



## **Test auf Mittelwert-Differenzen**

**t-Test**  
**Median-Test**  
**Ridit-Test**

Johann Bacher / Kurt Holm

Almo Statistik-System  
[www.almo-statistik.de](http://www.almo-statistik.de)  
[holm@almo-statistik.de](mailto:holm@almo-statistik.de)  
[kurt.holm@jku.at](mailto:kurt.holm@jku.at)

Der Programmteil, der die Zellenmittelwerte und -streuungen berechnet und ausgibt wurde von Kurt Holm geschrieben, der in P18.3 beschriebene Programmteil, der verschiedene Tests durchführt, von Johann Bacher.

Im Text wird häufig auf das Dokument **P0** Bezug genommen. Dabei handelt es sich um das Almo-Dokument "Arbeiten mit Almo.PDF" (Dokument 0).

## Weitere Almo-Dokumente

Die folgenden Dokumente können alle von der Handbuchseite in [www.almo-statistik.de](http://www.almo-statistik.de) heruntergeladen werden

0. Arbeiten\_mit\_Almo.PDF (1 MB)
- 1a. Eindimensionale Tabellierung.PDF (1.8 MB)
- 1b. Zwei- und drei-dimensionale Tabellierung.PDF (1.1 MB)
2. Beliebig-dimensionale Tabellierung.PDF (1.7 MB)
3. Nicht-parametrische Verfahren.PDF (0.9 MB)
4. Kanonische Analysen.PDF (1.8 MB)  
Diskriminanzanalyse.PDF (1.8 MB)  
enthält: Kanonische Korrelation, Diskriminanzanalyse, bivariate Korrespondenzanalyse, optimale Skalierung
5. Korrelation.PDF (1.4 MB)
6. Allgemeine multiple Korrespondenzanalyse.PDF (1.5 MB)
7. Allgemeines ordinales Rasch-Modell.PDF (0.6 MB)
- 7a. Wie man mit Almo ein Rasch-Modell rechnet.PDF (0.2 MB)
8. Tests auf Mittelwertsdifferenz, t-Test.PDF (1,6 MB)
9. Logitanalyse.pdf (1,2MB) enthält Logit- und Probitanalyse
10. Koeffizienten der Logitanalyse.PDF (0,06 MB)
11. Daten-Fusion.PDF (1,1 MB)
12. Daten-Imputation.PDF (1,3 MB)
13. ALM Allgemeines Lineares Modell.PDF (2.3 MB)
- 13a. ALM Allgemeines Lineares Modell II.PDF (2.7 MB)
14. Ereignisanalyse: Sterbetafel-Methode, Kaplan-Meier-Schätzer, Cox-Regression.PDF (1,5 MB)
15. Faktorenanalyse.PDF (1,6 MB)
16. Konfirmatorische Faktorenanalyse.PDF (0,3 MB)
17. Clusteranalyse.PDF (3 MB)
18. Pisa 2012 Almo-Daten und Analyse-Programme.PDF (17 KB)
19. Guttman- und Mokken-Skalierung.PFD (0.8 MB)
20. Latent Structure Analysis.PDF (1 MB)
21. Statistische Algorithmen in C (80 KB)
22. Conjoint-Analyse (PDF 0,8 MB)
23. Ausreisser entdecken (PDF 170 KB)
24. Statistische Datenanalyse Teil I, Data Mining I
25. Statistische Datenanalyse Teil II, Data Mining II
26. Statistische Datenanalyse Teil III, Arbeiten mit Almo-Datenanalyse-System
27. Mehrfachantworten, Tabellierung von Fragen mit Mehrfachantworten (0.8 MB)
28. Metrische multidimensionale Skalierung (MDS) (0,4 MB)
29. Metrisches multidimensionales Unfolding (MDU) (0,6 MB)
30. Nicht-metrische multidimensionale Skalierung (MDS) (0,5 MB)
31. Pfadanalyse.PDF (0,7 MB)

## Inhaltsverzeichnis

P18.0 Eingabe und Ausgabe .....	4
P18.0.1 Eingabe mit Programm-Maske .....	5
P18.2.1 Programm-Maske mit Optionen Prog18m2 .....	18
P18.3 Tests auf Differenzen zwischen ordinalen und quantitativen Variablen.....	23
P18.3.0 t-Test, Median-Test, Ridit-Test für unabhängige Stichproben .....	23
P18.3.1. t-Test und Vorzeichentest für abhängige Stichproben.....	31
P18.3.2 t-Test für einen theoretischen Mittelwert.....	38
P18.3.3 Test auf Varianzhomogenität bei der Varianzanalyse für den uni- und multivariaten Fall .....	44
P18.3.4 Die Teststatistiken für den multivariaten Fall.....	49
P18.3.5 Test aus Varianzhomogenität bei der Kovarianzanalyse für den univariaten Fall.....	50
Literatur .....	54

## P18.0 Eingabe und Ausgabe

Wir wollen beispielsweise wissen, welches durchschnittliches Einkommen besitzen Männer und Frauen. Wir haben also 1 unabhängige nominale Variable, das Geschlecht, und 1 abhängige quantitative Variable, das Einkommen. Almo liefert folgende Ausgabe:

Mittelwerte	
	Einkommen
Geschlecht weiblich	3323.53
Geschlecht männlich	4185.19

Standardabweichung	
	Einkommen
Geschlecht weiblich	2271.57
Geschlecht männlich	2653.26

Häufigkeit	
	Einkommen
Geschlecht weiblich	34
Geschlecht männlich	27

Almo rechnet in diesem Fall noch einen t-Test; ermittelt also, ob die Differenz zwischen Männer und Frauen signifikant ist. Wir werden das später zeigen.

Oder: Wir wollen wissen, welches durchschnittliche Einkommen Menschen mit der Merkmalskombination weiblich/Arbeiter besitzen. Der Sachverhalt ist also folgender: Wir haben 2 **nominale** Variable

Geschlecht: männlich, weiblich

Beruf: Arbeiter, Angestellter, Sonstiger

Und wir haben eine **quantitative** Variable: Einkommen.

Die Zahl der nominalen Variablen (die die Merkmalskombinationen bilden) ist in Almo nicht beschränkt.

Almo berechnet für alle Merkmalskombinationen der beiden nominalen Variablen die Häufigkeiten sowie den Mittelwert und die Standardabweichung der quantitativen Variablen

Mittelwerte		
Geschlecht	Beruf	Einkommen
weiblich	Arbeiter	2875.00
	Angestellter	2833.33
	Selbständiger	4875.00
männlich	Arbeiter	3875.00
	Angestellter	4272.73
	Selbständiger	4375.00
Gesamtmittel		3704.92

Standardabweichung

Geschlecht	Beruf	Einkommen
weiblich	Arbeiter	1832.86
	Angestellter	1802.78
	Selbständiger	2847.70
männlich	Arbeiter	2847.70
	Angestellter	2259.96
	Selbständiger	2912.80

Häufigkeit

Geschlecht	Beruf	Häufigkeit
weiblich	Arbeiter	8
	Angestellter	18
	Selbständiger	8
männlich	Arbeiter	8
	Angestellter	11
	Selbständiger	8

Almo rechnet noch einen t-Test (den wir später zeigen werden). Dabei werden die Merkmalskombinationen wie die Ausprägungen *einer* Variablen behandelt. Allerdings ist ein t-Test in diesem Falle (wenn 2 oder mehrere nominale Variable kombiniert werden) nicht angemessen. Hier sollte man mit Prog 20 eine Varianzanalyse rechnen.

### ***P18.0.1 Eingabe mit Programm-Maske***

Prog18mb.Msk  
t-Test fuer Mittelwerts-Differenz quantitativer Variabler  
Mediantest für ordinale Variable

---

Quantitative Variable

Beispiel: Welches mittlere Einkommen besitzen Frauen u. Männer ?  
Unterscheiden sie sich signifikant in ihrem mittleren Einkommen?

Prog18 berechnet die Mittelwerte und Streuungen der quantitativen Variablen "Einkommen" je Ausprägung der nominalen Variablen "Geschlecht" und prüft mit dem t-Test, ob die Differenz signifikant ist.

Besitzt die nominale Variable 3 oder mehr Ausprägungen (z.B. Beruf: Arbeiter, Angestellter, Selbständiger, Sonstig), dann wird

- (1) ein gewöhnlicher paarweiser t-Test gerechnet,
- (2) ein t-Test für paarweise Vergleiche (Kontraste) und
- (3) der Scheffe-Test

In diesem Fall wäre das Allgemeine Lineare Modell besser geeignet

Auch für die Ausprägungskombinationen von 2 oder mehreren nominalen Variablen, z.B. weiblich/Arbeiter, weiblich/Angestellter etc. kann ein gewöhnlicher (paarweiser) t-Test gerechnet werden  
Hier wäre eine Varianzanalyse mit Prog 20 vorzuziehen

Ordinale Variable

Anstelle der quantitativen Variablen oder zusätzlich zu ihnen können auch ordinale Variable eingesetzt werden. In diesem Fall wird ein Ridittest oder ein Mediantest gerechnet

Beispiel:

Die Schädigung von Bäumen wird durch Forstfachleute auf einer Skala von 1 (=leichte Schäden) bis 4 (=schwere Schäden) eingestuft  
Wie unterscheiden sich die Schädigung von Fichten und Kiefern ?  
Ist der Unterschied in der Schädigung signifikant ?

Was ist ein Kurzprogramm ? -->   
Bedienung -->

1

Vereinbare Variable=  ;

2  Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

"C:\Almo?\TESTDAT\Uarnamen.nam"

     zeige = Namensdatei in Output zeigen  
leer = nicht

4

5 **Datei aus der gelesen wird**   
 bei Datei-Problemen  
 **"C:\Almo7\Testdat\TESTDAT.FRE"**  
 **frei**  **Format der Daten**  
  **U1:20** **der Datensatz enthält diese Variablen**  
**Bei Format DIREKT schreiben Sie: alle\_U**

6  **Wenn Dateiformat FIX oder Nicht-Standard-FREI**

7 **Analyse-Variable**  
**nominale Variable**   
  **Geschlecht**  
 Für die Ausprägungen dieser nominalen Variablen werden die Mittelwerte der (nachfolgend angegebenen) quantitativen Variablen ermittelt und geprüft, ob ihre Differenz signifikant ist.  
 Für die ordinalen Variablen wird die Riditdifferenz bzw. Mediandifferenz und deren Signifikanz errechnet.  
 Wenn Sie 2 oder mehrere nominale Variable angeben, dann werden deren Ausprägungskombinationen gebildet

---

**quantitative Variable**  
 Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse  
  **Einkommen, Bewertung**

---

**ordinale Variable**   
 Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse  
  **Leistung**

8  **Option: Umkodierungen und Kein-Wert-Angaben**

## Erläuterungen zu den Boxen:

**Box 1:** Speicher für x Variable

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.1.

**Box 2:** Weitere Vereinbarungen

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.2.

**Box 3:** Datei der Variablennamen

Siehe P0.3.

**Box 4:** Freie Namensfelder

Siehe P0.3.

**Box 5:** Datei aus der gelesen wird

**Box 6:** Wenn Dateiformat FIX oder Nicht-Standard-FREI

Siehe P0.4.

**Box 7:** Analyse-Variable

**Analyse-Variable**

**nominale Variable**

**Geschlecht**

Für die Ausprägungen dieser nominalen Variablen werden die Mittelwerte der (nachfolgend angegebenen) quantitativen Variablen ermittelt und geprüft, ob ihre Differenz signifikant ist.

Für die ordinalen Variablen wird die Riditdifferenz bzw. Mediandifferenz und deren Signifikanz errechnet.

Wenn Sie 2 oder mehrere nominale Variable angeben, dann werden deren Ausprägungskombinationen gebildet

---

**quantitative Variable**

Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse

**Einkommen, Bewertung**

---

**ordinale Variable**

Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse

**Leistung**

Sie können die Analyse-Variablen „von Hand“ in die Eingabefelder schreiben oder Sie klicken auf den Knopf mit den 2 kleinen symbolischen Fenstern. Almo öffnet dann die Dialogbox „Variable für Analyse auswählen“. In dieser können Sie die Variable, die in die Eingabefelder eingeschrieben werden sollen per Mausklick selektieren. Siehe die ausführliche Beschreibung dieser Dialogbox in P0.11.

1. Eingabefeld: Nominale Variable  
 Siehe hierzu auch P19.0.1.

Wir haben hier Geschlecht und Beruf als nominale Variable eingegeben. Für die Ausprägungen dieser nominalen Variablen werden die Mittelwerte der (im 2. Eingabefeld angegebenen) quantitativen Variablen ermittelt und geprüft, ob ihre Differenz signifikant ist.

Wenn Sie, wie in unserem Beispiel 2 oder mehrere nominale Variable angeben, dann werden deren Ausprägungskombinationen gebildet, also :

```
männlich - Arbeiter
männlich - Angestellter
.
.
weiblich - Beamte
```

Die nominale Variable darf mehr als 2 Ausprägungen besitzen

Nominale Variable werden normalerweise ganzzahlig mit Schrittweite 1 kodiert sein.  
 Beispiel:

Beruf	Codeziffer
-----	-----
Arbeiter	1
Angestellter	2
Beamter	3
Bauer	4
.	.
.	.
.	.

Im vorliegenden Maskenprogramm ist es jedoch zulässig, daß sie auch Dezimalwerte besitzen. Beispiel:

Beruf	Codeziffer
-----	-----
Arbeiter	10.5
Angestellter	12
Beamter	13.99
Bauer	8.1
.	.
.	.
.	.

Als Codeziffern sind hier etwa Qualifikationspunkte verwendet worden.

Almo kodiert diese Werte zwangsweise um, und zwar in folgender Weise.

alte Codeziffer	neue Codeziffer
-----	-----
8.1	1
10.5	2
12	3
13.99	4
.	.
.	.
.	.

Es wird bei 1 begonnen und mit Schrittweite 1 aufsteigend umkodiert. Beachte: Bauer hat mit 8.1 die niedrigste Codeziffer. Almo weist ihm die neue niedrigste Codeziffer 1 zu.

Wenn der Benutzer diese zwangsweise Umkodierung vermeiden will, weil er anders als Almo umkodieren würde, dann muß er in der Box

"Kein-Wert-Angabe und Umkodierungen"

selbst umkodieren, etwa so:

Beruf(8.1=1; 10.5=2; 12, 13.9=3)

Hier hat der Benutzer Angestellte und Beamte in der Art ihrer Tätigkeit als sehr ähnlich erachtet und sie deswegen in einer Kategorie zusammengefaßt.

## 2. Eingabefeld: quantitative Variable

Siehe hierzu auch P19.0.1.

Wir haben hier "Einkommen" als quantitative Variable angegeben. Sie können auch mehrere quantitative Variable angeben, z.B.

Einkommen, Alter

Almo rechnet dann für jede quantitative Variable eine separate Analyse, d.h. es berechnet den Mittelwert aus Einkommen und Alter für alle Merkmalskombinationen der beiden nominalen Variablen Geschlecht und Beruf.

## 3. Eingabefeld:

Siehe hierzu auch P19.0.1.

Ordinale Variable werden normalerweise ganzzahlig mit Schrittweite 1 kodiert sein. Beispiel:

Schulbildung	Codeziffer
-----	-----
Volksschule	1
Hauptschule	2
Gymnasium	3
Fachschule	4
Universität	5

Im vorliegenden Maskenprogramm ist es jedoch zulässig, daß sie auch Dezimalwerte besitzen. Beispiel:

Schulbildung	Codeziffer
-----	-----
Volksschule	8.2
Hauptschule	10.5
Gymnasium	13.1
Fachschule	13.9
Universität	19.22

Als Codeziffern sind etwa die durchschnittlichen Ausbildungsjahre verwendet worden.

Almo kodiert diese Werte zwangsweise um, und zwar in folgender Weise

alte Codeziffer	neue Codeziffer
-----	-----
8.2	1
10.5	2
13.1	3
13.9	4

Es wird bei 1 begonnen und mit Schrittweite 1 aufsteigend umkodiert.

Wenn der Benutzer diese zwangsweise Umkodierung vermeiden will, weil er anders als Almo umkodieren würde, dann muß er in der Box

"Kein-Wert-Angabe und Umkodierungen"

selbst umkodieren, etwa so:

Schulbildung(8.2=1; 10.5=2; 13.1, 13.9=3; 19.22=4)

Hier hat der Benutzer Gymnasium und Fachschule für gleichrangig erachtet und in einer Kategorie zusammengefasst

**Box 8:** Kein-Wert-Angabe und Umkodierungen  
Siehe P0.5.

### P18.0.3 Ausgabe aus Prog18mb

Almo liefert aus dem Maskenprogramm Prog18m1 folgende Ausgabe:

Fuer Analyse ausgewaehlte Variable

V1      Geschlecht: männlich weiblich  
V5      Leistung  
V7      Einkommen  
V20     Bewertung

V1 wird auch bezeichnet mit A  
die Auspraegungen (bzw.Dummies) mit  
A1 =männlich  
A2 =weiblich

Zellenmittelwerte der  
quantitativen/ordinalen Variablen  
(arithmetisches Mittel bei quantitativen Variablen)  
(Mediane bei ordinalen Variablen)

	Leistung V5	Einkommen V7	Bewertung V20
Geschlec männlich A1	3.0000	3.3235	4.3824
Geschlec weiblich A2	4.0000	4.1852	4.2963

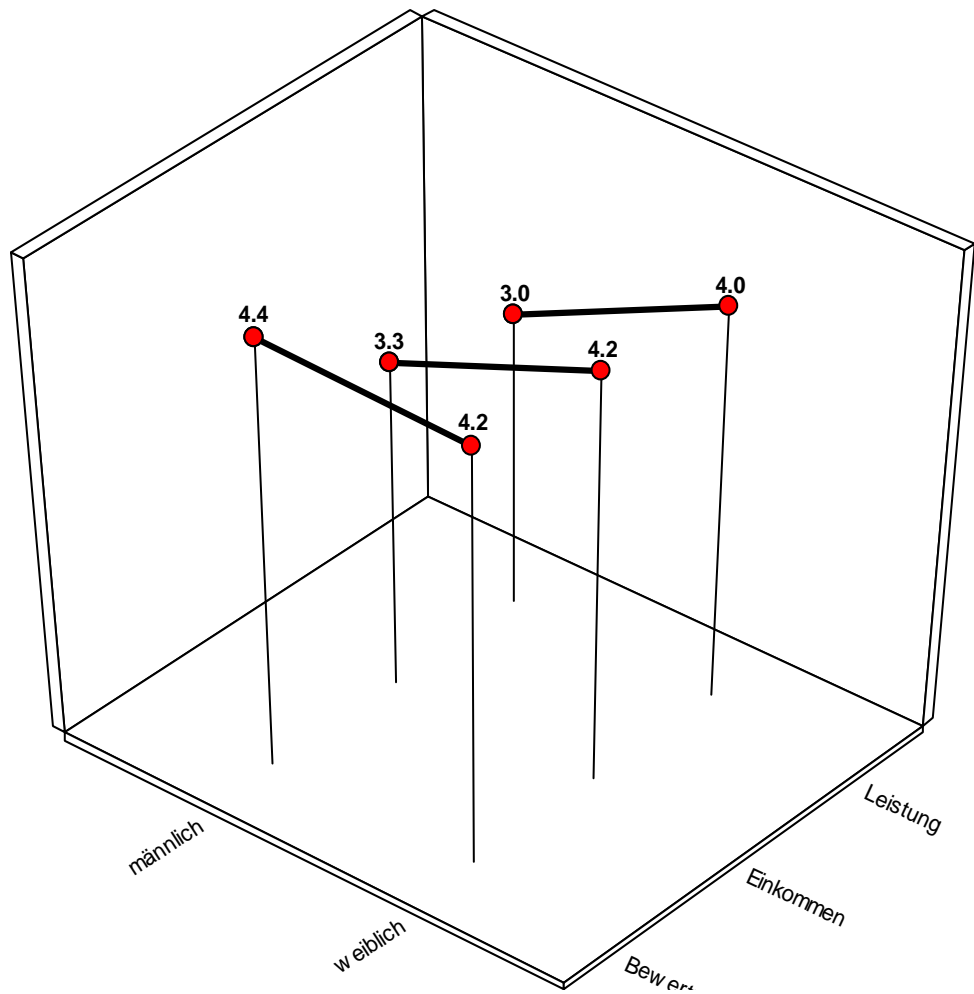
#### \*\*\*\*\* Erläuterung:

Männer haben in der quantitativen Variablen des Einkommens einen Mittelwert von 3.3235 etc.

In der ordinalen Variablen der Leistung ist der Median für die Männer 3.0 und für die Frauen 4.0.

Almo gibt hier folgende Grafik aus:

Mittelwerte



Streuung der  
quantitativen/ordinalen Variablen  
je Zelle  
Standardabweichung bei quantitativen Variablen  
(Mittlerer Quartilsabstand bei ordinalen Variablen)  
(Varianz bzw. Standardabweichung der quantitativen Variablen  
ist mit n, nicht mit n-1 dividiert)

	Leistung V5	Einkommen V7	Bewertung V20
Geschlec männlich A1	1.0000	2.2716	2.5897
Geschlec weiblich A2	1.5000	2.6533	2.7055

=====

die Zellenmittelwerte und Streuungen  
 der quantitativen und ordinalen Variablen  
 beruhen auf folgenden Besetzungszahlen (Zellenhaeufigkeiten)

	Leistung Einkomme Bewertung		
	V5	V7	V20
Geschlec männlich A1	34	34	34
Geschlec weiblich A2	27	27	27

=====

Zahl der insgesamt eingelesenen Einheiten 61

Zahl der in die Analyse einbezogenen Einheiten 61

=====

Median-Test fuer V5 Leistung

\*\*\*\*\* WARNUNG  
 Anwendung des Median-Test problematisch  
 Viele Beobachtungsfaelle (24.59%) gleich dem Median

Basistabelle für Median-Test für Gruppe 2 mit Gruppe 1

Gruppe	Zahl der Fälle			...Median
	kleiner	gleich	größer	
2	10	5	12	
1	18	10	6	
Gesamt	28	15	18	

unteres Dreieck der Matrix: Chi-Quadrat-Werte aus Mediantest  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der Chi-Quadrat-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mediandifferenz

	Geschlec Geschlec männlich weiblich A1 A2	
	Geschlec männlich A1	-
Geschlec weiblich A2	5.1954	-

=====

Test auf Varianzheterogenitaet fuer V7 Einkommen

unteres Dreieck der Matrix: F-Werte aus Varianzheterogenitaetstest  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der F-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht eine  
 signifikante Varianzheterogenitaet

	Geschlec maennlich A1	Geschlec weiblich A2
Geschlec maennlich A1	-	79.6782
Geschlec weiblich A2	1.3751	-

Freiheitsgrade fuer Teststatistiken fuer Varianzheterogenitaet  
 unteres Dreieck der Matrix: Freiheitsgrade der ersten Gruppe (z.B. A1)  
 oberes Dreieck der Matrix: Freiheitsgrade der zweiten Gruppe (z.B. A2)

	Geschlec maennlich A1	Geschlec weiblich A2
Geschlec maennlich A1	-	26
Geschlec weiblich A2	33	-

=====

Test auf Varianzheterogenitaet fuer V20 Bewertung

unteres Dreieck der Matrix: F-Werte aus Varianzheterogenitaetstest  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der F-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht eine  
 signifikante Varianzheterogenitaet

	Geschlec maennlich A1	Geschlec weiblich A2
Geschlec maennlich A1	-	59.4752
Geschlec weiblich A2	1.1000	-

Freiheitsgrade fuer Teststatistiken fuer Varianzheterogenitaet  
 unteres Dreieck der Matrix: Freiheitsgrade der ersten Gruppe (z.B. A1)  
 oberes Dreieck der Matrix: Freiheitsgrade der zweiten Gruppe (z.B. A2)

	Geschlec maennlich A1	Geschlec weiblich A2
Geschlec maennlich A1	-	26
Geschlec weiblich A2	33	-

=====

Mittelwertsdifferenzen fuer V7 Einkommen

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	-0.8617
Geschlec weiblich A2	-	-

=====

t-Test fuer Mittelwertsdifferenzen fuer V7 Einkommen

unteres Dreieck der Matrix: t-Werte aus gewoehnlichem t-Test  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	81.5499
Geschlec weiblich A2	1.3430	-

Freiheitsgrade fuer gewoehnliche t-Testwerte  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte  
 fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur von  
 einem Freiheitsgrad abhaengt.

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	-
Geschlec weiblich A2	59	-

=====

apriori t-Test (LSD) fuer Mittelwertsdifferenzen fuer V7 Einkommen

unteres Dreieck der Matrix: t-Werte  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	81.5499
Geschlec weiblich A2	1.3430	-

Freiheitsgrade: 59

=====

Zum apriori t-Test siehe Bortz: Statistik, 1993, S.246

Mittelwertsdifferenzen fuer V20 Bewertung

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	0.0861
Geschlec weiblich A2	-	-

=====

t-Test fuer Mittelwertsdifferenzen fuer V20 Bewertung

unteres Dreieck der Matrix: t-Werte aus gewoehnlichem t-Test  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	9.9592
Geschlec weiblich A2	0.1243	-

Freiheitsgrade fuer gewoehnliche t-Testwerte  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte  
 fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur von  
 einem Freiheitsgrad abhaengt.

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	-
Geschlec weiblich A2	59	-

=====

apriori t-Test (LSD) fuer Mittelwertsdifferenzen fuer V20 Bewertung

unteres Dreieck der Matrix: t-Werte  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	9.9592
Geschlec weiblich A2	0.1243	-

Freiheitsgrade: 59

=====

t-Test nach WELCH fuer Mittelwertsdifferenzen fuer V7 Einkommen

unteres Dreieck der Matrix: t-Werte aus t-Test nach WELCH  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	80.6650
Geschlec weiblich A2	1.3184	-

Freiheitsgrade fuer t-Testwerte nach WELCH  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte  
 fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur  
 von einem Freiheitsgrad abhaengt.

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	-
Geschlec weiblich A2	51	-

=====

t-Test nach WELCH fuer Mittelwertsdifferenzen fuer V20 Bewertung

unteres Dreieck der Matrix: t-Werte aus t-Test nach WELCH  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht  
 eine signifikante Mittelwertsdifferenz

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	9.9042
Geschlec weiblich A2	0.1236	-

Freiheitsgrade fuer t-Testwerte nach WELCH  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte  
 fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur  
 von einem Freiheitsgrad abhaengt.

	Geschlec männlich A1	Geschlec weiblich A2
Geschlec männlich A1	-	-
Geschlec weiblich A2	55	-

### ***P18.2.1 Programm-Maske mit Optionen Prog18m2***

Sie finden das Programm durch Klick auf den Knopf „Verfahren / Mittelwerts-Differenz“

5 **Datei aus der gelesen wird** **Hilfe**  
 bei Datei-Problemen  
  
 **Format der Daten** **Hilfe**  
 **der Datensatz enthält diese Variablen**  
 Bei Format DIREKT schreiben Sie: alle\_U

6 **Wenn Dateiformat FIX oder Nicht-Standard-FREI** **Hilfe**

7 **Analyse-Variablen**  
**nominale Variable** **Hilfe**  
  
 Für die Ausprägungen dieser nominalen Variablen werden die Mittelwerte der (nachfolgend angegebenen) quantitativen Variablen ermittelt und geprüft, ob ihre Differenz signifikant ist.  
 Für die ordinalen Variablen wird die Riditdifferenz bzw. Mediandifferenz und deren Signifikanz errechnet.  
 Wenn Sie 2 oder mehrere nominale Variable angeben, dann werden deren Ausprägungskombinationen gebildet

**quantitative Variable**  
 Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse

**ordinale Variable** **Hilfe**  
 Wenn Sie mehrere angeben, dann rechnet Almo für jede eine separate Analyse

8 **Option: Ein- und Ausschliessen von Untersuchungseinheiten**

9 **Option: Umkodierungen und Kein-Wert-Angaben**

10 **Option: Untersuchungseinheiten gewichten**

11 **verschiedene Optionen**

12 **Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix**

13 **Grafik-Optionen**

14 **Programmende**

**Erläuterungen zu den Boxen:**

Wir erläutern im folgenden nur jene Boxen, die gegenüber dem in P18.0.1 schon

dargestellten Maskenprogramm verändert bzw. neu sind.

**Box 8:** Ein- und Ausschließen von Untersuchungseinheiten  
Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.7.

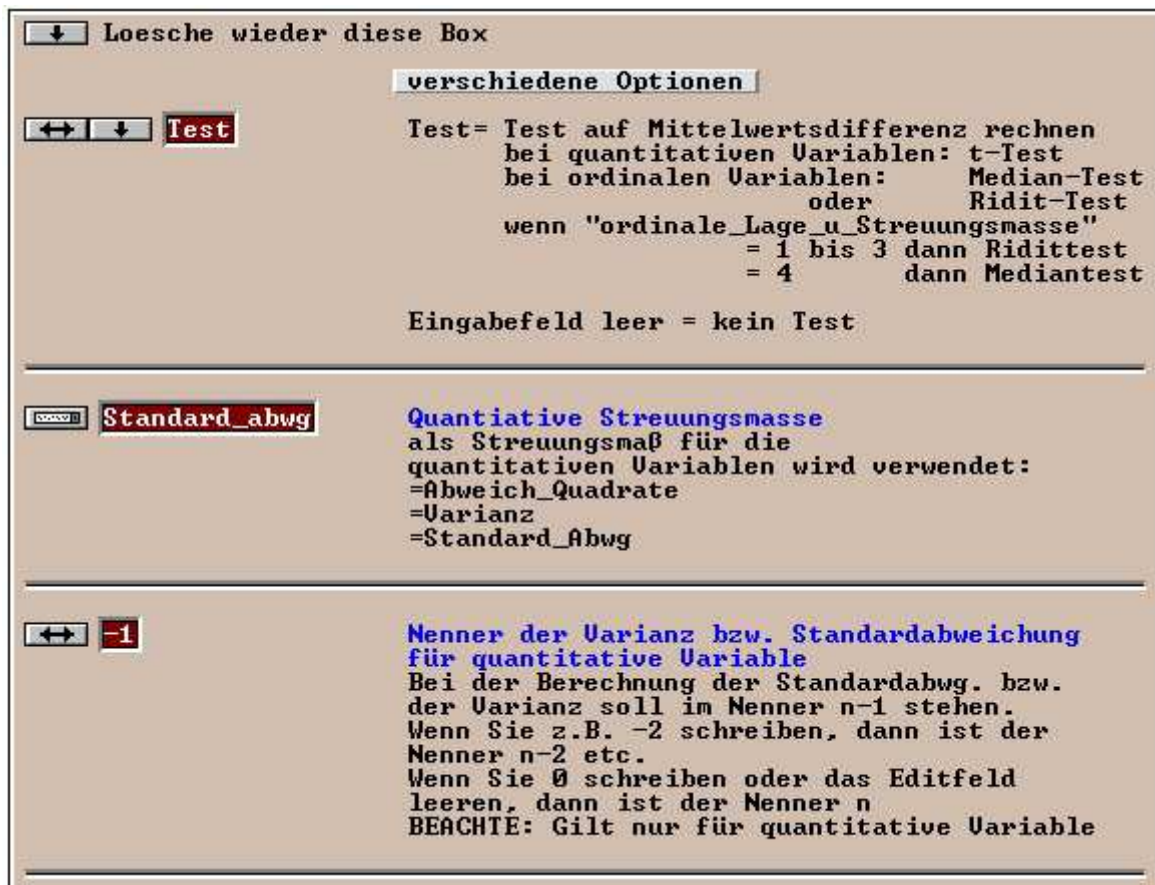
**Box 9:** Kein-Wert-Angabe und Umkodierungen  
Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.5.

**Box 10 :** Gewichtung  
Siehe P0.8.

**Box 11:** Verschiedene Optionen



Optionsbox geöffnet:





### 1. Eingabefeld: Test

Wird durch Klick auf den nach unten weisenden Pfeil das Wort "Test" eingesetzt, dann rechnet ALMO einen Test auf Mittelwertsdifferenz, allgemein auf "Lagedifferenz". D.h. es überprüft, ob die Differenzen zwischen den Mittelwerten der verschiedenen Ausprägungskombinationen signifikant sind.

**Bei quantitativen Variablen** werden die Mittelwerte und Streuungen der quantitativen Variablen (z.B. des Einkommens) je Ausprägung der nominalen Variablen (z.B. des Geschlechts) ermittelt und mit dem t-Test überprüft, ob die Differenz signifikant ist.

Besitzt die nominale Variable 3 oder mehr Ausprägungen (z.B. Beruf: Arbeiter, Angestellter, Selbständiger, Sonstig), dann wird

- (1) ein gewöhnlicher paarweiser t-Test gerechnet
- (2) ein t-Test für paarweise Vergleiche (Kontraste) und
- (3) der Scheffe-Test

In diesem Fall wäre auch eine Varianzanalyse mit Prog 20 möglich.

Auch für die Ausprägungskombinationen von 2 oder mehreren nominalen Variablen, z.B. weiblich/Arbeiter, weiblich/Angestellter etc. kann ein gewöhnlicher (paarweiser) t-Test gerechnet werden. Hier wäre eine Varianzanalyse mit Prog 20 vorzuziehen.

**Bei ordinalen Variablen** wird ein Ridittest oder ein Mediantest gerechnet. Welcher Test gerechnet wird hängt davon ab, welches ordinale Streuungsmaß im 4. Eingabefeld eingesetzt wurde.

Wir werden später in Abschnitt P18.3 auf den t-Tests, den Median- und Ridit-Test ausführlich eingehen.

## 2. Eingabefeld: Quantitative Streuungsmaße

Der Benutzer kann festlegen, welches Streuungsmaß für die quantitativen Variablen errechnet werden soll. Dabei gibt es folgende Alternativen:

Abweich_quadrate	Für jede Zelle wird für die quantitativen Variablen die Summe der Abweichungsquadrate als Streuungsmaß berechnet.
Varianz	Für jede Zelle wird für die quantitativen Variablen die Varianz als Streuungsmaß berechnet.
Standard_Abwg	Für jede Zelle wird für die quantitativen Variablen die Standardabweichung als Streuungsmaß berechnet.

Unabhängig vom gewählten Streuungsmaß rechnet Almo einen t-Test für Mittelwertsdifferenzen.

## 3. Eingabefeld: Nenner der Varianz bzw. Standardabweichung bei quantitativen Variablen

Bei der Berechnung der Standardabwg. bzw. der Varianz setzt Almo im Nenner  $n$  ein ( $n$ =Gesamtzahl aller Untersuchungseinheiten). Soll im Nenner  $n-1$  stehen, dann schreiben Sie in das Eingabefeld  $-1$ . Wenn Sie z.B.  $-2$  schreiben, dann ist der Nenner  $n-2$  etc. Wenn Sie  $0$  schreiben oder das Editfeld leeren, dann ist der Nenner  $n$ .

## 4. Eingabefeld: Ordinale Lage- und Streuungsmasse Ridittest oder Mediantest

Der Benutzer kann festlegen, welches Streuungsmaß für die ordinalen Variablen und welcher Test errechnet werden soll. Dabei gibt es folgende Alternativen:

## Eingabe

- =1 Für jede Zelle werden für die ordinale Variablen "Ridits" als Lagemaße und nach dem tau-b-Kalkül (DENZ, 1977) die Summe der Abweichungsquadrate als Streuungsmaße berechnet.
- =2 Für jede Zelle werden für die ordinale Variablen "Ridits" als Lagemaße und nach dem tau-b-Kalkül (DENZ, 1977) die Varianz als Streuungsmaße berechnet. (Dies ist die Voreinstellung)
- =3 Für jede Zelle werden für die ordinale Variablen "Ridits" als Lagemaße und nach dem tau-b-Kalkül (DENZ, 1977) die Standardabweichung als Streuungsmaße berechnet.
- =4 Für jede Zelle werden für die ordinale Variablen Mediane als Lagemaße und der mittlere Quartilabstand als Streuungsmaße berechnet. Bei der Berechnung der Mediane wird wie folgt verfahren: Bei gerader Fallzahl wird der Mittelwert der beiden mittleren Beobachtungen verwendet, bei ungeraden die mittlere Beobachtung.

Bei 1 - 3 wird der Ridittest gerechnet.

Bei 4 wird der Mediantest gerechnet.

**Box 13:** Optionen, die das "Aussehen" der Ergebnistabellen steuern  
Siehe P0.9.

**Box 14:** Grafik-Optionen  
Siehe P0.10.

## ***P18.3 Tests auf Differenzen zwischen ordinalen und quantitativen Variablen***

### **P18.3.0 t-Test, Median-Test, Ridit-Test für unabhängige Stichproben**

**Test bei ordinalen Variablen :** Test auf paarweise Riditdifferenzen (siehe dazu Abschnitt P10.5 Punkt f). Im Programm wird die Teststatistik der paarweisen Riditdifferenzen (absoluter z-Wert der Standardnormalverteilung) und das Signifikanzniveau der Teststatistik ausgegeben.

Wird zusätzlich "Ordinale\_Lage\_u\_Streungsmasse = 4;" im Programmparameter-Block angegeben (siehe P18.2.4) dann führt das Programm an Stelle des Tests auf paarweise Riditdifferenzen einen Test auf paarweise Mediandifferenzen durch (vgl. KRIZ, 1978,S.179). Median und Quantile werden dabei nach der Methode des "mittleren Wertes" berechnet. Siehe P4.2.

**Test bei quantitativen Variablen:** t-Test auf paarweise Mittelwertdifferenzen. Es werden folgende Teststatistiken berechnet und ausgegeben:

- Teststatistik für Heterogenität der Varianzen der Zellen - Signifikanzniveau der Teststatistik für die Heterogenität der Varianzen.
- Teststatistik der absoluten Mittelwertdifferenzen. Diese entspricht dem bekannten t-Test. Seine Anwendung ist nur dann sinnvoll, wenn keine Varianzheterogenität vorliegt. Signifikanzniveau der absoluten Mittelwertdifferenzen
- Teststatistik für absolute Mittelwertdifferenzen bei Varianzheterogenität. Es wird

der t-Test nach WELCH gerechnet (CLAUSS/ EBNER, 1975, S. 213).  
Signifikanzniveau des t-Tests nach WELCH.

Besitzt die unabhängige nominale Variable drei oder mehr Ausprägungen, dann wird noch berechnet

- der t-Test für paarweise Vergleiche (Kontraste)
- der Scheffé-Test

Betrachten wir ein Beispiel:

Es soll der Erfolg einer neuen Unterrichtsmethoden geprüft werden. Dazu wurden zwei Gruppen von Schülern gebildet. Die eine Gruppe wird unter sonst gleiche Voraussetzungen mit der alten Methode (Methode A) unterrichtet, die zweite Gruppe unter der Methode B. Am Ende des Versuchs wurde ein Wissenstest durchgeführt. Das Beispiel ist CLAUSS, G. & H. EBNER, 1974: Grundlagen der der Statistik, Berlin, S. 212ff entnommen. Die Daten sind im Almo-Ordner Testdat unter dem Namen „Clausebn.fre“ enthalten

Das Almo-Maskenprogramm Prog18m2, aus Abschnitt P18.2.1 auf dieses Beispiel angewendet, lautet, wobei wir nur Box 4, 5, 7, 10, 11

#### Box 4:




Freie Namensfelder Hilfe

↔ Name 1=Gruppe:GruppeA,GruppeB;  
↔ Name 2=Punkte;  
↔ Name 3=Fallzahl;

[...] erzeuge zusätzliche Namensfelder

#### Box 5:



Datei aus der gelesen wird Hilfe

bei Datei-Problemen

Format der Daten Hilfe

↔  der Datensatz enthält diese Variablen  
Bei Format DIREKT schreiben Sie: alle\_U

**Box 7:**

**Analyse-Variable**

**nominale Variable**

**Gruppe**

Für die Ausprägungen dieser nominalen Variablen werden die Mittelwerte der (nachfolgend angegebenen) quantitativen Variablen ermittelt und geprüft, ob ihre Diferenz signifikant ist.

Für die ordinalen Variablen wird die Riditdifferenz bzw. Mediandifferenz und deren Signifikanz errechnet.

Wenn Sie 2 oder mehrere nominale Variable angeben, dann werden deren Ausprägungskombinationen gebildet

---

**quantitative Variable**

Wenn Sie mehrere angeben, dann rechnet Also für jede eine separate Analyse

**Punkte**

---

**ordinale Variable**

Wenn Sie mehrere angeben, dann rechnet Also für jede eine separate Analyse

Box 11:

↓ Loesche wieder diese Box

**verschiedene Optionen**

↔ ↓ **Test**

Test= Test auf Mittelwertsdifferenz rechnen  
bei quantitativen Variablen: t-Test  
bei ordinalen Variablen: Median-Test  
oder Ridit-Test  
wenn "ordinale\_Lage\_u\_Streuungsmaesse"  
= 1 bis 3 dann Ridittest  
= 4 dann Mediantest

Eingabefeld leer = kein Test

---

☐ **Standard\_abwg**

**Quantitative Streuungsmaesse**  
als Streuungsmaesse fuer die  
quantitativen Variablen wird verwendet:  
=Abweich\_Quadrate  
=Varianz  
=Standard\_Abwg

---

↔ **-1**

**Nenner der Varianz bzw. Standardabweichung  
fuer quantitative Variable**  
Bei der Berechnung der Standardabwg. bzw.  
der Varianz soll im Nenner n-1 stehen.  
Wenn Sie z.B. -2 schreiben, dann ist der  
Nenner n-2 etc.  
Wenn Sie 0 schreiben oder das Editfeld  
leeren, dann ist der Nenner n  
BEACHTEN: Gilt nur fuer quantitative Variable

---

↑↓ **2**

**ordinale Lage\_u\_Streuungsmaesse  
fuer ordinale Variable**  
als Lage- und Streuungsmaesse fuer die  
ordinalen Variablen werden verwendet:

- 1= Ridits als Lagemaesse. Abweichungsquadrate  
nach dem Kendall'schen tau-b Kalkuul
- 2= Ridits als Lagemaesse. Varianzen nach dem  
Kendallschen tau-b Kalkuul.Uoreinstellung
- 3= Ridits als Lagemaesse. Standardabweichungen  
nach dem Kendallschen tau-b Kalkuul
- 4= Mediane als Lagemaesse. Mittlerer Quartils-  
abstand als Streuungsmaesse

BEACHTEN:  
Bei 1 bis 3 wird der Ridittest gerechnet  
Bei 4 wird der Mediantest gerechnet

---

↑↓ **0**

**Mittelwerte bzw. Ridits u. Besetzungszahlen  
in die Datei ".\Progs\P18\_Mittw.fre" speichern**  
1= ja 0=nein

---

↑↓ **0**

**Basis-Statistiken ausgeben**  
1= ja 0=nein

ALMO liefert folgende Ausgabe (gekürzt):

Fuer Analyse ausgewaehlte Variable

V1      Gruppe: GruppeA GruppeB  
V2      Punkte

V1 wird auch bezeichnet mit A  
die Auspraegungen (bzw.Dummies) mit  
A1 =GruppeA  
A2 =GruppeB

Zellenmittelwerte der  
quantitativen/ordinalen Variablen

			Punkte
			V2
Gruppe	GruppeA	A1	11.2273
Gruppe	GruppeB	A2	13.2000

Streuung (Standardabweichung) der  
quantitativen/ordinalen Variablen  
je Zelle  
(Varianz bzw. Standardabweichung  
ist mit n, nicht mit n-1 dividiert)

			Punkte
			V2
Gruppe	GruppeA	A1	2.9105
Gruppe	GruppeB	A2	1.8193

die Zellenmittelwerte und Streuungen der abhaengigen Variablen  
beruhen auf folgenden Besetzungszahlen (Zellenhaeufigkeiten)

			Punkte
			V2
Gruppe	GruppeA	A1	22
Gruppe	GruppeB	A2	20

Zahl der insgesamt eingelesenen Einheiten      14

Zahl der in die Analyse einbezogenen Einheiten  
unter Beruecksichtigung der Gewichtung:      42

=====

Einfaktorielle Varianzanalyse fuer quantitaive Variable V2 Punkte

Gesamtmittelwert (fuer n= 42) = 12.167  
 erklarte Streuung in Prozent = 13.899  
 multipler Korrelationskoeffizient = 0.373

Streuungszerlegung:  
 (Streuung in Abweichungsquadraten)

Gesamt = 293.3333 df-Gesamt = 41  
 zwischen den Auspraegungen (=erklarte Streuung) = 40.7697 df-zwischen = 1  
 innerhalb der Auspraegungen (=Fehlerstreuung) = 252.5636 df-innerh. = 40  
 F-Wert = 6.4569  
 Signifikanz 100\*(1-p) = 98.4981

=====

Test auf Varianzheterogenitaet fuer V2 Punkte

unteres Dreieck der Matrix: F-Werte aus Varianzheterogenitaetstest  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der F-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht eine signifikante Varianzheterogenitaet

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	97.7527
Gruppe	GruppeB	A2	2.5471	-

Freiheitsgrade fuer Teststatistiken fuer Varianzheterogenitaet  
 unteres Dreieck der Matrix: Freiheitsgrade der ersten Gruppe (z.B. A1)  
 oberes Dreieck der Matrix: Freiheitsgrade der zweiten Gruppe (z.B. A2)

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	19
Gruppe	GruppeB	A2	21	-

=====

Mittelwertsdifferenzen fuer V2 Punkte

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	-1.9727
Gruppe	GruppeB	A2	-	-

=====

t-Test fuer Mittelwertsdifferenzen fuer V2 Punkte

unteres Dreieck der Matrix: t-Werte aus gewoehnlichem t-Test  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht eine signifikante Mittelwertsdifferenz

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	98.4981
Gruppe	GruppeB	A2	2.5411	-

Freiheitsgrade fuer gewoehnliche t-Testwerte  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur von einem Freiheitsgrad abhaengt.

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	-
Gruppe	GruppeB	A2	40	-

=====

apriori t-Test (LSD) fuer Mittelwertsdifferenzen fuer V2 Punkte

unteres Dreieck der Matrix: t-Werte  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100  
 Werte ueber ca. 95 bedeuten: Es besteht eine signifikante Mittelwertsdifferenz

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	98.4981
Gruppe	GruppeB	A2	2.5411	-

Freiheitsgrade: 40

=====

t-Test nach WELCH fuer Mittelwertsdifferenzen fuer V2 Punkte

unteres Dreieck der Matrix: t-Werte aus t-Test nach WELCH  
 oberes Dreieck der Matrix: Signifikanz (1-p)\*100 der t-Werte  
 Werte ueber ca. 95 bedeuten: Es besteht eine signifikante Mittelwertsdifferenz

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe	GruppeA	A1	-	98.6468
Gruppe	GruppeB	A2	2.5957	-

Freiheitsgrade fuer t-Testwerte nach WELCH  
 unteres Dreieck der Matrix: Freiheitsgrade der t-Werte  
 fuer Gruppenvergleiche  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur  
 von einem Freiheitsgrad abhaengt.

			Gruppe GruppeA A1	Gruppe GruppeB A2
Gruppe Gruppe	GruppeA GruppeB	A1 A2	- 36	- -

In dem Beispiel ist die Annahme der Varianzhomogenität bei einem Signifikanzniveau von 95 % nicht erfüllt, da sich für den berechneten F-Wert eine Signifikanz von 97.8% ergibt. Das heißt, die  $H_0$ -Hypothese der Gleichheit beider Varianzen (Varianzhomogenität) kann mit einem Fehler von 2.2% verworfen werden. Da die Annahme der Varianzhomogenität nicht erfüllt ist, muß strenggenommen der t-Test nach WELCH verwendet werden. In dem Beispiel unterscheiden sich aber die Testgrößen des gewöhnlichen t-Tests und des t-Tests nach Welch kaum.

Anmerkung: Der gewöhnliche t-Test und der t-Test nach WELCH unterscheiden sich (a) in der Berechnung der Varianz  $s^2_{MW}$  der Mittelwertdifferenzen und (b) in der Berechnung der Freiheitsgrade df.

Für den gewöhnlichen t-Test gilt:

$$s^2_{MW} = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2} * \frac{n_1 + n_2}{n_1 * n_2}$$

$$df = n_1 + n_2 - 2$$

Beim t-Test nach WELCH werden  $s^2_{MW}$  und df wie folgt berechnet:

$$s^2_{MW} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

$$df - 1 = \frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1} \quad \text{mit} \quad c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

### P18.3.1. t-Test und Vorzeichentest für abhängige Stichproben

Beim t-Test für abhängige Stichproben werden folgende Teststatistiken berechnet und ausgegeben:

**Ordinale Variable:** Das Programm berechnet zwei Testgrößen: den Vorzeichentest ohne Berücksichtigung der Bindungen (KRIZ 1978:190) und einen modifizierten Vorzeichentest, bei dem Bindungen berücksichtigt werden. Die Teststatistik für den Vergleich von zwei wiederholten Beobachtungen X1 und X2 ist definiert als:

$t = d/s_d$  mit

$$d = \sum \text{sign}(x_{1i} - x_{2i}) / n_{12} \quad s_d^2 = \frac{\sum (\text{sign}(x_{1i} - x_{2i}))^2 - d^2 / n_{12}}{n_{12}(n_{12} - 1)}$$

wobei  $\text{sign}(x)$  die Vorzeichenfunktion ist.

Die Verteilung der Teststatistik wird durch eine t-Verteilung mit  $n_{12} - 1$  Freiheitsgraden approximiert.

**Quantitative Variable:** Es wird der t-Test für abhängige Stichproben (vgl. CLAUSS/EBNER, 1974, S. 240ff) zur Prüfung der Mittelwertdifferenzen und der t-Test nach FERGUSON (CLAUSS/EBNER, 1974, S. 244f) zur Prüfung der Varianzheterogenität berechnet.

Die beiden Teststatistiken sind definiert als:

t-Test nach FERGUSON zur Überprüfung der Varianzheterogenität:

$$t^2 = \frac{(s_1^2 - s_2^2)^2 * (n_{12} - 1)}{4 * (s_1^2 - s_2^2) * (1 - r_{12})^2}$$

mit

$s_i^2$  = Varianz der Variablen i

$r_{12}$  = Korrelation der Variablen 1 und 2

$n_{12}$  = Anzahl der (gültigen) Beobachtungen in den Variablen 1 und 2

Die Teststatistik  $t$  ( $\sqrt{t^2}$ ) ist t-verteilt mit  $n_{12}-2$  Freiheitsgraden.

t-Test für Mittelwertdifferenzen in abhängigen Stichproben:

$t = d/s_d$

mit

$$d = \sum (x_{1i} - x_{2i}) / n_{12} \quad s_d^2 = \frac{\sum (x_{1i} - x_{2i})^2 - d^2 / n_{12}}{n_{12}(n_{12} - 1)}$$

Die Teststatistik  $t$  ist t-verteilt mit  $n_{12}-1$  Freiheitsgraden.

#### Eingabe mit Programm-Maske Prog18m3

Sie finden das Programm durch Klick auf „Verfahren/Mittelwerts-Differenzen“.



5   bei Datei-Problemen  
  
 Format der Daten   
 der Datensatz enthält diese Variablen  
Bei Format DIREKT schreiben Sie: alle\_U

---

höchste Variablen-Nummer des Datensatzes

6  Wenn Dateiformat FIX oder Nicht-Standard-FREI

7  entweder quantitative oder ordinale Variable - nicht beides  
 quantitative Variable  
 ordinale Variable

8  Option: Ein- und Ausschliessen von Untersuchungseinheiten

9  Option: Umkodierungen und Kein-Wert-Angaben

10  Option: Untersuchungseinheiten gewichten

11  **Quantitative Streuungsmaße**  
Als Streuungsmaß für die quantitativen Variablen wird verwendet:  
=Abweich\_Quadrate  
=Varianz  
=Standard\_Abwg  
 **Ordinale Lage\_u\_Streuungsmaße**  
Als Lage- und Streuungsmaße für die ordinalen Variablen werden verwendet:  
1= Ridits als Lagemaß. Abweichungsquadrate nach dem Kendall'schen tau-b Kalkül  
2= Ridits als Lagemaß. Varianzen nach dem Kendall'schen tau-b Kalkül.Voreinstellung  
3= Ridits als Lagemaß. Standardabweichungen nach dem Kendall'schen tau-b Kalkül  
4= Mediane als Lagemaß. Mittlerer Quartilsabstand als Streuungsmaß  
 **Basis-Statistiken ausgeben**  
1= ja 0=nein

12  Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

13  Grafik-Optionen

14

## Erläuterungen zu den Boxen:

**Box 1:** Speicher für x Variable

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.1.

**Box 2:** Weitere Vereinbarungen

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.2.

**Box 3:** Datei der Variablennamen

Siehe P0...

**Box 4:** Freie Namensfelder

Siehe P0.3.

**Box 5:** Datei aus der gelesen wird

**Box 6:** Wenn Dateiformat FIX oder Nicht-Standard-FREI

Siehe P0.4.

**Box 7:** Analyse-Variable

Messwiederholungs-Variable  
entweder quantitative oder ordinale Variable - nicht beides

[↔] [□] [Messung1, Messung2, Messung3] \ quantitative Variable

[↔] [□] [ ] \ ordinale Variable

Hilfe

Geben Sie hier die Variable an, die Sie analysieren wollen. Die Variablen müssen entweder als quantitativ oder als ordinal angegeben werden. Es ist nicht möglich einen Teil der Variablen als quantitativ und einen Teil als ordinal zu deklarieren. Für quantitative Variable rechnet Almo den t-Test für abhängige Stichproben. Für ordinale Variable rechnet Almo den Vorzeichenstest für abhängige Stichproben.

Sie können die Analyse-Variablen „von Hand“ in die Eingabefelder schreiben oder Sie klicken auf den Knopf mit den 2 kleinen symbolischen Fenstern. Almo öffnet dann die Dialogbox „Variable für Analyse auswählen“. In dieser können Sie die Variable, die in die Eingabefelder eingeschrieben werden sollen per Mausklick selektieren. Siehe die ausführliche Beschreibung dieser Dialogbox in P0.11.

Zur Kodierung der ordinalen Variablen siehe Erläuterungen zu Prog18mb, Box 7.

**Box 8:** Ein- und Ausschließen von Untersuchungseinheiten

Siehe P0.7.

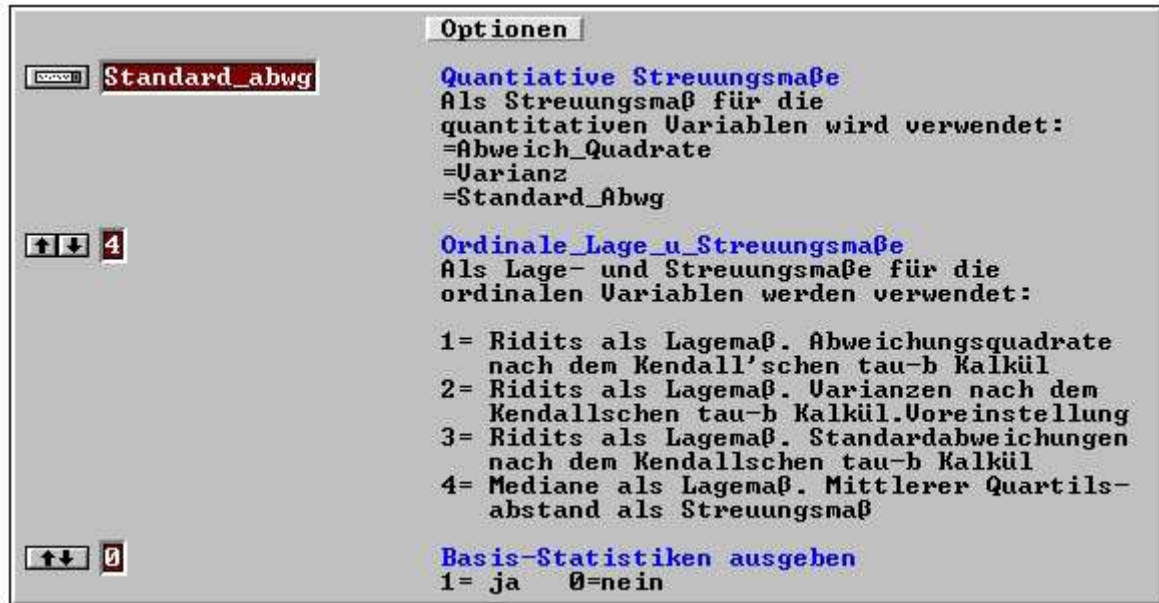
**Box 9:** Kein-Wert-Angabe und Umkodierungen

Siehe P0.5.

**Box 10:** Gewichtung

Siehe P0.8.

## Box 11: Optionen



Der Benutzer kann wählen, welches Lagemaß und welches Streuungsmaß er für die quantitativen bzw. ordinalen Variablen erhalten möchte.

Für quantitative Variable gibt es folgende Auswahlmöglichkeit für die Streuungsmaße:

Abweich\_Quadrate (Abweichungsquadrate)  
Varianz  
Standard\_Abwg (Standardabweichung)

Als Lagemaß wird immer der Mittelwert ausgegeben

Klicken Sie auf den Knopf und selektieren Sie dann das entsprechende Streuungsmaß.

Für ordinale Variable gibt es folgende Auswahlmöglichkeit:

- 1 Ridits als Lagemaß. Abweichungsquadrate nach dem Kendall'schen tau-b Kalkül
- 2 Ridits als Lagemaß. Varianzen nach dem Kendallschen tau-b Kalkül
- 3 Ridits als Lagemaß. Standardabweichungen nach dem Kendallschen tau-b Kalkül
- 4 Mediane als Lagemaß. Mittlerer Quartilsabstand als Streuungsmaß

Klicken Sie auf den Knopf mit dem nach unten weisenden Pfeil - so oft, bis die Kennziffer erscheint, die Sie haben wollen.

**Box 12:** Optionen, die das "Aussehen" der Ergebnistabellen steuern  
Siehe P0.9.

## Ausgabe

Das Maskenprogramm Prog18m3 (ohne Ein- bzw. Ausschließen von Untersuchungseinheiten und ohne Gewichtung) liefert folgende Ergebnisse (gekürzt):

Fuer Analyse ausgewaehlte Variable

V21 FiktivenominaleV:  
 V5 Messung1  
 V6 Messung2  
 V7 Messung3

V21 wird auch bezeichnet mit A  
 die Auspraegungen (bzw.Dummies) mit  
 A1

Zellenmittelwerte der  
 quantitativen/ordinalen Variablen

	Messung1	Messung2	Messung3
	V5	V6	V7
Fiktive nominale Var	3.8852	3.9344	3.7049

**\*\*\*\*\* Erläuterung:**

Hier und in den folgenden Tabellen erscheint als Name für die Zeilenvariable die Bezeichnung "Fiktive nominale Var". Es ist in Almo aus technischen Gründen notwendig beim Test für abhängige Stichproben eine fiktive Zeilenvariable einzuführen. Sie ist ohne Bedeutung. Löschen Sie diese einfach.

Streuung (Varianz) der  
 quantitativen/ordinalen Variablen  
 je Zelle  
 (Varianz bzw. Standardabweichung  
 ist mit n, nicht mit n-1 dividiert)

	Messung1	Messung2	Messung3
	V5	V6	V7
Fiktive nominale Var	3.6426	4.8809	6.1752

die Zellenmittelwerte und Streuungen  
 der quantitativen Variablen  
 beruhen auf folgenden Besetzungszahlen (Zellenhaeufigkeiten)

	Messung1	Messung2	Messung3
	V5	V6	V7
Fiktive nominale V	61	61	61

```
=====
Zahl der insgesamt eingelesenen Einheiten      61
Zahl der in die Analyse einbezogenen Einheiten 61
=====
```

**\*\*\*\*\* Erläuterung:**

Da in unserem Beispiel nicht 2 sondern 3 Messungen miteinander verglichen werden, wird im folgenden ein "paarweiser" t-Test für abhängige Stichproben durchgeführt. In diesen Falle wäre ein Messwiederholungs-Design mit dem

Allgemeinen Linearen Modell angemessener. Siehe Maskenprogramm Prog20m2.  
Bei nur 2 Messungen liefern beide Verfahren dasselbe Ergebnis.

Teststatistiken fuer Varianzheterogenitaet in abhaengigen Stichproben  
unteres Dreieck der Matrix: t-Werte aus t-Test nach FERGUSON  
oberes Dreieck der Matrix: Signifikanz  $(1-p)*100$  der t-Werte  
Werte ueber ca. 95 bedeuten: Es besteht eine  
signifikante Varianzheterogenitaet

		Messung1	Messung2	Messung3
		V5	V6	V7
Messung1	V5	-	73.6445	95.5426
Messung2	V6	1.1287	-	63.6436
Messung3	V7	2.0515	0.9153	-

**\*\*\*\*\* Erläuterung:**

In der Zelle V5-V7 (im rechten oberen Eck) steht ein Signifikanzwert, der größer als 95 ist. Die Streuung von Messung1 und Messung3 sind also signifikant verschieden. Sehr streng genommen dürfte dann der anschließende t-test für abhängige Stichproben nicht gerechnet werden. Hier kann der Benutzer aber durchaus "großzügig" sein. Der t-Test gilt als durchaus robust gegen Verletzungen der Bedingung der Varianzhomogenität.

Teststatistiken fuer absolute Mittelwertdifferenzen in abhaengigen Stichproben  
unteres Dreieck der Matrix: t-Werte aus t-Test fuer abh. Stichproben  
oberes Dreieck der Matrix: Signifikanz  $(1-p)*100$  der t-Werte  
Werte ueber ca. 95 bedeuten: Es besteht  
eine signifikante Mittelwertsdifferenz

		Messung1	Messung2	Messung3
		V5	V6	V7
Messung1	V5	-	10.6166	33.9218
Messung2	V6	0.1329	-	38.1088
Messung3	V7	0.4406	0.4995	-

**\*\*\*\*\* Erläuterung:**

Kein t-Wert (im oberen Dreieck) ist signifikant. D.h. die 3 Messungen sind in ihren Mittelwerten nicht signifikant verschieden.

Freiheitsgrade fuer t-Testwerte in abhaengigen Stichproben  
unteres Dreieck der Matrix: Freiheitsgrade der t-Werte nach FERGUSON  
oberes Dreieck der Matrix: Freiheitsgrade der t-Werte fuer t-Test

		Messung1	Messung2	Messung3
		V5	V6	V7
Messung1	V5	-	60	60
Messung2	V6	59	-	60
Messung3	V7	59	59	-

## **P18.3.2 t-Test für einen theoretischen Mittelwert**

### **Eingabe mit Programm-Maske Prog18m4**

Sie finden das Programm durch Klick auf „Verfahren/Mittelwerts-Differenzen“.

Prog18m4.Msk Kurzprogramm  
t-Test fuer einen theoretischen Mittelwert

---

Beispiel: 61 Untersuchungspersonen erbringen in einem Test eine mittlere Leistung von 3.89 Punkten. Aus theoretischen Ueberlegungen heraus dürfte ihre mittlere Leistung nur 3.5 Punkte betragen.  
Frage: Weichen sie signifikant vom theoretischen Mittelwert ab ?

Siehe auch Prog08m3: Wilcoxon Vorzeichenrangtest fuer einen theoretischen Median

Was ist ein Kurzprogramm ? -->   
Bedienung -->

1 Speicher fuer x Variable   
Vereinbare Variable=  ; höchste Variablen-Nummer + 1

2  Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3 Namen für die Analyse-Variable  
 Name 5=Leistung;

4 Datei aus der gelesen wird   
 "C:\Almo7\Testdat\TESTDAT.FRE" bei Datei-Problemen  
 frei Format der Daten   
  U1:20 der Datensatz enthält diese Variablen  
Bei Format DIREKT schreiben Sie: alle\_U  


---

 20 höchste Variablen-Nummer des Datensatzes

5  Wenn Dateiformat FIX oder Nicht-Standard-FREI

6 Analyse-Variable  
  Leistung quantitative Analyse-Variable  
 3.5 Wert des theoretischen Mittelwerts

7  Option: Ein- und Ausschliessen von Untersuchungseinheiten

8  Option: Umkodierungen und Kein-Wert-Angaben



### Erläuterungen zu den Boxen:

**Box 1:** Speicher für x Variable

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.1.

**Box 2:** Weitere Vereinbarungen

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.2.

**Box 3:** Datei der Variablennamen

Siehe P0.3.

**Box 4:** Freie Namensfelder

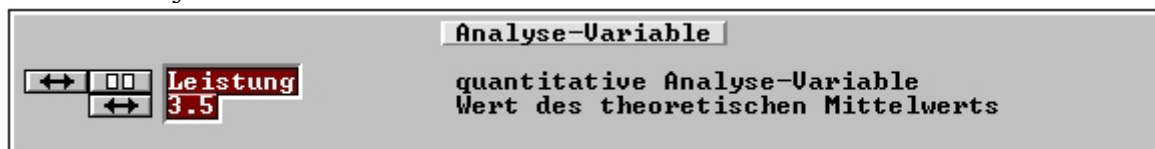
Siehe P0.3.

**Box 5:** Datei aus der gelesen wird

**Box 6:** Wenn Dateiformat FIX oder Nicht-Standard-FREI

Siehe P0.4.

**Box 7:** Analyse-Variable



Geben Sie hier die Variable an, die Sie analysieren wollen, sowie den vermuteten Wert des theoretischen Mittelwerts. Die Variable wird von Almo als quantitative begriffen.

**Box 8:** Kein-Wert-Angabe und Umkodierungen

Siehe P0.5.

## Box 9: Optionen



### 1. Eingabefeld:

Der Benutzer kann wählen, welches Streuungsmaß er für die Analyse-Variable erhalten möchte.

Dabei gibt es folgende Auswahlmöglichkeit:

Abweich\_Quadrate (Abweichungsquadrate)  
Varianz  
Standard\_Abwg (Standardabweichung)

Klicken Sie auf den Knopf und selektieren Sie dann das entsprechende Streuungsmaß.

### 2. Eingabefeld:

Nenner der Varianz bzw. Standardabweichung

Bei der Berechnung der Standardabwg. bzw. der Varianz setzt Almo im Nenner standardmäßig  $n$  ein. Wollen Sie, dass dort  $n-1$  stehen soll, dann schreiben Sie in das Eingabefeld  $-1$ . Wenn Sie z.B.  $-2$  schreiben, dann ist der Nenner  $n-2$  etc. Wenn Sie  $0$  schreiben oder das Editfeld leeren, dann ist der Nenner  $n$ .

## Ausgabe

Almo liefert aus dem Maskenprogramm Prog18m4 und obigem „selbst geschriebenem“ Almo-Programm folgendes Ergebnis (gekürzt):

```
Fuer Analyse ausgewaehlte Variable

V21   Fiktive_nominale_V: theoretGruppe empirGruppe
V5    Leistung

V21 wird auch bezeichnet mit A
die Auspraegungen (bzw.Dummies) mit
A1 =theoretGruppe
A2 =empirGruppe
```

Zellenmittelwerte der  
quantitativen/ordinalen Variablen

	Leistung V5
Fiktive nom Var theoretGruppe	3.5000
Fiktive nom Var empirGruppe	3.8852

Streuung (Varianz) der  
quantitativen/ordinalen Variablen  
je Zelle  
(Varianz bzw. Standardabweichung  
ist mit n-1 dividiert)

	Leistung V5
Fiktive nom Var theoretGruppe	0
Fiktive nom Var empirGruppe	3.7033

die Zellenmittelwerte und Streuungen der abhaengigen Variablen  
beruhen auf folgenden Besetzungszahlen (Zellenhaeufigkeiten)

	Leistung V5
Fiktive nom V theoretGruppe	1
Fiktive nom V empirGruppe	61

```
=====
Zahl der insgesamt eingelesenen Einheiten      62
Zahl der in die Analyse einbezogenen Einheiten 62
=====
```

t-Test der Differenz des theoretischen zum empirischen Mittelwert  
fuer Variable V5 Leistung

unteres Dreieck der Matrix: t-Wert  
oberes Dreieck der Matrix: Signifikanz (1-p)\*100 des t-Werts  
Werte ueber ca. 95 bedeuten: Es besteht  
eine signifikante Mittelwertsdifferenz

	Fiktiven theoretG A1	Fiktiven empirGru A2
Fiktiven theoretG A1	-	87.6683
Fiktiven empirGru A2	1.5635	-

**\*\*\*\*\* Erläuterung:**

Im oberen rechten Eck lesen wir eine Signifikanz ab von 87.6683. Dieser Wert liegt unter der üblichen Signifikanzgrenze von 95. Das bedeutet: Der empirische Mittelwert weicht nicht signifikant vom theoretischen ab.

Freiheitsgrade fuer t-Testwert  
 unteres Dreieck der Matrix: Freiheitsgrade des t-Werts  
 oberes Dreieck der Matrix: leer, da t-Verteilung nur  
 von einem Freiheitsgrad abhaengt

	Fiktiven theoretG A1	Fiktiven empirGru A2
Fiktiven theoretG A1	-	-
Fiktiven empirGru A2	60	-

**Beachte:**

Die in den Abschnitten P18.3.0 bis P18.3.2 behandelten Teststatistiken werden unter folgenden Bedingungen nicht berechnet.

Testgröße	wird unter folgenden Bedingungen nicht berechnet
<b>unabhängige Stichproben:</b>	
Riditanalyse	Wenn eine oder beide Besetzungszahlen der Vergleichsgruppen kleiner 1 ist.
Mediantest	Wenn alle Ausprägungen einer Vergleichsgruppe größer oder kleiner dem Gesamtmedian der beiden Vergleichsgruppen ist.
Varianzhomogenitätstest	Wenn eine oder beide der Varianzen der Vergleichsgruppen gleich 0 ist
gewöhnlicher t-Test	Wenn die Summe der Varianzen in beiden Vergleichsgruppen gleich 0 ist
t-Test nach Welch	Wenn eine Besetzungszahl in jeder Vergleichsgruppe gleich 0 ist
<b>abhängige Stichproben:</b>	
Vorzeichentest	Wenn die Besetzungszahl der Vergleichsmessungen größer 0 ist
modifizierter Vorzeichentest	Wenn die Varianz der Vorzeichendifferenzen größer 0 ist
Varianzhomogenitätstest	Wenn die Varianzen der beiden Vergleichsmessungen größer 0 sind und/oder der Absolutbetrag der Korrelation größer 1 ist
t-Test	Wenn die Varianz der Vergleichsmessungen größer 0 ist.
Vergl. mit theor. Mittelwerten: t-Test	Wenn die Varianz des empirischen Mittelwertes = 0 ist.

Vergleichsgruppen sind die jeweils verglichenen Ausprägungen der nominalen Variablen. Vergleichsmessungen die jeweils verglichenen Meßwiederholungen. Ist eine Berechnung aus den angegebenen Gründen nicht möglich, wird im Programm das Minuszeichen "-" angegeben.

### P18.3.3 Test auf Varianzhomogenität bei der Varianzanalyse für den uni- und multivariaten Fall

Für die Signifikanztests im Allgemeinen Linearen Modell gilt die Voraussetzung der „Varianzhomogenität“. Siehe dazu P20.19.

Die Annahme der Varianzhomogenität bedeutet, daß in jeder Zelle, die aus der Kombination der unabhängigen nominalen Variablen gebildet werden kann, die abhängige quantitative Variable dieselbe Varianz besitzen muß.

Betrachten wir ein Beispiel:

Die unabhängigen nominalen Variablen sind Geschlecht und Beruf. Die abhängige quantitative Variable ist die Leistung. Also gibt folgende Tabelle aus:

Varianz der quantitativen Variablen Leistung  
je Merkmalskombination der unabh. nominalen Variablen

Geschlec	Beruf	Leistung
männlich	Arbeiter	3.7344
	Angestel	3.9506
	Selbstän	3.6094
weiblich	Arbeiter	4.1094
	Angestel	3.4298
	Selbstän	3.8594

Die Varianzen in den Zellen dieser Tabelle sind verschieden. Die Frage ist: Sind sie signifikant verschieden ?

Um diese Frage zu beantworten rechnen wir mit Programm 18 einen Varianzhomogenitätstest.

Die Frage nach der Varianzhomogenität stellt sich auch im Falle der Kovarianzanalyse, wenn also auf Seiten der unabhängigen Variablen sich zusätzlich noch quantitative Variable befinden. In diesem Falle muß die Wirkung der unabhängigen quantitativen Variablen auf die abhängige Variable eliminiert werden bevor der eigentliche Kalkül des Varianzhomogenitätstest gerechnet wird.

In Also kann die Varianzhomogenität auch im multivariaten Fall getestet werden, wenn also mehrere abhängige Variable vorhanden sind. Dies geht gegenwärtig allerdings nur bei der Varianzanalyse. Bei der Kovarianzanalyse kann nur der univariate Fall behandelt werden.

#### **P18.3.3.1 Eingabe mit Programm-Maske Prog18me**

Sie finden das Programm nach Klick auf „Verfahren/Homogenitätstest“.

**Prog18me.Msk**  
**Varianzhomogenitätstest bei Varianzanalyse**  
**für den univariaten und multivariaten Fall**

**Beispiel:** Der Einfluss der nominalen Variablen Geschlecht und Beruf auf die Leistung in einem Test (=abhängige quantitative Variable) soll ermittelt werden.

Mit Prog18me soll untersucht werden, ob die Annahme der Varianzhomogenität erfüllt ist. Diese Annahme bedeutet, dass in jeder Zelle, die aus der Kombination der unabhängigen nominalen Variablen gebildet werden kann, die abhängige quantitative Variable dieselbe Varianz besitzen muss

Das Programm kann auch für den multivariaten Fall verwendet werden, d.h. wenn 2 oder mehrere abhängige quantitative Variable vorhanden sind

Siehe Handbuch, Abschnitt P18.3.3

Was ist ein Kurzprogramm ? -->

Bedienung -->

1

Vereinbare Variable=  ;

2  Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

"C:\Almo7\TESTDAT\Uarnamen.nam"

     zeige = Namensdatei in Output zeigen  
leer = nicht

4

5

"C:\Almo7\Testdat\TESTDAT.PRE"      bei Datei-Problemen

     Format der Daten

     der Datensatz enthält diese Variablen  
Bei Format DIREKT schreiben Sie: alle\_U

6  Wenn Dateiformat FIX oder Nicht-Standard-FREI

7

8

9

10

11

**Erläuterungen zu den Boxen:**

**Box 1:** Speicher für x Variable

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.1.

**Box 2:** Weitere Vereinbarungen

Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.2.

**Box 3:** Datei der Variablennamen

Siehe P0.3.

**Box 4:** Freie Namensfelder

Siehe P0.3.

**Box 5:** Datei aus der gelesen wird

**Box 6:** Wenn Dateiformat FIX oder Nicht-Standard-FREI

Siehe P0.4.

**Box 7:** Analyse-Variable

Sie können die Analyse-Variablen „von Hand“ in die Eingabefelder schreiben oder Sie klicken auf den Knopf mit den 2 kleinen symbolischen Fenstern. Almo öffnet dann die Dialogbox „Variable für Analyse auswählen“. In dieser können Sie die Variable, die in die Eingabefelder eingeschrieben werden sollen per Mausclick

selektieren. Siehe die ausführliche Beschreibung dieser Dialogbox in P0.11.

### 1. und 2. Eingabefeld

Geben Sie hier eine oder mehrere unabhängige nominale Variable an, sowie eine oder mehrere abhängige quantitative Variable. Die nominalen Variablen dürfen auch Dezimalwerte annehmen. Sie werden dann von Almo zwangsweise auf Ganzzahligkeit umkodiert. Siehe dazu auch P18.01, Erläuterung zu Box 7.

### 3. Eingabefeld

Wenn Sie nur 1 abhängige Variable in das 2. Eingabefeld eingeschrieben haben, dann schreiben Sie durch Klick auf den Knopf mit Pfeil nach unten eine 1. Almo rechnet dann einen univariaten Homogenitätstest. Wenn Sie 2 oder mehrere abhängige Variable angegeben haben, dann setzen Sie 2 ein.

**Box 8:** Ein- und Ausschließen von Untersuchungseinheiten  
Siehe P0.7.

**Box 9:** Kein-Wert-Angabe und Umkodierungen  
Siehe P0.5.

**Box 10:** Grafik-Optionen  
Siehe P0.10.

Wenn nur 1 abhängige Variable vorhanden ist, dann kann mit der Anweisung "**Test = Homogenitaet\_univariat;**" im Programmparameter\_Block die Annahme der Varianzhomogenität des allgemeinen linearen Modells (siehe Holm, 1979) überprüft werden. Wenn 2 oder mehrere abhängige Variable vorhanden sind, dann liegt der multivariate Fall vor. In diesem Fall lautet die Anweisung

Test=Homogenitaet\_multivariat;

### P18.3.3.3 Die Teststatistiken für den univariaten Fall

Das Programm berechnet folgende Teststatistiken, denen die  $H_0$ -Hypothese zugrunde liegt, daß sich die Streuungen der quantitativen/ordinalen Variablen je Merkmalskombination der nominalen Variablen rein zufällig unterscheiden:

**Box-Bartlett-Test:** Siehe dazu Winer, 1971, S. 208f.

**F-Max-Statistik:** Die Testgröße

$$b = \frac{s_{\max}^2}{(QS_G - QS_{\max}) / df_G - df_{\max}}$$

ist F-verteilt mit  $(df_{\max}, df_G - df_{\max})$  Freiheitsgraden. Die Annahme der Varianzhomogenität kann zum 5 %-Niveau ( $= \alpha$ -Fehler) verworfen werden; wenn  $(1-\alpha) \cdot 100 < 95$  ist. Die Größen der Gleichung sind wie folgt definiert:

$n_k$  Besetzungszahl der k-ten Zelle

$QS_k$  ..... Summe der Abweichungsquadrate der k-ten Zelle

$df_k = n_k - 1$  ..... Freiheitsgrade für  $QS_k$

$s_k^2 = QS_k / df_k$  ... Varianz der k-ten Zelle

$s_{(k)}^2$  ..... die der Größe nach geordneten Varianzen,

also  $s^2_{(\min)} \leq s^2_{(2)} \leq \dots \leq s^2_{(\max)}$

$QS_G = \sum_k QS_k$  .. Gesamtsumme der Abweichungsquadrate

$$df_G = \sum_k QS_k \dots \text{Freiheitsgrade für QSG}$$

**F-Min-Statistik:** Die Testgröße

$$b^{-1} = \frac{s_{\min}^2}{(QS_G - QS_{\min}) / df_G - df_{\min}}$$

ist ebenfalls F-verteilt mit  $(df_{\min}, df_G - df_{\min})$  Freiheitsgraden.

**Harrison-Mc.Cabe Test:** Die Testgröße für den Harrison-Mc.Cabe Test unterscheidet sich von den bisher besprochenen Tests dadurch, daß im Zähler die Gesamtstreuung steht.

Die Testgröße ist:  $b = \frac{QS_1 / df_1}{QS_G / df_G}$  mit  $QS_1 = \sum s_{(i)}^2 df_{(i)}$   
 $df_1 = \sum df_{(i)}$

Summiert wird, bis  $\sum n_{(i)} \geq m$  ist. Die Zahl m kann der Benutzer durch "**OPTION 17 = m**" festsetzen. Die Voreinstellung für OPTION 17 ist  $n/2$ , da für  $m=n/2$  der Test die größte Mächtigkeit besitzt. Der Test stellt eine Approximation durch die F-Verteilung dar und liefert nicht immer eine eindeutige Entscheidung. Folgende Konstellationen sind möglich:

		H <sub>0</sub> - verwerfen		
		ja	nein	
H <sub>0</sub> - akzeptieren	ja		2	≤ 1-α
	nein	1	3	≥ 1-α
		≥ 1-α	≤ 1-α	

- (1) H<sub>0</sub> kann zum Niveau  $(1 - ) * 100$  verworfen werden.  
Das Signifikanzniveau ist z.B. für beide Tests größer 95 %.
- (2) H<sub>0</sub> kann zum Niveau  $(1 - ) * 100$  akzeptiert werden.  
Das Signifikanzniveau ist z.B. für beide Tests kleiner 95 %.
- (3) unbestimmter Bereich.  
Das Signifikanzniveau ist z.B. für den Test, ob H<sub>0</sub> verworfen werden kann, kleiner 95 % und für den Test, ob H<sub>0</sub> akzeptiert werden kann, größer 95 %.

#### P18.3.3.4 Ausgabe im univariaten Fall

Das nachfolgende Beispiel enthält die Ergebnisse einer Untersuchung der Varianzheterogenität. Die untersuchte nominale Variable bestand aus vier Ausprägungen.

##### Varianzhomogenitätstest fuer Variable 2

Von den 4 Zellen muessen fuer den Varianzhomogenitäts- Test  
0 eliminiert werden.

Gesamtstreuung	318.519	Freiheitsgrade	27
F-Max-Statistik	4.008	Sign.niveau 100(1-p)	99.361
F-Min-Statistik	5.25	Sign.niveau 100(1-p)	90.163

Harrison-Mc.Cabe-Test fuer Varianz.- bzw. Kovarianzanalyse

Test, ob H<sub>0</sub> verworfen werden kann:

Teststatistik	3.375	Sign.niveau 100(1-p)	98.650
---------------	-------	----------------------	--------

Test, ob H<sub>0</sub> akzeptiert werden kann:

Teststatistik	7.425	Sign.niveau 100(1-p)	99.948
---------------	-------	----------------------	--------

```

Box-Bartlett-Test
  Teststatistik    13.347    Sign.niveau 100(1-p)    99.553
Anwendungsvoraussetzungen f. Box-Bartlett-Test
  Faelle, wo n > 2 (in %) 100
  Faelle, wo n > 5 (in %) 75

```

In dem obigen Beispiel kann die  $H_0$ -Hypothese der Varianzhomogenität beim Box-Bartlett-Test und beim F-Max Test zu einem Fehlerniveau von weniger als 1% verworfen werden. Beim Harrison-Mc.Cabe Test wäre das nur möglich, wenn beide Teststatistiken ein Signifikanzniveau größer 99% bzw. ein Fehlerniveau kleiner 1% besitzen. Legt man also ein Signifikanzniveau von 99% bzw. ein Fehlerniveau von 1% zugrunde, liegt die als Fall 3 (unbestimmter Bereich) beschriebene Situation vor. Bei einem Niveau von 95% kann dagegen die  $H_0$ -Hypothese verworfen werden. Die F-Min-Statistik schließlich führt nur zu einer Verwerfung, wenn ein Fehlerniveau von 10% zugrundegelegt wird. In dem Beispiel ist dieser Unterschied dadurch bedingt, daß die Zelle mit der kleinsten Streuung nur sehr schwach besetzt ist (in der Zelle befinden sich nur vier Beobachtungen) und deshalb die Streuung sehr unzuverlässig gemessen wird. Den beim Box-Bartlett-Test ausgegebenen Anwendungsvoraussetzungen liegen die Empfehlungen von WINER (1971, S. 95) zugrunde. Nach WINER sollten die Besetzungszahlen aller Zellen nicht kleiner 3 und die der meisten Zellen größer 5 sein.

### P18.3.4 Die Teststatistiken für den multivariaten Fall

Für den multivariaten Fall werden zwei Teststatistiken berechnet:

1. Box-Bartlett-Test für die Gleichheit der Varianz-Kovarianzmatrizen der Zellen (vgl. WINER, 1971, S. 370).
2. Box-Bartlett-Test für verbundene Symmetrie der Varianz-Kovarianzmatrizen der Zellen (vgl. WINER, 1971, S. 371).

Verbundene Symmetrie von Varianz-Kovarianzmatrizen liegt dann vor, wenn alle quantitativen Variablen in jeder Zelle dieselben Varianzen und Kovarianzen besitzen. Die Annahme der verbundenen Symmetrie geht beispielsweise in die Signifikanztest von Meßwiederholungsdesign ein.

#### Ausgabe für den multivariaten Fall (gekürzt):

```

Box-Bartlett-Test
  Teststatistik    6.402    Sign.niveau 100(1-p) 61.940
Box-Bartlett-Test fuer verbundene Symmetrie
  Teststatistik    0.566    Sign.niveau 100(1-p) 3.604
Anwendungsvoraussetzungen f. Box-Bartlett-Test
Verhaeltnisse Besetzungszahlen zur
Zahl der quantitativen Variablen in %

  Verhaeltnis kleiner 2                100.000
  Verhaeltnis zwischen 2 und 3         0.000
  Verhaeltnis zwischen 3 und 5         0.000
Faustregel (alle Besetzungszahlen größer 20, Zahl der
Zellen kleiner 6 und Zahl der quantitativen Variablen
kleiner 6 ist erfüllt hinsichtlich
- der Besetzungszahlen mit(in %) 0
- der Zahl der Zellen:           ja
- der Zahl der quant. Var.:      ja

```

In dem Beispiel kann die  $H_0$ -Hypothese der Gleichheit der Varianz-Kovarianzmatrizen und die der verbundenen Symmetrie nicht verworfen werden. Ein Verwerfen der  $H_0$ -Hypothese der Gleichheit der Varianz-Kovarianzmatrizen wäre nur bei einem Fehlerniveau von 40% bzw. einem Signifikanzniveau von 60% möglich. Bei der Interpretation ist aber Vorsicht geboten, da die Anwendungsvoraussetzung nach WINER (1971, S. 370), daß die Besetzungszahl jeder Zelle im Verhältnis zur Zahl der untersuchten Variablen relativ groß sein soll, nicht erfüllt ist. In beiden untersuchten Zellen liegt das Verhältnis der Besetzungszahl zu der Zahl der quantitativen Variablen unter 2.0. Auch die Faustregel (FAHRMEIR/HAMERLE, 1984, S. 74f), daß alle Besetzungszahlen größer 20 sein sollen, ist nicht erfüllt.

### **P18.3.5 Test aus Varianzhomogenität bei der Kovarianzanalyse für den univariaten Fall**

In Almo ist gegenwärtig nur der Test für den univariaten Fall enthalten.

#### **Eingabe mit Programm-Maske Prog18mh**

Sie finden das Programm durch Klick auf „Verfahren/Homogenitätstest“.



7

**Analyse-Variable - unabhängige Variable**

unabhängige nominale Variable

↔   **Geschlecht, Beruf**

↑↓   **2** Interaktionen x. Ordnung zwischen den  
unabh. nominalen Variablen miteinbeziehen  
0 =keine Interaktionen bilden

---

unabhängige quantitative Variable  
(=Kovariate)

↔   **Alter, Bildung**

↑↓   **2** Zahl der Kovariaten

8

**Analyse-Variable - abhängige Variable**

abhängige quantitative Variable  
(nur eine möglich)

↔   **Leistung**

9

↓ **Option: Ein- und Ausschliessen von Untersuchungseinheiten**

10

↓ **Option: Umkodierungen und Kein-Wert-Angaben**

11

↓ **Option: Verfahren**

12

↓ **Grafik-Optionen**

13

**Programmende**

**Erläuterungen zu den Boxen:**

**Box 1:** Speicher für x Variable  
Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.1.

**Box 2:** Weitere Vereinbarungen  
Siehe Almo-Dokument Nr.0 "Arbeiten\_mit\_Almo", Abschnitt P0.2.

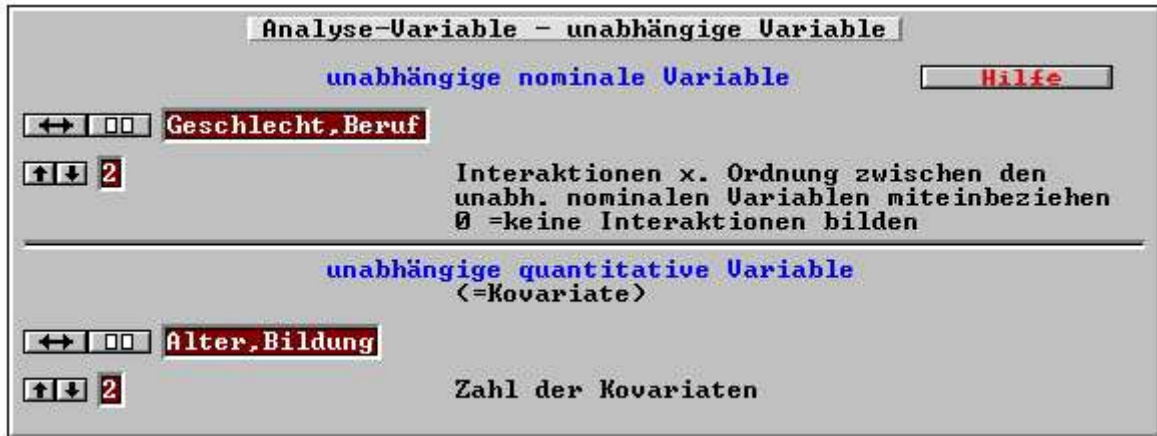
**Box 3:** Datei der Variablennamen  
Siehe P0.3.

**Box 4:** Freie Namensfelder  
Siehe P0.3.

**Box 5:** Datei aus der gelesen wird

**Box 6:** Wenn Dateiformat FIX oder Nicht-Standard-FREI  
Siehe P0.4.

**Box 7:** Analyse-Variable – unabhängige Variable



*1. Eingabefeld*

Geben Sie hier eine oder mehrere unabhängige nominale Variable an. Die nominalen Variablen dürfen auch Dezimalwerte annehmen. Sie werden dann von Almo zwangsweise auf Ganzzahligkeit umkodiert. Siehe dazu auch P18.0.1, Erläuterungen zu Box 7.

*2. Eingabefeld*

Sie können Interaktionen zwischen den unabhängigen nominalen Variablen bilden. Siehe P19.1.4, Erläuterungen zu Box 7.

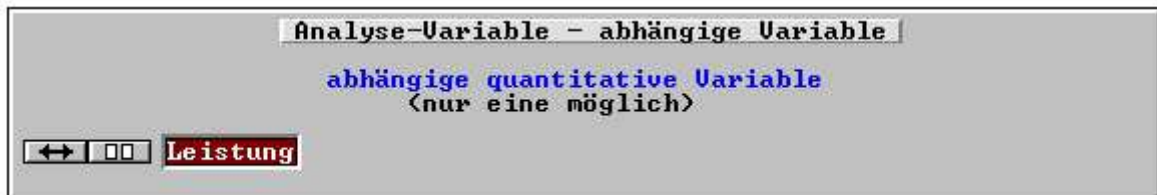
*3. Eingabefeld*

Geben Sie hier eine oder mehrere unabhängige quantitative Variable (= kovariate) an.

*4. Eingabefeld*

Geben Sie hier die Zahl der von Ihnen im 3. Eingabefeld selektierten Kovariaten an.

**Box 8:** Analysevariable – abhängige Variable



Geben Sie hier die abhängige quantitative Variable an. Es ist nur eine möglich.

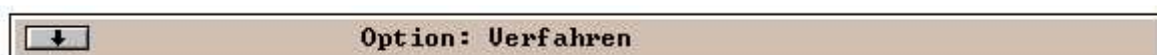
**Box 9:** Ein- und Ausschließen

Siehe P0.7.

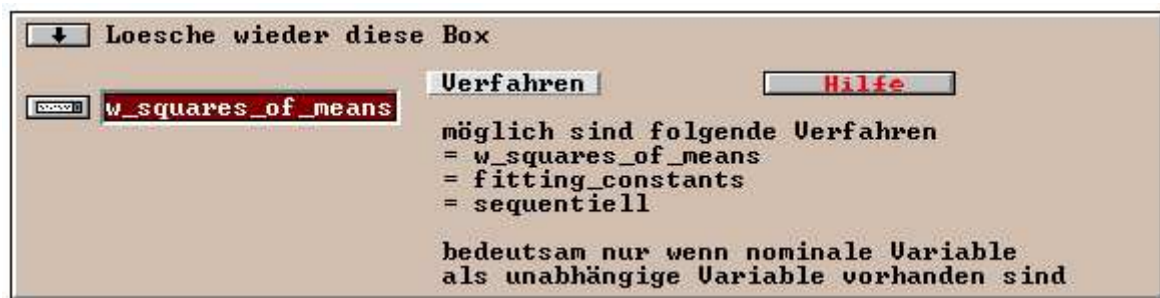
**Box 10:** Kein-Wert-Angabe und Umkodierungen

Siehe P0.5.

**Box 11:** Verfahren



Optionsbox geöffnet:



Siehe dazu die ausführliche Darstellung in P20.7.3. Wir empfehlen „w\_squares\_of\_means“; sollte nur eine unabhängige nominale Variable vorhanden sein, dann „fitting\_constants“.

**Box 12:** Grafik-Optionen  
Siehe P0.10.

## Literatur

In fast allen Statistik-Lehrbüchern werden t-Test und Median-Test behandelt.