

# Statistik I

---

## Einheit 10: Verfahren für Nominaldaten - $\chi^2$ – Test

25.06.2025 | Prof. Dr. Stephan Goerigk

# Verfahren für Nominaldaten

## Kurzvorstellung

### Voraussetzungen bisher gelernter Tests:

- Viele Hypothesentests nutzen Kombinationen aus numerischen vs. kategorialen UVs

Hypothesentest	AV	UVs	Fragestellung	Teststatistik
Ein-Stichproben t-Test	Intervallskaliert	Keine UV, nur Referenzwert	Unterschied zwischen Stichprobenmittelwert und Referenzwert?	t-Wert
Unabhängiger t-Test	Intervallskaliert	1 kategoriale UV, 2 Stufen	Unterschied zwischen 2 Gruppen?	t-Wert
Abhängiger t-Test	Intervallskaliert	1 UV Messwiederholung, 2 Messungen	Unterschied zwischen 2 Messzeitpunkten?	t-Wert
Einfaktorielle ANOVA	Intervallskaliert	1 kategoriale UV, $\geq 2$ Stufen	Unterschied zwischen $\geq 2$ Gruppen?	F-Wert
ANOVA mit Messwiederholung	Intervallskaliert	1 UV Messwiederholung, $\geq 2$ Messungen	Unterschied zwischen $\geq 2$ Messzeitpunkten?	F-Wert
Einfache Regression	Intervallskaliert	1 kategoriale UV oder 1 stetige UV	Kann UV die AV vorhersagen?	t-Wert (Steigung) oder F-Wert (Omnibus)
Mehrfaktorielle ANOVA	Intervallskaliert	2 kategoriale UVs	Unterschiede zwischen den Stufen der Faktoren? Besteht Interaktion?	F-Wert
Multiple Regression	Intervallskaliert	2 kategoriale oder stetige UVs	Können UVs die AV vorhersagen? Besteht Interaktion?	t-Wert (Steigung) oder F-Wert (Omnibus)
Mixed ANOVA	Intervallskaliert	2 UVs, davon 1 kategoriale UV und eine Messwiederholung	Unterschiede zwischen Stufen und Zeitpunkten? Besteht Interaktion?	F-Wert

# Verfahren für Nominaldaten

## Kurzvorstellung

### Voraussetzungen bisher gelernter Tests:

- ABER: Alle bislang kennengelernten statistischen Tests beinhalten **intervallskalierte AVs**
- Was können wir tun, wenn wir eine **nominalskalierte AV** haben?

Zunächst: nominalskalierte Variablen mit 2 Merkmalsausprägungen (dichotom aka. binär):

- Klassische Beispiele:
  - richtig vs. falsch
  - krank vs. gesund
  - rückfällig vs. nicht rückfällig
  - tot vs. lebendig

# Verfahren für Nominaldaten

## Kurzvorstellung - Zur Erinnerung:

- Nominalskalierte Variablen sind eine Art von Variablen, bei denen die Werte Kategorien oder Namen repräsentieren
- Die Kategorien oder Namen haben keine natürliche Ordnung oder Rangfolge (z.B. Geschlecht, Nationalität oder Augenfarbe).
- Es können nur Aussagen über die Gleichheit oder Ungleichheit der Kategorien gemacht werden.
- Es ist nicht möglich, Aussagen über die Größe oder den Abstand zwischen den Kategorien zu treffen.
- Die Umwandlung in andere Skalenniveaus wie ordinal oder metrisch ist nicht sinnvoll, da die Informationen über die Rangfolge oder die Abstände zwischen den Kategorien nicht vorhanden sind.
- Bei der Darstellung nominalskalierter Variablen werden häufig Balkendiagramme verwendet, um die Häufigkeit oder Verteilung der einzelnen Kategorien zu veranschaulichen.

# Verfahren für Nominaldaten

## Deskriptivstatistiken

- Lage- und Streuungsmaße lassen sich nicht berechnen
- ABER: Kategorien können ausgezählt werden (Häufigkeiten)

### Absolute Häufigkeiten (n):

```
table(data$Behandlungserfolg)
```

```
##
##   ja  nein
##    6    4
```

### Relative Häufigkeiten (%):

```
prop.table(table(data$Behandlungserfolg))
```

```
##
##   ja  nein
##  0.6  0.4
```

## Beispiel - nominalskalierte Variablen:

Behandlung	Behandlungserfolg
Psychotherapie	ja
Psychotherapie	ja
Psychotherapie	nein
Psychotherapie	ja
Psychotherapie	ja
Warteliste	nein
Warteliste	nein
Warteliste	nein
Warteliste	ja
Warteliste	ja

# Verfahren für Nominaldaten

## Kreuztabelle (aka. Kontingenztafel)

- Tabelle zur Verteilung von zwei oder mehr nominalskalierten Variablen.
- Werte beider Variablen in Zeilen und Spalten aufgeteilt
- Zellen enthalten Kombinationen beider Variablen.

### Absolute Häufigkeiten (n):

```
table(data$Behandlungserfolg, data$Behandlung)
```

```
##
##      Psychotherapie Warteliste
##   ja              4           2
##   nein            1           3
```

### Relative Häufigkeiten (%):

```
prop.table(table(data$Behandlungserfolg, data$Behandlung))
```

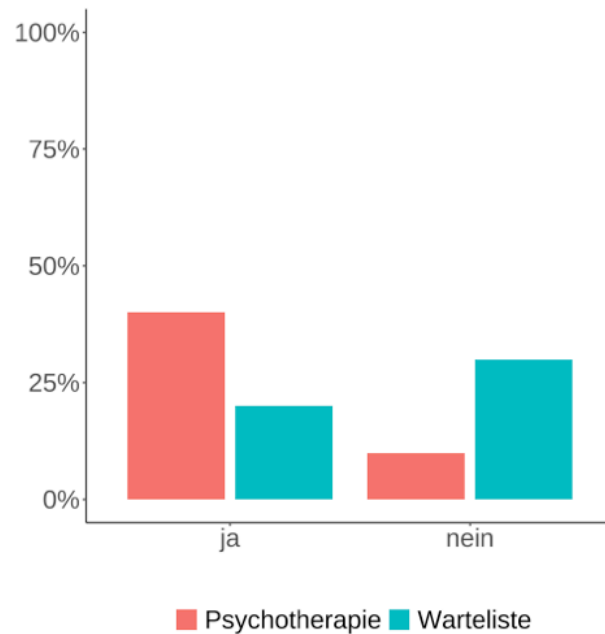
```
##
##      Psychotherapie Warteliste
##   ja              0.4         0.2
##   nein            0.1         0.3
```

## Beispiel - nominalskalierte Variablen:

Behandlung	Behandlungserfolg
Psychotherapie	ja
Psychotherapie	ja
Psychotherapie	nein
Psychotherapie	ja
Psychotherapie	ja
Warteliste	nein
Warteliste	nein
Warteliste	nein
Warteliste	ja
Warteliste	ja

# Verfahren für Nominaldaten

## Visualisierung im Balkendiagramm (Wiederholung)



## Beispiel - nominalskalierte Variablen:

Behandlung	Behandlungserfolg
Psychotherapie	ja
Psychotherapie	ja
Psychotherapie	nein
Psychotherapie	ja
Psychotherapie	ja
Warteliste	nein
Warteliste	nein
Warteliste	nein
Warteliste	ja
Warteliste	ja

# Verfahren für Nominaldaten

## $\chi^2$ -Tests

- Relative Häufigkeit in Stichprobe dient als Schätzer für Auftretenswahrscheinlichkeit in Population

Logik ähnlich wie t-Test ( $H_0$  des t-Tests: Mittelwerte sind gleich)

- $H_0$  : Beispiel: Kategorien sind **gleich verteilt** (Gleichverteilungshypothese).
- Die unter der  $H_0$  erwarteten Häufigkeiten werden mit beobachteten Häufigkeiten (Stichprobe) verglichen
- Testverteilung:  $\chi^2$ -Verteilung → hat eigene Tabelle
- Entscheidungslogik:
  - Vergleich empirischer  $\chi^2$ -Wert (Berechnung aus Daten) vs. kritischer  $\chi^2$ -Wert (aus Tabelle)
  - Wenn  $\chi_{emp}^2 > \chi_{krit}^2$  ist der Test signifikant



# Verfahren für Nominaldaten

## Eindimensionaler $\chi^2$ -Test

- Prüft Hypothesen über die Verteilung **einer** kategorialen Variablen
- Versuchspersonen hinsichtlich Merkmal mit  $k$  Stufen kategorisiert
- Stichprobe: Es liegt Verteilung mit absoluten Häufigkeiten vor (beobachtete Häufigkeiten)
- Aufgabe  $\chi^2$ -Test: Ermitteln, ob Verteilung in Stichprobe Annahme über Population entspricht

## Beispiel: Suizidraten bei Männern und Frauen

- Frage: Entsprechen in einer Stichprobe beobachtete Suizidraten von Männern und Frauen einer theoretisch erwarteten Verteilung?

# Verfahren für Nominaldaten

## $\chi^2$ -Tests

### Nullhypothese:

- Entscheidung für  $H_1$  über Ablehnung der  $H_0 \rightarrow$  Wenn  $H_0$  ausreichend unwahrscheinlich, wird  $H_1$  angenommen.
- $\chi^2$ -Test prüft, ob beobachtete Häufigkeiten von erwarteten Häufigkeiten abweichen.
- Erwartete Häufigkeiten entsprechen der  $H_0$  des Tests.
- Besonderheit  $\chi^2$ -Test: Jede Annahme über Verteilung kann als  $H_0$  dienen.

Gleichverteilungsannahme (häufig):

- $H_0$  : "Verteilung der Geschlechter in Population ist 50% vs. 50%."
- $H_1$  : "Geschlechter sind ungleich verteilt."

Nicht gleich verteilte Annahmen (denkbar):

- $H_0$  : "Verteilung der Geschlechter in Population ist 30% vs. 70%."
- $H_1$  : "Verteilung der Geschlechter weicht signifikant von dieser Annahme ab."

# Verfahren für Nominaldaten

## $\chi^2$ -Tests

### Nullhypothese:

Gleichverteilungsannahme:

- Häufigkeiten ( $f$ ) sind über alle Stufen des Merkmals hinweg gleich
- erwartete Häufigkeit jeder Zelle:

$$f_{e1} = f_{e2} = \dots f_{ek} = \frac{N}{k}$$

Beispiel - Suizidrate nach Geschlecht:

In Stichprobe:

Frauen	Männer	Summe
101	223	324

Erwartete Werte unter Gleichverteilungsannahme:

Frauen	Männer	Summe
162	162	324

# Verfahren für Nominaldaten

## $\chi^2$ -Tests

### Nullhypothese:

Nicht gleich verteilte Annahmen:

- Verteilung der Häufigkeiten ( $f$ ) entspricht theoretischen Vorüberlegungen (begründete Festlegung der Verteilung)
- erwartete Häufigkeit jeder Zelle: Multiplikation mit angenommener Auftretenswahrscheinlichkeit ( $p_i$ )

$$f_{ei} = N \cdot p_i$$

Beispiel - Suizidrate nach Geschlecht:

In Stichprobe:

Frauen	Männer	Summe
101	223	324

Erwartete Werte nach theoretischer Vorüberlegung (Männer begehen 3x häufiger Suizid):

Frauen	Männer	Summe
$324 \cdot 0.25 = 81$	$324 \cdot 0.75 = 243$	324

# Verfahren für Nominaldaten

## $\chi^2$ -Wert

- Entscheidung über signifikante Unterschiede zwischen beobachteten und erwarteten Häufigkeiten erfolgt über  $\chi^2$ -Wert
- Folgt der  $\chi^2$ -Verteilung → Wahrscheinlichkeit empirischer Werte unter Annahme der  $H_0$  bestimmbar

## Berechnung

- Was wird benötigt: beobachtete und erwartete absolute Häufigkeiten für alle Merkmalstufen
- $\chi^2$ -Wert gibt Abweichung der beobachteten von den erwarteten Häufigkeit an

$$\chi^2 = \sum_{i=1}^k \frac{(f_{bi} - f_{ei})^2}{f_{ei}}$$

mit:

- $k$  : Anzahl der Merkmalskategorien (Index  $i$ )
- $f_{bi}$  beobachtete absolute Häufigkeit von Kategorie  $i$
- $f_{ei}$  unter  $H_0$  erwartete absolute Häufigkeit von Kategorie  $i$

# Verfahren für Nominaldaten

## $\chi^2$ -Wert

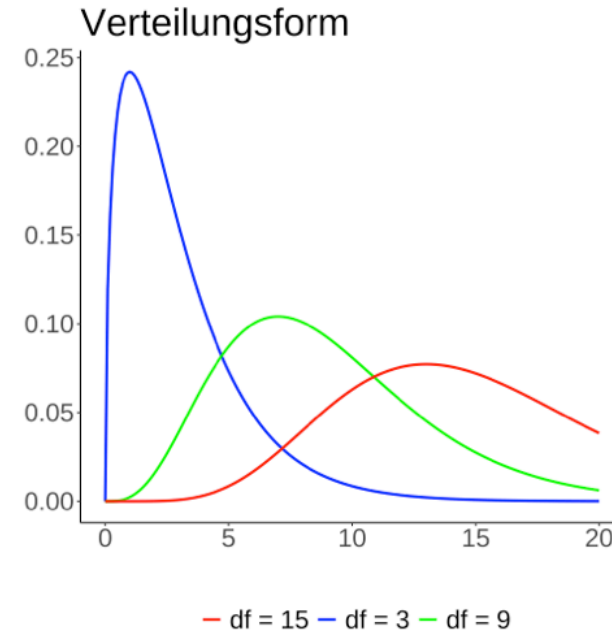
### Eigenschaften des $\chi^2$ -Werts

- $\chi^2 = 0$ , wenn beobachtete und erwartete Häufigkeiten in allen Zellen genau übereinstimmen
- Je größer die Abweichung der beobachteten von erwarteten Häufigkeiten, desto größer wird  $\chi^2$
- $\chi^2$  kann aufgrund der Quadrierung in der Formel nur positive Werte annehmen
  - Informationen über Richtung der Abweichung geht verloren
  - unspezifischer Test → es können keine gerichteten Hypothesen getestet werden
  - Ausnahme: eindimensionaler  $\chi^2$ -Test mit 2 Stufen

# Verfahren für Nominaldaten

## $\chi^2$ -Verteilung

- $\chi^2$  kann aufgrund der Quadrierung in der Formel nur positive Werte annehmen
- Wertebereich von 0 bis  $\infty$
- Form ist abhängig von Anzahl der Freiheitsgrade ( $df$ )
- Fläche unter der Kurve gibt an, wie wahrscheinlich ein  $\chi^2$ -Wert ist



# Verfahren für Nominaldaten

## $\chi^2$ -Wert

### Berechnung im Beispiel:

$$\chi^2 = \sum_{i=1}^k \frac{(f_{bi} - f_{ei})^2}{f_{ei}} = \frac{(101 - 81)^2}{81} + \frac{(223 - 243)^2}{243} = 6.58$$

- Für die  $H_0$ , dass die Häufigkeiten den theoretischen Annahmen entsprechen erhalten wir  $\chi^2 = 6.58$
- Zur Interpretation dieses Werts benötigen wir noch die Freiheitsgrade ( $df$ )

Notation	Frauen	Männer
$f_{bi}$	101	223
$f_{ei}$	$324 \cdot 0.25 = 81$	$324 \cdot 0.75 = 243$
$f_{bi} - f_{ei}$	$101 - 81 = 20$	$223 - 243 = -20$
$(f_{bi} - f_{ei})^2$	$20^2 = 400$	$-20^2 = 400$
$\frac{(f_{bi} - f_{ei})^2}{f_{ei}}$	$\frac{400}{81} = 4.94$	$\frac{400}{243} = 1.65$



# Verfahren für Nominaldaten

## $\chi^2$ -Wert

### Bestimmung der Freiheitsgrade:

- Anzahl der Summanden in der Formel die unabhängig voneinander variieren können
- Für den eindimensionalen  $\chi^2$ -Test:

$$df = k - 1$$

Berechnung im Beispiel:

$$df = k - 1 = 2 - 1 = 1$$

# Verfahren für Nominaldaten

$\chi^2$ -Wert

## Signifikanzprüfung:

- $\chi^2_{emp} = 6.58$
- $df = k - 1$
- $\alpha = .05$

Ablesen von  $\chi^2_{krit}$  aus der Tabelle:

- $\chi^2_{df=1} = 3.84$

Vergleich  $\chi^2_{emp}$  vs.  $\chi^2_{krit}$  :

- $6.58 > 3.84 \rightarrow$  Test ist signifikant

Fläche df	0,750	0,900	0,950
1	1,32330	2,70554	3,84146
2	2,77259	4,60517	5,99147
3	4,10835	6,25139	7,81473
4	5,38527	7,77944	9,48773
5	6,62568	9,23635	11,0705
6	7,84080	10,6446	12,5916
7	9,03715	12,0170	14,0671
8	10,2188	13,3616	15,5073
9	11,3887	14,6837	16,9190
10	12,5489	15,9871	18,3070
11	13,7007	17,2750	19,6751
12	14,8454	18,5494	21,0261
13	15,9839	19,8119	22,3621
14	17,1170	21,0642	23,6848
15	18,2451	22,3072	24,9958
16	19,3688	23,5418	26,2962
17	20,4887	24,7690	27,5871
18	21,6049	25,9894	28,8693
19	22,7178	27,2036	30,1435

# Verfahren für Nominaldaten

## Eindimensionaler $\chi^2$ -Test in R

```
# Stichprobendaten erstellen (1 Vektor und keine Tabelle, da eindimensional)
```

```
x = c(rep("Männer", 223), rep("Frauen", 101))  
table(x)
```

```
## x  
## Frauen Männer  
##    101    223
```

```
chisq.test(table(x), p = c(0.25, 0.75))
```

```
##  
##      Chi-squared test for given probabilities  
##  
## data:  table(x)  
## X-squared = 6.5844, df = 1, p-value = 0.01029
```

# Verfahren für Nominaldaten

## Gerichteter eindimensionaler $\chi^2$ -Test

- $\chi^2$ -Test ist normalerweise ungerichtet (wegen Quadrierung der Abweichungen)
- Spezialfall eindimensionaler  $\chi^2$ -Test bei Variable mit genau 2 Stufen
- In diesem Fall ist die Richtung der Abweichung eindeutig
- Interpretation: "Merkmalsstufe 1 tritt öfter auf als Merkmalsstufe 2."
- VORSICHT: Signifikanzniveau ( $\alpha$ ) wird dann verdoppelt (z.B. wenn  $\alpha = .05$  wird in der Tabelle  $\alpha = .10$  angenommen)
- Folge:  $\chi^2_{krit}$  verringert sich, Test wird eher signifikant (Teststärke nimmt zu)

# Verfahren für Nominaldaten

## Eindimensionaler $\chi^2$ -Test

### Effektstärke:

- Wie immer: Standardisiertes Maß für Größe des systematischen Unterschieds
- Gängige Effektstärke beim  $\chi^2$ -Test:  $w^2$
- Schätzung von  $w^2$  aus Stichprobendaten:

$$\hat{w}^2 = \frac{\chi^2}{N}$$

- Im Beispiel:  $\hat{w}^2 = \frac{\chi^2}{N} = \frac{6.58}{324} = 0.02$
- VORSICHT: G\*Power nutzt unquadrierte Größe  $w$  (aka "Phi":  $\varphi$ )

Konventionen nach Cohen (1988):

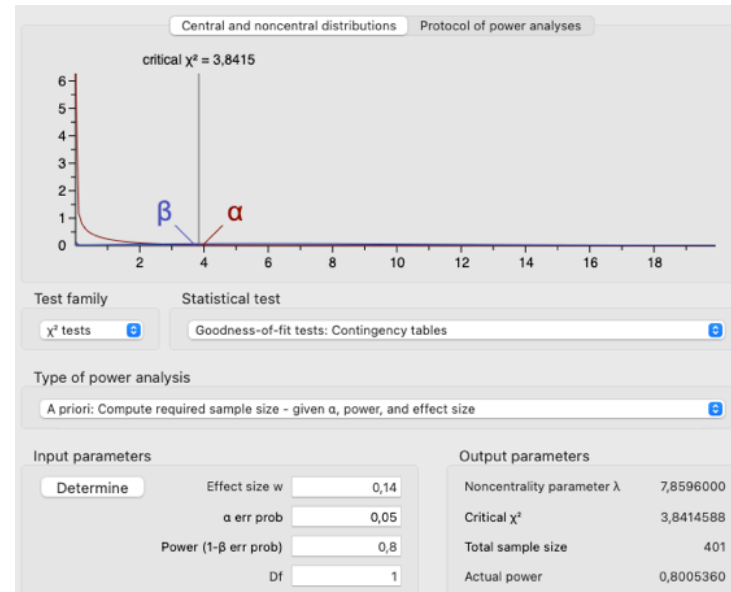
Effektstärke	Interpretation
0.01	kleiner Effekt
0.09	mittlerer Effekt
0.25	großer Effekt

→ nach Cohen handelt es sich in unserem Beispiel um eine kleine Effektstärke.

# Verfahren für Nominaldaten

## Eindimensionaler $\chi^2$ -Test

### Stichprobenumfangsplanung:



# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

- Erweiterung um eine weitere kategoriale Variable ( $\geq 2$  Stufen)
- Darstellung in Kreuztabelle (auch  $k \times l - \chi^2$ -Test)
  - Zeilen: Merkmal 1 mit  $k$  Stufen
  - Spalten: Merkmal 2 mit  $l$  Stufen)
- Wie zuvor: Vergleich theoretisch erwartete vs. beobachtete Häufigkeiten
- Klassische Anwendung: **Kontingenzanalyse**
  - Frage: Besteht ein stochastischer Zusammenhang zwischen den Merkmalen
  - z.B. "Sind Behandlungserfolge [nein vs. ja] zwischen Therapiemodalitäten [Therapie vs. Warteliste] gleich verteilt?"

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Hypothesenpaar $H_0$ vs. $H_1$ :

- Theoretisch unendlich viele  $H_0$  (alle denkbaren theoretischen Annahmen)
- Bei Kontingenzanalyse:
  - $H_0$  : Merkmale sind stochastisch unabhängig.
  - $H_1$  : Es besteht irgendeine Art von Zusammenhang zwischen den Stufen des einen Merkmals und den Stufen des anderen Merkmals



# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Berechnung erwarteter Häufigkeiten unter Annahme der $H_0$ :

- Zur Berechnung müssen wir wissen, wie sich jedes Merkmal alleine verteilen würde (Ignorieren des anderen Merkmals)
- Dies schätzen wir über die sogenannten **Randhäufigkeiten**
- relative Randhäufigkeiten (in %) dienen als Schätzer für die Wahrscheinlichkeit einer Merkmalsstufe in der Population ( $p_i$  bzw.  $p_j$ )

$$p_i = \frac{n_i}{N}, \text{ bzw. } p_j = \frac{n_j}{N}$$

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

**Berechnung erwarteter Häufigkeiten unter Annahme der  $H_0$  :**

Berechnung der Randhäufigkeiten:

```
addmargins(table(data$Behandlungserfolg, data$Behandlung), FUN = sum)
```

```
## Margins computed over dimensions
## in the following order:
## 1:
## 2:
```

```
##
##      Psychotherapie Warteliste sum
##   ja              4            2  6
##   nein             1            3  4
##   sum              5            5 10
```

**Beispiel - Therapieerfolg ( $N = 10$ ) :**

Behandlung	Behandlungserfolg
Psychotherapie	ja
Psychotherapie	ja
Psychotherapie	nein
Psychotherapie	ja
Psychotherapie	ja
Warteliste	nein
Warteliste	nein
Warteliste	nein
Warteliste	ja
Warteliste	ja

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

Berechnung erwarteter Häufigkeiten unter Annahme der  $H_0$  :

Berechnung der Randhäufigkeiten:

```
addmargins(table(data$Behandlungserfolg, data$Behandlung), FUN = sum)
```

```
## Margins computed over dimensions
## in the following order:
## 1:
## 2:
```

```
##
##      Psychotherapie Warteliste sum
##   ja           36           13  49
##   nein          14           37  51
##   sum           50           50 100
```

Beispiel - Therapieerfolg in größerer Stichprobe ( $N = 100$ ) :

- 45 Patient:innen hatten einen Therapieerfolg:  
 $p_{ja} = \frac{49}{100} = 0.49$
- Gegenwahrscheinlichkeit:  $p_{nein} = 1 - 0.49 = 0.51$
- 50 Patient:innen hatten einen Wartelistentherapie:  
 $p_{Warteliste} = \frac{50}{100} = 0.50$
- Gegenwahrscheinlichkeit:  $p_{Psychotherapie} = 1 - 0.50 = 0.50$

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

- Was bedeutet Stochastische Unabhängigkeit?
  - Merkmale beeinflussen einander nicht
  - Ranghäufigkeiten müssen sich in jeder einzelnen Stufe des Merkmals widerspiegeln
  - Ändert sich Verhältnis ist  $H_0$  verletzt

### Berechnung erwarteter Häufigkeiten unter Annahme der $H_0$ :

$$f_{eij} = p_i \cdot n_j \text{ bzw. } f_{eij} = p_j \cdot n_i$$

mit:

- $f_{eij}$  : erwartete Häufigkeit in Zelle  $ij$  der Kreuztabelle
- $p_i$  Randwahrscheinlichkeit der Merkmalsausprägung  $i$  von Merkmal 1
- $p_j$  Randwahrscheinlichkeit der Merkmalsausprägung  $j$  von Merkmal 2
- $n_i$  Randhäufigkeit  $i$  von Merkmal 1
- $n_j$  Randhäufigkeit  $j$  von Merkmal 2

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

Berechnung erwarteter Häufigkeiten unter Annahme der  $H_0$  :

$$f_{eij} = p_i \cdot n_j \text{ bzw. } f_{eij} = p_j \cdot n_i$$

Beide Formeln führen zum selben Ergebnis (lassen sich ineinander überführen):

$$f_{eij} = \frac{n_i \cdot n_j}{N} \rightarrow \frac{\text{Zeilensumme} \cdot \text{Spaltensumme}}{N}$$

```
## Margins computed over dimensions
## in the following order:
## 1:
## 2:
```

```
##
##      Psychotherapie Warteliste sum
##   ja           36           13  49
##  nein           14           37  51
##   sum           50           50 100
```

- Verhältnis von Therapieerfolg ist 49% zu 51%
- Falls Merkmale stochastisch unabhängig, muss dieses Verhältnis sich in beiden Behandlungen zeigen

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

Berechnung erwarteter Häufigkeiten unter Annahme der  $H_0$  :

$$f_{eij} = p_i \cdot n_j \text{ bzw. } f_{eij} = p_j \cdot n_i$$

Beide Formeln führen zum selben Ergebnis (lassen sich ineinander überführen):

$$f_{eij} = \frac{n_i \cdot n_j}{N} \rightarrow \frac{\text{Zeilensumme} \cdot \text{Spaltensumme}}{N}$$

```
## Margins computed over dimensions
## in the following order:
## 1:
## 2:
```

```
##
##      Psychotherapie Warteliste sum
##   ja              36          13  49
##   nein             14          37  51
##   sum              50          50 100
```

- $f_{eTh/ja} = 50 \cdot 0.49 = 24.5$
- $f_{eTh/nein} = 50 \cdot 0.51 = 25.5$
- $f_{eWa/ja} = 50 \cdot 0.49 = 24.5$
- $f_{eWa/nein} = 50 \cdot 0.51 = 25.5$

→ Beobachtete Werte weichen scheinbar von erwarteten ab. Ist diese Abweichung signifikant?

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

Berechnung von  $\chi^2_{emp}$  :

$$\chi^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{(f_{bij} - f_{eij})^2}{f_{eij}}$$

mit:

- $k$  : Anzahl der Kategorien von Merkmal 1 (Index  $i$ )
- $l$  : Anzahl der Kategorien von Merkmal 2 (Index  $j$ )
- $f_{bij}$  beobachtete absolute Häufigkeit von Merkmalskombination  $ij$
- $f_{eij}$  unter  $H_0$  erwartete absolute Häufigkeit von Merkmalskombination  $ij$

und:

$$df = (k - 1) \cdot (l - 1)$$

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Signifikanzprüfung:

```
## Margins computed over dimensions
## in the following order:
## 1:
## 2:
```

```
##
##      Psychotherapie Warteliste sum
##   ja              36          13  49
##   nein             14          37  51
##   sum              50          50 100
```

- $f_{eTh/ja} = 50 \cdot 0.49 = 24.5$
- $f_{eTh/nein} = 50 \cdot 0.51 = 25.5$
- $f_{eWa/ja} = 50 \cdot 0.49 = 24.5$
- $f_{eWa/nein} = 50 \cdot 0.51 = 25.5$

$$\chi_{emp}^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{(f_{bij} - f_{eij})^2}{f_{eij}} = \frac{(36 - 24.5)^2}{24.5} + \frac{(13 - 25.5)^2}{25.5} + \frac{(14 - 24.5)^2}{24.5} + \frac{(37 - 25.5)^2}{25.5} = 21.17$$

mit:

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



# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Signifikanzprüfung:

- $\chi^2_{emp} = 21.17$
- $df = 1$

Ablesen von  $\chi^2_{krit}$  aus der Tabelle:

- $\chi^2_{df=1} = 3.84$

Vergleich  $\chi^2_{emp}$  vs.  $\chi^2_{krit}$  :

- $21.17 > 3.84 \rightarrow$  Test ist signifikant

Fläche df	0,750	0,900	0,950
1	1,32330	2,70554	3,84146
2	2,77259	4,60517	5,99147
3	4,10835	6,25139	7,81473
4	5,38527	7,77944	9,48773
5	6,62568	9,23635	11,0705
6	7,84080	10,6446	12,5916
7	9,03715	12,0170	14,0671
8	10,2188	13,3616	15,5073
9	11,3887	14,6837	16,9190
10	12,5489	15,9871	18,3070
11	13,7007	17,2750	19,6751
12	14,8454	18,5494	21,0261
13	15,9839	19,8119	22,3621
14	17,1170	21,0642	23,6848
15	18,2451	22,3072	24,9958
16	19,3688	23,5418	26,2962
17	20,4887	24,7690	27,5871
18	21,6049	25,9894	28,8693
19	22,7178	27,2036	30,1435

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test - R

Unser Ergebnis (händisch):

```
chisq.test(table(data$Behandlungserfolg, data$Behandlung), correct = FALSE)
```

```
##  
##      Pearson's Chi-squared test  
##  
## data:  table(data$Behandlungserfolg, data$Behandlung)  
## X-squared = 21.168, df = 1, p-value = 0.000004206
```

R führt standardmäßig für 2x2 Kreuztabellen die so genannte Yates-Kontinuitätskorrektur durch (Ergebnis etwas genauer):

```
chisq.test(table(data$Behandlungserfolg, data$Behandlung))
```

```
##  
##      Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  table(data$Behandlungserfolg, data$Behandlung)  
## X-squared = 19.368, df = 1, p-value = 0.00001078
```

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Effekstärke $w^2$ :

- Auch hier lässt sich  $w^2$  verwenden
- Schätzung von  $w^2$  aus Stichprobendaten:

$$\hat{w}^2 = \frac{\chi^2}{N}$$

- Beispiel:  $\hat{w}^2 = \frac{21.17}{100} = 0.21 \rightarrow$  nach Cohen ein mittlerer Effekt.

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Effekstärke Cramers Phi-Koeffizient (Cramers Index, CI)

- Empirisches Effektstärkemaß, baut auf  $w^2$  auf
- Vorteil: Darf direkt wie Korrelationsmaß interpretiert werden
  - Wertebereich zwischen 0 und 1
  - 0 = stochastische Unabhängigkeit
  - 1 = perfekter Zusammenhang

$$CI = \sqrt{\frac{\chi^2}{N \cdot (R - 1)}}$$

mit:

- $R = \min(k; l)$

# Verfahren für Nominaldaten

## Zweidimensionaler $\chi^2$ -Test

### Effekstärke Cramers Phi-Koeffizient (Cramers Index, CI)

$$CI = \sqrt{\frac{\chi^2}{N \cdot (R - 1)}} = \sqrt{\frac{21.17}{100 \cdot (2 - 1)}} = 0.46$$

Konventionen nach Cohen (1988):

Effekstärke	Interpretation
0.1	kleiner Effekt
0.3	mittlerer Effekt
0.5	großer Effekt

→ nach Cohen handelt es sich in unserem Beispiel um eine mittlere Effekstärke.

# Verfahren für Nominaldaten

## Der Vierfelder $\chi^2$ -Test

- Spezialfall des zweidimensionalen  $\chi^2$ -Tests
- Beide Merkmale haben genau 2 Merkmalsstufen (2x2 Kontingenztabelle)

Dann kann Formel vereinfacht werden (Vierfelder-Tafel):

	B1	B2
A1	a	b
A2	c	d

$$\chi^2 = \frac{N \cdot (a \cdot d - b \cdot c)^2}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}$$

mit:

$$df = (k - 1) \cdot (l - 1)$$

- Vorteil: Keine Berechnung von Randhäufigkeiten etc. notwendig

# Verfahren für Nominaldaten

## Der Vierfelder $\chi^2$ -Test

	B1	B2
A1	a	b
A2	c	d

```
table(data$Behandlungserfolg, data$Behandlung)
```

```
##
##      Psychotherapie Warteliste
##   ja           36           13
##   nein          14           37
```

$$\chi_{emp}^2 = \frac{N \cdot (a \cdot d - b \cdot c)^2}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)} = \frac{100 \cdot (36 \cdot 37 - 13 \cdot 14)^2}{(36 + 13) \cdot (14 + 37) \cdot (36 + 14) \cdot (13 + 37)} = 21.17$$

# Verfahren für Nominaldaten

## Der Vierfelder $\chi^2$ -Test

### Effektstärke Phi-Koeffizient

- Effektstärke  $\phi$  entspricht Korrelation von 2 dichotomen Variablen

$$\phi = \frac{a \cdot d - b \cdot c}{\sqrt{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}}$$



# Verfahren für Nominaldaten

## Voraussetzungen $\chi^2$ -Tests

- $\chi^2$ -Tests haben nur relativ wenige Voraussetzungen
  1. Einzelbeobachtungen sind unabhängig voneinander
  2. Jede Person kann eindeutig einer Kategorie (oder Kombination von Kategorien) zugeordnet werden
  3. Erwartete Häufigkeiten in den Zellen größer als 5 (sonst analoger Alternativtest "Exakter Test nach Fisher")

## Take-aways

- $\chi^2$ -Tests beinhalten die **Analyse von Häufigkeiten**.
- Prinzip: Vergleich von **beobachteten vs. theoretisch erwarteten** Häufigkeiten.
- Für  $\chi^2$ -Tests existieren theoretisch **unendlich viele  $H_0$**  Möglichkeiten (häufig: Gleichverteilungsannahme)
- **Eindimensionaler  $\chi^2$ -Test** prüft, ob sich Verteilung der Kategorien einer nominalskalierten Variable unterscheiden.
- **Zweidimensionaler  $\chi^2$ -Test** prüft Verteilung von 2 nominalskalierten Variablen (stochastische Unabhängigkeit).
- Spezialfall: **Vierfelder  $\chi^2$ -Test** wenn 2 dichotome Merkmale vorliegen.
- **Effektstärkebestimmung** über  $w^2$  oder Phi-Koeffizient (wie Korrelation interpretierbar)

