

Aufgabenblatt 2 Ökonometrie MSc im Sommersemester

Aufgabe 1

Betrachten Sie den Wooldridge Datensatz `sleep75`. Der Datensatz enthält Informationen zu $n = 706$ Arbeitnehmern.

- Bestimmen Sie das durchschnittliche Alter und die durchschnittliche Schlafdauer der Arbeitnehmer aus dem Datensatz.
- Wie viele 23-jährige befinden sich im Datensatz? Wie alt ist der älteste Arbeitnehmer?
- Zeichnen Sie ein Streudiagramm, in welchem Sie die Variable `sleep` (Schlafdauer) auf der y-Achse gegen die Variable `age` (Alter) auf der x-Achse abtragen. Lässt sich ein Zusammenhang erkennen?
- Schätzen Sie die Parameter des Regressionsmodells

$$\text{sleep}_i = \beta_0 + \beta_1 \text{age}_i + u_i$$

mittels OLS und zeichnen Sie die geschätzte Regressionsgerade zusammen mit dem Streudiagramm in eine Grafik. Interpretieren Sie das Ergebnis.

- Welche Schlafdauer ist im Durchschnitt bei einem 25-jährigen Arbeitnehmer zu erwarten? Welche bei einem 60-jährigen Arbeitnehmer?
- Schätzen Sie die Parameter des Regressionsmodells erneut für den auf die ersten zehn Beobachtungen eingeschränkten Datensatz. Gibt es große Unterschiede zu den auf der gesamten Stichprobe basierenden Schätzungen?
- Der Datensatz beinhaltet auch die Variablen `educ` (Anzahl der Bildungsjahre) und `lhrwage` (logarithmierter Stundenlohn in \$). Schätzen Sie die Regressionsgleichung

$$\ln(\text{hrwage}_i) = \beta_0 + \beta_1 \text{educ}_i + u_i,$$

dabei bezeichnet `hrwage` den (unlogarithmierten) Stundenlohn. Um wieviel Prozent steigt der erwartete Stundenlohn, wenn sich die Anzahl der Bildungsjahre um ein Jahr erhöht?

Aufgabe 2

Öffnen Sie den Wooldridge Datensatz `wage1`. Dieser Datensatz umfasst 526 Beobachtungen bzgl. Stundenlöhnen und einer Reihe weiterer Variablen von Personen aus den USA, gemessen in Preisen aus dem Jahr 1976.

- Bestimmen Sie die geringste, mittlere und höchste Bildungsdauer sowie die entsprechenden Kennziffern für den Stundenlohn. Wie beurteilen Sie den mittleren Stundenlohn?
- Der Konsumentenpreisindex für die USA (Basisjahre 1982–1984 entsprechen dem Wert 100) betrug im Jahr 1976 56.9 und im Jahr 2016 240.00. Berechnen Sie den mittleren Stundenlohn im Preisniveau von 2016 ausgedrückt. Wie schätzen Sie diesen jetzt ein?

c) Schätzen Sie die folgenden Modelle mittels OLS:

$$\begin{aligned} \text{wage}_i &= \beta_0 + \beta_1 \text{educ}_i + u_i \\ \ln(\text{wage}_i) &= \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{exper} + u_i \end{aligned}$$

und gehen Sie auf die Anpassung sowie Interpretation der Modelle, insbesondere der Parameter, ein.

Aufgabe 3

Der Wooldridge Datensatz `bwght` enthält Geburtendaten aus den Vereinigten Staaten, unter anderem das Geburtsgewicht in Unzen¹ (`bwght`) und die durchschnittliche Anzahl der von der Mutter in der Schwangerschaft gerauchten Zigaretten (`cigs`) pro Tag. Die folgende einfache Regression wurde anhand von 1388 Geburtsdaten mittels OLS geschätzt:

$$\widehat{\text{bwght}} = 119.77 - 0.514 \text{ cigs}.$$

- Wie groß ist das prognostizierte Geburtsgewicht, wenn die Mutter während der Schwangerschaft gar nicht raucht (`cigs = 0`)? Wie schaut es aus, falls die Mutter im Schnitt eine Schachtel (à 20 Zigaretten) pro Tag raucht? Interpretieren Sie Ihre Beobachtung.
- Lässt sich der Mittelwert von `cigs` sinnvoll interpretieren? Nennen Sie Gründe.
- Wie groß müsste `cigs` sein um ein Geburtsgewicht von 40 bzw. 125 Unzen zu prognostizieren? Was folgt daraus für die Güte der linearen Anpassung?
- Ist die Anzahl pro Tag gerauchter Zigaretten Ihrer Meinung nach die einzige für das Geburtsgewicht relevante Einflussgröße? Falls nein, nennen Sie weitere relevante Größen.
- Inwieweit könnte die Bedingung $\text{Cov}(x, u) = 0$ im Regresionsmodell

$$\text{bwght}_i = \beta_0 + \beta_1 \text{cigs}_i + u_i$$

verletzt sein? Diskutieren Sie im Hinblick auf Aufgabenteil d).

¹Eine Unze (von lateinisch *uncia*) ist eine nichtmetrische Maßeinheit der Masse. Eine Unze entspricht etwa 28,35 Gramm.