

**Aufgabe 37**

Welches ist der passende Test? Begründen Sie jeweils Ihre Antwort und formulieren Sie auch die entsprechende Null- und Alternativhypothese.

- a) Sie wollen untersuchen, ob die Lebenszufriedenheit in Bayern größer ist als in Hessen (die Lebenszufriedenheit wird auf einer recht feinen Skala gemessen).
- b) Sie wollen untersuchen, ob ein (normalverteilter) Aggressionsscore von 17 straffälligen Jugendlichen nach einem Deeskalationstraining gesunken ist.
- c) Sie wollen untersuchen, ob sich die Körpergröße von Mädchen und Jungen in einer Krippengruppe (Gruppengröße 12 Kinder) unterscheidet.
- d) Sie wollen untersuchen, ob sich die Zustimmung (gemessen auf einer 13-stufigen Skala) der Bewohner eines Stadtviertels zu einem umstrittenen Bauprojekt nach einer Infoveranstaltung geändert hat.
- e) Sie wollen untersuchen, ob sich die Scheidungsraten zwischen zwei Städten unterscheiden.
- f) Sie wollen untersuchen, ob eine Gruppe von Personen die Nachrichten in ARD oder ZDF als aktueller einstuft.
- g) Sie wollen untersuchen, ob es einen Zusammenhang zwischen höchsten Bildungsabschluss und Wohnviertel in München gibt.
- h) Sie wollen untersuchen, ob sich die Beurteilung von „Die Berufsaussichten für Bachelorabsolventen sind schlechter als für Master-Absolventen“ gemessen anhand einer fünfstufigen Skala (stimme zu – stimme nicht zu) zwischen Studierenden der Geistes- und der Naturwissenschaften unterscheidet.

### Aufgabe 38

In der folgenden Tabelle ist die Anzahl aller vollzeit arbeitenden Mütter (das jüngste Kind ist unter 15 Jahre alt) getrennt nach Ost- und Westdeutschland gegeben. Die Daten beziehen sich auf eine Stichprobe und stammen vom Mikrozensus 2010 <sup>1</sup>.

	Vollzeit	
	ja	nein
Ostdeutschland	401	343
Westdeutschland	699	2456

Gibt es einen zum Niveau  $\alpha = 0.01$  signifikanten Zusammenhang zwischen diesen beiden Variablen?

### Aufgabe 39

Betrachten Sie den SPSS-Output zur linearen Regression über die Nettomiete (nm) von Münchner Wohnungen in Abhängigkeit von folgenden Variablen:

- rooms: Anzahl Zimmer
- zh0: Zentralheizung vorhanden (0: ja, 1: nein)

**Koeffizienten<sup>a</sup>**

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	201,099	65,530		3,069	,003
	rooms	164,664	22,980	,625	7,166	,000
	zh0	-213,234	71,988	-,258	-2,962	,004

a. Abhängige Variable: nm

- Interpretieren Sie die Regressionskoeffizienten.
- Welche Hypothesen werden bei der Überprüfung der Signifikanz der Variablen rooms und zh0 gegeneinander getestet?
- Erläutern Sie kurz, was der p-Wert angibt und wie man ihn interpretiert.

<sup>1</sup>Quelle:

<http://www.bpb.de/nachschlagen/zahlen-und-fakten/soziale-situation-in-deutschland/61606/erwerbstaetigkeit-nach-alter-des-juengsten-kindes>

- d) Testen Sie, ob die Variablen `rooms` und `zh0` jeweils einen signifikanten ( $\alpha = 0.05$ ) Einfluss haben.
- e) Berechnen Sie ein Konfidenzintervall ( $\gamma = 0.95$ ) für den Parameter der Einflußgröße `rooms` ( $n = 2053$ ).

### Aufgabe 40

In einer Studie zu Lesefähigkeiten von Schülern (Christa Kieferle, Pädagogik, LMU) wurden unter anderem die folgenden Variablen von 180 Schülern aus 8 Klassen (3. und 4. Klasse) erhoben:

- Fehlerzahl: Anzahl der Fehler bei einem Lesetest
- Lesezeitmin: Leseförderzeit in Minuten (Leseförderung in der Schule)
- Jahrgang: Dummyvariable mit Ausprägungen 3. Klasse = 1, 4. Klasse = 0
- Geschlecht: Dummyvariable mit Ausprägungen männlich = 1, weiblich = 0
- Lesen: Variable zur Frage „Wie oft liest du außerhalb der Schule?“ mit Ausprägungen oft, mittel, selten.

Es wurde eine Regressionsanalyse durchgeführt mit Fehleranzahl als Zielgröße und den anderen Variablen im Datensatz als Einflussgrößen:

```
Call:
lm(formula = Fehlerzahl ~ Lesezeitmin + Jahrgang + Geschlecht +
    Lesenoft + Lesenselten)

Residuals:
    Min       1Q   Median       3Q      Max
-14.354  -5.574  -1.361   4.344  25.291

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  14.80650    1.76753   8.377 1.75e-14 ***
Lesezeitmin  -0.08726    0.03272  -2.667  0.00838 **
Jahrgang      5.90591    1.17928   5.008 1.34e-06 ***
Geschlecht    3.16466    1.22758   2.578  0.01077 *
Lesenoft     -2.93325    1.38026  -2.125  0.03499 *
Lesenselten   4.94820    1.90625   2.596  0.01024 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.859 on 174 degrees of freedom
Multiple R-squared:  0.3081, Adjusted R-squared:  0.2882
F-statistic: 15.5 on 5 and 174 DF,  p-value: 1.34e-12
```

- a) Wie wurde die Variable **Lesen** umkodiert um sie im Regressionsmodell berücksichtigen zu können?
- b) Interpretieren Sie die Schätzungen der Regressionskoeffizienten.
- c) Welche der Variablen haben einen signifikanten ( $\alpha = 0.05$ ) Einfluss?
- d) Berechnen Sie ein Konfidenzintervall zum Sicherheitsgrad von 99% für den Parameter der Einflussgröße **Geschlecht**.