Analyse von Messdaten

Wissenschaftliche Methodik I

Björn Voß

Institut für Biomedizinische Genetik - RNA Biologie & Bioinformatik

Daten messen und auswerten

Größe/Merkmal und Einheit
Eine **Größe** oder ein **Merkmal** ist eine meßbare Eigenschaft.

- Höhe eines Baumes (in m)
- Schuppenfarbe eines Fisches (Wellenlänge in nm)
- Expressionshöhe eines Gens (Fluoreszenzintensität in w.E. *)

Messwert und **Einheit** zusammen stellen das Ergebnis einer Messung dar.

Im Allgemeinen
Jede Messung unterliegt unvermeidbaren Unsicherheiten.

*willkürliche Einheit

Messunsicherheiten - Beispiel

Sie wollen die Körpergröße eines/r Freundes/in bestimmen:

- Sie schätzen aufgrund Ihrer Erfahrung; Ergebnis: 175 cm
  Unsicherheit: Größe liegt zwischen 170 und 180 cm.
- Sie verwenden einen Zollstock und messen freistehend; Ergebnis: 174cm
  Unsicherheit: Größe zwischen 173 und 175 cm.
- Beim Arzt: gerader Rücken, Winkel auf dem Kopf; Ergebnis 175,5 cm
  Unsicherheit: Größe zwischen 175,3 und 175,7 cm.

Es gibt in diesem Beispiel zwei Faktoren die Einfluß auf das Messergebnis haben

- die Art wie gemessen wird (aufrecht oder gebeugt, Kopfneigung, ...)
- das Messgerät (Auge, Zollstock, ...)

Messabweichung

Im Allgemeinen gibt es noch viel mehr Faktoren die Messergebnisse beeinflussen können, so dass die Messwerte von den wahren Werten abweichen.

Die Betrachtung dieses Problems wurde früher als **Fehlerbetrachtung** bezeichnet und obwohl mittlerweile der Begriff Messabweichung als passender angesehen wird, wird der Begriff auch heute noch verwendet.

Ziel einer Messung
Möglichst genauen Schätzwert unter Angabe der Unsicherheit (Messabweichung) für ein Merkmal ermitteln.

$\Rightarrow$ Fehlerbetrachtung essentiell

Fehler

Die Verwendung des Begriffes „Fehler“ ist üblich, jedoch leider etwas irreführend. Wie bereits erwähnt sollte lieber von Messabweichungen gesprochen werden, da

- es sich hierbei nicht um ein Fehlverhalten des Messenden handelt
- das Ergebnis nicht um genau diesen Betrag falsch ist.

Größe der Abweichung

Warum ist es wichtig die Größe der Abweichung zu kennen?
Um herauszufinden ob zwei Biersorten den gleichen Alkoholgehalt haben führen zwei Gruppen Messungen an beiden Sorten durch.

Gruppe 1 ermittelt:	Sorte A: $4,8 \pm 0,1$ % vol.	Sorte B: $5,1 \pm 0,1$ % vol.
Gruppe 2 ermittelt:	Sorte A: $4,8 \pm 0,3$ % vol.	Sorte B: $5,1 \pm 0,3$ % vol.

Obwohl beide Gruppen die gleichen mittleren Alkoholkonzentrationen bestimmt haben, kann nur die erste mit genügender Sicherheit sagen, dass sich die Sorten in ihrem Alkoholgehalt unterscheiden.

Ursachen für Messabweichungen

Man kann verschiedene Ursachen unterscheiden:

- Grobe Fehler
- Systematische Abweichungen
- Zufällige Unsicherheiten

Grobe Fehler

Diese können der Person die die Messungen vorgenommen hat zugeschrieben werden.

Beispiele

- Fehler bei der Ablesung (falscher Betrachtungswinkel) und Bedienung (falscher Messbereich) des Messgerätes.
- Protokollierung falscher Daten oder Einheiten
- Fehler bei der Auswertung (Vorzeichen-, Rechen- oder Programmierfehler)

Maßnahmen

- Verhinderung durch sorgfältige Planung und gewissenhaftes Arbeiten!
- Nachvollziehbar/Korrigierbar durch detaillierte Protokollierung

Systematische Abweichungen

Beispiele

- Nichtbetrachtung von Einflüssen (Temperatur, Luftdruck, ...)
- Verunreinigungen in Substanzen (z.B. Spuren von Quecksilber in Nährmedien)
- Einfluss der Messung auf das Messobjekt (Thermometer entzieht Körper Wärme)
- Mängel an den Messgeräten (Alterung, Eichfehler, Undichtigkeiten, Konstruktionsmängel)

Problem
Können auch durch wiederholte Messungen nicht entdeckt werden.

Zufällige Unsicherheiten

Beispiele

- Zufällige Schwankungen der Messbedingungen
- Menschliche Unzulänglichkeiten (Sehschärfe beim Ablesen, „ruhige Hand“)

Typischerweise
Zufällige Abweichungen treten gleichermaßen in positiver wie in negativer Weise auf.

Salopp gesagt: „Sie heben sich auf!“

Zufällige Abweichungen können durch geeignete statistische Methoden in der Auswertung berücksichtigt werden.

Zufällige Unsicherheiten

Beispiel: Messung der Fallzeit eines Apfels aus 3m Höhe mit einer Stoppuhr.

- Fehlerquelle: Mensch - zufällige Schwankungen der Reaktionszeit
  $\Rightarrow$ zufällige Abweichungen in den gemessenen Zeiten
- Fehlerquelle: Uhr - geht zu langsam
  $\Rightarrow$ systematische Abweichung der Messdaten

Beispiel: Volumenmessung mittels Meßzylinder
Fehlerquelle: Parallaxe

- Das Auge ist nie in der exakt gleichen Position
  $\Rightarrow$ **zufällige** Abweichungen
- Durch den Versuchsaufbau bedingt schaut der Experimentator immer schräg auf das Messgerät
  $\Rightarrow$ **systematische** Abweichung

Angabe von Messabweichungen

Nehmen wir an, dass wir die Messabweichung δ_x für die uns interessierende Größe x in einem Experiment quantifizieren können. Dann geben wir das Messergebnis folgendermaßen an:

$$\text{Messwert} = \text{Bestwert} \pm \text{Abweichung}$$
$$x \quad = \quad x_{best} \quad \pm \quad \delta_x$$

Beispiel: Fallzeit des Apfels aus 3m Höhe

- Bestwert ($\approx$ Durchschnittszeit) $T_{best} = 0,78s$
- wahrscheinlicher Bereich $0,76 \leq T \leq 0,80$
  $T = 0,78s \pm 0,02s = (0,78 \pm 0,02)s$

Signifikante Stellen – sinnvolles Runden

Die Messabweichung δ_x ist ein Schätzwert und bestimmt die signifikante Stelle auf die das Messergebnis gerundet wird.

Beispiel: Messung der Erdbeschleunigung

$$g_{best} = 9,823765\,\frac{m}{s^2},\ \delta_g = 0,023769\,\frac{m}{s^2}$$

$$\Rightarrow 9,82\,\frac{m}{s^2} \pm 0,023769\,\frac{m}{s^2} \qquad \textbf{oder} \qquad 9,824\,\frac{m}{s^2} \pm 0,024\,\frac{m}{s^2}$$

Rundung der Messabweichung auf erste Stelle $z \neq 0$. Sollte $z = 1$ oder $z = 2$ sein, dann wird auf die nächste Stelle gerundet. Die Messabweichung wird immer **aufgerundet**! (DIN 1333)

Messwert	8,579617	8,579617
Abweichung	0,00383	0,00163
Rundestelle	1. Ziffer $\neq 0$ ist die 3	1. Ziffer $\neq 0$ ist die 1
	3. Nachkommastelle	4. Nachkommastelle
Messwert gerundet	$8,580$	$8,5796$
Messergebnis	$8,580 \pm 0,004$	$8,5796 \pm 0,0017$

Signifikante Stellen – sinnvolles Runden

Die Messabweichung δ_x ist ein Schätzwert und bestimmt die signifikante Stelle auf die das Messergebnis gerundet wird.

Beispiel: Messung der Erdbeschleunigung

$$g_{best} = 9,823765\frac{m}{s^2}, \delta_g = 0,023769\frac{m}{s^2}$$
$$\Rightarrow 9,82\frac{m}{s^2} \pm 0,023769\frac{m}{s^2} \qquad \textbf{oder} \qquad 9,824\frac{m}{s^2} \pm 0,024\frac{m}{s^2}$$

Rundung der Messabweichung auf erste Stelle $z \neq 0$. Sollte $z = 1$ oder $z = 2$ sein, dann wird auf die nächste Stelle gerundet. Die Messabweichung wird immer **aufgerundet**! (DIN 1333)

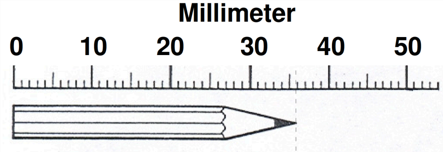
Messwert	8,579617	8,579617
Abweichung	0,00383	0,00163
Rundestelle	1. Ziffer $\neq 0$ ist die 3	1. Ziffer $\neq 0$ ist die 1
	3. Nachkommastelle	4. Nachkommastelle
Messwert gerundet	$8,580$	$8,5796$
Messergebnis	$8,580 \pm 0,004$	$8,5796 \pm 0,0017$

Runden

- Berechnungen sollten mit auf mehr Stellen gerundeten Werten durchgeführt werden
- Nur das Endergebnis wird gerundet wie beschrieben
- Die Messabweichung wird immer aufgerundet
- Die signifikante Stelle muss keine Nachkommastelle sein.
  Beispiel: 40.036,392 km $\pm$ 137,653 km $\rightarrow$ 40.040 km $\pm$ 140 km

Bestimmung zufälliger Messabweichungen

Schätzung
Beispiel: Messung mit einem Lineal



Annahme: Lineal ist zuverlässig
$\rightarrow$ **kein systematischer Fehler**

Spitze am nächsten zu Teilstrich bei 36 mm.

$l_{best} = 36$ mm, $\delta_l =??$ $\Rightarrow$ Wie genau kann ich mit dem Lineal messen?
Da die Teilstriche 1mm voneinander entfernt sind schätze ich die zufällige Abweichung auf $\pm 0,5$ mm.

$$\Rightarrow l = 36,0\text{mm} \pm 0,5\text{mm}$$

Schätzung durch wiederholte Messung

Mehrfache Messung am selben Objekt
Beispiel: Fallzeit eines Apfels aus 3m Höhe

Annahmen:

- zuverlässige Uhr $\rightarrow$ kein systematischer Fehler
- nur zufällige Abweichungen; Grund: Reaktionszeit bei Start und Stop der Uhr
- 5-malige Wiederholung des Versuchs

Messreihe: $0,77s \quad 0,80s \quad 0,79s \quad 0,76s \quad 0,78s$

Bestwert $T_{best} =$ Mittelwert $\bar{T} = \frac{0{,}77s+0{,}80s+0{,}79s+0{,}76s+0{,}78s}{5} = 0,78s$

Die Unsicherheit der Messung können wir aus der Streuung der Daten schätzen.

wahrscheinlicher Bereich: $0,76s - 0,80s \Rightarrow \delta_T = 0,02s$

Bestimmung der Messabweichung

Die bisher vorgestellten Methoden sind

subjektiv	welche Genauigkeit bei der Messung traue ich mir zu
labil	einzelne Ausreißer führen zu großen Messabweichungen
kontraintuitiv	je mehr Wiederholungen ich mache umso größer wird δ

Ideal
Methode zur Bestimmung der Messunsicherheit die **objektiv** und **robust** ist, sowie deren Genauigkeit durch erhöhten Aufwand gesteigert werden kann.

$\Rightarrow$ **Statistische Betrachtung**

Nimmt man an, dass nur zufällige Faktoren einen Einfluss auf die Messunsicherheit haben, dann ist eine statistische Herangehensweise möglich. Weiterhin wäre es ideal Messungen am selben Objekt zu wiederholen (technische Replikate)*.

Beispiel: Messung einer Länge
N=5 Messungen mit folgenden Ergebnissen für die **Länge l in mm: 71, 72, 72, 73, 71**
Es ist naheliegend anzunehmen, dass der Mittelwert dem wahren Wert am nächsten kommt:

$$\Rightarrow l_{best} = \bar{l} = \frac{71+72+72+73+71}{5}\mathrm{mm} = 71,8\mathrm{mm}$$

*zerstörungsfreie Messung

Statistische Betrachtung der Messabweichung

Wie gut waren unsere Messungen?
Am einfachsten wäre ein Vergleich mit dem wahren Wert, da dieser jedoch meistens* unbekannt ist verwenden wir den Bestwert (im Beispiel der Mittelwert) der Messung.

Messung Nr.	Messwert l_i [mm]	Abweichung $l_i - \bar{l}$ [mm]	$(l_i - \bar{l})^2$ [mm^2]
1	71	-0,8	0,64
2	72	0,2	0,04
3	72	0,2	0,04
4	73	1,2	1,44
5	71	-0,8	0,64
	$\bar{l} = 71,8$	$\sum(l_i - \bar{l}) = 0$	$\sum(l_i - \bar{l})^2 = 2,80$

$\Rightarrow$ **Nur die Summe der quadratischen Abweichungen taugt als Maß**

*außer bei der Eichung/Kalibrierung

Statistische Betrachtung der Messabweichung

Problem
Die Summe der quadratischen Abweichungen steigt mit der Anzahl der Messungen.

Mittlere quadratische Abweichung - Varianz
Normalisierung auf die Anzahl der Messungen

$$\mathbf{Var}(l) = \frac{1}{N} \sum_{i=1,\ldots,N} (l_i - \bar{l})^2$$

Standardabweichung
Um nun noch ein besser „interpretierbares“ Maß (gleiche Einheit wie die Messwerte) zu bekommen wird noch die Wurzel gezogen

$$\tilde{s}(l) = \sqrt{\frac{1}{N} \sum_{i=1,\ldots,N} (l_i - \bar{l})^2}$$

Statistische Betrachtung der Messabweichung

Messung Nr.	Messwert l_i [mm]	Abweichung $l_i - \bar{l}$ [mm]	$(l_i - \bar{l})^2$ [mm^2]
1	71	-0,8	0,64
2	72	0,2	0,04
3	72	0,2	0,04
4	73	1,2	1,44
5	71	-0,8	0,64
	$\bar{l} = 71,8$	$\sum(l_i - \bar{l}) = 0$	$\sum(l_i - \bar{l})^2 = 2,80$

$$\text{Varianz:}\ \ \mathbf{Var}(l) = \frac{1}{N} \sum_{i=1,\dots,N} (l_i - \bar{l})^2 = \frac{1}{5} \cdot 2,80 = 0,56$$

$$\text{Standardabweichung:}\ \ \tilde{s}(l) = \sqrt{\mathbf{Var}(l)} = \sqrt{0,56} = 0,748331$$

Statistische Betrachtung der Messabweichung

Messung Nr.	Messwert l_i [mm]	Abweichung $l_i - \bar{l}$ [mm]	$(l_i - \bar{l})^2$ [mm^2]
1	71	-0,8	0,64
2	72	0,2	0,04
3	72	0,2	0,04
4	73	1,2	1,44
5	71	-0,8	0,64
	$\bar{l} = 71,8$	$\sum(l_i - \bar{l}) = 0$	$\sum(l_i - \bar{l})^2 = 2,80$

$$\text{Varianz: } \mathbf{Var}(l) = \frac{1}{N} \sum_{i=1,\dots,N} (l_i - \bar{l})^2 = \frac{1}{5} \cdot 2,80 = 0,56$$

$$\text{Standardabweichung: } \tilde{s}(l) = \sqrt{\mathbf{Var}(l)} = \sqrt{0,56} = 0,748331$$

Interpretation?

Bedeutung
Viele in der Anwendung auftretende Verteilungen können durch eine Normalverteilung gut genähert werden.

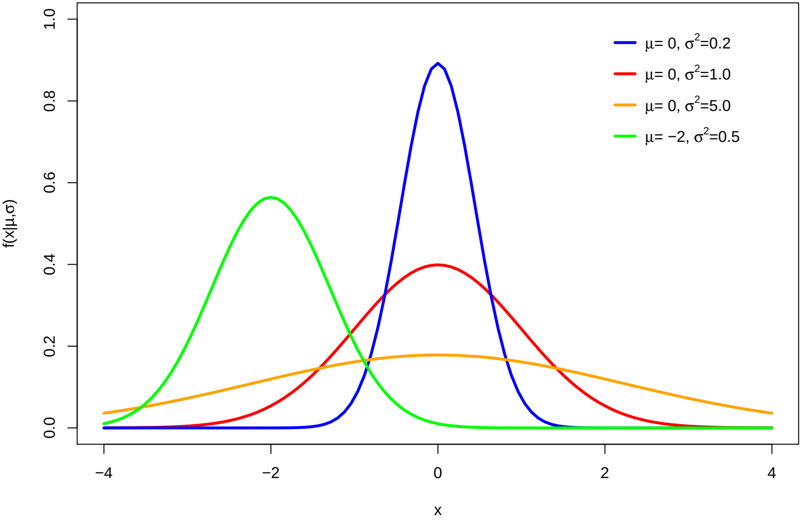
Sind $x_1, \ldots, x_n$ Beobachtungen eines solchen Merkmals, dann approximiert man:

$$\begin{aligned} \mu &\approx \bar{x} \\ \sigma &\approx s \quad \text{(oder } \tilde{s}\text{)} \end{aligned} \tag{1}$$

Dichte der Normalverteilung

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \qquad x \in \mathbb{R}$$

$\mu \in \mathbb{R}$ heißt Mittelwert und $\sigma > 0$ Standardabweichung von $f(x|\mu,\sigma)$.
(Details folgen)



$\mu \rightarrow$ Lage, $\sigma \rightarrow$ Form

Die Standardnormalverteilung

Die Normalverteilung mit $\mu = 0$ und $\sigma = 1$ wird Standardnormalverteilung genannt.

Jede normal verteilte Variable X kann durch folgende Umformung in die standardnormalverteilte Variable Z, mit $\mu = 0$ und $\sigma = 1$ überführt werden:

$$Z = \frac{X - \mu}{\sigma}$$

Dichtekurve der Standardnormalverteilung

$$\phi = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right)$$

Dichtekurve & Verteilungsfunktion

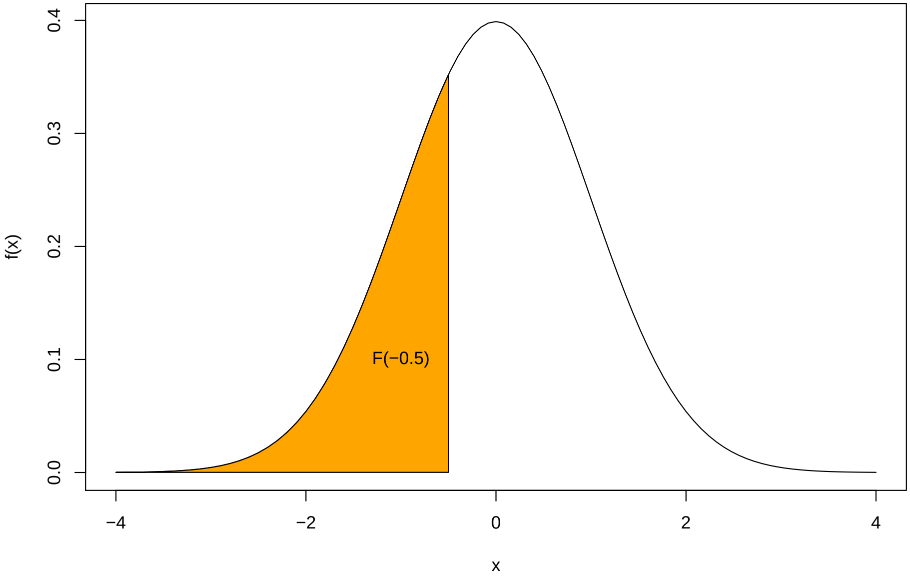
Dichte
Eine stetige Funktion f ist eine Dichte(kurve), wenn

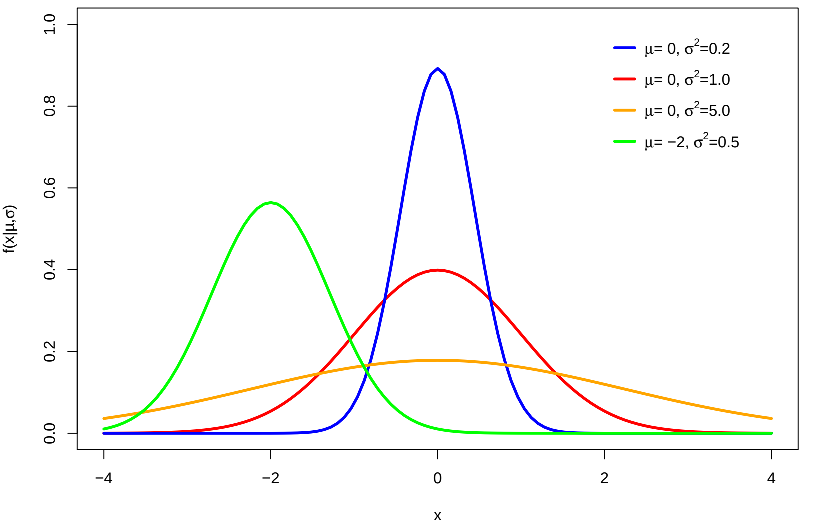
- $f(x) \geq 0$
- $\int_{\mathbb{R}} f(x)dx = 1$

Verteilungsfunktion
Die Verteilungsfunktion ist die Stammfunktion dieser Dichte:

$$F(x) = \int_{-\infty}^{x} f(y)dy$$

Mit ihr berechnet man die Fläche unter der Dichte(kurve) im Bereich $[-\infty, x]$.


F(−0.5)
f(x)
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x



$\mu \rightarrow$ Lage, $\sigma \rightarrow$ Form

Verteilungsfunktion der Normalverteilung

Dichte:

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}$$

Verteilungsfunktion:

$$F(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Überführen in die Standardnormalverteilung $Z = \frac{X-\mu}{\sigma}$

Dichte:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}z^2\right)}$$

Verteilungsfunktion:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$

Verteilungsfunktion der Normalverteilung

Dichte:

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}$$

Verteilungsfunktion:

$$F(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Überführen in die Standardnormalverteilung $Z = \frac{X-\mu}{\sigma}$

Dichte:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}z^2\right)}$$

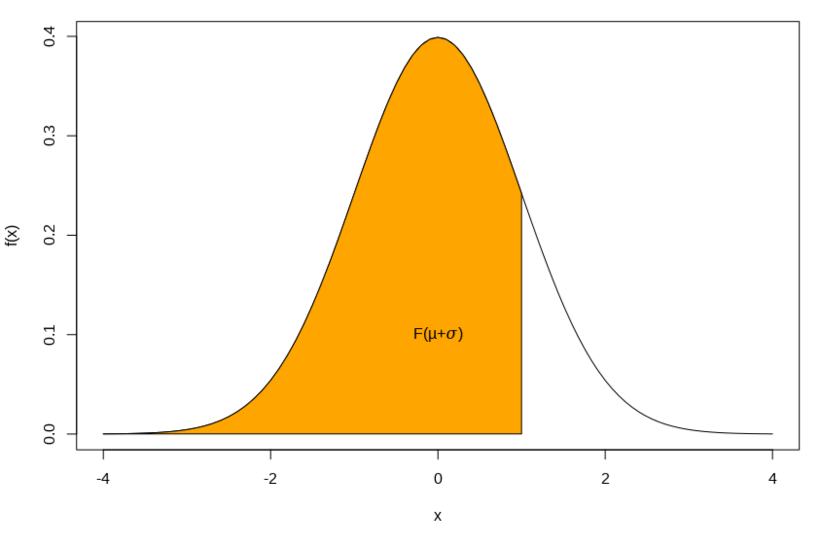
Verteilungsfunktion:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt = \Phi(x)$$

$$F(x|\mu,\sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

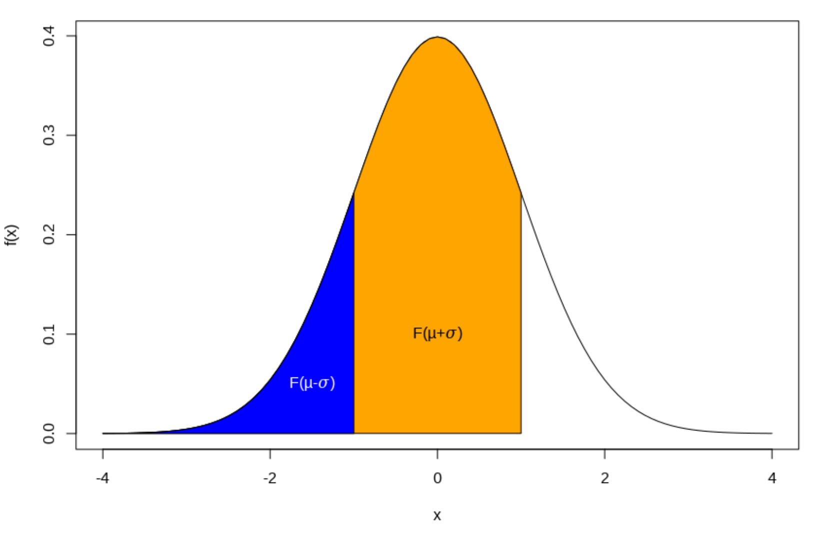
Sigmaumgebung

Wie groß ist der Bereich $\mu \pm \sigma$?



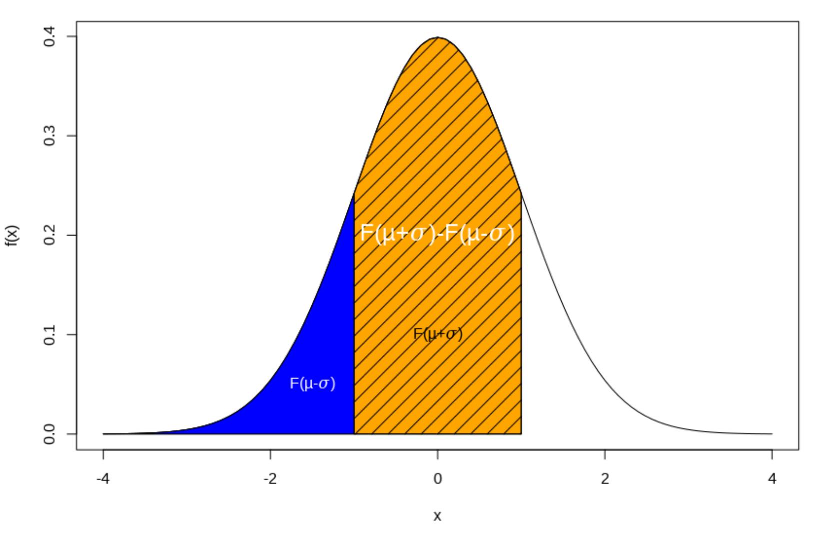
Sigmaumgebung

Wie groß ist der Bereich $\mu \pm \sigma$?



Sigmaumgebung

Wie groß ist der Bereich $\mu \pm \sigma$?

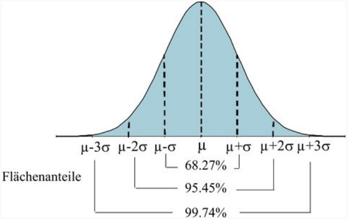


$$
\begin{aligned}
F(\mu \pm \sigma) = F(\mu + \sigma) - F(\mu - \sigma) &= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\
&= \Phi(1) - \Phi(-1) = 0.6826895
\end{aligned}
$$

$\Rightarrow$ Die Sigmaumgebung ist immer gleich groß (unabhängig von μ und σ)!

3-σ-Regel

68,3% der Beobachtungen liegen im Intervall $\mu \pm \sigma$
95,4% der Beobachtungen liegen im Intervall $\mu \pm 2\sigma$
99,7% der Beobachtungen liegen im Intervall $\mu \pm 3\sigma$



Angabe der Messunsicherheit

Die 3σ-Regel
Über die Standardabweichung lässt sich angeben mit welcher Wahrscheinlichkeit ein neuer Messwert in einem Bereich um den Mittelwert liegt.

Intervall	Wahrscheinlichkeit
$\bar{x} \pm \tilde{s}(x)$	68,3%
$\bar{x} \pm 2\tilde{s}(x)$	95,4%
$\bar{x} \pm 3\tilde{s}(x)$	99,7%

Messabweichung einer Einzelmessung
Für eine zukünftige Einzelmessung kann man nun sagen, dass sie mit einer Wahrscheinlichkeit von 95,4% im Bereich $\bar{x} \pm 2\tilde{s}(x)$ liegen wird.
In unserem Beispiel: $71,8\text{mm} \pm 2\sqrt{0,56}\text{mm} = 71,8\text{mm} \pm 1,5\text{mm}$

Anders interpretiert: Bei einer zukünftigen Messung eines anderen Objektes beträgt die Wahrscheinlichkeit 95,4%, dass der wahre Wert in einem Bereich von $\pm 1,5$mm um den Messwert liegt.

Was ist eigentlich mit dem Mittelwert?

Der aus den Messdaten berechnete Mittelwert entspricht sehr wahrscheinlich nicht dem wahren Wert für die Länge. Aber wie genau ist denn der Mittelwert?

Abweichung des Mittelwertes
Die Abweichung des Mittelwertes wird kleiner, je mehr Messungen durchgeführt werden. Die Standardabweichung des Mittelwertes ergibt sich zu:

$$\tilde{s}(\bar{x}) = \frac{\tilde{s}(x)}{\sqrt{N}} \qquad \text{(sog. } \textbf{Standardfehler}\text{)}$$

Die 3σ-Regel behält natürlich ihre Gültigkeit auch für Mittelwerte. Für eine Messung mit technischen Replikaten wird daher folgende Angabe verwendet:

- 68,3%-Vertrauensintervall: $\bar{x} \pm \frac{\tilde{s}(x)}{\sqrt{N}}$ $71,8\text{mm} \pm 0,3\text{mm}$
- 95,4%-Vertrauensintervall: $\bar{x} \pm 2\frac{\tilde{s}(x)}{\sqrt{N}}$ $71,8\text{mm} \pm 0,7\text{mm}$
- 99,7%-Vertrauensintervall: $\bar{x} \pm 3\frac{\tilde{s}(x)}{\sqrt{N}}$ $71,8\text{mm} \pm 1\text{mm}$

Zwischenfazit

Durchführung von Messungen

- Vermeidung von groben Fehlern und systematischen Abweichungen
- Minimierung der zufälligen Abweichungen

Bestimmung der Messabweichung

- Herstellerangabe zum Messgerät
- Abschätzung über Intuition, Erfahrung, …
- Statistische Betrachtung, wenn wiederholte Messung unter identischen Bedingungen möglich (techn. Replikate)
    - Standardabweichung
    - Standardfehler

Zwischenfazit

Angabe von Messergebnissen

- Messabweichung gibt signifikante Stelle vor
- Runden des Messwerts und der Messabweichung auf die gleiche signifikante Stelle
- Bei statistischer Betrachtung basierend auf dem Mittelwert $\bar{x}$ und der Standardabweichung $\tilde{s}(x)$, sowie unter Angabe des Vertrauensniveaus

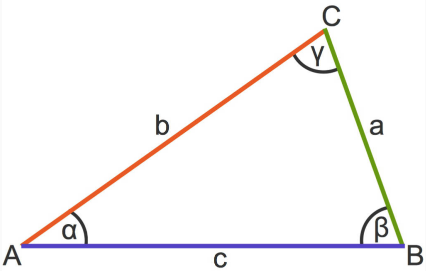
Vertrauensniveau	68,3%	95,4%	99,7%
Einzelmessung	$\bar{x} \pm \tilde{s}(x)$	$\bar{x} \pm 2\tilde{s}(x)$	$\bar{x} \pm 3\tilde{s}(x)$
Mittelwert	$\bar{x} \pm \frac{\tilde{s}(x)}{\sqrt{N}}$	$\bar{x} \pm 2\frac{\tilde{s}(x)}{\sqrt{N}}$	$\bar{x} \pm 3\frac{\tilde{s}(x)}{\sqrt{N}}$

Indirekte Messung

Kann man eine interessierende Größe nicht direkt messen, sondern nur indirekt über die Messung anderer Größen und deren Verrechnung, so stellt sich die Frage wie die Abweichungen berücksichtigt werden können.

Beispiel: Umfang eines Dreiecks
Der Umfang eines Dreiecks ist die Summe aller Kanten

$$U = a + b + c$$



U kann nicht direkt gemessen werden, sondern muss aus den Messwerten für a, b und c berechnet werden.
Die Messungen von a, b und c haben die zufälligen Unsicherheiten δ_a, δ_b und δ_c.

$\Rightarrow$ **Messunsicherheit** δ_U?

Größtfehlerabschätzung

Unter Einbeziehung der Unsicherheiten gilt:

$$U \pm \delta_U = a \pm \delta_a + b \pm \delta_b + c \pm \delta_c$$

$+$-Fall

$$U + \delta_U = a + \delta_a + b + \delta_b + c + \delta_c$$
$$\delta_U = \delta_a + \delta_b + \delta_c$$

$-$-Fall

$$U - \delta_U = a - \delta_a + b - \delta_b + c - \delta_c$$
$$\delta_U = \delta_a + \delta_b + \delta_c$$

Messunsicherheiten addieren sich bei der Addition/Subtraktion von Messergebnissen!

Größtfehlerabschätzung - Multiplikation/Division

Fläche eines Rechtecks

$$F \pm \delta_F = a \pm \delta_a \cdot b \pm \delta_b$$

Beispiel:

$$a = 100 \pm 1 \text{ cm}$$
$$b = 5 \pm 1\text{cm}$$
$$F_{min} = 99 \cdot 4 = 396 \text{ cm}^2$$
$$F_{max} = 101 \cdot 6 = 606 \text{ cm}^2$$
$$F_{best} = 100 \cdot 5 = 500 \text{ cm}^2$$

$$\Rightarrow F \approx 500 \pm 105$$

Relative Abweichung:

$$\frac{\delta_a}{a} = \frac{1}{100} = 0,01 \rightarrow \text{ genaue Messung}$$
$$\frac{\delta_b}{b} = \frac{1}{5} = 0,2 \qquad \rightarrow \text{grobe Messung}$$

Die große Unsicherheit der 2. Messung muss sich in der Gesamtunsicherheit niederschlagen.

Größtfehlerabschätzung - Multiplikation/Division

Messwertangabe mit relativer Abweichung
Beispiel:

$$\begin{aligned}
a &= 100 \cdot (1 \pm 0,01)\text{cm} \\
b &= 5 \cdot (1 \pm 0,2)\text{cm} \\
F_{max} &= a \cdot b = (100 \cdot (1 + 0,01)) \cdot (5 \cdot (1 + 0,2)) \\
&= 500 \cdot ((1 + 0,01) \cdot (1 + 0,2)) \\
&= 500 \cdot (1 + 0,01 + 0,2 + 0,002) = 500 \cdot (1 + 0,212)\ \text{cm}^2 \\
F_{min} &= a \cdot b = (100 \cdot (1 - 0,01)) \cdot (5 \cdot (1 - 0,2)) \\
&= 500 \cdot ((1 - 0,01) \cdot (1 - 0,2)) \\
&= 500 \cdot (1 - 0,01 - 0,2 + 0,002) = 500 \cdot (1 - 0,208)\ \text{cm}^2
\end{aligned}$$

$$\Rightarrow F \approx 500 \cdot (1 \pm 0,21)\ \text{cm}^2$$

Größtfehlerabschätzung - Multiplikation/Division

Allgemein:

$$x = p \cdot \left(1 \pm \frac{\delta_x}{|p|}\right), \qquad y = q \cdot \left(1 \pm \frac{\delta_y}{|q|}\right) \qquad z = x \cdot y$$

$$z = p \cdot \left(1 \pm \frac{\delta_x}{|p|}\right) \cdot q \cdot \left(1 \pm \frac{\delta_y}{|q|}\right)$$

$$z_{max} = pq \cdot \left(1 + \frac{\delta_x}{|p|}\right) \cdot \left(1 + \frac{\delta_y}{|q|}\right) = pq \cdot \left(1 + \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|} + \underbrace{\frac{\delta_x}{|p|}\frac{\delta_y}{|q|}}_{\text{sehr klein!}}\right)$$

$$z_{max} \approx pq \cdot \left(1 + \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}\right) \Rightarrow \text{analog: } z_{min} \approx pq \cdot \left(1 - \frac{\delta_x}{|p|} - \frac{\delta_y}{|q|}\right)$$

$$z = pq \cdot \left(1 \pm \left(\frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}\right)\right) \Rightarrow \frac{\delta_z}{|r|} = \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}$$

Bei Multiplikation und Division addieren sich die relativen Unsicherheiten!

Größtfehlerabschätzung - Multiplikation/Division

Allgemein:

$$x = p \cdot \left(1 \pm \frac{\delta_x}{|p|}\right), \qquad y = q \cdot \left(1 \pm \frac{\delta_y}{|q|}\right) \qquad z = x \cdot y$$

$$z = p \cdot \left(1 \pm \frac{\delta_x}{|p|}\right) \cdot q \cdot \left(1 \pm \frac{\delta_y}{|q|}\right)$$

$$z_{max} = pq \cdot \left(1 + \frac{\delta_x}{|p|}\right) \cdot \left(1 + \frac{\delta_y}{|q|}\right) = pq \cdot \left(1 + \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|} + \underbrace{\frac{\delta_x}{|p|}\frac{\delta_y}{|q|}}_{\text{sehr klein!}}\right)$$

$$z_{max} \approx pq \cdot \left(1 + \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}\right) \Rightarrow \text{analog: } z_{min} \approx pq \cdot \left(1 - \frac{\delta_x}{|p|} - \frac{\delta_y}{|q|}\right)$$

$$z = pq \cdot \left(1 \pm \left(\frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}\right)\right) \Rightarrow \frac{\delta_z}{|r|} = \frac{\delta_x}{|p|} + \frac{\delta_y}{|q|}$$

Bei Multiplikation und Division addieren sich die relativen Unsicherheiten!

Größtfehlerabschätzung - Zusammenfassung

Funktion	Gesamtunsicherheit	Kommentar
$q = C \cdot x$	$\delta_q = C \cdot \delta_x$	Dicke von 100 Blättern: Blatt $B = 0,012 \pm 0,002$ cm Stapel $S = 1,2 \pm 0,2$ cm
$q = x + y$ $q = x - y$	$\delta_q = \delta_x + \delta_y$	Bei Addition/Subtraktion addieren sich die **absoluten** Abweichungen!
$q = x \cdot y$ $q = x/y$	$\left\lvert\frac{\delta_q}{\lvert q\rvert}\right\rvert = \left\lvert\frac{\delta_x}{\lvert x\rvert}\right\rvert + \left\lvert\frac{\delta_y}{\lvert y\rvert}\right\rvert$	Bei Multiplikation/Division addieren sich die **relativen** Abweichungen!
$q = x^n y^m$	$\left\lvert\frac{\delta_q}{\lvert q\rvert}\right\rvert = \left\lvert n\frac{\delta_x}{\lvert x\rvert}\right\rvert + \left\lvert m\frac{\delta_y}{\lvert y\rvert}\right\rvert$	

Kompensation der Unsicherheiten
Wenn

- Messgrößen **unabhängig** voneinander gemessen werden und
- die Abweichungen **zufällig** sind

dann können sie sich gegenseitig kompensieren.

Für eine interessierende Größe $q(x, y, \ldots, z)$ kann die Standardabweichung folgendermaßen berechnet werden:

$$\delta_q = \sqrt{\left(\frac{dq}{dx}\delta_x\right)^2 + \left(\frac{dq}{dy}\delta_y\right)^2 + \ldots + \left(\frac{dq}{dz}\delta_z\right)^2}$$

Gaußsche Fehlerfortpflanzung

Für bestimmte Arten von Funktionen für q kann dies stark vereinfacht werden.

Funktion	Nach Gauß	Zum Vergleich: Größtfehler
$q = C \cdot x$	$\delta_q = C \cdot \delta_x$	$\delta_q = C \cdot \delta_x$
$q = x + / - y$	$\delta_q = \sqrt{(\delta_x)^2 + (\delta_y)^2}$	$\delta_q = \delta_x + \delta_y$
$q = x \cdot y$	$\frac{\delta_q}{q} = \sqrt{\left(\frac{\delta_x}{x}\right)^2 + \left(\frac{\delta_y}{y}\right)^2}$	$\left\lvert\frac{\delta_q}{q}\right\rvert = \left\lvert\frac{\delta_x}{x}\right\rvert + \left\lvert\frac{\delta_y}{y}\right\rvert$
$q = x^n y^m$	$\frac{\delta_q}{q} = \sqrt{\left(n\frac{\delta_x}{x}\right)^2 + \left(m\frac{\delta_y}{y}\right)^2}$	$\left\lvert\frac{\delta_q}{q}\right\rvert = \left\lvert n\frac{\delta_x}{x}\right\rvert + \left\lvert m\frac{\delta_y}{y}\right\rvert$

Beispiel: Bestimmung der spezifischen Wachstumsrate im Batch-Prozess

$$x_1 = 1,00 \pm 0,01 \text{ g/l},\ x_2 = 1,50 \pm 0,015 \text{ g/l},\ \Delta t = 120 \pm 1 \text{ min}$$

$$\mu = \frac{dx}{\Delta t} \cdot \frac{1}{\bar{x}} = \frac{x_2 - x_1}{\Delta t} \cdot \frac{1}{\frac{x_1 + x_2}{2}} = \frac{x_2 - x_1}{\Delta t} \cdot \frac{2}{x_1 + x_2} = \frac{0,5}{120} \cdot \frac{2}{2,5} = \frac{1}{300} \text{ min}^{-1}$$

Vereinfachung nicht möglich, daher über die partiellen Ableitungen:

$$\delta_\mu = \sqrt{\left(\frac{d\mu}{dx_1}\delta_{x_1}\right)^2 + \left(\frac{d\mu}{dx_2}\delta_{x_2}\right)^2 + \left(\frac{d\mu}{d\Delta t}\delta_{\Delta t}\right)^2}$$

Gaußsche Fehlerfortpflanzung

Fortsetzung

$$
\begin{aligned}
\frac{d\mu}{dx_1} &= -\frac{4x_2}{\Delta t(x_1+x_2)^2} = -\frac{4\cdot 1.5}{120\cdot 2,5^2} = -\frac{1}{125} \\
\frac{d\mu}{dx_2} &= \frac{4x_1}{\Delta t(x_1+x_2)^2} = \frac{4\cdot 1}{120\cdot 2,5^2} = \frac{2}{375} \\
\frac{d\mu}{d\Delta t} &= \frac{2(x_2-x_1)}{x_2+x_1)\Delta t^2} = \frac{2\cdot 0,5}{2,5\cdot 120^2} = \frac{1}{36.000} \\
\delta_\mu &= \sqrt{\left(\frac{1}{125}\cdot 0.01\right)^2 + \left(\frac{2}{375}0.015\right)^2 + \left(\frac{1}{36.000}\cdot 1\right)^2} \\
&= 0.0001164972 \rightsquigarrow 0.00012 \\
\Rightarrow \mu &= 0.00333 \pm 0.00012 \text{ min}^{-1} = 0.00333\cdot(1\pm 0.036)
\end{aligned}
$$

Die Abweichungen der Einzelmessungen waren $\leq 1\%$. Für den daraus berechneten Wert liegt die Abweichung bei 3,6%!

SynBioTeam