1 Definitionen

Mittelwert $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$

Varianz $\mathrm{VAR}(x)=s_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

Covarianz $\mathrm{COV}(x,y)=c_{xy}$
$=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$

Sample Regression Function $\hat{y}=b_1+b_2x$

Fittet Values / Vorhersage $\hat{y}_i=b_1+b_2x_i$

Residual $\hat{u}_i=y_i-\hat{y}_i$

Fehlerterm $u_i=y_i-E(y_i\mid x_i)$

Bestimmtheitsmaß $R^2=\frac{\mathrm{var}(\hat{y})}{\mathrm{var}(y)}=1-\frac{\mathrm{var}(\hat{u})}{\mathrm{var}(y)}$

Geschätzter Standardfehler: $\hat{\sigma}_{b_k}=\sqrt{\mathrm{var}(b_k)}$

2 OLS

2.1 Herleitungen der Parameter

2.1.1 Momentenmethode

$$\begin{aligned}
u &= y-\beta_1-\beta_2x\\
E(u) &= \mathrm{E}(y-\beta_1-\beta_2x)=0\\
&\Rightarrow \frac{1}{n}\sum(y_i-b_1-b_2x_i)=0\\
\mathrm{E}(x\cdot u) &= E[x\cdot(y-\beta_1-\beta_2x)]=0\\
&\Rightarrow \frac{1}{n}\sum[x_i\cdot(y_i-b_1-b_2x_i)]=0
\end{aligned}$$

$$\begin{aligned}
b_1 &= \bar{y}-b_2\bar{x}\\
b_2 &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{S_{xy}}{S_x^2}
\end{aligned}$$

2.1.2 Minimierungsproblem

$$\min_{b_1,b_2}\sum_{i=1}^{N}\hat{e}_i^2=\min_{b_1,b_2}\sum_{i=1}^{N}(y_i-b_1-b_2x_i)^2$$

$$\begin{aligned}
\frac{\partial\sum\hat{e}^2}{\partial b_1} &= -2\sum\hat{e}=0\\
&\Rightarrow \sum_N(y-b_1-b_2x)=0\\
&\Leftrightarrow \sum y=\sum b_1+b_2\sum x\\
&\Leftrightarrow \bar{y}N=Nb_1+b_2\sum\bar{x}N\\
&\Leftrightarrow b_1=\bar{y}-b_2\bar{x}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\sum\hat{e}^2}{\partial b_2} &= -2\sum\hat{e}x=0\\
&\Rightarrow \sum(y-b_1-b_2x)x=0\\
&\Leftrightarrow \sum(y-(\bar{y}-b_2\bar{x})-b_2x)x=0\\
&\Leftrightarrow \sum(y-\bar{y}-b_2(x-\bar{x}))x=0\\
&\Leftrightarrow \sum(y-\bar{y})_x=b_2\sum(x-\bar{x})x\\
&\Leftrightarrow b_2=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}=\frac{S_{xy}}{S_x^2}
\end{aligned}$$

Substitutionen:

$$\begin{aligned}
\bar{x} &= \frac{1}{N}\sum_N x\\
\bar{x}N &= \sum_N x\\
\sum(x-\bar{x})x &= \sum(x-\bar{x})^2\\
\sum(y-\bar{y})x &= \sum(y-y)(x-\bar{x})
\end{aligned}$$

2.2 Annahmen OLS

1. Linearität von $y=\beta_2+\beta_2x+e$
2. Stichprobenvarianz
3. Zufallsstichprobe [1] (x_i,y_i) ist unabhängig von (x_j,y_j)
4. Konditionaler Erwartungswert = 0 = $E(e\mid x)$
5. Homoskedastizität $\mathrm{var}(e_i\mid X_i)=\sigma^2$
6. Die Verteilung des Störterms, konditional auf x_i, ist normal: $e_i\mid x_i\sim\mathcal{N}(0,\sigma^2)$

[1] Impliziert das eine unabhängige und identische Verteilung vorliegt (independet and identical distribution (iid)). Bei beobachteten Daten ist diese Annahme sehr Kritisch. Hier lässt sich nicht sicherstellen das der datangenerierende Prozess wirklich zufällig ist.

2.3 Eigenschaften OLS

1. Summe der OLS-Residuen = 0
   $\sum_i \hat{e}_i = \overline{\hat{e}} = 0$
2. gefitterer Mittelwert = beobachteter Mittelwert

$$\hat{e}_i = y_i - \hat{y}_i \rightarrow \sum_i \hat{e}_i = \sum_i y_i - \sum_i \hat{y}_i$$
$$\frac{1}{N}\sum_i \hat{e}_i = \frac{1}{N}\sum_i y_i - \frac{1}{N}\sum_i \hat{y}_i$$
$\overline{\hat{e}} = \bar{y} - \overline{\hat{y}} = 0$ wenn $\overline{\hat{e}} = 0$ (siehe 1.) gilt weiter
$\overline{\hat{y}} = \bar{y}$

(gilt nur mit Interzept)

3. $\sum_i \hat{e}_i x_i = 0 \vee \operatorname{cov}(x_i, \hat{e}_i) = 0$

$\sum_i (x_i - \bar{x})\left(\hat{e}_i - \overline{\hat{e}}\right) = \sum_i x_i\hat{e}_i - \bar{x}\sum_i \hat{e}_i = 0$ wenn

$\sum_i \hat{e}_i = \overline{\hat{e}} = 0$

4. $\operatorname{cov}(\hat{y}_i, \hat{e}_i) = 0$

$$\begin{aligned}\operatorname{cov}(\hat{y}_i, \hat{e}_i) &= \operatorname{cov}([b_1 + b_2 x_i], \hat{e}_i) \\ &= \operatorname{cov}(b_1, \hat{e}_i) + \operatorname{cov}(b_2 x_i, \hat{e}_i) \\ &= 0 + b_2 \operatorname{cov}(x_i, \hat{e}_i) \\ &= 0\end{aligned}$$

(mit Interzept)

2.4 Qualität OLS

Der OLS-Schätzer liefert für verschiedene Stichproben verschieden Werte. Die Schätzer selbst sind damit Zufallsvariablen.

2.4.1 Eigenschafen

- Erwartungstreue

$$E[b_2] = \beta_2$$

- Effizienz (Geringe Stichprobenvariation)

$$\operatorname{var}[b_2] < \operatorname{var}[b_2^*]$$

- Konsistenz (Ein Schätzer wird mit zunehmender Stichprobengröße genauer)

$$\lim_{N\to\infty} P[|b_N - \beta| < \delta] = 1 \quad \delta > 0$$

2.4.2 Unverzerrtheit

Erwartungstreue/Unverzerrtheit $E(\hat{\theta}) = \theta$
Konsistenz $\operatorname{var}(\hat{\theta}) \to 0,$ für $n \to \infty$

Relation zwischen β_2 und b_2 herstellen

$$\begin{aligned}b_2 &= \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \\ &= \frac{\sum_i (x_i - \bar{x})(\beta_1 + \beta_2 x_i + e_i - \beta_1 - \beta_2\bar{x} - \bar{e})}{\sum_i (x_i - \bar{x})^2} \\ &= \frac{\sum_i (x_i - \bar{x})\beta_2(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} + \frac{\sum_i (x_i - \bar{x})(e_i - \bar{e})}{\sum_i (x_i - \bar{x})^2} \\ &= \beta_2 + \frac{\sum_i (x_i - \bar{x})(e_i - \bar{e})}{\sum_i (x_i - \bar{x})^2}\end{aligned}$$

Substitutionen:

$$\sum (x_i - \bar{x}) = 0$$
$$\sum x_i - n \cdot \bar{x} = 0$$

Erwartungswert Bilden

$$\begin{aligned}E(b_2) &= \beta_2 + \sum_i E\left(\frac{(x_i - \bar{x})(e_i - \bar{e})}{\sum_i (x_i - \bar{x})^2}\right) \\ &= \beta_2 + \sum_i \frac{E((x_i - \bar{x})(e_i - 0))}{\sum_i (x_i - \bar{x})^2} \\ &= \beta_2 + \sum_i \frac{E(cov(x_i, e_i))}{\sum_i (x_i - \bar{x})^2}\end{aligned}$$

Unverzerrtheit von b_1
Unter der Annahme, dass b_2 unverzerrt ist gilt:

$$E(b_1) = E\left(\left[\beta_1 + \beta_2\bar{X}\right] - b_2\bar{X}\right) = \beta_1$$

Beispiel

$$T_1 = \bar{X}, \quad X \sim \mathcal{N}\left(\mu, \sigma^2\right), \quad E(X) = \mu, \quad \operatorname{var}(X) = \sigma^2$$

$$\begin{aligned} E\left(T_1\right) = E(\bar{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} n\mu = \mu \end{aligned}$$

$$\begin{aligned} \operatorname{var}\left(T_1\right) = \operatorname{var}(\bar{X}) &= \operatorname{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{var}\left(X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \to 0 \end{aligned}$$

2.4.3 Varianz und Kovarianz von OLS

$$\begin{aligned} \operatorname{var}(b_1) &= E\left(\left[b_1 - E\left(b_1\right)\right]^2\right) = \sigma^2\left[\frac{\sum_i x_i^2}{N\sum_i\left(x_i - \bar{x}\right)^2}\right] \\ \operatorname{var}(b_2) &= \frac{\sigma^2}{\sum_i\left(x_i - \bar{x}\right)^2} \\ \operatorname{cov}\left(b_1, b_2\right) &= E\left(\left[b_1 - E\left(b_1\right)\right]\left[b_2 - E\left(b_2\right)\right]\right) \\ &= E\left(\left[b_1 - \beta_1\right]\left[b_2 - \beta_2\right]\right) \\ &= -\bar{x} E\left(\left[b_2 - \beta_2\right]^2\right) \\ \operatorname{cov}(b_1, b_2) &= -\bar{x}\operatorname{var}\left(b_2\right) \end{aligned}$$

Hereitung $\operatorname{var}(b_2)$

$$\begin{aligned} b_2 &= \beta_2 + \frac{\sum_i\left(x_i - \bar{x}\right)e_i}{\sum_i\left(x_i - \bar{x}\right)^2} \\ b_2 &= \beta_2 + \frac{1}{\sum\left(x_i - \bar{x}\right)^2}\sum\left(x_i - \bar{x}\right)e_i \\ \operatorname{var} &= \operatorname{var}\left(\beta_2\right) + \operatorname{var}\left(\frac{1}{\sum(1 - \bar{x})^2}\sum\left(x_i - \bar{x}\right)_{e_i}\right) \\ &= \left(\frac{1}{\sum\left(x_1 - \bar{x}\right)}^2\right)^2 \operatorname{var}\left(\sum\left(x_i - \bar{x}\right)e_i\right) \\ \operatorname{var}\left(b_2\right) &= \frac{\sum_i\left(x_i - \bar{x}\right)^2\operatorname{var}\left(e_i\right)}{\left(\sum_i\left(x_i - \bar{x}\right)^2\right)^2} \\ \operatorname{var}(b_2) &= \frac{\sigma^2}{\sum_i\left(x_i - \bar{x}\right)^2} \end{aligned}$$

Substitutionen:

$$\begin{aligned} \operatorname{var}(ax) &= a^2\operatorname{var}(x) \\ \operatorname{var}(ax + by) &= a^2\operatorname{var}(x) + b^2\operatorname{var}(y) + 2ab\operatorname{cov}(x, y) \end{aligned}$$

2.4.4 Schätzer $\hat{\sigma}^2$

$$\begin{aligned} \sigma^2 &= \frac{E\left[\sum_i \hat{e}_i^2\right]}{N - 2} \\ \hat{\sigma}^2 &= \frac{\sum_i \hat{e}_i^2}{N - 2} \\ \hat{\sigma}^2 &= \frac{\sum_i \hat{e}_i^2}{N - K} \leftarrow \text{für K Parameter} \end{aligned}$$

$$\hat{\sigma} = \sqrt{\frac{\sum_i \hat{e}_i^2}{N - K}}$$

2.5 Bestimmtheitsmaß R^2

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS}$$

2.5.1 Herleitung

$$
\begin{aligned}
y_i &= \hat{y}_i + \hat{e}_i \\
\operatorname{var}(y) &= \operatorname{var}(\hat{y}) + \operatorname{var}(\hat{e}) + 2\operatorname{cov}(\hat{y},\hat{e}) \\
&= \operatorname{var}(\hat{y}) + \operatorname{var}(\hat{e}) \\
\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})^2 &= \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i-\overline{\hat{y}}\right)^2 + \frac{1}{N}\sum_{i=1}^{N}\left(\hat{e}_i-\overline{\hat{e}}\right)^2 \\
\underbrace{\sum_{i=1}^{N}(y_i-\bar{y})^2}_{\text{TSS total sum of squares}} &= \underbrace{\sum_{i=1}^{N}\left(\hat{y}_i-\overline{\hat{y}}\right)^2}_{\text{ESS explained sum squared}} + \underbrace{\sum_{i=1}^{N}\left(\hat{e}_i-\overline{\hat{e}}\right)^2}_{\text{SSR sum of squard residuals}}
\end{aligned}
$$

2.6 Gauss-Markov-Theorem

> Unter den Gauss'schen Annahmen des klassischen linearen Regressionsmodells hat der OLS-Schätzer innerhalb der Klasse aller linearen und erwartungstreuen Schätzfunktionen die leinste Varianz, oder in anderen Worten, er ist BLUE d.h. ein **B**est **L**inear **U**nbiased **E**stimator.

2.7 Asymptotische Eigenschaften

Untersuchen das Verhalten von Zufallsvariablen die gegen $\infty (\approx 150)$ gehen.

2.7.1 Konsistenz

> b_N ist eine konsistente Schätzfunktion für β wenn gilt:
>
> $$\lim_{N\to\infty} P\left[|b_N-\beta|<\delta\right]=1 \quad \delta>0$$
>
> das heißt, dass die Wahrscheinlichkeit (P), dass mit steigendem Stichprobenumfang der Absolutbetrag der Differenz zwischen b_N und β kleiner als eine beliebig kleine Zahl δ wird, gegen 1 konvergiert.

2.7.2 Asymptotische Normalverteilung

Schätzer aus einer Grundgesamtheit konvergieren bei genügend Stichproben gegen Normalverteilung. Egal, ob die Grundgesamtheit normalverteilt ist oder nicht.

2.7.3 Asymptotische Effizienz

- b sei ein Schätzer für β.
- Die Varianz der asymptotischen Verteilung von b heißt asymptotische Varianz von b.
- Wenn b konsistent ist und die asymptotische Varianz kleiner ist als die aller anderen konsistenten Schätzer, dann heißt b asymptotisch effizient.

3 Intervallschätzung und Hypothesentests

Annahmen über die Verteilung des Fehlerterms. Die Annahme 6. der Gauss'schen Annahmen sorgt dafür, dass der OLS Schätzer die kleinste Varianz innerhalb der unverzerrten Schätzer besitzt. Die Annahme ist nicht kritisch, da der Schätzer asymptotisch gegen die Normalverteilung strebt.

3.1 Intervalschätzer

3.1.1 Schritte

1. Berechnung des Punktschätzers b_k und eines Schätzers $\hat{\sigma}^2$
2. Festlegung von α und bestimmen von $t^c_{\alpha/2}$ mit $N-K$ Freiheitsgraden
3. Berechnung des Intervalschätzers

$$P\left[b_k - t^c_{\alpha/2}\hat{\sigma}_{b_k} \le \beta_k \le b_k + t^c_{\alpha/2}\hat{\sigma}_{b_k}\right] = 1-\alpha$$

4. Interpretation des Intervalschätzers: $(1-\alpha)\times 100\%$ der Konfidenzintervalle enthalten β

Konfidenzintervalle sind enger wenn:

- je größer α
- je kleiner σ^2
- je größer N

Standardfehler

$$
\begin{aligned}
\operatorname{var}(e_i) &\equiv \sigma^2 \\
\operatorname{var}(b_2) &= \sigma^2_{b_2} = \sigma^2 / \sum_i (x_i-\bar{x})^2
\end{aligned}
$$

Punktschätzer dienen zur Schätzung eines unbekannten Parameters einer Grundgesamtheit auf grundlage einer Einyelnen Stichprobe. Vermitteln allerdings keine Informationen über die Unsicherheit.

3.1.2 Normalverteilung

$$\begin{aligned}
&P[z \leq -1.96] = P[z \geqslant +1,96] = 0.025 \\
&P[-1.96 \leq z \leq +1.96] = 1 - 0.05 = 0.95 \\
&P\left[-1.96 \leq \frac{b_k - \beta_k}{\sigma_{b_k}} \leq +1.96\right] = 0.95 \\
&P\left[-z_{\alpha/2} \leq \frac{b_k - \beta_k}{\sigma_{b_k}} \leq +z_{\alpha/2}\right] = 1 - \alpha
\end{aligned}$$

3.1.3 T-Verteilung

> Der Wert der ausgegenenen t-Statistik ist einfach Koeffizient dividiert durch Stadardfehler:
>
> $$t - Stat(b_k) = \frac{b_k}{\hat{\sigma}\ _{b_k}}$$

$$\begin{aligned}
&P\left[-t^c_{\alpha/2} \leq \frac{b_k - \beta_k}{\hat{\sigma}_{b_k}} \leq +t^c_{\alpha/2}\right] = 1 - \alpha \\
&P\left[b_k - t^c_{\alpha/2}\hat{\sigma}_{b_k} \leq \beta_k \leq b_k + t^c_{\alpha/2}\hat{\sigma}_{b_k}\right] = 1 - \alpha
\end{aligned}$$

$$\begin{aligned}
H_0 : \mu = \mu_0, &\quad H_1 : \mu \neq \mu_0 \\
H_0 : \mu \geq \mu_0, &\quad H_1 : \mu < \mu_0 \\
H_0 : \mu \leq \mu_0, &\quad H_1 : \mu > \mu_0
\end{aligned}$$

3.2 Einfache Hypothesentests

1. Ausgangssituation: Formuliere die Annahmen und nicht zu testenden Hintergrundhypothesen (z.B. b ist normalverteilt mit …)
2. Formuliere die Alternativ- und Nullhypothese z.B. $H_1 : \beta = 0, H_0 : \beta = 0$
3. Wähle die Teststatistik (z.B. t-Stat $= b/\hat{\sigma}_b$)
4. Bestimme die Verteilung der Teststatistik unter der Annahme, dass die Nullhypothese gilt z.B. $b/\hat{\sigma}_b \sim t_{N-2}$.
5. Wähle das Signifikanzniveau und bestimme den Annahme- und Verwerfungsbereich (z.B. behalte die Nullhypothese wenn $-1.96 \leq b/\hat{\sigma}_b \leq +1.96$; anderenfalls verwirf die Nullhypothese).

3.3 Zweiseitige Hypothesentests

$$\begin{aligned}
&P\left[\beta^0 - t^c_{\alpha/2}\hat{\sigma}_b \leqslant b \leqslant \beta^0 + t^c_{\alpha/2}\hat{\sigma}_b\right] = 1 - \alpha \\
&\left[\beta^0 - t^c_{\alpha/2}\hat{\sigma}_b; \beta^0 + t^c_{\alpha/2}\hat{\sigma}_b\right]
\end{aligned}$$

3.4 Fehlerarten

1. Fehler 1. Art: Nullhypothese wird fälschlicherweise Verworfen
2. Fehler 2. Art: Nullhypothese wird fälschlicherweise Akzeptiert

4 Logarithmen

$$\begin{aligned}
y &= a \cdot x^{\beta_2} \cdot v \\
\log(y) &= \log\left(a \cdot x^{\beta_2} \cdot v\right) = \underbrace{\log(a)}_{\beta_1} + \beta_2 \cdot \log(x) + \underbrace{\log(v)}_{u} \\
\Rightarrow \log(y) &= \beta_1 + \beta_2 \log(x) + u
\end{aligned}$$

$$\frac{\partial y}{\partial x} = \beta_2 \frac{y}{x} \quad \Leftrightarrow \quad \beta_2 = \frac{\partial y}{\partial x}\frac{x}{y}$$

Level-level	y	x	$\Delta y =$	$\beta_2 \Delta x$
Level-log	y	$\log(x)$	$\Delta y =$	$(\beta_2/100)\,\%\Delta x$
Log-level	$\log(y)$	x	$\%\Delta y =$	$(100\beta_2)\,\Delta x$
Log-log	$\log(y)$	$\log(x)$	$\%\Delta y =$	$\beta_2 \%\Delta x$

5 Multivariates Modell

Die Annahme, dass alle Variablen gleich und unabhängig Verteilt sind ist bei realen Daten oft nicht gegeben. Variablen können sich statistisch beeinflussen, ohne das ein kausaler Zusammenhang besteht.

5.1 Heteroskedastizität

> Ungleiche Varianz der Störterme.

Die Varianz der einzelnen Fehlerterme ist definiert als:

$$\mathrm{var}(e_i) = \sigma_i^2$$

Die Annahmen $e_i \sim iid(0, \sigma_i^2)$, $E(e_i) = 0$ und $\mathrm{cov}(e_i, e_j) = 0$ gelten weiterhin.

Der OLS-Schätzer bleibt erwartungstreu und konsitent.

5.1.1 Maßnahmen gegen Heteroskedastizität

- Heteroskedastiekonsistente Standardfehler
  - $\mathrm{var}(b_2) = \sum_i^N \frac{(x_i - \bar{x})^2 \sigma_i^2}{\left(\sum_i^N (x_i - \bar{x})^2\right)^2}$
- weighted - generalized Least Squares (WLS and GLS)