

Sommersemester 2009, Statistik mit SPSS



Bivariate Datenanalyse, Überblick bis Freitag

heute	Mittelwertvergleiche
heute	Kreuztabellen
Donnerstag	Kreuztabellen II und Zusammenhangsmaße
Donnerstag	Korrelationsanalyse und Streudiagramme, im Anschluss: Übungsklausur
Freitag	Regressionsanalyse
Freitag	Wiederholung, Klausurvorbereitung

Überblick Mittelwertvergleiche

1. Das Verfahren, Anwendungsbeispiele
2. Übersicht T-Tests
3. T-Test für unabhängige Stichproben

1. Verfahren T-Test, mögliche Fragestellungen

Der T-Test ist ein Verfahren der Inferenzstatistik, mit dem man überprüfen kann, ob sich die Mittelwerte zweier Stichproben überzufällig unterscheiden

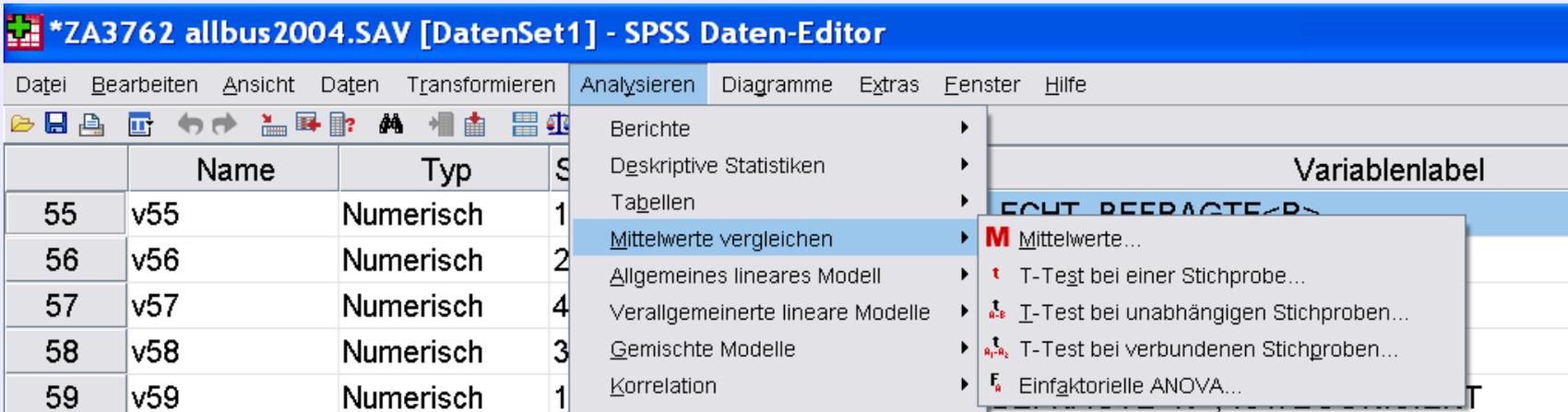
Typische Beispiele aus der Sozialforschung:

- Verdienen Männer mehr als Frauen?
- Verbringen Studenten mehr Zeit im Internet als Nicht-Studenten?
- Stufen sich Ostdeutsche weiter links ein als Westdeutsche?

Voraussetzungen für die Berechnung von Mittelwertunterschieden:

- Metrische Variablen
- Stichproben mit normalverteilten Werten (bei kleinen Stichproben; bei Ordinaldaten oder fehlender Normalverteilung: Nichtparametrische Tests)

Übersicht T-Tests



The screenshot shows the SPSS Daten-Editor window for the file '*ZA3762 allbus2004.SAV [DatenSet1]'. The 'Analysieren' menu is open, and the 'Mittelwerte vergleichen' option is selected. The sub-menu for 'Mittelwerte vergleichen' is also open, showing the following options:

- Mittelwerte...
- T-Test bei einer Stichprobe...
- T-Test bei unabhängigen Stichproben...
- T-Test bei verbundenen Stichproben...
- Einfaktorielle ANOVA...

The background table in the screenshot has the following data:

	Name	Typ	S
55	v55	Numerisch	1
56	v56	Numerisch	2
57	v57	Numerisch	4
58	v58	Numerisch	3
59	v59	Numerisch	1

- **T-Test bei einer Stichprobe:** prüft, ob der Mittelwert einer einzelnen Variablen von einer angegebenen Konstanten abweicht (z.B. schauen die Deutschen im Durchschnitt mehr als 2 Stunden TV?).
- **T-Test bei unabhängige Stichproben:** Vergleicht die Mittelwerte einer Variablen für zwei Fallgruppen (z.B. Unterschiede zw. Männern und Frauen im Einkommen).
- **T-Test bei verbundenen Stichproben:** Vergleicht den Mittelwert zweier Variablen für eine einzelne Gruppe (z.B. Blutdruckwerte in Zeitreihe).

Syntax T-Tests

T-TEST

One-sample tests:

```
T-TEST TESTVAL n /VARIABLE=varlist
```

Independent-samples tests:

```
T-TEST GROUPS=varname ({1,2**      }) /VARIABLES=varlist  
                        {value      }  
                        {value,value}
```

Paired-samples tests:

```
T-TEST PAIRS=varlist [WITH varlist [(PAIRED)]] [/varlist ...]
```

All types of tests:

```
[/MISSING={ANALYSIS**} [INCLUDE]]  
          {LISTWISE  }  
  
[/CRITERIA=CI ({0.95**)  
              {value }]
```

Examples

```
T-TEST GROUPS=WORLD(1,3) /VARIABLES=NTCPRI NTCSAL NTCPUR.
```

```
T-TEST PAIRS=TEACHER CONSTRUC MANAGER.
```

T-Test für unabhängige Stichproben

Beispiel: Es soll untersucht werden, ob zwischen Männern und Frauen ein signifikanter Unterschied bei der Bewertung Merkels besteht (Sympathie Merkel, Variablen f057_b) besteht.

H0: Zwischen Männern und Frauen besteht kein Unterschied in den Mittelwerten der Bewertung von Merkel

H1: Zwischen Männern und Frauen besteht ein Unterschied in den Mittelwerten der Bewertung von Merkel

→ H0 wird getestet

→ Zeigen sich signifikante Unterschiede zwischen den jeweiligen Gruppen in dieser Stichprobe, dann kann man davon ausgehen, dass diese Unterschiede auch in der Grundgesamtheit bestehen.

T-Test für unabhängige Stichproben

```
***Beispiel 1, Mittelwertunterschiede Sympathie Merkel nach
Geschlecht.
t-test groups geschlecht (0 1)
/var f057_b.
```

Gruppierungsvariable, UV

Testvariable, AV

SPSS Prozessor ist bereit

Ausgabe: Gruppenstatistiken

T-Test

[DatenSet2] C:\Dokumente und Einstellungen\anja.

Gruppenstatistiken



	geschlecht Ges...	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
f057_b Merkel: sympathisch	0 Mann	1274	2,85	1,731	,048
	1 Frau	1266	3,04	1,880	,053

Ausgabe: **Levene-Test** und T-Test

		Test bei unabhängigen Stichproben								
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
f057_b Merkel: sympathisch	Varianzen sind gleich	9,755	,002	2,671	2528	,008	,192	,072	,332	,051
	Varianzen sind nicht gleich			-2,670	2518,003	,008	-,192	,072	-,332	-,051

Levene-Test zur Überprüfung Varianzgleichheit:

H0: Varianzen sind gleich

H1: Varianzen sind nicht gleich

- ➔ Der Signifikanzwert von .002 besagt, dass bei einer Ablehnung der Nullhypothese (= die Varianzen sind gleich) die Irrtumswahrscheinlichkeit bei 0,2% liegt.
- ➔ Üblicherweise wird die H0 verworfen, wenn die Irrtumswahrscheinlichkeit unter 5% liegt (Sig. ≤ 0.05). D.h. hier kann die H0 verworfen werden: **die Varianzen sind wahrscheinlich nicht gleich.**
- ➔ Es werden somit die Testergebnisse in der Zeile „**Varianzen sind nicht gleich**“ herangezogen.

Ausgabe: Levene-Test und T-Test

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
f057_b Merkel: sympathisch	Varianzen sind gleich	9,755	,002	2,671	2528	,002	,192	,072	,332	,051
	Varianzen sind nicht gleich			-2,670	2518,003	,008	-,192	,072	-,332	-,051

T-Test für die Mittelwertgleichheit

- ➔ Der Mittelwertunterschied von $-0,192$ ist signifikant (p -Wert = $0,08$ d.h. Irrtumswahrscheinlichkeit bei Ablehnung der Nullhypothese: $0,8\%$).
- ➔ Das 95% - Konfidenzintervall: Das Intervall zwischen $-0,051$ und $-0,332$ enthält mit einer 95%igen Wahrscheinlichkeit den unbekanntem Populationsparameter (also hier den Mittelwertunterschied) D.h. der Mittelwertunterschied zwischen Männern und Frauen liegt in der GG (alle volljährigen Deutschen) vermutlich zwischen $-0,051$ und $-0,332$.

Irrtumswahrscheinlichkeit

<u>Irrtumswahrscheinlichk.</u>	<u>Bedeutung</u>	<u>Symbol</u>	
$P > 0,05$	(>5%)	nicht signifikant	n.s.
$P \leq 0,05$	(<= 5%)	signifikant	*
$P \leq 0,01$	(<= 1%)	sehr signifikant	**
$P \leq 0,001$	(<= 0,1%)	höchst signifikant	***

Weitere Beispiele