Ökonometrie Grundlagen, Übung 6

Henry Haustein

Aufgabe 1

Für einen unverzerrten Schätzer gilt $\mathbb{E}(\hat{\beta}) = \beta$.

$$\begin{aligned}
\mathbb{E}(\hat{\beta}) &= \mathbb{E}((X'X)^{-1}X'y) \\
&= \mathbb{E}(X^{-1}\underbrace{X'^{-1}X'}_{I}y) \\
&= \mathbb{E}(X^{-1}y) \\
&= X^{-1}\mathbb{E}(y) \\
&= X^{-1}\mathbb{E}(X\beta) \\
&= \underbrace{X^{-1}X}_{I}\beta \\
&= \beta
\end{aligned}$$

Aufgabe 2

Wir werden zeigen, dass sich $\Sigma_{\hat{\beta}}$ auch als $\sigma_u^2(X'X)^{-1}$ schreiben lässt.

$$\begin{aligned}
\Sigma_{\hat{\beta}} = \mathrm{Var}(\hat{\beta}) &= \mathbb{E}(\hat{\beta}^2) - \underbrace{\mathbb{E}(\hat{\beta})}_{\beta}\underbrace{\mathbb{E}(\hat{\beta}')}_{\beta} \\
&= \mathbb{E}\left([(X'X)^{-1}X'y]^2\right) - \beta^2 \\
&= \mathbb{E}\left([(X'X)^{-1}X'(X\beta + u)]^2\right) - \beta^2 \\
&= \mathbb{E}\left([\underbrace{(X'X)^{-1}(X'X)}_{I}(\beta + (X'X)^{-1}X'u)]^2\right) - \beta^2 \\
&= \mathbb{E}\left([\beta + (X'X)^{-1}X'u]^2\right) - \beta^2 \\
&= \beta^2 + \mathbb{E}\left([(X'X)^{-1}X'u]^2\right) - \beta^2 \\
&= [(X'X)^{-1}X']^2 \cdot \underbrace{\mathbb{E}(u^2)}_{\sigma_u^2} \\
&= \underbrace{(X'X)^{-1}(X'X)}_{I}(X'X)'^{-1} \cdot \sigma_u^2 \\
&= \underbrace{(X'X)'^{-1}}_{(X'X)^{-1}\text{, da symmetrisch}} \cdot \sigma_u^2 \\
&= (X'X)^{-1} \cdot \sigma_u^2
\end{aligned}$$

Aufgabe 3

Für einen unverzerrten Schätzer gilt $\mathbb{E}(\widehat{\sigma_u^2}) = \sigma_u^2$.

$$\begin{aligned}
\mathbb{E}(\widehat{\sigma_u^2}) &= \mathbb{E}\left(\frac{\hat{u}'\hat{u}}{T-k}\right) \\
&= \mathbb{E}\left(\frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{T-k}\right) \\
&= \frac{1}{T-k}\mathbb{E}(y'y - y'X\hat{\beta} - (X\hat{\beta})'y + (X\hat{\beta})'(X\hat{\beta})) \\
&= \frac{1}{T-k}(\mathbb{E}((X\hat{\beta}+u)'(X\hat{\beta}+u)) - \mathbb{E}((X\hat{\beta})X\hat{\beta}) - \mathbb{E}((X\hat{\beta})'(X\hat{\beta}+u)) + \hat{\beta}'X'X\hat{\beta}) \\
&= \frac{1}{T-k}(\mathbb{E}(\hat{\beta}'X'X\hat{\beta} + (X\hat{\beta})'u + u'(X\hat{\beta}) + u'u) - \mathbb{E}(X\hat{\beta}X\hat{\beta} + uX\hat{\beta}) - \mathbb{E}((X\hat{\beta})'X\hat{\beta} + (X\hat{\beta})'u) + \hat{\beta}'X'X\hat{\beta}) \\
&= \frac{1}{T-k}(\hat{\beta}'X'X\hat{\beta} + (X\hat{\beta})'\underbrace{\mathbb{E}(u)}_{0} + (X\hat{\beta})\underbrace{\mathbb{E}(u')}_{0} + \mathbb{E}(u'u) - \hat{\beta}'X'X\hat{\beta} - X\hat{\beta}\underbrace{\mathbb{E}(u)}_{0} - (X\hat{\beta})'X\hat{\beta} - (X\hat{\beta})'\underbrace{\mathbb{E}(u)}_{0} + \hat{\beta}'X'X\hat{\beta}) \\
&= \frac{1}{T-k}(\hat{\beta}'X'X\hat{\beta} + \mathbb{E}(u'u) - \hat{\beta}'X'X\hat{\beta} - \hat{\beta}'X'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta}) \\
&= \frac{1}{T-k}\mathbb{E}(u'u) \\
&= \mathbb{E}\left(\frac{u'u}{T-k}\right) \\
&= \sigma_u^2
\end{aligned}$$

Aufgabe 4

Ich weiß nicht wirklich, was ich hier beweisen soll. Das Gauß-Markow-Theorem sagt, dass ein Schätzer $\hat{\vartheta}$ genau dann die BLUE-Eigenschaft (*Best linear unbiased estimator*), wenn die zugehörigen Residuen

- den Erwartungswert 0 haben
- eine endliche Varianz haben
- unkorreliert sind

Alle diese 3 Eigenschaften sind durch die Annahmen (A1) und (A3) für alle Fehler (und damit insbesondere für den durch die Schätzung von $\hat{\beta}$ erzeugten Fehler) gegeben. Mit dem Theorem folgt also dass $\hat{\beta}$ die BLUE-Eigenschaft hat.

Aufgabe 5

Angenommen wir beobachten die Daten

$$y = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

und schätzen diese mit einem linearen Modell ($k = 1$) durch

$$\hat{y} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

Es ist $T = 2$ und für R^2 gilt dann:

$$\begin{aligned} R^2 &= \frac{\text{ESS}}{\text{RSS}} = \frac{(\hat{y}_1 - \bar{y})^2 + (\hat{y}_2 - \bar{y})^2}{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2} \\ &= \frac{2+2}{1+1} \\ &= 2 \end{aligned}$$

Für Theils adjustiertes Bestimmheitsmaß und Amemiyas adjustiertes Bestimmtheitsmaß gilt dann

$$\begin{aligned} \bar{R}^2 &= 1 - (1 - R^2)\frac{T+1}{T-k} \\ &= 1 - (1-2)\frac{2+1}{2-1} \\ &= 4 \\ \tilde{R}^2 &= 1 - (1 - R^2)\frac{T+k}{T-k} \\ &= 1 - (1-2)\frac{2+1}{2-1} \\ &= 4 \end{aligned}$$

Es gilt also $\bar{R}^2 \notin [0, 1]$ und $\tilde{R}^2 \notin [0, 1]$.

Aufgabe 6

(a) ökonomisches Modell: price $= \beta_0 + \beta_1 \cdot$ speed $+ \beta_2 \cdot$ hd $+ \beta_3 \cdot$ ram $+ \beta_4 \cdot$ screen
ökonometrisches Modell: $\text{price}_t = \beta_0 + \beta_1 \cdot \text{speed}_t + \beta_2 \cdot \text{hd}_t + \beta_3 \cdot \text{ram}_t + \beta_4 \cdot \text{screen}_t + u_t$

(b) in R

```
datensatz = Computers
modell = lm(price ~ speed + hd + ram + screen, data = datensatz)
summary(modell)
```

ergibt

	Estimate	Pr(>|t|)
(Intercept)	10.33311	0.907
speed	5.24930	< 2e-16
hd	-0.57936	< 2e-16
ram	76.74545	< 2e-16
screen	105.52592	< 2e-16

Der F-Test ergibt einen p-Value $< 2.2 \cdot 10^{-16}$ und es sind auch alle Parameter (bis auf β_0) signifikant von Null verschieden. Zudem ist $\bar{R}^2 = 0.4583$, was auf kein gutes Modell der Daten hindeutet.

(c) Praktisch gesehen macht es keinen Sinn den Preis von Computern nur von einem Parameter abhängig zu machen. Computer bestehen aus vielen Komponenten, das sollte sich auch im Modell wiederfinden. Die Devise ist also hier: mehr Parameter statt weniger. Schaut man sich den $\bar{R}^2$ für das volle Modell an, so ergibt sich

```
Fullmodell = lm(price ~ ., data = datensatz)
summary(Fullmodell)
```

Hier ist $\bar{R}^2 = 0.7752$.

Aufgabe 7

(a) ökonomisches Modell: $Z = \beta_0 + \beta_1 P + \beta_2 E_i$
ökonometrisches Modell: $Z_i = \beta_0 + \beta_1 P_i + \beta_2 E_i + u_i$

(b) Die Dummy-Variable k_i soll 0 werden, wenn die Familie Kinder hat, ansonsten ist $k_i = 0$.

$$Z_i = \beta_0 + \beta_1 k_i + \beta_2 P_i + \beta_3 k_i P_i + \beta_4 E_i + \beta_5 k_i E_i + u_i$$

mit Restriktion $R\beta = r$.

(c) in Blockmatrixform:

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_T \end{pmatrix} = \begin{pmatrix} 1 & k_1 & P_1 & k_1 P_1 & E_1 & k_1 E_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & k_T & P_T & k_T P_T & E_T & k_T E_T \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_5 \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_T \end{pmatrix}$$

mit Restriktion $R\beta = r$.

(d) Zielfunktion: $Q_T = (Z - X\beta)'(Z - X\beta) \to \min$ unter der Nebenbedingung $R\beta = r$. Schätzung der Parameter über

$$\hat{\beta} = (X'X)^{-1}X'Z - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R(X'X)^{-1}X'Z - r)$$

(e) Wenn gemeinsam auf die Signifikanz der Steigung getestet werden soll, sollte ein F-Test verwendet werden. Der t-Test testet nur auf die Signifikanz eines Parameters.

(f) Alle p-Values sind kleiner als $\alpha = 0.05$, damit sind alle Parameter signifikant von Null verschieden.