

Statistik – Methoden zum Mittelwertvergleich von mehreren Gruppen

Kruskal-Wallis-Test

Kruskal-Wallis-Test

Die einfaktorielle ANOVA kann nur eingesetzt werden, wenn bestimmte Voraussetzungen erfüllt sind:

- Mindestens intervallskalierte abhängige Variable
- Merkmalsausprägungen müssen unabhängig voneinander sein
- Normalverteilung der abhängigen Variable innerhalb aller Gruppen
- Varianzhomogenität der Gruppen

Sollten einzelne oder alle Voraussetzungen nicht erfüllt sein, muss ggf. zu einem nicht-parametrischen Verfahren gewechselt werden

Als Ersatz für die einfaktorielle ANOVA steht hier der **Kruskal-Wallis-Test** zu Verfügung

Kruskal-Wallis-Test

Voraussetzungen für den Kruskal-Wallis-Test

- Mindestens ordinal skalierte abhängige Variable
- Die zu vergleichenden Gruppen sind von einander unabhängig (ansonsten kann auf einen Friedman-Test zurückgegriffen werden)

Kruskal-Wallis-Test

Der Aufbau des Kruskal-Wallis-Tests ist vergleichbar mit dem Mann-Whitney-U-Test und wird entsprechend hier nicht weiter beschrieben

Hypothesen

H_0 Die Zentrallagen der einzelnen Gruppen unterscheiden sich nicht voneinander

H_1 Mindestens zwei Gruppen unterscheiden sich in der Zentrallage

Soll eine Aussage gemacht werden, wo Unterschiede liegen, ist ein entsprechendes Post-hoc-Verfahren erforderlich

Kruskal-Wallis-Test

Teststatistik des Kruskal-Wallis-Tests

- Alle Stichprobenwerte werden zusammengefasst, geordnet und mit Rängen versehen
- Für die einzelnen Gruppen werden nun die Rangsummen S_h berechnet

Kruskal-Wallis-Test

Teststatistik des Kruskal-Wallis-Tests

$$H = \frac{12}{n(n+1)} \sum_h \frac{S_h^2}{n_h} - 3(n+1)$$

Im Falle von Bindungen bestimmt sich die Teststatistik wie folgt:

$$H = \frac{\frac{12}{n(n+1)} \sum_h \frac{S_h^2}{n_h} - 3(n+1)}{1 - \frac{1}{(n^3 - n)} \sum t_{r(i)}^3 - t_{r(i)}}$$

mit $t_{r(i)}$: Zahl der gebundenen Beobachtungen mit Rang i

Kruskal-Wallis-Test

Die Teststatistik kann nun nach dem schon bekannten Muster mit einer χ^2 -Verteilung verglichen werden

Die erforderliche Anzahl der Freiheitsgrade bestimmt sich zu:

$$df = k - 1 \quad \text{mit } k = \text{Anzahl der Gruppen}$$

Für $\chi_{emp.}^2 > \chi_{krit.}^2$ verwerfen wir die Nullhypothese und wechseln zur Alternativhypothese

Für $h = 3$ und $n_h < 6$ gilt die χ^2 -Verteilung nicht, es ist auf geeignetes Tabellenwerk zurückzugreifen

Kruskal-Wallis-Test

- Der Kruskal-Wallis-Test dient als nicht-parametrischer Ersatz für die einfaktorielle ANOVA
- Eine mehrfaktorielle ANOVA kann bei Nichterfüllung der Voraussetzungen durch den nicht-parametrischen Scheirer-Ray-Hare-Test ersetzt werden (im *rcompanion*-Paket in RStudio enthalten)
- Der Kruskal-Wallis-Test macht im Falle der Alternativhypothese keine Aussage darüber, welche Gruppen sich unterscheiden, ein paarweiser Vergleich steht in RStudio zur Verfügung (`pairwise.wilcox.test`)

Kruskal-Wallis-Test

Beispielrechnung Beispiel_BK_KW.xlsx

Kruskal-Wallis rank sum test

data: variable by factor

Kruskal-Wallis chi-squared = 8.1204, df = 2, p-value = 0.01725

Ergebnisse aus Beispiel_BK_KW.xlsx

Rangsummen	Sh	Sh^2
A	1449	2099601
B	1602	2566404
C	1044	1089936
nh	30	
n	90	
Keine Bindungen		
H	8,12043956	