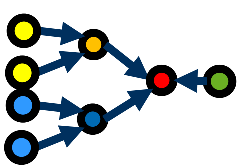
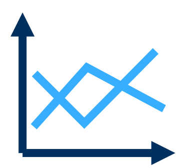
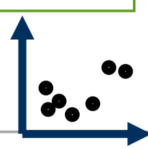
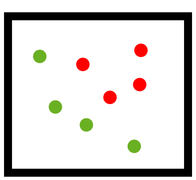
Hypothesen für statistische Tests

Es gibt hunderte von statistischen Tests

- Frage
  - Daten beschreiben
    - Zahl, Prozent, Verteilung
  - **Hypothese** testen
    - Unterschiede
      - ... im Mittelwert, Prozent, Verteilung
      - ...
    - Beziehung
      - Richtungslos: Korrelation
      - Gerichtet: Regression
      - ...
  - **Hypothesen** entdecken
    - z.B. Faktorenanlyse
    - ...

Wahl eines statistischen Tests

Wahl eines statistischen Testverfahren verlangt

- … Überblick über die grundlegenden statistischen Verfahren um daraus zu wählen
- … Kenntnisse über die Voraussetzungen und Anforderungen der stat. Tests
- … Kreativität, um aus den vorhanden Variablen neue zu bilden um die Daten für einen statistischen Test vorzubereiten.

Was die Wahl eines statistischen Test beeinflusst:

- die **Fragestellung und <u>Hypothesen</u>**
  - Deskriptive Auswertung
  - Hypothesen testen
    - Unterschiedshypothese: Unterschied zwischen welchen statistischen Kennwerten
    - Zusammenhangshypothese
  - Hypothesen entdecken
- das **<u>Skalenniveau</u>** (nominal-, ordinal-, metrischskaliert)
- die **Anzahl** der zu analysierenden **Variablen** (uni-, bi- oder multivariate Analysen)
- die **<u>Stichprobe</u>** (Größe; Art (unabhängig/abhängig))
- praktischen Problemen: Vorhandene Variablen/Daten
- Eigenschaften der Verteilung der Werte einer Variablen

Was die Wahl eines statistischen Test auch noch beeinflusst:

- dem zu untersuchendem Kennwerten/Verteilungen: Mittelwert, Häufigkeit, Korrelation, Rangsummen, Anteilen, Varianzen …
- der Kausalität des Zusammenhangs:
  - nur unabhängige Variablen
  - abhängig und unabhängige Variablen
- der Form des Zusammenhangs: linear, nichtlinear
- der Anzahl der der Gruppen: 1, 2 oder viele
- den Eigenschaften der Verteilung der Werte einer Variablen
- praktischen Problemen: Qualität der erhobenen Daten und des Erhebungsinstruments
- vielem mehr:
  - Wechselwirkungen zwischen Variablen
  - Welche Statistikprogramme stehen zur Verfügung.
  - Was versteht der Leser, was will er haben.
  - …

Hypothesen

Hypothesen sind abgeleitet von Forschungsfragen und eng verknüpft mit einem statistischen Test.

Von den verschiedene Hypothesenarten, braucht nicht jede Statistik (Es-gibt-Hypothesen). Manche Hypothesenart sind in den Sozialwissenschaften selten (Wenn-Dann) andere sehr typisch (Unterschied- Zusammenhangshypothesen).

Hypothesenarten

Hypothese	Erklärung	Beispiel
Es-gibt-Hypothesen	Bestätigt wenn ein Fall gefunden wurde.	
Wenn A dann B	Strenge Kausalität	
Unterschiedshypothesen	**Unterschiede in zwei (oder mehr) Gruppen**	
Zusammenhangshypothesen	**Stetiger Zusammenhang**	
Fälle Gruppieren	Ähnliche Fällen finden	
Variablen gruppieren	Ähnliche Variablen finden	

Von der Hypothesen über die Operationalisierung zum statistischen Test

Hypothese

„Je mehr jemand verdient, desto intelligenter ist er."

Je mehr
Geld
Desto mehr
IQ

Operationalisierung

Beispiel: Fragebogen

„Geld" messen mit Frage:

„IQ" messen mit:

Statistische Test

z.B. Korrelation zwischen

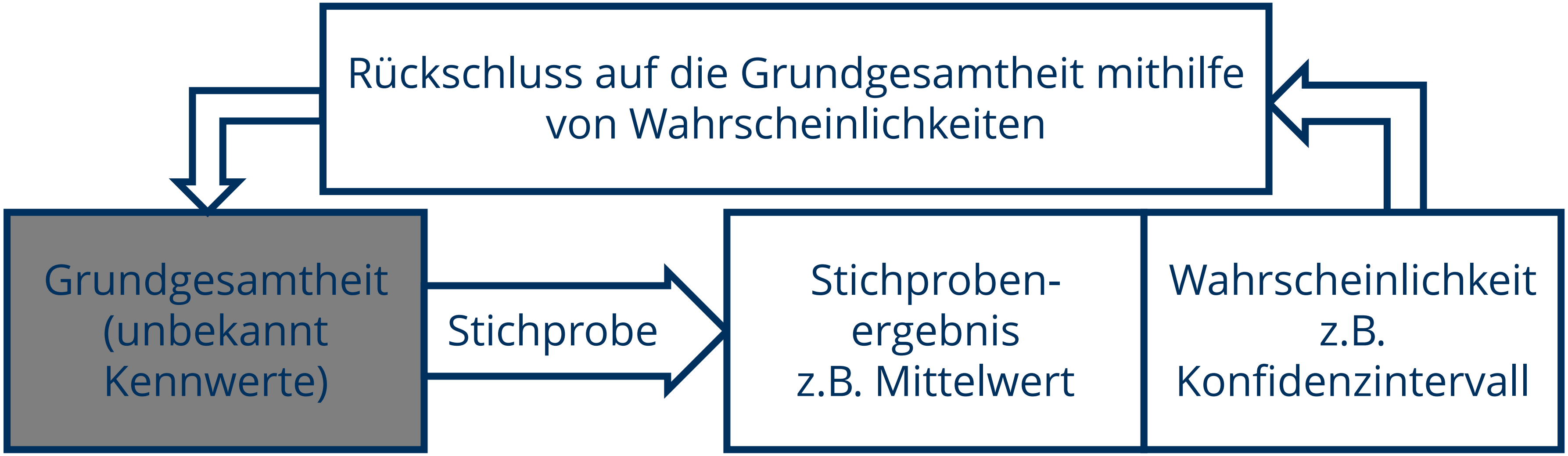
und

Übungen: Von der Frage zum Auswertungsverfahren

Frage	Variablen	Auswertungs-verfahren
Welche Einstellungen haben Mitarbeiter zur flexiblen Arbeitszeit?		
Empfinden Kunden einen oder zwei Ansprechpartner besser?		
Größere Anbieter haben höhere Preise		
Meetings die kürzer und besser vorbereitet sind erfolgreicher.		

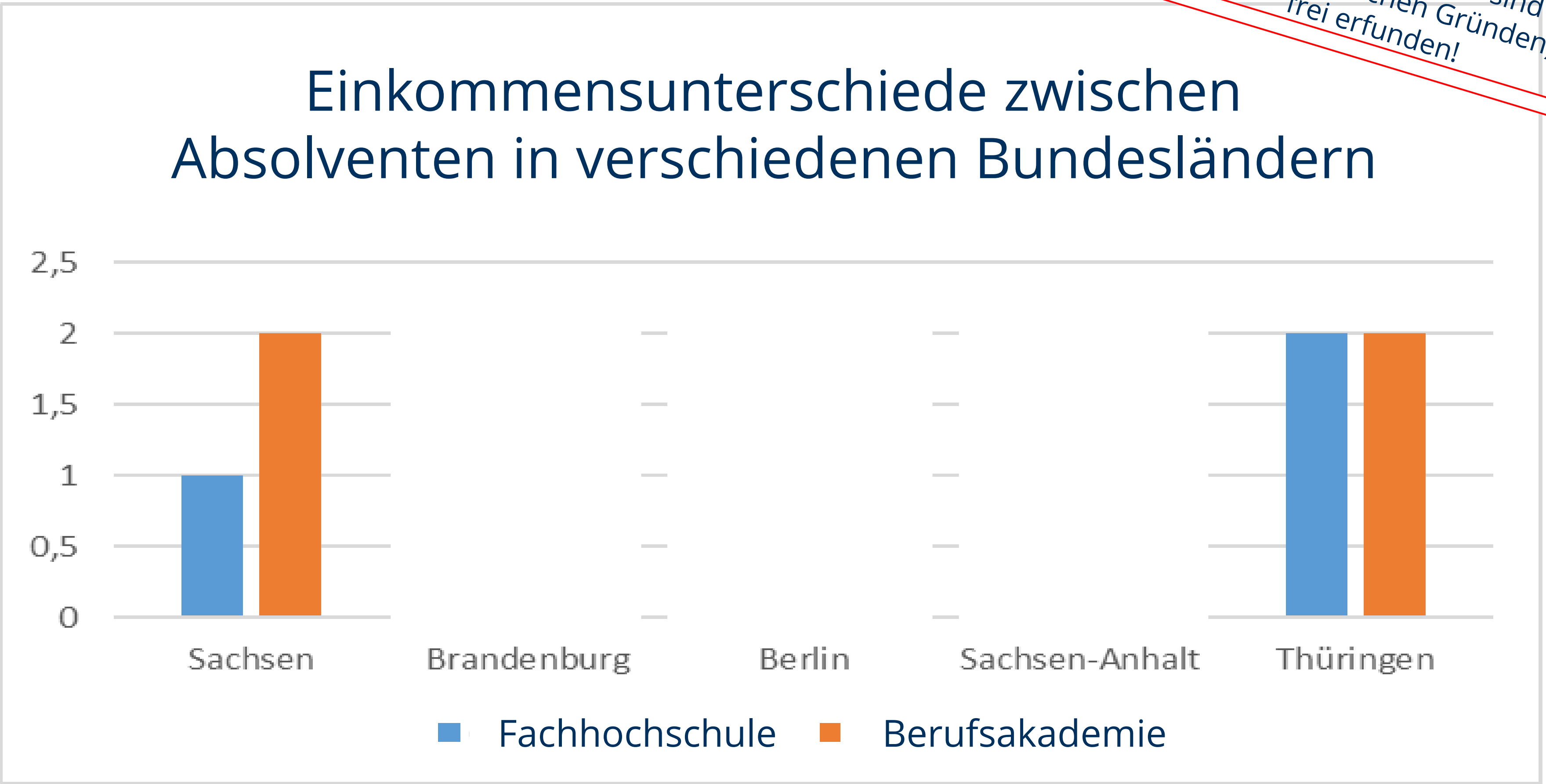
Hypothesen- oder Signifikanztests

Schließende Statistik



Wo ist der signifikante Unterschied?

Achtung: Die Daten sind (aus didaktischen Gründen) frei erfunden!



Nur nach Augenschein sehen wir …

sicher in Sachsen gibt es Unterschiede und in Thüringen gar nicht.

in Brandenburg gibt es **vermutlich** Unterschiede und in Sachsen-Anhalt **eher nicht**

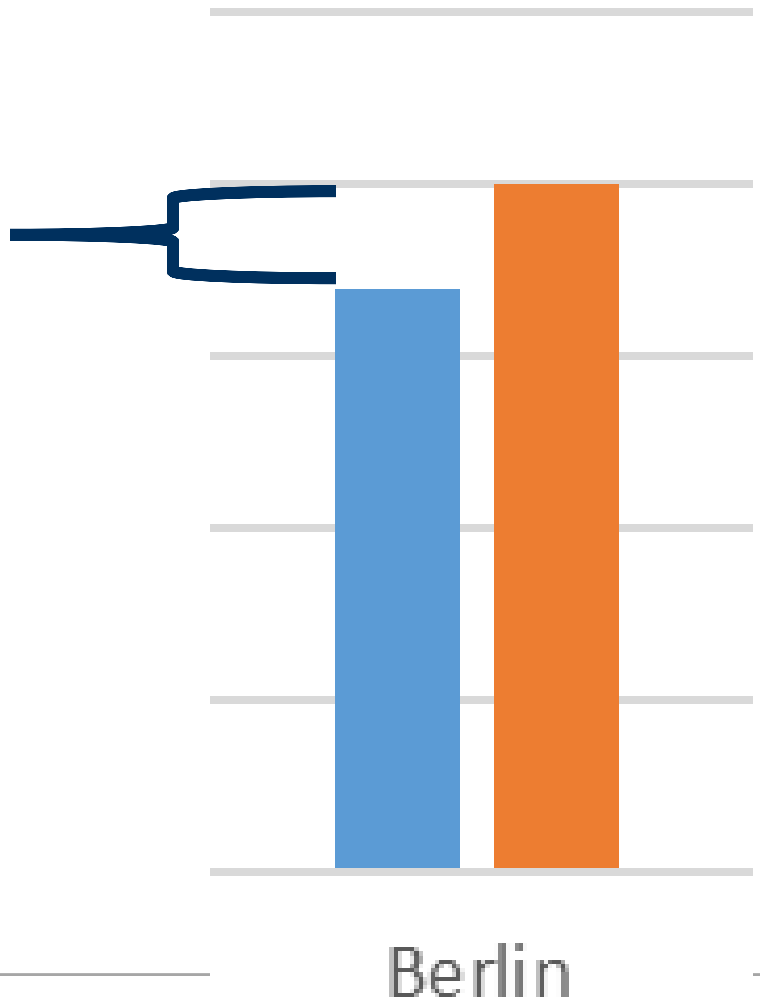
… und was ist nun mit Berlin? Wir brauchen einen Siginifikanztest.

Die Ausgangsfrage bei Signifikanztests

Zur Info! Signifikanz wird hier am Beispiel Mittelwerte verglichen. Weil dies anschaulicher ist als andere Fragestellungen.

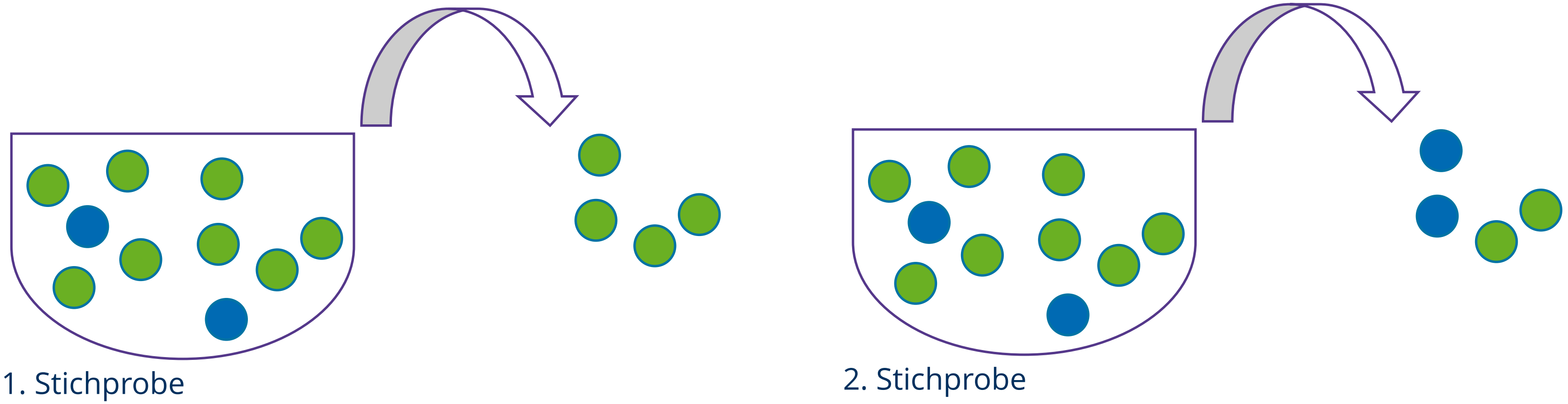
Ist der Unterschied wesentlicher Natur, oder ist das Ergebnis zufällig zustande gekommen?

Das ist DIE FRAGE, die statistische Tests beantworten.



Erinnerung: Warum gibt es zufällige Ereignisse?

Weil wir Stichproben untersuchen.



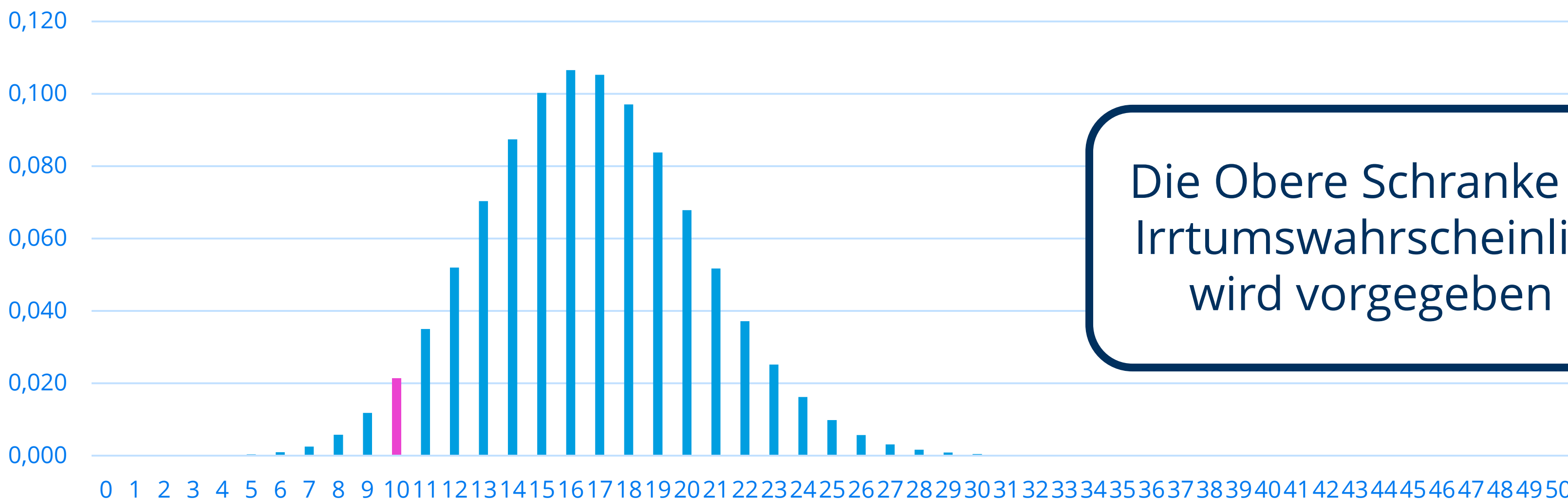
Irrtumswahrscheinlichkeit festlegen. Ein Beispiel für den Einstieg.

Sie möchten wissen, ob ein Würfel fair ist. Sie würfeln 100 Mal und haben 10 mal die 6.

Die Wahrscheinlichkeit dafür ist laut Binominalverteilung 4,3 %. (Vgl. Binominalverteilung)

Wir definieren, dass Ereignisse mit einer Wahrscheinlichkeit unter 5 % zu gering sind.

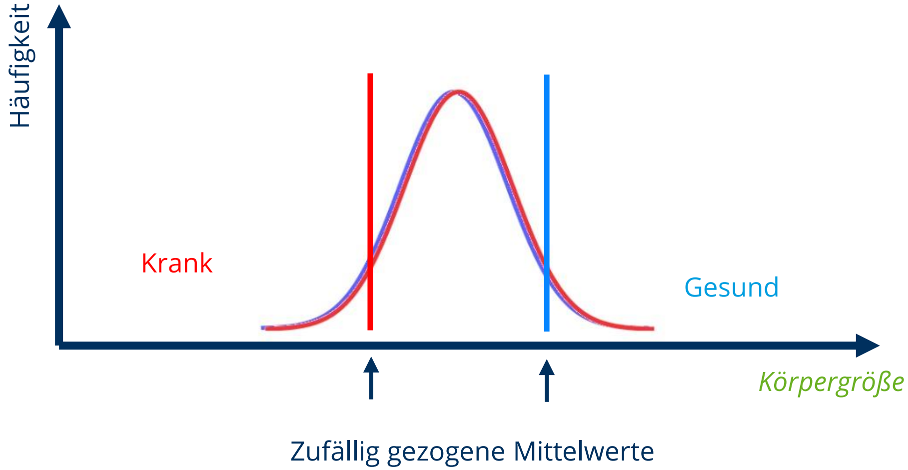
Wir schlussfolgern, dass der Würfel gezinkt ist. (Wissen tuen wir es aber nicht).



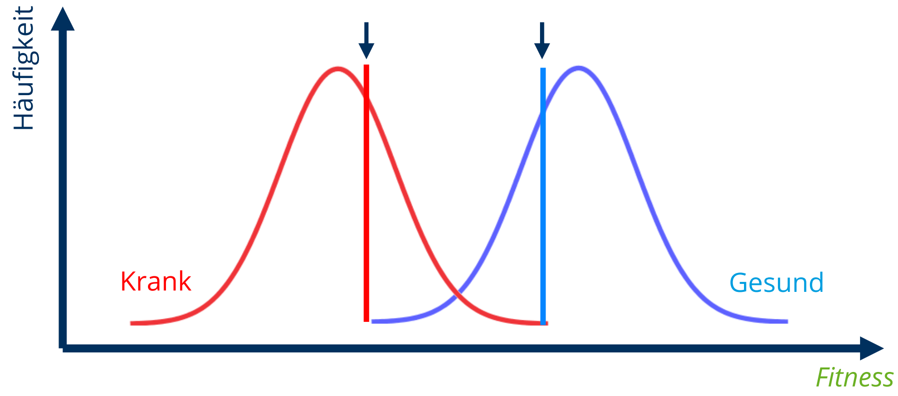
Die Obere Schranke für die Irrtumswahrscheinlichkeit, wird vorgegeben 5 %.

So ein Ergebnis kann prinzipiell auftreten, wenn der Würfel fair ist. Es ist nur sehr unwahrscheinlich.

Zufällige Ereignisse bei zwei Stichproben



Ergibt sich eine Differenz der Mittelwerte in zwei Gruppen kann nicht sofort geschlossen werden, dass die beiden Gruppen sich „wirklich“ von einander unterscheiden.



Ob ein Unterschied „signifikant“ ist kann nur ein Signifikanztest beantworten.

Dieser Test ermittelt auf Grund der Kennwerte der Verteilung und der Stichprobengröße, ob die gefundene Differenz signifikant ist.

Was brauchen statistische Tests und was liefern sie?

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.

Input	Statistische Test	Output
Datenwerte aus der Stichprobe (z.B. Mittelwert, Standard-abweichung und Stichproben-größe)	Berrechnungs-vorschriften (Sinn und Voraussetzungen des Tests müssen bekannt sein.)	Prüfgröße und andere Werte (Anhand dieser Werte erkennt man, ob z.B. ein Unterschied signifikant ist.)

Was machen statistische Tests?

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.

Statistische Test

Aufgrund eines Kennwertes (z.B. Mittelwerte zweier Stichproben, bzw. Gruppen) wird die Nullhypothese untersucht.

Beispiel für die Nullhypothese: „Die Mittelwerte der beiden Stichproben unterscheiden sich nicht voneinander.“

Statistische Tests: Null und Alternativhypothese

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.

Nullhypothese

- Ein statistischer Test prüft, ob kein Unterschied vorhanden ist/kein Zusammenhang da ist.
- H_0 = Nullhypothese:
- Bsp.: Es gibt **keinen** Unterschied in der Ausprägung einer Variable zwischen verschieden Gruppen.
- Als Formel: $\mu_0 = \mu_1$

Achtung!
Die Tests überprüfen die Nullhypothese. Die Alternativhypothese wird **nicht** von den Tests überprüft.

Alternativhypothese

- Die Alternativhypothese ist das Gegenteil der Nullhypothese.
- H_1 = Alternativhypothese:
- Bsp.: Es gibt **einen** Unterschied in der Ausprägung einer Variable zwischen verschieden Gruppen.
- Als Formel: $\mu_0 \neq \mu_1$

Hypothesen im Überblick

Forschungshypothese:

Im Durchschnitt produziert Maschine A längere Nudeln als Maschine B.

Nullhypothese H0: Der Mittelwert der Nudeln unterscheidet sich nicht.

$(\mu_a = \mu_b)$

Alternativhypothese allgemein H1: Die Mittelwerte unterscheiden sich signifikant.

$(\mu_a \neq \mu_b)$

Was sagt der Output eines statistischen Tests

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.

Statistische Test

Allg.

Statistischer Prüfgröße: z.B. t-Wert
Weiter Werte
z.B. Freiheitsgrade

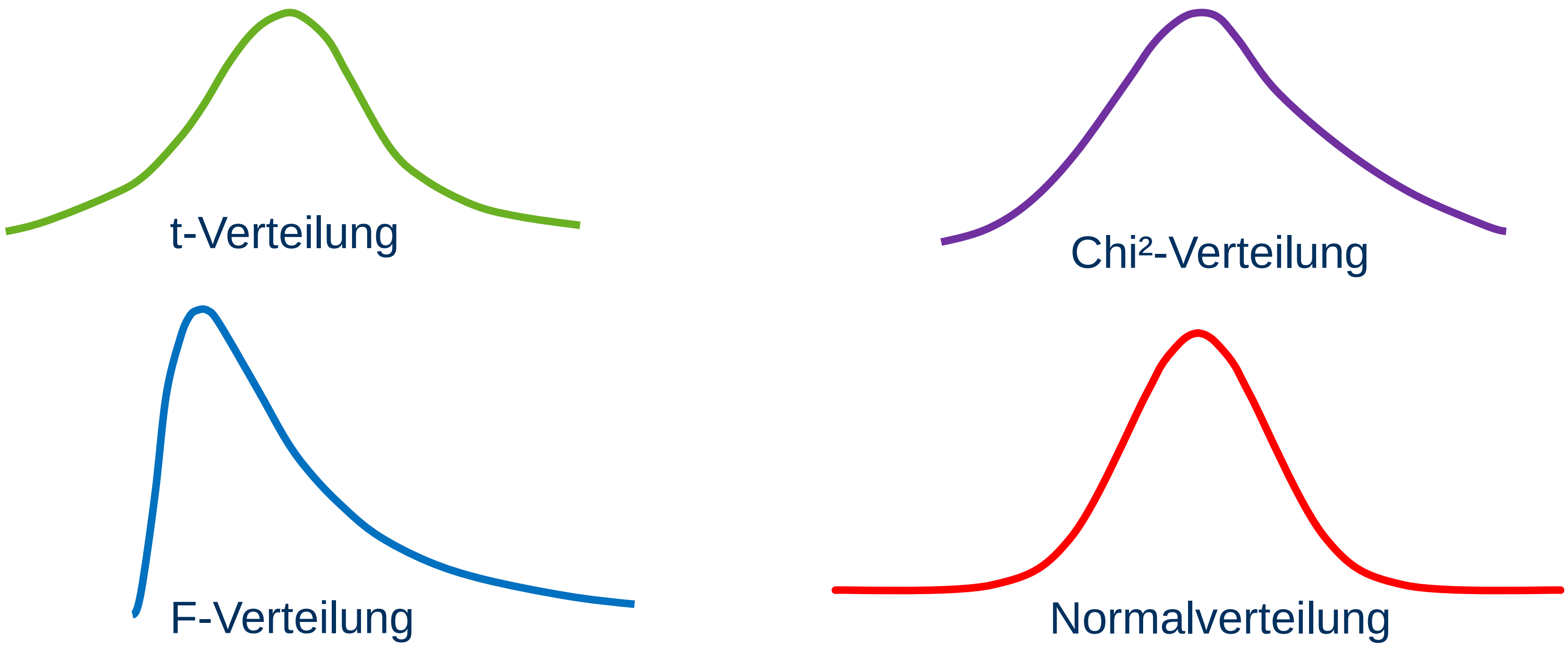
(Ermittlung des zugehörigen Signifkanzniveaus siehe folgende Folien)

Soft ware

In statistischer Software wird auch der statistische Wert angezeigt aber es wird sofort das zutreffende **Signifikanzniveau** ermittelt.

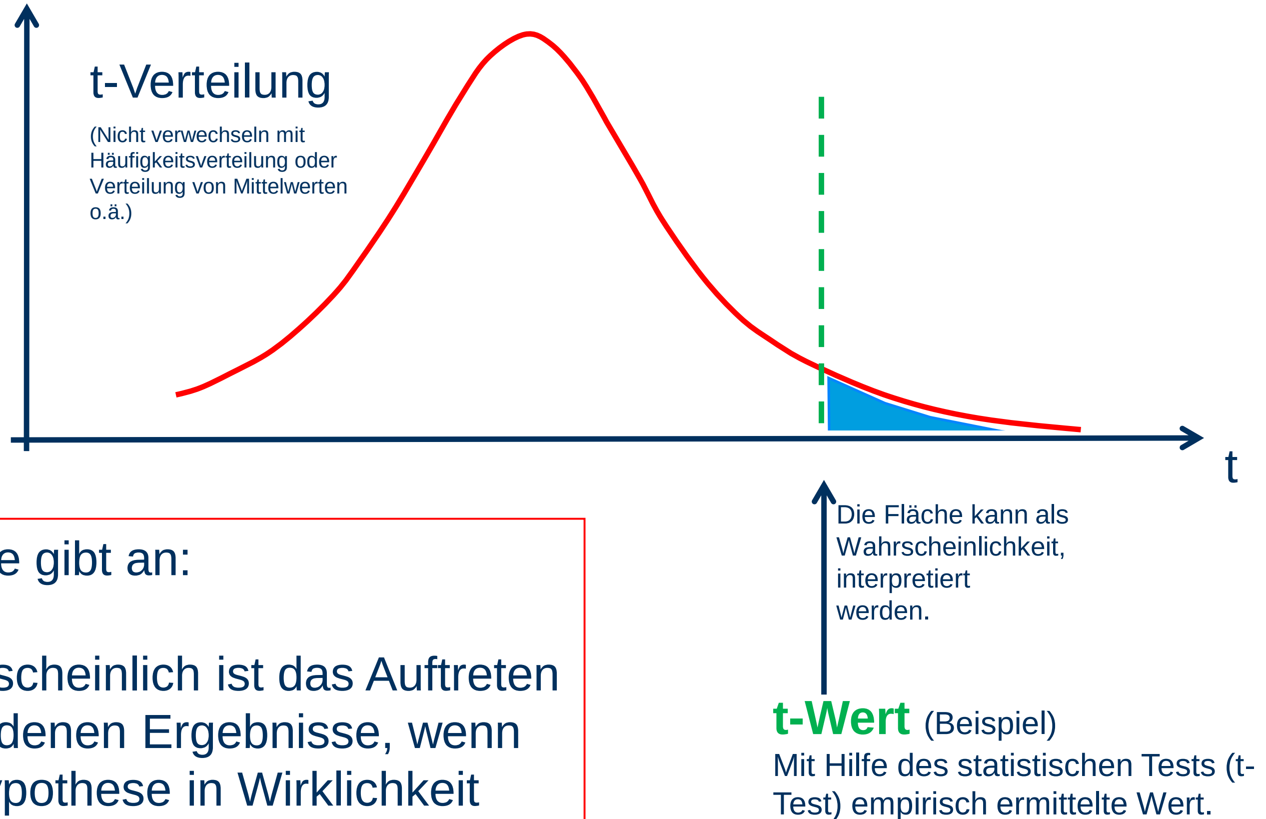
Hintergrund für statistische Tests

Jeder Wert eines statistischen Test kann – einer für den Test bekannten – Verteilungsform zugeordnet werden.



Bedeutung des t-Wertes

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.



Die Fläche gibt an:

Wie wahrscheinlich ist das Auftreten der gefundenen Ergebnisse, wenn die Nullhypothese in Wirklichkeit zutrifft.

Arbeiten mit der Tabelle der Verteilungsfunktion

4

Umrechnung:

$1 - 0{,}95 = 0{,}05$

Das ist das relevante Signifikanzniveau.

Beispiel:

$df = 16$

$t\text{-Wert} = 2{,}0$

1

2

Die nächst kleinere Zahl zum gefundenen t-Wert suchen.

3

Ablesen der zugehörigen Fläche (Fläche von links).

Freiheitsgrade:

$df = n_1 + n_2 - 2$

Tabelle D. Verteilungsfunktion der t-Verteilungen und zweiseitige Signifikanzgrenzen für Produkt-Moment-Korrelationen (zit. nach Glass, G.V., Stanley, J.C.: Statistical methods in education and psychology, p. 521. New Jersey: Prentice-Hall, Englewood Cliffs 1970)

Fläche* df	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,990	0,995	0,9995	$r_{0,05}$	$r_{0,01}$
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	[illegible]	12,706	31,821	63,657	636,619	0,997	1,000
2	0,142	0,289	0,445	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4,303	6,965	9,925	31,598	0,950	0,990
3	0,137	0,277	0,424	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3,182	4,541	5,841	12,941	0,878	0,959
4	0,134	0,271	0,414	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,776	3,747	4,604	8,610	0,811	0,917
5	0,132	0,267	0,408	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,571	3,365	4,032	6,859	0,754	0,874
6	0,131	0,265	0,404	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,447	3,143	3,707	5,959	0,707	0,834
7	0,130	0,263	0,402	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,365	2,998	3,499	5,405	0,666	0,798
8	0,130	0,262	0,399	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,896	3,355	5,041	0,632	0,765
9	0,129	0,261	0,398	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3,250	4,781	0,602	0,735
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	[illegible]	[illegible]	[illegible]	3,169	4,587	0,576	0,708
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	[illegible]	2,201	2,718	3,106	4,437	0,553	0,684
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	[illegible]	2,179	2,681	3,055	4,318	0,532	0,661
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	[illegible]	2,160	2,650	3,012	4,221	0,514	0,641
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	[illegible]	2,145	2,624	2,977	4,140	0,497	0,623
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	[illegible]	2,131	2,602	2,947	4,073	0,482	0,606
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	$\boxed{1{,}746}$	2,120	2,583	2,921	4,015	0,468	0,590
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	[illegible]	1,740	2,110	2,567	2,898	3,965	0,456	0,575
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922	0,444	0,561
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883	0,433	0,549
20	0,127	0,257	0,391	[illegible]	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850	0,423	0,537
21	0,127	[illegible]	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819	0,413	0,526
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792	0,404	0,515
[illegible]	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767	0,396	0,505
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745	0,388	0,496
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725	0,381	0,487
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707	0,374	0,478
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690	0,367	0,470
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674	0,361	0,463
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659	0,355	0,456
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646	0,349	0,449
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551	0,304	0,393
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460	0,250	0,325
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373	0,178	0,232
z	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291		

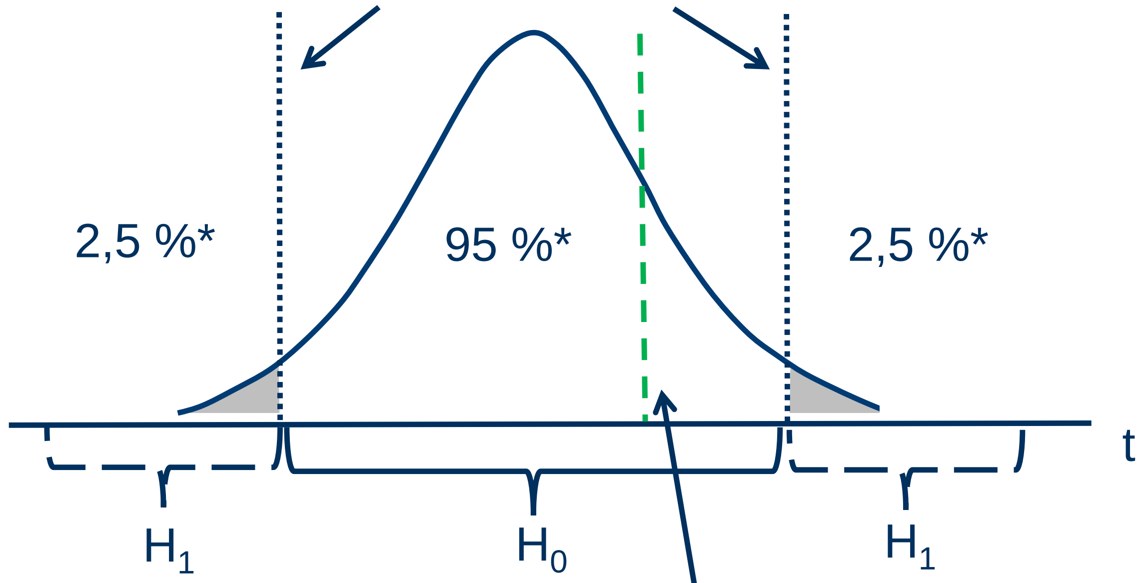
* Die Flächenanteile für negative t-Werte ergeben sich nach der Beziehung $p(-t_{df}) = 1 - p(t_{df})$

Quelle der Abbildung: Bortz, J. (1999): Statistik für Sozialwissenschaftler. Berlin u.a.: Springer, S. 775

Festlegung des Signifikanzniveaus

Zur Info! Statistische Tests werden hier am Beispiel des Mittelwertvergleichs erläutert.

(Willkürliche) Festlegung eines „kritischen“ Wertes ab dem der statistische Wert als „unwahrscheinlich“ gelten soll (Signifikanzniveau) z.B. insg. 0,05 (zwei mal 2,5)



* Die Prozente beschreiben, wie wahrscheinlich ein Ergebnisse wäre, falls H_0 in Wirklichkeit gilt.

t-Wert (Beispiel)
Mit Hilfe des statistischen Tests (t-Test) empirisch ermittelte Wert.

Unterschied gerichtete und ungerichtete Alternativhypothesen

Bekannt:

Alternativhypothese allgemein H1: Die Mittelwerte unterscheiden sich signifikant. ($\mu_a \neq \mu_b$)

Neu:

1. gerichtete Alternativhypothese H1: Der Mittelwert der Nudeln aus Maschine A ist signifkant größer als aus Maschine B.

   ($\mu_a > \mu_b$)

2. gerichtete Alternativhypothese H1: Der Mittelwert der Nudeln aus Maschine A ist signifkant kleiner als aus Maschine B.

   ($\mu_a < \mu_b$)

Es macht einen Unterschied, ob Sie eine gerichtete oder eine ungerichtete Alternativhypothese haben. Bei einer ungerichteten Hypothese

Allgemeines Vorgehen bei einem Signifikanztest

Daten vorbereitet

Hypothese formulieren

Heraussuchen des richtigen Tests

Ermittlung der Testgröße und des Signifikanzniveau (mit Hilfe von Software).

Entscheidung fällen ob H_0 beibehalten wird oder zugunsten der H_1 zurückgewiesen wird.

Ergebnisse formulieren (siehe folgende Folien)

Formulierung der Ergebnisse

Signifikant:

- Allgemein: Die Nullhypothese wird zugunsten der Alternativhypothese zurückgewiesen.
- Beispiel: „Es konnte ein signifikanter Unterschied im IQ zwischen Hauptschülern und Gymnasiasten festgestellt werden."

Nicht signifikant:

- Allgemein: Die Nullhypothese wird beibehalten/kann nicht zurückgewiesen werden.
- Beispiel: „Lehrer die sich in ihrer Freizeit mit Wirtschaft beschäftigen, zeigen kein signifikant höheres Interesse am Fach Wirtschaft als anderer Lehrer."

Interpretation von Hypothesentests

Wenn signifikantes Ergebnis ($p < 0{,}05$) → **Ablehnung/Zurückweisung** der Nullhypothese zugunsten der Alternativhypothese

Wenn kein signifikantes Ergebnis → **Beibehaltung** der Nullhypothese

Übung: Interpretation von signifikanten / nicht signifikanten Ergebnissen

Formulieren Sie folgende drei Ergebnis ausführlich:

Vergleich des Mittelwerts der Variable: „Motivation“

Gruppe 1: Auszubildende die selbstgesteuert lernen dürfen

Gruppe 2: Auszubildende mit einem festen Lehrplan

Ergebnis: **signifikant**

Vergleich des Mittelwerts der Variable: „Chancen für ein Job“

Gruppe 1: Schüler mit sechs Wochen Praktikum

Gruppe 2: Schüler mit einer Woche Praktikum

Ergebnis: **signifikant**

Vergleich des Mittelwerts der Variable: „Noten im Fach Statistik.“

Gruppe 1: Studentengruppe 1

Gruppe 2: Studentengruppe 2

Ergebnis: **insignifikant**

Übung Signifikanz

Was bedeutet Folgendes.

- In einem Artikel heißt es „Die Ergebnisse unserer Analyse sind signifikant (p = 0,001)“
- Drei Wissenschaftsteam machen alle drei jeweils eine Studie zur gleichen Frage. Sie erheben unabhängig Daten voneinander. Team A hat ein nicht signifikantes Ergebnis, Team B ein signifikantes Ergebnis (p = 0,05) und Team C ein signifikantes Ergebnis mit einem Signifikanzniveau von p = 0,01.
- Eine Forscherin erzählt, dass der Zusammenhang zwischen X und Y hochsignifikant ist.

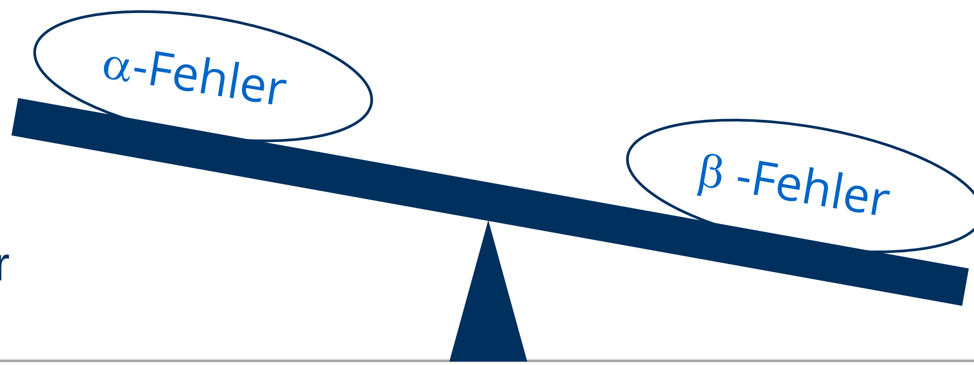
Hypothesen Quiz

	a	b
Die Forschungshypothese stimmt meist mit a.) der Nullhypothese oder b.) der Alternativhypothese überein.		
Richtig formuliert lautet eine Nullhypothese a.) $\mu_a = \mu_b$ oder b.) $\mu_a \neq \mu_b$		
Die Formulierung der Forschungshypothese kommt a.) vor oder b.) nach dem Aufstellen der Nullhypothese.		
Wenn Sie wissen wollen, ob zwei Sorten Reis sich in der Korngröße unterscheiden heißt die Alternativhypothese a.) Sorte A unterscheidet sich signifikant von B oder b.) Sorte A und B unterscheiden sich nicht signifikant voneinander.		

Fehler bei Signifikanztests

Für H1 entschieden, obwohl H0 gilt.

		In der Population/"Wahrheit" gilt	
		H_0	H_1
Entscheidung aufgrund der Stichprobe zugunsten	H_0	Richtige Entscheidung	β-Fehler Fehler 2. Art
	H_1	α-Fehler Fehler 1. Art (Hängt von Signifikanzniveau ab)	Richtige Entscheidung



Bei Verringerung des α-Niveaus (Signifikanzniveau) wird die Wahrscheinlichkeit für einen β-Fehler größer.

Übung Fehler 1. und 2. Art

Normalerweise kennen wir die Wahrheit nicht, welchen Fehler wir machen, daher hier konstruierte Beispiele:

Eine Personalabteilung einer großen Firma weiß genau, dass bei der Einstellung von neuen Mitarbeiterinnen und Mitarbeitern keine Gehaltsunterschiede zwischen Frauen und Männern gemacht werden. Nun gibt es eine Studie des Betriebsrat, die zu einem anderen Ergebnis kommt. Welcher Art des Fehler wurde hier gemacht?

Eine Firma produziert handgefertigte Porzellanteller. Je teurer, desto aufwendiger die Malerei. Sie schauen sich die Teller an und kommen zu dem Schluss, dass es keinen Zusammenhang zwischen Preis und Aufwand gibt. Welche Art des Fehlers haben Sie gemacht?

Wir wissen alle, die Kinder kommen nicht von den Störchen. Aber es gibt diese falsche Annahme von einen Zusammenhang. Kann hier ein Fehler 2. Art gemacht wurden sein?

Fehler 1. Art und 2. Art sind unterschiedlich gefährlich, je nach Kontext

Sie messen die durchschnittliche Anzahl von Streichhölzern. Und wollen wissen, ob die Anzahl signifikant von den versprochenen 38 abweicht. Nun hängt es vom Kontext ab, wie gefährlich ein Fehler 1. und 2. Art ist.

A. Sie machen diese Untersuchung privat, nur aus Neugier. Dann ist jeder Art von Fehler ________________

B. Sie machen diese Untersuchung für die Streichholzproduktion, dann wird jede Art des Fehlers ________________

Fehler 1. Art = sie finden einen Unterschied wo keiner ist: eventuell kommt es teuren Neuanschaffungen in der Produktion.

Fehler 2. Art = Sie finden keinen Unterschied wo eigentlich einer ist: Die Firma „wirbt" mit 38 Zündhölzern und später wird sie vom Verbraucherschutz angezeigt, weil weniger drin sind.

C. Sie machen diese Untersuchung für eine Antarktisexpedition, wo jedes Streichholz zählt. Dann könnte es im Extremfall sogar _____________ sein.

Übungsbeispiel Interpretation von signifikanten Ergebnissen

Gruppe	n	Mittelwert der Gelenkigkeit (1-10)
3 – 22	65	9,4
23 – 42	234	9,5
43 – 62	645	8,3
63 – 82	26	5,5

Ein Forscherteam hat untersucht, ob es einen Zusammenhang zwischen dem Alter und der Gelenkigkeit gibt, bei Menschen, die intensiv klettern. Sie haben insgesamt 3000 Leute untersucht. Sie haben alle Teilnehmer in Gruppen geteilt und alle Gruppen miteinander verglichen.

Sie finden keine signifikanten Unterschiede zwischen allen Gruppen.

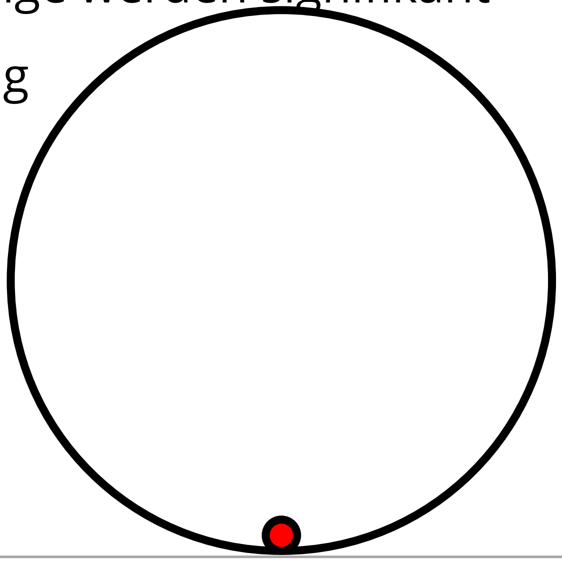
1. Gibt es wirklich keine Unterschiede?
2. Warum wurden keine Unterschiede gefunden?
3. Wie könnte man das anderes untersuchen?

Achtung die Daten sind frei erfunden.

Beachten bei statistischen Tests: Größer der Stichprobe

- **Zu kleine Stichproben:** Selbst starke Zusammenhänge werden von statistischen Test nicht als signifikant angesehen.
  - → größer Stichprobe
  - → Wiederholung der Studie

- **Größere Stichproben**: schwächste Zusammenhänge werden signifikant
  - → Relevanz ist neben Signifikanz auch wichtig



Beachten bei statistischen Tests: Theorie / Voraussetzungen

- Gefahr: Alle möglichen Zusammenhänge mal testen und dann erst Theorie
  - → Erst Theorie, dann testen
  - → Aber: BigDataAnalysen gehen einen anderen Weg

- Jeder statistische Test hat verschiedene **Voraussetzungen**.
- Viele dieser Voraussetzungen sind einleuchtend oder auf einen Blick zu überprüfen (Anzahl der Variablen/Anzahle der Fälle)
- Andere Voraussetzungen muss man erst mit bestimmten statistischen Prozeduren testen (Test auf Normalverteilung)

- Ein statistischer Test reicht nicht um die „Wahrheit" zu erfahren → Wiederholung der Studie ist notwendig

Siehe auch Quatember, Andreas (2008): Statistik ohne Angst vor Formeln. Studienbuch für Wirtschafts- und Sozialwissenschaftler. Pearson Studium. 2. aktualisierte Auflage, S. 185ff.

Übung: Aus meiner Praxis:

Ein Team von Forschern untersucht den Unterschied von 10 Variablen. Sie vergleicht jede Variable mit jeder und macht so 100 Vergleichen. In diesen Vergleichen erhält sie 2 signifikante Ergebnisse. Diese signifikanten Ergebnisse werden nun genauer untersucht. Sie finden jedoch keine Erklärung.

Was passiert hier?

Warnung vor zu vielen Signifikanztest (nicht vor zu viel einzelnen Studien)

Das-Niveau sagt uns: Falls In Wirklichkeit kein Zusammenhang/Unterschied vorhanden ist, kann das gefundene Ergebnis trotzdem mal zufällig auftreten.

Bsp.: Dazu wird 20 Mal eine Population, in der **kein** Zusammenhang zwischen zwei Merkmalen vorhanden ist, untersucht. Dazu ziehen wir 20 mal eine Stichprobe aus dieser Population.

Das Signifikanzniveau wird auf 5 % festgelegt.

- Beim 20. Mal erbringt der Test ein signifikantes Ergebnis. (Obwohl in Wirklichkeit kein Zusammenhang da ist)
- Dies ist auch ein Grund, warum man wissenschaftliche Ergebnisse immer mehrmals überprüfen sollte.

Stichprobe	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Ergebnis signifikant	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	JA

Vorsicht bei der Interpretation von (Nicht-) Signifikanten Ergebnissen

Beispiel: Sie untersuchen den Unterschied in Anzahl der beruflichen Kontakte zwischen zwei Mitarbeiterteams (Außendient / Innendienst). Sie finden keinen Unterschied. Gibt es wirklich keinen Unterschied?

Student t-Test

Zwei Unterschiedshypothesen

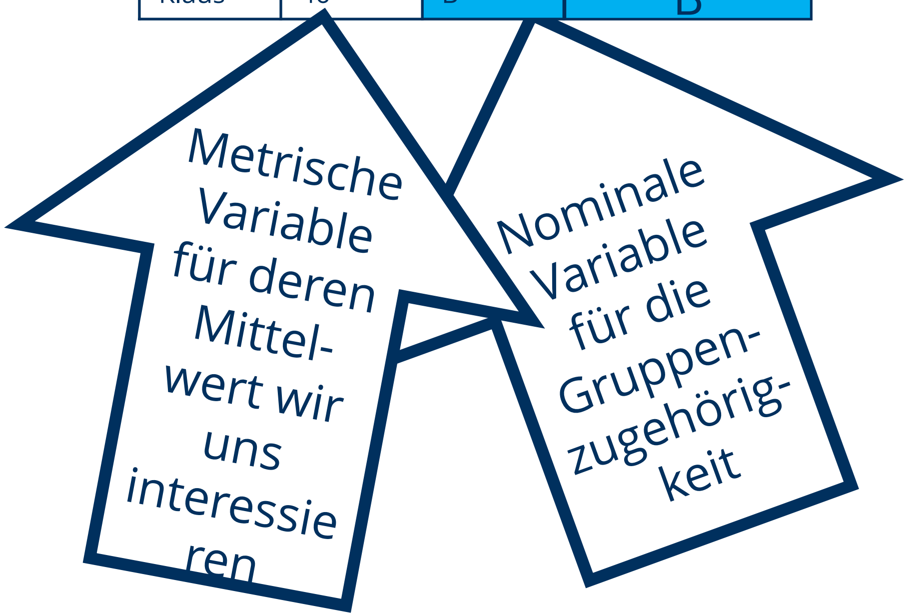
t-Test bei unabhängigen Stichproben

- Unabhängige (ungepaarte) Stichprobe
- t-Test bei unabhängigen Stichproben (metrisch Daten)

Name	Leistungstest	**Gruppe**	
Ilsa	33	A	Gruppe: A
Lisa	44	A	
Peter	22	B	Gruppe: B
Klaus	46	B	

t-Test bei zwei unabhängigen/ungepaarten Stichproben

Vornam e	Leistun gstest	Gruppe	
Ilsa	33	A	Gruppe: A
Peter	44	A	
Lisa	22	B	Gruppe: B
Klaus	46	B	



Frage des t-Test: Sind die Mittelwerte in zwei Gruppen unterschiedlich groß?

H0 = Mittelwerte in den zwei untersuchten Gruppen sind gleich

Skalenniveau:

abhängige Variable: metrisch;

unabhängige Variable: nominal/ordinal mit zwei Ausprägungen (z. B. Geschlecht; Gruppe A/B)

t-Test bei zwei unabhängigen/ungepaarten Stichproben: Die Formel

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

Diese Formel könnten Sie in Excel selbst modellieren, für die Interpretation benötigen Sie jedoch Tabellen für den t-Test. Als Signifikanzniveau wird allgemein 0,05 festgelegt.

Freiheitsgrad: $F = n_x + n_y - 2$ Den Freiheitsgrad braucht es für die Tabelle.

Siehe: https://de.wikipedia.org/wiki/Studentsche_t-Verteilung

Wenn der ermittelte t-Wert größer ist als der t-Wert in der Tabelle bei dem ermittelten Freiheitsgrad, dann kann von einem signifikanten Ergebnis ausgegangen werden.

	0.95
n=1	6,314
10	1,812
11	1,796
12	1,782
13	1,771
14	1,761
15	1,753
16	1,746
17	1,740
18	1,734
19	1,729
20	1,725
21	1,721
22	1,717
23	1,714
24	1,711
25	1,708
26	1,706
27	1,703
28	1,701
29	1,699
30	1,697

Ausschnitt aus der Tabelle der t-Quantile für 0,05 https://de.wikipedia.org/wiki/Studentsche_t-Verteilung

t-Test: Formulierung der Ergebnisse

- Sehr exakte Formulierung: *Die Nullhypothese, dass beide Geschlechter das Gleiche Gewicht haben, kann nicht aufrechterhalten werden. Sie wird zugunsten der Alternativhypothese – das Gewicht unterscheidet sich zwischen Männern und Frauen – zurückgewiesen. Der zum t-Test zugehörige p-Wert (0,014) liegt unter dem Signifikanzniveau von 0,05*
- Umgangssprachliche Formulierung ohne Erwähnung von H_0 und $H_{1)}$: *Die Ergebnisse weisen darauf hin, dass Frauen und Männer signifikant unterschiedliche viel Gewicht auf die Waage bringen. Der Mittelwert des Gewichts von Männer (84,58 Kilogramm) ist signifikant höher als der von Frauen (71,33 Kilogramm).*

Übung t-Test

Sie haben folgende Werte von Wettkampfteilnehmern. Es handelt sich um die erzielten Wettkampfpunkte in zwei Trainingsgruppen (A, B). Gruppe A hat Hallentraining gemacht und Gruppe B Outdoor-Training.

- Überprüfen Sie, ob die Werte der beiden Gruppen sich signifikant voneinander unterscheiden (Signifikanzniveau von 0,05).
- Formulieren Sie richtige Null- und Alternativhypothesen.
- Errechnen Sie in Zwischenschritten die Mittelwerte, Standardabweichung, den t-Wert und die Freiheitsgrade.
- Interpretieren Sie das Ergebnis inhaltlich. Welche Schlussfolgerung für die Trainingspraxis lassen sich ableiten?

Gruppe A	Gruppe B
56	12
45	34
74	85
84	75
34	54
57	37
64	56
78	37
64	65
35	42
97	78
65	68
43	98

t-Test bei zwei unabhängigen/ungepaarten Stichproben in Excel I

2 Variablen sind notwendig

1. Variable enthält die Gruppen (z.B. Frau/Mann), das gesamte Datenblatt muss danach sortiert werden
2. metrische Variable, deren Mittelwert uns interessiert

→ Daten → Datenanalyse → Zweistichproben-t-Test (Gleicher Varianzen/Unterschiedlicher Varianzen)

Ob die Varianzen gleich sind müsste vorher getestet werden

… Fortsetzung …

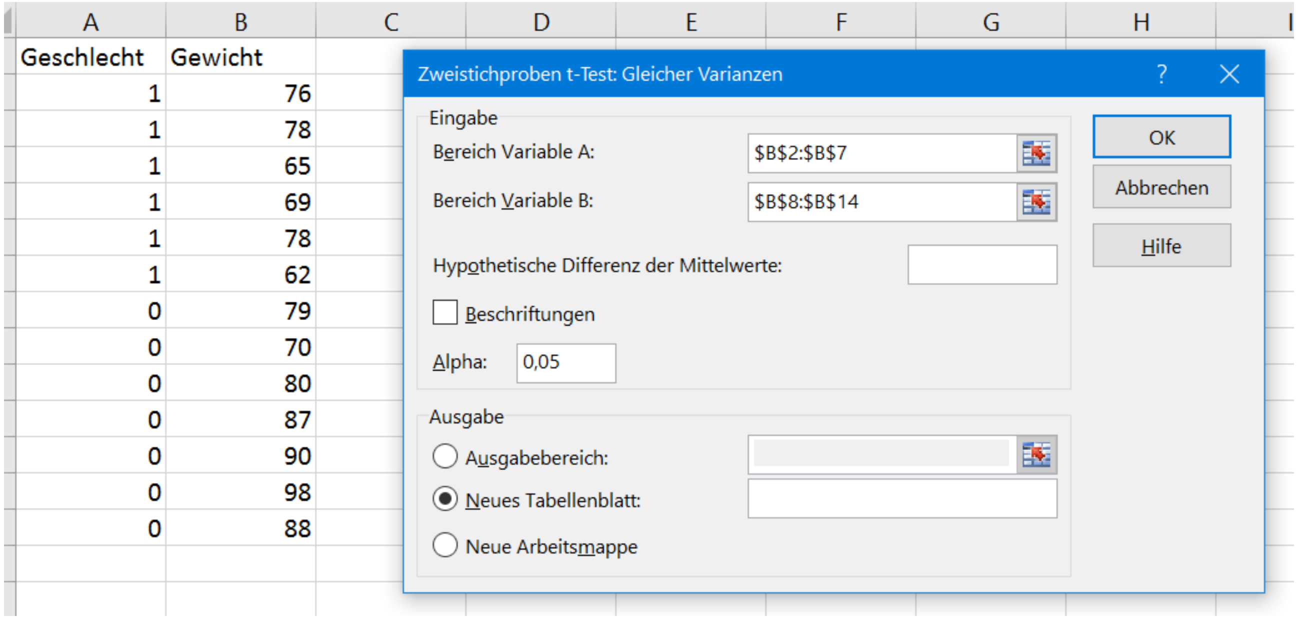
Das Excel-Add-In Datenanalyse muss aktiviert sein

t-Test bei zwei unabhängigen/ungepaarten Stichproben in Excel II

In „Bereich Variable A“ die Daten der 2. Variable ziehen, die alle zur 1. Gruppe (z.B. Frauen) gehören.

In „Bereich Variable B“ die Daten der 2. Variable ziehen, die zur 2. Gruppe gehören.

→ Ok



t-Test: Output

Zweistichproben t-Test unter der Annahme gleicher Varianzen		
	Variable 1	*Variable 2*
Mittelwert	71,333	84,571
Varianz	48,667	81,952
Beobachtungen	6	7
Gepoolte Varianz	66,823	
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	11	
t-Statistik	-2,911	
P(T<=t) einseitig	0,007	
Kritischer t-Wert bei einseitigem t-Test	1,796	
P(T<=t) zweiseitig	0,014	
Kritischer t-Wert bei zweiseitigem t-Test	2,201	

Mittelwerte beider Gruppen

p-Wert. hier liegt der p-Wert unter (unserem festgelegtem) Signifikanzniveau von 0,05 → signifikante Unterschiede

In der Verschriftlichung unbedingt die Mittelwerte und den p-Wert veröffentlichen. Profis veröffentlichen auch t-Statistik und Freiheitsgrade.

t-Test: Verletzung der Voraussetzung. Praktischer Hinweis

Bei dem vorangegangen Beispiel hatten wir angenommen die Varianzen des Gewichts bei Frauen und Männern sind gleich!

Tatsächlich scheinen sie nicht gleich gewesen zu sein, weshalb unter Datenanalyse eigentlich der Test „Zweistichproben t-Test: Unterschiedlicher Varianzen“ durchgeführt hätte werden sollen!

… Aber: der t-Test ist sehr robust gegen Verstöße der Voraussetzungen und meist kommen beide t-Testvarianten zu sehr ähnlichen Ergebnissen, so dass nur in Zweifelsfällen die Missachtung der Voraussetzungen die Ergebnisse beeinflusst.

(Zum Vergleich: Funktion T.TEST)

=T.TEST(...)

Geschlecht	Gewicht			
1	76			
1	78			
1	65			
1	69			
1	78			
1	62			
0	79			
0	70			
0	80			
0	87			
0	90			
0	98			
0	88			
		=T.TEST(B2:B7;B8:B14;2;2)		

T.TEST(Matrix1; Matrix2; Seiten; Typ)

Gerichtete Alternativhypothesen und einseitiger Signifikanztest.

Die Standardalternativhypothese lautet: Die Mittelwerte unterscheiden sich nicht.

Die gerichtete Alternativhypothese ist etwas genauer und behauptet nun, dass der Mittelwert der einen Stichprobe größer/kleiner als der Mittelwert der anderen Stichprobe ist.

Bei gerichteten Alternativhypothesen kann der p-Wert halbiert werden. Es handelt sich dann um einen einseitigen Signifikanztest. Der Standardfall ist der zweiseitige Signifikanztest.

Beispiel mit p-Wert von 0,07:

- 0,07 / 2 = 0,035
- Zweiseitig: 0,07 liegt über 0,05 und ist nicht signifikant
- Einseitig: 0,035 liegt unter 0,05 und ist signifikant

Haben wir ordinal skalierte Daten aber die gleiche Frage wie beim t-Test, verwendet man einen nichtparametrische Test: Man-Whitney U-Test (unabhängige Stichproben) Wilcoxon-Rangsummen-Test (abhängige Stichproben) die in Excel leider kompliziert umzusetzen sind.

Übung gerichtete Alternativhypothese

Betrachten Sie die Ausgabe für einen t-Test in Excel unten. Welchen Wert interpretieren Sie wenn sie eine gerichtete Hypothese haben? Wie ist noch mal der Zusammenhang zwischen einseitigen und zweiseitigen p-Wert?

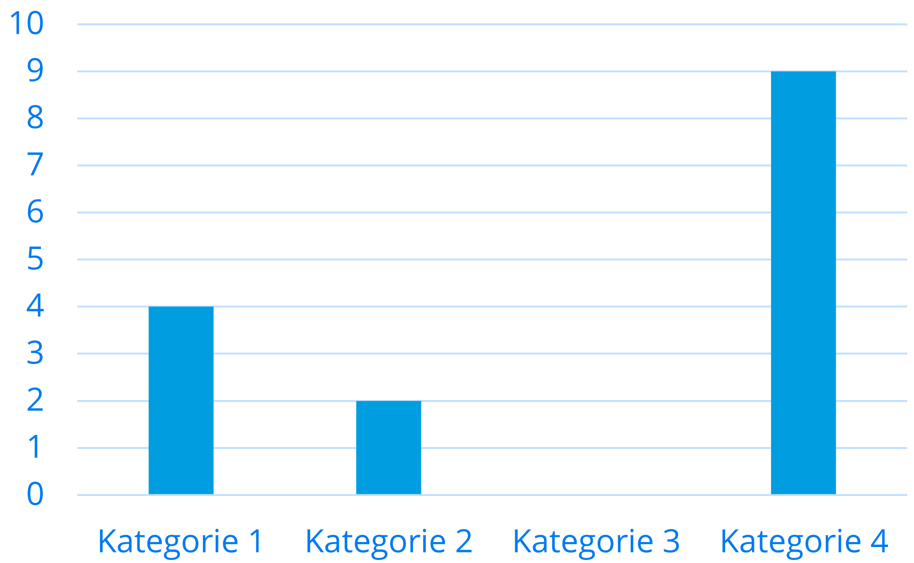
Zweistichproben t-Test unter der Annahme gleicher Varianzen		
	Variable 1	*Variable 2*
Mittelwert	71,333	84,571
Varianz	48,667	81,952
Beobachtungen	6	7
Gepoolte Varianz	66,823	
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	11	
t-Statistik	-2,911	
P(T<=t) einseitig	0,007	
Kritischer t-Wert bei einseitigem t-Test	1,796	
P(T<=t) zweiseitig	0,014	
Kritischer t-Wert bei zweiseitigem t-Test	2,201	

Sie wetten mit Ihrer Nachbarin, dass im Durchschnitt die gelben Rosen im Garten größer werden als die roten Rosen. Mit Ihrer Tante wetten Sie nur, dass die Länge der gelben Rosen sich signifikant von den roten Rosen unterscheiden. Mit wem haben sie eine gerichtete und mit wem eine ungerichtete Hypothese ausgemacht?

Besondere Voraussetzung beim t-Test

(Annähernd) Normalverteilung der abhängigen Variable

Sinn: Ansonsten könnte es sein, dass der Mittelwert keine sinnvolle Aussage über die Datenverteilung macht. (Siehe Beispiel in den Daten)



Weitere Voraussetzung beim t-Test

Gleichheit der Varianzen

Wie testet man auf Gleichheit der Varianzen?

Mit dem Levene-Test:

Sinn: Sind die Varianzen beider Stichproben gleich?

H0 = Die Varianzen beider Stichproben sind gleich.

Skalenniveau: metrisch

Beispiel in den Abbildungen:
Hier sind die Varianzen nicht gleich.
Die unterer Stichprobe hat viel kleinere Varianz

T-Test Online

Aufgabe: Öffnen Sie eine interaktive Seite zum t-Test und führen sie einen t-Test dort durch, vergleichen Sie mit den Ergebnissen aus Excel. https://statpages.info/

Übersicht weitere Tests

Es gibt verschiedene Arten von t-Tests

1. Intensiv hier behandelt: der t-Test bei unabhängigen Stichproben
2. Vergleich eines vorgegeben Mittelwerts/Anteilswerts aus der Grundgesamtheit mit dem Mittelwert/Anteilswert aus der Stichprobe (für Mittelwert: Gaußtest)
3. Vergleich von Differenzen von Anteilswerten
4. Vergleich von gepaarten/abhängigen Stichproben

Ordnen Sie die Beispiele zu:

a. Partei A hat signifikant weniger Sitze im Parlament als Partei B.
b. Ein Geschäft misst vor und nach der Einführung einer Werbung die Anzahl an täglichen Kunden im Geschäft.
c. Sie wollen wissen, ob dunkle oder Zartbitterschokoladenkugeln Ihnen mehr schmeckt.
d. In der Bevölkerung hat einen IQ von 100. Eine Direktorin möchte nun wissen, ob ihre Schüler über diesem Wert liegen.

Zweistichprobentest bei gepaarten / abhängigen Stichproben

t-Test bei gepaarten/abhängigen Stichproben

Sinn: Unterscheidet sich der Mittelwert von zwei Variablen.

H_0 = <u>Mittelwerte</u> in den <u>zwei</u> Variablen sind gleich.

Skalenniveau: Beide Variablen metrisch

Name	**Leistungstest vorher**	**Leistungstest nachher**
Ilsa	2	1
Peter	4	6
Lisa	9	7
Hans	2	5

Zweistichproben Test bei gepaarten/abhängigen Stichproben

$$Z = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

$\bar{d} =$
$Mittelwert\ der\ Differnz\ aller\ Paare$

$s_d =$
$Standardabweichung\ der\ Differenzen$

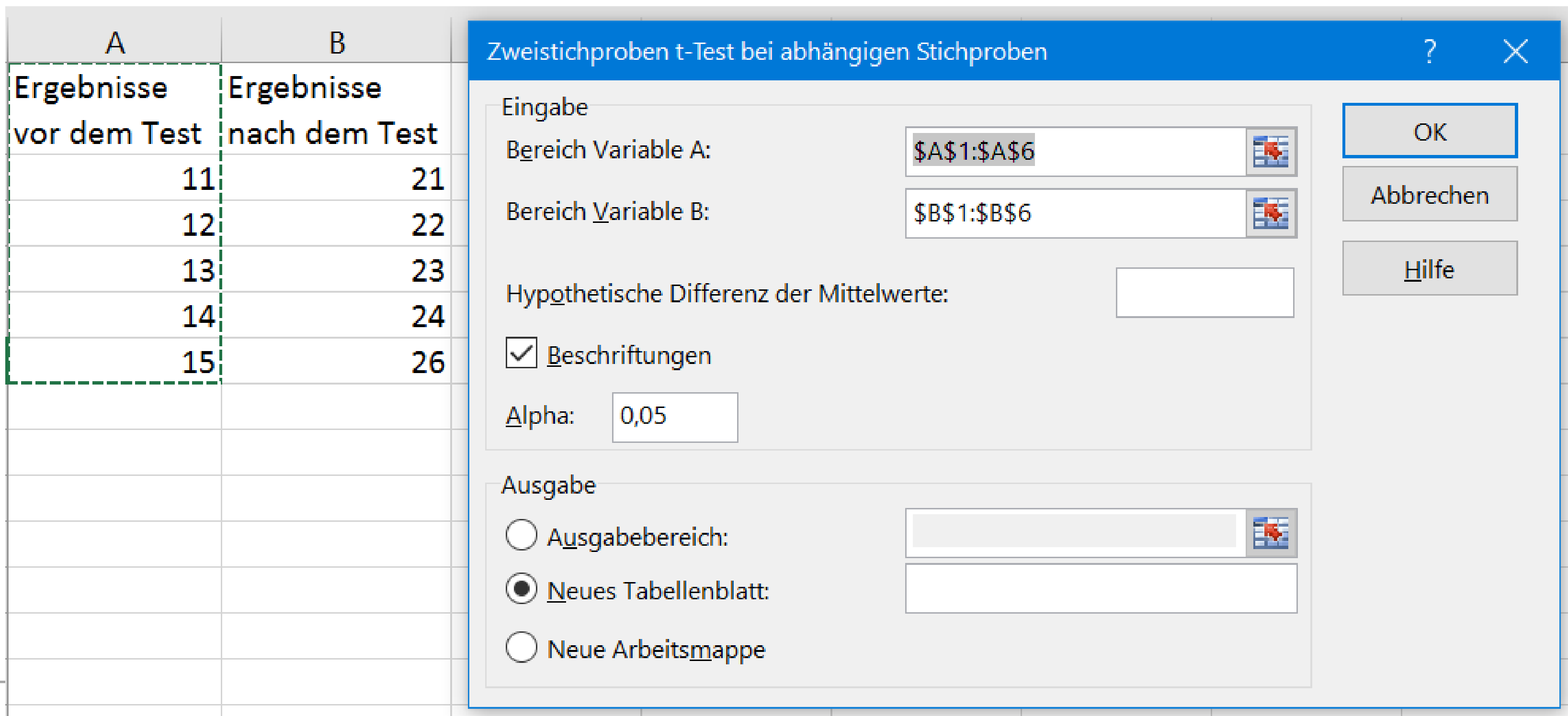
Name	**Leistungstest vorher**	**Leistungstest nachher**	**Differenz**
Ilsa	2	1	1
Peter	4	6	-2
Lisa	9	7	2
Hans	2	5	-3
Mittelwert $\bar{d}$			-0,5
Standardabweichung			2,38
Z-Wert			-0,420

Der ermittelte Z-Wert wird nun wieder in einer Tabelle – hier Z-Verteilung, für die Standardnormalverteilung nach geschaut. Er entspricht ca. 34,46 % Perzentil und liegt damit über den Signifikanzniveau von 0,05 d.h. die Differenzen sind nicht signifikant. Es gibt keinen signifikanten Unterschied zwischen dem Leistungstest „vorher“ und „nachher“. Das „Ereignis“ hatte keine Auswirkung auf die Leistung.

t-Test bei gepaarten/abhängigen Stichproben: in Excel

Das Excel-Add-In Datenanalyse muss aktiviert sein

→ Daten → Datenanalyse → Zweistichproben-t-Test bei abhängigen Stichproben



t-Test bei gepaarten/abhängigen Stichproben: Interpretation

Zweistichproben t-Test bei abhängigen Stichproben (Paarvergleichstest)		
	Ergebnisse vor dem Test	*Ergebnisse nach dem Test*
Mittelwert	13	23,2
Varianz	2,5	3,7
Beobachtungen	5	5
Pearson Korrelation	0,986	
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	4	
t-Statistik	-51	
P(T<=t) einseitig	0,000	
Kritischer t-Wert bei einseitigem t-Test	2,132	
P(T<=t) zweiseitig	0,000	
Kritischer t-Wert bei zweiseitigem t-Test	2,776	

Die Mittelwerte beider Variablen (Stichproben)

Der p-Wert des t-Tests

Die H_0 wird abgelehnt, da der zum t-Test zugehörige p-Wert unter dem Signifikanzniveau von 0,05 liegt.