

Datenverarbeitung im P1-Praktikum

Roger Wolf

24. Oktober 2024

Vorwort

- **Grundlagen zur Datenauswertung** im P1:
 - Statistische & systematische Unsicherheiten
 - Fehlerfortpflanzung
 - Darstellung von Messergebnissen
 - Verträglichkeit von Messergebnissen
 - Parameterschätzung
 - Verträglichkeit von Modellen
- **Vertiefende Ausführungen** (für diejenigen, die es wissen wollen)
 - Freitag 25.10. 16:00 – 17:30 (SR 9-1)
- **Schriftlich und im eigenen Tempo** durch das:
 - Diese Folien auf den Webseiten des P1 [[hier](#)]
 - Anleitung und Hinweise zum Versuch Datenverarbeitung [[hier](#)]
 - Textdokumente und Jupyter-notebooks zum ausprobieren [[hier](#)]

Was macht wissenschaftliches Messen aus?



Was macht wissenschaftliches Messen aus?

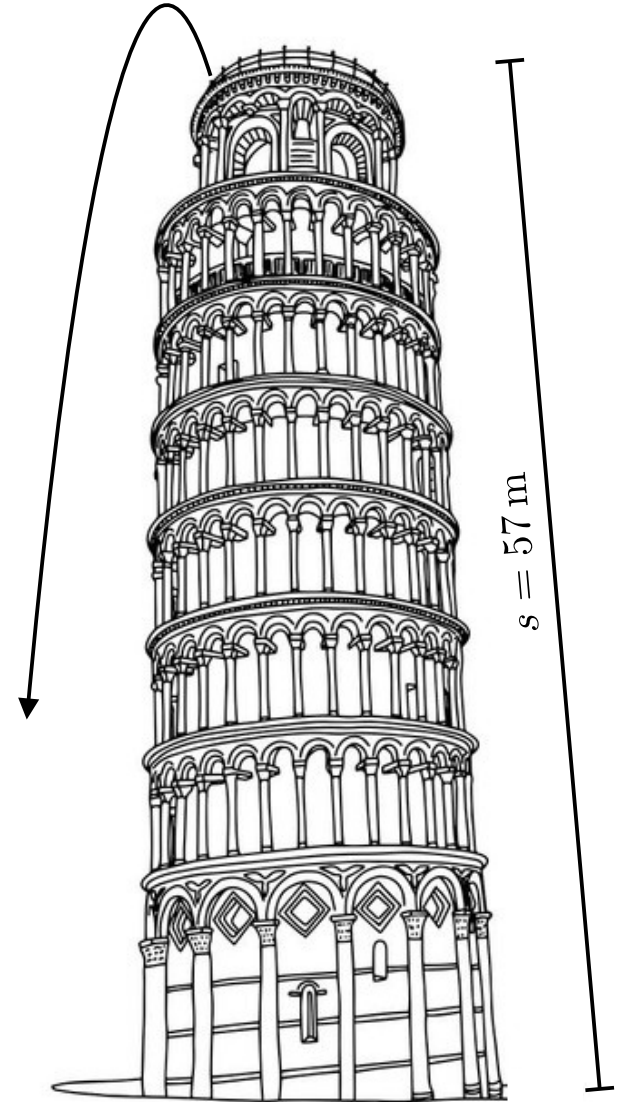


... Reproduzierbarkeit

Messung der Erdbeschleunigung g

$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2}$$

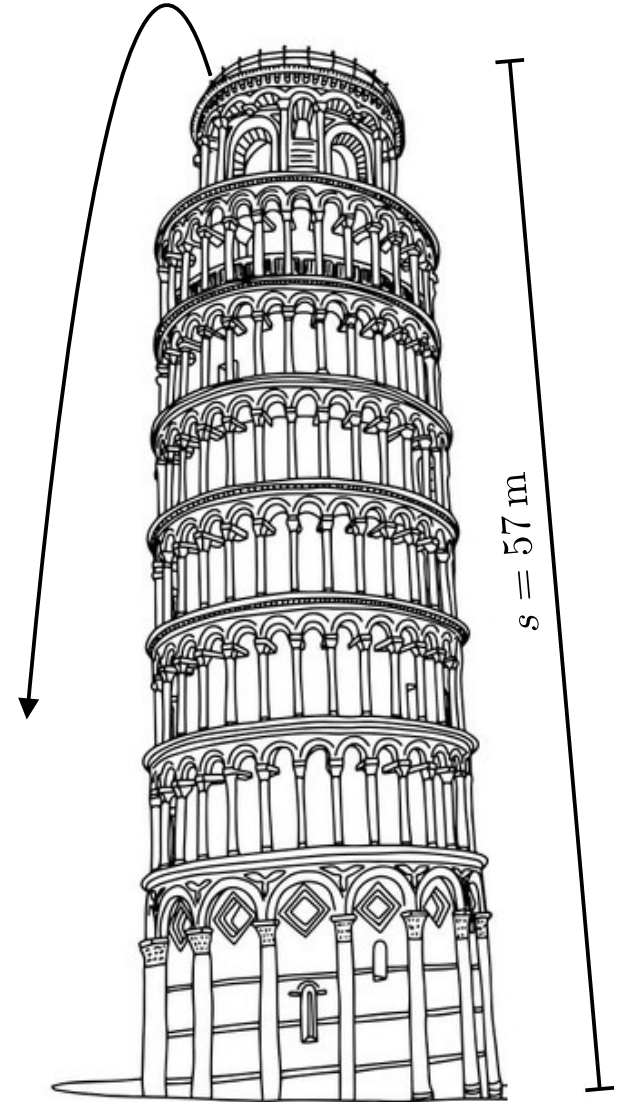
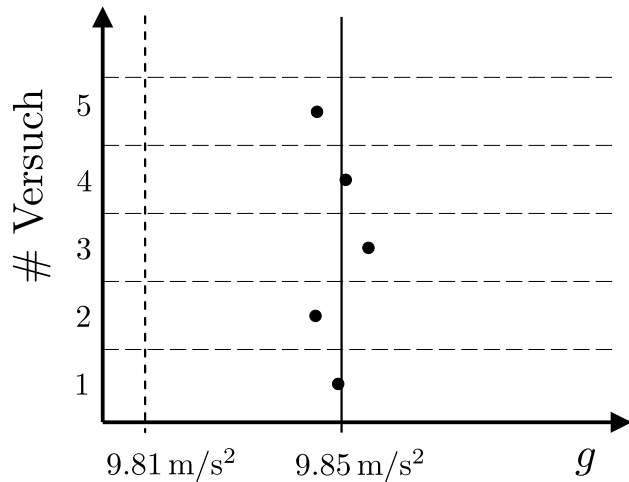
$$g(\ell, t) = \frac{2 \cdot 57 \text{ m}}{(3.40 \text{ s})^2} = 9.86 \text{ m/s}^2$$



Messung der Erdbeschleunigung g

$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2}$$

$$g(\ell, t) = \frac{2 \cdot 57 \text{ m}}{(3.40 \text{ s})^2} = 9.86 \text{ m/s}^2$$



Messung und Wahrheit

- **Wissenschaftliches Experiment** → nur innerhalb bestimmter Grenzen reproduzierbar (→ **Unsicherheiten**)
- Konfidenz- oder Vertrauensintervall
- Ohne zusätzliche Angabe eines Vertrauensintervalls ist jede wissenschaftliche Messung ...

Messung und Wahrheit

- **Wissenschaftliches Experiment** → nur innerhalb bestimmter Grenzen reproduzierbar (→ **Unsicherheiten**)
- Konfidenz- oder Vertrauensintervall
- Ohne zusätzliche Angabe eines Vertrauensintervalls ist jede wissenschaftliche Messung ...

W E R T L O S !

Messung und Wahrheit

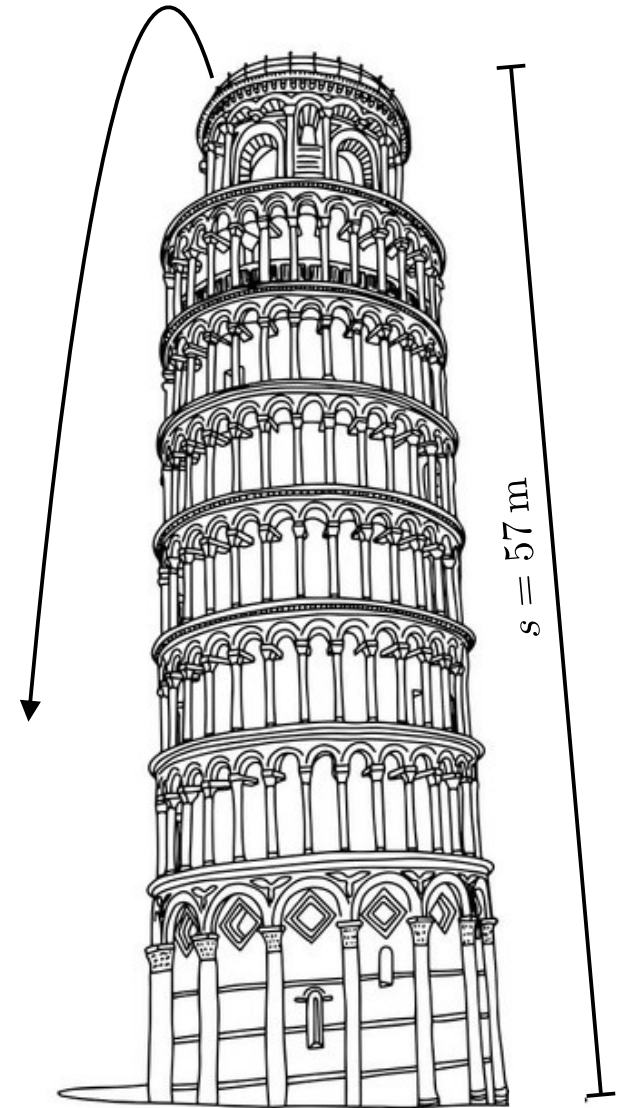
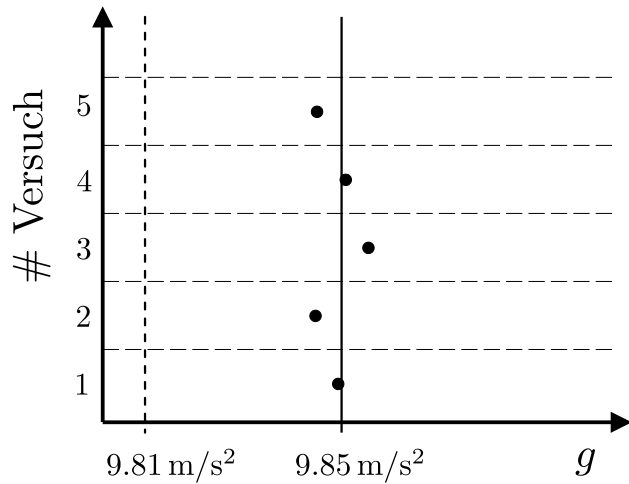
- **Wissenschaftliches Experiment** → nur innerhalb bestimmter Grenzen reproduzierbar (→ **Unsicherheiten**)
- Konfidenz- oder Vertrauensintervall
- Ohne zusätzliche Angabe eines Vertrauensintervalls ist jede wissenschaftliche Messung ...

**Kein Messwert ohne Angabe
eines Konfidenzintervalls**



Statistisch, systematisch, falsch!

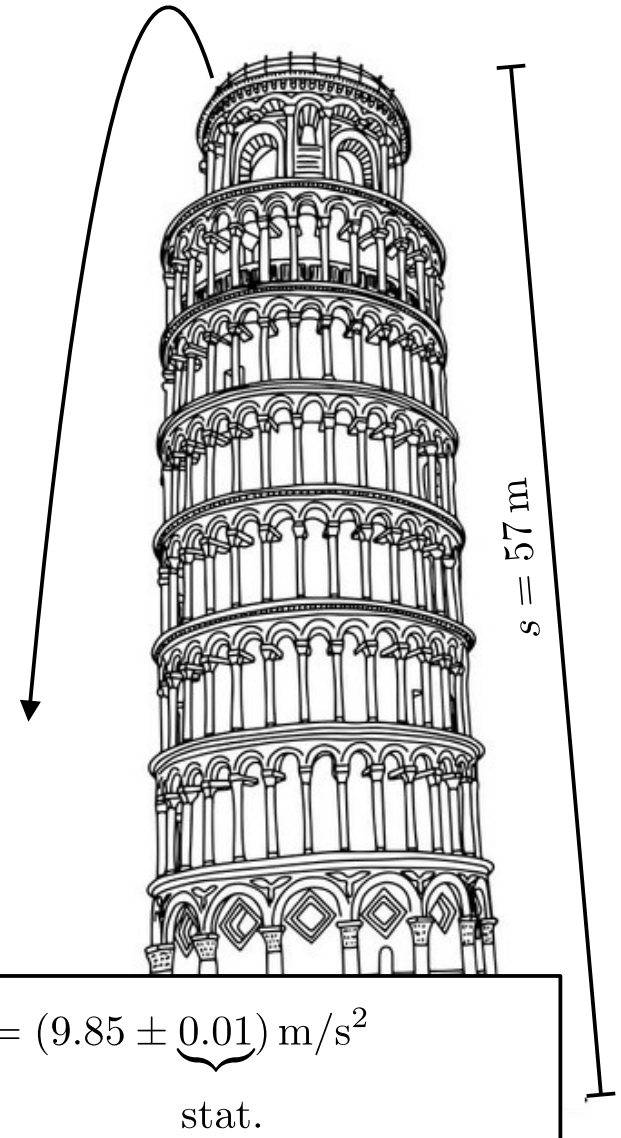
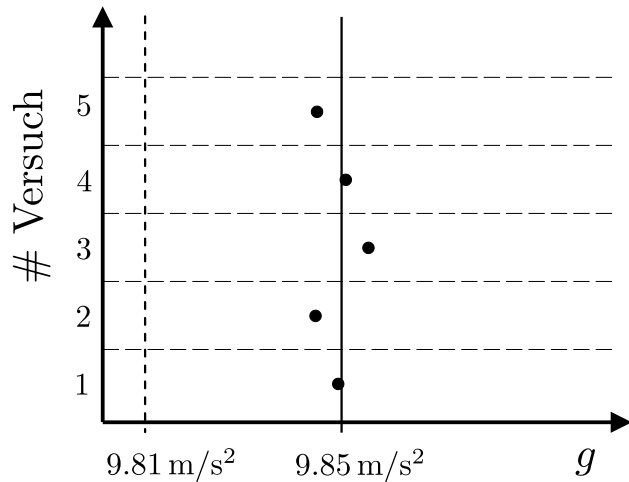
$$l = \frac{1}{2} g t^2 \quad g = \frac{2l}{t^2}$$



Statistisch, systematisch, falsch!

$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2}$$

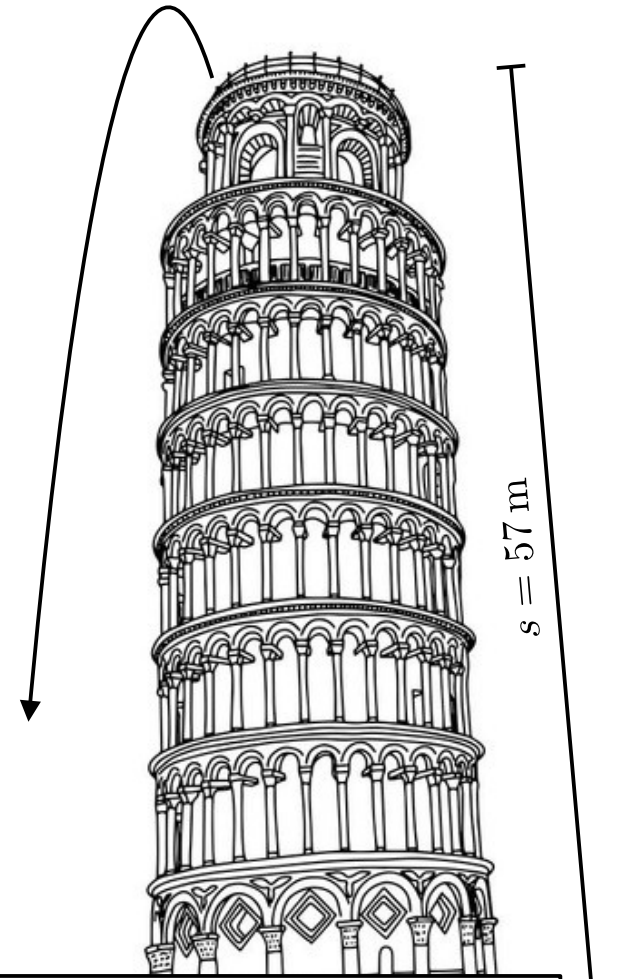
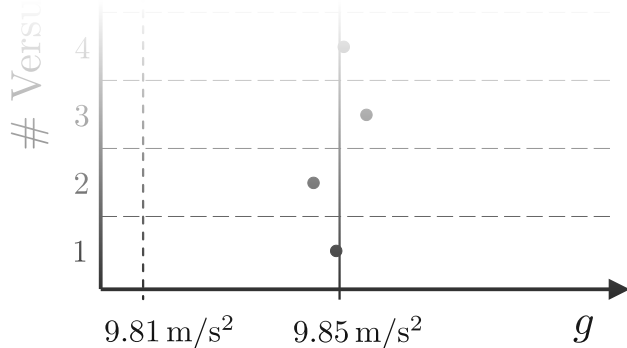
- Reaktionszeit beim Drücken der Stoppuhr unterliegt dem Zufall → **statistische Unsicherheit**.



Statistisch, systematisch, falsch!

$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2}$$

- Reaktionszeit beim Drücken der Stoppuhr unterliegt dem Zufall → **statistische Unsicherheit**.
- Turm vielleicht nicht 57 m, gewiss aber zwischen 54–60 m hoch → **systematische Unsicherheit**.

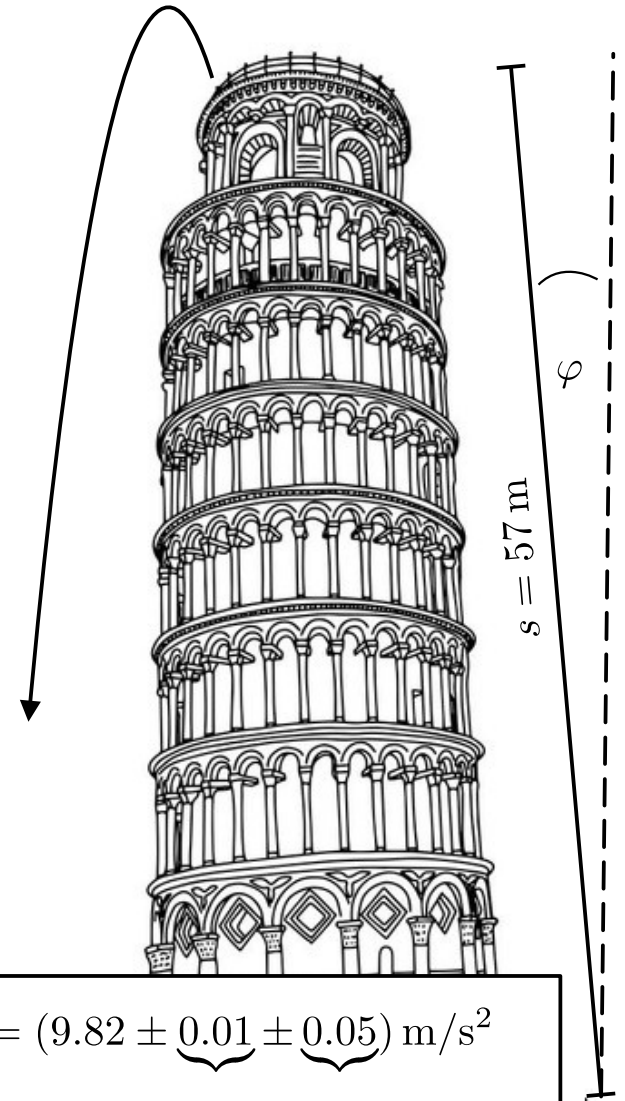
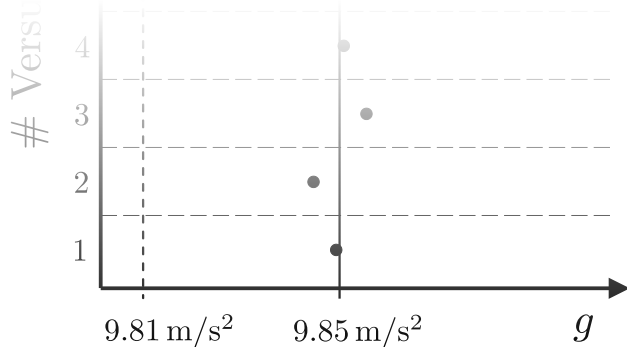


$$g = (9.85 \pm \underbrace{0.01}_{\text{stat.}} \pm \underbrace{0.05}_{\text{syst.}}) \text{ m/s}^2$$

Statistisch, systematisch, falsch!

$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2} \cos \varphi$$

- Reaktionszeit beim Drücken der Stoppuhr unterliegt dem Zufall → **statistische Unsicherheit**.
- Turm vielleicht nicht 57 m, gewiss aber zwischen 54–60 m hoch → **systematische Unsicherheit**.
- Turm steht schief! → **Fehler in meiner Annahme**.



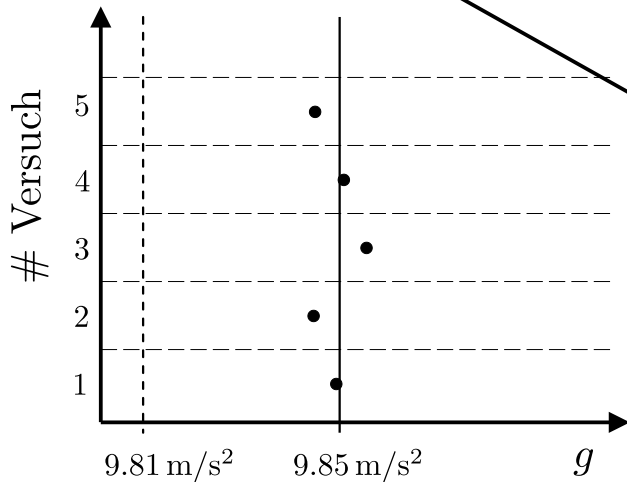
$$g = (9.82 \pm \underbrace{0.01}_{\text{stat.}} \pm \underbrace{0.05}_{\text{syst.}}) \text{ m/s}^2$$

Statistische Unsicherheiten

- Stichprobe der Länge n

Mittelwert der Stichprobe

$$\bar{g} = \frac{1}{n} \sum_{i=1}^n g_i$$



Varianz der Stichprobe

$$s_g^2 = \frac{1}{n-1} \sum_{i=1}^n (g_i - \bar{g})^2$$

$$\Delta \bar{g} = \sqrt{\frac{s_g^2}{n}}$$

$$g = (9.85 \pm \underbrace{0.01}_{\text{stat.}}) \text{ m/s}^2$$

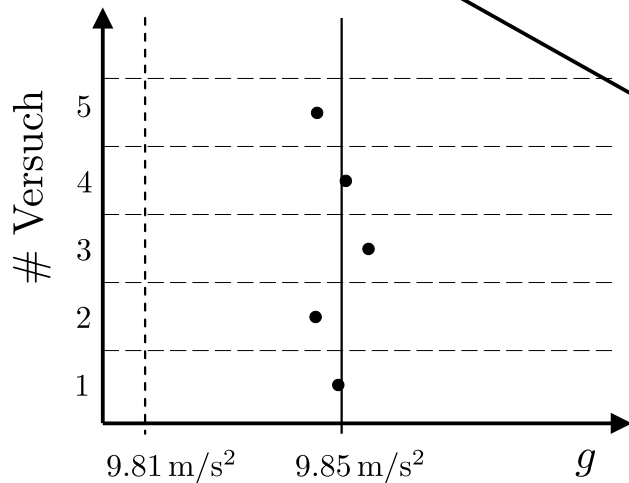
Statistische Unsicherheiten

- Stichprobe der Länge n

Unsicherheit auf \bar{g} geht mit $\frac{1}{\sqrt{n}}$

Mittelwert der Stichprobe

$$\bar{g} = \frac{1}{n} \sum_{i=1}^n g_i$$



Varianz der Stichprobe

$$s_g^2 = \frac{1}{n-1} \sum_{i=1}^n (g_i - \bar{g})^2$$

$$\Delta \bar{g} = \sqrt{\frac{s_g^2}{n}}$$

$$g = (9.85 \pm 0.01) \text{ m/s}^2$$

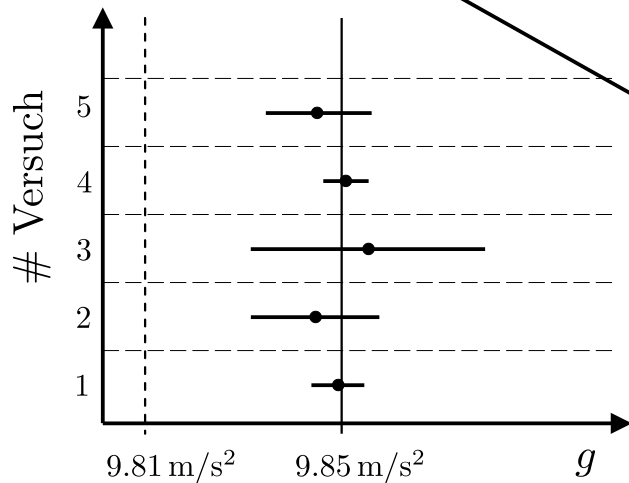
stat.

Statistische Unsicherheiten (continued)

- Stichprobe der Länge n
- Messungen **haben bereits Unsicherheiten?**

Gewichtetes Mittel der Stichprobe

$$\bar{g} = \frac{\sum w_i g_i}{\sum w_i}; \quad w_i = \frac{1}{\Delta g_i^2}$$



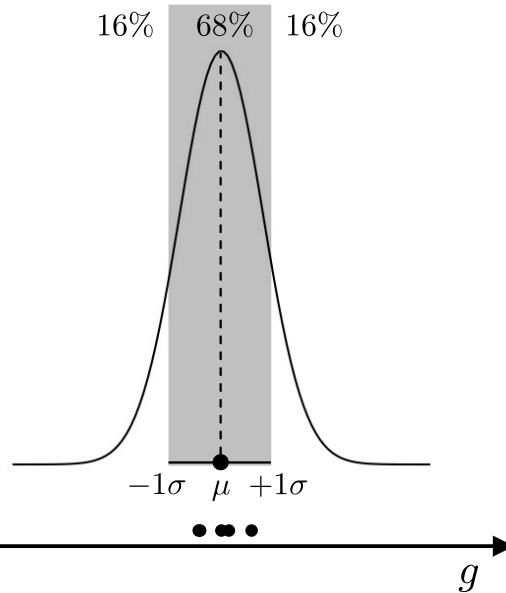
Varianz der Stichprobe

$$s_g^2 = \frac{1}{\sum w_i}$$

$$\Delta \bar{g} = \sqrt{s_g^2}$$

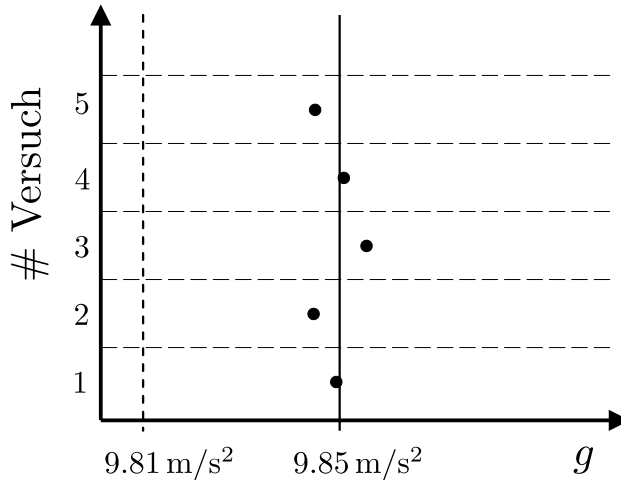
$$g = (9.85 \pm \underbrace{0.01}_{\text{stat.}}) \text{ m/s}^2$$

Normalverteilung



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

$x \simeq g; \quad \mu \simeq \bar{g}; \quad \sigma \simeq \Delta g$



Konfidenzintervalle

- $x \in [\mu - \sigma; \mu + \sigma]$ **68%** aller Messungen
- $x \in [\mu - 2\sigma; \mu + 2\sigma]$ **95%** aller Messungen
- $x \in [\mu - 3\sigma; \mu + 3\sigma]$ **99%** aller Messungen

Systematische Unsicherheiten

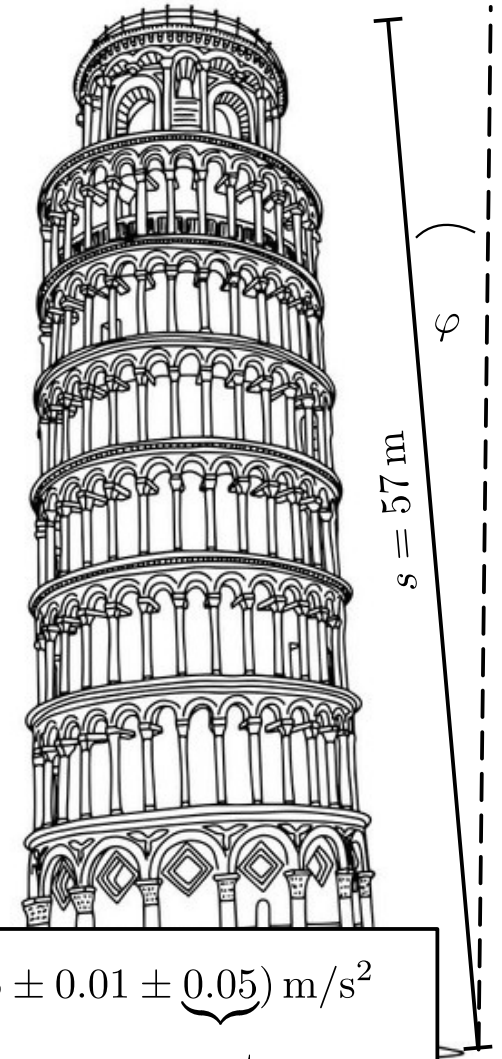
$$\ell = \frac{1}{2} g t^2 \quad g = \frac{2\ell}{t^2} \cos \varphi$$

- Hier: $\ell = (57 \pm 3) \text{ m}$
- Unsicherheit aufgrund **mangelnder Kenntnis äußerer Parameter**

Alle Parameter, die ich nicht selbst messe

Wie quantifiziere ich mangelnde Kenntnis?

$$g = (9.85 \pm 0.01 \pm \underbrace{0.05}_{\text{sys.}}) \text{ m/s}^2$$



Methode-1: Autorität anderer

METRA HIT 22 ... 26S/M Analog-Digital-Multimeter mit Signalgenerator

3-349-026-01
8/11.04

- METRA HIT 22/23/24/25/26**
- Präzisionsmultimeter (V, dB, Ω , F, Hz, °C/°F, V \rightarrow)
 - Auflösung: 10 μ V, 10 m Ω
 - Integrierte Quarzuhr für echtzeitbezogene MIN-MAX-Registrierung
 - Rechtecksignalgeneratorfunktionen
 - Infrarot-Datenschnittstelle
 - DKD-Kalibrierschein

- METRA HIT 23/24/25/26**
- Strommessung (10 A) direkt oder über Zangenstromwandler: ein Übersetzungsverhältnis von 1000:1 oder 10000:1 wird in der Anzeige berücksichtigt
 - METRA HIT 23S: 16 A-Messbereich (ungesichert) speziell für Messungen an Stromwandlern

- METRA HIT 22M/26M**
- Großer Messdatenspeicher für bis zu 100000 Messwerten
 - Echtzeitbezogener, quartzgesteuerter Datenlogger

- METRA HIT 25/26**
- Echteffektivwertmessung TRMS



Referenzbedingungen

Umgebungstemperatur	+23 °C \pm 2 K
Relative Feuchte	40 ... 60%
Frequenz der Messgröße	45 ... 65 Hz
Kurvenform der Messgröße	Sinus
Batteriespannung	3 V \pm 0,1 V

Technische Kennwerte

Messfunktion	Messbereich	Auflösung bei Messbereichsendwert		Eingangsimpedanz		Eigenabweichung der höchsten Auflösung bei Referenzbedingungen		Überlastbarkeit ⁷⁾		Messrate		
		30000 ¹⁾	3000	—	~ / ∞	\pm (... % v. M. + ... D)	\pm (... % v. M. + ... D)	Wert	Zeit	—	~ / ∞	~
V ⁴⁾	300 mV	10 μ V		> 20 M Ω	5 M Ω // < 50 pF	0,05 + 3 ¹⁾	0,5 + 30 (> 300 D)	1000 V DC AC eff Sinus	dauernd	50 ms (22M/ 26M: 1 ms)	0,5 s	1 s
	3 V	100 μ V		11 M Ω	5 M Ω // < 50 pF	0,05 + 3	0,2 + 30 (> 300 D)					
	30 V	1 mV		10 M Ω	5 M Ω // < 50 pF	0,05 + 3	0,2 + 30 (> 300 D)					
	300 V	10 mV		10 M Ω	5 M Ω // < 50 pF	0,05 + 3	0,2 + 30 (> 300 D)					
	1000 V	100 mV		10 M Ω	5 M Ω // < 50 pF	0,05 + 3	0,2 + 30 (> 300 D)					

- 1) Anzeige: 4% Stellen; für die Speicherung und Übertragung von Messwerten ist eine andere Auflösung und Abtastrate einstellbar im Menü rATE.
- 2) Stoppuhr: Format: mm:ss:h mit m=Minute, s=Sekunde und h=Hundertstelsekunde, max.: 99:59.9; nur über Tasten bedienbar
- 3) niedrigste messbare Frequenz bei sinusförmigem Messsignal symmetrisch zum Nullpunkt
- 4) METRA HIT 26S/M und 25S: Echte Effektivwertmessung TRMS
- 5) ohne 16 A-Sicherung

- 6) Anzeige bis max. 1,8 V, darüber Überlauf „OL“.
- 7) bei 0 ° ... + 40 °C
- 8) Werte < 100 Digit werden unterdrückt
- 9) 15 (20) ... 45 ... 65 Hz ... 20 (1) kHz Sinus. Einflüsse siehe Seite 4.
- 10) 12 A – 5 min, 16 A – 30 s, METRA HIT 23S: 16 A 10 min.
- 11) bei Funktion „Nullpunkteinstellung“ aktiv, Anzeige ZERO
- 11) die Amplitude der Eingangsspannung darf folgende Werte nicht unterschreiten/überschreiten:

Beispiel: **Datenblatt** im P1-Versuch Datenverarbeitung

Methode-2: Eigene Einschätzung

- Bona fides, **gesunder Menschenverstand**
- Anhaltspunkte:



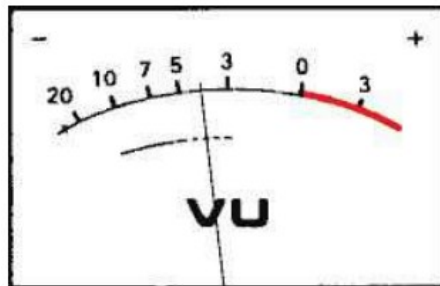
$$L = (82 \pm 0,5) \text{ mm}$$



$$U = (13.06 \pm 0.005) \text{ V}$$

⊕ Angaben aus Datenblatt

⊕ Anzeigefluktuationen



$$U = (4.0 \pm 0.2) \text{ V}$$

Methode-3: Syst. → stat.

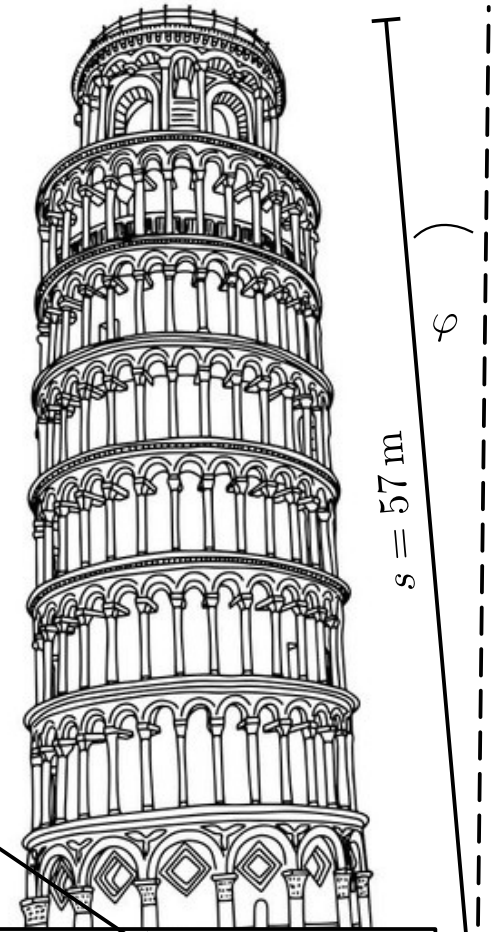
- Sie machen **aus einer systematischen eine statistische Unsicherheit!**
- Wie? – Geben Sie mir die Antwort...

Fehlerfortpflanzung

$$l = \frac{1}{2} g t^2 \quad g = \frac{2l}{t^2} \cos \varphi$$

- Wie komme ich von Δl auf Δg ?

$$\Delta g_l = \left| \frac{\partial}{\partial l} g \right| \Delta l = \frac{2}{t^2} \cos \varphi \Delta l$$



$$g = (9.85 \pm 0.01 \pm \underbrace{0.05}_{\text{syst.}}) \text{ m/s}^2$$

Fehlerfortpflanzung

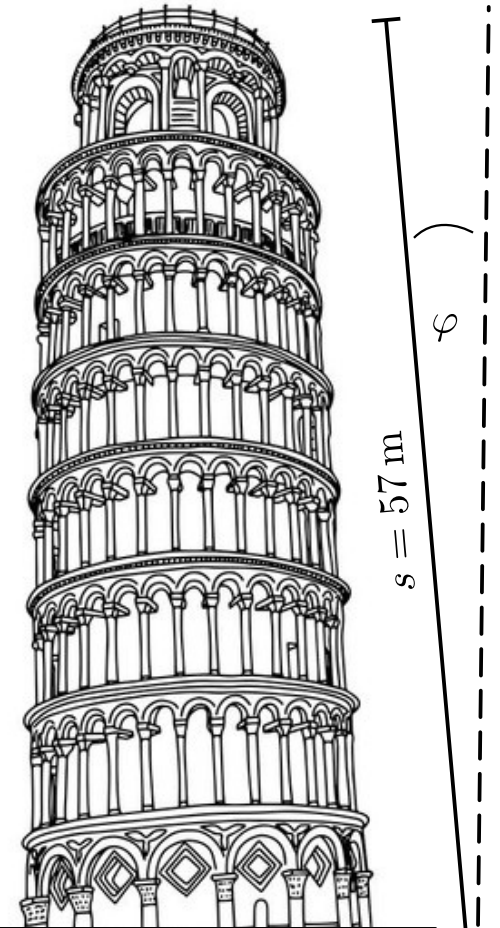
$$l = \frac{1}{2} g t^2 \quad g = \frac{2l}{t^2} \cos \varphi$$

- Wie komme ich von Δl auf Δg ?

$$\Delta g_l = \left| \frac{\partial}{\partial l} g \right| \Delta l = \frac{2}{t^2} \cos \varphi \Delta l$$

$$\Delta g_\varphi = \left| \frac{\partial}{\partial \varphi} g \right| \Delta \varphi = \frac{2l}{t^2} \sin \varphi \Delta \varphi$$

$$\Delta g_{\text{syst.}} = \sqrt{\Delta g_l^2 + \Delta g_\varphi^2}$$



$$g = (9.85 \pm 0.01 \pm \underbrace{0.05}_{\text{syst.}}) \text{ m/s}^2$$

Zwei einfache Beispiele

- **Summe/Differenz** zweier Größen:

$$f(x_0, x_1) = x_0 + x_1;$$

$$\Delta \hat{f} = \sqrt{\left(\frac{\partial f}{\partial x_0} \Delta \hat{x}_0\right)^2 + \left(\frac{\partial f}{\partial x_1} \Delta \hat{x}_1\right)^2} = \sqrt{\Delta \hat{x}_0^2 + \Delta \hat{x}_1^2},$$

“Fehler addieren sich quadratisch”

- **Produkt** zweier Größen:

$$f(x_0, x_1) = x_0 x_1;$$

$$\Delta \hat{f} = \sqrt{\left(\frac{\partial f}{\partial x_0} \Delta \hat{x}_0\right)^2 + \left(\frac{\partial f}{\partial x_1} \Delta \hat{x}_1\right)^2} = \sqrt{\hat{x}_1 \Delta \hat{x}_0^2 + \hat{x}_0 \Delta \hat{x}_1^2};$$

“Relative Fehler addieren sich quadratisch”

$$\frac{\Delta \hat{f}}{\hat{f}} = \sqrt{\left(\frac{\Delta \hat{x}_0}{\hat{x}_0}\right)^2 + \left(\frac{\Delta \hat{x}_1}{\hat{x}_1}\right)^2},$$

Darstellung von Messergebnissen

- **Im Text:**

$$g = (9.85 \pm 0.01 \pm 0.05) \text{ m/s}^2$$

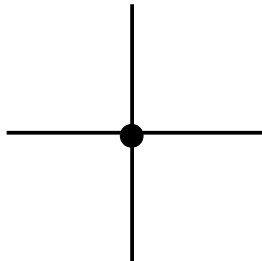


$$g = (9.8523685 \pm 0.012436 \pm 0.054758) \text{ m/s}^2$$



- **Signifikante (Nachkomma-)Stellen:** Nicht mehr Stellen, als als die erste von Null verschiedene führende Ziffer der größten Unsicherheit
- Bis dorthin wird **gerundet**
- Ausnahme: **Verwendung in weiteren Rechnungen**

- **Graphisch:**

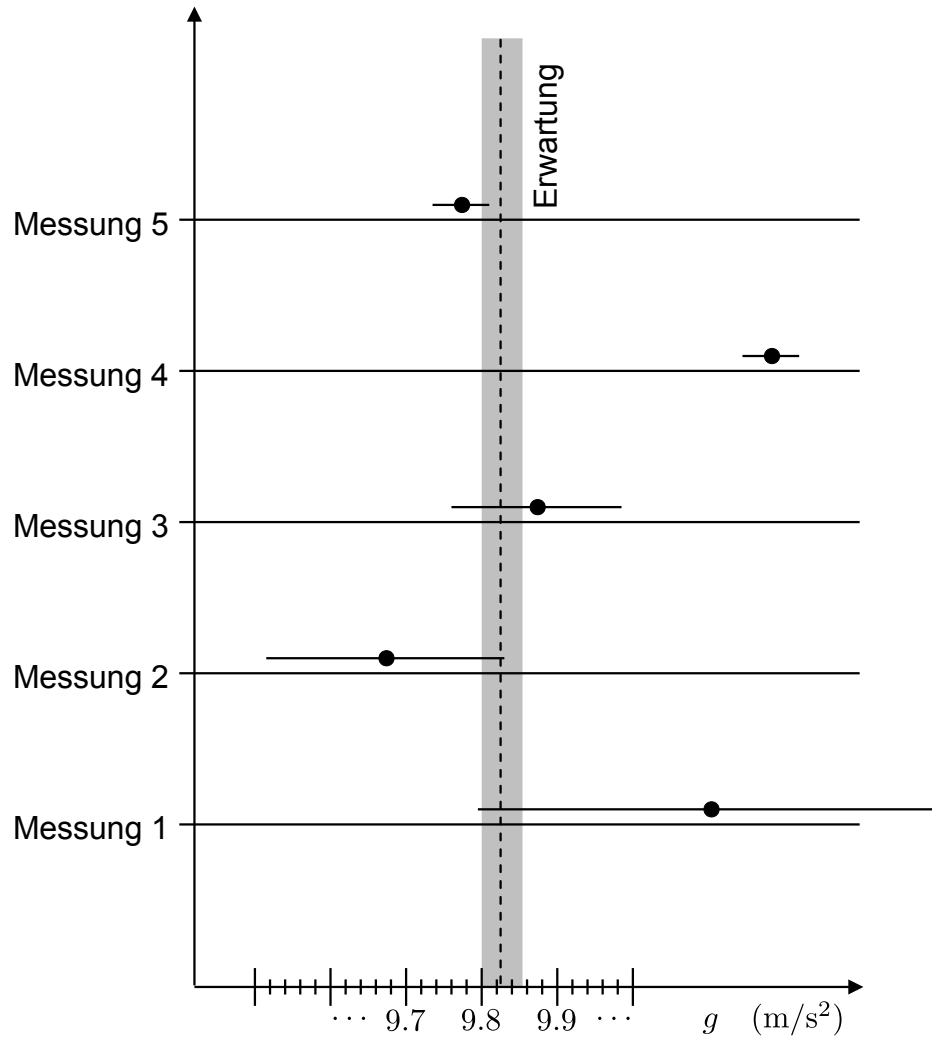


- Als **Punkt mit Fehlerbalken** (i.a. in x- UND y-Richtung)
- Länge des Balkens $\simeq \pm\sigma$
- Manchmal auch als Ellipse/Kontour

Interpretation von $g \pm \Delta g$

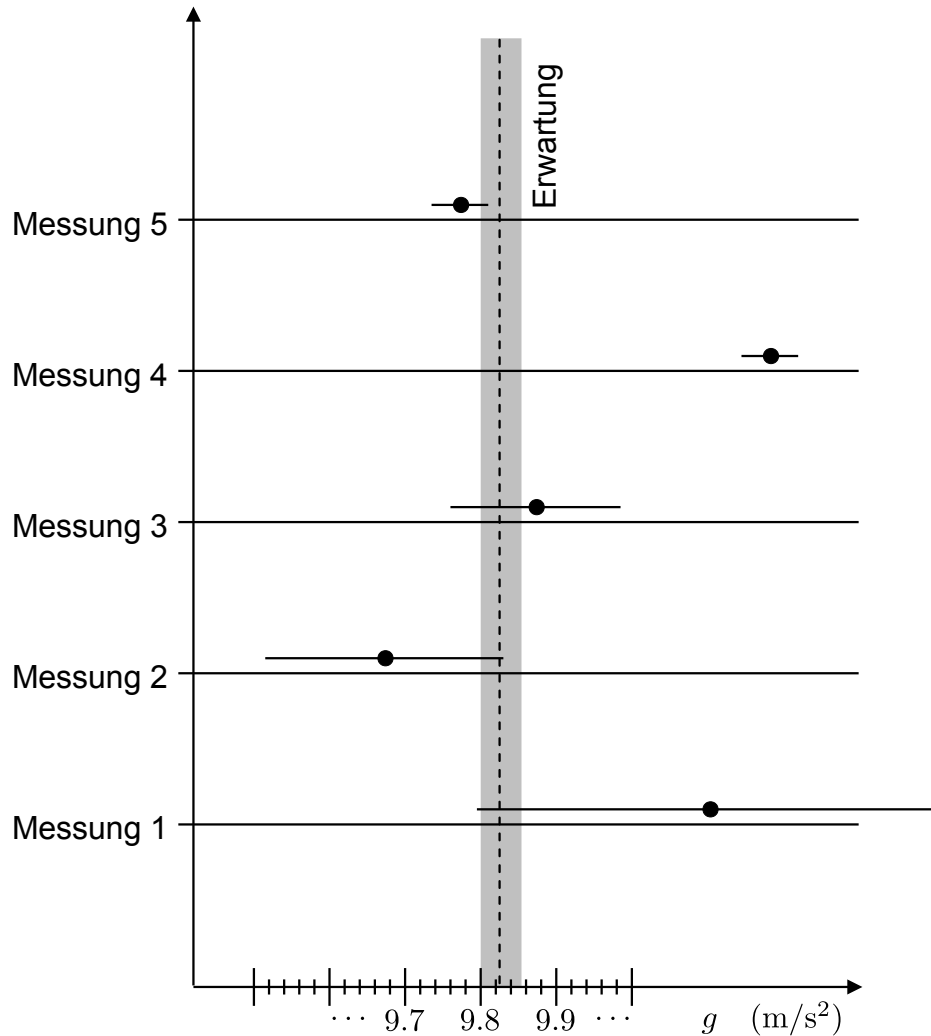
- Mit 68% Wahrscheinlichkeit liegt der Messpunkt **im Intervall** $[g - \Delta g; g + \Delta g]$ **um den wahren Wert** g (\rightarrow Inklusionsschluss)
- Mit 68% Wahrscheinlichkeit überdecken ein **Messwert** \hat{g} **den wahren Wert** g **mit den Vertrauensintervall** $[\hat{g} - \Delta g; \hat{g} + \Delta g]$ (\rightarrow Repräsentationsschluss)

Kompatibilität zweier Messungen



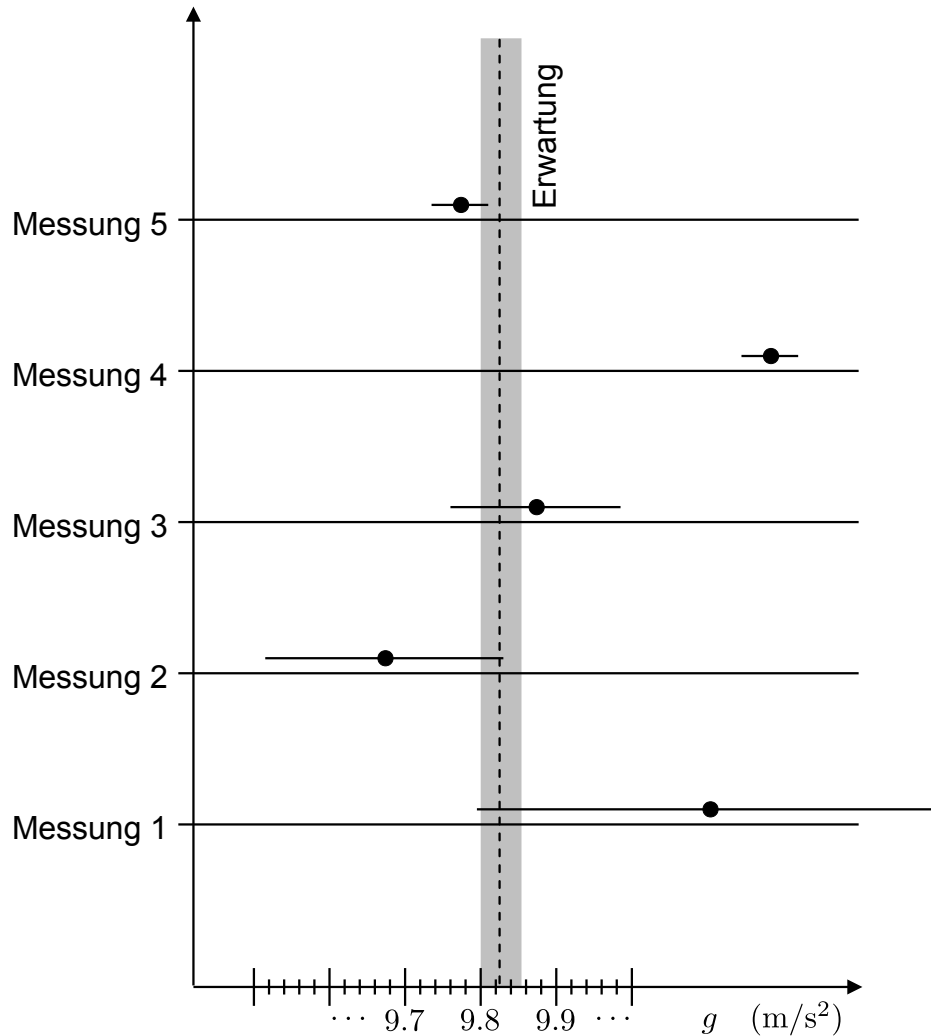
- Welche Messung(en) halten Sie für gut?

Kompatibilität zweier Messungen



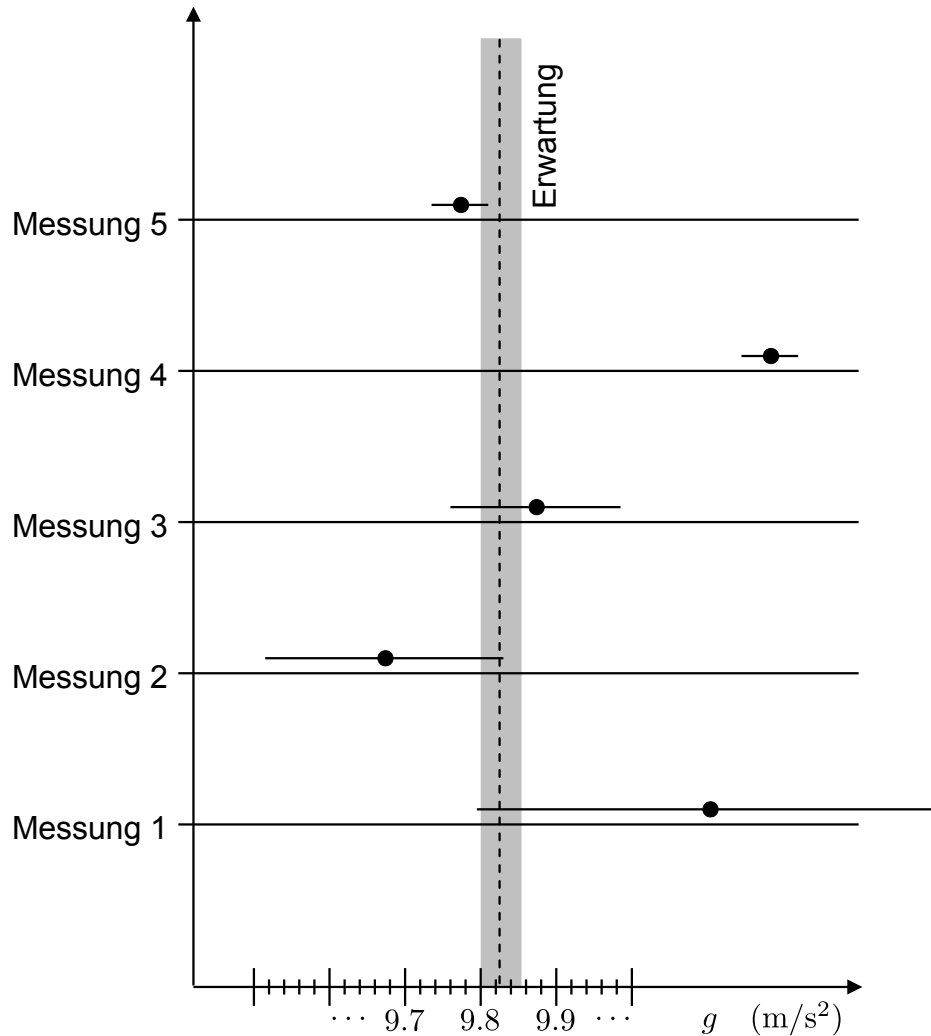
- Welche Messung(en) halten Sie für gut?
- Welche Messungen halten Sie für **kompatibel mit der Erwartung**?

Kompatibilität zweier Messungen



- Welche Messung(en) halten Sie für gut?
- Welche Messungen halten Sie für **kompatibel mit der Erwartung**?
- Welche Messungen halten Sie für **kompatibel miteinander**?

Kompatibilität zweier Messungen



- Welche Messung(en) halten Sie für gut?
- Welche Messungen halten Sie für **kompatibel mit der Erwartung**?
- Welche Messungen halten Sie für **kompatibel miteinander**?
- Wie beurteilen Sie also Kompatibilität?

Pull/Spannung

- Wir betrachten zwei unabhängige Messungen der gleichen Größe als **verträglich oder kompatibel**, wenn ihre Differenz mit 0 kompatibel ist:

$$f(x_0, x_1) = x_0 - x_1;$$

$$\Delta \hat{f} = \sqrt{\Delta \hat{x}_0^2 + \Delta \hat{x}_1^2},$$

- Wir bestimmen die Kompatibilität zweier Messungen **IMMER relativ zu den angegebenen Unsicherheiten**.
- Die Größe

$$\delta(\hat{x}_0, \hat{x}_1) \equiv \frac{\hat{x}_0 - \hat{x}_1}{\sqrt{\Delta \hat{x}_0^2 + \Delta \hat{x}_1^2}}$$

bezeichnen wir als Spannung/Pull zwischen den Messungen \hat{x}_0 und \hat{x}_1

Pull/Spannung (continued)

- Wenn Ihre Messung normalverteilt ist und Sie Ihre Unsicherheiten richtig als $x \in [\mu - \sigma; \mu + \sigma]$ Intervalle abgeschätzt haben ist die Größe

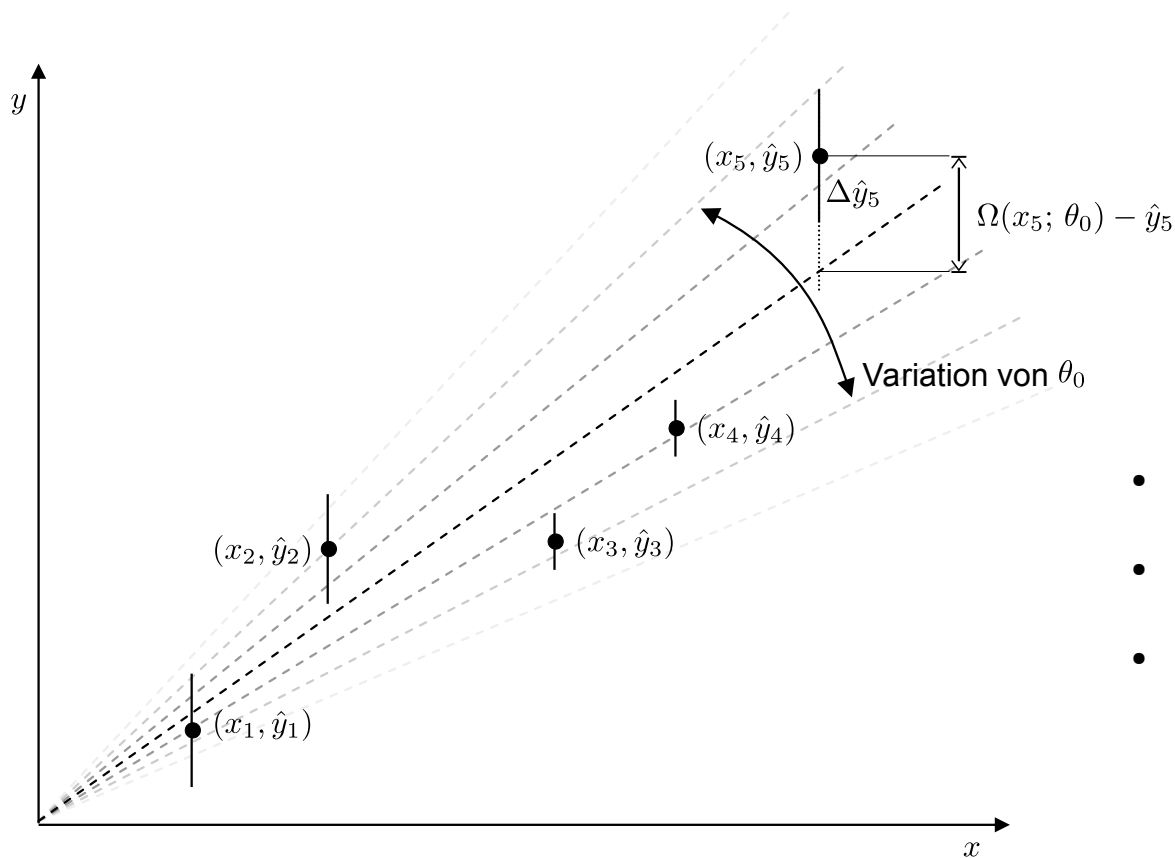
$$\delta(\hat{x}_0, \hat{x}_1) \equiv \frac{\hat{x}_0 - \hat{x}_1}{\sqrt{\Delta\hat{x}_0^2 + \Delta\hat{x}_1^2}}$$

nach einer Standardnormalverteilung $\varphi(\delta, 0, 1)$ verteilt!

Wenn \hat{x}_0 und \hat{x}_1 WIRKLICH kompatibel sind, dann gilt:

- $\delta \in [-1; 1]$ mit **68%** Wahrscheinlichkeit
- $\delta \in [-2; 2]$ mit **95%** Wahrscheinlichkeit
- $\delta \in [-3; 3]$ mit **99%** Wahrscheinlichkeit

Bestimmung von Modellparametern



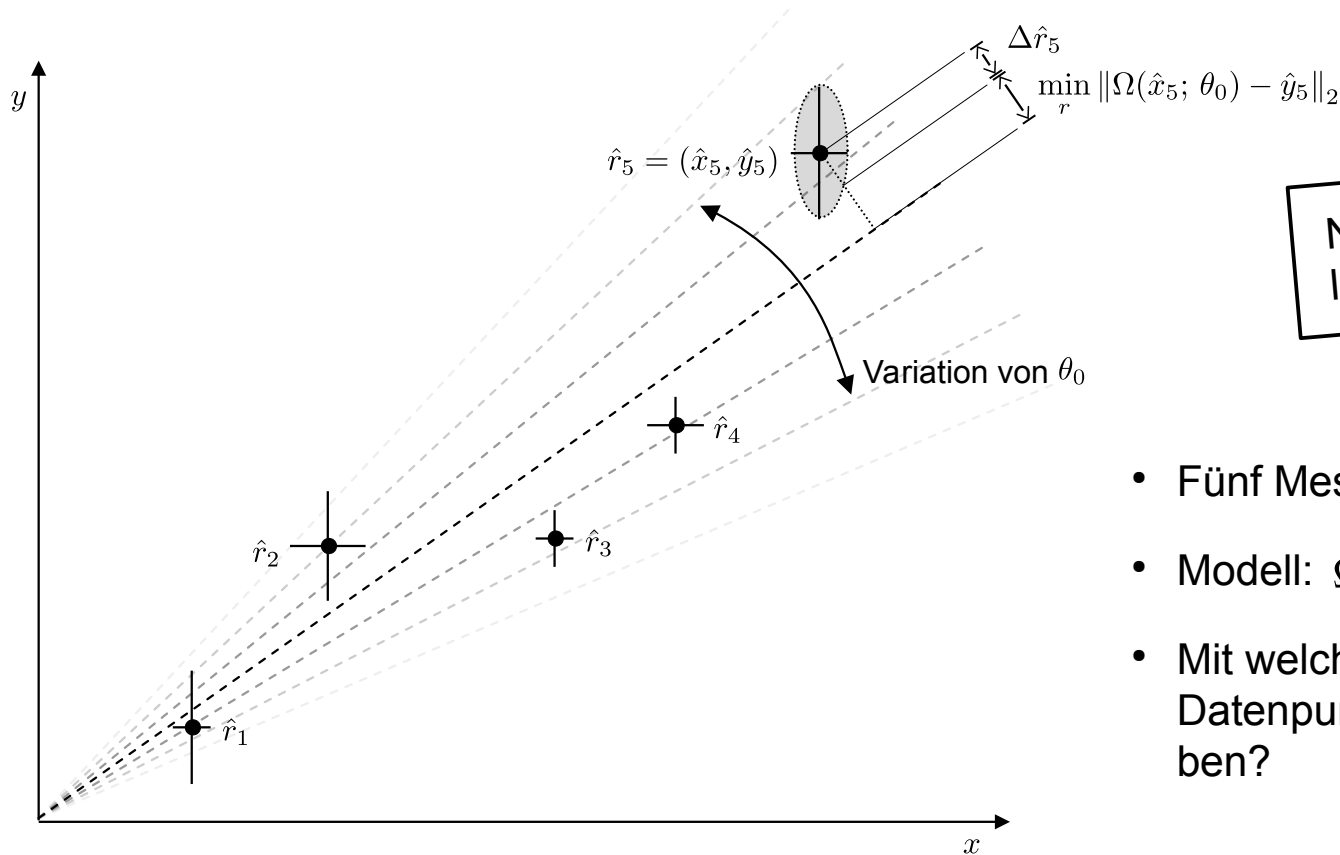
Analytisch
lösbar

- Fünf Messpunkte $\hat{r}_i = (\hat{x}_i; \hat{y}_i)$
- Modell: $\Omega : y = \theta_0 x$
- Mit welchem Wert $\hat{\theta}_0$ kann Ω die Datenpunkte am besten beschreiben?

$$z(\theta_0, \{x_i\}; \{\hat{y}_i\}) = \sum_{i=1}^n \left(\underbrace{\frac{\Omega(x_i; \theta_0) - \hat{y}_i}{\Delta \hat{y}_i}}_{\equiv \delta(\Omega, \hat{y}_i)} \right)^2 = \sum_{i=1}^n \left(\frac{\theta_0 x_i - \hat{y}_i}{\Delta \hat{y}_i} \right)^2$$

Minimiere Spannung
zwischen Modell und
Datenpunkten

Bestimmung von Modellparametern



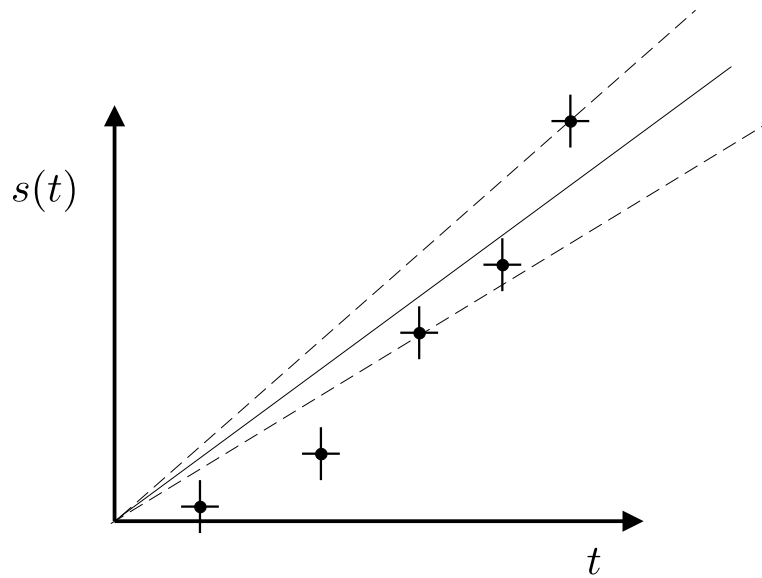
Nur numerisch lösbar!

- Fünf Messpunkte $\hat{r}_i = (\hat{x}_i; \hat{y}_i)$
- Modell: $\Omega : y = \theta_0 x$
- Mit welchem Wert $\hat{\theta}_0$ kann Ω die Datenpunkte am besten beschreiben?

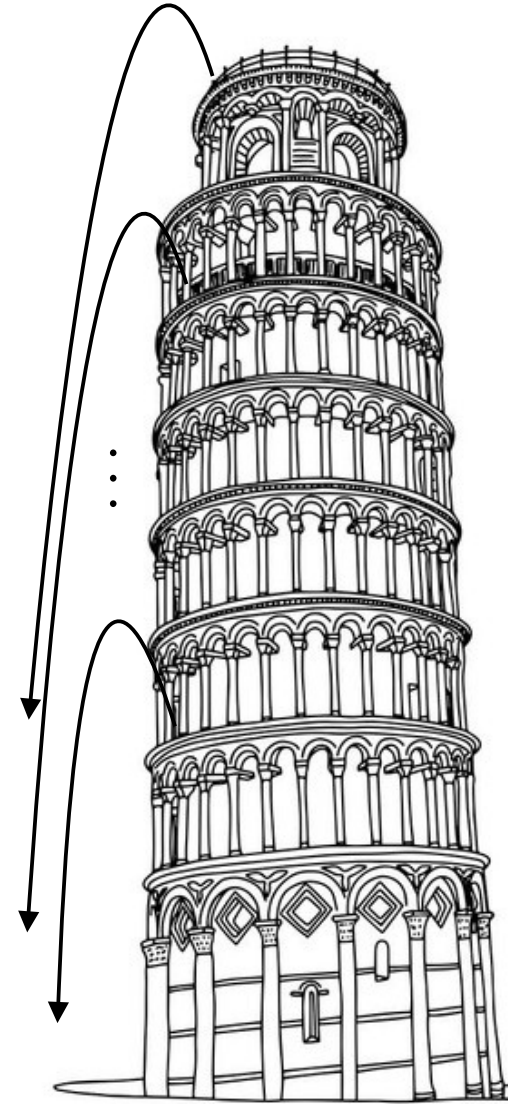
$$z(\theta_0; r, \{\hat{r}_i\}) = \sum_{i=1}^n \left(\underbrace{\frac{\min_r \|\Omega(r; \theta_0) - \hat{r}_i\|_2}{\Delta \hat{r}_i}}_{\equiv \delta(\Omega, \hat{r}_i)} \right)^2.$$

Minimiere Spannung zwischen Modell und Datenpunkten

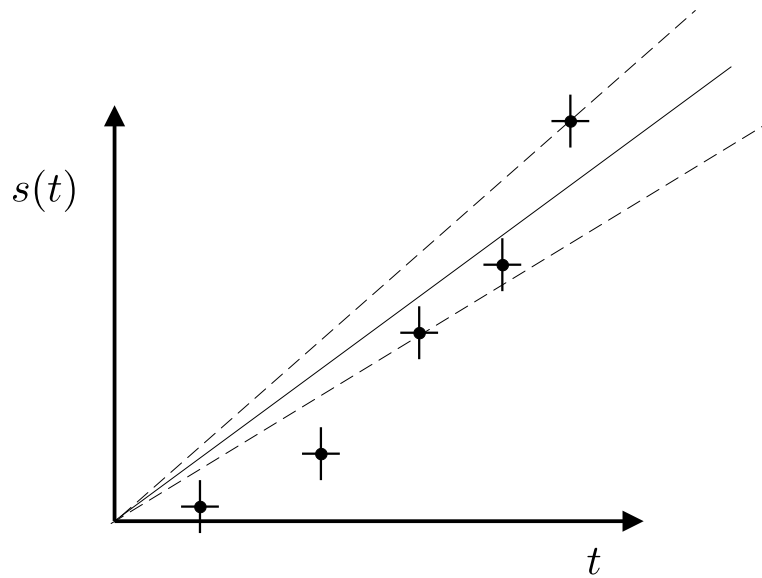
Anwendung auf unser Beispiel



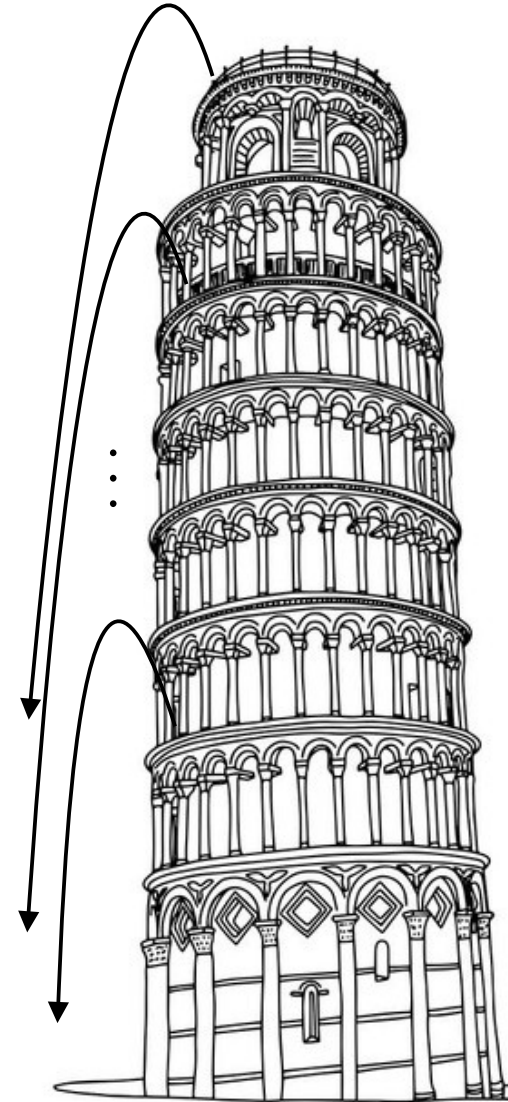
- Q: Wie groß ist die Steigung der Geraden?



Anwendung auf unser Beispiel

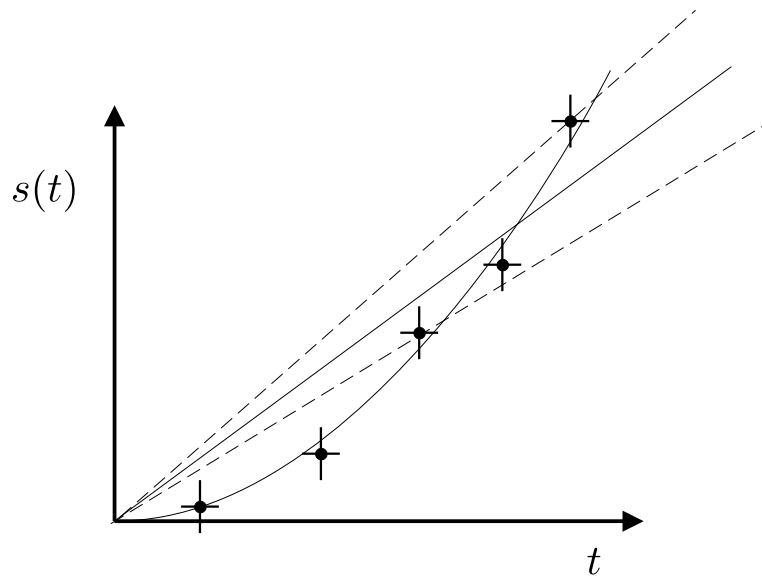


- Q: Wie groß ist die Steigung der Geraden? – **Falsche Frage!**

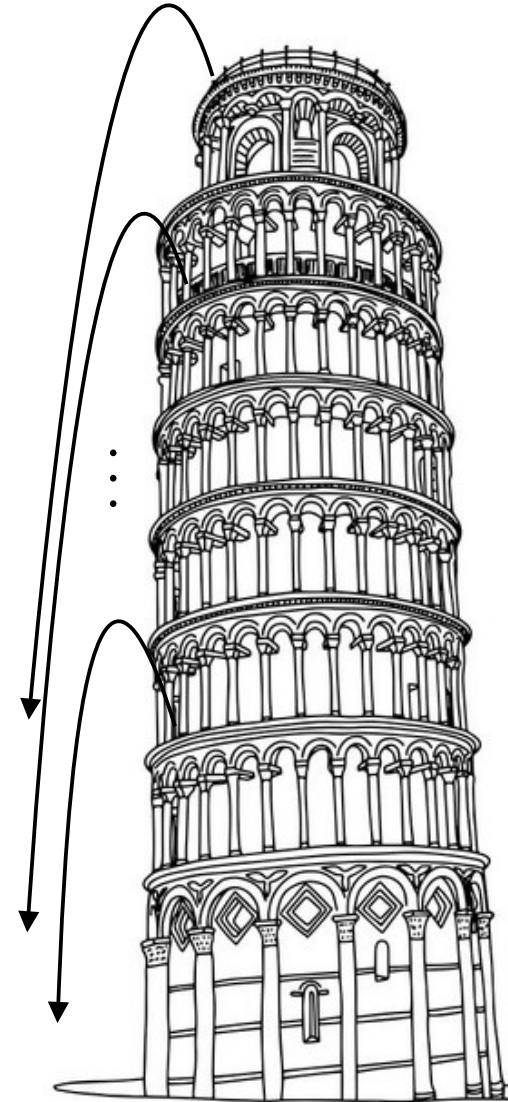


Anwendung auf unser Beispiel

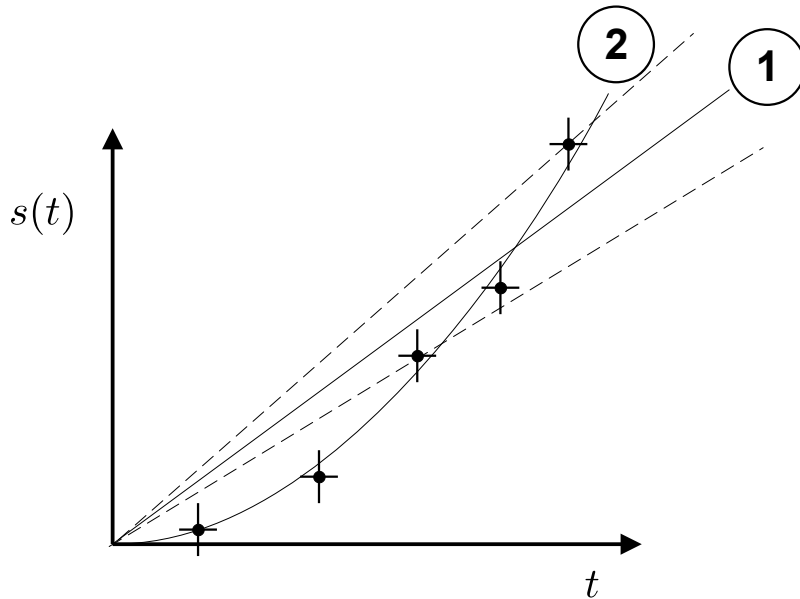
$$l = \frac{1}{2} g t^2$$



- Q: Wie groß ist die Steigung der Geraden? – **Falsche Annahme!**



- **Wie kompatibel ist mein Modell** selbst bei optimaler Wahl der Parameter mit den Daten?
- Quantifiziert durch den Wert von $\hat{z}(\theta_0, \{x_i\}; \{\hat{y}_i\})$ im Minimum

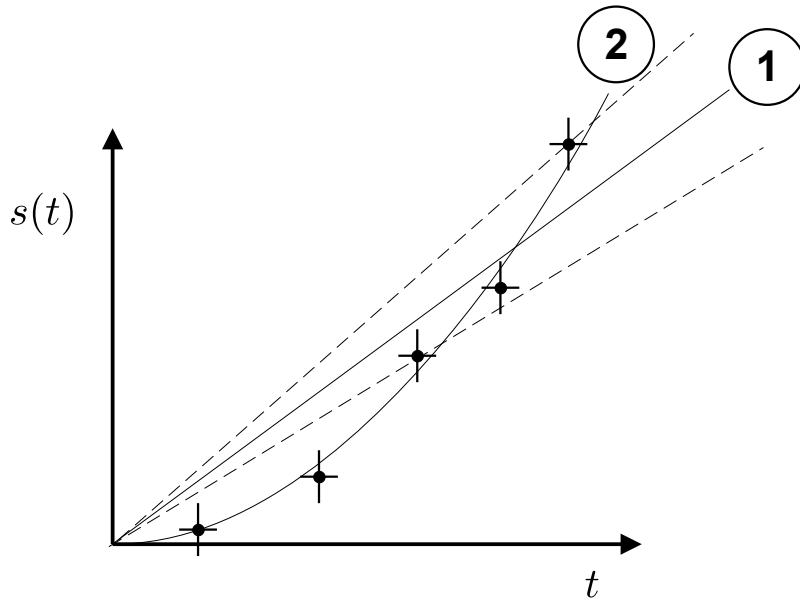


① $\hat{z}^{(1)}(\theta_0, \{x_i\}; \{\hat{y}_i\}) = 7.35$

② $\hat{z}^{(2)}(\theta_0, \{x_i\}; \{\hat{y}_i\}) = 4.05$

Welches Modell ist besser mit den Daten kompatibel?

- **Wie kompatibel ist mein Modell** selbst bei optimaler Wahl der Parameter mit den Daten?
- Quantifiziert durch den Wert von $\hat{z}(\theta_0, \{x_i\}; \{\hat{y}_i\})$ im Minimum



① $\hat{z}^{(1)}(\theta_0, \{x_i\}; \{\hat{y}_i\}) = 7.35$

② $\hat{z}^{(2)}(\theta_0, \{x_i\}; \{\hat{y}_i\}) = 4.05$

Welches Modell ist besser mit den Daten kompatibel?

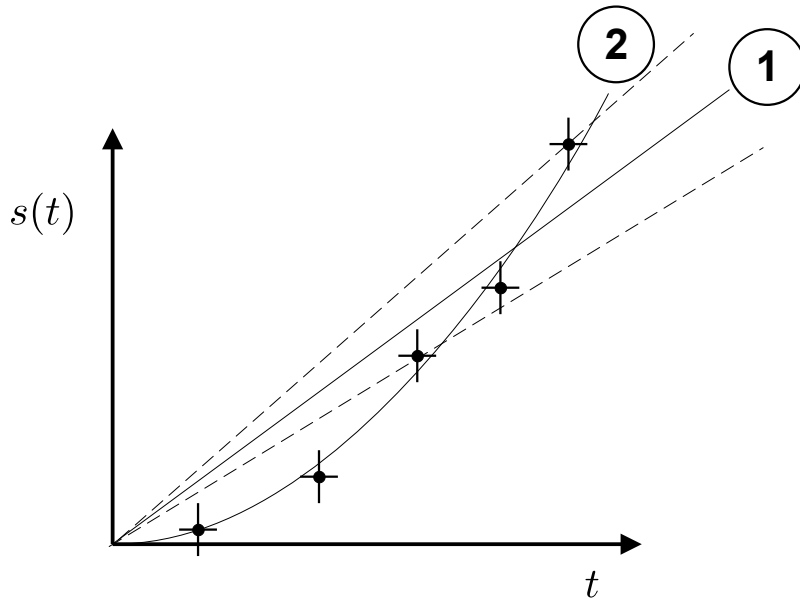
Für ein festgelegtes Modell mit $N(= 5)$ Datenpunkten gibt

$$\hat{z}/N$$

die **mittlere Spannung² pro Datenpunkt** an!

χ^2 -Test – Freiheitsgrad

- Enthält ein Modell freie Parameter muss N auf die **Anzahl der Freiheitsgrade (α)**, die zur Anpassung zu Verfügung stehen reduziert werden

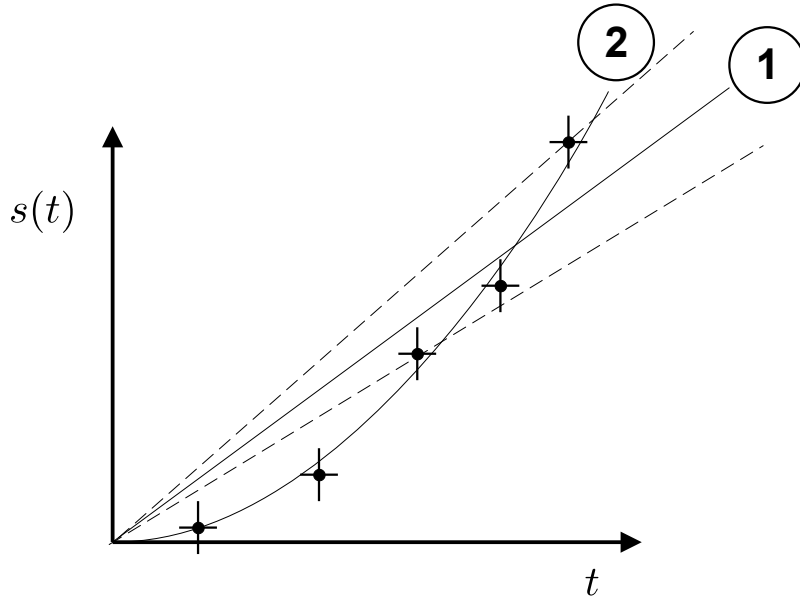


- 1 Durch zwei Punkte wird eine Gerade in der Ebene festgelegt.
- 2 Durch drei Punkte eine Parabel.

Freiheitsgrad

$$\alpha = (\# \text{ Punkte} - \# \text{ Parameter})$$

- Angewand auf unser Beispiel:



Ist $\hat{z}/\alpha \gg 1$ ist das angepasste Modell nicht mit den Daten kommpatibel!

- ① $\hat{z}^{(1)}(\theta_0, \{x_i\}; \{\hat{y}_i\}) / \alpha = 1.84$ ❌
- ② $\hat{z}^{(2)}(g, \{x_i\}; \{\hat{y}_i\}) / \alpha = 1.01$ ✅

Für ein Modell mit α Freiheitsgraden gibt

$$\hat{z}/\alpha$$

die **mittlere Spannung² pro Freiheitsgrad** an!

The end...



Mit diesem Rüstzeug sollten Sie gut fürs P1-Praktikum gewappnet sein.

(c) karlsruhe-insider.de

Backup
