

Aufgabe 1

In einer Umfrage wurden Personen zu ihrer subjektiven Lebenszufriedenheit befragt. Die abhängige Variable ZUFRIEDENHEIT wurde als Antwort auf die Frage „Wie zufrieden sind Sie gegenwärtig, alles in allem, mit Ihrem Leben?“ mit einer 11-stufigen Likert-Skala gemessen (0 entspricht ganz und gar unzufrieden und 10 entspricht ganz und gar zufrieden). Folgende unabhängige Variablen wurden betrachtet:

- das Haushaltseinkommen (logarithmiert, Variable LNEINKOMMEN),
- die Anzahl der im Haushalt wohnenden Personen (logarithmiert, Variable LNPERSONEN),
- der Arbeitslosenstatus (Variable ARBEITSLOS, 0 = arbeitstätig, 1 = arbeitslos),
- der Gesundheitszustand (Variable GESUNDHEIT, 0 = gesund, 1 = krank),
- und das Geschlecht (Variable GESCHLECHT, 0 = männlich, 1 = weiblich)

Die Variablen ZUFRIEDENHEIT, LNEINKOMMEN und LNPERSONEN wurden als metrisch vorausgesetzt. Die restlichen Variablen wurden als dichotome Variablen behandelt. Es wurde ein lineares Modell an die Daten angepasst. Die R^1 -Ausgabe des erstellten Modells ist auf der nächsten Seite angegeben.

Call:

```
lm(formula = ZUFRIEDENHEIT ~ LNEINKOMMEN + LNPERSONEN + ARBEITSLOS +  
    GESUNDHEIT + GESCHLECHT)
```

Residuals:

Min	1Q	Median	3Q	Max
-6.2506	-0.8052	0.2262	1.0478	4.7521

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.971588	0.592824	6.699	2.64e-11	***
LNEINKOMMEN	0.673659	0.077370	8.707	< 2e-16	***
LNPERSONEN	-0.418004	0.082887	-5.043	4.95e-07	***
ARBEITSLOS	-1.283929	0.127883	-10.040	< 2e-16	***
GESUNDHEIT	-0.744085	0.035989	-20.676	< 2e-16	***
GESCHLECHT	0.006614	0.066265	0.100	0.92	

¹<http://www.r-project.org/>

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.565 on 2245 degrees of freedom
(459 observations deleted due to missingness)

Multiple R-squared: 0.2501, Adjusted R-squared: 0.2484

F-statistic: 149.7 on 5 and 2245 DF, p-value: < 2.2e-16

- a) Interpretieren Sie die **R**-Ausgabe. Gehen Sie dabei insbesondere auf den Einfluss der Variable GESCHLECHT auf die Zufriedenheit ein.
- b) Diskutieren Sie die Annahme des metrischen Skalenniveaus der Variable ZUFRIEDENHEIT.

Aufgabe 2

Aus der Studentenstatistik seien folgende Anteile an Studienfächern bekannt:

- Der Anteil der Studierenden mit Hauptfach Soziologie beträgt 10%.
- Der Anteil der Studierenden mit Hauptfach BWL beträgt 15%.
- Der Anteil der Studierenden mit Hauptfach katholische Theologie beträgt 3%.
- In diesem hypothetischen Beispiel gibt es kein Doppelstudium, d.h. jeder Student hat nur ein Hauptfach.

Für eine Untersuchung wird aus allen Studierenden der LMU ein(e) Student(in) anhand seiner Matrikelnummer ausgelost (Annahme: Alle Studierende haben die gleiche Wahrscheinlichkeit gezogen zu werden).

Von der ausgelosten Person interessieren uns besonders folgende Merkmale:

- Hauptfach Soziologie? Ja/Nein
- Hauptfach BWL? Ja/Nein
- Hauptfach Kath. Theologie? Ja/Nein

- a) Begründen Sie, warum Sie von den Anteilen auf Wahrscheinlichkeiten schließen können.
- b) Berechnen Sie für die gezogene Person folgende Wahrscheinlichkeiten und kennzeichnen Sie die zugrundeliegenden Mengen in einem Venn-Diagramm:
 - i) $P(\text{nicht Hauptfach Soziologie})$
 - ii) $P(\text{Hauptfach Soziologie} \cup \text{Hauptfach BWL})$
 - iii) $P(\text{Hauptfach Soziologie} \cup \text{Hauptfach BWL} \cup \text{Hauptfach Kath. Theol.})$