Rangsummen-Test von Wilcoxon

Curdin Derungs

Unterlagen: https://github.com/curdon/zhaw

Lernziele

1. Sie haben eine Eselsbrücke für den Wilcoxon-Test
2. Sie können den Wilcoxon-Test konzeptionell erklären
3. Sie kennen die hauptsächlichen Unterschiede zum T-Test
4. Sie können den Wilcoxon-Test mit der R-Software auf den Einstichprobenfall und auf gepaarte Stichproben anwenden
5. Sie können den Output der entsprechenden R-Funktion interpretieren und in Ihre Worte übersetzen

Anwendungsfall

- Nicht-parmetrisches Pendant zum T-Test
  - Stichproben sind nicht normalverteilt -> sind sie fast nie!

- Kann für Einstichproben und für gepaarte Stichproben verwendet warden. Nicht für ungepaarte Stichproben (siehe Mann-Whitney Test).

- Anforderung an Stichprobe(n):
  - Messungen sind i.i.d.
    - Unabhängige: X_i sagt nicht über X_{i+n} aus
    - Identisch: es gibt keinen Trend, z.B. $X_i > X_{i+n} > X_{i+n+m}$

  - Skalaniveau der Messungen mindestens ordinal, sicherer ist interval
    - Ordinal: Abstände zwischen Rängen sollten möglichst vergleichbar sein

  - Messungen sind symetrisch um den Erwartungswert verteilt

Einstichproben vs. gepaarte Stichproben

ID	Präferenz
1	5
2	3
3	8
…	…

- Nur eine Stichprobe
- Beispielfrage: Haben die Teilnehmer eine positive Grundeinstellung zum Thema Statistik?

→ Details siehe Wandtafel

ID	Vorher	Nachher
1	5	3
2	3	3
3	8	6
…	…	…

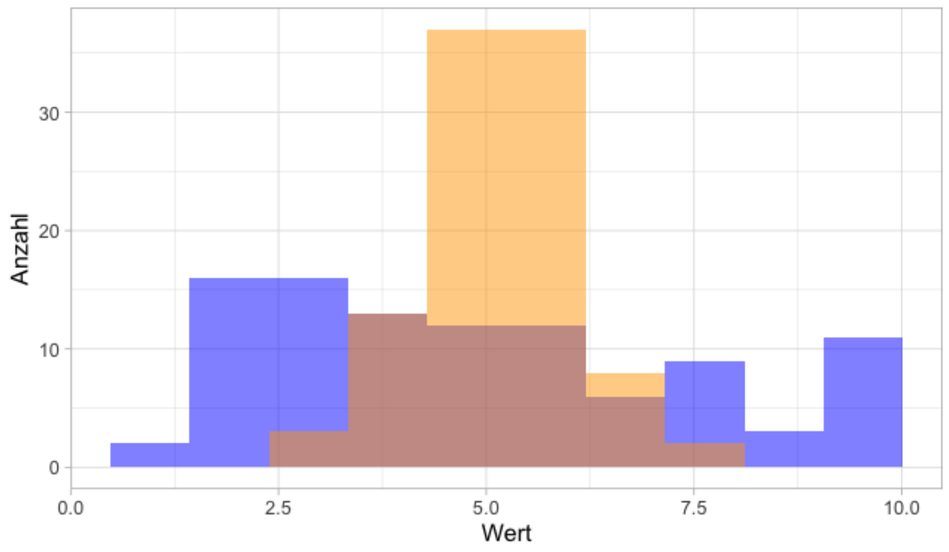
- Zwei Stichproben
- Zwei Messungen pro ID (Subjekt)
- Beispielfrage: Unterscheidet sich die Einstellung zum Produkt vor und nach der Info-Veranstaltung?

→ Details siehe R-Folien

Parametrisch vs. Nicht-Parametrisch

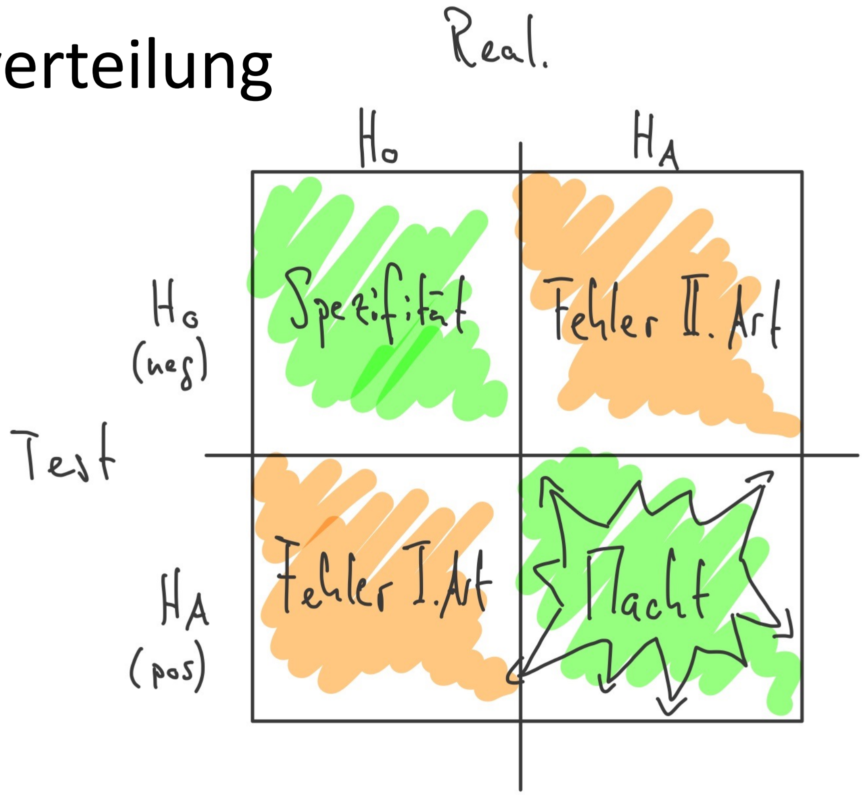
- Nicht-Parametrisch → nicht normalverteilt (Wann Wilcoxon?):
  - kleine Stichprobe (n < 20)
  - deutliche Ausreisser
  - nicht symetrisch um den Erwartungswert verteilt
  - Ordinale Messskala
- Mit Visualisierung testen (z.B. Histogram, QQ-Plot)

```
library(ggplot2)

set.seed(1)

#Beispiel Normal- vs. Nicht-Normalverteilt
normal <- rnorm(mean = 5, sd = 1, n = 100)
uniform <- runif(min=1, max=10, n = 100)

ggplot()+
geom_histogram(aes(x=uniform), bins = 10, alpha = .5, fill='blue')+
geom_histogram(aes(x=normal), bins = 10, alpha = .5, fill='orange')+
xlab("Wert")+
ylab("Anzahl")+
theme_light()
```



Parametrisch vs. Nicht-Parametrisch II

- Bei Normalverteilung haben parametrische Tests die grössere Macht als nicht-parametrische
  - Es wird mehr Information in den Daten verwendet, nicht bloss Ränge
- (Meist) kleinere Macht aber bei nicht-Normalverteilung
  - [Link]() zu Vergleichsstudie



Rechenbeispiel: Wilcoxon RST für gepaarte SP

ID	Vorher	Nachher
1	5	2
2	3	4
3	8	6
4	7	7
5	4	5

Diff
-3
1
-2
0
1

Rang
4
1.5
3
--
1.5

- Positive RS: $V_{+}: 1.5 + 1.5 = 3$
- Negative RS: $V_{-}: 4 + 3 = 7$
- RS Total: $V_{+} + V_{-} = \frac{n \times (n+1)}{2} = \frac{4 \times 5}{2} = 10$
- Erwartete RS unter H_0: $\mu_0: \frac{V_{+} + V_{-}}{2} = \frac{10}{2} = 5$
- Beobachtete RS: $V = \max(V_{+}; V_{-})$
- Test Stat.: $z = \frac{V - \mu_0}{\sigma_0} = \frac{7-5}{?} = ?$

Ein bisschen kompliziert

In Wilcoxon RST Tabelle nachschlagen für P-Value

H_0: Es gibt keinen Unterschied zwischen Vorher und Nachher
(+) Rangsummen = (-) Rangsummen

H_A: Nachher sind die Messwerte tiefer (z.B. Blutdruck)
(+) Rangsummen < (-) Rangsummen

Test Stat.: $z = \frac{beobachtet - erwartet}{Standardfehler}$

Wilcoxon-Test in R

R-Markdown

- HTML: https://github.com/curdon/zhaw/blob/main/rcode_cderungs.html
- R-Code: https://github.com/curdon/zhaw/blob/main/rcode_cderungs.Rmd

Links & Literatur

- Methodenberatung UZH: https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/wilkoxon.html
- MarinStatsLectures: https://www.youtube.com/watch?v=v4ZHlTbTOK8
- Fahrmeir et al.: Statistik: Der Weg zur Datenanalyse, S. 406ff
- Vergleichsstudie zur Macht des Wilcoxon RST bei unterschiedlichen Verteilungen: Imam et al. 2014