

UNIVERSITÄT KARLSRUHE (TH)
Fakultät für Physik
Physikalisches Praktikum
Leitung: Dr. Peter Blüm

76128 Karlsruhe
Engesser-Str. 7
☎ 0721-608-3577
email: peter.blum@cern.ch

Montag, 19. August 2002

Einführung zur Fehlerrechnung im Praktikum

Diese Abhandlung ist eine Einführung in die Fehleranalyse zur Verwendung im physikalischen Anfängerpraktikum. Sie soll die grundlegenden Werkzeuge für die Fehlerrechnung, wie sie in den Versuchsprotokollen durchgeführt werden soll, bereitstellen. Eine Begründung der verwendeten Formeln wird für normalverteilte Messwerte durchgeführt. Die für die Fehlerberechnung im Praktikum relevanten Formeln sind nummeriert und können im Praktikumsprotokoll unter Verweis auf diese Nummern zitiert werden.

Inhaltsverzeichnis

1	DARSTELLUNG VON MESSERGEBNISSEN	3
1.1	SIGNIFIKANTE STELLENDES MESSERGEBNISSES	3
2	ZUFÄLLIGE MESSABWEICHUNGEN	4
2.1	FEHLERQUELLEN FÜR ZUFÄLLIGE MESSABWEICHUNGEN	4
2.2	SCHÄTZUNG VON UN SICHERHEITEN	4
3	SYSTEMATISCHE MESSABWEICHUNGEN	6
3.1	FEHLERQUELLEN FÜR SYSTEMATISCHE MESSABWEICHUNGEN	6
3.2	ERMITTLUNG SYSTEMATISCHER FEHLER	6
4	FEHLERFORTPFLANZUNG	7
5	STATISTISCHE ANALYSE ZUFÄLLIGER UN SICHERHEITEN	9
5.1	HÄUFIGKEITSVERTEILUNG	9
5.2	WAHRSCHEINLICHKEITSVERTEILUNGEN	10
5.2.1	Normalverteilung	10
5.2.2	Binomial- oder Bernoulli-Verteilung	12
5.2.3	Poissonverteilung	14
5.3	RECHFERTIGUNG DER WAHL DER BESTWERTE	15
5.4	STANDARDABWEICHUNG DES MITTELWERTS	16
5.5	VERWERFEN VON DATEN	16
5.6	DISKREPANZ VON MESSERGEBNISSEN	17
6	AUSGLEICHSRECHNUNG	17
6.1	DER GEWICHTETE MITTELWERT	17
6.2	LINEARE REGRESSION	18
6.3	GEWICHTETE LINEARE REGRESSION	20
6.4	DER LINEARE KORRELATIONS Koeffizient	20
6.5	LINEARE REGRESSION G RAPHISCHE LÖSUNG	20
7	DER χ^2 TEST FÜR EINE VERTEILUNG	21
8	EMPFEHLENSWERTE BÜCHER ZUR FEHLERRECHNUNG	23

EINLEITUNG

Im Rahmen des Praktikums soll der Student das *Experimentieren* erlernen. Dazu zählt zum einen das Kennenlernen der Messtechnik und der Technik des Messens und zum andern aber auch das Erlernen der Bewertung eines Messergebnisses. Denn um aus den Ergebnissen eines Experimentes schliessen zu können, ob ein theoretisches Modell gültig ist oder nicht, muss die Qualität und Aussagekraft der Messung bekannt sein. Alle Messungen, wie sorgfältig und wissenschaftlich sie geplant und durchgeführt werden, unterliegen *Messunsicherheiten*. Sie zu untersuchen, ihre Grösse und Ursachen zu bestimmen, sind Gegenstand der **Fehleranalyse**.

Man vermeidet gern das Wort *Fehler* und benützt den Begriff *Messunsicherheit* (gemäss deutscher Industrienorm (DIN1319) werden alle Fehler bei einer Messung als *Messabweichung* bezeichnet). Dies unterstreicht, dass mit dem Wort *Fehler* im Zusammenhang mit Messergebnissen nicht ein *falsches Ergebnis* gemeint ist sondern *die Streuung* der Messwerte um den "wahren" Wert der Messgrösse, der leider unbekannt ist. Der "wahre" Wert liegt mit einer gewissen Wahrscheinlichkeit innerhalb des durch die Messunsicherheit definierten Bereichs.

Grobe Fehler bzw. Irrtümer, wie sie sich aus Missverständnissen oder Fehlüberlegungen bei der Bedienung der Messapparatur, aus falscher Protokollierung von Messdaten oder auch aus Programmfehlern in Auswertprogrammen ergeben, werden *nicht* als Messunsicherheiten betrachtet. In diesem Fall sind die Messungen oder Auswertungen falsch und müssen wiederholt werden. Das Vorhandensein *grober Fehler* erkennt man *nur* durch kritisches Überprüfen und Kontrollieren der Ergebnisse. Vermeiden kann man sie durch sorgfältiges Experimentieren.

Wie wichtig es ist, die Messabweichungen zu kennen, zeigt das folgende Beispiel.

Beispiel: Eine Messung der Lichtgeschwindigkeit ergibt $c = (3.09 \pm 0.15) \cdot 10^8$ m/s. Dieses Ergebnis ist in Übereinstimmung mit dem "wahren" Wert, für den wir den Literaturwert $c = 2.99792458 \cdot 10^8$ m/s betrachten.

Hätte man beim gleichen Messwert einen Messfehler von $0.01 \cdot 10^8$ m/s bestimmt, so wäre das Ergebnis zwar als präziser zu bewerten aber im Widerspruch zum "wahren" Wert. Man hätte dann zu folgern, dass man entweder eine sensationelle Entdeckung gemacht hat oder, was wahrscheinlich eher zutrifft, dass die Messung falsch ist oder die Fehleranalyse eine wichtige Fehlerquelle nicht berücksichtigt.

Man sieht, die Angabe eines Messwertes allein reicht nicht aus. Die Angabe der Messunsicherheit ist unbedingt notwendig, um auf die Aussagekraft der Messung zu schliessen. Provokativ kann man sagen:

Die Messung einer physikalischen Grösse ohne Angabe der Messunsicherheit ist wertlos.

Das Beispiel beleuchtet zwei grundlegende, unbedingt zu unterscheidende Begriffe beim Bewerten von Messergebnissen: **Präzision** und **Genauigkeit**. Der erste Begriff beschreibt, wie gut eine Messung durchgeführt wurde bzw. wie reproduzierbar der Messwert ist, der zweite gibt dagegen an, wie nahe der Messwert dem "wahren" Wert ist. Es ist einsichtig, dass man bei einer Messung sowohl nach Präzision als auch nach Genauigkeit streben muss. Abbildung 1 demonstriert die beiden Eigenschaften *präzise* und *genau*.

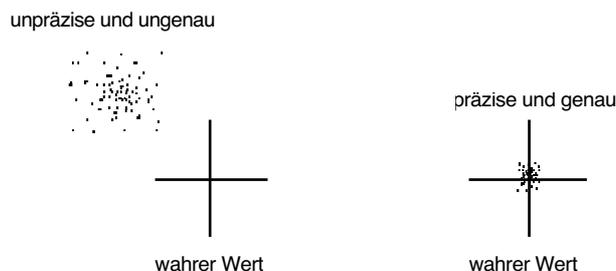


Abbildung 1: Demonstration der Begriffe *genau* und *präzise* (Die Punkte geben die einzelnen Messwerte an)

Diese beiden Eigenschaften werden durch zwei unterschiedliche Arten von Messabweichungen bestimmt. Zum einen beeinflussen **systematische Effekte**, hervorgerufen z.B. durch unberücksichtigte Umweltein-

flüsse oder fehlerhaft geeichte Messinstrumente, etc. die Genauigkeit einer Messung und zum andern begrenzen **zufällige Effekte**, wie sie sich als Summe vieler, kleiner und variierender Störungen einstellen können, die Präzision eines Experimentes. Um diese unterschiedlichen Aussagen bei einem Messergebnis erkennen zu können, werden im allgemeinen die statistischen und systematischen Fehler getrennt angegeben.

Die folgende kurze Abhandlung über Fehlerrechnung kann natürlich nicht alle Aspekte der Fehleranalyse erschöpfend behandeln, sie hat zum Ziel, den Studenten in die Lage zu versetzen, die notwendigen Aufgaben im Praktikum zu bewältigen und ihn dazu anzuregen, spezielle Fragen in der Literatur genauer zu studieren.

1 Darstellung von Messergebnissen

Messung einer physikalischen Grösse bedeutet, experimentell für diese Grösse die Masszahl zu einer gegebenen Einheit zu ermitteln. Eine vollständige Angabe von Messergebnissen beinhaltet daher Messwert, Messunsicherheit und Masseinheit. Man unterscheidet zwei Schreibweisen. Benützt man die *absolute Messunsicherheit*, so schreibt man den besten Schätzwert (*Bestwert*) der Messgrösse, der, wie später noch gezeigt wird, durch den **arithmetischen Mittelwert** der gemessenen Werte gegeben ist und die *Messabweichung*, die den Bereich kennzeichnet, in dem der "wahre" Wert mit einer Wahrscheinlichkeit von 68% liegt. Diesen Bereich nennt man auch **Vertrauensbereich**.

$$\begin{aligned} \text{Messwert} &= (\text{Bestwert} \pm \text{Unsicherheit}) [\text{Masseinheit}] \\ &= (\bar{x} \pm \delta x) [\text{Masseinheit}] \end{aligned}$$

Die Angabe der Unsicherheit δx zeigt zwar die Zuverlässigkeit des Messergebnisses an, doch lässt sich die Qualität der Messung schneller und eindeutiger an dem Quotienten aus Unsicherheit und Bestwert, *relative Messunsicherheit*, erkennen. Diese Grösse nennt man häufig auch die **Präzision der Messung**.

$$\text{Messwert} = (\text{Bestwert}) [\text{Masseinheit}] \pm \frac{\delta x}{\bar{x}} [\%]$$

Während man Messungen mit einem relativen Fehler von 10% als grob betrachtet, kennzeichnen relative Fehler von 1-2% ziemlich genaue Messungen. Die besten Ergebnisse im Praktikum liegen so zwischen 5-10%.

1.1 Signifikante Stellen eines Messergebnisses

Für die Angabe der Messunsicherheit im Praktikum wird folgende Regel angewandt.

Messunsicherheiten werden auf eine signifikante Stelle gerundet.

Das folgende Beispiel verdeutlicht diese Forderung.

Beispiel: Eine Messung der Erdbeschleunigung liefert das rechnerische Ergebnis 9.8243 m/s^2 und die Berechnung der Messunsicherheit $\pm 0.02385 \text{ m/s}^2$.

Das Ergebnis für die Messunsicherheit wird auf $\pm 0.02 \text{ m/s}^2$ gerundet. Eine Ausnahme von dieser Regel macht man im allgemeinen, wenn die erste signifikante Stelle eine 1 ist, da man dann einen relativ grossen Rundungsfehler begehen würde.

Für die Angabe des Messwertes wird eine ähnliche Regel angewandt.

Die letzte signifikante Stelle eines angegebenen Messwertes sollte dieselbe Grössenordnung besitzen wie die Messunsicherheit.

Das heisst, die Messunsicherheit bestimmt die Angabe eines Messergebnisses. Für das obige Beispiel bedeutet dies:

$$g = (9.82 \pm 0.02) \text{ m/s}^2$$

Wäre die Messunsicherheit 0.2 m/s^2 , so lautete die sinnvolle Angabe des Messergebnisses:

$$g = (9.8 \pm 0.2) \text{ m/s}^2$$

Werden allerdings Messwerte benützt, um die gesuchte Grösse zu berechnen, so müssen mindestens zwei signifikante Stellen der Messunsicherheit mitgeführt werden, um Rundungsfehler möglichst klein zu halten.

2 Zufällige Messabweichungen

Wir haben gesagt: Messung einer physikalischen Grösse bedeutet, experimentell für diese Grösse die Masszahl zu einer gegebenen Einheit zu ermitteln. Den "wahren" Wert dieser Masszahl kann man nicht ermitteln, da beim Messvorgang immer Fehlerquellen, systematischer und/oder zufälliger Art, vorhanden sind. Daher führen Messwiederholungen, selbst wenn sie unter identischen Bedingungen durchgeführt werden, nicht immer zu demselben Messwert, sondern es treten Abweichungen auf.

Sind diese Abweichungen unterschiedlich in Grösse und Richtung, so bezeichnet man sie als **zufällige Messabweichungen**. Sie sind "zufällig" in dem Sinn, dass man ihre Ursachen nicht im Einzelnen verfolgen kann. Aufgrund dieses stochastischen Verhaltens der Messergebnisse schreibt man den gemessenen Werten einen *Wahrscheinlichkeitscharakter* zu. Aus den Messwerten kann man einen besten Schätzwert für den "wahren" Wert der Messgrösse ermitteln. Der "wahre" Wert ist mit einer bestimmten Wahrscheinlichkeit in dem durch die Messabweichung definierten Bereich. Üblicherweise wählt man in der Physik den Bereich, in dem der wahre Wert mit einer Wahrscheinlichkeit von 68% liegt.

2.1 Fehlerquellen für zufällige Messabweichungen

Verschiedene Effekte können Ursachen für zufällige Fehler sein:

Statistische Messgrösse: Die Messgrösse selbst besitzt einen stochastischen Charakter, z.B. der radioaktive Zerfall von Atomkernen.

Unzulänglichkeit des Experimentators: Schätzungen und Interpolationen auf Messskalen (Parallaxenfehler) verursachen statistisch schwankende Abweichungen. Beim Messen einer Zeitdifferenz mit der Stoppuhr führt z.B. die Reaktionszeit des Experimentators zu zufälligen Abweichungen, oder bei der Längenmessung mit einem Massband wird von Messung zu Messung ein Rollmassband mit unterschiedlicher Kraft gezogen, so dass es sich unterschiedlich durchbiegt.

Äussere Einflüsse: Zufällige und unvorhersehbare äussere Einflüsse z.B. wechselnde Luftströmungen, kurzzeitige Temperaturschwankungen, etc.

2.2 Schätzung von Unsicherheiten

Die Bestimmung von Messabweichungen kann sich als ziemlich kompliziert erweisen. Bei vielen Messungen lassen sich jedoch recht einfach Abschätzungen für Unsicherheiten durchführen. Beim Ablesen einer Skala, zum Beispiel, vergleicht man die Lage eines Messpunkts oder eines Zeigers mit den Teilstrichen der Skala. Handelt es sich um eine fein unterteilte Skala, z.B. Millimeterunterteilung eines Lineals, so nimmt man den nächst gelegenen Teilstrich als Bestwert und schätzt die Unsicherheit mit \pm (halbe Intervallbreite) ab. Bei einer grob geteilten Skala wird man die Lage des Zeigers zwischen zwei Teilstrichen interpolieren, in dem man die Zentelbruchteile abschätzt und als Dezimalstelle angibt. Die Unsicherheit wird zu \pm (ein Zentelbruchteil) angenommen.

Bei Digitalanzeigen beträgt der Schätzfehler aufgrund des unbekanntem Rundungsverfahrens ± 1 in der letzten Stelle der Anzeige.

Bei Messungen, bei denen Unsicherheiten schwieriger zu schätzen sind, helfen Messwiederholungen die Messunsicherheit zu bestimmen. Das folgende Beispiel erklärt die Methode:

Beispiel: Zu messen sei die Schwingungsdauer eines Pendels mit der Stoppuhr. Infolge der Reaktionszeit beim Starten und Stoppen der Uhr kommt es zu zufälligen Messabweichungen. Folgende Messreihe wurde gemessen:

Δt in s 2.6 2.3 2.5 2.3 2.6 2.4 2.2 2.3 2.4 2.5 2.6 2.8 2.7

Es bieten sich verschiedene Möglichkeiten an, einen besten Schätzwert zu ermitteln. Man könnte, zum Beispiel, den *Meridian* (2.5) verwenden. Dabei handelt es sich um den Wert bei dem genau die Hälfte der Messwerte unterhalb bzw. oberhalb liegen. Eine bessere Wahl für die gesuchte Grösse ist aber, den **arithmetischen Mittelwert** der Messwerten anzunehmen:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

In Kapitel 5.3 wird bewiesen, dass für normalverteilte Messwerte x_i der arithmetische Mittelwert in der Tat den besten Schätzwert wiedergibt.

$$\bar{x} = 2.48 \text{ s}$$

Auch bei der Abschätzung der Messunsicherheit sind verschiedene Lösungswege vorstellbar. So könnte man etwas naiv argumentieren, dass der "wahre" Wert mit grosser Wahrscheinlichkeit zwischen dem höchsten (2.8 s) und dem niedrigsten Wert (2.2 s) liegt:

$$\Delta x = \pm 0.3 \text{ s}$$

Eine bessere Beschreibung der Messunsicherheit für die Messwerte wird auch hier durch eine Art Mittelwertbildung geliefert. Dabei muss man aber berücksichtigen, dass die Abweichungen positiv und negativ sein können, weshalb man die Quadratwurzel aus der mittleren quadratischen Abweichung verwendet (im Englischen nennt man dies *root mean square deviation*). Man spricht auch vom mittleren quadratischen Messfehler bzw. **Standardabweichung** s_x .

$$\Delta x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\Delta x = 0.172 \text{ s} \quad \Rightarrow \quad \Delta x = 0.17 \text{ s}$$

Bei einer korrekten statistischen Behandlung muss man berücksichtigen, dass dieselben Daten auch den Mittelwert bestimmen, d.h. die Anzahl der Freiheitsgrade um 1 reduziert ist.

$$s_x = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

$$s_x = 0.179 \text{ s} \quad \Rightarrow \quad s_x = 0.18 \text{ s}$$

Innerhalb des Bereiches $\bar{x} - s_x$ und $\bar{x} + s_x$ liegt mit 68% Sicherheit der nächste Messwert der gesuchten Grösse. Man beachte, dass das bedeutet, dass 32% der Messungen ausserhalb liegen können.

Der Mittelwert ist genauer bestimmt. Die Abweichung wird gegeben durch:

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} \quad (3)$$

Für das Beispiel :

$$s_{\bar{x}} = \frac{0.179 \text{ s}}{\sqrt{13}} \quad \Rightarrow \quad s_{\bar{x}} = 0.05 \text{ s}$$

Das endgültige Ergebnis für das Beispiel lautet also:

$$x_{\text{best}} = (2.48 \pm 0.05) \text{ s}$$

d.h. der „wahre“ Wert liegt mit 68% Sicherheit innerhalb von $\bar{x} - s_{\bar{x}}$ und $\bar{x} + s_{\bar{x}}$. Auch hier beträgt die *Irrtumswahrscheinlichkeit* 32%.

Man beachte, dass sich infolge von Messwiederholungen die Messabweichung des Einzelereignisses s_x nicht verändert (bis auf Schwankungen), die Genauigkeit von \bar{x} aber mit \sqrt{n} sich verbessert.

Wenn man eine Aussage mit einer grösseren statistischen Sicherheit machen möchte, so muss man als Messunsicherheit die doppelte oder gar dreifache Standardabweichung verwenden. Dann beträgt die statistische Sicherheit 95,5% bzw. 99,7%.

3 Systematische Messabweichungen

Systematische Messabweichungen zeigen bei identischen Messbedingungen immer *um den gleichen Betrag in die gleiche Richtung*. Sie können deshalb auch nicht durch Messwiederholungen erkannt oder beseitigt werden. Die bei Messwiederholung beobachteten Schwankungen werden durch zufällige Effekte verursacht, die sich systematischen Fehlern überlagern. Nur eine kritische Analyse des Messvorgangs kann helfen, solche systematische Einflüsse zu erkennen.

3.1 Fehlerquellen für systematische Messabweichungen

Es existieren eine Reihe von möglichen Verursachern, die man so weit als möglich ausschalten oder zumindest minimieren muss.

Umwelteinflüsse: Um solche Effekte zu eliminieren, sollte das zu untersuchende System möglichst von allen störenden äusseren Einflüssen isoliert werden. Ist es aus experimentellen Gründen nicht möglich, die Einflüsse auszuschalten, so muss man darauf rechnerisch korrigieren.

Unvollkommenheit der Messinstrumente: Diese können verschiedene Ursachen haben. Eine fehlerhafte Produktion z.B. nicht lineare Linearteilung oder Fehljustierung bzw. Fehleichung eines Messinstrumentes führt immer zu einer Messabweichung in eine Richtung.

Rückwirkung von Messinstrumenten auf die Messung: Diese Effekte lassen sich nie ganz ausschliessen. Ein bekanntes Beispiel ist die Bestimmung des ohmschen Widerstandes als Quotient einer Strom- und Spannungsmessung. Hier müssen die Innenwiderstände der Messinstrumente rechnerisch berücksichtigt werden.

Unzulänglichkeit des Experimentators: Auch der Experimentator kann Ursache systematischer Fehlerquellen sein. So können ungenügend verstandene Messvorgänge oder stark vorgeprägte Erwartungen für den Messwert eine unkritische Haltung dem Messergebnis gegenüber bewirken (z.B. bei notwendiger Interpolation von Skalen). Auch die Verwendung von Näherungsformeln (z.B. Feld einer 'langen Spule' bei endlicher Spulenlänge oder $\sin(x) \approx x$ für kleine x) führen zu systematischen Abweichungen, sofern sie nicht rechnerisch berücksichtigt werden. Eine häufige Quelle für Messfehler sind Beobachtungsfehler. Diese können sich ebenfalls zu systematischen Messfehlern entwickeln.

Beispiel: Beim Messen der Schwingungsdauer eines Pendels verwendet man das Durchschwingen der Ruhelage als Zeitmarke und markiert sie als Strich an der Wand. Es hängt nun von der Lage des Kopfs des Experimentators ab, wann der Pendel und der Strich fluchten. Das ist ein Parallaxefehler, der sich als systematischer Fehler auswirken kann, wenn der Experimentator schräg zur Schwingungsebene steht. Ähnliches kann beim Ablesen von Zeigerinstrumenten ohne Spiegelskala auftreten.

3.2 Ermittlung systematischer Fehler

Eine Methode zur Erkennung systematischer Messabweichungen besteht in einer gezielten Abänderung der Messbedingungen. Hierbei sollten stets die Parameter geändert werden, die keinen Einfluss auf die unbekanntes Messgrösse besitzen. Als Beispiel diene die Bestimmung des ohmschen Widerstands als Quotient einer Strom- und Spannungsmessung. Hier kann man systematische Messabweichungen erkennen durch:

- Verwendung anderer Messinstrumente (Innenwiderstände der Messgeräte, Überprüfung der Eichung)
- Messung bei verschiedenen Strömen und Spannungen (Kontaktwiderstand)
- Umpolung oder Verwendung von Wechselstrom (thermoelektrische Störpotentiale)

Eine noch wirkungsvollere Methode ist die Wahl einer grundsätzlich anderen Messmethode. Ein **Beispiel** aus dem Praktikum: Bestimmung der spezifischen Ladung e/m des Elektrons mit dem Fadenstrahlrohr und der Methode nach Busch. Zeigen die Messergebnisse keine *Diskrepanz*, d.h. stimmen sie innerhalb der Messgenauigkeit überein, so kann man davon ausgehen, dass mit grosser Wahrscheinlichkeit keine oder vernachlässigbare systematische Messfehler vorliegen.

Besonders geeignet ist die Wahl eines Messverfahrens, das systematische Fehler vermeidet, z.B. Messung elektrischer Spannungen durch Kompensation (Vergleich der unbekanntes Spannung mit einem geeichten Spannungsnorm).

Der Einfluss einer erkannten systematischen Fehlerquelle kann in der Regel minimalisiert werden. Oft kann man den systematischen Fehler durch nachträgliche rechnerische Korrektur eliminieren. Dazu sind entweder spezielle Hilfsexperimente (Eichmessungen) nötig oder die Korrekturfunktion ist theoretisch bekannt (z.B. Zählratenverlust durch Totzeit oder Korrektur für Wärmeübertragung bei Temperaturmessungen, etc.).

4 Fehlerfortpflanzung

Wenn eine zu messende Grösse q nicht direkt gemessen werden kann, sondern als Funktion von anderen Messgrössen $\{x_1, x_2, \dots\}$ berechnet werden muss, so pflanzen sich die Messfehler der direkt gemessenen Grössen x_i in das Ergebnis q fort. Das bedeutet, man misst anstelle der gesuchten Grösse $q(x_1, x_2, \dots)$ die Grösse $q(x_1+h_1, x_2+h_2, \dots)$. Geht man davon aus, dass die Messfehler klein sind im Vergleich zu den Messwerten, so kann die Funktion in eine Taylorreihe entwickelt und die Reihe nach dem linearen Glied h_k abgebrochen werden.

$$q(x_1 + h_1, x_2 + h_2, \dots) = q(x_1, x_2, \dots) + \frac{1}{1!} \sum_{k=1}^n \frac{\partial q}{\partial x_k} \cdot h_k$$

Sind die Fehler h_k in Grösse und Vorzeichen bekannt, so erhält man einen nach Vorzeichen und Grösse definierten Gesamtfehler Δq

$$\Delta q = q(x_1 + h_1, x_2 + h_2, \dots) - q(x_1, x_2, \dots) = \sum_{k=1}^n \frac{\partial q}{\partial x_k} \cdot h_k$$

Im allgemeinen sind aber die Fehler nicht bekannt, sondern man hat bestenfalls Schätzwerte für die gemessenen Grössen (Mittelwert, Standardabweichung). Wie kann man aus den Messabweichungen der Messwerte die Unsicherheit der Grösse q abschätzen? Wir gehen davon aus, dass der beste Wert für q gegeben ist durch:

$$\bar{q} = q(\bar{x}_1, \bar{x}_2, \dots)$$

Eine Abschätzung für Δq erhält man aus der Streuung der q_i , die man aus den einzelnen Messungen x_i berechnet. Im Grenzfall unendlich vieler Messungen gibt dann die Varianz der Verteilung der q_i einen Schätzwert für σ_q .

$$\sigma_q^2 = \lim \left[\frac{1}{n} \sum (q_i - \bar{q})^2 \right]$$

Benützt man nun

$$q_i - \bar{q} \cong (x_{1i} - \bar{x}_1) \frac{\partial q}{\partial x_{1i}} + (x_{2i} - \bar{x}_2) \frac{\partial q}{\partial x_{2i}} + \dots$$

so erhält man für die Varianz von q

$$\sigma_q^2 = \lim \frac{1}{n} \sum \left[(x_{1i} - \bar{x}_1)^2 \cdot \left(\frac{\partial q}{\partial x_{1i}} \right)^2 + (x_{2i} - \bar{x}_2)^2 \cdot \left(\frac{\partial q}{\partial x_{2i}} \right)^2 + 2(x_{1i} - \bar{x}_1) \cdot (x_{2i} - \bar{x}_2) \cdot \left(\frac{\partial q}{\partial x_{1i}} \right) \cdot \left(\frac{\partial q}{\partial x_{2i}} \right) + \dots \right]$$

Mit $n \rightarrow \infty$

$$\sigma_{x_1}^2 = \lim \frac{1}{n} \sum (x_{1i} - \bar{x}_1)^2 \quad \text{und} \quad \sigma_{x_2}^2 = \lim \frac{1}{n} \sum (x_{2i} - \bar{x}_2)^2$$

$$\text{so wie der Kovarianz} \quad \sigma_{x_1 x_2} = \lim \frac{1}{n} \sum (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)$$

folgt dann die **allgemeine Fehlerfortpflanzungsgleichung**:

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x_1} \right)^2 \sigma_{x_1}^2 + \left(\frac{\partial q}{\partial x_2} \right)^2 \sigma_{x_2}^2 + 2\sigma_{x_1 x_2} \left(\frac{\partial q}{\partial x_1} \right) \left(\frac{\partial q}{\partial x_2} \right) + \dots$$

Die ersten beiden Terme sind die mittleren quadratischen Abweichungen der x_i . Sie bilden den Hauptbeitrag zu der Varianz von q . Der dritte Term dieser Gleichung berücksichtigt eventuell vorhandene Korrelationen zwischen den x_i . Im allgemeinen sind die Messgrößen jedoch nicht korreliert und dann kann der dritte Term vernachlässigt werden und man findet das bekannte **Gauss'sche Fehlerfortpflanzungsgesetz**:

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x_1} \right)^2 \sigma_{x_1}^2 + \left(\frac{\partial q}{\partial x_2} \right)^2 \sigma_{x_2}^2 + \dots} \quad (4)$$

Anhand dieser allgemeinen Formel lassen sich viele spezielle Formeln zu Summen, Produkten, Exponentialausdrücken etc. berechnen.

Beispiel: $a = b \pm c$ $\sigma_a^2 = \sigma_b^2 + \sigma_c^2$

d.h. bei Addition oder Subtraktion addieren sich die absoluten Fehler.

$$f = x \cdot y \quad \text{oder} \quad f = x/y \quad \left(\frac{\sigma_f}{f} \right)^2 = \left(\frac{\sigma_x}{x} \right)^2 + \left(\frac{\sigma_y}{y} \right)^2$$

bei Multiplikation oder Division addieren sich die relativen Fehler.

$$f = x^r \quad \left(\frac{\sigma_f}{f} \right)^2 = r^2 \cdot \left(\frac{\sigma_x}{x} \right)^2$$

bei Potenzen vervielfachen sich die relativen Fehler um den Exponenten.

Kann man eine statistische Unabhängigkeit der Messabweichungen nicht voraussetzen, so verwendet man das arithmetische Fehlerfortpflanzungsgesetz:

$$\Delta q = \left| \frac{\partial q}{\partial x_1} \right| \Delta x_1 + \left| \frac{\partial q}{\partial x_2} \right| \Delta x_2 + \dots \quad (5)$$

Es entspricht einer Grösstfehlerabschätzung der Messung.

5 Statistische Analyse zufälliger Unsicherheiten

In Kapitel 3 wurde behauptet, dass durch Messwiederholung die Präzision eines Messergebnisses verbessert werden kann. In diesem Abschnitt wollen wir die theoretische Begründung für diese Eigenschaft behandeln.

Während man sich bei kleinen Datenmengen durch die Darstellung der Messergebnisse in Tabellenform noch leicht einen Überblick verschaffen und die Berechnung des Mittelwerts und der Standardabweichung durchführen kann, muss man bei grossen Datenmengen andere Verfahren benützen. Ein geeignetes Verfahren ist die Darstellung der Daten in einem Histogramm als Häufigkeitsverteilung. Diese wird mit der zugrundeliegenden **Grenzverteilung/Stammverteilung** verglichen. Unter Grenzverteilung versteht man die Verteilungsfunktion, die durch einen unendlich grossen Datensatz unabhängiger Messergebnisse $\{x_1, \dots, x_N\}$ beobachtet würde. In dieser Näherung gehen dann die Parameter der Verteilung des Datensamples über in die Parameter der Grenzverteilung.

5.1 Häufigkeitsverteilung

Zur Darstellung einer grossen Menge von Messergebnissen in Form einer Häufigkeitsverteilung wird die Anzahl der in einem Intervall Δx gefundenen Messwerte n_k über den Messwerten x_k aufgetragen. Diese Methode vermittelt einen direkten visuellen Überblick über die Messergebnisse. Der Schwerpunkt der Verteilung gibt den besten Schätzwert und die Breite der Verteilung ist ein Mass für die Streuung des gesuchten Wertes. In vielerlei Hinsicht ist es zweckmässig, mit relativen Häufigkeiten zu arbeiten. Dazu bestimmt man zunächst die Gesamtzahl der Messwerte

$$\sum n_k = N$$

und erhält dann die relative Häufigkeit

$$f_k(x_k) = \frac{n_k}{N}$$

Da die Summe über alle f_k gleich 1 ist, spricht man auch von einer **normierten Verteilung**.

Die Einführung des Begriffs der relativen Häufigkeit führt auch zwanglos zu der Wahrscheinlichkeitsinterpretation von Messergebnissen.

$$f_k(x) \cdot \Delta x = \text{Wahrscheinlichkeit ein Messergebnis in } \Delta x \text{ zu finden}$$

Beim Übergang zum Grenzfall unendlich vieler Beobachtungen können die Intervalle Δx infinitesimal klein werden und die diskrete Verteilung $f_k(x)$ geht in eine stetige Funktion $f(x)$ über, welche die Grenzverteilung beschreibt.

Ist $f(x)$ bekannt, so lassen sich Mittelwert und Standardabweichung relativ leicht als Momente der Verteilung berechnen. Das erste Moment, das die Lage der Verteilung charakterisiert, gibt den Mittelwert an.

$$\bar{x} = \int_{-\infty}^{+\infty} x f(x) dx \quad (6)$$

Das zweite Moment kennzeichnet die Breite der Verteilung und gibt ihre Varianz an

$$s^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2 f(x) dx \quad (7)$$

Die Wurzel aus der Varianz wird Standardabweichung s genannt. Sie ist ein direktes Mass für die Breite der Verteilung, d.h. für die Streuung der Messwerte.

Aus der Gleichung für die Varianz kann die folgende wichtige Beziehung hergeleitet werden.

$$\begin{aligned} s^2 &= \int x^2 f(x) dx - 2\bar{x} \int x f(x) dx + \bar{x}^2 \int f(x) dx \\ &= \overline{x^2} - \bar{x}^2 \end{aligned}$$

In der Statistik bezeichnet man die Gesamtheit aller unter gleichen Bedingungen möglichen Messungen als Grundgesamtheit. Sie wird durch die Grenzverteilung beschrieben. Da die erforderliche Datenmenge unendlich gross ist, kann man nie eine Realisierung der Grenzverteilung erreichen, sondern man muss sie durch eine Stichprobe vom Umfang n annähern.

5.2 Wahrscheinlichkeitsverteilungen

Unterschiedliche Messgrössen haben auch unterschiedliche Grenzverteilungen. So werden z.B. Messungen, die vielen kleinen und zufälligen Abweichungen unterliegen, durch eine **Gauss- oder Normalverteilung** beschrieben. Die Normalverteilung ist ein Grenzfall der **Binomialverteilung**. Andere Beispiele für physikalische Verteilungen sind in der Theorie der Wärme die **Maxwell'sche Geschwindigkeitsverteilung** oder beim Zählen von Ereignissen aus dem radioaktiven Zerfall die **Poissonverteilung**.

5.2.1 Normalverteilung

Die Gaussverteilung hat eine grosse Bedeutung, da viele physikalische Messungen exakt oder angenähert eine Wahrscheinlichkeitsdichte dieser Form haben. Das Maximum der Gausskurve gibt den besten Schätzwert x_w des "wahren" Wertes der gesuchten Grösse x an. Die Kurve ist symmetrisch um das Maximum x_w . Mathematisch ist die Gausskurve beschrieben durch

$$e^{-(x-x_w)^2/2\sigma^2}; \text{ mit } \sigma = \text{Breiteparameter}$$

Der Begriff "Breiteparameter" wird aus der Abbildung 2, die zwei Gausskurven mit unterschiedlichen Parametern zeigt, unmittelbar klar. Je kleiner der Breitenparameter σ ist, desto schmaler wird die Kurve und umso präziser ist die Messung.

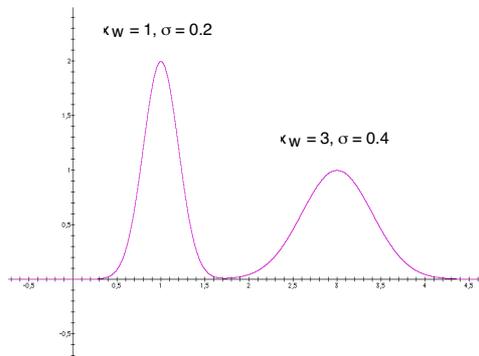


Abbildung 2: Zwei Normalverteilungen mit unterschiedlichen Breiten und Lagen

Um als Stammverteilung verwendbar zu sein, sollte die Verteilung normiert werden.

$$\text{Normierung: } \int_{-\infty}^{+\infty} N e^{-(x-x_w)^2/2\sigma^2} dx = 1$$

$$\text{Substitution: } y = \frac{x-x_w}{\sigma} \text{ und } dy = \frac{1}{\sigma} dx$$

$$N\sigma \underbrace{\int_{-\infty}^{+\infty} e^{-y^2/2} dy}_{\sqrt{2\pi}} \equiv 1 \Rightarrow N = \frac{1}{\sigma\sqrt{2\pi}}$$

Damit erhalten wir die normierte Gaussfunktion:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-x_w)^2/2\sigma^2} \quad (8)$$

Da σ im Nenner des Normierungsfaktors der Gausskurve steht, wird das Maximum bei kleiner werdendem σ grösser.

Mit der normalisierten Gaussfunktion kann nun der Mittelwert als erstes Moment der Verteilung berechnet werden:

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-(x-x_w)^2/2\sigma^2} dx$$

mit der Substitution $y=x-x_w$ und $dy = dx$ folgt

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \left(\underbrace{\int_{-\infty}^{+\infty} y e^{-y^2/2\sigma^2} dy}_{=0} + x_w \underbrace{\int_{-\infty}^{+\infty} e^{-y^2/2\sigma^2} dy}_{=\sigma\sqrt{2\pi}} \right)$$

$$\Rightarrow \bar{x} = x_w \quad (\text{wie erwartet})$$

Exakt richtig ist dieses Ergebnis allerdings nur für unendlich viele Messwerte, doch auch bei einer hinreichend grossen Anzahl von Messungen (30 und mehr) liegt der Mittelwert schon recht nahe bei x_w .

Entsprechend kann man die Standardabweichung beweisen.

$$s_x^2 = \frac{1}{\sigma\sqrt{2\pi}} \int (x - x_w)^2 e^{-(x-x_w)^2/2\sigma^2} dx$$

mit der Substitution $y = \frac{x - x_w}{\sigma}$ und $dy = \frac{1}{\sigma} dx$

$$s_x^2 = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-y^2/2} dy$$

Das Integral löst man mit partieller Integration: $\int u dv = uv - \int v du$

$$s_x^2 = \frac{\sigma^2}{\sqrt{2\pi}} \left(\underbrace{-y e^{-y^2/2}}_{=0} \Big|_{-\infty}^{+\infty} + \underbrace{\int_{-\infty}^{+\infty} e^{-y^2/2} dy}_{=\sqrt{2\pi}} \right)$$

$$s_x^2 = \sigma^2$$

Der Breite Parameter σ der Gaussverteilung ist gleich der Standardabweichung, die man nach unendlich vielen Messungen erzielen würde. Auch hier fängt die "Unendlichkeit" bei etwa 30 Messwerten an.

Auch die Vertrauensgrenze, die durch die Standardabweichung gegeben ist, kann man nun berechnen.

$$P(\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_w-\sigma}^{x_w+\sigma} e^{-(x-x_w)^2/2\sigma^2} dx \quad (9)$$

Substitution: $y = \frac{x - x_w}{\sigma}$; $dy = \frac{1}{\sigma} dx$; $y_u = -1$; $y_o = 1$

$$P(\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-y^2/2} dy = 68\%$$

Das Integral (9) ist ein Standardintegral der mathematischen Physik und wird häufig als Fehlerfunktion $\text{erf}(x)$ bezeichnet. Es lässt sich allerdings nicht analytisch lösen. Abbildung 3 zeigt die numerisch berechnete Wahrscheinlichkeit als Funktion des Abstandes vom "wahren" Wert.

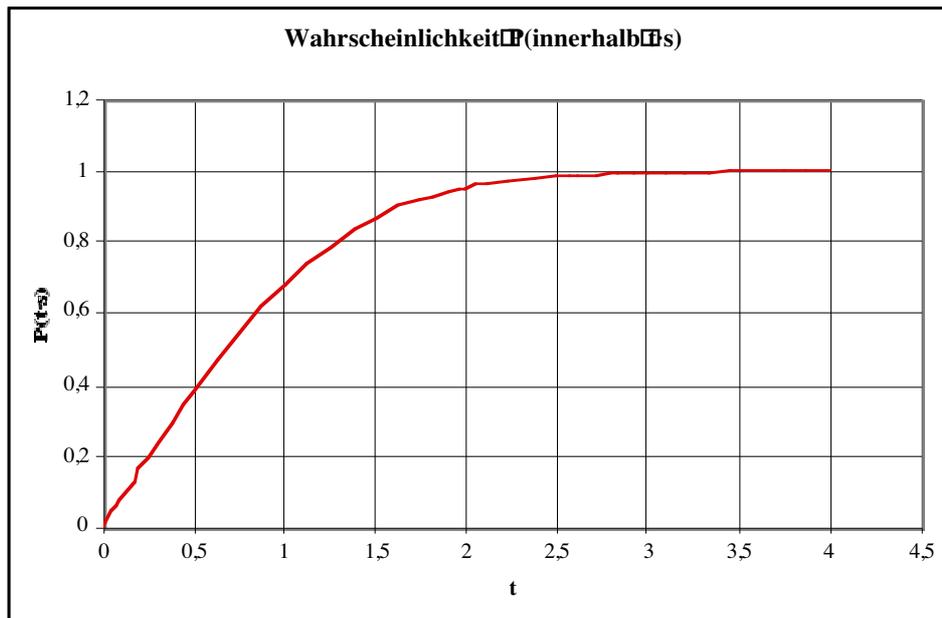


Abbildung 3: Die Wahrscheinlichkeit P, dass ein Messwert innerhalb $t \sigma$ liegt.

Die Wahrscheinlichkeit, dass eine Messung ein Ergebnis innerhalb von σ liefert, beträgt 68%. Aus der Abbildung 3 ist auch zu erkennen, dass die statistische Sicherheit mit der Wahl der Messunsicherheit von $\pm 2\sigma$ bei 95,5% und von $\pm 3\sigma$ bei 99,7% liegt.

Bei einem endlichen Umfang der Stichprobe $\{x_i, i=1, \dots, n < 30\}$ gilt für den Vertrauensbereich S der Messreihe:

$$x = \bar{x} \pm t_{n,S} \cdot \frac{s_x}{\sqrt{n}} \quad (10)$$

Der Korrekturfaktor t folgt aus der Studentverteilung. Und kann aus der nachfolgenden Tabelle entnommen werden.

n\S	68,3%	95,5%	99,7%
2	1,84	12,71	235,80
3	1,32	4,30	19,21
4	1,20	3,18	9,22
5	1,15	2,78	6,62
6	1,11	2,57	5,51
8	1,08	2,37	4,53
10	1,06	2,26	4,09
20	1,03	2,09	3,45
30	1,02	2,05	3,28

5.2.2 Binomial- oder Bernoulli-Verteilung

Hierbei handelt es sich um die wichtigste diskrete Verteilung. Sie beschreibt die Wahrscheinlichkeit, bei n Versuchen x erfolgreiche Ergebnisse zu finden, wobei jeder Versuch eine Erfolgchance p besitzt. Sie ist gegeben durch:

$$B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \quad (11)$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-x+1)}{1 \cdot 2 \cdot \dots \cdot x}$$

Beispiel: Wir werfen 4 Münzen und fragen, wie oft finden wir x = 0-, 1-, 2-, 3-, 4-mal Kopf oben. p ist in diesem Fall 1/2. Wir wiederholen den Versuch 4 mal, d.h. n = 4.

$$B(x; 4, \frac{1}{2}) = \binom{4}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$$

Erfolgswahrscheinlichkeit [%]	6,25	25	37,5	25	6,25
für x =	0	1	2	3	4

Wie erwartet findet man, dass die wahrscheinlichste Anzahl von ‚Kopf oben‘ x = 2 ist. Die Symmetrie der Verteilung liegt in p = 1/2 begründet.

Einige Eigenschaften dieser Verteilung: B(x;n,p) ist normiert. Der Mittelwert berechnet sich folgendermassen:

$$\bar{x} = \sum_{x=0}^n x \cdot B(x;n,p) = p \cdot n$$

Beweis:

$$(p+q)^n = \sum \binom{n}{x} p^x q^{n-x} \quad / \quad \frac{\partial}{\partial p}$$

$$n(p+q)^{n-1} = \sum x \binom{n}{x} p^{x-1} q^{n-x} \quad / \quad \cdot p \text{ und } p+q=1$$

$$n \cdot p = \underbrace{\sum x \binom{n}{x} p^x q^{n-x}}_{=\bar{x}}$$

Ähnlich kann man den Ausdruck für die Standardabweichung herleiten:

$$\sigma^2 = \sum (x - \bar{x})^2 B(x;n,p) = n \cdot p \cdot (1-p)$$

Beweis:

$$(p+q)^n = \sum \binom{n}{x} p^x q^{n-x} \quad / \quad \frac{\partial}{\partial p} \text{ zweimal anwenden}$$

$$n \cdot (n-1) \cdot (p+q)^{n-2} = \sum x \cdot (x-1) \binom{n}{x} p^{x-2} q^{n-x} \quad / \quad \cdot p^2 \text{ und } p+q=1$$

$$n(n-1)p^2 = \sum (x^2 - x) B(x;n,p)$$

$$= \overline{x^2} - \bar{x}$$

mit $\bar{x} = n \cdot p \Rightarrow \overline{x^2} = n \cdot (n-1) \cdot p^2 + n \cdot p$

Einsetzen in: $\sigma_x^2 = \overline{x^2} - \bar{x}^2$

$$= n(n-1)p^2 + np - n^2 p^2$$

$$= np(1-p)$$

$$\sigma_x = \sqrt{np(1-p)}$$

Zum Schluss sei noch bemerkt, dass die Binomialverteilung für $n \rightarrow \infty$ und nicht zu kleine p in die Gaussverteilung übergeht ($n > 30$ und $p > 0.05$).

5.2.3 Poissonverteilung

Wird in der Binomialverteilung die Wahrscheinlichkeit p sehr klein (<0.05) und die Anzahl der Versuche ist gross, so geht die Verteilung in eine Poissonverteilung über.

$$P(x) = e^{-\bar{x}} \frac{\bar{x}^x}{x!} \quad (12)$$

Beweis:

$$\begin{aligned} P(x) &= \binom{n}{x} p^x q^{n-x} \quad \text{für } n \rightarrow \infty \text{ und } p \rightarrow 0 \text{ aber } \bar{x} = n \cdot p = \text{const} \\ &= \frac{n!}{x!(n-x)!} \cdot \left(\frac{\bar{x}}{n}\right)^x \cdot \frac{\left(1 - \frac{\bar{x}}{n}\right)^n}{\left(1 - \frac{\bar{x}}{n}\right)^x} \\ &= \frac{\bar{x}^x}{x!} \cdot \frac{n \cdot (n-1) \cdot \dots \cdot (n-x+1)}{n^x} \cdot \frac{\left(1 - \frac{\bar{x}}{n}\right)^n}{\left(1 - \frac{\bar{x}}{n}\right)^x} \\ &= \frac{\bar{x}^x}{x!} \cdot \underbrace{\left(1 - \frac{\bar{x}}{n}\right)^n}_{\substack{= e^{-\bar{x}} \\ \text{für } n \rightarrow \infty}} \cdot \frac{n^x}{n^x} \cdot \underbrace{\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{x-1}{n}\right)}_{\substack{= 1 \\ \text{für } n \rightarrow \infty}} \left(1 - \frac{\bar{x}}{n}\right)^{-x} \end{aligned}$$

Die Poissonverteilung ist nur durch einen Parameter, den Mittelwert \bar{x} , bestimmt. Sie eignet sich besonders zur Beschreibung der Zerfälle instabiler Kerne. Auch die Poissonverteilung ist normiert.

Das erste Moment liefert den Mittelwert:

$$\begin{aligned} \bar{x} &= \sum_{x=0}^{\infty} xP(x) = \sum_{x=0}^{\infty} x e^{-\bar{x}} \frac{\bar{x}^x}{x!} \\ \bar{x} &= \bar{x} e^{-\bar{x}} \sum_{x=0}^{\infty} \frac{\bar{x}^{x-1}}{(x-1)!} \\ &= \bar{x} e^{-\bar{x}} \underbrace{\sum_{x=0}^{\infty} \frac{\bar{x}^{x-1}}{(x-1)!}}_{= e^{\bar{x}}} \\ \Rightarrow \bar{x} &= \bar{x} \end{aligned}$$

Das zweite Moment liefert die Varianz der Verteilung:

$$\sigma_x^2 = \sum_{x=0}^{\infty} (x - \bar{x})^2 P(x) = \bar{x}$$

Für die Standardabweichung gehen wir wieder von der Beziehung $\sigma_x^2 = \overline{x^2} - \bar{x}^2$ aus. Ähnlich wie im Fall des Mittelwertes kann man $\overline{x^2}$ berechnen und findet:

$$\begin{aligned} \overline{x^2} &= \bar{x}^2 + \bar{x} \quad \Rightarrow \quad \sigma_x^2 = \bar{x} \\ \text{bzw. } \sigma_x &= \sqrt{\bar{x}} \end{aligned} \quad (13)$$

d.h. bei der Poissonverteilung ist auch die Standardabweichung durch den Mittelwert bestimmt.

Für grosse Werte von \bar{x} nähert sich die Poissonverteilung der Normalverteilung an.

5.3 Rechtfertigung der Wahl der Bestwerte

Im allgemeinen beinhaltet eine Messung nur endlich viele Messwerte und man muss aus den N Messwerten, die besten Schätzwerte für x_w und σ finden. Kennt man die Grenzverteilung, so lassen sich diese Werte relativ einfach bestimmen. Wir nehmen an, dass wir es mit Messungen zu tun haben, die vielen kleinen, zufälligen Abweichungen unterliegen. Dann ist die Grenzverteilung durch die Gauss-Verteilung gegeben. Die Wahrscheinlichkeit, einen Messwert x_i zu finden, ist gegeben durch

$$P(x_i) \propto \frac{1}{\sigma} e^{-(x_i - x_w)^2 / 2\sigma^2}$$

Die Wahrscheinlichkeit, dass man genau das Datensample $\{x_1, \dots, x_N\}$ findet, ist für unabhängige Messungen gegeben durch das Produkt der Einzelwahrscheinlichkeiten.

$$\begin{aligned} P(x_1, \dots, x_N) &\propto \prod_{i=1}^N \frac{1}{\sigma} e^{-(x_i - x_w)^2 / 2\sigma^2} \\ &\propto \frac{1}{\sigma^N} e^{-\sum (x_i - x_w)^2 / 2\sigma^2} \end{aligned}$$

x_w und σ , der "wahre" Wert von x und der Breiteparameter der Verteilung, sind leider unbekannt. Die besten Schätzwerte für x_w und σ erhält man nach dem Prinzip der grössten Wahrscheinlichkeit (**Maximum Likelihood Prinzip**). Für den gegebenen Fall bedeutet das: "Für die N beobachteten Werte sind die besten Schätzungen für x_w und σ diejenigen Werte, für die $\{x_1, \dots, x_N\}$ die wahrscheinlichsten Messwerte sind". Das bedeutet $P(x_1, \dots, x_N)$ muss maximal bzw. $\sum_{i=1}^N (x_i - x_w)^2 / 2\sigma^2$ muss minimal werden.

$$\begin{aligned} \frac{\partial}{\partial x_w} \sum (x_i - x_w)^2 / 2\sigma^2 &= 0 \\ \sum (x_i - x_w) &= 0 \\ \Rightarrow x_w &= \frac{\sum x_i}{N} \equiv \bar{x} \end{aligned}$$

Wir finden als besten Schätzwert den arithmetischen Mittelwert der N Messwerte. Ähnlich kann man auch den Bestwert für den Breiteparameter σ finden:

$$\begin{aligned} \frac{\partial P}{\partial \sigma} &= 0 \\ -N \frac{1}{\sigma^{N+1}} e^{-\sum (x_i - x_w)^2 / 2\sigma^2} + \frac{1}{\sigma^N} \left(-\frac{\sum (x_i - x_w)^2}{-\sigma \cdot \sigma^2} \right) e^{-\sum (x_i - x_w)^2 / 2\sigma^2} &= 0 \\ \frac{1}{\sigma^{N+3}} \left[\sum (x_i - x_w)^2 - N\sigma^2 \right] e^{-\sum (x_i - x_w)^2 / 2\sigma^2} &= 0 \\ \Rightarrow \sum (x_i - x_w)^2 - N\sigma^2 &= 0 \\ \text{bzw. } \sigma^2 &= \frac{1}{N} \sum (x_i - x_w)^2 = s^2 \end{aligned}$$

Ersetzt man x_w durch seinen besten Schätzwert, so geht ein Freiheitsgrad verloren und man findet als Standardabweichung:

$$s = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}$$

5.4 Standardabweichung des Mittelwerts

Wir haben gesehen, dass der Mittelwert eines Datensamples $\{x_1, \dots, x_N\}$ den besten Schätzwert liefert

$$\bar{x} = \frac{x_1 + \dots + x_N}{N}$$

Die Unsicherheit dieses Schätzwertes wird durch die Standardabweichung des Mittelwertes gegeben. Wir nehmen an, die x_i sind normalverteilt um den "wahren" Wert und die Verteilung hat die Breite s_{x_i} . Die Messunsicherheit des Mittelwerts berechnet sich dann nach der Fehlerfortpflanzung:

$$s_{\bar{x}} = \sqrt{\left(\frac{\partial \bar{x}}{\partial x_1} s_{x_1}\right)^2 + \dots + \left(\frac{\partial \bar{x}}{\partial x_N} s_{x_N}\right)^2}$$

mit $s_1 = \dots = s_N = s_x$ und $\frac{\partial \bar{x}}{\partial x_i} = \frac{1}{N}$ folgt:

$$\begin{aligned} s_{\bar{x}} &= \sqrt{\left(\frac{1}{N} s_x\right)^2 + \dots + \left(\frac{1}{N} s_x\right)^2} = \sqrt{\frac{N \cdot s_x^2}{N^2}} \\ &= \frac{s_x}{\sqrt{N}} \end{aligned}$$

5.5 Verwerfen von Daten

Ein heikles Thema ist die Frage, ob man auffällige Messwerte verwerfen darf. Einerseits könnten sie zu falschen Ergebnissen führen, andererseits vernichtet man eventuell einen wichtigen, bisher unbekanntem Effekt. Generell sollte die Manipulation von Messwerten so restriktiv wie möglich gehandhabt werden. Folgendes Beispiel soll das Problem verdeutlichen.

Beispiel: Man misst die Schwingungsdauer eines Pendels und findet:

$$T \text{ in sec } \quad 3.8, 3.7, 3.5, 3.9, 3.7 \text{ und } 1.8$$

Zugegebenermaßen ist der Wert 1.8 sec sehr auffällig. Darf man ihn aber verwerfen? Seine Auswirkung wird durch den Vergleich der Mittelwerte und Standardabweichungen deutlich:

$$\bar{x} = 3.4 \text{ sec} \quad s_x \pm 0.8 \text{ sec}$$

im Vergleich zu den Werten ohne den verdächtigen Messwert: 3.7 sec und $s_x = \pm 0.2$ sec.

Die ehrlichste Methode in einem solchen Fall ist sicherlich, die Messung sehr viele Male zu wiederholen, so dass sich ein Ausreißer nicht sonderlich stark auswirkt. Wenn das nicht geht, hilft einem das **Chauvenetsche Kriterium**, eine nachvollziehbare Entscheidung bzgl. der Verwendung eines Messwertes zu treffen. Es besagt:

Ein verdächtiger Messwert in einem Datensatz $\{x_1, \dots, x_N\}$ ist dann zu verwerfen, wenn die Wahrscheinlichkeit von Messwerten, die mindestens so schlecht sind wie der verdächtige Wert, kleiner als 0.5 ist

Man berechnet also zunächst, wieviele Standardabweichungen liegt der verdächtige Wert vom Mittelwert weg:

$$t_{\text{verd}} = \frac{|x_{\text{verd}} - \bar{x}|}{s_x}$$

Für das obige Zahlenbeispiel berechnet man $t_{\text{verd}} = 2$. Die Wahrscheinlichkeit für einen Messwert ausserhalb $t_{\text{verd}} s_x$ ist: $P(\text{ausserhalb } 2\sigma) = 0.05$. Diesen Wert multipliziert man mit der Anzahl von Messwerten N und erhält das Kriterium:

$$n_{\text{ch}} = N \cdot P(\text{ausserhalb } t_{\text{verd}} s_x)$$

In unserem Beispiel findet man $n_{ch} = 0.3$. Dieser Wert ist kleiner als 0.5 und man darf daher den Wert 1.8 ausschliessen. Das korrekte Ergebnis für die Schwingungsdauer des Pendels ist also:

$$\bar{x} = (3.7 \pm 0.2)\text{sec}$$

5.6 Diskrepanz von Messergebnissen

Falls zwei Messwerte derselben Grösse nicht übereinstimmen, spricht man von **Diskrepanz**. Zahlenmässig ist Diskrepanz die Differenz der beiden Messwerte. Zur Beurteilung ihrer Signifikanz muss man die Messunsicherheiten betrachten. Im allgemeinen bewertet man eine Diskrepanz als signifikant, wenn die Messwerte mehr als das 2-fache der Messunsicherheit auseinander liegen. Man spricht auch von der **2 σ Grenze**. Die Wahrscheinlichkeit, dass ein Messwert ausserhalb der 2 σ Grenze liegt, beträgt 4.5%.

Das Auftreten einer solch grossen Abweichung deutet oft auf eine nicht entdeckte Quelle systematischer Abweichung hin. Es ist dann die Aufgabe des Experimentators, durch sorgfältige Überprüfung der Messapparatur und des Messverfahrens diese Quelle zu finden und zu eliminieren.

6 Ausgleichsrechnung

Bei Ausgleichsrechnungen unterscheidet man zwei grundlegende Fälle:

- *Ausgleich direkter Beobachtungen*: Hierbei wird für eine fehlerhafte Grösse, die in unterschiedlichen Messungen mit unterschiedlicher Genauigkeit bestimmt wurde, der beste Schätzwert ermittelt.
- *Ausgleich von vermittelnden Beobachtungen*: Hierbei ist die gesuchte Grösse durch einen funktionalen Zusammenhang aus den fehlerbehafteten Messwerten zu bestimmen.

6.1 Der gewichtete Mittelwert

Die Gauss'sche Verteilungsfunktion gestattet die Berechnung der Wahrscheinlichkeit, dass ein Messwert x_i innerhalb eines gegebenen Intervalls liegt. Jedem Messwert ist eine Varianz σ_i^2 zugeordnet. In der Regel sind sie für die einzelnen Messergebnisse unterschiedlich, da es sich um von einander unabhängige Messungen handelt (Verwendung eines anderen Messverfahrens oder einer anderen Messreihe). Die Wahrscheinlichkeit, gerade den beobachteten Satz von Messergebnissen zu erhalten, ist dann gegeben durch:

$$P(x_1, \dots, x_N) = \frac{1}{\sigma_1 \sigma_2 \dots \sigma_N \sqrt{(2\pi)^N}} e^{-\sum (x_i - x_w)^2 / 2\sigma_i^2}$$

Den besten Schätzwert für x_w erhält man nach der Methode der maximalen Wahrscheinlichkeit, d.h. $P(x_1, \dots, x_N)$ muss maximal sein bzw. der Exponent minimal.

$$\begin{aligned} \chi^2 = \sum_{i=1}^N \frac{(x_i - x_w)^2}{\sigma_i^2} &\Rightarrow \frac{\partial \chi^2}{\partial x_w} = 2 \sum_{i=1}^N \frac{(x_i - x_w)}{\sigma_i^2} = 0 \\ &\Rightarrow x_{best} = \frac{\sum \frac{x_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} \end{aligned}$$

Dieses Ergebnis schreibt sich eleganter, wenn man die reziproke Varianz als Gewicht einführt.

$$\text{mit } w_i = \frac{1}{\sigma_i^2} \Rightarrow x_{best} = \frac{\sum w_i x_i}{\sum w_i} \quad (14)$$

Falls die w_i gleich gross sind, vereinfacht sich der gewichtete Mittelwert auf den arithmetischen Mittelwert. Da das Gewicht einer Einzelmessung gleich der reziproken Varianz ist, trägt ein weniger präzises

gemessener Messwert viel weniger zum Endergebnis bei. Um die Unsicherheit in dem gewichteten Mittelwert zu bestimmen, benützt man das Fehlerfortpflanzungsgesetz.

$$\sigma_{x_{best}}^2 = \sum \left(\frac{\partial x_{best}}{\partial x_i} \sigma_{x_i} \right)^2$$

mit $\frac{\partial x_{best}}{\partial x_i} = \frac{w_i}{\sum w_i}$ und $\sigma_{x_i} = \frac{1}{\sqrt{w_i}}$

$$\Rightarrow \sigma_{x_{best}} = \sqrt{\frac{1}{\sum w_i}} \quad (15)$$

6.2 Lineare Regression

Im folgenden geht es darum, unbekannte Parameter in einer als bekannt vorausgesetzten Funktion, die mehrere gemessene Grössen mit einander verknüpft, zu bestimmen. Die Messfehler der gemessenen Grössen übertragen sich auf die Unbekannten. Der funktionale Zusammenhang kann entweder aus der Theorie folgen oder aber als zu prüfende Hypothese aufgestellt werden. Auch hier wird das Prinzip der maximalen Wahrscheinlichkeit in der Form der kleinsten Fehlerquadrate (**least square fit**) verwendet, um die unbekannt Parameter zu bestimmen.

Die folgende Diskussion beschränkt sich auf einen linearen Zusammenhang. Gegebenenfalls muss eine Linearisierung der Funktion vorgenommen werden. Zum Beispiel:

- Lineare Funktion, elastische Deformation $F=k \cdot x$; dargestellt werden $\{F, x\}$
- Einfach logarithmischer Zusammenhang, Temperaturabhängigkeit der Leitfähigkeit eines Halbleiters

$$\sigma = \sigma_0 e^{-\frac{E}{2kT}} \quad \Leftrightarrow \quad \ln \sigma = \ln \sigma_0 - \frac{E}{2kT}; \text{ dargestellt werden } \{\ln \sigma, 1/T\}$$

Auch wenn man fest entschlossen ist, einen rechnerischen Ausgleich durchzuführen, ist eine graphische Darstellung der Messwerte immer hilfreich, da sie einen schnellen Überblick über die Qualität der Messungen (etwaige Ausreisser) gibt. Für einen linearen Zusammenhang $y=A+B \cdot x$ lassen sich einfache Formeln für die Parameter und ihre Standardabweichungen angeben. Wir nehmen an, dass nur die y-Werte zu berücksichtigende Messunsicherheiten aufweisen und diese alle gleich gross sind. Dann beträgt die Wahrscheinlichkeit, einen Messwert y_i zu beobachten:

$$P(y_i) \propto \frac{1}{\sigma_y} e^{-(y_i - A - B \cdot x_i)^2 / 2\sigma_y^2}$$

Die Wahrscheinlichkeit, den gegebenen Satz von y-Messwerten zu finden, ist das Produkt der Einzelwahrscheinlichkeiten.

$$P_{tot} \propto \frac{1}{\sigma_y^N} e^{-\sum (y_i - A - B \cdot x_i)^2 / 2\sigma_y^2}$$

Die besten Schätzwerte für A, B erhält man nach der maximum likelihood Methode. Dazu muss der Exponent minimalisiert werden.

$$\chi^2 = \frac{\sum_{i=1}^N (y_i - A - B \cdot x_i)^2}{\sigma_y^2}$$

$$\frac{\partial \chi^2}{\partial A} = 0 = -\frac{2}{\sigma_y^2} \sum_{i=1}^N (y_i - A - B \cdot x_i)$$

$$\frac{\partial \chi^2}{\partial B} = 0 = -\frac{2}{\sigma_y^2} \sum_{i=1}^N x_i (y_i - A - B \cdot x_i)$$

Diese Gleichungen lassen sich in ein Gleichungssystem umformen, aus dem A, B leicht zu bestimmen ist.

$$\begin{aligned} A \cdot N + B \cdot \sum x_i &= \sum y_i \\ A \cdot \sum x_i + B \cdot \sum x_i^2 &= \sum y_i x_i \end{aligned}$$

$$\Rightarrow A = \frac{(\sum x_i^2) \cdot (\sum y_i) - (\sum x_i) \cdot (\sum x_i y_i)}{\Delta} \quad \text{mit } \Delta = N \cdot (\sum x_i^2) - (\sum x_i)^2$$

$$B = \frac{N \cdot (\sum x_i y_i) - (\sum x_i) \cdot (\sum y_i)}{\Delta} \quad (16)$$

Als nächstes sollen nun die Unsicherheiten in den Parametern A und B bestimmt werden. Betrachten wir zunächst die Standardabweichung der y-Werte. Wir gehen davon aus, dass die Abweichungen $(y_i - A - B \cdot x_i)$ normalverteilt sind und $(A + B \cdot x_i)$ den "wahren" Wert y_w beschreibt. Dann ergibt sich die Varianz:

$$\sigma_y^2 = \frac{1}{N} \sum (y_i - A - B \cdot x_i)^2$$

Die gleichen Argumente, die bei der Einführung der Varianz in Kapitel 3 angeführt wurden, müssen hier wiederholt werden. Durch die zwei zu bestimmenden Parameter ist die Anzahl der Freiheitsgrade um 2 reduziert und wir müssen den Faktor $1/N$ durch den Faktor $1/(N-2)$ ersetzen, d.h.

$$\sigma_y^2 = \frac{1}{N-2} \sum (y_i - A - B \cdot x_i)^2$$

Nun können wir die Unsicherheiten in den Parametern A und B mit Hilfe der Fehlerfortpflanzungsgesetz berechnen:

$$\begin{aligned} \sigma_A^2 &= \sum \left(\frac{\partial A}{\partial y_i} \cdot \sigma_{y_i} \right)^2 \\ &= \sum \left(\frac{(\sum x_i^2) - x_j \sum x_i}{\Delta} \cdot \sigma_{y_i} \right)^2 \\ &= \sum_{j=1}^N \left(\frac{\sigma_{y_j}^2}{\Delta^2} \left[(\sum x_i^2)^2 - 2 \cdot (\sum x_i^2) \cdot x_j \cdot (\sum x_i) + x_j^2 \cdot (\sum x_i)^2 \right] \right) \\ &= \frac{\sigma_{y_j}^2}{\Delta^2} \left[N \cdot (\sum x_i^2)^2 - 2 \cdot (\sum x_i^2) \cdot (\sum x_i)^2 + (\sum x_i^2) \cdot (\sum x_i)^2 \right] \\ &= \frac{\sigma_y^2}{\Delta^2} \cdot \sum x_i^2 \underbrace{\left[N \cdot \sum x_i^2 - (\sum x_i)^2 \right]}_{\Delta} \\ &\Rightarrow \sigma_A^2 = \frac{\sigma_y^2}{\Delta} \cdot \sum x_i^2 \quad (17) \end{aligned}$$

Ganz entsprechend lässt sich die Varianz von B berechnen:

$$\sigma_B^2 = \frac{\sigma_y^2}{\Delta} \cdot N \quad (18)$$

6.3 Gewichtete lineare Regression

Bisher wurde stillschweigend angenommen, dass jedes Wertepaar die gleiche Bedeutung für die Anpassung hat. Möchte man aber weniger genau gemessene Werte mit geringerer Wertung berücksichtigen, so muss man eine ‚gewichtete Regression‘ durchführen. Als Wichtungsfaktoren werden die reziproken Varianzen verwendet.

$$\begin{aligned} \Rightarrow A &= \frac{\left(\sum w_i x_i^2\right) \cdot \left(\sum w_i y_i\right) - \left(\sum w_i x_i\right) \cdot \left(\sum w_i x_i y_i\right)}{\Delta} \\ B &= \frac{\sum w_i \cdot \left(\sum w_i x_i y_i\right) - \left(\sum w_i x_i\right) \cdot \left(\sum w_i y_i\right)}{\Delta} \quad (19) \\ \text{mit } \Delta &= \sum w_i \cdot \left(\sum w_i x_i^2\right) - \left(\sum w_i x_i\right)^2 \end{aligned}$$

und für die Varianzen der Parameter findet man:

$$\Rightarrow \sigma_A^2 = \frac{\sigma_y^2}{\Delta} \cdot \sum w_i x_i^2 \quad \text{bzw.} \quad \sigma_B^2 = \frac{\sigma_y^2}{\Delta} \cdot \sum w_i \quad (20)$$

6.4 Der lineare Korrelationskoeffizient

Die Qualität der Anpassung $y=A+b \cdot x$ an die Wertepaare $\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_N, y_N\}$ kann mit Hilfe des linearen Korrelationskoeffizienten getestet werden. Er ist definiert als:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (21)$$

mit den Definitionen

$$\begin{aligned} \sigma_x^2 &= \frac{1}{N} \sum (x_i - \bar{x})^2, \quad \sigma_y^2 = \frac{1}{N} \sum (y_i - \bar{y})^2 \quad \text{und} \quad \sigma_{xy} = \frac{1}{N} \sum (x_i - \bar{x}) \cdot (y_i - \bar{y}) \\ \Rightarrow r &= \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (22) \end{aligned}$$

Dieser Wert liegt zwischen -1 und 1. Liegt er sehr nahe bei ± 1 , so bedeutet das, dass die Gerade die Wertepaare sehr gut beschreibt. $R > 0,9$ kann als Beweis eines linearen Zusammenhang angesehen werden. Liegt r nahe bei 0, so sind x, y nicht korreliert.

6.5 Lineare Regression - graphische Lösung

Die graphische Darstellung einer Messreihe $\{(x_1, y_1), \dots, (x_N, y_N)\}$ dient nicht nur dazu, eine Veranschaulichung des funktionellen Zusammenhangs zweier Messgrößen zu vermitteln, sondern kann auch zur quantitativen Auswertung verwendet werden. Beim Zeichnen der graphischen Darstellung beachte man:

- Verwendung von Millimeterpapier und einem mittelharten Bleistift.
- Wahl einer geeigneten Skalierung, die den gesamten Wertebereich ausnützt.
- Beschriftung der Achsen und Einfügen eines Titels.
- Sorgfältige Eintragung der Messpunkte als kleine Kreuze.
- Bei bekannten Messabweichungen werden die Messpunkte mit Fehlerbalken versehen.

Man zeichnet dann eine Gerade, die am besten die Messpunkte beschreibt.

Zur quantitativen Auswertung wählt man dann zwei Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ auf der Ausgleichsgeraden in etwa auf der Höhe des ersten bzw. des letzten Messpunktes. Dann ergibt sich die Steigung zu

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{und die Achsenabschnitte } a_x \text{ und } a_y \text{ berechnen sich aus}$$

$$a_x = x_1 - \frac{y_1}{m} \quad \text{bzw.} \quad a_y = y_1 - m \cdot x_1.$$

Für die Angabe der Unsicherheit in den Größen m , a_x , a_y benötigt man zunächst eine Abschätzung des absoluten Fehlers der Einzelmessung. Man kann dazu die grösste Abweichung eines Messpunktes Δx bzw. Δy von der gezeichneten Gerade verwenden. Nimmt man an, dass alle Messpunkte mit der gleichen Unsicherheit behaftet sind, dann ergeben sich die folgenden Unsicherheiten.

$$\frac{\delta m}{m} = \frac{2\Delta x}{x_2 - x_1} \quad \text{bzw.} \quad \frac{\delta m}{m} = \frac{2\Delta y}{y_2 - y_1}$$

$$\text{bzw.} \quad a_x = \frac{y_1 + y_2}{2} \cdot \frac{\delta m}{m^2} \quad \text{und} \quad a_y = \frac{x_1 + x_2}{2} \cdot \delta m$$

7 Der χ^2 Test für eine Verteilung

Bei der Auswertung von Experimenten stösst man häufig auf das Problem, den experimentellen Messpunkten theoretische Zahlenwerte bzw. geschlossene Kurven zuzuordnen. Auf Grund der statistischen Schwankungen der gemessenen Größen sind in der Regel verschiedene Werte der theoretischen Parameter mit dem Experiment verträglich. Zur Entscheidung, welche Kurve die experimentellen Ergebnisse am besten beschreibt, bedient man sich eines statistischen Prüfverfahrens, dem χ^2 -Test.

Die χ^2 - Funktion ist definiert als:

$$\chi^2 = \sum \left(\frac{x_i^{\text{exp}} - x_i^{\text{th}}}{\sigma_i} \right)^2$$

Die Summation läuft über alle unabhängigen Messwerte. χ^2 ist eine Funktion der theoretischen Parameter. Die beste Übereinstimmung zwischen Messwert und Theorie entspricht dem Minimum von χ^2 . Daher werden die theoretischen Parameter solange variiert, bis man den Minimalwert der χ^2 Verteilung erhält. Die zugeordneten Parameterwerte sind dann die besten Schätzwerte. Die Abbildung zeigt eine typische χ^2 - Minimalisierungskurve für einen Anpassungstest mit einem Parameter.

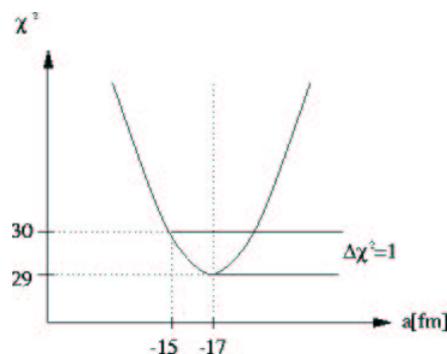


Abbildung: χ^2 Minimalisierung für die Winkelverteilung bei Neutron-Neutron-Streuung. Einziger Parameter ist die Stärke der Neutron-Neutron-Kraft (Streulänge a).

Bei der Auswertung von Experimenten interessiert neben den Schätzwerten der Parameter ebenfalls die mögliche Unsicherheit dieser Schätzungen. Man definiert daher für das Prüfverfahren einen statistischen Fehler des geschätzten Parameters. Er ergibt sich aus der Erhöhung von $\chi^2(\text{min})$ auf $\chi^2(\text{min})+1$.

Weiterhin von Bedeutung ist die Wahrscheinlichkeit, bei einer Wiederholung des Experimentes trotz identischer Versuchsbedingungen einen anderen χ^2 Wert zu erhalten. In der Statistik wird gezeigt, daß für einen grossen Stichprobenumfang n die zu χ^2 gehörende Wahrscheinlichkeitsverteilung die Form hat:

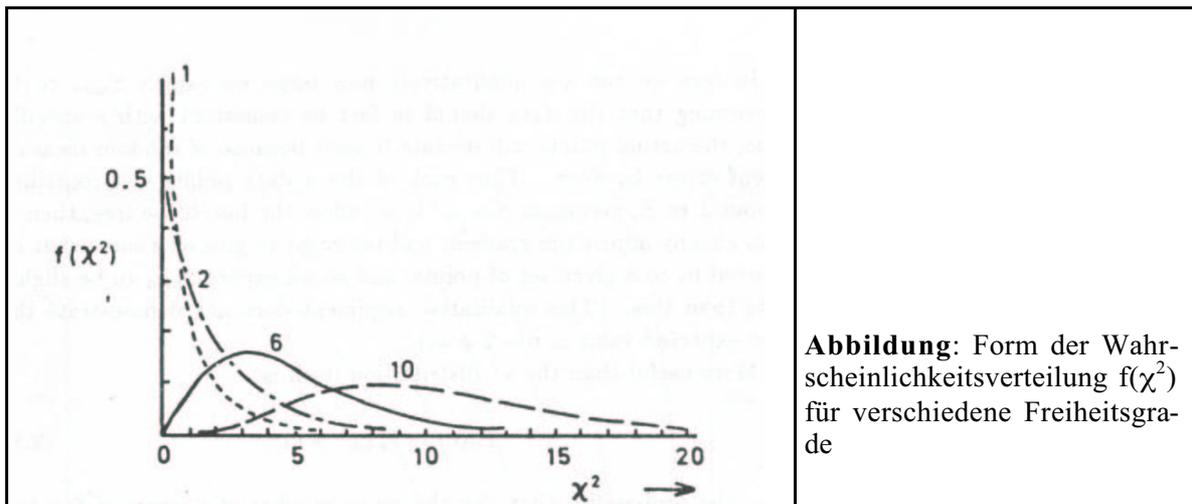
$$f(\chi^2) = \frac{1}{2^{\frac{\nu}{2}} \cdot (\frac{\nu-2}{2})!} \cdot (\chi^2)^{\frac{\nu-2}{2}} \cdot e^{-\frac{\chi^2}{2}}$$

Die Anzahl der Freiheitsgrade ν ist gleich der Anzahl der beobachteten Daten minus der aus den Daten zu berechnenden Parametern, d.h. gleich der Anzahl der unabhängigen Bestimmungsgleichungen $G(c,y)=0$ zwischen dem Parametervektor c und dem Beobachtungsvektor y . Dies sei am folgenden Beispiel erläutert:

Beispiel: Es werde eine Zählrate unter identischen Bedingungen beliebig oft aufgenommen. Für die Streuung der Messwerte werde eine Unterteilung in 20 fest gewählte Intervalle ΔN_i gemacht. Es wird die Häufigkeit registriert, mit der sich die Meßwerte auf die einzelnen Intervalle verteilen. Man sagt, für das Prüfverfahren stehen 20 Stützpunkte oder auch Klassen zur Verfügung. Die theoretische Verteilungsfunktion hat p freie Parameter, durch deren Variation die Anpassung an die gemessene Verteilung optimal gemacht werden kann. Bestimmt man ferner den Faktor für die Normierung der theoretische Verteilungsfunktion auf die Anzahl der Meßwerte, so beträgt in diesem Beispiel die Anzahl der Freiheitsgrade schließlich

$$\nu = 20 - p - 1 = 17$$

Die Anzahl der unabhängigen Bestimmungsmöglichkeiten erhält man also, wenn man von der Anzahl der Stützstellen die Zahl der variierten unabhängigen Parameter abzieht.

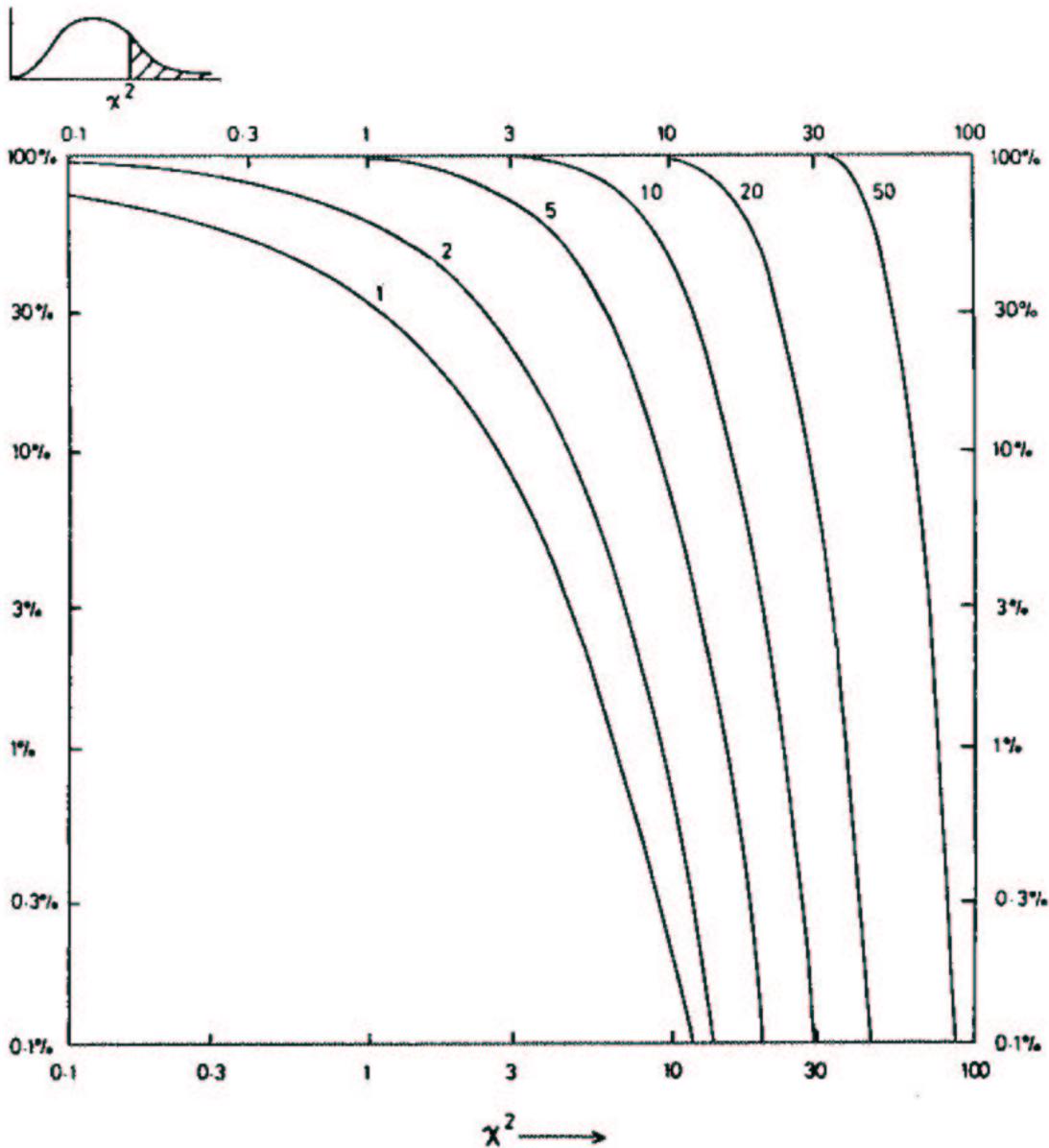


Den Verlauf der Wahrscheinlichkeitsverteilung $f(\chi^2)$ für $\nu = 1, 2, 6$ und 10 zeigt die Abbildung. Die Kurven $f(\chi^2)$ haben ihr Maximum etwas unterhalb des Wertes $\chi^2 = \nu$, d.h. wenn unser Ergebnis $\chi^2 \gg \nu$ lautet, ist es höchst unwahrscheinlich, dass unser vorhergesagtes theoretisches Modell richtig ist.

Es ist bequemer mit der Grösse

$$F(\chi_0^2) = \int_{\chi_0^2}^{\infty} f(\chi^2) d\chi^2$$

zuarbeiten. Das Integral gibt die Wahrscheinlichkeit an, bei einer Wiederholung des Experimentes einen Wert $\chi^2 > \chi_0^2$ zu finden. $F(\chi_0^2)$ ist daher ein direktes Maß für die Güte einer Hypothese. Sehr kleine Werte, $F(\chi_0^2) < 0,1$ deuten eine schlechte Übereinstimmung zwischen Experiment und Theorie an. Der Zusammenhang $\chi_0^2 \rightarrow F(\chi_0^2)$ als Funktion des Freiheitsgrades ν ist aus einem entsprechenden Tafelwerk oder der folgenden Abbildung zu entnehmen.



Man spricht bei einer Wahrscheinlichkeit von kleiner 0,05 davon, dass die Abweichung Experiment - Theorie signifikant sei. Ist die Wahrscheinlichkeit sogar kleiner 0,01, dann wird die Abweichung als hoch signifikant bezeichnet.

8 Empfehlenswerte Bücher zur Fehlerrechnung

Die folgende Liste gibt eine Auswahl an empfehlenswerten Büchern zu diesem Thema:

W.H. Gränicher, *Messung beendet - was nun?*,
Teubnerverlag Stuttgart, 1996

B.P. Roe, *Probability and Statistics in Experimental Physics*,
Springerverlag Berlin, 1992

P.R. Bevington and D.K. Robinson, *Data Reduction and Error Analysis for the Physical Science*, Mc-
Graw-Hill, 1992

J.R. Taylor, *Fehleranalyse*,
VCH-Verlag, 1988