

Biostatistik, WS 2013/2014

Chi-Quadrat-Test

Matthias Birkner

<http://www.mathematik.uni-mainz.de/~birkner/Biostatistik1314/>

10.1.2014



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Inhalt

- 1 Wiederholung: Wilcoxon's Rangsummentest
 - Motivation
 - Wilcoxon-Test für unabhängige Stichproben
- 2 Kontingenztafeln
 - χ^2 -Anpassungstest für eine vorgegebene Verteilung
 - χ^2 -Test auf Homogenität bzw. Unabhängigkeit
 - Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Inhalt

1 Wiederholung: Wilcoxon's Rangsummentest

- Motivation
- Wilcoxon-Test für unabhängige Stichproben

2 Kontingenztafeln

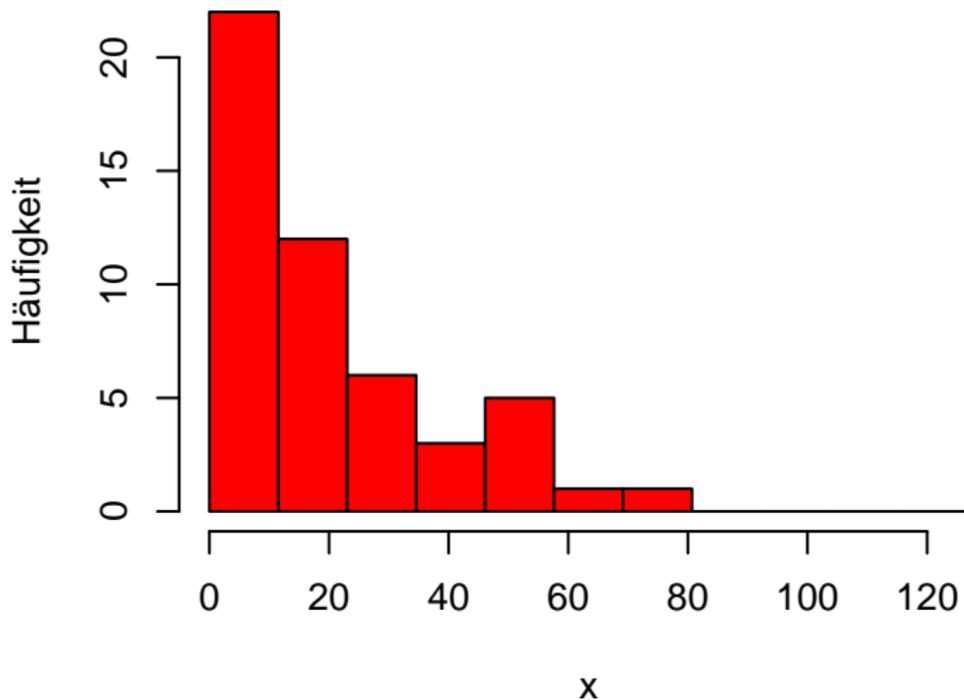
- χ^2 -Anpassungstest für eine vorgegebene Verteilung
- χ^2 -Test auf Homogenität bzw. Unabhängigkeit
- Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Bei (ungefähr) glockenförmigen und symmetrisch verteilten Beobachtungen oder wenn die Stichprobenumfänge genügend groß sind können wir den t -Test benutzen, um die Nullhypothese $\mu_1 = \mu_2$ zu testen: Die t -Statistik ist (annähernd) Student-verteilt.

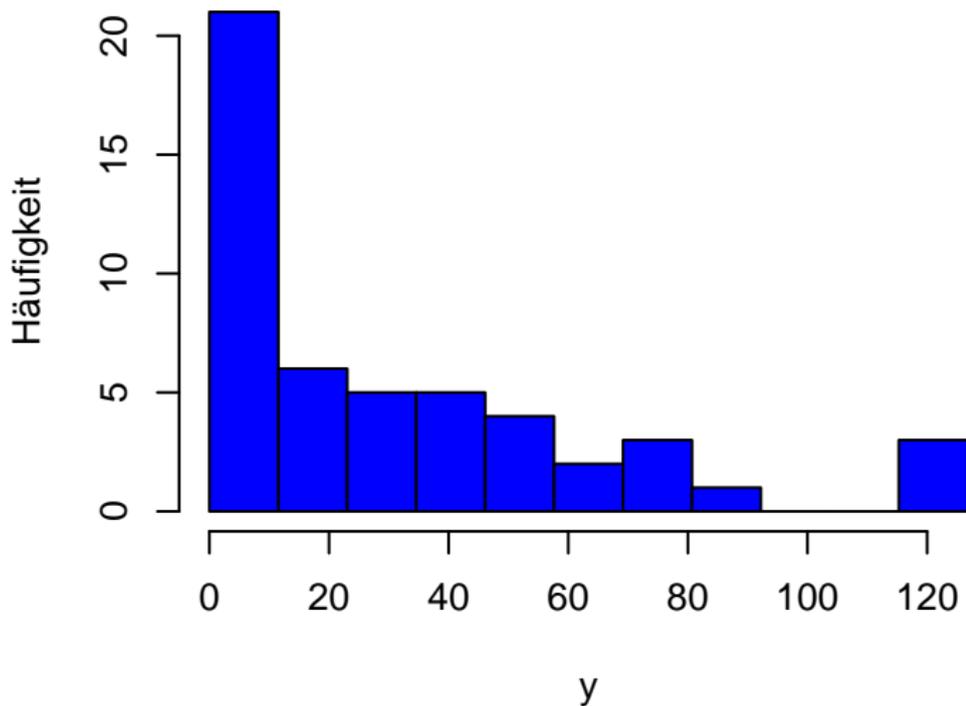
Bei (ungefähr) glockenförmigen und symmetrisch verteilten Beobachtungen oder wenn die Stichprobenumfänge genügend groß sind können wir den t -Test benutzen, um die Nullhypothese $\mu_1 = \mu_2$ zu testen: Die t -Statistik ist (annähernd) Student-verteilt.

Besonders bei sehr asymmetrischen und langschwänzigen Verteilungen kann das anders sein

Nehmen wir an, wir sollten folgende Verteilungen vergleichen:



Nehmen wir an, wir sollten folgende Verteilungen vergleichen:



Beispiele

- Wartezeiten
- Ausbreitungsentfernungen
- Zelltypenhäufigkeiten

Beispiele

- Wartezeiten
- Ausbreitungsentfernungen
- Zelltypenhäufigkeiten

Gesucht:

ein „verteilungsfreier“ Test,
mit dem man die Lage zweier Verteilungen
zueinander testen kann

Inhalt

1 Wiederholung: Wilcoxon's Rangsummentest

- Motivation
- Wilcoxon-Test für unabhängige Stichproben

2 Kontingenztafeln

- χ^2 -Anpassungstest für eine vorgegebene Verteilung
- χ^2 -Test auf Homogenität bzw. Unabhängigkeit
- Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Beobachtungen: Zwei Stichproben

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

Beobachtungen: Zwei Stichproben

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

Wir möchten die **Nullhypothese**:
 X und Y aus derselben Population
(X und Y haben **diesselbe Verteilung**)
testen

Beobachtungen: Zwei Stichproben

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

Wir möchten die **Nullhypothese**:
 X und Y aus derselben Population
(X und Y haben **diesselbe Verteilung**)
testen

gegen die **Alternative**:

Die beiden Verteilungen sind gegeneinander verschoben.

Beobachtungen: Zwei Stichproben

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

Wir möchten die **Nullhypothese**:
 X und Y aus derselben Population
(X und Y haben **diesselbe Verteilung**)
testen

gegen die **Alternative**:

Die beiden Verteilungen sind gegeneinander verschoben.

Wir sind also in einer Situation, die wir schon beim t -Test getroffen haben: Die zwei Verteilungen sind möglicherweise gegeneinander verschoben (haben insbesondere möglicherweise unterschiedliche Mittelwerte), aber wir möchten *nicht* die implizite Annahme treffen, dass es sich dabei (wenigstens ungefähr) um Normalverteilungen handelt.

Idee

Beobachtungen:

$X : x_1, x_2, \dots, x_m$

$Y : y_1, y_2, \dots, y_n$

Idee

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.

Idee

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.

Idee

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.
- Wenn die Nullhypothese zutrifft, sind die m X -Ränge eine rein zufällige Wahl aus $\{1, 2, \dots, m + n\}$.

Idee

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.
- Wenn die Nullhypothese zutrifft, sind die m X -Ränge eine rein zufällige Wahl aus $\{1, 2, \dots, m + n\}$.
- Berechne die Summe der X -Ränge, prüfe, ob dieser Wert untypisch groß oder klein.

Wilcoxon's Rangsummenstatistik

Beobachtungen:

$X : x_1, x_2, \dots, x_m$

$Y : y_1, y_2, \dots, y_n$



Frank Wilcoxon,
1892–1965

$W = \text{Summe der } X\text{-Ränge} - (1 + 2 + \dots + m)$
heißt

Wilcoxon's Rangsummenstatistik

Die Normierung ist so gewählt, dass $0 \leq W \leq m n$.

Wilcoxon's Rangsummenstatistik

Bemerkung 1:

$$W = \text{Summe der } X\text{-Ränge} - (1 + 2 + \dots + m)$$

Wir könnten auch die Summe der Y -Ränge benutzen,

Wilcoxon's Rangsummenstatistik

Bemerkung 1:

$$W = \text{Summe der } X\text{-Ränge} - (1 + 2 + \dots + m)$$

Wir könnten auch die Summe der Y -Ränge benutzen, denn

Summe der X -Ränge + Summe der Y -Ränge

$$= \text{Summe aller Ränge}$$

$$= 1 + 2 + \dots + (m + n) = \frac{(m + n)(m + n + 1)}{2}$$

Wilcoxon's Rangsummenstatistik

Bemerkung 1:

$$W = \text{Summe der } X\text{-Ränge} - (1 + 2 + \dots + m)$$

Wir könnten auch die Summe der Y -Ränge benutzen, denn

Summe der X -Ränge + Summe der Y -Ränge

$$= \text{Summe aller Ränge}$$

$$= 1 + 2 + \dots + (m + n) = \frac{(m + n)(m + n + 1)}{2}$$

Bemerkung 2:

Der Wilcoxon-Test heißt auch Mann-Whitney-Test, die Rangsummenstatistik auch Mann-Whitney Statistik U , sie unterscheidet sich (je nach Definition) von W um eine Konstante.

(In der Literatur sind beide Bezeichnungen üblich, man prüfe vor Verwendung von Tabellen, etc. die verwendete Konvention.)

Ein kleines Beispiel

- Beobachtungen:

X : 1,5; 5,6; 35,2

Y : 7,9; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

Ein kleines Beispiel

- Beobachtungen:

X : 1,5; 5,6; 35,2

Y : 7,9; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

- Lege Beobachtungen zusammen und sortiere:

1,5; 5,6; 7,9; 35,2; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

Ein kleines Beispiel

- Beobachtungen:

X : 1,5; 5,6; 35,2

Y : 7,9; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

- Lege Beobachtungen zusammen und sortiere:

1,5; 5,6; 7,9; 35,2; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

- Bestimme Ränge:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Ein kleines Beispiel

- Beobachtungen:

X : 1,5; 5,6; 35,2

Y : 7,9; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

- Lege Beobachtungen zusammen und sortiere:

1,5; 5,6; 7,9; 35,2; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8

- Bestimme Ränge:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10

- Rangsummenstatistik hier: $W = 1 + 2 + 4 - (1 + 2 + 3) = 1$

Interpretation von W

X -Population kleiner $\implies W$ klein:

Interpretation von W

X -Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

Interpretation von W

X -Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

Interpretation von W

X -Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

Interpretation von W

X -Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W =$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W =$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 20$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 20$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W =$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 20$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 19$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 20$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 19$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W =$

Interpretation von W

X-Population kleiner $\implies W$ klein:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 0$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 1$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 2$

X-Population größer $\implies W$ groß:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 21$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 20$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 19$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10 $W = 19$

Signifikanz

Nullhypothese:
 X -Stichprobe und Y -Stichprobe
stammen aus
derselben Verteilung

Signifikanz

Nullhypothese:
X-Stichprobe und Y-Stichprobe
stammen aus
derselben Verteilung

Die 3 Ränge der X-Stichprobe
1 2 3 4 5 6 7 8 9 10

Signifikanz

Nullhypothese:
 X -Stichprobe und Y -Stichprobe
stammen aus
derselben Verteilung

Die 3 Ränge der X -Stichprobe
1 2 3 4 5 6 7 8 9 10

hätten genausogut irgendwelche 3 Ränge

Signifikanz

Nullhypothese:
 X -Stichprobe und Y -Stichprobe
stammen aus
derselben Verteilung

Die 3 Ränge der X -Stichprobe

1 2 3 4 5 6 7 8 9 10

hätten genausogut irgendwelche 3 Ränge

1 2 3 4 5 6 7 8 9 10

sein können.

Signifikanz

Nullhypothese:
X-Stichprobe und Y-Stichprobe
stammen aus
derselben Verteilung

Die 3 Ränge der X-Stichprobe

1 2 3 4 5 6 7 8 9 10

hätten genausogut irgendwelche 3 Ränge

1 2 3 4 5 6 7 8 9 10

sein können.

Es gibt $\frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$ Möglichkeiten.

Signifikanz

Nullhypothese:
 X-Stichprobe und Y-Stichprobe
 stammen aus
 derselben Verteilung

Die 3 Ränge der X-Stichprobe

1 2 3 4 5 6 7 8 9 10

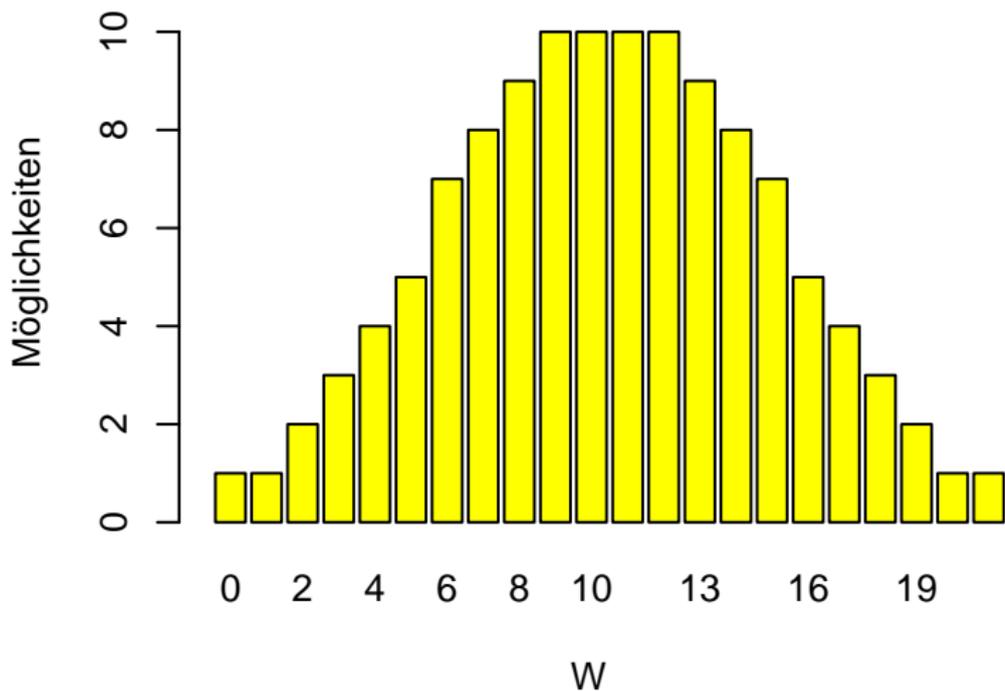
hätten genausogut irgendwelche 3 Ränge

1 2 3 4 5 6 7 8 9 10

sein können.

Es gibt $\frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$ Möglichkeiten.

(Allgemein: $\frac{(m+n)(m+n-1)\dots(n+1)}{m(m-1)\dots 1} = \frac{(m+n)!}{n!m!} = \binom{m+n}{m}$ Möglichkeiten)

Verteilung der Wilcoxon-Statistik ($m = 3, n = 7$)

Unter der Nullhypothese sind alle Rangbelegungen gleich
wahrscheinlich, also

$$\mathbb{P}(W = w) = \frac{\text{Anz. Möglichkeiten mit Rangsummenstatistik } w}{120}$$

Unter der Nullhypothese sind alle Rangbelegungen gleich
wahrscheinlich, also

$$\mathbb{P}(W = w) = \frac{\text{Anz. Möglichkeiten mit Rangsummenstatistik } w}{120}$$

Wir beobachten in unserem Beispiel:

1,5, 5,6; 7,9; 35,2; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8
somit $W = 1$

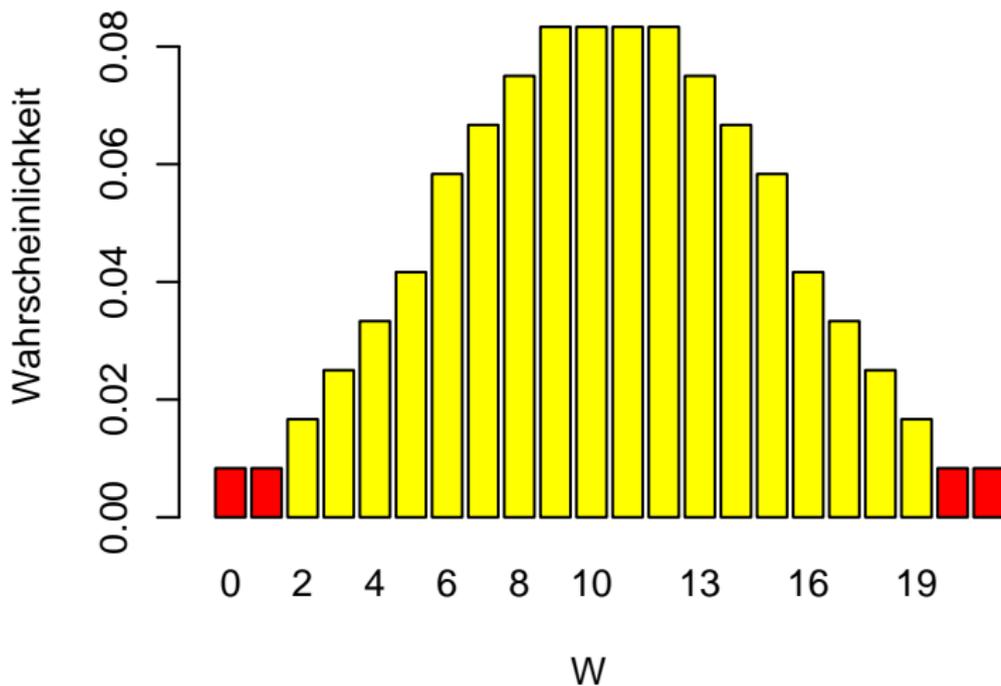
Unter der Nullhypothese sind alle Rangbelegungen gleich
wahrscheinlich, also

$$\mathbb{P}(W = w) = \frac{\text{Anz. Möglichkeiten mit Rangsummenstatistik } w}{120}$$

Wir beobachten in unserem Beispiel:

1,5, 5,6; 7,9; 35,2; 38,1; 41,0; 56,7; 112,1; 197,4; 381,8
somit $W = 1$

$$\begin{aligned} & \mathbb{P}(W \leq 1) + \mathbb{P}(W \geq 20) \\ &= \mathbb{P}(W = 0) + \mathbb{P}(W = 1) + \mathbb{P}(W = 20) + \mathbb{P}(W = 21) \\ &= \frac{1+1+1+1}{120} \doteq 0,033 \end{aligned}$$

Verteilung der Wilcoxon-Statistik ($m = 3, n = 7$)

Für unser Beispiel ($W = 1$) also:

$$p\text{-Wert} = \mathbb{P}(\text{ein so extremes } W) = 4/120 = 0,033$$

Für unser Beispiel ($W = 1$) also:

$$p\text{-Wert} = \mathbb{P}(\text{ein so extremes } W) = 4/120 = 0,033$$

Wir **lehnen** die **Nullhypothese**,
dass die Verteilungen
von X und Y
identisch sind,
auf dem 5%-Niveau **ab**.

Hinweis

Hinweis

Wenn der Wilcoxon-Test Signifikanz anzeigt, so kann das daran liegen, dass die zu grunde liegenden Verteilungen verschiedene Formen haben.

Hinweis

Hinweis

Wenn der Wilcoxon-Test Signifikanz anzeigt, so kann das daran liegen, dass die zu grunde liegenden Verteilungen verschiedene Formen haben.

Der Wilcoxon-Test kann beispielsweise Signifikanz anzeigen, **selbst wenn die Stichproben-Mittelwerte übereinstimmen!**

Vergleich von t -Test und Wilcoxon-Test

Beachte:

Sowohl der t -Test als auch
der Wilcoxon-Test können verwendet werden,
um eine vermutete Verschiebung der Verteilung zu stützen.

Vergleich von t -Test und Wilcoxon-Test

Beachte:

Sowohl der t -Test als auch der Wilcoxon-Test können verwendet werden, um eine vermutete Verschiebung der Verteilung zu stützen.

Der t -Test testet „nur“ auf Gleichheit der Erwartungswerte.

Vergleich von t -Test und Wilcoxon-Test

Beachte:

Sowohl der t -Test als auch der Wilcoxon-Test können verwendet werden, um eine vermutete Verschiebung der Verteilung zu stützen.

Der t -Test testet „nur“ auf Gleichheit der Erwartungswerte.
Der Wilcoxon-Test dagegen testet auf Gleichheit der gesamten Verteilungen.

Vergleich von t -Test und Wilcoxon-Test

Beachte:

Sowohl der t -Test als auch der Wilcoxon-Test können verwendet werden, um eine vermutete Verschiebung der Verteilung zu stützen.

Der t -Test testet „nur“ auf Gleichheit der Erwartungswerte.
Der Wilcoxon-Test dagegen testet auf Gleichheit der gesamten Verteilungen.

In den meisten Fällen liefern beide Tests dasselbe Ergebnis.
Im Allgemeinen ist für Lagetests der t -Test empfehlenswerter.

Vergleich von t -Test und Wilcoxon-Test

Beachte:

Sowohl der t -Test als auch der Wilcoxon-Test können verwendet werden, um eine vermutete Verschiebung der Verteilung zu stützen.

Der t -Test testet „nur“ auf Gleichheit der Erwartungswerte.
Der Wilcoxon-Test dagegen testet auf Gleichheit der gesamten Verteilungen.

In den meisten Fällen liefern beide Tests dasselbe Ergebnis.
Im Allgemeinen ist für Lagetests der t -Test empfehlenswerter.

In besonderen Fällen

- Verteilungen sind asymmetrisch
- Stichprobenlänge ist klein

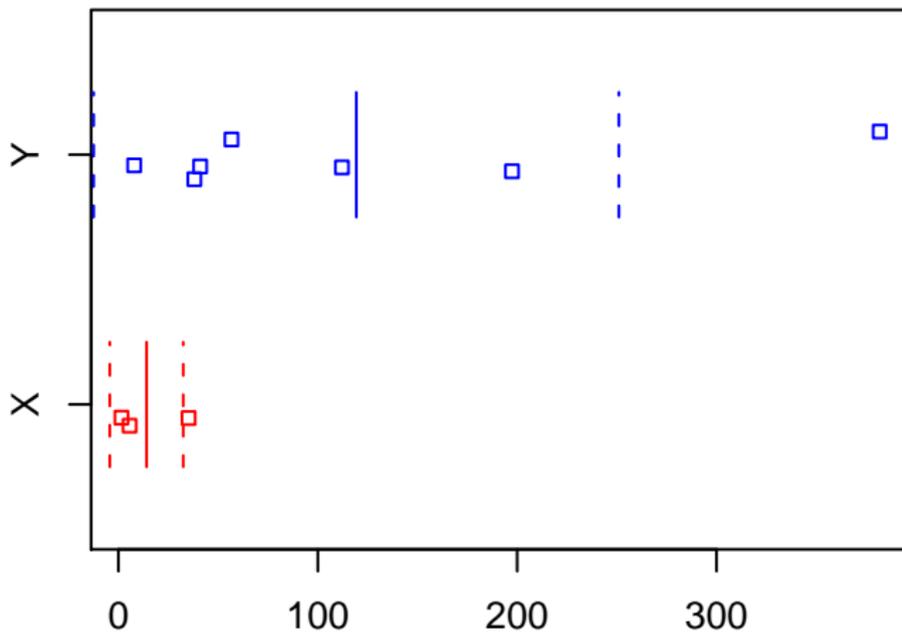
hat der Wilcoxon-Test eine höhere Testpower.

Vergleichen wir (spaßeshalber) mit dem t -Test (hier mit dem Statistikprogramm **R**, <http://www.r-project.org> ausgeführt):

```
> x
[1] 1.5 5.6 35.2
> y
[1] 7.9 38.1 41.0 56.7 112.1 197.4 381.8
> t.test(x,y,var.equal=TRUE)
```

Two Sample t-test

```
data: x and y
t = -1.3319, df = 8, p-value = 0.2196
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -287.30632 76.93489
sample estimates:
mean of x mean of y
 14.1000 119.2857
```



Inhalt

- 1 Wiederholung: Wilcoxon's Rangsummentest
 - Motivation
 - Wilcoxon-Test für unabhängige Stichproben
- 2 Kontingenztafeln
 - χ^2 -Anpassungstest für eine vorgegebene Verteilung
 - χ^2 -Test auf Homogenität bzw. Unabhängigkeit
 - Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Inhalt

- 1 Wiederholung: Wilcoxon's Rangsummentest
 - Motivation
 - Wilcoxon-Test für unabhängige Stichproben
- 2 Kontingenztafeln
 - χ^2 -Anpassungstest für eine vorgegebene Verteilung
 - χ^2 -Test auf Homogenität bzw. Unabhängigkeit
 - Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Mendels Erbsenexperiment

Betrachte zwei Merkmale:

Farbe: grün (rezessiv) vs. gelb (dominant)

Form: rund (dominant) vs. runzlig (rezessiv)

Mendels Erbsenexperiment

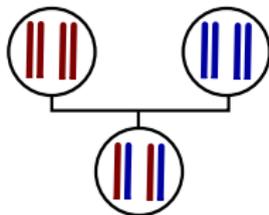
Betrachte zwei Merkmale:

Farbe: grün (rezessiv) vs. gelb (dominant)

Form: rund (dominant) vs. runzlig (rezessiv)

Beim Kreuzen von Doppelhybriden erwarten wir folgende Phänotypwahrscheinlichkeiten unter Mendelscher Segregation:

	grün	gelb
runzlig	$\frac{1}{16}$	$\frac{3}{16}$
rund	$\frac{3}{16}$	$\frac{9}{16}$



Mendels Erbsenexperiment

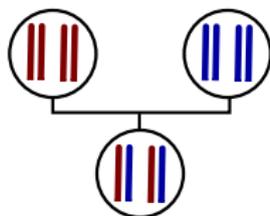
Betrachte zwei Merkmale:

Farbe: grün (rezessiv) vs. gelb (dominant)

Form: rund (dominant) vs. runzlig (rezessiv)

Beim Kreuzen von Doppelhybriden erwarten wir folgende Phänotypwahrscheinlichkeiten unter Mendelscher Segregation:

	grün	gelb
runzlig	$\frac{1}{16}$	$\frac{3}{16}$
rund	$\frac{3}{16}$	$\frac{9}{16}$



Im Experiment beobachtet ($n = 556$ Versuche):

	grün	gelb
runzlig	32	101
rund	108	315

Mendels Erbsenexperiment

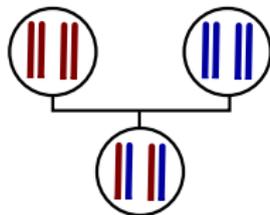
Betrachte zwei Merkmale:

Farbe: grün (rezessiv) vs. gelb (dominant)

Form: rund (dominant) vs. runzlig (rezessiv)

Beim Kreuzen von Doppelhybriden erwarten wir folgende Phänotypwahrscheinlichkeiten unter Mendelscher Segregation:

	grün	gelb
runzlig	$\frac{1}{16}$	$\frac{3}{16}$
rund	$\frac{3}{16}$	$\frac{9}{16}$



Im Experiment beobachtet ($n = 556$ Versuche):

	grün	gelb
runzlig	32	101
rund	108	315

Frage:

Passen die Beobachtungen zu den theoretischen Erwartungen?

Relative Häufigkeiten:

	grün/runzlig	gelb/runzlig	grün/rund	gelb/rund
erwartet	0,0625	0,1875	0,1875	0,5625
beobachtet	0,0576	0,1942	0,1816	0,5665

Relative Häufigkeiten:

	grün/runzlig	gelb/runzlig	grün/rund	gelb/rund
erwartet	0,0625	0,1875	0,1875	0,5625
beobachtet	0,0576	0,1942	0,1816	0,5665

bzw. in absoluten Häufigkeiten ($n = 556$):

	grün/runzlig	gelb/runzlig	grün/rund	gelb/rund
erwartet	34,75	104,25	104,25	312,75
beobachtet	32	101	108	315

Relative Häufigkeiten:

	grün/runzlig	gelb/runzlig	grün/rund	gelb/rund
erwartet	0,0625	0,1875	0,1875	0,5625
beobachtet	0,0576	0,1942	0,1816	0,5665

bzw. in absoluten Häufigkeiten ($n = 556$):

	grün/runzlig	gelb/runzlig	grün/rund	gelb/rund
erwartet	34,75	104,25	104,25	312,75
beobachtet	32	101	108	315

Können diese Abweichungen plausibel durch
Zufallsschwankungen erklärt werden?

Wir messen die Abweichungen durch die χ^2 -Statistik:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

wobei E_i = erwartete Anzahl in Klasse i und O_i = beobachtete (engl. *observed*) Anzahl in Klasse i .

(im Beispiel durchläuft i die vier möglichen Klassen grün/runzlig, gelb/runzlig, grün/rund, gelb/rund.)

Wieso teilen wir dabei $(O_i - E_i)^2$ durch $E_i = \mathbb{E}O_i$?

Wieso teilen wir dabei $(O_i - E_i)^2$ durch $E_i = \mathbb{E}O_i$?

Sei n die Gesamtzahl und p_i die Wahrscheinlichkeit (unter der Nullhypothese) jeder Beobachtung, zu O_i beizutragen.

Wieso teilen wir dabei $(O_i - E_i)^2$ durch $E_i = \mathbb{E}O_i$?

Sei n die Gesamtzahl und p_i die Wahrscheinlichkeit (unter der Nullhypothese) jeder Beobachtung, zu O_i beizutragen.

Unter der Nullhypothese ist O_i binomialverteilt:

$$\mathbb{P}(O_i = k) = \binom{n}{k} p_i^k (1 - p_i)^{n-k}.$$

Also

$$\mathbb{E} O_i = n p_i, \quad \mathbb{E}(O_i - E_i)^2 = \text{Var}(O_i) = n p_i (1 - p_i)$$

Wieso teilen wir dabei $(O_i - E_i)^2$ durch $E_i = \mathbb{E}O_i$?

Sei n die Gesamtzahl und p_i die Wahrscheinlichkeit (unter der Nullhypothese) jeder Beobachtung, zu O_i beizutragen.

Unter der Nullhypothese ist O_i binomialverteilt:

$$\mathbb{P}(O_i = k) = \binom{n}{k} p_i^k (1 - p_i)^{n-k}.$$

Also

$$\mathbb{E} O_i = n p_i, \quad \mathbb{E}(O_i - E_i)^2 = \text{Var}(O_i) = n p_i (1 - p_i)$$

und

$$\mathbb{E} \left[\frac{(O_i - E_i)^2}{E_i} \right] = \frac{\text{Var}(O_i)}{\mathbb{E} O_i} = 1 - p_i$$

(was gar nicht von n abhängt).

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556
beob. (O)	32	101	108	315	556

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556
beob. (O)	32	101	108	315	556
$O - E$	-2.75	-3.25	3.75	2.25	

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556
beob. (O)	32	101	108	315	556
$O - E$	-2.75	-3.25	3.75	2.25	
$(O - E)^2$	7.56	10.56	14.06	5.06	

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556
beob. (O)	32	101	108	315	556
$O - E$	-2.75	-3.25	3.75	2.25	
$(O - E)^2$	7.56	10.56	14.06	5.06	
$\frac{(O-E)^2}{E}$	0.22	0.10	0.13	0.02	0.47

$$\chi^2 = 0.47$$

Für das Erbsenbeispiel finden wir:

	gr/runz	ge/runz	gr/rund	ge/rund	Summe
theor. Ant.	0.0625	0.1875	0.1875	0.5625	
erw. (E)	34.75	104.25	104.25	312.75	556
beob. (O)	32	101	108	315	556
$O - E$	-2.75	-3.25	3.75	2.25	
$(O - E)^2$	7.56	10.56	14.06	5.06	
$\frac{(O-E)^2}{E}$	0.22	0.10	0.13	0.02	0.47

$$\chi^2 = 0.47$$

Ist ein Wert von $\chi^2 = 0.47$ ungewöhnlich?

Die (asymptotische) Verteilung von χ^2 hängt ab von der sog. Anzahl der Freiheitsgrade **df** (eng. *degrees of freedom*), anschaulich gesprochen die Anzahl der Dimensionen, in denen man von der Erwartung abweichen kann.

Die (asymptotische) Verteilung von χ^2 hängt ab von der sog. Anzahl der Freiheitsgrade **df** (eng. *degrees of freedom*), anschaulich gesprochen die Anzahl der Dimensionen, in denen man von der Erwartung abweichen kann.

In diesem Fall: Es gibt vier Klassen, die Summe der Beobachtungen muss die Gesamtzahl $n = 556$ ergeben.

Die (asymptotische) Verteilung von χ^2 hängt ab von der sog. Anzahl der Freiheitsgrade **df** (eng. *degrees of freedom*), anschaulich gesprochen die Anzahl der Dimensionen, in denen man von der Erwartung abweichen kann.

In diesem Fall: Es gibt vier Klassen, die Summe der Beobachtungen muss die Gesamtzahl $n = 556$ ergeben.

↪ wenn die ersten Zahlen 32, 101, 108 gegeben sind, ist die letzte bestimmt durch

$$315 = 556 - 32 - 101 - 108.$$

Die (asymptotische) Verteilung von χ^2 hängt ab von der sog. Anzahl der Freiheitsgrade **df** (eng. *degrees of freedom*), anschaulich gesprochen die Anzahl der Dimensionen, in denen man von der Erwartung abweichen kann.

In diesem Fall: Es gibt vier Klassen, die Summe der Beobachtungen muss die Gesamtzahl $n = 556$ ergeben.

↪ wenn die ersten Zahlen 32, 101, 108 gegeben sind, ist die letzte bestimmt durch

$$315 = 556 - 32 - 101 - 108.$$

$$\Rightarrow \text{df} = 3$$

Die (asymptotische) Verteilung von χ^2 hängt ab von der sog. Anzahl der Freiheitsgrade **df** (eng. *degrees of freedom*), anschaulich gesprochen die Anzahl der Dimensionen, in denen man von der Erwartung abweichen kann.

In diesem Fall: Es gibt vier Klassen, die Summe der Beobachtungen muss die Gesamtzahl $n = 556$ ergeben.

↪ wenn die ersten Zahlen 32, 101, 108 gegeben sind, ist die letzte bestimmt durch

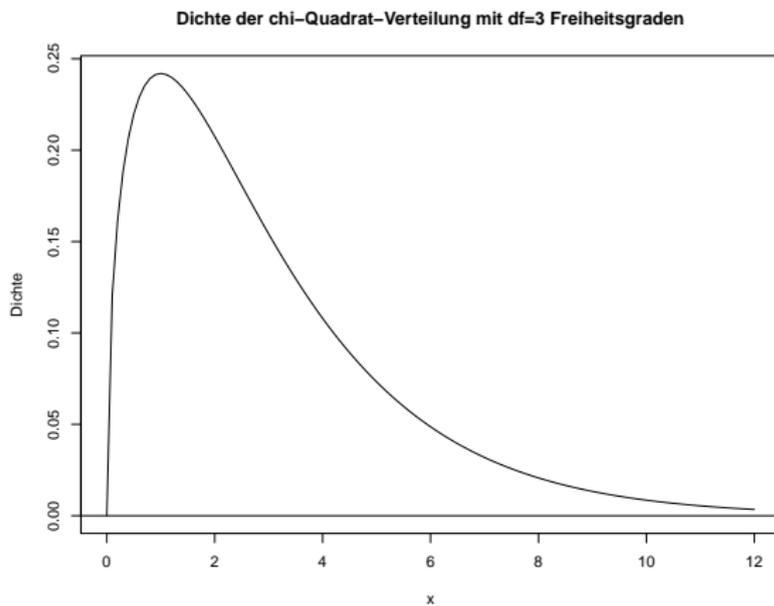
$$315 = 556 - 32 - 101 - 108.$$

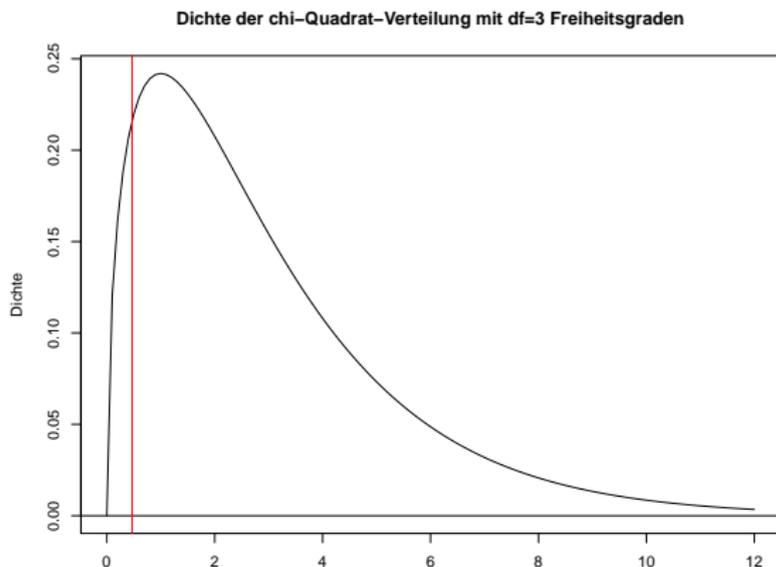
$$\Rightarrow \text{df} = 3$$

Merkregel

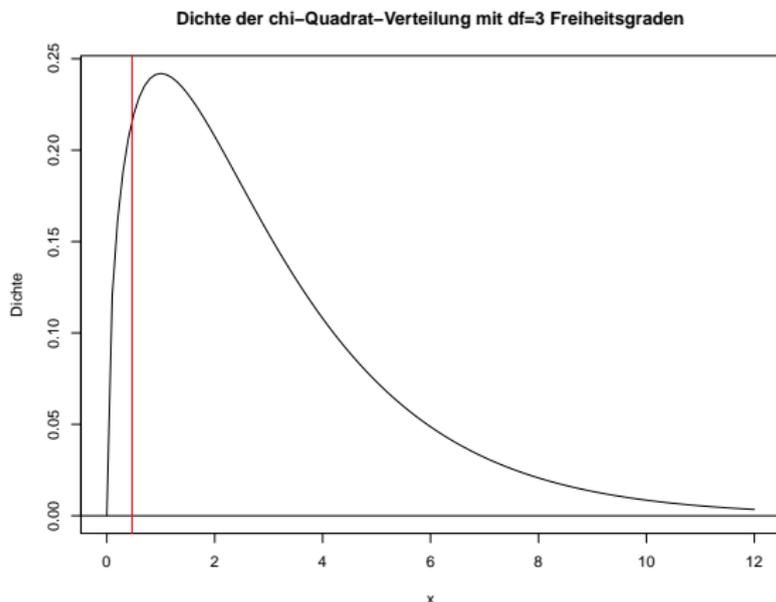
Allgemein gilt beim Chi-Quadrat-Anpassungstest mit k Klassen (wenn das Modell voll spezifiziert ist, d.h. keine Parameter geschätzt werden)

$$\text{df} = k - 1.$$





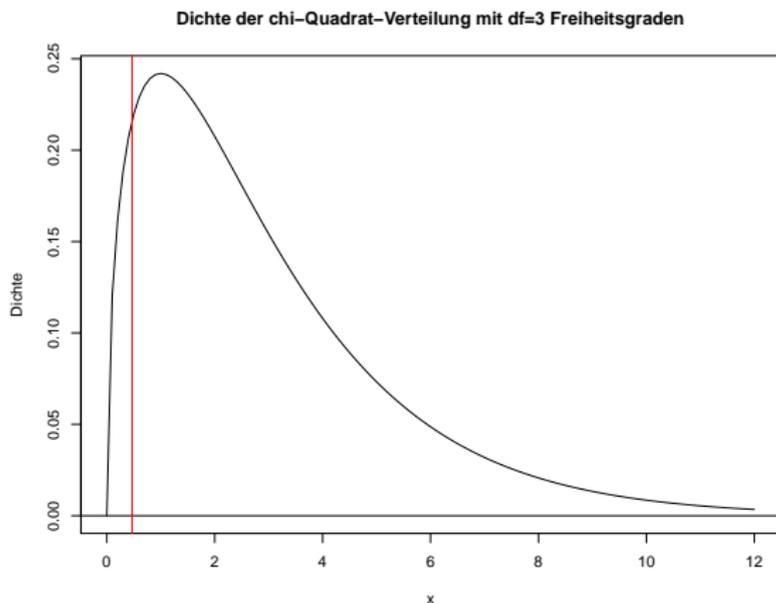
Wir hatten im Erbsenbeispiel gesehen: $\chi^2 = 0.47$ mit $df=3$ Freiheitsgraden.



Wir hatten im Erbsenbeispiel gesehen: $\chi^2 = 0.47$ mit $df=3$ Freiheitsgraden.

Für eine χ^2 mit 3 Freiheitsgraden-verteilte ZV X (man schreibt oft auch χ_3^2 -verteilt) gilt

$$\mathbb{P}(X \leq 0.47) \doteq 0,075$$



Wir hatten im Erbsenbeispiel gesehen: $\chi^2 = 0.47$ mit $df=3$ Freiheitsgraden.

Für eine χ^2 mit 3 Freiheitsgraden-verteilte ZV X (man schreibt oft auch χ_3^2 -verteilt) gilt

$\mathbb{P}(X \leq 0.47) \doteq 0,075$ (und somit ist der p -Wert $\mathbb{P}(X \geq 0.47) = 0.93$), demnach zeigt der χ^2 -Test keine signifikante Abweichung.

Inhalt

- 1 Wiederholung: Wilcoxon's Rangsummentest
 - Motivation
 - Wilcoxon-Test für unabhängige Stichproben
- 2 Kontingenztafeln
 - χ^2 -Anpassungstest für eine vorgegebene Verteilung
 - χ^2 -Test auf Homogenität bzw. Unabhängigkeit
 - Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Der Kuhstärling ist ein Brutparasit des Oropendola.

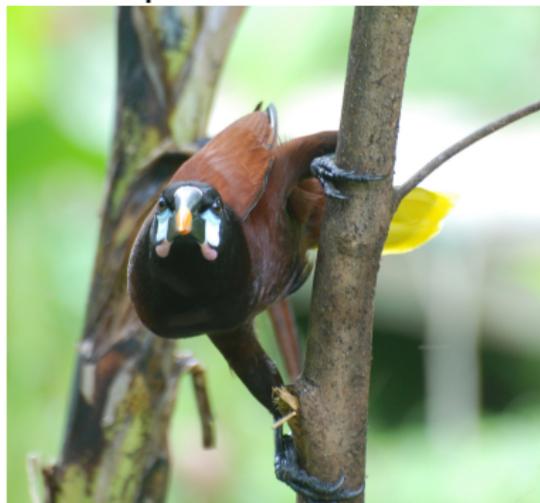


photo (c) by J. Oldenettel



N.G. Smith (1968) The advantage of being parasitized.
Nature, **219(5155)**:690-4

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.
- Normalerweise entfernen Oropendolas alles aus ihrem Nest, was nicht genau nach ihren Eiern aussieht.

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.
- Normalerweise entfernen Oropendolas alles aus ihrem Nest, was nicht genau nach ihren Eiern aussieht.
- In einigen Gegenden sind Kuhstärling-Eier gut von Oropendola-Eiern zu unterscheiden und werden trotzdem nicht aus den Nestern entfernt.

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.
- Normalerweise entfernen Oropendolas alles aus ihrem Nest, was nicht genau nach ihren Eiern aussieht.
- In einigen Gegenden sind Kuhstärling-Eier gut von Oropendola-Eiern zu unterscheiden und werden trotzdem nicht aus den Nestern entfernt.
- Wieso?

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.
- Normalerweise entfernen Oropendolas alles aus ihrem Nest, was nicht genau nach ihren Eiern aussieht.
- In einigen Gegenden sind Kuhstärling-Eier gut von Oropendola-Eiern zu unterscheiden und werden trotzdem nicht aus den Nestern entfernt.
- Wieso?
- Mögliche Erklärung: Junge Oropendolas sterben häufig am Befall durch Dasselfliegenlarven.

- Kuhstärling-Eier sehen Oropendola-Eiern meist sehr ähnlich.
- Normalerweise entfernen Oropendolas alles aus ihrem Nest, was nicht genau nach ihren Eiern aussieht.
- In einigen Gegenden sind Kuhstärling-Eier gut von Oropendola-Eiern zu unterscheiden und werden trotzdem nicht aus den Nestern entfernt.
- Wieso?
- Mögliche Erklärung: Junge Oropendolas sterben häufig am Befall durch Dasselfliegenlarven.
- Nester mit Kuhstärling-Eier sind möglicherweise besser vor Dasselfliegenlarven geschützt.

Anzahlen von Nestern, die von Dasselfliegenlarven befallen sind

Anzahl Kuhstärling-Eier	0	1	2
befallen	16	2	1
nicht befallen	2	11	16

Anzahlen von Nestern, die von Dasselfliegenlarven befallen sind

Anzahl Kuhstärling-Eier	0	1	2
befallen	16	2	1
nicht befallen	2	11	16

		Anzahl Kuhstärling-Eier	0	1	2
In Prozent:	befallen		89%	15%	6%
	nicht befallen		11%	85%	94%

- Anscheinend ist der Befall mit Dassel fliesenlarven reduziert, wenn die Nester Kuhstärlingeier enthalten.

- Anscheinend ist der Befall mit Dasselfliegenlarven reduziert, wenn die Nester Kuhstärlingeier enthalten.
- statistisch signifikant?

- Anscheinend ist der Befall mit Dasselfliegenlarven reduziert, wenn die Nester Kuhstärlingeier enthalten.
- statistisch signifikant?
- Nullhypothese: Die Wahrscheinlichkeit eines Nests, mit Dasselfliegenlarven befallen zu sein hängt nicht davon ab, ob oder wieviele Kuhstärlingeier in dem Nest liegen.

Anzahlen der von Dasselfliegenlarven befallenen Nester

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	16	2	1	19
nicht befallen	2	11	16	29
Σ	18	13	17	48

Welche Anzahlen würden wir unter der Nullhypothese erwarten?

Anzahlen der von Dasselfliegenlarven befallenen Nester

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	16	2	1	19
nicht befallen	2	11	16	29
Σ	18	13	17	48

Welche Anzahlen würden wir unter der Nullhypothese erwarten?

Das selbe Verhältnis $19/48$ in jeder Gruppe.

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen				19
nicht befallen				29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} =$$

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13			19
nicht befallen				29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13$$

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13			19
nicht befallen				29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} =$$

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15		19
nicht befallen				29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13$$

$$13 \cdot \frac{19}{48} = 5.15$$

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15		19
nicht befallen				29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} = 5.15$$

Alle anderen Werte sind nun festgelegt durch die **Summen**.

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15		19
nicht befallen	10.87			29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} = 5.15$$

Alle anderen Werte sind nun festgelegt durch die **Summen**.

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15		19
nicht befallen	10.87	7.85		29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} = 5.15$$

Alle anderen Werte sind nun festgelegt durch die **Summen**.

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15	6.72	19
nicht befallen	10.87	7.85		29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} = 5.15$$

Alle anderen Werte sind nun festgelegt durch die **Summen**.

Erwartete Anzahlen von Dasselfliegenlarven befallener Nester, bedingt auf die Zeilen- und Spaltensummen:

Anzahl Kuhstärling-Eier	0	1	2	Σ
befallen	7.13	5.15	6.72	19
nicht befallen	10.87	7.85	10.28	29
Σ	18	13	17	48

$$18 \cdot \frac{19}{48} = 7.13 \quad 13 \cdot \frac{19}{48} = 5.15$$

Alle anderen Werte sind nun festgelegt durch die **Summen**.

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen				
	nicht befallen				
	Σ				

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87			
	nicht befallen				
	Σ				

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15		
	nicht befallen				
	Σ				

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15		0
	nicht befallen				0
	Σ	0	0	0	0

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15		0
	nicht befallen	-8.87			0
	Σ	0	0	0	0

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15		0
	nicht befallen	-8.87	-3.15		0
	Σ	0	0	0	0

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15	-5.72	0
	nicht befallen	-8.87	-3.15		0
	Σ	0	0	0	0

beobachtet (O, observed):	befallen	16	2	1	19
	nicht befallen	2	11	16	29
	Σ	18	13	17	48

erwartet: (E):	befallen	7.13	5.15	6.72	19
	nicht befallen	10.87	7.85	10.28	29
	Σ	18	13	17	48

O-E:	befallen	8.87	-3.15	-5.72	0
	nicht befallen	-8.87	-3.15	5.72	0
	Σ	0	0	0	0

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} = 29.5$$

- Wenn die Zeilen- und Spaltensummen gegeben sind, bestimmen bereits 2 Werte in der Tabelle alle anderen Werte
- $\Rightarrow df=2$ für Kontingenztafeln mit zwei Zeilen und drei Spalten.
- Allgemein gilt für n Zeilen und m Spalten:

$$df = (n - 1) \cdot (m - 1)$$

Wir haben den Wert $\chi^2 = 29.5$ beobachtet.

Wir haben den Wert $\chi^2 = 29.5$ beobachtet.

Unter der Nullhypothese „die Wahrscheinlichkeit, mit der ein Nest von Dasselfliegenlarven befallen wird, hängt nicht von der Anzahl Kuhstärling-Eier ab“ ist die Teststatistik (approximativ) χ^2 -verteilt mit $2 = (2 - 1) \cdot (3 - 1)$ Freiheitsgraden.

Wir haben den Wert $\chi^2 = 29.5$ beobachtet.

Unter der Nullhypothese „die Wahrscheinlichkeit, mit der ein Nest von Dasseliegenlarven befallen wird, hängt nicht von der Anzahl Kuhstärling-Eier ab“ ist die Teststatistik (approximativ) χ^2 -verteilt mit $2 = (2 - 1) \cdot (3 - 1)$ Freiheitsgraden.

Das 99%-Quantil der χ^2 -Verteilung mit $df=2$ ist 9.21 (<29.5), wir können also die Nullhypothese zum Signifikanzniveau 1% ablehnen.

Wir haben den Wert $\chi^2 = 29.5$ beobachtet.

Unter der Nullhypothese „die Wahrscheinlichkeit, mit der ein Nest von Dasselfliegenlarven befallen wird, hängt nicht von der Anzahl Kuhstärling-Eier ab“ ist die Teststatistik (approximativ) χ^2 -verteilt mit $2 = (2 - 1) \cdot (3 - 1)$ Freiheitsgraden.

Das 99%-Quantil der χ^2 -Verteilung mit $df=2$ ist 9.21 (<29.5), wir können also die Nullhypothese zum Signifikanzniveau 1% ablehnen.

(Denn wenn die Nullhypothese zutrifft, so würden wir in weniger als 1% der Fälle einen so extremen Wert der χ^2 -Statistik beobachten.)

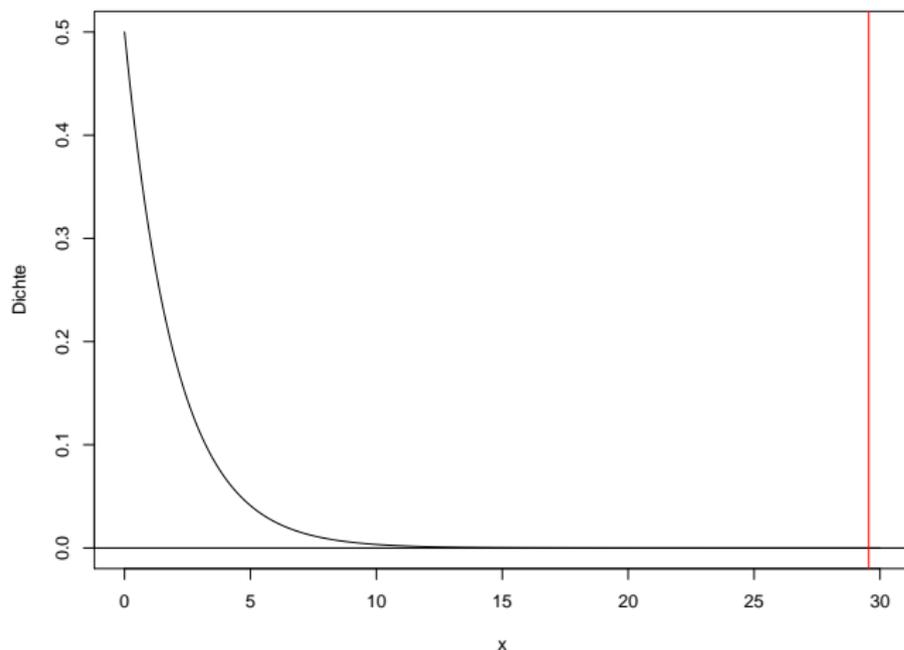
Wir haben den Wert $\chi^2 = 29.5$ beobachtet.

Unter der Nullhypothese „die Wahrscheinlichkeit, mit der ein Nest von Dasselbliegenlarven befallen wird, hängt nicht von der Anzahl Kuhstärling-Eier ab“ ist die Teststatistik (approximativ) χ^2 -verteilt mit $2 = (2 - 1) \cdot (3 - 1)$ Freiheitsgraden.

Das 99%-Quantil der χ^2 -Verteilung mit $df=2$ ist 9.21 (<29.5), wir können also die Nullhypothese zum Signifikanzniveau 1% ablehnen.

(Denn wenn die Nullhypothese zutrifft, so würden wir in weniger als 1% der Fälle einen so extremen Wert der χ^2 -Statistik beobachten.)

Faustregel: Die χ^2 -Approximation ist akzeptabel, wenn alle Erwartungswerte $E_i \geq 5$ erfüllen, was in dem Beispiel erfüllt ist. (Siehe die folgenden Folien für die mit dem Computer bestimmten exakten p -Werte.)

Dichte der chi-Quadrat-Verteilung mit $df=2$ Freiheitsgraden

Bemerkung 1: Genauere Rechnung ergibt: Für ein χ_2^2 -verteiltes X gilt $\mathbb{P}(X \geq 29.6) = 3.74 \cdot 10^{-7}$ (was hier wörtlich der p -Wert des χ^2 -Tests auf Unabhängigkeit wäre, in dieser Genauigkeit für statistische Zwecke allerdings sinnlos ist).

Bemerkung 2: Um die Gültigkeit der χ^2 -Approximation (und der Faustregel) in diesem Beispiel einzuschätzen, könnten wir einen Computer beauftragen, durch vielfach wiederholte Simulation den p -Wert zu schätzen.

Mit **R** funktionierte das beispielsweise folgendermaßen:

```
> M <- matrix(c(16,2,2,11,1,16),nrow=2)
> M
      [,1] [,2] [,3]
[1,]   16    2    1
[2,]    2   11   16
> chisq.test(M,simulate.p.value=TRUE,B=50000)
```

```
      Pearson's Chi-squared test with simulated p-value
      (based on 50000 replicates)
```

```
data:  M
X-squared = 29.5544, df = NA, p-value = 2e-05
```

Wir sehen: Der empirisch geschätzte p -Wert $2 \cdot 10^{-5}$ stimmt zwar nicht mit dem aus der χ^2 -Approximation überein, aber beide sind hochsignifikant klein (und in einem Bereich, in dem der exakte Wert sowieso statistisch „sinnlos“ ist). Insoweit ist die Faustregel hier bestätigt.

Inhalt

- 1 Wiederholung: Wilcoxon's Rangsummentest
 - Motivation
 - Wilcoxon-Test für unabhängige Stichproben
- 2 Kontingenztafeln
 - χ^2 -Anpassungstest für eine vorgegebene Verteilung
 - χ^2 -Test auf Homogenität bzw. Unabhängigkeit
 - Exkurs: χ^2 -Test für Modelle mit angepassten Parametern

Gegeben sei eine Population im *Hardy-Weinberg-Gleichgewicht* und ein Gen-Locus mit zwei möglichen Allelen A und B mit Häufigkeiten p und $1 - p$.

↪ Genotyp-Häufigkeiten

$$\begin{array}{c|c|c} \text{AA} & \text{AB} & \text{BB} \\ \hline p^2 & 2 \cdot p \cdot (1 - p) & (1 - p)^2 \end{array}$$

Beispiel: M/N Blutgruppen; Stichprobe: $n = 6129$ Amerikaner
europäischer Abstammung

beobachtet:	MM	MN	NN
	1787	3037	1305

Beispiel: M/N Blutgruppen; Stichprobe: $n = 6129$ Amerikaner
europäischer Abstammung

beobachtet:

MM	MN	NN
1787	3037	1305

Geschätzte Allelhäufigkeit von M:

$$\hat{p} = \frac{2 \cdot 1787 + 3037}{2 \cdot 6129} = 0.5393$$

Beispiel: M/N Blutgruppen; Stichprobe: $n = 6129$ Amerikaner
europäischer Abstammung

beobachtet:

MM	MN	NN
1787	3037	1305

Geschätzte Allelhäufigkeit von M:

$$\hat{p} = \frac{2 \cdot 1787 + 3037}{2 \cdot 6129} = 0.5393$$

↪ Erwartete Werte:

MM	MN	NN
\hat{p}^2	$2 \cdot \hat{p} \cdot (1 - \hat{p})$	$(1 - \hat{p})^2$

Beispiel: M/N Blutgruppen; Stichprobe: $n = 6129$ Amerikaner
europäischer Abstammung

beobachtet:

MM	MN	NN
1787	3037	1305

Geschätzte Allelhäufigkeit von M:

$$\hat{p} = \frac{2 \cdot 1787 + 3037}{2 \cdot 6129} = 0.5393$$

↪ Erwartete Werte:

MM	MN	NN	
\hat{p}^2	$2 \cdot \hat{p} \cdot (1 - \hat{p})$	$(1 - \hat{p})^2$	
0.291	0.497	0.212	(Anteile)
$n \cdot \hat{p}^2$	$n \cdot 2 \cdot \hat{p} \cdot (1 - \hat{p})$	$n \cdot (1 - \hat{p})^2$	

Beispiel: M/N Blutgruppen; Stichprobe: $n = 6129$ Amerikaner
europäischer Abstammung

beobachtet:

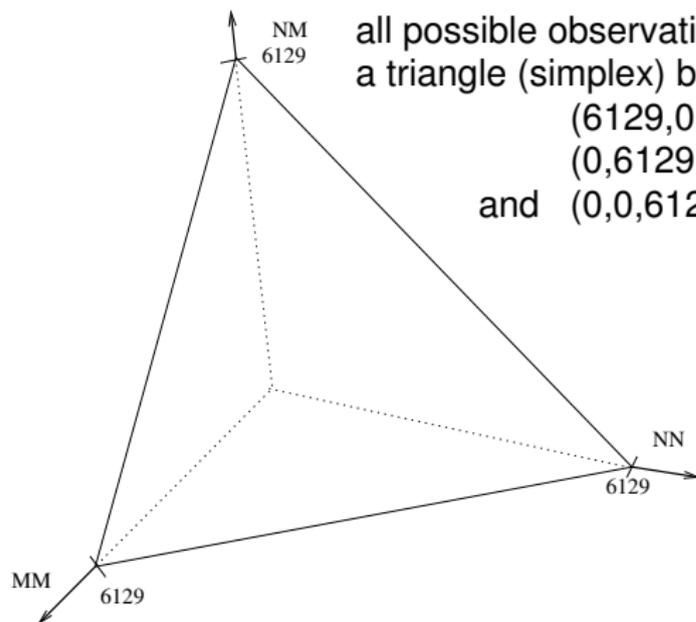
MM	MN	NN
1787	3037	1305

Geschätzte Allelhäufigkeit von M:

$$\hat{p} = \frac{2 \cdot 1787 + 3037}{2 \cdot 6129} = 0.5393$$

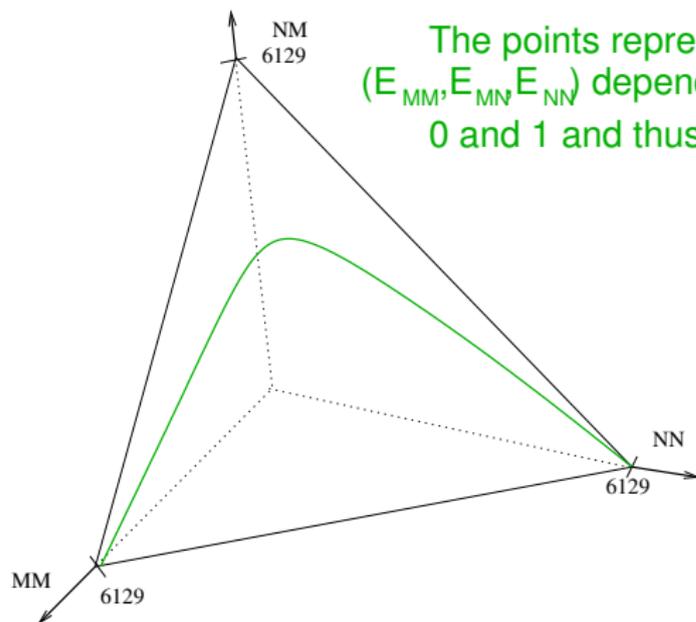
↪ Erwartete Werte:

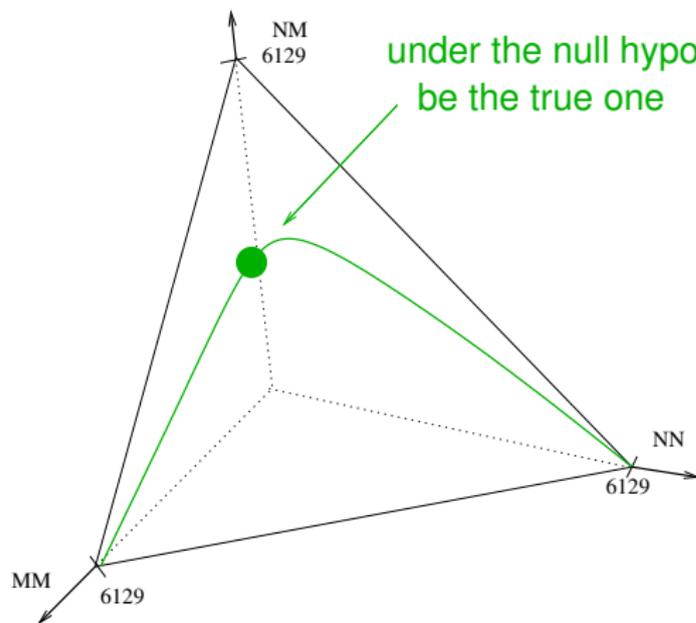
MM	MN	NN	
\hat{p}^2	$2 \cdot \hat{p} \cdot (1 - \hat{p})$	$(1 - \hat{p})^2$	
0.291	0.497	0.212	(Anteile)
$n \cdot \hat{p}^2$	$n \cdot 2 \cdot \hat{p} \cdot (1 - \hat{p})$	$n \cdot (1 - \hat{p})^2$	
1782.7	3045.6	1300.7	(Häufigkeiten)

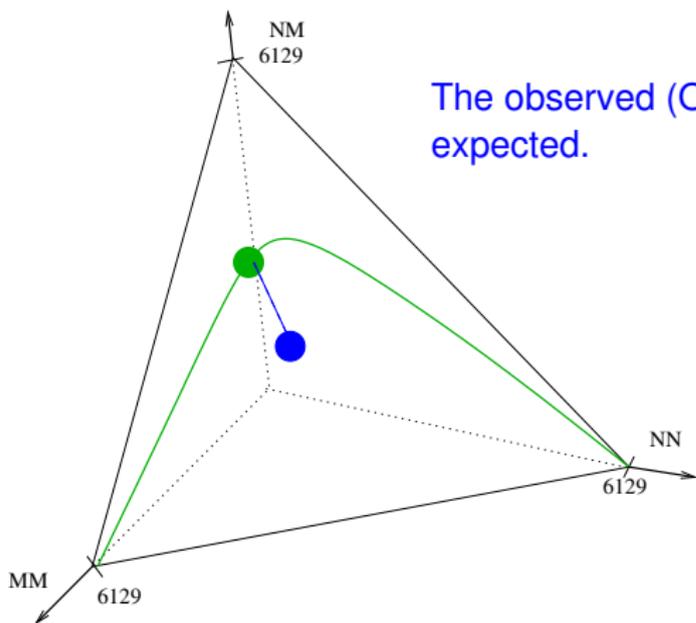


all possible observations (O_{MM}, O_{MN}, O_{NN}) are located on a triangle (simplex) between
 $(6129, 0, 0)$
 $(0, 6129, 0)$
and $(0, 0, 6129)$

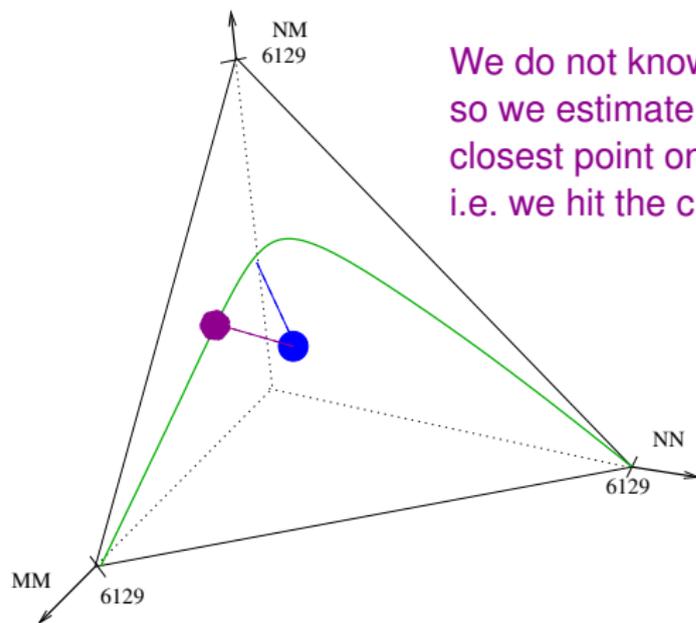
The points representing the Expected Values (E_{MM}, E_{MN}, E_{NN}) depend on one parameter p between 0 and 1 and thus form a curve in the simplex.



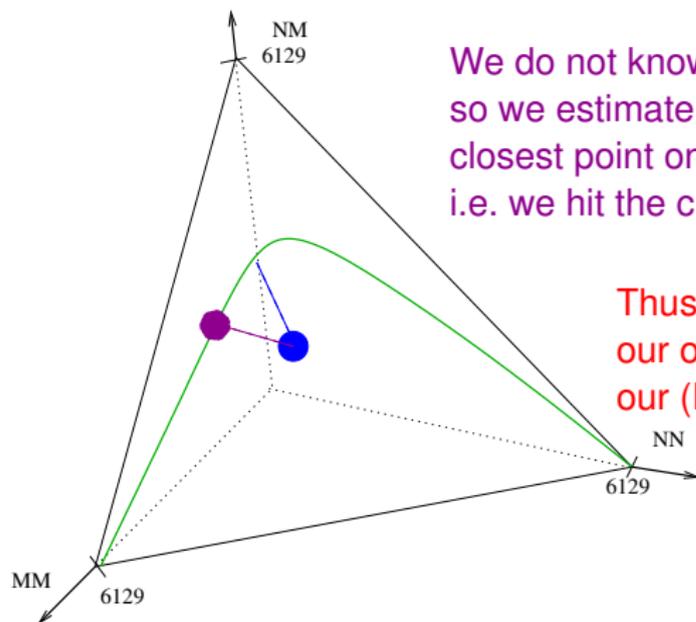




The observed (O_{MM}, O_{NM}, O_{NN}) will deviate from the expected.



We do not know the true expectation values so we estimate $(E_{MM} E_{MN} E_{NN})$ by taking the closest point on the curve of possible values, i.e. we hit the curve in a right angle.



We do not know the true expectation values so we estimate (E_{MM}, E_{NM}, E_{NN}) by taking the closest point on the curve of possible values, i.e. we hit the curve in a right angle.

Thus, deviations between our observations (O_{MM}, O_{NM}, O_{NN}) and our (E_{MM}, E_{NM}, E_{NN}) can only be in one dimension: perpendicular to the curve.

Für die Anzahl Freiheitsgrade im χ^2 -Test mit angepassten Parametern gilt

$$df = k - 1 - m$$

k = Anzahl Gruppen ($k=3$ Genotypen)

m = Anzahl Modellparameter ($m=1$ Parameter p)

Für die Anzahl Freiheitsgrade im χ^2 -Test mit angepassten Parametern gilt

$$df = k - 1 - m$$

k = Anzahl Gruppen ($k=3$ Genotypen)

m = Anzahl Modellparameter ($m=1$ Parameter p) im

Blutgruppenbeispiel also:

$$df = 3 - 1 - 1 = 1$$

Der Wert der χ^2 -Statistik ist

$$\frac{(1787 - 1782.7)^2}{1782.7} + \frac{(3037 - 3045.6)^2}{3045.6} + \frac{(1305 - 1300.7)^2}{1300.7} = 0.049.$$

Der Wert der χ^2 -Statistik ist

$$\frac{(1787 - 1782.7)^2}{1782.7} + \frac{(3037 - 3045.6)^2}{3045.6} + \frac{(1305 - 1300.7)^2}{1300.7} = 0.049.$$

Dieser Wert gibt keinen Anlass, an der Nullhypothese „die Population ist bezüglich des M/N-Blutgruppensystems im HW-Gleichgewicht“ zu zweifeln: 0.049 liegt zwischen dem 10%- und dem 30%-Quantil der χ^2 -Vert. mit einem Freiheitsgrad, wir könnten also eine solche oder noch größere Abweichung zwischen Beobachtung und Erwartung in ca. 80% der Fälle erwarten (der p -Wert ist 0.83).