

Übung zur Vorlesung „Empirische Ökonomie 1“

Übungsblatt 1: Lineares Regressionsmodell I

Aufgabe 1:

Der Datensatz `Andy.gdt` enthält Informationen über Umsätze und Werbeausgaben von 75 Filialen einer großen Fast-Food-Kette. Die Variablen sind:

`sales`: monatlicher Umsatz in \$1,000

`advert`: monatliche Ausgaben für Werbung in \$1,000.

Unter „Grundlegende Statistiken“ erhalten Sie in GRETL den folgenden Output :

Grundlegende Statistiken, benutze die Beobachtungen 1 - 75

	arith. Mittel	Median	Minimum	Maximum
<code>sales</code>	77,38	76,500	62,400	91,200
<code>advert</code>	1,84	1,8000	0,50000	3,1000

	Std. Abw.	Var'koeff.	Schiefe	Überwölbung
<code>sales</code>	6,49	0,083859	-0,010631	-0,74467
<code>advert</code>	0,83	0,45102	0,037087	-1,2951

Außerdem beträgt die Stichproben-Korrelation zwischen `sales` und `advert` $r = 0,22$. (Die Korrelation $r_{XY} = \frac{s_{XY}}{s_X s_Y}$ finden Sie in GRETL unter „Ansicht“ \Rightarrow „Korrelationsmatrix“.)

Betrachten Sie ein Regressionsmodell der Form

$$\text{sales}_i = \beta_0 + \beta_1 \cdot \text{advert}_i + u_i,$$

wobei der Index $i = 1, \dots, 75$ die Fast-Food-Restaurants kennzeichnet.

- Bestimmen Sie den OLS-Schätzer $\hat{\beta}_1$ für den Modellparameter β_1 .
- Bestimmen Sie den OLS-Schätzer $\hat{\beta}_0$ für den Modellparameter β_0 .
- Filialleiter Mike Meatloaf hat bisher \$1,500 für Werbung ausgegeben und überlegt nun, die Werbeausgaben auf \$2,000 zu steigern. Um wie viel würde sich sein vorhergesagter Umsatz durch diese Maßnahme erhöhen?

Aufgabe 2:

In dieser Aufgabe beschäftigen wir uns mit dem Zusammenhang zwischen der Anzahl abgeschlossener Bildungsjahre einer Person und der Entfernung der besuchten Highschool zum nächsten College. Dazu verwenden wir Daten einer Zufallsstichprobe mit 3796 ehemaligen Highschoolschülern (`colldist.gdt`). Den Datensatz finden Sie auf der Homepage dieser Veranstaltung. Sie können ihn auf Ihrer Festplatte speichern und dann in Gretl unter `Datei` → `Öffne Daten` → `Benutzerdatei` öffnen.

Die vorhandenen Variablen sind:

`dist` Entfernung der Highschool zum nächsten College in 10 Meilen
`ed` Anzahl der abgeschlossenen Bildungsjahre

Betrachten Sie ein Regressionsmodell der Form

$$ed_i = \beta_0 + \beta_1 dist_i + u_i, \quad i = 1, \dots, n$$

und schätzen Sie dieses in Gretl.

- Wie lautet die inhaltliche Interpretation des geschätzten Koeffizienten $\hat{\beta}_1$ in dieser Regression?
- Wie lautet die quantitative Interpretation des geschätzten Koeffizienten $\hat{\beta}_0$ in dieser Regression?
- Wie hoch ist die vorhergesagte Anzahl abgeschlossener Bildungsjahre einer Person, deren Highschool 20 Meilen vom nächsten College entfernt ist?
- Erklärt die Entfernung der Highschool zum nächsten College einen großen Anteil der Varianz der abgeschlossenen Bildungsjahren?

Aufgabe 3:

Sie verfügen über einen Datensatz mit Angaben von 129 Teilnehmern, die im Rahmen einer Sportmesse freiwillig an einer Befragung zum Zusammenhang zwischen Bierkonsum und Fitness-Leidenschaft teilgenommen haben. Darunter befindet sich auch ein Bodybuilder. In der Befragung wurde lediglich erfasst, wieviele Liter Bier die Befragte Person pro Woche durchschnittlich zu sich nimmt und wie viele Stunden sie im Monat Sport treibt.

Nennen und erläutern Sie (genau) drei Probleme, die in dieser Anwendung dazu führen können, dass eine OLS-Schätzung des Effekts der Fitness-Leidenschaft auf den Bierkonsum nicht angebracht ist.