Hypothesenprüfung

Grundfrage

Wie können die postulierten Eigenschaften der Grundgesamtheit (Theorie) durch stichprobenartig erhobene Daten (Empirie) überprüft werden?

Wissenschaftliche vs. statistische Hypothese

H_1 **Alternativhypothese** → **wissenschaftliche Hypothese**

Behauptung eines Zusammenhangs/Unterschieds

H_0 **Nullhypothese** → **statistische Hypothese**

Falsifikation

Arten von Alternativhypothesen

Unterschiedshypothese

Männer und Frauen **unterscheiden sich** hinsichtlich ihrer Schokoladenutzung während des Fernsehens. — ***ungerichtet, unspezifisch***

Männer essen beim Fernsehen **mehr** Schokolade als Frauen. — ***gerichtet, unspezifisch***

Männer essen **doppelt so viel** Schokolade beim Fernsehen als Frauen. — ***gerichtet, spezifisch***

Zusammenhangshypothese

Das Interesse an empirischen Forschungsmethoden **differiert je nach** Interesse an Mathematik. — ***ungerichtet, unspezifisch***

Je größer das Interesse an Mathematik, **desto geringer** das Interesse an empirischen Forschungsmethoden. — ***gerichtet, unspezifisch***

Im gleichen Maße, wie das Interesse an Mathematik steigt, sinkt das Interesse an empirischen Forschungsmethoden. — ***gerichtet, spezifisch***

Definition

Die **Nullhypothese** (H_0) ist eine formale **Gegenhypothese**, die behauptet, dass der in der Alternativhypothese postulierte Unterschied bzw. Zusammenhang **in der Grundgesamtheit nicht vorhanden** ist.

Alternativhypothese (H_1) und Nullhypothese (H_0) sind zueinander komplementäre Aussagen.

Hypothesentest

Beim statistischen Hypothesentest prüft man einen empirischen Sachverhalt gegen die Zufälligkeit einer Verteilung:

Gilt die in der Alternativhypothese formulierte Aussage oder nicht?

Um diese Frage zu beantworten, wird mit Hilfe von statistischen Tests versucht, **die Nullhypothese zu widerlegen**.

Prinzip der Falsifikation

“Wann immer wir nämlich glauben, die Lösung eines Problems gefunden zu haben, sollten wir unsere Lösung nicht verteidigen, sondern mit allen Mitteln versuchen, sie selbst umzustoßen.”

Karl Popper (Logik der Forschung)

α-Fehler und β-Fehler

		In der Population gilt die…	
Entscheidung aufgrund der Stichprobe zugunsten der…		H0	H1
	H0	richtige Entscheidung	β-Fehler
	H1	α-Fehler	richtige Entscheidung

α-Fehler (oder: Fehler 1. Art)

- Die H_0 wird verworfen, obwohl sie richtig ist.
- Man nimmt ein Phänomen an, dass in der GG nicht da ist.

β-Fehler (oder: Fehler 2. Art)

- Die H_1 wird verworfen, obwohl sie richtig ist.
- Man ignoriert ein Phänomen, dass in der GG eigentlich da ist.

Irrtumswahrscheinlichkeit

Definition

Die Wahrscheinlichkeit, mit der das gefundene Ergebnis oder extremere Ergebnisse **bei Gültigkeit von H_0** eintreten, bezeichnet man als **α-Fehlerwahrscheinlichkeit** oder **Irrtumswahrscheinlichkeit**.

Oder anders gesagt:

Wie hoch ist die Wahrscheinlichkeit, dass ich mich irre, wenn ich aufgrund meiner Stichprobendaten einen bestimmten Sachverhalt annehme?

Beträgt die Wahrscheinlichkeit hierfür weniger als 5%, so spricht man von einem **signifikanten**, bei weniger als 1% von einem **hochsignifikanten** Ergebnis.

Es gilt:

1 – Vertrauenswahrscheinlichkeit = Irrtumswahrscheinlichkeit

Signifikanzniveau

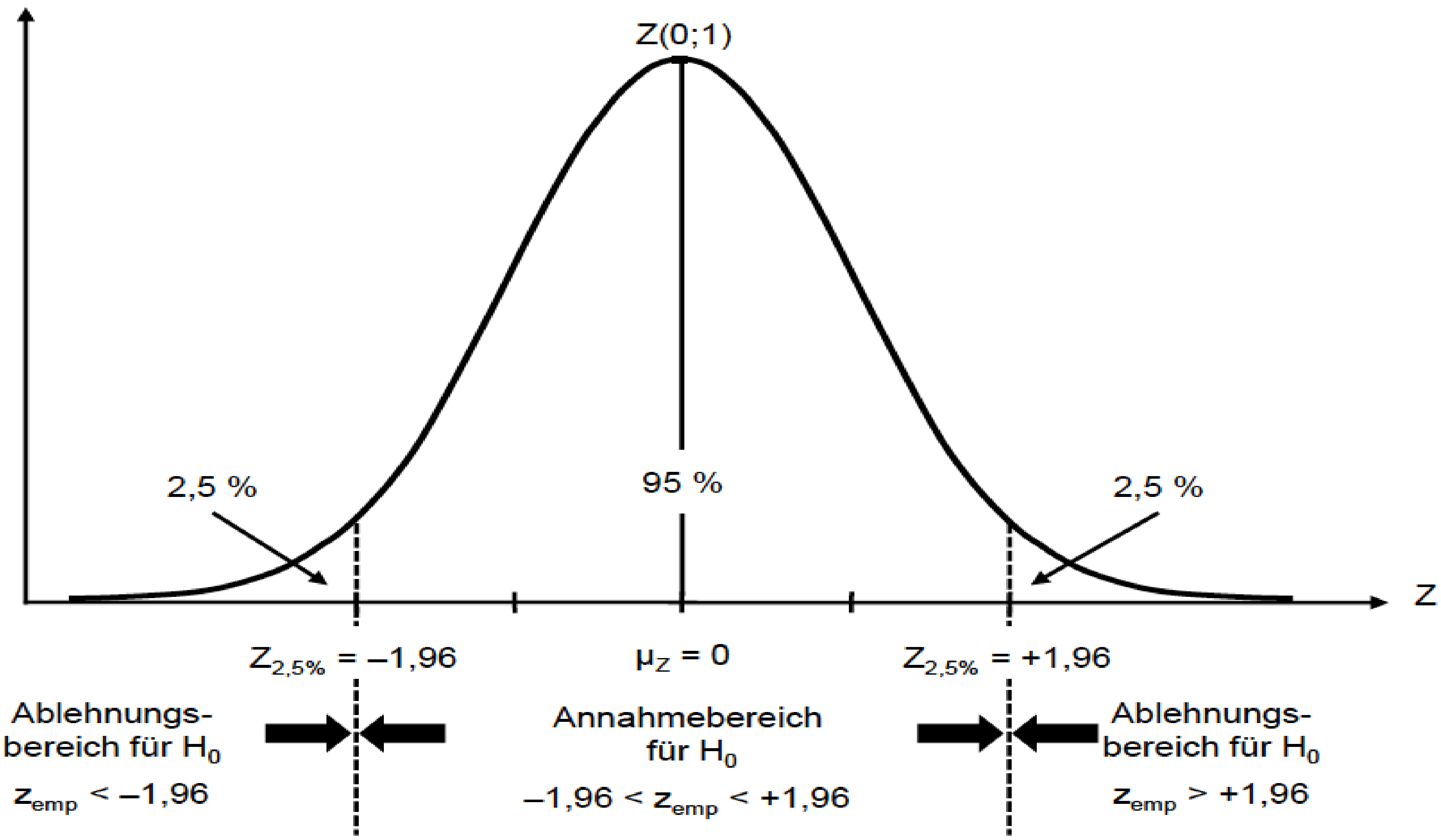
Definition

Signifikanz liegt dann vor, wenn anhand eines **Signifikanztest** davon ausgegangen werden kann, dass der in der Stichprobe gefundene **Zusammenhang** zwischen Variablen (oder die **Unterschiede** zwischen Gruppen) auch in der Grundgesamtheit **nicht zufällig** zu finden sind.

Das **Signifikanzniveau** gibt die maximal erlaubte **Irrtumswahrscheinlichkeit** an, unterhalb der man bereit ist, die **Alternativhypothese** anzunehmen.

Irrtumswahrscheinlichkeit p(α)	Bezeichnung des Stichprobenergebnisses als…	Kennzeichnung des Stichprobenergebnisses
$p > 0{,}05$	nicht signifikant	- *oder* n.s.
$p \leq 0{,}05$	signifikant	*
$p \leq 0{,}01$	sehr signifikant	**
$p \leq 0{,}001$	höchst signifikant	***

Beispiel ungerichtete Hypothese



Überblick Hypothesentests

Statistische Tests

Chi²-Test (χ^2 - Test)

T-Test

F-Test

z-Test

U-Test

Korrelationstest

...

Auswahlkriterien

Skalenniveau der Variablen

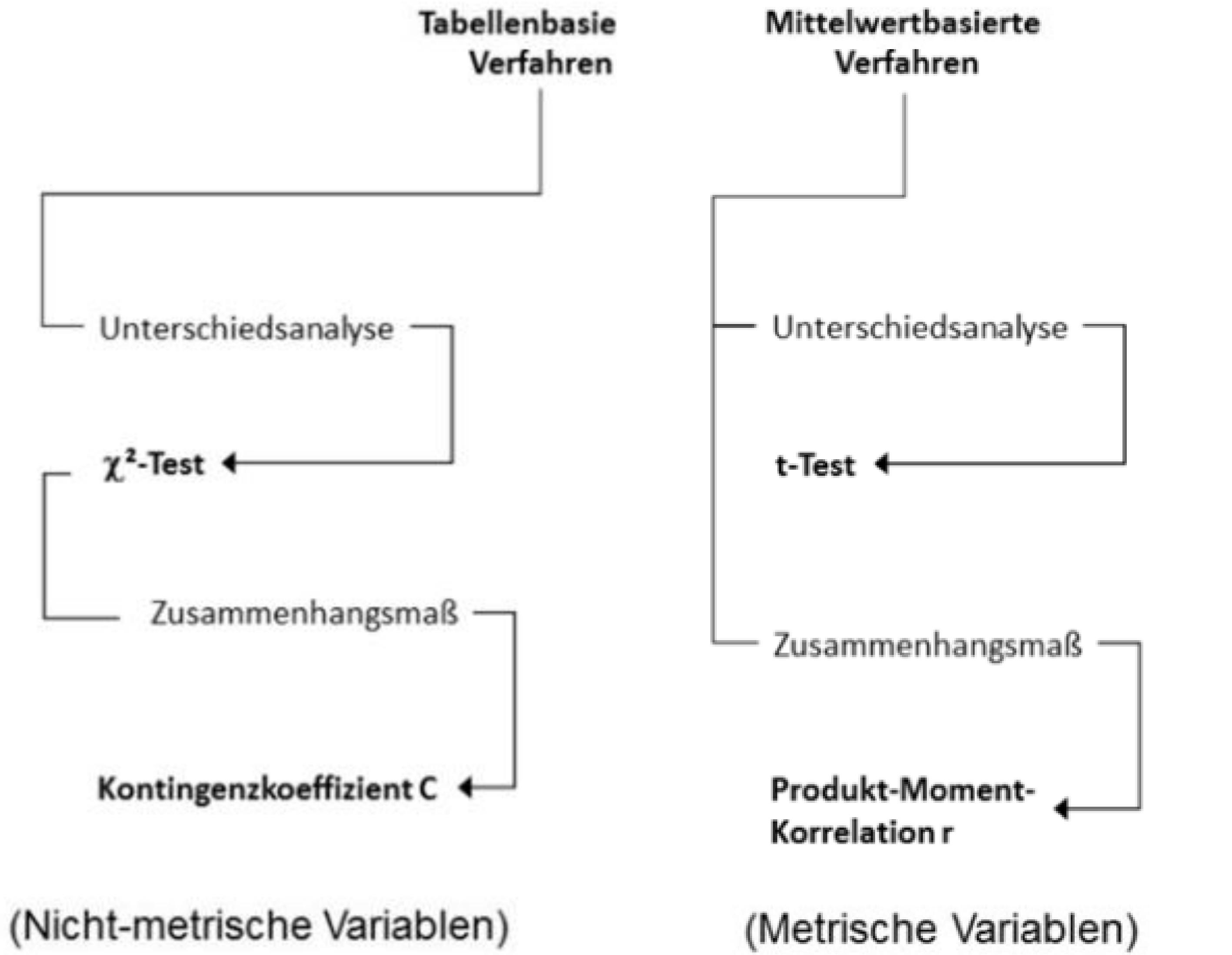
Verteilungseigenschaften der Variablen

Anzahl der Ausprägungen

Art der Hypothese

Stichprobenumfang

Übersicht



Tabellenbasie Verfahren
Mittelwertbasierte Verfahren
Unterschiedsanalyse
χ^2-Test
Zusammenhangsmaß
Kontingenzkoeffizient C
Unterschiedsanalyse
t-Test
Zusammenhangsmaß
Produkt-Moment-Korrelation r
(Nicht-metrische Variablen)
(Metrische Variablen)

Ein alter Bekannter: CHI²-Test

Logik

Beobachtete Häufigkeit in der Kreuztabelle werden mit erwarteten Häufigkeiten verglichen.

Erwartete Häufigkeiten sind die Zahlenwerte, die aufgrund der Randverteilung zu erwarten wären, wenn die Nullhypothese („Gleichverteilungshypothese“) zutrifft.

Wir erinnern uns:

CHI² ist ein nicht-standardisiertes Zusammenhangsmaß für nominalskalierte Variablen.

Die Tabelle mit den erwarteten Häufigkeiten nennt man *Indifferenztabelle* = „keine Differenz zwischen den Variablen“.

Die Prüfgröße Chi2 (χ^2) ist bei ausreichend großen (Zellen-) Häufigkeiten annähernd Chi²-verteilt bei m – 1 Freiheitsgraden (m = Merkmalsklassen).

Wenn die Nullhypothese zutrifft, sollte der Unterschied zwischen der beobachteten und der erwarteten Häufigkeit nur klein sein.

Vorgehensweise

(1) Hypothesen formulieren / Signifikanzniveau bestimmen

(2) Kontingenztabelle errechnen

Kreuztabelle mit beobachteten Häufigkeiten (f_b) in absoluten Zahlen

(3) Indifferenztabelle berechnen

Erwartete Häufigkeiten f_e) für jede Zelle berechnen

$$f_e = \frac{\text{Zeilen (n)} * \text{Spalten(n)}}{\text{Gesamt (n)}}$$

(4) Für jede Zelle Differenz f_b / f_e berechnen

$$\frac{(f_b - f_e)^2}{f_e}$$

(5) Addition der Werte aus (4) zum Chi²-Wert

(6) Berechnung der Freiheitsgrade

df = (Zeilenanzahl - 1) * (Spaltenanzahl - 1) **(df = degrees of freedom)**

(7) Vergleich des empirischen Chi²-Wertes aus (5) mit dem theoretischen Wert (-> Tabelle)

$$\chi^2_{theor} \geq \chi^2_{emp} = H_0 \text{ wird beibehalten}$$

$$\chi^2_{theor} < \chi^2_{emp} = H_0 \text{ wird verworfen}$$

!

Tabelle der Chi^2-Verteilung

(Auszug)

Vertrauenswahrscheinlichkeit

Freiheitsgrade

	1−α					
f	**0,900**	**0,950**	**0,975**	**0,990**	**0,995**	**0,999**
1	2,71	3,84	5,02	6,63	7,88	10,83
2	4,61	5,99	7,38	9,21	10,60	13,82
3	6,25	7,81	9,35	11,34	12,84	16,27
4	7,78	9,49	11,14	13,28	14,86	18,47
5	9,24	11,07	12,83	15,09	16,75	20,52
6	10,64	12,59	14,45	16,81	18,55	22,46
7	12,02	14,07	16,01	18,48	20,28	24,32
8	13,36	15,51	17,53	20,09	21,95	26,12
9	14,68	16,92	19,02	21,67	23,59	27,88
10	15,99	18,31	20,48	23,21	25,19	29,59

Was sind Freiheitsgrade?

Zufriedenheit mit der Demokratie in Deutschland	Nutzerkreis der Tageszeitung			Gesamt
	Nicht-Nutzer	Weitester Leserkreis	Stammnutzer	
eher zufrieden	23	50	?	321
mittel	66	130	?	979
eher unzufrieden	?	?	?	702
Gesamt	**128**	**266**	**1608**	**2002**

Wie viele Informationen brauche ich bei gegebenen Randverteilungen, um die letzten fehlenden Informationen zu erschließen?

df = (Zeilenanzahl – 1) ∗ (Spaltenanzahl – 1)

1. H_1 Es besteht <u>ein Zusammenhang</u> zwischen Geschlecht und Überlebenschance.
   H_0 Die Variablen sind unabhängig voneinander.

2. **Signifikanzniveau: 5%-ige Irrtumswahrscheinlichkeit**

3. **Kontingenztabelle**: **Überlebende der Titanic-Katastrophe (beobachtete Häufigkeiten)**

	nicht überlebt	überlebt	Gesamt
Frau	126	344	470
Mann	1364	367	1731
Gesamt	1490	711	2201

4. **Indifferenztabelle: Überlebende der Titanic-Katastrophe (erwartete Häufigkeiten)**

	nicht überlebt	überlebt	Gesamt
Frau	318,17	151,83	470
Mann	1171,83	559,17	1731
Gesamt	1490	711	2201

$$f_e = \frac{Zeilen\ (n) \cdot Spalten\ (n)}{Gesamt\ (n)}$$

Berechnung des Chi2-Wertes

Berechnung der Freiheitsgrade

Vergleich dieses Chi2-Wertes mit dem theoretisch erwartbaren → Tabelle

Ergebnis / Interpretation

→ vollständige Ausformulierung mit bezug auf die Hypothese

5. Chi²-Wert pro Zelle:

	nicht überlebt	überlebt
Frau	116,07	243,24
Mann	31,52	66,05

$$\frac{(fb - fe)^2}{f_e}$$

$$\sum = \chi^2 = 116{,}07 + 31{,}52 + 243{,}24 + 66{,}05 = \mathbf{456{,}87}$$

Chi² = 0: Die Variablen sind unabhängig voneinander.

Chi² ≠ 0: Es besteht ein Zusammenhang zwischen den Variablen.

6. Berechnung der Freiheitsgrade

$$df = (Zeilenanzahl - 1) * (Spaltenanzahl - 1)$$

$$df = (2 - 1) * (2 - 1) = 1 * 1 = 1$$

7. Vergleich des empirischen Chi²-Wertes (siehe 5.) mit dem theoretischen Chi²-Wert aus der Tabelle:

$x^2_{theor} \geq x^2_{emp}$ = H_0 wird beibehalten

$x^2_{theor} < x^2_{emp}$ = H_0 wird verworfen

!

x^2_{emp} = 456,87

df = 1

1 - ✂ = 0,95

↓

x^2_{theor} = 3,84

	1−α					
f	0,900	0,950	0,975	0,990	0,995	0,999
1	2,71	3,84	5,02	6,63	7,88	10,83
2	4,61	5,99	7,38	9,21	10,60	13,82
3	6,25	7,81	9,35	11,34	12,84	16,27
4	7,78	9,49	11,14	13,28	14,86	18,47
5	9,24	11,07	12,83	15,09	16,75	20,52
6	10,64	12,59	14,45	16,81	18,55	22,46
7	12,02	14,07	16,01	18,48	20,28	24,32

7. **Vergleich der unter 5. berechneten Summe mit dem theoretischen Wert aus der Tabelle. Wenn theoretischer Chi²-Wert größer ⇨ H_0 wird beibehalten!**

$$\chi^2 theo = 3{,}84 < \chi^2 emp = 456{,}87$$

⇨ H_0 wird verworfen und H_1 wird angenommen.

⇨ Mit **95%-iger Sicherheit** gibt es einen Zusammenhang zwischen dem Geschlecht und der Überlebenschance in der **Grundgesamtheit**.

Vorgehen zusammengefasst

1. **Hypothesen formulieren**
   a. Wissenschaftliche Hypothese (H1)
   b. Statistische Hypothese ableiten (H0)

2. **α-Fehler/Irrtumswahrscheinlichkeit bestimmen**

3. **Stichprobenparameter und Test-Verteilung bestimmen** (Normal, t, Chi^2, F)

4. **Prüfgröße berechnen**

5. **Vergleich des empirischen mit dem theoretischen Wert**

6. **Nullhypothese verwerfen oder beibehalten**

7. **Wiss. Hypothese beibehalten oder verwerfen**

T-Verteilung

Definition

Die T-Verteilung (eigentlich: **Student-T-Verteilung**) ist eine Wahrscheinlichkeitsverteilung.

Wenn die Standardabweichung **σ** der Grundgesamtheit unbekannt ist (Regelfall), erlaubt sie die Berechnung der Verteilung der Differenz vom Mittelwert der Stichprobe zum wahren Wert der Grundgesamtheit.

Der einzige Parameter, den die der T-Verteilung benötigt, ist ν = **Freiheitsgrade**. Sie beziehen sich auf die Größe der Stichprobe.

Die Verteilung wird mit wachsendem ν schmaler und geht für ν **->** ∞ in eine Standardnormalverteilung über.

Einsatz

Schätzung von unbekannten Parametern (z.B. **μ),** die fehlerbehaftet sind. Zur Erinnerung: tatsächlicher Wert = Messwert + Fehlerwert.

Da die Standardabweichung dieser Fehler in der Grundgesamtheit in der Regel unbekannt ist, muss sie geschätzt werden. -> T-Verteilung

Ist sie bekannt, wird eher die Normalverteilung verwendet.

Fun Fact

Der Erfinder war Mitarbeiter der Guinness-Brauerei, Dublin.

T-Test 1

Voraussetzung

Metrisch skalierte Daten

Normalverteilte Daten (bei $n \geq 200$)

Fragestellung

Unterscheiden sich die Mittelwerte x_1 und x_2…

bei **abhängigen** Stichproben

einer zweifach gemessenen Variable?

zweier Variablen?

bei **unabhängigen** Stichproben

zweier Gruppen?

Hypothesen

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

Testgröße

t-Wert

$df = n_1 + n_2 - 2$

$t_{theor} \geq t_{emp}$ = H_0 **wird beibehalten**

$t_{theor} < t_{emp}$ = H_0 **wird verworfen**

Abhängige und unabhängige Stichproben

Abhängige Stichproben

 $\dot{x}_{t1} =?$  $\dot{x}_{t2} =?$

z.B.: Verändert sich der Mittelwert einer Variable zu 2 Messzeitpunkten?

Unabhängige Stichproben

CDU SPD

z.B.: Unterscheidet sich der Mittelwert zwischen zwei Gruppen?

Beispiel

Vorhaben Test von 4 unterschiedlichen Versionen einer Webseite
Vorgabe eines Tasks

Messgröße SUCCESS erfolgreich erledigt: ja/nein

Frage Welche Version funktioniert am besten?

→ sichere, d.h. statistisch signifikante Ergebnisse

BETWEEN-GROUPS

SUCCESS	Version 1	Version 2	Version 3	Version 4
ja	42	47	51	39
nein	58	53	49	61
N	100	100	100	100

WITHIN-GROUP

SUCCESS	Version 1	Version 2	Version 3	Version 4
ja	32	36	41	34
nein	18	14	9	16
N	50			

Beispiel

BETWEEN-GROUPS

unabhängige Usergruppen

statistisches Rauschen maskiert u.U. Effekte

→ vergleichbare Zusammensetzung der Gruppen essentiell

↓

größeres Sample

keine Ermüdung, kein Lerneffekt

WITHIN-GROUP

personenidentische Usergruppe

weniger statistisches Rauschen

↓

höhere Wahrscheinlichkeit, „wahre“ Effekte zu finden

↓

kleineres Sample

Motivationsproblem, Langeweile, Lerneffekt

Art Test auf Signifikanz von Mittelwertunterschieden

Einsatz z.B. Gruppenvergleich von Tests; Mittelwerte TIME ON TASK in Sekunden

		Version 1	Version 2
ToT	Ø	154,5	135,6
	N	25	35

Mittelwerte: signifikant verschieden oder nicht?

Mittelwertvergleich → z.b. T-Test BETWEEN-GROUPS

PAIRED T-Test WITHIN-GROUP

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\overbrace{\bar{X}_1}^{\text{Mittelwert der ersten Stichprobe}} - \overbrace{\bar{X}_2}^{\text{Mittelwert der zweiten Stichprobe}}}{\underbrace{\sqrt{\frac{\overbrace{\left(\sum X_1^2 - \frac{(\sum X_1)^2}{n_1}\right)}^{\text{geschätzte Varianz von } X_1} + \overbrace{\left(\sum X_2^2 - \frac{(\sum X_2)^2}{n_2}\right)}^{\text{geschätzte Varianz von } X_2}}{\underbrace{n_1 + n_2 - 2}_{\text{Freiheitsgrade}}} \cdot \overbrace{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}^{\text{Korrektur um in den Standardfehler der Stichprobe umzurechnen}}}}_{\text{Standardfehler der Differenz des Mittelwerts zweier Stichproben}}}, \quad s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Definition

Für unabhängige Stichproben:

$$t_{emp} = \frac{(\bar{x}_1 - \bar{x}_2)}{\boxed{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}}}$$

Standardfehler der
Mittelwertdifferenzen

$$\boxed{\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)}} = \sqrt{\frac{\hat{\sigma}_1^2 * (n_1 - 1) + \hat{\sigma}_2{}^2 * (n_2 - 1)}{(n_1 - 1) + (n_2 - 1)} * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Für abhängige Stichproben:

$$t_{emp} = \frac{\bar{x}_d}{\hat{\sigma}_{\bar{x}_d}}$$

(***d = Differenz***)

T-Tabelle

Anzahl Freiheitsgrade *n*	**P für zweiseitigen Vertrauensbereich**							
	0,5	**0,75**	**0,8**	**0,9**	**0,95**	**0,98**	**0,99**	**0,998**
	P für einseitigen Vertrauensbereich							
	0,75	**0,875**	**0,90**	**0,95**	**0,975**	**0,99**	**0,995**	**0,999**
1	1,000	2,414	3,078	6,314	12,706	31,821	63,657	318,309
2	0,816	1,604	1,886	2,920	4,303	6,965	9,925	22,327
3	0,765	1,423	1,638	2,353	3,182	4,541	5,841	10,215
4	0,741	1,344	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,301	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,273	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,254	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,240	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,230	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,221	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,214	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,209	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,204	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,200	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,197	1,341	1,753	2,131	2,602	2,947	3,733

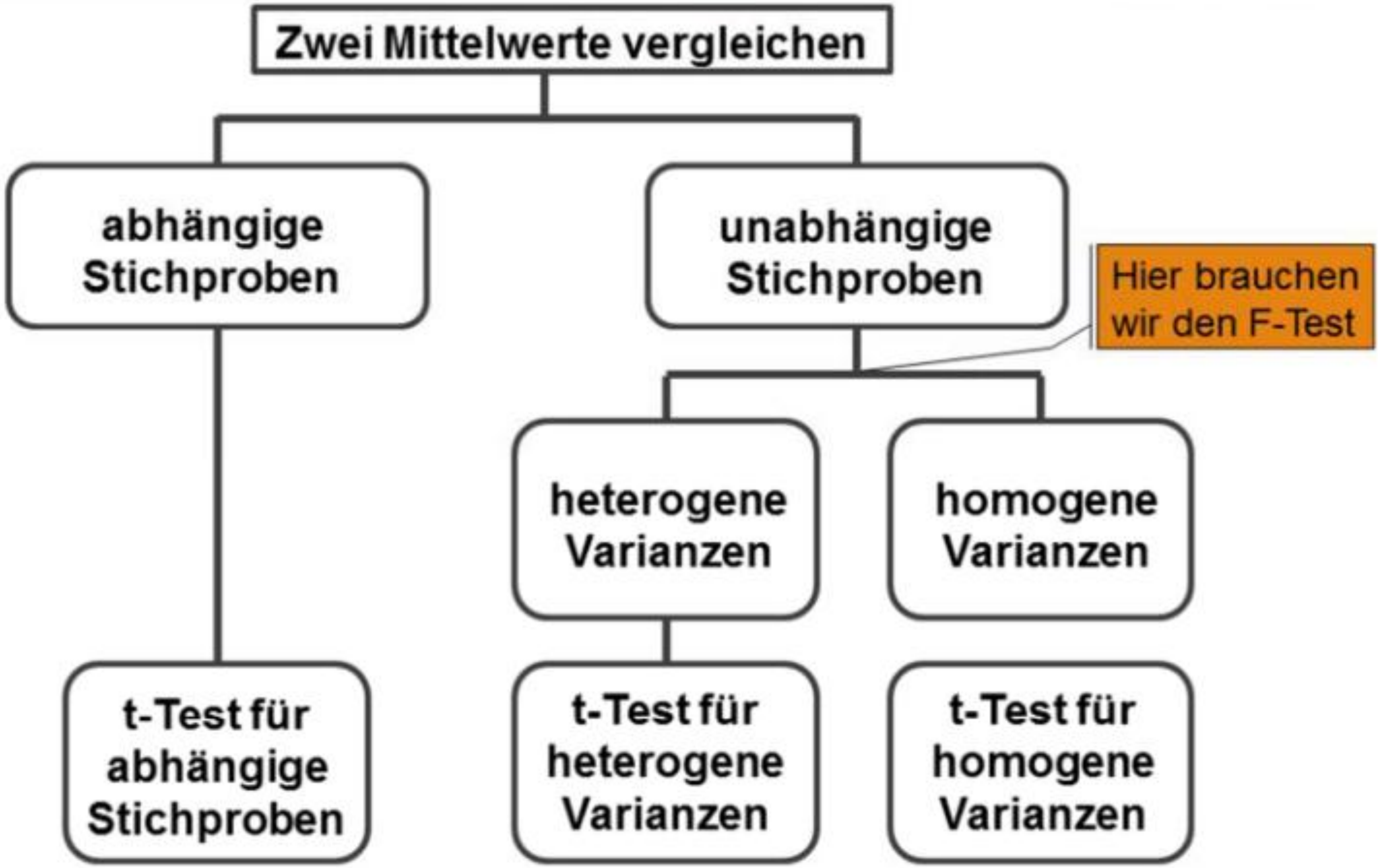
(Auszug)

Zur Erinnerung:

zweiseitig = ungerichtete Hypothese

einseitig = gerichtete Hypothese

T-Test: Entscheidungsbaum



Und noch eine Verteilung...

Definition

Die F-Verteilung (eigentlich Fischer-Verteilung) ist eine stetige Wahrscheinlichkeitsverteilung.

Eine F-verteilte Zufallsvariable ergibt sich als Quotient zweier jeweils durch die zugehörige Anzahl der Freiheitsgrade geteilter Chi^2-verteilter Zufallsvariablen. Die F-Verteilung besitzt zwei unabhängige Freiheitsgrade als Parameter.

Einsatz

Die F-Verteilung wird häufig in einem Test verwendet, um festzustellen, ob der Unterschied zweier Stichprobenvarianzen auf statistischer Schwankung beruht oder ob er auf unterschiedliche Grundgesamtheiten hinweist.

Auch im Rahmen der Varianzanalyse wird mit einer F-Statistik auf signifikante Unterschiede zwischen Grundgesamtheiten (Gruppen) getestet.

F-TEST

Voraussetzung

Zwei von einander unabhängige Stichproben

Normalverteilung der Kennwerte

Fragestellung

Sind die Varianzen zweier Stichproben heterogen oder homogen?

Oder anders: Entstammen die beiden Stichproben der gleichen Grundgesamtheit?

Hypothesen

$H_0: \sigma^2_1 = \sigma^2_2$

$H_1: \sigma^2_1 \neq \sigma^2_2$

Testgröße

$$F = \frac{s^2_1}{s^2_2}$$

In den Zähler wird der größere der beiden Werte gesetzt. Dann gilt F = 1, wenn beide Varianzen gleich groß sind, und: je größer F desto ungleicher (=heterogener) die Varianzen.

$df_1 = (n_1 - 1)$ Freiheitsgrade im Zähler

$df_2 = (n_2 - 1)$ Freiheitsgrade im Nenner

$f_{theor} \geq f_{emp}$ = H_0 **wird beibehalten**

$f_{theor} < f_{emp}$ = H_0 **wird verworfen**

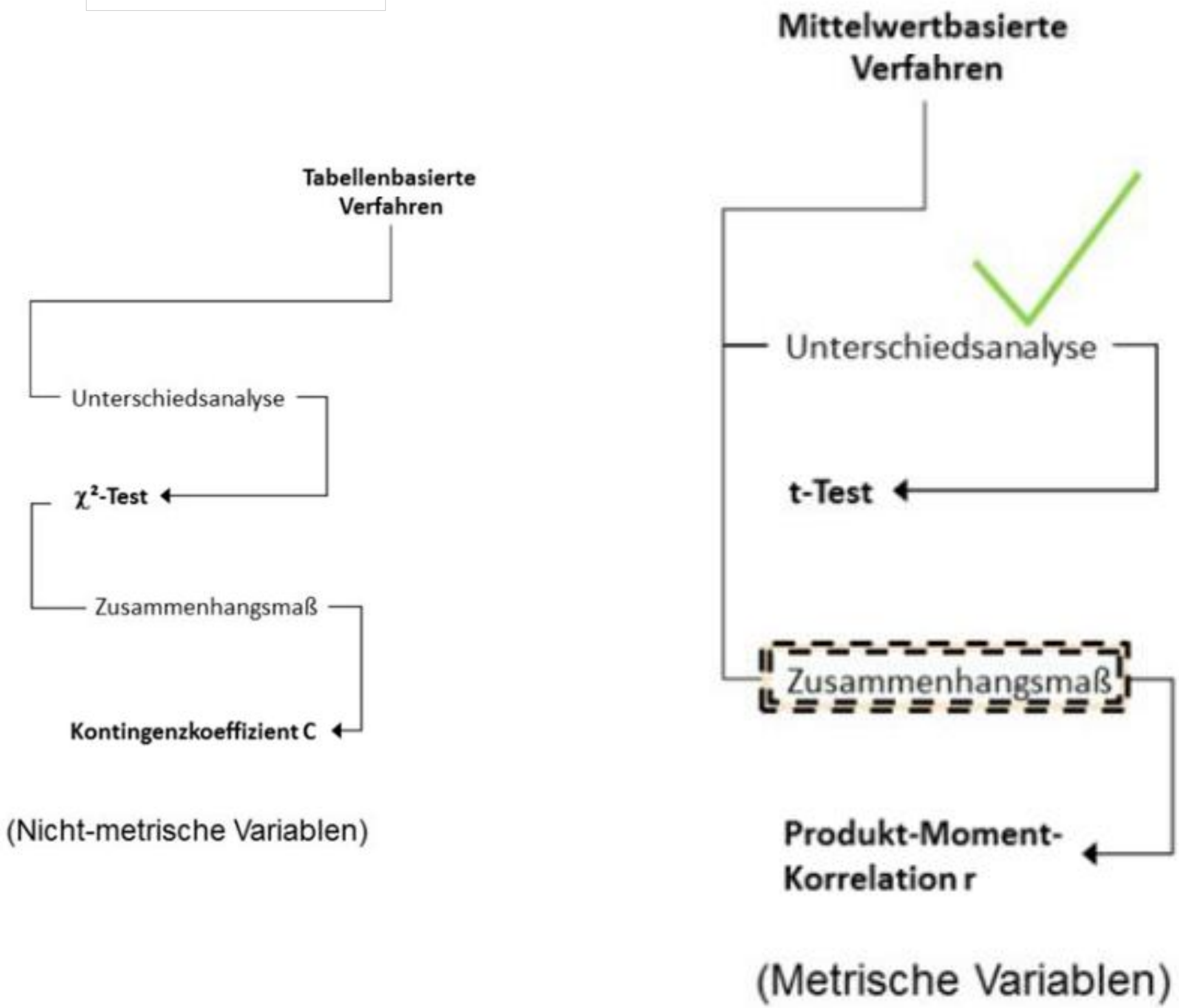
F-Tabelle

F-Verteilung für p = .95 bzw. α = .05

Freiheitsgrade Zähler waagerecht; Nenner senkrecht

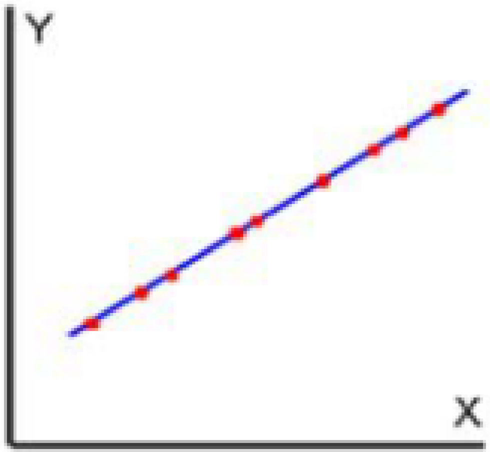
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Übersicht



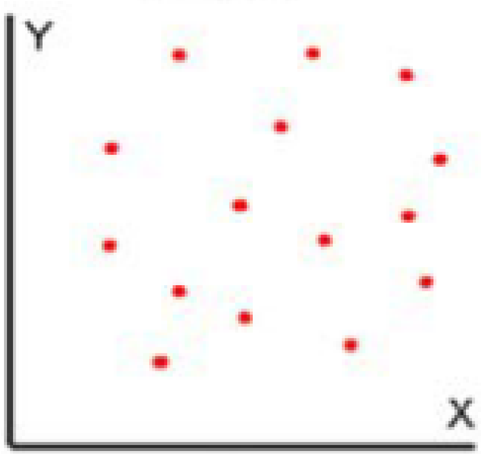
Tabellenbasierte Verfahren
Unterschiedsanalyse
χ^2-Test
Zusammenhangsmaß
Kontingenzkoeffizient C
(Nicht-metrische Variablen)
Mittelwertbasierte Verfahren
Unterschiedsanalyse
t-Test
Zusammenhangsmaß
Produkt-Moment-Korrelation r
(Metrische Variablen)

Noch ein alter Bekannter: Pearson‘s r

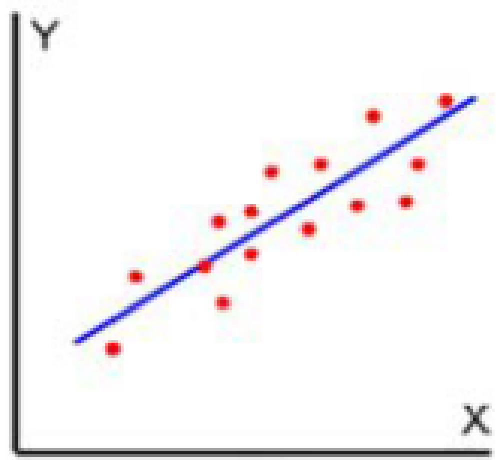


r = + 1.0

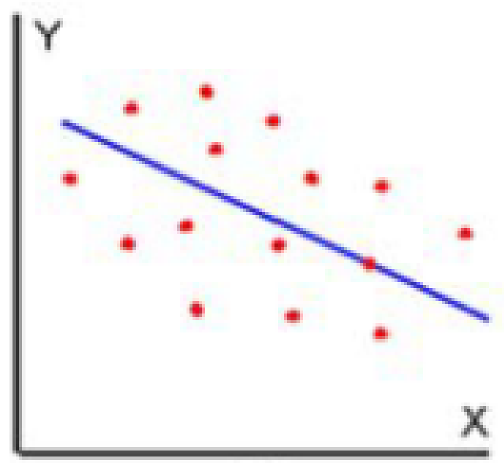
$$r_{xy} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y})}{s_x \cdot s_y}$$



r = 0.0



r = + 0.9



r = - 0.5

Pearson‘s r

Fragestellung

Hängen zwei Variablen X und Y signifikant (linear) zusammen? → Korrelation

Hypothesen

H_0: r = 0

H_1: r ≠ 0

Testgröße

Für den Signifikanztest wird der Stichprobenkennwert r in eine t-verteilte Prüfgröße umgewandelt.

$$t_{emp} = \frac{r * \sqrt{n-2}}{\sqrt{(n-r)^2}}$$

$t_{theor} \geq t_{emp}$ = H_0 wird beibehalten

$t_{theor} < t_{emp}$ = H_0 wird verworfen

Und wieder T-Tabelle...

Anzahl Freiheitsgrade *n*	***P* für zweiseitigen Vertrauensbereich**							
	0,5	**0,75**	**0,8**	**0,9**	**0,95**	**0,98**	**0,99**	**0,998**
	***P* für einseitigen Vertrauensbereich**							
	0,75	**0,875**	**0,90**	**0,95**	**0,975**	**0,99**	**0,995**	**0,999**
1	1,000	2,414	3,078	6,314	12,706	31,821	63,657	318,309
2	0,816	1,604	1,886	2,920	4,303	6,965	9,925	22,327
3	0,765	1,423	1,638	2,353	3,182	4,541	5,841	10,215
4	0,741	1,344	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,301	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,273	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,254	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,240	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,230	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,221	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,214	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,209	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,204	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,200	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,197	1,341	1,753	2,131	2,602	2,947	3,733

(Auszug)

Zur Erinnerung:

zweiseitig = ungerichtete Hypothese

einseitig = gerichtete Hypothese

Strichprobengröße

Wie viele Fälle brauche ich, um statistisch sichere Ergebnisse zu erhalten?

$$\frac{[z^2 * p(1-p)] / e^2}{1+[z^2 * p(1-p)]/e^2 * N}$$

N — Größe der Grundgesamtheit

e — Fehlerbereich (Genauigkeit) akzeptierte Fehlertoleranz

5% (e = 0,05), 10% (e = 0,10)
je größer die akzeptierte Fehlertoleranz, desto kleiner die Stichprobe

z — Konfidenzniveau (Vertrauenswahrscheinlichkeit)

üblich 95% Wahrscheinlichkeit, dass der Stichprobenwert dem in der Grundgesamtheit entspricht: z-Wert von 1,96

p — Standardabweichung (Mittelwertstreuung)

üblich bei ungekanntem Mittelwert: 50% (p = 0,5)
Worst Case ⟶ führt zu größerer Stichprobe