LU01b – Standardabweichung und die 68–95–99.7-Regel

Ziel: Du kannst die Standardabweichung als Streuungsmass erklären, die 68–95–99.7-Regel anwenden und Werte mithilfe des **z-Scores standardisieren**.

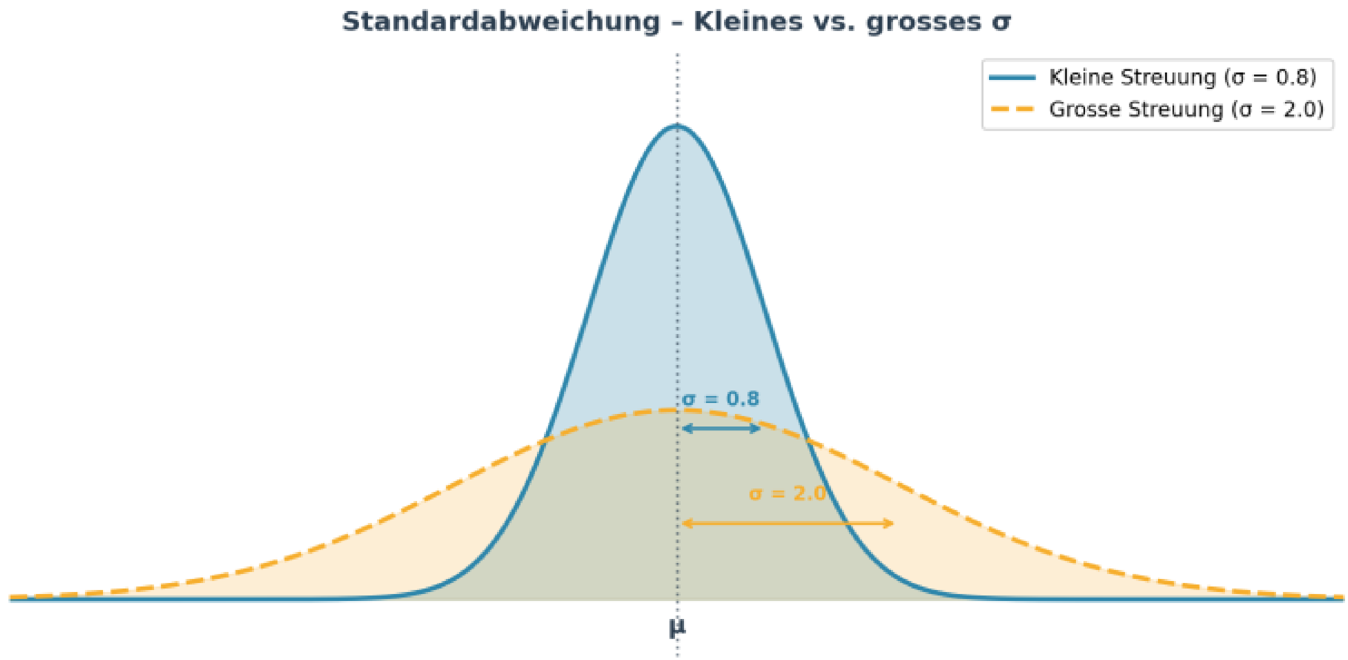
Kurztheorie (Merksätze)

- **σ (Sigma)** = Standardabweichung → misst, wie stark Werte um den Mittelwert streuen.
- **68 %** aller Werte liegen innerhalb von **± 1σ** vom Mittelwert.
- **95 %** aller Werte liegen innerhalb von **± 2σ** vom Mittelwert.
- **99.7 %** aller Werte liegen innerhalb von **± 3σ** vom Mittelwert.
- Der **z-Score** gibt an, wie viele Standardabweichungen ein Wert vom Mittelwert entfernt ist: \(z = \frac{x - \mu}{\sigma}\)

---

1) Was ist die Standardabweichung?

Die **Standardabweichung σ** (Sigma) misst, wie stark die Werte eines Datensatzes um den **Mittelwert µ** herum streuen.



Sigma	Bedeutung	Kurvenform
Kleines σ	Werte liegen eng beieinander	Schmale, hohe Glocke
Grosses σ	Werte sind weit gestreut	Breite, flache Glocke

> **Merksatz:** Ein kleines σ bedeutet **Präzision** (z.B. eine gut kalibrierte Maschine). Ein grosses σ

> bedeutet **grosse Streuung** (z.B. unterschiedliche Testergebnisse in einer heterogenen Klasse).

Die Formel für die Standardabweichung lautet:

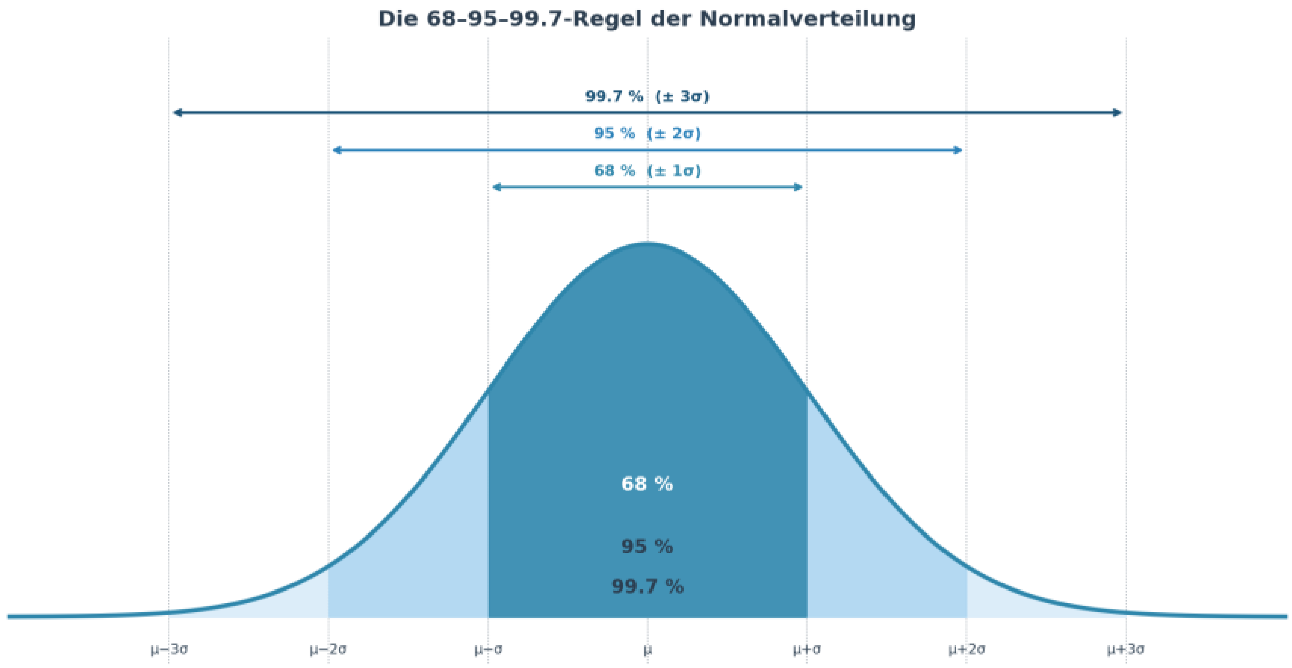
$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n}(x_i - \mu)^2}$$

> **Hinweis:** In der Praxis verwenden wir oft einen Taschenrechner oder Software (Excel, Python) für die Berechnung.

---

2) Die 68–95–99.7-Regel

Bei einer normalverteilten Datenmenge gilt immer folgende **Faustregel**:



Bereich	Anteil der Werte	Bedeutung
$\mu \pm 1\sigma$	**68 %**	Wahrscheinlich – 68 von 100 Werten fallen hier rein
$\mu \pm 2\sigma$	**95 %**	Sehr wahrscheinlich – 95 von 100
$\mu \pm 3\sigma$	**99.7 %**	Fast sicher – 997 von 1000

Diese Regel wird auch **Empirische Regel** genannt und ist eines der wichtigsten Werkzeuge der Statistik.

2.1 Anwendungsbeispiel: Körpergrössen

In einer Schule sind **95 % der Schülerinnen und Schüler zwischen 1.1 m und 1.7 m gross**.

Angenommen, die Daten sind normalverteilt – gesucht: Mittelwert μ und Standardabweichung σ.

Schritt 1 – Mittelwert berechnen: $$\mu = \frac{1.1 + 1.7}{2} = 1.4 \text{ m}$$

Schritt 2 – Standardabweichung berechnen:

95 % entsprechen ± 2σ → der Gesamtbereich von 1.1 bis 1.7 m umfasst 4σ: $$\sigma = \frac{1.7 - 1.1}{4} = \frac{0.6}{4} = 0.15 \text{ m}$$

Kontrolle:

- 1σ-Bereich: 1.25 m – 1.55 m → ca. **68 %** der Schüler
- 2σ-Bereich: 1.10 m – 1.70 m → ca. **95 %** der Schüler ✓
- 3σ-Bereich: 0.95 m – 1.85 m → ca. **99.7 %** der Schüler

---

3) Der z-Score (Standardwert)

Der **z-Score** gibt an, wie viele Standardabweichungen ein bestimmter Wert x vom Mittelwert μ entfernt ist.

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

Dabei gilt:

Symbol	Bedeutung
z	z-Score (Standardwert)
x	der zu standardisierende Wert
μ	Mittelwert der Verteilung
σ	Standardabweichung

3.1 Vorgehen

1. Mittelwert subtrahieren: $x - \mu$
2. Durch Standardabweichung dividieren: $\div\; \sigma$

> **Ergebnis:** Ein z-Score von **+2** bedeutet: der Wert liegt 2 Standardabweichungen **über** dem Mittelwert. Ein z-Score von **−1** bedeutet: 1 Standardabweichung **unter** dem Mittelwert.

3.2 Beispiel: Reisezeiten

Eine Umfrage über tägliche Reisezeiten (in Minuten) ergab:

```
26, 33, 65, 28, 34, 55, 25, 44, 50, 36, 26, 37, 43, 62, 35, 38, 45, 32, 28, 34
```

- Mittelwert: $\mu = 38.8$ Minuten
- Standardabweichung: $\sigma = 11.4$ Minuten

Die ersten drei Werte standardisiert:

Originalwert x	Berechnung	z-Score	Interpretation
26	$\frac{26 - 38.8}{11.4}$	**−1.12**	1.12σ unter dem Mittelwert

Originalwert x	Berechnung	z-Score	Interpretation
33	$\frac{33 - 38.8}{11.4}$	**−0.51**	0.51σ unter dem Mittelwert
65	$\frac{65 - 38.8}{11.4}$	**+2.30**	2.30σ über dem Mittelwert

3.3 Wozu dient der z-Score?

- **Vergleichbarkeit:** Werte aus verschiedenen Verteilungen können direkt verglichen werden.
- **Tabellen:** Mit einer einzigen Standardnormaltabelle lassen sich Wahrscheinlichkeiten für beliebige Normalverteilungen ablesen.
- **Ausreisser erkennen:** Werte mit |z| > 3 gelten als statistisch ungewöhnlich.

3.4 Praxisbeispiel: Faire Benotung

Ein Lehrer bewertet einen Test (max. 60 Punkte):

```
20, 15, 26, 32, 18, 28, 35, 14, 26, 22, 17
```

Der Test war zu schwer – die meisten Schüler bestehen nicht. Der Lehrer standardisiert die Noten und erklärt alle z-Scores unter −1 als „ungenügend„:

- $\mu = 23$, $\sigma = 6.6$
- z-Scores: −0.45, **−1.21**, 0.45, 1.36, −0.76, 0.76, 1.82, **−1.36**, 0.45, −0.15, −0.91

Nur **2 Schüler** (fett) gelten als ungenügend → deutlich fairer!

---

4) Prozentzonen der Standardnormalverteilung

Die Standardnormalverteilung (μ = 0, σ = 1) lässt sich in Zonen aufteilen:

Zone	Anteil in diesem Bereich	Kumulierter Anteil (ab −∞)
unter −3σ	0.1 %	0.1 %
−3σ bis −2σ	2.1 %	2.3 %
−2σ bis −1σ	13.6 %	15.9 %
−1σ bis 0	34.1 %	50.0 %
0 bis +1σ	34.1 %	84.1 %
+1σ bis +2σ	13.6 %	97.7 %
+2σ bis +3σ	2.1 %	99.7 %
über +3σ	0.1 %	99.9 %

Lesbeispiel

Dein Testergebnis liegt **0.5 Standardabweichungen über dem Durchschnitt** (z = +0.5). Wie viele Personen schnitten schlechter ab?

- Anteil unter dem Mittelwert ($z < 0$): **50 %**
- Anteil zwischen 0 und +0.5σ: **19.1 %**
- Total: **69.1 %** schnitten schlechter ab als du.

---

5) Praxisbeispiel: Zuckerbeutel-Produktion

Eine Abfüllmaschine hat: $\mu = 1010$ g, $\sigma = 20$ g.

Problem: Etwa 31 % der Beutel enthalten weniger als 1000 g!

Ziel: 1000 g soll bei **−2.5σ** liegen (dann nur noch 0.6 % zu leichte Beutel).

Option A – Mittelwert erhöhen: $$ \mu_{neu} = 1000 + 2.5 \times 20 = 1050 \text{ g} $$

Option B – Streuung verringern: $$ \sigma_{neu} = \frac{10}{2.5} = 4 \text{ g} $$

> **Fazit:** Mit der 68–95–99.7-Regel und dem z-Score können Produktionsprozesse **quantitativ gesteuert** werden – entweder durch Anpassen des Mittelwerts oder durch Verringern der Streuung.

Verständnisfragen

1. Ein Datensatz hat $\mu = 50$ und $\sigma = 10$. Welcher Anteil der Werte liegt zwischen **30 und 70**?
2. Berechne den z-Score für $x = 75$, wenn $\mu = 60$ und $\sigma = 8$.
3. Was bedeutet ein z-Score von **−2.5** anschaulich?
4. Warum ist die Standardisierung nützlich, wenn man Werte aus **verschiedenen Verteilungen** vergleichen will?
5. Eine Maschine produziert Teile mit $\mu = 100$ mm und $\sigma = 2$ mm. Wie gross ist der Anteil der Teile, die **kürzer als 96 mm** sind?

Bearbeitet nach mathsisfun.com