Statistik II, SoSe 23

26.6.23, Aktuelles, Statistisches Testen, t-Test, Teil 2

Simone Abendschön

Inhalte heute

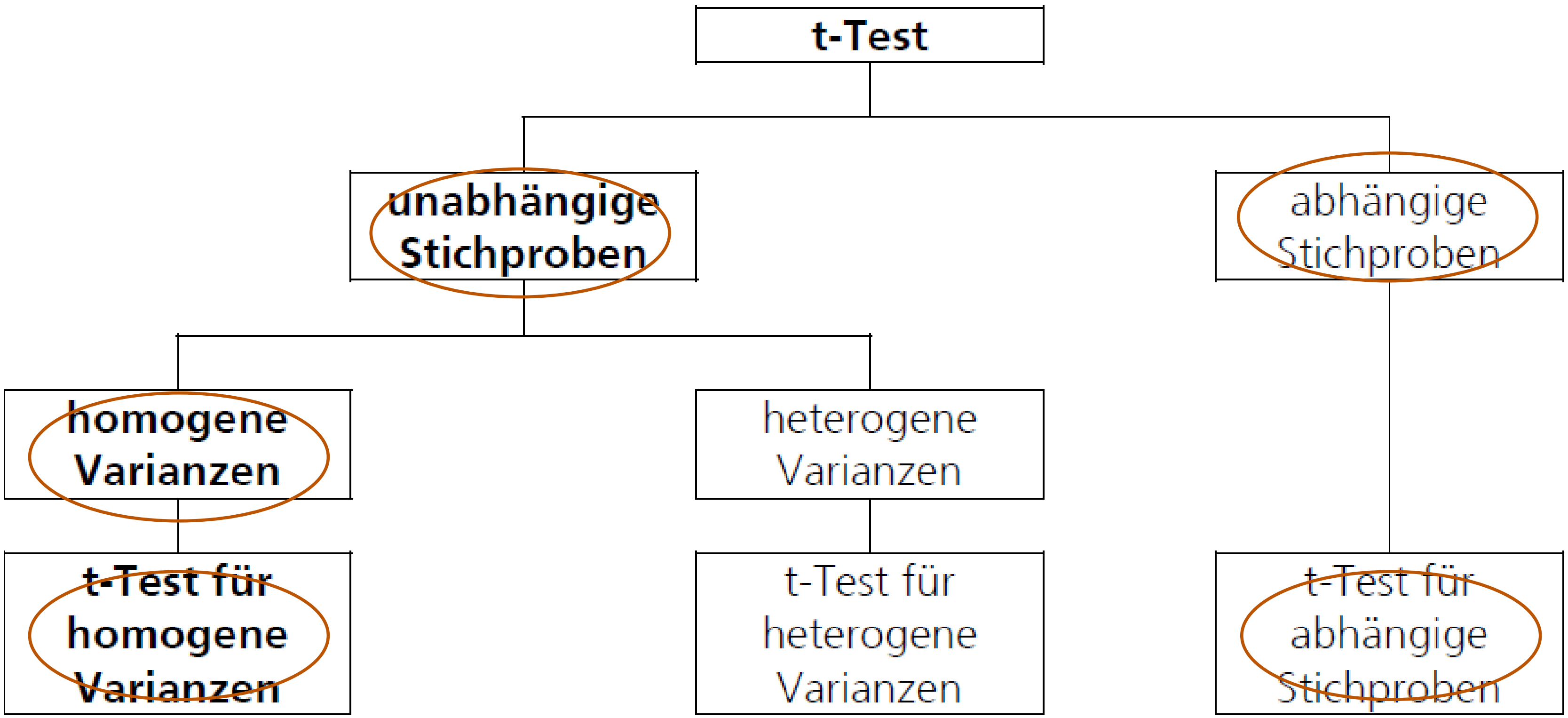
JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Abschluss Hypothesentests:
  - T-Test für abhängige Stichproben
  - Signifikanzniveaus
  - Fehlerarten
- Lineare Regression, Teil 1
  - Wiederholung bivariates Regressionsmodell
- Schriftliche Evaluation der Veranstaltung

Lernziele

- Sie können t-Tests für abhängige Stichproben durchführen
- Sie verstehen, wozu man eine lineare Regressionsanalyse macht
- Sie verstehen, wie ein (bivariates) Regressionsmodell geschätzt wird

Varianten des t-Tests



Quelle: Eigene Darstellung

Übungsfragen (HA/Tutorium)

3. Durchführung t-Test (Annahme homogene Varianzen) für ein fiktives Beispiel: mittlere **Lebenszufriedenheit von Männern und Frauen in einer Befragung von 42 Leuten (gemessen von 0 gar nicht zufrieden bis 10 sehr zufrieden)**

Ist der Unterschied statistisch signifikant?

	Frauen	Männer
n	20	22
$\bar{x}$	7,5	6,5
s^2	1,2	1,3

Übungsfrage 3 - Lösung

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

1.Hypothese und Signifikanzniveau

H_0: Es gibt keine Mittelwertdifferenz = Die Lebenszufriedenheit von Männern und Frauen ist gleich hoch

H_1: Es gibt eine Mittelwertdifferenz zwischen 2 Gruppen in der Population = Die Lebenszufriedenheit von Männern und Frauen unterscheidet sich

$$H_0: \mu_1 - \mu_2 = 0 \quad \Rightarrow \quad \mu_1 = \mu_2$$
$$H_1: \mu_1 - \mu_2 \neq 0 \quad \Rightarrow \quad \mu_1 \neq \mu_0$$

2.Ablehnungsbereich

Frage nach einem Unterschied $\Rightarrow$ ungerichtete Hypothese:

$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0,025$

$1 - \frac{\alpha}{2} = 1 - 0,025 = 0,975$

Freiheitsgrade t-test:

$Df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2 = 20 + 22 - 2 = 40$

In t-Tabelle nachschauen nach bei 0,975 und 40: $t_{krit} = \pm 2,021$

Übungsfrage - Lösung

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

3.Berechnung der Prüfgröße

$$t_{40} = \frac{(\bar{x}_1 - \bar{x}_2)}{\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)}}$$

$$\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= \sqrt{\frac{(20-1)\cdot 1{,}2 + (22-1)\cdot 1{,}3}{20+22-2}} \cdot \sqrt{\frac{1}{20} + \frac{1}{22}} \approx 1,829$$

$$t_{40} = \frac{(\bar{x}_1 - \bar{x}_2)}{1{,}829} = \frac{(7{,}5-6{,}5)}{1{,}829} = \frac{1}{1{,}829} \approx 0,55$$

Übungsfrage - Lösung

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

4.Interpretation der Ergebnisse

$t_{40} = 0,55$

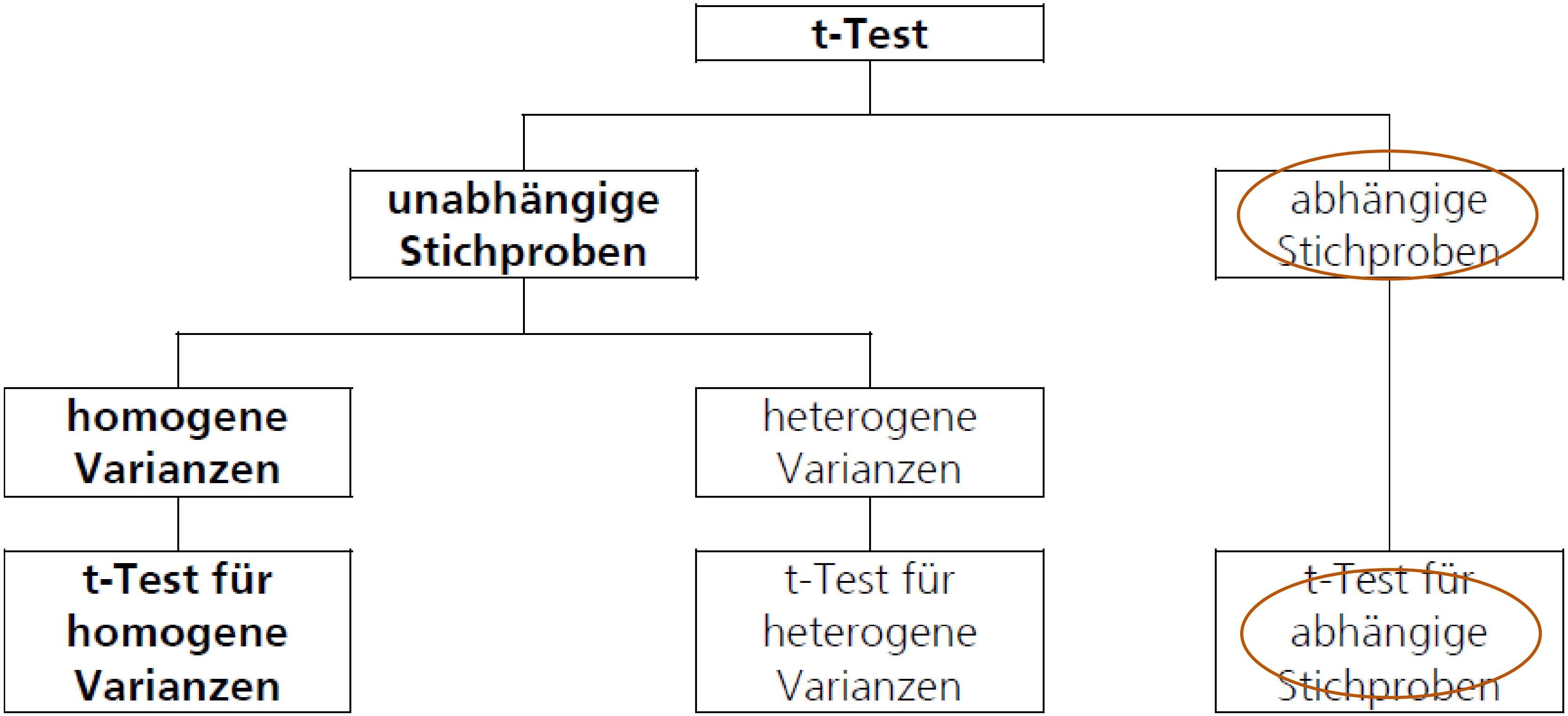
$t_{krit} = 2,021$

$0,55 < 2,021$

H_0 bleibt weiterhin bestehen

Es gibt keinen statistisch signifikanten Unterschied zwischen der mittleren Lebenszufriedenheit von Frauen im Vergleich zu Männern.

Varianten des t-Tests



Quelle: Eigene Darstellung

Ergänzung: t bei Varianzungleichheit

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

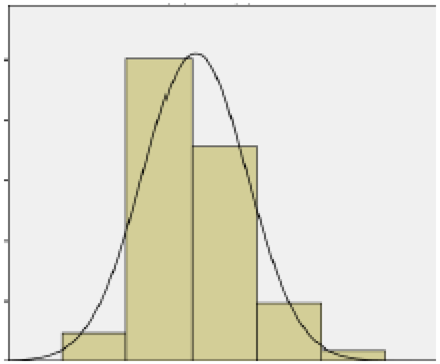
- (NICHT KLAUSURRELEVANT)
- Annahme, dass Varianzen in beiden GG identisch (Varianzhomogenität), ansonsten Korrektur notwendig (kann man ebenfalls testen, Levene-Test als F-Test)
- Bei heterogenen Varianzen wird ein statistisch angepasster t-Test berechnet (mit Welch-Korrektur im Nenner):
- $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

T-test bei abhängigen Stichproben

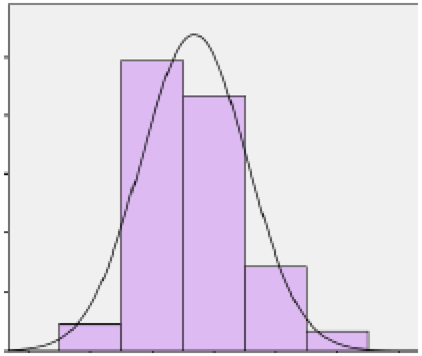
- Abhängige (oder gepaarte bzw. verbundene) Daten, „Stichproben“: von jeder Person liegen zwei Messungen vor → Berechnung von t berücksichtigt die Differenz der Wertepaare
- Bei *abhängigen* Messungen muss man davon ausgehen, dass die einzelnen Messungen *korrelieren*
- Durch korrelierte Messungen können Verzerrungen bei der Berechnung der Prüfgrößen entstehen, deshalb eigenes Verfahren

T-Test abhängige Stichproben

Stichprobe (n)



Ist die Veränderung im Mittelwert überzufällig?



Mittelwert: $\bar{x}_A$
Varianz s_A^2

Mittelwert: $\bar{x}_B$
Varianz s_B^2

T-test bei abhängigen Stichproben

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Beispiel: Wissen über gesunde Ernährung von Schüler*innen vor und nach einer Infoveranstaltung (Skala von 0 bis 10)

ID	Wert vor der Veranstaltung	Wert nach der Veranstaltung	Differenz
1	6	8	2
2	4	7	3
3	5	10	5
4	5	8	3
5	4	7	3
6	6	8	2
	Mittelwert: 5	Mittelwert: 8	Mittelwert der Differenzen: 3

Quelle: Eigene Darstellung

Wh. Wie wird Prüfgröße t berechnet?

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Abhängige Stichproben

t= $\dfrac{\bar{x}_d - \mu_d}{\hat{\sigma}_{\bar{x}_d}} = \dfrac{\bar{x}_d - 0}{\hat{\sigma}_{\bar{x}_d}} = \dfrac{\bar{x}_d}{\hat{\sigma}_{\bar{x}_d}}$

Standardfehler d. Mittelwert der Differenzen muss aus den Daten geschätzt werden

T-Test bei abhängigen Stichproben

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

1. Formulierung der Hypothesen
2. Bestimmung des Ablehnungsbereiches
3. Bestimmung der Prüfgröße
4. Interpretation / Schlussfolgerung

1. Formulierung der Hypothesen

Nullhypothese: (MW-)Unterschied ist zufällig zustande gekommen

(Es gibt keinen Unterschied zwischen dem mittleren Wissen vor und nach der Veranstaltung)

$$H_0: \mu_1 = \mu_2$$

Alternativ-/Forschungshypothese: (MW-)Unterschied ist überzufällig zustande gekommen (Es gibt einen Unterschied im mittleren Wissen vor und nach der Veranstaltung)

$$H_A\ \mu 1 \neq \mu 2$$

- Signifikanzniveau: 5%

2. Bestimmung Ablehnungsbereich

Signifikanzniveau wird auf 0,05 festgelegt (Das Risiko, H_0 auf Grund des Stichprobenergebnisses abzulehnen, obwohl diese in der Grundgesamtheit gilt, wird mit 5% festgelegt)

→ d.h. in unserem Beispiel: t_{krit} (2,571, siehe t-tabelle für df=5)

3. Bestimmung Prüfgröße

$$t = \frac{\bar{x}_d - \mu_d}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d - 0}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d}{\hat{\sigma}_{\bar{x}_d}}$$

Grundannahme Nullhypothese

Geschätzter Standardfehler d. Mittelwerts der Differenzen

3. Bestimmung Prüfgröße

$$t= \frac{\bar{x}_d - \mu_d}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d - 0}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d}{\hat{\sigma}_{\bar{x}_d}}$$

Standardfehler d. Mittelwerts der Differenzen:

$$\hat{\sigma}_{\bar{x}_d} = \frac{\hat{\sigma}_d}{\sqrt{n}}$$

Geschätzte Standardabweichung der mittleren Differenz

Schätzung Standardabweichung Mittelwerts der Differenzen:

$$\hat{\sigma}_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{x}_d)^2}{n-1}}$$

$$= \sqrt{\frac{(2-3)^2+(3-3)^2+(5-3)^2+(3-3)^2+(3-3)^2+(2-3)^2}{6-1}} = \sqrt{\frac{1+0+4+0+0+1}{5}} = \sqrt{\frac{6}{5}} = 1{,}10$$

3. Bestimmung Prüfgröße

$$t = \frac{\bar{x}_d - \mu_d}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d - 0}{\hat{\sigma}_{\bar{x}_d}} = \frac{\bar{x}_d}{\hat{\sigma}_{\bar{x}_d}}$$

Standardfehler d. Mittelwerts der Differenzen:

$$\hat{\sigma}_{\bar{x}_d} = \frac{\hat{\sigma}_d}{\sqrt{n}}$$

Schätzung Standardabweichung Mittelwerts der Differenzen:

$$\hat{\sigma}_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{x}_d)^2}{n-1}}$$

$$= \sqrt{\frac{(2-3)^2+(3-3)^2+(5-3)^2+(3-3)^2+(3-3)^2+(2-3)^2}{6-1}} = \sqrt{\frac{1+0+4+0+0+1}{5}} = \sqrt{\frac{6}{5}} = 1{,}10$$

$$t = \frac{\bar{x}_d}{\frac{\hat{\sigma}_d}{\sqrt{n}}} \rightarrow t = \frac{3}{\frac{1{,}1}{\sqrt{6}}} = 6{,}67$$

4. Interpretation

Da $t_{emp} > t_{krit}$ (2,57, siehe t-tabelle für df=5) kann die Nullhypothese abgelehnt werden. Forschungshypothese wird akzeptiert.

Das heißt, wir gehen in unserem Beispiel davon aus, dass die mittlere Differenz von 3 Skalenpunkten nach der Infoveranstaltung nicht zufällig entstanden sein wird und dass eine durchschnittliche Veränderung auch für die Grundgesamtheit erwartet wird.

Voraussetzung t-Tests

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Zufallsstichprobe
- (pseudo-)metrisches Skalenniveau der Variablen (→ für nominales Skalenniveau: Chi-Quadrat-Test)
- (annähernde) Normalverteilung des Merkmals in der Grundgesamtheit

- ermöglicht Vergleich zwischen *2 Gruppen* (→ für Vergleiche zwischen mehr Gruppen: Varianzanalyse mit F-Test - gleiche Logik, ähnliches Vorgehen)

Hypothesentests und Fehler

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Wir nutzen Stichprobendaten, um eine statistische Entscheidung im Rahmen von Hypothesentests zu treffen, die ihrerseits die Grundlage für inhaltliche Schlussfolgerungen darstellen
- Aber: Da unsere Stichprobendaten Zufallsauswahlen repräsentieren, sind zufällige Abweichungen und damit Fehlentscheidungen möglich

→ Fehler 1. Art (Alpha-Fehler) und Fehler 2. Art (Beta-Fehler)

Fehler 1. und 2. Art

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

	In Grundgesamtheit gilt die		
Entscheidung aufgrund der Stichprobe		H_0	H_A
	H_0	Richtige Entscheidung	β-Fehler
	H_A	α-Fehler	Richtige Entscheidung

Übung Signifikanzniveaus, p-Werte

Entscheidung über H_0 beruht auf berechneter Prüfgröße (z.B. z-Statistik, t-Statistik…)

Was heißt das? Ergänzen Sie:

- Hypothesentest "signifikant": H_0 wird …?
- Hypothesentest „nicht signifikant" (n.s.): H_0 wird …?

Wh. Signifikanzniveaus und p-Werte

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Entscheidung über H_0 beruht auf berechneter Prüfgröße (z.B. z-Statistik, t-Statistik, etc.)

Was heißt das?

- “signifikant”: H_0 wird abgelehnt
- „nicht signifikant“ (n.s.): H_0 wird beibehalten

Übungen Signifikanzniveaus, p-Werte

- α und p-Wert verweisen auf Irrtumswahrscheinlichkeit bzw. die Wahrscheinlichkeit eines Fehlers 1. Art/Typ-I-Fehler, also die Nullhypothese fälschlicherweise.....?
- ***p*< .05 :** Die Wahrscheinlichkeit für den empirischen Befund, wenn die … gültig ist, ist kleiner als 5 Prozent

Übungen Signifikanzniveaus und p-Werte

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- α und p-Wert verweisen auf Irrtumswahrscheinlichkeit bzw. die Wahrscheinlichkeit eines Typ-I Fehlers, also die Nullhypothese fälschlicherweise abzulehnen
- ***p*< .05 :** Die Wahrscheinlichkeit für den empirischen Befund, wenn die Nullhypothese gültig ist, ist kleiner als 5 Prozent

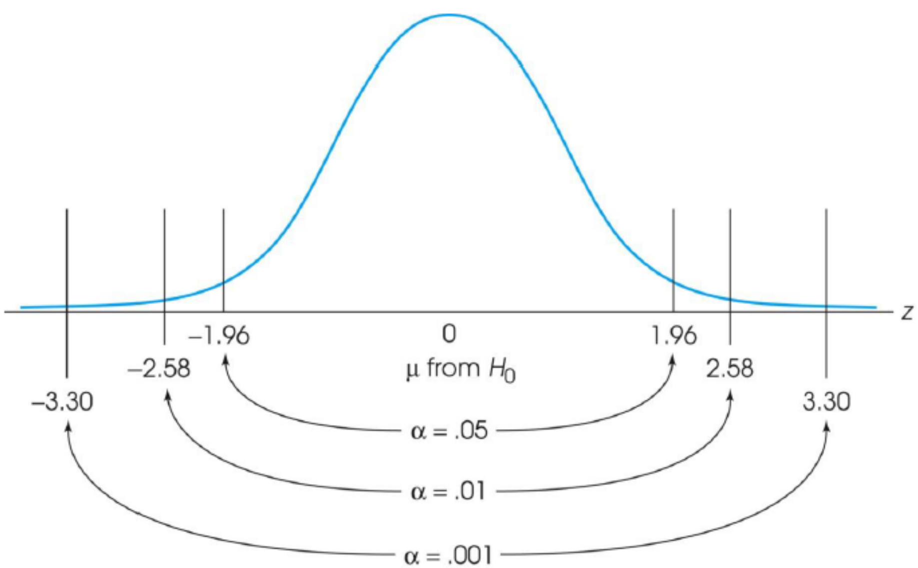
Übungen

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Treffen die folgenden Aussagen zu oder nicht?

- Wenn sich ein Stichprobenmittelwert für $\alpha = 0.01$ im Ablehnungsbereich befindet, befindet sich dieser Stichprobenmittelwert *immer* auch für $\alpha = 0.05$ im Ablehnungsbereich.
- Wenn sich ein Stichprobenmittelwert für $\alpha = 0.05$ im Ablehnungsbereich befindet, befindet sich dieser Stichprobenmittelwert *immer* auch für $\alpha = 0.01$ im Ablehnungsbereich.

Übung



1. Wenn sich ein Stichprobenmittelwert für α = 0.01 im Ablehnungsbereich befindet, befindet sich dieser Stichprobenmittelwert *immer* auch für α = 0.05 im Ablehnungsbereich. Richtig, denn ein Stichprobenmittelwert > 2.58 ist immer auch > 1.96

2. Wenn sich ein Stichprobenmittelwert für α = 0.05 im Ablehnungsbereich befindet, befindet sich dieser Stichprobenmittelwert *immer* auch für α = 0.01 im Ablehnungsbereich. Falsch, denn ein Stichprobenmittelwert > 1.96 ist nicht notwendigerweise auch > 2.58 (er könnte z.B. den Wert 2.4 annehmen)

Einheit heute, 2. Teil

- Lineare Regression, Teil 1 (Wdh. von Statistik I)
  - Um was geht‘s bei Regressionsanalysen überhaupt? (Grundlagen, Logik, Möglichkeiten)
  - Bivariates Regressionsmodell

Ergänzende Materialien:

- **Folien aus Stat I**
- **Lehrbrief Kapitel 4**
- **Lernmodul 4**

Zwischenstand

Was können wir mit bivariaten Analysen leisten?

Zusammenhänge zwischen 2 Variablen beschreiben und auf statistische Signifikanz prüfen

- Kreuztabellen und die damit verbundenen Maße (Chi-Quadrat, Cramers V) können Auskunft über Zusammenhänge zwischen zwei (nicht-metrischen) Variablen liefern
- Korrelationskoeffizienten (z.B. Pearsons r) beschreiben die Stärke eines linearen Zusammenhangs zweier Variablen
- Mittelwertvergleiche mit T-Testverfahren können testen, ob der Unterschied zwischen zwei Mittelwerten statistisch signifikant ist

Zusammenfassung

Was können bisherige Analyseverfahren **nicht** leisten?

- Überprüfen, inwieweit ein abhängiges Merkmal womöglich noch von weiteren Variablen abhängig ist
- Kausalitätsrichtung zwischen Variablen untersuchen

→ Multivariate Verfahren, v.a. Regressionsanalyse

Multivariate Analysen

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Multivariate Analysen beziehen mehr als 2 Variablen in Analysen ein
- Klassische Unterscheidung zwischen

(1) Strukturentdeckenden/dimensionsreduzierenden Verfahren, z.B. Faktoren-/Hauptkomponentenanalyse und Clusteranalyse

(2) Strukturprüfenden Verfahren, z.B. Regressionsanalyse-verfahren

- Können für die Einflüsse möglicher Drittvariablen kontrollieren
- Können eine abhängige Variable durch mehrere unabhängige Variablen „erklären“

Regressionsanalyse: Einführung

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

- Mit Regressionsanalysen adressiert man zwei Fragen:

  (1) Wie gut erklären bestimmte Faktoren (unabhängige Variablen) die Varianz einer abhängigen Variable?

  (2) Welchen Einfluss üben die einzelnen Faktoren auf diese abhängige Variable aus (unter Konstanthalten des Einflusses der anderen unabhängigen Variablen)?

➢ Damit lassen sich (theoretisch entwickelte) Hypothesen über die Beeinflussungsstruktur bestimmter Variablen auf andere Variablen prüfen

➢ Beispiel?

Regressionsanalyse: Einführung

- Korrelationsanalyse behandelt X und Y „gleichwertig“, d.h.
  - Unterscheidung *Explanans* (uV) und *Explanandum* (aV) spielt statistisch keine Rolle
- Regressionsanalytische Terminologie: Y ist Explanandum/Kriterium, X ist Explanans/Prädiktor
- Y soll durch X erklärt bzw. vorhergesagt bzw. auf X „zurückgeführt“ werden

Regressionsanalyse: Einführung

Tabelle 46: Unterschiedliche Bezeichnungen für Variablen der Regressionsanalyse

Abhängige Variable (y)	Unabhängige Variable (x)
Erklärte Variable	Erklärende Variable
Kriteriums(variable)	Prädiktor(variable)
Endogene Variable	Exogene Variable
Regressand	Regressor

Quelle: Eigene Darstellung

Tabelle aus Lehrbrief, Kapitel 4

Regressionsanalyse: Einführung

- Hauptfunktion:

1) Statistische Methode, um den Einfluss mehrerer unabhängiger Variablen (uV) auf eine abhängige Variable (aV) zu untersuchen
2) Informationsreduktion dieser Untersuchung auf wenige Kennzahlen (z.B. ∝, β und R^2)

- Formal:

1) Misst den Einfluss von (einer) uV (X) auf eine aV (Y)
2) Die einzelnen Ausprägungen von Y hängen funktional von den jeweiligen Ausprägungen der X-Variable(n) ab:
   - $Y = f(X)$

Regressionsanalyse: Einführung

Was ist:

- die Richtung
- die Stärke
- die statistische Signifikanz

… des Einflusses von X auf Y?