

UNIVERSITÄT LEIPZIG

Medizinische Fakultät

QSB 1: Epidemiologie, Medizinische Biometrie und Medizinische Informatik

Übung 7
Statistisches Testen III (χ^2-Test)

Wintersemester 2026/27

imise.

Lernziele

Die Studierenden sollen

- aus einer Beschreibung Kontingenztafeln aufstellen,
- Tests auf Unabhängigkeit bzw. den Vergleich von Häufigkeiten selbstständig durchführen,
- Häufigkeitsdifferenz bzw. Odds Ratio als Effektmaße berechnen und
- statistische Signifikanz und klinische Relevanz sicher unterscheiden können.

Merkmale

Quantitative (metrische) Merkmale

- Kalorienverbrauch/Tag
- Hämoglobin im Blut (mg/dl)

Qualitative (nominale / ordinale) Merkmale

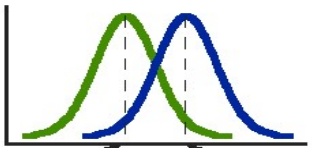
- Geschlecht (m/w/d)
- Blutgruppe
- NYHA-Stadium

Mögliche Fragestellungen

Welche Tests sind für welche Fragestellung geeignet?

Vergleich von Mittelwerten

Z. B.: Mittlere Änderung der Herz-frequenz in den Gruppen A und B



z. B. **t-Test**

Vergleich von Häufigkeiten

sportlicher Betätigung (j | n)
je nach Raucherstatus (j | n)

oder Test auf Unabhängigkeit

z. B. χ^2**-Test**

Ablaufschema bei Hypothesentests

vgl. Skriptenheft Kapitel 5.6, S. 59

1. Wissenschaftliche Fragestellung → statistisches Modell
2. Ziehen der Stichprobe und Erfassung der Daten
3. Formulierung der Hypothesen H_0 / H_A
4. Wahl des Signifikanzniveaus α
5. Wahl des Tests / der Prüfgröße
6. Prüfung der Testvoraussetzungen
7. Ermittlung des Akzeptanzbereiches
8. Berechnung der Prüfgröße
9. Testentscheidung
10. Interpretation der Ergebnisse im Kontext der Fragestellung (inkl. Angabe eines Effektschätzers)

χ^2-Test - Beispiel 1

Herzsportstudie (Übung 1)

Fragestellung:

Gibt es einen Zusammenhang zwischen Rauchen und sportlicher Betätigung?

1. Einordnung:

- zwei binäre Merkmale
- Prüfung eines Zusammenhangs

→ Chi² –Test

Beispiel 1 - Hypothesen H_0 und H_A

Nullhypothese H_0:

Zwischen Rauchen und sportlicher Betätigung besteht kein Zusammenhang.

Alternativhypothese H_A:

Zwischen Rauchen und sportlicher Betätigung besteht ein Zusammenhang.

Die Hypothese der Gleichheit der Wahrscheinlichkeit von Rauchen in den beiden Gruppen:

$$H_0: p_{Sport} = p_{keinSport}$$

$$H_A: p_{Sport} \neq p_{keinSport}$$

führt zu demselben Test.

Beispiel 1 - Signifikanzniveau α

Als Signifikanzniveau wird $\alpha = 0{,}05$ festgelegt.

Kreuztabelle (Vierfeldertafel) aus der Stichprobe:

Kreuztabelle

		regelmäßig Sport		
		nein	ja	Gesamt
Raucher	nein	130	63	193
	ja	47	21	68
Gesamt		177	84	261

Beispiel 1 - Testvoraussetzungen prüfen

Testvoraussetzungen bei 2×2-Kontingenztafel:

- $n \geq 40$
- alle Erwartungswerte $e_{ij} \geq 5$

		regelmäßig Sport		
		nein	ja	Gesamt
Raucher	nein	130	63	$n_{1.}$
	ja	47	21	$n_{2.}$
Gesamt		$n_{.1}$	$n_{.2}$	n

$$e_{ij} = \frac{n_{i.} \cdot n_{.j}}{n}$$

Berechnen Sie die 4 Erwartungswerte!

Beispiel 1 - Ermittlung des Akzeptanzbereiches

→ Quantil $\chi^2_{f;1-\alpha}$ aus Tabelle ablesen

Signifikanzniveau: $\alpha = 0{,}05$

Freiheitsgrade (FG): $f = (\text{Zeilenzahl} - 1) \cdot (\text{Spaltenzahl} - 1)$

$= (2-1) \cdot (2-1) = 1$

Quantil: $\chi^2_{1;1-0{,}05} = 3{,}84$

Akzeptanzbereich: $[0; 3{,}84)$ ist komplementär zum

Ablehnbereich: $[3{,}84; \infty)$

	Merkmal 2		
Merk mal 1	a	b	a+b
	c	d	c+d
	a+c	b+d	n

Beispiel 1 - Berechnung der Prüfgröße

Für jede Zelle wird berechnet: $\frac{(\text{beobachtete Anzahl } b_{ij} - \text{erwartete Anzahl } e_{ij})^2}{\text{erwartete Anzahl } e_{ij}}$

Und summiert:

$$X^2 = \sum_{i,j} \frac{(b_{ij} - e_{ij})^2}{e_{ij}}$$

Spezialfall Vierfeldertafel:

$$X^2 = \frac{(a \cdot d - b \cdot c)^2}{(a+c) \cdot (b+d) \cdot (a+b) \cdot (c+d)} \cdot n \sim \chi_1^2$$

Einsetzen der Werte:

$$X^2 = \frac{(130 \cdot 21 - 63 \cdot 47)^2}{177 \cdot 84 \cdot 193 \cdot 68} \cdot 261 = 0{,}07$$

Beispiel 1 - Testentscheidung: H_0 oder H_A ?

Handrechnung: Vergleiche den berechneten Wert für X^2 mit dem Quantil der Prüfverteilung

$$X^2 = 0{,}07 \quad < \quad \chi^2_{1;1-0{,}05} = 3{,}84$$

Software: Vergleiche den p-Wert mit dem Signifikanzniveau α

$$p = 0{,}789 \quad > \quad \alpha = 0{,}05$$

Die Nullhypothese wird auf einem Signifikanzniveau von 5% nicht abgelehnt (beibehalten).

Beispiel 1 - Interpretation

		regelmäßig Sport		
		nein	ja	Gesamt
Raucher	nein	130	63	193
	ja	47	21	68
Gesamt		177	84	261

Effektmaß: z. B. Differenz der relativen Häufigkeiten:

$\frac{47}{177} - \frac{21}{84} = 0{,}016$, d. h. 1,6% .

In dieser Kohorte gibt es unter Studenten, die nicht regelmäßig Sport treiben, 1,6% mehr Raucher als unter denen, welche regelmäßig Sport treiben. Dieser Unterschied ist nicht signifikant ($\alpha = 5\%$).

Aufgabe 1 - Literaturbeispiel

Table 2. Efficacy and Safety.*

Variable	Canagliflozin	Placebo
Efficacy		
Primary composite outcome	245/2202	340/2199
Doubling of serum creatinine level	118/2202	188/2199
End-stage kidney disease	116/2202	165/2199

Typ-II-Diabetiker mit eingeschränkter Nierenfunktion wurden mit Canagliflozin bzw. Placebo behandelt.

Studienfrage:

Gibt es Unterschiede in der Häufigkeit von Patienten, die das Kreatinin-Level im Laufe der Studie verdoppelten?

Signifikanzniveau 0,1%.

Aufgabe 1

- Überführen Sie die klinische Frage in ein mathematisches Problem!
  - Formulieren Sie die Hypothesen!
  - Stellen Sie die Vierfeldertafel auf!
  - Berechnen Sie die relativen Häufigkeiten pro Arm!
- Führen Sie den Test durch und treffen Sie die Testentscheidung. Die Testvoraussetzungen sind erfüllt.
- Interpretieren Sie das Ergebnis. Wählen Sie als Effektmaß die Differenz der Häufigkeiten!

Aufgabe 1 - Interpretation

In der Placebogruppe gibt es 3,19% mehr Patienten, deren Kreatinin-Level sich verdoppelte. Dieser Unterschied ist signifikant ($\alpha = 0{,}001$).

Aus der Häufigkeitsdifferenz kann man unmittelbar die "Number Needed to Treat" berechnen:

$$NNT = \frac{1}{Häufigkeitsdifferenz} = \frac{1}{0{,}0319} \approx 31$$

Number needed to treat for benefit: $NNT \approx 31$

D. h. im Mittel profitiert von 31 mit Canagliflozin Behandelten einer hinsichtlich der Verdopplung des Kreatinin-Levels.

Hinweise

Wenn Testvoraussetzungen nicht erfüllt (erwartete Häufigkeiten klein):

⇨ Fisher‘s exakter Test

Bei abhängigen Beobachtungen:

⇨ McNemar – Test

Beispiel: Schmerzen vor Operation vs.
Schmerzen nach Operation

Hausaufgaben

Table 2. Efficacy and Safety.*

Variable	Canagliflozin	Placebo
Efficacy		
Primary composite outcome	245/2202	340/2199
Doubling of serum creatinine level	118/2202	188/2199
End-stage kidney disease	116/2202	165/2199

1. Berechnen Sie für die experimentelle Gruppe vs. Placebo das Odds Ratio für verdoppeltes Kreatinin!
2. Prüfen Sie die Hypothese, dass sich die Behandlungsarme hinsichtlich der Häufigkeit von Nierenversagen (end-stage kidney disease) unterscheiden (Testvoraussetzungen erfüllt, $\alpha = 0{,}001$)! Berechnen Sie als Effektmaß die Differenz der relativen Häufigkeiten von Nierenversagen!

Hausaufgabe 3

In einer Studie des deutschen SepNet (VISEP Studie) wurde die konventionelle und die intensivierte Insulintherapie bei Patienten mit schwerer Sepsis oder septischem Schock untersucht. Primärer Endpunkt war die 28-Tage-Mortalität (Brunkhorst et al. N Engl J Med 2008; 358:125-39.).

Beurteilen Sie auf einem Signifikanzniveau von 5%, ob sich die 28 Tage Mortalität bei Patienten mit schwerer Sepsis oder septischem Schock nach intensivierter und konventioneller Insulintherapie unterscheidet!

Testvoraussetzungen sind erfüllt.

Hausaufgabe 3 - Studienergebnisse

Table 2. Primary and Secondary Outcomes.*

Variable		Insulin Therapy		
	All Patients (N=537)	Conventional (N=290)	Intensive (N=247)	P Value†
Death				
At 28 days§				0.74
No./total no.	136/536	75/289	61/247	
Percent (95% CI)	25.4 (21.7–29.1)	26.0 (20.9–31.0)	24.7 (19.3–30.1)	

- Treffen Sie anhand der gegebenen Daten die Testentscheidung (Signifikanzniveau 5%) und interpretieren Sie das Ergebnis!

Hausaufgabe 4

In der VISEP-Studie wurde auch die Gabe von kristalloidem Volumenersatz gegenüber der Gabe von kolloidalem Volumenersatz getestet (zweifaktorielles Studiendesign).

Die 90-Tage-Mortalität nach kristalloidem Volumenersatz lag bei 33,9% (93 / 274 Todesfälle) und nach kolloidalem Volumenersatz bei 41,0% (107 / 261 Todesfälle; p = 0,09).

- Treffen Sie die Testentscheidung ($\alpha = 5\%$) und interpretieren Sie das Ergebnis anhand der Häufigkeitsdifferenz! (Testvoraussetzungen sind erfüllt.)

Volumenersatztherapie: Anwendung in der Notfall- und Intensivmedizin zur Therapie einer Hypovolämie oder zur Volumensubstitution bei Schock angewendet (Gabe als Infusion)

Hausaufgabe 5

Für den kolloidalen Volumenersatz wurde eine maximale Tagesdosis im Prüfplan vorgegeben, die jedoch bei 100 von 262 Patienten an mindestens einem Studientag um mehr als 10 Prozent überschritten wurde.

Die Auswirkung des Protokollverstoßes auf die 90-Tages-Mortalität wurde in einer Sensitivitätsanalyse untersucht.

Beurteilen Sie auf einem Signifikanzniveau von 1%, ob sich die 90 Tage Mortalität in den 3 Gruppen unterscheidet!

(Testvoraussetzungen sind erfüllt)

Hausaufgabe 5 - Studienergebnisse, Prüfgröße

	Volumenersatz			
	Kristalloid	Kolloidal		total
		Überschreitung Tagesdosis		
90-Tages-Mortalität		nein	ja	
überlebend	181 (66%)	112 (69%)	42 (42%)	335 (63%)
verstorben	93 (34%)	50 (31%)	57 (58%)	200 (37%)
total	274(1)	162	99(1)	535

(1) 90 Tage Mortalität jeweils für einen Patienten nicht bekannt

Der Wert der Teststatistik X^2 beträgt 21,572.