ln L(a) = -\sum_{i=1}^n \ln f(x_i; a)$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 2

Methode der kleinsten Quadrate

Einsetzen der Gauß-Funktion

$$\begin{aligned}\ell(a) &= -\sum_{i=1}^n \ln \left(\frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i - a)^2}{2\sigma_i^2}} \right) \\ &= \underbrace{-\sum_{i=1}^n \ln \frac{1}{\sqrt{2\pi\sigma_i^2}}}_{= \text{const}} + \frac{1}{2} \underbrace{\sum_{i=1}^n \left(\frac{x_i - a}{\sigma_i} \right)^2}_{\equiv \chi^2(a)}\end{aligned}$$

⇒ Wichtiger Zusammenhang zwischen dem Logarithmus der Likelihood- und der χ^2 -Funktion

$$\ell(a) = \text{const} + \frac{1}{2} \chi^2(a)$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 3

Methode der kleinsten Quadrate

Aus diesem Zusammenhang folgt nun sofort für die Parameterschätzung mit der χ^2 Methode

Schätzung des Parameters a

Der Maximum-Likelihood-Methode, d.h. der Minimierung von $\ell(a)$, entspricht die Minimierung der χ^2 -Funktion (kleinste Quadrate)

Bestimmung des Parameterfehler

Fehler können wiederum über Differenzen der χ^2 -Funktion zu ihrem Minimum bestimmt werden

Aber: Der Faktor $\frac{1}{2}$ vor dem χ^2 bedeutet, dass die Abstände zum Minimum, die bestimmten Konfidenzintervallen entsprechen, im Vergleich zur Likelihood-Methode verdoppelt werden müssen

Beispiel: die 1- σ Fehler (68% Konfidenzintervalle) ergeben sich aus der Bedingung

$$\chi^2(a \pm \sigma) = \chi_{\min}^2 + 1$$

was äquivalent ist zu

$$\ell(a \pm \sigma) = \ell_{\min} + \frac{1}{2}$$

Achten sie darauf, wenn sie z.B. MINUIT verwenden



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 4

Methode der kleinsten Quadrate

- Wir hatten die χ^2 -Funktion eingeführt (Vorlesung 3) als Quadratsumme

$$\chi^2 = \sum_{i=1}^n x_i^2$$

der Zufallsvariablen x_i , die einer standardisierten Gauß-Funktion ($\langle x \rangle = 0, \sigma^2 = 1$) folgen

- Wie man durch Parametertransformation leicht zeigen kann, ist die allgemeine Definition

$$\chi^2 = \sum_{i=1}^n \left(\frac{x_i - \langle x \rangle}{\sigma_i} \right)^2$$

d.h. die Summe der Quadrate der mit dem reziproken Fehler gewichteten Abweichungen vom Mittel- oder Erwartungswert

- Man bezeichnet die Differenzen $x_i - \langle x \rangle$ auch als Residuen



Einfaches Beispiel zur χ^2 -Methode

- Wir haben n gaußverteilte Messungen x_i einer Größe vorliegen
- Die Messungen sollen unkorreliert sein, d.h. die x_i sind unabhängige Zufallsvariablen, und ihre Fehler σ_i können unterschiedlich sein
- Den besten Schätzwert für den Mittelwert μ erhält man durch Minimierung von

$$\chi^2(\mu) = \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma_i} \right)^2$$

$$\Rightarrow \frac{d}{d\mu} \chi^2(\mu) = -2 \sum_{i=1}^n \frac{x_i - \mu}{\sigma_i^2} = 0$$

$$\Rightarrow \sum_{i=1}^n \frac{x_i}{\sigma_i^2} = \mu \sum_{i=1}^n \frac{1}{\sigma_i^2} \Rightarrow \mu = \frac{\sum_{i=1}^n \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

Im Mittelwert werden die Einzelmessungen mit der reziproken Varianz (Fehlerquadrat) gewichtet

Bei gleichen Fehlern ($\sigma_i = \sigma$) gilt

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$



Einfaches Beispiel zur χ^2 -Methode

Was erwarten wir für den Fehler des Mittelwertes σ_μ ?

Wir können die Gleichung $\mu = \sum_{i=1}^n \frac{x_i}{\sigma_i^2} / \sum_{i=1}^n \frac{1}{\sigma_i^2}$ auch als

Parametertransformation des Zufallsvektors $\vec{x}$ auf die Zufallsvariable μ auffassen

⇒ Fehlerfortpflanzung

$$V'(\mu) = B V(\vec{x}) B^T \quad \text{mit} \quad B = \left(\frac{\partial \mu}{\partial x_1} \quad \frac{\partial \mu}{\partial x_2} \quad \dots \quad \frac{\partial \mu}{\partial x_n} \right) = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \left(\frac{1}{\sigma_1^2} \quad \frac{1}{\sigma_2^2} \quad \dots \quad \frac{1}{\sigma_n^2} \right)$$

$$V'(\mu) = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \left(\frac{1}{\sigma_1^2} \quad \frac{1}{\sigma_2^2} \quad \dots \quad \frac{1}{\sigma_n^2} \right) \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix} \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \begin{pmatrix} \frac{1}{\sigma_1^2} \\ \frac{1}{\sigma_2^2} \\ \vdots \\ \frac{1}{\sigma_n^2} \end{pmatrix}$$

$$= \left(\frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \right)^2 \left(\frac{1}{\sigma_1^2} \quad \frac{1}{\sigma_2^2} \quad \dots \quad \frac{1}{\sigma_n^2} \right) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} = \sigma_\mu^2$$

$$\frac{1}{\sigma_\mu^2} = \sum_{i=1}^n \frac{1}{\sigma_i^2}$$

Diesen Fehler sollte man auch durch $\chi_{\min}^2 + 1$ erhalten (im Mittel!) wenn die x_i tatsächlich gaußverteilt sind



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 7

Verallgemeinerung der χ^2 -Methode

Man kann mit der χ^2 -Methode nicht nur Mittelwerte schätzen, sondern allgemein Parameter, die über eine Funktion mit der Messgröße verknüpft sind

- Seien x_i die n gaußverteilten Messungen mit den Fehlern σ_i und $X_i = X_i(a)$ die in einem Modell erwarteten Werte als Funktion des zu schätzenden Parameters a
- Dann lautet die zu minimierende χ^2 -Funktion

$$\chi^2(a) = \sum_{i=1}^n \left(\frac{x_i - X_i(a)}{\sigma_i} \right)^2$$

- Beispiel:

Wir betrachten ein Experiment zum Fallgesetz $s = \frac{1}{2}gT^2$ um die Erdbeschleunigung g zu bestimmen

Dazu werden n Messungen s_i des Fallweges zu bestimmten Zeiten T_i gemacht mit den unkorrelierten Fehlern σ_{s_i}

(die T_i seien in diesem einfachen Fall beliebig genau bekannt)

$$\Rightarrow \chi^2(g) = \sum_{i=1}^n \left(\frac{s_i(T_i) - S(T_i; g)}{\sigma_{s_i}} \right)^2 = \sum_{i=1}^n \left(\frac{s_i(T_i) - \frac{1}{2}gT_i^2}{\sigma_{s_i}} \right)^2$$

aus der durch Minimierung der Schätzwert für g mit dem Fehler σ_g folgt



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 8

Verallgemeinerung der χ^2 -Methode

- Die Messungen x_i müssen nicht notwendigerweise die gleichen physikalischen Größen sein

Natürlich müssen die Erwartungen $X_i(a)$ für die Messwerte x_i dann mit den entsprechenden Funktionen des physikalischen Modells berechnet werden

- Z.B. könnten wir zum χ^2 des Fallversuches Messungen eines Fadenpendel-Experimentes dazufügen, um die Bestimmung von g zu verbessern

⇒ Die χ^2 -Methode erlaubt auf einfache Weise Messungen zu kombinieren

- Alles bisher zur χ^2 -Methode gesagte ist einfach auf mehrere Parameter a_i zu verallgemeinern

- Bemerkung:
Auch mit der Maximum-Likelihood-Methode kann man natürlich mehrere, beliebige Parameter schätzen und Messungen kombinieren



Matrix-Schreibweise

Im allgemeinen Fall der Bestimmung von mehreren Parametern aus mehreren Messungen empfiehlt sich die Benutzung der Matrix-Schreibweise

- Wir fassen die Messwerte x_i und die Erwartungen $X_i(\vec{a})$ als n -komponentige und die Parameter a_i als m -komponentigen Spaltenvektor auf:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \vec{X}(\vec{a}) = \begin{pmatrix} X_1(\vec{a}) \\ X_2(\vec{a}) \\ \vdots \\ X_n(\vec{a}) \end{pmatrix} \quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

- Die Zufallsvariablen x_i sollen zunächst unkorreliert sein d.h. die Kovarianzmatrix $V(\vec{x})$ und deren inverse Matrix $V^{-1}(\vec{x})$ sind diagonal:

$$V(\vec{x}) = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix} \quad V^{-1}(\vec{x}) = \begin{pmatrix} 1/\sigma_1^2 & & & \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ & & & 1/\sigma_n^2 \end{pmatrix}$$

- Damit können wir die χ^2 -Funktion schreiben als

$$\chi^2(\vec{a}) = (\vec{x} - \vec{X}(\vec{a}))^T V^{-1}(\vec{x}) (\vec{x} - \vec{X}(\vec{a}))$$



Kovarianzmatrix der gefitteten Parameter

Maximum-Likelihood-Methode:

Erlaubt nicht nur Berechnung der Fehler der Parameter $\vec{a}$, sondern auch der zugehörigen vollständigen Kovarianzmatrix $V(\vec{a})$

In der parabolischen Näherung gilt:

$$G = V^{-1}(\vec{a}) \quad \text{mit} \quad G_{ij} = \frac{\partial^2}{\partial a_i \partial a_j} \ell(\vec{a})$$

χ^2 -Methode:

Aus dem Zusammenhang zwischen negativen Logarithmus der Likelihood-Funktion und χ^2 -Funktion

$$\ell(\vec{a}) = \text{const} + \frac{1}{2} \chi^2(\vec{a})$$

folgt, dass auch hier die Kovarianzmatrix berechnet werden kann:

$$G_{ij} = \frac{1}{2} \frac{\partial^2}{\partial a_i \partial a_j} \chi^2(\vec{a})$$

MINUIT berechnet in der parabolischen Näherung numerisch die komplette Kovarianzmatrix sowohl für Maximum-Likelihood als auch χ^2 -Fits

Skalierung erfolgt durch geeignetes Setzen von SET ERRdef

Maximum-Likelihood: 0.5 χ^2 -Methode: 1.0



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 11

Korrelierte Messungen

Bislang haben wir nur unkorrelierte Messungen behandelt

Mit anderen Worten: die x_i sind unabhängige Zufallsvariablen und die Kovarianzmatrix diagonal

Ein Vorteil der χ^2 Methode ist, dass auch ganz oder teilweise korrelierte Messungen einfach analysiert werden können

Denn es gilt weiterhin

$$\chi^2(\vec{a}) = \left(\vec{x} - \vec{X}(\vec{a}) \right)^T V^{-1}(\vec{x}) \left(\vec{x} - \vec{X}(\vec{a}) \right)$$

wobei die Kovarianzen in den Nicht-Diagonalelementen der Matrix $V(\vec{x})$ enthalten sind

Beweis: Wir haben in Vorlesung 5 gesehen, dass man immer eine orthogonale Transformation $\vec{y} = U\vec{x}$ finden kann, so dass die Zufallsvariablen y_i unkorreliert sind. Für $\vec{y}$ und $\vec{Y} = U\vec{X}$ gilt dann aber

$$\begin{aligned} \chi^2(\vec{a}) &= \left(\vec{y} - \vec{Y}(\vec{a}) \right)^T V^{-1}(\vec{y}) \left(\vec{y} - \vec{Y}(\vec{a}) \right) \\ &= \left(\vec{x} - \vec{X}(\vec{a}) \right)^T U^T V^{-1}(\vec{y}) U \left(\vec{x} - \vec{X}(\vec{a}) \right) \end{aligned}$$

$$U^T V^{-1}(\vec{y}) U = U^{-1} V^{-1}(\vec{y}) \left(U^T \right)^{-1} = \left(U^T V(\vec{y}) U \right)^{-1} = V'^{-1}(\vec{x})$$



Universität Hamburg

Datenanalyse in der Physik

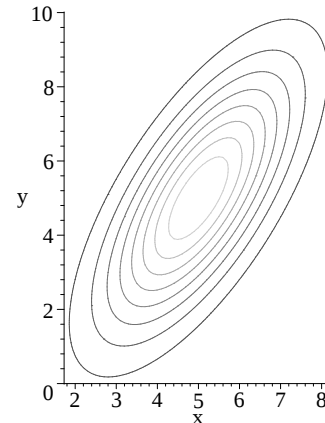
Vorlesung 8 – p. 12

Korrelierte Messungen

Was bedeuten *korrelierte Messungen* bzw. *Fehler*?

- Wir betrachten nochmals das Beispiel aus Vorlesung 4 für zwei Zufallsvariablen x, y mit $\text{cov}(x, y) > 0$

⇒ wenn ein Wert $x > \langle x \rangle$ gefunden wird, ist die Wahrscheinlichkeit auch einen Wert $y > \langle y \rangle$ zu finden erhöht (und umgekehrt)



- Übertragen auf unsere Messwerte $\vec{x}$ heißt das:

Wenn die Messung x_i größer ist als der wahre Wert $\hat{x}$, so hat auch die Messung x_j eine höhere Wahrscheinlichkeit größer als $\hat{x}$ zu sein und umgekehrt

⇒ $\text{cov}(x_i, x_j) > 0$

- Entsprechendes gilt natürlich für antikorrelierte Messungen $\text{cov}(x_i, x_j) < 0$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 13

Beispiel: Korrelationen bei systematischen Fehlern

Ein wichtiges Anwendungsbeispiel für Fehlerkorrelationen sind systematische Fehler

- Statistische Fehler sind ja i.a. unkorreliert
Man betrachte nur unser Beispiel der wiederholten Messungen von Zerfallsraten
- Bei systematischen Fehler ist das anders
Sie können nicht, teilweise oder gänzlich korreliert sein (oder anti-korreliert)
- Es ist die Aufgabe des Physikers, verschiedene Fehlerquellen einer Messung auf ihre Korrelation hin zu untersuchen

Wir wollen hier ein einfaches, besonders einleuchtendes Beispiel betrachten

- Gegeben seien n Messungen s_i einer Strecke oder Länge
- Alle Messungen seien mit dem gleichen Maßstab durchgeführt worden, der einen Gauß'schen Fehler Δ haben soll
- Es ist evident, dass alle Messungen s_i zu 100% korreliert sind
Denn wenn der Maßstabs z.B. so beschaffen wäre, dass er einen Messwert liefert, der größer ist als die wahre Länge, so werden alle Längen s_i um den gleichen Betrag zu groß gemessen



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 14

Beispiel: Korrelationen bei systematischen Fehlern

Wie würde nun die Kovarianzmatrix $V(\vec{s})$ aussehen, die für einen χ^2 -Fit aufgestellt werden muss?

- Aus der Definition des Korrelationskoeffizienten folgt:

$$\rho_{ij} = \frac{\text{cov}(s_i, s_j)}{\sigma_i \sigma_j} \implies \text{cov}(s_i, s_j) = \rho_{ij} \sigma_i \sigma_j = \Delta^2$$

wegen $\rho_{ij} = 1$ und $\sigma_i = \sigma_j = \Delta$

$$\implies V(\vec{s}) = \begin{pmatrix} \Delta^2 & \Delta^2 & \dots & \Delta^2 \\ \Delta^2 & \Delta^2 & \dots & \Delta^2 \\ \vdots & \vdots & \ddots & \vdots \\ \Delta^2 & \Delta^2 & \dots & \Delta^2 \end{pmatrix}$$

Achtung: dieses Problem ist nicht für einen χ^2 -Fit geeignet, denn wegen $\det(V) = 0$ existiert V^{-1} nicht!

- I.a. gibt es noch andere Fehlerquellen bei den Messungen, z.B. Ablesefehler σ_i , die unkorreliert sind und quadratisch zu dem Skalenfehler addiert werden können

$$\implies V(\vec{s}) = \begin{pmatrix} \sigma_1^2 + \Delta^2 & \Delta^2 & \dots & \Delta^2 \\ \Delta^2 & \sigma_2^2 + \Delta^2 & \dots & \Delta^2 \\ \vdots & \vdots & \ddots & \vdots \\ \Delta^2 & \Delta^2 & \dots & \sigma_n^2 + \Delta^2 \end{pmatrix}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 15

Nachtrag Korrelationen in χ^2 -Fits

- Wir haben in unserem Beispiel gesehen, dass die Kovarianzmatrix für n Messungen s_i mit den unabhängigen (statistischen) Fehlern σ_i und dem 100% korrelierten Fehlern Δ so aussieht:

$$V(\vec{s}) = \begin{pmatrix} \sigma_1^2 + \Delta^2 & \Delta^2 & \dots & \Delta^2 \\ \Delta^2 & \sigma_2^2 + \Delta^2 & \dots & \Delta^2 \\ \vdots & \vdots & \ddots & \vdots \\ \Delta^2 & \Delta^2 & \dots & \sigma_n^2 + \Delta^2 \end{pmatrix}$$

$\implies$ der Schätzwert $s_m \pm \sigma_s$

- In diesem speziellen Fall könnte man auch den χ^2 Fit mit der Kovarianzmatrix der unkorrelierten Fehler ausführen

$$V'(\vec{s}) = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}$$

und anschließend den 100% korrelierten systematischen Fehler quadratisch dazu addieren

$\implies$ der Schätzwert $s_m \pm \sigma'_s \pm \Delta$ bzw. $s_m \pm \sqrt{\sigma'^2_s + \Delta^2}$

- Aufstellen der kompletten Kovarianzmatrix ist allgemeiner



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 16

Gegenüberstellung Maximum-Likelihood- und χ^2 -Methode

Wichtige Eigenschaften & Unterschiede der Methoden zur Parameterschätzung

Maximum-Likelihood:

- Erfordert Kenntnis der Wahrscheinlichkeitsdichteverteilung
- Kein Binnens der Daten erforderlich
⇒ Kein Informationsverlust, besonders wichtig bei kleinen Datensätzen
- Liefert Fehler inklusive vollständiger Kovarianzmatrix der Parameter

χ^2 -Methode:

- Ist für gaußverteilte Zufallszahlen äquivalent zur Maximum-Likelihood-Methode
Zur Parameterschätzung ist χ^2 -Methode auch für andere pdf anwendbar
allerdings gelten dann die Formeln für Fehler und Kovarianzmatrix nicht mehr
- Erfordert Binnens falls eine allgemeine Verteilung an die Daten angepasst werden soll
- Einfache Möglichkeit um Korrelationen der Messungen zu berücksichtigen
insbesondere systematische Fehler
- Erlaubt einfachen Test der Güte der Anpassungsrechnung (Fit) → χ^2 -Test



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 17

Prüfung von Hypothesen

Eine sehr wichtige Aufgabenstellung eines Physikers ist die Interpretation von Messdaten im Rahmen eines physikalischen Modells

Dazu gehören

1. Aufstellen einer Hypothese (physikalisches Modell)
2. Bestimmung der Parameter des Modells
3. Überprüfung der Hypothese anhand der Messdaten

Problem:

Wie kann man die Übereinstimmung von Messdaten und Modell quantifizieren?

Es existieren eine Reihe von Methoden:

χ^2 -Test, Studentscher t -Test, F -Test, Kolmogorov-Smirnov-Test, ...
mit verschiedenen Eigenschaften, Vor- und Nachteilen

Wir behandeln hier nur den χ^2 -Test



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 18

χ^2 -Verteilung

Wir haben bereits in Vorlesung 3 die χ^2 -Verteilung kennengelernt:

Seien $x_1, x_2, \dots, x_n$ unabhängige, gaußverteilte Zufallsvariablen mit den Varianzen σ_i und X_i deren Erwartungswerte so folgt die Summe der Quadrate

$$u = \chi^2 = \sum_{i=1}^n \left(\frac{x_i - X_i}{\sigma_i} \right)^2$$

einer χ^2 -Verteilung mit n -Freiheitsgraden:

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

Bemerkungen:

- Dies gilt natürlich auch für korrelierten Zufallsvariablen, wenn wir das χ^2 über die Kovarianzmatrix definieren
- Wir nehmen zunächst an, dass die Erwartungen X_i von außen, d.h. vom Modell, vorgeben sind und nicht aus den Daten selbst bestimmt werden

Dann entspricht die Zahl der Freiheitsgrade genau der Zahl der Zufallsvariablen (Messungen)



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 19

χ^2 -Verteilung

Für den Mittelwert der χ^2 -Verteilung gilt:

$$\langle u \rangle = \langle \chi^2 \rangle = \int_0^\infty u f_n(u) du = n$$

bzw. der Mittelwert des χ^2 pro Freiheitsgrad n ist:

$$\langle \chi^2 \rangle / n = 1$$

Daraus folgt:

Falls die Hypothese zutrifft, d.h. das physikalische Modell richtig ist, so sollte man im Mittel $\chi^2/n = 1$ finden

z.B. wenn die Messreihe sehr oft wiederholt wird

- Wenn man also in einer Messreihe einen Wert $\chi^2/n \approx 1$ findet, ist das ein Zeichen dafür, dass die Hypothese korrekt ist
- Man kann die Güte der Hypothese, also die Übereinstimmung zwischen Messdaten und physikalischem Modell, mit der integrierten χ^2 -Funktion quantifizieren



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 20

Beispiel

Gegeben sei eine Messreihe mit 10 Wertepaaren zum Fallgesetz

t/s	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
s/m	0,1	0,5	0,6	1,0	0,9	2,0	2,8	3,4	4,3	4,5

Alle Messdaten sollen hier den gleichen, unkorrelierten Fehler $\sigma = 0,3$ m haben

Die Messdaten werden mit zwei Hypothesen verglichen:

A $s(t) = \frac{1}{2}gt^2 \implies \chi^2/n = 9,2/10$

B $s(t) = \frac{1}{2}gt \implies \chi^2/n = 69,1/10$

In beiden Fällen ist der Wert $g = 9,81 \text{ m/s}^2$ vorgegeben

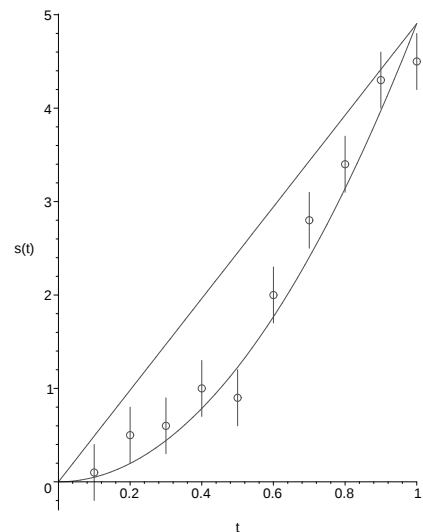
Die Wahrscheinlichkeiten diesen oder einen größeren χ^2 -Wert zu finden ($n=10$)

$$P = 1 - F_n(\chi^2) = 1 - \int_0^{\chi^2} f_n(u) du = \int_{\chi^2}^{+\infty} f_n(u) du = \int_{\chi^2}^{+\infty} \frac{\frac{1}{2} \left(\frac{u}{2}\right)^{\frac{n}{2}-1} e^{-\frac{u}{2}}}{\Gamma(\frac{n}{2})} du$$

ist für diese beiden Hypothesen

$$P_A = 0,52$$

$$P_B = 6,5 \cdot 10^{-11}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 21

Bedeutung des χ^2 -Wertes



$\chi^2/n \gg 1$

ist ein Zeichen dafür, dass

- die Hypothese falsch ist (Modell, Parameter)
- die Fehler der Messungen unterschätzt werden
- bei Kombinationen die Daten inkompatibel sind (messen nicht das gleiche)



$\chi^2/n \ll 1$

ist meist ein Zeichen dafür, dass Fehler überschätzt werden

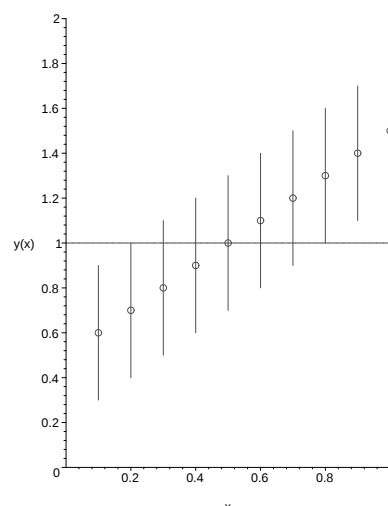


$\chi^2/n \approx 1$

unterstützt die Hypothese

ABER ein guter Wert des χ^2/n entbindet nicht von kritischer Betrachtung der Messungen!

Siehe das Beispiel rechts, dass einen Wert $\chi^2/n = 9,4/10$ liefert, obwohl die Hypothese $y(x) = 1$ wohl offensichtlich falsch ist



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 22

Zahl der Freiheitsgrade

- Bis hier war das Modell mit allen Parametern im χ^2 -Test vorgegeben
 $\Rightarrow$ Zahl der Freiheitsgrade ist gleich der Zahl der Messwerte n

Wie bestimmt man die Zahl der Freiheitsgrade, falls ein oder mehrere Parameter in dem χ^2 -Fit selbst ermittelt werden?

- Wenn in dem χ^2 -Fit

$$\chi^2(\vec{a}) = \sum_{i=1}^n \left(\frac{x_i - X_i(\vec{a})}{\sigma_i} \right)^2$$

m Parameter $a_1, a_2, \dots, a_m$ (bzw. $\vec{a}$) bestimmt werden soll, so folgen aus der Minimierungsbedingung m Gleichungen

$$\frac{\partial}{\partial a_j} \chi^2(\vec{a}) = 0 \quad j = 1, \dots, m$$

- Jede dieser Gleichungen reduziert den statistischen Variationsspielraum der Messwerte gegenüber der Vorhersage des Modells und verringert die Zahl der Freiheitsgrade um 1



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 23

Zahl der Freiheitsgrade

- Beispiele zur Verdeutlichung:

- $n = 1$ und $m = 1$:

Das Minimum des χ^2 wird offenbar für $x = X(a)$ erreicht, d.h. $\chi^2 = 0$

Das gilt allgemein für $n = m$

- $n = 2$ und $m = 1$:

Ein Freiheitsgrad wird verbraucht durch Anpassung des Parameters

Vorstellung: Parameter wird durch einen Messwert bestimmt und der zweite kann um die dann absolute Voraussage statistisch schwanken

- usw.

- Zahl der Freiheitsgrade ist (engl.: degrees of freedom *dof*) ist gleich Zahl der Messwerte n minus der daraus zu bestimmenden Parameter m

$$dof = n - m$$

- Funktionswert am Minimum des χ^2 -Fits kann zur Überprüfung der Hypothese benutzt werden χ^2 -Test

$$\chi^2_{\min}/dof$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 24

Beispiel

Wir betrachten wieder die Messreihe zum Fallgesetz

χ^2 -Fits zur Bestimmung des Parameters g liefern folgende Ergebnisse

A $s(t) = \frac{1}{2}gt^2$

$$g = 10,04 \pm 0,38 \text{ m/s}^2$$

$$\chi^2_{\min}/dof = 8,8/9$$

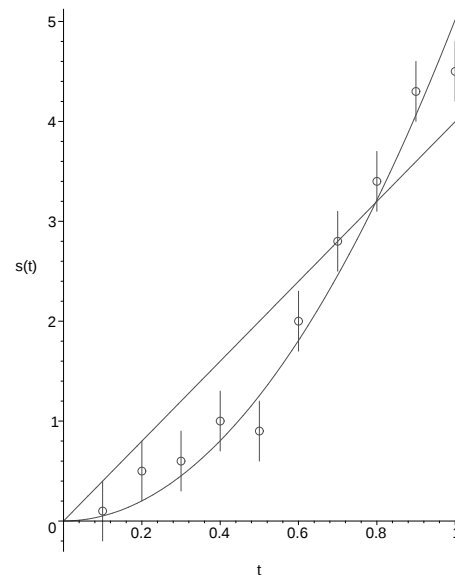
$$P = 1 - F_n(\chi^2) = 0,46$$

B $s(t) = \frac{1}{2}gt$

$$g = 7,99 \pm 0,31 \text{ m/s}$$

$$\chi^2_{\min}/dof = 33,9/9$$

$$P = 1 - F_n(\chi^2) = 9 \cdot 10^{-5}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 25

Fehler, Fehler, Fehler, ...

Ein paar zusammenfassende Bemerkungen zu Fehlern und Ihrer Bestimmung:
Wir betrachten zunächst Zählexperimente, also Experimente, deren Ausgang eine bestimmte Zahl von Ereignissen ist (radioaktive Zerfälle, Experimente an Beschleunigern etc.)

● Einmalige Messung mit dem Ausgang x_m

Die zu Grunde liegende pdf ist eine Gauß-Verteilung (oder Poisson).

In diesem Fall gilt $\sigma^2 = \mu$, wobei $\mu = \hat{x}$ der wahre Mittelwert der Gauß-Verteilung ist.

Falls wir μ kennen, z.B. aus anderen, genaueren Experimenten, einer Theorie etc., folgt das Ergebnis:

$$x_m \pm \sqrt{\mu}$$

$\sigma_m = \sqrt{\mu}$ ist der Fehler der Einzelmessung. In dem so definierten Intervall ist der wahre Wert $\hat{x}$ mit 68% Wahrscheinlichkeit (Konfidenz) enthalten. Falls man den wahren Wert nicht kennt, so nehme man x_m als beste Schätzung, also

$$x_m \pm \sqrt{x_m}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 26

Fehler, Fehler, Fehler, ...



Mehrere unabhängige Messungen x_i der gleichen Größe:

Annahme ist, dass alle Messungen der gleichen pdf unterliegen (zu überprüfen!) . Die beste Schätzung des wahren Wertes ist (siehe χ^2 -Methode):

$$x_m = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{bzw.} \quad x_m = \frac{\sum_{i=1}^n \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

Falls Gauß-Verteilung zu Grunde liegt hat jede Einzelmessung den Fehler $\sigma_i = \sqrt{\hat{x}}$ bzw. $\sigma_i = \sqrt{x_m}$, denn x_m ist eine bessere Schätzung als die x_i . Für den Fehler der Messung (Fehler des Mittelwertes) σ_m gilt dann

$$\frac{1}{\sigma_m^2} = \sum_{i=1}^n \frac{1}{\sigma_i^2} = \frac{1}{x_m} \sum_{i=1}^n 1 = \frac{n}{x_m} \quad \Rightarrow \quad \sigma_m = \frac{\sigma_i}{\sqrt{n}} = \sqrt{\frac{x_m}{n}}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 27

Fehler, Fehler, Fehler, ...

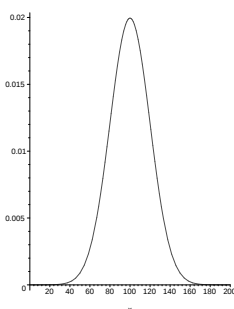
Allgemein:



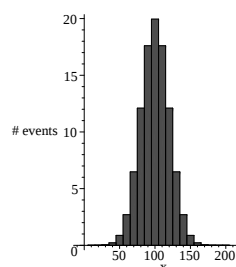
Messe wiederholt (n -mal) eine Größe x , die gleicher pdf mit Varianz σ^2 unterliegt

Mittelwert $x_m = \langle x \rangle = \frac{1}{n} \sum_{i=1}^n x_i$

Fehler auf Mittelwert $\sigma_m = \frac{\sigma}{\sqrt{n}}$



pdf mit Mittelwert μ und Varianz σ^2



n Messungen mit Mittelwert x_m und Varianz σ^2 und Fehler auf dem Mittelwert $\sigma_m = \frac{\sigma}{\sqrt{n}}$



Beweis:

1.) Fehlerfortpflanzung (siehe Beispiel zur χ^2 -Methode)

2.) Zentraler Grenzwertsatz ist andere Möglichkeit:

Betrachte Summe s der Messwerte $s = \sum_{i=1}^n x_i = nx_m$

Die Varianz dieser Summe ist $\sigma_s^2 = \sum_{i=1}^n \sigma^2 = n\sigma^2$ oder $\sigma_s = \sqrt{n}\sigma$

Daraus folgt sofort $\sigma_m = \frac{\sqrt{n}\sigma}{n} = \frac{\sigma}{\sqrt{n}}$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 28

Fehler, Fehler, Fehler, ...



Anderes Beispiel: Impulsmessung durch Spurkammer im B -Feld

(eigentlich $1/p_T$, weil das die gaußverteilte Größe ist)

Wir messen den Impuls eines Teilchens zu x_m . Den Fehler dieser Messung σ_m können wir nur angeben, wenn wir die Breite der Gauß-Verteilung kennen (z.B. durch vorherige Testmessungen, Ausmessen der pdf durch Messung vieler Teilchen etc.)

Wir machen nun Messung von n Teilchen mit gleichem Impuls und erhalten eine Verteilung, die sich der zu Grunde liegenden wahren Gauß-Verteilung annähert. Die Messung von x verbessert sich durch Mittelung zu

$$x_m = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{und Fehler} \quad \frac{\sigma_m}{\sqrt{n}}$$

Dabei ist nur der statistische Fehler berücksichtigt.

Es mag noch systematische Fehler geben (Unsicherheit im B -Feld, Position der Drähte etc.), die zwischen allen Messungen korreliert sind und durch Mittelung nicht reduziert werden.



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 29

Fehler, Fehler, Fehler, ...



Anpassungsrechnungen: Maximum-Likelihood oder χ^2 -Fit

Bietet sich an, wenn

- Mittelung nicht auf einfache Art und Weise durchgeführt werden kann
- Korrelationen zwischen Fehlern wichtig sind (χ^2)
- Daten an andere Verteilungen angepasst werden müssen (siehe Beispiel der Myonpaarasymmetrie)



Neben den besten Schätzwerten können auch Fehler ermittelt werden

- 1- σ -Fehler, 68% Wahrscheinlichkeitsinhalt

$$l_{min} + 1/2 \quad \text{Maximum-Likelihood}$$

$$\chi_{min}^2 + 1 \quad \chi^2\text{-Methode}$$



Universität Hamburg

Datenanalyse in der Physik

Vorlesung 8 – p. 30