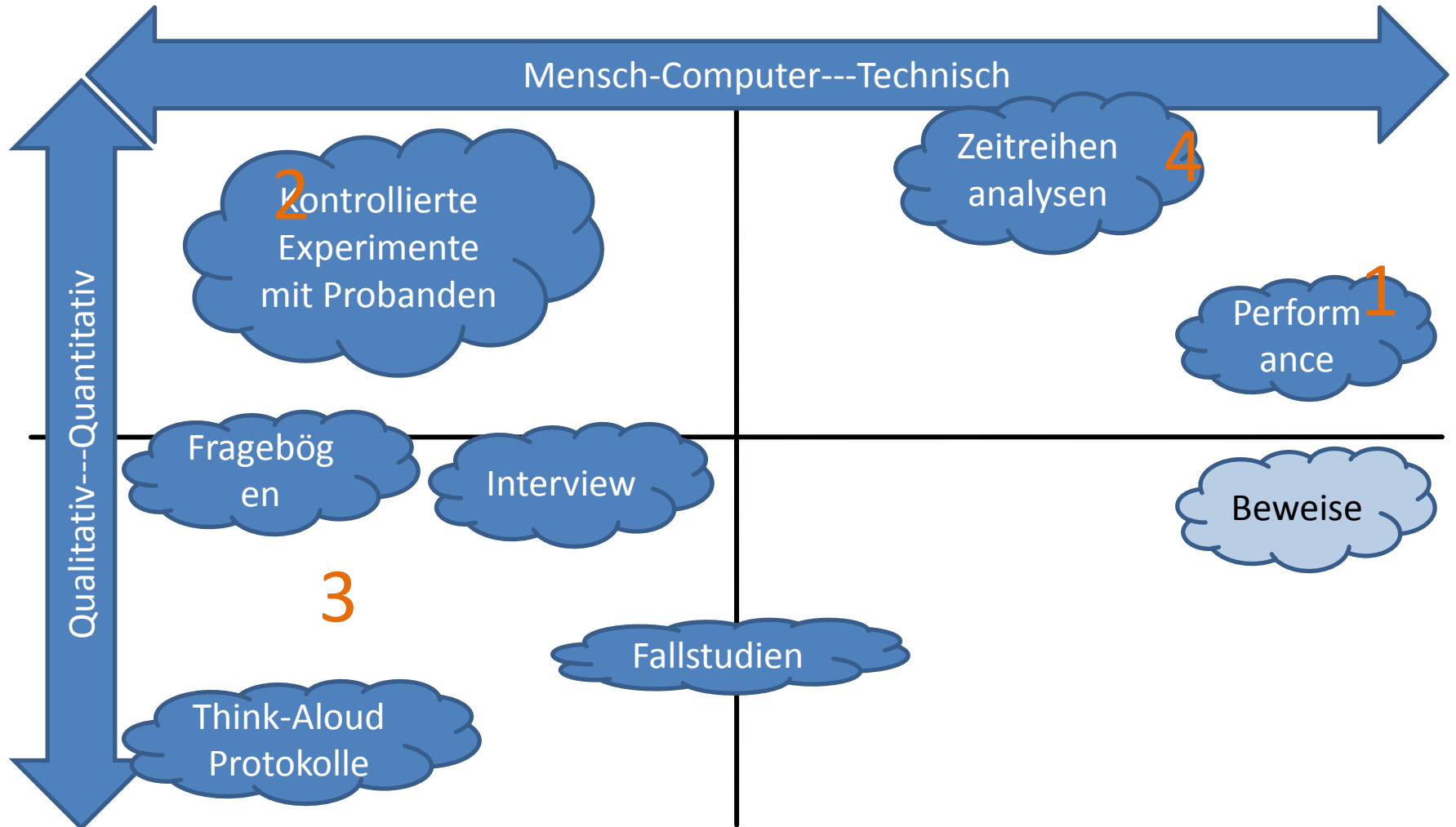


Performance Messungen

Einordnung



Lernziele

- Schwierigkeiten von Performance-Analysen verstehen
- Performance-Analysen bewerten können



Warum Performanceanalyse?

Warum Performanceanalyse?

- Alternativen vergleichen
- Einfluss eines Features
- System Tuning
- Relative Performance erkennen (über Zeit)
- Absolute Performance für ausgewählte Fälle
- Erwartungen setzen
- Analyse von Systemverhalten

Analysetechniken

- Messen
 - keine vereinfachenden Annahmen
 - i.d.R. am glaubwürdigsten
 - inflexibel, spezielles System

Analysetechniken

- Messen
 - keine vereinfachenden Annahmen
 - i.d.R. am glaubwürdigsten
 - inflexibel, spezielles System
- Simulation
 - Abstraktion
 - Flexibel

Analysetechniken

- Messen
 - keine vereinfachenden Annahmen
 - i.d.R. am glaubwürdigsten
 - inflexibel, spezielles System
- Simulation
 - Abstraktion
 - Flexibel
- Analytisches Modellieren
 - Mathematische Beschreibung des Systems
 - Starke Abstraktion, i.d.R. kaum glaubwürdig
 - Insbesondere zur frühen Validierung

Benchmark

Benchmark

- Ausführen realer Programme/Hardwarekomponenten in realen Umgebungen (keine analytische Simulation)
- Messen von Performance, Speicherverbrauch, usw.
- Automatisierbar
- Kein menschlicher Einfluss

Benchmark - Beispiele

Benchmark - Beispiele

- 3DMark (Grafikkarte/System)
- TCP-H (Datawarehouse)
- TCP-C (On-line transaction processing)
- Sintel (Video-Encoder)

Was messen?

Was messen?

- Ausführungszeit
- CPU-Zyklen
- MIPS (Million instructions per second)
- MFLOPS (Million floating-point operations per second)
- SPEC (System Performance Evaluation Cooperative)
- QUIPS (Quality improvements per second)
- Transaktionen pro Sekunde

Aufgabe

- Welche Kriterien sollte eine gute Metrik erfüllen?
- Sind die vorgestellten Metriken gute Metriken nach Ihren Kriterien?

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität							
Reliabilität							
Wiederhol- barkeit							
Einfache Messbarkeit							
Konsistenz							
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität							
Wiederhol- barkeit							
Einfache Messbarkeit							
Konsistenz							
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit							
Einfache Messbarkeit							
Konsistenz							
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit	+	+	+	+	+	+	+
Einfache Messbarkeit							
Konsistenz							
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit	+	+	+	+	+	+	+
Einfache Messbarkeit	+	+	+	+	+	+	+
Konsistenz							
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit	+	+	+	+	+	+	+
Einfache Messbarkeit	+	+	+	+	+	+	+
Konsistenz	+	+	-	-	+	+	+
Unabhängig- keit							

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit	+	+	+	+	+	+	+
Einfache Messbarkeit	+	+	+	+	+	+	+
Konsistenz	+	+	-	-	+	+	+
Unabhängig- keit	+	+	+	-	-	+	+

Kriterien

Kriterium	Ausführungs-zeit	CPU Zykle n	MIPS	MFLOPS	SPEC	QUIPS	Transactions/ second
Linearität	+	-	-	+	-	+	+
Reliabilität	+	-	-	-	-	-	+
Wiederhol- barkeit	+	+	+	+	+	+	+
Einfache Messbarkeit	+	+	+	+	+	+	+
Konsistenz	+	+	-	-	+	+	+
Unabhängig- keit	+	+	+	-	-	+	+

Beispiel für Prüfungsfrage: Welche Metrik(en) würden Sie benutzen, um den schnellsten Sortieralgorithmus zu bestimmen?

Störvariablen

Störvariablen

- Beeinflussen das Messergebnis systematisch oder unsystematisch

Störvariablen

- Beeinflussen das Messergebnis systematisch oder unsystematisch
- Beispiele:
 - Hintergrundprozesse
 - Hardwareunterschiede
 - Temperaturunterschiede
 - Eingabedaten, zufällig?
 - Heap-Size
 - Hardware-Plattform
 - System-Interrupts
 - Parallelität in Single- und Multicore-Systemen
 - Garbage Collection

Aufgabe

- Wie kann man den Einfluss dieser Störvariablen kontrollieren?

Typisches Vorgehen: Bester Wert

- Wiederholen
- Bester, zweitbester oder schlechtester Wert
- Bsp: Antwortzeiten für Programmieraufgabe
- R: Daten einlesen
 - `data <- read.csv("rt.csv", header=TRUE, sep = ";", dec = ".")`
 - header: gibt an, ob Variablen/Spaltennamen in der ersten Zeile stehen
 - sep: Separator für Datensätze in der selben Zeile
 - dec: Dezimaltrennzeichen
 - `rt <- data[, 'time']`
 - `min(rt)/max(rt)`

Typisches Vorgehen: Mittelwert

- Messung wiederholen
- Mittelwert bilden

$$\bar{x}_{arithm} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- R:
 - mean(rt)

Median

Median

- Wert, der in der Mitte liegt
- Robust gegen Ausreißer
- R:
 - `median(rt)`
- Bei gerader Anzahl an Messwerten:
 - Arithmetisches Mittel der beiden mittleren Werte
 - Einen der beiden mittleren Werte angeben

Median oder Mittelwert?

Median oder Mittelwert?

- Median statt arithmetisches Mittel, wenn
 - Ordinale Daten*
- *Skalenniveaus
 - Nominal (z.B. Geschlecht)
 - Ordinal (z.B. Platzierungen)
 - Metrisch (z.B. Temperatur, Antwortzeit)

Median oder Mittelwert?

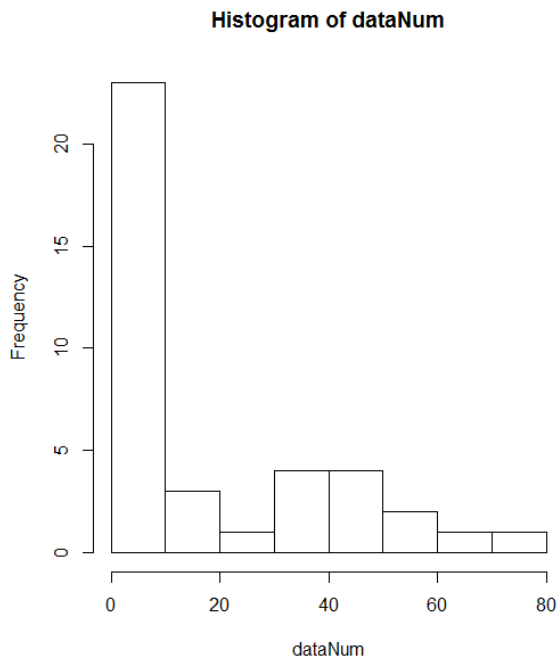
- Median statt arithmetisches Mittel, wenn
 - Ordinale Daten*
 - Wenig Messwerte
 - Asymmetrische Verteilung
 - Ausreißer
- *Skalenniveaus
 - Nominal (z.B. Geschlecht)
 - Ordinal (z.B. Platzierungen)
 - Metrisch (z.B. Temperatur, Antwortzeit)

Daten anschauen

- Überblick verschaffen
- Verteilung und Ausreißer einschätzen

Histogramme

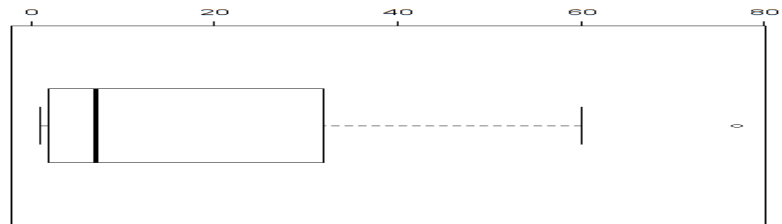
- Häufigkeit von Messwerten in festgelegten Bereichen



- R
 - `rtNum <- as.numeric(unlist(rt))`
 - `hist(rt)`

Boxplots

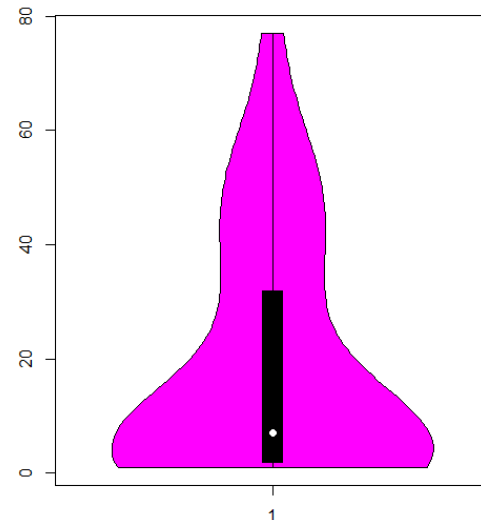
- Boxplot zeigt
 - Median als breite Linie
 - Quartile als Box (50% aller Werte in der Box)
 - Whiskers
 - Ausreisser als Punkte
- Graphische Darstellung von Verteilungen



- R: `boxplot(rt)`

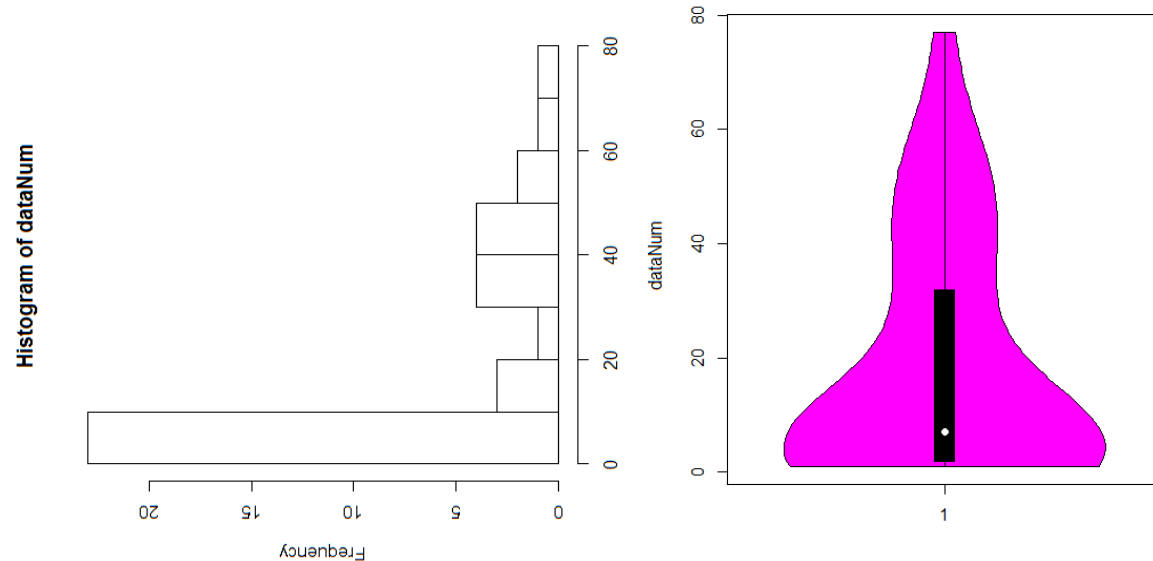
Violin-Plot

- Zeigt zusätzlich zu Boxplot die Verteilung der Daten
- R:
 - `install.packages("vioplot")`
 - `library(vioplot)`
 - `vioplot(rtNum)`



Violin-Plot

- Zeigt zusätzlich zu Boxplot die Verteilung der Daten
- R:
 - `install.p`
 - `library(\`
 - `vioplot(`



Messmodel

- $y = \tau + \varepsilon$
- y : beobachteter Wert
- τ : wahrer Wert
- ε : Fehler
- Population: griechische Buchstaben
- Stichprobe: deutsche Buchstaben

Fehlermodell

- Echter Mittelwert: 10
- 1 zufälliger Fehler, Einfluss ± 1
- Messwerte: 9 (50%) und 11 (50%)
- 2 zufällige Fehler, je ± 1
- Messwerte: 8 (25%), 10 (50%) und 12 (25%)
- 3 zufällige Fehler, je ± 1
- Messwerte: 7 (12.5%), 9 (37.5%), 11 (37.5%), 13 (12.5%)
- N zufällige Fehler, je ± 1
- Normalverteilung

Fehlermodell

- Echter Mittelwert: 10
- 1 zufälliger Fehler, Einfluss ± 1
- Messwerte: 9 (50%) und 11 (50%)

Fehlermodell

- Echter Mittelwert: 10
- 1 zufälliger Fehler, Einfluss ± 1
- Messwerte: 9 (50%) und 11 (50%)
- 2 zufällige Fehler, je ± 1
- Messwerte: 8 (25%), 10 (50%) und 12 (25%)

Fehlermodell

- Echter Mittelwert: 10
- 1 zufälliger Fehler, Einfluss ± 1
- Messwerte: 9 (50%) und 11 (50%)
- 2 zufällige Fehler, je ± 1
- Messwerte: 8 (25%), 10 (50%) und 12 (25%)
- 3 zufällige Fehler, je ± 1
- Messwerte: 7 (12.5%), 9 (37.5%), 11 (37.5%), 13 (12.5%)

Fehlermodell

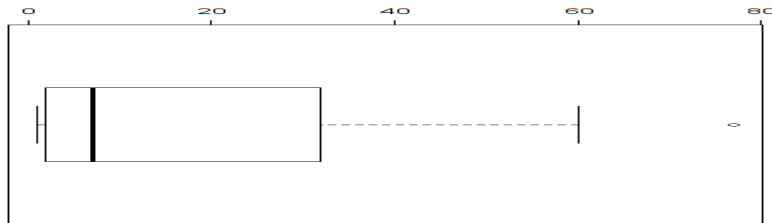
- Echter Mittelwert: 10
- 1 zufälliger Fehler, Einfluss ± 1
- Messwerte: 9 (50%) und 11 (50%)
- 2 zufällige Fehler, je ± 1
- Messwerte: 8 (25%), 10 (50%) und 12 (25%)
- 3 zufällige Fehler, je ± 1
- Messwerte: 7 (12.5%), 9 (37.5%), 11 (37.5%), 13 (12.5%)
- N zufällige Fehler, je ± 1
- Normalverteilung

Normalverteilung



Streuung

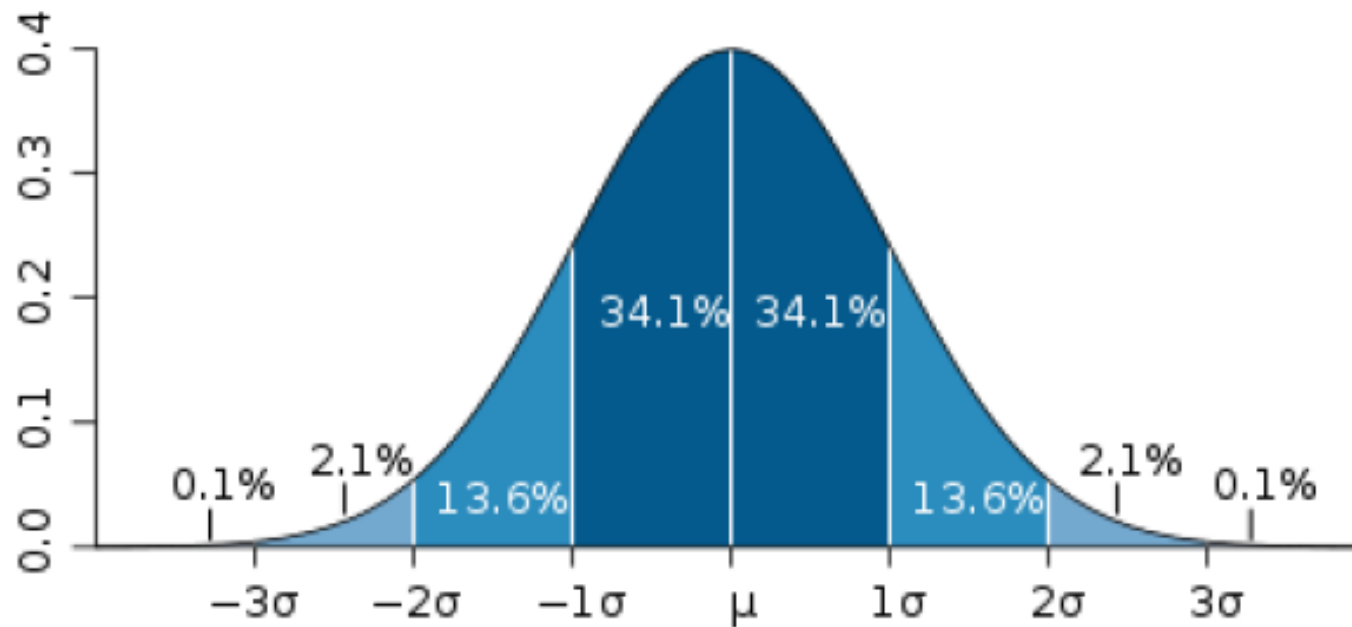
- Mittelwert: 45,55
- Boxplot



Standardabweichung

- R:
 - `sd(rtNum)`
 - 21,55
- Mittelwert: 45,55
- 24 → 45,55 (34 % der Messwerte)
- 45,55 → 67,1 (34% der Messwerte)

Standardabweichung



Standardabweichung: Anwendung

Standardabweichung: Anwendung

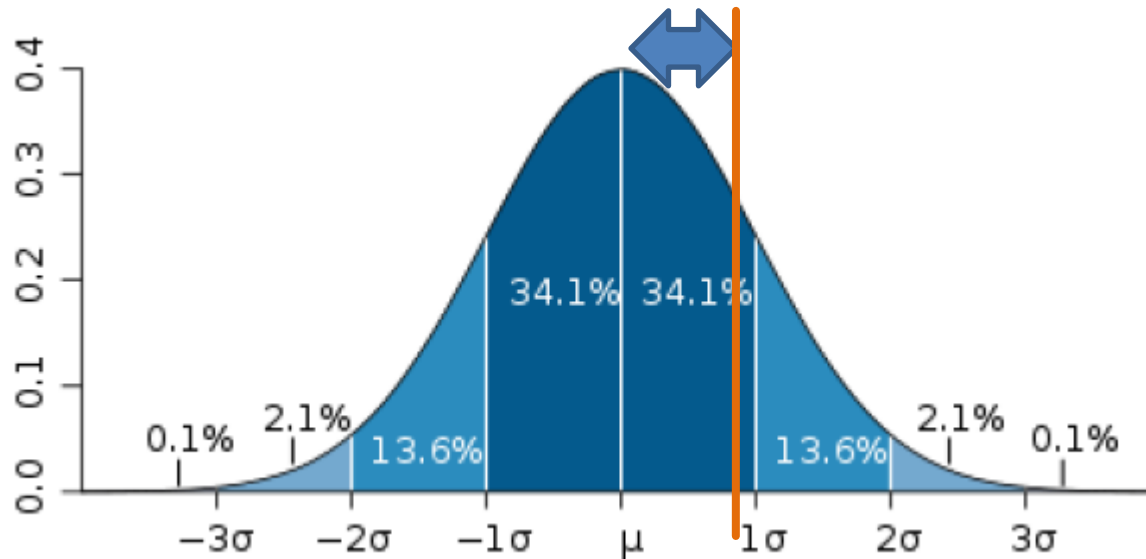
- Ausreißer definieren
- Hochbegabung definieren
- Entdeckung des Higgs-Boson verkünden

Genauigkeit vs. Präzision

Genauigkeit:

Abweichung beobachteter Mittelwerte vom wahren Mittelwert

Wichtig bei Zeitmessungen

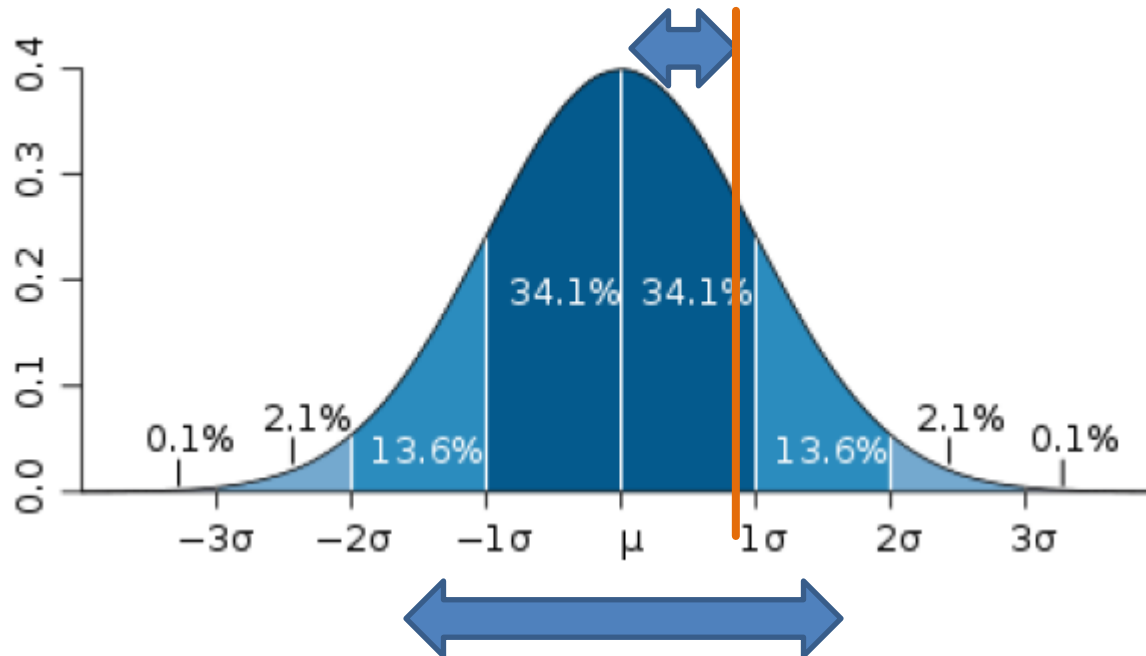


Genauigkeit vs. Präzision

Genauigkeit:

Abweichung beobachteter Mittelwerte vom wahren Mittelwert

Wichtig bei Zeitmessungen



Präzision:

Streuung um Stichprobenmittelwert

Ursache von Messfehlern unklar

Zufällige vs. Systematische Fehler

- Systematische Fehler: Fehler des Experiments/der Messmethode
 - CPU Speed: Messung bei unterschiedliche Temperaturen
 - Zustand nicht zurückgesetzt für zweite Messung
 - Geringe Varianz, bis konstant über alle Messungen
 - Im Design ausschließen, braucht Erfahrung
 - Genauigkeit
- Zufällige Fehler
 - Nicht kontrollierbar
 - Stochastische Methoden
 - Präzision

Lernziele

- Schwierigkeiten von Performance-Analysen verstehen
- Performance-Analysen bewerten können

Literatur

- David Lilja. *Measuring Computer Performance: A practitioner's guide*. Cambridge University Press. 2000.
- Performance-Paper
- Beliebiges Statistikbuch

Hausaufgabe

- Folgende Paper auszugsweise lesen:
 - How Do Professional Developers Comprehend Software? (Abschnitt II, Abschnitt III überfliegen)
 - An Experiment About Static and Dynamic Type Systems (Abschnitt 4, Abschnitt 5 überfliegen)
- Experiment-Aufbau bewerten:
 - Was würden Sie genauso machen? Warum?
 - Was würden Sie anders machen? Warum?