3 Vergleich zweier unverbundener Stichproben

3.1 Der Zweistichproben t-Test

Es wird vorausgesetzt, dass die beiden Teilstichproben $x_1, x_2, ..., x_{n_1}$ und $y_1, y_2, ..., y_{n_2}$ jeweils aus (voneinander unabhängigen) normalverteilten Grundgesamtheiten stammen. D.h. die erste Teilstichprobe besteht aus Realisierungen unabhängiger und identisch normalverteilter Zufallsvariablen $X_1, X_2, ..., X_{n_1}$ mit dem Erwartungswert μ_1 (und der Varianz σ_1^2) und die zweite analog aus Realisierungen von unabhängig und identisch normalverteilten Zufallsvariablen $Y_1, Y_2, ..., Y_{n_2}$ mit dem Erwartungswert μ_2 (und der Varianz σ_2^2), wobei die Zufallsvariablen beider Teilstichproben auch unabhängig sind. Wir beschreiben in diesem Kapitel den t-Test für gleiche und auch für ungleiche Varianzen σ_1^2 und σ_2^2. Man kann dann anhand der beiden Teilstichproben Hypothesentests bezüglich der Erwartungswerte durchführen. Wir testen die folgenden Hypothesen:

H_0: $\mu_1 = \mu_2$

gegen

H_1: $\mu_1 \neq \mu_2$

Kommen wir nun zu den Daten für unser Beispiel. Wir verwenden zwei Teilstichproben mit gleichem Umfang, wobei die Umfänge hier auch unterschiedlich groß sein können.

v1	v2
170	181
175	185
178	181
172	184

Nach der Dateneingabe können Sie den Button →**Vergleich zweier unverbundener Teilstichproben** anklicken, wobei Sie zuvor die beiden Variablen v1 und v2 unter diesem Menüpunkt auswählen müssen. Danach können Sie →**t-Test unverbundene Teilstichproben** auswählen und erhalten den folgenden Output:

t-Test für zwei unverbundene Stichproben und gleichen Varianzen

	Stichprobenumfänge (n1 und n2)	geschätzte Varianzen	geschätzte Standardabweichungen	arithmetische Mittel
Stichprobe 1	4	12.25	3.5	173.75
Stichprobe 2	4	4.25	2.061552812809	182.75

Prüfgröße t (Freiheitsgrade der t-Verteilung: 6)	-4.431293675256
p-Wert	0.0044

Test bezüglich der Gleichheit der Varianzen beider Stichproben

H0: sigma1 = sigma2
gegen
H1: sigma1 <> sigma2

Prüfgröße für F-Test (Freiheitsgrade der F-Verteilung: m = 3 und n = 3)	2.8823529411765
p-Wert	0.4078

---

t-Test für zwei unverbundene Stichproben und ungleiche Varianzen

H0: mue1 = mue2
gegen
H1: mue1 <> mue2

Prüfgröße t (Freiheitsgrade der t-Verteilung für Behrens-Fisher-Problem: 4.8579925651)	-4.431293675256
p-Wert	0.0114

Für die Prüfung auf Signifikanz wurde die Zahl der Freiheitsgrade auf 4 abgerundet!

Kommen wir zum Output. Im ersten Teil wird ein t-Test für gleiche Varianzen durchgeführt, d.h. es wird vorausgesetzt, dass die Varianzen σ_1^2 und σ_2^2 der Zufallsvariablen $X_1, X_2, ..., X_{n_1}$ und $Y_1, Y_2, ..., Y_{n_2}$ gleich sind. Für den t-Test im dritten Teil des Outputs können die Varianzen auch verschieden sein.

Im mittleren Teil des Outputs wird ein Test bezüglich der Gleichheit der Varianzen σ_1^2 und σ_2^2 durchgeführt. Damit dürfte der t-Test im ersten Teil nicht verwendet werden, wenn die Hypothese der Gleichheit beider Varianzen verworfen werden kann.

Die Prüfgröße des t-Tests für gleiche Varianzen ist:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}\cdot(1/n_1 + 1/n_2)}}$$

Dabei ist $\overline{x}_i$ der Mittelwert und s_i^2 die empirische Varianz der i-ten Teilstichprobe. Die Prüfgröße t ist (unter H_0) Realisierung einer mit $\upsilon = n_1 + n_2 - 2$ Freiheitsgraden t-verteilten Zufallsvariablen. Der p-Wert für den zweiseitigen Test ist dann:

$$\text{p-Wert} = 2(1 - F_{t_\upsilon}(|t|))$$

Im Beispiel ist t = -4.43129… und der p-Wert ≈ 0.0044.

Die Prüfgröße des t-Tests für ungleiche Varianzen ist:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Es ist zu beachten, dass die Freiheitsgrade für diesen Test nicht unbedingt eine natürliche Zahl ergeben. Die Prüfgröße t ist eine Realisierung einer (unter H_0) asymptotisch t-verteilten Zufallsvariablen T. Die Freiheitsgrade ν berechnen sich hier wie folgt:

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

Im Beispiel ist t = -4.43129…, ν = 4,85799… und der p-Wert ≈ 0.0114. Die Freiheitsgrade wurden bei der Berechnung des p-Wertes abgerundet. Dadurch wird der p-Wert vergrößert. Das der Wert für die Prüfgröße mit dem für gleiche Varianzen übereinstimmt liegt an den gleichen Stichprobenumfängen n_1 und n_2 im Beispiel.

Die Nullhypothese könnte in beiden t-Tests auf einem Niveau von 5% verworfen werden, da beide p-Werte ≤ 0,05 sind.

Verwendet man die Freiheitsgrade ν bei der t-Verteilung, so würde sich beim Test für ungleiche Varianzen ein p-Wert von 0,007297… ergeben.

Kommen wir zum Schluss zum mittleren Test. Hier kann man bezüglich der Varianz folgendes testen:

H_0: $\sigma_1^2 = \sigma_2^2$

gegen

H_1: $\sigma_1^2 \neq \sigma_2^2$

Wenn man bei diesem Test die Prüfgröße so definiert, dass man die größere der empirischen Varianzen in den Zähler schreibt, dann muss man nur eine Seite überprüfen, obwohl es sich um einen zweiseitigen Test handelt.

Falls $s_1^2 > s_2^2$ ist

$$f = s_1^2 / s_2^2 ,$$

sonst gilt:

$$f = s_2^2 / s_1^2$$

Diese Definition findet man in verschiedenen Büchern, hier sollte man aber aufpassen, dass der Nenner nicht zu klein wird, andernfalls sollte man die linke Seite überprüfen. Ist die Prüfgröße wie oben definiert, dann ergibt sich folgender p-Wert:

$$\text{p-Wert} = 1 - F_{F_{n_1-1,n_2-1}}(f)$$

Also ist F die Verteilungsfunktion der F-Verteilung mit n_1 - 1 und n_2 - 1 Freiheitsgraden.

Bei der F-Verteilung wird bei allen Verfahren vom System im Output der p-Wert nur dann exakt berechnet, wenn für die beiden Freiheitsgrade der F-Verteilung m_1 und m_2 die Bedingung $m_1 + m_2 \leq 240$ erfüllt ist, oder wenn $m_1 = 1$ und m_2 beliebig ist, oder wenn $m_2 = 1$ und m_1 beliebig ist. Sonst wird für $m_1 + m_2 > 240$ die F-Verteilung über die Chi-Quadrat-Verteilung approximiert.

Im Online-Modul ist die Bedingung $m_1 + m_2 \leq 240$ bei diesem Test immer erfüllt, da die beiden Stichprobenumfänge jeweils höchsten den Wert 100 haben können.

Im Beispiel gilt:

$$f = 2.882352\ldots \text{ und p-Wert} \approx 0{,}4078.$$

Somit kann die Nullhypothese (der Gleichheit beider Varianzen) auf keinem gängigen Signifikanzniveau verworfen werden (selbst nicht auf einem für Anpassungstest üblichen von bis zu 25%).

Bemerkung:
Bei unabhängigen Stichproben hätte man die Daten auch im großen Menü in der folgenden Form eingeben können: In der Variable v1 wird die Gruppenzugehörigkeit gespeichert, z.B. mit dem Wert 1 für die erste und 2 für die zweite Gruppe, man könnte aber auch Buchstaben für die Gruppen verwenden. In der Variablen v2 wird dann die entsprechende Größe gespeichert.

v1	v2
1	170
1	175
1	178
1	172
2	181
2	185
2	181
2	184

Danach kann man unter dem Button „Vergleich unverbundener Teilstichproben“ die Variable v1 als Faktorvariable und die Variable v2 als abhängige Variable auswählen und dann auf →**Vergleich unverbundener Teilstichproben** und →**Berechnung** klicken. Hier werden die entsprechenden Verfahren angeboten (je nachdem ob zwei Gruppen oder mehr als zwei Gruppen vorliegen).

Umsetzung mit SAS:

```
data dat1;
input x y;
datalines;
1 170
1 175
1 178
1 172
2 181
2 185
2 181
2 184

run;
proc ttest data=dat1;
class x;
var y;
run;
```

Das SAS System

Die Prozedur TTEST
Variable: y

x	N	Mittelwert	Std.abw.	Std.fehler	Minimum	Maximum
1	4	173.8	3.5000	1.7500	170.0	178.0
2	4	182.8	2.0616	1.0308	181.0	185.0
Diff (1-2)		-9.0000	2.8723	2.0310		

x	Methode	Mittelwert	95% CL Mittelwert		Std.abw.	95% CL Std Dev	
1		173.8	168.2	179.3	3.5000	1.9827	13.0499
2		182.8	179.5	186.0	2.0616	1.1678	7.6866
Diff (1-2)	**Gepoolt**	-9.0000	-13.9697	-4.0303	2.8723	1.8509	6.3250
Diff (1-2)	**Satterthwaite**	-9.0000	-14.2671	-3.7329			

Methode	Varianzen	DF	t-Wert	Pr > |t|
Gepoolt	Gleich	6	-4.43	0.0044
Satterthwaite	Ungleich	4.858	-4.43	0.0073

Gleichheit der Varianzen

Methode	Zähler Freiheits-grade	Nenner Freiheits-grade	F-Statistik	Pr > F
Folded F	3	3	2.88	0.4078