Rangvarianzanalyse

Die Rangvarianzanalyse, auch *Kruskal-Wallis*-Test genannt, ist die Erweiterung des U-Tests auf $k \geq 2$ Stichproben und stellt somit eine parameterfreie Alternative zur einfachen VA für unabhängige Stichproben dar.

ZIEL

Analog zur VA ist der Zweck des *Kruskal-Wallis*-Tests der Vergleich mehrerer Stichproben dahingehend, ob sich zumindest zwei von ihnen in der Größe ihrer Messwerte signifikant unterscheiden.

VORAUSSETZUNGEN

Analog zum U-Test lauten die Bedingungen:

1.) Das untersuchte Merkmal muss stetig sein.
2.) Die Daten müssen mindestens rangskaliert sein.
3.) Die Stichproben müssen unabhängig sein.

VORGEHENSWEISE

Aufstellen der Hypothesen

$H_0 \ldots$ rein zufällige Abweichungen, es handelt sich um Stichproben aus einer gemeinsamen Population (oder aus identischen Populationen)

$H_1 \ldots$ Abweichungen nicht nur Zufallsschwankungen, mindestens ein Paar von Stichproben ist aus unterschiedlichen Populationen

Rangsummen

Alle Messwerte werden der Größe nach geordnet und mit ihrer Rangzahl versehen. Bei mehreren gleichen Werten („Bindung“) wird jedem die mittlere Rangzahl zugeordnet.

Anschließend wird die Rangsumme $r_{.j}$ jeder Stichprobe berechnet, um in weiterer Folge die Testgröße h berechnen zu können.

Prüfgröße

$$h = \frac{12}{N(N+1)} \cdot \sum_{j=1}^{k} \frac{r_{\cdot j}^2}{n_j} - 3(N+1)$$

Signifikanzprüfung

h muss mit dem kritischen Wert aus der χ^2-Tabelle verglichen werden.

Achtung:
Da die H_1 gerichtet ist ($\rightarrow$ *große* Unterschiede), handelt es sich - wie bei der VA - immer um eine einseitige Fragestellung.

Ist die Prüfgröße größer als der kritische Wert ($h > \chi^2_{krit.}$), so ist der Test signifikant ($\Rightarrow H_0$ verwerfen, H_1 annehmen).

Bindungen

Eigentlich wird ein stetiges Merkmal vorausgesetzt. Die Wahrscheinlichkeit für zwei identische Werte müsste also gleich 0 sein. In der Praxis ist das Auftreten von Bindungen aber kaum zu vermeiden. Falls dies der Fall ist, muss h korrigiert werden:

$$h_{korr.} = \frac{1}{1-m} \cdot h$$

$m = \frac{1}{N^3-N} \cdot \sum\limits_{l} (t_l^3 - t_l)$

$N = \sum n_j \ldots$ Gesamtzahl der Messwerte
$t_l \ldots$ Anzahl der Werte innerhalb der l-ten Bindung

Beachte:
h muss nur korrigiert werden, wenn es nicht signifikant ist.

Begründung:
Die korrigierte Form ist stets größer als die ursprüngliche Testgröße ($\frac{1}{1-m} > 1 \quad \Rightarrow \quad h_{korr.} > h$). Die Korrektur für ein signifikantes h, also $h > \chi^2_{krit.}$, ist somit nicht notwendig, denn das Ergebnis würde sich mit Sicherheit nicht ändern: $h_{korr.} > h > \chi^2_{krit.}$.