

# Der $U$ -Test von Mann & Whitney

## ZIEL

Analog zum  $t$ -Test ist der Zweck des  $U$ -Tests der Vergleich von zwei Stichproben dahingehend, ob sie sich in der Größe der Messwerte signifikant unterscheiden.

## VORAUSSETZUNGEN

Auch dieser Test kommt nicht völlig ohne Voraussetzungen aus. Im Gegensatz zum  $t$ -Test sind es aber relativ schwache Bedingungen:

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

## VORGEHENSWEISE

### *Ermitteln von $u$ und $u'$*

Alle Messwerte werden geordnet. Man überprüft, wie stark sich die beiden Gruppen dabei vermischen, indem man  $u$ , die Anzahl der Permutationen (Vertauschungen) verglichen mit dem Extremfall (alle Werte  $a$  der ersten Stichprobe sind kleiner als alle Werte  $b$  der zweiten) ermittelt.

### *Beispiel:*

$a \dots$  Werte der 1. Stichprobe,  $n_1 = 4$

$b \dots$  Werte der 2. Stichprobe,  $n_2 = 3$

$a a a a b b b : u = 0$

$a a b a a b b : u = 2$

$a a b b a b a : u = 5$

Um von der Ausgangssituation (oben) auf den mittleren Fall zu kommen, sind 2 Vertauschungen  $ab \rightarrow ba$  notwendig (vertauscht werden nur nebeneinanderliegende Werte), für den unteren sind es schon 5.

Um etwas einfacher auf die Zahl der Permutationen zu kommen, kann man für jedes  $a$  abzählen, wieviele  $b$ s auf der „falschen“ (hier: linken) Seite liegen und diese Anzahlen addieren (oder umgekehrt: für jedes  $b$  die rechts davon liegenden  $a$ s zusammenzählen).

Der der obigen Ausgangssituation entgegengesetzte Fall ist, dass alle  $a$  größer als alle  $b$  sind. Die Anzahl der Permutationen, um vom einen Extremfall zum anderen zu kommen, ist genau das Produkt der Stichprobenumfänge.

$$b b b a a a a: u = 4 \cdot 3 = 12$$

Bezeichnet man die Anzahl der Permutationen, um die sich die vorliegende Anordnung von diesem Extremfall unterscheidet, mit  $u'$ , so ergeben sich für unsere Beispiele folgende Werte:

	$u$	$u'$
$a \ a \ a \ a \ b \ b \ b$	0	12
$a \ a \ b \ a \ a \ b \ b$	2	10
$a \ a \ b \ b \ a \ b \ a$	5	7
$b \ b \ b \ a \ a \ a \ a$	12	0

***vereinfachte Berechnung von  $u$  und  $u'$***

Weist man jedem Wert die Zahl zu, die seiner Position entspricht und addiert diese für beide Stichproben, so erhält man die sog. Rangsummen  $r_1$  und  $r_2$ :

1	2	3	4	5	6	7	$r_1$	$r_2$
$a$	$a$	$a$	$a$	$b$	$b$	$b$	10	18
$a$	$a$	$b$	$a$	$a$	$b$	$b$	12	16
$a$	$a$	$b$	$b$	$a$	$b$	$a$	15	13
$b$	$b$	$b$	$a$	$a$	$a$	$a$	22	6

Durch folgende Transformationen kann man aus  $r_2$ ,  $n_1$  und  $n_2$   $u$  bzw. aus  $r_1$ ,  $n_1$  und  $n_2$   $u'$  berechnen:

$$u = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - r_2$$

$$u' = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - r_1$$

### ***Prüfung auf Signifikanz***

Beim  $U$ -Test gibt es zwei Methoden, die Prüfgröße  $u$  (bzw.  $u'$ ) auf Signifikanz zu testen: Für kleine Stichproben ( $n_1 \leq 10$  und  $n_2 \leq 10$ ) sind die kritischen Werte tabelliert und die exakte Vorgehensweise kann angewendet werden. Falls zumindest eine der Stichproben zu groß ist, wird mit NV-Approximation gearbeitet.

#### ***(A) exakte Methode***

Der Signifikanztest besteht wie bei den bisher gelernten Verfahren in einem Vergleich mit dem passenden kritischen Wert aus der Tabelle. Die Prüfgröße  $u_{min}$  ist hierbei der kleinere der beiden Werte  $u$  und  $u'$ :

*Formal:*

$$u_{min} = \min(u, u')$$

**VORSICHT:** Der Unterschied zwischen den beiden Stichproben ist umso größer, je kleiner  $u_{min}$  ist und umgekehrt. Deshalb ist das Ergebnis signifikant, falls  $u_{min} \leq U_{krit.}$ ! (Im Gegensatz dazu sind die meisten Tests signifikant, wenn die Prüfgröße größer als der kritische Wert ist.)

### **(B) NV-Approximation**

Für  $n_1 \rightarrow \infty$  oder  $n_2 \rightarrow \infty$  ist  $U$  asymptotisch normalverteilt mit  $E(U) = \frac{n_1 n_2}{2}$  und  $\sigma^2(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$ .

Dementsprechend erhält man durch Standardisierung von  $u$  eine (asymptotisch) standardnormalverteilte Testgröße:

$$z = \frac{u - E(U)}{\sigma(U)} = \frac{u - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Diese ist mit dem kritischen  $z$ -Wert (am besten aus der  $t$ -Tabelle bei  $df = \infty$  abzulesen) zu vergleichen.

### **Bindungen**

Da vorausgesetzt wird, dass den ermittelten Daten ein stetiges Merkmal zugrunde liegt, dürften keine identischen Messwerte („Bindungen“) auftreten. In der Praxis ist dies aber aus verschiedenen Gründen nicht immer vermeidbar. Falls Bindungen vorhanden sind, werden den Werten die mittleren Rangzahlen zugewiesen.

Im Fall der NV-Approximation muss außerdem die Varianz korrigiert werden, indem statt der oben angegebenen folgende Formel verwendet wird:

$$\sigma_{\text{kor.}}^2(U) = \frac{n_1 n_2}{12N(N-1)} [N^3 - N - \sum_{l=1}^k (t_l^3 - t_l)]$$

*Erläuterungen zu dieser Formel:*

$N = n_1 + n_2 \dots$  Gesamtzahl der Messwerte

$k \dots$  Anzahl der Bindungen

$t_l \dots$  Anzahl der Werte innerhalb der  $l$ -ten Bindung

*Beispiel für Bindungen:*

$A : 8, 10, 10, 11, 11, 15$

$B : 9, 10, 11, 11, 13, 15$

$N = 12$

$k = 3$  (3 Bindungen: 10, 11 und 15)

$t_1 = 3, t_2 = 4, t_3 = 2$  (denn:  $3 \times 10, 4 \times 11$  und  $2 \times 15$ )