Aufgaben zur Ökonometrie I

3. Signifikanztests und Konfidenzintervalle

3.1 Wie groß ist der Standardfehler der Regression von GASV auf VEINKR (≙ Energiemodell Ib, s. Ergebnisse in Aufgabe 2.8) (mit Interpretation)?

Lösung:

Der **Standardfehler der Regression** (SER) ist gleich der **Wurzel aus der Residualvarianz**, bei der die Summe der quadrierten Residuen auf die Anzahl der Freiheitsgrade bezogen ist:

$$\text{SER} = \hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k}}$$

Arbeitstabelle:

t	y_t	$\hat{y}_t$	$\hat{u}_t$	$\hat{u}_t^2$
1	1,0	0,995	0,005	0,000025
2	0,8	0,995	-0,195	0,038025
3	0,7	0,638	0,062	0,003844
4	0,9	0,876	0,024	0,000576
5	1,2	1,114	0,086	0,007396
6	1,1	1,114	-0,014	0,000196
7	1,4	1,352	0,048	0,002304
Σ	7,1	≈7,1		0,051826

Residualvarianz:

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \frac{0{,}051826}{7-2} = \frac{0{,}051826}{5} = 0{,}010365$$

Standardfehler der Regression (SER)**:**

$$\text{SER} = \hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{0{,}010365} = 0{,}102$$

<u>Interpretation:</u>

Der Standardfehler der Regression ist die Wurzel aus der erwartungstreu geschätzten Varianz der Störgröße. Er gibt die mittlere Abweichung der y-Werte von den Regressionswerten $\hat{y}_t$ an. Bei der Regressionsschätzung des Gasverbrauchs in Abhängigkeit vom verfügbaren Einkommen wird danach ein durchschnittlicher Fehler von 0,102 Einheiten gemacht.

3.2 Die OLS-Schätzer der Regressionskoeffizienten β_1 und β_2 der Nachfragefunktion für Erdgas (≙ Energiemodell Ia) lauten (n=7): $\hat{\beta}_1 = 5{,}809$ und $\hat{\beta}_2 = -0{,}368$. Die Summe der quadrierten Residuen beträgt 0,067, die Summe der Gaspreise 91,2 und die Summe der quadrierten Gaspreise 1190,28.

Testen Sie die OLS-Schätzer der Regressionskoeffizienten auf Signifikanz (α=0,01)!

Lösung:

A. Signifikanztests (zweiseitig) über $\hat{\beta}_1$

1. Schritt	Hypothesenformulierung: H_0: $ß_1 = 0$, H_1: $ß_1 \neq 0$
2. Schritt	Festlegung des Signikanzniveaus: $\alpha = 0{,}01$
3. Schritt	Berechnen der Prüfgröße: Residualvarianz: $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \frac{0{,}067}{7-2} = 0{,}0134$ Geschätzte Varianz von $\hat{\beta}_1$: $\hat{Var}(\hat{\beta}_1) = \hat{\sigma}^2 \frac{\sum x_t^2}{n\sum x_t^2 - (\sum x_t)^2} = 0{,}0134 \cdot \frac{1190{,}28}{7 \cdot 1190{,}28 - 91{,}2^2}$ $= 0{,}0134 \cdot 81{,}975 = 1{,}0985$ Prüfgröße: $t = \frac{\hat{\beta}_1}{\sqrt{\hat{Var}(\hat{\beta}_1)}} = \frac{5{,}809}{\sqrt{1{,}0985}} = \frac{5{,}809}{1{,}048} = 5{,}543$
4. Schritt	Kritischer Wert: $t_{n-k;1-\alpha/2} = t_{7-2;0{,}995} = t_{5;0{,}995} = 4{,}032$
5. Schritt	Testentscheidung: $\lvert t \rvert = 5{,}543 > t_{5;0{,}995} = 4{,}032$ → Nullhypothese ablehnen

B. Signifikanztest (zweiseitig) über $\hat{\beta}_2$:

1. Schritt	Hypothesenformulierung: H_0: $ß_2 = 0$, H_1: $ß_2 \neq 0$
2. Schritt	Festlegung des Signikanzniveaus: $\alpha = 0{,}01$
3. Schritt	Berechnen der Prüfgröße: Residualvarianz: $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \frac{0{,}067}{7-2} = 0{,}0134$ Geschätzte Varianz von $\hat{\beta}_2$: $\hat{Var}(\hat{\beta}_2) = \hat{\sigma}^2 \frac{n}{n\sum x_t^2 - (\sum x_t)^2} = 0{,}0134 \cdot \frac{7}{7 \cdot 1190{,}28 - 91{,}2^2}$ $= 0{,}0134 \cdot 0{,}4821 = 0{,}00646$ Prüfgröße: $t = \frac{\hat{\beta}_2}{\sqrt{\hat{Var}(\hat{\beta}_2)}} = \frac{-0{,}368}{\sqrt{0{,}00646}} = \frac{-0{,}368}{0{,}0804} = -4{,}577$
4. Schritt	Kritischer Wert: $t_{n-k;1-\alpha/2} = t_{7-2;0{,}995} = t_{5;0{,}995} = 4{,}032$
5. Schritt	Testentscheidung: $\lvert t \rvert = 4{,}577 > t_{5;0{,}995} = 4{,}032$ → Nullhypothese ablehnen

3.3 Gegeben ist die OLS-geschätzte Nachfragefunktion (≙ Energiemodell II)

$$\widehat{GASV}_t = 2{,}706 + 0{,}199 \cdot GASP_t + 0{,}074 \cdot VEINK_t$$

und die Inverse der Produktmatrix **X'X**

$$(\mathbf{X'X})^{-1} = \begin{pmatrix} 255{,}7066 & -15{,}7459 & -4{,}1520 \\ -15{,}7459 & 0{,}997745 & 0{,}2262 \\ -4{,}1520 & 0{,}2262 & 0{,}099228 \end{pmatrix}.$$

Die Summe der quadrierten Residuen beträgt 0,0127.

Testen Sie den Einfluss des Gaspreises (GASP) und des verfügbaren Einkommens (VEINK) auf die Nachfrage nach Erdgas auf Signifikanz (α=0,05)!

Lösung:

A. Signifikanztest (zweiseitig) über $\hat{\beta}_2$ (Einfluss des Gaspreises)

1. Schritt	Hypothesenformulierung: H_0: $ß_2 = 0$, H_1: $ß_2 \neq 0$
2. Schritt	Festlegung des Signikanzniveaus: $\alpha = 0{,}05$
3. Schritt	Berechnen der Prüfgröße: Residualvarianz: $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \frac{0{,}0127}{7-3} = 0{,}003175 \rightarrow \hat{\sigma} = \sqrt{0{,}003175} = 0{,}056347$ Prüfgröße: $t = \frac{\hat{\beta}_2}{\hat{\sigma} \cdot \sqrt{x^{22}}} = \frac{0{,}199}{0{,}056347 \cdot \sqrt{0{,}997745}} = \frac{0{,}199}{0{,}056283} = 3{,}536$
4. Schritt	Kritischer Wert: $t_{n-k;1-\alpha/2} = t_{7-3;0{,}975} = t_{4;0{,}975} = 2{,}776$
5. Schritt	Testentscheidung: $\lvert t \rvert = 3{,}536 > t_{4;0{,}975} = 2{,}776 \rightarrow$ Nullhypothese ablehnen

B. Signifikanztest (zweiseitig) über $\hat{\beta}_3$ (Einfluss des verfügbaren Einkommens)

1. Schritt	Hypothesenformulierung: H_0: $ß_3 = 0$. H_1: $ß_3 \neq 0$
2. Schritt	Festlegung des Signikanzniveaus: $\alpha = 0{,}05$
3. Schritt	Berechnen der Prüfgröße: Residualvarianz: $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \frac{0{,}0127}{7-3} = 0{,}003175 \rightarrow \hat{\sigma} = \sqrt{0{,}003175} = 0{,}056347$ Prüfgröße: $t = \frac{\hat{\beta}_2}{\hat{\sigma} \cdot \sqrt{x^{22}}} = \frac{0{,}074}{0{,}056347 \cdot \sqrt{0{,}099228}} = \frac{0{,}074}{0{,}017750} = 4{,}169$
4. Schritt	Kritischer Wert: $t_{n-k;1-\alpha/2} = t_{7-3;0{,}975} = t_{4;0{,}975} = 2{,}776$
5. Schritt	Testentscheidung: $\lvert t \rvert = 4{,}169 > t_{4;0{,}975} = 2{,}776 \rightarrow$ Nullhypothese ablehnen

3.4 Bestimmen Sie die 95%-Konfidenzintervalle unbekannten Regressionskoeffizienten β_1 und β_2 des Energiemodells Ia (s. hierzu Angaben in Aufgabe 3.2)!

Lösung:

A. 95%-Konfidenzintervall für β_1:

1. Schritt	Festlegung des Konfidenzniveaus 1-α: $1-\alpha = 0{,}95$
2. Schritt	95%-Konfidenzintervall für β_1 $P\left(\hat{\beta}_j - t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1} \le \beta_j \le \hat{\beta}_j + t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1}\right) = 1-\alpha$ mit $\hat{\sigma}^2_{\hat{\beta}_1} = \hat{\sigma}^2 \dfrac{\sum x_t^2}{n\sum x_t^2 - (\sum x_t)^2}$ und $\hat{\sigma}^2 = \dfrac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \dfrac{0{,}067}{7-2} = 0{,}0134 \rightarrow \hat{\sigma} = \sqrt{0{,}0134} = 0{,}1158$
3. Schritt	(1-α/2)-Quantil der t-Verteilung mit n-k FG $t_{n-k;1-\alpha/2} = t_{7-2;0{,}975} = t_{5;0{,}975} = 2{,}571$
4. Schritt	Konkretes 95%-Konfidenzintervall $\left[\hat{\beta}_1 - t_{5;0{,}975} \cdot \hat{\sigma}_{\hat{\beta}_1};\ \hat{\beta}_1 + t_{5;0{,}975} \cdot \hat{\sigma}_{\hat{\beta}_1}\right] =$ $= \left[\hat{\beta}_1 - t_{5;0{,}975} \cdot \hat{\sigma}\sqrt{\dfrac{\sum x_t^2}{n\sum x_t^2 - (\sum x_t)^2}};\ \hat{\beta}_1 + t_{5;0{,}975} \cdot \hat{\sigma}\sqrt{\dfrac{\sum x_t^2}{n\sum x_t^2 - (\sum x_t)^2}}\right] =$ $\left[5{,}809 - 2{,}571 \cdot 0{,}1158 \cdot \sqrt{\dfrac{1190{,}28}{7 \cdot 1190{,}28 - 91{,}2^2}};\ 5{,}809 + 2{,}571 \cdot 0{,}1158 \cdot \sqrt{\dfrac{1190{,}28}{7 \cdot 1190{,}28 - 91{,}2^2}}\right]$ $= \left[5{,}809 - 2{,}571 \cdot 0{,}1158 \cdot \sqrt{81{,}975};\ 5{,}809 + 2{,}571 \cdot 0{,}1158 \cdot \sqrt{81{,}975}\right]$ $= \left[5{,}809 - 2{,}571 \cdot 1{,}048;\ 5{,}809 + 2{,}571 \cdot 1{,}048\right]$ $= \left[5{,}809 - 2{,}694;\ 5{,}809 + 2{,}694\right] = [3{,}115;\ 8{,}503]$

B. 95%-Konfidenzintervall für β_2:

1. Schritt	Festlegung des Konfidenzniveaus 1-α: $1-\alpha = 0{,}95$
2. Schritt	95%-Konfidenzintervall für β_1 $P\left(\hat{\beta}_j - t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1} \le \beta_j \le \hat{\beta}_j + t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1}\right) = 1-\alpha$ mit $\hat{\sigma}^2_{\hat{\beta}_1} = \hat{\sigma}^2 \dfrac{n}{n\sum x_t^2 - (\sum x_t)^2}$ und $\hat{\sigma}^2 = \dfrac{\sum_{t=1}^{n} \hat{u}_t^2}{n-k} = \dfrac{0{,}067}{7-2} = 0{,}0134 \rightarrow \hat{\sigma} = \sqrt{0{,}0134} = 0{,}1158$
3. Schritt	(1-α/2)-Quantil der t-Verteilung mit n-k FG $t_{n-k;1-\alpha/2} = t_{7-2;0{,}975} = t_{5;0{,}975} = 2{,}571$
4. Schritt	Konkretes 95%-Konfidenzintervall $\left[\hat{\beta}_2 - t_{5;0{,}975} \cdot \hat{\sigma}_{\hat{\beta}_2}; \hat{\beta}_2 + t_{5;0{,}975} \cdot \hat{\sigma}_{\hat{\beta}_2}\right] =$ $\left[\hat{\beta}_2 - t_{5;0{,}975} \cdot \hat{\sigma} \dfrac{n}{n\sum x_t^2 - (\sum x_t)^2};\ \hat{\beta}_2 + t_{5;0{,}975} \cdot \hat{\sigma} \dfrac{n}{n\sum x_t^2 - (\sum x_t)^2}\right] =$ $\left[-0{,}368 - 2{,}571 \cdot 0{,}1158 \cdot \sqrt{\dfrac{7}{7 \cdot 1190{,}28 - 91{,}2^2}};\ -0{,}368 + 2{,}571 \cdot 0{,}1158 \cdot \sqrt{\dfrac{7}{7 \cdot 1190{,}28 - 91{,}2^2}}\right]$ $= \left[-0{,}368 - 2{,}571 \cdot 0{,}1158 \cdot 0{,}6943;\ -0{,}368 + 2{,}571 \cdot 0{,}1158 \cdot 0{,}6943\right]$ $= \left[-0{,}368 - 2{,}571 \cdot 0{,}0804;\ -0{,}368 + 2{,}571 \cdot 0{,}0804\right]$ $= \left[-0{,}368 - 0{,}207;\ -0{,}368 + 0{,}207\right] = [-0{,}575;\ -0{,}161]$

3.5 Berechnen Sie für das Energiemodell II (s. Aufgabe 2.5) die beiden Residuen für das erste und letzte Jahr des Stützbereichs! Welche Aussage über die Güte der Anpassung an den Rändern lässt sich aus ihnen gewinnen?

Lösung:

OLS-geschätzte Nachfragefunktion (= Energiemodell II)

$$\underbrace{\widehat{GASV}_t}_{\hat{y}_t} = 2{,}706 - 0{,}199 \cdot \underbrace{GASP_t}_{x_{1t}} + 0{,}074 \cdot \underbrace{VEINK_t}_{x_{2t}}$$

Regressionswerte für das erste und letzte Jahr:

$$\hat{y}_1 = 2{,}706 - 0{,}199 \cdot (x_{11} = 12{,}9) + 0{,}074 \cdot (x_{21} = 12) = 2{,}706 - 2{,}567 + 0{,}888 = 1{,}027$$

$$\hat{y}_7 = 2{,}706 - 0{,}199 \cdot (x_{17} = 12{,}5) + 0{,}074 \cdot (x_{27} = 15) = 2{,}706 - 2{,}488 + 1{,}110 = 1{,}328$$

Residuen für das erste und letzte Jahr:

$$\hat{u}_1 = y_1 - \hat{y}_1 = 1{,}0 - 1{,}027 = -0{,}027$$

$$\hat{u}_7 = y_7 - \hat{y}_7 = 1{,}4 - 1{,}328 = 0{,}072$$

Die Güte der Anpassung wird über die prozentuale Verhältniszahl

$$\frac{|\hat{u}_t|}{|y_t|} \cdot 100\%$$

ermittelt. Die Verhältniszahl ist bei einer optimalen Anpassung des Datenpunktes null, da dann der y-Wert mit dem Regressionswert $\hat{y}_t$ übereinstimmen würde. Je geringer der Wert, desto besser ist die Anpassung.

Güte der Anpassung bei t = 1:

$$\frac{|\hat{u}_1|}{|y_1|} \cdot 100\% = \frac{|-0{,}027|}{|1{,}0|} \cdot 100\% = 0{,}027 \cdot 100\% = 2{,}7\%$$

Der relative Fehler in der ersten Periode berägt 2,7%.

Güte der Anpassung bei t = 7:

$$\frac{|\hat{u}_7|}{|y_7|} \cdot 100\% = \frac{|0{,}072|}{|1{,}4|} \cdot 100\% = 0{,}051 \cdot 100\% = 5{,}1\%$$

Der relative Fehler in der ersten Periode beträgt 5,1%.

Die Anpassung ist in der ersten Periode deutlich besser als in der letzten Periode, d.h. der Gasverbrauch lässt sich in der ersten Periode erheblich besser durch den Gaspreis und das verfügbare Einkommen erklären als in der letzten Periode.

3.6 Gegeben ist die OLS-geschätzte Nachfragefunktion für Erdgas

$$\hat{GASV}_t = 5{,}809 - 0{,}368 \cdot GASP_t,\ t=1,2,...,7.$$

a) Geben Sie die Quellen der Streuung der Nachfrage nach Erdgas in Form einer ANOVA-Tabelle wieder!

Lösung:

Abhängige Variable (y): Nachfrage nach Erdgas (GASV)
Unabhängige Variable (x): Erdgaspreis (GASP)

Regressionswerte:

t=1: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_1 = 12{,}9) = 1{,}062$

t=2: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_2 = 13{,}8) = 0{,}731$

t=3: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_3 = 13{,}7) = 0{,}767$

t=4: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_4 = 13{,}2) = 0{,}951$

t=5: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_5 = 12{,}2) = 1{,}319$

t=6: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_6 = 12{,}9) = 1{,}062$

t=7: $\hat{y}_1 = 5{,}809 - 0{,}368 \cdot (x_7 = 12{,}5) = 1{,}209$

Arithmetisches Mittel der abhängigen Variablen:

$$\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t = \frac{1}{7} 7{,}1 = 1{,}014$$

Arbeitstabelle:

t	x_t	y_t	$\hat{y}_t$	$\hat{u}_t = y_t - \hat{y}_t$	$(y_t - \bar{y})^2$	$(\hat{y}_t - \bar{y})^2$	$\hat{u}_t^2$
1	12,9	1,0	1,062	-0,062	$(-0{,}014)^2 = 0{,}000$	$0{,}048^2 = 0{,}002$	0,004
2	13,8	0,8	0,731	0,069	$(-0{,}214)^2 = 0{,}046$	$(-0{,}283)^2 = 0{,}080$	0,005
3	13,7	0,7	0,767	-0,067	$(-0{,}314)^2 = 0{,}099$	$(-0{,}247)^2 = 0{,}061$	0,004
4	13,2	0,9	0,951	-0,051	$(-0{,}114^2 = 0{,}013$	$(-0{,}063)^2 = 0{,}004$	0,003
5	12,2	1,2	1,319	-0,119	$0{,}186^2 = 0{,}035$	$0{,}305^2 = 0{,}093$	0,014
6	12,9	1,1	1,062	0,038	$0{,}086^2 = 0{,}007$	$0{,}048^2 = 0{,}002$	0,001
7	12,5	1,4	1,209	0,191	$0{,}386^2 = 0{,}149$	$0{,}195^2 = 0{,}038$	0,036
Σ	91,2	7,1	7,101	-0,001	0,349	0,280	0,067

ANOVA-Tabelle:

Quelle der Streuung	Summe der Abweichungsquadrate	Freiheits-grade	Mittlere Abweichungsquadratsumme
Regression (erklärt)	$\sum(\hat{y}_t - \bar{y})^2 = 0{,}280$	$k-1 = 2-1 = 1$	$\sum(\hat{y}_t - \bar{y})^2 / (k-1) = 0{,}280/1 = 0{,}280$
Residuen (unerklärt)	$\sum \hat{u}_t^2 = 0{,}067$	$n-k = 7-2 = 5$	$\sum \hat{u}_t^2 / (n-k) = 0{,}067/5 = 0{,}013$
gesamt	$\sum(y_t - \bar{y})^2 = 0{,}349$	$n-1 = 7-1 = 6$	

b) Testen Sie den Gesamtzusammenhang des Nachfragemodells unter Verwendung der in der ANOVA-Tabelle ermittelten Abweichungsquadratsummen (α=0,05)!

Lösung:

→ F-Test über den Gesamtzusammenhang

1. Schritt	Formulierung der Nullhypothese H_0: $\beta_2 = 0$, H_1: $\beta_2 \neq 0$
2. Schritt	Festlegung des Signifikanzniveaus $\alpha = 0{,}05$
3. Schritt	Berechnen der Prüfgröße: $F = \frac{\sum(\hat{y}_t - \bar{y})^2 /(k-1)}{\sum \hat{u}_t^2 / (n-k)} = \frac{0{,}280/(2-1)}{0{,}067/(7-2)} = \frac{0{,}280}{0{,}013} = 21{,}538$
4. Schritt	Kritischer Wert: $F_{k-1;n-k;1-\alpha} = F_{2-1;7-2;1-0{,}01} = F_{1;5;0{,}95} = 6{,}61$
5. Schritt	Testentscheidung: F=21,538 > $F_{1;5;0{,}95} = 6{,}61$ → Nullhypothese ablehnen

c) Testen Sie den Gesamtzusammenhang unter Verwendung des Bestimmtheitsmaßes (α=0,05)!

Lösung:

→ F-Test über den Gesamtzusammenhang

1. Schritt	Formulierung der Nullhypothese H_0: $\beta_2 = 0$, H_1: $\beta_2 \neq 0$
2. Schritt	Festlegung des Signikanzniveaus $\alpha = 0{,}05$
3. Schritt	Berechnen der Prüfgröße $R^2 = \frac{\sum(\hat{y}_t - \bar{y})^2}{\sum(y_t - \bar{y})^2} = \frac{0{,}280}{0{,}349} = 0{,}802$ $F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0{,}802/(2-1)}{(1-0{,}802)/(7-2)} = \frac{0{,}802}{0{,}0396} = 20{,}253$
4. Schritt	Kritischer Wert: $F_{k-1;n-k;1-\alpha} = F_{2-1;7-2;1-0{,}01} = F_{1;5;0{,}95} = 6{,}61$
5. Schritt	Testentscheidung: F=20,253 > $F_{1;5;0{,}95} = 6{,}61$ → Nullhypothese ablehnen

3.7 Unter Verwendung der Zeitreihendaten über die Investitionen (I), den Nominalzins (i) und die Inflationsrate (π)

t	I_t	i_t	π_t
1	100	6	2
2	108	5,8	2,2
3	122	5	2,6
4	128	4,8	3,2
5	134	4	3,5

soll die Investitionsfunktion

$$I_t = \beta_1 + \beta_2 \cdot i_t + \beta_3 \cdot \pi_t + u_t$$

ökonometrisch geschätzt und getestet werden. Die Störvariable u_t erfüllt hierbei die Standardannahmen.

a) Schätzen Sie die Investitionsfunktion mit der gewöhnlichen Methode der kleinsten Quadrate unter Verwendung des Matrizenkalküls!

Hinweis: Die Inverse $(\mathbf{X'X})^{-1}$ lautet:

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 423{,}482 & -49{,}786 & -62{,}361 \\ -49{,}786 & 5{,}918 & 7{,}217 \\ -62{,}361 & 7{,}217 & 9{,}411 \end{bmatrix}.$$

Lösung:

Gesucht ist der OLS-Schätzvektor $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$. Zur Berechnung des Produkts $\mathbf{X'y}$ ist die Beobachtungsmatrix $\mathbf{X}$,

$$\mathbf{X} = \begin{bmatrix} 1 & 6 & 2 \\ 1 & 5{,}8 & 2{,}2 \\ 1 & 5 & 2{,}6 \\ 1 & 4{,}8 & 3{,}2 \\ 1 & 4 & 3{,}5 \end{bmatrix},$$

zu transponieren:

$$\mathbf{X'} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 6 & 5{,}8 & 5 & 4{,}8 & 4 \\ 2 & 2{,}2 & 2{,}6 & 3{,}2 & 3{,}5 \end{bmatrix}.$$

Mit dem Vektor der abhängigen Variablen,

$$\mathbf{y} = \begin{bmatrix} 100 \\ 108 \\ 122 \\ 128 \\ 134 \end{bmatrix}$$

erhält man

$$\mathbf{X'y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 6 & 5{,}8 & 5 & 4{,}8 & 4 \\ 2 & 2{,}2 & 2{,}6 & 3{,}2 & 3{,}5 \end{bmatrix} \cdot \begin{bmatrix} 100 \\ 108 \\ 122 \\ 128 \\ 134 \end{bmatrix} = \begin{bmatrix} 592 \\ 2986{,}8 \\ 1633{,}4 \end{bmatrix}.$$

Der Vektor der geschätzten Regressionskoeffizienten, $\hat{\boldsymbol{\beta}}$, lautet damit

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y} = \begin{bmatrix} 423{,}482 & -49{,}786 & -62{,}361 \\ -49{,}786 & 5{,}918 & 7{,}217 \\ -62{,}361 & 7{,}217 & 9{,}411 \end{bmatrix} \cdot \begin{bmatrix} 592 \\ 2986{,}8 \\ 1633{,}4 \end{bmatrix} = \begin{bmatrix} \hat{\beta}_1 = 138{,}299 \\ \hat{\beta}_2 = -9{,}215 \\ \hat{\beta}_3 = 10{,}104 \end{bmatrix}.$$

b) In der makroökonomischen Theorie wird unterstellt, dass die Investitionen I allein vom Realzins i-π abhängig sind und von der Inflationsrate π_t sonst keine weiteren Effekte ausgehen.

Geben Sie die hinter dieser Aussage stehende lineare Restriktion in skalarer Form und Matrixnotation an!

Lösung:

- Lineare Restriktion in skalarer Form:

Realzins: i_t-π_t, Effekt auf Investitionen I_t: $\beta_2 \cdot (i_t - \pi_t)$

Investitionsfunktion mit Realzins:

$$I_t = \beta_1 + \beta_2 \cdot (i_t - \pi_t) + u_t = \beta_1 + \beta_2 \cdot i_t - \beta_2 \cdot \pi_t + u_t$$
$$\Rightarrow \beta_3 = -\beta_2 \quad \text{oder} \quad \beta_2 + \beta_3 = 0 \quad \text{(lineare Restriktion)}$$

- Lineare Restriktion in Matrixnotation:

$$\mathbf{r'}\boldsymbol{\beta} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 0$$

c) Testen Sie die in Teil b) angegebene lineare Restriktion mit dem t-Test (α=0,05) unter Verwendung des erwartungstreuen Schätzers $\hat{\sigma}^2$ der Störvarianz von 16,950! Interpretieren Sie das Testergebnis!

Lösung:

1. Schritt	**Hypothesenformulierung** Nullhypothese H_0: $\beta_2 + \beta_3 = 0$ Alternativhypothese H_1: $\beta_2 + \beta_3 \neq 0$
2. Schritt	**Festlegung des Signifikanzniveaus** Signifikanzniveau $\alpha = 0{,}05$
3. Schritt	**Berechnen der Prüfgröße des t-Tests** Prüfgröße: $t = \dfrac{\mathbf{r}'\hat{\boldsymbol{\beta}} - c}{\hat{\sigma}\sqrt{\mathbf{r}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{r}}}$ mit c=0 $\hat{\sigma}^2 = 16{,}950 \quad \rightarrow \quad \hat{\sigma} = 4{,}117$ $\mathbf{r}'\hat{\boldsymbol{\beta}} = [0\ 1\ 1]\cdot\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{bmatrix} = [0\ 1\ 1]\cdot\underbrace{\begin{bmatrix}138{,}299\\ -9{,}215\\ 10{,}104\end{bmatrix}}_{\text{s. Teil a)}} = 0{,}889$ $\mathbf{r}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{r} = [0\ \ 1\ \ 1]\cdot\underbrace{\begin{bmatrix}423{,}482 & -49{,}786 & -62{,}361\\ -49{,}786 & 5{,}918 & 7{,}217\\ -62{,}361 & 7{,}217 & 9{,}411\end{bmatrix}}_{\text{s. Teil a)}}\cdot\begin{bmatrix}0\\1\\1\end{bmatrix} = [-112{,}147\ \ 13{,}135\ \ 16{,}628]\cdot\begin{bmatrix}0\\1\\1\end{bmatrix} = 29{,}763$ Wert der Prüfgröße: $t = \dfrac{0{,}889 - 0}{4{,}117\cdot\sqrt{29{,}763}} = \dfrac{0{,}889}{4{,}117\cdot 5{,}456} = \dfrac{0{,}889}{22{,}462} = 0{,}040$
4. Schritt	**Kritischer Wert** 0,975-Quantil der t-Verteilung mit n-k = 5-3 = 2 Freiheitsgraden: $t_{2;0{,}975} = 4{,}303$
5. Schritt	**Testentscheidung** $\lvert t \rvert = 0{,}040 < t_{2;0{,}975} = 4{,}303 \quad \rightarrow$ Nullhypothese annehmen <u>Interpretation:</u> Die Annahme, dass allein der Realzins r-π Einfluss die Investitionen und die Inflationsrate π keine darüber hinausgehenden Effekte auf sie hat, kann nicht widerlegt werden.

d) Testen Sie die in Teil b) angegebene lineare Restriktion mit dem F-Test (α=0,05)!

Lösung:

1. Schritt	**Hypothesenformulierung** Nullhypothese H_0: $\beta_2 + \beta_3 = 0$ Alternativhypothese H_1: $\beta_2 + \beta_3 \neq 0$
2. Schritt	**Festlegung des Signifikanzniveaus** Signifikanzniveau $\alpha = 0{,}05$
3. Schritt	**Berechnen der Prüfgröße des t-Tests** Prüfgröße: $F = \dfrac{(\mathbf{r}'\hat{\boldsymbol{\beta}} - c)^2}{\hat{\sigma}^2 \cdot \mathbf{r}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{r}}$ mit c=0 Wert der Prüfgröße (Werte des Zählers und Nenners s. Teil c)): $F = \dfrac{(0{,}889 - 0)^2}{16{,}950 \cdot 29{,}763} = \dfrac{0{,}790}{504{,}483} = 0{,}00156$
4. Schritt	**Kritischer Wert** 0,95-Quantil der F-Verteilung mit 1 und 5-3 = 2 Freiheitsgraden: $F_{1;2;0,95} = 18{,}5$
5. Schritt	**Testentscheidung** $F = 0{,}00156 < F_{1;2;0,975} = 18{,}5 \rightarrow$ Nullhypothese annehmen

3.8 Die OLS-Schätzung der Nachfrage nach Erdgas (GASV) in Abhängigkeit vom Erdgaspreis (GASPR), dem Preis für Fernwärme (FERNWPR) und dem verfügbaren Einkommen (VEINKR) (Energiemodell III) lautet wie folgt:

Dependent Variable: GASV
Method: Least Squares
Sample: 1980 1995
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.670524	5.240572	-1.654500	0.1239
GASPR	-4.510243	2.910106	-1.549855	0.1471
FERNWPR	16.21776	4.035146	4.019125	0.0017
VEINKR	0.006006	0.001900	3.160758	0.0082

R-squared	0.795878	Mean dependent var	12.65000
Adjusted R-squared	0.744847	S.D. dependent var	1.402379
S.E. of regression	0.708379	Akaike info criterion	2.360643
Sum squared resid	6.021608	Schwarz criterion	2.553790
Log likelihood	-14.88514	F-statistic	15.59609
Durbin-Watson stat	1.461440	Prob(F-statistic)	0.000193

a) Welche Informationen über die Gesamtanpassung der Nachfrage nach Erdgas lassen sich der Output-Tabelle entnehmen?

Lösung:

- Bestimmtheitsmaß (Determinationskoeffizient)

$R^2 = 0{,}796$

79,6% der Streuung des Gasverbrauchs wird durch den Gaspreis, den Fern-wärmepreis und das verfügbare Einkommen erklärt.

- Korrigiertes Bestimmtheitsmaß

$\overline{R}^2 = 0{,}745$

Bestimmtheit unter Berücksichtigung der Freiheitsgrade

Anwendbar beim Vergleich der Güte der Anpassung unterschiedlicher Modelle

- F-Test über den Gesamtzusammenhang

H_0: $\beta_2 = \beta_3 = \beta_4 = 0$

Signifikanzniveau α=0,01

$F=15{,}596 > F_{k-1;n-k;0,99} = F_{3;12;0,99} = 5{,}95$ → H_0 ablehnen

oder

$p = 0{,}000193 < \alpha = 0{,}01$ → H_0 ablehnen

b) Zeigen Sie die Beziehung zwischen „S.E. of regression“ und „Sum squared resid“ auf!

Lösung:

S.E. of regression (standard error of regression	Sum squared resid
$SER = \sqrt{\frac{1}{n-k}\sum_{t=1}^{16}\hat{u}_t^2} == 0{,}708379$	$SSR = \sum_{t=1}^{16}\hat{u}_t^2 = 6{,}021608$

$$\begin{aligned} SSR &= SER^2 \cdot (n-k) \\ &= 0{,}708379^2 \cdot (16-4) \\ &= 0{,}501801 \cdot 12 = 6{,}021612 \end{aligned}$$

c) Bestimmen Sie 99%-Konfidenzintervalle für die Regressionskoeffizienten der erklärenden Variablen (ohne Scheinvariable)! Welche Aussagen über die Signifikanz der geschätzten Regressionskoeffizienten können Sie daraus machen?

Lösung:

99%-Konfidenzintervall für β_j:

$$P\left(\hat{\beta}_j - t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_j} \le \beta_j \le \hat{\beta}_j + t_{n-k;1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_j}\right) = 1-\alpha$$

1-α = 0,99 → 99,5%-Quantil der t-Verteilung mit n-k=16-4=12 Freiheitsgraden:

$t_{n-k;0,995} = t_{12;0,995} = 3{,}055$

- Konkretes 99%-Konfidenzintervall für β_2 (Regressionskoeff. des Gaspreises)

$$\left[\hat{\beta}_2 - t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_2};\ \hat{\beta}_2 + t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_2}\right] =$$

$$[-4{,}510 - 3{,}055 \cdot 2{,}910;\ -4{,}510 + 3{,}055 \cdot 2{,}910] =$$

$$[-4{,}510 - 8{,}890;\ -4{,}510 + 8{,}890] = [-13{,}400;\ 4{,}380]$$

Keine Signifikanz bei α=0,01, da das Konfidenzintervall den Wert 0 überdeckt

- Konkretes 99%-Konfidenzintervall für β_3 (Regressionskoeff. des Fernwärmepreises)

$$\left[\hat{\beta}_3 - t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_3};\ \hat{\beta}_3 + t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_3}\right] =$$

$$[16{,}218 - 3{,}055 \cdot 4{,}035;\ 16{,}218 + 3{,}055 \cdot 4{,}035]$$

$$[16{,}218 - 12{,}327;\ 16{,}218 + 12{,}327] = [3{,}891;\ 28{,}545]$$

Signifikanz bei α=0,01, da das Konfidenzintervall den Wert 0 nicht überdeckt

- Konkretes 99%-Konfidenzintervall für β_4 (Regressionskoeff. des verfügbaren Einkommnes)

$$\left[\hat{\beta}_4 - t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_4};\ \hat{\beta}_4 + t_{12;0,995} \cdot \hat{\sigma}_{\hat{\beta}_4}\right] =$$

$$[0{,}0060 - 3{,}055 \cdot 0{,}0019;\ 0{,}0060 + 3{,}055 \cdot 0{,}0019]$$

$$[0{,}0060 - 0{,}0058;\ 0{,}0060 + 0{,}0058] = [0{,}0002;\ 0{,}0118]$$

Signifikanz bei α=0,01, da das Konfidenzintervall den Wert 0 nicht überdeckt