Statistik – Methoden zum Vergleich von zwei Gruppen

t-Test für den Mittelwertunterschied von zwei unabhängigen/verbundenen Stichproben

t-Test – zwei unabhängige Stichproben

- Überprüfung von Unterschiedshypothesen hinsichtlich des Mittelwerts
- Gibt es einen signifikanten Unterschied im **Mittelwert** zwischen zwei Gruppen
- Mindestens intervallskalierte Variablen

Beispiele:
Wirkt eine Blutdrucksenker besser, wenn auch die Ernährung umgestellt wird?

Produziert Werk A besser als Werk B

t-Test – zwei unabhängige Stichproben

Beispiel Blutdrucksenker

Es werden zwei Testgruppen untersucht, die eine erhält den Blutdrucksenker, die andere Gruppe ernährt sich zusätzlich anders

Die Hypothesen:

H_0: $\bar{x}_1 = \bar{x}_2$ Kein Unterschied bei den Blutdruckwerten zwischen den beiden Gruppen

H_1: $\bar{x}_1 \neq \bar{x}_2$ Es gibt einen Unterschied bei den Blutdruckwerten zwischen den beiden Gruppen

t-Test – zwei unabhängige Stichproben

- Basis des t-Tests : Die t-Verteilung
- t-Verteilung hat in Abhängigkeit von der Anzahl der Freiheitgrade verschiedene Ausprägungen
- Mit wachsender Anzahl an Freiheitsgraden nähert sich die t-Verteilung immer mehr der Standardnormalverteilung
- Für n ≥ 30 gehen beide Verteilungen in einander über

t-Test – zwei unabhängige Stichproben

Berechnung der Freiheitsgrade bei zwei Stichproben

$$df = n_1 + n_2 - 2$$

df **Degree of Freedom**

n_i **Stichprobengrößen**

Diese Berechnung gilt im Grunde nur für gleich große Stichproben, für stark unterschiedliche Werte ergibt sich eine deutlich komplexere Bestimmung, auf die hier nicht näher eingegangen wird

t-Test – zwei unabhängige Stichproben

Grundlegende Arten des t-Tests

- **Test für unabhängige Stichproben**
  Die untersuchten Gruppen sind nicht von einander abhängig, siehe Beispiel Blutdrucksenker

- **Test für abhängige Stichproben**
  Die untersuchten Gruppen sind nicht von einander unabhängig, z.B. wiederkehrende Untersuchung der selben Personen

- **Test für eine Stichprobe**
  Kennen Sie schon

t-Test – zwei unabhängige Stichproben

Voraussetzungen für den t-Test für zwei unabhängige Stichproben

- Mindestens intervallskaliert
- Keine extremen Ausreißer
- Normalverteilung beider Stichproben (bzw. $n_i \geq 30$)
- Keine signifikanten Streuungsunterschiede

Sind diese Voraussetzungen nicht erfüllt: Ausweichen auf andere Testverfahren (i.a. nicht-parametrische Tests)

t-Test – zwei unabhängige Stichproben

- Für den t-Test sind erhebliche Voraussetzungen zu erfüllen, er belohnt aber mit einer hohen Stabilität und bietet die größte Power (Wahrscheinlichkeit, einen Unterschied zu erkennen, der auch real vorhanden ist)

- **Unabhängigkeit der Stichproben**
  - Die Messwerte der einen Gruppe hängen nicht von den Werten der anderen Gruppe ab
  - Untersuchung unterschiedlicher Individuen

t-Test – zwei unabhängige Stichproben

- **Normalverteilung** in beiden Grundgesamtheiten
  - Muss getestet werden, wenn n < 30!
  - Grundlage Zentraler Grenzwertsatz

- **Varianzgleichheit**
  - Die Streuung in den untersuchten Gruppen darf nicht signifikant voneinander abweichen
  - Bei Zweifel ist zu Testen! (Homogenität der Varianz: F-Test bzw. Bartlett-Test)

t-Test – zwei unabhängige Stichproben

Alternativen bei Nichterfüllung

Voraussetzung	Alternativen
Intervallskalierung	Mann-Whitney-U-Test Wilcoxon-Rangsummen-Test Chi-Quadrat-Test
Normalverteilung	Mann-Whitney-U-Test Wilcoxon-Rangsummen-Test
Varianzgleichheit	Welch-Test Mann-Whitney-U-Test Wilcoxon-Rangsummen-Test
Unabhängigkeit	T-Test für verbundene Stichproben

t-Test – zwei unabhängige Stichproben

Die Prüfgröße

Der t-Wert für unsere Daten bestimmt sich aus:

$$\boldsymbol{t} = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled} * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \text{mit} \qquad \boldsymbol{s_{pooled}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Dieser t-Wert wird im folgenden mit einem zu bestimmenden kritischen t-Wert verglichen

Das Ergebnis dieses Vergleichs entscheidet über die Schluss letztendlich gültige Hypothese

t-Test – zwei unabhängige Stichproben

Mögliche Hypothesen für den t-Test bei zwei unabhängigen Stichproben

	ungerichtet	gerichtet	gerichtet
H_0	$\bar{x}_1 = \bar{x}_2$	$\bar{x}_1 \leq \bar{x}_2$	$\bar{x}_1 \geq \bar{x}_2$
H_1	$\bar{x}_1 \neq \bar{x}_2$	$\bar{x}_1 > \bar{x}_2$	$\bar{x}_1 < \bar{x}_2$
$t_{kritisch}$	$\pm t_{1-\alpha/2,df}$	$t_{1-\alpha,df}$	$t_{1-\alpha,df}$
	zweiseitig	einseitig	einseitig
H_0verwerfen	$\|t\| > t_{kritisch}$	$t > t_{kritisch}$	$t < t_{kritisch}$

t-Test – zwei unabhängige Stichproben

Beispiel *Beispiel_Warteschlange.xlsx*

An zwei Supermarktkassen wird die durchschnittliche Wartezeit der Kunden ermittelt
Für die zwei Gruppen gilt:

$n_1 = 50$ $\bar{x}_1 = 8{,}18\, min$ $s_1 = 1{,}3810\, min$
$n_2 = 50$ $\bar{x}_2 = 6{,}24\, min$ $s_2 = 1{,}6114\, min$

Wir verzichten hier auf eine komplexe Handrechnung und nutzen stattdessen R

Für das Beispiel gehen wir von gleichen Varianzen aus

t-Test – zwei unabhängige Stichproben

Beispiel

Signifikanzniveau α = 5%

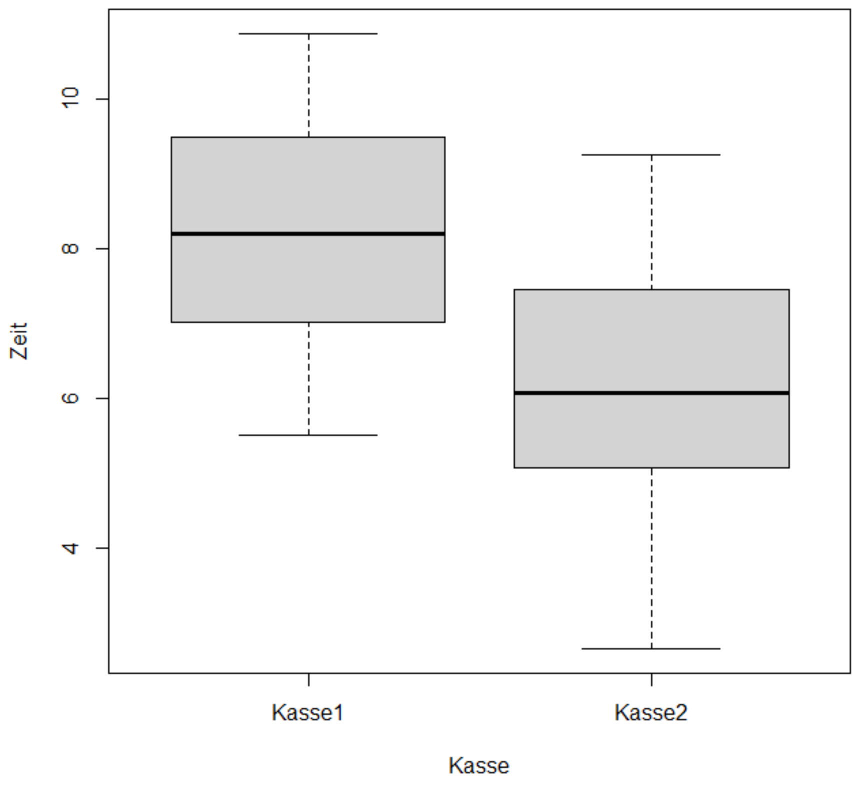
H_0: $\bar{x}_1 = \bar{x}_2$ Kein Unterschied bei den Wartezeiten der beiden Kassen

H_1: $\bar{x}_1 \neq \bar{x}_2$ Es gibt einen Unterschied bei den Wartezeiten der Kassen

t-Test – zwei unabhängige Stichproben

Beispiel

Ein Boxplot zeigt Unterschiede in Lage und Streuung



t-Test – zwei unabhängige Stichproben

Beispiel

```
Two Sample t-test

data:  Wartezeit by Index
t = 6.4826, df = 98, p-value = 0.000000003657
alternative hypothesis: true difference in means
between group Kasse1 and group Kasse2 is not equal to 0
95 percent confidence interval:
 1.349993 2.541163
sample estimates:
mean in group Kasse1 mean in group Kasse2
            8.181439             6.235861
```

p < α: Wechsel zur Alternativhypothese, unterschiedliche Wartezeiten

t-Test – gepaarte Stichproben

- t-Test für gepaarte Stichproben (auch als Test für abhängige / verbundene Stichproben bekannt)

- Ziel ist auch hier herauszufinden, ob sich die beiden Gruppen signifikant unterscheiden

- Gepaarte Gruppen:

  - Wiederholte Messungen, z.B. Medikamententests
  - Natürliche Paare, z.B. Zwillinge
  - Matching, zwei Individuen sind aufgrund einer Drittvariablen verbunden

t-Test – gepaarte Stichproben

Geändertes Vorgehen

- Kein Vergleich der Mittelwerte $\bar{x}_1, \bar{x}_2$ zweier unabhängiger Stichproben, sondern Untersuchung eines „Differenzdatensatzes“ $x_i = x_{i,SP1} - x_{i,SP2}$
- Der nachfolgende t-Test für gepaarte Stichproben ist vergleichbar mit dem t-Test für eine Stichprobe gegen einen Vorgabewert 0

t-Test – gepaarte Stichproben

Mögliche Hypothesen für den t-Test bei gepaarten Stichproben

	ungerichtet	gerichtet	gerichtet
H_0	$\bar{x}_d = 0$	$\bar{x}_d \leq 0$	$\bar{x}_d \geq 0$
H_1	$\bar{x}_d \neq 0$	$\bar{x}_d > 0$	$\bar{x}_d < 0$
$t_{kritisch}$	$\pm t_{1-\alpha/2, df}$	$t_{1-\alpha, df}$	$t_{1-\alpha, df}$
	zweiseitig	einseitig	einseitig
H_0verwerfen	$\|t\| > t_{kritisch}$	$t > t_{kritisch}$	$t < t_{kritisch}$

t-Test – gepaarte Stichproben

Voraussetzungen für den t-Test für zwei verbundene Stichproben

- Abhängigkeit der Stichproben (Korrelation?)

- Für die zu bildenden Differenzpaare ist gefordert
  - Mindestens intervallskaliert

  - Normalverteilung (bzw. $n \geq 30$)

  - Möglichst positive Korrelation (sonst kann die Power leiden)

t-Test – gepaarte Stichproben

Die Prüfgröße für gepaarte Stichproben

Der t-Wert für unsere Daten bestimmt sich aus:

$$\boldsymbol{t = \frac{\overline{x}_d}{s_d}} \qquad \text{mit} \qquad \boldsymbol{s_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \overline{x}_d)^2}{n(n-1)}}}$$

Dieser t-Wert wird im folgenden mit einem zu bestimmenden kritischen t-Wert verglichen

Das Ergebnis dieses Vergleichs entscheidet über die Schluss letztendlich gültige Hypothese

t-Test – gepaarte Stichproben

Beispiel *Beispiel_Pressure.xlsx*

Es wird die Wirkung eines Blutdrucksenkers an einer Testgruppe (n = 50) untersucht. Zu Beginn der Untersuchung wird der Blutdruck ohne vorherige Medikamentengabe gemessen, nachfolgend erhält dieselbe Gruppe einen Blutdrucksenker und wird noch einmal vermessen. Untersucht wird der systolische Wert.

Wir verzichten hier auf eine komplexe Handrechnung und nutzen stattdessen R

t-Test – gepaarte Stichproben

Beispiel

Signifikanzniveau α = 5%

$x_d = x_{Nachher} - x_{Vorher}$

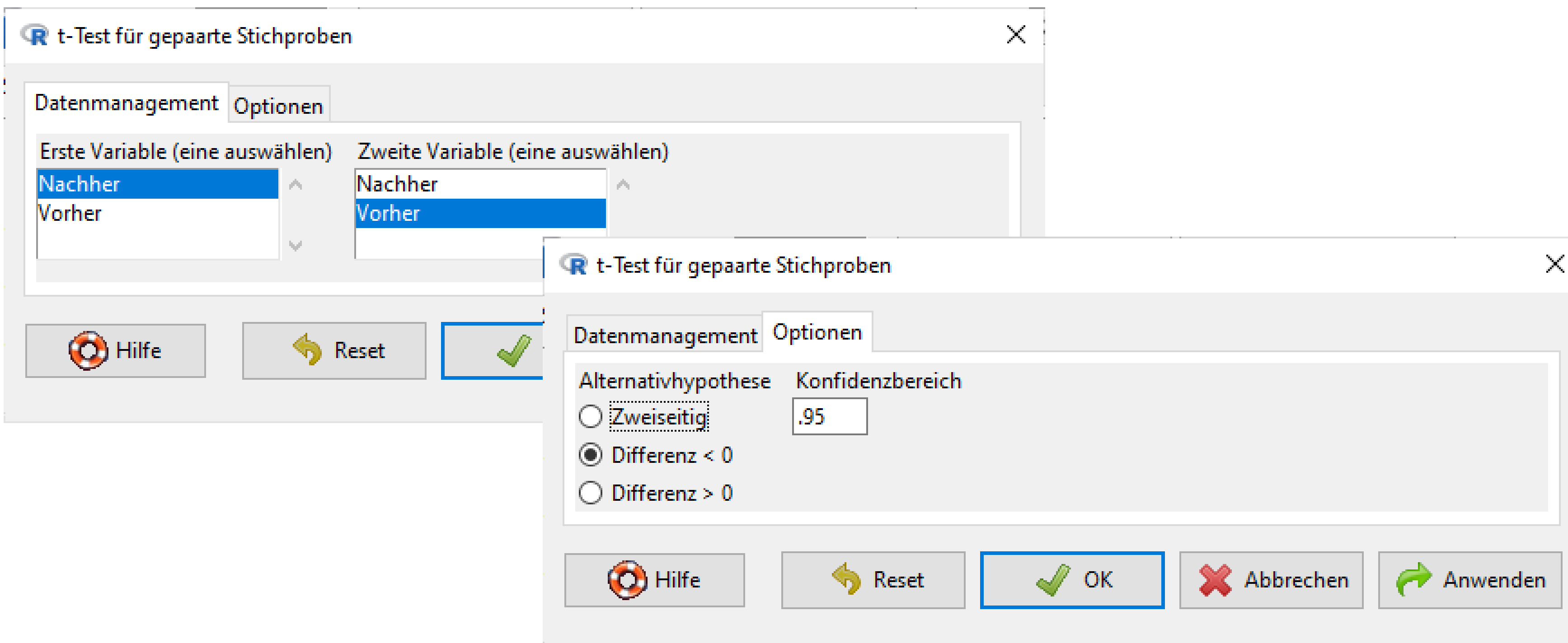
H_0: $\bar{x}_d \geq 0$ Der Blutdruck bleibt gleich bzw. steigt trotz Medikament

H_1: $\bar{x}_d < 0$ Der Blutdruck sinkt nach Medikamentengabe

t-Test – gepaarte Stichproben

Beispiel Eingabe in R

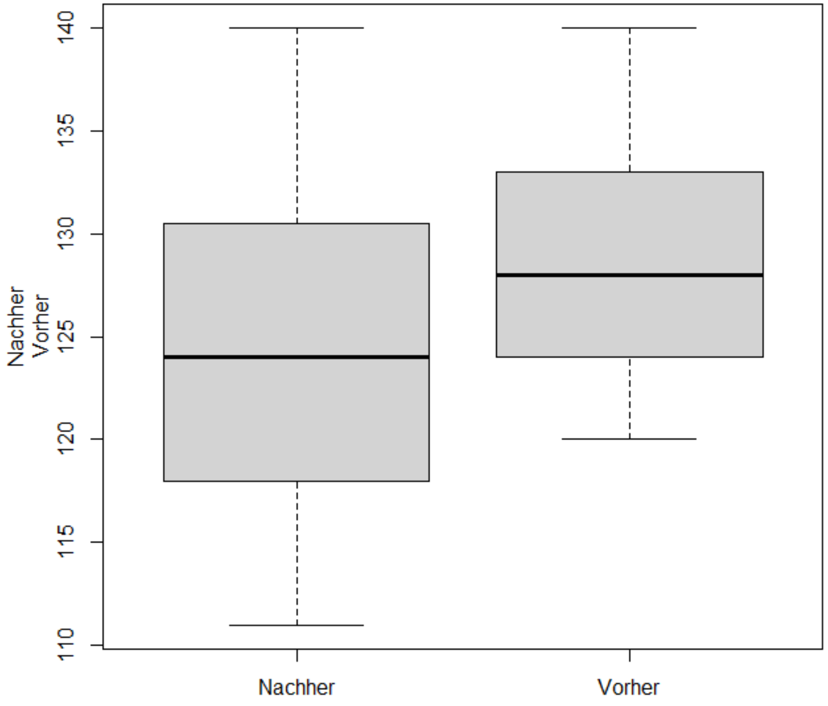
Bei der Eingabe in R ist nicht unmittelbar erkennbar, wie die Differenzbildung erfolgt: x_d = Erste Variable – Zweite Variable



t-Test – gepaarte Stichproben

Beispiel

Ein Boxplot zeigt Unterschiede in Lage und Streuung



t-Test – gepaarte Stichproben

Beispiel

```
Paired t-test

data:  Nachher and Vorher
t = -7.9293, df = 49, p-value = 1.225e-10
alternative hypothesis: true mean difference is less
than 0
95 percent confidence interval:
      -Inf -3.390818
sample estimates:
mean difference
           -4.3
```

p < α: Wechsel zur Alternativhypothese, Blutdruck sinkt