# Grundlagen der Biostatistik und Informatik

## Hypothesenprüfungen III.
## Nichtparametrische Methoden

dr László Smeller
Semmelweis Universität
2012

## Übersicht der Teste

| Verteilung / Stichproben | Normalverteilte Daten | Die Verteilung der Daten ist unbekannt |
|---|---|---|
| Eine Stichprobe | Einstichproben t-Test | Vorzeichnentest Wilcoxon Test |
| Zwei Stichproben | Zweistichproben t-test | Mann-Whittney U-Test |
| Mehrere Stichproben | ANOVA (Varianzanalyse) | Kruskal-Wallis Test |



## Nichtparametrische Methoden

Als Erinnerung: Bedingungen der t-Teste:
- kontinuierliches Merkmal (z.B. Körperhöhe, Körpertemperatur...)
- die Daten müssen eine Normalverteilung folgen!

Was passiert, wenn die Vorbedingungen nicht gelten?

- nur ordinale Daten (Ordinalskala)
- keine Normalverteilung

→ **Nichtparametrische Methoden**

z. B. Schmerzmittel – Wie es schmerzt? Kann nur auf einer ordinalen Skala gemessen werden:

1, 2, 3, 4, 5 oder



## Nichtparametrische Metoden

**Vorteile:**
- Verteilungsunabhängigkeit
- Ordinal-, Intervall-, Verhältnisskalen

**Nachteile:**
- Datenreduktion, Informationsverlust
- größere Wahrscheinlichkeit der Fehler 2. Art
- Nur größere Unterschiede können detektiert werden als bei den parametrischen Teste

# Eine Stichprobe: Vorzeichentest

Datenpaaren, d.h. gepaarter Test (Änderung oder Unterschied)

Datenreduktion:

Alternativen (einander ausschließende Ereignisse)

z.B.: Verbesserung oder Verschlechterung des Krankheitszustandes

Erfolg- Misserfolg

$\Rightarrow$ **Binomialverteilung**

Hat das Medikament eine Wirkung? D.h.: Sind signifikant mehr Fallen mit Verbesserung als mit Verschlechterung? Haben die zwei Alternativen unterschiedliche Wahrscheinlichkeit?

$H_0$: Die zwei Alternativen haben dieselbe Wahrscheinlichkeit.

Analogie: Münzenexperiment: Kopf oder Zahl



# Vorzeichentest: Beispiel des Kopfschmerzes

Kopfschmerzen vor und nach der Einnahme des Medikamentes werden an einer relativen Skala gegeben.

Der Kopfschmerz erniedrigt in ***k*** aus ***n*** Fällen (in ***n-k*** Fällen es erhöht sich. Die „keine Änderung" Fällen werden nicht beachtet.)

Ist die Änderung signifikant?

$H_0$: Das Medikament ist unwirksam, d.h. Erniedrigung und Erhöhung des Kopfschmerzes sind gleich wahrscheinlich.

Beispiel1.: Kopfschmerzen sinkt bei 9 aus 10 Patienten.

Beispiel2.: Kopfschmerzen sinkt bei 7 aus 9 Patienten.

Bei Gültigkeit der Nullhypothese: Analogie mit dem Münzenexperiment:

Experiment: *k*-mal Kopf aus *n* Versuche.

Analogie für Beispiel 1.: 9-mal Kopf aus 10 Versuche

Analogie für Beispiel2.: 7-mal Kopf aus 9 Versuche



# Vorzeichentest: Analogie mit dem Münzenexperiment

Bei Münzenexperiment kann man die Wahrscheinlichkeit der unterschiedlichen Fällen ausrechnen (Binomialverteilung!):

Anzahl von Experimenten

| Anzahl von Kopf | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 50.0% | 25.0% | 12.5% | 6.3% | 3.1% | 1.6% | 0.8% | 0.4% | 0.2% | 0.1% |
| 1 | 50.0% | 50.0% | 37.5% | 25.0% | 15.6% | 9.4% | 5.5% | 3.1% | 1.8% | 1.0% |
| 2 | | 25.0% | 37.5% | 37.5% | 31.3% | 23.4% | 16.4% | 10.9% | 7.0% | 4.4% |
| 3 | | | 12.5% | 25.0% | 31.3% | 31.3% | 27.3% | 21.9% | 16.4% | 11.7% |
| 4 | | | | 6.3% | 15.6% | 23.4% | 27.3% | 27.3% | 24.6% | 20.5% |
| 5 | | | | | 3.1% | 9.4% | 16.4% | 21.9% | 24.6% | 24.6% |
| 6 | | | | | | 1.6% | 5.5% | 10.9% | 16.4% | 20.5% |
| 7 | | | | | | | 0.8% | 3.1% | 7.0% | 11.7% |
| 8 | | | | | | | | 0.4% | 1.8% | 4.4% |
| 9 | | | | | | | | | 0.2% | 1.0% |
| 10 | | | | | | | | | | 0.1% |

Binomialverteilung

$$p_k = \binom{n}{k} p^k q^{n-k}$$



# Vorzeichentest: Anwendung der Binomialverteilung

Bei Gültigkeit der $H_0$ gibt dieselbe Tabelle die Wahrscheinlichkeit der entsprechenden Fällen:

Anzahl von Patienten

| Anzahl der Verbesserungen | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 50.0% | 25.0% | 12.5% | 6.3% | 3.1% | 1.6% | 0.8% | 0.4% | 0.2% | 0.1% |
| 1 | 50.0% | 50.0% | 37.5% | 25.0% | 15.6% | 9.4% | 5.5% | 3.1% | 1.8% | 1.0% |
| 2 | | 25.0% | 37.5% | 37.5% | 31.3% | 23.4% | 16.4% | 10.9% | 7.0% | 4.4% |
| 3 | | | 12.5% | 25.0% | 31.3% | 31.3% | 27.3% | 21.9% | 16.4% | 11.7% |
| 4 | | | | 6.3% | 15.6% | 23.4% | 27.3% | 27.3% | 24.6% | 20.5% |
| 5 | | | | | 3.1% | 9.4% | 16.4% | 21.9% | 24.6% | 24.6% |
| 6 | | | | | | 1.6% | 5.5% | 10.9% | 16.4% | 20.5% |
| 7 | | | | | | | 0.8% | 3.1% | 7.0% | 11.7% |
| 8 | | | | | | | | 0.4% | 1.8% | 4.4% |
| 9 | | | | | | | | | 0.2% | 1.0% |
| 10 | | | | | | | | | | 0.1% |

Binomialverteilung

$$p_k = \binom{n}{k} p^k q^{n-k}$$

Irrtumswahrscheinlichkeit=5% (2,5%+2,5%)

## Vorzeichentest: Beispiel 2, Überlebenszeit

Überlebenszeit bei behandelten Ratten mit einem Tumor. (Tage)

168, 190, 280, 221, 110, 165, 179, 250, 195, 276

Es ist bekannt dass die Überlebenszeit der nicht behandelten Ratten mit dieser Tumorart 170 Tage beträgt (Median!)

Kann der Median der Überlebenszeiten der behandelten Ratten 170 Tage sein?

$H_0$: Median der Überlebenszeiten der behandelten Ratten beträgt 170 Tage.

168, 150, 280, 221, 230, 165, 179, 250, 195, 276

\-   -   +   +   +   -   +   +   +   +

Bei Gültigkeit der $H_0$ die Daten sind >170 Tage zu 50% Wahrsch.

< 170 Tage zu 50% Wahrsch.



## Vorzeichentest: Anwendung der Binomialverteilung

Bei Gültigkeit der $H_0$ gibt diese Tabelle die Wahrscheinlichkeit der entsprechenden Fällen:

| Anzahl der Ratten mit Überelebenszeit > 170 Tage | Anzahl von Ratten: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 50.0% | 25.0% | 12.5% | 6.3% | 3.1% | 1.6% | 0.8% | 0.4% | 0.2% | 0.1% |
| 1 | 50.0% | 50.0% | 37.5% | 25.0% | 15.6% | 9.4% | 5.5% | 3.1% | 1.8% | 1.0% |
| 2 | | 25.0% | 37.5% | 37.5% | 31.3% | 23.4% | 16.4% | 10.9% | 7.0% | 4.4% |
| 3 | | | 12.5% | 25.0% | 31.3% | 31.3% | 27.3% | 21.9% | 16.4% | 11.7% |
| 4 | | | | 6.3% | 15.6% | 23.4% | 27.3% | 27.3% | 24.6% | 20.5% |
| 5 | | | | | 3.1% | 9.4% | 16.4% | 21.9% | 24.6% | 24.6% |
| 6 | | | | | | 1.6% | 5.5% | 10.9% | 16.4% | 20.5% |
| 7 | | | | | | | 0.8% | 3.1% | 7.0% | 11.7% |
| 8 | | | | | | | | 0.4% | 1.8% | 4.4% |
| 9 | | | | | | | | | 0.2% | 1.0% |
| 10 | | | | | | | | | | 0.1% |

Binomialverteilung

$$p_k = \binom{n}{k} p^k q^{n-k}$$

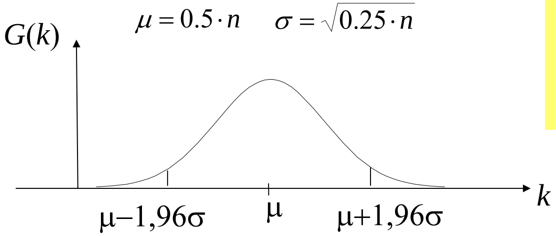
Irrtumswahrscheinlichkeit=5% (2,5%+2,5%)



## Vorzeichentest: Annäherung bei n > 20

Annäherung bei $n > 20$ mit Normalverteilung:

$\mu = 0.5 \cdot n \qquad \sigma = \sqrt{0.25 \cdot n}$




Siehe: Binomial-verteilung:

$\boldsymbol{\mu = p \cdot n}$

$\boldsymbol{\sigma = \sqrt{p \cdot q \cdot n}}$

z. B. 100 Patienten, 56 Verbesserungen, 34 Verschlechterungen

$H_0$: ?; $\mu_0$=? $\mu$ = ?; $\sigma$ = ?; Entscheidung?

Analogie zu Einstichproben t-Test

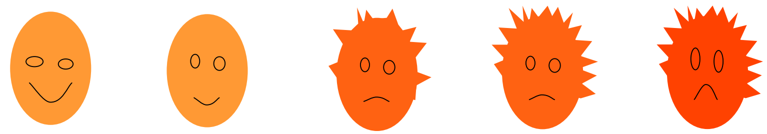
(Lösung: 56+34=90 μ=45 σ =Wurzel(90/4)=4,74 μ+1,96*σ=45+9,3=54,3 <56 => signifikant (5% Irrt.w.)!



## Prinzip der Rang Teste

Rang: Position eines Wertes innerhalb einer nach der Größe sortierten Wertereihe

z.B. Kopfschmerzen:



1 2 3 4 5

Mit Hilfe der Ränge führt man eine Gleichverteilung ein!

# Rang Test Methode – Gekoppelte Ränge

Wenn zwei oder mehrere ursprüngliche Daten gleich sind:

| | |
|---|---|
| originale Daten | 3, 7, 1, 13, 13, 16 |
| geordnete Daten | 1, 3, 7, 13, 13, 16 |
| Ränge | 1, 2, 3, 4.5, 4.5, 6 |

gekoppelte Ränge:

die bekommen den Durschnittsrang



# Durchschnitt der Ränge

| | |
|---|---|
| In steigende Reihe geordnete Daten: | $x_1$ , $x_2$ … $x_{(n-1)/2}$ , $\boxed{x_{(n+1)/2}}$ …. $x_{n-1}$ ,$x_n$ |
| Ränge: | 1, 2 … (n-1)/2 (n+1)/2 … n-1, n |

(n ist ungerade)

Durchschnitt der Ränge: $\bar{R} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$

Durchschnittlicher Rang = Rang des Medians

Wenn n ist gerade:

Median= $(x_{n/2} + x_{n/2+1})/2$

Durchschnittlicher Rang= (n+1)/2

Rangteste testen den Median!



# Eine Stichprobe: Wilcoxon-Vorzeichen Rangtest

Eine Stichprobe (Gepaarte Test)

Ordinale Daten

Ist der Median der Datenreihe gleich Null (oder ein bestimmter Wert)?

$H_0$: Der Median der Daten ist Null (oder ein bestimmter Wert).

Die Ränge bekommen ein Vorzeihen.

Der Durchschnitt der Ränge wird geprüft.

Wenn die Nullhypothese gültig ist, es sind gleich viele und gleich große positive und negative Ränge, Durchschnitt der Rängre ist Null!



# Wilcoxon-Vorzeichen Rangtest: Einführung mit einem Beispiel

Überlebenszeit der Ratten:

168, 150, 280, 221, 230, 165, 179, 250, 195, 276

Ist der Median der Überlebenszeiten unterschiedlich von 170 Tage?

$H_0$: Der Median der Überlebenszeiten beträgt 170 Tage.

**Überlebenszeitenunterschiede der Ratten im Vergleich zur 170 Tage:**

**-2, -20, +110, +51, +60, -5, +9, +80, +25, +106**

**Geordnet nach Betrag der Änderung:**

**-2, -5, +9, -20, +25, +51, +60, +80, +106, +110,**

**Ränge (nach betrag der Änderung):**

**1, 2, 3, 4, 5, 6, 7, 8, 9, 10**

**Ränge mit Vorzeichnen:**

**-1, -2, +3, -4, +5, +6, +7, +8, +9, +10**

Durchschnitt: 4.10

Standardabw.: 4.91

## Wilcoxon Vorzeichen Rangtest: Beispiel der Überlebenszeiten der Ratten

Der Durchschnitt folgt einer Normalverteilung, wenn genug viele Daten sind (Zentraler Grenzwertsatz)

Anwendung der t-Verteilung (Annäherung!):

$$\boldsymbol{t}_{\boldsymbol{n}-1} = \frac{\bar{\boldsymbol{R}}}{\boldsymbol{s} / \sqrt{\boldsymbol{n}}}$$

Durchschnitt der Ränge

Standardabweichung der Ränge

Anzahl der Daten

Freiheitsgrad

Entscheidung: wie beim Einstichproben ***t***-Test

**Ränge mit Vorzeichnen:**

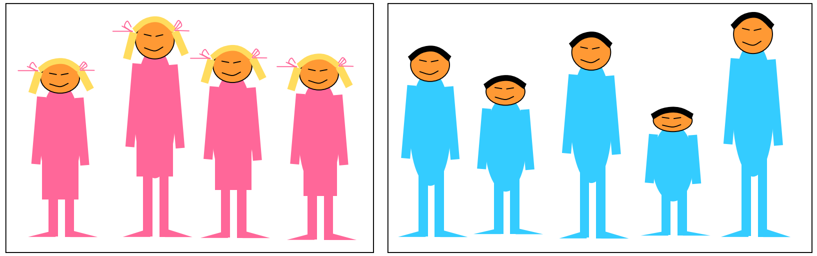
**-1, -2, +3, -4, +5, +6, +7, +8, +9, +10** $\longrightarrow$ Durchschnitt: 4.10 Standardabw.: 4.91

$$t_9 = \frac{4{,}10}{4{,}91/\sqrt{10}} = 2{,}64 \quad \Rightarrow \quad t_9 > t_{9;5\%} \quad \Rightarrow \quad H_0 \text{ is abgelehnt}$$

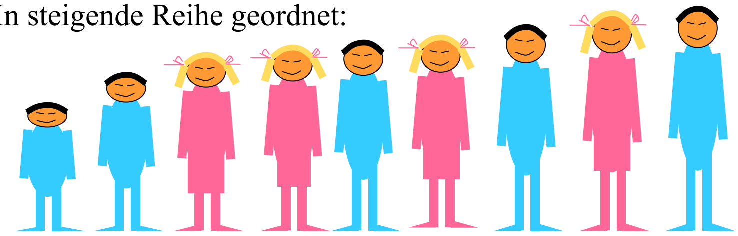
$t_{9;5\%}$=2,26 (aus der Tabelle) $p < 5\%$ (mit Excel)



## Vergleich von zwei Stichproben



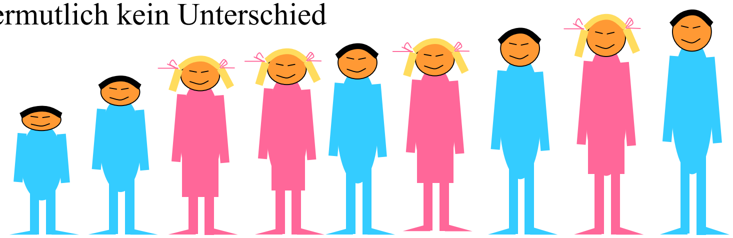
In steigende Reihe geordnet:



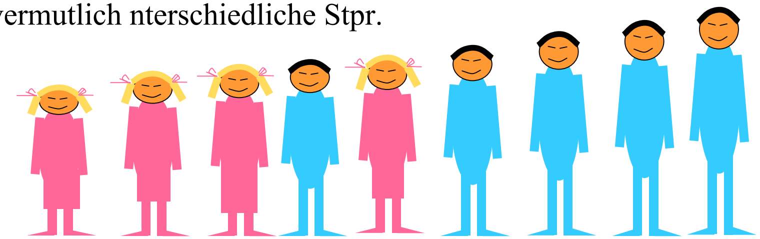
Ränge: 1 2 3 4 5 6 7 8 9



vermutlich kein Unterschied



vermutlich nterschiedliche Stpr.



1 2 3 4 5 6 7 8 9



## Mann – Whitney *U* Test

(Auch als Wilcoxon Rank Summe Test genannt)

Vergleich von zwei Stichproben ($n_1$, $n_2$)

$H_0$: Die zwei Stichproben stammen aus der selben Grundgesamtheit

Für klein $n_1$, $n_2$:

1. Zuordnung der Ränge der in den zwei zusammengeordneten Stichproben.
2. Bestimmung die Summen der Ränge in beiden Gruppen ($T_1$,$T_2$).
3. $\boldsymbol{U}_1 = \boldsymbol{n}_1\boldsymbol{n}_2 + \frac{\boldsymbol{n}_1(\boldsymbol{n}_1+1)}{2} - \boldsymbol{T}_1 \quad \boldsymbol{U}_2 = \boldsymbol{n}_1\boldsymbol{n}_2 + \frac{\boldsymbol{n}_2(\boldsymbol{n}_2+1)}{2} - \boldsymbol{T}_2$
4. Bestimmung des ***U*** wertes: Min{$U_1U_2$}
5. Vergleich von $U$ mit $U_{p,n1,n2}$ (Aus der *U*-Tabelle)

$U < U_{p,n1,n2}$ $H_0$ ablehnen $U > U_{p,n1,n2}$ $H_0$ nicht ablehnen;

# Mann - Whitney $U$ Test: Beispiel

Idee: Wenn die zwei Stichproben aus der selben Grundgesamtheit stammen, die Ränge verteilen sich zufällig.

Beispiel: Zeit der Tumorausbildung (Wochen) in zwei Gruppen der unterschiedlich behandelten Ratten.

Daten

| | |
|---|---|
| 1. Stichprobe | **11, 12, 19, 15, 17** |
| 2. Stichprobe | **18, 14, 13, 21, 16** |

| | |
|---|---|
| Geordnete Daten | **11, 12, 13, 14, 15, 16, 17, 18, 19, 21** |
| Ränge | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 |



# Mann – Whitney $U$ Test

| | 1. Stichprobe | 2. Stichprobe |
|---|---|---|
| Elemente | **11, 12, 19, 15, 17** | **18, 14, 13, 21, 16** |
| Ränge | 1, 2, 9, 5, 7 | 8, 4, 3, 10 , 6 |
| Summe der Ränge | **24** | **31** |

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_1 \quad U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_2$$

$U_1$=5x5+5x6/2-24=16 $U_2$=5x5+5x6/2-31=9

$U_1$=16 $U_2$=9 $\Rightarrow$ $U = 9,$



# Die U-Tabelle

**Tabelle.** Kritische Werte von $U$ für den Test von Wilcoxon, Mann und Whitney für den einseitigen Test: $\alpha = 0{,}025$; zweiseitigen Test: $\alpha = 0{,}05$

| m \ n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | - | | | | | | | | | | | | | | | | | | | |
| 2 | - | - | | | | | | | | | | | | | | | | | | |
| 3 | - | - | - | | | | | | | | | | | | | | | | | |
| 4 | - | - | - | 0 | | | | | | | | | | | | | | | | |
| 5 | - | - | 0 | 1 | 2 | | | | | | | | | | | | | | | |
| 6 | - | - | 1 | 2 | 3 | 5 | | | | | | | | | | | | | | |
| 7 | - | - | 1 | 3 | 5 | 6 | 8 | | | | | | | | | | | | | |
| 8 | - | 0 | 2 | 4 | 6 | 8 | 10 | 13 | | | | | | | | | | | | |
| 9 | - | 0 | 2 | 4 | 7 | 10 | 12 | 15 | 17 | | | | | | | | | | | |
| 10 | - | 0 | 3 | 5 | 8 | 11 | 14 | 17 | 20 | 23 | | | | | | | | | | |
| 11 | - | 0 | 3 | 6 | 9 | 13 | 16 | 19 | 23 | 26 | 30 | | | | | | | | | |
| 12 | - | 1 | 4 | 7 | 11 | 14 | 18 | 22 | 26 | 29 | 33 | 37 | | | | | | | | |
| 13 | - | 1 | 4 | 8 | 12 | 16 | 20 | 24 | 28 | 33 | 37 | 41 | 45 | | | | | | | |
| 14 | - | 1 | 5 | 9 | 13 | 17 | 22 | 26 | 31 | 36 | 40 | 45 | 50 | 55 | | | | | | |
| 15 | - | 1 | 5 | 10 | 14 | 19 | 24 | 29 | 34 | 39 | 44 | 49 | 54 | 59 | 64 | | | | | |
| 16 | - | 1 | 6 | 11 | 15 | 21 | 26 | 31 | 37 | 42 | 47 | 53 | 59 | 64 | 70 | 75 | | | | |
| 17 | - | 2 | 6 | 11 | 17 | 22 | 28 | 34 | 39 | 45 | 51 | 57 | 63 | 69 | 75 | 81 | 87 | | | |
| 18 | - | 2 | 7 | 12 | 18 | 24 | 30 | 36 | 42 | 48 | 55 | 61 | 67 | 74 | 80 | 86 | 93 | 99 | | |
| 19 | - | 2 | 7 | 13 | 19 | 25 | 32 | 38 | 45 | 52 | 58 | 65 | 72 | 78 | 85 | 92 | 99 | 106 | 113 | |
| 20 | - | 2 | 8 | 14 | 20 | 27 | 34 | 41 | 48 | 55 | 62 | 69 | 76 | 83 | 90 | 98 | 105 | 112 | 119 | 127 |
| 21 | - | 3 | 8 | 15 | 22 | 29 | 36 | 43 | 50 | 58 | 65 | 73 | 80 | 88 | 96 | 103 | 111 | 119 | 126 | 134 |
| 22 | - | 3 | 9 | 16 | 23 | 30 | 38 | 45 | 53 | 61 | 69 | 77 | 85 | 93 | 101 | 109 | 117 | 125 | 133 | 141 |
| 23 | - | 3 | 9 | 17 | 24 | 32 | 40 | 48 | 56 | 64 | 73 | 81 | 89 | 98 | 106 | 115 | 123 | 132 | 140 | 149 |
| 24 | - | 3 | 10 | 17 | 25 | 33 | 42 | 50 | 59 | 67 | 76 | 85 | 94 | 102 | 111 | 120 | 129 | 138 | 147 | 156 |
| 25 | - | 3 | 10 | 18 | 27 | 35 | 44 | 53 | 62 | 71 | 80 | 89 | 98 | 107 | 117 | 126 | 135 | 145 | 154 | 163 |
| 26 | - | 4 | 11 | 19 | 28 | 37 | 46 | 55 | 64 | 74 | 83 | 93 | 102 | 112 | 122 | 132 | 141 | 151 | 161 | 171 |
| 27 | - | 4 | 11 | 20 | 29 | 38 | 48 | 57 | 67 | 77 | 87 | 97 | 107 | 117 | 127 | 137 | 147 | 158 | 168 | 178 |
| 28 | - | 4 | 12 | 21 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 101 | 111 | 122 | 132 | 143 | 154 | 164 | 175 | 186 |
| 29 | - | 4 | 13 | 22 | 32 | 42 | 52 | 62 | 73 | 83 | 94 | 105 | 116 | 127 | 138 | 149 | 160 | 171 | 182 | 193 |

# Mann – Whitney $U$ Test

| | 1. Stichprobe | 2. Stichprobe |
|---|---|---|
| Elemente | **11, 12, 19, 15, 17** | **18, 14, 13, 21, 16** |
| Ränge | 1, 2, 9, 5, 7 | 8, 4, 3, 10 , 6 |
| Summe der Ränge | **24** | **31** |

$U_1$=16 $U_2$=9 $U = 9, \; > \; U_{5\%,5,45} = 2$

$\Downarrow$

$H_0$ kann nicht abgelehnt werden

Wenn $n_1$; $n_2$ < 20 **$U$**-Tabelle,

für große $n_1$ und $n_2$ Annäherung



Eine Web-Implementation: http://faculty.vassar.edu/lowry/utest.html

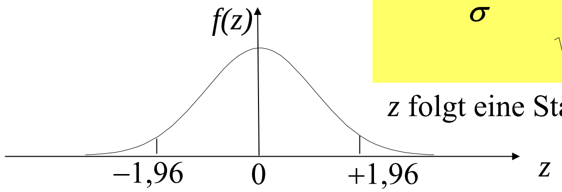
## Mann – Whitney *U* Test: Annäherung

Bei Gültigkeit der Nullhypothese folgen die Daten der Gruppe 1 eine Gleichverteilung, mit möglichen werten von $1\ldots n_1+n_2$)

Erwartungswert und die theoretische Streuung von $T_1$ können berechnet werden:

$$\mu = \frac{n_1(n_1+n_2+1)}{2} \qquad \sigma = \sqrt{n_1 n_2 (n_1+n_2+1)/12}$$

$$z = \frac{T_1-\mu}{\sigma} = \frac{T_1 - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}}$$




*z* folgt eine Standard-Normalverteilung



## Kruskal – Wallis Test

- Vergleich von mehreren Stichproben
- Stichproben sollen unabhängig voneinander ausgewählt werden

1. Bestimmung der Ränge für alle Ergebnisse zusammen
2. Ausrechnung von Rangsummen für alle Stichproben
3. $$H = \left(\frac{12}{N(N+1)} \sum_{j=1}^{h} \frac{R_j^2}{n_j}\right) - 3(N+1)$$

*N*: Anzahl von Messergebnisse     *h*: Anzahl von Stichproben

$R_j$: Rangsumme in der *j*-ten Stichprobe

$n_j$: Anzahl von Messergebnisse in der *j*-sten Stichprobe



## Kruskal – Wallis Test

Korrektion des *H* – wertes in Fall der gekoppelten Ränge:

$E_i = e_i^3 - e_i$

$e_i$: Anzahl der Elemente die in dem *i*-ten Stichprobe gleich sind

$$H_{korr} = \frac{H}{1 - \frac{\sum_{i=1}^{k} E_i}{N^3 - N}}$$



## Kruskal – Wallis Test

Für $n_j < 5 \rightarrow$ Tabelle

$n_j > 5$:

$\chi^2$ Verteilung; Freiheitsgrad: $h$-1

Wenn $H < \chi^2_{krit}$ $H_0$ behalten

Wenn $H > \chi^2_{krit}$ $H_0$ ablehnen; es gibt einen Unterschied (irgendwo) zwischen den Stichproben.

Man kann nicht bestimmen mit dem Kruskal – Wallis Test ob es zwischen welchen Sticproben den Unterschied gibt.

→ Paarweise vergleichen (**Nach dem Kruskal – Wallis Test)**

## Ende der Rangteste

Wo nur die Ränge zählen!

## Bemerkung: Vergleich von Hypothesenprüfungen und Schätzungen

zB.: Blutdrucksenker

Blutdruckänderungen (mmHg):

-13, 5, -29, -22, 13, -8, -19, -12

Durchschnitt: -10,625 mmHg

Standardfehler: 4,917 mmHg

Schätzung: Konfidenintervall:

$\bar{x} \pm 2s_{\bar{x}}$ -10,6±9,8 mmHg -20,4 … – 0,8 mmHg

enthält Null nicht! => Blutdrucksänkender Effekt!

t-Test:

$t$ = -10,625/4,917=-2,161 $|t|<t_{FG=7\,;5\%}=2{,}365$


kein signifikanter Effekt!



## Lösung des Problems: Genaues Konfidenzintervall

$\bar{x} \pm 2s_{\bar{x}}$ ist nur eine grobe Annäherung des Konfidenzintervalles.

Das genaue Konfidenzintervall für 95% Konfidenzniveau ist:

$$\bar{x} \pm t_{n-1;5\%} s_{\bar{x}}$$

Es zählt nur bei kleinen Stichproben (n<20)

In dem Beispiel des Blutdrucksenkers:

$\bar{x} \pm t_{n-1;5\%} s_{\bar{x}}$ (-10,6±2,365•4,917)mmHg =

(-10,6±11,6) mmHg -22,2 … 0,8 mmHg

| FG \ p | 0.05 | 0.02 | 0.01 |
|---|---|---|---|
| 2 | 4.30266 | 6.96455 | 9.92499 |
| 3 | 3.18245 | 4.54071 | 5.84085 |
| 4 | 2.77645 | 3.74694 | 4.60408 |
| 5 | 2.57058 | 3.36493 | 4.03212 |
| 6 | 2.44691 | 3.14267 | 3.70743 |
| 7 | 2.36462 | 2.99795 | 3.49948 |
| 8 | 2.30601 | 2.89647 | 3.35538 |
| 9 | 2.26216 | 2.82143 | 3.24984 |
| 10 | 2.22814 | 2.76377 | 3.16926 |
| 11 | 2.20099 | 2.71808 | 3.10582 |
| 12 | 2.17881 | 2.68099 | 3.05454 |
| 13 | 2.16037 | 2.65030 | 3.01228 |
| 14 | 2.14479 | 2.62449 | 2.97685 |
| 15 | 2.13145 | 2.60248 | 2.94673 |
| 20 | 2.08596 | 2.52798 | 2.84534 |
| 50 | 2.00856 | 2.40327 | 2.67779 |
| 70 | 1.99444 | 2.38080 | 2.64790 |
| 100 | 1.98397 | 2.36421 | 2.62589 |
| unendlich | 1.95996 | 2.32635 | 2.57583 |