Permutationstests

1 Wilcoxon-Rangsummentest (exakter Permutationstest)

Gegeben sei das Zweistichprobenproblem aus Tabelle 1.

Tabelle 1: *Beispiel 1*

	x_1	x_2	x_3	y_1	y_2	y_3	y_4	y_5
Beobachtung	3.8	4.0	4.1	0.1	1.6	0.8	2.7	1.1
Rang	6	7	8	1	4	2	5	3

Hierbei könnte es nun von Interesse sein, ob die Lage der Verteilungen von X und Y übereinstimmen. Für diese Nullhypothese kann untersucht werden, ob es in ihrem Sinne unplausibel ist, dass die drei größten Werte der vereinigten Stichprobe zu X gehören. Unter H_0 hätte jede der acht Beobachtungen mit gleicher Wahrscheinlichkeit die größte Beobachtung sein können, ebenso die zweitgrößte, usw.. Betrachtet man nun den Rangvektor, so gibt es 8! verschiedene Möglichkeiten die Einträge anzuordnen, welche unter H_0 gleich wahrscheinlich sind. Um die Verteilung der Ränge mittels einer Teststatistik zu charakterisieren, kann die Rangsumme der X_i, $1 \leq i \leq 3$, gebildet werden, so dass in dem Beispiel nur die Ränge r_1, r_2 und r_3 von Bedeutung sind. Für jede Kombination (r_1, r_2, r_3) gibt es genau $3! \cdot 5!$ Permutationen, so dass jede Kombination die Wahrscheinlichkeit $\frac{3!5!}{8!} = \frac{1}{56}$ aufweist. Die Verteilung für die Situation aus Beispiel 1 ist in Tabelle 2 dargestellt.

Daraufhin kann die beobachtete Rangsumme mit dem entsprechenden Quantil der Verteilung von W_N verglichen werden, um eine Testentscheidung zu treffen. Dieses Beispiel ist in ähnlicher Form in Büning, Trenkler (1994) zu finden.

Tabelle 2: *Verteilung der Teststatistik W_N*

Teststatistik: $W_N = w$	Ränge der X_i	$P(W_N = w)$
21	(6,7,8)	$\frac{1}{56}$
20	(5,7,8)	$\frac{1}{56}$
19	(4,7,8), (5,6,8)	$\frac{2}{56}$
18	(3,7,8), (4,6,8), (5,6,7)	$\frac{3}{56}$
17	(2,7,8), (3,6,8), (4,6,7), (4,5,8)	$\frac{4}{56}$
16	(1,7,8), (2,6,8), (3,5,8) (3,6,7), (4,5,7)	$\frac{5}{56}$
15	(1,6,8), (2,5,8), (2,6,7), (3,5,7), (3,4,8), (4,5,6)	$\frac{6}{56}$
14	(1,6,7), (1,5,8), (2,5,7), (2,4,8), (3,4,7), (3,5,6)	$\frac{6}{56}$
13	(1,4,8), (1,5,7), (2,3,8), (2,4,7), (2,5,6), (3,4,6)	$\frac{6}{56}$
12	(1,3,8), (1,4,7), (1,5,6), (2,3,7), (2,4,6), (3,4,5)	$\frac{6}{56}$
11	(1,2,8), (1,3,7), (1,4,6), (2,3,6), (2,4,5)	$\frac{5}{56}$
10	(1,2,7), (1,3,6), (1,4,5), (2,3,5)	$\frac{4}{56}$
9	(1,2,6), (1,3,5), (2,3,4)	$\frac{3}{56}$
8	(1,2,5), (1,3,4)	$\frac{2}{56}$
7	(1,2,4)	$\frac{1}{56}$
6	(1,2,3)	$\frac{1}{56}$

2 Wilcoxon-Rangsummentest (approximativer Permutationstest)

Für große Stichprobenumfänge kann es sehr aufwendig werden, die exakte Verteilung einer Permutationsstatistik unter H_0 analog zu Abschnitt 1 zu bestimmen. Hier kann es hilfreich sein, die Verteilung der Permutationsstatistik mittels einer Monte-Carlo-Simulation zu approximieren. Dafür geht man für den Wilcoxon-Rangsummentest wie folgt vor:

1. Permutiere den Beobachtungsvektor (bzw. permutiere den Rangvektor).
2. Berechne die Teststatistik.
3. Wiederhole die Schritte 1 und 2 insgesamt n-mal und bestimme die empirische Verteilungsfunktion der Teststatistik (wähle zum Beispiel $n = 10000$).
4. Vergleiche den Wert der Teststatistik (aus den Originaldaten) mit dem $(1 - \alpha)$-Quantil

der empirischen Verteilungsfunktion.

Vergleiche Good (2005, Kapitel 1.3 und Kapitel 14.2) und Kunert (2002).

3 χ^2-Test für Kontingenztafeln (approximativer Permutationstest)

In diesem Abschnitt wird der Permutationstest auf die Situation, dass eine Kontingenztafel vorliegt, übertragen. Es ist somit zu untersuchen, wie plausibel eine gegebene Kontingenztafel ist unter der Annahme, dass die Nullhypothese wahr ist. Das Vorgehen ist sowohl für den Unabhängigkeitstest als auch für den Homogenitätstest geeignet. Zu bestimmen ist die Verteilung der Teststatistik

$$V = \sum_{i=1}^{m} \sum_{j=1}^{q} \frac{(n_{ij} - \tilde{n}_{ij})^2}{\tilde{n}_{ij}},$$

vergleiche Genschel und Becker (2005, Seite 303). Diese Statistik nimmt somit im Folgenden die Rolle von W_N aus Abschnitt 2 ein. Zunächst wird angenommen, dass die Randverteilungen gegeben sind. Die vorliegende Kontingenztafel wird nun als Realisierung eines Zufallsvektors aufgefasst. Für eine gegebene (3x2)-Kontingenztafel (vgl. Tabelle 3) resultiert beispielhaft der Vektor

$$b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Dabei wird jeder Erfolg mit einer "1" kodiert, jeder Misserfolg mit einer "0". Der Vektor b enthält zunächst n_{11}-mal den Eintrag "0", dann n_{12}-mal den Eintrag "1", n_{21}-mal den Eintrag "0", n_{22}-mal den Eintrag "1", n_{31}-mal den Eintrag "0" und n_{32}-mal den Eintrag "1". Die ersten $n_{11} + n_{12}$ Einträge charakterisieren somit die Verteilung der Gruppe 1, die darauf folgenden $n_{21} + n_{22}$ Einträge die Verteilung der Gruppe 2 und die letzten $n_{31} + n_{32}$ Einträge beschreiben die Verteilung der Gruppe 3. Zudem enthält der Vektor b insgesamt $n_{\cdot 1}$ mal den Eintrag "0" und $n_{\cdot 2}$ mal den Eintrag "1". Das weitere Vorgehen entspricht dem Vorgehen aus Abschnitt 2. Das heißt, die Einträge des Vektors b werden permutiert und für jede Permutation wird die zugehörige Kontingenztafel bestimmt, indem ausgezählt wird, wie viele "0" und "1" in den ersten $n_{11} + n_{12}$ Einträgen des neuen Vektors zu finden sind, wie viele "0" und "1" in den nächsten $n_{21} + n_{22}$ Einträgen des neuen Vektors zu finden sind und wie viele "0" und "1" in den letzten $n_{31} + n_{32}$ Einträgen des neuen Vektors zu finden sind. Daraufhin wird die Teststatistik V berechnet. Es folgen schließlich Schritt

3 und 4 aus Abschnitt 2. Dieses Vorgehen garantiert, dass nach jeder Permutation die Randhäufigkeiten gleich bleiben. Abschließend sei angemerkt, dass der Permutationstest für Kontingenztafeln beliebiger Größe möglich ist.

Tabelle 3: *(3x2)-Kontingenztafel*

	Erfolg		
	nein	ja	
Gruppe 1	n_{11}	n_{12}	$n_{1\cdot}$
Gruppe 2	n_{21}	n_{22}	$n_{2\cdot}$
Gruppe 3	n_{31}	n_{32}	$n_{3\cdot}$
	$n_{\cdot 1}$	$n_{\cdot 2}$	n

Literatur

[1] Büning, H. und Trenkler, G. (1994): *Nichtparametrische statistische Methoden*, 2. Auflage, de Gruyter, Berlin.

[2] Genschel, U. und Becker, C. (2005): *Schließende Statistik*, Springer, Berlin.

[3] Good, P. (2005): *Permutation, parametric and bootstrap tests of hypotheses*, 3. Auflage, Springer, New York.

[4] Kunert, J. (2002): *Statistical Methods to Examine Differences in the Rating of Soft-Drinks Among Different Groups of Consumers*. Food Quality and Preference 13, 555 - 559.