

Übung zur Vorlesung „Empirische Ökonomie 1“

Übungsblatt 6: Multiple Regression und Hypothesentests

Allgemeiner Hinweis: Bitte verwenden Sie für alle Regressionen „robuste“ Standardfehler. Öffnen Sie den Beispiel-Datensatz von Greene `greene22_2` („Fair data on extra-marital affairs“) in GRETL. Bei den Beobachtungseinheiten handelt es sich um eine Zufallsstichprobe von 601 verheirateten Amerikanern aus der uns interessierenden Grundgesamtheit. Die Variablen sind:

- Y Anzahl der außerehelichen Affären im letzten Jahr
- Z1 „Dummyvariable“ für das Geschlecht [0 = weiblich, 1 = männlich]
- Z2 Alter [Jahre]
- Z3 Dauer der Ehe [Jahre]
- Z4 „Dummyvariable“ für gemeinsame Kinder [0 = nein, 1 = ja]
- Z5 Religiosität [Skala von 1 = anti-religiös bis 5 = sehr religiös]
- Z6 Bildung [Jahre]
- Z7 „Prestige-Score“ des Berufes [Skala von 1 = wenig bis 7 = viel Prestige]
- Z8 Selbstbewertung der Ehe [Skala von 1 = sehr unglücklich bis 5 = sehr glücklich]

Aufgabe 1:

Sie möchten untersuchen, ob das Alter eines Menschen die Anzahl seiner außerehelichen Affären beeinflusst. Betrachten Sie zunächst ein Regressionsmodell der Form

$$Y_i = \beta_0 + \beta_1 Z2_i + u_i$$

und nehmen Sie zunächst an, dass die Annahmen A1–A3 für dieses Regressionsmodell erfüllt sind.

- (a) Schätzen Sie dieses Modell in GRETL. Wie hoch ist der geschätzte Einfluss des Alters?
- (b) Testen Sie die Nullhypothese „Das Alter hat keinen Einfluss auf die Anzahl der außerehelichen Affären“ bei einem Signifikanzniveau von 5%. Was ist das Ergebnis?
- (c) Testen Sie die Nullhypothese „Bei einem um zehn Jahre älteren Menschen ist die Anzahl der außerehelichen Affären durchschnittlich um 0,1 höher“ bei einem Signifikanzniveau von 5%. Was ist das Ergebnis?

- (d) Ein Kollege behauptet, dass die Anzahl der außerehelichen Affären mit der Zahl der Ehejahre steige. Falls er recht haben sollte: Was würde dies für die Gültigkeit der Annahmen A1–A3 sowie für die Konsistenz und Erwartungstreue des Schätzers $\hat{\beta}_1$ bedeuten?
- (e) Nehmen Sie nun die Zahl der Ehejahre in die Schätzung auf. Schätzen Sie also ein Modell der Form

$$Y_i = \beta_0 + \beta_1 Z2_i + \beta_2 Z3_i + u_i$$

und nehmen Sie an, dass die Annahmen M1–M4 für dieses Regressionsmodell erfüllt sind. Testen Sie die Nullhypothese „Das Alter hat keinen Einfluss auf die Anzahl der außerehelichen Affären“ bei einem Signifikanzniveau von 5%. Was ist das Ergebnis?

Aufgabe 2:

Betrachten Sie weiterhin das Modell aus Aufgabe 1 (e):

$$Y_i = \beta_0 + \beta_1 Z2_i + \beta_2 Z3_i + u_i$$

Gehen Sie davon aus, dass die Annahmen M1–M4 des linearen Regressionsmodells mit mehreren erklärenden Variablen erfüllt sind. Um Hypothesen in Modellen mit mehreren Regressoren zu testen, ist es nützlich, die Matrixschreibweise zu verwenden. Ausgedrückt werden die Hypothesen dann gemäß der Notation aus der Vorlesung in der Form $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$, wobei $\boldsymbol{\beta} = (\beta_0 \ \beta_1 \ \beta_2)'$ der Spaltenvektor der Koeffizienten ist. Der Spaltenvektor \mathbf{r} enthält lediglich Zahlen. Mit welcher Matrix \mathbf{R} können die folgenden Hypothesen in dieser Form ausgedrückt werden?

- (a) H_0 : Weder das Alter noch die Anzahl der Ehejahre hat einen Einfluss auf die Anzahl der Affären.
- (b) H_0 : Eine 40-jährige Person, die 20 Jahre verheiratet ist, hat im Schnitt zwei Affären mehr, als eine 30-Jährige Person, die 5 Jahre verheiratet ist.

Aufgabe 3:

Testen Sie nun die Hypothesen aus Aufgabe 2 jeweils auf dem 5% Signifikanzniveau. Wie lautet das Ergebnis?

- (a) H_0 : Weder das Alter noch die Anzahl der Ehejahre hat einen Einfluss auf die Anzahl der Affären.
- (b) H_0 : Eine 40-jährige Person, die 20 Jahre verheiratet ist, hat im Schnitt zwei Affären mehr, als eine 30-Jährige Person, die 5 Jahre verheiratet ist.

Hinweis: Solche Tests lassen sich in GRETL einfach durchführen: In dem Fenster mit den Regressionsergebnissen wählen Sie im Menü „Tests“ den Menüpunkt „lineare Restriktionen“. In diesem Eingabefeld können Sie nun mehrere Restriktionen eingeben – eine pro Zeile. Dabei verwenden Sie für die Parameter die Bezeichnungen **b1** bis **b3**.

Bitte beachten Sie dabei, dass der Parameter β_0 mit **b1** angesprochen wird, der Parameter β_1 mit **b2** usw. Beispiel: Wenn Sie die Nullhypothese $\beta_1 = 1$ testen wollen, lautet der entsprechende Ausdruck in GRETL: **b2=1**. Am Einfachsten interpretieren Sie die Ergebnisse, indem Sie auf den p -Wert achten.