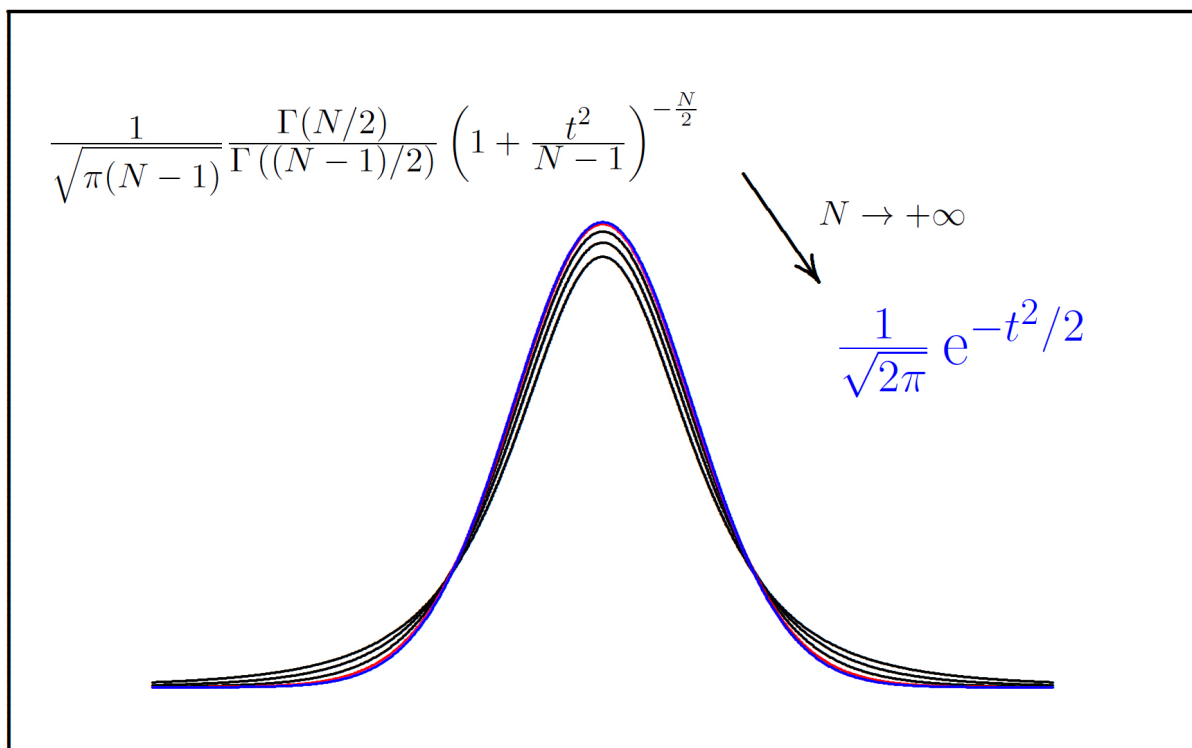
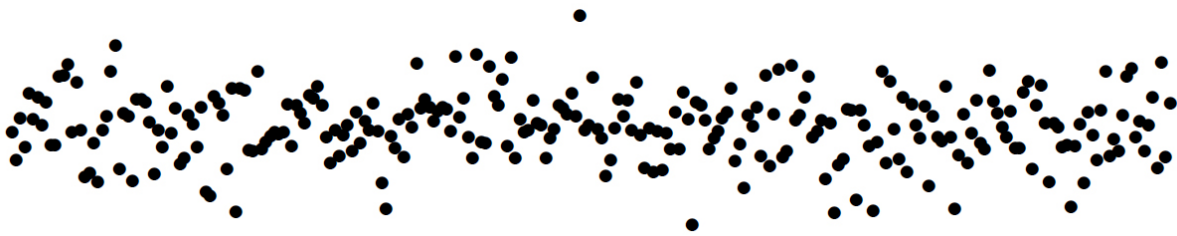


# Messabweichungen im physikalischen Experiment

Bernd Pompe

Fassung vom 25. Oktober 2017



## Zusammenfassung

Es wird eine Einführung in die Behandlung von Messdaten und die Schätzung ihrer Abweichungen vom unbekanntem wahren Wert gegeben. Sie richtet sich vor allem an Studenten verschiedener Studienrichtungen, die in einem physikalischen Praktikum die „Kunst des Messens“ erlernen wollen. Dem Experimentator wird in kompakter Form das nötige Handwerkszeug für den sicheren Umgang mit experimentell gewonnenen Daten bereitgestellt. Die Darstellung bemüht sich um Anschaulichkeit einiger grundlegender Begriffe und Verfahren. Wer jedoch die streng mathematischen Grundlagen erlernen möchte, muss auf die weiterführende Literatur zur Wahrscheinlichkeitsrechnung und Statistik verwiesen werden.

## Inhaltsverzeichnis

<b>1</b>	<b>Grundbegriffe</b>	<b>3</b>
1.1	Messen, Maßzahl und Einheit . . . . .	3
1.2	Messreihe . . . . .	3
1.3	Wahrer Wert . . . . .	5
1.4	Verteilungsfunktion . . . . .	5
1.5	Verteilungsdichte . . . . .	6
1.6	Normalverteilung . . . . .	6
1.7	Quantile und Boxplot . . . . .	7
<b>2</b>	<b>Analyse von Messabweichungen</b>	<b>9</b>
2.1	Grobe Messabweichung (Ausreißer) . . . . .	9
2.2	Mittelwert . . . . .	10
2.3	Stationarität . . . . .	10
2.4	Zufällige Messabweichungen . . . . .	11
2.5	Systematische Messabweichung . . . . .	11
2.6	Standardabweichung der Einzelmessung . . . . .	12
2.7	Normierung . . . . .	13
2.8	Vertrauensbereich . . . . .	13
2.9	$t$ -Verteilung . . . . .	14
2.10	Gesamte Messabweichung . . . . .	18
2.11	Beispiel . . . . .	18
2.12	Anmerkungen . . . . .	20
<b>3</b>	<b>Fehlerfortpflanzung</b>	<b>20</b>
3.1	Problemstellung . . . . .	20

---

3.2	Statistische Fehlerfortpflanzung . . . . .	21
<b>4</b>	<b>Regression</b>	<b>22</b>
4.1	Problemstellung . . . . .	22
4.2	Lineare Regression . . . . .	23
<b>A</b>	<b>Anhang: Lineare Regression</b>	<b>24</b>
A.1	1. Fall: $x$ fehlerfrei, $y$ mit festem Fehler $\sigma_y$ . . . . .	24
A.2	2. Fall: $x_n$ und $y_n$ mit Fehler $\sigma_x$ bzw. $\sigma_y$ . . . . .	25
A.3	3. Fall: $x_n$ fehlerfrei, $y_n$ mit variablem Fehler $\sigma_{y;n}$ . . . . .	25
A.4	4. Fall: $x_n$ und $y_n$ mit variablem Fehler $\sigma_{x;n}$ bzw. $\sigma_{y;n}$ . . . . .	26

---

# 1 Grundbegriffe

## 1.1 Messen, Maßzahl und Einheit

Naturwissenschaften wie die Physik entwickeln Theorien über Phänomene und Vorgänge in der Natur, deren Wahrheitsgehalt durch Beobachtungen und Experimente beurteilt wird. Der erste Schritt hierzu sind Messungen, bei denen eine physikalische Größe  $G$  bestimmt wird, für die eine *Maßeinheit*  $[G]$  und eine Messvorschrift wohl definiert ist. Messen bedeutet, das Vielfache  $\{G\}$  der Maßeinheit zu ermitteln.  $\{G\}$  heißt *Maßzahl*. Das Messergebnis wird geschrieben als

$$G = \{G\}[G] .$$

Gelte beispielsweise für die Messung einer Länge,

$$L = 5 \text{ m} ,$$

dann sind  $\{L\} = 5$  die Maßzahl und  $[L] = \text{m}$  die Maßeinheit, das *Meter*. Folgende Schreibweisen sind ebenso zulässig,<sup>1)</sup>

$$L/\text{m} = \frac{L}{\text{m}} = 5 .$$

Beträgt eine Maßzahl null, so lässt man in der Regel die Einheit weg, man schreibt also zum Beispiel  $L = 0$  statt  $L = 0 \text{ m}$ . Eine Ausnahme stellen Temperaturangaben wie  $\vartheta = 0 \text{ }^\circ\text{C}$  dar, um Verwechslungen von Temperaturskalen, etwa der Kelvin- mit der Celsius-Skala, zu vermeiden.

In der Physik werden nach dem Internationalen Einheitensystem 7 Basisgrößen und zwei ergänzende Einheiten definiert. Das sind

Basisgröße und Einheit	
die Länge $l$	das (auch der) Meter m
die Masse $m$	das Kilogramm kg
die Zeit $t$	die Sekunde s
die elektrische Stromstärke $I$	das Ampere A
die Temperatur $T$	das Kelvin K
die Stoffmenge	das Mol mol
die Lichtstärke	die Candela cd
ergänzende Größe und Einheit	
ebener Winkel $\varphi$	der Radiant rad
räumlicher Winkel $\Omega$	der Steradian sr

Alle anderen physikalischen Größen sind daraus abgeleitet, wie zum Beispiel die kinetische Energie  $E_{\text{kin}}$  bei der Translation eines Körpers der Masse  $m$  ( $[m] = \text{kg}$ ) und der Geschwindigkeit  $v$ ,  $[v] = \text{m/s}$ ,<sup>2)</sup>

$$E_{\text{kin}} = \frac{1}{2}mv^2 , \quad \text{mit der Maßeinheit} \quad [E_{\text{kin}}] = \frac{\text{kg} \cdot \text{m}^2}{\text{s}^2} .$$

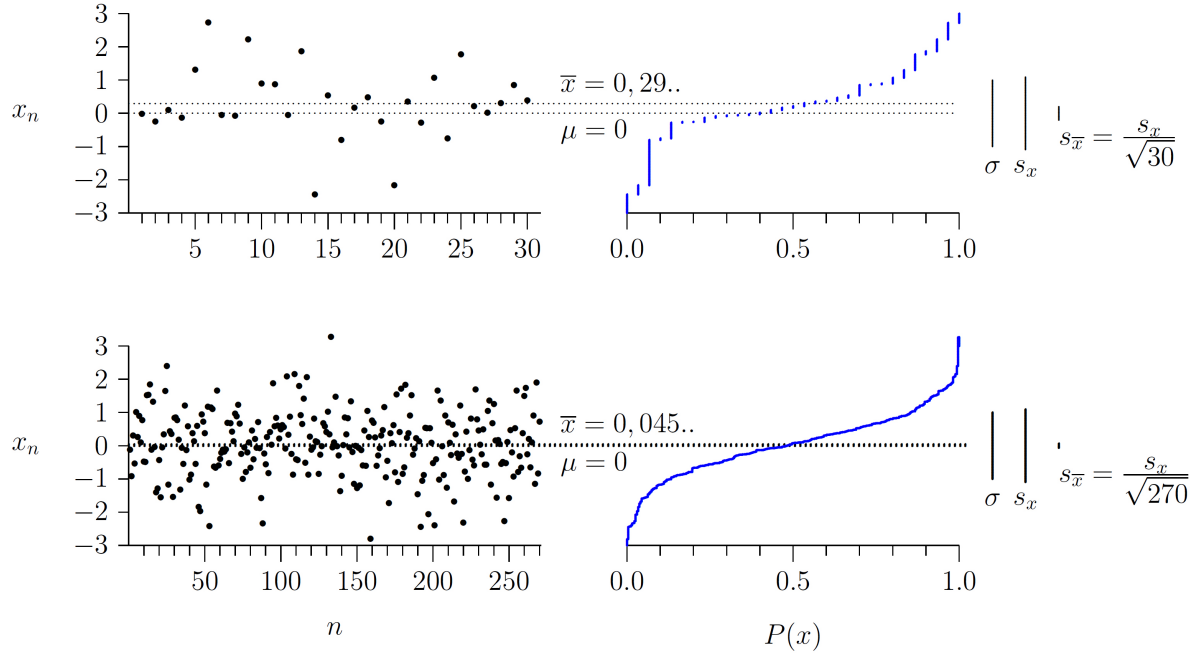
## 1.2 Messreihe

In einem Experiment soll eine physikalische Größe gemessen werden, beispielsweise die Lufttemperatur. Wird eine solche Messung mehrmals ausgeführt, so wird in der Regel nicht immer der selbe Messwert

---

<sup>1)</sup> Man beachte, dass beim exakten Formelsatz die physikalische Größe kursiv und die Maßeinheit steil zu setzen sind. Folglich bezeichnet zum Beispiel m die Maßeinheit Meter und  $m$  die physikalische Größe Masse.

<sup>2)</sup> Beim exakten Formelsatz wird eine physikalische Größe mit nur einem Buchstaben bezeichnet. Möchte man beispielsweise verschiedene Energieformen wie die kinetische und die potentielle unterscheiden, so macht man dies mit Indizes, die steil gesetzt werden, man schreibt also  $E_{\text{kin}}$  und  $E_{\text{pot}}$ . Als Index sind möglichst Kürzel zu verwenden oder besser nur ein Buchstabe. Man sollte sich jedoch an die übliche Notation halten.



**Abb. 1:** Konstruktion der empirischen Verteilungsfunktion  $P(x)$  für Stichproben  $\{x_n\}_{n=1}^N$  unterschiedlicher Längen,  $N = 30$  (oben) und  $N = 270$  (unten). Die Daten stammen aus einer Computersimulation. Sie sind normalverteilt und voneinander unabhängig, mit dem Mittelwert  $\mu = 0$  und der Standardabweichung  $\sigma = 1$ . Eingetragen sind auch jeweils der empirische Mittelwert  $\bar{x}$ , die Standardabweichung der Einzelmessung  $s_x$  sowie der Vertrauensbereich  $s_{\bar{x}}$ .

erhalten. Die Folge der  $N$  Messwerte, die *Messreihe* oder auch *Stichprobe*,

$$\{x_1, x_2, \dots, x_n, \dots, x_N\} \equiv \{x_n\}_{n=1}^N \quad (1)$$

schwankt also im Allgemeinen. Folglich stellt sich die Frage, wie man einen Wert findet, der dem unbekanntem wahren Wert  $x_w$  möglichst nahe kommt.

Abbildung 1 (links) zeigt zwei Beispiele für Messreihen der Länge  $N = 30$  bzw.  $N = 270$ .<sup>3)</sup> Die Werte schwanken um den als unbekannt angenommenen wahren Wert  $x_w$ .<sup>4)</sup> Allerdings sind die meisten Werte dichter als 3 beim wahren Wert.

Wegen der zufälligen Schwankungen fasst man eine jede Messung als Ergebnis eines *Zufallsexperimentes* auf. Das Ergebnis einer Messung  $x$  wird dann als Realisierung einer *Zufallsgröße*  $X$  angesehen. Eine Zufallsgröße ordnet einem zufälligen Ereignis  $E$  (etwa eine bestimmte Zeigerstellung des Messgerätes) eine reelle Zahl  $x$  zu,  $X : E \rightarrow x$ . Das Ereignis  $E$  besteht hier also darin, dass das Messinstrument einen bestimmten Messwert anzeigt. Die reelle Zahl ist dann die Maßzahl der Messgröße. Die zugehörige Maßeinheit ergibt sich aus der Beschriftung oder anderen Angaben zum Messgerät.

<sup>3)</sup> Für die folgenden Betrachtungen ist die Maßeinheit  $[x]$  nicht von Bedeutung und wird deshalb weggelassen, es werden also nur die Maßzahlen  $\{x\}$  betrachtet. Man kann sich auch auf den Standpunkt stellen, die Messgröße  $x$  habe die Einheit  $[x] = 1$ . Man sagt dann auch, „ $x$  ist einheitenlos“. Im Englischen schreibt man auch  $[x] = a.u.$ , was für „arbitrary unit“ steht.

<sup>4)</sup> Der wahre Wert ist in der Abbildung konkret angegeben,  $x_w = 0$ . Er ist hier bekannt, weil die beiden Datensätze aus einer Computersimulation stammen. In der Messpraxis ist  $x_w$  jedoch immer unbekannt.

---

### 1.3 Wahrer Wert

Die Begriffsbildung *wahrer Wert* ist problematisch, weil sie davon ausgeht, dass es diesen Wert tatsächlich gibt. Aber schon das einfache Beispiel der Abmessungen eines mehr oder weniger unregelmäßigen realen Rundstabes lehrt uns, dass dies streng genommen unzulässig ist. Nur in der *mathematischen Idealisierung* etwa als Zylinder kann dem Stab eine bestimmte Länge zugeordnet werden. Auf immer kleineren Skalen entspricht ein realer Stab aber immer weniger seinem Ideal. Wenn dennoch von einer „wahren Größe“ ausgegangen wird, so wird darunter eine über die geometrischen Unregelmäßigkeiten gemittelte Größe verstanden. Wiederholte Längenmessungen an einem Stab etwa mit einer Schieblehre, die immer wieder erneut angesetzt wird, liefern im Allgemeinen unterschiedliche Messergebnisse, sofern die Präzision der Schieblehre ausreicht, die Unregelmäßigkeiten des Stabes aufzulösen.

Ein wahrer Wert ist somit als *ideeller Wert* anzusehen, der in der Regel nicht bekannt ist. Ausnahmen ergeben sich beispielsweise, wenn die Messgröße nur diskrete Werte annehmen kann, wie etwa beim Würfeln.

Abweichungen einer Messgröße vom wahren Wert heißen *Messabweichung*, veraltet auch *Messfehler*.<sup>5)</sup>

### 1.4 Verteilungsfunktion

Um die zufälligen Schwankungen der Messwerte in einer Messreihe zu beschreiben, konstruiert man die *empirische Verteilungsfunktion*

$$P(x) = \frac{N_x}{N}, \quad \text{mit } N_x = \text{Anzahl der Messwerte für die } x_n < x \text{ gilt.} \quad (2)$$

Somit ist  $P(x)$  die relative Häufigkeit, mit der Messwerte kleiner als  $x$  in der Messreihe (1) zu finden sind. Sie hat Werte zwischen 0 und 1. Abbildung 1 (rechts) zeigt zwei Beispiele. Es ist leicht einzusehen, dass  $P(x)$  mit wachsendem  $x$  nicht abfallen kann. An den Werten  $x_n$  der Messreihe hat  $P$  Sprungstellen der Höhe  $1/N$ .

Sind die Messwerte alle verschieden, so können sie eindeutig der Größe nach sortiert werden. Die so geordnete Messreihe sei

$$x_1 < x_2 < \dots < x_n < \dots < x_N. \quad (3)$$

Es gilt dann,

$$P(x) = \begin{cases} 0 & : & x \leq x_1 \\ (n-1)/N & : & x_{n-1} < x \leq x_n, & \text{für } n = 2, 3, \dots, N \\ 1 & : & x_N < x \end{cases}$$

Die Verteilungsfunktion ist eine sog. Treppenfunktion. Sie hat Werte zwischen 0 und 1 und genau  $N$  Treppenstufen gleicher Höhe  $P(x_n) - P(x_{n-1}) = 1/N$ . Diese Stufen werden also bei längeren Messreihen kleiner, vgl. Abb. 1 (oben und unten). Unter der Voraussetzung, dass alle Messwerte verschieden sind, ergibt sich für  $N \rightarrow +\infty$  eine stetige Funktion. Auch wenn dieser Grenzübergang praktisch nicht vollzogen werden kann, können doch niemals unendlich viele Messungen aufgenommen werden, so ist dies doch ein sinnvolles Konstrukt, um schließlich zu einer Theorie der Messfehler zu gelangen.

---

<sup>5)</sup> Das *Deutsche Institut für Normierung* (DIN) legt mit der „DIN 1319“ Normen für die Messtechnik fest. Die Schrift besteht aus 4 Teilen. Der erste Teil DIN 1319-1 hat den Titel „Grundbegriffe der Messtechnik: Messen, Zählen, Prüfen“ und wurde zuletzt 1995 aktualisiert.

---

## 1.5 Verteilungsdichte

Sei  $P(x)$  eine Verteilungsfunktion für den Grenzfall unendlich vieler Messwerte, also für  $N \rightarrow +\infty$ . Darüber hinaus sei sie differenzierbar. Dann heißt die Ableitung

$$p(x) = \frac{dP(x)}{dx}$$

*Verteilungsdichte* oder auch *Wahrscheinlichkeitsdichte*. Umgekehrt kann aus der Dichte die Verteilungsfunktion berechnet werden,

$$P(x) = \int_{-\infty}^x p(z) dz .$$

Für die folgenden Betrachtungen sind drei Kennzahlen der Verteilungsfunktion wichtig.

**Mittelwert (Erwartungswert):**

$$\mu \equiv \int_{-\infty}^{+\infty} x p(x) dx . \quad (4)$$

**Streuung (Varianz):**

$$\sigma^2 \equiv \int_{-\infty}^{+\infty} (x - \mu)^2 p(x) dx . \quad (5)$$

**Standardabweichung:**

$$\sigma \equiv +\sqrt{\sigma^2} . \quad (6)$$

## 1.6 Normalverteilung

Abbildung 2 zeigt die Dichte der sogenannten *Normalverteilung* oder auch *Gaussverteilung*,

$$p_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (7)$$

Wegen ihrer markanten Form wird sie auch *Glockenkurve* genannt. Die Parameter  $\mu$  und  $\sigma$  sind der Mittelwert bzw. die Standardabweichung. Die Normalverteilung spielt bei der Behandlung von Messabweichungen wie allgemein in der Wahrscheinlichkeitsrechnung bzw. Statistik eine herausragende Rolle, was auf den *Zentralen Grenzwertsatz* der Wahrscheinlichkeitsrechnung zurückgeht,

*Immer dann, wenn die zufällige Messabweichung einer Einzelmessung als additive Überlagerung einer Vielzahl von zufälligen, untereinander unabhängigen und in ihrer Variabilität (Varianz) beschränkten Größen angesehen werden kann, ist die resultierende Abweichung annähernd normalverteilt. Die Näherung gilt umso besser, je mehr zufällige Einzelgrößen einwirken.*

Diese Aussage ist umso bemerkenswerter, als die Verteilung der einzelnen Einflüsse oder gar ihr physikalischer Ursprung nicht bekannt sein müssen.<sup>6)</sup> Folglich wird bei der Behandlung von Messdaten häufig davon ausgegangen, dass sie normalverteilt sind. Die Beispiele in Abbildung 1 (links) sind normalverteilte Zufallszahlen aus einer Computersimulation mit dem Mittelwert  $\mu = 0$  und der Varianz  $\sigma^2 = 1$ . Die entsprechende Dichte  $p_{0,1}$  heißt *standardnormalverteilt*.<sup>7)</sup>

---

<sup>6)</sup> Im Zusammenhang mit der Behandlung von Messfehlern wurde die Normalverteilung von den Mathematikern ADRIEN MARIE LEGENDRE (1752–1833) und CARL FRIEDRICH GAUSS (1777–1855) eingeführt.

<sup>7)</sup> Ist eine Zufallsgröße  $X$  normalverteilt mit der Dichte  $p_{\mu,\sigma}$ , so schreibt man hierfür  $X \sim \mathcal{N}(\mu, \sigma)$ . Folglich bedeutet  $X \sim \mathcal{N}(0, 1)$ , dass  $X$  standardnormalverteilt ist.

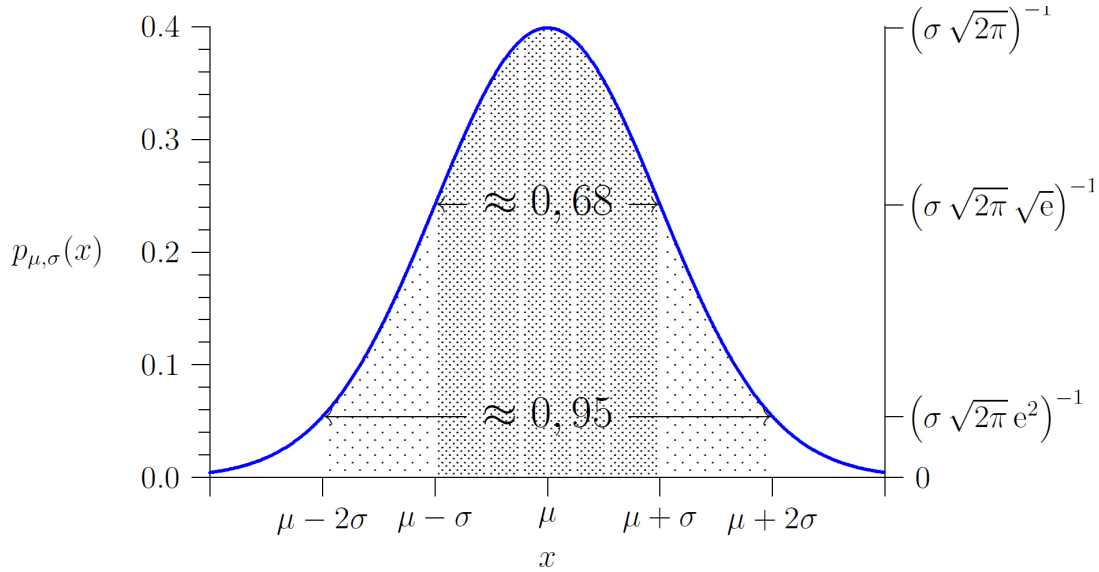


Abb. 2: Dichte der Normalverteilung mit Mittelwert  $\mu$  und Standardabweichung  $\sigma$ , Gl. (7)

## 1.7 Quantile und Boxplot

Eine Verteilungsfunktion  $P(x)$  die streng monoton wächst, hat eine eindeutige Umkehrfunktion  $P^{-1}$ . Folglich kann zu einer vorgegebenen Wahrscheinlichkeit  $\alpha$ ,  $0 < \alpha < 1$ , ein Wert  $\tilde{x}_\alpha = P^{-1}(\alpha)$  angegeben werden, der  $\alpha$ -Quantil genannt wird, Abb. 3. Somit ist  $\alpha$  die Wahrscheinlichkeit, dass die Messgröße (Zufallsgröße) einen Wert kleiner als  $\tilde{x}_\alpha$  annimmt.  $\alpha$  ergibt sich auch aus der Fläche unter der Wahrscheinlichkeitsdichte  $p(x)$  in den Grenzen von  $-\infty$  und  $\tilde{x}_\alpha$ ,

$$\alpha = \int_{-\infty}^{\tilde{x}_\alpha} p(x) dx .$$

Für endliche Messreihen wächst die Verteilungsfunktion nicht kontinuierlich, sondern ist eine Treppenfunktion wie in den Beispielen von Abb. 1. Quantile einer endlichen Messreihe (1) werden deshalb anders definiert. Seien die Messwerte schon der Größe nach sortiert, wobei hier allerdings auch die Gleichheit von Messwerten zugelassen wird,

$$x_1 \leq x_2 \leq \dots \leq x_n \leq \dots \leq x_N . \quad (8)$$

Für eine beliebige reelle Zahl  $\alpha$ , mit  $0 < \alpha < 1$ , ist das  $\alpha$ -Quantil definiert durch

$$\tilde{x}_\alpha \equiv \begin{cases} \frac{1}{2} (x_{\alpha N} + x_{\alpha N + 1}) & : \alpha N \text{ ganzzahlig} \\ x_{\lceil \alpha N \rceil} & : \text{andernfalls} \end{cases} \quad (9)$$

Darin bezeichnet  $\lceil \dots \rceil$  die Aufrundungsfunktion (Gauss-Klammer),  $\lceil \alpha N \rceil \equiv \min\{k \in \mathbb{Z} \mid k \geq \alpha N\}$ . Beispielsweise gilt  $\lceil 4,3 \rceil = \lceil 4,7 \rceil = \lceil 5,0 \rceil = 5$ .

Zumindest der relative Anteil  $\alpha$  aller  $N$  Messwerte hat einen Wert  $\leq \tilde{x}_\alpha$ . Einige Quantile haben besondere Namen. Der *Median* ist identisch mit dem 0,5-Quantil  $\tilde{x}_{0,5}$ . Weitere spezielle Quantile sind in der Tab. 1 bezeichnet.

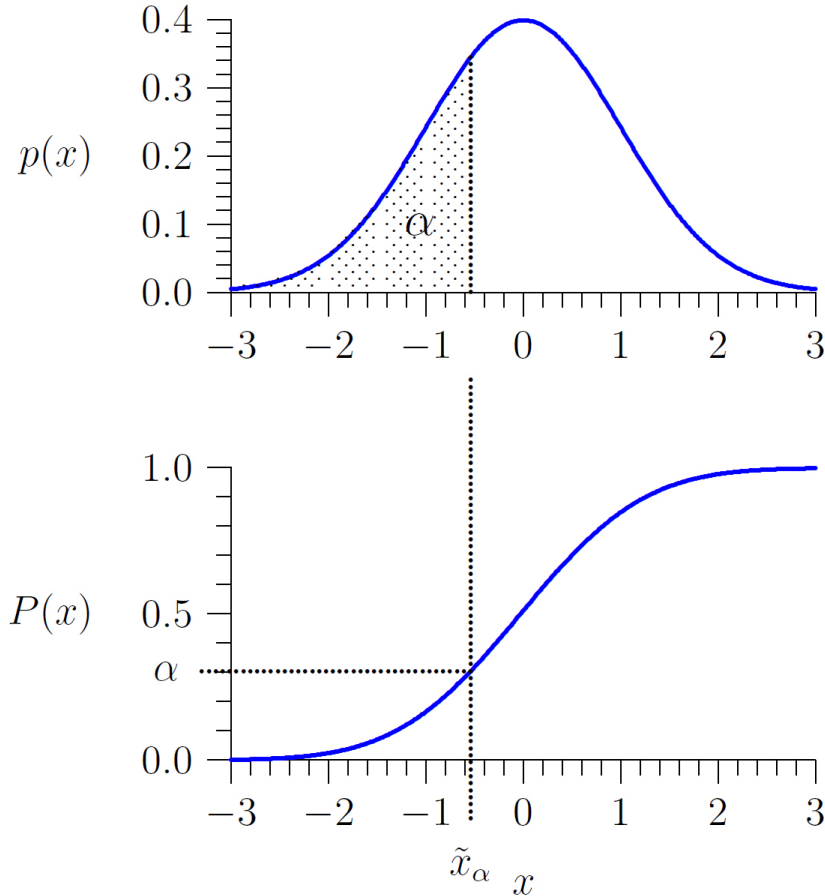
Beispielsweise hat die Messreihe

$$\{x_n\}_{n=1}^8 = \{2, 4, 7, 8, 12, 14, 19, 32\} \quad (10)$$



**Tab. 1: Bezeichnungen für spezielle  $\alpha$ -Quantile.**

$\alpha$	Quantil-Bezeichnung
1/2	Median
1/3; 2/3	Terzil
1/4; 2/4; 3/4	Quartil
1/5; 2/5; 3/5; 4/5	Quintil
1/10; 2/10; ... 9/10	Dezil
1/100; 2/100; ... 99/100	Perzentil



**Abb. 3:** An der Stelle des  $\alpha$ -Quantils  $\tilde{x}_\alpha$  hat die kontinuierliche Verteilungsfunktion  $P$  den Wert  $\alpha$ ,  $P(\tilde{x}_\alpha) = \alpha$ . Dies ist zugleich die Fläche unter der Wahrscheinlichkeitsdichte  $p(x)$  in den Grenzen  $-\infty$  und  $\tilde{x}_\alpha$ .

den Median  $\tilde{x}_{0,5} = (x_4 + x_5)/2 = 10$ . Würde der größte Wert  $x_8$  fehlen, so erhielte man den Median  $\tilde{x}_{0,5} = x_{[0,5,7]} = x_4 = 8$ .

Mit einem sogenannten *Boxplot* werden wesentliche Eigenschaften der Verteilungsfunktion veranschaulicht, Abb. 4. Unter *Box* versteht man das Quartilintervall  $[\tilde{x}_{0,25}, \tilde{x}_{0,75}]$ . Ein Boxplot setzt voraus, dass zumindest  $N = 5$  Daten vorliegen. Im Fall, dass  $N = 5$  gilt und die Daten schon geordnet vorliegen, also für  $x_1 \leq x_2 \leq x_3 \leq x_4 \leq x_5$ , gelten,

$$x_{\min} = x_1 \quad , \quad \tilde{x}_{0,25} = x_2 \quad , \quad \tilde{x}_{0,50} = x_3 \quad , \quad \tilde{x}_{0,75} = x_4 \quad , \quad x_{\max} = x_5 \quad .$$

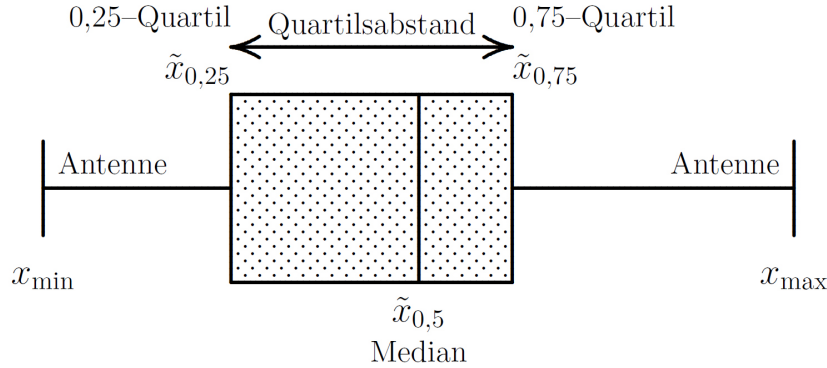


Abb. 4: Definition eines Boxplots

## 2 Analyse von Messabweichungen

### 2.1 Grobe Messabweichung (Ausreißer)

Werte der Messreihe, die um mehr als das 1,5-fache des *Quartilsabstands*  $\tilde{x}_{0,75} - \tilde{x}_{0,25}$  vom unteren oder oberen Quartil entfernt sind, heißen *Ausreißer*. Für einen Ausreißer  $x^*$  gilt also

$$\begin{aligned}
 x^* &< \tilde{x}_{0,25} - 1,5 \cdot (\tilde{x}_{0,75} - \tilde{x}_{0,25}) = 2,5 \cdot \tilde{x}_{0,25} - 1,5 \cdot \tilde{x}_{0,75} \\
 &\text{oder} \\
 x^* &> \tilde{x}_{0,75} + 1,5 \cdot (\tilde{x}_{0,75} - \tilde{x}_{0,25}) = 2,5 \cdot \tilde{x}_{0,75} - 1,5 \cdot \tilde{x}_{0,25}
 \end{aligned}
 \tag{11}$$

Bei Verwendung des Boxplots, Abb. 4, erkennt man einen Ausreißer daran, dass eine Antenne länger als das 1,5-Fache der Boxbreite ist. Hat man in einem Datensatz einen Ausreißer, so wird dieser eliminiert und erneut ein Boxplot für die verbliebenen Daten gebildet. Dieses Prozedere wird fortgesetzt, bis schließlich kein Ausreißer mehr vorhanden ist oder die Datenanzahl auf die minimal notwendige Anzahl  $N = 5$  reduziert ist.

Beipielsweise erhält man für die Messreihe (10) die Quartile

$$\tilde{x}_{0,25} = x_{[0,25 \cdot 8]} = x_2 = 4 \quad \text{und} \quad \tilde{x}_{0,75} = x_{[0,75 \cdot 8]} = x_6 = 14 .$$

Die 1,5-fache Boxgröße ist somit  $1,5 \cdot (x_6 - x_2) = 15$ . Der kleinste Wert  $x_1 = 2$  liegt dichter als 15 am unteren Quartil  $\tilde{x}_{0,25} = 4$  und wird somit nicht als Ausreißer betrachtet. Der größte Wert  $x_8 = 32$  liegt weiter als 15 vom oberen Quartil  $\tilde{x}_{0,75} = 14$  entfernt und ist somit ein Ausreißer. Für die weitere Datenanalyse ist er zu eliminieren.

Die Grenzen für die Festlegung als Ausreißer ist recht subjektiv. Deshalb sollte in der Laborpraxis die Entscheidung, einen Messwert wegzulassen, gut begründet sein. Dies trifft vor allem für Messwerte zu, die nur mit großem Aufwand gewonnen wurden, so dass die Erhebung weiterer Messwerte nicht einfach zu bewerkstelligen ist. Eine Sichtung der Messreihe auf Ausreißer ist jedoch notwendig, weil in der Messpraxis aus den verschiedensten Gründen grobe Fehler auftreten können, etwa bei einer unachtsamen Zerstörung der Kalibrierung eines Messgerätes.

Aber ganz so willkürlich, wie es zunächst scheint, ist die obige Festlegung eines Ausreißers doch nicht. Würden die Messwerte normalverteilt sein, was durch den zentralen Grenzwertsatz häufig naheliegt, dann würde bei vielen Messwerten ( $N \gtrsim 30$ ) ein einzelner Messwert mit einer Wahrscheinlichkeit von weniger als 0,01 irrtümlich als Ausreißer klassifiziert. Dies zeigt man durch Integration über die Glockenkurve über den durch (11) festgelegten Bereich.<sup>8)</sup>

<sup>8)</sup> Für die Standardnormalverteilung (Varianz 1 und Mittelwert 0) gelten  $\tilde{x}_{0,25} = -0,675(5)$  und  $\tilde{x}_{0,75} = 0,675(5)$ .

---

## 2.2 Mittelwert

Im Folgenden wird aufgezeigt, wie man einen Wert  $\bar{x}$  bestimmt, der eine möglichst gute Schätzung von  $x_w$  ist, und es wird eine Angabe gemacht, mit welcher Wahrscheinlichkeit sich  $\bar{x}$  von  $x_w$  um weniger als  $\Delta x$  unterscheidet. Die Vorgehensweise wird in der Wahrscheinlichkeitsrechnung bzw. Statistik tiefer begründet. Hier werden nur die wichtigsten Ergebnisse zusammengefasst, die gewissermaßen als „Kochrezept“ für den Experimentator anzusehen sind.

Der *empirische Mittelwert* einer Messreihe (1) ist wie folgt definiert,

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n \quad (12)$$

Man kann zeigen, dass  $\bar{x}$  in einem wohl definiertem Sinn die „beste“ Schätzung für den unbekanntem wahren Wert  $x_w$  ist, denn er ist *erwartungstreu* und *konsistent*. Darunter versteht man Folgendes:

**Erwartungstreu:** Würde man den Messaufwand deutlich vergrößern, indem man wiederholt Messreihen der Länge  $N$  ermittelt, z. B.  $M$ -mal, so würde man im Allgemeinen nicht für eine jede Messreihe den gleichen empirischen Mittelwert erhalten. Die entsprechenden Mittelwerte

$$\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m, \dots, \bar{x}_M \quad (13)$$

schwanken folglich. Man betrachtet nun von diesen empirischen Mittelwerten  $\bar{x}_m$  wiederum den Mittelwert,

$$\langle \bar{x} \rangle \equiv \lim_{M \rightarrow +\infty} \frac{1}{M} \sum_{m=1}^M \bar{x}_m .$$

Wenn  $\langle \bar{x} \rangle$  gleich dem wahren Mittelwert  $\mu$  der Messreihe ist, dann nennt man den Schätzer (12) *erwartungstreu*.

**Konsistenz:** Würde man den Messaufwand wiederum deutlich vergrößern, nun aber aus allen  $N$  Daten  $\bar{x}$  berechnen, so hofft man, dass im Grenzfalle der wahre Mittelwert der Messreihe  $\mu$  erhalten wird, dass also

$$\lim_{N \rightarrow +\infty} \bar{x} = \mu$$

gilt. Ist dies der Fall, so nennt man den Schätzer *konsistent*.

In der Messpraxis stehen jedoch immer nur endlich viele Messdaten zur Verfügung, aus welchen nur der empirische Mittelwert  $\bar{x}$  berechnet werden kann. Somit steht die Frage, wie dicht  $\bar{x}$  dem unbekanntem Mittelwert  $\mu$  kommt. Weiter unten wird hierauf eine Antwort gegeben.

In der Abb. 1 sind die empirischen Mittelwerte für die beiden Messreihen angegeben. Wie erwartet liegen sie dicht beim wahren Mittelwert der Messreihe  $\mu = 0$ . Dieser ist hier bekannt, weil die Messreihen aus einer Computersimulation stammen. Weiter unten wird genauer ausgeführt, dass bei einer längeren Messreihe die Wahrscheinlichkeit dafür steigt, dass  $\bar{x}$  dichter bei  $\mu$  liegt.

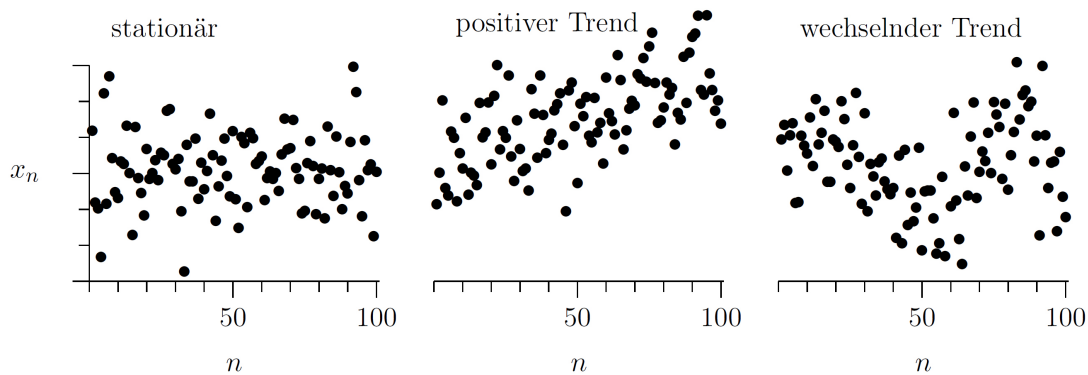
## 2.3 Stationarität

Bei der obigen Schätzung von Mittelwerten aus Messreihen wie auch bei der Schätzung anderer Kenngrößen (s.u.) muss *Stationarität* vorausgesetzt werden. Dies bedeutet, dass alle  $N$  Messwerte unter gleichen

Folglich sind  $2,5 \cdot \tilde{x}_{0,25} - 1,5 \cdot \tilde{x}_{0,75} \approx -2,7$  und  $2,5 \cdot \tilde{x}_{0,75} - 1,5 \cdot \tilde{x}_{0,25} \approx 2,7$ . Die Irrtumswahrscheinlichkeit liegt also bei

$$\frac{1}{\sqrt{2\pi}} \int_{|x|>2,7} e^{-x^2/2} dx \approx 0,007 .$$

Messbedingungen gewonnen wurden. Die „Kunst des Messens“ besteht u.a. darin, Messbedingungen wenigstens näherungsweise unverändert zu belassen. Beispiele für offenbar nichtstationäre Messreihen zeigt Abb. 5. Aber nicht immer sind Instationaritäten so leicht zu erkennen, wie in diesem Bild. Es gibt jedoch statistische Verfahren, welche die Stationarität testen, was hier jedoch nicht weiter ausgeführt wird.



**Abb. 5:** Beispiele für eine stationäre Messreihe (links) im Unterschied zu instationären Daten mit positivem Trend (mittel) und Trendschwankungen (rechts). Für die Fehleranalyse sind stationäre Messreihen erforderlich. Darüber hinaus müssen die Messwerte voneinander unabhängig sein.

## 2.4 Zufällige Messabweichungen

Unter einer *zufälligen Messabweichung* (veraltet auch *zufälliger Fehler*) versteht man die Abweichung eines Messwertes vom wahren Wert, welche von Messung zu Messung unterschiedlich groß ist und auch im Vorzeichen schwankt. Auch bei Wiederholung der Messung unter vermeintlich gleichen Bedingungen streuen die Messwerte infolge von Zufallseinflüssen.

Der Experimentator kann zufällige Messabweichungen in der Regel nicht verhindern. Zuweilen kann er jedoch ihre Stärke durch Veränderung des Messvorgangs verringern. Insbesondere wird ihr Einfluss auf den nach (12) berechneten Mittelwert bei wachsender Anzahl  $N$  von Messwerten immer geringer.

Für endlich viele Messwerte kann man unter recht allgemeinen Annahmen Wahrscheinlichkeiten angeben, mit denen ein Messwert oder auch ein empirischer Mittelwert in einem bestimmten Intervall um den wahren Wert liegt, sofern die Messabweichung nur zufällig sind (s.u.). In diesem Sinne machen zufällige Messabweichungen das Messergebnis unsicher, aber nicht „falsch“.

## 2.5 Systematische Messabweichung

Ist die Erwartungstreue des Mittelwertschätzers (12) gegeben, so würde der Mittelwert der Messreihe  $\mu$  im Allgemeinen immer noch vom wahren Wert  $x_w$  abweichen. Diese Abweichung nennt man *systematischer Fehler* der Messung,  $\Delta_{\text{sys}}$ . Den wahren Wert erhält man aus

$$x_w = \mu - \Delta_{\text{sys}} .$$

Systematische Fehler können verschiedene Ursachen haben, beispielsweise Fehleichungen des Messinstrumentes. Sie können durch das Messverfahren begründet sein. Im Unterschied zu zufälligen Fehlern haben sie immer dasselbe Vorzeichen, d.h. sie machen alle Messwerte systematisch um den gleichen Wert zu groß ( $\Delta_{\text{sys}} > 0$ ) oder zu klein ( $\Delta_{\text{sys}} < 0$ ). Folglich kann ein systematischer Fehler durch wiederholte Messungen und Mittelwertbildung nicht verringert werden. In diesem Sinne machen systematische Messabweichungen das Messergebnis streng genommen immer „falsch“.

Zuweilen kann man jedoch einen Teil der systematischen Gesamtabweichung berücksichtigen, z. B. durch *Nullpunktkorrektur* eines Messgerätes. Häufig kann man bestimmte systematische Fehler zwar prinzipiell verringern, ist aber der nötige Aufwand unangemessen groß, verzichtet man darauf. Praktisch können systematische Fehler niemals ganz vermieden werden, da es keine vollkommenen Messgeräte gibt.

Solide Hersteller geben die Präzision ihrer Geräte an, indem sie die größt mögliche systematische Messabweichung abschätzen. Eine Messung kann nie genauer sein, als diese Abweichung. Beispielsweise gilt für übliche Schieblehren zur Längenmessung  $|\Delta_{\text{sys}}| \lesssim 0,05 \text{ mm}$ . Besonders hochwertige Instrumente für anspruchsvolle Messaufgaben müssen zur Minimierung systematischer Abweichungen von Zeit zu Zeit neu kalibriert werden.

## 2.6 Standardabweichung der Einzelmessung

Die *Standardabweichung*  $\sigma$  einer Zufallsgröße  $X$  mit der Verteilungsdichte  $p(x)$  wurde bereits in (6), S. 6, eingeführt.  $X$  beschreibt eine einzelne Messung und die Kenngröße  $\sigma$  ist ein Maß für die mittleren Abweichungen der Werte einer Messreihe (1) vom Mittelwert  $\mu$ .

In der Praxis kennt man jedoch im Allgemeinen  $\sigma$  nicht. Man möchte nun aus einer (bekannten) Messreihe (1)  $\sigma$  schätzen. Man kann allgemein zeigen, dass die beste Schätzung für  $\sigma$  die sogenannte *empirische Standardabweichung* ist,

$$s_x = + \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2} \quad (14)$$

Darin ist  $\bar{x}$  der empirische Mittelwert (12), der auch aus derselben Messreihe berechnet wird.

Auf den ersten Blick mag es verwundern, dass hier als Vorfaktor der Summe nicht  $1/N$  sondern  $1/(N-1)$  steht. Aber nur so ist  $s_x$  ein erwartungstreuer Schätzer für  $\sigma$ . Würde allerdings in (14) anstelle des empirischen Mittelwertes  $\bar{x}$  der (unbekannte) wahre Mittelwert  $\mu$  verwendet, dann würde der Vorfaktor  $1/N$  einen erwartungstreuen Schätzer liefern. Bei langen Messreihen  $N > 30$  ist der Unterschied jedoch zumeist vernachlässigbar.

In der Abb. 1 sind die empirischen Standardabweichungen für die beiden Messreihen angegeben. Wie erwartet liegen sie dicht beim wahren Wert  $\sigma = 1$ , der hier bekannt ist, weil die Messreihen aus einer Computersimulation stammen.

Für die Aussage, dass  $s_x$  ein erwartungstreuer und konsistenter Schätzer von  $\sigma$  ist, müssen all jene Voraussetzungen gemacht werden, wie schon bei der Mittelwertschätzung (S. 10). Insbesondere muss die Messreihe wiederum stationär sein. Hier ist jedoch eine weitere Annahme zu machen:

**Unkorreliertheit:** Die Daten der Messreihe müssen *unkorreliert* sein. Fasst man den Messwert als  $x_n$ ,  $n = 1, 2, \dots, N$ , als Realisierung einer Zufallsgröße  $X_n$  auf,  $n = 1, 2, \dots, N$ , so muss der Mittelwert des Produktes  $(X_{n_1} - \mu)(X_{n_2} - \mu)$  für alle  $n_1 \neq n_2$  null sein.

Dies ist eine Bedingung, deren Bedeutung nicht leicht einzusehen ist, allerdings durch eine einfache Rechnung gezeigt werden kann. Für die Messpraxis kann eine einfache Regel formuliert werden, die darauf aufbaut, dass die Messdaten *statistisch unabhängig* sein müssen, woraus die Unkorreliertheit folgt. Unter Unabhängigkeit versteht man vereinfachend, dass vorherige Messergebnisse keinen Einfluss auf das aktuelle Ergebnis haben dürfen.<sup>9)</sup> Praktisch bedeutet dies insbesondere, dass man die Einstellungen der

<sup>9)</sup> In der Wahrscheinlichkeitsrechnung wird der Begriff der statistischen Unabhängigkeit präzise gefasst. Sind beispielsweise zwei Zufallsgrößen  $X_1$  und  $X_2$  durch die Wahrscheinlichkeitsdichten  $p_1(x_1)$  bzw.  $p_2(x_2)$  beschrieben, so heißen sie unabhängig, wenn die Verbundwahrscheinlichkeitsdichte  $p_{12}(x_1, x_2)$  des Zufallsvektors  $(X_1, X_2)$  das Produkt der Randdichten

Messapparatur, etwa das Anlegen einer Schieblehre an einen starren Körper, von Messung zu Messung erneut machen muss und dass der Messvorgang das Messobjekt nicht verändert, also immer dieselben Messbedingungen herrschen. Beispiele für Abhängigkeit sind die instationären Messreihen in Abb. 5 (mittel und rechts).<sup>10)</sup>

## 2.7 Normierung

Die Messwerte  $x_n$  einer Messreihe (1) können wie folgt normiert werden,

$$x_n \rightarrow y_n \equiv \frac{x_n - \bar{x}}{s_x} , \quad (15)$$

Darin sind  $\bar{x}$  der empirische Mittelwert (12) und  $s_x$  die empirische Standardabweichung (14). Die so normierte Messreihe

$$\{y_1, y_2, \dots, y_n, \dots, y_N\} \quad (16)$$

hat den empirischen Mittelwert  $\bar{y} = 0$  und die empirische Standardabweichung  $s_y = 1$ .

Hat eine Zufallsgröße  $X$  dem Mittelwert  $\mu_X$  und die Standardabweichung  $\sigma_X > 0$ , so hat die abgeleitete Größe

$$Y = \frac{X - \mu_X}{\sigma_X}$$

den Mittelwert  $\mu_Y = 0$  und die Standardabweichung  $\sigma_Y = 1$ .

## 2.8 Vertrauensbereich

Wäre die Standardabweichung  $\sigma$  einer Einzelmessung bekannt, so könnte man leicht die Standardabweichung des Mittelwertschätzers  $\bar{x}$  angeben,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} , \quad (17)$$

was aus einer einfachen Rechnung folgt, hier aber nicht weiter ausgeführt wird.

In der Messpraxis werden anstelle der unbekanntenen Werte  $x_w$  und  $\sigma$  die entsprechenden empirischen Werte (Schätzwerte)  $\bar{x}$  bzw.  $s_x$  verwendet. Dies führt auf die *Standardabweichung des (empirischen) Mittelwertes*, die auch kurz *Vertrauensbereich* genannt wird,

$$s_{\bar{x}} = \frac{s_x}{\sqrt{N}} = + \sqrt{\frac{1}{N(N-1)} \sum_{n=1}^N (x_n - \bar{x})^2} \quad (18)$$

Der Vertrauensbereich fällt mit zunehmender Datenanzahl  $N$ , allerdings nur mit der Wurzel. Wollte man ihn halbieren, müsste man gleich 4-mal so viele Messungen ausführen, beispielsweise  $N = 40$  statt  $N = 10$ .

Die Bedeutung des Vertrauensbereichs liegt darin, dass man mit ihm Intervalle angeben kann, in denen der wahre Wert  $x_w$  mit einer bestimmten Wahrscheinlichkeit liegt, was nun genauer ausgeführt wird.

ten ist,  $p_{12}(x_1, x_2) = p_1(x_1) \cdot p_2(x_2)$ . Aus der statistischen Unabhängigkeit folgt die Unkorreliertheit. Die Umkehrung gilt jedoch im Allgemeinen nicht, wohl aber bei im Verbund normalverteilten Zufallsgrößen.

<sup>10)</sup> In der statistischen Theorie wird oftmals eine Messreihe vorausgesetzt, die keinerlei statistische Abhängigkeiten aufweist. Man spricht dann von einer *mathematischen Stichprobe*.

---

## 2.9 $t$ -Verteilung

Gegeben sei eine Messreihe (1) von  $N > 3$  normalverteilten Daten. Der unbekannte wahre Wert sei  $\mu$ , der durch den empirischen Mittelwert  $\bar{x}$ , Gl. (12), geschätzt wird, mit  $s_{\bar{x}}$  als Vertrauensbereich (18). Es wird nun die folgende normierte Größe betrachtet, der sogenannte  $t$ -Wert,

$$t = \frac{\mu - \bar{x}}{s_{\bar{x}}} . \quad (19)$$

Der  $t$ -Wert verändert sich zufällig, abhängig von der konkreten Stichprobe. Er ist also eine Zufallsgröße, für die man in der Wahrscheinlichkeitstheorie die sogenannte  $t$ -Verteilungsdichte zum Freiheitsgrad  $N - 1$  berechnet, <sup>11) 12)</sup>

$$p_{N-1}(t) = \frac{1}{\sqrt{\pi(N-1)}} \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)} \left(1 + \frac{t^2}{N-1}\right)^{-N/2} . \quad (20)$$

Darin bezeichnet  $\Gamma$  die Gamma-Funktion. Für jede ganze Zahl  $n > 0$  gilt,

$$\Gamma(n+1) = n! , \quad \Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{n! 4^n} \sqrt{\pi} .$$

Die  $t$ -Verteilung  $p_{N-1}(t)$  und ihre Wahrscheinlichkeitsdichte  $p_{N-1}(t)$  sind in Abb. 6 für einige ausgewählte Freiheitsgrade dargestellt. Konkrete Zahlenwerte für einige  $\alpha$ -Quantile finden sich in Tab. 2, S. 17. Die Dichten ähneln der Gaußschen Glockenkurve (vgl. Abb. 2, S.7). Insbesondere sind die Dichten auch symmetrisch um den Mittelwert  $\mu_{\Gamma} = 0$ , somit gilt also  $p_{N-1}(t) = p_{N-1}(-t)$ . Für die Standardabweichung der  $t$ -Verteilung mit dem Freiheitsgrad  $N - 1 > 2$  erhält man,

$$\sigma_{\Gamma, N-1} = \sqrt{\frac{N-1}{N-3}} .$$

Für  $N < \infty$  gilt somit  $\sigma_{\Gamma} > 1$  und im Grenzfall  $N = +\infty$  folgt  $\sigma_{\Gamma} = 1$ . Für niedrige Freiheitsgrade fällt die Dichte der  $t$ -Verteilung für kleine bzw. große  $t$ -Werte deutlich langsamer als die Normalverteilung ab. Für viele Messwerte  $N$  geht jedoch  $p_{N-1}(t)$  in die Dichte der Standardnormalverteilung über,

$$\lim_{N \rightarrow \infty} p_{N-1}(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} .$$

Für  $N - 1 = 100$  (blaue Kurve) sind die Unterschiede zur Normalverteilung im Rahmen der Zeichengenauigkeit nicht wahrnehmbar. Häufig wird in der Praxis bereits ab  $N \approx 30$  (rote Kurve) mit der Normalverteilung gerechnet.

Somit kann man folgende Aussage machen:

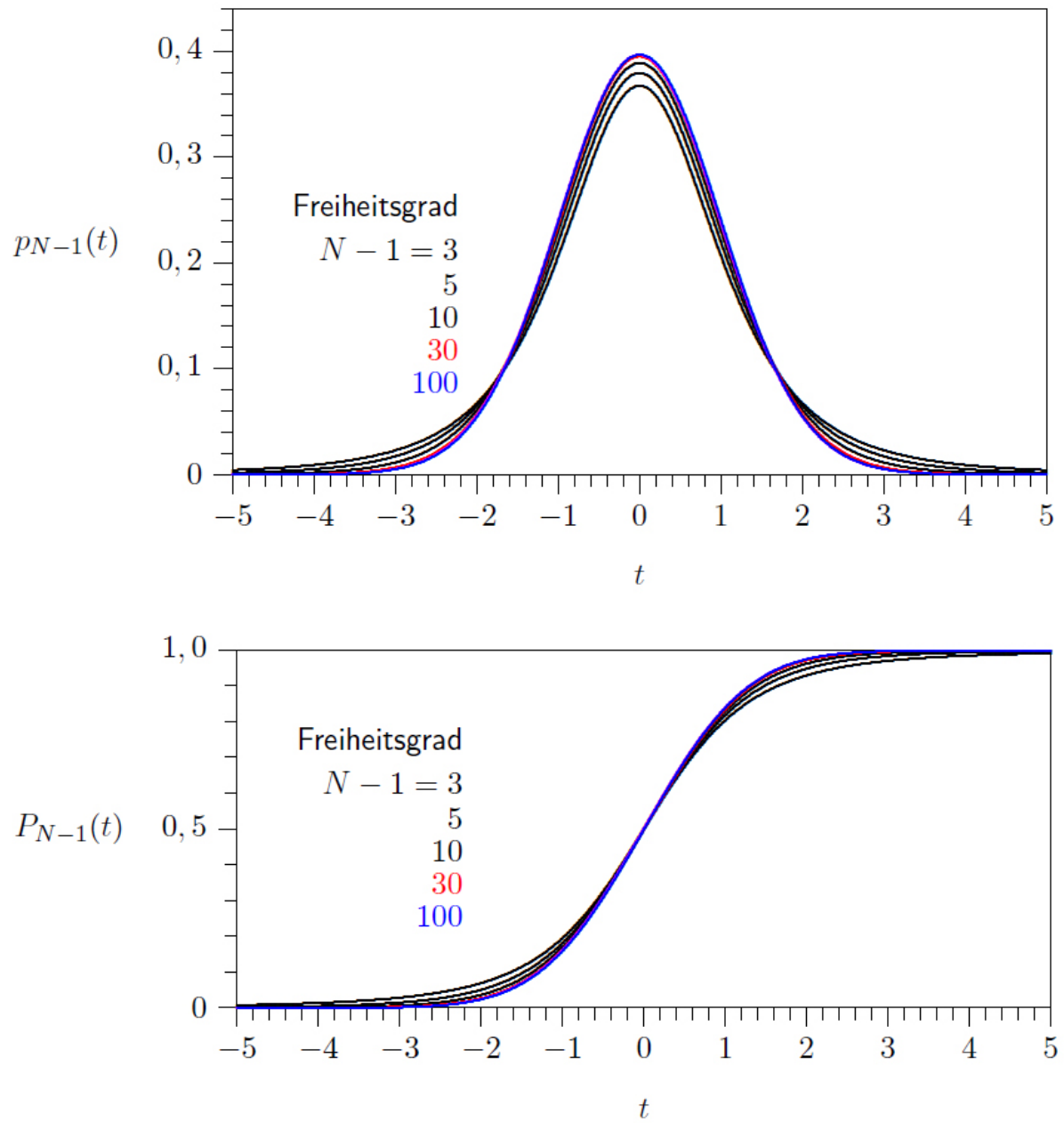
*Der empirische Mittelwert  $\bar{x}$  von  $N$  normalverteilten Messwerten liegt mit der Wahrscheinlichkeit  $1 - 2\alpha$  im Intervall*

$$\left[ \mu - |\tilde{t}_{\alpha}| \cdot s_{\bar{x}} , \quad \mu + |\tilde{t}_{\alpha}| \cdot s_{\bar{x}} \right] .$$

---

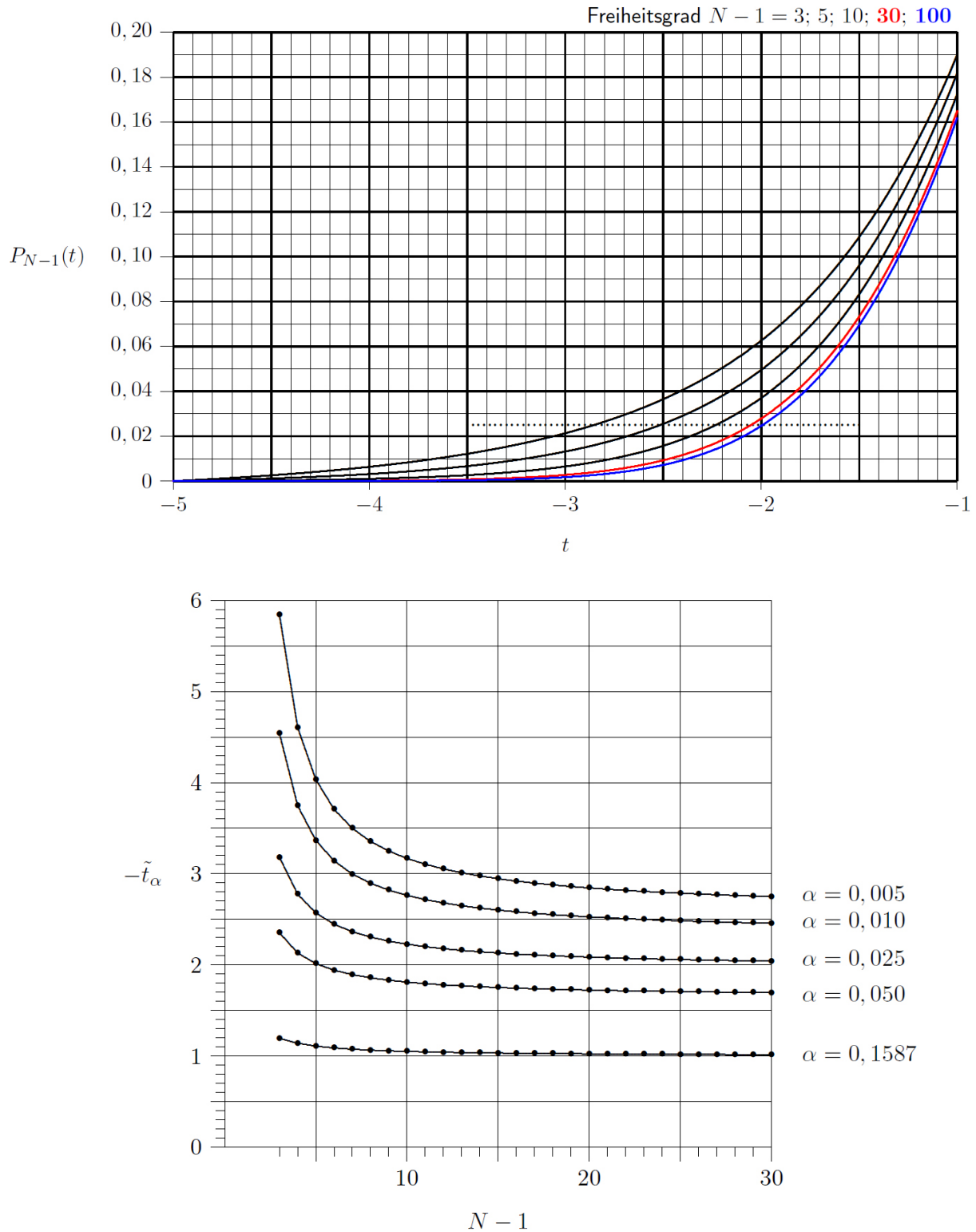
<sup>11)</sup> Die  $t$ -Verteilung wird auch *Student-Verteilung* genannt. Der englische Statistiker W. S. GOSSET hat sie im Jahre 1901 als erster auf empirischem Weg gefunden und unter dem Pseudonym „Student“ publiziert.

<sup>12)</sup> Es sei angemerkt, dass die Größe  $(\mu - \bar{x})$  für eine beliebige Anzahl  $N$  von Messwerten normalverteilt ist, sofern die einzelnen Messwerte normalverteilt sind. Der  $t$ -Wert  $(\mu - \bar{x})/s_{\bar{x}}$  ist jedoch nicht normalverteilt, weil der Divisor  $s_{\bar{x}}$  keine feste Größe ist, da er aus der selben Messreihe berechnet wird wie  $\bar{x}$ . Beide Größen  $\bar{x}$  und  $s_{\bar{x}}$  sind also zufälligen Schwankungen unterworfen, die nicht unabhängig voneinander sind. Dies führt letztlich dazu, dass der  $t$ -Wert nicht normal- sondern  $t$ -verteilt ist.



**Abb. 6:** Wahrscheinlichkeitsdichten  $p_{N-1}$  (oben) und entsprechende Verteilungsfunktionen  $P_{N-1}$  (unten) der  $t$ -Verteilung für verschiedene Freiheitsgrade  $N - 1$ . Für  $N \gtrsim 30$  (rot und blau) kann  $p_{N-1}$  gut durch die Normalverteilungsdichte  $p_\infty$  genähert werden.





**Abb. 7:** Oben:  $t$ -Verteilung für verschiedene Freiheitsgrade  $N - 1$ . Vergrößerung der Abb. 6 (unten) mit zusätzlichen Werten für den Freiheitsgrad  $N - 1$ . Unten:  $\alpha$ -Quantile der  $t$ -Verteilung als Funktion des Freiheitsgrades  $N - 1$  für ausgewählte Werte der Wahrscheinlichkeit  $\alpha$ .

Tab. 2:  $\alpha$ -Quantile  $\tilde{t}_\alpha$  der  $t$ -Verteilung verschiedener Freiheitsgrade  $N - 1$ . (Aufgetragen ist  $-\tilde{t}_\alpha$  mit der maximalen Abweichung  $\pm 0,001$ ).  
Für  $N \rightarrow +\infty$  gilt  $-\tilde{t}_\alpha \rightarrow 1$ , falls  $\alpha = 0,1587\dots$ . Die  $t$ -Verteilung strebt hierbei gegen die Normalverteilung mit der Standardabweichung 1 und dem Mittelwert 0. Im Intervall  $[-1, 1]$  liegt die Zufallsgröße dann mit der Wahrscheinlichkeit  $1 - 2\alpha \approx 0,6826$ .

$N - 1$	$\alpha$				
	0,1587	0,0500	0,0250	0,0100	0,0050
	$(1 - 2\alpha) \cdot 100\%$				
	68,26%	90%	95%	98%	99%
3	1,197	2,353	3,182	4,541	5,841
4	1,141	2,132	2,776	3,747	4,604
5	1,110	2,015	2,571	3,365	4,032
6	1,090	1,943	2,447	3,143	3,707
7	1,077	1,895	2,365	2,998	3,499
8	1,066	1,860	2,306	2,896	3,355
9	1,059	1,833	2,262	2,821	3,250
10	1,052	1,812	2,228	2,764	3,169
11	1,047	1,796	2,201	2,718	3,106
12	1,043	1,782	2,179	2,681	3,055
13	1,040	1,771	2,160	2,650	3,012
14	1,037	1,761	2,145	2,624	2,977
15	1,034	1,753	2,131	2,602	2,947
16	1,032	1,746	2,120	2,583	2,921
17	1,030	1,740	2,110	2,567	2,898
18	1,028	1,734	2,101	2,552	2,878
19	1,027	1,729	2,093	2,539	2,861
20	1,025	1,725	2,086	2,528	2,845
21	1,024	1,721	2,080	2,518	2,831
22	1,023	1,717	2,074	2,508	2,819
23	1,022	1,714	2,069	2,500	2,807
24	1,021	1,711	2,064	2,492	2,797
25	1,020	1,708	2,060	2,485	2,787
26	1,019	1,706	2,056	2,479	2,779
27	1,019	1,703	2,052	2,473	2,771
28	1,018	1,701	2,048	2,467	2,763
29	1,017	1,699	2,045	2,462	2,756
30	1,017	1,697	2,042	2,457	2,750
31	1,016	1,696	2,040	2,453	2,744
35	1,014	1,690	2,030	2,438	2,724
40	1,012	1,684	2,021	2,423	2,704
45	1,011	1,679	2,014	2,412	2,690
50	1,010	1,676	2,009	2,403	2,678
60	1,008	1,671	2,000	2,390	2,660
70	1,007	1,667	1,994	2,381	2,648
80	1,006	1,664	1,990	2,374	2,639
90	1,005	1,662	1,987	2,369	2,632
100	1,005	1,660	1,984	2,364	2,626
200	1,002	1,653	1,972	2,345	2,601
500	1,001	1,648	1,965	2,334	2,586

---

Man schreibt hierfür auch,

$$\mu = \bar{x} \pm |\tilde{t}_\alpha| \cdot s_{\bar{x}} \quad ((1 - 2\alpha) \cdot 100\%) .$$

Man gibt sich beispielsweise die Wahrscheinlichkeit  $2\alpha = 0,05$  vor<sup>13)</sup> und bestimmt dann den Betrag des Quantils  $\tilde{t}_\alpha$ . Dazu werden in entsprechenden Büchern Tafeln angegeben, man kann die Werte aber auch mit hinreichender Genauigkeit aus der Abb. 7 ablesen. Für  $N = 11$  erhält man  $\tilde{t}_\alpha \approx 2,3$  (Kurve für  $\alpha = 0,025$  an der Stelle  $N - 1 = 10$ ). Somit gilt,

$$\mu = \bar{x} \pm 2,3 \cdot s_{\bar{x}} \quad (95\%) .$$

Würde man den empirischen Mittelwert  $\bar{x}$  und den Vertrauensbereich  $s_{\bar{x}}$  aus mehr als  $N = 30$  Messwerten berechnet haben, so könnte man in guter Näherung anstelle der  $t$ -Verteilung die Standardnormalverteilung verwenden. Man erhält dann  $|t_{\alpha=0,025}| \approx 2$  und folglich  $\mu = \bar{x} \pm 2 \cdot s_{\bar{x}} \quad (95\%)$ . Dies entspricht der Tatsache, dass das Integral über die Normalverteilungsdichte in den Grenzen von  $\pm 2\sigma$  den Wert  $0,95 \dots$  liefert (Abb. 2, S. 7).

## 2.10 Gesamte Messabweichung

Um letztlich eine Aussage zur gesamten Messabweichung zu machen, muss auch der systematische Fehler berücksichtigt werden, der sich auch bei vielen Messwerten nicht „herausmittelt“. Man erhält somit folgende Angabe,

$$x_w = \bar{x} \pm (|t_\alpha| \cdot s_{\bar{x}} + \Delta_{\text{sys}}) \quad (> (1 - 2\alpha) \cdot 100\%)$$

(21)

Für sehr viele Messwerte  $N$  kann der Vertrauensbereich  $s_{\bar{x}}$  beliebig klein gemacht werden (s. Gl.(17), S.13). Kann dann der zufällige Fehler  $|t_\alpha|s_{\bar{x}}$  gegenüber dem systematischen  $\Delta_{\text{sys}}$  vernachlässigt werden, gilt also  $|t_\alpha|s_{\bar{x}} \ll \Delta_{\text{sys}}$ , vereinfacht sich die Fehlerangabe zu,

$$x_w = \bar{x} \pm \Delta_{\text{sys}} .$$

Das bedeutet, im Intervall  $[\mu - \Delta_{\text{sys}}, \mu + \Delta_{\text{sys}}]$  liegt der empirische Mittelwert  $\bar{x}$  quasi sicher.

## 2.11 Beispiel

**Physikalische Messgröße:** Länge  $L$

**Messgerät:** Präzision-Schieblehre, Auflösung  $0,01$  mm

**Systematischer Fehler:** Größtfehler des Messinstrumentes,  $\Delta_{\text{sys}} = 0,02$  mm

**Messreihe:** Anzahl Messwerte  $N = 12$

#	1	2	3	4	5	6	7	8	9	10	11	12
$l/\text{mm}$	10,19	9,99	9,90	10,05	10,01	10,12	9,87	9,94	10,00	10,04	9,93	9,23

---

<sup>13)</sup> In diesem physikalischen Praktikum wird grundsätzlich mit der Signifikanz von 95% gearbeitet.

---

**Sortierte Messreihe zur Sichtung grober Fehler (von Ausreißern)**

$n$	1	2	3	4	5	6	7	8	9	10	11	12
$L_n/\text{mm}$	9,23	9,87	9,90	9,93	9,94	9,99	10,00	10,01	10,04	10,05	10,12	10,19

---

Quantile nach Gl. 9, S. 7:

---

kleinster Wert:	$L_{\min} = L_1 = 9,23 \text{ mm}$
unteres Quartil:	$\tilde{l}_{1/4} = (L_{N/4} + L_{1+N/4})/2 = (L_3 + L_4)/2 = 9,915 \text{ mm}$
Median:	$\tilde{l}_{1/2} = (L_{N/2} + L_{1+N/2})/2 = (L_6 + L_7)/2 = 9,995 \text{ mm}$
oberes Quartil:	$\tilde{l}_{3/4} = (L_{3N/4} + L_{1+3N/4})/2 = (L_9 + L_{10})/2 = 10,045 \text{ mm}$
größter Wert:	$L_{\max} = L_{12} = 10,19 \text{ mm}$
Quartilsabstand:	$\Delta\tilde{l} = \tilde{l}_{3/4} - \tilde{l}_{1/4} = 0,13 \text{ mm}$

---

Entscheidung (s. Gl. 11, S. 9):  $L_{\min}$  ist Ausreißer, weil

$$\tilde{l}_{1/4} - L_{\min} = 0,685 \text{ mm} > 1,5 \cdot \Delta\tilde{l} = 0,195 \text{ mm}$$

$L_{\max}$  ist kein Ausreißer, weil

$$L_{\max} - \tilde{l}_{3/4} = 0,145 \text{ mm} < 1,5 \cdot \Delta\tilde{l} = 0,195 \text{ mm}$$

**Weiterarbeit mit korrigierter (reduzierter) Messreihe:** neue Anzahl Messwerte  $N = 11$

$n$	1	2	3	4	5	6	7	8	9	10	11
$L_n/\text{mm}$	9,87	9,90	9,93	9,94	9,99	10,00	10,01	10,04	10,05	10,12	10,19

**Empirischer Mittelwert:** nach Gl. 12, S. 10,  
 $\bar{L}/\text{mm} = 10,00363.. \approx 10,004$

**Empirische Standardabweichung der Einzelmessung** nach Gl. 14, S. 12:  
 $s_L/\text{mm} = 0,09489.. \approx 0,095$

**Vertrauensbereich** nach Gl. 18, S. 13:  
 $s_{\bar{L}}/\text{mm} = 0,02861.. \approx 0,029$

**Festlegung des Signifikanzniveaus**  $2\alpha = 0,05$ , d.h., 95%-tige Sicherheit

$\alpha$ -Quantile der t-Verteilung zum Freiheitsgrad  $N - 1 = 10$ : Ablesen des Quantils  $\tilde{t}_\alpha$  aus Abb. 7, S. 16 (Quantilsbetrag  $|\tilde{t}_\alpha|$  eher nach oben runden),

$$|\tilde{t}_{0,025}| \lesssim 2,3 .$$

**Aussage zum zufälligen Fehler:** Der empirischen Mittelwert  $\bar{L}$  der Messreihe liegt mit einer Wahrscheinlichkeit von zumindest 95% nicht weiter als

$$\tilde{t}_{0,025} \cdot s_{\bar{L}} \approx 2,3 \cdot 0,029 \text{ mm} \approx 0,067 \text{ mm}$$

---

vom (unbekannten) wahren Wert  $\mu_L$  der Messreihe entfernt, oder kürzer,

$$\mu_L = (10,004 \pm 0,067) \text{ mm} \quad (95\%) .$$

Man schreibt hierfür auch eleganter,

$$\mu_L = 10,004(67) \text{ mm} \quad (95\%) .$$

Je nach Bedeutung dieser Längenmessung, könnte man auch stärker runden,

$$\mu_L = 10,00(7) \text{ mm} \quad (95\%) .$$

**Aussage zur wahren Länge:** Der systematische Fehler  $\Delta_{\text{sys}} = 0,02 \text{ mm}$  liegt in der gleichen Größenordnung wie der Vertrauensbereich  $s_{\bar{L}} \approx 0,03 \text{ mm}$ . Deshalb wird auch er zur Gesamtfehleraussage herangezogen. Nach Gl. (21), S. 18 erhält man schließlich die folgende Aussage zur wahren Länge,

$L_w = 10,00(9) \text{ mm} \quad (> 95\%)$	(22)
--	------

**Interpretation:** Der Gesamtfehlers von  $0,09 \text{ mm}$  setzt sich aus dem zufälligen Anteil von etwa  $0,07 \text{ mm}$  und dem systematischen Anteil von  $0,02 \text{ mm}$  zusammen. Es würde sich also lohnen, noch mehr Messwerte in die Rechnung einzubeziehen. Bei 9-mal so vielen Messwerten, also für etwa  $N = 100$  Messwerte, würde z. B. der zufällige Anteil auf etwa  $0,07 \text{ mm}/\sqrt{9} \approx 0,02 \text{ mm}$  fallen, also mit dem systematischen Fehler etwa gleich auf liegen. Erst bei etwa 250-mal so vielen Messwerten würde der zufällige Fehler gegenüber dem systematischen vernachlässigbar sein, also bei  $N \approx 2500$  Messwerten. Dieser hohe Anzahl ist manuell kaum realisierbar, möglicherweise aber mit automatisierten Messplätzen.

## 2.12 Anmerkungen

Die obigen Darlegungen haben gezeigt, dass eine umfassende Analyse der zufälligen Messungenauigkeiten sehr aufwändig sein kann. Mit zunehmender Erfahrung des Experimentators und unter Zuhilfenahme von entsprechenden Computerprogrammen kann der Zeitaufwand jedoch deutlich reduziert werden. Komfortablen Computerprogrammen braucht man nur noch die Messwerte und den systematischen Maximalfehler eingeben.

Aus didaktischen Gründen werden im Anfängerkurs solche Programme zur Analyse von Messungenauigkeiten jedoch nicht bereitgestellt, denn dem ungeübten Experimentator bleiben leicht die Hintergründe der Analyse verborgen, und er kann grobe Fehler bei der Analyse leicht übersehen.

Erkennt man, dass die zufälligen Messabweichungen deutlich kleiner als der systematische Maximalfehler (Präzision) des Messinstrumentes ist, dann kann auf eine Analyse der zufälligen Messungenauigkeiten verzichtet werden. Wie oben schon beschrieben, ist dann die gesamte Messungenauigkeit allein durch diesen systematischen Maximalfehler gegeben.

## 3 Fehlerfortpflanzung

### 3.1 Problemstellung

Häufig interessiert man sich für eine Größe  $y$ , die nicht direkt gemessen werden kann. Kennt man allerdings einen analytischen Zusammenhang zwischen  $y$  und anderen bekannten Größen  $x_1, x_2, \dots, x_n, \dots, x_N$ , die

---

ihrerseits direkt oder auch indirekt gemessen sein können, so rechnet man  $y$  einfach aus,<sup>14)</sup>

$$y = f(x_1, x_2, \dots, x_N) .$$

Sind die Größen  $x_n$  mit einer Messabweichung  $\Delta x_n > 0$  behaftet, so pflanzen sich diese Abweichungen auf  $y$  fort. Folglich steht die Frage, wie groß dann die Messabweichung  $\Delta y$  ist. Diese Frage ist generell nicht einfach zu beantworten, weil dazu im Allgemeinen die Verteilungsfunktion einer Zufallsgröße  $Y$  zu bestimmen ist, deren Realisierung  $y$  ist und die eine Funktion von anderen Zufallsgrößen  $X_1, X_2, \dots, X_N$ , deren Realisierungen die Werte  $x_1, x_2, \dots, x_N$  sind, und deren Verteilung als bekannt anzunehmen ist,

$$Y = f(X_1, X_2, \dots, X_N) .$$

Unter bestimmten Voraussetzungen gibt es hierfür eine Lösung, die zu bestimmten Rechenvorschriften führt, was nun genauer ausgeführt wird.

### 3.2 Statistische Fehlerfortpflanzung

Die Messwerte  $x_1, \dots, x_N$  werden als Realisierung normalverteilter und unkorrelierter<sup>15)</sup> Zufallsgrößen  $X_1, \dots, X_N$  aufgefasst, mit den Mittelwerten und Standardabweichungen

$$\bar{X}_1, \dots, \bar{X}_N \quad \text{bzw.} \quad \Delta X_1, \dots, \Delta X_N .$$

Der berechnete Wert

$$y = f(x_1, \dots, x_n, \dots, x_N)$$

ist dann eine Realisierung der Zufallsgröße

$$Y = f(X_1, \dots, X_n, \dots, X_N) .$$

Ist die Funktion  $f$  stetig nach  $x_n$  (partiell) differenzierbar, so ist für kleine Standardabweichungen  $\Delta X_n$  die Zufallsgröße

$$Y_n = f(\bar{X}_1, \dots, \bar{X}_{n-1}, X_n, \bar{X}_{n+1}, \dots, \bar{X}_N)$$

ebenfalls normalverteilt, mit dem Mittelwert

$$\bar{Y}_n = f(\bar{X}_1, \dots, \bar{X}_{n-1}, \bar{X}_n, \bar{X}_{n+1}, \dots, \bar{X}_N)$$

und der Standardabweichung

$$\Delta Y_n \approx \left| \frac{\partial f}{\partial x_n} \right|_{\bar{X}_1, \dots, \bar{X}_n, \dots, \bar{X}_N} \cdot \Delta X_n . \quad (23)$$

Dies möge nun für alle  $n = 1, 2, \dots, N$  zutreffen. Dann kann man die Zufallsgröße  $Y$  als additive Überlagerung normalverteilter und statistisch unabhängiger Zufallsgrößen  $Y_n$  auffassen, mit dem Mittelwert

$$\bar{Y} = f(\bar{X}_1, \dots, \bar{X}_{n-1}, \bar{X}_n, \bar{X}_{n+1}, \dots, \bar{X}_N)$$

und der Standardabweichung<sup>16)</sup>

$$\Delta Y \approx \sqrt{\sum_{n=1}^N \Delta Y_n^2} .$$

---

<sup>14)</sup> Man sagt hier, dass  $y$  indirekt gemessen wurde. Die Sprechweise,  $y$  sei berechnet ist häufig irreführend, weil dies suggerieren kann, dass  $y$  aus rein theoretischen Rechnungen stammt.

<sup>15)</sup> Für normalverteilte Zufallsgrößen ist die Unkorreliertheit äquivalent zur statistischen Unabhängigkeit. Im Allgemeinen folgt jedoch aus der Unkorreliertheit nicht die Unabhängigkeit. Allerdings gilt die Umkehrung immer.

<sup>16)</sup> Allgemein gilt: Seien zwei Zufallsgrößen normalverteilt und unkorreliert, mit den Mittelwerten  $m_1$  bzw.  $m_2$  und den Standardabweichungen  $\sigma_1$  bzw.  $\sigma_2$ . Dann ist die Summe dieser Zufallsgrößen ebenfalls normalverteilt, mit dem Mittelwert  $m_1 + m_2$  und der Standardabweichung  $\sqrt{\sigma_1^2 + \sigma_2^2}$ . Eine analoge Aussage gilt für mehr als zwei Zufallsgrößen.

Ersetzt man hier  $\Delta Y_n$  nach (23) und verwendet für die Standardabweichungen  $\Delta X_n$  die Schätzwerte  $\Delta x_n$ , so erhält man die sogenannte *GAUSSsche Regel der Fehlerfortpflanzung*,

$$\Delta y = \sqrt{\sum_{n=1}^N \left( \frac{\partial f}{\partial x_n} \Big|_{\bar{x}_1, \dots, \bar{x}_n, \dots, \bar{x}_N} \cdot \Delta x_n \right)^2} \quad (24)$$

Zusammenfassend gilt also Folgendes: Für die einzelnen Messwerte  $x_1$  bis  $x_N$  seien folgende Voraussetzungen erfüllt,

1. Für jeden Messwerte  $x_n$ ,  $n = 1, 2, \dots, N$ , sei ein empirischer Mittelwert  $\bar{x}_n$  und die empirische Standardabweichung  $\Delta x_n$  bekannt. Der Mittelwert sei (zumindest näherungsweise) normalverteilt. (Dies ist nach dem zentralen Grenzwertsatz der Fall ist, wenn  $\bar{x}_n$  aus vielen Einzelmessungen durch arithmetrische Mittelung gewonnen wurde.)
2. Die Funktion  $y = f(x_1, \dots, x_N)$  ändere sich im Intervall  $[\bar{x}_n - \Delta x_n, \bar{x}_n + \Delta x_n]$ , für alle  $n = 1, 2, \dots, N$  annähernd linear.

Unter diesen Bedingungen ist die berechnete Größe  $y$  auch (näherungsweise) normalverteilt, mit dem Mittelwert  $\bar{y} = f(\bar{x}_1, \dots, \bar{x}_N)$  und der Standardabweichung  $\Delta y$ , die sich nach (24) näherungsweise berechnen lässt.

Diese traditionelle Vorgehensweise impliziert viele Ungenauigkeiten, weil die Voraussetzungen wie Normalverteilung und Linearität in der Regel (niemals) streng erfüllt sind. Mit der heutigen Rechentechnik lassen sich jedoch im konkreten Einzelfall meist recht leicht Simulationen durchführen, bei denen in Sekundenschnelle viele Tausend Realisierungen der Zufallsgröße  $Y$  berechnen werden, so dass man deren Wahrscheinlichkeitsverteilung recht genau approximieren kann und somit die gewünschten Aussagen über die Größe  $y$  bekommt. Deshalb wird der Weg der Simulation heutzutage häufig beschritten.

Jedoch hat die obige klassische GAUSSsche Regel der Fehlerfortpflanzung auch heute ihre Berechtigung, liefert sie doch einen schnellen Überblick zur Messabweichung von der gesuchten Größe  $y$ . Darüber hinaus kann der Term

$$E_n \equiv \left| \frac{\partial f}{\partial x_n} \right|$$

als *Empfindlichkeit* der Größe  $y$  bei Änderung der Messgröße  $x_n$  angesehen werden. Gilt nun zum Beispiel  $E_1 \cdot \Delta x_1 \ll E_2 \cdot \Delta x_2$ , so würde man beim Experimentieren anstreben, den Wert  $x_2$  häufiger zu messen, so dass der Vertrauensbereich von  $\Delta x_2$  kleiner wird. Somit geben die Empfindlichkeiten  $E_n$  wichtige Hinweise für die Versuchsplanung.

## 4 Regression

### 4.1 Problemstellung

Es seien zwei (physikalische) Größen  $x$  und  $y$  gegeben. Aus theoretischen Überlegungen sei bekannt, dass zwischen ihnen ein funktionaler Zusammenhang einer bestimmten Form (Funktionen-Klasse) besteht,

$$y = f_{a,b,\dots}(x) .$$

Darin seien  $a, b, \dots$  gewisse Parameter, deren genauer Wert zunächst unbekannt ist. Um diese Werte zu ermitteln, werden nun für verschiedene Messwerte  $x_n$  von  $x$  die zugehörigen Messwerte  $y_n$  von  $y$  experimentell bestimmt, möglicherweise auch indirekt. Man erhält dann zwei Messreihen, die als *verbundene*

---

Stichprobe angesehen werden,

$$\{(x_n, y_n)\}_{n=1,2,\dots,N} .$$

Die Messwerte sind im Allgemeinen mit Messabweichungen behaftet, folglich wird in der Regel

$$y \neq f_{a,b,\dots}(x_n)$$

gelten.<sup>17)</sup> Mit der sogenannten *Regression* (oder *Ausgleichsrechnung*) werden die Parameter  $a, b, \dots$  aus den gegebenen Messwerten „möglichst gut“ geschätzt. Dabei wird häufig die Methode der kleinsten quadratischen Abweichung angewandt, indem die Parameter gerade so gewählt werden, dass die Größe

$$\chi^2(a, b, \dots) \equiv \sum_{n=1}^N (y_n - f_{a,b,\dots}(x_n))^2$$

minimal wird. Für normalverteilte Messabweichungen ist diese Methode äquivalent zur Maximum-Likelihood-Methode, die für große  $N$  beste (erwartungstreue, effiziente) Schätzungen für die Parameter liefert (s. Anhang, S. 24ff.).

## 4.2 Lineare Regression

Im Spezialfall, dass zwischen den Größen  $x$  und  $y$  ein linearer Zusammenhang besteht,

$$y = a + b \cdot x ,$$

spricht man von einer *linearen Regression*. Je nachdem, ob nur  $y$  oder auch  $x$  Messabweichungen haben, versehen ist, werden verschiedene Fälle betrachtet (s. Anhang, S. ??). Die beste Schätzung der beiden Parameter erhält man wie folgt:

$$b = \frac{\overline{xy} - \bar{x} \bar{y}}{S_x^2} \quad (25)$$

mit

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n , \quad \bar{y} = \frac{1}{N} \sum_{n=1}^N y_n , \quad \overline{xy} = \frac{1}{N} \sum_{n=1}^N x_n y_n , \quad \overline{x^2} = \frac{1}{N} \sum_{n=1}^N x_n^2 , \quad (26)$$

und

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

sowie

$$a = \bar{y} - b \cdot \bar{x} \quad (27)$$

Die Standardabweichungen von  $b$  und  $a$  sind

$$s_b = \sqrt{\frac{1}{N} \frac{\sigma_y^2}{S_x^2}} , \quad \text{bzw.} \quad s_a = s_b \cdot \sqrt{\overline{x^2}} . \quad (28)$$

---

<sup>17)</sup> Diese Ungleichheit kann auch dadurch entstehen, dass eine zu enge Funktionen-Klasse gewählt wurde.



## A Anhang: Lineare Regression

### A.1 1. Fall: $x$ fehlerfrei, $y$ mit festem Fehler $\sigma_y$

Die Messwerte  $x_n$ ,  $n = 1, 2, \dots, N$ , seien fehlerfrei, und die Messfehler  $y_n$  seien bei beliebigem  $n$  normalverteilt mit der Standardabweichung  $\sigma_y$ . Dann ist die Log-Likelihood-Funktion gegeben durch

$$\begin{aligned} \ln L(y_1, \dots, y_N) &= \ln \prod_{n=1}^N \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{[y_n - (a + bx_n)]^2}{2\sigma_y^2}\right) \\ &= \sum_{n=1}^N \left[ \ln\left(\frac{1}{\sqrt{2\pi}\sigma_y}\right) + \left(-\frac{[y_n - (a + bx_n)]^2}{2\sigma_y^2}\right) \right] \\ &= \sum_{n=1}^N \ln\left(\frac{1}{\sqrt{2\pi}\sigma_y}\right) - \frac{\chi^2}{2}. \end{aligned}$$

darin haben wir

$$\chi^2(a, b) \equiv \frac{1}{\sigma_y^2} \sum_{n=1}^N (y_n - a - b \cdot x_n)^2 \quad (29)$$

gesetzt. Die Likelihood-Funktion nimmt ihren Maximalwert beim Minimum von  $\chi^2(a, b)$  an. Dies ist die sogenannte *Methode der kleinsten Fehlerquadrate*,

$$\chi^2(a, b) = \min!$$

Aus

$$\frac{\partial \chi^2(a, b)}{\partial a} = 0 \quad \text{und} \quad \frac{\partial \chi^2(a, b)}{\partial b} = 0$$

erhalten wir folgende Schätzwerte für die Parameter  $a$  und  $b$  unserer Regressionsgeraden:

$$\hat{b} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x^2} \quad (30)$$

mit

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n, \quad \bar{y} = \frac{1}{N} \sum_{n=1}^N y_n, \quad \overline{xy} = \frac{1}{N} \sum_{n=1}^N x_n y_n, \quad \overline{x^2} = \frac{1}{N} \sum_{n=1}^N x_n^2, \quad (31)$$

und

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

sowie

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x} \quad (32)$$

Die Standardabweichung von  $\hat{b}$  ist

$$s_b = \sqrt{\frac{1}{N} \frac{\sigma_y^2}{S_x^2}}. \quad (33)$$

Wenn die Gerade gut den Zusammenhang zwischen  $x$  und  $y$  widerspiegelt, dann gilt  $\chi^2/(N-2) \approx 1$ , wobei  $\chi^2$  in (29) an der Stelle  $a = \hat{a}$  und  $b = \hat{b}$  zu nehmen ist. Folglich gilt dann

$$\begin{aligned} \sigma_y^2 \approx s_y^2 &= \frac{1}{N-2} \sum_{n=1}^N (y_n - \hat{b} \cdot x_n - \hat{a})^2 \\ &= \frac{N}{N-2} \left( \overline{y^2} - \bar{y}^2 - \hat{b}^2 (\overline{x^2} - \bar{x}^2) \right) \\ &= \frac{N}{N-2} S_y^2 (1 - \rho_{xy}^2), \end{aligned}$$

mit dem Korrelationskoeffizienten

$$\varrho_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x S_y}, \quad S_y^2 = \overline{y^2} - \bar{y}^2, \quad \overline{y^2} = \frac{1}{N} \sum_{n=1}^N y_n^2.$$

Setzt man dies in (33) ein, folgt schließlich,

$$s_b \approx \sqrt{\frac{1}{N-2} \cdot \frac{S_y^2}{S_x^2} \cdot (1 - \varrho_{xy}^2)} \quad (34)$$

Auf der rechten Seite stehen nur noch Größen, die aus den Messwertpaaren  $(x_n, y_n)$  berechnet werden können.

Für die Standardabweichung von  $\hat{a}$  erhält man

$$s_a = s_b \cdot \sqrt{\overline{x^2}} \quad (35)$$

Der nach (30) geschätzte Anstieg  $\hat{b}$  liegt mit ca. 68% iger Wahrscheinlichkeit dichter als  $s_b$  am (unbekannten) wahren Anstieg  $b$ ,

$$b = \hat{b} \pm s_b \quad (68\%) .$$

Eine analoge Aussage gilt für den Schnittpunkt  $a$ ,

$$a = \hat{a} \pm s_a \quad (68\%) .$$

Ersetzt man in den Gleichungen  $\sigma_y$  durch  $2 \cdot \sigma_y$ , erhöht sich die Signifikanz auf etwa 95%. Es sei jedoch erinnert, dass hierbei nur zufällige Fehler der Messwerte  $y_n$  erfasst werden.

## A.2 2. Fall: $x_n$ und $y_n$ mit Fehler $\sigma_x$ bzw. $\sigma_y$

Die Messwerte  $x_n$  und  $y_n$  seien mit dem Fehler (Standardabweichung)  $\sigma_x$  bzw.  $\sigma_y$  gemessen, für jede Messung  $n$  also gleich. Man kann dies näherungsweise auf den ersten Fall zurückführen, indem man  $x_n$  wiederum als fehlerfrei ansieht, dafür aber die Varianz von  $y_n$  um  $\hat{b}^2 \sigma_x^2$  erhöht. Die Parameter  $a$  und  $b$  sowie deren Unsicherheiten werden also wieder nach (30), (32), (34) und (35) geschätzt, wobei in (33)

$$\sigma_y^2 \longrightarrow \sigma_{y^*}^2 = \sigma_y^2 + \hat{b}^2 \sigma_x^2$$

zu ersetzen ist. Insbesondere gilt hier für den Fall, dass die Gerade die Daten gut approximiert,

$$\sigma_y^2 + \hat{b}^2 \sigma_x^2 \approx s_{y^*}^2 = \frac{1}{N-2} \sum_{n=1}^N (y_n - \hat{b} \cdot x_n - \hat{a})^2 .$$

Folglich kann die Ungenauigkeit des Anstiegs genau wie im ersten Fall nach (34) bestimmt werden.

## A.3 3. Fall: $x_n$ fehlerfrei, $y_n$ mit variablem Fehler $\sigma_{y;n}$

Die Messwerte  $x_n$  seien fehlerfrei, und die Messfehler  $y_n$ ,  $n = 1, 2, \dots, N$ , seien normalverteilt mit der Standardabweichung  $\sigma_{y;n}$ .

Die Methode der kleinsten gewichteten Fehlerquadrate,

$$\chi^2 = \sum_{n=1}^N \frac{(y_n - b \cdot x_n - a)^2}{\sigma_{y;n}^2} = \min! \quad (36)$$

liefert dann dieselben Gleichungen wie (30), (32), (34) und (35), nur mit dem Unterschied, dass anstelle der gewöhnlichen (31) nun gewichtete Mittelwerte treten,

$$\bar{x} = \sum_{n=1}^N w_n x_n, \quad \bar{y} = \sum_{n=1}^N w_n y_n, \quad \overline{xy} = \sum_{n=1}^N w_n x_n y_n, \quad \overline{x^2} = \sum_{n=1}^N w_n x_n^2, \quad (37)$$

mit den Gewichten

$$w_n = \frac{\frac{1}{\sigma_{y;n}^2}}{\sum_{m=1 \dots N} \frac{1}{\sigma_{y;m}^2}} = \left( \sum_{m=1 \dots N} \frac{\sigma_{y;n}^2}{\sigma_{y;m}^2} \right)^{-1}.$$

Darüber hinaus ist in (34)  $\sigma_y^2$  durch

$$\overline{\sigma_y^2} = \frac{N}{\sum_{m=1 \dots N} \frac{1}{\sigma_{y;m}^2}}$$

zu ersetzen.

Sind alle Fehler gleich  $\sigma_{y;n}^2 = \sigma_y^2$ , folgt  $w_n = 1/N$ , was wieder der gewöhnlichen Mittelung (31) entspricht, sowie  $\overline{\sigma_y^2} = \sigma_y^2$ .

Ist die Ungenauigkeit  $\sigma_{y;n}$  ein gewisser Bruch  $\alpha$  des Messwertes  $y_n$ , z.B.  $\alpha = 0,005$  bei einem relativen Fehler von 0,5%, also  $\sigma_{y;n} = \alpha \cdot y_n$ , dann kann für die Gewichte wie folgt geschrieben werden,

$$w_n = \left( \sum_{m=1 \dots N} \frac{y_n^2}{y_m^2} \right)^{-1}.$$

#### A.4 4. Fall: $x_n$ und $y_n$ mit variablem Fehler $\sigma_{x;n}$ bzw. $\sigma_{y;n}$

Die Messwerte  $x_n$  und  $y_n$  seien mit dem Fehler (Standardabweichung)  $\sigma_{x;n}$  bzw.  $\sigma_{y;n}$  gemessen. Man kann dies näherungsweise auf den obigen Fall zurückführen, indem man  $x_n$  wiederum als fehlerfrei ansieht, dafür aber die Varianz von  $y_n$  um  $\hat{b}^2 \sigma_{x;n}^2$  erhöht. Die Parameter  $a$  und  $b$  sowie deren Unsicherheiten werden also wieder nach (30), (32), (34) und (35) geschätzt, wobei gewichtete Mittelwerte (37) zu verwenden sind, nun aber mit den Gewichten

$$w_n = \frac{\frac{1}{\sigma_{y;n}^2 + \hat{b}^2 \sigma_{x;n}^2}}{\sum_{m=1 \dots N} \frac{1}{\sigma_{y;m}^2 + \hat{b}^2 \sigma_{x;m}^2}}.$$

Damit hängt aber die rechte Seite von (30) auch von  $\hat{b}$  ab, und eine Auflösung nach  $\hat{b}$  ist im Allgemeinen schwierig.

# Index

Ausgleichsrechnung, 24  
Ausreißer, 10

Box, 9  
Boxplot, 9

Dezil, 9

Erwartungstreue, Mittelwertschätzung, 11

Fehler, gesamt, 19  
Fehler, zufällig, 12  
Fehlerfortpflanzung, ff., 21  
Fehlerfortpflanzung, nach Gaus, 23  
Fehlerfortpflanzung, statistisch, 22

Gaussverteilung, 7  
Glockenkurve, 7

Median, 9  
Messabweichung, 6  
Messabweichung, grob, 10  
Messabweichung, systematisch, 12  
Messabweichung, zufällig, 12  
Messfehler, 6  
Messreihe, 4, 5

Normalverteilung, 7

Perzentil, 9

Quantil, 8  
Quantile, 8  
Quartil, 9  
Quintil, 9

Regression, lineare, 24

Schätzer, erwartungstreuer, 11  
Schätzer, konsistenter, 11  
Standardabweichung des Mittelwertes, 14  
Standardabweichung, empirische, 13  
Stichprobe, 5  
Stichprobe, mathematische, 14  
Stichprobe, verbundene, 24  
Student-Verteilung, 15

t-Verteilung, 15  
t-Wert, 15  
Terzil, 9

Verteilungsdichte, 7  
Vertrauensbereich, 14

wahrer Wert, 6  
Wahrscheinlichkeitsdichte, 7

Zufallsgröße, 5