Inferenz im multiplen Regressionsmodell

Kapitel 4, Teil 1

Ökonometrie I
Michael Hauser

Inhalt

- Annahme normalverteilter Fehler
- Stichprobenverteilung des OLS Schätzers
- t-Test und Konfidenzintervall für Parameter
- t-Verteilung im Anhang

Genauso wie man eine Realisation des Stichprobenmittels $\overline{x}_n$ für eine gezogene Stichprobe erhält, liefert die OLS Schätzung eine Realisation für die Koeffizienten der Regressionsgleichung.

Wir diskutieren hier die Stichprobenverteilung der OLS Schätzer.

Annahme normal verteilter Fehler

Annahme normal verteilter Fehler

Zur Ermittlung der Stichprobenverteilung, benötigen wir in kleinen Stichproben eine weitere Annahme. (Alternativ könnte man alle Überlegungen auch asymptotisch, analog zum ZGWS, durchführen.)

MLR.6 Der Fehler u ist unabhängig von den erklärenden Variablen $x_1, \ldots, x_k$ und normal verteilt

$$u \sim \mathrm{N}(0, \sigma^2)$$

Das Mittel von u ist 0, seine Varianz σ^2.

Wir verwenden die Abkürzung *iid* für *unabhängig, identisch verteilt*, bzw independently, identically distributed.

Bem: Bei multivariater Normalvtlg fallen Unkorreliertheit und Unabhängigkeit zusammen.

Annahme normal verteilter Fehler

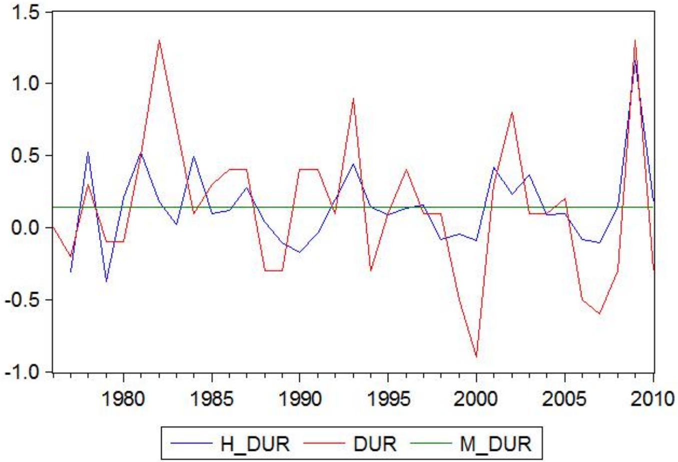
Unser Modell lautet nun

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u \qquad \text{mit} \qquad u \;\text{ iid }\; \mathrm{N}(0, \sigma^2)$$

Graphik:
Streudiagramm, Regressionsgerade, N-Vtlg der Fehler zentriert um die Regressionsgerade

Die bedingte Verteilung von y

Unbedingte und bedingte Erwartung: Okun



DUR . . . Δu_t, M_DUR . . . Mittelwert von Δu_t,
H_DUR . . . $\widehat{\Delta u}_t = \widehat{\mu}_{\Delta u|\rho}$

Unbedingte und bedingte Erwartung, Bsp

Bsp: Okun's Modell lautet $\Delta u = 0.500 - 0.165\,\rho + \widehat{\epsilon}$

- Der Mittelwert der Veränderung der Al-Rate ist

$$\overline{\Delta u} = 0.140$$

  Er ist konstant über die Zeit. Wir nennen ihn *unbedingtes Mittel* von Δu.

- Wenn wir die Information über die Wachstumsrate des BIP, ρ, verwenden, können wir Δu besser beschreiben.

$$\widehat{\Delta u}_t = 0.500 - 0.165\,\rho_t$$

  Wir bezeichnen $\widehat{\Delta u}_t$ als das *bedingte Mittel* von Δu gegeben ρ:

$$\widehat{\Delta u} = \widehat{\mu}_{\Delta u|\rho}$$

Modelle zu unbedingter und bedingter Erwartung

- Das Modell zum *unbedingten* Mittel ist

$$\Delta u_t = \beta_0 + \epsilon_t$$

Wenn ϵ_t normal verteilt ist ($\sigma_\epsilon^2 = \sigma_{\Delta u}^2$)

$$\Delta u \;\sim\; \mathrm{N}(\beta_0, \sigma_{\Delta u}^2)$$

- Das Modell zum *bedingten* Mittel ist

$$\Delta u_t = \beta_0 + \beta_1\,\rho_t + \epsilon_t$$

Die Mittelung findet hier nur über den Fehler ϵ statt.

$$\Delta u|\rho \;\sim\; \mathrm{N}(\beta_0 + \beta_1\,\rho\,,\, \sigma_\epsilon^2)$$

Skip: Unbedingte und bedingte Erwartung allgemein

Wir unterscheiden

$$\mathrm{E}(y) = \mu_y \qquad \text{die } \textit{unbedingte Erwartung}$$

und die *bedingte Erwartung* bei gegebenen x-Variablen

$$\mathrm{E}(y|\boldsymbol{X}) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k = \mu_{y|\boldsymbol{x}}$$

Man schreibt auch um die Abhängigkeit von y von den x-Variablen anzuzeigen

$$y|\boldsymbol{X} = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$$

Skip: $y|\boldsymbol{X}$ ist normal verteilt

- Der **Erwartungswert von** $y|\boldsymbol{X}$ ist eine Funktion der x-Variablen:

$$\mathrm{E}(y|\boldsymbol{X}) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k = \mu_{y|\boldsymbol{x}}$$

- $$y|\boldsymbol{X} = \mu_{y|\boldsymbol{x}} + u$$

- Die **Varianz von** $y|\boldsymbol{X}$ ist $\mathrm{V}(u) = \sigma^2$:

$$\mathrm{V}(y|\boldsymbol{X}) = \mathrm{E}[(y|\boldsymbol{X} - \mu_{y|\boldsymbol{x}})^2] = \mathrm{E}(u^2) = \sigma^2$$

- $y|\boldsymbol{X}$ **ist normal verteilt**, da u normal verteilt ist:

$$y|\boldsymbol{X} \;\sim\; \mathrm{N}(\mu_{y|\boldsymbol{x}}, \sigma^2)$$

Skip: $y|\boldsymbol{X}$ ist normal verteilt

Zwischenrechnung:

$$\begin{aligned} \mathrm{E}(y|\boldsymbol{X}) &= \mathrm{E}(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u \mid \boldsymbol{X}) \\ &= \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \mathrm{E}(u|\boldsymbol{X}) \\ &= \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k \end{aligned}$$

da $\mathrm{E}(x_j|\boldsymbol{X}) = \mathrm{E}(x_j|1, x_1, \ldots, x_k) = x_j$.

$$y|\boldsymbol{X} = \mu_{y|\boldsymbol{x}} + u$$

Verteilung des OLS Schätzers

Der OLS Schätzer ist normal verteilt

- Unter unseren Annahmen ist der **OLS-Schätzer normal verteilt**

$$b_j \sim \mathrm{N}(\beta_j, \sigma^2[(\boldsymbol{X}'\boldsymbol{X})^{-1}]_{jj}) \qquad \text{mit} \qquad j = 0, \ldots, k$$

  Dies ermöglicht nun Tests für die Koeffizienten anzugeben.

Herleitung:

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \qquad \text{und} \qquad \boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$$

Einsetzen gibt

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'[\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}] = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}$$

$\mathrm{E}(\boldsymbol{b}) = \boldsymbol{\beta}$. Da $\boldsymbol{u}$ normal vtlt ist, ist auch $\boldsymbol{b}$ bei gegebenem $\boldsymbol{X}$ normal vtlt mit der bekannten Varianz-Kovarianz-Matrix, $\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$.

Der OLS Schätzer ist normal verteilt

Wir verwenden im Weiteren den Standardfehler des geschätzten Koeffizienten b_j

$$se(b_j) = \sqrt{\widehat{\sigma}^2[(\boldsymbol{X}'\boldsymbol{X})^{-1}]_{jj}}$$

Inferenz

t-Test für $\beta_j = 0$

Die Test-Statistik für (2-seitiger Test)

$$H_0 : \beta_j = 0 \qquad \text{und} \qquad H_1 : \beta_j \neq 0$$

lautet

$$t = \frac{b_j}{se(b_j)} \sim t_{n-k-1}$$

Sie ist t-verteilt mit $(n - k - 1)$ Freiheitsgraden.

Die Nullhypothese sagt nicht nur, dass

- der Parameter $\beta_j = 0$ ist, sondern auch, dass
- die vermeintlich erklärende Variable x_j mit null multipliziert wird, und daher *in die Gleichung nicht aufgenommen werden soll*.

t-Test, allgemeine H_0

Etwas allgemeiner (2-seitiger Test)

$$H_0 : \beta_j = \beta_j^{(0)} \qquad H_1 : \beta_j \neq \beta_j^{(0)}$$

Die Test-Statistik ist

$$t = \frac{b_j - \beta_j^{(0)}}{se(b_j)} \sim t_{n-k-1}$$

Sie ist t-verteilt mit $(n - k - 1)$ Freiheitsgraden.

Konfidenzintervall für β_j

Analog zum **Konfidenzintervall**, KI, für das Mittel wird hier ein KI für β_j konstruiert:

$$[\, b_j - t_{n-k-1,\alpha/2} se(b_j),\ b_j + t_{n-k-1,1-\alpha/2} se(b_j)\,]$$

Mit $\alpha = 0.05$ ($n - k - 1 > 30$)

$$[\, b_j - 1.96\, se(b_j),\ b_j + 1.96\, se(b_j)\,]$$

Das KI überdeckt den wahren Parameter β_j mit der Wahrscheinlichkeit $(1 - \alpha) = 0.95$.

Okun's Gesetz: t-Test

$$\widehat{\Delta u_t} = \underset{(0.116)}{0.500} - \underset{(0.042)}{0.165}\,\rho_t \qquad \begin{aligned} R^2 &= 0.322 \\ RSS &= 5.575 \end{aligned}$$

- Wir unterstellen in der Beziehung zw Wachstum und Veränderung der AI-Rate einen negativen Zusammenhang. Das Vorzeichen vor ρ_t passt in der Schätzgleichung.
- Der t-Test mit $H_0 : \beta_1 = 0$ ergibt bei FG=$34 - 2 = 32 > 30$

$$t = \frac{-0.165}{0.042} = -3.928 < -1.96$$

  Damit kann man die H_0 mit $\alpha = 0.05$, beidseitig, ablehnen.
  b_1 ist signifikant von null verschieden.
- Die Variable ρ_t kann in der Gleichung verbleiben.

Okun's Gesetz: *t*-Test und Konfidenzintervall

Ang wir vermuten, dass Δu stärker auf ρ reagiert. Etwa mit

- $\beta_1^{(0)} = -0.20$: Kann diese H_0 verworfen werden?

  $$t = \frac{-0.165 - (-0.20)}{0.042} = 0.833 < 1.96$$

  Die Hypo $\beta_1 = -0.20$ kann nicht verworfen werden, und *passt* daher zu unseren Daten. Oder mit

- $\beta_1^{(0)} = -0.25$:

  $$t = \frac{-0.165 - (-0.25)}{0.042} = 2.024 > 1.96$$

  Die Hypo $\beta_1 = -0.25$ wird verworfen. Ein Effekt von ρ auf Δu dieser Stärke wird *nicht bestätigt*.

Okun's Gesetz: *t*-Test und Konfidenzintervall

Das 95% Konfidenzintervall lautet

$$[-0.165 - 1.96 \cdot 0.042,\ -0.165 + 1.96 \cdot 0.042] = [-0.247, -0.083]$$

Ein Vergleich der beiden Tests mit dem KI ergibt:

- -0.20 liegt im KI. Die $H_0 : \beta_1 = -0.20$ wird nicht verworfen,
- -0.25 außerhalb des KIs. Die $H_0 : \beta_1 = -0.25$ wird verworfen.

Okun's Gesetz: einseitiger t-Test

Der 2-seitige t-Test ist nicht ausreichend, da ein $\beta_1 > 0$ ökonomisch nicht plausibel ist. Daher

$$H_0 : \beta_1 \geq 0 \text{ (!)} \qquad \text{und} \qquad H_1 : \beta_1 < 0$$

$$t = \frac{-0.165}{0.042} = -3.928 < -1.645 = t_c$$

bei einem $\alpha = 0.05$, ein- bzw *links*-seitig.

$t_c = -1.96$ entspricht einem $\alpha = 0.025$ einseitig.

Ein 1-seitiger Test *entspricht* einem 2-seitigen, wenn $\alpha_{(1)} = \alpha_{(2)}/2$ und das Vorzeichen des Koeffizienten passt. ($\alpha_{(1)}$, $\alpha_{(2)}$ sind die α's des 1- bzw 2-seitigen Tests.)

Übungen und Referenzen

Referenzen

Hackl 4.4, 5
Wooldridge 4

Übungen

1 Wählen sie das Datenfile `twoyear.wf1` bzw `twoyear_R.txt`.
Erzeugen sie die Variable `wage`: `Generate wage = exp(lwage)`
(a) Berechnen sie Mittel, Varianz, etc. mittels `Descriptive Statistics`.
Plotten und beschreiben sie das Histogramm.
(b) Schätzen sie die Gleichung

$$wage_i = \beta_0 + u_i$$

Vergleichen sie b_0 mit dem Stichprobenmittel.
(c) Schätzen sie die Gleichung

$$wage_i = \beta_0 + \beta_1 female_i + u_i$$

Da für Männer `female = 0` ist, gibt b_0 den mittleren Stundenlohn für Männer und $(b_0 + b_1)$ den mittleren Lohn der Frauen. Wie groß ist der Unterschied?
(d) Schätzen sie die Modelle aus (a) und (b) mit der abhängigen Variablen `exper`.
(e) Schätzen und interpretieren sie

$$exper = \beta_0 + \beta_1 female + u \quad \text{und} \quad wage = \beta_0 + \beta_1 female + \beta_2 exper + u$$

Übungen

2 Wählen sie das Datenfile `okuns_law.wf1` bzw `okuns_law_R.txt`.
Schätzen sie die Gleichung

$$\Delta u_t = \beta_0 + \beta_1 \rho_t + \epsilon_t$$

mit OLS.
(a) Führen sie t-Tests für b_0 und b_1 auf null durch.
(b) Berechnen sie beide KIs für $\alpha = 0.05$.
(c) Testen sie die 3 Nullhypothesen
H_0: $\beta_1 = -0.15$, H_0: $\beta_1 = -0.10$ und H_0: $\beta_1 = -0.05$.
(d) Schätzen sie die umgekehrte funktionale Beziehung, berechnen sie das KI für die Steigung und vergleichen sie mit dem KI aus der Schätzung für die USA, Kap 3, Üb 4. Gibt es eine Überlappung?

Übungen

3 Wählen sie das Datenfile `international_inequality.wf1` bzw `international_inequality_R.txt`. Untersuchen sie den Zusammenhang zwischen `imprisonment_log` und `income_inequality`.
(a) Geben sie an, um welche Beobachtungseinheiten es sich handelt. Wie groß ist der Stichprobenumfang.
(b) Wie sind die beiden Variablen definiert. In welchen Einheiten werden sie gemessen.
(c) Stellen sie die Daten graphisch dar, zB in einem Streudiagramm.
(c) Regressieren sie `imprisonment_log` auf `income_inequality`.
(d) Entscheiden sie, ob sie eine Konstante ins Modell aufnehmen.
(e) Geben sie ein KI für β_1 an, und interpretieren sie es.