universität innsbruck

Eigenschaften von OLS Schätzfunktionen

Grundlagen der Ökonometrie

herbert.stocker@uibk.ac.at

www.hsto.info/econometrics

Eigenschaften von OLS Schätzfunktionen

- Es gibt viele Schätzfunktionen (OLS, ML, GMM, . . .).
  Wie können diese verglichen/beurteilt werden?
  Was unterscheidet ‘gute’ von ‘schlechten’ Schätzfunktionen?

Eigenschaften von OLS Schätzfunktionen

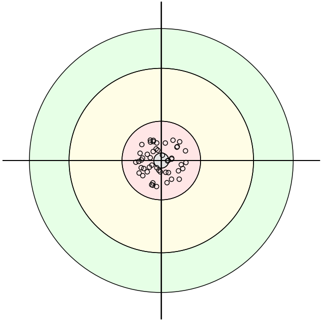
- Es gibt viele Schätzfunktionen (OLS, ML, GMM, . . .).
  Wie können diese verglichen/beurteilt werden?
  Was unterscheidet ‘gute’ von ‘schlechten’ Schätzfunktionen?
- Eigenschaften von Schätzfunktionen:
  - Erwartungstreue
  - Effizienz (Varianzminimalität)
  - Konsistenz
  - usw.
- Wie können diese Eigenschaften gezeigt/bewiesen werden?

Eigenschaften von OLS Schätzfunktionen

- Es gibt viele Schätzfunktionen (OLS, ML, GMM, . . .).
  Wie können diese verglichen/beurteilt werden?
  Was unterscheidet ‘gute’ von ‘schlechten’ Schätzfunktionen?
- Eigenschaften von Schätzfunktionen:
  - Erwartungstreue
  - Effizienz (Varianzminimalität)
  - Konsistenz
  - usw.
- Wie können diese Eigenschaften gezeigt/bewiesen werden?
- *Frage: Welche Annahmen müssen erfüllt sein, damit OLS-Schätzfunktionen diese Eigenschaften besitzen?*
- z.B. für Erwartungstreue und Effizienz: *Gauss-Markov Annahmen*

Eigenschaften von OLS Schätzfunktionen

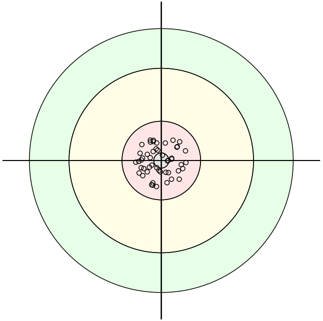
1)



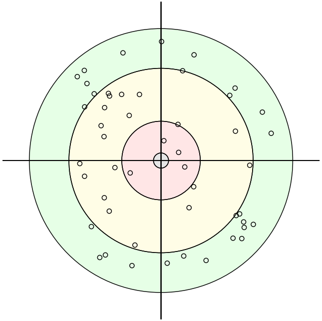
1) erwartungstreu

Eigenschaften von OLS Schätzfunktionen

1)



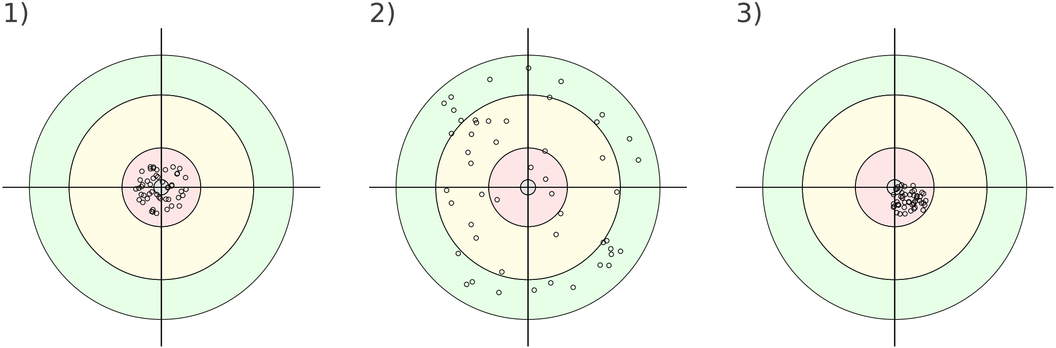
2)



1) erwartungstreu und effizient,

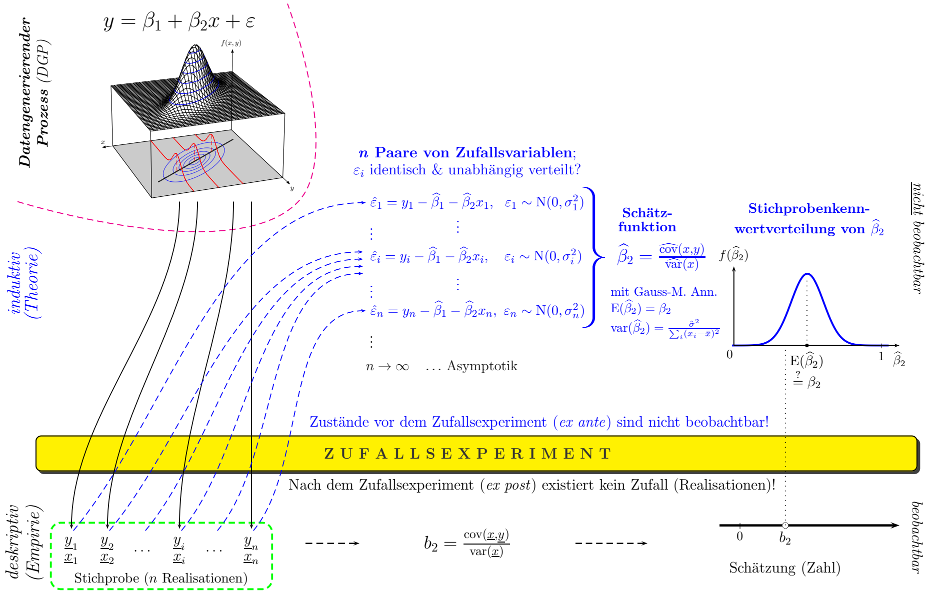
2) erwartungstreu, aber *nicht* effizient,

Eigenschaften von OLS Schätzfunktionen



1) erwartungstreu und effizient,
2) erwartungstreu, aber *nicht* effizient,
3) verzerrt, also auch nicht effizient.

Stochastische Regressionsanalyse



Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \stackrel{?}{=} \beta_2$

- trifft die Schätzfunktion *‘im Durchschnitt’* richtig?

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \stackrel{?}{=} \beta_2$

- trifft die Schätzfunktion *'im Durchschnitt'* richtig?
- Beispiel: OLS Steigungskoeffizient:
  wir benötigen Zusammenhang zwischen $\widehat{\beta}_2$ und β_2

$$\widehat{\beta}_2 = \frac{\widehat{\mathrm{cov}}(x,y)}{\widehat{\mathrm{var}}(x)}$$

$$y = \beta_1 + \beta_2 x + \varepsilon \qquad \text{(PRF)}$$

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \stackrel{?}{=} \beta_2$

- trifft die Schätzfunktion *'im Durchschnitt'* richtig?
- Beispiel: OLS Steigungskoeffizient:
  wir benötigen Zusammenhang zwischen $\widehat{\beta}_2$ und β_2

$$\widehat{\beta}_2 = \frac{\widehat{\mathrm{cov}}(x,y)}{\widehat{\mathrm{var}}(x)}$$

$$y = \beta_1 + \beta_2 x + \varepsilon \qquad \text{(PRF)}$$

- Einsetzen der PRF in OLS Schätzfunktion

$$\begin{aligned}
\widehat{\beta}_2 &= \frac{\widehat{\mathrm{cov}}(x,y)}{\widehat{\mathrm{var}}(x)} = \frac{\widehat{\mathrm{cov}}\left[x, (\beta_1 + \beta_2 x + \varepsilon)\right]}{\widehat{\mathrm{var}}(x)} \\
&= \frac{1}{\widehat{\mathrm{var}}(x)} \left[\underbrace{\widehat{\mathrm{cov}}(x,\beta_1)}_{=0} + \beta_2 \underbrace{\widehat{\mathrm{cov}}(x,x)}_{=\widehat{\mathrm{var}}(x)} + \widehat{\mathrm{cov}}(x,\varepsilon) \right]
\end{aligned}$$

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \stackrel{?}{=} \beta_2$

$$\widehat{\beta}_2 = \beta_2 + \frac{\widehat{\mathrm{cov}}(x,\varepsilon)}{\widehat{\mathrm{var}}(x)}$$

Deshalb

$$\boxed{\mathrm{E}(\widehat{\beta}_2) = \beta_2 \quad \text{wenn und nur wenn} \;\; \mathrm{E}(\widehat{\mathrm{cov}}(x,\varepsilon)) = 0}$$

$\Rightarrow$ OLS Schätzfunktion ist erwartungstreu, wenn

A1: PRF linear und korrekt spezifiziert

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \overset{?}{=} \beta_2$

$$\widehat{\beta}_2 = \beta_2 + \frac{\widehat{\mathrm{cov}}(x, \varepsilon)}{\widehat{\mathrm{var}}(x)}$$

Deshalb

$$\boxed{\mathrm{E}(\widehat{\beta}_2) = \beta_2 \quad \text{wenn und nur wenn} \quad \mathrm{E}(\widehat{\mathrm{cov}}(x, \varepsilon)) = 0}$$

$\Rightarrow$ OLS Schätzfunktion ist erwartungstreu, wenn

A1: PRF linear und korrekt spezifiziert

A2: $n \geq k$ und keine perfekte Multikollinearität

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue: $\mathrm{E}(\widehat{\beta}_2) \stackrel{?}{=} \beta_2$

$$\widehat{\beta}_2 = \beta_2 + \frac{\widehat{\mathrm{cov}}(x,\varepsilon)}{\widehat{\mathrm{var}}(x)}$$

Deshalb

$$\boxed{\mathrm{E}(\widehat{\beta}_2) = \beta_2 \quad \text{wenn und nur wenn} \quad \mathrm{E}(\widehat{\mathrm{cov}}(x,\varepsilon)) = 0}$$

$\Rightarrow$ OLS Schätzfunktion ist erwartungstreu, wenn

A1: PRF linear und korrekt spezifiziert

A2: $n \geq k$ und keine perfekte Multikollinearität

A3: $\mathrm{E}(\widehat{\mathrm{cov}}(x,\varepsilon)) = 0$

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: wenn $\mathrm{E}(\widehat{\mathrm{cov}}(x,\varepsilon)) \neq 0$ ist ε eine Funktion der x, also $\varepsilon(x)$

PRF: $y = \beta_1 + \beta_2 x + \varepsilon(x)$

mit dem marginalen Effekt

$$\frac{dy}{dx} = \underbrace{\beta_2 + \frac{d\varepsilon}{dx}}_{\beta_2^*}$$

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: wenn $\mathrm{E}(\widehat{\mathrm{cov}}(x, \varepsilon)) \neq 0$ ist ε eine Funktion der x, also $\varepsilon(x)$

PRF: $y = \beta_1 + \beta_2 x + \varepsilon(x)$

mit dem marginalen Effekt

$$\frac{dy}{dx} = \underbrace{\beta_2 + \frac{d\varepsilon}{dx}}_{\beta_2^*}$$

- Wir können mit OLS nur die Summe β_2^* schätzen, nicht den interessierenden Koeffizienten β_2!

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: wenn $\mathrm{E}(\widehat{\mathrm{cov}}(x,\varepsilon)) \neq 0$ ist ε eine Funktion der x, also $\varepsilon(x)$

$$\text{PRF:} \quad y = \beta_1 + \beta_2 x + \varepsilon(x)$$

mit dem marginalen Effekt

$$\frac{dy}{dx} = \underbrace{\beta_2 + \frac{d\varepsilon}{dx}}_{\beta_2^*}$$

- Wir können mit OLS nur die Summe β_2^* schätzen, nicht den interessierenden Koeffizienten β_2!
- Identifikationsproblem – *Endogenität* (endogene Regressoren)

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: wenn $\mathrm{E}(\widehat{\mathrm{cov}}(x, \varepsilon)) \neq 0$ ist ε eine Funktion der x, also $\varepsilon(x)$

$$\text{PRF:} \quad y = \beta_1 + \beta_2 x + \varepsilon(x)$$

mit dem marginalen Effekt

$$\frac{dy}{dx} = \underbrace{\beta_2 + \frac{d\varepsilon}{dx}}_{\beta_2^*}$$

- Wir können mit OLS nur die Summe β_2^* schätzen, nicht den interessierenden Koeffizienten β_2!
- Identifikationsproblem – *Endogenität* (endogene Regressoren)
- $\Rightarrow$ OLS liefert verzerrte Ergebnisse!

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: Endogene Regressoren: Kernproblem der Ökonometrie

Typische Fälle:

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: Endogene Regressoren: Kernproblem der Ökonometrie

Typische Fälle:

- fehlenden relevanten Regressoren (*omitted variables bias*), unbeobachtete Heterogenität und Selektionsprobleme,

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: Endogene Regressoren: Kernproblem der Ökonometrie

Typische Fälle:

- fehlenden relevanten Regressoren (*omitted variables bias*), unbeobachtete Heterogenität und Selektionsprobleme,
- Simultaner Abhängigkeit in interdependenten Systemen (*feed-back Mechanismen*; *omitted equation bias*).

Eigenschaften von OLS Schätzfunktionen

Erwartungstreue:

ad A3: Endogene Regressoren: Kernproblem der Ökonometrie

Typische Fälle:

- fehlenden relevanten Regressoren (*omitted variables bias*), unbeobachtete Heterogenität und Selektionsprobleme,
- Simultaner Abhängigkeit in interdependenten Systemen (*feed-back Mechanismen*; *omitted equation bias*).
- Messfehler in den erklärenden Variablen.

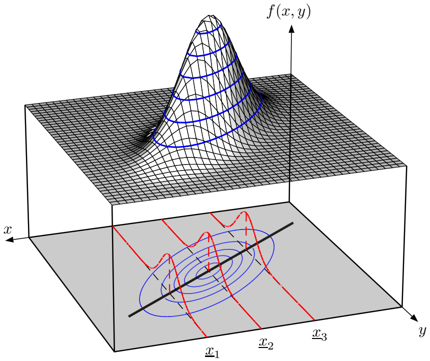
Eigenschaften von OLS Schätzfunktionen

Stochastische Regressoren: $\rightarrow$ Sampling

Für *stochastische* Regressoren werden 2 weitere Annahmen benötigt:

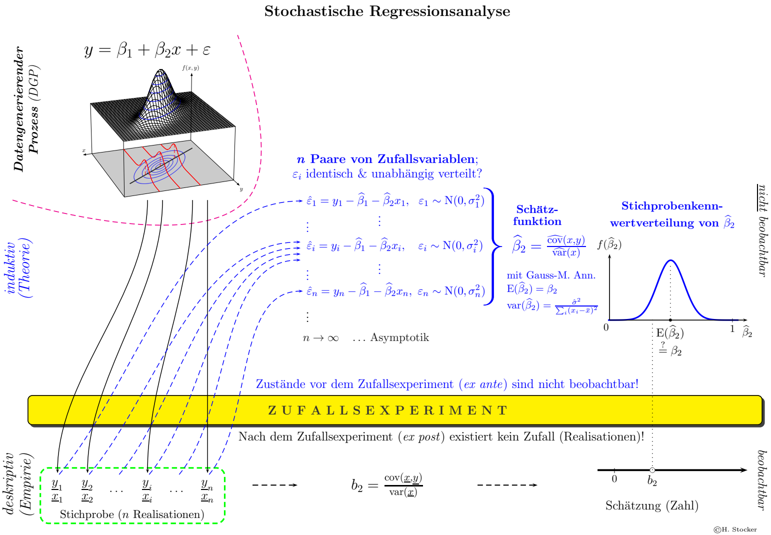
AS 1: Die (x_i, y_i) Paare sind i.i.d. $\forall\, i = 1, \ldots n$,

AS 2: Große Ausreißer sind unwahrwahrscheinlich:
$0 < \mathrm{E}(x_i^4) < \infty$ und
$0 < \mathrm{E}(y_i^4) < \infty$



Resultate bleiben auch für stochast. Regressoren weitgehend gültig.

OLS Standardfehler



OLS Standardfehler

$$\widehat{\beta}_2 = \frac{\widehat{\mathrm{cov}}(x,y)}{\widehat{\mathrm{var}}(x)} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\, y_i$$

weil

$$\begin{aligned}
\sum_i (x_i - \bar{x})(y_i - \bar{y}) &= \sum_i (x_i - \bar{x}) y_i - \sum_i (x_i - \bar{x})\bar{y} \\
&= \sum_i (x_i - \bar{x}) y_i - \bar{y} \sum_i (x_i - \bar{x}) \\
&= \sum_i (x_i - \bar{x}) y_i \qquad \text{weil } \sum_i (x_i - \bar{x}) = 0
\end{aligned}$$

OLS Standardfehler

$$\widehat{\beta}_2 = \frac{\widehat{\operatorname{cov}}(x,y)}{\widehat{\operatorname{var}}(x)} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\, y_i$$

weil

$$\begin{aligned} \sum_i (x_i - \bar{x})(y_i - \bar{y}) &= \sum_i (x_i - \bar{x}) y_i - \sum_i (x_i - \bar{x}) \bar{y} \\ &= \sum_i (x_i - \bar{x}) y_i - \bar{y} \sum_i (x_i - \bar{x}) \\ &= \sum_i (x_i - \bar{x}) y_i \qquad \text{weil } \sum_i (x_i - \bar{x}) = 0 \end{aligned}$$

Also: OLS Schätzfunktion ist linear in y

$$\widehat{\beta}_2 = \sum_{i=1}^{n} w_i y_i \qquad \text{mit} \quad w_i := \frac{(x_i - \bar{x})}{\sum_j (x_j - \bar{x})^2}$$

OLS Standardfehler

Für die Gewichte w_i gilt: (siehe Skript)

1. $\sum_i w_i = 0$
2. $\sum_i w_i^2 = 1/\sum_i (x_i - \bar{x})^2$
3. $\sum_i w_i x_i = 1$

Einsetzen der PRF in die OLS Schätzfunktion $\widehat{\beta}_2$ gibt

$$\begin{aligned}
\widehat{\beta}_2 &= \sum_i w_i y_i = \sum_i w_i(\beta_1 + \beta_2 x_i + \varepsilon_i) \\
&= \beta_1 \sum_i w_i + \beta_2 \sum_i w_i x_i + \sum_i w_i \varepsilon_i \\
&= \beta_2 + \sum_i w_i \varepsilon_i
\end{aligned}$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist definiert als

$$\mathrm{var}(\widehat{\beta}_2) = \mathrm{E}[\widehat{\beta}_2 - \mathrm{E}(\widehat{\beta}_2)]^2$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist definiert als

$$\begin{aligned}\operatorname{var}(\widehat{\beta}_2) &= \mathrm{E}[\widehat{\beta}_2 - \mathrm{E}(\widehat{\beta}_2)]^2 \\ &= \mathrm{E}[\widehat{\beta}_2 - \beta_2]^2 \qquad (\text{wenn } \mathrm{E}(\widehat{\beta}_2) = \beta_2)\end{aligned}$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist definiert als

$$\begin{aligned}
\text{var}(\widehat{\beta}_2) &= \mathrm{E}[\widehat{\beta}_2 - \mathrm{E}(\widehat{\beta}_2)]^2 \\
&= \mathrm{E}[\widehat{\beta}_2 - \beta_2]^2 \qquad (\text{wenn } \mathrm{E}(\widehat{\beta}_2) = \beta_2) \\
&= \mathrm{E}\left(\sum_i w_i \varepsilon_i\right)^2 \qquad (\text{weil } \widehat{\beta}_2 = \beta_2 + \sum w_i \varepsilon_i, \text{ s.o.}) \\
&= \mathrm{E}\big(w_1^2 \varepsilon_1^2 + w_2^2 \varepsilon_2^2 + \cdots + w_n^2 \varepsilon_n^2 + \cdots \\
&\quad \cdots + 2 w_1 w_2 \varepsilon_1 \varepsilon_2 + \cdots + 2 w_{n-1} w_n \varepsilon_{n-1} \varepsilon_n\big)
\end{aligned}$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist definiert als

$$\begin{aligned}
\operatorname{var}(\widehat{\beta}_2) &= \mathrm{E}[\widehat{\beta}_2 - \mathrm{E}(\widehat{\beta}_2)]^2 \\
&= \mathrm{E}[\widehat{\beta}_2 - \beta_2]^2 \qquad \text{(wenn } \mathrm{E}(\widehat{\beta}_2) = \beta_2\text{)} \\
&= \mathrm{E}\left(\sum_i w_i \varepsilon_i\right)^2 \qquad \text{(weil } \widehat{\beta}_2 = \beta_2 + \sum w_i \varepsilon_i, \text{ s.o.)} \\
&= \mathrm{E}\left(w_1^2\varepsilon_1^2 + w_2^2\varepsilon_2^2 + \cdots + w_n^2\varepsilon_n^2 + \cdots \right. \\
&\qquad \left. \cdots + 2w_1w_2\varepsilon_1\varepsilon_2 + \cdots + 2w_{n-1}w_n\varepsilon_{n-1}\varepsilon_n\right) \\
&= \underbrace{\mathrm{E}\left(\sum_{i=1}^n w_i^2\varepsilon_i^2\right)}_{\substack{= \sigma^2 \sum_i w_i^2 \text{ wenn} \\ \text{homoskedastisch}}} + \underbrace{\mathrm{E}\left(\sum_{i=1}^n \sum_{\substack{j=2 \\ j>i}}^n 2w_iw_j\varepsilon_i\varepsilon_j\right)}_{\substack{= 0 \text{ wenn keine} \\ \text{Autokorrelation}}}
\end{aligned}$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist also

$$\begin{aligned}\operatorname{var}(\widehat{\beta}_2) &= \mathrm{E}\left(w_1^2\varepsilon_1^2 + w_2^2\varepsilon_2^2 + \cdots + w_n^2\varepsilon_n^2 + \cdots\right. \\ &\quad \left.\cdots + 2w_1w_2\varepsilon_1\varepsilon_2 + \cdots + 2w_{n-1}w_n\varepsilon_{n-1}\varepsilon_n\right)\end{aligned}$$

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist also

$$\begin{aligned}\operatorname{var}(\widehat{\beta}_2) &= \operatorname{E}\left(w_1^2\varepsilon_1^2 + w_2^2\varepsilon_2^2 + \cdots + w_n^2\varepsilon_n^2 + \cdots\right.\\ &\quad \left.\cdots + 2w_1w_2\varepsilon_1\varepsilon_2 + \cdots + 2w_{n-1}w_n\varepsilon_{n-1}\varepsilon_n\right)\end{aligned}$$

Mit zusätzlicher Annahme **A4**: $\boxed{\varepsilon_i \sim \text{i.i.d.}\left(0, \sigma^2\right)}$

d.h. wenn $\operatorname{var}(\varepsilon_i) := \operatorname{E}(\varepsilon_i)^2 = \sigma^2$ und $\operatorname{cov}(\varepsilon_i, \varepsilon_j) := \operatorname{E}(\varepsilon_i, \varepsilon_j) = 0$ für $i \neq j$:

OLS Standardfehler

Die Varianz der Zufallsvariable $\widehat{\beta}_2$ ist also

$$\begin{aligned}\operatorname{var}(\widehat{\beta}_2) &= \mathrm{E}\left(w_1^2\varepsilon_1^2 + w_2^2\varepsilon_2^2 + \cdots + w_n^2\varepsilon_n^2 + \cdots\right.\\ &\quad \left.\cdots + 2w_1w_2\varepsilon_1\varepsilon_2 + \cdots + 2w_{n-1}w_n\varepsilon_{n-1}\varepsilon_n\right)\end{aligned}$$

Mit zusätzlicher Annahme **A4**: $\boxed{\varepsilon_i \sim \text{i.i.d.}\left(0, \sigma^2\right)}$

d.h. wenn $\operatorname{var}(\varepsilon_i) := \mathrm{E}(\varepsilon_i)^2 = \sigma^2$ und $\operatorname{cov}(\varepsilon_i, \varepsilon_j) := \mathrm{E}(\varepsilon_i, \varepsilon_j) = 0$ für $i \neq j$:

$$\begin{aligned}\operatorname{var}(\widehat{\beta}_2) &= \mathrm{E}\left(w_1^2\sigma^2 + w_2^2\sigma^2 + \cdots + w_n^2\sigma^2 + \cdots\right.\\ &\quad \left.\cdots + 2w_1w_2 0 + \cdots + 2w_{n-1}w_n 0\right)\\ &= \sigma^2 \sum_i w_i^2 \qquad \text{mit } \sum_i w_i^2 = 1/\sum_i (x_i - \bar{x})^2, \text{(s.o.)}\\ &= \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2}\end{aligned}$$

OLS Standardfehler

Varianz des OLS Steigungskoeffizienten $\widehat{\beta}_2$:

$$\boxed{\operatorname{var}(\widehat{\beta}_2) = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2}}$$

- kann aber nicht geschätzt werden, da σ^2 ein unbeobachtbarer Parameter des DGP ist!

OLS Standardfehler

Varianz des OLS Steigungskoeffizienten $\widehat{\beta}_2$:

$$\boxed{\operatorname{var}(\widehat{\beta}_2) = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2}}$$

- kann aber nicht geschätzt werden, da σ^2 ein unbeobachtbarer Parameter des DGP ist!
- Eine unverzerrte Schätzfunktion für σ^2 ist

$$\boxed{\hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}}$$

mit $\hat{\varepsilon}_i$: OLS Residuen; $n-k$: Freiheitsgrade

OLS Standardfehler

Freiheitsgrade:

- Für jeden der insgesamt k zu schätzenden Parameter wird eine Bedingung 1. Ordnung (FOC) benötigt!

OLS Standardfehler

Freiheitsgrade:

- Für jeden der insgesamt k zu schätzenden Parameter wird eine Bedingung 1. Ordnung (FOC) benötigt!
- z.B. $\sum_i \hat{\varepsilon}_i = 0$ und $\sum_i \hat{\varepsilon}_i x_i = 0$

OLS Standardfehler

Freiheitsgrade:

- Für jeden der insgesamt k zu schätzenden Parameter wird eine Bedingung 1. Ordnung (FOC) benötigt!
- z.B. $\sum_i \hat{\varepsilon}_i = 0$ und $\sum_i \hat{\varepsilon}_i x_i = 0$
- Jede FOC determiniert *ein* Residuum (*‘verliert Freiheit’*)

OLS Standardfehler

Freiheitsgrade:

- Für jeden der insgesamt k zu schätzenden Parameter wird eine Bedingung 1. Ordnung (FOC) benötigt!
- z.B. $\sum_i \hat{\varepsilon}_i = 0$ und $\sum_i \hat{\varepsilon}_i x_i = 0$
- Jede FOC determiniert *ein* Residuum (*‘verliert Freiheit’*)
- Für drei Residuen und FOC $\sum_{i=1}^{3} \hat{\varepsilon}_i = 0$:
  wenn $\hat{\varepsilon}_1 = -3$, $\hat{\varepsilon}_2 = 5$ folgt $\hat{\varepsilon}_3 = -2$!

OLS Standardfehler

Freiheitsgrade:

- Für jeden der insgesamt k zu schätzenden Parameter wird eine Bedingung 1. Ordnung (FOC) benötigt!
- z.B. $\sum_i \hat{\varepsilon}_i = 0$ und $\sum_i \hat{\varepsilon}_i x_i = 0$
- Jede FOC determiniert *ein* Residuum (*‘verliert Freiheit’*)
- Für drei Residuen und FOC $\sum_{i=1}^{3} \hat{\varepsilon}_i = 0$:
  wenn $\hat{\varepsilon}_1 = -3$, $\hat{\varepsilon}_2 = 5$ folgt $\hat{\varepsilon}_3 = -2$!
- $\Rightarrow$ wenn k Koeffizienten geschätzt werden verbleiben $n - k$ Freiheitsgrade!

OLS Standardfehler

Standardfehler der Koeffizienten (Wurzel aus Varianz)

$$\widehat{\mathrm{se}}(\widehat{\beta}_2) := \hat{\sigma}_{\widehat{\beta}_2} = \sqrt{\frac{\hat{\sigma}^2}{\sum [x_i - \bar{x}]^2}}, \qquad \widehat{\mathrm{se}}(\widehat{\beta}_1) := \hat{\sigma}_{\widehat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2 \sum x_i^2}{n \sum [x_i - \bar{x}]^2}}$$

mit

$$\hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$$

- Standardfehler (*standard error*) := Standardabweichung einer Schätzfunkt.
- Kennzahl für die Genauigkeit eines geschätzten Koeffizienten in Bezug auf die Stichprobenvariabilität!
- Werden für Hypothesentests benötigt.
- Sollten *immer* (!) unter (bzw. neben) Koeffizienten angegeben werden!

OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

(‘was macht eine Schätzung genau’?)

$$\widehat{\mathrm{se}}(\widehat{\beta}_2) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \qquad \text{mit } \hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$$

Ceteris paribus ist der Standardfehler umso kleiner, d.h. die Schätzung genauer, ...

1. je kleiner die Varianz der Störterme σ^2 (bzw. deren Schätzung $\hat{\sigma}^2$) ist,

OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

(‘was macht eine Schätzung genau’?)

$$\widehat{\text{se}}(\widehat{\beta}_2) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad \text{mit } \hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$$

Ceteris paribus ist der Standardfehler umso kleiner, d.h. die Schätzung genauer, ...

1. je kleiner die Varianz der Störterme σ^2 (bzw. deren Schätzung $\hat{\sigma}^2$) ist,
2. je *größer* die Streuung der x ist, d.h., je größer $\sum_i (x_i - \bar{x})^2$) ist,

OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

('was macht eine Schätzung genau'?)

$$\widehat{\mathrm{se}}(\widehat{\beta}_2) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad \text{mit } \hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$$

Ceteris paribus ist der Standardfehler umso kleiner, d.h. die Schätzung genauer, ...

1. je kleiner die Varianz der Störterme σ^2 (bzw. deren Schätzung $\hat{\sigma}^2$) ist,
2. je *größer* die Streuung der x ist, d.h., je größer $\sum_i (x_i - \bar{x})^2$) ist,
3. je größer n ist, da der Nenner $\sum_{i=1}^{n}(x_i - \bar{x})^2$ mit n zunimmt.

OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

(‘was macht eine Schätzung genau’?)

$$\widehat{\mathrm{se}}(\widehat{\beta}_2) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \qquad \text{mit } \hat{\sigma}^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$$

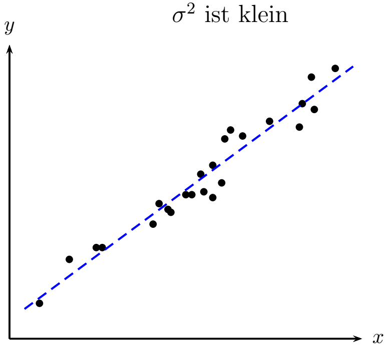
Ceteris paribus ist der Standardfehler umso kleiner, d.h. die Schätzung genauer, . . .

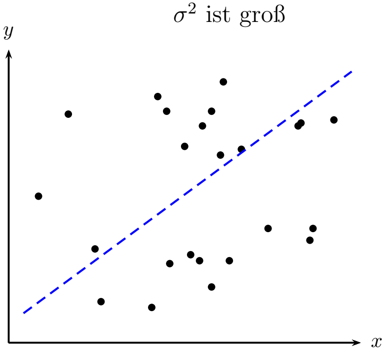
1. je kleiner die Varianz der Störterme σ^2 (bzw. deren Schätzung $\hat{\sigma}^2$) ist,
2. je *größer* die Streuung der x ist, d.h., je größer $\sum_i (x_i - \bar{x})^2$) ist,
3. je größer n ist, da der Nenner $\sum_{i=1}^{n}(x_i - \bar{x})^2$ mit n zunimmt.
4. Multiple Regression: je weniger die x untereinander korreliert sind.

OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

1) Regressionen mit unterschiedlicher Varianz von ε (σ^2):



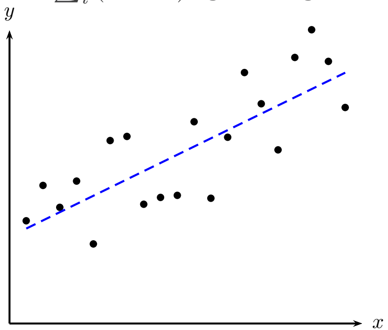


OLS Standardfehler

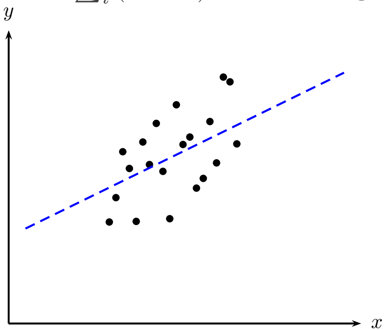
Determinanten der Standardfehler der Koeffizienten

2) Unterschiedliche Streuung der erklärenden x Variable:

$\sum_i (x_i - \bar{x})^2$ groß $\Rightarrow$ genau



$\sum_i (x_i - \bar{x})^2$ klein $\Rightarrow$ ungenau



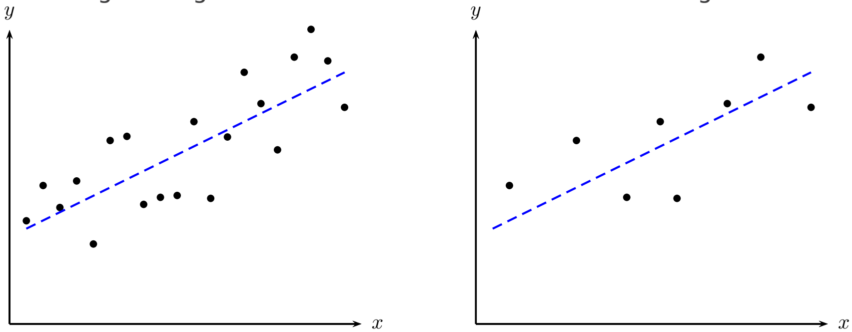
OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

3) Unterschiedliche Größe der Stichprobe (n):

n groß $\Rightarrow$ genau

n klein $\Rightarrow$ ungenau



OLS Standardfehler

Determinanten der Standardfehler der Koeffizienten

4) Multiple Regression:

$$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_{i2} + \widehat{\beta}_3 x_{i3} + \cdots + \widehat{\beta}_h x_{ih} + \cdots + \widehat{\beta}_k x_{ik} + \hat{\varepsilon}_i$$

Standardfehler des Koeffizienten h:

$$\widehat{\mathrm{se}}(\widehat{\beta}_h) = \sqrt{\frac{\hat{\sigma}^2}{(1 - R_h^2) \sum_i (x_{ih} - \bar{x}_h)^2}}$$

mit R_h^2 Bestimmtheitsmaß der Hilfsregression:

$$x_{ih} = \widehat{\alpha}_1 + \widehat{\alpha}_2 x_{i2} + \cdots + \widehat{\alpha}_{h-1} x_{ih-1} + \widehat{\alpha}_{h+1} x_{ih+1} + \cdots + \widehat{\alpha}_k x_{ik} + \nu_i \quad \rightarrow \quad R_h^2$$

$\Rightarrow$ je größer die lineare Abhängigkeit der x untereinander, je ungenauer die Schätzung!!! (Multikollinearität)

OLS Standardfehler

Robuste Standardfehler der Koeffizienten:

- Wenn

  A4: $\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$

  verletzt sind OLS Standardfehler *falsch* (d.h. verzerrt)!!!

OLS Standardfehler

Robuste Standardfehler der Koeffizienten:

- Wenn

  A4: $\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$

  verletzt sind OLS Standardfehler *falsch* (d.h. verzerrt)!!!
- Für solche Fälle wurden **robuste Standardfehler** (White, 1980) entwickelt.

OLS Standardfehler

Robuste Standardfehler der Koeffizienten:

- Wenn

$$\textbf{A4}: \qquad \varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$$

  verletzt sind OLS Standardfehler *falsch* (d.h. verzerrt)!!!
- Für solche Fälle wurden **robuste Standardfehler** (White, 1980) entwickelt.
- Falls A4 erfüllt sind diese zwar deutlich ungenauer,
  aber wenn A4 verletzt sind diese zumindest *konsistent* (d.h. werden mit zunehmender Stichprobengröße genauer).

OLS Standardfehler

Robuste Standardfehler der Koeffizienten:

- Wenn

  A4: $\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$

  verletzt sind OLS Standardfehler *falsch* (d.h. verzerrt)!!!
- Für solche Fälle wurden **robuste Standardfehler** (White, 1980) entwickelt.
- Falls A4 erfüllt sind diese zwar deutlich ungenauer,
  aber wenn A4 verletzt sind diese zumindest *konsistent* (d.h. werden mit zunehmender Stichprobengröße genauer).
- kommt später bei *Heteroskedastizität* ausführlicher . . .

Effizienz: Gauss-Markov Theorem

Gauss-Markov Theorem

Unter den Gauss-Markov Annahmen A1 – A4 des ‘klassischen linearen Regressionsmodells’ haben OLS-Schätzfunktionen innerhalb der Klasse aller linearen und erwartungstreuen Schätzfunktionen die kleinste Varianz, oder in anderen Worten, sie sind **BLUE**, d.h. OLS ist ein **B**est **L**inear **U**nbiased **E**stimator.

Effizienz: Gauss-Markov Theorem

Gauss-Markov Theorem

Unter den Gauss-Markov Annahmen A1 – A4 des ‘klassischen linearen Regressionsmodells’ haben OLS-Schätzfunktionen innerhalb der Klasse aller linearen und erwartungstreuen Schätzfunktionen die kleinste Varianz, oder in anderen Worten, sie sind **BLUE**, d.h. OLS ist ein **B**est **L**inear **U**nbiased **E**stimator.

Gauss-Markov Annahmen:

1. PRF ist linear und korrekt spezifiziert: $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon$
2. $n \geq k$ und keine perfekte Multikollinearität
3. Störterme und Regressoren sind stochastisch unabhängig: $\mathrm{E}(\varepsilon_i|\mathbf{X}) = \mathrm{E}(\varepsilon_i) = 0$
4. Störterme sind i.i.d.: $\varepsilon_i \sim \mathrm{i.i.d.}(0, \sigma^2)$

Effizienz: Gauss-Markov Theorem

OLS-Schätzfunktionen sind effizient, d.h. $\mathrm{var}(\widehat{\beta}_2^{\mathsf{OLS}}) \leq \mathrm{var}(\tilde{\beta}_2)$
($\tilde{\beta}_2$ ist eine beliebige lineare erwartungstreue Schätzfunktion für β_2)

Beweisidee:

1. Starte mit beliebiger linearen Schätzfunktion, z.B.
   $\tilde{\beta} = \sum_i c_i \, y_i$

Effizienz: Gauss-Markov Theorem

OLS-Schätzfunktionen sind effizient, d.h. $\mathrm{var}(\widehat{\beta}_2^{\text{OLS}}) \leq \mathrm{var}(\tilde{\beta}_2)$
($\tilde{\beta}_2$ ist eine beliebige lineare erwartungstreue Schätzfunktion für β_2)

Beweisidee:

1. Starte mit beliebiger linearen Schätzfunktion, z.B.
   $\tilde{\beta} = \sum_i c_i \, y_i$
2. Ermitteln notwendige Bedingungen für Erwartungstreue:
   $\mathrm{E}(\tilde{\beta}) = \beta$ wenn $\sum_i c_i = 0$ und $\sum_i c_i x_i = 1$ (siehe Appendix)

Effizienz: Gauss-Markov Theorem

OLS-Schätzfunktionen sind effizient, d.h. $\operatorname{var}(\widehat{\beta}_2^{\mathsf{OLS}}) \leq \operatorname{var}(\tilde{\beta}_2)$
($\tilde{\beta}_2$ ist eine beliebige lineare erwartungstreue Schätzfunktion für β_2)

Beweisidee:

1. Starte mit beliebiger linearen Schätzfunktion, z.B.
   $\tilde{\beta} = \sum_i c_i \, y_i$
2. Ermitteln notwendige Bedingungen für Erwartungstreue:
   $\mathrm{E}(\tilde{\beta}) = \beta$ wenn $\sum_i c_i = 0$ und $\sum_i c_i x_i = 1$ (siehe Appendix)
3. Minimieren Varianz dieser linearen Schätzfunktion unter Nebenbedingung der Erwartungstreue:
   $\min$: $\operatorname{var}(\tilde{\beta})$ unter NB: $\sum_i c_i = 0$ und $\sum_i c_i x_i = 1$; $\quad \rightarrow$ Lagrange

Effizienz: Gauss-Markov Theorem

OLS-Schätzfunktionen sind effizient, d.h. $\mathrm{var}(\widehat{\beta}_2^{\text{OLS}}) \leq \mathrm{var}(\tilde{\beta}_2)$
($\tilde{\beta}_2$ ist eine beliebige lineare erwartungstreue Schätzfunktion für β_2)

Beweisidee:

1. Starte mit beliebiger linearen Schätzfunktion, z.B.
   $\tilde{\beta} = \sum_i c_i\, y_i$
2. Ermitteln notwendige Bedingungen für Erwartungstreue:
   $\mathrm{E}(\tilde{\beta}) = \beta$ wenn $\sum_i c_i = 0$ und $\sum_i c_i x_i = 1$ (siehe Appendix)
3. Minimieren Varianz dieser linearen Schätzfunktion unter Nebenbedingung der Erwartungstreue:
   $\min$: $\mathrm{var}(\tilde{\beta})$ unter NB: $\sum_i c_i = 0$ und $\sum_i c_i x_i = 1$; $\rightarrow$ Lagrange
4. Lösung dieses Minimierungsproblems: $c_i = w_i$,
   also ist OLS Schätzfunktion *varianzminimal*!
   Beweisführung erfordert wieder Annahmen A1 – A4!

Zusammenfassung

Gauss-Markov Annahmen:

1. PRF ist linear und korrekt spezifiziert: $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon$
2. $n \geq k$ und keine perfekte Multikollinearität
3. Störterme und Regressoren sind stochastisch unabhängig:
   $\mathrm{E}(\varepsilon_i|\mathbf{X}) = \mathrm{E}(\varepsilon_i) = 0$
4. Störterme sind $\mathrm{i.i.d.}$: $\varepsilon_i \sim \mathrm{i.i.d.}(0, \sigma^2)$

- **Erwartungstreue:** A1 – A3 müssen erfüllt sein.
- **Effizienz:** A1 – A4 müssen erfüllt sein.

  Falls A1 – A3 erfüllt aber A4 verletzt ist sind die OLS Standardfehler verzerrt!
  $\Rightarrow$ Hypothesentests und Konfidenzintervalle ungültig!!!

Zusammenfassung

Gauss-Markov Annahmen:

1. PRF ist linear und korrekt spezifiziert: $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon$
2. $n \geq k$ und keine perfekte Multikollinearität
3. Störterme und Regressoren sind stochastisch unabhängig: $\mathrm{E}(\varepsilon_i|\mathbf{X}) = \mathrm{E}(\varepsilon_i) = 0$
4. Störterme sind $\mathrm{i.i.d.}$: $\varepsilon_i \sim \mathrm{i.i.d.}(0, \sigma^2)$

- **Erwartungstreue:** A1 – A3 müssen erfüllt sein.
- **Effizienz:** A1 – A4 müssen erfüllt sein.

  Falls A1 – A3 erfüllt aber A4 verletzt ist sind die OLS Standardfehler verzerrt!
  $\Rightarrow$ Hypothesentests und Konfidenzintervalle ungültig!!!

Thanx . . .