

7 Vergleich mehrerer verbundener Stichproben

7.1 Friedman Rang-Varianzanalyse

Der Test von Friedman ist ein Analogon zur Varianzanalyse, nur dass dieser Test nichtparametrisch ist und deshalb im Gegensatz zur Varianzanalyse keine Normalverteilung voraussetzt. Außerdem geht man hier von g Teilstichproben gleichen Umfangs n aus, die verbunden sein können. Dabei könnte es sich beispielsweise um eine Studie zur Untersuchung der Gewichtsveränderung von n Personen handeln, von denen die Körpergewichte zu g Zeitpunkten vorliegen. Wir gehen allgemein vom Modell der doppelten Varianzanalyse aus:

$$X_{ij} = \mu + \alpha_i + \beta_j + E_{ij}, \text{ mit } i = 1, 2, \dots, n \text{ und } j = 1, 2, \dots, g.$$

X_{ij} ist die i -te Beobachtung der j -ten Teilstichprobe. E_{ij} sind die Fehlervariablen, die nicht, wie bei der Varianzanalyse, normalverteilt sein müssen. Die Zufallsvariablen $E_{1j}, E_{2j}, \dots, E_{nj}$ werden als unabhängig und identisch stetig verteilt vorausgesetzt, d.h. die Zufallsvariablen E_{ij} sind innerhalb der j -ten Teilstichprobe unabhängig und identisch stetig verteilt. α_i ist der Effekt der i -ten Beobachtung und β_j der Effekt der j -ten Gruppe. Wir wollen nun auf einen Unterschied zwischen den Gruppen testen und formulieren die Hypothesen:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_g$$

gegen

$$H_1: \text{mindestens ein } \beta_j \text{ ist verschieden von } \beta_k$$

Im Beispiel verwenden wir die folgenden Daten:

v1	v2	v3	v4
2	7	4	1
8	24	11	4
6	16	9	1
25	15	5	1

Die Rangzahlen werden für jede Datenzeile $x_{i1}, x_{i2}, \dots, x_{ig}$ separat bestimmt (siehe Output). Im Folgenden sind $r_{i1}, r_{i2}, \dots, r_{ig}$ die Rangzahlen der i -ten Zeile. D.h. im Beispiel ist $r_{11} = 2, r_{12} = 4, r_{13} = 3$ und $r_{14} = 1$.

Wenn Bindungen auftreten, was in unserem Beispiel nicht der Fall ist (kein Wert kommt innerhalb einer Zeile doppelt vor), so müssen diese bei der Berechnung der Prüfgröße wieder berücksichtigt werden. Dazu benötigen wir diesmal pro Zeile eine Liste der Häufigkeiten des Auftretens der Beobachtungen. Es seien $t_{i1}, t_{i2}, \dots, t_{ik_i}$ die Häufigkeiten in der i -te Zeile.

Im Beispiel sind alle $t_{ij} = 1$ und $k_i = 4$ für $j = 1, 2, \dots, k_i$ und $i = 1, 2, \dots, n$. Würde man die erste Zeile modifizieren, z.B. $x_{11} = 2, x_{12} = 7, x_{13} = 1$ und $x_{14} = 1$, dann würden sich die Rangzahlen $r_{11} = 3, r_{12} = 4, r_{13} = 1,5$ und $r_{14} = 1,5$ für die erste Zeile ergeben. In diesem Fall wäre $t_{11} = 2, t_{12} = 1, t_{13} = 1$ und somit $k_1 = 3$ (siehe Kapitel 1.5).

Zur Durchführung des Tests müssen nach der Dateneingabe die Variablen v1, v2, v3 und v4 ganz rechts im Menü unter „Welche Spalten sollen verglichen werden.“ ausgewählt werden. Danach kann **→Vergleich mehrerer verbundener Teilstichproben →Friedman Rangvarianzanalyse** gewählt werden.

Somit ergibt sich der Output:

Friedman

H0: Es gibt keinen Unterschied zwischen den Effekten der Gruppen
gegen

H1: Mindestens zwei Gruppen unterscheiden sich hinsichtlich der Effekte

Die Daten:

	Teilstichprobe 1 (Rangzahl in Klammern)	Teilstichprobe 2 (Rangzahl in Klammern)	Teilstichprobe 3 (Rangzahl in Klammern)	Teilstichprobe 4 (Rangzahl in Klammern)
Beobachtung 1	2 (2)	7 (4)	4 (3)	1 (1)
Beobachtung 2	8 (2)	24 (4)	11 (3)	4 (1)
Beobachtung 3	6 (2)	16 (4)	9 (3)	1 (1)
Beobachtung 4	25 (4)	15 (3)	5 (2)	1 (1)
Rangsummen	10	15	11	4

Prüfgröße v	9.3
Prüfgröße v bei Bindungen	9.3
approximativer p-Wert (Freiheitsgrade Chi-Quadrat-Verteilung: 3)	0.0256

Es sei r_j die Rangsumme der j -ten Spalte, also:

$$r_j = \sum_{i=1}^n r_{ij}$$

Im Beispiel ist $r_1 = 10$, $r_2 = 15$, $r_3 = 11$ und $r_4 = 4$.

Für den Mittelwert dieser Rangsummen gilt (dahinter sehen Sie den Wert im Beispiel):

$$\bar{r} = \frac{1}{g} \sum_{j=1}^g r_j = \frac{n(g+1)}{2} = 10$$

Treten keine Bindungen auf, so ist v die Prüfgröße:

$$v = \frac{12}{ng(g+1)} \sum_{j=1}^g (r_j - \bar{r})^2 = \frac{12}{ng(g+1)} \sum_{j=1}^g r_j^2 - 3n(g+1)$$

Im Beispiel ist $v = 9,3$.

Bei Bindungen wird die Prüfgröße wie folgt berechnet:

$$v^* = \frac{12 \sum_{j=1}^g (r_j - \bar{r})^2}{ng(g+1) - \frac{1}{g-1} \sum_{i=1}^n \left(\left(\sum_{j=1}^{k_i} t_{ij}^3 \right) - g \right)}$$

Wenn - wie im Beispiel - keine Bindungen auftreten, haben beide Prüfgrößen denselben Wert (d.h. hier gilt dann $v^* = v$).

Die Prüfgröße ist (unter H_0) eine Realisierung einer asymptotisch Chi-Quadrat-verteilten Zufallsvariablen mit $g - 1$ Freiheitsgraden. Die Nullhypothese kann dann auf einem Signifikanzniveau von α beim asymptotischen Test verworfen werden, wenn:

$$\text{p-Wert} = 1 - F_{\chi_{g-1}^2}(v^*) \leq \alpha$$

Im Beispiel gilt: p-Wert $\approx 0,0256$

Wir könnten hier die Nullhypothese auf einem Signifikanzniveau von 5% verwerfen ($p\text{-Wert} \leq 0,05$) und somit einen signifikanten Unterschied zwischen den Teilstichproben nachweisen.

In Büchern wie z.B. in [3], [8] und [9] findet man Tabellen der exakten Verteilung für den Fall, dass keine Bindungen vorliegen. Wir wollen noch mal zum Schluss die exakte Verteilung (siehe erste Spalte) in unserem Beispiel in einer Tabelle ausgeben:

v	$P(V \leq v)$	$P(V = v)$	$P(V \geq v) = 1 - P(V < v)$
0	0,007595	0,007596	1
0,3	0,071832	0,064236	0,992405
0,6	0,099609	0,027778	0,928168
0,9	0,200304	0,100694	0,900391
1,2	0,246094	0,045790	0,799696
1,5	0,323351	0,077257	0,753906
1,8	0,351128	0,027778	0,676649
2,1	0,476128	0,125	0,648872
2,4	0,492405	0,016276	0,523872
2,7	0,567925	0,075521	0,507595
3,0	0,610749	0,042824	0,432075
3,3	0,645472	0,034722	0,389251
3,6	0,675854	0,030382	0,354528
3,9	0,758319	0,082465	0,324146
4,5	0,799986	0,041667	0,241681
4,8	0,810258	0,010272	0,200014
5,1	0,841508	0,03125	0,189742
5,4	0,858869	0,017361	0,158492
5,7	0,894748	0,035880	0,141131
6,0	0,905599	0,010851	0,105252
6,3	0,92296	0,017361	0,094401
6,6	0,932219	0,009259	0,077040

v	$P(V \leq v)$	$P(V = v)$	$P(V \geq v) = 1 - P(V < v)$
6,9	0,946108	0,013889	0,067781
7,2	0,948278	0,002170	0,053892
7,5	0,963614	0,015336	0,051722
7,8	0,967086	0,003472	0,036386
8,1	0,980975	0,013889	0,032914
8,4	0,985894	0,004919	0,019025
8,7	0,988498	0,002604	0,014106
9,3	0,993128	0,004630	0,011502
9,6	0,993779	0,000651	0,006872
9,9	0,997251	0,003472	0,006221
10,2	0,998409	0,001157	0,002749
10,8	0,99906	0,000651	0,001591
11,1	0,999928	0,000868	0,000940
12,0	1	0,000072	0,000072

Der exakte p-Wert beträgt: $P(V \geq 9,3) = 0,0115\dots$

Mit diesem p-Wert könnten wir auch hier die Nullhypothese auf einem Signifikanzniveau von 5% verwerfen.

Umsetzung mit SAS:

```

data dat1;
input b x y;
datalines;
1 2
1 8
1 6
1 25
2 7
2 24
2 16
2 15
3 4
3 11
3 9
3 5
4 1
4 4
4 1
4 1
run;

```

```
proc sort data =dat1;
by x;
run;

proc rank data = dat1 out=dat2;
by x;
var z;
ranks rz;
run;

proc sort data = dat2;
by y;
run;

proc univariate data = dat2 noprint;
var rz;
by y;
output out = temp sum = s;
run;

data temp;
set temp;
sq = s**2;
run;

proc univariate data = temp;
var sq;
output out = temp1 sum = ssq;
run;

data output;
set temp1;    n=4;g=4; /* n: Anzahl der Beobachtungen
und g: Anzahl der Gruppen*/
v=12/(n*g*(g+1))*ssq-3*n*(g+1);
prob_v = 1-probchi(v,g-1);
run;

proc print data = output;
run;
```

SAS-Output der Prozedur Print (Datei output):

Beob.	ssq	n	g	v	prob_v
1	462	4	4	9.3	0.025557

Oder in einer kurzen Variante mit der **Prozedur Freq**:

```
data dat1xy;
input b x y;
cards;
1 1 2
2 1 8
3 1 6
4 1 25
1 2 7
2 2 24
3 2 16
4 2 15
1 3 4
2 3 11
3 3 9
4 3 5
1 4 1
2 4 4
3 4 1
4 4 1
run;

proc freq data=dat1xy;
tables b*x*y /
cmh2 scores=rank noprint;
run;
```

SAS-Output der Prozedur Freq:

Das SAS System

Die Prozedur FREQ

Beschreibende Statistiken für x nach y
Kontrolliert für b

Cochran-Mantel-Haenszel-Statistiken (RANK-Werte)

Statistik	Alternative Hypothesis	DF	Wert	Prob
1	Korrelation ungleich 0	1	3.6300	0.0567
2	Zeilenmittel ungleich	3	9.3000	0.0256

Gesamtstichprobengröße = 16