

PROC NPAR1WAY

zum Durchführen des U-Tests für zwei unverbundene Stichproben
(für **quantitative nicht-normalverteilte** Merkmale)

Allgemeine Form:

```
PROC NPAR1WAY DATA=name Optionen ;  
  VAR variablenliste ;  
  CLASS vergleichsvariable ;  
RUN ;
```

Beispiel und Beschreibung der Programm-Statements:

```
TITLE1 'U-Test für zwei unverbundene Stichproben nach Rauchen' ;  
TITLE2 '-----' ;  
PROC NPAR1WAY DATA=beispiel WILCOXON ;  
  VAR gewicht ;  
  CLASS rauchen ;  
RUN ;
```

Die Prozedur beginnt mit PROC endet mit RUN ;

Das ; am Ende eines Befehls beendet eine Anweisung innerhalb der Prozedur.

NPAR1WAY ist der Name der Prozedur.

DATA = beispiel legt das zu verwendende Datenfile fest (im Beispiel **beispiel**).

VAR ist das gehörende **Schlüsselwort** und steht für die Liste der Merkmalsnamen.

Ohne **VAR** - Statement berechnet SAS die statistischen Maßzahlen für **alle numerischen** Merkmale in der Datendatei.

Die Aufzählung der Merkmalsnamen darf in einem **VAR**-Statement auch über mehrere Zeilen gehen. Das Semikolon kommt dann nach der letzten Variablen.

variablenliste wird mit den Namen der zu verarbeitenden Merkmale überschrieben. Trennen Sie die **Merkmalsnamen** mit **Leerzeichen!!!**

Mit der Option **WILCOXON** bewirken Sie, dass SAS den U-Test nach Mann-Whitney-Wilcoxon auf die Daten anwendet.

Mit dem **CLASS**-Statement definieren Sie **eine** Gruppenvariable (Vergleichsmerkmal). SAS vergleicht die quantitativen Merkmale in der Variablenliste nach den Ausprägungen dieses Merkmals. **Die Gruppenvariable muss qualitativ sein.**

vergleichsvariable ist der Name des **qualitativen** Merkmals, nach dessen Ausprägungen der Vergleich stattfinden soll.

Mit dem `TITLE`-Befehl stellen Sie den Ergebnissen Überschriften voran. Der Befehl ist nicht zwingend notwendig, bringt aber Struktur in die Auswertung. **Achtung!** - SAS übernimmt Titel in nachfolgende Prozeduren, wenn dort das `TITLE`-Statement fehlt.

SAS gibt die berechneten (Sum of Scores) und erwarteten Rangsummen (Expected under H0), die erwartete Standardabweichung (Std Dev Under H0), die mittleren Rangsummen (Mean Score), die standardisierte Werte für den Rangsummentest (Wilcoxon Two-Sample Test) und das Testergebnis für den U-Test aus (unter Kruskal-Wallis-Test).

Starten Sie die Programmzeilen aus dem Beispiel, zeigt SAS im Output-Fenster folgende Informationen:

U-Test fuer zwei unverbundene Stichproben nach Rauchen

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable **Gewicht**
Classified by Variable **rauchen**

rauchen	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
ja	25	1032.0	887.50	81.503725	41.280000
nein	45	1453.0	1597.50	81.503725	32.288889

Average scores were used for ties.

Wilcoxon Two-Sample Test

Statistic **1032.0000**

Normal Approximation

Z **1.7668**
One-Sided Pr > Z **0.0386**
Two-Sided Pr > |Z| **0.0773**

t Approximation

One-Sided Pr > Z **0.0408**
Two-Sided Pr > |Z| **0.0817**

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square **3.1433**
DF **1**
Pr > Chi-Square **0.0762**

Beschreibung des Outputs:

Die berechneten Werte beziehen sich auf die Gesamtzahl der Beobachtungen abzüglich der Datensätze mit fehlenden Werten.

Variable zeigt die Namen der verwendeten Merkmale, im Beispiel: **Gewicht**. **rauchen** ist der Name der Vergleichsvariablen, SAS gibt hier die Ausprägungen **ja** und **nein** aus.

N = Anzahl der Beobachtungen: 25 Raucher, 45 Nichtraucher

Sum of Scores = berechneten Rangsummen.

Sie werden berechnet nach der Formel $((1 + n_1 + n_2) / 2) \cdot n_1$, wobei Sie für n_1 die Fallzahl der einen und für n_2 die Fallzahl der anderen Gruppe in die Formel einsetzen.

ja	nein
1032,0	1453,0

Expected Under H0 = erwartete Rangsummen.

Sie werden berechnet nach der Formel $((1 + n_1 + n_2) \cdot n_1) / 2$ bzw. $((1 + n_1 + n_2) \cdot n_2) / 2$, wobei Sie für n_1 die Fallzahl der einen und für n_2 die Fallzahl der anderen Gruppe in die Formel einsetzen.

ja	nein
887,50	1597,50

Std Dev Under H0 = erwartete Standardabweichung.

Standardabweichung der Rangsummen bei der Nullhypothese.

ja	nein
81,503725	81,503725

Mean Score = mittlere Rangsummen.

Mittelwerte der Rangsummen nicht der Originalwerte!!!

ja	nein
81,503725	81,503725

Average scores were used for ties.

Haben zwei oder gar mehr Werte den gleichen Rang werden die Rangzahlen addiert und durch die Anzahl der Werte geteilt. Dies wird als Bindung bezeichnet.

Wilcoxon Two-Sample Test.

Gibt standardisierte Werte für den Rangsummentest aus: SAS nähert die Verteilung der vorliegenden Daten einer Normalverteilung (Normal Approximation) bzw. t-Verteilung (t Approximation) an. Da die durchgeführten Berechnungen nicht dem eigentlichen Wilcoxon-Rangsummen-Test entsprechen, kann dieser **Ergebnisteil ignoriert werden!!!**

Testgröße	Normalverteilung		t-Verteilung	
	einseitig p-Wert	zweiseitig p-Wert	einseitig p-Wert	zweiseitig p-Wert
1,7668	0,0386	0,0773	0,0408	0,0817

Kruskal-Wallis Test.

Der Kruskal-Wallis-Test stellt die Verallgemeinerung des Wilcoxon-Rangsummen-Tests beim Vergleich von mehr als zwei Gruppen dar und ist beim Vergleich zweier Gruppen gleichwertig zum U-Test. DF ist die Abkürzung für degrees of freedom und bezeichnet die Freiheitsgrade.

Testgröße	p-Wert
3,1433	0,0762

Beschreibung eines weiteren nützlichen Statements:

- **Statement BY:** Unterscheidet die Berechnungen nach den Ausprägungen einer weiteren qualitativen Gruppenvariablen.

```
TITLE1 'U-Test für zwei unverbundene Stichproben nach Rauchen' ;
TITLE2 '-----' ;
PROC NPAR1WAY DATA=beispiel WILCOXON ;
  BY geschlecht ;
  VAR gewicht ;
  CLASS rauchen ;
RUN ;
```

- ⇒ SAS gibt Rangsummen und Teststatistiken getrennt nach Rauchern und Nichtrauchern aus:

U-Test fuer zwei unverbundene Stichproben nach Rauchen

Geschlecht=maennlich

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable **gewicht**
Classified by Variable **rauchen**

rauchen	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
ja	11	149.50	132.0	16.215934	13.590909
nein	12	126.50	144.0	16.215934	10.541667

Average scores were used for ties.

Wilcoxon Two-Sample Test

Statistic 149.5000

Normal Approximation

Z 1.0484
One-Sided Pr > Z 0.1472
Two-Sided Pr > |Z| 0.2945

t Approximation

One-Sided Pr > Z 0.1529
Two-Sided Pr > |Z| 0.3059

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square 1.1646
DF 1
Pr > Chi-Square 0.2805

U-Test fuer zwei unverbundene Stichproben nach Rauchen

Geschlecht=weiblich

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable **Gewicht**
Classified by Variable **rauchen**

rauchen	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
nein	33	767.50	792.0	42.913742	23.257576
ja	14	360.50	336.0	42.913742	25.750000

Average scores were used for ties.

Wilcoxon Two-Sample Test

Statistic **360.5000**

Normal Approximation

Z **0.5593**
One-Sided Pr > Z **0.2880**
Two-Sided Pr > |Z| **0.5760**

t Approximation

One-Sided Pr > Z **0.2893**
Two-Sided Pr > |Z| **0.5787**

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square **0.3259**
DF **1**
Pr > Chi-Square **0.5681**