
Ökonometrie – eine Einführung

Ludwig von Auer (Universität Trier)

Stand: April 2016

1 Einleitung

- Braucht man Ökonometriker?
- Was ist Ökonometrie?
- Die vier Aufgaben der Ökonometrie
 - Spezifikation
 - Schätzung
 - Hypothesentest
 - Prognose
- Aufbau der Lehrveranstaltung
- Datenmaterial

2 Spezifikation

- A-Annahmen
 - Erster Schritt: Formulierung eines plausiblen linearen Modells
 - Zweiter und dritter Schritt: Beobachtungsindex und Störgröße
 - Formulierung der A-Annahmen
- Statistisches Repetitorium I
 - Zufallsvariable und Wahrscheinlichkeitsverteilung
 - Erwartungswert einer Zufallsvariable
 - Varianz einer Zufallsvariable
 - Bedingte und gemeinsame Wahrscheinlichkeitsverteilung
 - Kovarianz zweier Zufallsvariablen
 - Rechenregeln für Erwartungswert, Varianz und Kovarianz
 - Die Normalverteilung
- B-Annahmen
 - Begründungen für die Existenz der Störgröße
 - Störgrößen wiederholter Stichproben
 - Formulierung der B-Annahmen

- Statistisches Repetitorium II
 - Stichproben-Mittelwert einer Variable
 - Stichproben-Varianz einer Variable
 - Stichproben-Kovarianz einer Variable
- C-Annahmen

3 Schätzung I: Punktschätzung

- KQ-Methode – eine Illustration
- KQ-Methode – eine algebraische Formulierung
 - Summe der Residuenquadrate
 - Herleitung der Schätzformeln
- Interpretation der KQ-Schätzer
- Bestimmtheitsmaß
 - Grafische Veranschaulichung
 - Definition des Bestimmtheitsmaßes
 - Berechnung des Bestimmtheitsmaßes

- 4 Indikatoren für die Qualität von Schätzverfahren
 - Statistischer Hintergrund
 - Warum ist y_t eine Zufallsvariable?
 - Warum sind die KQ-Schätzer Zufallsvariablen?
 - Zwei Kriterien: Unverzerrtheit und Effizienz
 - Unverzerrtheit und Effizienz der KQ-Methode
 - Statistisches Repetitorium III
 - Standard-Normalverteilung?
 - Chi-Quadrat-Verteilung?
 - t-Verteilung?
 - F-Verteilung
 - Wahrscheinlichkeitsverteilungen der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$
 - Wahrscheinlichkeitsverteilung von \mathbf{y}_t
 - Wahrscheinlichkeitsverteilungen von $\hat{\alpha}$ und $\hat{\beta}$

5 Schätzung II: Intervallschätzer

- Intervallschätzer und ihre Interpretation
- Intervallschätzer für β bei bekanntem σ^2
- Intervallschätzer für β bei unbekanntem σ^2
 - Herleitung des Intervallschätzers
 - Interpretation des Intervallschätzers
 - Aussagekraft von Intervallschätzern
- Intervallschätzer für α

6 Hypothesentest

- Zweiseitiger Hypothesentest
 - Ein grafisches Entscheidungsverfahren
 - Ein analytisches Entscheidungsverfahren
 - Zusammenhang zwischen analytischem und grafischem Vorgehen
 - Zusammenhang zwischen zweiseitigem Hypothesentest und Intervallschätzer
- Einseitiger Hypothesentest
 - Ein grafisches Entscheidungsverfahren
 - Ein analytisches Entscheidungsverfahren
- p -Wert
- Wahl der geeigneten Nullhypothese und des geeigneten Signifikanzniveaus
 - Strategie A: Nullhypothese behauptet Gegenteil der Anfangsvermutung
 - Strategie B: Nullhypothese stimmt mit Anfangsvermutung überein

- Trennschärfe von Tests
- Anmerkungen zu zweiseitigen Tests

7 Prognose

- Punktprognose
 - Berechnung der Punktprognose
 - Verlässlichkeit der Punktprognose
- Prognoseintervall

8 Spezifikation

- A-Annahmen
 - Erster Schritt
 - Zweiter und dritter Schritt
 - Formulierung der A-Annahmen
- B-Annahmen
 - Formulierung der B-Annahmen
 - Interpretation der B-Annahmen
- C-Annahmen

9 Schätzung

- Punktschätzer $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$
- Interpretation der Schätzer $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$
 - Formale Interpretation
 - Ökonomische Interpretation
- Autonome Variation der exogenen Variablen
 - Korrelation zwischen den exogenen Variablen
 - Berechnung der autonomen Variation
- Informationsverarbeitung der KQ-Methode und Bestimmtheitsmaß R^2
 - Definition des Bestimmtheitsmaßes
 - Berechnung des Bestimmtheitsmaßes
 - Bestimmtheitsmaß und Venn-Diagramme
 - KQ-Methode als zweistufiger Prozess
 - Partielles Bestimmtheitsmaß
- Unverzerrtheit und Effizienz der KQ-Methode
 - Erwartungswert und Varianz der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$

- Interpretation der Formeln
- Schätzformeln für $var(\hat{\alpha})$, $var(\hat{\beta}_k)$ und $var(\hat{\beta}_1, \hat{\beta}_2)$
- BLUE- bzw. BUE-Eigenschaft
der KQ-Schätzer
- Wahrscheinlichkeitsverteilungen der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$
 - Wahrscheinlichkeitsverteilung der y_t
 - Wahrscheinlichkeitsverteilungen der Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$
- Intervallschätzer

10 Hypothesentest

- Testen einer Linearkombination von Parametern: t -Test
 - Zweiseitiger t -Test
 - Einseitiger t -Test
- Simultaner Test mehrerer Linearkombinationen von Parametern: F -Test
 - Eine wichtige Nullhypothese
 - Test einer allgemeinen Nullhypothese
- Zusammenhang zwischen t -Test und F -Test bei $L=1$
 - Zweiseitiger F -Test einer einzelnen Linearkombination
 - Probleme des F -Tests bei einseitigen Hypothesen
- Zusammenhang zwischen t -Test und F -Test bei $L=2$
 - Numerisches Beispiel
 - Unterschied zwischen individuellen und simultanen Parametertests

11 Prognose

- Punktprognose
 - Prognosewert und Prognosefehler
 - Verlässlichkeit der Punktprognose
- Prognoseintervall

- 12 Präsentation der Schätzergebnisse und deren computergestützte Berechnung
 - Computergestützte ökonometrische Analyse
 - Ökonometrische Software
 - Interpretation des Computeroutputs
 - Präsentation von Schätzergebnissen

13 Annahme A1: Variablenauswahl

- Konsequenzen der Annahmeverletzung
 - Auslassen relevanter Variablen
 - Verwendung irrelevanter Variablen
- Diagnose und Neu-Spezifikation
 - Korrigiertes Bestimmtheitsmaß \bar{R}^2
 - Weitere Kennzahlen: AIC, SC und PC
 - F -Test
 - t -Test
 - Zusammenhang zwischen korrigiertem Bestimmtheitsmaß, F -Test und t -Test
 - Ungenesteter F -Test
- Spezifikations-Methodologien
 - Steinmetz- versus Maurer-Methodologie
 - Wichtiges Problem bei der Variablenauswahl

14 Annahme A2: Funktionale Form

- Konsequenzen der Annahmeverletzung
- Einige alternative Funktionsformen
 - Semi-logarithmisches Modell (Linlog-Modell)
 - Inverses Modell
 - Exponential-Modell (Loglin-Modell)
 - Logarithmisches Modell (Loglog-Modell)
 - Log-inverses Modell
 - Quadratisches Modell
 - Eine vergleichende Anwendung
- Diagnose und Neu-Spezifikation
 - Regression Specification Error Test (RESET)
 - Bestimmtheitsmaß R^2
 - Box-Cox-Test

15 Annahme A3: Konstante Parameterwerte

- Konsequenzen der Annahmeverletzung
 - Ein geeignetes Strukturbruchmodell
 - Schätzung und Interpretation der Parameter des Strukturbruchmodells
 - Getrennte Schätzung der zwei Phasen
 - Eine mögliche alternative Formulierung des Strukturbruchmodells
 - Komplexere Strukturbrüche
 - Konsequenzen aus einer Vernachlässigung des Strukturbruchs
- Diagnose
 - F -Test
 - t -Test
 - Prognostischer Chow-Test
 - Unbekannter Zeitpunkt des Strukturbruchs
- Stetige Veränderung von Parameterwerten
- Exkurs: Qualitative exogene Variablen

- Einführung einer Dummy-Variable
- Ein allgemeines Dummy-Variablen-Modell

16 Annahme B1: Erwartungswert der Störgröße

- Konsequenzen der Annahmeverletzung
 - Konstanter Messfehler bei der Erfassung der endogenen Variable
 - Konstanter Messfehler bei der Erfassung einer exogenen Variable
 - Funktionale Modelltransformation
 - Gestutzte endogene Variable
- Diagnose
 - Überprüfung der Datenerhebung
 - Überprüfung auf Basis der Daten
- Anwendbare Schätzverfahren

17 Annahme B2: Homoskedastizität

- Konsequenzen der Annahmeverletzung
 - Konsequenzen für die Punktschätzung
 - Konsequenzen für Intervallschätzung und Hypothesentest
- Diagnose
 - Grundidee der Tests auf Heteroskedastizität
 - Goldfeld-Quandt-Test
 - Breusch-Pagan-Test
 - White-Test
- Anwendbare Schätzverfahren
 - VKQ-Methode
 - GVKQ-Methode
 - KQ-Methode mit Whites HK-Schätzer

18 Annahme B3: Freiheit von Autokorrelation

- Konsequenzen der Annahmeverletzung
 - Erwartungswert von u_t
 - Varianz von u_t
 - Kovarianz von u_t und u_{t-1}
 - Konsequenzen für die Punktschätzung
 - Konsequenzen für Intervallschätzung und Hypothesentest
- Diagnose
 - Grafische Analyse
 - Schätzer für ρ
 - Durbin-Watson-Test
 - Durbin-Watson-Test für negative Autokorrelation
 - Nachteile des Durbin-Watson-Tests
- Anwendbare Schätzverfahren
 - Ermittlung von x_1^* und y_1^*
 - VKQ-Methode von Hildreth und Lu
 - GVKQ-Methode von Cochrane und Orcutt

19 Annahme B4: Normalverteilte Störgrößen

- Konsequenzen der Annahmeverletzung
- Diagnose
 - Grafische Analyse
 - Jarque-Bera-Test

- 20 Annahme C1: Zufallsunabhängige exogene Variablen
- Weitere Qualitätskriterien für Schätzer: Konsistenz und asymptotische Effizienz
 - Konsistenz
 - Rechenregeln für Wahrscheinlichkeitsgrenzwerte
 - Asymptotische Effizienz
 - Konsequenzen der Annahmeverletzung
 - Fall 1: Störgrößen und Beobachtungen der exogenen Variable unabhängig
 - Fall 2: Störgrößen und Beobachtungen der exogenen Variable kontemporär unkorreliert
 - Eine mögliche Ursache für Fall 2: y_{t-1} als „exogene Variable“
 - Fall 3: Störgrößen und Beobachtungen der exogenen Variable kontemporär korreliert
 - Eine mögliche Ursache für Fall 3: Probleme bei der Erfassung der exogenen Variable
 - Anwendbare Schätzverfahren

- IV-Schätzung mit der ZSKQ-Methode
- Auswahl der Instrumentvariablen
- ZSKQ-Schätzung in der multiplen Regression
- Konsistenz der ZSKQ-Schätzer
- Wahrscheinlichkeitsverteilung und Varianz der ZSKQ-Schätzer
- Fazit der ZSKQ-Schätzung
- Diagnose
 - Vorüberlegungen
 - Spezifikationstest von Hausman

21 Annahme C2: Multikollinearität

- Konsequenzen der Annahmeverletzung
 - Grafische Veranschaulichung
 - Konsequenzen perfekter Multikollinearität für Punkt-, Intervallschätzung und Hypothesentests
 - Konsequenzen imperfekter Multikollinearität für Punkt-, Intervallschätzung und Hypothesentests
- Diagnose
 - Diagnose von Multikollinearität
 - Hohe Schätzvarianz der Punktschätzer: Multikollinearität oder Fehlspezifikation?
- Angemessener Umgang mit Multikollinearität
 - Verfahren zur Eindämmung des Multikollinearitätsproblems
 - Verwendung zusätzlicher Informationen

22 Dynamische Modelle

- Stochastische Prozesse und Stationarität
 - Stochastische Prozesse
 - Stationarität von stochastischen Prozessen
 - $I(1)$ -Prozesse
- Interpretation dynamischer Modelle
 - Interpretation einzelner Parameter
 - Kurzfristiger und langfristiger Multiplikator
 - Median-Lag
- Allgemeine Schätzprobleme dynamischer Modelle
 - Zwei zentrale Schätzprobleme
 - Mögliche Lösungsstrategien
- Modelle mit geometrischer Lag-Verteilung
 - Geometrische Lag-Verteilungen
 - Koyck-Modell
 - Ein Verwandter des Koyck-Modells: Partielles Anpassungsmodell
 - Ein weiterer Verwandter des Koyck-Modells: Modell adaptiver

Erwartungen

- Modelle mit rationaler Lag-Verteilung und ihre Fehlerkorrektur-Formulierung
 - Langfristige Gleichgewichtsbeziehung
 - Fehlerkorrektur-Formulierung des ADL(1,1)-Modells
 - Schätzung des Fehlerkorrekturmodells
 - Fehlerkorrekturmodell und ökonomische Theorie

23 Interdependente Gleichungssysteme

- Nicht-Konsistenz der KQ-Schätzer
- Indirekte KQ-Methode (IKQ-Methode)
 - Strukturelle Form versus reduzierte Form
 - Schätzung der Parameter der reduzierten Form
 - Schätzung der Parameter der strukturellen Form
- Identifikationsproblem
 - Ein verkleinertes Gleichungssystem
 - Ein erweitertes Gleichungssystem
 - Ordnungskriterium
- Zweistufige KQ-Methode (ZSKQ-Methode)
 - ZSKQ-Schätzung mit Hilfe der reduzierten Form
 - ZSKQ-Schätzung im Überblick
- Weitere Beispiele interdependenter Gleichungssysteme
 - Gleichungssysteme mit Lag-Variablen
 - Keynesianisches Makromodell
 - Partielles Marktgleichgewichtsmodell

Kapitel 1

Einleitung

1.1 Braucht man Ökonometriker

- Die Menge an verfügbaren Daten hat kontinuierlich zugenommen.
- Die Auswertung solcher Daten erfordert geeignete empirische Methoden.
- Für wirtschaftswissenschaftliche Fragen wird dabei vor allem die Ökonometrie benötigt.

1.2 Was ist Ökonometrie?

- Ökonomische Theorien werden der empirischen Realität gegenübergestellt.
- Es existieren zwei Richtungen der empirischen Forschung:
 - experimentelle Empirie
 - historische Empirie

Definition: Ökonometrie

Die Ökonometrie analysiert anhand von beobachtbaren Daten (ökonomische Realität) ökonomische Wirkungszusammenhänge (ökonomische Theorie). Dabei greift sie auf Methoden zurück, die in der statistischen Theorie entwickelt wurden.

1.3 Die vier Aufgaben der Ökonometrie

Beispiel Trinkgeld

Es soll der numerische Zusammenhang zwischen der Höhe des Rechnungsbetrages x und der Höhe des Trinkgeldes y untersucht werden.

Das *ökonomische Modell* lautet:

$$y = f(x)$$

1.3.1 Spezifikation

- Die Spezifikation umfasst drei Komponenten:

a) Funktionale Spezifikation

$$y = \beta x \quad (1.1)$$

$$y_t = \beta x_t \quad (1.2)$$

$$y_t = \beta x_t + u_t . \quad (1.3)$$

b) Störgrößen-Spezifikation

c) Variablen-Spezifikation

Definition: Ökonometrisches Modell

Das vollständig spezifizierte Modell (1.3) ist das *ökonometrische Modell*.

1.3.2 Schätzung

Numerische Illustration 1.1

Es seien zwei Gäste beobachtet worden. Dabei bezeichnet x_t den Rechnungsbetrag und y_t das Trinkgeld (beides in Euro):

$$\text{Gast 1} \quad : \quad (x_1 = 10, y_1 = 2)$$

$$\text{Gast 2} \quad : \quad (x_2 = 30, y_2 = 3) .$$

- Es wurde im ökonometrischen Modell unterstellt, dass beide Gäste den gleichen Wert β besitzen.

Numerische Illustration 1.2

Frage: Sind bei den beiden Gästen Störeinflüsse wirksam geworden?

Antwort:

Frage: Welche Werte sind für β plausibel?

Antwort:

- Der Schätzwert für β wird durch $\hat{\beta}$ bezeichnet.
- Das *geschätzte Modell* (Variante I) lautet:

$$\hat{y}_t = \hat{\beta}x_t . \quad (1.4)$$

- Die geschätzte Störung (genannt: *Residuum*) lautet:

$$\begin{aligned} \hat{u}_t &= y_t - \hat{y}_t \\ &= y_t - \hat{\beta}x_t . \end{aligned} \quad (1.5)$$

- Umstellen liefert das geschätzte Modell (Variante II):

$$y_t = \hat{\beta}x_t + \hat{u}_t . \quad (1.6)$$

Numerische Illustration 1.3

Für $\hat{\beta} = 0,15$ lautet das geschätzte Modell:

$$\hat{y}_t = 0,15 \cdot x_t \quad (\text{Variante I}) \quad (1.7)$$

$$y_t = 0,15 \cdot x_t + \hat{u}_t \quad (\text{Variante II})$$

Frage: Berechnen Sie die Werte von \hat{y}_1 , \hat{y}_2 , \hat{u}_1 und \hat{u}_2 .

Antwort:

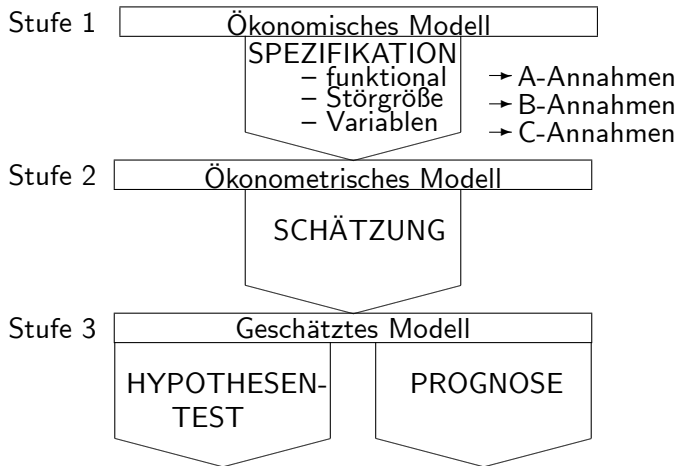


Abbildung 1.1: Die vier Aufgaben ökonometrischer Analyse.

1.3.3 Hypothesentest

- Auf Basis des geschätzten Modells lassen sich verschiedene Hypothesen überprüfen.

Beispiel 1: „Der wahre Wert von β beträgt 0,05“

Beispiel 2: „Das ökonometrische Modell ist nicht korrekt spezifiziert“

1.3.4 Prognose

Numerische Illustration 1.4

Frage: Welches Trinkgeld würde ein Gast mit einem Rechnungsbetrag von $x_0 = 40$ Euro schätzungsweise geben?

Antwort:

1.4 Aufbau der Lehrveranstaltung

Die Annahmen der

funktionalen Spezifikation werden als *A-Annahmen*,

die der

Störgrößen-Spezifikation als *B-Annahmen*

und die der

Variablen-Spezifikation als *C-Annahmen*

bezeichnet.

Der Kurs setzt sich aus vier Teilen zusammen:

- Teil I *einfaches lineares Regressionsmodell*
(Kapitel 2 bis 7).
- Teil II *multiple lineares Regressionsmodell*
(Kapitel 8 bis 12).
- Teil III Probleme, die sich aus den verschiedenen
möglichen Annahmeverletzungen ergeben
(Kapitel 13 bis 21).
- Teil IV zwei weiterführende Bereiche der Ökonometrie
(Kapitel 22 und 23).

1.5 Datenmaterial

Es existieren drei Typen von Daten:

- Zeitreihendaten
- Querschnittsdaten
- Paneldaten.

Tabelle 1.1: Datenpaare (x_t = Rechnungsbetrag, y_t = Trinkgeld; beides in Euro) von 9 beobachteten Gäste.

	Stammgast 1	Stammgast 2	Stammgast 3
1. Abend	$(x_1, y_1) = (10, 2)$	$(x_2, y_2) = (30, 3)$	$(x_3, y_3) = (50, 7)$
2. Abend	$(x_4, y_4) = (20, 2)$	$(x_5, y_5) = (35, 3)$	$(x_6, y_6) = (41, 6)$
3. Abend	$(x_7, y_7) = (25, 4)$	$(x_8, y_8) = (14, 2)$	$(x_9, y_9) = (17, 2)$

Kapitel 2

Spezifikation

Beispiel Trinkgeld (mit 20 Beobachtungen)

Die Daten der 20 Gäste sind in der folgenden Tabelle wiedergegeben.

Tabelle 2.1: Die Daten von 20 beobachteten Gästen.

t	x_t	y_t	t	x_t	y_t
1	10,00	2,00	11	60,00	7,00
2	30,00	3,00	12	47,50	5,50
⋮	⋮	⋮	⋮	⋮	⋮
10	12,50	1,00	20	20,00	2,50

2.1 A-Annahmen

- Mit Hilfe der *A-Annahmen* erfolgt die *funktionale* Spezifikation des ökonometrischen Modells.

2.1.1 Erster Schritt: Formulierung eines plausiblen linearen Modells

- Das ökonomische Modell lautet

$$y = f(x) \tag{2.1}$$

- Wenn $f(x)$ als lineare Funktion spezifiziert wird, ergibt sich

$$y = \alpha + \beta x . \tag{2.2}$$

- **Frage:** Für welche Parameterwerte wird aus dem linearen Zusammenhang zwischen x und y ein proportionaler Zusammenhang?

Antwort:

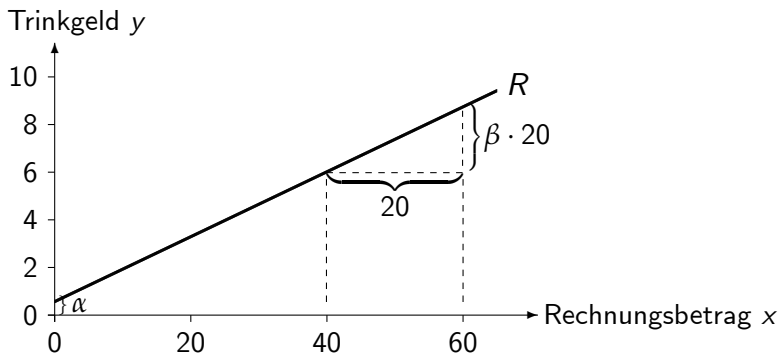


Abbildung 2.1: Der „wahre“ Zusammenhang zwischen Rechnungsbetrag x und Trinkgeld y .

2.1.2 Zweiter und dritter Schritt: Beobachtungsindex und Störgröße

- Die Hinzufügung eines Beobachtungsindex liefert:

$$y_t = \alpha + \beta x_t \quad \text{für } t = 1, 2, \dots, 20 . \quad (2.3)$$

- **Frage:** Wie müsste die Punktwolke in Abbildung 2.2 aussehen, wenn keine Störeinflüsse aufgetreten wären?

Antwort:

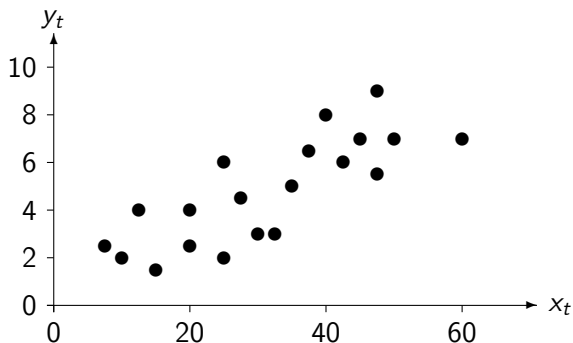


Abbildung 2.2: Die Daten des Trinkgeldbeispiels in grafischer Form.

- Das zum ökonomischen Modell korrespondierende ökonometrische Modell lautet:

$$y_t = \alpha + \beta x_t + u_t \quad \text{für } t = 1, 2, \dots, 20 . \quad (2.4)$$

- Dabei werden die Parameter α und β als *Regressionsparameter* bezeichnet.
- Die Variable u_t ist als eine *Störgröße* definiert.
- Ihre grafische Interpretation ist in Abbildung 2.3 wiedergegeben.

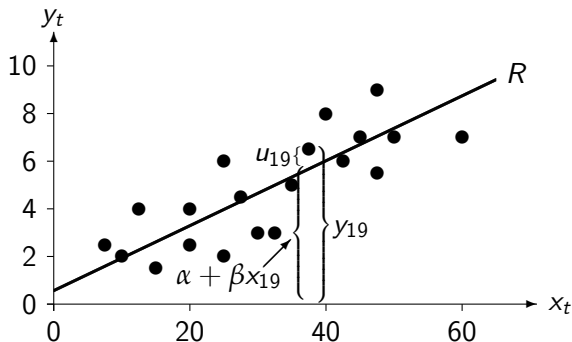


Abbildung 2.3: Der Zusammenhang zwischen beobachtetem Wert y_t , Störgröße u_t und „ungestörtem“ Einfluss $\alpha + \beta x_t$.

2.1.3 Formulierung der A-Annahmen

- Das ökonometrische Modell lautet:

$$y_t = \alpha + \beta x_t + u_t \quad (2.4)$$

Annahme a1

Es fehlen keine relevanten exogenen Variablen und die exogene Variable x_t ist nicht irrelevant.

Annahme a2

Der wahre Zusammenhang zwischen x_t und y_t ist linear.

Annahme a3

Die Parameter α und β sind für alle T Beobachtungen (x_t, y_t) konstant.

- **Frage:** Welche A-Annahmen sind in den Punktwolken der nachfolgenden Abbildungen verletzt?

Antwort:

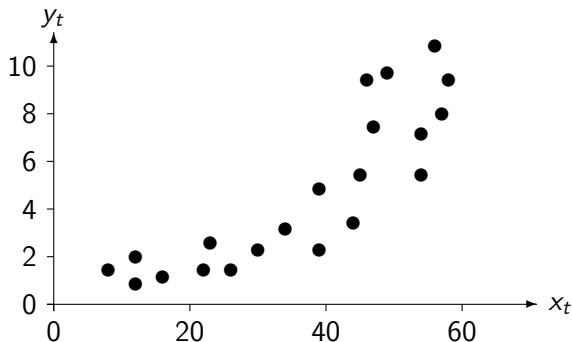


Abbildung 2.4: Eine Punktwolke, die auf eine Verletzung der A-Annahmen hindeutet.

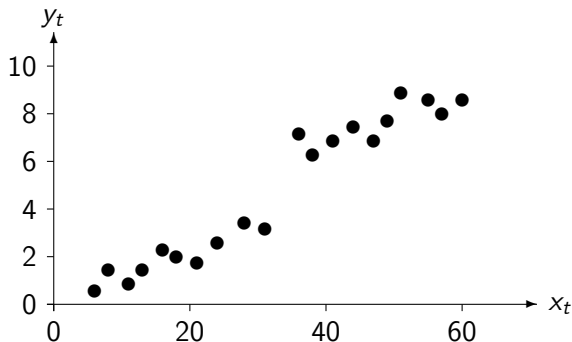


Abbildung 2.5: Eine weitere Punktwolke, die auf eine Verletzung der A-Annahmen hindeutet.

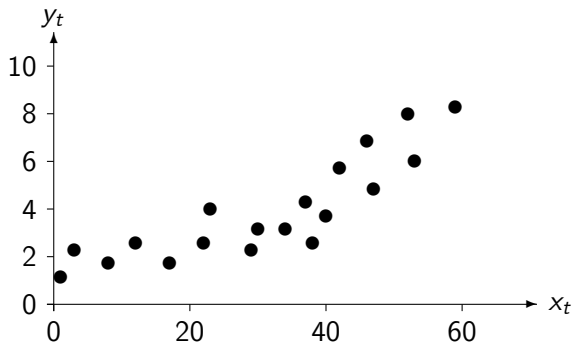


Abbildung 2.6: Eine weitere Punktwolke, die auf eine Verletzung der A-Annahmen hindeutet.

2.2 Statistisches Repetitorium I

- Dieses Repetitorium beschäftigt sich ausschließlich mit dem Thema *Zufallsvariablen*.
- Es werden nur diejenigen Aspekte wiederholt, welche später im Kurs benötigt werden.

2.2.1 Zufallsvariable und Wahrscheinlichkeitsverteilung

- **Frage:** Wie viele mögliche Ausprägungen besitzt die Zufallsvariable:

$u_1 =$ "Geworfene Augenzahl bei einmaligem Würfeln" ?

Antwort:

- **Frage:** Welche Wahrscheinlichkeit besitzt jede einzelne Ausprägung?

Antwort:

- **Frage:** Wie viele mögliche Ausprägungen besitzt die Zufallsvariable:

u_2 = "Summe der geworfenen Augenzahlen
bei zweimaligem Würfeln" ?

Antwort:

- Die Wahrscheinlichkeit für die Ausprägung “2” beträgt bei u_2 :

$$f(2) = (1/6) \cdot (1/6) = 1/36 .$$

- Die Wahrscheinlichkeit für die Ausprägung “3” beträgt

$$f(3) = 2 \cdot (1/6) \cdot (1/6) = 2/36 .$$

- Jeder der möglichen Ausprägungen von u_2 kann eine Wahrscheinlichkeit ihres Auftretens zugeordnet werden.
- Man bezeichnet diese Zuordnung als die *Wahrscheinlichkeitsverteilung* der Zufallsvariable u_2 .

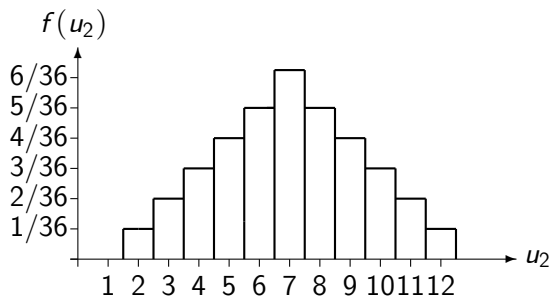


Abbildung 2.7: (Teil a) Die Wahrscheinlichkeitsverteilung der diskreten Zufallsvariable u_2 .

- Man unterscheidet zwischen *diskreten* Zufallsvariablen und *stetigen* Zufallsvariablen.
- Ein weiteres Beispiel für eine diskrete Zufallsvariable ist:

$u_3 =$ „Summe der geworfenen Augenzahlen
bei 100.000 mal Würfeln“

- **Frage:** Wie viele Ausprägungen besitzt diese Zufallsvariable?
Antwort:

- Ein Beispiel für eine stetige Zufallsvariable ist:

$u_4 =$ “eine reelle Zahl aus dem Intervall $[0,1]$ ” .

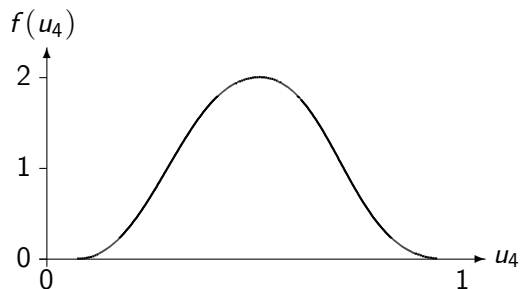


Abbildung 2.7: (Teil b) Eine mögliche Wahrscheinlichkeitsverteilung der stetigen Zufallsvariable u_4 .

2.2.2 Erwartungswert einer Zufallsvariable

- $f(u_i)$ bezeichnet die Wahrscheinlichkeit, mit der die Ausprägung i der Zufallsvariable u beobachtet wird.
- Der *Erwartungswert* der Zufallsvariable u lautet:

$$E(u) = \sum_{i=1}^N f(u_i) \cdot u_i . \quad (2.5)$$

- Wahrscheinlichkeiten müssen sich zu 1 addieren:

$$\sum_{i=1}^N f(u_i) = 1 .$$

- Der Erwartungswert $E(u)$ ist das mit den Wahrscheinlichkeiten $f(u_i)$ gewichtete Mittel aller möglichen Ausprägungen der Zufallsvariable u .

Numerische Illustration 2.1

Für die Zufallsvariable

$u =$ "Geworfene Augenzahl bei einmaligem Würfeln"

gilt $N = 6$.

Frage: Welchen Erwartungswert besitzt diese Zufallsvariable?

Antwort:

2.2.3 Varianz einer Zufallsvariable

- Die *Varianz* einer Zufallsvariable misst die Streuung der Zufallsvariable um ihren Erwartungswert:

$$\text{var}(u) = \sum_{i=1}^N f(u_i) (u_i - E(u))^2 . \quad (2.6)$$

Numerische Illustration 2.2

Da im Würfel Beispiel $E(u) = 3,5$, ergibt sich:

$$\begin{aligned} \text{var}(u) &= \frac{1}{6} \cdot 6,25 + \frac{1}{6} \cdot 2,25 + \frac{1}{6} \cdot 0,25 \\ &\quad + \frac{1}{6} \cdot 0,25 + \frac{1}{6} \cdot 2,25 + \frac{1}{6} \cdot 6,25 \\ &= 2,91666 . \end{aligned}$$

- Die quadrierte Abweichung $(u - E(u))^2$ kann ebenfalls als Zufallsvariable aufgefasst werden, wobei $E(u)$ zufallsunabhängig ist:

$$f(u_j) = f((u_j - E(u))^2) .$$

- Dabei sind im Falle von

$$(u_i - E(u))^2 = (u_j - E(u))^2$$

die beiden Ausprägungen $(u_i - E(u))^2$ und $(u_j - E(u))^2$ dennoch als *verschiedene* Ausprägungen zu interpretieren.

Numerische Illustration 2.3

Im Würfelbeispiel gilt $E(u) = 3,5$ und

$$f(6) = f((6 - 3,5)^2)$$

$$f(1) = f((1 - 3,5)^2) .$$

Obwohl

$$(6 - 3,5)^2 = (1 - 3,5)^2 = 6,25$$

sind $(6 - 3,5)^2$ und $(1 - 3,5)^2$ als unterschiedliche Ausprägungen zu interpretieren.

- Folglich kann die Varianz auch in der Form

$$\text{var}(u) = \sum_{i=1}^N f((u_i - E(u))^2) (u_i - E(u))^2 \quad (2.7)$$

ausgedrückt werden, bzw.

$$\text{var}(u) = E \left[(u - E(u))^2 \right] . \quad (2.8)$$

- Die Wurzel der Varianz bezeichnet man als *Standardabweichung* (engl.: *standard deviation*):

$$\text{sd}(u) = \sqrt{\text{var}(u)} .$$

2.2.4 Bedingte und gemeinsame Wahrscheinlichkeitsverteilung

- Gegeben seien die Zufallsvariablen

u_1 = „Geworfene Augenzahl bei einmaligem Würfeln“

u_6 = „Anzahl der natürlichen Zahlen, durch welche die geworfene Augenzahl teilbar ist“.

Tabelle 2.2: Ausprägungen der Zufallsvariablen u_1 und u_6 im Zufallsexperiment „einmaliges Würfeln“.

Zufallsvariable	Ausprägung					
u_1	1	2	3	4	5	6
u_6	1	2	2	3	2	4

- Für u_1 lässt sich eine *bedingte Wahrscheinlichkeitsverteilung* definieren: $f(u_{1i}|u_{6j})$.
- Sie ordnet bei gegebenem Wert u_{6j} jeder möglichen Ausprägung von u_1 die Wahrscheinlichkeit ihres Auftretens zu.
- Für vorgegebenes $u_{6j} = 2$ ergibt sich die folgende bedingte Wahrscheinlichkeitsverteilung:

$$f(u_{1i}=1|u_{6j}=2) = 0$$

$$f(u_{1i}=2|u_{6j}=2) = 1/3$$

$$f(u_{1i}=3|u_{6j}=2) = 1/3$$

$$f(u_{1i}=4|u_{6j}=2) = 0$$

$$f(u_{1i}=5|u_{6j}=2) = 1/3$$

$$f(u_{1i}=6|u_{6j}=2) = 0$$

- Die *gemeinsame Wahrscheinlichkeitsverteilung* $f(u_{1i}, u_{6j})$ ordnet jeder möglichen *Ausprägungskombination* (u_{1i}, u_{6j}) eine Wahrscheinlichkeit ihres Auftretens zu.
- Es gilt:

$$f(u_{1i}, u_{6j}) = f(u_{1i} | u_{6j}) \cdot f(u_{6j}) = f(u_{6j} | u_{1i}) \cdot f(u_{1i}) . \quad (2.9)$$

Numerische Illustration 2.4

Die Wahrscheinlichkeit für das Auftreten der Ausprägungskombination $u_{1i} = 3$ und $u_{6j} = 2$ ergibt sich gemäß der Gleichung (2.9) aus

$$\begin{aligned} f(u_{1i} = 3, u_{6j} = 2) &= f(u_{1i} = 3 | u_{6j} = 2) \cdot f(u_{6j} = 2) \\ &= 1/3 \cdot 1/2 = 1/6 \end{aligned}$$

oder alternativ aus

$$\begin{aligned} f(u_{1i} = 3, u_{6j} = 2) &= f(u_{6j} = 2 | u_{1i} = 3) \cdot f(u_{1i} = 3) \\ &= 1 \cdot 1/6 = 1/6 . \end{aligned}$$

(Forts.)

Tabelle 2.3: Gemeinsame Wahrscheinlichkeiten der Zufallsvariablen u_1 und u_6 .

		u_{1i}					
		1	2	3	4	5	6
u_{6j}	1	1/6	0	0	0	0	0
	2	0	1/6	1/6	0	1/6	0
	3	0	0	0	1/6	0	0
	4	0	0	0	0	0	1/6

2.2.5 Kovarianz zweier Zufallsvariablen

Unkorreliertheit

- Ein positiver Zusammenhang besteht, wenn tendenziell

$$(u_{1i} - E(u_1)) > 0 \quad \iff \quad (u_{6j} - E(u_6)) > 0$$

$$(u_{1i} - E(u_1)) < 0 \quad \iff \quad (u_{6j} - E(u_6)) < 0 .$$

- Ein negativer Zusammenhang besteht, wenn tendenziell

$$(u_{1i} - E(u_1)) > 0 \quad \iff \quad (u_{6j} - E(u_6)) < 0$$

$$(u_{1i} - E(u_1)) < 0 \quad \iff \quad (u_{6j} - E(u_6)) > 0 .$$

- Die *Kovarianz* der beiden Zufallsvariablen formalisiert diesen Zusammenhang:

$$\text{cov}(u_1, u_6) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_6} f(u_{1i}, u_{6j}) [(u_{1i} - E(u_1))(u_{6j} - E(u_6))] \quad (2.10)$$

$$= E [(u_1 - E(u_1))(u_6 - E(u_6))] . \quad (2.11)$$

- Beträgt die Kovarianz 0, dann üben die beiden Zufallsvariablen keinen linearen Einfluss aufeinander aus. Sie sind dann *unkorreliert*.

Numerische Illustration 2.5

Der Erwartungswert von u_1 besitzt den Wert $E(u_1) = 3,5$. Für u_6 erhält man:

$$E(u_6) = \frac{1}{6} \cdot 1 + \frac{3}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 = 2,333$$

und mit Hilfe von Tabelle 2.3

$$\begin{aligned} \text{cov}(u_1, u_6) &= \frac{1}{6} (-2,5) (-1,333) + \frac{1}{6} (-1,5) (-0,333) \\ &\quad + \frac{1}{6} (-0,5) (-0,333) + \frac{1}{6} (0,5) (0,666) \\ &\quad + \frac{1}{6} (1,5) (-0,333) + \frac{1}{6} (2,5) (1,666) \\ &= 1,333 . \end{aligned}$$

Korrelationskoeffizient

- Gegeben seien die Zufallsvariablen

u_5 = „Körpergröße eines erwachsenen Engländers“ ,

u_7 = „Schuhgröße eines erwachsenen Engländers“ .

- Der Korrelationskoeffizient der Zufallsvariablen u_5 und u_7 lautet:

$$\text{cor}(u_5, u_7) = \frac{\text{cov}(u_5, u_7)}{\text{sd}(u_5) \cdot \text{sd}(u_7)} . \quad (2.12)$$

- Es gilt immer $-1 \leq \text{cor}(u_5, u_7) \leq 1$.

Unabhängigkeit

- Zufallsvariablen, die weder einen linearen noch einen nicht-linearen Einfluss aufeinander ausüben, werden als statistisch *unabhängig* bezeichnet.
- Ein Beispiel:

$$u_1 = \text{„Geworfene Augenzahl bei Würfel 1“}$$

und

$$u_2 = \text{„Geworfene Augenzahl bei Würfel 2“}.$$

- Die Kenntnis der tatsächlich beobachteten Ausprägung der ersten Zufallsvariable gibt keinerlei zusätzliche Informationen über die Ausprägung der anderen Zufallsvariable:

$$f(u_{1i}|u_{2j}) = f(u_{1i}) . \quad (2.13)$$

- Es ergibt sich aus (2.9) und (2.13):

$$f(u_{1i}, u_{2j}) = f(u_{1i}|u_{2j}) \cdot f(u_{2j}) = f(u_{1i}) \cdot f(u_{2j}) . \quad (2.14)$$

2.2.6 Rechenregeln für Erwartungswert, Varianz und Kovarianz

Erwartungswert

- Es seien u_1 und u_2 zwei Zufallsvariablen und x_1 und x_2 zwei Konstanten. Dann gilt:

$$E(x_1) = x_1 \quad (2.15)$$

$$E(x_1 \cdot u_1) = x_1 \cdot E(u_1) \quad (2.16)$$

$$E(u_1 + u_2) = E(u_1) + E(u_2) \quad (2.17)$$

und damit

$$E(x_1 + x_2 \cdot u_2) = x_1 + x_2 \cdot E(u_2) . \quad (2.18)$$

- Im Regelfall gilt: $E(u_1 \cdot u_2) \neq E(u_1) \cdot E(u_2)$. Nur wenn u_1 und u_2 unkorreliert oder sogar voneinander unabhängig sind, gilt:

$$E(u_1 \cdot u_2) = E(u_1) \cdot E(u_2) . \quad (2.19)$$

- Es gilt immer:

$$E[E(u)] = E(u) .$$

Varianz

- Die Zufallsvariable u_3 ergebe sich aus den anderen Größen gemäß $u_3 = x_1 \cdot u_1 + x_2 \cdot u_2$.
- Es gilt folgende Regel:

$$\text{var}(u_3) = x_1^2 \text{var}(u_1) + x_2^2 \text{var}(u_2) + 2x_1x_2 \text{cov}(u_1, u_2) . \quad (2.20)$$

- Für den Spezialfall $u_3 = x_1 + x_2 \cdot u_2$ (also $u_1 = 1$) ergibt sich:

$$\text{var}(u_3) = x_2^2 \text{var}(u_2) . \quad (2.21)$$

Kovarianz

- Die Kovarianz der zuvor definierten Zufallsvariable u_3 und einer Zufallsvariable u_4 lautet:

$$\begin{aligned} \text{COV}(u_3, u_4) &= \text{COV}(x_1 \cdot u_1 + x_2 \cdot u_2, u_4) \\ &= \text{COV}(x_1 \cdot u_1, u_4) + \text{COV}(x_2 \cdot u_2, u_4) \\ &= x_1 \text{COV}(u_1, u_4) + x_2 \text{COV}(u_2, u_4) .(2.22) \end{aligned}$$

- Definition (2.11) zeigt, dass Kovarianzen immer symmetrisch sind:

$$\text{COV}(u_3, u_4) = \text{COV}(u_4, u_3) \quad (2.23)$$

2.2.7 Die Normalverteilung

- Die Gestalt der Normalverteilung ähnelt Abbildung 2.7.
- Die Gestalt hängt ausschließlich vom Erwartungswert und der Varianz der Zufallsvariable u ab:

$$u \sim N(E(u), var(u)) .$$

2.3 B-Annahmen

- Annahmen, die im Rahmen der *Störgrößen*-Spezifikation, also bezüglich der Variablen u_t getroffen werden, werden hier als *B-Annahmen* bezeichnet.

2.3.1 Begründungen für die Existenz der Störgröße

- 1 Die verwendeten Daten enthalten Erhebungs- und Messfehler.
- 2 Bestimmte erklärende Variablen sind nicht berücksichtigt.
- 3 Das menschliche Verhalten enthält selbst ein Zufallselement.

2.3.2 Störgrößen wiederholter Stichproben

- Die Stichprobe umfasst T Wertepaare (x_t, y_t) .
- Es wird im Rahmen der A-Annahmen unterstellt, dass

$$y_t = \alpha + \beta x_t + u_t .$$

- Dabei besitzen die Parameter α und β unbekannte aber *fixe* Werte.
- Auch die Werte der Variable x_t sind keine Zufallsvariablen.
- Für jede Beobachtung existiert eine eigene Störgröße (u_1, u_2, \dots, u_T) und jede dieser T Störgrößen ist eine eigene *Zufallsvariable*.

Definition: Wiederholte Stichprobe

Eine zweite Stichprobe *mit den gleichen x_t -Werten wie in der ersten Stichprobe* bezeichnet man als (gedanklich) *wiederholte Stichprobe*.

- Jede wiederholte Stichprobe würde neue Ausprägungen (Werte) für die T Zufallsvariablen (Störgrößen) u_1, u_2, \dots, u_T liefern.
- Der Erwartungswert $E(u_1)$ entspricht dem arithmetischen Mittel der unendlich vielen Ausprägungen der Zufallsvariable u_1 .
- Die Varianz $var(u_1)$ misst die Streuung dieser Ausprägungen.
- Analoges gilt für u_2, u_3, \dots, u_T .

Numerische Illustration 2.6

Tabelle 2.4: Mögliche Werte für die Störgrößen u_t bei wiederholten Stichproben.

	Stichpr. 1		Stichpr. 2		Stichpr. 3	
	x_t	u_t	x_t	u_t	x_t	u_t
Beob. 1	10	0,5	10	-0,2	10	-0,1
Beob. 2	30	-1,1	30	0,6	30	0
Beob. 3	50	0,3	50	0,1	50	-0,4
⋮							⋮

2.3.3 Formulierung der B-Annahmen

- Jede der T Störgrößen u_t ist eine eigene Zufallsvariable.
- Bezüglich dieser Zufallsvariablen werden vier Annahmen getroffen.

Annahme b1

Die Störgröße u_t hat für alle Beobachtungen t einen Erwartungswert von 0, das heißt,

$$E(u_t) = 0, \quad \text{für } t = 1, 2, \dots, T. \quad (2.24)$$

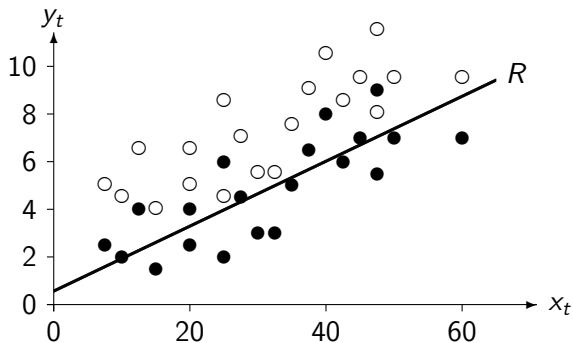


Abbildung 2.8: Konstanter Messfehler bei der Erfassung von y_t .

Annahme b2

Die Störgröße u_t hat für alle Beobachtungen t eine konstante Varianz, das heißt,

$$\text{var}(u_t) = \sigma^2, \quad \text{für } t = 1, 2, \dots, T. \quad (2.25)$$

- Falls Annahme b2 verletzt ist, spricht man von *heteroskedastischen* (oder nicht homoskedastischen) Störgrößen.

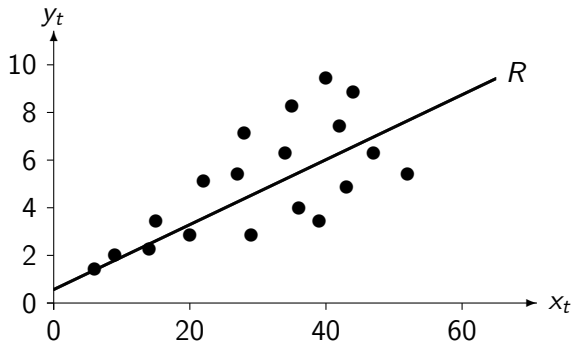


Abbildung 2.9: Eine Punktwolke, die auf heteroskedastische Störgrößen hindeutet.

Annahme b3

Die Störgrößen sind nicht miteinander korreliert, das heißt,

$$\text{cov}(u_t, u_s) = 0, \quad (2.26)$$

für alle $t \neq s$ sowie $t = 1, 2, \dots, T$ und $s = 1, 2, \dots, T$.

- Ist Annahme b3 erfüllt, dann sagt man, dass *keine Autokorrelation* vorliegt.

Annahme b4

Die Störgrößen u_t sind normalverteilt, das heißt,

$$u_t \sim N(E(u), \text{var}(u_t)), \quad \text{für } t = 1, 2, \dots, T.$$

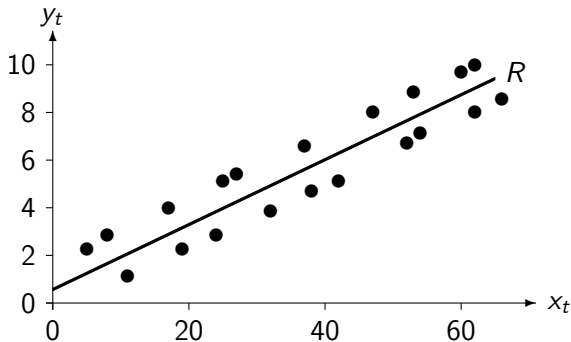


Abbildung 2.10: Eine Punktwolke, die auf Störgrößen hindeutet, welche nicht normalverteilt sind.

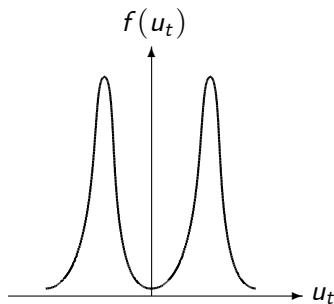


Abbildung 2.11: Eine mögliche Wahrscheinlichkeitsverteilung einer nicht normalverteilten Störgröße u_t .

- Wenn die Annahmen b1 bis b3 erfüllt sind, dann spricht man vom *weißen Rauschen* der Störgrößen.
- Wenn alle vier B-Annahmen erfüllt sind, dann hat jede der T Zufallsvariablen u_t die gleiche Wahrscheinlichkeitsverteilung:

$$u_t \sim UN(0, \sigma^2) , \quad (2.27)$$

für alle $t = 1, 2, \dots, T$.

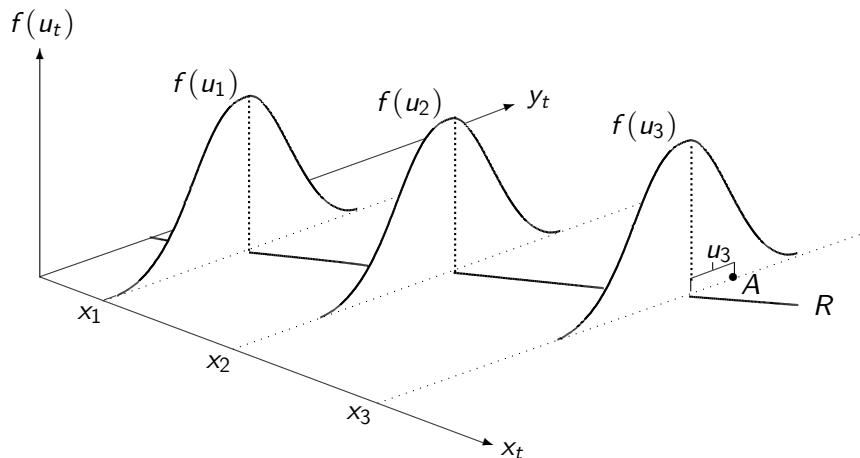


Abbildung 2.12: Eine Veranschaulichung des Annahmenkomplexes b1–b4.

2.4 Statistisches Repetitorium II

- Dieses Repetitorium beschäftigt sich ausschließlich mit dem Thema *Stichproben*.
- Es werden nur diejenigen Aspekte wiederholt, welche später im Kurs benötigt werden.

2.4.1 Stichproben-Mittelwert einer Variable

- Der arithmetische Stichproben-Mittelwert der T Beobachtungen x_1, x_2, \dots, x_T lautet:

$$\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t ,$$

2.4.2 Stichproben-Varianz einer Variable

- Der Durchschnitt der quadratischen Abweichungen vom Stichproben-Mittelwert \bar{x} beträgt:

$$\begin{aligned}\widehat{\text{var}}(x) &= \frac{1}{T-1} \sum_{t=1}^T (x_t - \bar{x})^2 \\ &= \frac{1}{T-1} S_{xx} ,\end{aligned}$$

wobei $S_{xx} = \sum_{t=1}^T (x_t - \bar{x})^2$ als *Variation* der Variable x bezeichnet wird.

- **Frage:** Warum wird durch $T - 1$ und nicht durch T dividiert?

Antwort:

- Die Stichproben-Standardabweichung lautet:

$$\widehat{sd}(x) = \sqrt{\widehat{var}(x)} .$$

- Gebräuchlich ist auch die Bezeichnung „Standardfehler“ und die Notation $se(x)$.

2.4.3 Stichproben-Kovarianz einer Variable

- Für die T Beobachtungspaare $(x_1, y_1), (x_2, y_2), \dots, (x_T, y_T)$ lautet die Stichproben-Kovarianz:

$$\begin{aligned}\widehat{\text{cov}}(x, y) &= \frac{1}{T-1} \sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y}) \\ &= \frac{1}{T-1} S_{xy} .\end{aligned}$$

- Dabei bezeichnet $S_{xy} = \sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})$ die *Kovariation*.

- Der *Stichproben-Korrelationskoeffizient* lautet:

$$\begin{aligned}\widehat{cor}(x,y) &= \frac{\widehat{cov}(x,y)}{\widehat{sd}(x) \cdot \widehat{sd}(y)} \\ &= \frac{S_{xy} / (T - 1)}{\sqrt{S_{xx} / (T - 1)} \cdot \sqrt{S_{yy} / (T - 1)}} \\ &= \frac{S_{xy}}{\sqrt{S_{xx}} \cdot \sqrt{S_{yy}}} .\end{aligned}$$

2.5 C-Annahmen

- Die *Variablen*-Spezifikation geschieht durch das Formulieren der *C-Annahmen*.

Annahme c1

Die exogene Variable x_t ist keine Zufallsvariable, sondern kann wie in einem Experiment kontrolliert werden.

Annahme c2

Die exogene Variable x_t weist nicht für alle Beobachtungen t den gleichen Wert auf: $S_{xx} > 0$.

- Für eine Schätzung der wahren Gerade R müssen mindestens zwei Beobachtungen vorliegen: $T \geq 2$.
- **Frage:** Warum ist diese Anforderung bei Erfüllung der C-Annahmen automatisch erfüllt?

Antwort:

Kapitel 3

Schätzung I: Punktschätzung

- Das ökonometrische Modell der Einfachregression lautet:

$$y_t = \alpha + \beta x_t + u_t . \quad (3.1)$$

Beispiel Trinkgeld

Im Trinkgeld-Beispiel stehen lediglich die nachfolgenden drei Beobachtungen zur Verfügung.

Tabelle 3.1: Rechnungsbetrag x_t und Trinkgeld y_t (beides in Euro) von drei beobachteten Gästen.

t	x_t	y_t
1	10	2
2	30	3
3	50	7

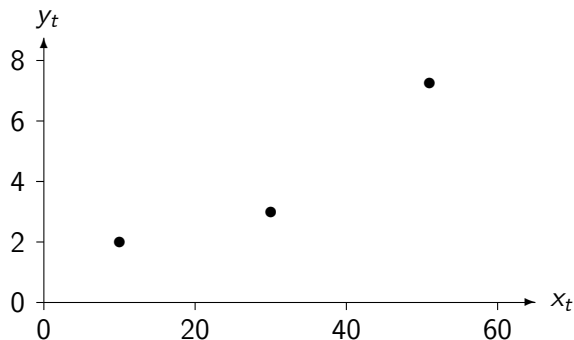


Abbildung 3.1: Die drei Beobachtungspunkte des Trinkgeld-Beispiels.

- *Schätzung*, ist die Berechnung von numerischen Schätzwerten $\hat{\alpha}$ und $\hat{\beta}$ für die wahren Parameter α und β .
- Schätzung ist also der Versuch in Abbildung 3.1 eine Regressionsgerade \hat{R} einzupassen, welche der wahren Gerade R möglichst nahe kommt.
- Um zwischen *wahrer* und *geschätzter* Sphäre zu unterscheiden, wird Letztere stets durch ein Dach markiert: $\hat{\alpha}$ bzw. $\hat{\beta}$.
- Im Folgenden werden $\hat{\alpha}$ bzw. $\hat{\beta}$ als *Schätzer* bezeichnet.

Definition: Geschätztes Modell

Die Gleichung

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t \quad (3.2)$$

wird als das *geschätzte Modell* (Variante I) bezeichnet.

- Dabei gibt die Variable \hat{y}_t zu jedem Rechnungsbetrag x_t den Trinkgeldbetrag an, den sich unseren Schätzern $\hat{\alpha}$ und $\hat{\beta}$ zufolge in einer Welt ohne Störeinflüsse ergeben müsste.
- Definiert man $\hat{u}_t = y_t - \hat{y}_t$, dann lässt sich das geschätzte Modell auch in der Variante II angeben:

$$y_t = \hat{\alpha} + \hat{\beta}x_t + \hat{u}_t .$$

3.1 KQ-Methode – eine Illustration

- Abbildung 3.2 zeigt drei unterschiedliche Regressionsgeraden.
- **Frage:** Welche dieser Geraden würde man als die „am besten passende Gerade“ erachten? Begründen Sie Ihre Wahl.

Antwort:

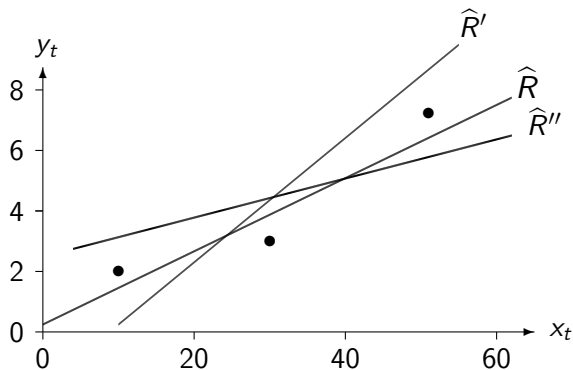


Abbildung 3.2: Drei verschiedene Regressionsgeraden.

- **Frage:** Welche Interpretation besitzen die vertikalen Abstände der Punkte zur Regressionsgerade?

Antwort:

Definition: Residuen

Die Störgrößen u_t sind nicht beobachtbar. Die geschätzten Werte dieser Störgrößen werden als Residuen \hat{u}_t bezeichnet.

- Abbildung 3.3 veranschaulicht den Unterschied zwischen Störgrößen u_t und Residuen \hat{u}_t .

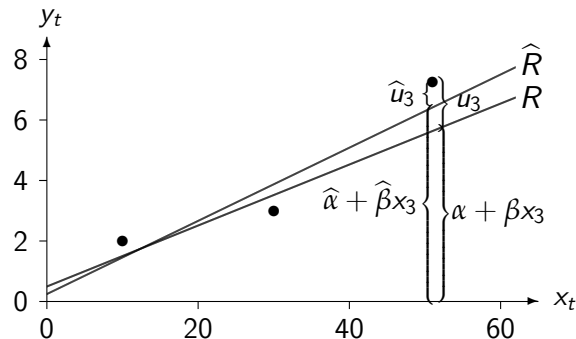


Abbildung 3.3: Der Zusammenhang zwischen Residuen \hat{u}_t und Störgrößen u_t .

- Die *Kleinstbetrag*-Methode minimiert die Summe der Residuenbeträge, also $\sum |\hat{u}_t|$.
- **Frage:** Welche Regressionsgerade würde sich in Abbildung 3.1 aus der Kleinstbetrag-Methode ergeben?
Antwort:

Definition: KQ-Methode

Die *Kleinstquadrat*-Methode (KQ-Methode) minimiert die Summe der *quadrierten* Residuen, also $\sum \hat{u}_t^2$.

- Die Residuenquadrate der KQ-Regressionsgerade sind in Abbildung 3.4 veranschaulicht.

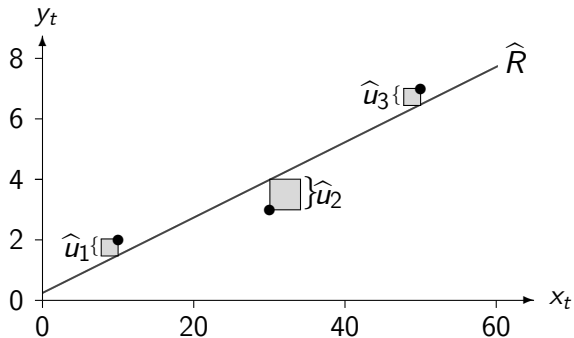


Abbildung 3.4: Die Residuenquadrate der Regressionsgerade \hat{R} .

- Die Kleinstbetrag-Methode und die KQ-Methode liefern unterschiedliche Regressionsgeraden, das heißt unterschiedliche Schätzwerte $\hat{\alpha}$ und $\hat{\beta}$.
- In der KQ-Methode werden die Schätzwerte $\hat{\alpha}$ und $\hat{\beta}$ so gewählt, dass insbesondere einzelne große Residuen vermieden werden.

3.2 KQ-Methode – eine algebraische Formulierung

- Es werden Formeln hergeleitet, welche direkt aus den beobachteten Daten jene Werte $\hat{\alpha}$ und $\hat{\beta}$ liefern, welche die Summe der Residuenquadrate minimieren.
- Diese Formeln werden als die *KQ-Schätzer* bezeichnet.

3.2.1 Summe der Residuenquadrate

- Die Summe der quadrierten Residuen lautet:

$$S_{\widehat{u}\widehat{u}} \equiv \sum_{t=1}^T \widehat{u}_t^2 . \quad (3.3)$$

- Die geschätzte Regressionsgerade \widehat{R} ist das grafische Pendant zu dem durch die Schätzer $\widehat{\alpha}$ und $\widehat{\beta}$ festgelegten *geschätzten Modell*, also zu der linearen Beziehung

$$\widehat{y}_t = \widehat{\alpha} + \widehat{\beta}x_t . \quad (3.4)$$

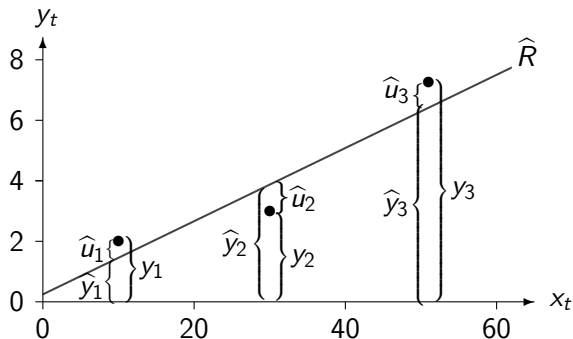


Abbildung 3.5: Der Zusammenhang zwischen beobachteten Trinkgeldbeträgen y_t , vermuteten Trinkgeldbeträgen \hat{y}_t und Residuen \hat{u}_t .

- Aus Abbildung 3.5 erkennt man unmittelbar, dass

$$\hat{u}_t = y_t - \hat{y}_t . \quad (3.5)$$

- Einsetzen der Gleichung (3.4) in (3.5) ergibt

$$\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta}x_t . \quad (3.6)$$

(Auflösen nach y_t würde Variante II des geschätzten Modells ergeben)

- Einsetzen von (3.6) in Gleichung (3.3) liefert

$$S_{\hat{u}\hat{u}} = \sum_{t=1}^T (y_t - \hat{\alpha} - \hat{\beta}x_t)^2 . \quad (3.7)$$

3.2.2 Herleitung der Schätzformeln

- Zur Vereinfachung der Notation werden im Folgenden die Summenzeichen Σ ohne Laufindex $t = 1, 2, \dots, T$ geschrieben, also $\Sigma = \Sigma_{t=1}^T$.
- Die Bedingungen erster Ordnung für ein Minimum von $S_{\hat{u}\hat{u}}$ lauten:

$$\partial S_{\hat{u}\hat{u}} / \partial \hat{\alpha} = \sum 2(y_t - \hat{\alpha} - \hat{\beta}x_t)(-1) = 0 \quad (3.8a)$$

$$\partial S_{\hat{u}\hat{u}} / \partial \hat{\beta} = \sum 2(y_t - \hat{\alpha} - \hat{\beta}x_t)(-x_t) = 0. \quad (3.8b)$$

- Nachdem beide Gleichungen durch -2 dividiert wurden, können sie unter Beachtung der Rechenregeln für Summenzeichen in folgender Form geschrieben werden:

$$\begin{aligned}\sum y_t - \sum \hat{\alpha} - \sum \hat{\beta} x_t &= 0 \\ \sum y_t x_t - \sum \hat{\alpha} x_t - \sum \hat{\beta} x_t x_t &= 0\end{aligned}$$

- Durch weiteres Umformen ergibt sich

$$\sum y_t = T\hat{\alpha} + \hat{\beta} \sum x_t \quad (3.9a)$$

$$\sum y_t x_t = \hat{\alpha} \sum x_t + \hat{\beta} \sum x_t^2. \quad (3.9b)$$

- Die Gleichungen (3.9) nennt man *Normalgleichungen*.
- Dividiert man (3.9a) durch T , erhält man

$$\bar{y} = \hat{\alpha} + \hat{\beta} \bar{x} \quad (3.10)$$

mit $\bar{y} = (1/T) \sum y_t$ und $\bar{x} = (1/T) \sum x_t$.

- Auflösen von (3.10) nach $\hat{\alpha}$ und Einsetzen in (3.9b) ergibt

$$\begin{aligned}\sum y_t x_t &= \sum x_t (\bar{y} - \hat{\beta} \bar{x}) + \hat{\beta} \sum x_t^2 & (3.11) \\ &= \sum x_t \bar{y} - \sum x_t \hat{\beta} \bar{x} + \hat{\beta} \sum x_t^2 \\ &= \bar{y} \sum x_t - \hat{\beta} \bar{x} \sum x_t + \hat{\beta} \sum x_t^2 \\ &= \bar{y} T \bar{x} - \hat{\beta} \bar{x} T \bar{x} + \hat{\beta} \sum x_t^2 \\ &= T \bar{x} \bar{y} - \hat{\beta} T \bar{x}^2 + \hat{\beta} \sum x_t^2\end{aligned}$$

und damit

$$\sum y_t x_t - T \bar{x} \bar{y} = \hat{\beta} (\sum x_t^2 - T \bar{x}^2) . \quad (3.12)$$

- Aus dem Statistischen Repetitorium II stammen die folgenden Definitionen:

$$S_{yy} \equiv \sum (y_t - \bar{y})^2 = \sum (y_t^2 - 2y_t\bar{y} + \bar{y}^2) \quad (3.13a)$$

$$= \sum y_t^2 - \sum 2y_t\bar{y} + \sum \bar{y}^2$$

$$= \sum y_t^2 - 2\bar{y} T \sum y_t / T + T \bar{y}^2$$

$$= \sum y_t^2 - 2\bar{y} T \bar{y} + T \bar{y}^2 = \sum y_t^2 - T \bar{y}^2$$

$$S_{xx} \equiv \sum (x_t - \bar{x})^2 = \sum x_t^2 - T \bar{x}^2 \quad (3.13b)$$

$$S_{xy} \equiv \sum (x_t - \bar{x})(y_t - \bar{y}) = \sum y_t x_t - T \bar{x} \bar{y}. \quad (3.13c)$$

- S_{yy} ist die *Variation* der endogenen Variable, S_{xx} die *Variation* der exogenen Variable und S_{xy} die *Kovariation*.

- Einsetzen von (3.13b) und (3.13c) in (3.12) ergibt

$$S_{xy} = \hat{\beta} S_{xx} . \quad (3.14)$$

Jeweils nach $\hat{\alpha}$ und $\hat{\beta}$ aufgelöst, liefern die Gleichungen (3.14) und (3.10) die gewünschten KQ-Schätzer für α und β .

Formel: KQ-Schätzer (Einfachregression)

$$\hat{\beta} = S_{xy} / S_{xx} \quad (3.15a)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} . \quad (3.15b)$$

Numerische Illustration 3.1

Aus den Daten der Tabelle 3.1 ergibt sich $\bar{x} = 30$ und $\bar{y} = 4$.
Die Variationen betragen

$$S_{yy} \equiv \sum (y_t - \bar{y})^2 = 4 + 1 + 9 = 14 \quad (3.16)$$

$$S_{xx} \equiv \sum (x_t - \bar{x})^2 = 400 + 0 + 400 = 800 . \quad (3.17)$$

Die Kovariation beträgt

$$\begin{aligned} S_{xy} &\equiv \sum (x_t - \bar{x})(y_t - \bar{y}) \\ &= (-20) \cdot (-2) + 0 \cdot 1 + 20 \cdot 3 = 100 . \quad (3.18) \end{aligned}$$

(Forts.)

Die Werte der Schätzer $\hat{\alpha}$ und $\hat{\beta}$ errechnen sich gemäß

$$\hat{\beta} = S_{xy} / S_{xx} = 100 / 800 = 0,125 \quad (3.19)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 4 - 0,125 \cdot 30 = 0,25 . \quad (3.20)$$

Das geschätzte Modell lautet somit

$$\hat{y}_t = 0,25 + 0,125 \cdot x_t .$$

(Forts.)

In einer Welt ohne Störeinflüsse würde man für die Rechnungsbeträge $x_1 = 10$, $x_2 = 30$ und $x_3 = 50$ die folgenden Trinkgelder erwarten:

$$\hat{y}_1 = 0,25 + 0,125 \cdot 10 = 1,5$$

$$\hat{y}_2 = 0,25 + 0,125 \cdot 30 = 4$$

$$\hat{y}_3 = 0,25 + 0,125 \cdot 50 = 6,5 .$$

(Forts.)

Die Residuen lauten folglich:

$$\hat{u}_1 = y_1 - \hat{y}_1 = 2 - 1,5 = 0,5$$

$$\hat{u}_2 = y_2 - \hat{y}_2 = 3 - 4 = -1$$

$$\hat{u}_3 = y_3 - \hat{y}_3 = 7 - 6,5 = 0,5$$

und die Summe der Residuenquadrate beträgt

$$S_{\hat{u}\hat{u}} = \sum \hat{u}_t^2 = 0,5^2 + (-1)^2 + 0,5^2 = 1,5 . \quad (3.21)$$

- Es existieren auch andere Schätzverfahren als die KQ-Methode (z.B. Kleinstbetrag-Methode).
- Wenn $\hat{\alpha}$ und $\hat{\beta}$ mit den KQ-Formeln (3.15) berechnet werden, dann sollte man genaugenommen von den *KQ-Schätzern* $\hat{\alpha}^{KQ}$ und $\hat{\beta}^{KQ}$ sprechen und nicht einfach nur von den *Schätzern* $\hat{\alpha}$ und $\hat{\beta}$.
- In diesem Kurs sind mit $\hat{\alpha}$ und $\hat{\beta}$ immer die KQ-Schätzer gemeint.

- Für die Summe der Residuen einer KQ-Schätzung gilt immer:

$$\sum \hat{u}_t = \sum (y_t - \hat{\alpha} - \hat{\beta}x_t) = \sum y_t - T\hat{\alpha} - \hat{\beta} \sum x_t = 0, \quad (3.22)$$

wobei der letzte Schritt aus Gleichung (3.9a) folgt.

- Damit gilt

$$\bar{\hat{u}} = \sum \hat{u}_t / T = 0.$$

- Folglich erhalten wir

$$S_{\hat{u}\hat{u}} = \sum \hat{u}_t^2 = \sum (\hat{u}_t - \bar{\hat{u}})^2.$$

- Dies ist eine Rechtfertigung für die Notation $S_{\hat{u}\hat{u}}$.

3.3 Interpretation der KQ-Schätzer

Interpretation von $\hat{\beta} = 0,125$

- Der Schätzer $\hat{\beta}$ ist der Steigungsparameter des geschätzten Modells ($\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$):

$$\hat{\beta} = \frac{d\hat{y}_t}{dx_t} .$$

- **Frage:** Was sagt $\hat{\beta} = 0,125$ folglich aus?
Antwort:

Interpretation von $\hat{\alpha} = 0,25$

- **Frage:** Was sagt $\alpha = 0,25$ aus?

Antwort:

- Der Parameter $\hat{\alpha}$ ist eher als eine *technische Größe* zu verstehen, die festlegt, auf welchem *Niveau* sich der beobachtete lineare Zusammenhang abspielt.

3.4 Bestimmtheitsmaß

- Das Bestimmtheitsmaß kann als Kriterium für die „Erklärungskraft“ einer KQ-Schätzung aufgefasst werden.

3.4.1 Grafische Veranschaulichung

- Die Variation der endogenen Variable lautet:

$$S_{yy} \equiv \sum (y_t - \bar{y})^2 .$$

- Folglich ist die Variation S_{yy} die Summe der drei Quadrate der Abbildung 3.6.
- S_{yy} beschreibt das Ausmaß der Schwankungen der endogenen Variable von y_t um ihren Mittelwert \bar{y} .
- Die Regressionsgerade \hat{R}_{KQ} ist der Versuch, diese Schwankungen in y_t durch die unterschiedlichen Werte x_t zu erklären.

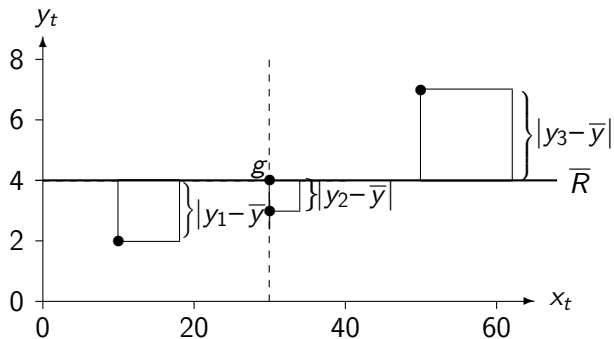


Abbildung 3.6: Die Variation der endogenen Variable y_t .

- Aus der Herleitung der KQ-Schätzformeln für den Niveauparameter α folgt:

$$\bar{y} = \hat{\alpha} + \hat{\beta} \bar{x}. \quad (3.10)$$

- Eine auf Grundlage der KQ-Methode gewonnene Regressionsgerade \hat{R}_{KQ} liefert folglich an der Stelle \bar{x} den Wert \bar{y} .
 - **Frage:** Was bedeutet dies für den Verlauf von \hat{R}_{KQ} in Bezug auf das „Gravitationszentrum“ $g = (\bar{x}, \bar{y})$?
- Antwort:**

- $S_{\widehat{u\hat{u}}}$ ist der „unerklärte Anteil“ an der Variation S_{yy} .
- Wäre $\widehat{R}_{KQ} = \bar{R}$, dann würde dies bedeuten, dass $y_t - \bar{y} = \widehat{u}_t$ und damit $S_{\widehat{u\hat{u}}} = S_{yy}$.
- **Frage:** Was würde dies für die Erklärungskraft der KQ-Schätzung bedeuten?

Antwort:

- Für die in Abbildung 3.7 wiedergegebene Gerade \widehat{R}_{KQ} haben die Residuenquadrate $S_{\widehat{u\hat{u}}}$ – die Summe der mittelgrauen Quadrate – ihr Minimum erreicht.
- Gegenüber \bar{R} verringert \widehat{R}_{KQ} (Abbildung 3.7) den unerklärten Anteil an der Variation der endogenen Variable (S_{yy}) von $S_{yy}/S_{yy} = 1$ auf $S_{\widehat{u\hat{u}}}/S_{yy}$.

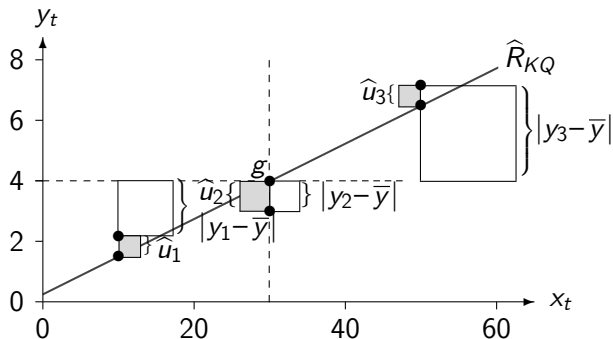


Abbildung 3.7: Die Variation der endogenen Variable y_t und der Anteil an dieser Variation, der auch durch \hat{R}_{KQ} nicht erklärt werden kann.

- Die Differenz aus S_{yy} und $S_{\widehat{u}\widehat{u}}$ ist die Variation der Variable y_t , die durch die Regressionsgerade \widehat{R}_{KQ} erklärt werden kann:

$$S_{yy} - S_{\widehat{u}\widehat{u}} = \text{„erklärte Variation“} . \quad (3.23)$$

- Es kann gezeigt werden, dass für den Fall der KQ-Schätzung folgende Beziehung gilt:

$$\text{„erklärte Variation“} = \sum (\widehat{y}_t - \bar{y})^2 \equiv S_{\widehat{y}\widehat{y}} , \quad (3.24)$$

wobei $\bar{y} = \overline{\widehat{y}}$.

- Grafisch entspricht die „erklärte Variation“ $S_{\widehat{y}\widehat{y}}$ der Summe der in Abbildung 3.8 dunkelgrau schraffierten Flächen.
- Aus den Ausdrücken (3.23) und (3.24) folgt:

$$S_{yy} = S_{\widehat{y}\widehat{y}} + S_{\widehat{u}\widehat{u}} .$$

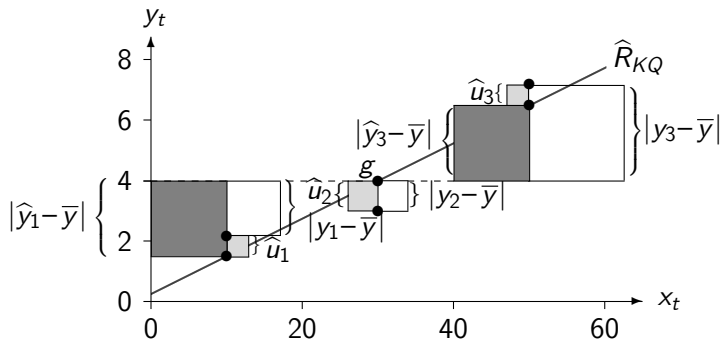


Abbildung 3.8: Die Variation der endogenen Variable y_t und der Anteil an dieser Variation, der durch \hat{R}_{KQ} erklärt werden kann.

3.4.2 Definition des Bestimmtheitsmaßes

Definition: Bestimmtheitsmaß

Das *Bestimmtheitsmaß* R^2 ist folgendermaßen definiert:

$$\begin{aligned} R^2 &= \frac{\text{„erklärte Variation“}}{\text{„gesamte Variation“}} \\ &= \frac{S_{yy} - S_{\hat{u}\hat{u}}}{S_{yy}} = \frac{S_{\hat{y}\hat{y}}}{S_{yy}}. \end{aligned} \quad (3.25)$$

- **Frage:** Welcher Wert ergibt sich für das Bestimmtheitsmaß, wenn die Punktwolke derjenigen in Abbildung 3.9 entspricht?

Antwort:

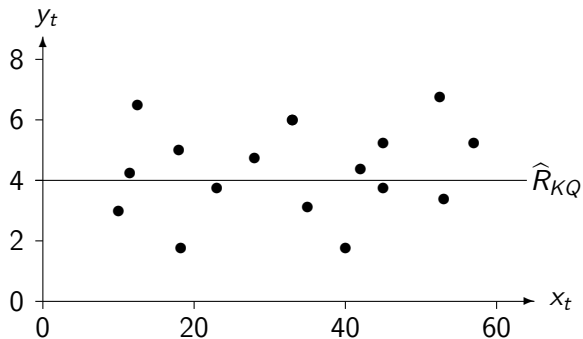


Abbildung 3.9: Eine mögliche Punktwolke.

3.4.3 Berechnung des Bestimmtheitsmaßes

- Es lässt sich zeigen, dass R^2 direkt aus den Beobachtungsdaten mit Hilfe der folgenden Formel berechnet werden kann:

$$R^2 = \frac{\hat{\beta} S_{xy}}{S_{yy}} \quad (3.26)$$

- Wegen $\hat{\beta} = S_{xy} / S_{xx}$ ergibt sich aus (3.26) folgende Berechnungsformel für R^2 :

Formel: Bestimmtheitsmaß

$$R^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} \quad (3.27)$$

Numerische Illustration 3.2

Aus den Ergebnissen (3.16), (3.17) und (3.18) lässt sich das Bestimmtheitsmaß unmittelbar via Gleichung (3.27)

berechnen:

$$R^2 = \frac{100^2}{800 \cdot 14} \approx 89,3\% .$$

Da in (3.21) bereits berechnet wurde, dass $S_{\hat{u}\hat{u}} = 1,5$, könnte man R^2 auch aus Definition (3.25) berechnen:

$$R^2 = \frac{S_{yy} - S_{\hat{u}\hat{u}}}{S_{yy}} = \frac{14 - 1,5}{14} \approx 89,3\% .$$

- Das Bestimmtheitsmaß einer Einfachregression steht in enger Beziehung zum Korrelationskoeffizienten der Variablen x_t und y_t :

$$\text{cor}(x,y) = \frac{S_{xy}}{\sqrt{S_{xx}} \cdot \sqrt{S_{yy}}} = \pm \sqrt{\frac{S_{xy}^2}{S_{xx} S_{yy}}} = \pm \sqrt{R^2}$$

(siehe Statistisches Repetitorium II).

Kapitel 4

Indikatoren für die Qualität von Schätzverfahren

4.1 Statistischer Hintergrund

- Die KQ-Methode schreibt die folgenden Formeln für die Berechnung der Größen $\hat{\alpha}$ und $\hat{\beta}$ vor:

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}}$$
$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}.$$

- Die KQ-Schätzer werden weiterhin durch $\hat{\alpha}$ und $\hat{\beta}$ anstatt durch $\hat{\alpha}^{KQ}$ und $\hat{\beta}^{KQ}$ bezeichnet.
- Grundlage ist wieder das Konzept der wiederholten Stichprobe.

- Jede der T Beobachtungen besteht aus einem Wertepaar (x_t, y_t) .
- Es wird unterstellt, dass die Beziehung zwischen jedem Wertepaar x_t und y_t durch das ökonometrische Modell

$$y_t = \alpha + \beta x_t + u_t \quad (4.1)$$

beschrieben werden kann.

4.1.1 Warum ist y_t eine Zufallsvariable?

- Für jede wiederholte Stichprobe ergäben sich für alle u_t normalerweise neue Werte.
- Jede der T Störgrößen u_t ist deshalb eine eigene Zufallsvariable.
- Da

$$y_t = \alpha + \beta x_t + u_t ,$$

sind die Werte von y_t lediglich eine „mechanische Transformation“ der Zufallsvariable u_t .

Numerische Illustration 4.1

Tabelle 4.1: Mögliche Werte für die Störgrößen u_t und Trinkgelder y_t bei wiederholten Stichproben.

	Stichprobe 1			Stichprobe 2			Stichprobe 3		
	x_t	u_t	y_t	x_t	u_t	y_t	x_t	u_t	y_t
Beob. 1	10	0,5	2,0	10	-0,2	1,3	10	-0,1	1,4
Beob. 2	30	-1,1	3,0	30	0,6	4,7	30	0	4,1
Beob. 3	50	0,3	7,0	50	0,1	6,8	50	-0,4	6,3
$\hat{\alpha}$ und $\hat{\beta}$	0,25/0,125			0,142/0,138			0,258/0,123			

(Forts.)

Es sei $\alpha = 0,2$ und $\beta = 0,13$. Das wahre Modell lautet folglich:

$$y_t = 0,2 + 0,13 \cdot x_t + u_t .$$

Die Zufallsvariablen u_1 , u_2 und u_3 haben in den wiederholten Stichproben die in Tabelle 4.1 aufgeführten (eigentlich nicht beobachtbaren) Werte angenommen.

Frage: Überprüfen Sie in Tabelle 4.1 die Werte von y_1 .

Antwort:

4.1.2 Warum sind die KQ-Schätzer Zufallsvariablen?

- Es gilt $\bar{y} = (\sum y_t) / T$ und

$$\begin{aligned}
 \hat{\beta} &= \frac{S_{xy}}{S_{xx}} \\
 &= \frac{\sum x_t y_t - T \bar{x} \bar{y}}{S_{xx}} \\
 [\text{aus (3.13b)}] &= \frac{\sum x_t y_t - \sum y_t \bar{x}}{S_{xx}} \tag{4.2}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\sum y_t (x_t - \bar{x})}{S_{xx}} \\
 &= \sum y_t \frac{x_t - \bar{x}}{S_{xx}} . \tag{4.3}
 \end{aligned}$$

- Ferner gilt:

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}. \quad (4.4)$$

Numerische Illustration 4.2

Tabelle 4.1 zeigt, dass sich in jeder wiederholten Stichprobe neue Werte für die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ einstellen würden.

- Die KQ-Schätzer

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

sind Zufallsvariablen.

- Die KQ-Schätzer besitzen folglich jeweils eine eigene Wahrscheinlichkeitsverteilung (jeweils mit Erwartungswert und Varianz).
- Auch bei anderen Schätzverfahren wären die Schätzer für α und β Zufallsvariablen mit einer Wahrscheinlichkeitsverteilung.

- └ 4. Indikatoren für die Qualität von Schätzverfahren
- └ 4.2. Zwei Kriterien: Unverzerrtheit und Effizienz

4.2 Zwei Kriterien: Unverzerrtheit und Effizienz

Definition: Unverzerrtheit

Ein Schätzer $\hat{\beta}^A$ heißt *unverzerrt*, wenn die aus wiederholten Stichproben ermittelten Werte $\hat{\beta}^A$ im Mittel den wahren Wert β treffen würden, das heißt, wenn gilt:

$$E(\hat{\beta}^A) = \beta .$$

Definition: Effizienz

Ein unverzerrter Schätzer $\hat{\beta}^A$ ist *effizient*, wenn er innerhalb der Klasse der unverzerrten Schätzer die kleinste Streuung $\text{var}(\hat{\beta}^A)$ aufweist.

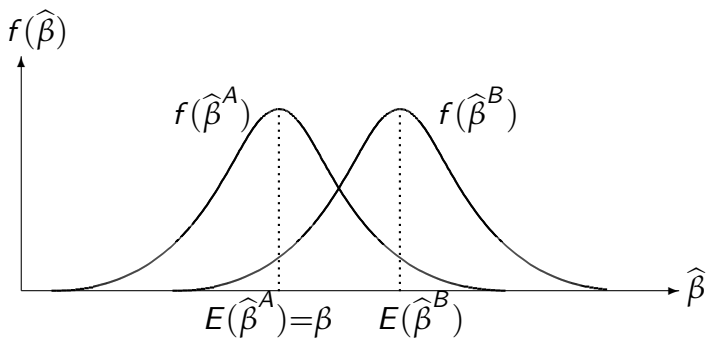


Abbildung 4.1: Ein unverzerrter ($\hat{\beta}^A$) und ein verzerrter ($\hat{\beta}^B$) Schätzer.

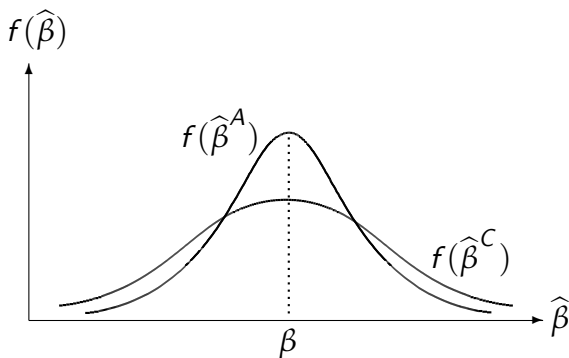


Abbildung 4.2: Zwei unverzerrte Schätzer mit unterschiedlicher Varianz.

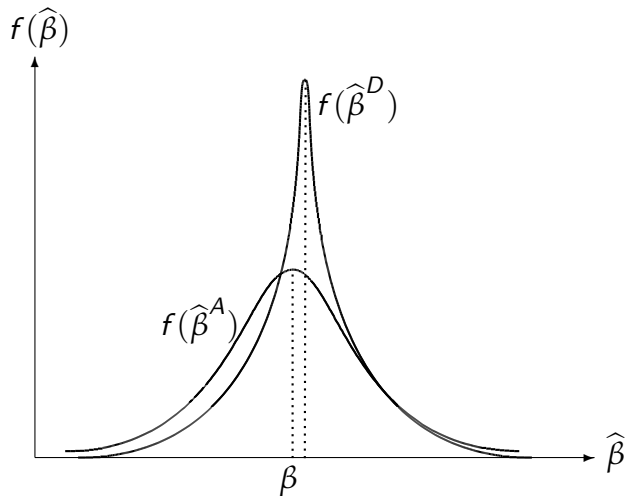


Abbildung 4.3: Ein verzerrter Schätzer mit geringer Varianz ($\hat{\beta}^D$) und ein unverzerrter Schätzer mit größerer Varianz ($\hat{\beta}^A$).

4.3 Unverzerrtheit und Effizienz der KQ-Methode

- Es lässt sich für die KQ-Schätzer zeigen, dass unter den getroffenen A-, B- und C-Annahmen (Annahmen b2 bis b4 sind allerdings nicht erforderlich):

$$E(\hat{\alpha}) = \alpha \quad \text{und} \quad E(\hat{\beta}) = \beta . \quad (4.5)$$

KQ-Schätzer sind unverzerrt

Unter den getroffenen A-, B- und C-Annahmen (Annahmen b2 bis b4 nicht erforderlich) sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ unverzerrt.

- Ferner lässt sich zeigen, dass

$$\text{var}(\hat{\beta}) = \sigma^2 / S_{xx} \quad (4.6a)$$

und

$$\text{var}(\hat{\alpha}) = \sigma^2 (1/T + \bar{x}^2 / S_{xx}) \quad (4.6b)$$

$$\text{cov}(\hat{\alpha}, \hat{\beta}) = -\sigma^2 (\bar{x} / S_{xx}) . \quad (4.6c)$$

- Die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ sind *lineare Schätzer*.

- └ 4. Indikatoren für die Qualität von Schätzverfahren
 - └ 4.3. Unverzerrtheit und Effizienz der KQ-Methode

- **Frage:** Welchen Einfluss besitzen die Größen σ^2 und S_{xx} auf die Varianz des KQ-Schätzers $\hat{\beta}$?

Antwort:

- └ 4. Indikatoren für die Qualität von Schätzverfahren
 - └ 4.3. Unverzerrtheit und Effizienz der KQ-Methode

- Man kann zeigen, dass die Varianzen (4.6a) und (4.6b) die geringst möglichen innerhalb der Klasse der unverzerrten linearen Schätzer darstellen.

KQ-Schätzer sind BLUE

Unter den getroffenen A-, B- und C-Annahmen (Annahme b4 nicht erforderlich) sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ innerhalb der Klasse der unverzerrten linearen Schätzer effizient (BLUE-Eigenschaft).

- Wenn auch Annahme b4 (Störterme u_t sind normalverteilt) hinzugenommen wird, dann lässt sich zeigen:

KQ-Schätzer sind BUE

Unter den getroffenen A-, B- und C-Annahmen sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ innerhalb der Klasse aller unverzerrten Schätzer, nicht-lineare Schätzer eingeschlossen, effizient (BUE-Eigenschaft).

- Ein und dieselbe Varianz ist mit sehr verschiedenen Verteilungen vereinbar (siehe Abbildung 4.4).

- └ 4. Indikatoren für die Qualität von Schätzverfahren
 - └ 4.3. Unverzerrtheit und Effizienz der KQ-Methode

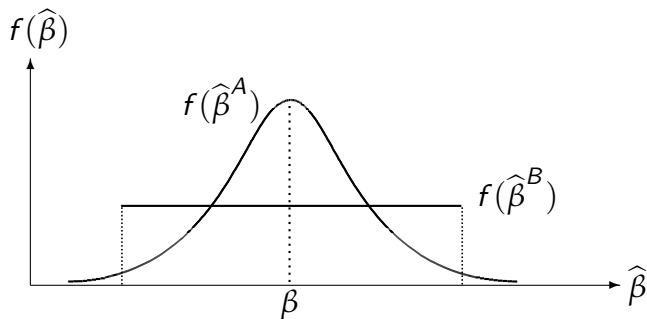


Abbildung 4.4: Zwei unverzerrte Schätzer mit identischer Varianz.

4.4 Statistisches Repetitorium III

- Es werden die folgenden Wahrscheinlichkeitsverteilungen kurz vorgestellt:
 - Standard-Normalverteilung
 - χ^2 -Verteilung
 - t -Verteilung
 - F -Verteilung

4.4.1 Standard-Normalverteilung

- Die *Standard-Normalverteilung* ist eine Normalverteilung mit einem Erwartungswert von 0 und einer Varianz von 1.
- Jede normalverteilte Zufallsvariable lässt sich durch geeignete Transformation in eine standard-normalverteilte Zufallsvariable überführen.
- Es sei u eine normalverteilte Zufallsvariable:

$$u \sim N(E(u), var(u))$$

mit $E(u) = \mu$ und $var(u) = \sigma^2$ (d.h. $sd(u) = \sigma$).

- Die Zufallsvariable u wird linear transformiert:

$$z = \frac{1}{\sigma}u - \frac{\mu}{\sigma} = \frac{u - \mu}{\sigma}. \quad (4.7)$$

- Die Zufallsvariable z ist ebenfalls normalverteilt, denn die lineare Transformation einer normalverteilten Zufallsvariable folgt ebenfalls einer Normalverteilung.
- Für den Erwartungswert $E(z)$ erhält man:

$$\begin{aligned} E(z) &= E\left(\frac{1}{\sigma}u - \frac{\mu}{\sigma}\right) \\ [\text{aus (2.16), (2.17)}] &= \frac{1}{\sigma}E(u) - E\left(\frac{\mu}{\sigma}\right) \\ [\text{aus (2.15)}] &= \frac{1}{\sigma}\mu - \frac{\mu}{\sigma} = 0. \end{aligned}$$

- Für die Varianz $var(z)$ ergibt sich:

$$\begin{aligned} var(z) &= var\left(\frac{1}{\sigma}u - \frac{\mu}{\sigma}\right) \\ [aus (2.21)] \quad &= \frac{1}{\sigma^2}var(u) = \frac{1}{\sigma^2}\sigma^2 = 1 . \end{aligned}$$

- Folglich gilt:

$$z \sim N(0,1) .$$

- Man bezeichnet die Transformation (4.7) als *Standardisierung der Zufallsvariable u* .

4.4.2 χ^2 -Verteilung

- Betrachtet wird die folgende Zufallsvariable:

$$\begin{aligned}c &= \text{“Summe aus } v \text{ quadrierten standard-normalverteilten} \\ &\quad \text{unabhängigen Zufallsvariablen”} . \\ &= z_1^2 + z_2^2 + \dots + z_v^2 .\end{aligned}$$

- Diese zusammengesetzte Zufallsvariable c folgt einer $\chi^2_{(v)}$ -Verteilung:

$$c \sim \chi^2_{(v)} .$$

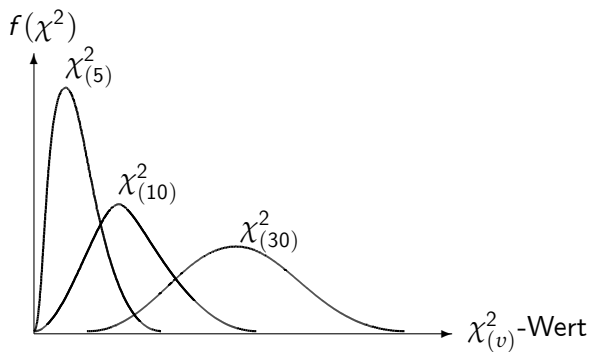


Abbildung 4.5: Drei exemplarische $\chi^2_{(v)}$ -Verteilungen mit unterschiedlichen Freiheitsgraden v .

4.4.3 t -Verteilung

- Es sei c eine $\chi^2_{(v)}$ -verteilte Zufallsvariable und z eine von c unabhängige, standard-normalverteilte Zufallsvariable.
- Aus diesen beiden Zufallsvariablen lässt sich folgende Zufallsvariable konstruieren:

$$t = \frac{z}{\sqrt{c/v}} .$$

- Diese Zufallsvariable folgt einer t -Verteilung mit v Freiheitsgraden:

$$t \sim t_{(v)} .$$

- Für die $t_{(v)}$ -Verteilung gilt

$$E(t) = 0 .$$

- Die $t_{(v)}$ -Verteilung mit unendlich vielen Freiheitsgraden ($v \rightarrow \infty$) entspricht der Standard-Normalverteilung.

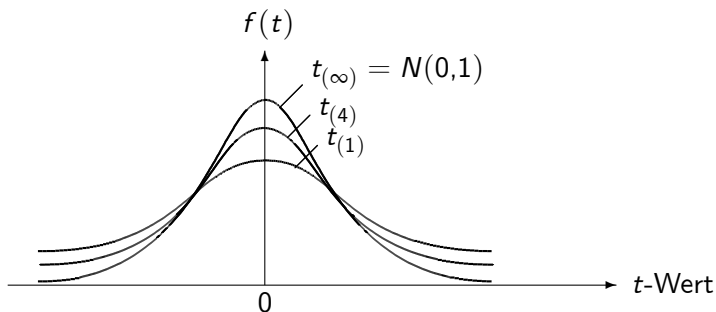


Abbildung 4.6: Drei exemplarische $t_{(v)}$ -Verteilungen mit unterschiedlichen Freiheitsgraden v .

4.4.4 F -Verteilung

- Es sei c_1 eine $\chi^2_{(v_1)}$ -verteilte Zufallsvariable und c_2 eine von c_1 unabhängige, $\chi^2_{(v_2)}$ -verteilte Zufallsvariable.
- Aus c_1 und c_2 lässt sich die folgende Zufallsvariable bilden:

$$F = \frac{c_1/v_1}{c_2/v_2} = \frac{c_1}{c_2} \cdot \frac{v_2}{v_1} .$$

- Diese Zufallsvariable ist F -verteilt:

$$F \sim F_{(v_1, v_2)} .$$

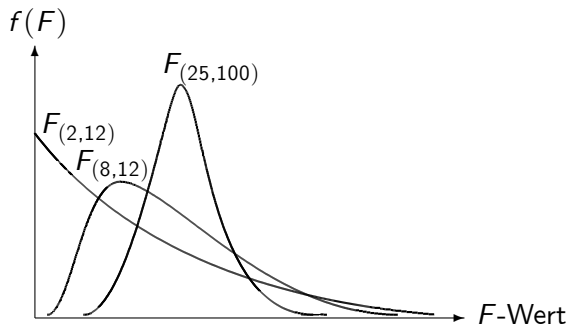


Abbildung 4.7: Drei exemplarische $F_{(v_1, v_2)}$ -Verteilungen mit unterschiedlichen Freiheitsgraden v_1 und v_2 .

4.5 Wahrscheinlichkeitsverteilungen der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$

- Die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ sind lineare Transformationen der Zufallsvariablen y_t .

4.5.1 Wahrscheinlichkeitsverteilung von y_t

- Aus den Annahmen b1 bis b4 ergibt sich:

$$u_t \sim UN(0, \sigma^2), \quad \text{für alle } t = 1, 2, \dots, T.$$

- Da die Störgröße u_t normalverteilt ist, gilt dies auch für ihre lineare Transformation

$$y_t = \alpha + \beta x_t + u_t.$$

- Außerdem gilt:

$$\begin{aligned} E(y_t) &= E[\alpha + \beta x_t + u_t] \\ &= E(\alpha) + E(\beta x_t) + E(u_t) \\ &= \alpha + \beta x_t + 0 \end{aligned}$$

- Ferner gilt:

$$\begin{aligned} \text{var}(y_t) &= E \left[[y_t - E(y_t)]^2 \right] \\ &= E \left[[y_t - \alpha - \beta x_t]^2 \right] \\ &= E \left[u_t^2 \right] \\ &= E \left[[u_t - \underbrace{E(u_t)}_{=0}]^2 \right] && (= \text{var}(u_t)) \\ &= \sigma^2 . \end{aligned}$$

- Da die Unabhängigkeits-Eigenschaft der u_t durch die Transformation nicht berührt wird, sind auch die y_t voneinander unabhängig.
- Diese Resultate implizieren, dass

$$y_t \sim UN(\alpha + \beta x_t, \sigma^2), \quad \text{für alle } t = 1, 2, \dots, T.$$

4.5.2 Wahrscheinlichkeitsverteilung von $\hat{\alpha}$ und $\hat{\beta}$

- Die gewichtete Summe unabhängig normalverteilter Zufallsvariablen ist ebenfalls normalverteilt.
- Gleichungen (4.3) und (4.4) zeigen, dass sich die Formeln der KQ-Schätzer (Zufallsvariablen) $\hat{\alpha}$ und $\hat{\beta}$ als eine gewichtete Summe der unabhängig normalverteilten Zufallsvariablen y_t darstellen lassen.

- Unter Ausnutzung der Ergebnisse (4.5), (4.6a) und (4.6b) ergibt sich das folgende Resultat:

KQ-Schätzer sind normalverteilt

Die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ sind normalverteilte Zufallsvariablen mit

$$\begin{aligned}\hat{\alpha} &\sim N(\alpha, \sigma^2(1/T + \bar{x}^2/S_{xx})) \\ \hat{\beta} &\sim N(\beta, \sigma^2/S_{xx}) .\end{aligned}\tag{4.8}$$

Kapitel 5

Schätzung II: Intervallschätzer

- Die Schätzer $\hat{\alpha}$ und $\hat{\beta}$ werden häufig als *Punktschätzer* bezeichnet.
- Es wird im Folgenden eine Schätzformel entwickelt, die aus den Daten einer Stichprobe ein *Intervall* berechnet.
- Die Formel bezeichnet man als *Intervallschätzer*.
- Grundlage für die Konstruktion der Formel ist die Wahrscheinlichkeitsverteilung des Schätzers $\hat{\beta}$:

$$\hat{\beta} \sim N(\beta, \sigma^2 / S_{xx}) . \quad (5.1)$$

- Sowohl die *Position* als auch die *Gestalt* dieser Normalverteilung sind unbekannt.

5.1 Intervallschätzer und ihre Interpretation

- Der KQ-Punktschätzer für β lautet:

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} . \quad (5.2)$$

- Die Formel für den (KQ-)Intervallschätzer soll durch

$$[\hat{\beta} - k_{a/2} ; \hat{\beta} + k_{a/2}] \quad (5.3)$$

beschrieben werden können.

- Sie soll so beschaffen sein, dass sie mit vorgegebener Wahrscheinlichkeit $1-a$ Intervalle $[\hat{\beta} - k_{a/2} ; \hat{\beta} + k_{a/2}]$ liefert, die den wahren Wert β überdecken:

$$\Pr \left\{ \hat{\beta} - k_{a/2} \leq \beta \leq \hat{\beta} + k_{a/2} \right\} = 1 - a . \quad (5.4)$$

- Den Wert a bezeichnet man als *Signifikanzniveau*.
- Das aus der Stichprobe tatsächlich berechnete Intervall $\left[\hat{\beta} - k_{a/2} ; \hat{\beta} + k_{a/2} \right]$ bezeichnet man als *Konfidenzintervall*.
- Dieser Begriff wird in der Statistik immer für Intervalle verwendet, die für unbekannte aber zufallsunabhängige Parameter (z.B. α, β, σ^2) berechnet worden sind.
- Jede wiederholte Stichprobe würde ein neues Konfidenzintervall für β erzeugen.

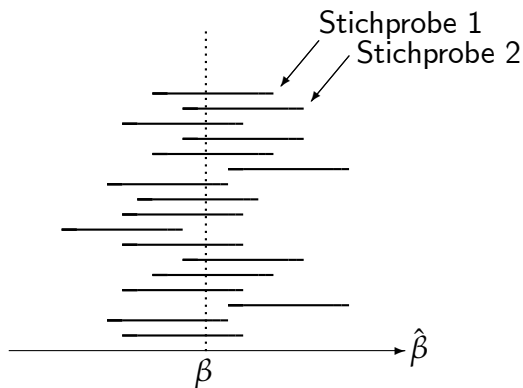


Abbildung 5.1: Konfidenzintervalle, die ein Intervallschätzer bei wiederholten Stichproben und bekanntem σ^2 liefern könnte.

5.2 Intervallschätzer für β bei bekanntem σ^2

Schritt 1: Standardisierung von $\hat{\beta}$

- Aus $\text{var}(\hat{\beta}) = \sigma^2 / S_{xx}$ folgt:

$$\text{sd}(\hat{\beta}) = \sqrt{\sigma^2 / S_{xx}} , \quad (5.5)$$

wobei unterstellt ist, dass σ^2 bekannt ist.

- Die normalverteilte Zufallsvariable $\hat{\beta}$ wird standardisiert:

$$z = \frac{\hat{\beta} - E(\hat{\beta})}{\text{sd}(\hat{\beta})} = \frac{\hat{\beta} - \beta}{\text{sd}(\hat{\beta})} . \quad (5.6)$$

- Die Zufallsvariable z ist demnach standard-normalverteilt:

$$z \sim N(0,1) .$$

- **Frage:** Warum gehen wir den Umweg über z statt direkt mit $\hat{\beta}$ zu arbeiten?

Antwort:

Schritt 2: Ermittlung eines $z_{\alpha/2}$ -Wertes

- Die Wahrscheinlichkeitsverteilung einer standard-normalverteilten Zufallsvariablen z ist bekannt.
- Die Standard-Normalverteilung ist symmetrisch um 0.
- **Frage:** Es gilt $\Pr \{z \leq 0,44\} = 0,67$. Wie groß ist die Wahrscheinlichkeit, dass z in das Intervall $[-0,44 ; 0,44]$ fällt?

Antwort:

- Für eine standard-normalverteilte Zufallsvariable z lassen sich deshalb für jedes Signifikanzniveau a feste Intervallgrenzen $-z_{a/2}$ und $z_{a/2}$ angeben, so dass

$$\Pr \{-z_{a/2} \leq z \leq z_{a/2}\} = 1 - a . \quad (5.7)$$

- In Worten: Bei einem Anteil von $1-a$ der wiederholten Stichproben liegt der jeweilige Wert für die standard-normalverteilte Zufallsvariable z im (festen) Intervall $[-z_{a/2} ; z_{a/2}]$.

Numerische Illustration 5.1

Es sei ein Signifikanzniveau von 5% vorgegeben, also $\alpha = 0,05$ und damit $\alpha/2 = 0,025$.

Frage: Welche Werte $-z_{0,025}$ und $z_{0,025}$ lassen sich aus Tabelle T.1 ablesen?

Antwort:

Schritt 3: Formulierung des Intervallschätzers

- Setzt man (5.6) in Gleichung (5.7) ein, so ergibt sich

$$\Pr \left\{ -z_{a/2} \leq \frac{\hat{\beta} - \beta}{sd(\hat{\beta})} \leq z_{a/2} \right\} = 1 - a. \quad (5.8)$$

- Die zulässigen mathematischen Umformungen der Ungleichungen innerhalb der geschweiften Klammern, verändern die Aussage der Gleichung nicht.

Numerische Illustration 5.2

X sei die Zufallsvariable „geworfene Augenzahl bei einmaligem Würfeln“.

Frage: Für welche Ausprägungen ist $\{3 \leq X \leq 5\}$ erfüllt und wie groß ist $\Pr \{3 \leq X \leq 5\}$?

Antwort:

Frage: Für welche Ausprägungen ist $\{2 \leq X - 1 \leq 4\}$ erfüllt?

Antwort:

Frage: Wie groß ist also $\Pr \{2 \leq X - 1 \leq 4\}$?

Antwort:

- Man kann also in (5.8) innerhalb der geschweiften Klammern nach β auflösen:

$$\begin{aligned}\Pr\left\{-\hat{\beta} - z_{a/2} \cdot sd(\hat{\beta}) \leq -\beta \leq -\hat{\beta} + z_{a/2} \cdot sd(\hat{\beta})\right\} &= 1-a \\ \Rightarrow \Pr\left\{\hat{\beta} + z_{a/2} \cdot sd(\hat{\beta}) \geq \beta \geq \hat{\beta} - z_{a/2} \cdot sd(\hat{\beta})\right\} &= 1-a \\ \Rightarrow \Pr\left\{\hat{\beta} - z_{a/2} \cdot sd(\hat{\beta}) \leq \beta \leq \hat{\beta} + z_{a/2} \cdot sd(\hat{\beta})\right\} &= 1-a. \quad (5.9)\end{aligned}$$

- Gleichung (5.9) legt erneut ein Intervall fest:

$$\left[\hat{\beta} - z_{a/2} \cdot sd(\hat{\beta}) ; \hat{\beta} + z_{a/2} \cdot sd(\hat{\beta})\right]. \quad (5.10)$$

- In Worten: Bei einem Anteil von $1-a$ der wiederholten Stichproben deckt das jeweils gemäß (5.10) berechnete Konfidenzintervall den festen wahren Wert β ab.

- Ein Vergleich der Beziehungen (5.9) und (5.4) bzw. der Intervalle (5.10) und (5.3) zeigt unmittelbar, dass

$$k_{\alpha/2} = z_{\alpha/2} \cdot sd(\hat{\beta}) . \quad (5.11)$$

- Das Intervall $\left[\hat{\beta} - z_{\alpha/2} \cdot sd(\hat{\beta}) ; \hat{\beta} + z_{\alpha/2} \cdot sd(\hat{\beta}) \right]$ ist der Intervallschätzer des Parameters β .

Numerische Illustration 5.3

Es wurde bereits $z_{\alpha/2} = 1,96$ ermittelt. Es wird (unzulässigerweise) angenommen, die Varianz der Störgröße u_t sei bekannt und betrage $\sigma^2 = 2$. Folglich ergibt sich aus Gleichung (4.6a):

$$sd(\hat{\beta}) = \sqrt{\sigma^2 / S_{xx}} = \sqrt{2/800} = 0,05 .$$

(Forts.)

Der Intervallschätzer (5.10) wird folglich zu

$$\left[\hat{\beta} - 1,96 \cdot 0,05 ; \hat{\beta} + 1,96 \cdot 0,05 \right] .$$

Aus (3.19) ist bereits bekannt, dass $\hat{\beta} = 0,125$. Einsetzen liefert

$$[0,125 - 1,96 \cdot 0,05 ; 0,125 + 1,96 \cdot 0,05]$$

und damit das Konfidenzintervall $[0,027 ; 0,223]$.

5.3 Intervallschätzer für β bei unbekanntem σ^2

- In der Realität ist σ^2 unbekannt.
- Folglich muss σ^2 in einem vorgelagerten Schritt geschätzt werden.

5.3.1 Herleitung des Intervallschätzers

Schritt 1: Schätzung von σ^2 und $sd(\hat{\beta})$

Numerische Illustration 5.4

Tabelle 5.1 wiederholt die in Tabelle 4.1 wiedergegebenen – eigentlich nicht beobachtbaren – Störgrößen (gedanklich wiederholter) Stichproben.

(Forts.)

Tabelle 5.1: Mögliche Werte für die Störgrößen u_t bei wiederholten Stichproben.

	Stichprobe		
	1	2	3
Beobachtung 1	0,5	-0,2	-0,1
Beobachtung 2	-1,1	0,6	0
Beobachtung 3	0,3	0,1	-0,4

Mit Hilfe der KQ-Methode wurden aus Stichprobe 1 die Werte $\hat{\alpha} = 0,25$ und $\hat{\beta} = 0,125$ berechnet und hieraus wiederum die Residuen $\hat{u}_1 = 0,5$, $\hat{u}_2 = -1$ und $\hat{u}_3 = 0,5$.

(Forts.)

Alle Störgrößenwerte der Tabelle 5.1 werden durch den gleichen „Zufallsgenerator“ erzeugt.

Frage: Warum kann dies für die Schätzung von σ^2 ausgenutzt werden?

Antwort:

- Wären die Störgrößenwerte direkt beobachtbar, könnte als Schätzformel für σ^2 einfach die Stichproben-Varianz

$$\hat{\sigma}^2 = \frac{S_{uu}}{T - 1}$$

verwendet werden.

- Es sind allerdings lediglich die Residuen bekannt.

- Folgendes Resultat kann gezeigt werden:

Formel: Schätzer der Störgrößenvarianz

Wenn Annahmen b1 bis b3 erfüllt sind, dann ist im Rahmen der Einfachregression die Formel

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T - 2} \quad (5.12)$$

ein unverzerrter Schätzer für σ^2 .

- Ist $\hat{\sigma}^2$ ermittelt, kann auch

$$\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}} \quad (5.13)$$

berechnet werden.

Numerische Illustration 5.5

Aus (3.21) ist bekannt, dass $S_{\hat{u}\hat{u}} = 1,5$. Die Schätzung der unbekanntenen Störgrößenvarianz σ^2 beträgt

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 2) = 1,5 / 1 = 1,5 . \quad (5.14)$$

Der Wert $S_{xx} = 800$ ist aus Resultat (3.17) bekannt. Die Standardabweichung des Schätzers $\hat{\beta}$ errechnet sich gemäß

$$\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}} = \sqrt{1,5 / 800} \approx 0,0433 . \quad (5.15)$$

Schritt 2: Standardisierung von $\hat{\beta}$

- Die Zufallsvariable $\hat{\beta}$ wird der folgenden Transformation unterzogen:

$$t = \frac{\hat{\beta} - E(\hat{\beta})}{\widehat{sd}(\hat{\beta})} = \frac{\hat{\beta} - \beta}{\widehat{sd}(\hat{\beta})} . \quad (5.16)$$

- Die Zufallsvariable t folgt einer t -Verteilung mit $(T-2)$ Freiheitsgraden:

$$t \sim t_{(T-2)} .$$

Schritt 3: Ermittlung eines $t_{a/2}$ -Wertes

- Aus der statistischen Theorie ist bekannt, dass sich für eine $t_{(T-2)}$ -verteilte Zufallsvariable t für ein vorgegebenes Signifikanzniveau a feste Intervallgrenzen $-t_{a/2}$ und $t_{a/2}$ angeben lassen, so dass

$$\Pr \{ -t_{a/2} \leq t \leq t_{a/2} \} = 1 - a . \quad (5.17)$$

- In Worten: Bei einem Anteil von $1-a$ der wiederholten Stichproben fällt der jeweilige Wert t ins Intervall $[-t_{a/2} ; t_{a/2}]$.

Numerische Illustration 5.6

Aus Tabelle T.2 können die Werte $-t_{a/2}$ und $t_{a/2}$ abgelesen werden.

Frage: Welche Werte ergeben sich für $a = 0,05$ und $T = 3$?

Antwort:

Schritt 4: Formulierung des Intervallschätzers

- Setzt man (5.16) in Gleichung (5.17) ein, und löst nach β auf, so ergibt sich

$$\Pr \left\{ \hat{\beta} - t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \leq \beta \leq \hat{\beta} + t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \right\} = 1 - a . \quad (5.18)$$

- Gleichung (5.18) definiert den Intervallschätzer.

Formel: Intervallschätzer

$$\left[\hat{\beta} - t_{a/2} \cdot \widehat{sd}(\hat{\beta}) ; \hat{\beta} + t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \right] \quad (5.19)$$

- Der Vergleich der Intervalle (5.19) und (5.3) zeigt, dass

$$k = t_{\alpha/2} \cdot \widehat{sd}(\widehat{\beta}) . \quad (5.20)$$

- Sämtliche numerischen Werte des Intervallschätzers (5.19) sind bekannt.

Numerische Illustration 5.7

In Schritt 1 wurde $\widehat{sd}(\widehat{\beta}) \approx 0,0433$ errechnet. In Schritt 3 ergab sich $t_{\alpha/2} = 12,7062$. Aus Resultat (3.19) ist bekannt, dass $\widehat{\beta} = 0,125$. Einsetzen der bisherigen Ergebnisse in den Intervallschätzer (5.19) liefert

$$[0,125 - 12,7062 \cdot 0,0433 ; 0,125 + 12,7062 \cdot 0,0433]$$

und damit das Konfidenzintervall $[-0,4252 ; 0,6752]$.

- Da in der Praxis σ^2 niemals bekannt ist, benutzt man den Intervallschätzer (5.19) und nicht den Intervallschätzer (5.10).

5.3.2 Interpretation des Intervallschätzers

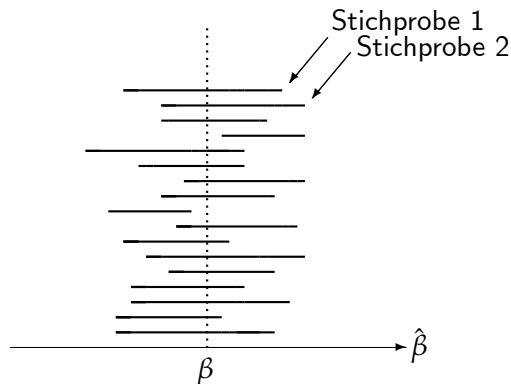


Abbildung 5.2: Konfidenzintervalle, die ein Intervallschätzer bei wiederholten Stichproben und unbekanntem σ^2 liefern könnte.

Numerische Illustration 5.8

Das Konfidenzintervall des Trinkgeld-Beispiels beträgt $[-0,4252 ; 0,6752]$.

Dieses Konfidenzintervall sagt nicht, dass β mit 95% Wahrscheinlichkeit innerhalb dieses Intervalls liegt!

Wahrscheinlichkeitsaussagen beziehen sich immer auf Zufallsvariablen und nicht auf ihre tatsächlichen Ausprägungen.

5.3.3 Aussagekraft von Intervallschätzern

- Der KQ-Punktschätzer $\hat{\beta}$ ist eine Mutmaßung bezüglich des wahren Wertes β .
- Der Intervallschätzer (5.19), dessen Zentrum $\hat{\beta}$ ist, liefert ein Signal für die Verlässlichkeit der Mutmaßung.
- Die Verlässlichkeit ist umso größer, je enger das Konfidenzintervall.
- Ein enges Konfidenzintervall erfordert einen genügend hohen Informationsgehalt in den Daten.

- **Frage:** Welchen Einfluss hat der Beobachtungsumfang T auf die Intervallbreite?

Antwort:

- **Frage:** Welchen Einfluss hat bei gegebener Standardabweichung $\widehat{sd}(\widehat{\beta})$ und gegebenen Freiheitsgraden $T-2$ das Signifikanzniveau α auf die Intervallbreite?

Antwort:

5.4 Intervallschätzer für α

- Ausgangspunkt ist die Gleichung

$$\Pr \{ \hat{\alpha} - k \leq \alpha \leq \hat{\alpha} + k \} = 1 - a .$$

- Folgt man den beschriebenen vier Schritten, so ergibt sich schließlich

$$\Pr \left\{ \hat{\alpha} - t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) \leq \alpha \leq \hat{\alpha} + t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) \right\} = 1 - a ,$$

wobei mit Blick auf Ergebnis (4.6b)

$$\widehat{sd}(\hat{\alpha}) = \sqrt{\widehat{\sigma}^2 (1/T + \bar{x}^2/S_{xx})} . \quad (5.21)$$

- Der Intervallschätzer für $\hat{\alpha}$ lautet somit

$$\left[\hat{\alpha} - t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) ; \hat{\alpha} + t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) \right] . \quad (5.22)$$

Numerische Illustration 5.9

Im Trinkgeld-Beispiel ist bereits bekannt, dass

$$\bar{x} = 30, \quad S_{xx} = 800 \quad \text{und} \quad \hat{\sigma}^2 = 1,5 .$$

Damit ergibt sich gemäß (5.21) eine geschätzte Standardabweichung von

$$\widehat{sd}(\hat{\alpha}) = \sqrt{1,5 (1/3 + 30^2/800)} \approx 1,4790 .$$

Der Wert von $t_{\alpha/2}$ beträgt weiterhin 12,7062 und der Schätzer $\hat{\alpha}$ besitzt gemäß (3.20) einen Wert von 0,25. Einsetzen in (5.22) liefert ein Konfidenzintervall von $[-18,5424 ; 19,0424]$.

Kapitel 6

Hypothesentest

6.1 Zweiseitiger Hypothesentest

- Jeder Test beginnt mit einer Ausgangsbehauptung, der sogenannten *Nullhypothese* $H_0 : \beta = q$.
- Eine *Alternativhypothese* H_1 besagt das exakte Gegenteil der Nullhypothese, also $H_1 : \beta \neq q$.
- Ein konkretes Beispiel: $H_0 : \beta = 0,7$ und $H_1 : \beta \neq 0,7$.
- Im Hypothesentest wird kontrolliert, ob $\hat{\beta}$ in der Nähe von q liegt
- Im Beispiel würde kontrolliert werden, ob $\hat{\beta}$ in der Nähe von $0,7$ liegt, z.B. innerhalb des Intervalls $[0,6 ; 0,8]$.
- Liegt $\hat{\beta}$ außerhalb des Intervalls, wird H_0 abgelehnt.

6.1.1 Ein grafisches Entscheidungsverfahren

- Vereinfachend wird zunächst angenommen, σ^2 und damit $var(\hat{\beta})$ wären bekannt.
- Wenn die Nullhypothese $H_0 : \beta = q$ wahr wäre (und auch σ^2 und damit $var(\hat{\beta})$ bekannt wären), dann würde man die entsprechende Wahrscheinlichkeitsverteilung des Schätzers $\hat{\beta}$ genau angeben können:

$$\hat{\beta} \sim N(q, var(\hat{\beta})) . \quad (6.1)$$

- In Abbildung 6.1 ist diese Normalverteilung dargestellt.

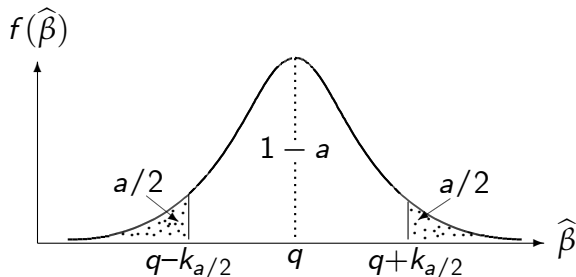


Abbildung 6.1: Die Wahrscheinlichkeitsverteilung des Schätzers $\hat{\beta}$ (gültig nur unter der Prämisse, dass $H_0 : \beta = q$ wahr ist).

- Die Wahrscheinlichkeit, dass *bei Gültigkeit der Nullhypothese* der Wert von $\hat{\beta}$ im Intervall $[q - k_{a/2}; q + k_{a/2}]$ liegt, beträgt $1 - a$.
- Formal:

$$\Pr \left\{ q - k_{a/2} \leq \hat{\beta} \leq q + k_{a/2} \right\} = 1 - a. \quad (6.2)$$

- Liegt $\hat{\beta}$ im Intervall $[q - k_{a/2}; q + k_{a/2}]$, dann wird die Nullhypothese nicht verworfen.
- Das Intervall $[q - k_{a/2}; q + k_{a/2}]$ wird als Akzeptanzintervall bezeichnet.
- Das Wort „Akzeptanz“ bedeutet nicht „wahr“ oder „korrekt“, sondern lediglich, dass es zu keiner Ablehnung der Nullhypothese gekommen ist.
- Passender wäre deshalb der Begriff „*Toleranzintervall*“.

- Liegt $\hat{\beta}$ im linken oder rechten Bereich außerhalb des Akzeptanzintervalls, dann lehnt man H_0 ab.
- **Frage:** Welche Rechtfertigung kann man für diese Ablehnung heranziehen?
Antwort:
- Man sagt dann, der geschätzte Wert $\hat{\beta}$ ist statistisch *signifikant* von dem Wert der Nullhypothese entfernt.

- Die Bereiche außerhalb des Akzeptanzintervalls werden hier als „*Ablehnungsbereiche*“ bezeichnet.
- Die Wahrscheinlichkeitsmasse in den Ablehnungsbereichen beträgt insgesamt immer α .
- Diese Wahrscheinlichkeitsmasse wird als das *Signifikanzniveau* des Hypothesentests bezeichnet.
- Es handelt sich um einen *zweiseitigen Hypothesentest* (Ablehnungsbereiche auf beiden Seiten).

- Abbildung 6.1 basiert auf der Annahme, dass H_0 wahr ist.
- **Frage:** Warum wird das Signifikanzniveau α oftmals als *Irrtumswahrscheinlichkeit* bezeichnet?
Antwort:

6.1.2 Ein analytisches Entscheidungsverfahren

- Da σ^2 nicht bekannt ist, wird nicht $\hat{\beta}$ mit $[q - k_{a/2} ; q + k_{a/2}]$ verglichen, sondern es wird ein „transformierter Vergleich“ angestellt.
- Bei diesem transformierten Vergleich wird t (das transformierte $\hat{\beta}$) mit $-t_{a/2}$ und $t_{a/2}$ (den transformierten Werten $q - k_{a/2}$ und $q + k_{a/2}$) verglichen.

Schritt 1: Formulierung von H_0 und H_1 und Festlegung des Signifikanzniveaus a

- Ausgangspunkt sind die Null- und Alternativhypothese:

$$H_0 : \beta = q \quad \text{und} \quad H_1 : \beta \neq q .$$

- Sie sollen auf einem Signifikanzniveau von a getestet werden.

Numerische Illustration 6.1

H_0 lautet im Trinkgeld-Beispiel: $\beta = 0,7$ und die Alternativhypothese H_1 lautet $\beta \neq 0,7$. Das Signifikanzniveau betrage $a = 0,05$.

Schritt 2: Schätzung von $sd(\hat{\beta})$

- Aus den Gleichungen (5.13) und (5.12) ist bereits bekannt, dass $\widehat{sd}(\hat{\beta})$ aus

$$\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}} \quad (6.3)$$

ermittelt werden kann, wobei

$$\hat{\sigma}^2 = S_{\widehat{uu}} / (T - 2) . \quad (6.4)$$

Numerische Illustration 6.2

Es liegen $T = 3$ Beobachtungen vor. Aus den Resultaten (3.21) und (3.17) ist bekannt, dass $S_{\hat{u}\hat{u}} = 1,5$ und $S_{xx} = 800$. Die Störgrößenvarianz σ^2 kann aus

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 2) = 1,5 / 1 = 1,5$$

geschätzt werden und die Standardabweichung des Schätzers $\hat{\beta}$ aus

$$\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}} = \sqrt{1,5 / 800} \approx 0,0433 . \quad (6.5)$$

Schritt 3: Ermittlung des t -Wertes aus der Stichprobe

- Die Zufallsvariable $\hat{\beta}$ wird transformiert:

$$t = \frac{\hat{\beta} - q}{\widehat{sd}(\hat{\beta})}. \quad (6.6)$$

- **Frage:** Warum kann t in Bezug auf die Nullhypothese H_0 als ein Widerspruchsmaß interpretiert werden und warum wäre $\hat{\beta} - q$ für sich noch kein sinnvolles Maß?

Antwort:

Numerische Illustration 6.3

Aus der Stichprobe ergab sich in (3.19) $\hat{\beta} = 0,125$ und Schritt 2 lieferte $\widehat{sd}(\hat{\beta}) \approx 0,0433$. Somit erhält man:

$$t = \frac{\hat{\beta} - q}{\widehat{sd}(\hat{\beta})} = \frac{0,125 - 0,7}{0,0433} \approx -13,2794 .$$

- Es kann gezeigt werden, dass *bei Gültigkeit der Nullhypothese* die standardisierte Zufallsvariable t einer t -Verteilung mit $T-2$ Freiheitsgraden folgt:

$$t \sim t_{(T-2)} .$$

Schritt 4: Ermittlung des kritischen Wertes $t_{a/2}$

- Für eine $t_{(T-2)}$ -verteilte Zufallsvariable lassen sich bei vorgegebenem Signifikanzniveau a feste Intervallgrenzen $-t_{a/2}$ und $t_{a/2}$ angeben, so dass:

$$\Pr \{ -t_{a/2} \leq t \leq t_{a/2} \} = 1 - a . \quad (6.7)$$

Numerische Illustration 6.4

Aus Tabelle T.2 können die kritischen Werte (Intervallgrenzen $-t_{a/2}$ und $t_{a/2}$) abgelesen werden.

Frage: Welche kritischen Werte ergeben sich im Hypothesentest des Trinkgeld-Beispiels?

Antwort:

Numerische Illustration 6.5

Frage: Der mit Formel (6.6) berechnete Wert t ist bei Gültigkeit der Nullhypothese $H_0 : \beta = 0,7$ eine $t_{(T-2)}$ -verteilte Zufallsvariable. Wie sollte Gleichung (6.7) in Bezug auf den Test der Nullhypothese folglich interpretiert werden?

Antwort:

Schritt 5: Vergleich von $t_{\alpha/2}$ und t

- Falls $|t| > t_{\alpha/2}$, wird die Nullhypothese verworfen.

Numerische Illustration 6.6

Frage: Zu welcher Testentscheidung gelangt man im Hypothesentest des Trinkgeld-Beispiels?

Antwort:

- Die fünf Schritte eines zweiseitigen t -Tests lassen sich folgendermaßen zusammenfassen:

t -Test

Um die Nullhypothese $H_0 : \beta = q$ auf einem Signifikanzniveau von α zu testen, wird der t -Wert

$$t = \frac{\hat{\beta} - q}{\widehat{sd}(\hat{\beta})}$$

berechnet, wobei $\hat{\beta} = S_{xy} / S_{xx}$ und $\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}}$ mit $\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 2)$. Liegt der t -Wert außerhalb des Akzeptanzintervalls $[-t_{\alpha/2}, t_{\alpha/2}]$, dann wird H_0 verworfen.

6.1.3 Zusammenhang zwischen analytischem und grafischem Vorgehen

- Der Vergleich von $\hat{\beta}$ mit den Intervallgrenzen $q - t_{a/2} \cdot \widehat{sd}(\hat{\beta})$ bzw. $q + t_{a/2} \cdot \widehat{sd}(\hat{\beta})$ und der Vergleich von t mit den Intervallgrenzen $-t_{a/2}$ und $t_{a/2}$ sind letztlich äquivalent.
- Einsetzen des Ausdrucks (6.6) in Gleichung (6.7) liefert:

$$\Pr \left\{ -t_{a/2} \leq (\hat{\beta} - q) / \widehat{sd}(\hat{\beta}) \leq t_{a/2} \right\} = 1 - a .$$

$$\Rightarrow \Pr \left\{ q - t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \leq \hat{\beta} \leq q + t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \right\} = 1 - a. \quad (6.8)$$

- Liegt $\hat{\beta}$ außerhalb des Akzeptanzintervalls $\left[q - t_{a/2} \cdot \widehat{sd}(\hat{\beta}) ; q + t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \right]$, wird H_0 verworfen.
- Dies ergibt sich genau dann, wenn t außerhalb des Intervalls $[-t_{a/2} ; t_{a/2}]$ liegt, denn (6.8) und (6.7) sind unmittelbar ineinander überführbar.

Numerische Illustration 6.7

Als Akzeptanzintervall ergibt sich hier $[0,1498 ; 1,2502]$. Da der in der Stichprobe beobachtete Wert von $\hat{\beta} = 0,125$ außerhalb dieses Intervalls liegt, wird $H_0 : \beta = 0,7$ auch auf Basis des Vergleiches von $\hat{\beta}$ mit $q = 0,7$ abgelehnt.

- Das Zentrum des Akzeptanzintervalls $\left[q - t_{a/2} \cdot \widehat{sd}(\widehat{\beta}) ; q + t_{a/2} \cdot \widehat{sd}(\widehat{\beta}) \right]$ ist q . Es ist durch H_0 festgelegt.
- **Frage:** Warum ist die Breite des Akzeptanzintervalls aber zufallsabhängig?
Antwort:
- **Frage:** Wie muss Gleichung (6.8) folglich interpretiert werden?
Antwort:

6.1.4 Zusammenhang zwischen zweiseitigem Hypothesentest und Intervallschätzer

- Es besteht ein enger Zusammenhang zwischen dem t -Test einer Nullhypothese der Form $H_0 : \beta = q$ und dem mittels Intervallschätzer berechneten Konfidenzintervall.
- Löst man den Ausdruck in den geschweiften Klammern der Gleichung (6.8) nach q auf, so ergibt sich

$$\Pr \left\{ \hat{\beta} - t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \leq q \leq \hat{\beta} + t_{a/2} \cdot \widehat{sd}(\hat{\beta}) \right\} = 1 - a .$$

- **Frage:** Wie ist diese Gleichung zu interpretieren?
Antwort:

- Es genügt also, den Wert q der Nullhypothese mit dem im Rahmen der Intervallschätzung berechneten Konfidenzintervall zu vergleichen.
- Liegt q außerhalb des Konfidenzintervalls, dann wird die Nullhypothese verworfen.

Numerische Illustration 6.8

An früherer Stelle ergab sich für β für ein Signifikanzniveau von 5% das Konfidenzintervall $[-0,4252 ; 0,6752]$.

Der in der Nullhypothese $H_0 : \beta = 0,7$ postulierte Wert wird von diesem Intervall nicht abgedeckt.

Die Nullhypothese wird folglich abgelehnt.

Intervallschätzer und Hypothesentest

Jede Nullhypothese $H_0 : \beta = q$, deren q -Wert nicht vom berechneten Konfidenzintervall abgedeckt wird, wird verworfen. Wird der q -Wert hingegen abgedeckt, wird die Nullhypothese nicht verworfen.

6.2 Einseitiger Hypothesentest

- Anstelle von Nullhypothesen der Form $H_0 : \beta = q$ können auch Nullhypothesen der Form $H_0 : \beta \geq q$ oder $H_0 : \beta \leq q$ getestet werden.
- Es handelt sich dann um einen einseitigen Hypothesentest.

6.2.1 Ein grafisches Entscheidungsverfahren

- Abbildung 6.2 illustriert den Test einer Nullhypothese der Form

$$H_0 : \beta \leq q \quad (\text{und damit } H_1 : \beta > q) .$$

- Da der Ablehnungsbereich rechts ist, heißt ein solcher Test *rechtsseitiger* Hypothesentest.
- Im Vergleich zu einem zweiseitigen Hypothesentest muss bei gleichem Signifikanzniveau a (Irrtumswahrscheinlichkeit) die Wahrscheinlichkeitsmasse am rechten Rand verdoppelt werden, also $q + k_a < q + k_{a/2}$.

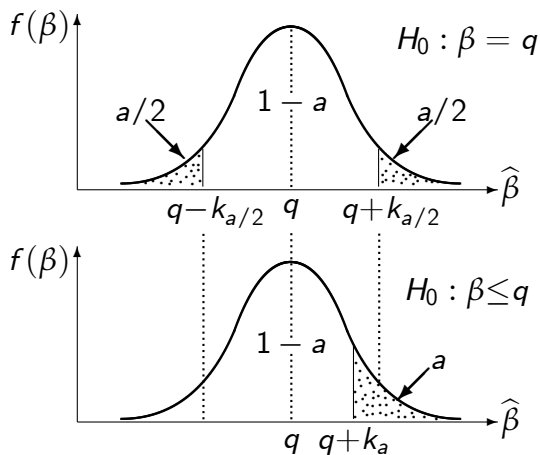


Abbildung 6.2: Zweiseitiger und rechtsseitiger Test mit identischem Signifikanzniveau a und folglich mit unterschiedlichen kritischen Werten.

6.2.2 Ein analytisches Entscheidungsverfahren

- Formal liegt dem *rechtsseitigen* Test ($H_0 : \beta \leq q$) folgende Beziehung zugrunde:

$$\Pr \left\{ \hat{\beta} \leq q + k_a \right\} = 1 - a . \quad (6.9)$$

- **Frage:** Wie ist Gleichung (6.9) zu interpretieren?
Antwort:

Schritt 1: Formulierung von H_0 und H_1 und Festlegung des Signifikanzniveaus α

- Im *rechtsseitigen* Test lauten die Null- und Alternativhypothese

$$H_0 : \beta \leq q \quad \text{und} \quad H_1 : \beta > q .$$

- Sie sollen auf einem Signifikanzniveau von α getestet werden.

Schritt 2: Schätzung von $sd(\hat{\beta})$

- Das Vorgehen ist identisch mit demjenigen des zweiseitigen Tests.

Schritt 3: Ermittlung des t -Wertes aus der Stichprobe

- Den t -Wert der Stichprobe errechnet man wie gewohnt aus

$$t = \frac{\hat{\beta} - q}{\widehat{sd}(\hat{\beta})}, \quad (6.10)$$

wobei q die in der Nullhypothese $H_0 : \beta \leq q$ formulierte Grenze zwischen H_0 und H_1 darstellt.

- Aus der Statistik ist bekannt, dass bei Gültigkeit von $\beta = q$:

$$t \sim t_{(T-2)} .$$

Schritt 4: Ermittlung des kritischen Wertes t_a

- Für eine $t_{(T-2)}$ -verteilte Zufallsvariable t kann bei vorgegebenem Signifikanzniveau a eine feste Intervallgrenze t_a angegeben werden, so dass:

$$\Pr \{t \leq t_a\} = 1 - a . \quad (6.11)$$

Schritt 5: Vergleich von t_a und t

- Fällt t größer als t_a aus, dann wird die Nullhypothese verworfen.

Numerische Illustration 6.9

Der Ökonometriker vermute, dass der Rechnungsbetrag einen positiven Einfluss auf das Trinkgeld ausübe. Wir verwenden das Gegenteil der Vermutung als Nullhypothese.

Schritt 1: $H_0 : \beta \leq 0$ und $H_1 : \beta > 0$ und Signifikanzniveau $\alpha = 5\%$.

Schritt 2: In Gleichung (6.5) wurde $\widehat{sd}(\widehat{\beta}) \approx 0,0433$ berechnet.

Schritt 3: Der t -Wert der Stichprobe errechnet sich aus

$$t = \frac{\widehat{\beta} - q}{\widehat{sd}(\widehat{\beta})} = \frac{0,125 - 0}{0,0433} \approx 2,8868 . \quad (6.12)$$

(Forts.)

Schritt 4: Aus Tabelle T.2 kann der t_a -Wert ermittelt werden.

Schritt 5: Der t -Wert wird mit dem t_a -Wert verglichen.

Frage: Zu welchem Testergebnis gelangen wir?

Antwort:

Frage: Zu welchem Testergebnis wären wir gelangt, wenn die Ergebnisse von $\hat{\beta} = 0,125$ und $\widehat{sd}(\hat{\beta}) \approx 0,0433$ auf einer Basis von 20 Beobachtungen (18 Freiheitsgraden) zustande gekommen wären?

Antwort:

- Die fünf Schritte eines einseitigen t -Tests lassen sich folgendermaßen zusammenfassen:

rechtsseitiger t -Test

Um die Nullhypothese $H_0 : \beta \leq q$ auf einem Signifikanzniveau von α zu testen (rechtsseitiger t -Test), wird der t -Wert

$$t = \frac{\hat{\beta} - q}{\widehat{sd}(\hat{\beta})}$$

berechnet, wobei $\hat{\beta} = S_{xy} / S_{xx}$ und $\widehat{sd}(\hat{\beta}) = \sqrt{\hat{\sigma}^2 / S_{xx}}$ mit $\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 2)$. Liegt dieser t -Wert außerhalb des Akzeptanzintervalls $(-\infty, t_\alpha]$, dann wird H_0 verworfen.

linksseitiger t -Test

Um die Nullhypothese $H_0 : \beta \geq q$ zu testen (linksseitiger t -Test), wird der t -Wert mit dem Akzeptanzintervall $[-t_a, \infty)$ verglichen.

6.3 p -Wert

- Es wird der rechtsseitige Hypothesentest $H_0 : \beta \leq q$ betrachtet.

Traditionelles Vorgehen

- Für die Ablehnung der Nullhypothese muss $t > t_a$ gelten.
- **Frage:** Zu welcher Testentscheidung gelangt man im oberen Teil von Abbildung 6.3?

Antwort:

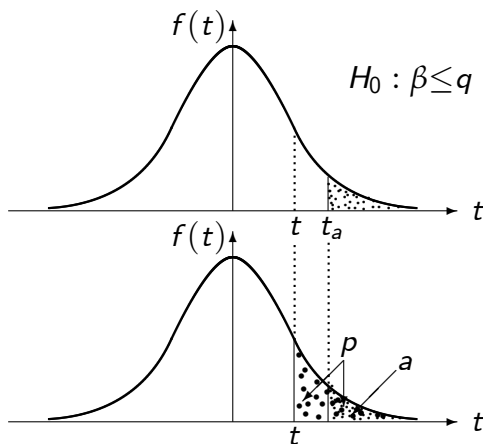


Abbildung 6.3: Die Gleichwertigkeit von Tests auf Basis des t - und p -Wertes.

Definition: p -Wert

Der p -Wert eines Tests von $H_0 : \beta \leq q$ ist die Wahrscheinlichkeitsmasse rechts vom berechneten t -Wert.

Der p -Wert eines Tests von $H_0 : \beta \geq q$ ist die Wahrscheinlichkeitsmasse links vom berechneten t -Wert.

Alternatives Vorgehen

- Für die Ablehnung der Nullhypothese muss $p < a$ gelten.
- Dies ist ein Vergleich von Wahrscheinlichkeitsmassen.
- **Frage:** Warum führen beide Vorgehen immer zur gleichen Testentscheidung?

Antwort:

Numerische Illustration 6.10

Im Trinkgeld-Beispiel lautete die Nullhypothese $H_0 : \beta \leq 0$.
Aus der Stichprobe wurde in (3.19) und (6.5) errechnet, dass $\hat{\beta} = 0,125$ und $\widehat{sd}(\hat{\beta}) \approx 0,0433$.

Daraus ergab sich in (6.12) ein t -Wert von 2,8868.

Es war $t_a = 6,3138$.

Ökonometrische Standard-Software liefert die Wahrscheinlichkeitsmasse (p -Wert), die außerhalb des durch den t -Wert definierten Intervalls $(-\infty ; 2,8868]$ liegt.

Für die betrachtete Stichprobe ergibt sich $p = 10,6\%$ und damit $p > a$.

Die Nullhypothese kann nicht verworfen werden.

6.4 Wahl der geeigneten Nullhypothese und des geeigneten Signifikanzniveaus

- Im Trinkgeld-Beispiel laute die Anfangsvermutung:
Der Rechnungsbetrag übt einen positiven Einfluss auf die Höhe des Trinkgelds aus ($\beta > 0$).
- Für die Untermauerung dieser Anfangsvermutung stehen zwei Strategien zur Verfügung:
 - A) Das Gegenteil der Vermutung wird als Nullhypothese verwendet ($H_0 : \beta \leq 0$).
 - B) Die Vermutung wird als Nullhypothese verwendet ($H_0 : \beta \geq 0$).

6.4.1 Strategie A: Nullhypothese behauptet Gegenteil der Anfangsvermutung

- Es wird $H_0 : \beta \leq 0$ getestet.
- Dies ist ein rechtsseitiger Test.
- Vier Fälle sind für das Ergebnis des Hypothesentests denkbar:
 - 1 Korrekte Ablehnung
 - 2 Irrtümliche Ablehnung
 - 3 Korrekte Akzeptanz
 - 4 Irrtümliche Akzeptanz

- **Frage:** Gegeben $H_0 : \beta \leq 0$ ist wahr (im Sinne von $\beta = 0$), wie groß ist die Wahrscheinlichkeit, dass im Hypothesentest der Fall 2 eintritt (siehe kritischer Wert k in Abbildung 6.4)?

Antwort:

- Das Signifikanzniveau α wird auch als *Typ I-Fehlerrisiko* (oder auch als α -Fehler) bezeichnet.

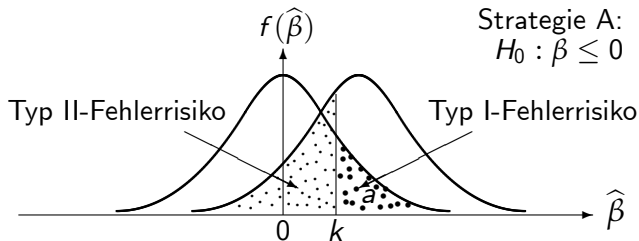


Abbildung 6.4: Die Bedeutung des Typ I-Fehlerrisikos (Signifikanzniveau α) in Teststrategie A.

- Es kann im Hypothesentest auch zur „irrtümlichen Akzeptanz“ der Nullhypothese $H_0 : \beta \leq 0$ kommen.
- **Frage:** Die rechte Wahrscheinlichkeitsverteilung in Abbildung 6.4 sei korrekt, also $H_0 : \beta \leq 0$ sei nicht wahr. Warum ist das eingezeichnete *Typ II-Fehlerrisiko* die Wahrscheinlichkeit, dass im Hypothesentest eine irrtümliche Akzeptanz eintritt?

Antwort:

- Es besteht ein Zielkonflikt zwischen niedrigem Typ I-Fehlerrisiko und niedrigem Typ II-Fehlerrisiko.
- **Frage:** Auf welches Fehlerrisiko sollte im Kontext der Strategie A mehr geachtet werden?

Antwort:

- **Frage:** Was bedeutet dies für die Wahl des Signifikanzniveaus?

Antwort:

6.4.2 Strategie B: Nullhypothese stimmt mit Anfangsvermutung überein

- Nun wird $H_0 : \beta \geq 0$ getestet.
- Dies ist ein linksseitiger Test.
- Auch hier sind die vier Fälle für das Ergebnis des Hypothesentests denkbar:
 - ① Korrekte Ablehnung
 - ② Irrtümliche Ablehnung
 - ③ Korrekte Akzeptanz
 - ④ Irrtümliche Akzeptanz

- **Frage:** Gegeben $H_0 : \beta \geq 0$ ist wahr (im Sinne von $\beta = 0$), wie groß ist die Wahrscheinlichkeit, dass im Hypothesentest eine irrtümliche Ablehnung eintritt (siehe kritischer Wert k in Abbildung 6.5)?

Antwort:

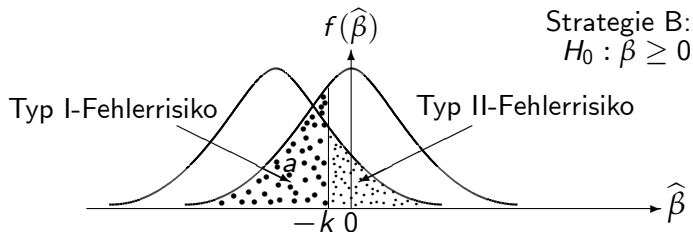


Abbildung 6.5: Die Bedeutung des Typ I-Fehlerrisikos (Signifikanzniveau α) in Teststrategie B.

- Es kann im Hypothesentest auch zur „irrtümlichen Akzeptanz“ der Nullhypothese $H_0 : \beta \geq 0$ kommen.
- **Frage:** Die linke Wahrscheinlichkeitsverteilung in Abbildung 6.5 sei korrekt, also $H_0 : \beta \geq 0$ sei nicht wahr. Warum ist das eingezeichnet Typ II-Fehlerrisiko die Wahrscheinlichkeit dafür, dass im Hypothesentest eine irrtümliche Akzeptanz eintritt?

Antwort:

- **Frage:** Auf welches Fehlerrisiko sollte im Kontext der Strategie B mehr geachtet werden?

Antwort:

- **Frage:** Welche Probleme bestehen bei Strategie B?

Antwort:

6.4.3 Trennschärfe von Tests

- Es gibt einen Zielkonflikt zwischen Typ I-Fehlerrisiko und Typ II-Fehlerrisiko.
- Die „Gesamtverlässlichkeit“ eines Tests wird als *Trennschärfe* (power) bezeichnet.
- Gemeint ist die „Ablehnungskraft“ im Sinne von Strategie A.
- Zusätzliche Informationen erhöhen die Trennschärfe eines Tests (kleineres Typ I-Fehlerrisiko bei gleichbleibendem Typ II-Fehlerrisiko oder auch umgekehrt).

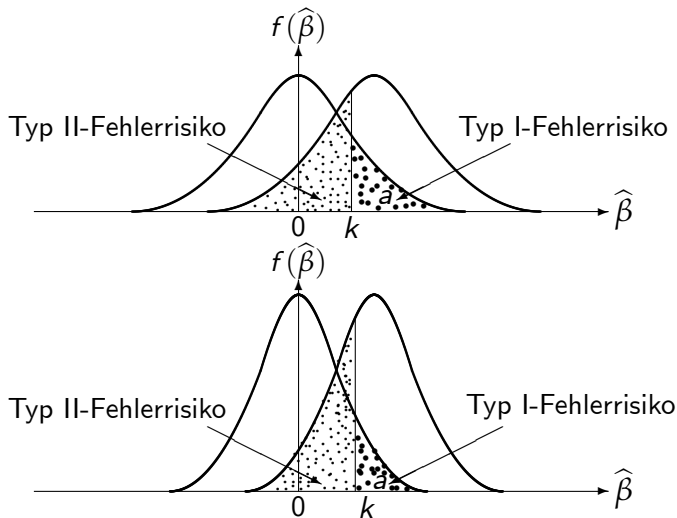


Abbildung 6.6: Senkung des Typ I-Fehlerrisikos bei konstantem Typ II-Fehlerrisiko.

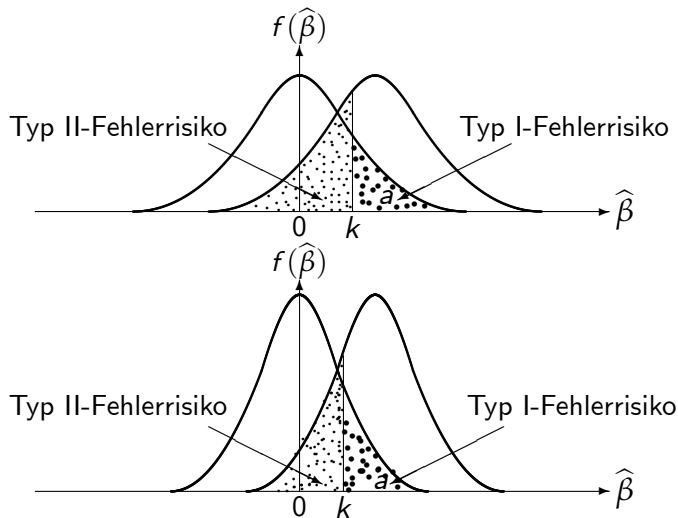


Abbildung 6.7: Senkung des Typ II-Fehlerrisikos bei konstantem Typ I-Fehlerrisiko.

6.4.4 Anmerkungen zu zweiseitigen Tests

- Ein Test mit $H_0 : \beta \neq q$ ist nicht möglich (egal ob die Anfangsvermutung $\beta = q$ oder $\beta \neq q$ lautet).
- Es kann immer nur $H_0 : \beta = q$ getestet werden.
- Lautet die Anfangsvermutung $\beta \neq q$, so läuft ein Test von $H_0 : \beta = q$ auf Strategie A hinaus (kleines Signifikanzniveau wählen).
- Lautet die Anfangsvermutung $\beta = q$, dann entspricht der Test von $H_0 : \beta = q$ der Strategie B (großes Signifikanzniveau wählen).

Kapitel 7

Prognose

7.1 Punktprognose

7.1.1 Berechnung der Punktprognose

- Es werden hier nur „*bedingte*“ Prognosen betrachtet.
- **Frage:** Warum ist die Basis für eine bedingte Prognose immer das geschätzte Modell:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t .$$

Antwort:

Numerische Illustration 7.1

Das geschätzte Modell lautet

$$\hat{y}_t = 0,25 + 0,125 \cdot x_t .$$

Frage: Welchen Wert besitzt die bedingte Punktprognose (\hat{y}_0) für $x_0 = 20$?

Antwort:

- Das allgemeine Vorgehen lautet: In das geschätzte Modell

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$$

setzt man einen bestimmten Wert x_0 für x_t ein und erhält damit die *Punktprognose*:

$$\hat{y}_0 = \hat{\alpha} + \hat{\beta}x_0 . \quad (7.1)$$

- Der tatsächliche Wert von y_0 wird aber nicht \hat{y}_0 , sondern

$$y_0 = \alpha + \beta x_0 + u_0$$

sein.

- Der Prognosefehler lässt sich gemäß

$$\begin{aligned}\hat{y}_0 - y_0 &= (\hat{\alpha} + \hat{\beta}x_0) - (\alpha + \beta x_0 + u_0) \\ &= (\hat{\alpha} - \alpha) + (\hat{\beta} - \beta)x_0 - u_0\end{aligned}\quad (7.2)$$

berechnen.

- Es kann demnach zwei Ursachen für einen Prognosefehler geben:
 - 1 **Stichprobenfehler:** Die Parameterschätzungen $\hat{\alpha}$ und $\hat{\beta}$ können von den wahren Werten α und β abweichen.
 - 2 **Störgrößenfehler:** Die Störgröße u_0 kann einen von 0 verschiedenen Wert annehmen.

7.1.2 Verlässlichkeit der Punktprognose

- Statt die Wahrscheinlichkeitsverteilung der Zufallsvariable \hat{y}_0 zu verwenden, wird üblicherweise die Wahrscheinlichkeitsverteilung des Prognosefehlers $(\hat{y}_0 - y_0)$ benutzt.
- Gemäß Gleichung (7.2) gilt für den Erwartungswert des Prognosefehlers:

$$\begin{aligned} E(\hat{y}_0 - y_0) &= E\left((\hat{\alpha} - \alpha) + (\hat{\beta} - \beta)x_0 - u_0\right) \\ &= E(\hat{\alpha}) - E(\alpha) + E(\hat{\beta}x_0) - E(\beta x_0) - E(u_0) \\ &= \alpha - \alpha + \beta x_0 - \beta x_0 - 0 = 0. \end{aligned} \quad (7.3)$$

- Es kann gezeigt werden (Anhang des Lehrbuches), dass die Varianz des Prognosefehlers

$$\text{var}(\hat{y}_0 - y_0) = \sigma^2 + \sigma^2 \left(\frac{1}{T} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right) \quad (7.4)$$

beträgt.

- Sie setzt sich aus dem Störgrößenfehler und dem Stichprobenfehler zusammen.
- Da auch hier wieder die Störgrößenvarianz σ^2 nicht bekannt ist, muss man sie durch den Schätzer $\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T-2)$ ersetzen:

$$\widehat{\text{var}}(\hat{y}_0 - y_0) = \hat{\sigma}^2 \left(1 + \frac{1}{T} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right), \quad (7.5)$$

wobei $\hat{\sigma}^2$ ausgeklammert wurde.

Numerische Illustration 7.2

Für $x_0 = 20$ wurde $\hat{y}_0 = 2,75$ Euro prognostiziert.

In Gleichung (3.21) wurde eine Störgrößenvarianz von $\hat{\sigma}^2 = 1,5$ geschätzt.

In Resultat (3.17) wurde $S_{xx} = 800$ berechnet und \bar{x} betrug 30.

Demnach ergibt sich die geschätzte Varianz des Prognosefehlers gemäß Formel (7.5) aus

$$\widehat{\text{var}}(\hat{y}_0 - y_0) = 1,5 \left[1 + 1/3 + (20 - 30)^2 / 800 \right] = 2,1875 . \quad (7.6)$$

Für $x_0 = 5$ ergäbe sich hingegen

$$\widehat{\text{var}}(\hat{y}_0 - y_0) = 1,5 \left[1 + 1/3 + (5 - 30)^2 / 800 \right] = 3,1719 .$$

7.2 Prognoseintervall

- Es soll ein Prognoseintervall für \hat{y}_0 berechnet werden.
- Es taucht erneut das Problem auf, dass σ^2 unbekannt ist.
- Aus Gleichung (7.4) wird offensichtlich, dass damit auch

$$sd(\hat{y}_0 - y_0) = \sqrt{\text{var}(\hat{y}_0 - y_0)}$$

unbekannt ist und geschätzt werden muss.

Schritt 1: Schätzung von $sd(\hat{y}_0 - y_0)$ und Festlegung des Signifikanzniveaus

- σ^2 wird durch den mit Formel (5.12) berechneten Schätzer $\hat{\sigma}^2$ ersetzt.
- Aus Gleichung (7.5) erhält man dann $\widehat{var}(\hat{y}_0 - y_0)$ und die Wurzel daraus entspricht $\widehat{sd}(\hat{y}_0 - y_0)$.
- Ferner wird das Signifikanzniveau α festgelegt.

Schritt 2: Standardisierung des Prognosefehlers ($\hat{y}_0 - y_0$)

- Der Prognosefehler ($\hat{y}_0 - y_0$) lässt sich wie gewohnt standardisieren:

$$t = \frac{(\hat{y}_0 - y_0) - \overbrace{E(\hat{y}_0 - y_0)}^{=0}}{\widehat{sd}(\hat{y}_0 - y_0)} = \frac{\hat{y}_0 - y_0}{\widehat{sd}(\hat{y}_0 - y_0)}. \quad (7.7)$$

- Die transformierte Zufallsvariable t folgt einer t -Verteilung mit $T-2$ Freiheitsgraden: $t \sim t_{(T-2)}$.

Schritt 3: Ermittlung des $t_{a/2}$ -Wertes

- Aus der Statistik ist bekannt, dass sich für eine $t_{(T-2)}$ -verteilte Zufallsvariable t für vorgegebenes Signifikanzniveau a feste Intervallgrenzen $-t_{a/2}$ und $t_{a/2}$ angeben lassen (siehe Tabelle T.2), so dass

$$\Pr \{ -t_{a/2} \leq t \leq t_{a/2} \} = 1 - a .$$

Schritt 4: Formulierung des Prognoseintervalls

- Der Wert von t fällt mit einer Wahrscheinlichkeit von $1 - a$ in das Konfidenzintervall $[-t_{a/2} ; t_{a/2}]$, also

$$\Pr \left\{ -t_{a/2} \leq \frac{\hat{y}_0 - y_0}{\widehat{sd}(\hat{y}_0 - y_0)} \leq t_{a/2} \right\} = 1 - a .$$

- Durch Auflösen nach y_0 ergibt sich:

$$\Pr \left\{ \hat{y}_0 - t_{a/2} \cdot \widehat{sd}(\hat{y}_0 - y_0) \leq y_0 \leq \hat{y}_0 + t_{a/2} \cdot \widehat{sd}(\hat{y}_0 - y_0) \right\} = 1 - a . \quad (7.8)$$

- Das gesuchte Prognoseintervall lautet folglich:

$$\left[\hat{y}_0 - t_{a/2} \cdot \widehat{sd}(\hat{y}_0 - y_0) ; \hat{y}_0 + t_{a/2} \cdot \widehat{sd}(\hat{y}_0 - y_0) \right] . \quad (7.9)$$

- Grafisch lässt sich für ein vorgegebenes Signifikanzniveau a und einen gegebenen Beobachtungsumfang T die Gestalt der Prognoseintervalle wie in Abbildung 7.1 beschreiben.

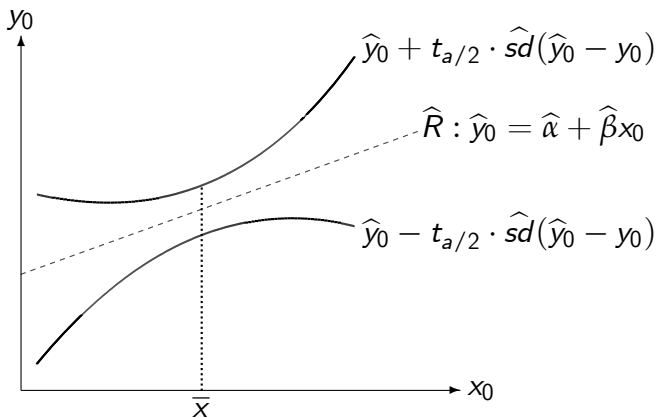


Abbildung 7.1: Die Abhängigkeit der Breite des Prognoseintervalls vom Abstand $(x_0 - \bar{x})$.

Numerische Illustration 7.3

Es wurde in Gleichung (7.6)

$$\widehat{var}(\widehat{y}_0 - y_0) = 2,1875$$

ermittelt. Die geschätzte Standardabweichung beträgt somit

$$\widehat{sd}(\widehat{y}_0 - y_0) = \sqrt{2,1875} \approx 1,4790 .$$

Bei einem Signifikanzniveau von $\alpha = 5\%$ und einem einzigen Freiheitsgrad ergibt sich aus der Tabelle T.2 ein kritischer Wert von $t_{\alpha/2} = 12,7062$. Bei einem bedingenden Wert von $x_0 = 20$ lautete die Punktprognose: $\widehat{y}_0 = 2,75$. Einsetzen dieser Werte in Formel (7.9) ergibt ein Prognoseintervall von

$$[-16,0425 ; 21,5425] .$$

Kapitel 8

Spezifikation

Beispiel Dünger

Auf einem Testareal mit 30 homogenen Beobachtungspartellen wurde ein Experiment durchgeführt, bei dem der Gerstenoutput für verschiedene Kombinationen von 'Phosphat' und 'Stickstoff' beobachtet wurde.

Tabelle 8.1: Phosphat p_t (in kg/ha), Stickstoff n_t (in kg/ha) und Gerstenoutput g_t (in dt/ha) auf 30 beobachteten Partellen.

t	p_t	n_t	g_t	t	p_t	n_t	g_t
1	22,00	40,00	38,36	16	25,00	110,00	59,55
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	25,00	100,00	57,27	30	29,00	100,00	64,39

- Das *ökonomische Modell* formalisiert den grundlegenden Wirkungszusammenhang zwischen Gersten-Output (g) und Düngemitteln (p und n):

$$g = f(p, n).$$

8.1 A-Annahmen

8.1.1 Erster Schritt

- Ein möglicher ökonomischer Wirkungszusammenhang:

$$g = \alpha + \beta_1 p + \beta_2 n .$$

- **Frage:** Warum ist diese Spezifikation vermutlich keine gute Beschreibung der Realität?

Antwort:

- Realistischer ist

$$g = Ap^{\beta_1} n^{\beta_2} , \quad (8.1)$$

wobei A , β_1 , und β_2 drei konstante Parameter sind.

- Logarithmieren beider Seiten liefert

$$\begin{aligned} \ln g &= \ln \left(Ap^{\beta_1} n^{\beta_2} \right) \\ &= \ln A + \ln \left(p^{\beta_1} \right) + \ln \left(n^{\beta_2} \right) \\ &= \ln A + \beta_1 \ln p + \beta_2 \ln n . \end{aligned}$$

- Definiert man $y = \ln g$, $x_1 = \ln p$ und $x_2 = \ln n$, ergibt sich das vertraute lineare Modell

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 , \quad (8.2)$$

wobei $\alpha = \ln A$.

- Das grafische Pendant der Gleichung (8.2) ist die „wahre Ebene“ R der Abbildung 8.1.
- **Frage:** Welches Vorzeichen besitzen in Abbildung 8.1 die Parameter α , β_1 und β_2 ?

Antwort:

- Gleichung (8.2) und damit die „wahre Ebene“ R beschreiben einen linearen Zusammenhang zwischen den „logarithmierten Düngermittelmengen“ und dem „logarithmierten Gersten-Output“.

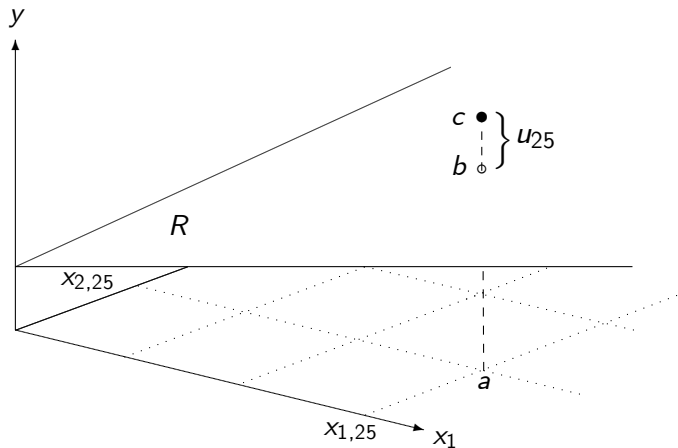


Abbildung 8.1: Die wahre Ebene für den Zusammenhang zwischen logarithmierten Düngemittelmengen (x_1 und x_2) und logarithmiertem Gersten-Output (y).

- Die zu verwendenden Daten sind deshalb nicht mehr p_t , n_t und g_t , sondern ihre logarithmierten Werte:

Tabelle 8.2: Logarithmierte Daten.

t	x₁ (= $\ln p_t$)	x₂ (= $\ln n_t$)	y_t (= $\ln g_t$)
1	3,0910	3,6889	3,6470
2	3,0910	4,0943	3,8924
⋮	⋮	⋮	⋮
30	3,3673	4,6052	4,1650

8.1.2 Zweiter und dritter Schritt

- Die Daten liegen als Datentripel (x_{1t}, x_{2t}, y_t) vor mit $t = 1, 2, \dots, 30$.
- Die Daten sind in Abbildung 8.2 als Punktwolke wiedergegeben.
- Die Punkte liegen nicht auf der wahren Ebene R .
- Es muss deshalb eine Störgröße u_t zum Modell hinzugefügt werden.
- Das ökonometrische Modell lautet dann:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t, \quad (8.3)$$

für $t = 1, 2, \dots, 30$.

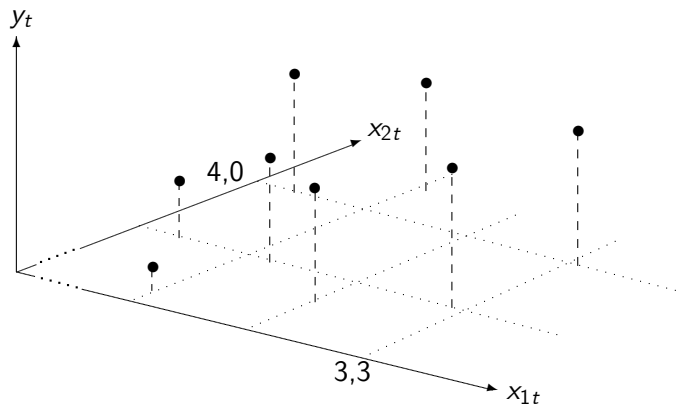


Abbildung 8.2: Eine Punktwolke im dreidimensionalen Raum.

- **Frage:** Was wird in Abbildung 8.1 anhand der Beobachtung $t = 25$ veranschaulicht?

Antwort:

- K steht immer für die Anzahl der exogenen Variablen.

Definition: Multiples Regressionsmodell

Das multiple Regressionsmodell lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t . \quad (8.4)$$

- Folglich beträgt die Zahl der Parameter des multiplen Regressionsmodells immer $K + 1$ (K Steigungsparameter β_k und der Niveauparameter α).
- Wenn im multiplen Regressionsmodell $K = 1$ gilt, dann ergibt sich das einfache Regressionsmodell.
- Bei $K = 2$ ergibt sich das Modell (8.3).

8.1.3 Formulierung der A-Annahmen

- Die funktionale Spezifikation vollzieht sich wie gewohnt im Rahmen der A-Annahmen.

Annahme A1

Im ökonometrischen Modell fehlen keine relevanten exogenen Variablen und die benutzten exogenen Variablen sind nicht irrelevant.

Annahme A2

Der wahre Zusammenhang zwischen $x_{1t}, x_{2t}, \dots, x_{Kt}$ und y_t ist linear.

Annahme A3

Die $K + 1$ Parameter $\alpha, \beta_1, \beta_2, \dots, \beta_K$ sind für alle T Beobachtungen $(x_{1t}, x_{2t}, \dots, x_{Kt}, y_t)$ konstant.

- Im Folgenden wird unterstellt, dass die A-Annahmen im Dünger-Beispiel erfüllt sind, also das ökonometrische Modell (8.3) funktional korrekt ist.

8.2 B-Annahmen

8.2.1 Formulierung der B-Annahmen

- Die Spezifikation der Störgrößen vollzieht sich wie gewohnt im Rahmen der B-Annahmen.

Annahme B1 (Erwartungswert)

Die Störgröße u_t hat für alle Beobachtungen t einen Erwartungswert von 0, das heißt,

$$E(u_t) = 0, \quad (8.5)$$

für $t = 1, 2, \dots, T$.

Annahme B2 (Homoskedastizität)

Die Störgröße u_t hat für alle Beobachtungen t eine konstante Varianz, das heißt,

$$\text{var}(u_t) = \sigma^2, \quad (8.6)$$

für $t = 1, 2, \dots, T$.

Annahme B3 (Freiheit von Autokorrelation)

Die Störgrößen sind nicht miteinander korreliert, das heißt,

$$\text{cov}(u_t, u_s) = 0, \quad (8.7)$$

für alle $t \neq s$ sowie $t = 1, 2, \dots, T$ und $s = 1, 2, \dots, T$.

Annahme B4 (Normalverteilung)

Die Störgröße u_t ist normalverteilt, das heißt

$$u_t \sim UN(0, \sigma^2) ,$$

für $t = 1, 2, \dots, T$.

- Zusammengenommen bedeuten die Annahmen B1 bis B4:

$$u_t \sim UN(0, \sigma^2) , \quad \text{für alle } t = 1, 2, \dots, T. \quad (8.8)$$

8.2.2 Interpretation der B-Annahmen

- Wiederholte Stichproben würden bedeuten, dass man für jede Düngerdosierung und damit für jede Parzelle t eine eigene Sequenz von Störgrößen u_t generiert.
- Die Annahmen B1, B2 und B4 beziehen sich auf die statistischen Eigenschaften dieser Sequenzen.
- In Annahme B3 geht es um die Beziehung zwischen den Sequenzen.

- **Frage:** Nennen Sie Gründe dafür, dass die Annahmen B1, B2 und B3 im Falle des Dünger-Beispiels möglicherweise nicht erfüllt sind.

Antwort:

- Wenn in der tatsächlichen und in den wiederholten Stichproben für jede Düngerdosierung die jeweilige Parzelle zufällig aus den 30 Parzellen ausgelost werden würde, dann könnte das Risiko einer Verletzung der Annahmen B1 und B2 weitgehend vermieden werden.
- Im Folgenden wird angenommen, dass im Dünger-Beispiel alle B-Annahmen erfüllt sind.

8.3 C-Annahmen

- Die Spezifikation der exogenen Variablen vollzieht sich wie gewohnt im Rahmen der C-Annahmen.

Annahme C1 (Exogene Variablen fix)

Die exogenen Variablen $x_{1t}, x_{2t}, \dots, x_{Kt}$ sind keine Zufallsvariablen, sondern können wie in einem Experiment kontrolliert werden.

- In der Einfachregression forderte Annahme c2: $S_{xx} > 0$.

Annahme C2 (Freiheit von Perfekter Multikollinearität)

Es existieren keine Parameterwerte $\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_K$ (wobei mindestens ein $\gamma_k \neq 0$), so dass zwischen den exogenen Variablen $x_{1t}, x_{2t}, \dots, x_{Kt}$ die lineare Beziehung

$$\gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \dots + \gamma_K x_{Kt} = 0 \quad (8.9)$$

für alle $t = 1, 2, \dots, T$ gilt.

- **Frage:** Handelt es sich bei den Daten der Tabelle 8.3 um perfekte Multikollinearität?

Tabelle 8.3: Der Unterschied zwischen perfekter und imperfekter Multikollinearität.

t	x_{1t}	x_{2t}	x_{3t}
1	12	4	2
2	17	6	3
3	32	12	6
4	10	2	4
\vdots	\vdots	\vdots	\vdots
T	20	8	5

- **Antwort:**

- Um zu erkennen, warum Annahme C2 notwendig ist, wird exemplarisch das ökonometrische Modell einer Zweifachregression betrachtet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t. \quad (8.3)$$

- Obige Gleichung (8.9) vereinfacht sich dann zu

$$\gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} = 0$$

und damit zu

$$x_{2t} = -(\gamma_0/\gamma_2) - (\gamma_1/\gamma_2) x_{1t} = \delta_0 + \delta_1 x_{1t} \quad (8.10)$$

wobei $\delta_0 = -(\gamma_0/\gamma_2)$ und $\delta_1 = -(\gamma_1/\gamma_2)$.

- Gleichung (8.3) degeneriert dann zu einer Einfachregression, denn

$$\begin{aligned}y_t &= \alpha + \beta_1 x_{1t} + \beta_2 (\delta_0 + \delta_1 x_{1t}) + u_t \\ &= \alpha' + \beta' x_{1t} + u_t ,\end{aligned}$$

wobei

$$\alpha' = \alpha + \beta_2 \delta_0$$

und

$$\beta' = \beta_1 + \beta_2 \delta_1 .$$

- **Frage:** Warum ist Annahme C2 eine Verallgemeinerung von Annahme c2?

Antwort:

- Wenn $T < K + 1$, ist Annahme C2 automatisch verletzt.
- Folglich fordert Annahme C2 implizit, dass $T \geq K + 1$.

Kapitel 9

Schätzung

- Das ökonometrische Modell der Mehrfachregression lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t \quad (9.1)$$

- Das ökonometrische Modell der Zweifachregression lautet

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \quad (9.2)$$

- Das geschätzte Modell lautet

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} \quad (9.3)$$

- Als Residuum bezeichnen wir wieder

$$\hat{u}_t = y_t - \hat{y}_t \quad (9.4)$$

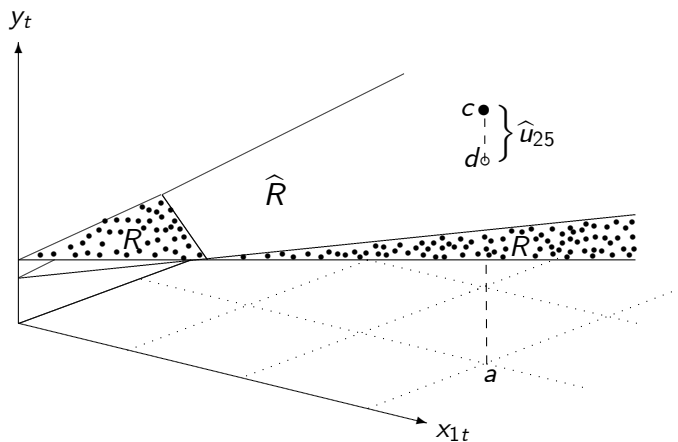


Abbildung 9.1: Der Zusammenhang zwischen wahrer Ebene R , Schätzebene \hat{R} und Residuen \hat{u}_t .

- **Frage:** Was wird in Abbildung 9.1 anhand der Beobachtung $t = 25$ veranschaulicht?

Antwort:

- Aus $\hat{u}_t = y_t - \hat{y}_t$ folgt

$$\hat{y}_t = y_t - \hat{u}_t$$

- Einsetzen in das geschätzte Modell

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t}$$

und nach y_t auflösen liefert die zweite Variante des geschätzten Modells:

$$y_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} + \hat{u}_t . \quad (9.5)$$

9.1 Punktschätzer $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$

- Die Summe der quadrierten Residuen lautet:

$$S_{\hat{u}\hat{u}} = \sum \hat{u}_t^2 . \quad (9.6)$$

- Das geschätzte Modell (9.5) nach \hat{u}_t auflösen:

$$\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta}_1 x_{1t} - \hat{\beta}_2 x_{2t} , \quad \text{für } t = 1, 2, \dots, T. \quad (9.7)$$

- Einsetzen in (9.6) liefert

$$S_{\hat{u}\hat{u}} = \sum \left(y_t - \hat{\alpha} - \hat{\beta}_1 x_{1t} - \hat{\beta}_2 x_{2t} \right)^2 . \quad (9.8)$$

- Diese Summe muss durch die Wahl von $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$ minimiert werden.

- Es müssen die folgenden drei Bedingungen erster Ordnung erfüllt sein:

$$\partial S_{\hat{u}\hat{u}} / \partial \hat{\alpha} = 0 \quad (9.9)$$

$$\partial S_{\hat{u}\hat{u}} / \partial \hat{\beta}_1 = 0 \quad (9.10)$$

$$\partial S_{\hat{u}\hat{u}} / \partial \hat{\beta}_2 = 0. \quad (9.11)$$

- Diese müssen anschließend nach $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$ aufgelöst werden.

- Notation in der Einfachregression:

$$\text{Variation } S_{xx} \equiv \sum (x_t - \bar{x})^2$$

$$\text{Kovariation } S_{xy} \equiv \sum (x_t - \bar{x})(y_t - \bar{y}) .$$

- Notation in der Zweifachregression:

$$S_{11} \equiv \sum (x_{1t} - \bar{x}_1)^2 = \sum x_{1t}^2 - T\bar{x}_1^2 \quad (9.12)$$

$$S_{22} \equiv \sum (x_{2t} - \bar{x}_2)^2 = \sum x_{2t}^2 - T\bar{x}_2^2 \quad (9.13)$$

$$S_{yy} \equiv \sum (y_t - \bar{y})(y_t - \bar{y}) = \sum y_t^2 - T\bar{y}^2 \quad (9.14)$$

$$S_{12} \equiv \sum (x_{1t} - \bar{x}_1)(x_{2t} - \bar{x}_2) = \sum x_{1t}x_{2t} - T\bar{x}_1\bar{x}_2 \quad (9.15)$$

$$S_{1y} \equiv \sum (x_{1t} - \bar{x}_1)(y_t - \bar{y}) = \sum x_{1t}y_t - T\bar{x}_1\bar{y} \quad (9.16)$$

$$S_{2y} \equiv \sum (x_{2t} - \bar{x}_2)(y_t - \bar{y}) = \sum x_{2t}y_t - T\bar{x}_2\bar{y} . \quad (9.17)$$

- Man erhält nach einer Reihe von Umformungen:

$$S_{1y} = \hat{\beta}_1 S_{11} + \hat{\beta}_2 S_{12}$$

$$S_{2y} = \hat{\beta}_1 S_{12} + \hat{\beta}_2 S_{22} .$$

- Nach $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$ aufgelöst, ergeben sich die folgenden Gleichungen

Formel: KQ-Schätzer (Zweifachregression)

$$\hat{\alpha} = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 \quad (9.18)$$

$$\hat{\beta}_1 = \frac{S_{22} S_{1y} - S_{12} S_{2y}}{S_{11} S_{22} - S_{12}^2} \quad (9.19)$$

$$\hat{\beta}_2 = \frac{S_{11} S_{2y} - S_{12} S_{1y}}{S_{11} S_{22} - S_{12}^2} \quad (9.20)$$

Numerische Illustration 9.1

Aus den Daten ergeben sich die Werte:

$$S_{11} = 0,22369 \quad (9.21)$$

$$S_{22} = 3,67920 \quad (9.22)$$

$$S_{yy} = 0,44629 \quad (9.23)$$

$$S_{12} = -0,00556 \quad (9.24)$$

$$S_{1y} = 0,13198 \quad (9.25)$$

$$S_{2y} = 0,96267 \quad (9.26)$$

Einsetzen dieser Werte in die Schätzformeln liefert die Schätzer der exogenen Variablen ($x_{1t} = \ln p_t$) und ($x_{2t} = \ln n_t$).

(Forts.)

Sie lauten

$$\hat{\beta}_1 = \frac{3,67920 \cdot 0,13198 + 0,00556 \cdot 0,96267}{0,22369 \cdot 3,67920 - [-0,00556]^2} = 0,59652 \quad (9.27)$$

$$\hat{\beta}_2 = \frac{0,22369 \cdot 0,96267 + 0,00556 \cdot 0,13198}{0,22369 \cdot 3,67920 - [-0,00556]^2} = 0,26255 \quad (9.28)$$

(Forts.)

Die Durchschnittswerte der drei Variablen lauten:

$$\bar{x}_1 = 3,22577, \bar{x}_2 = 4,32415 \text{ und } \bar{y} = 4,01386.$$

Zusammen mit den Werten für $\hat{\beta}_1$ und $\hat{\beta}_2$ ergibt sich:

$$\begin{aligned}\hat{\alpha} &= 4,01386 - 0,59652 \cdot 3,22577 - 0,26255 \cdot 4,32415 \\ &= 0,95432 .\end{aligned}$$

Das geschätzte Modell lautet demnach

$$\hat{y}_t = 0,95432 + 0,59652 \cdot x_{1t} + 0,26255 \cdot x_{2t} .$$

9.2 Interpretation der Schätzer $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$

9.2.1 Formale Interpretation

- Aus dem geschätzten Modell

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t}$$

erhält man die Ableitungen

$$\partial \hat{y}_t / \partial x_{1t} = \hat{\beta}_1 \quad \text{und} \quad \partial \hat{y}_t / \partial x_{2t} = \hat{\beta}_2. \quad (9.29)$$

- **Frage:** Welche formale Interpretation ergibt sich folglich?
Antwort:

9.2.2 Ökonomische Interpretation

Schätzer $\hat{\alpha} = 0,95432$:

- Bei Düngermengen $x_{1t} = \ln p_t = 0$ und $x_{2t} = \ln n_t = 0$ beträgt der erwartete *logarithmierte* Gersten-Output

$$\hat{y}_t = \hat{\alpha} = 0,95432$$

also

$$\widehat{\ln g_t} = \hat{\alpha} = 0,95432 .$$

- Wenn die Störgrößenvarianz σ^2 nicht zu groß ist, dann gilt

$$\widehat{\ln g_t} \approx \ln \hat{g}_t .$$

- Folglich erhalten wir

$$\ln \hat{g}_t = 0,95432$$

$$e^{\ln \hat{g}_t} = e^{0,95432}$$

$$[\text{da } e^{\ln \hat{g}_t} = \hat{g}_t] \quad \hat{g}_t = e^{0,95432} = 2,59690 . \quad (9.30)$$

- Man beachte, dass

$$x_{1t} = x_{2t} = 0$$

gleichbedeutend ist mit

$$n_t = p_t = 1.$$

Schätzer $\hat{\beta}_1 = 0,59652$ und $\hat{\beta}_2 = 0,26255$:

- Für $\hat{\beta}_1$ gilt

$$\hat{\beta}_1 = \frac{\partial \hat{y}_t}{\partial x_{1t}} = \frac{\partial (\ln \hat{g}_t)}{\partial (\ln p_t)}. \quad (9.31)$$

- Da

$$\frac{\partial \ln \hat{g}_t}{\partial \hat{g}_t} = \frac{1}{\hat{g}_t} \quad \text{und} \quad \frac{\partial \ln p_t}{\partial p_t} = \frac{1}{p_t}$$

können in Gleichung (9.31) $\partial (\ln \hat{g}_t)$ und $\partial (\ln p_t)$ durch $\partial \hat{g}_t / \hat{g}_t$ und $\partial p_t / p_t$ ersetzt werden:

$$\hat{\beta}_1 = \frac{\partial \hat{g}_t / \hat{g}_t}{\partial p_t / p_t} = \frac{100 \cdot (\partial \hat{g}_t / \hat{g}_t)}{100 \cdot (\partial p_t / p_t)} = \frac{\text{Prozentänderung von } \hat{g}_t}{\text{Prozentänderung von } p_t}.$$

- **Frage:** Welche ökonomische Interpretation besitzt demnach $\hat{\beta}_1$ im Düngerbeispiel?
Antwort: Es handelt sich um einen Spezialfall, nämlich um eine Elastizität. Sie zeigt an, um wie viel Prozent der Gerstenoutput steigt, wenn die Phosphatmenge um 1% erhöht wird und gleichzeitig der Stickstoff unverändert gelassen wird.
- Folgende allgemeine Interpretation kann immer verwendet werden:

Interpretation der geschätzten Steigungsparameter

Der KQ-Schätzwert $\hat{\beta}_k$ gibt an, um wie viel der Wert der endogenen Variable y_t ansteigt, wenn die exogene Variable x_k um eine Einheit erhöht wird und alle anderen exogenen Variablen konstant gehalten werden.

9.3 Autonome Variation der exogenen Variablen

9.3.1 Korrelation zwischen den exogenen Variablen

- Korrelationen zwischen den exogenen Variablen sind eine Besonderheit der multiplen Regressionsmodelle.
- Für den Fall einer Zweifachregression soll die Stärke der Korrelation gemessen werden.

- Es werden zwei *Hilfsregressionen* durchgeführt:

$$x_{1t} = \hat{\gamma}' + \hat{\delta}' x_{2t} + \hat{x}_{[1]t} \quad (9.32)$$

$$x_{2t} = \hat{\gamma}'' + \hat{\delta}'' x_{1t} + \hat{x}_{[2]t} . \quad (9.33)$$

- Die Residuen werden durch $\hat{x}_{[1]t}$ und $\hat{x}_{[2]t}$ bezeichnet.
- In Analogie zu $\hat{\beta} = S_{xy} / S_{xx}$ lauten die KQ-Schätzer

$$\hat{\delta}' = S_{21} / S_{22}$$

$$\hat{\delta}'' = S_{12} / S_{11} .$$

Numerische Illustration 9.2

Im Dünger-Beispiel gilt: $S_{11} = 0,22369$, $S_{22} = 3,67920$,
 $S_{12}(= S_{21}) = -0,00556$, $\bar{x}_1 = 3,22577$ und $\bar{x}_2 = 4,32415$.

Für die Parameter des Modells (9.32) ergeben sich die Schätzwerte

$$\begin{aligned}\hat{\delta}' &= \frac{S_{21}}{S_{22}} = \frac{-0,00556}{3,67920} & (9.34) \\ &= -0,00151\end{aligned}$$

$$\begin{aligned}\hat{\gamma}' &= \bar{x}_1 - \hat{\delta}'\bar{x}_2 = 3,22577 - 0,00151 \cdot 4,32415 & (9.35) \\ &= 3,23230.\end{aligned}$$

(Forts.)

Entsprechend erhält man für Modell (9.33) die Schätzwerte

$$\begin{aligned}\hat{\delta}'' &= \frac{S_{12}}{S_{11}} = \frac{-0,00556}{0,22369} \\ &= -0,02484\end{aligned}$$

$$\begin{aligned}\hat{\gamma}'' &= \bar{x}_2 - \hat{\delta}'' \bar{x}_1 = 4,32415 - (-0,02484) \cdot 3,22577 \\ &= 4,40428.\end{aligned}$$

Aus den beiden KQ-Schätzungen erhält man auch die Werte der Residuen $\hat{x}_{[1]t}$ und $\hat{x}_{[2]t}$.

- Das Bestimmtheitsmaß zu Modell (9.32) sei hier durch $R^2_{1.2}$ bezeichnet.
- Für das Modell $y_t = \alpha + \beta x_t + u_t$ errechnet man das Bestimmtheitsmaß aus

$$R^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}}$$

- Das Bestimmtheitsmaß zu Modell (9.32), $x_{1t} = \hat{\gamma}' + \hat{\delta}' x_{2t} + \hat{x}_{[1]t}$, lautet folglich

$$R^2_{1.2} = \frac{S_{21}^2}{S_{22} S_{11}} . \quad (9.36)$$

- Das Bestimmtheitsmaß zu Modell (9.33) ist identisch mit demjenigen des Modells (9.32), denn $S_{12} = S_{21}$ und damit

$$R_{2.1}^2 = \frac{S_{12}^2}{S_{11}S_{22}} = \frac{S_{21}^2}{S_{22}S_{11}} = R_{1.2}^2 . \quad (9.37)$$

- Es gilt

$$\widehat{cor}(x_{1t}, x_{2t}) = \pm \sqrt{R_{1.2}^2} ,$$

wobei das Vorzeichen dem Vorzeichen des Steigungsparameters $\widehat{\delta}'$ bzw. $\widehat{\delta}''$ entspricht.

Numerische Illustration 9.3

Aus den Ergebnissen (9.21), (9.22) und (9.24) erhält man:

$$R_{1.2}^2 = \frac{S_{12}^2}{S_{11}S_{22}} = \frac{[-0,00556]^2}{0,22369 \cdot 3,67920} = 0,00004 . \quad (9.38)$$

Der Korrelationskoeffizient beträgt folglich

$$\widehat{cor}(x_{1t}, x_{2t}) = -\sqrt{0,00004} = -0,00613 .$$

9.3.2 Berechnung der autonomen Variation

- In der Einfachregression

$$y_t = \hat{\alpha} + \hat{\beta}x_t + \hat{u}_t$$

sind die KQ-geschätzten Residuen \hat{u}_t jene Komponente in y_t , welche nicht durch x_t zu erklären ist.

- Folglich sind x_t und \hat{u}_t unkorreliert.

- **Frage:** Wie sind die Residuen $\hat{x}_{[1]t}$ des KQ-geschätzten Modells (9.32),

$$x_{1t} = \hat{\gamma}' + \hat{\delta}' x_{2t} + \hat{x}_{[1]t} ,$$

zu interpretieren?

Antwort:

- Wir können $\hat{x}_{[1]t}$ als die *autonome Komponente* in x_{1t} auffassen.
- Die Variation von $\hat{x}_{[1]t}$ wird hier als *autonome Variation* $S_{[11]}$ bezeichnet.
- Es lässt sich zeigen (siehe Lehrbuch-Anhang), dass

$$S_{[11]} = S_{11}(1 - R_{1.2}^2) . \quad (9.39)$$

- Aus der Gesamtvariation der exogenen Variable x_{1t} wird mit Hilfe des Korrekturfaktors $(1 - R_{1.2}^2)$ die *autonome Variation* von x_{1t} errechnet.
- **Frage:** Warum ist $S_{[11]}$ niemals größer als S_{11} ?

Antwort:

- Ganz analog gilt

$$S_{[22]} = S_{22}(1 - R_{2.1}^2) . \quad (9.40)$$

- Aus Gleichung (9.37) wissen wir, dass $R_{2.1}^2 = R_{1.2}^2$.

Numerische Illustration 9.4

Wir wissen bereits, dass $S_{11} = 0,22369$, $S_{22} = 3,67920$ und $R_{1.2}^2 = 0,00004$. Einsetzen in die Formeln (9.39) und (9.40) liefert:

$$\begin{aligned} S_{[11]} &= 0,22369 \cdot (1 - 0,00004) = 0,22368 \\ S_{[22]} &= 3,67920 \cdot (1 - 0,00004) = 3,67906 . \end{aligned}$$

9.4 Informationsverarbeitung der KQ-Methode und Bestimmtheitsmaß R^2

9.4.1 Definition des Bestimmtheitsmaßes

- Es gilt wie in der Einfachregression

$$\text{Gesamtvariation:} \quad S_{yy} \equiv \sum (y_t - \bar{y})^2$$

$$\text{unerklärte Variation:} \quad S_{\hat{u}\hat{u}} \equiv \sum (\hat{u}_t - \bar{\hat{u}})^2 = \sum \hat{u}_t^2$$

$$\text{erklärte Variation:} \quad S_{\hat{y}\hat{y}} \equiv \sum (\hat{y}_t - \bar{y})^2, \quad (9.41)$$

wobei

$$\bar{y} = \bar{\hat{y}} = (1/T) \sum \hat{y}_t.$$

- Damit ergibt sich genau wie bei der Einfachregression:

$$S_{yy} = S_{\widehat{yy}} + S_{\widehat{uu}} . \quad (9.42)$$

- Entsprechend ist das *Bestimmtheitsmaß* R^2 genau wie im Falle der Einfachregression definiert:

$$R^2 = \frac{S_{yy} - S_{\widehat{uu}}}{S_{yy}} = \frac{S_{\widehat{yy}}}{S_{yy}} . \quad (9.43)$$

9.4.2 Berechnung des Bestimmtheitsmaßes

- Für die Einfachregression konnte gezeigt werden, dass

$$S_{\widehat{y}y} = \widehat{\beta} S_{xy}$$

und damit

$$R^2 = \frac{S_{\widehat{y}y}}{S_{yy}} = \frac{\widehat{\beta} S_{xy}}{S_{yy}} .$$

- Für die Mehrfachregression kann gezeigt werden (siehe matrixalgebraischen Lehrbuch-Anhang), dass

$$S_{\widehat{y}y} = \sum_{k=1}^K \widehat{\beta}_k S_{ky} \quad (9.44)$$

und damit

$$R^2 = \frac{S_{\widehat{y}y}}{S_{yy}} = \frac{\sum_{k=1}^K \widehat{\beta}_k S_{ky}}{S_{yy}} . \quad (9.45)$$

Numerische Illustration 9.5

Wir wissen, dass

$$\hat{\beta}_1 = 0,59652 \quad \text{und} \quad \hat{\beta}_2 = 0,26255$$

sowie

$$S_{1y} = 0,13198, \quad S_{2y} = 0,96267 \quad \text{und} \quad S_{yy} = 0,44629 .$$

Einsetzen in Formel (9.45) ergibt

$$R^2 = \frac{0,59652 \cdot 0,13198 + 0,26255 \cdot 0,96267}{0,44629} = 74,3\% .$$

(Forts.)

Ökonometrische Standard-Software berechnet den Wert:

$$S_{\widehat{u}\widehat{u}} = 0,11481 . \quad (9.46)$$

Eine Berechnung von R^2 mit Hilfe von Formel (9.43) liefert folglich

$$R^2 = \frac{0,44629 - 0,11481}{0,44629} = 74,3\% ,$$

was mit dem ersten Ergebnis aus Formel (9.45) übereinstimmt.

9.4.3 Bestimmtheitsmaß und Venn-Diagramme

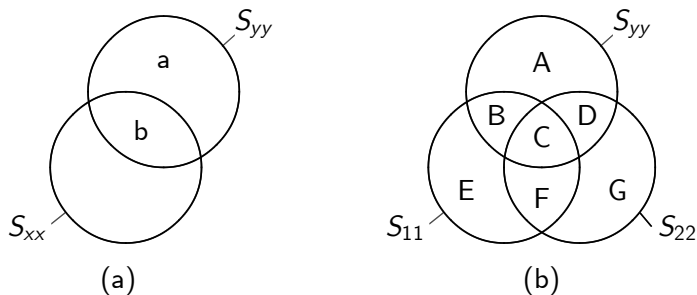


Abbildung 9.2: Grafische Veranschaulichung des Bestimmtheitsmaßes.

- In der Einfachregression gilt:

$$R^2 = \frac{b}{a + b}$$

- In der Zweifachregression gilt:

$$R^2 = \frac{B + C + D}{A + B + C + D}$$

9.4.4 KQ-Methode als zweistufiger Prozess

- Die autonomen Komponenten $\hat{x}_{[1]t}$ und $\hat{x}_{[2]t}$ ergaben sich als Residuen der beiden Hilfsregressionen (9.32) und (9.33):

$$\begin{aligned}x_{1t} &= \hat{\gamma}' + \hat{\delta}' x_{2t} + \hat{x}_{[1]t} \\x_{2t} &= \hat{\gamma}'' + \hat{\delta}'' x_{1t} + \hat{x}_{[2]t} .\end{aligned}$$

- Diese beiden Einfachregressionen stellen deshalb die erste Stufe des zweistufigen Prozesses dar.
- In der zweiten Stufe wird zunächst folgende Einfachregression durchgeführt:

$$y_t = \alpha_1' + \beta_1' \hat{x}_{[1]t} + v_{1t} . \quad (9.47)$$

- Dabei ist v_{1t} die Störgröße der Regression.

- Der KQ-Schätzer für den Steigungsparameter β_1' lautet wie gewohnt

$$\hat{\beta}_1' = \frac{S_{[1]y}}{S_{[11]}}, \quad (9.48)$$

wobei $S_{[1]y} = \sum(\hat{x}_{[1]t} - \bar{\hat{x}}_{[1]})(y_t - \bar{y})$ die Kovariation zwischen $\hat{x}_{[1]t}$ und y_t ist.

- $\hat{x}_{[1]t}$ ist mit y_t aber nicht mit x_{2t} korreliert.
- $S_{[11]}$ ist die autonome Variation der Variable x_{1t} .
- $S_{[11]}$ ist in Abbildung 9.2 durch die Flächen B und E repräsentiert.
- Folglich stellt sich die Einfachregression (9.47) grafisch wie in Abbildung 9.3 dar.

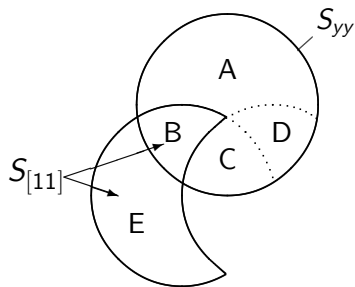


Abbildung 9.3: Das zur Einfachregression (9.47) korrespondierende Venn-Diagramm.

- Um die zweite Stufe des zweistufigen Prozesses abzuschließen, wird folgende Regression durchgeführt:

$$y_t = \alpha_2' + \beta_2' \hat{x}_{[2]t} + v_{2t} . \quad (9.49)$$

- **Frage:** Wie müsste Abbildung 9.2 verändert werden, um die Regression (9.49) darzustellen?

Antwort:

Numerische Illustration 9.6

Wir wissen, dass $S_{[11]} = 0,22368$ und $S_{[22]} = 3,67906$. Aus den Werten der Variablen $\hat{x}_{[1]t}$ und y_t ergibt sich eine Kovariation von $S_{[1]y} = 0,13343$. Die Kovariation der Variablen $\hat{x}_{[2]t}$ und y_t beträgt $S_{[2]y} = 0,96595$. Die KQ-Schätzungen der ökonometrischen Modelle (9.47) und (9.49) liefert deshalb

$$\hat{\beta}'_1 = \frac{S_{[1]y}}{S_{[11]}} = \frac{0,13343}{0,22368} = 0,59652 \quad (9.50)$$

$$\hat{\beta}'_2 = \frac{S_{[2]y}}{S_{[22]}} = \frac{0,96595}{3,67906} = 0,26255 \quad (9.51)$$

- Die Schätzwerte $\widehat{\beta}'_1$ und $\widehat{\beta}'_2$ stimmen genau mit den Schätzwerten $\widehat{\beta}_1$ und $\widehat{\beta}_2$ überein, die wir in unserer KQ-Schätzung des ökonometrischen Modells

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \quad (9.2)$$

ermittelt hatten:

$$\widehat{\beta}'_1 = \widehat{\beta}_1 \quad \text{und} \quad \widehat{\beta}'_2 = \widehat{\beta}_2 . \quad (9.52)$$

- Dies ist ein Teilergebnis des *Frisch-Waugh-Lovell-Theorems*.

- **Frage:** Welche intuitive Begründung kann für dieses Teilergebnis angeführt werden?

Antwort:

- Der Schätzwert $\widehat{\beta}'_1$ wird aus der KQ-Schätzung des Modells (9.47) gewonnen.
- Diese KQ-Schätzung war in Abbildung 9.3 dargestellt.
- Der Schätzwert $\widehat{\beta}'_1$ wird demnach allein auf Basis der Informationen in Fläche B ermittelt.
- Da $\widehat{\beta}'_1 = \widehat{\beta}_1$, basiert auch die Berechnung von $\widehat{\beta}_1$ in der Abbildung 9.2 (siehe nächste Folie) allein auf den Informationen in Fläche B .
- Genau die gleiche Argumentationslinie gilt für $\widehat{\beta}'_2$ und $\widehat{\beta}_2$.

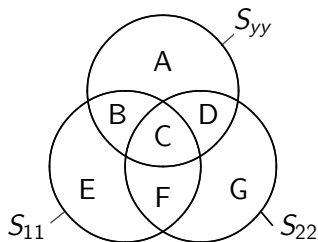


Abbildung 9.2: Grafische Veranschaulichung des Bestimmtheitsmaßes.

- Die Informationen in Fläche C sind dennoch nicht nutzlos.
- Gegeben die Flächengrößen B und D der Abbildung 9.2, führt eine Vergrößerung der Fläche C auf Kosten der Fläche A zu einem erhöhten Bestimmtheitsmaß.
- **Frage:** Warum erhöht dies auch die Genauigkeit von Prognosen?

Antwort:

9.4.5 Partielles Bestimmtheitsmaß

- Wenn man y_t allein auf x_{2t} regressiert, erhält man als Residuen die um die x_{2t} -Komponente bereinigten y_t -Werte.
- Diese Residuen seien durch $\hat{y}_{[1]t}$ bezeichnet.
- Man könnte die folgende Einfachregression durchführen:

$$\hat{y}_{[1]t} = \alpha_1'' + \beta_1'' \hat{x}_{[1]t} + w_{1t} . \quad (9.53)$$

- **Frage:** Wie würde das entsprechende Venn-Diagramm aussehen?

Antwort:

- Folglich würde auch $\widehat{\beta}_1''$, der KQ-Schätzer der Einfachregression (9.53), allein aus den Informationen in Fläche B ermittelt werden und es würde wieder zum gleichen Schätzergebnis kommen wie bei $\widehat{\beta}_1$, dem KQ-Schätzer der Zweifachregression (9.2).
- Analoges gilt für die Schätzung von β_2 .

Frisch-Waugh-Lovell-Theorem

Es gilt:

$$\widehat{\beta}_k'' = \widehat{\beta}_k' = \widehat{\beta}_k .$$

- Zur KQ-Schätzung des Modells (9.53) lässt sich auch das Bestimmtheitsmaß berechnen.
- Man könnte sagen, es misst im ökonometrischen Modell (9.2) die *partielle* Erklärungskraft der Variable x_{1t} in Bezug auf die Variable y_t .
- Deshalb bezeichnet man es als das *partielle Bestimmtheitsmaß* der Variable x_{1t} .

- **Frage:** Wie könnte man in Abbildung 9.2 das partielle Bestimmtheitsmaß der Variable x_{1t} grafisch beschreiben? Wie könnte dasjenige von x_{2t} beschrieben werden?

Antwort:

- Die Summe der partiellen Bestimmtheitsmaße addieren sich *nicht* zum Bestimmtheitsmaß R^2 .

9.5 Unverzerrtheit und Effizienz der KQ-Methode

9.5.1 Erwartungswert und Varianz der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$

- In der Zweifachregression ergeben sich für die KQ-Schätzer die Erwartungswerte

$$E(\hat{\beta}_1) = \beta_1 \quad (9.54)$$

$$E(\hat{\beta}_2) = \beta_2 \quad (9.55)$$

$$E(\hat{\alpha}) = \alpha . \quad (9.56)$$

Formel: Varianzen der KQ-Schätzer bei $K = 2$

$$\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{11} (1 - R_{1.2}^2)} = \frac{\sigma^2}{S_{[11]}} \quad (9.57)$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{S_{22} (1 - R_{1.2}^2)} = \frac{\sigma^2}{S_{[22]}} \quad (9.58)$$

$$\begin{aligned} \text{var}(\hat{\alpha}) &= \sigma^2 / T + \bar{x}_1^2 \text{var}(\hat{\beta}_1) + \bar{x}_2^2 \text{var}(\hat{\beta}_2) \\ &\quad + 2\bar{x}_1\bar{x}_2 \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \end{aligned} \quad (9.59)$$

$$\text{cov}(\hat{\beta}_1, \hat{\beta}_2) = \frac{-\sigma^2 R_{1.2}^2}{S_{12} (1 - R_{1.2}^2)}, \quad (9.60)$$

wobei

$$R_{1.2}^2 = \frac{S_{12}^2}{S_{11} S_{22}}.$$

9.5.2 Interpretation der Formeln

- Aus den Gleichungen (9.54) bis (9.56) folgt unmittelbar:

KQ-Schätzung ist unverzerrt

Bei Gültigkeit der A-, B-, und C-Annahmen (Annahmen B2 bis B4 nicht erforderlich) sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$ ($k = 1, \dots, K$) unverzerrt.

- **Frage:** Welchen Einfluss hat $R_{1.2}^2$ auf die Präzision, mit der wir $\hat{\beta}_1$ und $\hat{\beta}_2$ schätzen können?

Antwort:

- Der Quotient $1 / (1 - R_{1.2}^2)$ wird häufig als *Variance Inflation Factor* bezeichnet.

9.5.3 Schätzformeln für $var(\hat{\alpha})$, $var(\hat{\beta}_k)$ und $var(\hat{\beta}_1, \hat{\beta}_2)$

Formel: Schätzung der Störgrößenvarianz

Wenn Annahmen B1 bis B3 erfüllt sind, dann ist im Rahmen der Mehrfachregression mit K exogenen Variablen die Formel

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T - K - 1} \quad (9.61)$$

ein unverzerrter Schätzer der Störgrößenvarianz σ^2 .

- In der Zweifachregression lautet der Schätzer also

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T-3} \quad (9.62)$$

- In der Einfachregression lautet er

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T-2} .$$

Numerische Illustration 9.7

Aus $T = 30$ und $S_{\hat{u}\hat{u}} = 0,11481$ erhält man

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 3) = 0,11481 / 27 = 0,00425 . \quad (9.63)$$

Einsetzen in (9.57) bis (9.60) liefert

$$\widehat{\text{var}}(\hat{\beta}_1) = 0,01901 \quad (9.64)$$

$$\widehat{\text{var}}(\hat{\beta}_2) = 0,00116 \quad (9.65)$$

$$\widehat{\text{var}}(\hat{\alpha}) = 0,22037$$

$$\widehat{\text{cov}}(\hat{\beta}_1, \hat{\beta}_2) = 0,00003 . \quad (9.66)$$

9.5.4 BLUE- bzw. BUE-Eigenschaft der KQ-Schätzer

Gauss-Markov-Theorem

Auch in der Mehrfachregression besitzen die KQ-Schätzer die BLUE- bzw. BUE-Eigenschaft, sofern die A-, B- und C-Annahmen erfüllt sind (Annahme B4 ist für die BLUE-Eigenschaft nicht erforderlich).

9.6 Wahrscheinlichkeitsverteilungen der KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$

9.6.1 Wahrscheinlichkeitsverteilung der y_t

- Es gilt $y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$ und deshalb:

$$\begin{aligned} E(y_t) &= E[\alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t] \\ &= E(\alpha) + E(\beta_1 x_{1t}) + E(\beta_2 x_{2t}) + E(u_t) \\ &= \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + 0. \end{aligned}$$

- Ferner gilt

$$\begin{aligned} \text{var}(y_t) &= E \left[[y_t - E(y_t)]^2 \right] \\ &= E \left[[y_t - \alpha - \beta_1 x_{1t} - \beta_2 x_{2t}]^2 \right] \\ &= E \left[u_t^2 \right] \\ &= E \left[[u_t - E(u_t)]^2 \right] = \sigma^2 . \end{aligned}$$

- Man erhält also

$$y_t \sim UN(\alpha + \beta_1 x_{1t} + \beta_2 x_{2t}, \sigma^2),$$

für alle $t = 1, 2, \dots, T$.

9.6.2 Wahrscheinlichkeitsverteilungen der Schätzer $\hat{\alpha}$ und $\hat{\beta}_k$

- Die gewichtete Summe (d.h. eine lineare Funktion) unabhängig normalverteilter Zufallsvariablen ist ebenfalls normalverteilt.
- Die KQ-Schätzer (9.18) bis (9.20) sind gewichtete Summen der Zufallsvariablen y_t .
- Folglich gilt unter Ausnutzung von (9.54) bis (9.59) auch in der Zweifachregression das folgende Resultat.

KQ-Schätzer sind normalverteilt

Die KQ-Schätzer $\hat{\alpha}$, $\hat{\beta}_1$ und $\hat{\beta}_2$ sind normalverteilte Zufallsvariablen mit

$$\hat{\beta}_1 \sim N \left(\beta_1, \frac{\sigma^2}{S_{11} (1 - R_{1.2}^2)} \right) \quad (9.67)$$

$$\hat{\beta}_2 \sim N \left(\beta_2, \frac{\sigma^2}{S_{22} (1 - R_{1.2}^2)} \right)$$

$$\hat{\alpha} \sim N(\alpha, \text{var}(\hat{\alpha}))$$

9.7 Intervallschätzer

- Ausgangspunkt ist wie gewohnt

$$\Pr \left\{ \hat{\beta}_1 - k \leq \beta_1 \leq \hat{\beta}_1 + k \right\} = 1 - a . \quad (9.68)$$

- Auch hier ist σ^2 unbekannt und damit ist auch

$$sd(\hat{\beta}_1) = \sqrt{\text{var}(\hat{\beta}_1)} = \sqrt{\frac{\sigma^2}{S_{11} (1 - R_{1.2}^2)}}$$

unbekannt.

Schritt 1: Schätzung von σ^2 und $sd(\hat{\beta}_1)$

- Ein unverzerrter Schätzer für σ^2 :

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 3). \quad (9.62)$$

- Die geschätzte Standardabweichung der Zufallsvariable $\hat{\beta}_1$ lautet:

$$\hat{sd}(\hat{\beta}_1) = \sqrt{\widehat{\text{var}}(\hat{\beta}_1)} = \sqrt{\frac{\hat{\sigma}^2}{S_{11} (1 - R_{1 \cdot 2}^2)}} \quad (9.69)$$

Numerische Illustration 9.8

Aus $\widehat{\text{var}}(\hat{\beta}_1) = 0,01901$ ergibt sich unmittelbar:

$$\hat{sd}(\hat{\beta}_1) = \sqrt{\widehat{\text{var}}(\hat{\beta}_1)} = \sqrt{0,01901} = 0,13788. \quad (9.70)$$

Schritt 2: Standardisierung von $\hat{\beta}_1$

- Die Standardisierung lautet wie in der Einfachregression:

$$t = \frac{\hat{\beta}_1 - E(\hat{\beta}_1)}{\widehat{sd}(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - \beta_1}{\widehat{sd}(\hat{\beta}_1)} . \quad (9.71)$$

- Dabei folgt t einer t -Verteilung mit $(T - 3)$ Freiheitsgraden: $t \sim t_{(T-3)}$.

Schritt 3: Ermittlung eines $t_{a/2}$ -Wertes

- Für eine $t_{(T-3)}$ -verteilte Zufallsvariable t gilt:

$$\Pr \{-t_{a/2} \leq t \leq t_{a/2}\} = 1 - a \quad (9.72)$$

Numerische Illustration 9.9

Signifikanzniveau $a = 5\%$. Anzahl der Freiheitsgrade $30 - 3 = 27$. Aus Tabelle T.2: t liegt mit einer Wahrscheinlichkeit von 95% im Intervall $[-2,0518 ; 2,0518]$, also $t_{a/2} = 2,0518$.

Schritt 4: Formulierung des Intervallschätzers

- Einsetzen von (9.71) in (9.72) und nach β_1 auflösen liefert

$$\Pr \left\{ \hat{\beta}_1 - t_{a/2} \cdot \widehat{sd}(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{a/2} \cdot \widehat{sd}(\hat{\beta}_1) \right\} = 1 - a \quad (9.73)$$

- Es ergibt sich folglich der gleiche Intervallschätzer wie in der Einfachregression:

Formel: Intervallschätzer

$$\left[\hat{\beta}_k - t_{a/2} \cdot \widehat{sd}(\hat{\beta}_k) ; \hat{\beta}_k + t_{a/2} \cdot \widehat{sd}(\hat{\beta}_k) \right] \quad (9.74)$$

Numerische Illustration 9.10

Unsere Schätzungen lauteten $\hat{\beta}_1 = 0,59652$ und $\widehat{sd}(\hat{\beta}_1) = 0,13788$. Aus Tabelle T.2 hatten wir $t_{a/2} = 2,0518$ ermittelt. Einsetzen in (9.74) liefert das Intervall $[0,31360 ; 0,87945]$.

Intervallschätzer für $\hat{\alpha}$:

- Dieser lautet wie in der Einfachregression:

$$\left[\hat{\alpha} - t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) ; \hat{\alpha} + t_{a/2} \cdot \widehat{sd}(\hat{\alpha}) \right] . \quad (9.75)$$

Kapitel 10

Hypothesentest

10.1 Testen einer Linearkombination von Parametern: t -Test

10.1.1 Zweiseitiger t -Test

- In der Einfachregression wurden zweiseitige Hypothesen der Form

$$H_0 : \beta = q \quad \text{und} \quad H_1 : \beta \neq q \quad (10.1)$$

betrachtet.

- Im Rahmen einer Zweifachregression

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t, \quad (10.2)$$

laute die Null- und Alternativhypothese

$$H_0 : r_1 \beta_1 + r_2 \beta_2 = q \quad \text{und} \quad H_1 : r_1 \beta_1 + r_2 \beta_2 \neq q, \quad (10.3)$$

wobei r_1 , r_2 und q Konstanten sind, denen ein numerischer Wert zugeordnet ist.

- **Frage:** Für welche Werte r_1 und r_2 vereinfacht sich (10.3) zu (10.1)?

Antwort:

Schritt 1: Formulierung von H_0 und H_1 und Festlegung des Signifikanzniveaus α

- Die Null- und Alternativhypothese sind durch (10.3) gegeben.

Numerische Illustration 10.1

Ein besonders relevanter Test ist:

$$H_0 : \beta_1 + \beta_2 = 1 \quad \text{und} \quad H_1 : \beta_1 + \beta_2 \neq 1 .$$

Das Signifikanzniveau sei $\alpha = 5\%$.

Frage: Welche ökonomische Frage wird bei dieser Nullhypothese getestet?

Antwort:

Schritt 2: Schätzung von $sd(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)$

- In diesem Test ist $(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)$ die relevante Zufallsvariable und $\hat{sd}(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)$ ihre geschätzte Standardabweichung.
- Es gilt

$$\text{var}(r_1\hat{\beta}_1 + r_2\hat{\beta}_2) = r_1^2 \text{var}(\hat{\beta}_1) + r_2^2 \text{var}(\hat{\beta}_2) + 2r_1r_2 \text{cov}(\hat{\beta}_1, \hat{\beta}_2) .$$

- Folglich ergibt sich

$$\begin{aligned}
 \widehat{sd}(r_1\widehat{\beta}_1 + r_2\widehat{\beta}_2) &= \sqrt{\widehat{var}(r_1\widehat{\beta}_1 + r_2\widehat{\beta}_2)} \\
 &= \sqrt{r_1^2\widehat{var}(\widehat{\beta}_1) + r_2^2\widehat{var}(\widehat{\beta}_2) + 2r_1r_2\widehat{cov}(\widehat{\beta}_1, \widehat{\beta}_2)} \\
 &= \sqrt{r_1^2 \frac{\widehat{\sigma}^2}{S_{11}(1-R_{1.2}^2)} + r_2^2 \frac{\widehat{\sigma}^2}{S_{22}(1-R_{1.2}^2)} + 2r_1r_2 \frac{-\widehat{\sigma}^2 R_{1.2}^2}{S_{12}(1-R_{1.2}^2)}} \quad (10.4)
 \end{aligned}$$

wobei

$$\widehat{\sigma}^2 = S_{\widehat{u}\widehat{u}} / (T - 3) . \quad (10.5)$$

Numerische Illustration 10.2

Einsetzen der bisherigen Berechnungen in Formel (10.4) ergibt:

$$\begin{aligned} & \widehat{sd}(1 \cdot \widehat{\beta}_1 + 1 \cdot \widehat{\beta}_2) \\ = & \sqrt{\frac{0,00425}{0,22369(0,99996)} + \frac{0,00425}{3,67920(0,99996)} + 2 \frac{-0,00425 \cdot 0,00004}{-0,00556(0,99996)}} \\ = & \sqrt{0,01900 + 0,00116 + 0,00006} \\ = & 0,14220 . \end{aligned} \tag{10.6}$$

Schritt 3: Ermittlung des t -Wertes aus der Stichprobe

- Beim Test der einfachen Nullhypothese wurde $t = (\hat{\beta} - q) / \widehat{sd}(\hat{\beta})$ berechnet.
- Entsprechend ist nun $(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)$ die Zufallsvariable, $q (= r_1\beta_1 + r_2\beta_2)$ der Erwartungswert und $\widehat{sd}(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)$ die geschätzte Standardabweichung.
- Der t -Wert lautet deshalb

$$t = \frac{(r_1\hat{\beta}_1 + r_2\hat{\beta}_2) - q}{\widehat{sd}(r_1\hat{\beta}_1 + r_2\hat{\beta}_2)} \quad (10.7)$$

Numerische Illustration 10.3

Der beobachtete Wert für t beträgt

$$t = \frac{(1 \cdot 0,59652 + 1 \cdot 0,26255) - 1}{0,14220} \approx -0,991 . \quad (10.8)$$

- Wenn die Nullhypothese wahr ist, dann gilt: $t \sim t_{(T-3)}$.

Schritt 4: Ermittlung des kritischen Wertes $t_{a/2}$

- Für das vorgegebene Signifikanzniveau a muss jener Wert $t_{a/2}$ ermittelt werden, für den gilt:

$$\Pr \{-t_{a/2} \leq t \leq t_{a/2}\} = 1 - a .$$

Numerische Illustration 10.4

Frage: Aus Tabelle T.2 können die Werte $-t_{a/2}$ und $t_{a/2}$ ermittelt werden. Wie lauten sie?

Antwort:

Schritt 5: Vergleich von $t_{\alpha/2}$ und t

Numerische Illustration 10.5

Der berechnete t -Wert von $-0,991$ wird mit dem Wert $t_{\alpha/2}$ verglichen.

Frage: Zu welcher Testentscheidung gelangen wir?

Antwort:

10.1.2 Einseitiger t -Test

- Ein linksseitiger Test hätte folgende allgemeine Form:

$$H_0 : r_1\beta_1 + r_2\beta_2 \geq q \quad \text{und} \quad H_1 : r_1\beta_1 + r_2\beta_2 < q . \quad (10.9)$$

- Aus (10.7) ist ersichtlich, dass nur sehr negative t -Werte zu einer Ablehnung führen.

Numerische Illustration 10.6

Die Hypothese laute

$$H_0 : \beta_1 + \beta_2 \geq 1 \quad \text{und} \quad H_1 : \beta_1 + \beta_2 < 1 .$$

Frage: Was sagt diese Nullhypothese ökonomisch aus?

Antwort:

(Forts.)

Frage: Welcher t -Wert ergibt sich?

Antwort:

Frage: Welcher kritische Wert $-t_a$ ergibt sich?

Antwort:

Frage: Zu welcher Testentscheidung gelangen wir?

Antwort:

10.2 Simultaner Test mehrerer Linearkombinationen von Parametern: *F*-Test

- Für den Fall der Zweifachregression,

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t ,$$

sind

$$H_0 : \beta_1 = 0 \text{ und } H_1 : \beta_1 \neq 0$$

beziehungsweise

$$H_0 : \beta_2 = 0 \text{ und } H_1 : \beta_2 \neq 0$$

die einfachsten Null- und Alternativhypothesen.

- Entsprechend stellt der *simultane* Test der Nullhypothese

$$H_0 : \beta_1 = 0 \text{ und gleichzeitig } \beta_2 = 0 \quad (10.10)$$

gegen die Alternativhypothese

$$H_1 : \beta_1 \neq 0 \text{ und/oder } \beta_2 \neq 0 \quad (10.11)$$

den wichtigsten und einfachsten Fall simultan zu testender “Linearkombinationen” dar.

- **Frage:** Warum wurde in der Alternativhypothese „und/oder“ benutzt?

Antwort:

10.2.1 Eine wichtige Nullhypothese

Schritt 1: Formulierung von H_0 und H_1 und Festlegung des Signifikanzniveaus α

Numerische Illustration 10.7

Im Rahmen des Dünger-Beispiels soll die Nullhypothese

$$H_0 : \beta_1 = \beta_2 = 0 \quad (10.12)$$

gegen die Alternativhypothese

$$H_1 : \beta_1 \neq 0 \text{ und/oder } \beta_2 \neq 0$$

auf einem Signifikanzniveau von $\alpha = 5\%$ getestet werden.

Schritt 2: Berechnung der Summe der Residuenquadrate für das ökonometrische Modell ($S_{\widehat{u}\widehat{u}}$) und das Nullhypothesenmodell ($S_{\widehat{u}\widehat{u}}^0$)

- Das Modell der Zweifachregression lautet $y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t$ und seine Summe der Residuenquadrate beträgt $S_{\widehat{u}\widehat{u}}$.
- Das *Nullhypothesenmodell* lautet somit:

$$y_t = \alpha + 0 \cdot x_{1t} + 0 \cdot x_{2t} + u_t^0 = \alpha + u_t^0. \quad (10.13)$$

- Das entsprechende geschätzte Nullhypothesemodell lautet

$$y_t = \widehat{\alpha} + \widehat{u}_t^0.$$

- Umstellen liefert

$$\widehat{u}_t^0 = y_t - \widehat{\alpha}. \quad (10.14)$$

- Eine KQ-Schätzung ist hier also eine Minimierung von

$$\begin{aligned} S_{\hat{u}\hat{u}}^0 &= \sum (y_t - \hat{\alpha})^2 \\ &= \sum (y_t^2 + \hat{\alpha}^2 - 2y_t\hat{\alpha}) \\ &= \sum y_t^2 + T\hat{\alpha}^2 - 2\hat{\alpha} \sum y_t . \end{aligned}$$

- Ableiten dieses Ausdrucks nach $\hat{\alpha}$ und Nullsetzen ergibt

$$\partial S_{\hat{u}\hat{u}}^0 / \partial \hat{\alpha} = 2T\hat{\alpha} - 2 \sum y_t \stackrel{!}{=} 0 .$$

- Folglich ergibt sich

$$T\hat{\alpha} = \sum y_t \quad \text{und damit} \quad \hat{\alpha} = \bar{y} .$$

- Einsetzen in (10.14) liefert

$$\hat{u}_t^0 = y_t - \bar{y}$$

und damit

$$\sum (\hat{u}_t^0)^2 = \sum (y_t - \bar{y})^2$$

beziehungsweise

$$S_{\hat{u}\hat{u}}^0 = S_{yy} .$$

Numerische Illustration 10.8

Es war $S_{\hat{u}\hat{u}} = 0,11481$. Das Nullhypothesenmodell (10.13) liefert $S_{\hat{u}\hat{u}}^0 (= S_{yy}) = 0,44629$.

Schritt 3: Berechnung des F -Wertes

- Es wird die folgende Zufallsvariable berechnet:

$$F = \frac{(S_{\hat{u}\hat{u}}^0 - S_{\hat{u}\hat{u}}) / L}{S_{\hat{u}\hat{u}} / (T - K - 1)}, \quad (10.15)$$

wobei L die Anzahl der in der Nullhypothese formulierten Linearkombinationen ist und $T - K - 1$ die Anzahl der Freiheitsgrade.

- **Frage:** Wie viel zusätzliche Freiheitsgrade besitzt das Nullhypothesenmodell gegenüber dem unrestringierten Modell?

Antwort:

- **Frage:** Wie ist folglich der Zähler der Formel (10.15) zu interpretieren?
Antwort:

- **Frage:** Wie ist der Nenner zu interpretieren?
Antwort:

- Besteht zwischen der Realität und der Nullhypothese ein deutlicher Widerspruch, dann sollte der Zähler in Formel (10.15) deutlich größer als der Nenner ausfallen.

- Im Beispiel gilt $K = 2$ und $L = 2$ und damit:

$$F = \frac{(S_{\hat{u}\hat{u}}^0 - S_{\hat{u}\hat{u}}) / 2}{S_{\hat{u}\hat{u}} / (T - 3)} . \quad (10.16)$$

Numerische Illustration 10.9

Es liegen $T = 30$ Beobachtungen vor. Einsetzen der Ergebnisse aus Schritt 2 liefert den F -Wert:

$$F = \frac{(0,44629 - 0,11481) / 2}{0,11481 / (30 - 3)} \approx 38,98 .$$

- Wenn die Nullhypothese wahr ist, dann folgt F einer $F_{(v_1, v_2)}$ -Verteilung, wobei $v_1 = L$ und $v_2 = T - K - 1$:

$$F \sim F_{(L, T - K - 1)} .$$

Schritt 4: Ermittlung eines kritischen Wertes F_a

Numerische Illustration 10.10

Tabelle T.3 hat auf der Horizontalen die Freiheitsgrade L abgetragen (v_1) und auf der Vertikalen die Freiheitsgrade $T - K - 1$ (v_2). Die tabellierten $F_{0,05}$ -Werte besagen, dass für eine $F_{(v_1, v_2)}$ -verteilte Zufallsvariable 5% der Wahrscheinlichkeitsmasse rechts des jeweiligen $F_{0,05}$ -Wertes liegen.

Frage: Welcher $F_{0,05}$ -Wert ergibt sich in unserem Test?

Antwort:

Schritt 5: Vergleich von F_a und F

- Wenn $F > F_a$, wird die Nullhypothese abgelehnt.

Numerische Illustration 10.11

$F_{0,05}$, der kritische Wert für den Hypothesentest, wird mit dem berechneten F -Wert von 38,98 verglichen.

Frage: Welche Testentscheidung ergibt sich?

Antwort:

10.2.2 Test einer allgemeinen Nullhypothese

- Das Modell der Zweifachregression lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t .$$

- Der Test einer allgemeineren Nullhypothese wird anhand der konkreten Nullhypothese:

$$H_0 : \beta_1 + \beta_2 = 1 \quad \text{und gleichzeitig} \quad \alpha = 1 \quad (10.17)$$

erläutert.

- Die zur Nullhypothese (10.17) komplementäre Alternativhypothese lautet:

$$H_1 : \beta_1 + \beta_2 \neq 1 \quad \text{und/oder} \quad \alpha \neq 1 . \quad (10.18)$$

- Als Nullhypothesen-Modell ergibt sich:

$$\begin{aligned} y_t &= 1 + \beta_1 x_{1t} + (1 - \beta_1) x_{2t} + u_t^0 \\ &= 1 + \beta_1 (x_{1t} - x_{2t}) + x_{2t} + u_t^0 . \end{aligned} \quad (10.19)$$

Numerische Illustration 10.12

Schritt 1: Es wird die Nullhypothese (10.17) auf einem Signifikanzniveau von $\alpha = 5\%$ getestet.

Schritt 2: Wir wissen, dass $S_{\hat{u}\hat{u}} = 0,11481$. Für das Nullhypotesenmodell lässt sich mittels ökonomischer Standard-Software die Summe der Residuenquadrate $S_{\hat{u}\hat{u}}^0$ berechnen: $S_{\hat{u}\hat{u}}^0 = 0,87854$.

Schritt 3: Die Anzahl der Linearkombinationen L beträgt 2 und $T - K - 1$ ist wieder 27. Es ergibt sich somit

$$F = \frac{(0,87854 - 0,11481) / 2}{0,11481 / 27} \approx 89,80 .$$

(Forts.)

Wenn die Nullhypothese wahr ist, dann folgt die Zufallsvariable F einer $F_{(2,27)}$ -Verteilung.

Schritt 4: Aus Tabelle T.3 ist bereits bekannt, dass $F_{0,05} = 3,35$.

Schritt 5: Da $F > F_{0,05}$, muss die Nullhypothese verworfen werden.

- Zusammengefasst lautet das Vorgehen des F -Tests:

F -Test

Um eine aus L Linearkombinationen bestehende Nullhypothese (z.B. $H_0 : \beta_1 = q_1$ und $\beta_2 = q_1$) auf einem Signifikanzniveau von α zu testen, wird der F -Wert

$$F = \frac{(S_{\hat{u}\hat{u}}^0 - S_{\hat{u}\hat{u}}) / L}{S_{\hat{u}\hat{u}} / (T - K - 1)}$$

berechnet, wobei $S_{\hat{u}\hat{u}}^0$ die Summe der Residuenquadrate des Nullhypothesenmodells, $S_{\hat{u}\hat{u}}$ die Summe der Residuenquadrate des unrestringierten Modells und $T - K - 1$ seine Anzahl der Freiheitsgrade darstellt. Wenn $F > F_\alpha$, wird H_0 verworfen. Dabei wird F_α aus der F -Verteilung mit $\nu_1 = L$ und $\nu_2 = T - K - 1$ Freiheitsgraden ermittelt.

10.3.1 Zweiseitiger F -Test einer einzelnen Linearkombination

Numerische Illustration 10.13

Schritt 1: Signifikanzniveau $\alpha = 5\%$. Die Hypothese laute:

$$H_0 : \beta_1 + \beta_2 = 1 \quad \text{und} \quad H_1 : \beta_1 + \beta_2 \neq 1 .$$

Schritt 2: Als Nullhypothesenmodell ergibt sich dann

$$y_t = \alpha + \beta_1 (x_{1t} - x_{2t}) + x_{2t} + u_t^0 . \quad (10.20)$$

Ökonometrische Standard-Software errechnet für dieses Modell einen $S_{\hat{u}\hat{u}}^0$ -Wert von 0,11899. Für das ökonometrische Modell hatten wir $S_{\hat{u}\hat{u}} = 0,11481$.

(Forts.)

Schritt 3: Es gelten $L = 1$ und $T - K - 1 = 27$. Es ergibt sich somit

$$F = \frac{(0,11899 - 0,11481) / 1}{0,11481 / 27} \approx 0,982 .$$

Wenn die Nullhypothese wahr ist, dann folgt die Zufallsvariable F einer $F_{(1,27)}$ -Verteilung.

Schritt 4: Aus Tabelle T.3 kann der kritische Wert $F_{0,05}$ abgelesen werden.

Frage: Welcher $F_{0,05}$ -Wert ergibt sich?

Antwort:

(Forts.)

Schritt 5: F und $F_{0,05}$ werden miteinander verglichen.

Frage: Zu welcher Entscheidung gelangt man?

Antwort:

Beim t -Test kam man zu einem t -Wert von $-0,991$ und das Konfidenzintervall betrug $[-2,0518 ; 2,0518]$.

Frage: Wie verhält sich das Testergebnis zu demjenigen des F -Tests?

Antwort:

10.3.2 Probleme des F -Tests bei einseitigen Hypothesen

- Einseitiger Test der Null- und Alternativhypothese:

$$H_0 : \beta_1 + \beta_2 \geq 1 \quad \text{und} \quad H_1 : \beta_1 + \beta_2 < 1$$

- Einsetzen dieser Hypothese (genauer: $\beta_1 + \beta_2 = 1$) in das Modell der Zweifachregression,

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t ,$$

liefert das Nullhypothesenmodell

$$y_t = \alpha + \beta_1 (x_{1t} - x_{2t}) + x_{2t} + u_t^0 .$$

- Man erhält $S_{\hat{u}\hat{u}}^0$ und $S_{\hat{u}\hat{u}}$ und damit einen F -Wert.
- **Frage:** Warum stellt ein großer F -Wert nicht notwendigerweise einen Widerspruch zur Nullhypothese dar?

Antwort:

- Bei Signifikanzniveau α muss deshalb der $F_{2,\alpha}$ -Wert als kritischer Wert benutzt werden.

10.4 Zusammenhang zwischen t -Test und F -Test bei $L=2$

- Der Test der Nullhypothese

$$H_0 : \beta_1 = \beta_2 = 0,33 \quad (10.21)$$

gegen die Alternativhypothese

$$H_1 : \beta_1 \neq 0,33 \text{ und/oder } \beta_2 \neq 0,33$$

wurde als *simultaner* Parametertest bezeichnet.

- Für diesen Zweck wird ein F -Test eingesetzt.

- Bei *individuellen* Tests überprüft man die Null- und Alternativhypothese

$$H_0 : \beta_1 = 0,33 \quad \text{und} \quad H_1 : \beta_1 \neq 0,33 \quad (10.22)$$

separat von der Null- und Alternativhypothese

$$H_0 : \beta_2 = 0,33 \quad \text{und} \quad H_2 : \beta_2 \neq 0,33 . \quad (10.23)$$

- Für diesen Zweck werden t -Tests eingesetzt.

10.4.1 Numerisches Beispiel

Numerische Illustration 10.14

Individuelle t -Tests für $H_0 : \beta_1 = 0,33$ und $H_0 : \beta_2 = 0,33$

Es ist bereits bekannt, dass $\hat{\beta}_1 = 0,59652$, $\hat{\beta}_2 = 0,26255$,
 $\widehat{sd}(\hat{\beta}_1) = 0,13788$ und $\widehat{sd}(\hat{\beta}_2) = 0,03400$. Das
Signifikanzniveau betrage $\alpha = 5\%$. Es ergibt sich bei
Nullhypothese $H_0 : \beta_1 = 0,33$

$$t = \frac{0,59652 - 0,33}{0,13788} = 1,933$$

und bei Nullhypothese $H_0 : \beta_2 = 0,33$

$$t = \frac{0,26255 - 0,33}{0,03400} = -1,984 .$$

(Forts.)

Aus Tabelle T.2 kann man bei $\alpha = 5\%$ und 27 Freiheitsgraden den kritischen Wert $t_{\alpha/2}$ ablesen.

Frage: Welcher $t_{\alpha/2}$ -Wert ergibt sich und wie liegen die berechneten t -Werte relativ zum Intervall $[-t_{\alpha/2} ; t_{\alpha/2}]$?

Antwort:

(Forts.)

F -Test für $H_0 : \beta_1 = \beta_2 = 0,33$

Bereits bekannt ist, dass $S_{\widehat{uu}} = 0,11481$. Ökonometrische Software liefert $S_{\widehat{uu}}^0 = 0,14764$. Der F -Wert errechnet sich dann gemäß:

$$F = \frac{(0,14764 - 0,11481) / 2}{0,11481 / (30 - 3)} = 3,860 .$$

(Forts.)

Frage: Zu welchem $F_{0,05}$ -Wert und zu welchem Testergebnis gelangt man?

Antwort:

10.4.2 Unterschied zwischen individuellen und simultanen Parametertests

- Punkt A hat die Koordinaten $(\hat{\beta}_1, \hat{\beta}_2) = (0,59652, 0,26255)$.
- Ein t -Test vergleicht das Akzeptanzintervall $[q - t_{\alpha/2} \cdot \widehat{sd}(\hat{\beta}_k); q + t_{\alpha/2} \cdot \widehat{sd}(\hat{\beta}_k)]$ mit $\hat{\beta}_k$.
- Für $H_0 : \beta_1 = 0,33$ lautet dieses Intervall $[0,05; 0,61]$.
- Für $H_0 : \beta_2 = 0,33$ lautet es $[0,26; 0,40]$.
- **Frage:** Welche Testentscheidungen ergeben sich bei diesen beiden t -Tests?

Antwort:

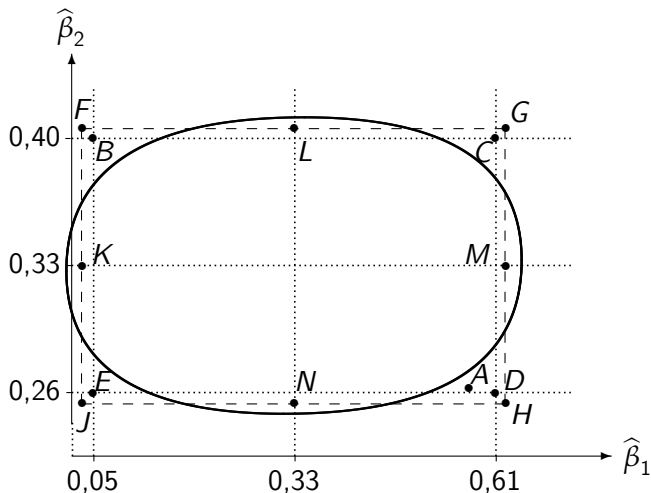


Abbildung 10.1: Der Vergleich der Toleranzintervalle zweier t -Tests mit der Toleranzregion eines F -Tests.

- Im F -Test wird dem beobachteten Wertepaar $(\hat{\beta}_1, \hat{\beta}_2)$ eine Akeptanzregion gegenübergestellt.
- Sie ist um den Punkt $(0,33, 0,33)$ zentriert und muss 95% Wahrscheinlichkeitsmasse abdecken.
- **Frage:** Deckt das Rechteck $BCDE$ 95% Masse ab?
Antwort:

- Der F -Test arbeitet mit der (flächenmäßig) *minimalen* Akzeptanzregion. Dies ist die eingezeichnete Ellipse.

- Wenn $cov(\hat{\beta}_1, \hat{\beta}_2) > 0$, besitzt die Ellipse eine ansteigende Schiefelage, bei $cov(\hat{\beta}_1, \hat{\beta}_2) < 0$ eine abfallende.
- **Frage:** Kann auch der Fall eintreten, dass die zwei individuellen t -Tests abgelehnt werden, der gemeinsame F -Test aber nicht?

Antwort:

Kapitel 11

Prognose

11.1 Punktprognose

11.1.1 Prognosewert und Prognosefehler

- Das geschätzte Modell der Zweifachregression lautet

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_{1t} + \hat{\beta}_2 x_{2t} . \quad (11.1)$$

- Die *Punktprognose* der endogenen Variable $\hat{y}_0 = \widehat{\ln g_0}$ ergibt sich aus

$$\hat{y}_0 = \hat{\alpha} + \hat{\beta}_1 x_{10} + \hat{\beta}_2 x_{20} . \quad (11.2)$$

Numerische Illustration 11.1

Das geschätzte Dünger-Modell lautet

$$\hat{y}_t = 0,95432 + 0,59652 \cdot x_{1t} + 0,26255 \cdot x_{2t} .$$

Für eine Parzelle mit 29 kg/ha Phosphat
($x_{10} = \ln 29 = 3,36730$) und 120 kg/ha Stickstoff
($x_{20} = \ln 120 = 4,78749$) ergibt sich aus dem geschätzten
Modell ein prognostizierter Wert von

$$\hat{y}_0 = \widehat{\ln g_0} = 4,21994 .$$

- Würden $\widehat{\ln g_0} = \ln \widehat{g_0}$ und damit $\widehat{y_0} = \ln \widehat{g_0}$ gelten, dann könnten wir, wegen $\widehat{g_0} = e^{\ln \widehat{g_0}}$, einfach die folgende Umrechnung vornehmen:

$$\widehat{g_0} = e^{\widehat{y_0}} . \quad (11.3)$$

- Wegen Jensens Ungleichung würde dieses Vorgehen jedoch zu einer systematischen Unterschätzung des Prognosewertes $\widehat{g_0}$ führen.
- Um die Verzerrung in $\widehat{g_0}$ zu vermeiden, muss man den Korrekturfaktor $e^{\widehat{\sigma}^2/2}$ aufnehmen:

$$\widehat{g_0} = e^{\widehat{\sigma}^2/2} e^{\widehat{y_0}} . \quad (11.4)$$

- **Frage:** Würde Formel (11.3) den Prognosewert über- oder unterschätzen?

Antwort:

Numerische Illustration 11.2

Im Dünger-Beispiel ist der Korrekturfaktor vernachlässigbar, denn $\hat{\sigma}^2 = 0,00425$. Der Wert des Korrekturfaktors weicht folglich nur minimal von 1 ab: $e^{0,00425/2} = 1,0021$. Mit Formel (11.4) erhalten wir

$$\hat{g}_0 = 1,0021 \cdot e^{4,2199} = 68,17 .$$

Der prognostizierte Output beträgt demnach 68,17 dt/ha.

11.1.2 Verlässlichkeit der Punktprognose

- Die Punktprognose

$$\hat{y}_0 = \hat{\alpha} + \hat{\beta}_1 x_{10} + \hat{\beta}_2 x_{20} \quad (11.5)$$

kann vom tatsächlich eintretenden Wert

$$y_0 = \alpha + \beta_1 x_{10} + \beta_2 x_{20} + u_0$$

abweichen.

- Der Prognosefehler beträgt somit

$$\hat{y}_0 - y_0 = \hat{\alpha} - \alpha + \hat{\beta}_1 x_{10} - \beta_1 x_{10} + \hat{\beta}_2 x_{20} - \beta_2 x_{20} - u_0 .$$

- Der Erwartungswert ergibt sich aus

$$\begin{aligned} E(\hat{y}_0 - y_0) &= E\left(\hat{\alpha} - \alpha + \hat{\beta}_1 x_{10} - \beta_1 x_{10} + \hat{\beta}_2 x_{20} - \beta_2 x_{20} - u_0\right) \\ &= E(\hat{\alpha} - \alpha) + E(\hat{\beta}_1 - \beta_1) x_{10} + E(\hat{\beta}_2 - \beta_2) x_{20} - E(u_0) \\ &= 0 + 0 + 0 - 0 = 0. \end{aligned} \quad (11.6)$$

- Die Varianz lautet

$$\begin{aligned} \text{var}(\hat{y}_0 - y_0) &= \sigma^2 \left(1 + 1/T\right) + (x_{10} - \bar{x}_1)^2 \text{var}(\hat{\beta}_1) \\ &\quad + (x_{20} - \bar{x}_2)^2 \text{var}(\hat{\beta}_2) \\ &\quad + 2(x_{10} - \bar{x}_1)(x_{20} - \bar{x}_2) \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \end{aligned} \quad (11.7)$$

Numerische Illustration 11.3

Wir wissen, dass

$$E(\hat{y}_0 - y_0) = 0$$

und dass

$$\hat{\sigma}^2 = 0,00425$$

$$\widehat{\text{var}}(\hat{\beta}_1) = 0,01900$$

$$\widehat{\text{var}}(\hat{\beta}_2) = 0,00116$$

$$\widehat{\text{cov}}(\hat{\beta}_1, \hat{\beta}_2) = 0,00003 .$$

(Forts.)

Formel (11.7) liefert folglich

$$\begin{aligned} & \widehat{\text{var}}(\widehat{y}_0 - y_0) \\ = & 0,00425 (1 + 1/30) \\ & + (3,36730 - 3,22577)^2 0,01900 \\ & + (4,78749 - 4,32415)^2 0,00116 \\ & + 2 (3,36730 - 3,22577) (4,78749 - 4,32415) 0,00003 \\ = & 0,00502 . \end{aligned}$$

11.2 Prognoseintervall

- Die Standardisierung des Prognosefehlers ergibt

$$t = \frac{\hat{y}_0 - y_0 - E(\hat{y}_0 - y_0)}{\widehat{sd}(\hat{y}_0 - y_0)} = \frac{\hat{y}_0 - y_0}{\widehat{sd}(\hat{y}_0 - y_0)},$$

wobei

$$\widehat{sd}(\hat{y}_0 - y_0) = \sqrt{\widehat{var}(\hat{y}_0 - y_0)}.$$

- Es gilt

$$t \sim t_{(T-3)} .$$

- Der t -Wert der Stichprobe fällt mit einer Wahrscheinlichkeit von $1 - a$ in das Konfidenzintervall $[-t_{a/2} ; t_{a/2}]$:

$$\Pr \left\{ -t_{a/2} \leq \frac{\hat{y}_0 - y_0}{\widehat{sd}(\hat{y}_0 - y_0)} \leq t_{a/2} \right\} = 1 - a .$$

- Auflösen nach y_0 liefert:

$$\Pr\left\{\hat{y}_0 - t_{a/2} \cdot \hat{sd}(\hat{y}_0 - y_0) \leq y_0 \leq \hat{y}_0 + t_{a/2} \cdot \hat{sd}(\hat{y}_0 - y_0)\right\} = 1 - a.$$

- Das gesuchte Prognoseintervall lautet folglich:

Formel: Prognoseintervall

$$\left[\hat{y}_0 - t_{a/2} \cdot \hat{sd}(\hat{y}_0 - y_0) ; \hat{y}_0 + t_{a/2} \cdot \hat{sd}(\hat{y}_0 - y_0) \right] \quad (11.8)$$

Numerische Illustration 11.4

Aus $\widehat{\text{var}}(\hat{y}_0 - y_0) = 0,00502$ folgt

$$\widehat{\text{sd}}(\hat{y}_0 - y_0) = \sqrt{0,00502} \approx 0,07088 .$$

Bei einem Signifikanzniveau von $\alpha = 5\%$ und 27 Freiheitsgraden ergibt sich aus Tabelle T.2 ein kritischer Wert von

$$t_{\alpha/2} = 2,0518 .$$

Unsere Punktprognose lautete

$$\hat{y}_0 = 4,21994 .$$

(Forts.)

Einsetzen dieser Werte in (11.8) ergibt ein Prognoseintervall von

$$[4,07452 ; 4,36536] .$$

Da

$$\ln 58,8 \approx 4,07452 \quad \text{und} \quad \ln 78,7 \approx 4,36536 ,$$

entspricht dies einem Gersten-Output, der zwischen 58,8 dt/ha und 78,7 dt/ha liegt.

Kapitel 12

Präsentation der Schätzergebnisse und deren computergestützte Berechnung

12.1 Computergestützte ökonomische Analyse

12.1.1 Ökonometrische Software

Tabelle 12.1: Eine Auswahl ökonomischer Softwarepakete.

<i>EViews</i>	www.eviews.com
<i>GAUSS</i>	www.aptech.com
<i>LIMDEP+NLOGIT</i>	www.limdep.com
<i>MICROFIT</i>	www.econ.cam.ac.uk/microfit
<i>PcGive</i>	www.pcgive.com
<i>R</i>	www.r-project.org/

Tabelle 12.1 (Forts.): Eine Auswahl ökonomischer Softwarepakete.

<i>RATS</i>	www.estima.com
<i>SAS/ETS</i>	www.sas.com
<i>SHAZAM</i>	http://econometrics.com/
<i>STATA</i>	www.stata.com
<i>STATISTICA</i>	www.statsoft.com
<i>TSP</i>	www.tspintl.com

12.1.2 Interpretation des Computeroutputs

Numerische Illustration 12.1

Tabelle 12.2: Typischer Computeroutput

Zahl der Beobachtungen	30			
Freiheitsgrade	27			
Bestimmtheitsmaß R^2	0,743			
Störgrößenvarianz $\hat{\sigma}^2$	0,00425			
F-Test: F -Wert	38,98			
p -Wert	0,00000001			
Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t -Wert	p -Wert
Konstante	0,95432	0,46943	2,0329	0,052
Phosphat	0,59652	0,13788	4,3265	<0,001
Stickstoff	0,26255	0,03400	7,7228	<0,001

12.2 Präsentation von Schätzergebnissen

Numerische Illustration 12.2

Die KQ-Schätzergebnisse des Dünger-Beispiels würde man in der folgenden Weise präsentieren:

$$\hat{y}_t = 0,95432 + 0,59652 x_1 + 0,26255 x_2 .$$

(0,46943) (0,13788) (0,0340) $\hat{sd}(\hat{\cdot})$ -Werte

(Forts.)

Oftmals wird auch statt der Standardabweichung $\widehat{sd}(\hat{\cdot})$ der jeweilige t -Wert angegeben.

Frage: Warum enthalten beide Darstellungsformen die gleichen Informationen?

Antwort:

- Regressionsergebnisse mit sehr vielen exogenen Variablen werden häufig in einer einfachen Tabellenform präsentiert.

Numerische Illustration 12.3

Für das Dünger-Beispiel würde sich die folgende tabellarische Präsentationsform anbieten.

Tabelle 12.1: KQ-Schätzergebnisse des Dünger-Beispiels in Tabellenform.

Variable	Koeffizient
Konstante	0,95432*
Phosphat	0,59652***
Stickstoff	0,26255***
<i>*$p < 10\%$, **$p < 5\%$, ***$p < 1\%$</i>	
F-Wert	38,98
Beobachtungen	30

Kapitel 13

Annahme A1: Variablenauswahl

- Das ökonometrische Modell einer Mehrfachregression lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t . \quad (13.1)$$

Annahme A1

In Gleichung (13.1) fehlen keine relevanten exogenen Variablen und die in Gleichung (13.1) benutzten exogenen Variablen sind nicht irrelevant.

- Eine Variablenauswahl ist gelungen, wenn das ökonometrische Modell
 - 1 keine relevanten exogenen Variablen auslässt und
 - 2 keine irrelevanten exogenen Variablen enthält.

13.1 Konsequenzen der Annahmeverletzung

Beispiel zu Kapitel 13

Es soll ermittelt werden, welche Faktoren in welchem Ausmaß für die Höhe des Lohnes y_t in einem Betrieb verantwortlich sind. Die Ausbildung x_{1t} , das Alter x_{2t} und die Dauer der Firmenzugehörigkeit x_{3t} werden als mögliche exogene Variablen erachtet.

(Forts.)

Tabelle 13.1: Lohnhöhe y_t , Ausbildungszeit x_{1t} , Alter x_{2t} und Dauer der Firmenzugehörigkeit x_{3t} der 20 Mitarbeiter eines Betriebes.

t	y_t	x_{1t}	x_{2t}	x_{3t}	t	y_t	x_{1t}	x_{2t}	x_{3t}
1	1250	1	28	12	11	1350	1	30	13
2	1950	9	34	8	12	1600	2	43	21
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	2000	4	58	30	20	1550	2	41	6

- Drei konkurrierende Modelle:

$$y_t = \alpha + \beta_1 x_{1t} + u'_t \quad (13.2)$$

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \text{ (korrektes Modell)} \quad (13.3)$$

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u''_t \quad (13.4)$$

wobei y_t die Höhe des Lohnes, x_{1t} die Ausbildungszeit, x_{2t} das Alter und x_{3t} die Firmenzugehörigkeit darstellen.

Numerische Illustration 13.1

Es werden für die drei konkurrierenden Modelle (13.2), (13.3) und (13.4) KQ-Schätzungen durchgeführt. Die Ergebnisse finden sich in Tabelle 13.2.

(Forts.)

Tabelle 13.2: Übersicht der Schätzergebnisse.

Modell	Variable	Koeff.	$\widehat{sd}(\cdot)$	t-Wert	p-Wert
(13.2)	Konstante	1354,7	94,2	14,377	< 0,001
	Ausbildung	89,3	19,8	4,505	< 0,001
(13.3)	Konstante	1027,8	164,5	6,249	< 0,001
	Ausbildung	62,6	21,2	2,953	0,009
	Alter	10,6	4,6	2,317	0,033
(13.4)	Konstante	1000,5	225,7	4,432	< 0,001
	Ausbildung	62,4	21,8	2,859	0,011
	Alter	12,4	10,7	1,159	0,263
	Firmenzug.	-2,6	14,3	-0,183	0,857

13.1.1 Auslassen relevanter Variablen

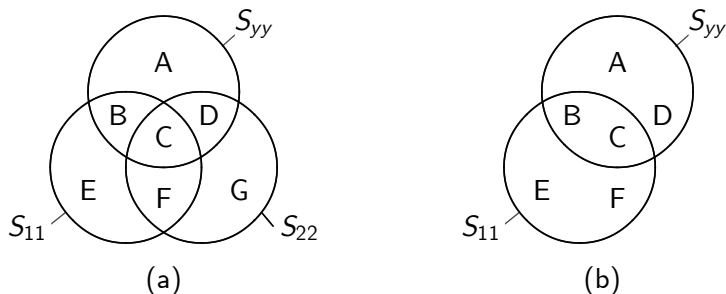


Abbildung 13.1: Der Wirkungszusammenhang des (a) korrekten und (b) unvollständigen Modells.

Auswirkungen auf den Erwartungswert der Störgrößen

- Für die Störgröße des fehlspezifizierten Modells u'_t gilt

$$u'_t = \beta_2 x_{2t} + u_t .$$

- Damit ergibt sich eine Verletzung der Annahme B1:

$$\begin{aligned} E(u'_t) &= E(\beta_2 x_{2t} + u_t) \\ &= \beta_2 x_{2t} + E(u_t) \\ &= \beta_2 x_{2t} + 0 \neq 0 . \end{aligned} \quad (13.5)$$

Konsequenzen für den Punktschätzer

- Für die KQ-Schätzung der Modelle (13.2) und (13.3) ergibt sich die Beziehung

$$\widehat{\beta}'_1 = \widehat{\beta}_1 + \widehat{\beta}_2 \frac{S_{12}}{S_{11}}. \quad (13.6)$$

- Folglich erhält man

$$\begin{aligned} E(\widehat{\beta}'_1) &= E\left(\widehat{\beta}_1 + \widehat{\beta}_2 \frac{S_{12}}{S_{11}}\right) \\ &= E(\widehat{\beta}_1) + E\left(\widehat{\beta}_2 \frac{S_{12}}{S_{11}}\right) \\ &= \beta_1 + \beta_2 \frac{S_{12}}{S_{11}}. \end{aligned} \quad (13.7)$$

Numerische Illustration 13.2

Im Lohn-Beispiel ergibt sich

$$\frac{S_{12}}{S_{11}} = \frac{448,9}{178,2} = 2,52 .$$

Aus Tabelle 13.2 ist bekannt, dass im korrekten Modell (13.3) $\hat{\beta}_1 = 62,6$ und $\hat{\beta}_2 = 10,6$. Folglich erhält man – wie in (13.6) behauptet –

$$\hat{\beta}'_1 = \hat{\beta}_1 + \hat{\beta}_2 \frac{S_{12}}{S_{11}} = 62,6 + 10,6 \cdot 2,52 = 89,3 .$$

- Für den KQ-Schätzer des Niveauparameters α erhält man:

$$E(\hat{\alpha}') = \alpha + \beta_2 \left(\bar{x}_2 - \frac{S_{12}}{S_{11}} \bar{x}_1 \right) .$$

Konsequenzen für Intervallschätzer

- Der Intervallschätzer lautet

$$\left[\hat{\beta}'_1 - t_{a/2} \cdot \widehat{sd}(\hat{\beta}'_1) ; \hat{\beta}'_1 + t_{a/2} \cdot \widehat{sd}(\hat{\beta}'_1) \right] .$$

- **Frage:** Warum ist dieser Schätzer verzerrt?

Antwort:

- Es muss $\text{var}(\hat{\beta}'_1)$ unverzerrt geschätzt werden, denn

$$\widehat{sd}(\hat{\beta}'_1) = \sqrt{\widehat{\text{var}}(\hat{\beta}'_1)} .$$

- Da

$$\text{var}(\hat{\beta}'_1) = \frac{\sigma^2}{S_{11}} , \quad (13.8)$$

benötigt man eine unverzerrte Schätzung von σ^2 .

- Die auf Basis des unvollständigen Modells (13.2) geschätzte Störgrößenvarianz lautet

$$\hat{\sigma}^2 = \frac{S_{\hat{u}'\hat{u}'}}{T - 2}. \quad (13.9)$$

- **Frage:** Warum liefert die Formel verzerrte Schätzungen?
Antwort:
- Da $\hat{\sigma}^2$ verzerrt ist, gilt dies auch für $\widehat{var}(\hat{\beta}'_1)$.

Numerische Illustration 13.3

Ökonometrische Software liefert für Modell (13.3):

$$S_{\hat{u}\hat{u}} = 957.698 \text{ und } \hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - 3) = 56.335.$$

Für das unvollständige Modell (13.2) erhält man

$$S_{\hat{u}'\hat{u}'} = 1.260.028 \text{ und damit aus (13.9) } \hat{\sigma}^2 = 70.002.$$

Da man anstelle von $\hat{\sigma}^2 = 56.335$ den verzerrten Wert $\hat{\sigma}^2 = 70.002$ in (13.8) einsetzt, ist $\widehat{var}(\hat{\beta}'_1)$ verzerrt.

Folglich ist auch der in Tabelle 13.2 angegebene Wert von $\widehat{sd}(\hat{\beta}'_1)$ unbrauchbar.

Konsequenzen für Hypothesentests

- Im t -Test wird der folgende t -Wert berechnet:

$$t = \frac{\widehat{\beta}'_1 - q}{\widehat{sd}(\widehat{\beta}'_1)}.$$

- Dabei ist $\widehat{\beta}'_1$ der Indikator für die Realität, also für das wahre β .
- Wird $\widehat{\beta}'_1$ aber verzerrt geschätzt, verliert er seine Indikatoreigenschaft.
- Dann sind t -Tests wertlos.
- Analoges gilt für F -Tests.

Relevante Variablen fehlen

Das Auslassen relevanter Variablen hat folgende Konsequenzen:

- 1 verzerre Punktschätzer,
- 2 verzerre Intervallschätzer und
- 3 wertlose Hypothesentests.

13.1.2 Verwendung irrelevanter Variablen

Auswirkungen auf den Erwartungswert der Störgröße

- Der Vergleich der Modelle (13.3) und (13.4) zeigt, dass

$$u_t = u_t'' + \beta_3 x_{3t} .$$

- Für die Störgröße u_t'' des fehlspezifizierten Modells gilt folglich

$$u_t'' = u_t - \beta_3 x_{3t} .$$

- **Frage:** Führt die irrelevante exogene Variable zu einer Verletzung der Annahme B1?

Antwort: Aufgrund von $\beta_3 = 0$ erhält man $u_t'' = u_t$ und damit $E(u_t'') = E(u_t) = 0$. Die Annahme bleibt unverletzt.

Konsequenzen für Punktschätzer

- Folglich bleiben die Punktschätzer trotz irrelevanter exogener Variable unverzerrt:

$$E(\hat{\alpha}_1'') = \alpha$$

$$E(\hat{\beta}_1'') = \beta_1$$

$$E(\hat{\beta}_2'') = \beta_2,$$

$$E(\hat{\beta}_3'') = \beta_3 = 0.$$

- Im korrekten Modell (13.3) beträgt die Varianz des KQ-Schätzers $\hat{\beta}_1$:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{11} (1 - R_{1.2}^2)} . \quad (13.10)$$

- Im fehlspezifizierten Modell (13.4) beträgt die Varianz des KQ-Schätzers $\hat{\beta}_1''$:

$$\text{var}(\hat{\beta}_1'') = \frac{\sigma^2}{S_{11} (1 - R_{1.23}^2)} . \quad (13.11)$$

- **Frage:** Warum haben irrelevante exogene Variablen also dennoch unerwünschte Nebenwirkungen?

Antwort:

Konsequenzen für Intervallschätzer und Hypothesentest

- Der Intervallschätzer lautet

$$\left[\hat{\beta}_1'' - t_{a/2} \cdot \widehat{sd}(\hat{\beta}_1'') ; \hat{\beta}_1'' + t_{a/2} \cdot \widehat{sd}(\hat{\beta}_1'') \right] .$$

- $\hat{\beta}_1''$ wird unverzerrt geschätzt.
- Es müssen auch $\text{var}(\hat{\beta}_1'')$ und damit σ^2 unverzerrt geschätzt werden.
- Für den Fall der Dreifachregression – wie Modell (13.4) – ist der Schätzer

$$\hat{\sigma}^2 = S_{\hat{u}''\hat{u}''} / (T - 4) . \quad (13.12)$$

unverzerrt.

- Einsetzen in (13.11) liefert einen unverzerrten Wert $\widehat{\text{var}}(\hat{\beta}_1'')$.
- Der Intervallschätzer ist folglich unverzerrt.

- Der Intervallschätzer ist allerdings nicht effizient, denn $var(\hat{\beta}_1'') > var(\hat{\beta}_1)$.
- Auch Hypothesentests sind zulässig, aber die Testschärfe ist unnötig gering.

Irrelevante Variablen sind enthalten

Die Aufnahme irrelevanter Variablen hat folgende Konsequenzen:

- 1 unverzerzte, aber ineffiziente Punktschätzer,
- 2 unverzerzte, aber ineffiziente Intervallschätzer und
- 3 verwendbare, aber unnötig unscharfe Hypothesentests.

13.2 Diagnose und Neu-Spezifikation

- Jede zusätzliche Variable
 - verringert das *Verzerrungsrisiko* und
 - erhöht die *Schätzvarianz*.
- Die nachfolgenden Kriterien der Variablenauswahl wägen diese beiden Effekte in unterschiedlicher Weise gegeneinander ab.

13.2.1 Korrigiertes Bestimmtheitsmaß \bar{R}^2

- Das Bestimmtheitsmaß

$$R^2 = \frac{S_{yy} - S_{\hat{u}\hat{u}}}{S_{yy}} = 1 - \frac{S_{\hat{u}\hat{u}}}{S_{yy}} \quad (13.13)$$

ist nur dann ein sinnvolles Kriterium, wenn folgende Bedingungen erfüllt sind:

- 1 die endogene Variable der Modelle ist identisch, das heißt ihre numerischen Werte stimmen überein,
 - 2 die Anzahl der exogenen Variablen ist in den Modellen identisch,
 - 3 die Modelle besitzen einen Niveauparameter α .
- **Frage:** Sind diese drei Bedingungen im Lohn-Beispiel erfüllt?

Antwort:

Numerische Illustration 13.4

Ökonometrische Software liefert für die drei Lohn-Modelle folgende Bestimmtheitsmaße und Schätzvarianzen für $\hat{\alpha}$ und $\hat{\beta}_1$:

Tabelle 13.3: Bestimmtheitsmaße und Schätzvarianzen der drei zu vergleichenden Modelle.

Modell	R² in %	$\widehat{\text{var}}(\hat{\alpha})$	$\widehat{\text{var}}(\hat{\beta}_1)$
(13.2)	52,99	8.877	392,824
(13.3)	64,27	27.052	449,043
(13.4)	64,34	50.953	477,745

- **Frage:** Warum ist der Anstieg des Bestimmtheitsmaßes bei zusätzlichen exogenen Variablen kein Zufall?

Antwort:

- Das Bestimmtheitsmaß ist folglich ungeeignet für den Vergleich von Modellen mit unterschiedlich vielen Variablen.

- Deshalb wurde ein modifiziertes Bestimmtheitsmaß entwickelt.

Formel: Korrigiertes Bestimmtheitsmaß

$$\bar{R}^2 = 1 - \frac{S_{\hat{u}\hat{u}} / (T - K - 1)}{S_{yy} / (T - 1)} \quad (13.14)$$

$$\begin{aligned} &= 1 - \frac{S_{\hat{u}\hat{u}} (T - 1)}{S_{yy} (T - K - 1)} \\ &= 1 - (1 - R^2) \frac{T - 1}{T - K - 1} \quad (13.15) \end{aligned}$$

- **Frage:** Wo zeigt sich beim korrigierten Bestimmtheitsmaß das Abwägen zwischen Verzerrungsrisiko und Schätzvarianz?

Antwort:

Numerische Illustration 13.5

Vergleich der drei Lohn-Modelle:

$$\text{Modell (13.2)} : \bar{R}^2 = 1 - (1 - 0,5299) \frac{20 - 1}{20 - 2} = 0,5038$$

$$\text{Modell (13.3)} : \bar{R}^2 = 1 - (1 - 0,6427) \frac{20 - 1}{20 - 3} = 0,6007$$

$$\text{Modell (13.4)} : \bar{R}^2 = 1 - (1 - 0,6434) \frac{20 - 1}{20 - 4} = 0,5766$$

Das korrigierte Bestimmtheitsmaß spricht folglich für Modell (13.3).

13.2.2 Weitere Kennzahlen: AIC, SC und PC

Weitere Auswahlkriterien

Für die Variablenauswahl können auch das Akaike-Kriterium (*AIC*), das Schwarz-Kriterium (*SC*) oder das Prognosekriterium (*PC*) herangezogen werden:

$$AIC = \ln \left(\frac{S_{\hat{u}\hat{u}}}{T} \right) + \frac{2(K+1)}{T} \quad (13.16)$$

$$SC = \ln \left(\frac{S_{\hat{u}\hat{u}}}{T} \right) + \frac{(K+1) \ln T}{T} \quad (13.17)$$

$$PC = \frac{S_{\hat{u}\hat{u}} [1 + (K+1)/T]}{T - K - 1} \quad (13.18)$$

- **Frage:** Wo zeigt sich in diesen Formeln das Abwägen zwischen Verzerrungsrisiko und Schätzvarianz?

Antwort:

- **Frage:** Sind gute Modelle bei diesen Kennzahlen durch kleine oder große Werte gekennzeichnet.

Antwort:

Numerische Illustration 13.6

Für die drei zu vergleichenden Modelle (13.2), (13.3) und (13.4) erhält man die $S_{\hat{u}\hat{u}}$ -Werte 1.260.028, 957.698 und 955.692. Die Formeln (13.16), (13.17) und (13.18) liefern dann folgende Ergebnisse:

Tabelle 13.4: AIC, SC und PC der drei zu vergleichenden Modelle.

Modell	AIC	SC	PC
(13.2)	11,251	11,350	77.001,715
(13.3)	11,077	11,226	61.968,701
(13.4)	11,174	11,374	71.676,877

Die Kennzahlen sprechen also erneut für Modell (13.3).

13.2.3 F -Test

- Um zwischen den Modellen (13.2) und (13.4) abzuwägen, könnte man einen F -Test für $H_0 : \beta_2 = \beta_3 = 0$ durchführen.
- Ein F -Wert berechnet sich immer gemäß

$$F = \frac{(S_{\hat{u}\hat{u}}^0 - S_{\hat{u}\hat{u}}) / L}{S_{\hat{u}\hat{u}} / (T - K - 1)} . \quad (13.19)$$

Numerische Illustration 13.7

Ein F -Test der Nullhypothese $H_0 : \beta_2 = \beta_3 = 0$ liefert im Modell (13.4):

$$F = \frac{(1.260.028 - 955.692) / 2}{955.692 / (20 - 4)} = 2,548 . \quad (13.20)$$

Frage: Zu welchem Testergebnis gelangen wir?

Antwort:

13.2.4 t -Test

- Um zwischen (13.3) und (13.4) abzuwägen, wird die Nullhypothese $H_0 : \beta_3 = 0$ getestet:

$$t = \frac{\hat{\beta}_3 - q}{\widehat{sd}(\hat{\beta}_3)} = \frac{\hat{\beta}_3 - 0}{\widehat{sd}(\hat{\beta}_3)} . \quad (13.21)$$

- Normalerweise sollte für sämtliche exogene Variablen der zu vergleichenden Modelle der jeweilige t -Wert ermittelt werden.

Numerische Illustration 13.8

Der kritische Wert $t_{\alpha/2}$ liegt bei einem Signifikanzniveau von $\alpha = 0,05$ für Modell (13.3) bei 2,1098 (17 Freiheitsgrade) und für Modell (13.4) bei 2,1199 (16 Freiheitsgrade). Die Ergebnisse in Tabelle 13.2 sprechen insgesamt für Modell (13.3), also für das korrekte Modell.

13.2.5 Zusammenhang zwischen korrigiertem Bestimmtheitsmaß, F -Test und t -Test

- Es sei \bar{R}^2 das korrigierte Bestimmtheitsmaß des unrestringierten Modells, z.B. Modell (13.4).
- \bar{R}^{0^2} sei das korrigierte Bestimmtheitsmaß des Nullhypothesenmodells, z.B. Modell (13.3).

F -Test und korrigiertes Bestimmtheitsmaß

Es gilt $\bar{R}^2 > \bar{R}^{0^2}$ genau dann, wenn $F > 1$.

- Wenn nur eine Linearkombination getestet wird ($H_0 : \beta_k = 0$), kann auch ein t -Test eingesetzt werden und es gilt $F = t^2$.
- Folglich gilt $\bar{R}^2 > \bar{R}^0{}^2$ genau dann, wenn $t^2 > 1$, also wenn t außerhalb des Intervalls $[-1,1]$ liegt.
- Bei den üblichen Signifikanzniveaus sind F_a und t_a immer größer als 1.
- Folglich neigt das Kriterium des korrigierten Bestimmtheitsmaßes stärker zur Aufnahme neuer Variablen als das Kriterium der F - und t -Tests.

13.2.6 Ungenesteter F -Test

- Die Restriktion $\beta_3 = 0$ macht aus Modell (13.4) das Modell (13.3).
- Man sagt dann, Modell (13.4) ist „eingebettet“ in Modell (13.3).
- Üblicher ist der aus dem Englischen abgeleitete Begriff „genestet“.
- Alternativkandidat zu Modell (13.3) sei das Modell:

$$y_t = \alpha + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t''' . \quad (13.22)$$

Numerische Illustration 13.9

Tabelle 13.5: Schätzergebnisse für Modell (13.22).

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	921,4	267,2	3,449	0,003
Alter	20,7	12,2	1,691	0,109
Firmenzugehör.	-4,1	17,0	-0,241	0,812

Bei einem Signifikanzniveau von 5% sind weder das Alter noch die Firmenzugehörigkeit signifikant.

- Der ungenestete F -Test geht in zwei Schritten vor:
 - ① Aus den beiden Modellen wird ein “Mega-Modell” gebildet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t$$

[entspricht Modell (13.4)]

- ② Für dieses Modell werden hintereinander die Nullhypothesen $H_0 : \beta_1 = 0$ und $H_0^* : \beta_3 = 0$ jeweils mit einem F -Test überprüft (t -Tests wären hier natürlich auch zulässig).

Numerische Illustration 13.10

Wenn $L = 1$, kann die Beziehung $F = t^2$ ausgenutzt werden.
Für $H_0 : \beta_1 = 0$ liefert das Mega-Modell

$$F = t^2 = 2,859^2 = 8,174 \quad (\text{siehe Tabelle 13.2})$$

und für $H_0 : \beta_3 = 0$

$$F = t^2 = -0,183^2 = 0,033 \quad (\text{siehe Tabelle 13.2}).$$

Der kritische F -Wert beträgt $F_{0,05} = 4,494$.

Frage: Welches Modell ziehen wir demnach vor?

Antwort:

13.3 Spezifikations-Methodologien

- Die *Maurer-Methodologie* beginnt mit einem schlanken Modell und fügt zusätzliche Variablen hinzu.
- Alternative Bezeichnungen sind:
 - bottom-up
 - specific-to-general
 - simple-to-general
- Die *Steinmetz-Methodologie* beginnt mit einem umfassenden Modell und entfernt anschließend Variablen.
- Alternative Bezeichnungen sind:
 - top-down
 - general-to-specific
 - general-to-simple.

13.3.1 Steinmetz- versus Maurer-Methodologie

- **Frage:** Warum wird die Steinmetz-Methodologie tendenziell bevorzugt?

Antwort:

13.3.2 Wichtiges Problem bei der Variablenauswahl

Numerische Illustration 13.11

Tabelle 13.2 sei die Entscheidungsgrundlage, um zwischen den drei konkurrierenden Modellen (13.2), (13.3) und (13.4) zu entscheiden.

Frage: Warum könnte die Steinmetz-Methodologie sowohl zu Modell (13.2) als auch zu Modell (13.3) führen?

Antwort:

Kapitel 14

Annahme A2: Funktionale Form

- Das ökonometrische Modell einer Mehrfachregression lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t .$$

Annahme A2

Der wahre Zusammenhang zwischen $x_{1t}, x_{2t}, \dots, x_{Kt}$ und y_t ist linear.

Beispiel zu Kapitel 14

Die Milchleistung (m) einer Kuh hängt bei Konstanthaltung aller anderen Inputs ausschließlich von der zugefütterten Menge des Kraftfutters (f) ab.

Das ökonomische Modell lautet demnach:

$$m = g(f) . \quad (14.1)$$

Die Daten finden sich in Tabelle 14.1 und die entsprechende Punktwolke in Abbildung 14.1.

(Forts.)

Tabelle 14.1: Kraftfutter f_t (in Zentner/Jahr wobei 1 Zentner = 50 kg) und Milchleistung m_t (in Liter/Jahr) von 12 beobachteten Kühen.

t	f_t	m_t	t	f_t	m_t
1	10	6525	7	8	5821
2	30	8437	8	14	7531
⋮	⋮	⋮	⋮	⋮	⋮
6	22	7236	12	28	8112

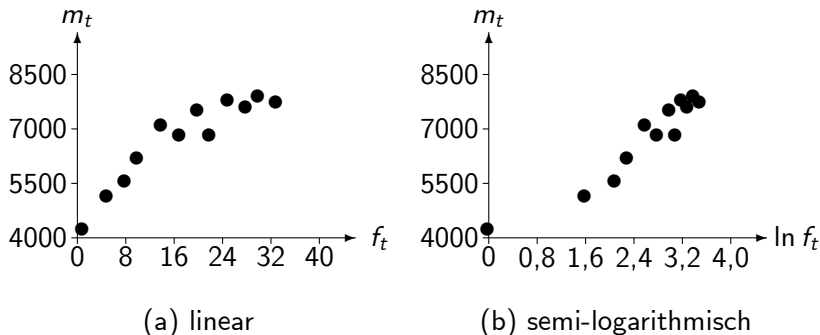


Abbildung 14.1: Die (a) Originaldaten und (b) ein Versuch, den Zusammenhang zwischen Kraftfuttereinsatz f_t und Milchleistung m_t in linearisierter Form abzubilden.

- Das zum ökonomischen Modell (14.1) korrespondierende lineare ökonometrische Modell lautet:

$$m_t = \alpha + \beta f_t + u_t . \quad (14.2)$$

- **Frage:** Welchen Eindruck vermittelt die Punktwolke hinsichtlich der Annahme A2? Was lässt sich über die Grenzerträge in der Milchproduktion aussagen?

Antwort:

14.1 Konsequenzen der Annahmeverletzung

- **Frage:** Warum sollte keine KQ-Schätzung des Modells (14.2) vorgenommen werden?

Antwort:

14.2 Einige alternative Funktionsformen

- Nicht-lineare Wirkungszusammenhänge lassen sich oftmals in linearer Form wiedergeben.
- Verschiedene nicht-lineare Beziehungsvarianten zwischen f_t und m_t werden nachfolgend als lineare Beziehung zwischen zwei Variablen formalisiert, die selbst jeweils Funktionen von f_t und m_t sind.

14.2.1 Semi-logarithmisches Modell (Linlog-Modell)

- Der lineare Zusammenhang zwischen $\ln f_t$ und m_t (siehe Abbildung 14.1) lautet

$$m_t = \alpha + \beta \ln f_t + u_t . \quad (14.3)$$

- Umbenennung von $m_t = y_t$ und $\ln f_t = x_t$ ergibt das lineare Modell

$$y_t = \alpha + \beta x_t + u_t . \quad (14.4)$$

- Abbildung 14.1 zeigt die entsprechende Punktwolke.
- **Frage:** Erfüllt Modell (14.3) Annahme A2?
Antwort:

- **Frage:** Welche Interpretation besitzt β im Modell (14.3)?

Antwort:

- β ist auch für die Beziehung zwischen f_t und m_t bedeutsam.
- Differenziert man $m_t = \alpha + \beta \ln f_t + u_t$ nach f_t , erhält man die marginale Veränderung

$$\frac{dm_t}{df_t} = \frac{dm_t}{d(\ln f_t)} \cdot \frac{d(\ln f_t)}{df_t} = \beta \cdot \frac{1}{f_t} . \quad (14.5)$$

- **Frage:** Was bedeutet dieses Resultat hinsichtlich der Grenzerträge der Milchproduktion?

Antwort:

- Umstellen von (14.5) liefert

$$\begin{aligned} dm_t &= \beta \cdot \frac{df_t}{f_t} \\ &= \frac{\beta}{100} \cdot \left[100 \cdot \frac{df_t}{f_t} \right] \\ &= \frac{\beta}{100} \cdot [\text{Prozentänderung in } f_t] . \end{aligned}$$

- Bei einer *einprozentigen* Erhöhung des Kraftfuttereinsatzes verändert sich der Milchoutput folglich um $\beta/100$ Liter.
- **Frage:** Um wie viel Liter würde sich der Milchoutput bei einer zehnpromzentigen Erhöhung des Kraftfuttereinsatzes verändern?

Antwort:

14.2.2 Inverses Modell

- Der lineare Zusammenhang zwischen den Variablen $1/f_t$ und m_t lautet

$$m_t = \alpha + \beta (1/f_t) + u_t . \quad (14.6)$$

- Umbenennung von $m_t = y_t$ und $1/f_t = x_t$ ergibt das lineare Modell (14.4).
- Die Punktwolke ist in Abbildung 14.2 (a) wiedergegeben.
- Der Parameter β scheint hier negativ zu sein.
- **Frage:** Wie ist β im Modell (14.6) zu interpretieren?
Antwort:

- **Frage:** Ist das inverse Modell (14.6) eine geeignete Beschreibung der Realität in linearer Form?

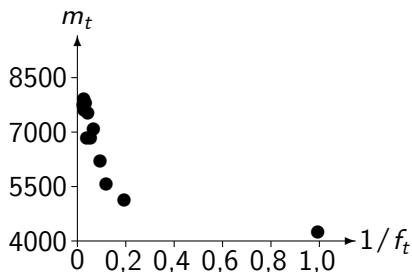
Antwort:

- Differenzieren des Modells (14.6) nach f_t , ergibt die marginale Veränderung

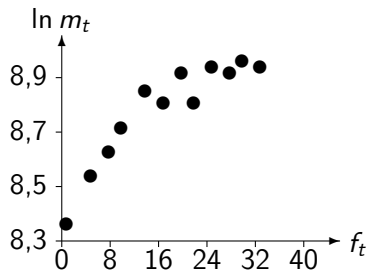
$$\frac{dm_t}{df_t} = -\frac{\beta}{f_t^2}.$$

- Wie beim semi-logarithmischen Modell hängt die marginale Veränderung sowohl von β als auch von f_t ab.
- **Frage:** Würden bei negativem β abnehmende oder zunehmende Grenzerträge vorliegen?

Antwort:



(a) invers



(b) exponential

Abbildung 14.2: Zwei weitere Versuche, den Zusammenhang zwischen Krafffuttereinsatz f_t und Milchleistung m_t in linearisierter Form abzubilden.

14.2.3 Exponential-Modell (Loglin-Modell)

- Der lineare Zusammenhang zwischen f_t und $\ln m_t$ lautet

$$\ln m_t = \alpha + \beta f_t + u_t . \quad (14.7)$$

- **Frage:** Ist angesichts der Punktwolke in Abbildung 14.2 (b) das Exponential-Modell (14.7) eine geeignete Beschreibung der Realität in linearer Form?

Antwort:

- Differenzieren nach f_t ergibt

$$\frac{d(\ln m_t)}{df_t} = \beta . \quad (14.8)$$

- Da

$$\frac{d(\ln m_t)}{dm_t} = \frac{1}{m_t} ,$$

gilt

$$d(\ln m_t) = \frac{dm_t}{m_t} . \quad (14.9)$$

- Ausdruck (14.8) ist deshalb identisch mit

$$\frac{dm_t}{m_t} = \beta df_t . \quad (14.10)$$

- Beidseitige Multiplikation mit 100 ergibt

$$100 \frac{dm_t}{m_t} = 100 \beta df_t$$

und damit

$$[\text{Prozenterhöhung in } m_t] = 100 \beta df_t .$$

- **Frage:** Wie ist diese Gleichung zu interpretieren?
Antwort:
- Oftmals wird ein solches Maß als *Semi-Elastizität* bezeichnet.

- In exponentieller Form geschrieben lautet Gleichung (14.7):

$$e^{\ln m_t} = e^{\alpha + \beta f_t} ,$$

wobei u_t gleich 0 gesetzt wurde.

- Da $e^{\ln m_t} = m_t$, erhält man

$$m_t = e^{\alpha + \beta f_t} .$$

- Die marginale Veränderung dm_t/df_t beträgt folglich

$$\frac{dm_t}{df_t} = \beta e^{\alpha + \beta f_t} .$$

- **Frage:** Was würde dieses Ergebnis bezüglich der Grenzerträge des Kraftfuttereinsatzes bedeuten?

Antwort:

14.2.4 Logarithmisches Modell (Loglog-Modell)

- Der lineare Zusammenhang zwischen den Variablen $\ln f_t$ und $\ln m_t$ lautet

$$\ln m_t = \alpha + \beta \ln f_t + u_t . \quad (14.11)$$

- **Frage:** Ist angesichts der Punktwolke in Abbildung 14.3 (a) das logarithmische Modell (14.7) eine geeignete Beschreibung der Realität in linearer Form?

Antwort:

- Differenzieren nach $\ln f_t$ ergibt

$$\frac{d(\ln m_t)}{d(\ln f_t)} = \beta .$$

- Es ergibt sich auch hier wieder die gewohnte Interpretation eines Steigungsparameters.

- Da

$$\frac{d(\ln m_t)}{dm_t} = \frac{1}{m_t} \quad \text{und} \quad \frac{d(\ln f_t)}{df_t} = \frac{1}{f_t},$$

erhält man

$$\begin{aligned} \beta &= \frac{dm_t/m_t}{df_t/f_t} && (14.12) \\ &= \frac{100 \cdot dm_t/m_t}{100 \cdot df_t/f_t} = \frac{[\text{Prozenterhöhung in } m_t]}{[\text{Prozenterhöhung in } f_t]}. \end{aligned}$$

- Umstellen liefert

$$[\text{Prozenterhöhung in } m_t] = \beta \cdot [\text{Prozenterhöhung in } f_t]$$

- **Frage:** Wie ist diese Gleichung zu interpretieren?

Antwort:

- Es handelt sich bei β also um die *Elastizität* der Milchleistung m_t in Bezug auf den Kraftfuttereinsatz f_t .
- Bildet man die exponentielle Form der Gleichung (14.11), wobei u_t gleich 0 gesetzt wird, so ergibt sich

$$\begin{aligned} e^{\ln m_t} &= e^{\alpha + \beta \ln f_t} \\ &= e^{\alpha} e^{\beta \ln f_t} \end{aligned}$$

- Wegen

$$e^{\ln m_t} = m_t \quad \text{und} \quad e^{\beta \ln f_t} = e^{\ln f_t^{\beta}} = f_t^{\beta},$$

erhält man

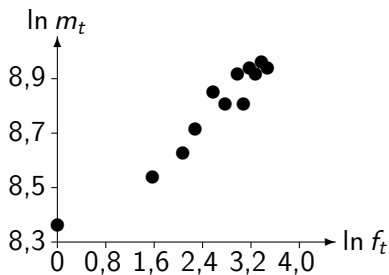
$$m_t = e^{\alpha} f_t^{\beta}. \quad (14.13)$$

- Auflösen von (14.12) nach dm_t/df_t und Ersetzen von m_t durch den Ausdruck in (14.13) liefert

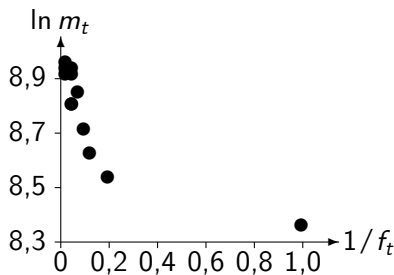
$$\frac{dm_t}{df_t} = e^\alpha \beta f_t^{\beta-1} .$$

- **Frage:** Was besagt dieses Ergebnis bezüglich der Grenzerträge des Kraftfuttereinsatzes?

Antwort:



(a) logarithmisch



(b) log-invers

Abbildung 14.3: Zwei weitere Versuche, den Zusammenhang zwischen Kraftfuttereinsatz f_t und Milchleistung m_t in linearisierter Form abzubilden.

14.2.5 Log-inverses Modell

- Der lineare Zusammenhang zwischen den Variablen $1/f_t$ und $\ln m_t$ lautet

$$\ln m_t = \alpha + \beta (1/f_t) + u_t . \quad (14.14)$$

- Abbildung 14.3 (b) spricht gegen die Verwendung dieses Modells.
- Setzt man u_t gleich 0, so lautet die exponentielle Form der Gleichung (14.14):

$$m_t = e^{\alpha + \beta(1/f_t)} .$$

- Differenzieren nach f_t liefert die marginale Veränderung:

$$\frac{dm_t}{df_t} = -\beta \frac{e^{\alpha + \beta(1/f_t)}}{f_t^2} .$$

14.2.6 Quadratisches Modell

- Der lineare Zusammenhang zwischen den Variablen f_t , f_t^2 und m_t lautet

$$m_t = \alpha + \beta_1 f_t + \beta_2 f_t^2 + u_t . \quad (14.15)$$

- Mit den Umbenennungen $m_t = y_t$, $f_t = x_{1t}$ und $f_t^2 = x_{2t}$ ergibt sich das lineare Modell

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t .$$

- Differenzieren der Gleichung (14.15) liefert

$$\frac{dm_t}{df_t} = \beta_1 + 2\beta_2 f_t .$$

- **Frage:** Was lässt sich für $\beta_1 > 0$ und $\beta_2 < 0$ bezüglich der Grenzerträge des Kraftfuttereinsatzes aussagen?

Antwort:

14.2.7 Eine vergleichende Anwendung

Tabelle 14.2: Verschiedene Funktionstypen in ihrer linearisierten Form und mit ihren marginalen Veränderungen.

Funktionstyp	lineare Funktion	marg. Veränd. (dm_t/df_t)
linear	$m_t = \alpha + \beta f_t + u_t$	β
semi-log.	$m_t = \alpha + \beta \ln f_t + u_t$	β / f_t
invers	$m_t = \alpha + \beta(1/f_t) + u_t$	$-\beta / f_t^2$
exponential	$\ln m_t = \alpha + \beta f_t + u_t$	$\beta e^{\alpha + \beta f_t}$
logarithmisch	$\ln m_t = \alpha + \beta \ln f_t + u_t$	$e^{\alpha} \beta f_t^{\beta-1}$
log-invers	$\ln m_t = \alpha + \beta(1/f_t) + u_t$	$-\beta e^{\alpha + \beta(1/f_t)} / f_t^2$
quadratisch	$m_t = \alpha + \beta_1 f_t + \beta_2 f_t^2 + u_t$	$\beta_1 + 2\beta_2 f_t$

Numerische Illustration 14.1

Tabelle 14.3: Eine Übersicht der Schätzergebnisse des Milch-Beispiels.

Modell	Funktionstyp	$\hat{\alpha}$	$\hat{\beta}$	R^2
(14.2)	linear	4985,270	118,914	85,6
(14.3)	semi-logarith.	3818,334	1268,803	90,5
(14.6)	invers	7657,231	-3762,321	59,5
(14.7)	exponential	8,524	0,018	82,0
(14.11)	logarithmisch	8,323	0,204	94,0
(14.14)	log-invers	8,943	-0,631	67,6
			$\hat{\beta}_1$ $\hat{\beta}_2$	
(14.15)	quadratisch	4109,445	271,393 -4,432	95,0

- **Frage:** Welche Interpretation besitzt $\hat{\beta}$ im linearen, im Exponential-, im semi-logarithmischen und im logarithmischen Modell?

Antwort: *lineares Modell:* Ein zusätzlicher Zentner Kraftfutter erhöht den Milchoutput um $\hat{\beta} = 118,914$ Einheiten (Liter).

Exponential-Modell: Ein zusätzlicher Zentner Kraftfutter erhöht den Milchoutput um $(100 \cdot \hat{\beta}) = (100 \cdot 0,018)$ Prozent, also um 1,8 Prozent.

semi-logarithmisches Modell: Eine einprozentige Erhöhung des Kraftfuttoreinsatzes erhöht den Milchoutput um $\hat{\beta}/100 = 12,68803$ Einheiten (Liter).

logarithmisches Modell: Eine einprozentige Erhöhung des Kraftfuttoreinsatzes erhöht den Milchoutput um $\hat{\beta} = 0,204$ Prozent (nicht 20,4 Prozent!).

Umgang mit nicht-linearen Wirkungszusammenhängen

Nicht-lineare Zusammenhänge können oftmals in ein lineares Modell transformiert und so einer KQ-Schätzung zugänglich gemacht werden.

Bei der Interpretation der Parameter ist große Sorgfalt geboten. Die Interpretation hängt vom gewählten Funktionstyp des Modells (z.B. logarithmisch) ab.

14.3 Diagnose und Neu-Spezifikation

- Nur manchmal macht die ökonomische Theorie Vorgaben in Bezug auf die funktionale Form.
- Für die Auswahl der besten funktionalen Form helfen grafische Verfahren nicht immer weiter.
- Es müssen deshalb Spezifikations- und Fehlspezifikationstests eingesetzt werden.

14.3.1 Regression Specification Error Test (RESET)

Grundidee des RESET-Verfahrens

- Die endogene Variable ist eine Funktion der exogenen Variable: $y_t = f(x_t)$.
- Eine Taylor-Expansion liefert:

$$y_t = f(x_t) \approx \alpha + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \dots \quad (14.16)$$

- Wäre der wahre Wirkungszusammenhang nicht linear und folglich in der Form (14.16) approximierbar und würde man dennoch das einfache lineare Modell

$$y_t = \alpha + \beta_1 x_t + u'_t \quad (14.17)$$

verwenden, dann würden in diesem Modell die relevanten Variablen x_t^2 , x_t^3 , usw. fehlen.

- Man kann das Modell (14.17) um Variablen ergänzen, in welchen Informationen über x_t^2 , x_t^3 , usw. stecken, und anschließend überprüfen, ob die ergänzten Variablen relevant sind.
- Es sei unterstellt, dass Modell (14.17) korrekt ist.
- Dann gilt

$$\hat{y}_t = \hat{\alpha} + \hat{\beta}_1 x_t .$$

- Die quadrierten geschätzten Werte ergeben sich aus

$$\hat{y}_t^2 = \left(\hat{\alpha} + \hat{\beta}_1 x_t \right)^2 = \hat{\alpha}^2 + \hat{\beta}_1^2 x_t^2 + 2\hat{\alpha}\hat{\beta}_1 x_t . \quad (14.18)$$

- Analoge Ausdrücke lassen sich für \hat{y}_t^3 , \hat{y}_t^4 , usw. angeben.

- Das um Informationen der Variablen x_t^2 , x_t^3 , usw. ergänzte Modell (14.17) lautet:

$$y_t = \alpha + \beta_1 x_t + \gamma_1 \hat{y}_t^2 + \gamma_2 \hat{y}_t^3 + \gamma_3 \hat{y}_t^4 + u_t^* . \quad (14.19)$$

- **Frage:** Mit welchem Hypothesentest könnte man nun überprüfen, ob das lineare Modell (14.17) korrekt ist?
Antwort:

Arbeitsschritte des RESET-Verfahrens

1. Man ermittelt für das lineare Modell die geschätzten Werte der endogenen Variable \hat{y}_t und die Summe der Residuenquadrate $S'_{\hat{u}\hat{u}}$. Ferner berechnet man die Werte der Variablen \hat{y}_t^2 , \hat{y}_t^3 und \hat{y}_t^4 .
2. Das lineare Modell wird erweitert zu Modell (14.19). Die Residuenquadrate werden berechnet ($S^*_{\hat{u}\hat{u}}$).
3. Für das erweiterte Modell (14.19) wird ein F -Test der Nullhypothese $H_0 : \gamma_1 = \gamma_2 = \gamma_3 = 0$ durchgeführt. Der entsprechende F -Wert errechnet sich aus

$$F = \frac{(S'_{\hat{u}\hat{u}} - S^*_{\hat{u}\hat{u}}) / L}{S^*_{\hat{u}\hat{u}} / (T - K^* - 1)} .$$

4. Falls die Nullhypothese wahr ist, folgt die Zufallsvariable F einer $F_{(L, T-K^*-1)}$ -Verteilung. F_a wird ermittelt. Wenn $F > F_a$, wird H_0 abgelehnt.

Numerische Illustration 14.2

1. Im Milch-Beispiel ergibt sich für das einfache lineare Modell, $m_t = \alpha + \beta_1 f_t + u'_t$, der Wert $S'_{\hat{u}\hat{u}} = 2.786.870$.

Aus $\hat{m}_t = \hat{\alpha} + \hat{\beta}_1 f_t$ lassen sich die entsprechenden \hat{m}_t -Werte berechnen und daraus die Werte für \hat{m}_t^2 , \hat{m}_t^3 und \hat{m}_t^4 .

(Forts.)

2. Erweitert man das einfache lineare Modell um die Variablen \hat{m}_t^2 , \hat{m}_t^3 und \hat{m}_t^4 , dann ergibt sich

$$m_t = \alpha + \beta_1 f_t + \gamma_1 \hat{m}_t^2 + \gamma_2 \hat{m}_t^3 + \gamma_3 \hat{m}_t^4 + u_t^* .$$

Für dieses Modell erhält man $S_{\hat{u}\hat{u}}^* = 932.014$.

3. Für die Nullhypothese $H_0 : \gamma_1 = \gamma_2 = \gamma_3 = 0$ ergibt sich folgender F -Wert:

$$F = \frac{(2.786.870 - 932.014) / 3}{932.014 / 7} \approx 4,644 .$$

(Forts.)

4. Der Vergleich des berechneten F -Wertes mit dem $F_{0,05}$ -Wert zeigt, ob die Nullhypothese abzulehnen ist.

Frage: Zu welcher Testentscheidung gelangen wir?

Antwort:

Aussagekraft des RESET-Verfahrens (und ähnlicher Testverfahren)

- Kommt es im F -Test des RESET-Verfahrens zu einer Ablehnung der Nullhypothese, dann ist dies ein sehr verlässliches Signal für eine Fehlspezifikation.
- Kommt es im F -Test des RESET-Verfahrens hingegen zu keiner Ablehnung der Nullhypothese, dann ist dies noch kein sehr verlässliches Signal für eine korrekte Spezifikation.
- Diese Aussage gilt ganz allgemein:

Evidenz für die Nichtverletzung der A-, B- und C-Annahmen

In Tests, welche die Nichtverletzung einer betrachteten Annahme als Nullhypothese verwenden, lässt sich nur schwache Evidenz für die Nichtverletzung sammeln.

14.3.2 Bestimmtheitsmaß R^2

- Das Bestimmtheitsmaß darf nur dann angewendet werden, wenn
 - ① die endogene Variable der Modelle identisch ist,
 - ② die Anzahl der exogenen Variablen in den Modellen identisch ist,
 - ③ die Modelle einen Niveauparameter α besitzen.

Numerische Illustration 14.3

Aus Tabelle 14.3 ist ersichtlich:

Innerhalb der Modellgruppe mit nicht-logarithmierter endogener Variable ist das semi-logarithmische Modell am besten geeignet.

Innerhalb der Gruppe mit logarithmierter endogener Variable ist das logarithmische Modell am besten geeignet.

14.3.3 Box-Cox-Test

Box-Cox-Test für Logarithmierung der endogenen Variable

- Diese einfache Variante des Box-Cox Tests ist geeignet für zwei Modelle, bei denen das eine die endogene Variable y_t besitzt und das andere die endogene Variable $\ln y_t$.
- Ein Beispiel ist:

$$y_t = \alpha + \beta \ln x_t + u_t \quad \text{und} \quad \ln y_t = \alpha + \beta \ln x_t + u_t .$$

- Der Test vollzieht sich in 9 Schritten.

1. KQ-Schätzung des Modells mit der endogenen Variable y_t (Summe der Residuenquadrate: $S_{\hat{u}\hat{u}}$).
2. KQ-Schätzung des Modells mit der endogenen Variable $\ln y_t$ (Summe der Residuenquadrate: $S_{\hat{u}\hat{u}}^*$).
3. Ermittlung des *geometrischen* Mittels der Variable y_t :

$$\begin{aligned}\tilde{y} &= (y_1 \cdot y_2 \cdot \dots \cdot y_T)^{(1/T)} \\ &= e^{\ln(y_1 \cdot y_2 \cdot \dots \cdot y_T)^{(1/T)}} \\ &= e^{(1/T) \cdot \ln(y_1 \cdot y_2 \cdot \dots \cdot y_T)} \\ &= e^{(1/T) \cdot \sum_{t=1}^T \ln y_t} .\end{aligned}$$

4. Würde man das Modell mit der endogenen Variable y_t durch \tilde{y} dividieren und anschließend eine KQ-Schätzung dieses transformierten Modells durchführen, so ergäbe sich als Summe der Residuenquadrate der Wert $S_{\hat{u}\hat{u}}/\tilde{y}^2$.
5. Es kann gezeigt werden, dass $S_{\hat{u}\hat{u}}/\tilde{y}^2$ und $S_{\hat{u}\hat{u}}^*$ direkt vergleichbar sind. Gilt beispielsweise $S_{\hat{u}\hat{u}}/\tilde{y}^2 < S_{\hat{u}\hat{u}}^*$, dann spricht dies für das Modell mit y_t als endogener Variable.
6. Die Ergebnisse der Schritte 1 bis 5 werden durch einen Test der Nullhypothese $H_0 : S_{uu}/\tilde{y}^2 = S_{uu}^*$ ergänzt.

7. Es gilt

$$I = \frac{T}{2} \left| \ln \left(\frac{S_{\hat{u}\hat{u}} / \tilde{y}^2}{S_{\hat{u}\hat{u}}^*} \right) \right| \sim \chi_{(1)}^2.$$

Frage: Warum gilt immer $I \geq 0$? Wann ergibt sich $I = 0$? Zeigen Sie, dass

$$\left| \ln \left(\frac{S_{\hat{u}\hat{u}} / \tilde{y}^2}{S_{\hat{u}\hat{u}}^*} \right) \right| = \left| \ln \left(\frac{S_{\hat{u}\hat{u}}^*}{S_{\hat{u}\hat{u}} / \tilde{y}^2} \right) \right|.$$

Antwort:

8. Liegt der beobachtete Wert der Zufallsvariable I oberhalb des kritischen Wertes, dann wird die Nullhypothese verworfen.
9. Wenn die Modelle empirisch nicht gleichwertig sind, dann wird dasjenige Modell ausgewählt, welches die geringere Summe der Residuenquadrate aufweist.

Numerische Illustration 14.4

Es wird ein Vergleich zwischen dem logarithmischen Modell

$$\ln m_t = \alpha + \beta \ln f_t + u_t \quad (14.11)$$

und dem semi-logarithmischen Modell

$$m_t = \alpha + \beta \ln f_t + u_t \quad (14.3)$$

durchgeführt. Im Milch-Beispiel ist $\sum_{t=1}^T \ln m_t = 106,18432$.

Das geometrische Mittel der endogenen Variable beträgt demnach

$$\tilde{m} = e^{(1/12)106,18432} \approx 6965,29 .$$

(Forts.)

Eine KQ-Schätzung der zu vergleichenden Modelle (14.11) und (14.3) liefert die folgenden Ergebnisse:

$$\begin{aligned} S_{\widehat{u\widehat{u}}}^* &= 0,02873 \\ S_{\widehat{u\widehat{u}}} / \tilde{m}^2 &= 1847012,9 / 48515209,7 = 0,03807 . \end{aligned}$$

Da $S_{\widehat{u\widehat{u}}}^* < S_{\widehat{u\widehat{u}}} / \tilde{m}^2$, scheint das logarithmische Modell (14.11) besser geeignet als das semi-logarithmische Modell (14.3).

(Forts.)

Für I ergibt sich:

$$I = \frac{12}{2} \left| \ln \left(\frac{0,03807}{0,02873} \right) \right| \approx 1,68877 .$$

Frage: Zu welchem Testergebnis gelangen wir bei einem Signifikanzniveau von 5%?

Antwort:

Ein allgemeineres Box-Cox-Verfahren

- Eine allgemeine Variablentransformation:

$$y_t(\lambda) = \frac{y_t^\lambda - 1}{\lambda} \quad (14.20)$$

$$x_t(\psi) = \frac{x_t^\psi - 1}{\psi}, \quad (14.21)$$

wobei λ und ψ Transformationsparameter darstellen.

- Eine sehr allgemeine Modellformulierung lautet:

$$\frac{y_t^\lambda - 1}{\lambda} = \alpha^* + \beta \frac{x_t^\psi - 1}{\psi} + u_t, \quad (14.22)$$

wobei $\lambda \neq 0$ und $\psi \neq 0$.

- Setzt man $\lambda = \psi = 1$, dann ergibt sich das lineare Modell (14.4), denn

$$\begin{aligned}y_t - 1 &= \alpha^* + \beta(x_t - 1) + u_t \\ \implies y_t &= \alpha^* + 1 - \beta + \beta x_t + u_t \\ \implies &= \alpha + \beta x_t + u_t ,\end{aligned}$$

wobei $\alpha = \alpha^* + 1 - \beta$.

- Ferner ergibt sich
 - für $\lambda \rightarrow 0$ und $\psi \rightarrow 0$ das logarithmische Modell,
 - für $\lambda = 1$ und $\psi \rightarrow 0$ das semi-logarithmische Modell,
 - für $\lambda \rightarrow 0$ und $\psi = 1$ das Exponential-Modell,
 - für $\lambda \rightarrow 0$ und $\psi = -1$ das log-inverse Modell und
 - für $\lambda = 1$ und $\psi = -1$ das inverse Modell.
- Das allgemeine Box-Cox-Verfahren vollzieht sich in 5 Schritten.

1. Man ermittelt das *geometrische* Mittel

$$\begin{aligned}\tilde{y} &= (y_1 \cdot y_2 \cdot \dots \cdot y_T)^{(1/T)} \\ &= e^{(1/T) \cdot \sum_{t=1}^T \ln y_t} .\end{aligned}$$

2. Im Box-Cox-Modell (14.22) wird y_t durch die transformierte Variable $y_t^* = y_t / \tilde{y}$ ersetzt:

$$\frac{(y_t^*)^\lambda - 1}{\lambda} = \alpha^* + \beta \frac{x_t^\psi - 1}{\psi} + u_t . \quad (14.23)$$

3. Es wird eine bestimmte λ - ψ -Kombination benutzt und die sich daraus ergebenden Werte der endogenen Variable $[(y_t^*)^\lambda - 1]/\lambda$ und der exogenen Variable $(x_t^\psi - 1)/\psi$ des Modells (14.23) berechnet. Auf Basis dieser Werte erfolgt eine KQ-Schätzung des Modells (14.23).
4. Schritt 3 wird für alle anderen möglichen λ - ψ -Kombinationen wiederholt.
5. Man wählt diejenige λ - ψ -Kombination, welche die kleinste Summe der Residuenquadrate generiert. Setzt man die Werte dieser „besten“ λ - ψ -Kombination in Modell (14.22) ein, so hat man die beste funktionale Form für das zu schätzende Modell.

Auswahl des geeigneten Funktionstyps

Das RESET-Verfahren kann anzeigen, ob beim betrachteten Modell eine Fehlspezifikation vorliegt.

Mit dem Bestimmtheitsmaß kann das Feld der Alternativkandidaten eingeengt werden. Verglichen werden können nur Modelle mit identischer endogener Variable und gleicher Anzahl an exogenen Variablen.

Der Box-Cox-Test erlaubt Vergleiche zwischen Modellen mit unterschiedlicher endogener Variable.

Kapitel 15

Annahme A3: Konstante Parameterwerte

- Das ökonometrische Modell einer Mehrfachregression lautet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_K x_{Kt} + u_t .$$

Annahme A3

Die $K + 1$ Parameter $\alpha, \beta_1, \beta_2, \dots, \beta_K$ sind für alle T Beobachtungen $(x_{1t}, x_{2t}, \dots, x_{Kt}, y_t)$ konstant.

Beispiel 1 zu Kapitel 15

Trotz wirtschaftlicher Wachstumsraten von etwa 2% stieg die Erwerbslosenquote in Deutschland während der frühen 1990er Jahre stetig an. Deshalb wurde die zu Beginn 2005 wirksam gewordene Reform der Sozialsysteme (Hartz IV Gesetz) durchgeführt.

Wir untersuchen den Einfluss des Wirtschaftswachstums auf die Veränderung der Erwerbslosenquote in Deutschland. Die Jahresdaten des Zeitraums 1992 bis 2014 sind in Tabelle 15.1 und in Abbildung 15.1 wiedergegeben.

(Forts.)

Tabelle 15.1: Wirtschaftswachstum x_t und Veränderung der Erwerbslosenquote y_t für die 23 Jahre des Zeitraums 1992 bis 2014 (jeweils in Prozentpunkten); Datenquelle: Online-Datenbank des Statistischen Bundesamtes.

Jahr/t	x_t	y_t	Jahr/t	x_t	y_t	Jahr/t	x_t	y_t
92/1	1,9	1,0	01/10	1,7	-0,1	10/19	4,1	-0,7
93/2	-1,0	1,2	02/11	0,0	0,7	11/20	3,6	-0,9
94/3	2,5	0,5	03/12	-0,7	1,0	12/21	0,4	-0,5
95/4	1,7	-0,2	04/13	1,2	0,6	13/22	0,1	-0,1
96/5	0,8	0,6	05/14	0,7	0,8	14/23	1,6	-0,2
97/6	1,8	0,6	06/15	3,7	-0,9			
98/7	2,0	-0,2	07/16	3,3	-1,5			
99/8	2,0	-0,8	08/17	1,1	-1,0			
00/9	3,0	-0,7	09/18	-5,6	0,2			

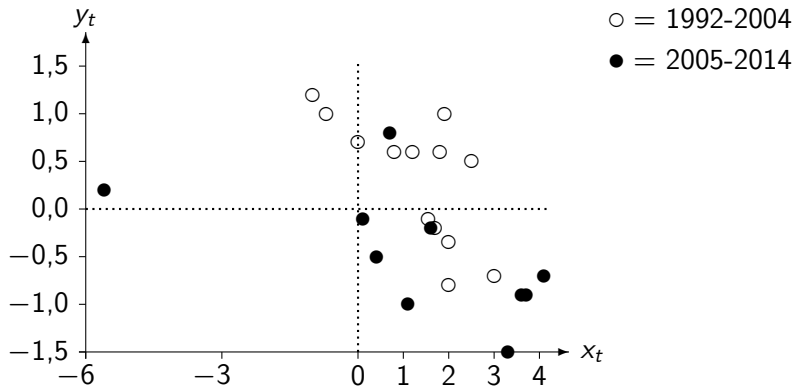


Abbildung 15.1: Die Datenpunkte des Arbeitsmarkt-Beispiels.

15.1 Konsequenzen der Annahmeverletzung

- Zwei getrennte Modelle für Phasen I und II:

$$\text{Phase I : } y_t = \alpha_I + \beta_I x_t + u_t \quad (15.1)$$

$$\text{Phase II : } y_t = \alpha_{II} + \beta_{II} x_t + u_t, \quad (15.2)$$

Fall A: $\alpha_I = \alpha_{II}, \beta_I = \beta_{II} \Rightarrow$ kein Strukturbruch

Fall B: $\alpha_I \neq \alpha_{II}, \beta_I = \beta_{II} \Rightarrow$ Strukturbruch in α

Fall C: $\alpha_I = \alpha_{II}, \beta_I \neq \beta_{II} \Rightarrow$ Strukturbruch in β

Fall D: $\alpha_I \neq \alpha_{II}, \beta_I \neq \beta_{II} \Rightarrow$ Strukturbruch in α und β

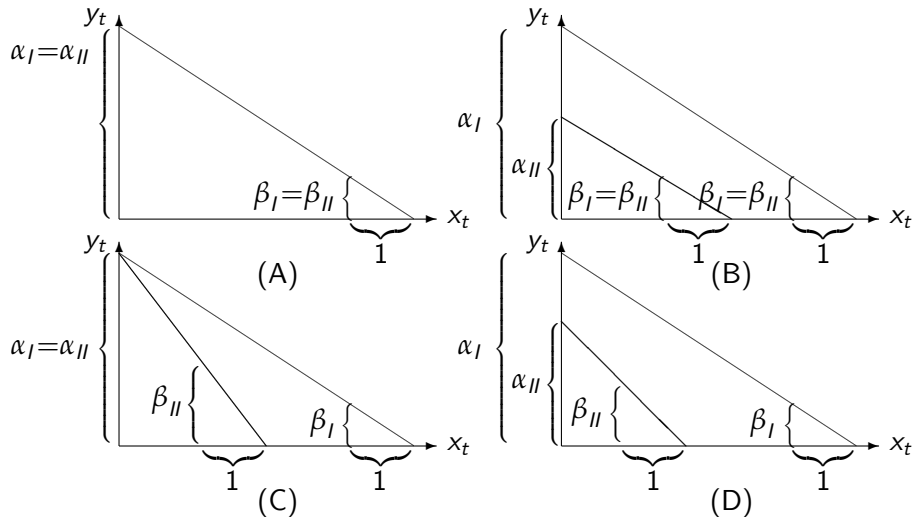


Abbildung 15.2: Vier mögliche Parameterkonstellationen.

15.1.1 Ein geeignetes Strukturbruchmodell

- Die Parameter der Phase II können in der Form

$$\alpha_{II} = \alpha_I + \gamma \quad (15.3)$$

$$\beta_{II} = \beta_I + \delta \quad (15.4)$$

geschrieben werden.

- Frage:** Wie sind die Parameter γ und δ demnach zu interpretieren?

Antwort:

- Das Modell der Phase II, (15.2), lautet dann:

$$\begin{aligned}y_t &= (\alpha_I + \gamma) + (\beta_I + \delta) x_t + u_t \\ &= \alpha_I + \gamma + \beta_I x_t + \delta x_t + u_t .\end{aligned}\tag{15.5}$$

- Das Modell der Phase I lautete:

$$y_t = \alpha_I + \beta_I x_t + u_t .\tag{15.1}$$

- Die Modelle der Phasen I und II, (15.1) und (15.5), können auch als Funktionen genau der gleichen Parameter geschrieben werden:

$$\text{Phase I: } y_t = \alpha_I + 0 \cdot \gamma + \beta_I x_t + 0 \cdot \delta x_t + u_t \quad (15.6)$$

mit $t = 1, 2, \dots, T_I$

$$\text{Phase II: } y_t = \alpha_I + 1 \cdot \gamma + \beta_I x_t + 1 \cdot \delta x_t + u_t \quad (15.7)$$

mit $t = T_I + 1, T_I + 2, \dots, T$,

wobei T_I die Anzahl der Beobachtungen der Phase I ist.

Definition: Strukturbruchmodell

Das Strukturbruchmodell

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + \delta D_t x_t + u_t, \quad (15.8)$$

mit

$$D_t = \begin{cases} 0 & \text{wenn } t = 1, 2, \dots, T_I \\ 1 & \text{wenn } t = T_I + 1, T_I + 2, \dots, T \end{cases}$$

gibt die in Gleichungen (15.6) und (15.7) beschriebenen Wirkungszusammenhänge der Phasen I und II in einer einzigen Gleichung wieder.

- Das Modell (15.8) ist Fall D und enthält die Fälle A bis C als Spezialfälle (jeweils $t = 1, 2, \dots, T$):

Fall A: $\gamma = 0$ und $\delta = 0$

$$y_t = \alpha_I + \beta_I x_t + u_t \quad (15.9)$$

Fall B: $\gamma \neq 0$ und $\delta = 0$

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + u_t \quad (15.10)$$

Fall C: $\gamma = 0$ und $\delta \neq 0$

$$y_t = \alpha_I + \beta_I x_t + \delta D_t x_t + u_t \quad (15.11)$$

Fall D: $\gamma \neq 0$ und $\delta \neq 0$

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + \delta D_t x_t + u_t . \quad (15.8)$$

15.1.2 Schätzung und Interpretation der Parameter des Strukturbruchmodells

- Das Strukturbruchmodell

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + \delta D_t x_t + u_t, \quad t = 1, 2, \dots, T \quad (15.8)$$

kann mit der KQ-Methode geschätzt werden.

- $D_t x_t$ ist eine *Interaktions-Dummy-Variable*.
- Es kann gezeigt werden, dass in die Schätzwerte für α_I und β_I allein die Beobachtungen der Phase I eingehen und in die Schätzwerte für γ und δ – und damit α_{II} und β_{II} – allein die Beobachtungen der Phase II.
- Es müssen deshalb sowohl für Phase I als auch für Phase II mindestens 2 Beobachtungen vorliegen.

Numerische Illustration 15.1

Im Arbeitsmarkt-Beispiel werden auf Basis der 23 Beobachtungen $\tilde{K} = 3$ Steigungsparameter und ein Niveauparameter geschätzt.

Die Freiheitsgrade dieser Schätzung betragen folglich $T - \tilde{K} - 1 = 19$.

Die Werte der Variablen des Strukturbruchmodells finden sich auszugsweise in Tabelle 15.2.

Die KQ-Schätzergebnisse sind in Tabelle 15.3 wiedergegeben.

(Forts.)

Tabelle 15.2: Veränderung der Erwerbslosenquote y_t , Wirtschaftswachstum x_t , Dummy-Variable D_t und Interaktions-Dummy $D_t x_t$ für die 23 Jahre des Zeitraums 1992 bis 2014.

	Jahr	t	y_t	x_t	D_t	$D_t x_t$
Phase I	1992	1	1,0	1,9	0	0
	1993	2	1,2	-1,0	0	0
	⋮	⋮	⋮	⋮	⋮	⋮
	2004	13	0,6	1,2	0	0
Phase II	2005	14	0,8	0,7	1	0,7
	2006	15	-0,9	3,7	1	3,7
	⋮	⋮	⋮	⋮	⋮	⋮
	2014	18	-0,2	1,6	1	1,6

(Forts.)

Tabelle 15.3: Schätzergebnisse für das Strukturbruchmodell (15.8) des Arbeitsmarkt-Beispiels.

Variable	Koeff.	$\widehat{sd}(\widehat{\cdot})$	t-Wert	p-Wert
Konstante	0,809	0,219	3,695	0,002
Wachstum	-0,374	0,126	-2,975	0,008
Dummy	-1,106	0,286	-3,862	0,001
Interaktions-Dummy	0,233	0,140	1,664	0,112

Unter Berücksichtigung zusätzlicher Dezimalstellen erhält man

$$\widehat{\alpha}_{II} = \widehat{\alpha}_I + \widehat{\gamma} = 0,809 - 1,106 = -0,296$$

$$\widehat{\beta}_{II} = \widehat{\beta}_I + \widehat{\delta} = -0,374 + 0,233 = -0,141 .$$

(Forts.)

Frage: Wie sind die oberen beiden Werte der Spalte „Koeff.“ in Tabelle 15.3 zu interpretieren?

Antwort:

(Forts.)

Frage: Wie sind die unteren beiden KQ-Schätzwerte zu interpretieren?

Antwort:

15.1.3 Getrennte Schätzung der zwei Phasen

Numerische Illustration 15.2

Tabelle 15.4: Schätzergebnisse bei getrennten KQ-Schätzungen der Modelle (15.1) und (15.2).

Modell	Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
(15.1)	Konstante	0,809	0,206	3,939	0,002
Phase I	Wachstum	-0,374	0,118	-3,171	0,009
(15.2)	Konstante	-0,296	0,199	-1,489	0,175
Phase II	Wachstum	-0,141	0,066	-2,129	0,066

(Forts.)

Frage: Was fällt Ihnen beim Vergleich mit den Ergebnissen der Tabelle 15.3 auf?

Antwort:

- **Frage:** Warum gilt bei beiden Schätzverfahren

$$sd(\hat{\beta}_1) = \sqrt{\sigma^2 / S_{xx}}?$$

Antwort:

- **Frage:** Warum weicht der *Schätzwert*

$\widehat{sd}(\widehat{\beta}_1) = \sqrt{\widehat{\sigma}^2 / S'_{xx}}$ dennoch in den beiden Verfahren voneinander ab?

Antwort:

15.1.4 Eine mögliche alternative Formulierung des Strukturbruchmodells

- Da $\gamma = \alpha_{II} - \alpha_I$ und $\delta = \beta_{II} - \beta_I$, kann das Strukturbruchmodell (15.8) auch in folgender Form geschrieben werden:

$$\begin{aligned}y_t &= \alpha_I + (\alpha_{II} - \alpha_I)D_t + \beta_I x_t + (\beta_{II} - \beta_I)D_t x_t + u_t \\ &= \alpha_I(1 - D_t) + \alpha_{II}D_t + \beta_I(1 - D_t)x_t + \beta_{II}D_t x_t + u_t .\end{aligned}$$

15.1.5 Komplexere Strukturbrüche

- Man möchte auch die Jahresdaten der Jahre 1970 bis 1991 einbeziehen
- Wegen der deutschen Wiedervereinigung im Übergang auf das Jahr 1990 muss ein weiterer Strukturbruch zugelassen werden.

- Das entsprechende Strukturbruchmodell lautet:

$$y_t = \alpha_I + \gamma_{II} D_{II t} + \gamma_{III} D_{III t} + (\beta_I + \delta_{II} D_{II t} + \delta_{III} D_{III t}) x_t + u_t . \quad (15.12)$$

- Phase I: $D_{II t} = D_{III t} = 0$

$$\Rightarrow y_t = \alpha_I + \beta_I x_t + u_t$$

- Phase II: $D_{II t} = 1$ und $D_{III t} = 0$

$$\Rightarrow y_t = \alpha_I + \gamma_{II} + (\beta_I + \delta_{II}) x_t + u_t$$

- Phase III: $D_{II t} = 0$ und $D_{III t} = 1$

$$\Rightarrow y_t = \alpha_I + \gamma_{III} + (\beta_I + \delta_{III}) x_t + u_t .$$

- Der Unterschied zwischen β_{II} und β_{III} beträgt

$$\beta_{III} - \beta_{II} = \beta_I + \delta_{III} - (\beta_I + \delta_{II}) = \delta_{III} - \delta_{II} . \quad (15.13)$$

15.1.6 Konsequenzen aus einer Vernachlässigung des Strukturbruchs

- Bei Vernachlässigung des Strukturbruchs würde man die Parameter des „falschen“ Modells

$$y_t = \alpha + \beta x_t + u_t \quad (15.14)$$

schätzen.

Numerische Illustration 15.3

Tabelle 15.5: Die KQ-Schätzergebnisse für Modell (15.14).

Variable	Koeff.	$\widehat{\text{sd}}(\widehat{\cdot})$	t-Wert	p-Wert
Konstante	0,216	0,169	1,277	0,216
Wachstum	-0,186	0,071	-2,613	0,016

15.2 Diagnose

- Das Strukturbruchmodell lautete:

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + \delta D_t x_t + u_t \quad (15.8)$$

Fall A: $\gamma = 0$ und $\delta = 0$

$$\Rightarrow y_t = \alpha_I + \beta_I x_t + u_t$$

Fall B: $\gamma \neq 0$ und $\delta = 0$

$$\Rightarrow y_t = \alpha_I + \gamma D_t + \beta_I x_t + u_t$$

Fall C: $\gamma = 0$ und $\delta \neq 0$

$$\Rightarrow y_t = \alpha_I + \beta_I x_t + \delta D_t x_t + u_t$$

Fall D: $\gamma \neq 0$ und $\delta \neq 0$

$$\Rightarrow \text{Strukturbruchmodell (15.8)}$$

15.2.1 *F*-Test

- Es ist zu überprüfen, ob überhaupt ein Strukturbruch stattgefunden hat.
- Dies geschieht durch einen *F*-Test der Nullhypothese H_0 :
 $\gamma = \delta = 0$.
- **Frage:** Welcher der vier Fälle kann ausgeschlossen werden, wenn diese Nullhypothese abgelehnt wird?
Antwort:

Numerische Illustration 15.4

Im Arbeitsmarkt-Beispiel ergibt sich für das Modell (15.8)

$$S_{\hat{u}\hat{u}} = 5,252.$$

Für das Nullhypothesen-Modell (15.14), also

$$y_t = \alpha + \beta x_t + u_t, \text{ ergibt sich } S_{\hat{u}\hat{u}}^0 = 9,664.$$

Damit erhält man

$$F = \frac{(9,664 - 5,252) / 2}{5,252 / 19} = 7,978 .$$

Frage: Zu welcher Testentscheidung gelangt man?

Antwort:

15.2.2 t -Test

- Fall A wurde im F -Test verworfen.
- Das Strukturbruchmodell

$$y_t = \alpha_I + \gamma D_t + \beta_I x_t + \delta D_t x_t + u_t \quad (15.8)$$

beschreibt Fall D (Strukturbruch in beiden Parametern).

- Ein t -Test von $H_0: \gamma = 0$ differenziert zwischen Fall C ($\gamma = 0, \delta \neq 0$) und D ($\gamma \neq 0, \delta \neq 0$).
- Ein t -Test von $H_0: \delta = 0$ differenziert zwischen Fall B ($\gamma \neq 0, \delta = 0$) und D ($\gamma \neq 0, \delta \neq 0$).

Numerische Illustration 15.5

Die t -Werte für $H_0: \gamma = 0$ und $H_0: \delta = 0$ finden sich in Tabelle 15.3.

Die Zahl der Freiheitsgrade beträgt 19.

Frage: Zu welchen Testentscheidungen gelangt man und welcher Fall erscheint folglich korrekt zu sein?

Antwort:

15.2.3 Prognostischer Chow-Test

- Das Modell der Phase I lautet wie gewohnt
$$y_t = \alpha_I + \beta_I x_t + u_t.$$
- Eine KQ-Schätzung dieses Modells auf Basis der Beobachtungen der Phase I liefert $S_{\hat{u}\hat{u}}^I$.
- T_{II} ist die Anzahl der Beobachtungen der Phase II.
- Eine KQ-Schätzung des Modells auf Basis sämtlicher verfügbarer Beobachtungen $T = T_I + T_{II}$ liefert $S_{\hat{u}\hat{u}}^*$.
- Eine solche Schätzung ist nur dann zulässig, wenn kein Strukturbruch vorliegt.

- Die Nullhypothese des prognostischen Chow-Tests lautet: kein Strukturbruch.
- Der F -Wert errechnet sich dann gemäß

$$F = \frac{(S_{\widehat{uu}}^* - S_{\widehat{uu}}^l) / T_{II}}{S_{\widehat{uu}}^l / (T_I - K - 1)} . \quad (15.15)$$

- Bei Gültigkeit der Nullhypothese „kein Strukturbruch“ folgt die Variable F einer $F_{(T_{II}, T_I - K - 1)}$ -Verteilung.

Numerische Illustration 15.6

Man befindet sich im Jahr 2006 und möchte untersuchen, ob das 2005 wirksam gewordene Hartz IV Gesetz sofort einen Strukturbruch ausgelöst hat.

Der Beobachtungsumfang beträgt $T = 14$ ($T_I = 13$ und $T_{II} = 1$).

Es ergeben sich die Werte $S_{\hat{u}\hat{u}}^* = 2,734$ und $S_{\hat{u}\hat{u}}^I = 2,676$.

Daraus ergibt sich

$$F = \frac{(2,734 - 2,676)/1}{2,676/(13 - 1 - 1)} = 0,239 .$$

Der kritische Wert beträgt $F_{0,05} = 4,84$ ($v_1 = 1$, $v_2 = 11$).

Die Nullhypothese „kein Strukturbruch“ kann auf einem Signifikanzniveau von 5% nicht abgelehnt werden.

15.2.4 Unbekannter Zeitpunkt des Strukturbruchs

- Ereignete sich der Strukturbruch im Übergang auf das Jahr 2004, 2005, 2006, 2007 oder 2008?
- Bislang wurde das Jahr 2005 unterstellt.
- Mit dem *Quandt likelihood ratio test* (abgekürzt: *QLR-Test*) kann man überprüfen, ob *und wann* ein Strukturbruch vorlag.
- Der QLR-Test setzt sich aus einer Sequenz der zuvor beschriebenen *F*-Tests zusammen.

Numerische Illustration 15.7

Für Modell (15.8) wird die Nullhypothese $H_0 : \gamma = \delta = 0$ getestet.

Der F -Test mit einem Strukturbruch im Übergang auf das Jahr 2004 liefert $S_{\hat{u}\hat{u}} = 6,223$ und $S_{\hat{u}\hat{u}}^0 = 9,664$ und damit

$$F = \frac{(9,664 - 6,223) / 2}{6,223 / 19} = 5,253 .$$

Die entsprechenden F -Tests für die nachfolgenden Jahre liefern:

2005: $F = 7,978$ (wissen wir bereits)

2006: $F = 15,216$

2007: $F = 14,735$

2008: $F = 12,774$.

- Der größte F -Wert ergibt sich für den Übergang auf das Jahr 2006.
- Dies ist der plausibelste Zeitpunkt für einen Strukturbruch.
- Wie testet man nun die Nullhypothese, dass sich zwischen 2004 und 2008 kein Strukturbruch ereignete?
- Es wurde eine ganze Sequenz an F -Tests durchgeführt.
- Deshalb ist es nicht zulässig, die F -Werte der einzelnen Jahre einfach mit dem kritischen Wert der $F_{(2,19)}$ -Verteilung (5% Signifikanzniveau) zu vergleichen.
- Eine Analogie kann dies veranschaulichen.

Beispiel 2 zu Kapitel 15

Die Qualität eines einzelnen frisch produzierten Spielwürfels wird überprüft. Anschließend wird ein Karton mit 100 frisch produzierten Würfeln überprüft.

Numerische Illustration 15.8

Die Nullhypothese lautet H_0 : „Würfel ist perfekt gefertigt“.

Frage: Wie wahrscheinlich ist es, bei einem perfekt gefertigten Würfel bei dreimaligem Würfeln jedesmal die gleiche Zahl zu erhalten?

Antwort:

(Forts.)

Die Nullhypothese lautet H_0 : „alle 100 Würfel sind perfekt gefertigt“.

Frage: Wie groß ist die Wahrscheinlichkeit, dass bei einem perfekt gefertigten Würfel das Ereignis „*nicht*-dreimal-die-gleiche-Zahl“ eintritt?

Antwort:

Frage: Wie groß ist die Wahrscheinlichkeit, dass bei 100 perfekt gefertigten Würfeln bei *allen* das Ereignis „*nicht*-dreimal-die-gleiche-Zahl“ eintritt?

Antwort:

(Forts.)

Frage: Wie groß ist die Wahrscheinlichkeit, dass bei 100 perfekt gefertigten Würfeln bei mindestens einem das Ereignis „dreimal-die-gleiche-Zahl“ eintritt?

Antwort:

Frage: Welche Testentscheidung muss also bei einem Signifikanzniveau von 5% getroffen werden, wenn einmal das Ereignis „dreimal-die-gleiche-Zahl“ eintritt?

Antwort:

- *Fazit:* Ein bestimmtes Stichprobenergebnis, welches bei einmaliger Durchführung eines Hypothesentests zur Ablehnung der Nullhypothese führen würde, verliert diese Eigenschaft, wenn der Einzeltest hinreichend oft wiederholt wird.
- Dies muss im Hypothesentest bei der Wahl des angemessenen kritischen Wertes berücksichtigt werden.
- Der QLR-Test liefert für die betrachtete Sequenz von F -Tests den korrekten kritischen Wert.
- Dieser Wert hängt nicht nur von L ab, sondern auch von dem als Anteil am gesamten Zeithorizont ausgedrückten Beginn und Ende des für einen Strukturbruch in Frage kommenden Zeitabschnitts (π_0 und π_1).

Numerische Illustration 15.9

Der erste F -Test aus der Sequenz von F -Tests bezieht sich auf das Jahr 2004, also auf die Periode (Beobachtung) $t_A = 13$ („A“ für „Anfang“).

In Beziehung gesetzt zur Gesamtzahl der Beobachtungen ($T = 18$) ergibt sich der Anteil

$$\pi_0 = t_A / T = 13 / 23 \approx 0,57 .$$

Der letzte F -Test bezieht sich auf das Jahr 2008, also auf $t_E = 17$ („E“ für „Ende“).

Es ergibt sich

$$\pi_1 = t_E / T = 17 / 23 \approx 0,74 .$$

(Forts.)

Diese Werte liefern

$$\lambda = \frac{\pi_1(1 - \pi_0)}{\pi_0(1 - \pi_1)} = \frac{0,74(1 - 0,57)}{0,57(1 - 0,74)} = 2,18 .$$

Frage: Warum gilt immer $\lambda > 1$?

Antwort:

In Tabelle T.5 kann abgelesen werden, in welchem Korridor der kritische Wert liegt.

(Forts.)

Frage: Zu welcher Testentscheidung gelangt man?

Antwort:

Identifikation von Strukturbrüchen

Für die Identifikation von Strukturbrüchen stehen verschiedene Verfahren zur Verfügung:

- t -Test,
- F -Test,
- Prognostischer Chow-Test,
- QLR-Test.

Alle diese Tests setzen auf dem Strukturbruchmodell auf.

15.3 Stetige Veränderung von Parameterwerten

- t ist hier eine *Trendvariable*. Ihre Werte finden sich in Tabelle 15.1.
- Der wahre Zusammenhang laute:

$$y_t = \alpha + \beta_t x_t + u_t, \quad (15.16)$$

wobei

$$\beta_t = \mu + \lambda t. \quad (15.17)$$

- Einsetzen von (15.17) in (15.16) ergibt:

$$\begin{aligned}y_t &= \alpha + (\mu + \lambda t)x_t + u_t \\ &= \alpha + \mu x_t + \lambda(tx_t) + u_t .\end{aligned}$$

- Die endogene Variable y_t wird auf die exogene Variable x_t und die *Interaktions-Variable* tx_t regressiert.
- **Frage:** Mit welchem Test könnte man überprüfen, ob eine stetige Veränderung der Parameterwerte vorliegt?

Antwort:

15.4 Exkurs: Qualitative exogene Variablen

Beispiel 3 zu Kapitel 15

Es wird das Lohn-Beispiel des Kapitels 13 aufgegriffen. Der Lohn y_t ist eine Funktion der Ausbildung x_{1t} und des Alters x_{2t} :

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t . \quad (15.18)$$

Einige Mitarbeiterinnen des Betriebes beklagen, dass sie beim Lohn diskriminiert werden. Um dies zu überprüfen, muss für jeden der zwanzig Mitarbeiter zusätzlich zu den Daten der Tabelle 13.1 das Geschlecht ermittelt werden. Es zeigt sich, dass die Mitarbeiter $t = 1,4,7,9,11,16,17$ und 19 Frauen sind.

15.4.1 Einführung einer Dummy-Variable

- Die Variable *Geschlecht* ist nicht quantifizierbar. Sie ist eine *qualitative* Variable.
- Es kann eine Regression für Männer und eine separate Regression für Frauen durchgeführt werden.
- Beide Regressionen können wie beim Strukturbruch in einem Dummy-Variablen-Modell zusammengefasst werden.
- Zu diesem Zweck wird die folgende Dummy-Variable definiert:

$$D_t = \begin{cases} 0 & \text{wenn Mitarbeiter } t \text{ männlich} \\ 1 & \text{wenn Mitarbeiter } t \text{ weiblich} . \end{cases} \quad (15.19)$$

15.4.2 Ein allgemeines Dummy-Variablen-Modell

- Das zu (15.18) korrespondierende Dummy-Variablen-Modell lautet:

$$y_t = \alpha + \gamma D_t + \beta_1 x_{1t} + \delta_1 D_t x_{1t} + \beta_2 x_{2t} + \delta_2 D_t x_{2t} + u_t . \quad (15.20)$$

- Männer $D_t = 0$:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t .$$

- Frauen $D_t = 1$:

$$\begin{aligned} y_t &= \alpha + \gamma + \beta_1 x_{1t} + \delta_1 x_{1t} + \beta_2 x_{2t} + \delta_2 x_{2t} + u_t \\ &= \alpha + \gamma + (\beta_1 + \delta_1) x_{1t} + (\beta_2 + \delta_2) x_{2t} + u_t . \end{aligned}$$

- **Frage:** Wie sind die Parameter γ , δ_1 und δ_2 zu interpretieren?

Antwort:

Numerische Illustration 15.10

Tabelle 15.6: Schätzergebnisse des Dummy-Variablen-Modells (15.20).

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	930,15	136,34	6,82	<0,001
Bildung	60,33	15,33	3,93	0,001
Alter	16,20	3,64	4,45	<0,001
Dummy	142,51	211,67	0,67	0,512
Interakt.-D. (Bildung)	-45,10	32,76	-1,38	0,190
Interakt.-D. (Alter)	-7,67	6,21	-1,24	0,237

(Forts.)

Frage: Ist die Diskriminierungsklage für 18-jährige Mitarbeiterinnen mit einjähriger Ausbildung berechtigt?

Antwort:

(Forts.)

Frage: Erscheint die Diskriminierungsklage für die anderen Mitarbeiterinnen berechtigt?

Antwort:

(Forts.)

Führt man für das Modell (15.20) einen F -Test der Nullhypothese $H_0 : \delta_1 = \delta_2 = 0$ durch, so wird diese auf einem Signifikanzniveau von 1% verworfen.

Die Parameterwerte, welche die Diskriminierungsklage stützen (δ_1 und δ_2), sind demnach gemeinsam signifikant.

Frage: Was kann bezüglich der Signifikanz des Parameterwerts festgestellt werden, der die Diskriminierungsklage in Frage stellt?

Antwort:

KQ-Schätzung bei qualitativen exogenen Variablen

Qualitative exogene Variablen können mit Hilfe von einer oder mehreren (nur notwendig bei mehr als zwei möglichen Merkmalsausprägungen) Dummy-Variablen berücksichtigt werden.

Die Modellierung erfolgt ganz analog zum Strukturbruchmodell.

Kapitel 16

Annahme B1: Erwartungswert der Störgröße

Annahme B1 (Erwartungswert)

Die Störgröße u_t hat für alle Beobachtungen t einen Erwartungswert von 0, das heißt,

$$E(u_t) = 0, \quad \text{für } t = 1, 2, \dots, T.$$

Beispiel 1 zu Kapitel 16

Ein Zulieferbetrieb der Automobilbranche stellt Kugellager her. Die Produktionsanlage muss regelmäßig gewartet werden.

Tabelle 16.1: Für 6 verschiedene Versuchssituationen der Anteil fehlerhafter Kugellager y_t (in Promillpunkten) und die Anzahl der Produktionsschichten x_t zwischen zwei Wartungsschichten.

t	y_t	x_t	t	y_t	x_t
1	14	12	4	20	24
2	20	16	5	28	28
3	22	20	6	26	32

(Forts.)

Frage: Wie sind die Zahlen der Beobachtung $t = 4$ zu interpretieren?

Antwort:

16.1 Konsequenzen der Annahmeverletzung

- Die schlimmste Konsequenz wäre eine Verzerrung in den KQ-Schätzern.
- Es wird im Folgenden überprüft, ob

$$E(\hat{\beta}) \neq \beta \quad \text{oder} \quad E(\hat{\alpha}) \neq \alpha \quad (16.1)$$

oder beides.

- In der Einfachregression gelten folgende Beziehungen:

$$E(\hat{\beta}) = \beta + \frac{\sum (x_t - \bar{x}) E(u_t)}{S_{xx}} \quad (16.2)$$

$$E(\hat{\alpha}) = \alpha - \bar{x} \left[E(\hat{\beta}) - \beta \right] + \frac{1}{T} \sum E(u_t) \quad (16.3)$$

- **Frage:** Welche Erwartungswerte ergeben sich, wenn Annahme B1 erfüllt ist?

Antwort:

- Welche der Schätzer im Falle einer B1-Verletzung verzerrt sind, hängt davon ab, ob

$$E(u_t) = \lambda \neq 0 \quad (16.4)$$

oder ob

$$E(u_t) = \lambda_t \neq 0. \quad (16.5)$$

16.1.1 Konstanter Messfehler bei der Erfassung der endogenen Variable

Konsequenzen für $E(u_t)$

- Das wahre Modell sei

$$y_t^* = \alpha + \beta x_t + u_t^* . \quad (16.6)$$

- Es erfülle alle A-, B- und C-Annahmen.
- Wenn nicht y_t^* beobachtet wird, sondern $y_t = y_t^* + \lambda$, dann kann Gleichung (16.6) auch in der folgenden Form geschrieben werden:

$$\begin{aligned} y_t - \lambda &= \alpha + \beta x_t + u_t^* \\ \Rightarrow y_t &= \alpha + \beta x_t + u_t , \end{aligned} \quad (16.7)$$

wobei

$$u_t = u_t^* + \lambda . \quad (16.8)$$

- Modell (16.7) liegt der KQ-Schätzung zugrunde.

- **Frage:** Was bedeutet das Ergebnis (16.8) für die Erfüllung von Annahme B1 im Modell (16.7)?

Antwort:

Numerische Illustration 16.1

Tabelle 16.2: Vergleich der Schätzergebnisse bei korrekt gemessener endogener Variable (Werte der Tabelle 16.1) und bei um 8 Promillpunkte zu hoch gemessener endogener Variable.

y_t-Mess.	Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
korrekt	Konstante	8,781	3,595	2,443	0,071
	Schichten	0,586	0,156	3,753	0,020
zu hoch	Konstante	16,781	3,595	4,668	0,010
	Schichten	0,586	0,156	3,753	0,020

Konsequenzen für $\hat{\alpha}$ und $\hat{\beta}$ – eine grafische Betrachtung

- **Frage:** Welche Konsequenzen ergeben sich für die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$?

Antwort:

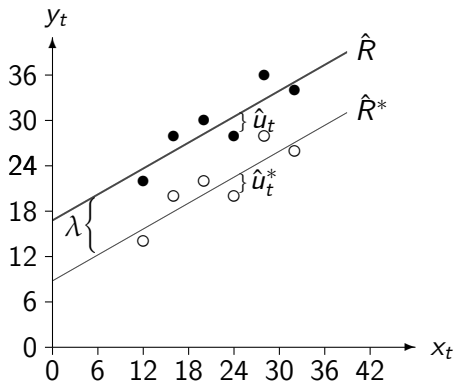


Abbildung 16.1: Konstanter Messfehler in Höhe von $\lambda = 8$ bei der Erfassung von y_t .

Konsequenzen für $\hat{\alpha}$ und $\hat{\beta}$ – eine formale Betrachtung

- Setzt man Gleichung (16.4) in Gleichung (16.2) ein, so ergibt sich

$$E(\hat{\beta}) = \beta + \frac{\lambda \sum (x_t - \bar{x})}{S_{xx}} = \beta, \quad (16.10)$$

wobei ausgenutzt wurde, dass

$$\begin{aligned} \sum (x_t - \bar{x}) &= \sum x_t - \sum \bar{x} = \sum x_t - T\bar{x} \\ &= \sum x_t - T \left[(1/T) \sum x_t \right] = 0. \end{aligned}$$

- Setzt man die Gleichungen (16.4) und (16.10) in Gleichung (16.3) ein, so ergibt sich

$$\begin{aligned} E(\hat{\alpha}) &= \alpha - \bar{x} [\beta - \beta] + \frac{1}{T} (T\lambda) \\ &= \alpha + \lambda . \end{aligned} \tag{16.11}$$

Konsequenzen für Intervallschätzer und Hypothesentest

- **Frage:** Welche Konsequenzen ergeben sich für den Intervallschätzer und für Hypothesentests, die auf den Niveauparameter α bezogen sind.

Antwort:

- Damit für β Intervallschätzer und Hypothesentests Gültigkeit besitzen, muss $\text{var}(\hat{\beta})$ unverzerrt geschätzt werden.

- Es lässt sich zeigen, dass

$$\text{var}(\hat{\beta}) = \text{var}(\hat{\beta}^*) = \sigma^2 / S_{xx} . \quad (16.12)$$

- Es muss σ^2 unverzerrt geschätzt werden.
- Die Äquivalenz von \hat{u}_t^* und \hat{u}_t besagt, dass $S_{\hat{u}\hat{u}} (= \sum \hat{u}_t^2)$ und $S_{\hat{u}\hat{u}}^* (= \sum \hat{u}_t^{*2})$ identische Werte besitzen.

- Folglich gilt:

$$E(\hat{\sigma}^2) = E[S_{\hat{u}\hat{u}} / (T - 2)] = E[S_{\hat{u}\hat{u}}^* / (T - 2)] . \quad (16.13)$$

- Da das korrekte Modell alle A-, B- und C-Annahmen erfüllt, gilt

$$E[S_{\hat{u}\hat{u}}^* / (T - 2)] = \sigma^2 . \quad (16.14)$$

- Folglich ist das auf Basis des fehlerhaften Modells (16.7) ermittelte $\hat{\sigma}^2$ unverzerrt.
- Damit ist auch $\widehat{var}(\hat{\beta}) = \hat{\sigma}^2 / S_{xx}$ ein unverzerrter Schätzer.
- Der Intervallschätzer bleibt folglich unverzerrt und Hypothesentests sind aussagekräftig.

Konsequenzen für $\hat{\beta}$ in Modellen ohne Niveauparameter

Beispiel 2 zu Kapitel 16

Die Aktie des Automobilzulieferers wird auf Basis des Capital Asset Pricing Modells (CAPM) untersucht. Es sei:

r_t die Rendite dieser Aktie in Periode t

r_t^m die über alle Aktien der Börse gemittelte Marktrendite

r_t^f die Rendite einer risikolosen Anlage.

- Der CAPM-Idee zufolge lautet das ökonometrische Modell:

$$\left(r_t - r_t^f \right) = \beta \left(r_t^m - r_t^f \right) + u_t .$$

- **Frage:** Welche ökonomische Interpretation besitzt der Ausdruck $\left(r_t - r_t^f \right)$ und welche der Ausdruck $\left(r_t^m - r_t^f \right)$?

Antwort:

- **Frage:** Was würde folglich $\beta = 1$ ökonomisch aussagen?

Antwort:

- Definiert man

$$\begin{aligned}y_t &= r_t - r_t^f \\x_t &= r_t^m - r_t^f ,\end{aligned}$$

ergibt sich das ökonometrische Modell

$$y_t = \beta x_t + u_t .$$

- Abbildung 16.2 veranschaulicht, dass die Verletzung der Annahme B1 diesmal auch auf die Schätzung des Steigungsparameters β durchschlägt.

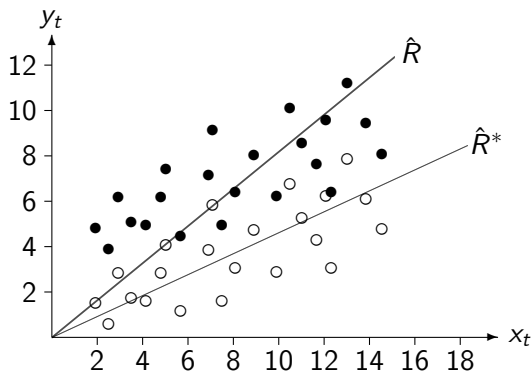


Abbildung 16.2: Konstanter Messfehler bei der Erfassung von y_t in einem Modell ohne Niveauparameter.

16.1.2 Konstanter Messfehler bei der Erfassung einer exogenen Variable

- In Abbildung 16.3 ist ein Messfehler von $\phi = 10$ wiedergegeben.
- **Frage:** Warum führt ein Messfehler von ϕ zu $E(u_t) = -\beta\phi$?

Antwort:

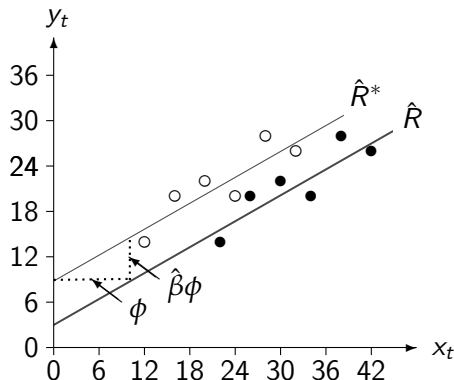


Abbildung 16.3: Konstanter Messfehler in der Erfassung von x_t .

Konstanter Messfehler

Ein konstanter Messfehler in der endogenen oder in einer exogenen Variable führt in Modellen mit Niveauparameter α zu

- 1 einem verzerrten Punktschätzer $\hat{\alpha}$,
- 2 verzerrten Intervallschätzern und
- 3 wertlosen Hypothesentests in Bezug auf α .

In Bezug auf die Steigungsparameter β_k ergeben sich jedoch keine Probleme (Ausnahme: Modelle ohne Niveauparameter α).

16.1.3 Funktionale Modelltransformation

Beispiel 3 zu Kapitel 16

Die Kugellagerproduktion des Automobilzulieferers sei durch die Produktionsfunktion

$$g = Ak^\beta . \quad (16.15)$$

charakterisiert. Dabei bezeichne:

g den Pro-Kopf-Output

k die Kapitalintensität (Kapital pro Arbeitseinheit) und

A und β seien konstante Parameter.

- Das ökonometrische Modell könnte beispielsweise

$$g_t = Ak_t^\beta v_t \quad (16.16)$$

lauten, wobei v_t eine Störgröße darstellt.

- Folglich erhält man

$$E(g_t) = Ak_t^\beta E(v_t) .$$

- Eine sinnvolle Interpretation der Produktionsfunktion (16.15) erfordert, dass $E(v_t) = 1$, dass also

$$E(g_t) = Ak_t^\beta .$$

- Annahmen A2 und B1 sind demnach verletzt.
- Logarithmieren vermeidet die A2-Verletzung. Es ergibt sich:

$$\ln g_t = \ln A + \beta \ln k_t + \ln v_t . \quad (16.17)$$

- Definiert man

$$\ln g_t = y_t$$

$$\ln A = \alpha$$

$$\ln k_t = x_t$$

$$\ln v_t = u_t$$

so erhält man

$$y_t = \alpha + \beta x_t + u_t . \quad (16.18)$$

Konsequenzen für $E(u_t)$

- Gemäß Jensens Ungleichung gilt:

$$E[\ln v_t] < \ln[E(v_t)] . \quad (16.19)$$

- Da

$$\ln v_t = u_t, \quad E(v_t) = 1 \quad \text{und} \quad \ln 1 = 0 ,$$

gilt

$$E(u_t) < 0$$

und folglich

$$E(u_t) = \lambda \neq 0.$$

- Annahme B1 ist folglich verletzt.

- Eine Produktionsfunktion der Form

$$g_t = Ak_t^\beta e^{u_t}$$

mit

$$E(u_t) = 0$$

würde zu

$$\ln g_t = \ln A + \beta \ln k_t + u_t .$$

führen.

- Es ergäbe sich also keine Verletzung der Annahme B1.

Konsequenzen für $\hat{\alpha}$ und $\hat{\beta}$

- Der Parameter A der Produktionsfunktion (16.15) kann als technisches Niveau der Produktion interpretiert werden.
- Sein Schätzwert ergibt sich beim ökonometrischen Modell (16.17) aus:

$$\hat{A} = e^{\ln \hat{A}} = e^{\hat{\alpha}} . \quad (16.20)$$

16.1.4 Gestutzte endogene Variable

Beispiel 4 zu Kapitel 16

Gibt es bei den Mitarbeitern einen linearen Zusammenhang zwischen der Häufigkeit des Zuspätkommens (y_t) und der Länge des Arbeitsweges (x_t)? Die Datenpunkte sind in Abbildung 16.4 wiedergegeben.

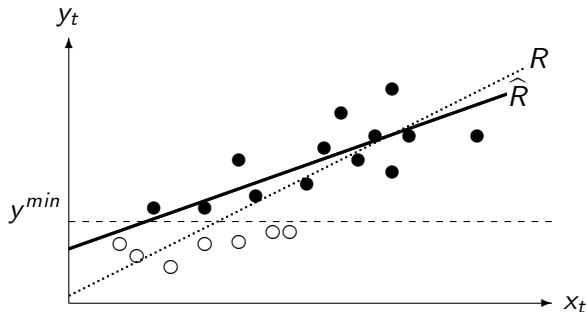


Abbildung 16.4: Die Datenpunkte für die Untersuchung des Zusammenhangs zwischen Arbeitsweg und Unpünktlichkeit.

Konsequenzen für $E(u_t)$ – eine grafische Veranschaulichung

- Abbildung 16.5 zeigt für drei Personen die jeweilige Wahrscheinlichkeitsverteilung der Störgröße u_t .
- $E(u_1)$ ist der Erwartungswert der Störgröße für eine *in der Stichprobe befindliche Person* mit Anfahrtsweg x_1 .
- $E(u_1)$ ergibt sich als Mittelwert der im nicht gepunkteten Bereich liegenden Störgrößen.
- Die Abbildung veranschaulicht, dass eine gestutzte endogene Variable zu einer B1-Verletzung in der Variante $E(u_t) = \lambda_t \neq 0$ führt.

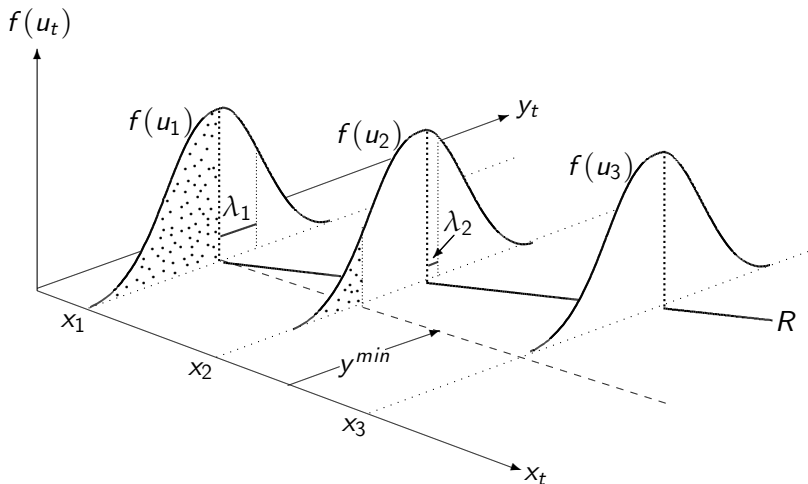


Abbildung 16.5: Die Wirkung einer begrenzten endogenen Variable auf den Erwartungswert der Störgröße.

Konsequenzen für $\hat{\alpha}$ und $\hat{\beta}$, Intervallschätzer und Hypothesentests

- Da $E(u_t) = \lambda_t \neq 0$ und nicht $E(u_t) = \lambda \neq 0$, kann der Niveauparameter die Verzerrung nicht mehr absorbieren.

Gestutze endogene Variable

Eine gestutze endogene Variable hat folgende Konsequenzen:

- 1 verzernte Punktschätzer,
- 2 verzernte Intervallschätzer und
- 3 wertlose Hypothesentests.

Zensierte und gestutzte Daten

- Wenn alle Personen mit $y_t < y^{\min}$ von der Untersuchung komplett ausgeschlossen werden, dann spricht man von einem *gestutzten Datensatz*.
- Wenn auch die Personen mit $y_t < y^{\min}$ in der Untersuchung berücksichtigt werden, aber bei ihnen, neben dem Wert der exogenen Variablen, lediglich bekannt ist, dass $y_t < y^{\min}$, dann spricht man von einem *zensierten Datensatz*.

16.2 Diagnose

16.2.1 Überprüfung der Datenerhebung

- Es ist zu überprüfen, ob Messfehler vorliegen könnten oder ob man es mit einem gestutzten oder zensierten Datensatz zu tun hat.
- Leichter zu prüfen ist, ob eine Modelltransformation vorgenommen wurde.

16.2.2 Überprüfung auf Basis der Daten

- Die Überprüfung der Annahme B1 auf Basis der Daten ist nur möglich, wenn die ökonomische Theorie ein Modell ohne Niveauparameter bereitstellt (z.B. CAPM).
- In diesem Fall kann man mit einem t -Test überprüfen, ob $\hat{\alpha}$ signifikant von 0 verschieden ist.
- Ist die Theorie korrekt, dann signalisiert ein signifikantes $\hat{\alpha}$ eine Verzerrung.

16.3 Anwendbare Schätzverfahren

- Falls $E(u_t) = \lambda$ und man nur an den Steigungsparametern interessiert ist, kann die KQ-Methode eingesetzt werden.
- Falls $E(u_t) = \lambda_t$, benötigt man andere Schätzverfahren.
- Bei gestutzten Datensätze nimmt man Verfahren, die auf der Idee der Maximum Likelihood Methode aufbauen (nicht Gegenstand dieses Kurses).
- Bei zensierten Daten könnte man auch das Heckman-Schätzverfahren einsetzen (ebenfalls nicht Gegenstand dieses Kurses).

Kapitel 17

Annahme B2: Homoskedastizität

Annahme B2 (Homoskedastizität)

Die Störgröße u_t hat für alle Beobachtungen t eine konstante Varianz, das heißt,

$$\text{var}(u_t) = \sigma^2, \quad \text{für } t = 1, 2, \dots, T.$$

Beispiel zu Kapitel 17

Die Höhe der durchschnittlichen Kaltmiete in einem Stadtviertel werde durch die Entfernung zum Stadtzentrum erklärt. Die Daten sind in Tabelle 17.1 und die Punktwolke in Abbildung 17.1 wiedergegeben.

(Forts.)

Tabelle 17.1: Entfernung vom Stadtzentrum x_t (in km) und Nettokaltmiete y_t (in Euro/m²) für 12 Stadtviertel.

t	x_t	y_t	t	x_t	y_t
1	0,50	16,80	7	3,10	12,80
2	1,40	16,20	8	4,40	12,20
3	1,10	15,90	9	3,70	15,00
4	2,20	15,40	10	3,00	13,60
5	1,30	16,40	11	3,50	14,10
6	3,20	13,20	12	4,10	13,30

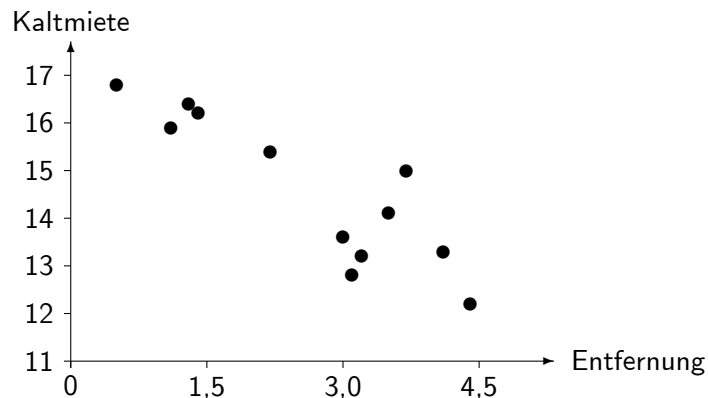


Abbildung 17.1: Die Datenpunkte des Mieten-Beispiels.

- Die Streuung der Störgrößen scheint in den peripheren Stadtvierteln größer auszufallen als in den zentralen.
- Die Störgröße des ökonometrischen Modells ist dann nicht *homoskedastisch*, sondern *heteroskedastisch*.

17.1 Konsequenzen der Annahmeverletzung

- Für das Mieten-Beispiel lautet das ökonometrische Modell:

$$y_t = \alpha + \beta x_t + u_t . \quad (17.1)$$

- Abbildung 17.1 signalisiert, dass $\sigma_t^2 \neq \sigma^2$.
- Exemplarisch wird der folgende Fall betrachtet:

$$\sigma_t^2 = \sigma^2 x_t^2 \quad \text{für } t = 1, 2, \dots, 12 . \quad (17.2)$$

- **Frage:** Was besagt Gleichung (17.2) bezüglich der Streuung der Störgrößen?

Antwort:

17.1.1 Konsequenzen für die Punktschätzung

- Da Annahme B1 erfüllt ist, gilt gemäß der Gleichungen (16.2) und (16.3):

$$E(\hat{\beta}) = \beta \quad \text{und} \quad E(\hat{\alpha}) = \alpha .$$

- Die KQ-Schätzer des ökonometrischen Modells (17.1) verlieren wegen der Heteroskedastizität ihre Effizienz-Eigenschaft.
- Um dies zu zeigen, wird das Modell (17.1) so transformiert, dass Annahme B2 wieder erfüllt ist.

Transformation des Modells

- Um die Gültigkeit der Annahme B2 wiederherzustellen wird (17.1) durch x_t dividiert:

$$\frac{y_t}{x_t} = \alpha \frac{1}{x_t} + \beta \frac{x_t}{x_t} + \frac{u_t}{x_t} . \quad (17.3)$$

- Die Varianz der neuen Störgröße u_t/x_t beträgt

$$\text{var} \left(\frac{u_t}{x_t} \right) = \left(\frac{1}{x_t} \right)^2 \text{var} (u_t) = \frac{1}{x_t^2} \sigma^2 x_t^2 = \sigma^2 . \quad (17.4)$$

- Ferner gilt:

$$E \left(\frac{u_t}{x_t} \right) = \frac{1}{x_t} E(u_t) = 0 \quad (17.5)$$

und

$$\frac{u_t}{x_t} \sim UN(0, \sigma^2) . \quad (17.6)$$

- Es werden folgende Umbenennungen vorgenommen:

$$y_t^* = \frac{y_t}{x_t}, \quad z_t^* = \frac{1}{x_t}, \quad \text{und} \quad u_t^* = \frac{u_t}{x_t} . \quad (17.7)$$

- Damit lautet das Modell:

$$y_t^* = \beta + \alpha z_t^* + u_t^* . \quad (17.8)$$

- Alle A-, B- und C-Annahmen sind erfüllt.
- Modell (17.8) ist eine Einfachregression mit dem Niveauparameter β und dem Steigungsparameter α .

- **Frage:** Wie sind die Parameter α und β zu interpretieren?
Antwort:

KQ-Schätzer des transformierten Modells

- Es gelten für Modell (17.8) die für Niveau- und Steigungsparameter üblichen KQ-Schätzformeln:

$$\hat{\alpha}^* = \frac{S_{z^*y^*}}{S_{z^*z^*}} \quad (17.9)$$

$$\hat{\beta}^* = \bar{y}^* - \hat{\alpha}^* \bar{z}^* . \quad (17.10)$$

Numerische Illustration 17.1

Eine Transformation der Daten des Mieten-Beispiels liefert die in Tabelle 17.2 aufgelisteten Daten. Eine KQ-Schätzung auf Basis dieser Daten liefert die Ergebnisse in Tabelle 17.3.

Tabelle 17.2: Die transformierten Daten des Mieten-Beispiels.

t	z_t^*	y_t^*	t	z_t^*	y_t^*
1	2,000	33,600	7	0,323	4,129
2	0,714	11,571	8	0,227	2,773
3	0,909	14,455	9	0,270	4,054
4	0,454	7,000	10	0,333	4,533
5	0,769	12,615	11	0,286	4,029
6	0,313	4,125	12	0,244	3,244

(Forts.)

Tabelle 17.3: Vergleich der KQ-Schätzergebnisse des Mieten-Beispiels bei Verwendung des transformierten Modells (17.8) und des heteroskedastischen Modells (17.1).

Modell	Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
(17.8)	transf. Konst.	17,347	0,162	107,218	<0,001
	transf. Entf.	-1,048	0,121	-8,658	<0,001
(17.1)	Konstante	17,393	0,527	32,995	<0,001
	Entfernung	-1,074	0,182	-5,906	<0,001

Warum sind die KQ-Schätzer eines Modells mit heteroskedastischer Störgröße ineffizient?

- Modell (17.8) erfüllt alle A-, B- und C-Annahmen.
- **Frage:** Warum ist dann der KQ-Schätzer des Modells (17.1) ineffizient?

Antwort:

- Modell (17.8) minimiert $\sum (\hat{u}_t^*)^2 = \sum (\hat{u}_t / x_t)^2$.
- Modell (17.1) minimiert hingegen $\sum \hat{u}_t^2$.
- Kleine x_t -Werte generieren relativ zu großen x_t -Werten „potenziell größere“ \hat{u}_t / x_t -Werte.
- Die $\hat{\alpha}^*$ und $\hat{\beta}^*$ -Werte werden so gewählt, dass sich insbesondere für Beobachtungen mit kleinen x_t -Werten kleine Residuen \hat{u}_t ergeben.
- Beobachtungen mit kleinen x_t -Werten erhalten also eine stärkere Bedeutung als diejenigen mit großen x_t -Werten.
- **Frage:** Warum ist dies vorteilhaft für die Qualität der Schätzung?

Antwort:

17.1.2 Konsequenzen für Intervallschätzung und Hypothesentest

- Im Modell (17.1) wird die Varianz des KQ-Schätzers $\hat{\beta}$ durch folgende Gleichung definiert:

$$\text{var}(\hat{\beta}) = \frac{1}{S_{xx}^2} \sum (x_t - \bar{x})^2 \sigma_t^2. \quad (17.11)$$

- Nur für $\sigma_t^2 = \sigma^2$ ergibt sich

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{S_{xx}^2} \sum (x_t - \bar{x})^2 = \frac{\sigma^2}{S_{xx}^2} S_{xx} = \frac{\sigma^2}{S_{xx}}. \quad (17.12)$$

- Da $\sigma_t^2 \neq \sigma^2$, erfordert eine unverzerrte Schätzung von $\text{var}(\hat{\beta})$, dass der jeweilige Wert $\sigma_t^2 (= \sigma^2 x_t^2)$ unverzerrt geschätzt wird, was wiederum eine unverzerrte Schätzung von σ^2 erfordert.

- Da das transformierte Modell (17.8) alle A-, B- und C-Annahmen erfüllt, ist

$$\hat{\sigma}^2 = \frac{S_{\hat{u}^* \hat{u}^*}}{T - 2} \quad (17.13)$$

ein unverzerrter Schätzer für σ^2 .

- Folglich wäre

$$\hat{\sigma}_t^2 = \hat{\sigma}^2 \cdot x_t^2 = \frac{S_{\hat{u}^* \hat{u}^*}}{T - 2} \cdot x_t^2 \quad (17.14)$$

eine unverzerrte Schätzung für σ_t^2 .

- Setzte man die gemäß (17.14) berechneten Werte $\hat{\sigma}_t^2$ in (17.11) ein, ergäbe sich eine unverzerrte Schätzung von $\text{var}(\hat{\beta})$.

- Unter Missachtung der Heteroskedastizität würde man σ^2 jedoch gemäß

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T-2} \quad (17.15)$$

schätzen und in (17.12) einsetzen.

- **Frage:** Warum weist dann der Schätzer $\widehat{var}(\hat{\beta})$ gleich zwei Verzerrungsquellen auf?

Antwort:

Folgen der Heteroskedastizität

Die KQ-Schätzung eines ökonometrischen Modells mit heteroskedastischer Störgröße führt zu

- 1 unverzerrten aber ineffizienten Punktschätzern,
- 2 verzerrten Intervallschätzern und
- 3 wertlosen Hypothesentests.

17.2 Diagnose

17.2.1 Grundidee der Tests auf Heteroskedastizität

- In (gedanklich) wiederholten Stichproben würde sich für jede der T Beobachtungen eine eigene Sequenz von Störgrößen ergeben.
- Im Falle homoskedastischer Störgrößen wäre die Streuung innerhalb der Störgrößensequenz bei allen T Beobachtungen einheitlich σ^2 .
- Bei heteroskedastischen Störgrößen, hätte jede Störgrößensequenz ihre eigene Varianz σ_t^2 .

- Diese Varianz ist der über alle wiederholten Stichproben gebildete Durchschnitt der Terme $[u_t - E(u_t)]^2 = u_t^2$.
- Der Durchschnitt aus den in der Störgrößensequenz der Beobachtung t aufgetretenen u_t^2 wäre deshalb ein guter Schätzer für σ_t^2 .
- Es steht aber nur die eine tatsächlich erfasste Stichprobe zur Verfügung.
- u_1^2 wäre für σ_1^2 ein sehr grober aber dennoch brauchbarer Schätzer (analog u_2^2 und u_3^2 für σ_2^2 und σ_3^2).

- Da u_t^2 unbekannt ist, muss man das jeweilige Residuenquadrat \hat{u}_t^2 verwenden.
- Die \hat{u}_t^2 liefern also Informationen über die Varianz der Störgrößen einzelner Beobachtungen oder einer Gruppe von Beobachtungen.
- *Erinnerung:* Wenn H_0 die Nichtverletzung der betrachteten Annahme postuliert (hier: „Homoskedastizität“), dann ist eine Ablehnung von H_0 ein starkes Signal für die Verletzung der Annahme, aber eine Nichtablehnung von H_0 ein recht schwaches Signal für die Erfüllung der untersuchten Annahmen.

17.2.2 Goldfeld-Quandt-Test

- Dieser Test lässt sich immer dann anwenden, wenn die Störgrößenvarianz in *monotoner* Weise positiv oder negativ von der Höhe einer *einzelnen* exogenen Variable abhängt.

1. Schritt: Neusortierung und Partition der Beobachtungen

- Es muss die exogene Variable x_{kt} identifiziert werden, welche für die Heteroskedastizität verantwortlich ist.
- Die Beobachtungen werden entsprechend der Höhe der x_{kt} -Werte neu sortiert.
- Anschließend werden die Beobachtungen in zwei Gruppen unterteilt.

Numerische Illustration 17.2

Die Stadtviertel werden in zwei Gruppen unterteilt.

Es bietet sich an, eine Gruppe *I* aus den fünf zentralen Stadtvierteln ($T_I = 5$) und eine Gruppe *II* aus den sieben peripheren Stadtvierteln zu bilden ($T_{II} = 7$).

(Forts.)

Tabelle 17.4: Entfernung vom Stadtzentrum x_t (in km) und Nettokaltmiete y_t (in Euro/ m^2) für zentrale ($t = 1,2,\dots,5$) und periphere ($t = 6,7,\dots,12$) Stadtviertel.

Gruppe I			Gruppe II		
t	x_t	y_t	t	x_t	y_t
1	0,50	16,80	6	3,00	13,60
2	1,10	15,90	7	3,10	12,80
3	1,30	16,40	8	3,20	13,20
4	1,40	16,20	9	3,50	14,10
5	2,20	15,40	10	3,70	15,00
			11	4,10	13,30
			12	4,40	12,20

2. Schritt: Formulierung der Null- und Alternativhypothese

- Wenn man vermutet, dass $\sigma_{II}^2 > \sigma_I^2$, dann wird die Nullhypothese $H_0 : \sigma_{II}^2 \leq \sigma_I^2$ getestet.
- Die Alternativhypothese lautet $H_1 : \sigma_{II}^2 > \sigma_I^2$.

Numerische Illustration 17.3

Im Mieten-Beispiel ist von $\sigma_{II}^2 > \sigma_I^2$ auszugehen. Die zu testende Nullhypothese lautet deshalb: $H_0 : \sigma_{II}^2 \leq \sigma_I^2$.

3. Schritt: Separate Schätzungen

- Für jede der beiden Gruppen wird eine separate KQ-Schätzung vorgenommen und die jeweilige Summe der Residuenquadrate $S_{\hat{u}\hat{u}}^I$ und $S_{\hat{u}\hat{u}}^{II}$ berechnet.

Numerische Illustration 17.4

Eine separate Schätzung für die beiden Gruppen liefert die in Tabelle 17.5 aufgelisteten Ergebnisse. Die entsprechenden Summen der Residuenquadrate lauten $S_{\widehat{uu}}^I = 0,246$ und $S_{\widehat{uu}}^{II} = 4,666$.

Tabelle 17.5: Übersicht der Ergebnisse bei getrennter Schätzung.

Modell	Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
I	Konstante	17,128	0,330	51,972	<0,001
$t = 1, \dots, 5$	Entfernung	-0,760	0,234	-3,253	0,047
II	Konstante	14,841	2,691	5,515	0,003
$t = 6, \dots, 12$	Entfernung	-0,387	0,747	-0,519	0,626

4. Schritt: Berechnung des F -Wertes und Testentscheidung

- Bei Gültigkeit der Nullhypothese (im Sinne von $\sigma_{II}^2 = \sigma_I^2$) wären

$$\hat{\sigma}_I^2 = \frac{S'_{\hat{u}\hat{u}}}{T_I - K - 1} \quad \text{und} \quad \hat{\sigma}_{II}^2 = \frac{S''_{\hat{u}\hat{u}}}{T_{II} - K - 1} \quad (17.16)$$

unverzerrte Schätzer für σ_I^2 und σ_{II}^2 .

- Goldfeld und Quandt (1965) haben gezeigt, dass dann

$$F = \frac{\hat{\sigma}_{II}^2}{\hat{\sigma}_I^2} \quad (17.17)$$

einer $F_{(T_{II}-K-1, T_I-K-1)}$ -Verteilung folgt.

- Wenn $F > F_a$, dann wird die Nullhypothese zu Gunsten der Alternativhypothese (Heteroskedastizität) verworfen.

Numerische Illustration 17.5

Aus den bisherigen Ergebnissen erhält man

$$\hat{\sigma}_I^2 = \frac{0,246}{3} = 0,082 \quad \text{und} \quad \hat{\sigma}_{II}^2 = \frac{4,666}{5} = 0,933 .$$

Einsetzen dieser Schätzwerte in Formel (17.17) liefert den F-Wert:

$$F = \frac{0,933}{0,082} = 11,378 .$$

Frage: Zu welcher Testentscheidung gelangt man?

Antwort:

- Der Goldfeld-Quandt-Test kann auch dann angewendet werden, wenn die Heteroskedastizität durch eine beobachtbare externe Variable w_t ausgelöst wird, also einer Variable, die nicht Teil des ökonometrischen Modells ist.
- Bei hinreichend großem Stichprobenumfang können im Goldfeld-Quandt-Test einige mittlere Beobachtungen weggelassen werden (maximal ein Drittel).

17.2.3 Breusch-Pagan-Test

- Sind mehrere Variablen in linearer Form für die Heteroskedastizität verantwortlich und sind diese Variablen bekannt, dann kann der *Breusch-Pagan-Test* eingesetzt werden.
- Er prüft, ob sich zwischen \hat{u}_t^2 – also den Schätzern für σ_t^2 – und den verwendeten exogenen Variablen ein systematischer Zusammenhang feststellen lässt.
- Exemplarisch wird folgende Zweifachregression betrachtet:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t . \quad (17.18)$$

- Der Test vollzieht sich in fünf Schritten:
 1. Die Nullhypothese besagt, dass die Störgrößen homoskedastisch sind.
 2. Man schätzt Gleichung (17.18) und errechnet die Residuen \hat{u}_t .

3. Man bildet die lineare „Hilfsregression“:

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + v_t, \quad (17.19)$$

wobei v_t eine neue Störgröße darstellt.

Gleichung (17.19) kann immer dann eingesetzt werden, wenn die Störgrößenvarianz σ_t^2 eine Funktion der rechten Seite der Gleichung (17.19) ist, z.B. auch im Falle von $\sigma_t^2 = \sqrt{\gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t}}$.

Wären auch externe Variablen für die Heteroskedastizität verantwortlich, so müssten auch diese in die Hilfsregression (17.19) in linearer Form aufgenommen werden.

4. Das R^2 der KQ-Schätzung der Hilfsregression (17.19) wird errechnet.

Frage: Welcher Wert R^2 müsste sich bei Gültigkeit der Nullhypothese (Homoskedastizität) ergeben?

Antwort:

5. Bei Gültigkeit der Nullhypothese folgt $R^2 T$ einer $\chi^2_{(v)}$ -Verteilung. Dabei bezeichnet v die Zahl der Steigungsparameter der Hilfsregression (17.19). Fällt $R^2 T$ größer aus als der kritische Wert der $\chi^2_{(v)}$ -Verteilung, wird die Nullhypothese (Homoskedastizität) abgelehnt.

- $R^2 T$ folgt nur bei großem Stichprobenumfang T einer $\chi^2_{(v)}$ -Verteilung.
- Bei kleinen Stichprobenumfängen ersetzen Praktiker deshalb Schritt 5 durch einen F -Test der Nullhypothese $H_0: \gamma_1 = \gamma_2 = 0$.
- Der Breusch-Pagan-Test ist aber auch bei großem Stichprobenumfang nur dann ein verlässlicher Indikator, wenn in der Hilfsregression genau die richtigen Variablen enthalten sind.

17.2.4 White-Test

- Wenn die Heteroskedastizität von keinen externen Variablen ausgelöst wird, aber gleichzeitig die genaue Form der Heteroskedastizität unklar ist, dann bietet sich der White-Test an.
- Er vollzieht sich in genau den gleichen fünf Schritten wie der Breusch-Pagan-Test.
- Er verwendet aber in Schritt 4 nicht Gleichung (17.19), sondern

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \gamma_3 x_{1t}^2 + \gamma_4 x_{2t}^2 + \gamma_5 x_{1t} x_{2t} + v_t \quad (17.20)$$

- Auch die Zuverlässigkeit des White-Tests ist bei kleinem Stichprobenumfang eingeschränkt.
- In einem solchen Fall kann man Schritt 5 durch einen F -Test der Nullhypothese $H_0: \gamma_1 = \gamma_2 = \dots = \gamma_5 = 0$ ersetzen.
- **Frage:** Warum bereitet der White-Test bei Modellen mit vielen exogenen Variablen Probleme?
Antwort:
- Der große Vorteil des White-Tests gegenüber dem Breusch-Pagan-Test besteht darin, dass er sehr allgemeine Heteroskedastizitätsformen zulässt.

Tests auf Heteroskedastizität

Wenn keine externen Variablen für die Heteroskedastizität verantwortlich sind, kann der White-Test verwendet werden.

Falls die Heteroskedastizität von einer einzigen uns bekannten (möglicherweise auch externen) Variable ausgelöst wird, kann der Goldfeld-Quandt-Test eingesetzt werden.

Falls mehrere uns bekannte (möglicherweise auch externe) Variablen in linearer Form für die Heteroskedastizität verantwortlich sind, kann der Breusch-Pagan-Test verwendet werden.

17.3 Anwendbare Schätzverfahren

- Im Mieten-Beispiel wurde das heteroskedastische Modell

$$y_t = \alpha + \beta x_t + u_t$$

in ein homoskedastisches Modell transformiert und für dieses die KQ-Schätzer ermittelt.

- Dieses Schätzverfahren hat in der Literatur einen eigenen Namen: VKQ-Schätzung (V steht für „verallgemeinerte“).
- Es werden noch weitere Schätzverfahren betrachtet.

17.3.1 VKQ-Methode

- Als VKQ-Schätzung bezeichnet man die KQ-Schätzung eines transformierten Modells, welches sämtliche A-, B- und C-Annahmen erfüllt.
- Die erforderliche Transformation des ursprünglichen Modells hängt vom Typ der Heteroskedastizität ab.

Fall 1: $\sigma_t^2 = \sigma^2 x_t^2$

- Dieser Fall wurde bereits analysiert.
- Es ergaben sich die KQ-Schätzer

$$\hat{\alpha}^{VKQ} = \frac{S_{x^* y^*}}{S_{x^* x^*}} \quad (17.21)$$

$$\hat{\beta}^{VKQ} = \bar{y}^* - \hat{\alpha}^{VKQ} \bar{x}^* \quad (17.22)$$

wobei y_t^* und z_t^* gemäß der Definitionen (17.7) gegeben waren.

Fall 2: $\sigma_t^2 = \sigma^2 x_t$

- In diesem Fall ist das Ausgangsmodell durch $\sqrt{x_t}$ zu dividieren:

$$\frac{y_t}{\sqrt{x_t}} = \alpha \frac{1}{\sqrt{x_t}} + \beta \frac{x_t}{\sqrt{x_t}} + \frac{u_t}{\sqrt{x_t}} .$$

- Definiert man

$$y_t^* = \frac{y_t}{\sqrt{x_t}}, \quad z_t^* = \frac{1}{\sqrt{x_t}}, \quad x_t^* = \frac{x_t}{\sqrt{x_t}} \quad \text{und} \quad u_t^* = \frac{u}{\sqrt{x_t}}, \quad (17.23)$$

so erhält man

$$y_t^* = \alpha z_t^* + \beta x_t^* + u_t^* . \quad (17.24)$$

- Die entsprechenden Schätzer lauten

$$\hat{\alpha}^{VKQ} = \frac{S'_{x^*x^*} S'_{z^*y^*} - S'_{z^*x^*} S'_{x^*y^*}}{S'_{z^*z^*} S'_{x^*x^*} - (S'_{z^*x^*})^2} \quad (17.25)$$

$$\hat{\beta}^{VKQ} = \frac{S'_{z^*z^*} S'_{x^*y^*} - S'_{z^*x^*} S'_{z^*y^*}}{S'_{z^*z^*} S'_{x^*x^*} - (S'_{z^*x^*})^2}, \quad (17.26)$$

wobei

$$S'_{z^*z^*} = \sum (z_t^*)^2 \quad (17.27)$$

$$S'_{x^*x^*} = \sum (x_t^*)^2 \quad (17.28)$$

$$S'_{z^*x^*} = \sum z_t^* x_t^* \quad (17.29)$$

$$S'_{z^*y^*} = \sum z_t^* y_t^* \quad (17.30)$$

$$S'_{x^*y^*} = \sum x_t^* y_t^* \quad (17.31)$$

- Setzt man die Definitionen (17.27) bis (17.31) in die Schätzformeln (17.25) und (17.26) ein, so ergeben sich nach ein paar Umformungen die folgenden Schätzformeln:

$$\hat{\alpha}^{VKQ} = \frac{\sum \frac{y_t}{x_t} (x_t - \bar{x})}{\sum \frac{1}{x_t} (x_t - \bar{x})} \quad (17.32)$$

$$\hat{\beta}^{VKQ} = \frac{\sum \frac{1}{x_t} (y_t - \bar{y})}{\sum \frac{1}{x_t} (x_t - \bar{x})} . \quad (17.33)$$

- Die neue Störgröße ist homoskedastisch, denn

$$\text{var}(u_t^*) = \text{var}\left(\frac{u_t}{\sqrt{x_t}}\right) = \frac{1}{x_t} \text{var}(u_t) = \frac{\sigma_t^2}{x_t} = \frac{\sigma^2 x_t}{x_t} = \sigma^2 . \quad (17.34)$$

- Auch diese VKQ-Schätzung ist unverzerrt und effizient.
- Falls $\sigma_t^2 = \sigma^2 w_t$ und w_t eine externe Variable darstellt, dividiert man das Ausgangsmodell durch $\sqrt{w_t}$ und falls $\sigma_t^2 = \sigma^2 w_t^2$, dividiert man es durch w_t .

17.3.2 GVKQ-Methode

Fall 3: Partitionierbarer Datensatz

- Abbildung 17.1 wäre auch mit folgendem Fall kompatibel:

$$\sigma_t^2 = \sigma_I^2, \quad \text{für } t = 1, \dots, 5, \quad (17.35)$$

$$\sigma_t^2 = \sigma_{II}^2, \quad \text{für } t = 6, \dots, 12. \quad (17.36)$$

- Man hat es dann mit zwei separaten ökonomischen Modellen zu tun:

$$\text{Gruppe I: } y_t = \alpha + \beta x_t + u_t, \quad \text{für } t = 1, \dots, 5, \quad (17.37)$$

$$\text{Gruppe II: } y_t = \alpha + \beta x_t + u_t, \quad \text{für } t = 6, \dots, 12. \quad (17.38)$$

- Beide Modelle erfüllen alle A-, B- und C-Annahmen.

Transformiertes Modell

- Modell (17.37) wird durch σ_I und Modell (17.38) durch σ_{II} geteilt:

$$\frac{y_t}{\sigma_I} = \alpha \frac{1}{\sigma_I} + \beta \frac{x_t}{\sigma_I} + \frac{u_t}{\sigma_I}, \quad \text{für } t = 1, \dots, 5 \quad (17.39)$$

$$\frac{y_t}{\sigma_{II}} = \alpha \frac{1}{\sigma_{II}} + \beta \frac{x_t}{\sigma_{II}} + \frac{u_t}{\sigma_{II}}, \quad \text{für } t = 6, \dots, 12. \quad (17.40)$$

- Die Störgrößenvarianzen dieser zwei Modelle lauten:

$$\text{var} \left(\frac{u_t}{\sigma_I} \right) = \frac{1}{\sigma_I^2} \text{var}(u_t) = \frac{1}{\sigma_I^2} \sigma_I^2 = 1$$

$$\text{var} \left(\frac{u_t}{\sigma_{II}} \right) = \frac{1}{\sigma_{II}^2} \text{var}(u_t) = \frac{1}{\sigma_{II}^2} \sigma_{II}^2 = 1.$$

- Die Modelle (17.39) und (17.40) lassen sich zusammenfassen zu

$$y_t^* = \alpha z_t^* + \beta x_t^* + u_t^* \quad \text{für } t = 1, \dots, 12, \quad (17.41)$$

wobei für $t = 1, \dots, 5$,

$$y_t^* = \frac{y_t}{\sigma_I}, \quad z_t^* = \frac{1}{\sigma_I}, \quad x_t^* = \frac{x_t}{\sigma_I}, \quad u_t^* = \frac{u_t}{\sigma_I}, \quad (17.42)$$

und für $t = 6, \dots, 12$,

$$y_t^* = \frac{y_t}{\sigma_{II}}, \quad z_t^* = \frac{1}{\sigma_{II}}, \quad x_t^* = \frac{x_t}{\sigma_{II}}, \quad u_t^* = \frac{u_t}{\sigma_{II}}. \quad (17.43)$$

Schätzung des transformierten Modells

- Zunächst wird aus Modell (17.37) die Varianz σ_I^2 und aus Modell (17.38) die Varianz σ_{II}^2 mit der üblichen Formel

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}} / (T - K - 1)$$

geschätzt.

- Mit Hilfe der Schätzwerte $\hat{\sigma}_I^2$ und $\hat{\sigma}_{II}^2$ werden aus (17.42) und (17.43) die Werte von y_t^* , z_t^* und x_t^* berechnet.
- Daraufhin kann eine KQ-Schätzung des transformierten Modells (17.41) erfolgen.
- Diese VKQ-Schätzung mit vorgelagerten $\hat{\sigma}^2$ -Schätzungen nennt man *GVKQ-Schätzung* (G steht für „Geschätzte“).

- Die gemäß (17.42) und (17.43) ermittelten Werte werden in die VKQ-Schätzformeln (17.25) und (17.26) eingesetzt, um die GVKQ-Schätzwerte zu erhalten.

Numerische Illustration 17.6

Die Schätzergebnisse für die getrennten Schätzungen der zentralen und peripheren Stadtviertel waren in Tabelle 17.5 wiedergegeben. Aus $S_{\hat{u}\hat{u}}^I = 0,246$ bzw. $S_{\hat{u}\hat{u}}^{II} = 4,666$ folgt:

$$\hat{\sigma}_I^2 = 0,246/3 = 0,082 \quad \text{und} \quad \hat{\sigma}_{II}^2 = 4,666/5 = 0,933 .$$

Damit lassen sich die Werte von y_t^* , z_t^* und x_t^* des Modells (17.41) berechnen.

(Forts.)

Eine KQ-Schätzung des Modells auf Basis der Datentransformationen (17.42) und (17.43) und der VKQ-Schätzformeln (17.25) und (17.26) liefert die Ergebnisse der Tabelle 17.6.

Tabelle 17.6: Schätzergebnisse des Mieten-Beispiels bei Partitionierung des Datensatzes in zwei Gruppen.

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
transf. Konstante	17,416	0,243	72,002	<0,001
transf. Entfernung	-1,014	0,135	-7,489	<0,001

Schätzeigenschaften der GVKQ-Schätzer

- Die GVKQ-Schätzer sind unverzerrt und ihre Schätzvarianz ist geringer als diejenige der entsprechenden KQ-Schätzer.
- Es kann nicht davon ausgegangen werden, dass sie effizient sind.
- Die GVKQ-Schätzer sind aber *asymptotisch effizient*.

17.3.3 KQ-Methode mit Whites HK-Schätzer

- Ist die Heteroskedastizitätsform bekannt, dann kann die VKQ-Methode bzw. GVKQ-Methode eingesetzt werden.
- Wenn die Heteroskedastizitätsform unbekannt ist, muss ein anderer Weg verfolgt werden.
- Die KQ-Schätzung des heteroskedastischen Ausgangsmodells ist zwar unverzerrt, aber (in geringem Umfang) ineffizient.
- Ferner werden $sd(\hat{\alpha})$ und $sd(\hat{\beta})$ verzerrt geschätzt.
- Man könnte deshalb an der KQ-Schätzung des ursprünglichen Modells festhalten, aber für die Schätzung von $sd(\hat{\alpha})$ und $sd(\hat{\beta})$ modifizierte Formeln einsetzen.

- Bei großem Stichprobenumfang T liefert dieses Vorgehen verlässliche Schätzwerte.
- Aus Gleichung (17.11) ist bekannt, dass:

$$\text{var}(\hat{\beta}) = \frac{1}{S_{xx}^2} \sum (x_t - \bar{x})^2 \sigma_t^2. \quad (17.11)$$

- Verwendet man für die Schätzung von $\text{var}(\hat{\beta})$ jeweils \hat{u}_t^2 als Schätzer für σ_t^2 , dann wird (17.11) zu Whites *heteroskedastizitäts-konsistenten (HK)* Schätzer:

Formel: Whites HK-Schätzer

$$\widehat{\text{var}}(\hat{\beta})^{HK} = \frac{1}{S_{xx}^2} \sum (x_t - \bar{x})^2 \hat{u}_t^2 \quad (17.44)$$

- Auf Basis dieses HK-Schätzers berechnet man

$$\widehat{sd}(\widehat{\beta})^{HK} = \sqrt{\widehat{var}(\widehat{\beta})^{HK}}.$$

- Man kann den KQ-Schätzer $\widehat{\beta}$ in Verbund mit dem HK-Schätzer $\widehat{var}(\widehat{\beta})^{HK}$ selbst dann einsetzen, wenn unklar ist, ob heteroskedastische Störgrößen vorliegen.
- Wenn externe Variablen für die Heteroskedastizität verantwortlich sind, kann der HK-Schätzer nicht verwendet werden.
- Der HK-Schätzer wird oftmals auch als *robuster Schätzer* bezeichnet.

Numerische Illustration 17.7

Im Mieten-Beispiel erhält man $S_{xx} = 18,223$ und $\sum (x_t - \bar{x})^2 \hat{u}_t^2 = 5,239$.

Einsetzen in den HK-Schätzer (17.44) liefert

$$\widehat{\text{var}}(\hat{\beta})^{HK} = \frac{1}{18,223^2} 5,239 = 0,016 .$$

Als geschätzte Standardabweichung ergibt sich folglich

$$\widehat{sd}(\hat{\beta})^{HK} = \sqrt{0,016} = 0,126 .$$

In Tabelle 17.3 wurde hingegen $\widehat{sd}(\hat{\beta}) = 0,182$ ausgewiesen.

Anwendbare Schätzverfahren

Wenn die Form der Heteroskedastizität bekannt ist, können je nach Heteroskedastizitätsform eine VKQ- oder eine GVKQ-Schätzung eingesetzt werden.

Wenn keine externen Variablen für die Heteroskedastizität verantwortlich sind, kann Whites HK-Schätzer verwendet werden. Kenntnis von der Heteroskedastizitätsform ist dann nicht erforderlich.

Kapitel 18

Annahme B3: Freiheit von Autokorrelation

Annahme B3 (Freiheit von Autokorrelation)

Die Störgröße ist nicht autokorreliert, das heißt,

$$\text{cov}(u_t, u_s) = 0 ,$$

für alle $t \neq s$ sowie $t = 1, 2, \dots, T$ und $s = 1, 2, \dots, T$.

Beispiel Trinkwasserfilter

Es soll eine Preisabsatzfunktion für den Marktführer von Trinkwasserfiltern ermittelt werden. Die Daten sind in Tabelle 18.1 und Abbildung 18.1 wiedergegeben.

Beispiel (Forts.)

Tabelle 18.1: Monatliche Absatzmenge y_t (in 1000 Stück) und Verkaufspreis x_t (in Euro) für die vergangenen 24 Monate.

t	x_t	y_t	t	x_t	y_t
1	24,2	1990	13	32,2	1700
2	25,5	1630	14	32,4	1450
3	26,8	1570	15	33,2	1480
⋮	⋮	⋮	⋮	⋮	⋮
10	28,4	1270	22	28,3	1660
11	29,8	1250	23	27,5	1500
12	31,3	1500	24	26,8	1810

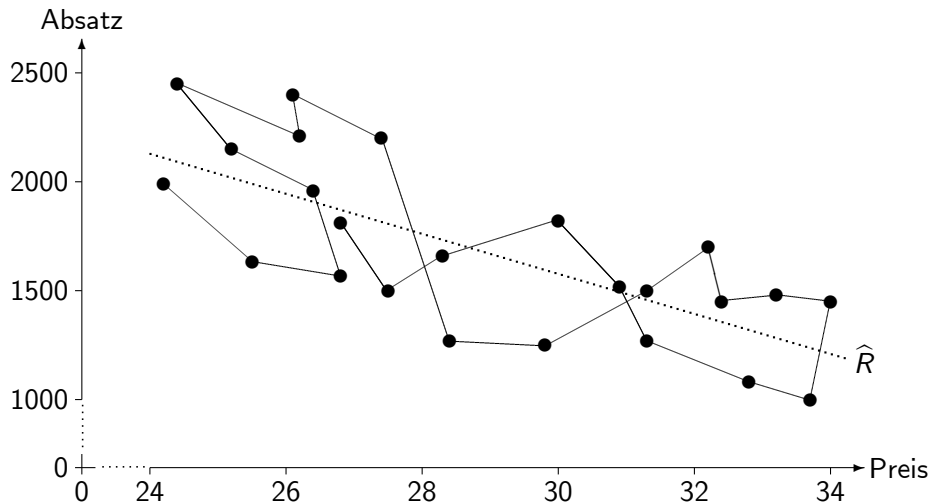


Abbildung 18.1: Die Datenpunkte des Filter-Beispiels.

- Die Preisabsatzfunktion wird durch das folgende ökonometrische Modell beschrieben:

$$y_t = \alpha + \beta x_t + u_t . \quad (18.1)$$

Numerische Illustration 18.1

Tabelle 18.2: Schätzergebnisse einer KQ-Schätzung des Modells (18.1) des Filter-Beispiels.

Variable	Koeff.	$\widehat{\text{sd}}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	4413,33	534,39	8,259	<0,001
Preis	-94,42	18,36	-5,144	<0,001

- Abbildung 18.1 zeigt auch die KQ-Regressionsgerade \hat{R} .
- **Frage:** Warum deutet Abbildung 18.1 auf autokorrelierte Störgrößen?
Antwort:

18.1 Konsequenzen der Annahmeverletzung

- Die Störgröße folge einem autoregressiven Prozess erster Ordnung:

$$u_t = \rho u_{t-1} + e_t . \quad (18.2)$$

- ρ bezeichnet dabei eine Konstante mit

$$-1 < \rho < 1 \quad (18.3)$$

- e_t ist eine Störgröße, welche die Annahmen B1 bis B4 erfüllt:

$$e_t \sim UN(0, \sigma_e^2) .$$

- Ein AR(2)-Prozess hätte die Form:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + e_t . \quad (18.4)$$

18.1.1 Erwartungswert von u_t

- Da

$$\begin{aligned}u_t &= \rho u_{t-1} + e_t \\ &= \rho(\rho u_{t-2} + e_{t-1}) + e_t \\ &= \rho^2 u_{t-2} + \rho e_{t-1} + e_t \\ &\vdots \\ &= \sum_{j=0}^{\infty} \rho^j e_{t-j},\end{aligned}\tag{18.5}$$

ist die Störgröße u_t die gewichtete Summe aus e_t und allen vergangenen Störgrößen e_{t-j} des AR(1)-Prozesses.

- Folglich gilt

$$\begin{aligned} E(u_t) &= E\left(\sum_{j=0}^{\infty} \rho^j e_{t-j}\right) \\ &= \sum_{j=0}^{\infty} \rho^j \underbrace{E(e_{t-j})}_{=0} \\ &= 0. \end{aligned} \tag{18.6}$$

18.1.2 Varianz von u_t

- Die Varianz von u_t lässt sich folgendermaßen schreiben:

$$\begin{aligned}
 \text{var}(u_t) &= E \left[\left[u_t - \overbrace{E(u_t)}^{=0} \right]^2 \right] = E [u_t^2] = E \left[\left[\sum_{j=0}^{\infty} \rho^j e_{t-j} \right]^2 \right] \\
 &= E (e_t^2 + \rho^2 e_{t-1}^2 + \rho^4 e_{t-2}^2 + \dots + 2\rho e_t e_{t-1} + 2\rho^2 e_t e_{t-2} + \dots) \\
 &= E (e_t^2) + \rho^2 E (e_{t-1}^2) + \rho^4 E (e_{t-2}^2) + \dots \\
 &\quad + 2\rho E (e_t e_{t-1}) + 2\rho^2 E (e_t e_{t-2}) + \dots \\
 &= \sigma_e^2 + \rho^2 \sigma_e^2 + \rho^4 \sigma_e^2 + \dots + 2\rho \cdot 0 + 2\rho^2 \cdot 0 + \dots \\
 &= \sigma_e^2 \sum_{j=0}^{\infty} \rho^{2j} .
 \end{aligned} \tag{18.7}$$

- Da

$$\sum_{j=0}^{\infty} \rho^{2j} = \frac{1}{1 - \rho^2} ,$$

erhält man aus Gleichung (18.7):

$$\text{var}(u_t) = \frac{\sigma_e^2}{1 - \rho^2} \geq \sigma_e^2 . \quad (18.8)$$

- Die Varianz von u_t sei wie gewohnt durch σ^2 symbolisiert, also

$$\text{var}(u_t) = \frac{\sigma_e^2}{1 - \rho^2} = \sigma^2 .$$

18.1.3 Kovarianz von u_t und u_{t-1}

- Die Kovarianz ergibt sich aus

$$\begin{aligned}
 \text{cov}(u_t, u_{t-1}) &= E \left[\left[u_t - \overbrace{E(u_t)}^{=0} \right] \left[u_{t-1} - \overbrace{E(u_{t-1})}^{=0} \right] \right] \\
 &= E [u_t u_{t-1}] \\
 &= E [(\rho u_{t-1} + e_t) u_{t-1}] \\
 &= \rho E (u_{t-1}^2) + E(e_t u_{t-1}) .
 \end{aligned}$$

- Frage:** Warum gilt $E(e_t u_{t-1}) = 0$?
- Antwort:**

- Da $E(e_t u_{t-1}) = 0$ und

$$E(u_{t-1}^2) = E(u_t^2) = \text{var}(u_t) = \sigma_e^2 / (1 - \rho^2) \equiv \sigma^2 ,$$

gilt

$$\text{cov}(u_t, u_{t-1}) = \rho \sigma^2 . \quad (18.9)$$

- Ferner kann gezeigt werden, dass

$$\text{cov}(u_t, u_{t-\tau}) = \rho^\tau \sigma^2 \neq 0 . \quad (18.10)$$

- Man beachte, dass

$$\text{cor}(u_t, u_{t-1}) = \frac{\text{cov}(u_t, u_{t-1})}{\text{sd}(u_t) \text{sd}(u_{t-1})} = \frac{\rho \sigma^2}{\sigma \sigma} = \rho .$$

18.1.4 Konsequenzen für die Punktschätzung

- Setzt man Gleichung (18.2), $u_t = \rho u_{t-1} + e_t$, in Beziehung (18.1), $y_t = \alpha + \beta x_t + u_t$, ein, so erhält man

$$y_t = \alpha + \beta x_t + \rho u_{t-1} + e_t . \quad (18.11)$$

- Da Beziehung (18.1) für alle Perioden t gilt, wissen wir, dass

$$y_{t-1} = \alpha + \beta x_{t-1} + u_{t-1}$$

und damit

$$u_{t-1} = y_{t-1} - \alpha - \beta x_{t-1} .$$

- Setzt man diesen Term in Gleichung (18.11) ein, so ergibt sich

$$y_t = \alpha + \beta x_t + \rho y_{t-1} - \rho \alpha - \rho \beta x_{t-1} + e_t .$$

- Umstellen liefert

$$y_t - \rho y_{t-1} = (1 - \rho) \alpha + \beta (x_t - \rho x_{t-1}) + e_t . \quad (18.12)$$

- Definiert man

$$y_t^* = y_t - \rho y_{t-1} \quad (18.13a)$$

$$z_t^* = 1 - \rho \quad (18.13b)$$

$$x_t^* = x_t - \rho x_{t-1} \quad (18.13c)$$

und

$$u_t^* = e_t , \quad (18.14)$$

dann wird Gleichung (18.12) zu

$$y_t^* = \alpha z_t^* + \beta x_t^* + u_t^* . \quad (18.15)$$

- **Frage:** Welche Schlussfolgerung könnte man daraus bezüglich der Effizienz der KQ-Schätzung des Modells (18.1) ziehen?

Antwort:

18.1.5 Konsequenzen für Intervallschätzung und Hypothesentest

- Die Varianz des KQ-Schätzers zu Modell (18.1) lautet:

$$\begin{aligned} \text{var}(\hat{\beta}) = & \frac{\sigma^2}{S_{xx}} \left[1 + \rho \frac{\sum_{t=1}^{T-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{S_{xx}} \right. \\ & + \rho^2 \frac{\sum_{t=1}^{T-2} (x_t - \bar{x})(x_{t+2} - \bar{x})}{S_{xx}} + \dots (18.16) \\ & \left. + \rho^{T-1} \frac{(x_1 - \bar{x})(x_T - \bar{x})}{S_{xx}} \right]. \end{aligned}$$

- Für $\rho=0$ würde der Klammerausdruck zu 1 und man erhalte

$$\text{var}(\hat{\beta}) = \frac{\sigma^2}{S_{xx}}. \quad (18.17)$$

- Eine unverzerrte Schätzung von $var(\hat{\beta})$ erfordert, dass (18.16) benutzt wird, wobei σ^2 durch eine unverzerrte Schätzung $\hat{\sigma}^2$ zu ersetzen ist.
- Da Modell (18.15) alle A-, B- und C-Annahmen erfüllt und $u_t^* = e_t$, ist der Schätzer

$$\hat{\sigma}_e^2 = \frac{S_{\hat{u}^* \hat{u}^*}}{T - 2} \quad (18.18)$$

unverzerrt.

- Einsetzen dieses Wertes in (18.8) würde eine unverzerrte Schätzung von σ^2 liefern.
- Diesen Schätzwert könnte man in (18.16) einsetzen und bekäme eine unverzerrte Schätzung für $var(\hat{\beta})$.

- Unter Missachtung der Autokorrelation würde man hingegen

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T-2} \quad (18.19)$$

benutzen und in Formel (18.17) einsetzen.

- Es kann gezeigt werden, dass der Erwartungswert für Formel (18.19) folgendermaßen lautet:

$$E(\hat{\sigma}^2) = \frac{\sigma^2}{T-2} \left[T - [2/(1-\rho)] - 2\rho \left(\sum_{t=1}^{T-1} (x_t - \bar{x})(x_{t+1} - \bar{x}) / S_{xx} \right) \right] \quad (18.20)$$

- **Frage:** Was folgt aus (18.20) in Bezug auf die Schätzung von σ^2 ? Vergleichen Sie $\rho = 0$ mit $\rho \neq 0$.

Antwort:

- In der Schätzung von $var(\hat{\beta})$ ergibt sich für $\rho \neq 0$ eine doppelte Verzerrung, denn σ^2 würd mit (18.19) statt mit (18.18) geschätzt werden und $var(\hat{\beta})$ mit (18.17) statt mit (18.16).

Folgen der Autokorrelation

Die KQ-Schätzung eines ökonometrischen Modells mit autokorrelierter Störgröße führt zu

- 1 unverzerrten aber ineffizienten Punktschätzern,
- 2 verzerrten Intervallschätzern und
- 3 wertlosen Hypothesentests.

18.2 Diagnose

Numerische Illustration 18.2

Tabelle 18.3: Residuen aus einer KQ-Schätzung des Modells (18.1).

t	\hat{u}_t	t	\hat{u}_t	t	\hat{u}_t	t	\hat{u}_t
1	-138,47	7	270,36	13	326,85	19	-188,12
2	-375,73	8	450,92	14	95,73	20	24,11
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6	340,41	12	41,88	18	-236,50	24	-73,00

- Die Residuen können als Grundlage für die Überprüfung der Autokorrelationsvermutung dienen.

18.2.1 Grafische Analyse

- Abbildung 18.1 war eine mögliche Analysemethode.
- Man kann auch die Residuenpaare $(\hat{u}_{t-1}, \hat{u}_t)$ der KQ-Schätzung des ursprünglichen Modells (18.1) auswerten.

Numerische Illustration 18.3

Tabelle 18.4: Residuenpaare $(\hat{u}_{t-1}; \hat{u}_t)$ aus einer KQ-Schätzung des Modells (18.1).

t	$(\hat{u}_{t-1}; \hat{u}_t)$	t	$(\hat{u}_{t-1}; \hat{u}_t)$
1	nicht bekannt	13	(41,88 ; 326,85)
2	(-138,47 ; -375,73)	14	(326,85 ; 95,73)
⋮	⋮	⋮	⋮
12	(-349,75 ; 41,88)	24	(-316,90 ; -73,00)

- Abbildung 18.2 stellt die 23 verfügbaren Residuenpaare $(\hat{u}_{t-1}; \hat{u}_t)$ als Punktwolke grafisch dar.

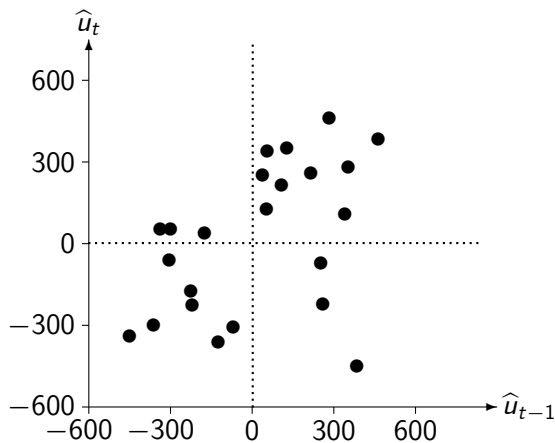


Abbildung 18.2: Grafische Analyse der Residuenpaare $(\hat{u}_{t-1}; \hat{u}_t)$.

18.2.2 Schätzer für ρ

- Folgende Gleichung wird mit der KQ-Methode geschätzt:

$$\hat{u}_t = \rho \hat{u}_{t-1} + e_t^* . \quad (18.21)$$

Formel: KQ-Schätzformel für ρ

$$\hat{\rho} = \frac{\sum_{t=2}^T \hat{u}_{t-1} \hat{u}_t}{\sum_{t=2}^T \hat{u}_{t-1}^2} = \frac{S_{\hat{u}_{t-1} \hat{u}_t}}{S_{\hat{u}_{t-1} \hat{u}_{t-1}}} . \quad (18.22)$$

- **Frage:** Warum beginnt die Summenbildung bei $t = 2$ statt $t = 1$?
Antwort:

- Für große Stichproben besitzt der Schätzer (18.22) gute statistische Eigenschaften.

Numerische Illustration 18.4

Aus den berechneten Residuen des Filter-Beispiels (Tabelle 18.3) ergeben sich für die KQ-Schätzung des Modells (18.21) die folgenden Resultate:

Tabelle 18.5: Schätzergebnisse einer KQ-Schätzung des Modells (18.21).

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
ρ	0,446	0,190	2,35	0,028

(Forts.)

Der $\widehat{sd}(\widehat{\rho})$ -Wert ergab sich aus

$$\widehat{sd}(\widehat{\rho}) = \sqrt{\frac{\widehat{\sigma}_{e^*}^2}{S_{\widehat{u}_{t-1}\widehat{u}_{t-1}}}}, \quad (18.23)$$

wobei $\widehat{\sigma}_{e^*}$ die geschätzte Varianz der Störgröße e_t^* des Modells (18.21) ist.

Der t -Wert für $H_0 : \rho = 0$ ergab sich wie gewohnt aus

$$t = \frac{\widehat{\rho} - 0}{\widehat{sd}(\widehat{\rho})} = \frac{0,446}{0,190} = 2,35 .$$

(Forts.)

Frage: Wie lautet der kritische Wert $t_{0,025}$ und zu welcher Testentscheidung gelangen wir?

Antwort:

- Es mussten in der KQ-Schätzung und im t -Test Residuen \hat{u}_t statt tatsächlich beobachtete Störgrößen u_t verwendet werden.
- Wooldridge (1991) zeigt, dass bei hinreichend großem Stichprobenumfang der t -Test dennoch eingesetzt werden kann.
- Für kleine Stichproben sollte aber ein anderes Verfahren angewendet werden.

18.2.3 Durbin-Watson-Test

- Es wird vermutet, dass in Gleichung (18.2) der Fall $\rho > 0$ vorliegt.

Schritt 1: Formulierung der Nullhypothese

- Die Nullhypothese formuliert das Gegenteil der Anfangsvermutung:

$$H_0 : \rho \leq 0 . \quad (18.24)$$

- Die Alternativhypothese lautet

$$H_1 : \rho > 0 . \quad (18.25)$$

- Das Signifikanzniveau α wird festgelegt.

Schritt 2: Berechnung des Durbin-Watson-Wertes d

Formel: Durbin-Watson-Wert

$$d = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2} . \quad (18.26)$$

Numerische Illustration 18.5

Im Filter-Beispiel wird $H_0 : \rho \leq 0$ getestet.

Die Residuen der KQ-Schätzung des ursprünglichen Modells (18.1) liefern

$$d = \frac{1.836.165}{1.675.152} = 1,096 .$$

- Zusammenhang zwischen d und $\hat{\rho}$:

$$d = \frac{\sum_{t=2}^T \hat{u}_t^2 + \sum_{t=2}^T \hat{u}_{t-1}^2 - 2 \sum_{t=2}^T \hat{u}_{t-1} \hat{u}_t}{\sum_{t=1}^T \hat{u}_t^2}$$

$$= \frac{\sum_{t=2}^T \hat{u}_t^2}{\sum_{t=1}^T \hat{u}_t^2} + \frac{\sum_{t=2}^T \hat{u}_{t-1}^2}{\sum_{t=1}^T \hat{u}_t^2} - \frac{2 \sum_{t=2}^T \hat{u}_{t-1} \hat{u}_t}{\sum_{t=1}^T \hat{u}_t^2} \quad (18.27)$$

$$\approx 1 + 1 - 2\hat{\rho} \quad (18.28)$$

$$= 2(1 - \hat{\rho}) . \quad (18.29)$$

- **Frage:** Welche d -Werte ergeben sich für $\hat{\rho} = -1$, $\hat{\rho} = 0$ und $\hat{\rho} = 1$?

Antwort:

Numerische Illustration 18.6

Aus $\hat{\rho} = 0,446$ und Gleichung (18.29) erhält man einen Wert von

$$\begin{aligned}d &= 1 + 1 - 2 \cdot 0,446 \\ &= 1,108 .\end{aligned}$$

Gleichung (18.26) bzw. (18.27) liefert hingegen einen Wert von

$$\begin{aligned}d &= 0,989 + 0,997 - 2 \cdot 0,445 \\ &= 1,096 .\end{aligned}$$

Schritt 3: Ermittlung des kritischen Wertes

- Je dichter der d -Wert bei 0 liegt ($\hat{\rho}$ dicht bei 1), umso unplausibler wird die Nullhypothese $H_0 : \rho \leq 0$.
- Das Prinzip ist in Abbildung 18.3 veranschaulicht.
- Die Wahrscheinlichkeitsverteilung von d muss mit ökonometrischer Software berechnet werden.
- Diese liefert auch den kritischen Wert d_a .

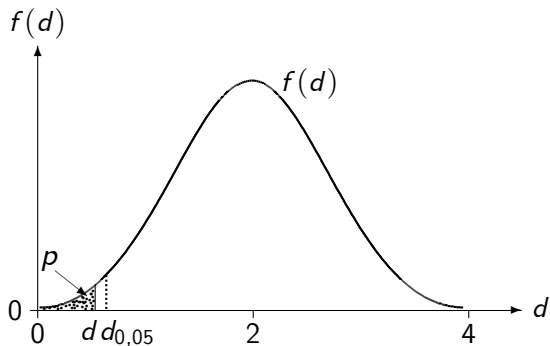


Abbildung 18.3: Eine mögliche Wahrscheinlichkeitsverteilung der Zufallsvariable d .

Schritt 4: Testentscheidung

- Wenn $d < d_a$, dann wird $H_0 : \rho \leq 0$ abgelehnt und $H_1 : \rho > 0$ akzeptiert.

Durbin-Watson-Test auf Basis des p -Wertes

- Ökonometrische Software liefert auch den p -Wert zum berechneten d -Wert.
- Wenn $p < \alpha$, dann wird $H_0 : \rho \leq 0$ abgelehnt und $H_1 : \rho > 0$ akzeptiert.

Numerische Illustration 18.7

Im Filter-Beispiel erhält man zum d -Wert von 1,096 einen p -Wert von 0,4%.

Folglich gilt $p < \alpha$ und die Nullhypothese (keine positive Autokorrelation) wird abgelehnt.

Die Störgrößen folgen anscheinend einem AR(1)-Prozess und somit ist Annahme B3 verletzt.

18.2.4 Durbin-Watson-Test für negative Autokorrelation

- Die Vermutung lautet, dass $\rho < 0$.
- **Frage:** Wie lautet die Nullhypothese und welche d -Werte sprechen gegen diese?

Antwort:

- **Frage:** Wann wird die Nullhypothese verworfen?

Antwort:

18.2.5 Nachteile des Durbin-Watson-Tests

- Der Test ist nicht zulässig, wenn y_t durch eigene Vergangenheitswerte $(y_{t-1}, y_{t-2}, \dots)$ erklärt wird.
- Der Test ist nur zum Testen eines AR(1)-Prozesses einsetzbar.
- Falls keine Software vorhanden ist, die $f(d)$ berechnen kann, können sich Situationen ergeben, in denen keine Testentscheidung getroffen werden kann.

Diagnose der Autokorrelation

Oftmals gibt schon die grafische Betrachtung der Punktwolke oder der Residuen einen Hinweis darauf, ob Autokorrelation vorliegt.

Ferner kann auf Basis der Residuen eine Schätzung des Parameters ρ vorgenommen werden (erfordert aber hinreichenden Stichprobenumfang).

Der Durbin-Watson-Test ist auch bei kleineren Stichprobenumfängen aussagekräftig.

18.3 Anwendbare Schätzverfahren

- Das ökonometrische Modell lautet:

$$y_t = \alpha + \beta x_t + u_t, \quad (18.1)$$

wobei $u_t = \rho u_{t-1} + e_t$.

- Dieses Modell wurde in die Form

$$y_t^* = z_t^* \alpha + x_t^* \beta + u_t^* \quad (18.15)$$

transformiert, wobei

$$y_t^* = y_t - \rho y_{t-1} \quad (18.13a)$$

$$z_t^* = 1 - \rho \quad (18.13b)$$

$$x_t^* = x_t - \rho x_{t-1} \quad (18.13c)$$

$$u_t^* = e_t. \quad (18.14)$$

- **Frage:** Warum ist eine VKQ-Schätzung nicht ohne weiteres durchführbar?

Antwort:

18.3.1 Ermittlung von x_1^* und y_1^*

- Für Beobachtung $t=1$ lautet das ökonometrische Modell:

$$y_1 = \alpha + \beta x_1 + u_1 .$$

- Multiplikation mit $\sqrt{1 - \rho^2}$ führt zu

$$\sqrt{1 - \rho^2} y_1 = \sqrt{1 - \rho^2} \alpha + \sqrt{1 - \rho^2} \beta x_1 + \sqrt{1 - \rho^2} u_1 .$$

- Definiert man

$$y_1^* = \sqrt{1 - \rho^2} y_1, \quad z_1^* = \sqrt{1 - \rho^2} \quad \text{und} \quad x_1^* = \sqrt{1 - \rho^2} x_1 \quad (18.30)$$

sowie

$$u_1^* = \sqrt{1 - \rho^2} u_1, \quad (18.31)$$

so ergibt sich

$$y_1^* = \alpha z_1^* + \beta x_1^* + u_1^*. \quad (18.32)$$

- Für u_1^* ergibt sich der Erwartungswert

$$\begin{aligned} E(u_1^*) &= E\left(\sqrt{1-\rho^2} u_1\right) \\ &= \sqrt{1-\rho^2} \cdot E(u_1) \\ &= 0 \end{aligned}$$

- Die Varianz von u_1^* beträgt

$$\begin{aligned} \text{var}(u_1^*) &= \text{var}\left(\sqrt{1-\rho^2} u_1\right) \\ &= (1-\rho^2) \cdot \text{var}(u_1) \\ &= (1-\rho^2) \frac{\sigma_e^2}{1-\rho^2} \\ &= \sigma_e^2. \end{aligned}$$

- **Frage:** Warum ist auch Annahme B3 für u_1^* erfüllt?
Antwort:

18.3.2 VKQ-Methode von Hildreth und Lu

- Es werden unterschiedliche ρ -Werte für die Datentransformationen (18.30) und (18.13) ausprobiert und jeweils eine KQ-Schätzung des transformierten Modells (18.15) durchgeführt.
- Derjenige ρ -Wert, welcher die geringste Summe der Residuenquadrate $\sum (\hat{u}_t^*)^2$ liefert, wird letztlich verwendet.

Numerische Illustration 18.8

Tabelle 18.6: Schätzergebnisse des Filter-Beispiels für unterschiedliche ρ -Werte.

ρ	0,40	0,44	0,45	0,46
$S_{\hat{u}^* \hat{u}^*}$	1.343.065	1.338.892	1.338.674	1.338.786
Konstante	4.386,47	4.385,91	4.385,93	4.386,02
Preis	-93,68	-93,69	-93,70	-93,71

18.3.3 GVKQ-Methode von Cochrane und Orcutt

- Man verfährt nach den folgenden Schritten:
 - 1 Es wird eine KQ-Schätzung des ursprünglichen Modells (18.1) durchgeführt.
 - 2 Es werden die entsprechenden Residuen gemäß $\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta}x_t$ ermittelt.
 - 3 Ein Schätzwert für ρ wird gemäß (18.22) berechnet.
 - 4 Benutzt man den Schätzwert $\hat{\rho}$ für die Variablenberechnung in (18.13) und (18.30), so erhält man numerische Werte für z_t^* , x_t^* und y_t^* .
 - 5 Auf Basis der Werte z_t^* , x_t^* und y_t^* wird eine KQ-Schätzung des transformierten Modells (18.15) vorgenommen.

- Beim *iterativen Cochrane-Orcutt-Verfahren* werden die Schritte 2 bis 5 mehrmals wiederholt.
- Für die Berechnung der Residuen (Schritt 2) werden dabei die aus der GVKQ-Methode ermittelten Schätzwerte $\hat{\alpha}^{GVKQ}$ und $\hat{\beta}^{GVKQ}$ benutzt:

$$\hat{u}_t^{**} = y_t - \hat{\alpha}^{GVKQ} - \hat{\beta}^{GVKQ} x_t . \quad (18.33)$$

Numerische Illustration 18.9

Tabelle 18.7: Schätzergebnisse des iterativen Cochrane-Orcutt-Verfahrens.

Variable	KQ-Koeff.	1. Iteration	2. Iteration
Konstante	4.413,33	4.385,91	4.385,92
Preis	-94,42	-93,69	-93,69
verwend. $\hat{\rho}$	—	0,446	0,447

Kapitel 19

Annahme B4: Normalverteilte Störgrößen

Annahme B4 (Normalverteilung)

Die Störgrößen u_t sind normalverteilt.

Beispiel Wirtschaftswachstum

In einem einfachen neoklassischen Solow-Wachstumsmodell lässt sich unter Annahme einer Cobb-Douglas-Produktionsfunktion das logarithmierte Pro-Kopf-Einkommen (y_t) eines Landes im langfristigen Gleichgewicht in Abhängigkeit der logarithmierten Sparquote (x_{1t}) und einer zweiten Variable (x_{2t}) darstellen, nämlich der logarithmierten Summe der Raten des Bevölkerungswachstums, des technischen Fortschritts und der Kapitalabschreibungen (Daten in Tabelle 19.1):

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t . \quad (19.1)$$

Beispiel (Forts.)

Tabelle 19.1: Logarithmiertes Pro-Kopf-Einkommen y_t , logarithmierte Sparquote x_{1t} und logarithmierte Summe aus den Raten des Bevölkerungswachstums, des technischen Fortschritts und der Kapitalabschreibungen x_{2t} für 75 Länder (wobei die letzten zwei Raten als für alle Länder identisch unterstellt werden).

t	y_t	x_{1t}	x_{2t}	t	y_t	x_{1t}	x_{2t}	t	y_t	x_{1t}	x_{2t}
1	8,38	3,18	-2,58	26	8,66	3,14	-2,50	51	8,41	2,69	-2,47
2	8,21	3,34	-2,50	27	7,68	2,50	-2,53	52	8,10	2,84	-2,54
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	8,47	3,10	-2,56	50	9,79	3,15	-2,66	75	9,42	3,11	-2,70

19.1 Konsequenzen der Annahmeverletzung

- Die KQ-Methode bewahrt die BLUE-Eigenschaft.
- Probleme ergeben sich nur bei Intervallschätzungen und Hypothesentests.

Numerische Illustration 19.1

Mankiw, Romer und Weil (1992) führen mit Hilfe der Daten der Tabelle 19.1 eine KQ-Schätzung des ökonometrischen Modells (19.1) durch:

Tabelle 19.2: KQ-Schätzergebnisse für das Wachstumsmodell (19.1).

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	-0,722	1,324	-0,545	0,587
log. Sparquote	1,318	0,171	7,708	<0,001
log. Bevölkerungsw.	-2,017	0,534	-3,778	<0,001

- Zufallsvariablen, auf die eine große Zahl von unabhängigen kleinen Einflussfaktoren wirkt, sind näherungsweise normalverteilt.
- Der *zentrale Grenzwertsatz* stellt dies sicher.
- In diesen Fällen ist Annahme B4 näherungsweise erfüllt.
- Selbst wenn Annahme B4 nicht erfüllt ist, sorgt der zentrale Grenzwertsatz dafür, dass bei hinreichend großem Stichprobenumfang T die KQ-Schätzer näherungsweise normalverteilt sind.

Folgen der Annahmeverletzung

Bei großem Stichprobenumfang führt die KQ-Schätzung eines ökonometrischen Modells mit einer nicht normalverteilten Störgröße nur selten zu Problemen, denn die KQ-Schätzer sind dennoch näherungsweise normalverteilt.

19.2 Diagnose

- Um etwas über die statistischen Eigenschaften der Störgrößen u_t zu erfahren, ist es immer sinnvoll die Residuen \hat{u}_t zu betrachten.

19.2.1 Grafische Analyse

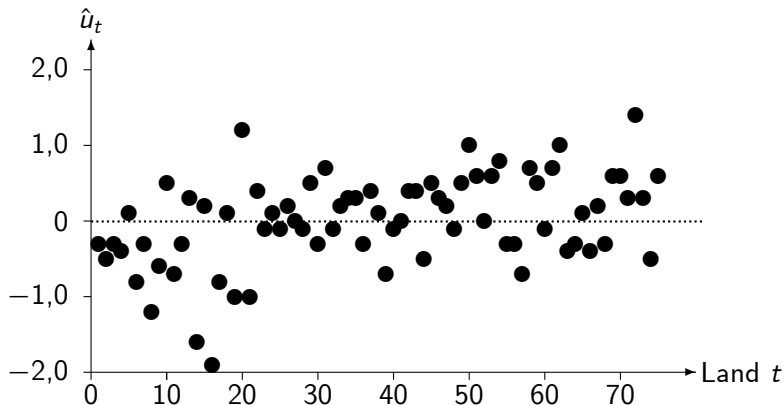


Abbildung 19.1: Residuen \hat{u}_t der 75 Länder aus einer KQ-Schätzung des Wachstumsmodells (19.1).

- Eine grafische Diagnose kann auch mit einem *Histogramm* vorgenommen werden.
- Die Standardabweichung der Störgrößen lautet:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{S_{\hat{u}\hat{u}}}{T - K - 1}}. \quad (19.2)$$

- Der Wertebereich der Residuen wird in Abschnitte unterteilt (vgl. Abbildung 19.2).
- Im Beispiel besitzen die Abschnitte eine Breite von $0,5\hat{\sigma}$.
- Die T Residuen werden den Abschnitten zugeordnet.

Numerische Illustration 19.2

Für die Standardabweichung ergibt sich:

$$\hat{\sigma} = \sqrt{26,848/72} = 0,611 .$$

Das Residuum der Beobachtung $t=1$ besitzt einen Wert von $\hat{u}_1 = -0,287$. Es liegt im Abschnitt $[-0,5\hat{\sigma} ; 0]$.

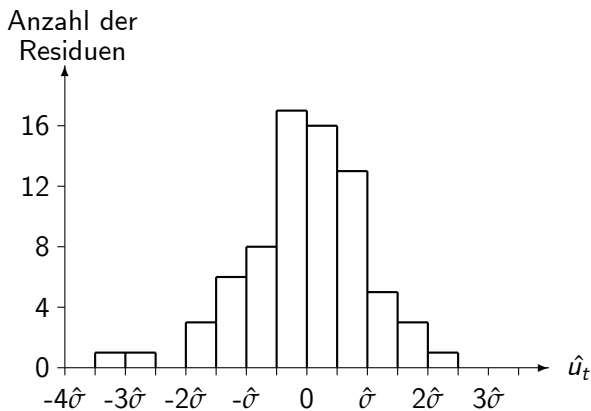


Abbildung 19.2: Histogramm der Residuen aus einer KQ-Schätzung des Wachstumsmodells (19.1).

19.2.2 Jarque-Bera-Test

- Die Varianz ist das *zentrale Moment zweiter Ordnung*:
 $E[(u_t - E(u_t))^2]$.
- Das zentrale Moment dritter Ordnung quantifiziert die Schiefe einer Verteilung.
- Ein verbreitetes Schiefemaß ist

$$asym(u_t) = \frac{E[(u_t - E(u_t))^3]}{[var(u_t)]^{3/2}} .$$

- **Frage:** Welcher Wert $asym(u_t)$ ergibt sich für die Normalverteilung?

Antwort:

- Das zentrale Moment vierter Ordnung quantifiziert die „Spitzigkeit“ (Kurtosis) einer Verteilung.
- Ein gebräuchliches Kurtosismaß ist

$$kur(u_t) = \frac{E \left[(u_t - E(u_t))^4 \right]}{[var(u_t)]^2} .$$

- Für die Normalverteilung gilt: $kur(u_t) = 3$.

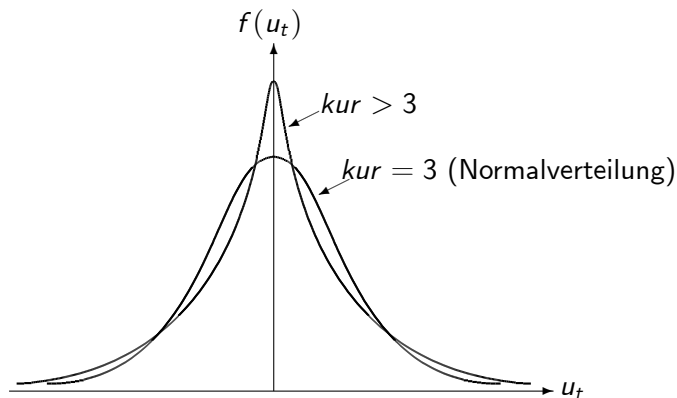


Abbildung 19.3: Kurtosis zweier Verteilungen mit ähnlichen Varianzen.

- Die Schätzer für $asym(u_t)$ und $kur(u_t)$ lauten:

$$\widehat{asym}(u_t) = \frac{\frac{1}{T} \sum \widehat{u}_t^3}{\left(\frac{1}{T} \sum \widehat{u}_t^2\right)^{3/2}} \quad (19.3)$$

$$\widehat{kur}(u_t) = \frac{\frac{1}{T} \sum \widehat{u}_t^4}{\left(\frac{1}{T} \sum \widehat{u}_t^2\right)^2} . \quad (19.4)$$

- Weichen die Schätzwerte $\widehat{sym}(u_t)$ und $\widehat{kur}(u_t)$ deutlich von 0 bzw. 3 ab, dann sind die Störgrößen eventuell nicht normalverteilt.

- Jarque und Bera (1987) leiten das folgende Resultat her:

Jarque-Bera-Test

Wenn die Störgrößen u_t normalverteilt sind, dann gilt:

$$JB = T \left[\frac{[\widehat{sym}(u_t)]^2}{6} + \frac{[\widehat{kur}(u_t) - 3]^2}{24} \right] \sim \chi^2_{(2)}. \quad (19.5)$$

- **Frage:** Warum deutet ein hoher JB -Wert auf eine Verletzung der Annahme B4 hin?

Antwort:

Numerische Illustration 19.3

Die Residuen \hat{u}_t aus der KQ-Schätzung des Modells (19.1) liefern:

$$\widehat{sym}(u_t) = \frac{\frac{1}{T} \sum \hat{u}_t^3}{\left(\frac{1}{T} \sum \hat{u}_t^2\right)^{3/2}} = \frac{-0,115}{(0,358)^{3/2}} = -0,537$$

$$\widehat{kur}(u_t) = \frac{\frac{1}{T} \sum \hat{u}_t^4}{\left(\frac{1}{T} \sum \hat{u}_t^2\right)^2} = \frac{0,488}{(0,358)^2} = 3,808 .$$

Es ergibt sich ein *JB*-Wert von 5,64985.

Bei einem Signifikanzniveau von 5% beträgt der kritische Wert der $\chi^2_{(2)}$ -Verteilung 5,99146.

Kapitel 20

Annahme C1: Zufallsunabhängige exogene Variablen

Annahme C1 (Zufallsunabhängige exogene Variablen)

Die exogenen Variablen $x_{1t}, x_{2t}, \dots, x_{Kt}$ sind keine Zufallsvariablen, sondern können wie in einem Experiment kontrolliert werden.

20.1 Weitere Qualitätskriterien für Schätzer: Konsistenz und asymptotische Effizienz

- Für den KQ-Schätzer einer Einfachregression wurde in Kapitel 4 die folgende Eigenschaft hergeleitet:

$$\hat{\beta} = \beta + \frac{\sum (x_t - \bar{x}) u_t}{S_{xx}} . \quad (20.1)$$

- Folglich gilt:

$$E(\hat{\beta}) = \beta + E \left[\frac{\sum (x_t - \bar{x}) u_t}{S_{xx}} \right]. \quad (20.2)$$

$$= \beta + E \left[\frac{(x_1 - \bar{x}) u_1}{S_{xx}} + \dots + \frac{(x_T - \bar{x}) u_T}{S_{xx}} \right]$$

$$= \beta + \left[E \frac{(x_1 - \bar{x}) u_1}{S_{xx}} + \dots + E \frac{(x_T - \bar{x}) u_T}{S_{xx}} \right]$$

$$= \beta + \sum \left[E \left[\frac{x_t - \bar{x}}{S_{xx}} \cdot u_t \right] \right]. \quad (20.3)$$

- Annahme C1 erlaubt, diese Gleichung umzuformen zu

$$\begin{aligned} E(\hat{\beta}) &= \beta + \sum \frac{x_t - \bar{x}}{S_{xx}} [E(u_t)] \\ &= \beta + \frac{\sum (x_t - \bar{x}) E(u_t)}{S_{xx}} . \end{aligned} \quad (20.4)$$

- Da $E(u_t) = 0$, ergibt sich

$$E(\hat{\beta}) = \beta . \quad (20.5)$$

- Auch die Effizienz der KQ-Schätzer konnte gezeigt werden.
- Es existieren aber noch weitere Qualitätskriterien.

20.1.1 Konsistenz

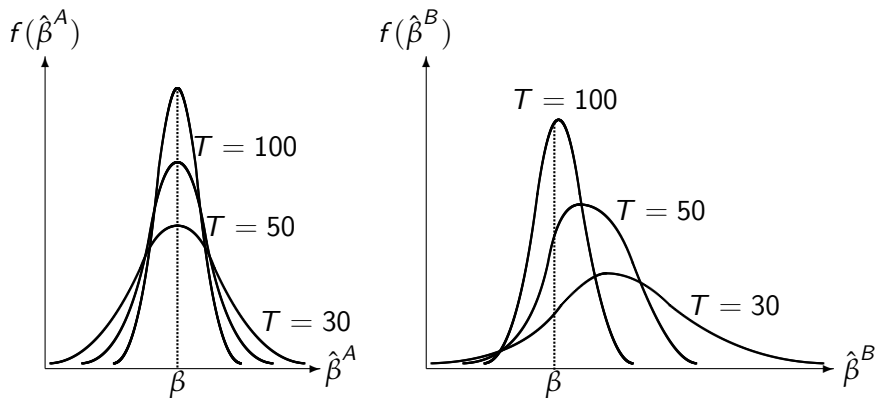


Abbildung 20.1: Die Wirkung steigender Stichprobenumfänge T auf die Wahrscheinlichkeitsverteilungen der Schätzer $\hat{\beta}^A$ und $\hat{\beta}^B$.

Definition: Konsistenz

Wenn der Beobachtungsumfang T gegen unendlich geht und dabei der Grenzwert von $\hat{\beta}^A$ dem wahren Wert β entspricht, dann ist der Schätzer $\hat{\beta}^A$ konsistent:

$$\text{plim}(\hat{\beta}^A) = \beta .$$

- Zwei weitere mögliche asymptotische Eigenschaften des Schätzers $\hat{\beta}^A$:

$$\lim_{T \rightarrow \infty} E(\hat{\beta}^A) = \beta \quad (20.6)$$

$$\lim_{T \rightarrow \infty} \text{var}(\hat{\beta}^A) = 0. \quad (20.7)$$

- (20.6) besagt, dass der Schätzer $\hat{\beta}^A$ *asymptotisch unverzerrt* ist.
- (20.7) besagt, dass die Varianz von $\hat{\beta}^A$ bei unendlich großem Beobachtungsumfang T gegen 0 geht.
- Die Eigenschaften (20.6) und (20.7) sind hinreichend, aber nicht notwendig für Konsistenz.

- **Frage:** Ist der Schätzer $\hat{\beta}^C$ (Abbildung 20.2) asymptotisch unverzerrt und konsistent?

Antwort:

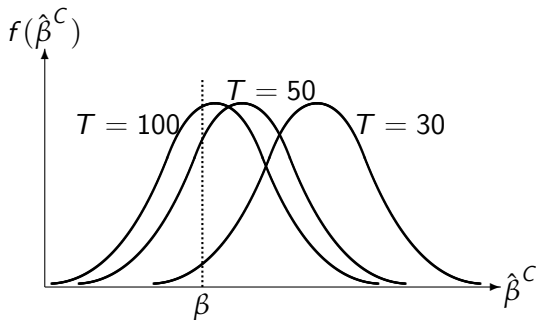


Abbildung 20.2: Die Wirkung steigender Stichprobenumfänge T auf die Wahrscheinlichkeitsverteilungen des asymptotisch unverzerrten Schätzers $\hat{\beta}^C$.

- Bei Gültigkeit sämtlicher A-, B- und C-Annahmen beträgt die Varianz des KQ-Schätzers

$$\text{var}(\hat{\beta}) = \sigma^2 / S_{xx} .$$

- **Frage:** Warum ist der KQ-Schätzer $\hat{\beta}$ konsistent?
Antwort:

20.1.2 Rechenregeln für Wahrscheinlichkeitsgrenzwerte

- Es seien $\hat{\beta}^A$ und $\hat{\beta}^B$ zwei beliebige konsistente Schätzer.
- Das heißt, $\text{plim}(\hat{\beta}^A)$ und $\text{plim}(\hat{\beta}^B)$ existieren und entsprechen den jeweiligen wahren β -Werten.
- Ferner sei c eine beliebige Konstante.
- Es gilt dann:

$$\begin{aligned}\text{plim}(c) &= c \\ \text{plim}(\hat{\beta}^A + \hat{\beta}^B) &= \text{plim}(\hat{\beta}^A) + \text{plim}(\hat{\beta}^B)\end{aligned}$$

- Ferner gilt:

$$\begin{aligned}\text{plim}(\hat{\beta}^A \cdot \hat{\beta}^B) &= \text{plim}(\hat{\beta}^A) \cdot \text{plim}(\hat{\beta}^B) \\ \text{plim}\left(\frac{\hat{\beta}^A}{\hat{\beta}^B}\right) &= \frac{\text{plim}(\hat{\beta}^A)}{\text{plim}(\hat{\beta}^B)}.\end{aligned}$$

20.1.3 Asymptotische Effizienz

- In Abbildung 20.1 gilt:

$$\text{var}(\hat{\beta}^A) < \text{var}(\hat{\beta}^B) \quad \text{für alle Stichprobenumfänge } T .$$

Definition: Asymptotische Effizienz

Wenn ein konsistenter Schätzer $\hat{\beta}^A$ schneller als *alle* anderen konsistenten Schätzer gegen den wahren Wert β konvergiert, dann wird er als *asymptotisch effizient* bezeichnet.

20.2 Konsequenzen der Annahmeverletzung

Fall 1 Die Störgrößen u_t und die Beobachtungen x_s sind unabhängig voneinander.

Unabhängigkeit impliziert, dass $cov(x_s, u_t) = 0$, für alle $s, t = 1, 2, \dots, T$.

Fall 2 Die Störgrößen u_t und die Beobachtungen x_s sind kontemporär (d.h. $s = t$) unkorreliert:

$$cov(x_t, u_t) = 0, \text{ für alle } t = 1, 2, \dots, T. \quad (20.8)$$

Fall 3 Die Störgrößen u_t und die Beobachtungen x_t sind kontemporär korreliert:

$$cov(x_t, u_t) \neq 0, \text{ für mind. eine Beobachtung } t. \quad (20.9)$$

20.2.1 Fall 1: Störgrößen und Beobachtungen der exogenen Variable unabhängig

- Ausgangspunkt ist wieder:

$$E(\hat{\beta}) = \beta + \sum \left[E \left[\frac{x_t - \bar{x}}{S_{xx}} \cdot u_t \right] \right] . \quad (20.3)$$

- Folglich erhält man

$$\begin{aligned} E(\hat{\beta}) &= \beta + \sum \left[E \left[\frac{x_t - \bar{x}}{S_{xx}} \right] \cdot \underbrace{E(u_t)}_{=0} \right] \\ &= \beta . \end{aligned} \quad (20.10)$$

Asymptotische Eigenschaften der Variablen x_t und u_t

- Es wird angenommen, dass

$$\lim_{T \rightarrow \infty} E(\bar{x}) = q_x \text{ und } \lim_{T \rightarrow \infty} E(\sum x_t^2 / T) = q_{xx} \quad (20.11)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{x}) = 0 \text{ und } \lim_{T \rightarrow \infty} \text{var}(\sum x_t^2 / T) = 0, \quad (20.12)$$

wobei q_x und q_{xx} feste Grenzwerte darstellen.

- Folglich gilt:

$$\text{plim}(\bar{x}) = q_x \text{ und } \text{plim}(\sum x_t^2 / T) = q_{xx}. \quad (20.13)$$

- Man erhält deshalb:

$$\begin{aligned}\text{plim}(S_{xx} / T) &= \text{plim} \left[(\sum x_t^2 - T\bar{x}^2) / T \right] \\ &= \text{plim}(\sum x_t^2 / T - \bar{x}^2) \\ &= \text{plim}(\sum x_t^2 / T) - \text{plim}(\bar{x}) \cdot \text{plim}(\bar{x}) \\ &= q_{xx} - (q_x)^2 \equiv \sigma_x^2, \quad (20.14)\end{aligned}$$

wobei σ_x^2 einen festen Grenzwert darstellt.

- Ferner wird angenommen, dass

$$\lim_{T \rightarrow \infty} \sigma_u^2 = \sigma_u^2, \quad (20.15)$$

wobei auch σ_u^2 einen festen Grenzwert darstellt.

Konsistenz der KQ-Schätzer

- Erweitert man den Quotienten in (20.1) mit $(1/T)$, so erhält man

$$\hat{\beta} = \beta + \frac{(1/T) \sum (x_t - \bar{x}) u_t}{(1/T) S_{xx}} . \quad (20.16)$$

- Es kann formal gezeigt werden, dass

$$\text{plim} \left[(1/T) \sum (x_t - \bar{x}) u_t \right] = 0 . \quad (20.17)$$

- **Frage:** Welche intuitive Erklärung kann hierfür angeführt werden?

Antwort:

- Aus (20.14) und (20.17) folgt, dass Zähler und Nenner in (20.16) einen Wahrscheinlichkeitsgrenzwert besitzen.

- Folglich liefert (20.16):

$$\begin{aligned}\text{plim}(\hat{\beta}) &= \text{plim}(\beta) + \text{plim} \left[\frac{(1/T) \sum (x_t - \bar{x}) u_t}{(1/T) S_{xx}} \right] \\ &= \beta + \frac{\text{plim} [(1/T) \sum (x_t - \bar{x}) u_t]}{\text{plim} [(1/T) S_{xx}]} . \quad (20.18)\end{aligned}$$

- Setzt man die Ergebnisse (20.14) und (20.17) in (20.18) ein, so erhält man

$$\text{plim}(\hat{\beta}) = \beta + \frac{0}{q_{xx} - q_x^2} = \beta .$$

Fazit

- Annahme c1 kann durch folgende schwächere Annahme ersetzt werden:

Annahme c1* Die Zufallsvariablen x_s ($s=1,2,\dots,T$) und die Störgrößen u_t ($t=1,2,\dots,T$) sind voneinander unabhängig und

$$\lim_{T \rightarrow \infty} E(\bar{x}) = q_x, \quad \lim_{T \rightarrow \infty} E(\sum x_t^2 / T) = q_{xx}, \quad (20.11)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{x}) = 0, \quad \lim_{T \rightarrow \infty} \text{var}(\sum x_t^2 / T) = 0, \quad (20.12)$$

$$\lim_{T \rightarrow \infty} \sigma^2 = \sigma_u^2. \quad (20.15)$$

KQ-Schätzung in Fall 1

Unter Annahme c1* bleiben die KQ-Schätzer unverzerrt und sie sind auch konsistent. Ferner kann gezeigt werden, dass sie effizient sind und σ^2 , $var(\alpha)$ und $var(\beta_k)$ unverzerrt geschätzt werden.

20.2.2 Fall 2: Störgrößen und Beobachtungen der exogenen Variable kontemporär unkorreliert

- Kontemporäre Unkorreliertheit bedeutet, dass

$$\text{cov}(x_t, u_t) = 0 \quad \text{für alle } t = 1, 2, \dots, T. \quad (20.8)$$

- Es ist aber nicht garantiert, dass auch

$$\text{cov}(x_s, u_t) = 0 \quad \text{für } s \neq t.$$

- Es kann also weiterhin

$$E(\hat{\beta}) = \beta + E \left[\frac{\sum (x_t - \bar{x}) u_t}{S_{xx}} \right] \quad (20.2)$$

in

$$E(\hat{\beta}) = \beta + \sum \left[E \left[\frac{(x_t - \bar{x})}{S_{xx}} u_t \right] \right] \quad (20.3)$$

überführt werden, aber (20.3) nicht in $E(\hat{\beta}) = \beta$.

- Unter den Annahmen (20.11), (20.12) und (20.15) ist Gleichung (20.18) auch hier gültig:

$$\text{plim}(\hat{\beta}) = \beta .$$

Fazit

- Annahme c1* kann durch folgende schwächere Annahme ersetzt werden:

Annahme c1** Die Werte der Zufallsvariablen x_t und die Störgrößenwerte u_t ($t = 1, 2, \dots, T$) sind *kontemporär* unkorreliert:

$$\text{cov}(x_t, u_t) = 0 \quad \text{für alle } t = 1, 2, \dots, T .$$

Ferner gilt

$$\lim_{T \rightarrow \infty} E(\bar{x}) = q_x , \quad \lim_{T \rightarrow \infty} E(\sum x_t^2 / T) = q_{xx} , \quad (20.11)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{x}) = 0 , \quad \lim_{T \rightarrow \infty} \text{var}(\sum x_t^2 / T) = 0 , \quad (20.12)$$

$$\lim_{T \rightarrow \infty} \sigma^2 = \sigma_u^2 . \quad (20.15)$$

KQ-Schätzung in Fall 2

Unter Annahme $c1^{**}$ sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ verzerrt, aber konsistent. Intervallschätzer und Hypothesentests bewahren bei hinreichend großem Stichprobenumfang ihre Gültigkeit.

20.2.3 Eine mögliche Ursache für Fall 2: y_{t-1} als „exogene Variable“

- Das klassische Beispiel für Fall 2 ist das dynamische Modell

$$y_t = \alpha + \beta y_{t-1} + u_t . \quad (20.19)$$

- Der für Periode t gültige Wert der „exogenen Variable“ (y_{t-1}) ist mit u_{t-1} , der Störgröße der Periode $t-1$ korreliert, denn

$$y_{t-1} = \alpha + \beta y_{t-2} + u_{t-1} .$$

- Im Falle einer solchen „exogenen Variable“ lauten die Forderungen (20.11) der Annahme $c1^{**}$:

$$\lim_{T \rightarrow \infty} E(\bar{y}) = q_y, \quad \lim_{T \rightarrow \infty} E(\sum y_t^2 / T) = q_{yy} .$$

- Wenn y_t im Zeitablauf stetig steigt, sind diese Forderungen nicht erfüllt und damit (20.14), also $\text{plim}(S_{yy} / T) = q_{yy} - q_y^2 \equiv \sigma_y^2$, nicht herleitbar.

20.2.4 Fall 3: Störgrößen und Beobachtungen der exogenen Variable kontemporär korreliert

- Es gilt $cov(x_t, u_t) \neq 0$ für mindestens eine Beobachtung t .
- Es werde der Fall $cov(x_t, u_t) > 0$ und $E(x_t) = \tilde{x}$ (für alle x_t) betrachtet, also

$$cov(x_t, u_t) = E \left[\underbrace{[x_t - E(x_t)]}_{= \tilde{x}} \underbrace{[u_t - E(u_t)]}_{= 0} \right] > 0 .$$

- Die Folgen sind in Abbildung 20.3 (linker Teil) veranschaulicht.
- Die Folgen von $cov(x_t, u_t) < 0$ sind im rechten Teil veranschaulicht.

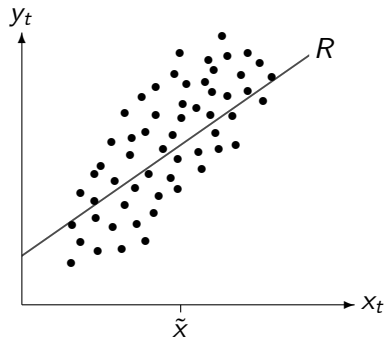
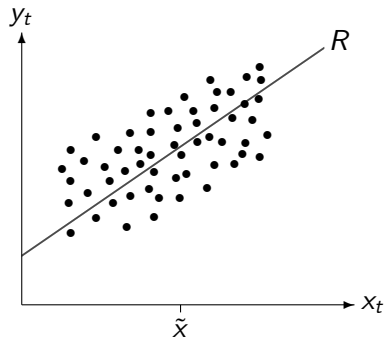
(a) $cov(x_t, u_t) > 0$ (b) $cov(x_t, u_t) < 0$

Abbildung 20.3: Die Lage der beobachteten Datenpunkte und der wahren Gerade R wenn (a) $cov(x_t, u_t) > 0$ bzw. wenn (b) $cov(x_t, u_t) < 0$.

- Formal folgt die Nicht-Konsistenz wieder aus der Gleichung

$$\hat{\beta} = \beta + \frac{(1/T) \sum (x_t - \bar{x}) u_t}{(1/T) S_{xx}} . \quad (20.16)$$

- Der Zähler ist ein geeigneter Schätzer für $cov(x_t, u_t)$.
- Es gilt deshalb

$$\text{plim}(1/T) \left[\sum (x_t - \bar{x}) u_t \right] \neq 0$$

und unter Beachtung von (20.14):

$$\text{plim}(\hat{\beta}) \neq \beta .$$

Fazit

Annahme c1*** Es gilt:

$$\lim_{T \rightarrow \infty} E(\bar{x}) = q_x, \quad \lim_{T \rightarrow \infty} E(\sum x_t^2 / T) = q_{xx}, \quad (20.11)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{x}) = 0, \quad \lim_{T \rightarrow \infty} \text{var}(\sum x_t^2 / T) = 0, \quad (20.12)$$

$$\lim_{T \rightarrow \infty} \sigma^2 = \sigma_u^2. \quad (20.15)$$

KQ-Schätzung in Fall 3

Unter Annahme c1*** sind die KQ-Schätzer $\hat{\alpha}$ und $\hat{\beta}$ verzerrt und nicht konsistent.

20.2.5 Eine mögliche Ursache für Fall 3: Probleme bei der Erfassung der exogenen Variable

- Ein unsystematischer Fehler bei der Erfassung der *exogenen Variable* ergibt sich beispielsweise dann, wenn mit einer *Proxyvariable* x_t gearbeitet werden muss, die mit der eigentlichen Variable x_t^* korreliert ist:

$$x_t = x_t^* + v_t . \quad (20.20)$$

- Dabei ist v_t als Erfassungsfehler zu interpretieren.
- Es wird im Folgenden unterstellt, dass

$$\text{cov}(x_t^*, v_t) = 0 \quad \text{für alle } t = 1, 2, \dots, T . \quad (20.21)$$

Beispiel Versicherungsverträge

Es soll die Rentabilität einer Mitarbeiterschulung für die 30 Außendienstmitarbeiter einer Versicherungsgesellschaft eingeschätzt werden. Zu diesem Zweck wird der Zusammenhang zwischen der allgemeinen Verkaufsbefähigung der Mitarbeiter und der Zahl der verkauften Verträge untersucht (Daten in Tabelle 20.1).

Beispiel (Forts.)

Tabelle 20.1: Abgeschlossene Versicherungsverträge y_t , Testergebnis im Assessmentcenter x_t , Anzahl der im Vorjahr abgeschlossenen Verträge z_t und tatsächliche Verkaufsbefähigung x_t^* für 30 Mitarbeiter.

t	y_t	x_t	z_t	x_t^*	t	y_t	x_t	z_t	x_t^*
1	11	40	14	32	16	24	94	24	81
2	19	44	18	43	17	11	49	18	38
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	31	66	25	73	30	29	82	28	80

Störgröße eines Modells mit fehlerhaft erfasster exogener Variable

- Der wahre Zusammenhang laute

$$y_t = \alpha + \beta x_t^* + e_t, \quad (20.22)$$

- Gemäß Gleichung (20.20) gelte

$$x_t^* = x_t - v_t.$$

- Ferner sei

$$\text{cov}(x_t, e_t) = 0 \quad \text{für alle } t = 1, 2, \dots, T. \quad (20.23)$$

- Es wird angenommen, dass

$$\text{cov}(e_s, v_t) = 0 \quad (20.24)$$

für alle $s = 1, 2, \dots, T$ und $t = 1, 2, \dots, T$.

- Ferner wird angenommen, dass

$$E(e_t) = 0 \quad E(v_t) = 0 \quad (20.25a)$$

$$\text{var}(e_t) = \sigma_e^2 \quad \text{var}(v_t) = \sigma_v^2 \quad (20.25b)$$

$$\text{cov}(e_s, e_t) = 0 \quad \text{cov}(v_s, v_t) = 0 \text{ für } t \neq s \quad (20.25c)$$

- Unter Rückgriff auf Rechenregel (2.20) und Annahme (20.21) ergibt sich für die Varianz der Proxyvariable x_t die folgende Zerlegung:

$$\begin{aligned}\sigma_x^2 &= \text{var}(x_t) \\ &= \text{var}(x_t^* + v_t) \\ &= \text{var}(x_t^*) + \text{var}(v_t) + 2\text{cov}(x_t^*, v_t) \\ &= \sigma_{x^*}^2 + \sigma_v^2.\end{aligned}$$

- Folglich gilt immer:

$$\sigma_v^2 \leq \sigma_x^2. \quad (20.26)$$

- Ersetzt man im Modell (20.22) x_t^* durch $(x_t - v_t)$, so ergibt sich

$$\begin{aligned}y_t &= \alpha + \beta x_t + e_t - \beta v_t \\ &= \alpha + \beta x_t + u_t\end{aligned}\quad (20.27)$$

wobei

$$u_t = e_t - \beta v_t . \quad (20.28)$$

- Gleichung (20.27) beschreibt den wahren Wirkungszusammenhang zwischen y_t und der beobachtbaren Variable x_t .
- Es wird nun untersucht, ob die Störgröße u_t die Annahmen B1 bis B3 erfüllt.

Annahme B1

- Aus (20.28) erhält man:

$$\begin{aligned} E(u_t) &= E(e_t - \beta v_t) \\ &= E(e_t) - \beta E(v_t) \\ &= 0. \end{aligned}$$

- **Frage:** Was bedeutet bei zufallsabhängigen x_t -Werten $E(u_t) = 0$ in grafischer Hinsicht (siehe Abbildung 20.3)?
Antwort:

- Abbildung 20.3 veranschaulicht, dass es im Falle einer C1-Verletzung trotz Erfüllung der Annahme B1 zu einer Verzerrung der KQ-Schätzer kommen kann.

Annahme B2

- Die Rechenregeln für Varianzen liefern:

$$\begin{aligned} \text{var}(u_t) &= \text{var}(e_t - \beta v_t) \\ &= \text{var}(e_t) + \beta^2 \text{var}(v_t) - 2\beta \text{cov}(e_t, v_t) \\ &= \sigma_e^2 + \beta^2 \sigma_v^2 . \end{aligned}$$

- **Frage:** Ist für u_t folglich Annahme B2 verletzt oder erfüllt?

Antwort:

Annahme B3:

- Die Kovarianz zwischen u_t und u_s lautet:

$$\begin{aligned} \text{cov}(u_t, u_s) &= E [[u_t - E(u_t)] [u_s - E(u_s)]] \\ &= E(u_t u_s) = E [(e_t - \beta v_t) (e_s - \beta v_s)] \\ &= E (e_t e_s) - \beta E(e_t v_s) - \beta E(e_s v_t) + \beta^2 E(v_t v_s) \\ &= \text{cov}(e_t, e_s) - \beta \text{cov}(e_t, v_s) - \beta \text{cov}(e_s, v_t) + \beta^2 \text{cov}(v_t, v_s) \\ &= 0 . \end{aligned}$$

- Also ist auch Annahme B3 erfüllt.

Verzerrung der KQ-Schätzer

- Gemäß (20.20) gilt

$$x_t = x_t^* + v_t$$

- Gemäß (20.28) gilt

$$u_t = e_t - \beta v_t.$$

- **Frage:** Warum ergibt sich dann bei positivem β der Fall $cov(x_t, u_t) < 0$?
- **Antwort:**
- Um formal zu zeigen, dass $cov(x_t, u_t) < 0$, wird zunächst $cov(x_t, v_t)$ hergeleitet.

- Unter Ausnutzung der Beziehung $x_t = x_t^* + v_t$, der Rechenregel (2.22), der Annahme (20.21) und der Beziehung

$$\sigma_v^2 = \text{var}(v_t) = \text{cov}(v_t, v_t)$$

ergibt sich

$$\begin{aligned} \text{cov}(x_t, v_t) &= \text{cov}(x_t^* + v_t, v_t) \\ &= \text{cov}(x_t^*, v_t) + \text{cov}(v_t, v_t) \\ &= 0 + \text{var}(v_t) \\ &= \sigma_v^2. \end{aligned} \tag{20.29}$$

- Rechenregeln (2.22) und (2.23) in Verbindung mit Annahme (20.23) und Ergebnis (20.29) erlauben es, die Kovarianz von x_t und u_t zu berechnen:

$$\begin{aligned} \text{cov}(x_t, u_t) &= \text{cov}(u_t, x_t) \\ &= \text{cov}(e_t - \beta v_t, x_t) \\ &= \text{cov}(e_t, x_t) - \beta \text{cov}(v_t, x_t) \\ &= \text{cov}(x_t, e_t) - \beta \text{cov}(x_t, v_t) \\ &= -\beta \sigma_v^2. \end{aligned} \tag{20.30}$$

Nicht-Konsistenz der KQ-Schätzer

- Ausgangspunkt ist wieder die Gleichung

$$\text{plim}(\hat{\beta}) = \beta + \frac{\text{plim} \left[(1/T) \sum (x_t - \bar{x}) u_t \right]}{\text{plim} \left[(1/T) S_{xx} \right]} . \quad (20.18)$$

- Der Nenner beträgt gemäß (20.14) σ_x^2 .
- Der Zähler entspricht $\text{cov}(x_t, u_t)$.
- Einsetzen des Ausdrucks (20.30) liefert:

$$\text{plim}(\hat{\beta}) = \beta - \frac{\beta \sigma_v^2}{\sigma_x^2} = \beta \left(1 - \frac{\sigma_v^2}{\sigma_x^2} \right) . \quad (20.31)$$

- **Frage:** Welcher Zusammenhang besteht zwischen dem Ausmaß der Nicht-Konsistenz und σ_v^2 ?
- **Antwort:**

Exkurs: Fehlerhafte Erfassung der endogenen Variable

- Der wahre Zusammenhang laute

$$y_t^* = \alpha + \beta x_t + e_t . \quad (20.32)$$

- Statt y_t^* sei aber y_t erfasst worden, wobei

$$v_t = y_t - y_t^*$$

den Erfassungsfehler darstellt und die Gleichungen (20.25a) bis (20.25c) erfüllt.

- Substituiert man in Modell (20.32) die Variable y_t^* durch den Term $(y_t - v_t)$, so erhält man

$$\begin{aligned} y_t &= \alpha + \beta x_t + e_t + v_t \\ &= \alpha + \beta x_t + u_t , \end{aligned}$$

wobei

$$u_t = e_t + v_t .$$

- Die neue Störgröße u_t erfüllt die B-Annahmen und $cov(x_t, u_t) = 0$.
- **Frage:** Besitzt u_t eine größere Varianz als e_t ?
Antwort:

20.3 Anwendbare Schätzverfahren

- Wenn Fall 3 vorliegt, kann die KQ-Methode nicht verwendet werden.
- Die Variable x_t^* des Modells (20.22) ist nicht beobachtbar.
- Die ersatzweise eingesetzte exogene Variable x_t ist mit der resultierenden Störgröße u_t kontemporär korreliert.
- Die KQ-Schätzer sind dann weder unverzerrt noch konsistent.
- Es muss deshalb innerhalb der exogenen Variable x_t die mit u_t korrelierte Komponente von der unkorrelierten Komponente getrennt werden.
- Dann kann in der KQ-Schätzung des Modells (20.22) an Stelle von x_t^* die unkorrelierte Komponente von x_t verwendet werden.

20.3.1 IV-Schätzung mit der ZSKQ-Methode

- Um die unkorrelierte Komponente von x_t zu ermitteln, wird eine *Instrumentvariable* z_t benötigt, für die

$$\begin{aligned} \text{cov}(z_t, x_t) &\neq 0 && \text{für alle } t = 1, 2, \dots, T, \\ \text{cov}(z_t, u_t) &= 0 && \text{für alle } t = 1, 2, \dots, T. \end{aligned} \quad (20.33)$$

- Wird eine solche Variable z_t in den Schätzprozess integriert, dann spricht man von einer IV-Schätzung.
- Die wichtigste IV-Schätzung ist die zweistufige KQ-Schätzung (ZSKQ-Schätzung).

- In der ersten Stufe der ZSKQ-Methode wird mit Hilfe von z_t jene Komponente in x_t herausgefiltert, die nicht mit u_t korreliert ist.
- In der zweiten Stufe wird die endogene Variable auf die herausgefilterte Komponente regressiert.
- Dieses Vorgehen liefert konsistente Schätzungen der Parameter α und β in Modell (20.22).

1. Stufe der ZSKQ-Schätzung

- Die Hilfsregression

$$x_t = \pi_0 + \pi_1 z_t + w_t, \quad (20.34)$$

zerlegt x_t in die zwei Komponenten $\pi_0 + \pi_1 z_t$ und die Störgröße w_t .

- **Frage:** Warum ist $\pi_0 + \pi_1 z_t$ jene Komponente in x_t , welche nicht mit u_t korreliert ist?

Antwort:

- Die Zahlenwerte von π_0 und π_1 sind unbekannt und damit auch die Zahlenwerte der Komponente $\pi_0 + \pi_1 z_t$.
- **Frage:** Wie kann dieses Problem gelöst werden?
Antwort:

- Man kann \hat{x}_t als einen Schätzwert für die unbekannt exogene Variable x_t^* im wahren Modell (20.22) auffassen.

2. Stufe der ZSKQ-Methode

- Es wird eine KQ-Schätzung des Modells

$$y_t = \alpha + \beta \hat{x}_t + e_t \quad (20.36)$$

durchgeführt.

- Diese KQ-Schätzung liefert zwar keine unverzerrten, aber doch immerhin konsistente Schätzungen der Parameter α und β .

- Grafisch bedeutet die Verwendung von \hat{x}_t , dass die Punktwolke nicht länger aus (x_t, y_t) -Beobachtungspunkten gebildet wird, sondern aus (\hat{x}_t, y_t) -Beobachtungspunkten.
- Bezogen auf den rechten Teil der Abbildung 20.3 werden alle Punkte horizontal näher an den Durchschnittswert \bar{x} herangerückt.
- Es ergibt sich eine Punktwolke, bei der die einzelnen Punkte entlang der gesamten wahren Gerade R gleichmäßig ober- und unterhalb dieser Gerade liegen.
- Die KQ-Schätzung des Modells (20.36) bedeutet grafisch, dass eine Regressionsgerade \hat{R} durch diese neue Punktwolke gelegt wird.

Verschmelzung der beiden Schätzstufen

- Formal lassen sich die beiden Schätzstufen in einer einzigen Schätzstufe bündeln (siehe Anhang des Lehrbuches).

ZSKQ-Schätzer (Einfachregression)

$$\hat{\beta}^{\text{ZSKQ}} = \frac{\sum (z_t - \bar{z})(y_t - \bar{y})}{\sum (z_t - \bar{z})(x_t - \bar{x})} \quad (20.37)$$

$$= \frac{S_{zy}}{S_{zx}} \quad (20.38)$$

und

$$\hat{\alpha}^{\text{ZSKQ}} = \bar{y} - \hat{\beta}^{\text{ZSKQ}} \bar{x}. \quad (20.39)$$

- Zum Vergleich: $\hat{\beta} = S_{xy} / S_{xx}$ und $\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}$.

Numerische Illustration 20.1

Im Versicherungs-Beispiel wurde versucht, die wahre Verkaufsbefähigung x_t^* durch die Proxyvariable x_t (Testergebnis im Assessmentcenter) zu erfassen.

Dies führte zu $u_t = e_t - \beta v_t$ und damit zu $cov(x_t, u_t) < 0$.

Als Instrumentvariable wird das Vorperiodenergebnis z_t (Daten in Tabelle 20.1) benutzt.

Frage: Sind z_t und $u_t = e_t - \beta v_t$ unkorreliert?

Antwort:

(Forts.)

Ökonometrische Software liefert $S_{zy} = 1042$ und $S_{zx} = 3637$ und damit

$$\hat{\beta}^{ZSKQ} = 0,286 .$$

Mit $\bar{y} = 22,8$ und $\bar{x} = 58,6$ ergibt sich aus (20.39):

$$\hat{\alpha}^{ZSKQ} = 22,8 - 0,286 \cdot 58,6 = 6,007 .$$

Zum Vergleich: Die KQ-Schätzung liefert $\hat{\beta} = 0,221$ und $\hat{\alpha} = 9,851$.

20.3.2 Auswahl der Instrumentvariablen

- Jede zusätzliche Instrumentvariable, die eine spürbare Korrelation mit x_t , aber keine Korrelation mit u_t aufweist, erhöht die Güte der Schätzwerte \hat{x}_t :

$$x_t = \pi_0 + \pi_1 z_{1t} + \dots + \pi_M z_{Mt} + w_t . \quad (20.40)$$

- Häufig besitzen die Instrumentvariablen eine zu geringe Korrelation mit x_t („schwache Instrumentvariablen“).
- Dies wirkt sich negativ auf die Verlässlichkeit des ZSKQ-Schätzers aus (siehe Abschnitt 20.3.5).
- Wenn sich für die Hilfsregression (20.40) in einem F -Test der Nullhypothese $H_0 : \pi_1 = \pi_2 = \dots = \pi_M = 0$ ein F -Wert von unter 10 ergibt, dann muss die Gruppe der verwendeten Instrumentvariablen $(z_{1t}, z_{2t}, \dots, z_{Mt})$ als schwach eingestuft werden.

Numerische Illustration 20.2

Im Versicherungs-Beispiel vereinfacht sich die Hilfsregression (20.40) zur Hilfsregression (20.34).

Ökonometrische Software berechnet für die Nullhypothese $H_0 : \pi_1 = 0$ einen t -Wert von 6,9927 und damit einen F -Wert von $t^2 = 48,898$.

Es wurde also eine sehr starke Instrumentvariable eingesetzt.

20.3.3 ZSKQ-Schätzung in der multiplen Regression

- Bei einer Zweifachregression, bei der allein x_{1t} mit der Störgröße u_t korreliert ist, benutzt man die Hilfsregression

$$x_{1t} = \pi_0 + \pi_1 z_{1t} + \dots + \pi_M z_{Mt} + \pi_{M+1} x_{2t} + w_t . \quad (20.41)$$

- Auf der rechten Seite erscheint x_{2t} .
- Wenn sowohl x_{1t} als auch x_{2t} mit u_t korreliert sind, muss jede exogene Variable ihre eigene Hilfsregression erhalten.
- Beide Hilfsregressionen haben dann die Form (20.40), mit x_{1t} bzw. x_{2t} an Stelle von x_t .
- In beiden Hilfsregressionen werden die gleichen Instrumentvariablen verwendet.
- Es muss mindestens eine Instrumentvariable pro korrelierter exogener Variable vorhanden sein.

20.3.4 Konsistenz der ZSKQ-Schätzer

- In Analogie zu den Annahmen bezüglich x_t muss folgendes gelten:

$$\lim_{T \rightarrow \infty} E(\bar{z}) = q_z \quad \text{und} \quad \lim_{T \rightarrow \infty} E(\sum z^2 / T) = q_{zz} \quad (20.42)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{z}) = 0 \quad \text{und} \quad \lim_{T \rightarrow \infty} \text{var}(\sum z_t^2 / T) = 0, \quad (20.43)$$

wobei q_z und q_{zz} feste Grenzwerte darstellen.

- Ferner muss für den Term S_{zx} / T ein Wahrscheinlichkeitsgrenzwert existieren:

$$\text{plim} (S_{zx} / T) = \sigma_{zx}^2. \quad (20.44)$$

- Die Schätzgleichung (20.37) kann umgeformt werden zu

$$\hat{\beta}^{ZSKQ} = \beta + \frac{\sum (z_t - \bar{z}) u_t}{S_{zx}}. \quad (20.45)$$

- Bildet man auf beiden Seiten den Wahrscheinlichkeitsgrenzwert, so erhält man:

$$\text{plim}(\hat{\beta}^{ZSKQ}) = \beta + \frac{\text{plim} [(1/T) \sum (z_t - \bar{z}) u_t]}{\text{plim} [(1/T) S_{zx}]}. \quad (20.46)$$

- Der Nenner beträgt gemäß (20.44) σ_{zx}^2 .
- Der Zähler ist ein Schätzer für $\text{cov}(z_t, u_t) = 0$. Folglich gilt:

$$\text{plim} [(1/T) \sum (z_t - \bar{z}) u_t] = 0. \quad (20.47)$$

- Folglich ergibt sich:

$$\text{plim}(\hat{\beta}^{ZSKQ}) = \beta .$$

- Es ist aber nicht möglich, (20.45) in

$$E(\hat{\beta}^{ZSKQ}) = \beta$$

zu überführen.

20.3.5 Wahrscheinlichkeitsverteilung und Varianz der ZSKQ-Schätzer

- Die Wahrscheinlichkeitsverteilung von $\hat{\beta}^{ZSKQ}$ konvergiert bei hinreichend großem Stichprobenumfang gegen $N(\beta, \text{var}(\hat{\beta}^{ZSKQ}))$.
- Die Störgrößenvarianz σ^2 kann durch

$$\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}^{ZSKQ}}{T - 2} \quad (20.48)$$

konsistent geschätzt werden, mit

$$\hat{u}_t = y_t - \hat{\alpha}^{ZSKQ} - \hat{\beta}^{ZSKQ} x_t . \quad (20.49)$$

- Ferner kann gezeigt werden, dass für $\text{var}(\hat{\beta}^{\text{ZSKQ}})$ der folgende Schätzer eingesetzt werden kann:

$$\begin{aligned} \widehat{\text{var}}(\hat{\beta}^{\text{ZSKQ}}) &= \frac{\hat{\sigma}^2 \sum (z_t - \bar{z})^2}{[\sum (z_t - \bar{z})(x_t - \bar{x})]^2} \\ &= \frac{\hat{\sigma}^2 S_{zz}}{S_{zx}^2} \end{aligned} \quad (20.50)$$

$$\begin{aligned} &= \frac{\hat{\sigma}^2 S_{xx} S_{zz}}{S_{xx} S_{zx}^2} \\ &= \frac{\hat{\sigma}^2}{S_{xx}} \cdot \frac{1}{R_{z \cdot x}^2} \\ &= \widehat{\text{var}}(\hat{\beta}) \cdot \frac{1}{R_{z \cdot x}^2} \end{aligned} \quad (20.51)$$

- Dabei ist $R_{z \cdot x}^2 = S_{zx}^2 / (S_{xx} S_{zz})$ das Bestimmtheitsmaß zur Hilfsregression (20.34).
- **Frage:** Warum offenbart Gleichung (20.51), dass sich schwache Instrumentvariablen negativ auf die Qualität der IV-Schätzung auswirken?
Antwort:

Numerische Illustration 20.3

Es wurde bereits berechnet, dass $S_{zx} = 3637$ und $\hat{\beta}^{ZSKQ} = 0,286$.

Aus Gleichung (20.49) ergeben sich die entsprechenden Residuen der ZSKQ-Schätzung.

Aus diesen erhält man:

$$S_{\hat{u}\hat{u}}^{ZSKQ} = \sum \hat{u}_t^2 = 670,240 .$$

Aus Formel (20.48) ergibt sich

$$\hat{\sigma}^2 = \frac{670,240}{28} = 23,937 . \quad (20.52)$$

Ferner lässt sich berechnen, dass $S_{zz} = 1375$.

(Forts.)

Einsetzen der Werte in Gleichung (20.50) liefert

$$\widehat{\text{var}}(\widehat{\beta}^{\text{ZSKQ}}) = \frac{23,937 \cdot 1375}{3637^2} = 0,00249 \quad (20.53)$$

und damit $\widehat{\text{sd}}(\widehat{\beta}^{\text{ZSKQ}}) = 0,050$.

Dividiert man den Schätzwert $\widehat{\beta}^{\text{ZSKQ}} = 0,286$ durch $\widehat{\text{sd}}(\widehat{\beta}^{\text{ZSKQ}}) = 0,050$, so erhält man einen Wert von 5,72. Dies ist ein approximativer t -Wert der Nullhypothese $H_0 : \beta = 0$.

Frage: Welche Testentscheidung ergibt sich?

Antwort:

20.3.6 Fazit der ZSKQ-Schätzung

- Besteht zwischen x_t und u_t kontemporäre Korrelation (Fall 3), dann ist die KQ-Schätzung unzulässig.
- Die ZSKQ-Schätzung bietet einen Ausweg.
- Die Instrumentvariable z_t muss die folgenden Eigenschaften besitzen:

$$\text{cov}(z_t, u_t) = 0 \quad \text{für alle } t = 1, 2, \dots, T, \quad (20.33)$$

$$\text{plim}(S_{zx} / T) = \sigma_{zx}^2, \quad (20.44)$$

$$\lim_{T \rightarrow \infty} E(\bar{z}) = q_z, \quad \lim_{T \rightarrow \infty} E(\sum z^2 / T) = q_{zz}, \quad (20.42)$$

$$\lim_{T \rightarrow \infty} \text{var}(\bar{z}) = 0, \quad \lim_{T \rightarrow \infty} \text{var}(\sum z_t^2 / T) = 0, \quad (20.43)$$

- Der ZSKQ-Schätzer $\hat{\beta}^{ZSKQ} = S_{zy} / S_{zx}$ ist dann konsistent.
- Seine Varianz kann mit Hilfe der Formel

$$\widehat{\text{var}}(\hat{\beta}^{ZSKQ}) = \hat{\sigma}^2 S_{zz} / S_{zx}^2 \quad (20.50)$$

geschätzt werden, wobei

$$\hat{\sigma}^2 = S_{\hat{u}\hat{u}}^{ZSKQ} / (T - 2) .$$

20.4 Diagnose

- Bevor wir eine KQ-Schätzung durchführen, sollten wir überprüfen, ob $cov(x_t, u_t) = 0$.
- Dies geschieht mit dem *Spezifikationstest von Hausman*.

20.4.1 Vorüberlegungen

- Wenn $cov(x_t, u_t) = 0$, wären sowohl die KQ- also auch die ZSKQ-Methode zulässig.
- $\hat{\beta}$ und $\hat{\beta}^{ZSKQ}$ weichen dann kaum voneinander ab.
- Wenn $cov(x_t, u_t) \neq 0$, dann weichen $\hat{\beta}$ und $\hat{\beta}^{ZSKQ}$ voneinander ab.

20.4.2 Spezifikationstest von Hausman

- Es wird die Nullhypothese

$$H_0 : \text{plim} \left[(1/T) \sum (x_t - \bar{x}) u_t \right] = 0$$

gegen die Alternativhypothese

$$H_1 : \text{plim} \left[(1/T) \sum (x_t - \bar{x}) u_t \right] \neq 0$$

getestet.

Spezifikationstest von Hausman

Bei Gültigkeit der Nullhypothese gilt

$$m = \frac{(\widehat{\beta}^{\text{ZSKQ}} - \widehat{\beta})^2}{\widehat{\text{var}}(\widehat{\beta}^{\text{ZSKQ}}) - \widehat{\text{var}}(\widehat{\beta})} \sim \chi^2_{(k)} \quad (20.54)$$

- Bei der Berechnung von $\widehat{\text{var}}(\widehat{\beta}^{\text{ZSKQ}})$ und $\widehat{\text{var}}(\widehat{\beta})$ werden entweder *einheitlich* die Residuen der KQ-Schätzung zugrunde gelegt oder *einheitlich* die Residuen der ZSKQ-Schätzung.

Numerische Illustration 20.4

Der auf Basis einer KQ-Schätzung berechnete Wert für $\hat{\sigma}^2$ beträgt 21,615 und $S_{xx} = 15123$.

Damit ergibt sich

$$\widehat{\text{var}}(\hat{\beta}) = \frac{\hat{\sigma}^2}{S_{xx}} = \frac{21,615}{15123} = 0,00143 .$$

Unter Verwendung des KQ-gestützten Schätzers für σ^2 und den Werten $S_{zz} = 1375$ sowie $S_{zx} = 3637$ ergibt sich:

$$\widehat{\text{var}}(\hat{\beta}^{\text{ZSKQ}}) = \frac{\hat{\sigma}^2 S_{zz}}{S_{zx}^2} = \frac{21,615 \cdot 1375}{3637^2} = 0,00225 .$$

(Forts.)

Mit $\hat{\beta} = 0,221$ und $\hat{\beta}^{ZSKQ} = 0,286$ erhält man

$$m = \frac{(0,286 - 0,221)^2}{0,00225 - 0,00143} = 5,25237 .$$

Bei einem Signifikanzniveau von 5% liegt der kritische Wert der $\chi^2_{(1)}$ -Verteilung bei 3,84146.

Kapitel 21

Annahme C2: Multikollinearität

Annahme C2 (Freiheit von Perfekter Multikollinearität)

Es existieren keine Parameterwerte $\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_K$ (wobei mindestens ein $\gamma_k \neq 0$), so dass zwischen den exogenen Variablen $x_{1t}, x_{2t}, \dots, x_{Kt}$ für alle $t = 1, 2, \dots, T$ die folgende lineare Beziehung gilt:

$$\gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \dots + \gamma_K x_{Kt} = 0. \quad (21.1)$$

Beispiel zu Kapitel 21

Die Preise von Laserdruckern (Februar 1992 bis August 2001) sollen durch deren unterschiedliche Qualität erklärt werden. Die Daten für diese sogenannte „hedonische Regressionsanalyse“ sind in Tabelle 21.1 wiedergegeben. Sie deuten auf eine positive Korrelation zwischen x_{1t} und x_{4t} .

(Forts.)

Tabelle 21.1: Preis y_t (in Euro), Druckgeschwindigkeit x_{1t} (in Seiten pro Minute), Druckergröße x_{2t} (in dm^3), Ersatzteilkosten x_{3t} (in Cent/pro Seite) und Zeitdifferenz zum ersten betrachteten Modell x_{4t} (in Monaten) von 44 postscriptfähigen Laserdruckern.

t	y_t	x_{1t}	x_{2t}	x_{3t}	x_{4t}	t	y_t	x_{1t}	x_{2t}	x_{3t}	x_{4t}
1	2652	4	28,6	3,45	0	23	1110	6	50,0	2,15	38
2	1504	4	83,2	2,75	0	24	1270	4	18,4	2,35	38
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
22	1745	12	39,1	1,95	38	44	690	10	46,1	2,10	114

21.1 Konsequenzen der Annahmeverletzung

21.1.1 Grafische Veranschaulichung

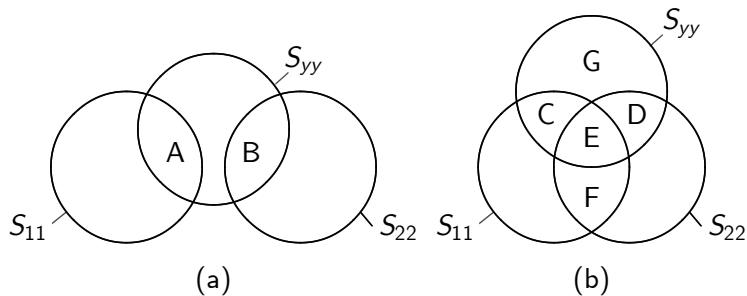


Abbildung 21.1: Zweifachregression (a) ohne und (b) mit Multikollinearität

- Die Fläche $C + E$ ist etwas größer als die Fläche A .
- **Frage:** Warum wird β_1 im Fall (a) dennoch mit größerer Präzision als in Fall (b) geschätzt.

Antwort:

- Analoges gilt für β_2 .
- Fläche E ist für Schätzung des Gesamteinflusses der beiden exogenen Variablen relevant.

21.1.2 Konsequenzen perfekter Multikollinearität für Punkt-, Intervallschätzung und Hypothesentests

- Der KQ-Schätzer einer Zweifachregression lautet:

$$\hat{\beta}_1 = \frac{S_{22}S_{1y} - S_{12}S_{2y}}{S_{11}S_{22} - S_{12}^2} . \quad (21.2)$$

- Die beiden exogenen Variablen seien perfekt multikollinear: $x_1 = \gamma x_2$.

- Für den Nenner folgt daraus:

$$\begin{aligned} S_{11}S_{22} - S_{12}^2 &= \sum (\gamma x_2 - \gamma \bar{x}_2)^2 \sum (x_2 - \bar{x}_2)^2 \\ &\quad - \left[\sum (\gamma x_2 - \gamma \bar{x}_2) (x_2 - \bar{x}_2) \right]^2 \\ &= \sum \gamma^2 (x_2 - \bar{x}_2)^2 \sum (x_2 - \bar{x}_2)^2 \\ &\quad - \left[\gamma \sum (x_2 - \bar{x}_2) (x_2 - \bar{x}_2) \right]^2 \\ &= \gamma^2 \left[\sum (x_2 - \bar{x}_2)^2 \right]^2 - \gamma^2 \left[\sum (x_2 - \bar{x}_2)^2 \right]^2 \\ &= 0. \end{aligned}$$

- **Frage:** Welche Konsequenzen hat dieses Resultat für die KQ-Schätzung?

Antwort:

21.1.3 Konsequenzen imperfekter Multikollinearität für Punkt-, Intervallschätzung und Hypothesentests

Konsequenzen für die Punktschätzung

- Die Varianz von β_1 ist in der Zweifachregression durch

$$\text{var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{11}(1 - R_{1.2}^2)} = \frac{\sigma^2}{S_{11}} \cdot \frac{1}{(1 - R_{1.2}^2)} \quad (21.3)$$

definiert, wobei

$$R_{1.2}^2 = S_{12}^2 / (S_{11}S_{22}) .$$

- Der Quotient

$$\frac{1}{(1 - R_{1.2}^2)}$$

wird als *Variance Inflation Factor (VIF)* bezeichnet.

- **Frage:** In welcher Hinsicht wirkt sich hohe Multikollinearität negativ auf die Punktschätzung von β_1 aus?

Antwort:

- **Frage:** Warum führt hohe Multikollinearität aber nicht zwangsläufig zu diesem Problem?

Antwort:

Konsequenzen für Intervallschätzung und Hypothesentest

- Multikollinearität bedeutet, dass der $sd(\hat{\beta}_k)$ -Wert tendenziell sehr groß ist.
- Folglich sind die geschätzten Intervalle sehr groß (unpräzise) und Hypothesentests sind unscharf.
- Wenn beispielsweise $H_0 : \beta_k = 0$ getestet werden soll, dann bedeutet ein großer $\hat{\beta}_k$ -Wert noch nicht, dass auch

$$t = \frac{\hat{\beta}_k - 0}{\widehat{sd}(\hat{\beta}_k)}$$

größer als $t_{\alpha/2}$ ist.

Konsequenzen der Multikollinearität

Eine hohe Korrelation der exogenen Variablen stellt keine Verletzung der Annahme C2 dar. Lediglich die Schätzung der individuellen Einflüsse der exogenen Variablen wird erschwert.

21.2 Diagnose

- Es stellen sich zwei *separate* Fragen:
 - ① In welchem Ausmaß sind die benutzten Daten multikollinear?
 - ② Falls die KQ-Schätzer insignifikant sind, in welchem Ausmaß kann dies auf Multikollinearität und in welchem Ausmaß auf andere Ursachen (z.B. Fehlspezifikation des Modells) zurückgeführt werden?

21.2.1 Diagnose von Multikollinearität

- Das folgende Beispiel zeigt: Nur bei Zweifachregressionen ist

$$R_{1.2}^2 = S_{12}^2 / (S_{11}S_{22})$$

ein geeigneter Indikator für das Ausmaß der Multikollinearität.

Numerische Illustration 21.1

Frage: Warum sind die Daten in Tabelle 21.2 perfekt multikollinear?

Antwort:

(Forts.)

Tabelle 21.2: Ein Beispiel für kleine Werte der Korrelationsmaße trotz perfekter Multikollinearität.

t	x_{1t}	x_{2t}	x_{3t}	Korrelationsmaße
1	-1	9	8	$R_{1.2}^2 = 0,261$
2	5	-3	2	$R_{1.3}^2 = 0,290$
3	9	3	12	$R_{2.3}^2 = 0,202$
4	4	-4	0	
5	-6	5	-1	

Frage: Für welche Parameterkonstellation $(\gamma_0, \gamma_1, \gamma_2, \gamma_3)$ ist also für alle $T = 5$ Beobachtungen die folgende Beziehung erfüllt: $\gamma_0 + \gamma_1 x_{1t} + \gamma_2 x_{2t} + \gamma_3 x_{3t} = 0$?

Antwort:

- Bei einer Mehrfachregression existiert keine eindeutige Maßzahl für Multikollinearität.
- Im Falle einer Vierfachregression,

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t, \quad (21.4)$$

kann die *Hilfsregression*

$$x_{1t} = \gamma + \mu_2 x_{2t} + \mu_3 x_{3t} + \mu_4 x_{4t} + e_t$$

durchgeführt und das entsprechende Bestimmtheitsmaß $R_{1.234}^2$ berechnet werden.

- Ganz analog sollten $R_{2.134}^2$, $R_{3.124}^2$ und $R_{4.123}^2$ ermittelt werden.
- Ergänzend sollten auch $R_{1.23}^2$, $R_{1.24}^2$... bis $R_{4.23}^2$ bzw. $R_{1.2}^2$, ... bis $R_{3.4}^2$ berechnet werden.

Numerische Illustration 21.2

Tabelle 21.3: Bestimmtheitsmaße für alle möglichen Hilfsregressionen des Laserdrucker-Beispiels.

\mathbf{x}_{1t} (Geschw.)	\mathbf{x}_{2t} (Größe)	\mathbf{x}_{3t} (Ersatzteil)	\mathbf{x}_{4t} (Zeitdiff.)
$R_{1.234}^2 = 0,691$	$R_{2.134}^2 = 0,390$	$R_{3.124}^2 = 0,319$	$R_{4.123}^2 = 0,580$
$R_{1.23}^2 = 0,432$	$R_{2.13}^2 = 0,235$	$R_{3.12}^2 = 0,313$	$R_{4.12}^2 = 0,576$
$R_{1.24}^2 = 0,649$	$R_{2.14}^2 = 0,364$	$R_{3.14}^2 = 0,290$	$R_{4.13}^2 = 0,473$
$R_{1.34}^2 = 0,496$	$R_{2.34}^2 = 0,003$	$R_{3.24}^2 = 0,224$	$R_{4.23}^2 = 0,226$
$R_{1.2}^2 = 0,172(+)$	$R_{2.1}^2 = R_{1.2}^2$	$R_{3.1}^2 = R_{1.3}^2$	$R_{4.1}^2 = R_{1.4}^2$
$R_{1.3}^2 = 0,258(-)$	$R_{2.3}^2 = 0,000(+)$	$R_{3.2}^2 = R_{2.3}^2$	$R_{4.2}^2 = R_{2.4}^2$
$R_{1.4}^2 = 0,449(+)$	$R_{2.4}^2 = 0,002(-)$	$R_{3.4}^2 = 0,224(-)$	$R_{4.3}^2 = R_{3.4}^2$

(Forts.)

Frage: Welche Korrelationsmuster lassen sich erkennen?

Antwort:

Frage: Die Schätzung welcher Steigungsparameter könnte möglicherweise unter Multikollinearität leiden?

Antwort:

21.2.2 Hohe Schätzvarianz der Punktschätzer: Multikollinearität oder Fehlspezifikation?

Vergleich der individuellen und gemeinsamen Signifikanz der Parameter

- Weisen individuelle Schätzer hohe Schätzvarianzen auf (und damit insignifikante t -Werte) und ist gleichzeitig R^2 groß (bzw. ein F -Test hochsignifikant), dann signalisiert dies Schätzprobleme, die auf Multikollinearität zurückzuführen sind.
- Die Informationen in den Daten sind nicht ausreichend, um die *individuellen* Parameter genau zu schätzen.
- Das ökonometrische Modell ist aber in Ordnung.

Numerische Illustration 21.3

Tabelle 21.4: Schätzergebnisse des Laserdrucker-Beispiels.

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	2021,058	310,146	6,516	<0,001
Druckgeschwind.	108,086	21,415	5,047	<0,001
Druckergröße	-0,512	2,611	-0,196	0,846
Ersatzteilkosten	-116,663	90,309	-1,292	0,204
Zeitdifferenz	-21,792	2,632	-8,280	<0,001

(Forts.)

Frage: Wie ist der Schätzwert zu $\hat{\beta}_4$ zu interpretieren?

Antwort:

Frage: Wurde die Schätzgenauigkeit der individuellen Parameter durch die Multikollinearität erheblich eingeschränkt?

Antwort:

(Forts.)

Der F -Test zu $H_0 : \beta_2 = \beta_3 = 0$ liefert $F = 0,944$.

Der kritische Wert ($v_1 = 2, v_2 = 41$) liegt bei etwa $F_{0,05} = 3,23$.

Frage: Zu welcher Schlussfolgerung gelangt man bezüglich der Relevanz von x_{2t} und x_{3t} ?

Antwort:

Der F -Test zu $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ liefert $F = 21,5$.

Der kritische Wert ($v_1 = 4, v_2 = 39$) liegt bei etwa

$F_{0,05} = 2,612$.

Das Bestimmtheitsmaß beträgt $R^2 = 68,8\%$.

Ausmaß der Multikollinearität

Wenn in der KQ-Schätzung des ökonometrischen Modells niedrige t -Werte mit hohen F -Werten einhergehen, signalisiert dies, dass die Variablen des Modells zwar multikollinear aber relevant sind.

Um das Ausmaß der Multikollinearität zu erkennen, können Hilfsregressionen durchgeführt werden, in denen jeweils eine exogene Variable x_{kt} auf eine oder mehrere andere exogene Variablen regressiert wird.

21.3 Angemessener Umgang mit Multikollinearität

21.3.1 Verfahren zur Eindämmung des Multikollinearitätsproblems

Auslassen von Variablen

- Geringe t -Werte müssen nicht durch Irrelevanz, sondern können auch durch Multikollinearität ausgelöst werden.
- **Frage:** Welche Konsequenzen, würden sich ergeben, wenn relevante Variablen aufgrund kleiner t -Werte aus dem Modell entfernt würden?

Antwort:

Numerische Illustration 21.4

Tabelle 21.5: Schätzergebnisse des Laserdrucker-Beispiels bei Auslassen der Variablen x_{2t} und x_{3t} .

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	1631,579	125,066	13,046	<0,001
Druckgeschwind.	112,105	16,000	7,007	<0,001
Zeitdifferenz	-20,895	2,295	-9,106	<0,001

Frage: Warum ist im Laserdrucker-Beispiel die Gefahr gering, dass durch das Auslassen der Variablen x_{2t} und x_{3t} eine Verzerrung in der KQ-Schätzung auftritt?

Antwort:

Differenzenbildung

- Das Ausgangsmodell lautete

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t . \quad (21.4)$$

- Für die Vorperiode lautet dieses Modell:

$$y_{t-1} = \alpha + \beta_1 x_{1t-1} + \beta_2 x_{2t-1} + \beta_3 x_{3t-1} + \beta_4 x_{4t-1} + u_{t-1} . \quad (21.5)$$

- Subtrahiert man Gleichung (21.5) von Gleichung (21.4), so erhält man

$$y_t - y_{t-1} = \beta_1 (x_{1t} - x_{1t-1}) + \beta_2 (x_{2t} - x_{2t-1}) \\ + \beta_3 (x_{3t} - x_{3t-1}) + \beta_4 (x_{4t} - x_{4t-1}) + u_t - u_{t-1} .$$

- Definiert man

$$\Delta y_t = y_t - y_{t-1}$$

$$\Delta x_{kt} = x_{kt} - x_{kt-1} \quad k = 1, 2, 3, 4$$

$$\Delta u_t = u_t - u_{t-1},$$

so ergibt sich die *erste Differenzenform* des Ausgangsmodells

$$\Delta y_t = \beta_1 \Delta x_{1t} + \beta_2 \Delta x_{2t} + \beta_3 \Delta x_{3t} + \beta_4 \Delta x_{4t} + \Delta u_t. \quad (21.6)$$

- Die exogenen Variablen dieses Modells weisen normalerweise eine deutlich geringere Multikollinearität auf.
- Beispielsweise beträgt im Laserdrucker-Beispiel die Korrelation zwischen den Variablen Δx_{1t} und Δx_{4t} lediglich 0,03, verglichen mit $R_{1.4}^2 = 0,449$.

- **Frage:** Welches neue Problem hat man sich aber durch die Differenzenbildung eingehandelt?

Antwort:

21.3.2 Verwendung zusätzlicher Informationen

Zusätzliche Beobachtungen

- **Frage:** Warum ist dieser Weg nicht immer möglich bzw. erfolgreich?

Antwort:

Zusätzliche externe Informationen: RKQ-Schätzung

- Es sei aus anderen Studien bekannt, dass

$$\beta_2 = 0 \quad (21.7)$$

$$\beta_3 = -2 \cdot \beta_1 . \quad (21.8)$$

- Setzt man dies in das Ausgangsmodell (21.4) ein, so erhält man

$$\begin{aligned} y_t &= \alpha + \beta_1 x_{1t} + 0 \cdot x_{2t} - 2 \cdot \beta_1 \cdot x_{3t} + \beta_4 \cdot x_{4t} + u_t \\ &= \alpha + \beta_1 x_{1t}^* + \beta_4 x_{4t} + u_t , \end{aligned} \quad (21.9)$$

wobei $x_{1t}^* = x_{1t} - 2 \cdot x_{3t}$.

- Die KQ-Schätzung dieses Modell wird als restringierte KQ-Schätzung (RKQ-Schätzung) bezeichnet.

Numerische Illustration 21.5

Im Laserdrucker-Beispiel liefert die RKQ-Schätzung die folgenden Werte:

$$\begin{aligned}\hat{\beta}_1^{RKQ} &= 97,024 & \text{und} & & \hat{\beta}_4^{RKQ} &= -21,546 \\ \widehat{\text{var}}(\hat{\beta}_1^{RKQ}) &= 182,709 & \text{und} & & \widehat{\text{var}}(\hat{\beta}_4^{RKQ}) &= 5,35 \quad (21.10)\end{aligned}$$

Damit ergibt sich für β_3 :

$$\begin{aligned}\hat{\beta}_3^{RKQ} &= -2 \cdot \hat{\beta}_1^{RKQ} = -194,049 \\ \widehat{\text{var}}(\hat{\beta}_3^{RKQ}) &= \widehat{\text{var}}(-2 \cdot \hat{\beta}_1^{RKQ}) \\ &= (-2)^2 \cdot \widehat{\text{var}}(\hat{\beta}_1^{RKQ}) = 730,836 .\end{aligned}$$

(Forts.)

Man erhält die folgenden Standardabweichungen:

$$\widehat{sd}(\widehat{\beta}_1^{RKQ}) = 13,517$$

$$\widehat{sd}(\widehat{\beta}_3^{RKQ}) = 27,034$$

$$\widehat{sd}(\widehat{\beta}_4^{RKQ}) = 2,315 .$$

- Ein Vergleich mit Tabelle 21.4 zeigt, dass die Schätzvarianz deutlich verringert wurde.
- Das Vorgehen ist allerdings nur dann zulässig, wenn (21.7) und (21.8) korrekt sind.

(Forts.)

Um Anhaltspunkte über das Ausmaß einer möglichen Verzerrung zu erhalten, sollte man in jedem Fall einen F -Test mit (21.7) und (21.8) als Nullhypothese durchführen, also

$$H_0 : \beta_2 = 0 \text{ und gleichzeitig } \beta_3 = -2 \cdot \beta_1 .$$

Es ergibt sich ein F -Wert von 0,405 ($L=2$ und $T-K-1=39$). Der kritische Wert beträgt $F_{0,05} = 3,238$.

Keine „Lösung“ des Multikollinearitätsproblems möglich

Auch bei starker Multikollinearität bleibt die KQ-Schätzung effizient (BUE-Eigenschaft). Eine Erhöhung der Schätzgenauigkeit der individuellen Parameter kann nur über die Einbeziehung korrekter zusätzlicher Informationen gelingen.

Kapitel 22

Dynamische Modelle

Beispiel zu Kapitel 22

Es wird die Anpassungszeit untersucht, mit der ein Software-Unternehmen seine Beschäftigtenzahl den Schwankungen in der Auftragslage anpasst. Die zur Verfügung stehenden Daten sind in Tabelle 22.1 wiedergegeben.

Tabelle 22.1: Monatliche Beschäftigtenzahl y_t und Zahl neuer Aufträge x_t während der vergangenen 36 Monate.

t	y_t	x_t	t	y_t	x_t	t	y_t	x_t
1	2163	1340	13	2159	1296	25	2646	2115
2	2050	1411	14	1942	1201	26	2306	2150
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
12	1665	1409	24	2428	2114	36	3521	2053

22.1 Stochastische Prozesse und Stationarität

22.1.1 Stochastische Prozesse

- Die Werte der Variablen $x_1, x_2, x_3, \dots, x_T$ sind Zufallsvariablen mit jeweils *eigener* Wahrscheinlichkeitsverteilung, Erwartungswert $E(x_t)$ und Varianz $var(x_t)$.
- $cov(x_t, x_{t+\tau})$ ist die Kovarianz zweier solcher Variablen, wobei τ für den zeitlichen Abstand zwischen diesen beiden Zufallsvariablen steht.
- Die Sequenz der Zufallsvariablen $x_1, x_2, x_3, \dots, x_T$ wird als *stochastischer Prozess* bezeichnet, die Sequenz aus *beobachteten Werten* als *Zeitreihe*.

22.1.2 Stationarität von stochastischen Prozessen

- Ein bestimmter stochastischer Prozess lässt sich durch die statistischen Eigenschaften (Erwartungswert, Varianz, Kovarianz) seiner Zufallsvariablen charakterisieren.
- Abbildung 22.1 zeigt einige Beispiele.

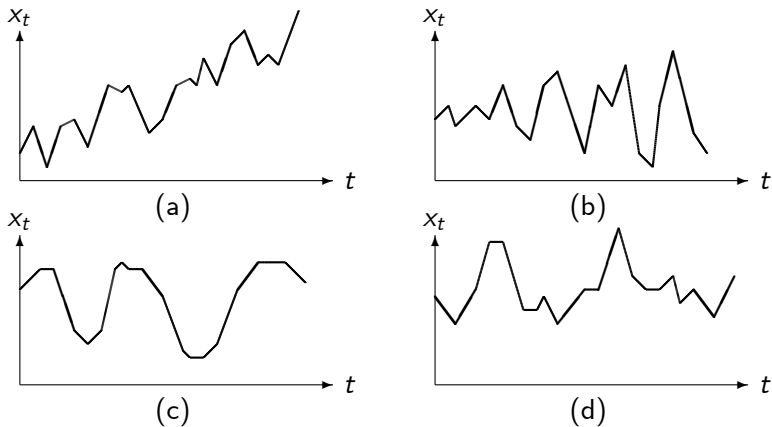


Abbildung 22.1: Zeitreihen verschiedener stochastischer Prozesse.

- Formal ist ein stationärer Prozess durch drei Eigenschaften definiert:
 - ① $E(x_t) = \mu$ für alle $t = 1, 2, \dots, T$;
 - ② $var(x_t) = \sigma_x^2$ für alle $t = 1, 2, \dots, T$;
 - ③ $cov(x_t, x_{t+\tau}) = \gamma_\tau$ für alle $t = 1, 2, \dots, T - \tau$
und alle $\tau = 1, 2, \dots, T - 1$.
- Dabei bezeichnen μ , σ_x^2 und $\gamma_1, \gamma_2, \dots$ jeweils Konstanten.
- Die dritte Eigenschaft fordert *nicht*, dass $cov(x_1, x_2) = cov(x_1, x_3)$.

- Der $AR(1)$ -Prozess mit $|\rho| < 1$ ist ein Beispiel für einen stationären stochastischen Prozess.
- Sobald eine der drei Stationaritäts-Bedingungen verletzt ist, bezeichnet man den stochastischen Prozess als *nicht-stationär*.
- Bei solchen Prozessen konvergiert der Term $(1/T)S_{xx}$ mit steigendem Stichprobenumfang (= Ausweitung des Zeithorizontes) normalerweise nicht gegen eine feste Schranke σ_x^2 .
- Damit sind weder Annahme $c1^{**}$ noch Annahme $c1^{***}$ erfüllt.

22.1.3 I(1)-Prozesse

- Aus den T Zufallsvariablen x_1, x_2, \dots, x_T lässt sich der stochastische Prozess $\Delta x_2, \Delta x_3, \dots, \Delta x_T$ definieren, wobei $\Delta x_2 = x_2 - x_1$, $\Delta x_3 = x_3 - x_2$ und $\Delta x_T = x_T - x_{T-1}$.
- Oftmals stellt die Sequenz x_1, x_2, \dots, x_T einen nicht-stationären Prozess dar, während die Sequenz $\Delta x_2, \Delta x_3, \dots, \Delta x_T$ die Stationaritätsbedingungen erfüllt.
- Man bezeichnet dann den nicht-stationären Prozess x_1, x_2, \dots, x_T als integriert vom Grade 1, oder kurz als I(1)-Prozess.
- Die Begriffe „stochastischer Prozess“ und „Zeitreihe“ werden oftmals synonym verwendet.
- Wir sprechen deshalb vereinfachend von einer stationären bzw. einer nicht-stationären Zeitreihe x_t .

22.2 Interpretation dynamischer Modelle

- Für das Beschäftigungs-Beispiel könnte folgende Modellform gewählt werden:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_K x_{t-K} + v_t, \quad (22.1)$$

wobei v_t eine Störgröße ist, welche sämtliche B-Annahmen erfüllt, und die $K + 1$ Parameter β_k Steigungsparameter darstellen.

- Die Variablen x_{t-1}, x_{t-2}, \dots und auch die Variable y_{t-1} werden häufig als *Lag-Variablen* bezeichnet.

- Die Werte der $K+1$ exogenen Variablen beruhen alle auf ein und derselben Variable (*Zahl der Aufträge*).
- Modell (22.1) wird als *bivariates dynamisches Modell* oder auch als *distributed lag model* bezeichnet.
- **Frage:** Wie viele Lag-Variablen enthält Modell (22.1)?
Antwort:

22.2.1 Interpretation einzelner Parameter

- Der Parameter α ist der Niveauparameter des linearen Zusammenhangs.
- Die Steigungsparameter β_k ergeben sich wie immer durch Differenzieren: $\partial y_t / \partial x_{t-k} = \beta_k$.
- **Frage:** Welche Interpretation besitzt β_2 in Modell (22.1)?
Antwort:

- Wegen (22.1) gilt auch:

$$y_{t+2} = \alpha + \beta_0 x_{t+2} + \beta_1 x_{t+1} + \beta_2 x_t + \dots + \beta_K x_{t-K+2} + v_{t+2}.$$

- **Frage:** Welche weitere Interpretation besitzt β_2 demnach?

Antwort:

22.2.2 Kurzfristiger und langfristiger Multiplikator

- Es sei angenommen, dass $v_t = v_{t-1} = \dots = 0$, und dass bis einschließlich Periode $t - 1$ die Zahl der Aufträge konstant \tilde{x} betragen habe.
- Damit ergibt sich aus Gleichung (22.1) für Periode $t - 1$ die folgende Beschäftigung:

$$\begin{aligned}y_{t-1} &= \alpha + \beta_0 \tilde{x} + \beta_1 \tilde{x} + \dots + \beta_K \tilde{x} \\ &= \alpha + \tilde{x} \sum_{k=0}^K \beta_k .\end{aligned}\tag{22.2}$$

Kurzfristiger Multiplikator

- In Periode t habe sich die Auftragslage gegenüber der Vergangenheit um eine Einheit auf $\tilde{x}+1$ verbessert.
- Die Beschäftigung in Periode t ergibt sich gemäß Gleichung (22.1) aus

$$\begin{aligned}y_t &= \alpha + \beta_0 (\tilde{x} + 1) + \beta_1 \tilde{x} + \dots + \beta_K \tilde{x} \\ &= \beta_0 + \alpha + \tilde{x} \sum_{k=0}^K \beta_k \\ &= \beta_0 + y_{t-1} .\end{aligned}$$

- Der Parameter β_0 misst demnach den *kurzfristigen* Effekt.
- Bei einer Auftragserhöhung um $\Delta \tilde{x}$ hätte der kurzfristige Effekt $\Delta \tilde{x} \cdot \beta_0$ betragen.
- Man nennt β_0 deshalb den *kurzfristigen Multiplikator*.

Langfristiger Multiplikator

- Wenn das neue Auftragsniveau $\tilde{x}+1$ auch in Periode $t+1$ gehalten werden kann, dann beträgt die Beschäftigung in Periode $t+1$:

$$y_{t+1} = \alpha + (\beta_0 + \beta_1) (\tilde{x} + 1) + \tilde{x} \sum_{k=2}^K \beta_k$$

$$= \beta_0 + \beta_1 + \alpha + \tilde{x} \sum_{k=0}^K \beta_k$$

$$[\text{aus (22.2)}] \quad = \beta_0 + \beta_1 + y_{t-1}$$

$$= \beta_1 + y_t . \quad (22.3)$$

- Wenn das neue Auftragsniveau $\tilde{x}+1$ dauerhaft gehalten werden kann, kommt es zu weiteren Beschäftigungsausweitungen (siehe Abbildung 22.2).

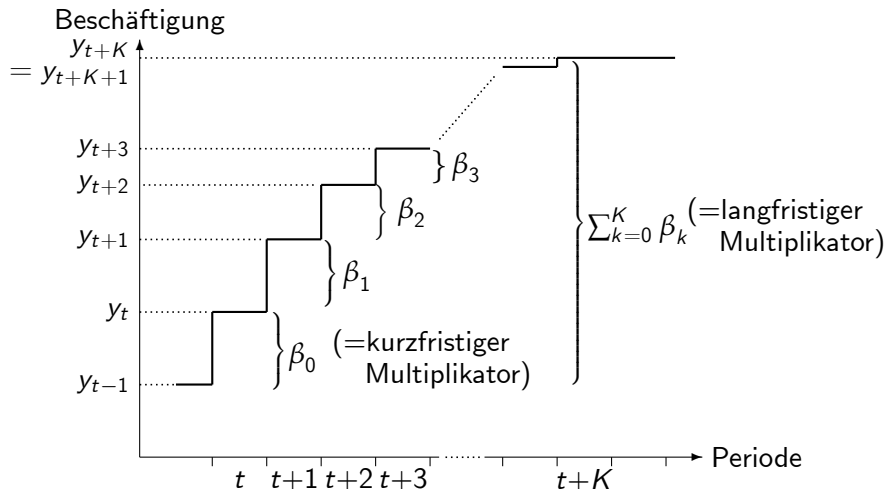


Abbildung 22.2: Der Anpassungsprozess der Beschäftigung nach einer dauerhaften Erhöhung der Zahl der Aufträge von \tilde{x} auf $\tilde{x} + 1$.

- Die Beschäftigung in Periode $t+K$ beträgt (wobei $x_t = \dots = x_{t+K} = \tilde{x} + 1$):

$$y_{t+K} = \alpha + (\tilde{x} + 1) \sum_{k=0}^K \beta_k \quad (22.4)$$

$$= \alpha + \tilde{x} \sum_{k=0}^K \beta_k + \sum_{k=0}^K \beta_k \quad (22.5)$$

$$\text{[aus (22.2)]} \quad = y_{t-1} + \sum_{k=0}^K \beta_k \quad (22.6)$$

$$= \beta_K + y_{t-1} + \sum_{k=0}^{K-1} \beta_k$$

$$= \beta_K + y_{t+K-1} .$$

- Gegenüber der Vorperiode $t+K-1$ erhöht sich die Beschäftigung in Periode $t+K$ um β_K .
- Wenn auch in der Folgeperiode $t+K+1$ die Auftragslage $\tilde{x}+1$ beträgt, dann ergibt sich für die Beschäftigung:

$$y_{t+K+1} = \alpha + (\tilde{x} + 1) \sum_{k=0}^K \beta_k$$

$$[\text{aus (22.4)}] \quad = y_{t+K} .$$

- Damit ist die Beschäftigungsausweitung zum Erliegen gekommen.

- Aus Gleichung (22.6) ist ersichtlich, dass gegenüber Periode $t-1$ die dauerhafte Niveausteigerung auf $\tilde{x} + 1$ einen Beschäftigungseffekt von $\sum_{k=0}^K \beta_k$ bewirkt.
- Bei einer Auftragserhöhung um $\Delta\tilde{x}$, hätte der langfristige Effekt $\Delta\tilde{x} \cdot \sum_{k=0}^K \beta_k$ betragen.
- Deshalb wird die Summe $\sum_{k=0}^K \beta_k$ als *langfristiger Multiplikator* bezeichnet.

22.2.3 Median-Lag

- Im dynamischen Modell (22.1) dauert es $K+1$ Perioden, bis sich die endogene Variable vollständig an ein neues um $\Delta\tilde{x}$ verändertes Niveau der exogenen Variable anpasst.
- Der Gesamteffekt der Anpassung beträgt $\Delta\tilde{x} \cdot \sum_{k=0}^K \beta_k$.
- Der Anteil an dieser Anpassung, der in der ersten Periode mit neuem \tilde{x} -Wert erfolgt, beträgt

$$w_1 = \frac{\beta_0 \Delta\tilde{x}}{\Delta\tilde{x} \cdot \sum_{k=0}^K \beta_k} = \frac{\beta_0}{\sum_{k=0}^K \beta_k} .$$

- Der Anteil, der während der zweiten Periode vollzogen wird, beträgt

$$w_2 = \frac{\beta_1}{\sum_{k=0}^K \beta_k} .$$

- Nach m Perioden ist demnach ein Anpassungsanteil von

$$w_1 + w_2 + \dots + w_m = \frac{\beta_0 + \beta_1 + \dots + \beta_{m-1}}{\sum_{k=0}^K \beta_k} \quad (22.7)$$

geleistet.

- Die Periodenzahl m , für die der Quotient in (22.7) – der Anpassungsanteil – den Wert 0,5 annimmt, wird als *Median-Lag* bezeichnet.

Multiplikatoren und Median-Lag

In einem dynamischen Modell der Form

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_K x_{t-K} + v_t$$

ist β_0 der kurzfristige Multiplikator und $\sum_{k=0}^K \beta_k$ der langfristige Multiplikator. Die Periodenzahl m , ab der

$$\frac{\beta_0 + \beta_1 + \dots + \beta_{m-1}}{\sum_{k=0}^K \beta_k} \geq 0,5$$

erfüllt ist, wird als Median-Lag (Hälfte der Anpassungszeit) bezeichnet.

22.3 Allgemeine Schätzprobleme dynamischer Modelle

22.3.1 Zwei zentrale Schätzprobleme

- Eine KQ-Schätzung des dynamischen Modells

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_K x_{t-K} + v_t \quad (22.1)$$

stößt auf zwei wesentliche Schwierigkeiten:

- 1 Es existieren nur $T-K$ *vollständige* Beobachtungen.
Für die Schätzung benötigen wir:

$$\text{Anzahl der vollst. Beobacht.} \geq \text{Zahl der Parameter}$$

$$\implies T - K \geq K + 2$$

$$\implies T \geq 2(K + 1) .$$

- 2 Es kann zu hoher Multikollinearität kommen.

22.3.2 Mögliche Lösungsstrategien

- Man kann versuchen, die Zahl der zu schätzenden Parameter zu reduzieren.
- Dies könnte zugleich auch das Multikollinearitätsproblem lindern.
- Koyck (1954) erkannte, dass die Zahl der Parameter dadurch verringert werden kann, dass man von einem funktionalen Zusammenhang zwischen den Parametern β_k ausgeht.
- Koyck unterstellte einen *geometrischen* Zusammenhang.
- Almon (1965) unterstellte einen *polynomialen* Zusammenhang zwischen den Parametern β_k .
- Almon-Modelle finden heute allerdings selten Verwendung.

22.4 Modelle mit geometrischer Lag-Verteilung

22.4.1 Geometrische Lag-Verteilungen

- Beide in Abbildung 22.3 dargestellten Lag-Verteilungen lassen sich formal folgendermaßen ausdrücken:

$$\beta_k = \beta_0 \lambda^k, \quad (22.8)$$

wobei $0 < \lambda < 1$.

- Der Parameter λ signalisiert die Trägheit der Parameterwerte.

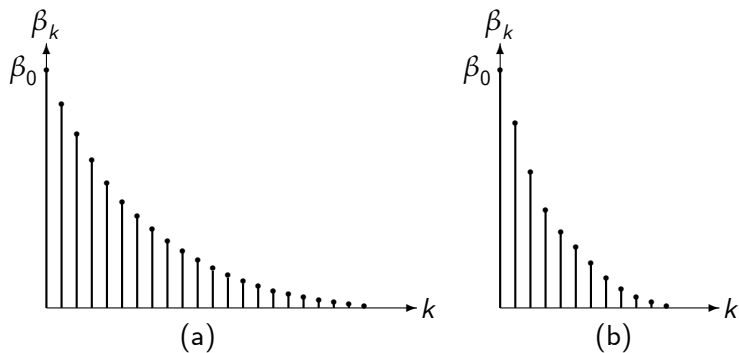


Abbildung 22.3: Zwei geometrische Lag-Verteilungen mit unterschiedlicher Trägheit.

- Eine Lag-Verteilung, die gemäß (22.8) formalisiert werden kann, bezeichnet man als *geometrische Lag-Verteilung*.
- Sie stellt einen Spezialfall *rationaler Lag-Verteilungen* dar.
- Rationale und damit auch geometrische Lag-Verteilungen eignen sich insbesondere für Modelle, bei denen von einem unendlichen Lag-Horizont auszugehen ist.
- In einem solchen Fall lautet das Ausgangsmodell:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + v_t . \quad (22.9)$$

- Ersetzt man die Parameter β_1 , β_2 , usw. durch ihre gemäß Gleichung (22.8) definierten Terme $\beta_0 \lambda$, $\beta_0 \lambda^2$, usw., so ergibt sich:

$$y_t = \alpha + \beta_0 x_t + \beta_0 \lambda x_{t-1} + \beta_0 \lambda^2 x_{t-2} + \dots + v_t . \quad (22.10)$$

- Der langfristige Multiplikator lautet demzufolge:

$$\begin{aligned}\sum_{k=0}^{\infty} \beta_0 \lambda^k &= \beta_0 (1 + \lambda + \lambda^2 + \dots) \\ &= \beta_0 \left(\frac{1}{1 - \lambda} \right) .\end{aligned}\quad (22.11)$$

- Der kurzfristige Multiplikator ist weiterhin durch β_0 gegeben.
- Es lässt sich zeigen, dass der Median-Lag m im Modell (22.10) durch

$$m = -\frac{\ln 2}{\ln \lambda} \quad (22.12)$$

berechnet werden kann.

22.4.2 Koyck-Modell

Koyck-Transformation

- Für die Vorperiode lautet Modell (22.10):

$$y_{t-1} = \alpha + \beta_0 x_{t-1} + \beta_0 \lambda x_{t-2} + \beta_0 \lambda^2 x_{t-3} + \dots + v_{t-1} . \quad (22.13)$$

- Multipliziert man diese Gleichung mit λ und subtrahiert sie anschließend von Gleichung (22.10), dann erhält man

$$y_t - \lambda y_{t-1} = [\alpha - \lambda \alpha] + \beta_0 x_t + v_t - \lambda v_{t-1}$$

und damit

$$y_t = \alpha_0 + \beta_0 x_t + \lambda y_{t-1} + u_t , \quad (22.14)$$

wobei

$$u_t = v_t - \lambda v_{t-1} \quad (22.15)$$

$$\alpha_0 = \alpha(1 - \lambda) . \quad (22.16)$$

Schätzung des Koyck-Modells

- Im Koyck-Modell (22.14) müssen lediglich α_0 , β_0 und λ geschätzt werden.
- Da

$$E(u_t) = E(v_t) - \lambda E(v_{t-1}) = 0 ,$$

bewahrt Annahme B1 auch im Koyck-Modell ihre Gültigkeit.

- Jedoch ist u_t autokorreliert und damit Annahme B3 verletzt:

$$\begin{aligned} \text{cov}(u_t, u_{t-1}) &= E [[u_t - E(u_t)] [u_{t-1} - E(u_{t-1})]] \\ &= E [(v_t - \lambda v_{t-1})(v_{t-1} - \lambda v_{t-2})] \\ &= \underbrace{E(v_t v_{t-1})}_{= 0} - \lambda \underbrace{E(v_t v_{t-2})}_{= 0} \\ &\quad - \lambda E(v_{t-1}^2) + \lambda^2 \underbrace{E(v_{t-1} v_{t-2})}_{= 0} \\ &= -\lambda \sigma^2 \neq 0. \end{aligned}$$

- Es handelt sich bei Gleichung (22.15) um einen sogenannten *Moving-Average-Prozess erster Ordnung*, abgekürzt MA(1)-Prozess.
- Auch Annahme C1 bereitet Probleme, denn die verzögerte endogene Variable y_{t-1} taucht als (zweite) „exogene Variable“ auf.
- Damit ist mindestens eine der exogenen Variablen zufallsabhängig.
- **Frage:** Warum gilt $cov(y_{t-1}, u_t) \neq 0$ (kontemporäre Korrelation)?
Antwort:

- Die ZSKQ-Schätzung bietet einen Ausweg.
- Als Instrumentvariable für y_{t-1} bietet sich x_{t-1} an (aber Multikollinearitäts-Problem).
- Aus Annahme $c1^{***}$ (bzw. Annahme $C1^{***}$) folgt aber, dass für die ZSKQ-Schätzung unter anderem $\text{plim}[(1/T) S_{xx}] = \sigma_x^2$ für alle exogenen Variablen gelten muss.
- Im Koyck Modell benötigt man also

$$\text{plim} [(1/T) S_{yy}] = \sigma_y^2 .$$

- Dies kann nur dann erfüllt sein, wenn die „exogene Variable“ y_{t-1} *stationär* ist.

- Falls x_t eine Zufallsvariable darstellt, muss sie ebenfalls stationär sein.
- Die im Koyck-Modell (22.14) getroffene Annahme $0 < \lambda < 1$ stellt sicher, dass bei stationärem x_t auch y_t stationär ist.
- Stationarität von y_t ist gleichbedeutend mit Stationarität von y_{t-1} .

Numerische Illustration 22.1

Wenn man im Beschäftigungs-Beispiel unterstellt, dass die Werte der Steigungsparameter β_k mit zunehmendem Lag k geringer ausfallen, kann für das Koyck-Modell,

$$y_t = \alpha_0 + \beta_0 x_t + \lambda y_{t-1} + u_t ,$$

auf Basis der Daten in Tabelle 22.1 eine ZSKQ-Schätzung durchgeführt werden (mit x_{t-1} als Instrumentvariable). Dabei wird angenommen, dass x_t stationär ist.

(Forts.)

Tabelle 22.2: Schätzergebnisse für eine ZSKQ-Schätzung (mit x_{t-1} als Instrumentvariable) des Koyck-Modells (22.14).

Variable	Koeff.	$\widehat{\text{sd}}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	-219,175	430,150	-0,510	0,614
Aufträge	0,499	0,233	2,144	0,040
Trägheit (λ)	0,749	0,194	3,854	0,001

(Forts.)

Aus Gleichung (22.8) erhält man:

$$\hat{\beta}_1^{ZSKQ} = 0,374, \quad \hat{\beta}_2^{ZSKQ} = 0,280, \quad \hat{\beta}_3^{ZSKQ} = 0,210, \dots$$

Aus Gleichung (22.16) folgt:

$$\hat{\alpha}^{ZSKQ} = -219,175 / (1 - 0,749) = -873,2 .$$

Der langfristige Multiplikator beträgt

$$\hat{\beta}_0^{ZSKQ} / (1 - \hat{\lambda}^{ZSKQ}) = 1,988 .$$

22.4.3 Ein Verwandter des Koyck-Modells: Partielles Anpassungsmodell

- Das Koyck-Modell lautete

$$y_t = \alpha_0 + \beta_0 x_t + \lambda y_{t-1} + u_t, \quad (22.14)$$

wobei

$$u_t = v_t - \lambda v_{t-1}. \quad (22.15)$$

- Es wurde durch eine Koyck-Transformation aus dem Ausgangsmodell

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + v_t \quad (22.13)$$

gewonnen.

- Es wird nun ein besser fundiertes Ausgangsmodell vorgestellt und ebenfalls in die Form (22.14) gebracht.
- Ausgangspunkt ist der *langfristige* Zusammenhang

$$y_t^* = \mu + \delta x_t, \quad (22.17)$$

wobei μ und δ zwei konstante Parameter sind und y_t^* den zu x_t passenden *langfristigen Gleichgewichtswert* bezeichnet.

- Es wird angenommen, dass die tatsächliche Anpassung $(y_t - y_{t-1})$ nur einen Teil der langfristig eigentlich notwendigen Anpassung $(y_t^* - y_{t-1})$ ausmacht:

$$y_t - y_{t-1} = (1 - \lambda)(y_t^* - y_{t-1}) + v_t, \quad (22.18)$$

wobei v_t eine Störgröße im Anpassungsprozess bezeichnet und $(1 - \lambda)$ den Anteil der sofort vollzogenen Anpassung festlegt.

- Setzt man Gleichung (22.17) in Gleichung (22.18) ein, so ergibt sich

$$y_t - y_{t-1} = (1 - \lambda)(\mu + \delta x_t - y_{t-1}) + v_t .$$

- Umstellen liefert

$$y_t = (1 - \lambda)\mu + (1 - \lambda)\delta x_t + \lambda y_{t-1} + v_t . \quad (22.19)$$

- Definiert man

$$\alpha_0 = (1 - \lambda)\mu$$

$$\beta_0 = (1 - \lambda)\delta$$

$$u_t = v_t ,$$

so erhält man das Koyck-Modell (22.14).

- **Frage:** Welcher Unterschied zum Koyck-Modell besteht jedoch?
Antwort:

- Sofern y_{t-1} und x_t stationär sind, liefert die KQ-Methode konsistente Schätzer.

- Gleichung (22.19) bedeutet für Periode $t - 1$:

$$y_{t-1} = (1 - \lambda)\mu + (1 - \lambda)\delta x_{t-1} + \lambda y_{t-2} + v_{t-1} .$$

- **Frage:** Warum zeigt diese Gleichung, dass auch im Modell adaptiver Erwartungen die KQ-Schätzer nicht unverzerrt sind?

Antwort:

22.4.4 Ein weiterer Verwandter des Koyck-Modells: Modell adaptiver Erwartungen

- Ausgangspunkt ist die folgende Annahme:

$$y_t = \mu + \delta x_{t+1}^e + v_t, \quad (22.20)$$

wobei μ und δ wieder zwei konstante Parameter sind und v_t eine Störgröße.

- **Frage:** Was bedeutet dieser Zusammenhang inhaltlich?
Antwort:

- Im *Modell adaptiver Erwartungen* geht man von folgender Regel aus:
 - wenn $x_t = x_t^e$, dann $x_{t+1}^e = x_t^e$.
 - wenn $x_t \neq x_t^e$, dann wird der Erwartungsfehler ($x_t - x_t^e$) als Korrekturterm in der Prognose für Periode $t+1$ berücksichtigt:

$$x_{t+1}^e = x_t^e + (1 - \lambda)(x_t - x_t^e), \quad (22.21)$$

wobei $0 < \lambda < 1$.

- Umformung der Gleichung (22.21) liefert

$$x_{t+1}^e = (1 - \lambda)x_t + \lambda x_t^e . \quad (22.22)$$

- Aus Gleichung (22.20) ist bekannt, dass

$$x_{t+1}^e = \frac{y_t - \mu - v_t}{\delta}$$

und damit auch

$$x_t^e = \frac{y_{t-1} - \mu - v_{t-1}}{\delta} .$$

- Setzt man diese beiden Beziehungen in Gleichung (22.22) ein, so ergibt sich

$$\frac{y_t - \mu - v_t}{\delta} = (1 - \lambda)x_t + \lambda \frac{y_{t-1} - \mu - v_{t-1}}{\delta} .$$

- Multiplikation mit δ und anschließendes Ausmultiplizieren liefert

$$y_t - \mu - v_t = (1 - \lambda)\delta x_t + \lambda y_{t-1} - \lambda\mu - \lambda v_{t-1}$$

und folglich

$$y_t = (1 - \lambda)\mu + (1 - \lambda)\delta x_t + \lambda y_{t-1} + v_t - \lambda v_{t-1} .$$

- Definiert man schließlich

$$\alpha_0 = (1 - \lambda)\mu$$

$$\beta_0 = (1 - \lambda)\delta$$

$$u_t = v_t - \lambda v_{t-1} ,$$

so erhält man

$$y_t = \alpha_0 + \beta_0 x_t + \lambda y_{t-1} + u_t .$$

- **Frage:** Vergleichen Sie dieses Modell mit dem Koyck-Modell (22.14).

Antwort:

22.5 Modelle mit rationaler Lag-Verteilung und ihre Fehlerkorrektur-Formulierung

- Wir hatten das Koyck-Modell

$$y_t = \alpha_0 + \beta_0 x_t + \lambda y_{t-1} + u_t \quad (22.14)$$

betrachtet, wobei $0 < \lambda < 1$ (*geometrisches Lag-Modell*).

- Modell (22.14) kann als Spezialfall der *rationalen Lag-Modelle* aufgefasst werden.
- Diese Klasse ist durch folgende allgemeine Gleichung definiert:

$$y_t = \alpha_0 + \beta_0 x_t + \mu_1 x_{t-1} + \mu_2 x_{t-2} + \dots + \mu_K x_{t-K} + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_M y_{t-M} + u_t. \quad (22.23)$$

- Die Klasse der rationalen Lag-Modelle ist in der Lage, jegliche Lag-Verteilung zu approximieren.
- Ein Beispiel ist in Abbildung 22.4 wiedergegeben.

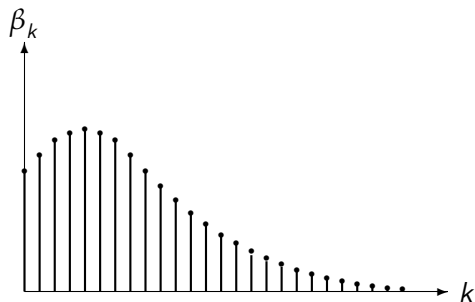


Abbildung 22.4: Eine rationale Lag-Verteilung.

- In der englischsprachigen Literatur hat sich der Begriff *autoregressive distributed lag model* durchgesetzt, oder kurz ADL(M, K)-Modell.
- Dabei gibt M die Anzahl der im Modell berücksichtigten Lag-Variablen der endogenen Variable y_t an und K die Anzahl der berücksichtigten Lag-Variablen der exogenen Variable x_t .
- Hier genügt die Betrachtung eines ADL(1,1)-Modells:

$$y_t = \alpha_0 + \beta_0 x_t + \mu x_{t-1} + \lambda y_{t-1} + u_t, \quad (22.24)$$

wobei weiterhin $0 < \lambda < 1$ als *Stabilitätsbedingung* unterstellt ist.

- Das geometrische Lag-Modell ist ein ADL(1,0)-Modell.

22.5.1 Langfristige Gleichgewichtsbeziehung

- Gleichung (22.24) lässt sich umstellen zu

$$y_t - \lambda y_{t-1} = \alpha_0 + \beta_0 x_t + \mu x_{t-1} + u_t . \quad (22.25)$$

- Im langfristigen Gleichgewicht mit $u_t = u_{t-1} = \dots = 0$ würde gelten, dass $x_t = x_{t-1} = x_{t-2} = \dots = x^*$ und damit auch $y_t = y_{t-1} = y_{t-2} = \dots = y^*$.
- Für diese langfristige Betrachtung liefert Gleichung (22.25):

$$\begin{aligned} y^* - \lambda y^* &= \alpha_0 + \beta_0 x^* + \mu x^* \\ \Rightarrow y^* &= \frac{\alpha_0}{1 - \lambda} + \frac{\beta_0 + \mu}{1 - \lambda} x^* \quad (22.26) \end{aligned}$$

- Gleichung (22.26) gibt an, welcher Wert y^* sich einstellt, wenn der Wert der exogenen Variable dauerhaft auf x^* verharret.
- Der langfristige Multiplikator beträgt demnach

$$\frac{dy^*}{dx^*} = \frac{\beta_0 + \mu}{1 - \lambda} .$$

22.5.2 Fehlerkorrektur-Formulierung des ADL(1,1)-Modells

- Wenn die Störgröße u_t autokorreliert ist, kommt es zu kontemporärer Korrelation mit der „dritten exogenen Variable“ y_{t-1} .
- Ist u_t nicht autokorreliert, so besteht immerhin eine *intertemporäre* Korrelation zwischen u_t und y_{t-1} .
- In diesem Fall erfordert eine konsistente KQ-Schätzung, dass $\text{plim}[(1/T) S_{xx}] = \sigma_x^2$ und $\text{plim}[(1/T) S_{yy}] = \sigma_y^2$.
- Letzteres erfordert, dass y_t (und damit y_{t-1}) sowie x_t (und folglich x_{t-1}) stationär sind.
- Oftmals sind y_t und x_t jedoch nicht stationär, aber immerhin integriert vom Grade 1: y_t und x_t sind $I(1)$.

- Subtrahiert man auf beiden Seiten des Modells (22.24) y_{t-1} und subtrahiert und addiert man auf der rechten Seite $\beta_0 x_{t-1}$, so ergibt sich

$$\begin{aligned}\Delta y_t &= \alpha_0 + \beta_0 x_t + \beta_0 x_{t-1} - \beta_0 x_{t-1} + \mu x_{t-1} - (1-\lambda)y_{t-1} + u_t \\ &= \beta_0 \Delta x_t - [(1-\lambda)y_{t-1} - \alpha_0 - (\beta_0 + \mu)x_{t-1}] + u_t \\ &= \beta_0 \Delta x_t - (1-\lambda) \left[y_{t-1} - \frac{\alpha_0}{1-\lambda} - \frac{\beta_0 + \mu}{1-\lambda} x_{t-1} \right] + u_t, \quad (22.27)\end{aligned}$$

wobei wie gewohnt

$$\begin{aligned}\Delta y_t &= y_t - y_{t-1} \\ \Delta x_t &= x_t - x_{t-1} .\end{aligned}$$

- Es kann demnach drei mögliche Ursachen für eine Veränderung in der endogenen Variable (Δy_t) geben:
 - ① Dritter Summand: Die Störgröße u_t weicht von 0 ab.
 - ② Erster Summand: Kurzfristiger Anpassungseffekt.
 - ③ Zweiter Summand: Aus der Vorperiode übernommener Anpassungsbedarf in der endogenen Variable y_t („Fehlerkorrektur“).

22.5.3 Schätzung des Fehlerkorrekturmodells

- Das Modell (22.27) wird als die „Fehlerkorrektur-Formulierung“ des Modells (22.24) oder einfach als *Fehlerkorrekturmodell* bezeichnet (*Error-Correction-Model*, *ECM*).
- Der Ausdruck in eckigen Klammern kann als *Ungleichgewichtsfehler* e_{t-1} interpretiert werden, welcher angibt, wie stark in Periode $t-1$ der Wert für y_{t-1} von seinem zu x_{t-1} passendem Gleichgewichtswert entfernt war:

$$e_{t-1} = y_{t-1} - \frac{\alpha_0}{1 - \lambda} - \frac{\beta_0 + \mu}{1 - \lambda} x_{t-1} . \quad (22.28)$$

- Wenn x_t und y_t (und damit x_{t-1} und y_{t-1}) jeweils $I(1)$ sind und eine lineare langfristige Beziehung – also die Beziehung (22.26) – zwischen diesen beiden Variablen besteht, dann ist die durch Gleichung (22.28) definierte Zeitreihe der Ungleichgewichtsfehler e_{t-1} stationär.
- In einem solchen Fall werden x_t und y_t (bzw. x_{t-1} und y_{t-1}) als *kointegriert* bezeichnet.
- Wenn im Fehlerkorrekturmodell (22.27) Δx_t und der eckige Klammersausdruck e_{t-1} stationär sind, bleibt Annahme $c1^{**}$ (genauer: Annahme $C1^{**}$) unverletzt und die KQ-Methode liefert für das Fehlerkorrekturmodell (22.27) konsistente Schätzer.

- Die Schätzung umfasst zumeist die folgenden vier Arbeitsschritte:
 - 1 Für jede der Zeitreihen x_t und y_t wird individuell überprüft, welchen Integrationsgrad d sie besitzen.
 - 2 Wenn x_t und y_t beide $I(1)$ sind, dann erfolgt eine KQ-Schätzung der nach y_{t-1} aufgelösten Beziehung (22.28). Diese Schätzung liefert Residuen \hat{e}_{t-1} .
 - 3 Man überprüft, ob die Residuen \hat{e}_{t-1} $I(0)$ sind, ob also x_t und y_t kointegriert sind.
 - 4 Falls x_t und y_t kointegriert sind (e_{t-1} stationär), erfolgt eine KQ-Schätzung des Fehlerkorrekturmodells (22.27). Dabei werden die Residuenwerte \hat{e}_{t-1} anstelle des nicht beobachtbaren Ungleichgewichtsfehlers e_{t-1} – also des eckigen Klammerterms – benutzt.

- Engle und Granger (1987) konnten zeigen, dass dieses Vorgehen konsistente und asymptotisch effiziente Schätzergebnisse liefert.

22.5.4 Fehlerkorrekturmodell und ökonomische Theorie

- In der empirischen Realität wird das langfristige Gleichgewichtsstreben ständig durch kurzfristiges Anpassungsverhalten überlagert.
- Die langfristige Komponente des Fehlerkorrekturmodells (22.27) besitzt einen soliden ökonomischen Unterbau, die kurzfristige Komponente dagegen nicht.
- In der empirischen Implementierung des kurzfristigen Anpassungsverhaltens verwendet man deshalb Fehlerkorrekturmodelle, die aus dem allgemeineren $ADL(M,K)$ -Modell (22.23) und nicht aus der einfachen Variante des $ADL(1,1)$ -Modells (22.24) gebildet wurden.

Kapitel 23

Interdependente Gleichungssysteme

- Bislang wurden nur ökonomische Zusammenhänge betrachtet, die sich durch eine *einzelne* Gleichung erfassen lassen.
- Dabei wurden die Werte der endogenen Variable y_t durch die Werte einer oder mehrerer exogener Variablen x_{kt} erklärt.

Beispiel zu Kapitel 23

Ein Pharma-Hersteller besitzt ein Patent für einen Wirkstoff zur Behandlung eines speziellen rheumatischen Leidens. Wir werden vom Hersteller beauftragt, den Einfluss seiner Werbemaßnahmen w_t auf den Absatz a_t zu ermitteln. Tabelle 23.1 zeigt den verfügbaren Datensatz.

(Forts.)

Tabelle 23.1: Absatz a_t (in 100 g), Werbeanzeigen w_t (in Doppelseiten), Preis des Wirkstoffes p_t (in Euro/100 g) und Anzeigenpreis q_t (in 1000 Euro/Doppelseite) während der vergangenen 24 Quartale.

t	a_t	w_t	p_t	q_t	t	a_t	w_t	p_t	q_t
1	9	67	409	29	13	267	153	219	33
2	107	113	339	27	14	234	161	249	35
3	148	152	309	27	15	205	80	269	32
4	256	186	289	25	16	222	148	269	35
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
12	226	166	259	30	24	156	76	299	40

- Man könnte das folgende ökonometrische Modell aufstellen:

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + u_t , \quad (23.1)$$

wobei die Störgröße u_t die Annahmen B1 bis B4 erfülle.

- Die Kausalität zwischen Werbeaufwand und Absatz läuft aber in beiden Richtungen.
- Dies könnte zu Problemen bei der KQ-Schätzung führen.

23.1 Nicht-Konsistenz der KQ-Schätzer

- Formal lässt sich die umgedrehte Kausalität in der folgenden Gleichung wiedergeben:

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + v_t . \quad (23.2)$$

- Die Störgröße v_t erfülle alle vier B-Annahmen.
- Es wird zugelassen, dass eine Korrelation zwischen v_t und u_t besteht. Dabei wird allerdings angenommen, dass

$$\begin{aligned} \text{cov}(u_t, v_t) &= \sigma_{uv} , & t = 1, 2, \dots, T , \\ \text{cov}(u_s, v_t) &= 0 , & \text{für } s \neq t \text{ und } s, t = 1, 2, \dots, T , \end{aligned}$$

wobei σ_{uv} eine Konstante ist.

- **Frage:** Warum tritt in den Gleichungen (23.1) und (23.2) kontemporäre Korrelation auf?

Antwort:

- Würde man diese Korrelationen missachten und die Gleichungen (23.1) und (23.2) separat mit KQ-Methode schätzen, so wären die Ergebnisse verzerrt und nicht einmal konsistent.

23.2 Indirekte KQ-Methode (IKQ-Methode)

23.2.1 Strukturelle Form versus reduzierte Form

- Um konsistente Schätzer zu erhalten, wird das durch (23.1) und (23.2) definierte Gleichungssystem

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + u_t \quad (23.3)$$

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + v_t \quad (23.4)$$

transformiert.

- In Gleichung (23.3) wird w_t durch den entsprechenden Term in Gleichung (23.4) substituiert:

$$a_t = \alpha + \beta_1(\gamma + \delta_1 a_t + \delta_2 q_t + v_t) + \beta_2 p_t + u_t .$$

- Umstellen liefert

$$(1 - \beta_1 \delta_1) a_t = \alpha + \beta_1 \gamma + \beta_2 p_t + \beta_1 \delta_2 q_t + \beta_1 v_t + u_t .$$

- Division durch $(1 - \beta_1 \delta_1)$ ergibt

$$a_t = \frac{\alpha + \beta_1 \gamma}{1 - \beta_1 \delta_1} + \frac{\beta_2}{1 - \beta_1 \delta_1} p_t + \frac{\beta_1 \delta_2}{1 - \beta_1 \delta_1} q_t + \frac{\beta_1 v_t + u_t}{1 - \beta_1 \delta_1} . \quad (23.5)$$

- Substituiert man in Gleichung (23.4) den Absatz a_t durch den entsprechenden Term in Gleichung (23.3) und löst anschließend nach w_t auf, so ergibt sich

$$w_t = \frac{\gamma + \delta_1 \alpha}{1 - \beta_1 \delta_1} + \frac{\beta_2 \delta_1}{1 - \beta_1 \delta_1} p_t + \frac{\delta_2}{1 - \beta_1 \delta_1} q_t + \frac{\delta_1 u_t + v_t}{1 - \beta_1 \delta_1} . \quad (23.6)$$

- Man definiert in Gleichung (23.5)

$$\pi_1 = \frac{\alpha + \beta_1 \gamma}{1 - \beta_1 \delta_1} , \quad \pi_2 = \frac{\beta_2}{1 - \beta_1 \delta_1} , \quad \pi_3 = \frac{\beta_1 \delta_2}{1 - \beta_1 \delta_1} . \quad (23.7)$$

und in Gleichung (23.6)

$$\pi_4 = \frac{\gamma + \delta_1 \alpha}{1 - \beta_1 \delta_1} , \quad \pi_5 = \frac{\beta_2 \delta_1}{1 - \beta_1 \delta_1} , \quad \pi_6 = \frac{\delta_2}{1 - \beta_1 \delta_1} . \quad (23.8)$$

- Es ergibt sich das Gleichungssystem

$$a_t = \pi_1 + \pi_2 p_t + \pi_3 q_t + u_t^* \quad (23.9)$$

$$w_t = \pi_4 + \pi_5 p_t + \pi_6 q_t + v_t^* , \quad (23.10)$$

wobei

$$u_t^* = \frac{\beta_1 v_t + u_t}{1 - \beta_1 \delta_1} \quad \text{und} \quad v_t^* = \frac{\delta_1 u_t + v_t}{1 - \beta_1 \delta_1} . \quad (23.11)$$

- **Frage:** Bleiben die vier B-Annahmen auch für die Störgrößen u_t^* und v_t^* gültig?

Antwort:

Strukturelle und reduzierte Form

Man bezeichnet die Gleichungen (23.9) und (23.10) als *reduzierte Form*.

Die Ausgangsgleichungen (23.3) und (23.4) werden als *strukturelle Form* bezeichnet.

- Welche Wirkung auf den Absatz a_t ergäbe sich, wenn man bei gegebenen Preisen p_t und q_t nicht die gemäß der Werbegleichung (23.4) passende Zahl der Werbeanzeigen w_t wählen würde, sondern – aus einer Laune heraus – eine Werbeanzeige mehr schalten würde?
- Dies kann wie eine in der Werbegleichung (23.4) aufgetretene positive Störung v_t interpretiert werden.
- Es würde sich zwischen w_t und a_t eine Kette wechselseitiger Effekte einstellen.
- Die Gesamtwirkung erhält man aus der Ableitung der Gleichung (23.11) nach v_t :

$$\partial u_t^* / \partial v_t = \beta_1 / (1 - \beta_1 \delta_1) .$$

23.2.2 Schätzung der Parameter der reduzierten Form

Numerische Illustration 23.1

Tabelle 23.2: Ergebnisse einer KQ-Schätzung der reduzierten Form.

	Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
(23.9)	Konstante	524,896	51,401	10,212	<0,001
	Wirk.-preis	-1,194	0,127	-9,408	<0,001
	Anz.-preis	-0,288	1,337	-0,216	0,831
(23.10)	Konstante	316,445	45,024	7,028	<0,001
	Wirk.-preis	-0,326	0,111	-2,932	0,008
	Anz.-preis	-2,809	1,171	-2,397	0,026

23.2.3 Schätzung der Parameter der strukturellen Form

- Aus (23.7) und (23.8) ergeben sich für die sechs Parameter der strukturellen Form die folgenden Bestimmungsgleichungen:

$$\alpha = \pi_1 - \frac{\pi_3\pi_4}{\pi_6}, \quad \beta_1 = \frac{\pi_3}{\pi_6}, \quad \beta_2 = \pi_2 - \frac{\pi_3\pi_5}{\pi_6}, \quad (23.12)$$

$$\gamma = \pi_4 - \frac{\pi_1\pi_5}{\pi_2}, \quad \delta_1 = \frac{\pi_5}{\pi_2}, \quad \delta_2 = \pi_6 - \frac{\pi_3\pi_5}{\pi_2}. \quad (23.13)$$

- Die beschriebene Schätzung der strukturellen Parameter bezeichnet man als *indirekte KQ-Methode (IKQ-Methode)* und die entsprechenden Schätzer als *IKQ-Schätzer*.
- Sie sind konsistent.

Numerische Illustration 23.2

In Tabelle 23.2 waren die Schätzwerte der Parameter π_1 bis π_6 aufgeführt. Einsetzen in die Gleichungen (23.12) und (23.13) liefert die indirekten KQ-Schätzer:

$$\begin{array}{lll} \hat{\alpha}^{IKQ} = 492,412 & \hat{\beta}_1^{IKQ} = 0,103 & \hat{\beta}_2^{IKQ} = -1,161 \\ \hat{\gamma}^{IKQ} = 173,140 & \hat{\delta}_1^{IKQ} = 0,273 & \hat{\delta}_2^{IKQ} = -2,730 . \end{array}$$

- Die Parameter $\hat{\beta}_1^{IKQ}$, $\hat{\beta}_2^{IKQ}$, $\hat{\delta}_1^{IKQ}$ und $\hat{\delta}_2^{IKQ}$ geben jeweils nur den *Primäreffekt* an.
- Die *Gesamteffekte* sind aus den Parametern der reduzierten Form ($\hat{\pi}_1$ bis $\hat{\pi}_6$) ablesbar.

Numerische Illustration 23.3

Um einen Eindruck vom Ausmaß der Nicht-Konsistenz einer KQ-Schätzung zu vermitteln, wird hier (unzulässigerweise) die strukturelle Form geschätzt:

$$\begin{array}{lll} \hat{\alpha}^{KQ} = 445,469 & \hat{\beta}_1^{KQ} = 0,297 & \hat{\beta}_2^{KQ} = -1,089 \\ \hat{\gamma}^{KQ} = 173,320 & \hat{\delta}_1^{KQ} = 0,272 & \hat{\delta}_2^{KQ} = -2,732 . \end{array}$$

Für die Parameter der Absatzgleichung (23.3) ergeben sich deutliche Unterschiede zur IKQ-Methode, für die Parameter der Werbegleichung (23.4) hingegen nicht.

23.3 Identifikationsproblem

- Aus den KQ-Schätzern der reduzierten Form ($\hat{\pi}_1$ bis $\hat{\pi}_6$) konnte für sämtliche Parameter der strukturellen Form (α , β_1 , β_2 , γ , δ_1 und δ_2) jeweils ein eindeutiger Schätzer errechnet werden.
- Diese Rückrechnung ist nicht für alle interdependenten Gleichungssysteme möglich.

23.3.1 Ein verkleinertes Gleichungssystem

- Es wird das folgende verkleinerte Pharma-Modell betrachtet:

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + u_t \quad (23.14)$$

$$w_t = \gamma + \delta_1 a_t + v_t . \quad (23.15)$$

- Die reduzierte Form lautet

$$a_t = \pi_1 + \pi_2 p_t + u_t^* \quad (23.16)$$

$$w_t = \pi_3 + \pi_4 p_t + v_t^* , \quad (23.17)$$

wobei

$$\pi_1 = \frac{\alpha + \beta_1 \gamma}{1 - \beta_1 \delta_1} , \quad \pi_2 = \frac{\beta_2}{1 - \beta_1 \delta_1} , \quad (23.18)$$

$$\pi_3 = \frac{\gamma + \delta_1 \alpha}{1 - \beta_1 \delta_1} , \quad \pi_4 = \frac{\beta_2 \delta_1}{1 - \beta_1 \delta_1} , \quad (23.19)$$

$$u_t^* = \frac{\beta_1 v_t + u_t}{1 - \beta_1 \delta_1} , \quad v_t^* = \frac{\delta_1 u_t + v_t}{1 - \beta_1 \delta_1} .$$

- Auch diese reduzierte Form kann mit der KQ-Methode problemlos geschätzt werden.
- Da die reduzierte Form vier Parameter aufweist, stehen für die Rückrechnung auf die *fünf* strukturellen Parameter lediglich *vier* Gleichungen zur Verfügung.
- Für welche strukturellen Parameter ist eine Berechnung möglich?
- Aus den Gleichungen für π_2 und π_4 ergibt sich

$$\delta_1 = \frac{\pi_4}{\pi_2}. \quad (23.20)$$

- Die π_1 -Beziehung in (23.18) lässt sich erweitern und umformen zu

$$\begin{aligned}\delta_1 \pi_1 &= \frac{\delta_1 \alpha}{1 - \beta_1 \delta_1} + \frac{\gamma}{1 - \beta_1 \delta_1} \beta_1 \delta_1 + \frac{\gamma}{1 - \beta_1 \delta_1} - \frac{\gamma}{1 - \beta_1 \delta_1} \\ &= \pi_3 + \frac{\gamma}{1 - \beta_1 \delta_1} (\beta_1 \delta_1 - 1) \\ &= \pi_3 - \gamma .\end{aligned}\tag{23.21}$$

- Substituiert man in Gleichung (23.21) δ_1 durch π_4 / π_2 und löst nach γ auf, so ergibt sich

$$\gamma = \pi_3 - \frac{\pi_1 \pi_4}{\pi_2} .\tag{23.22}$$

- Die Parameter α , β_1 und β_2 sind hingegen nicht berechenbar.
- Die IKQ-Methode führt im verkleinerten Gleichungssystem also nur partiell zu eindeutigen Schätzergebnissen.

23.3.2 Ein erweitertes Gleichungssystem

- Es wird das folgende erweiterte Pharma-Modell betrachtet:

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + \beta_3 t + u_t \quad (23.23)$$

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + v_t . \quad (23.24)$$

- Die Variable t entspricht dem jeweiligen Beobachtungsindex t . Sie wird oftmals als *Trendvariable* bezeichnet.
- Die reduzierte Form des Modells lautet:

$$a_t = \pi_1 + \pi_2 p_t + \pi_3 t + \pi_4 q_t + u_t^* \quad (23.25)$$

$$w_t = \pi_5 + \pi_6 p_t + \pi_7 t + \pi_8 q_t + v_t^* . \quad (23.26)$$

- In diesem Fall lauten die Bestimmungsgleichungen:

$$\pi_1 = \frac{\alpha + \beta_1 \gamma}{1 - \beta_1 \delta_1}, \quad \pi_2 = \frac{\beta_2}{1 - \beta_1 \delta_1}, \quad \pi_3 = \frac{\beta_3}{1 - \beta_1 \delta_1}, \quad \pi_4 = \frac{\beta_1 \delta_2}{1 - \beta_1 \delta_1} \quad (23.27)$$
$$\pi_5 = \frac{\gamma + \delta_1 \alpha}{1 - \beta_1 \delta_1}, \quad \pi_6 = \frac{\beta_2 \delta_1}{1 - \beta_1 \delta_1}, \quad \pi_7 = \frac{\beta_3 \delta_1}{1 - \beta_1 \delta_1}, \quad \pi_8 = \frac{\delta_2}{1 - \beta_1 \delta_1}. \quad (23.28)$$

- Es stehen *acht* Gleichungen für die Rückrechnung auf die *sieben* strukturellen Parameter zur Verfügung.
- Es können nicht für alle sieben strukturellen Parameter *eindeutigen* Lösungen gefunden werden.
- Aus den Gleichungen (23.27) und (23.28) ist ersichtlich, dass $\delta_1 = \pi_6 / \pi_2$ und gleichzeitig $\delta_1 = \pi_7 / \pi_3$.
- Wie man die Mehrdeutigkeit direkt aus der strukturellen Form erkennt, wird im Folgenden erläutert.

23.3.3 Ordnungskriterium

- Im Pharma-Modell (23.3) und (23.4) waren die strukturellen Parameter beider Gleichungen eindeutig bestimmbar.
- Im verkleinerten Modell (23.14) und (23.15) waren es nur diejenigen der Gleichung (23.15).
- Im erweiterten Modell (23.23) und (23.24) waren es nur diejenigen der Gleichung (23.23).

Identifikation von Gleichungen

Strukturelle Gleichungen, deren Parameter eindeutig bestimmbar sind, werden als *genau identifiziert* bezeichnet; z.B. Gleichung (23.15).

Strukturelle Gleichungen, deren Parameter aus Mangel an verbleibenden Bestimmungsgleichungen nicht bestimmbar sind, werden als *unteridentifiziert* bezeichnet; z.B. Gleichung (23.14).

Strukturelle Gleichungen, deren Parameter aufgrund eines Überflusses an verbleibenden Bestimmungsgleichungen nicht eindeutig bestimmbar sind, nennt man *überidentifiziert*; z.B. Gleichung (23.24).

- Um herauszubekommen, welche strukturellen Gleichungen genau identifiziert, welche unter- und welche überidentifiziert sind, existiert das sogenannte *Ordnungskriterium (Abzählkriterium)*.
- Das Ordnungskriterium unterscheidet zwischen *system-endogenen* und *system-exogenen (präterminierte oder auch vorherbestimmte)* Variablen.

- Im erweiterten Pharma-Beispiel (23.23) und (23.24) sind p_t , t und q_t system-exogene Variablen, weil ihre Werte außerhalb des Systems festgelegt werden.
- Am Niveauparameter kann man sich eine „imaginäre Variable“ x_{0t} vorstellen, deren Wert für alle Beobachtungen 1 beträgt.
- Auch sie kann den system-exogenen Variablen zugerechnet werden.
- Die Variablen a_t und w_t sind system-endogene Variablen.
- Die Entscheidung, ob es sich bei einer Variable um eine system-endogene oder eine system-exogene Variable handelt, ist nicht immer leicht.

- Es sei

\dot{K} = Anzahl der system-exogenen Variablen im gesamten Gleichungssystem.

K^* = Anzahl der system-exogenen Variablen in der betrachteten Gleichung.

M^* = Anzahl der system-endogenen Variablen in der betrachteten Gleichung.

Ordnungskriterium

Eine Gleichung ist

unteridentifiziert, wenn $\dot{K} - K^* < M^* - 1$,

genau identifiziert, wenn $\dot{K} - K^* = M^* - 1$,

überidentifiziert, wenn $\dot{K} - K^* > M^* - 1$.

- $\dot{K} - K^*$ ist die Zahl der in der betrachteten Gleichung fehlenden system-exogenen Variablen.
- $M^* - 1$ ist die Zahl der rechts des Gleichheitszeichens erscheinenden system-endogenen Variablen der Gleichung.
- Man muss also lediglich die Zahl der ausgeschlossenen system-exogenen Variablen mit der Zahl der rechtsseitigen system-endogenen Variablen vergleichen.

Numerische Illustration 23.4

Das erweiterte Pharma-Modell lautet

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + \beta_3 t + u_t \quad (23.23)$$

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + v_t . \quad (23.24)$$

Frage: Welche Ergebnisse liefert das Ordnungskriterium für diese beiden Gleichungen?

Antwort:

- Das Ordnungskriterium ist eine einfache Bedingung für die genaue Identifikation einer Gleichung.
- Genau genommen ist dies aber nur eine notwendige Bedingung.
- Zusätzlich müsste auch das sogenannte *Rangkriterium* erfüllt sein.
- Normalerweise ist es ausreichend, das Ordnungskriterium zu benutzen.

- Es war bei allen drei Gleichungssystemen möglich, die reduzierte Form zu bilden und ihre Parameter mit der KQ-Methode zu schätzen.
- Lediglich die Rückrechnung auf die strukturellen Parameter war nicht in allen drei betrachteten Gleichungssystemen möglich.
- Nur für genau identifizierte Gleichungen erhält man eindeutige strukturelle Parameterwerte.
- Die IKQ-Methode ist folglich nur für die Schätzung von Parametern genau identifizierter Gleichungen zu gebrauchen.

23.4 Zweistufige KQ-Methode (ZSKQ-Methode)

Berechnung struktureller Parameter

Enthält ein Gleichungssystem eine unteridentifizierte Gleichung, dann sind die strukturellen Parameter dieser Gleichung nicht konsistent schätzbar.

Enthält das System eine genau identifizierte Gleichung, so können sämtliche Parameter dieser Gleichung mit der IKQ-Methode konsistent geschätzt werden.

Für überidentifizierte Gleichungen liefert die IKQ-Methode keine eindeutigen Schätzergebnisse.

Verwendung der ZSKQ-Methode

Die *zweistufige KQ-Methode (ZSKQ-Methode)* liefert auch für den Fall der Überidentifikation eindeutige Ergebnisse und für genau identifizierte Gleichungen liefert sie die gleichen Ergebnisse wie die IKQ-Methode.

23.4.1 ZSKQ-Schätzung mit Hilfe der reduzierten Form

- Die ZSKQ-Methode ist für alle genau identifizierten und alle überidentifizierten Gleichungen eines Gleichungssystems anwendbar, also auch für die beiden Gleichungen des erweiterten Pharma-Modells:

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + \beta_3 t + u_t \quad (23.23)$$

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + v_t . \quad (23.24)$$

- Exemplarisch wird die überidentifizierte Gleichung (23.24) betrachtet.
- Eine KQ-Schätzung ist nicht möglich, aber eine ZSKQ-Schätzung kann vorgenommen werden.

- Welche Instrumentvariable bietet sich an, um die mit v_t kontemporär korrelierte Komponente in a_t von der nicht korrelierten Komponente zu trennen?
- Es wird eine lineare Kombination aus sämtlichen system-exogenen Variablen (x_{0t} , p_t , t und q_t) als Instrumentvariable verwendet:

$$a_t = \pi_1 + \pi_2 p_t + \pi_3 t + \pi_4 q_t + u_t^* . \quad (23.29)$$

- Dabei sollten die numerischen Werte der Parameter π_1 bis π_4 so gewählt sein, dass die Komponente $(\pi_1 + \pi_2 p_t + \pi_3 t + \pi_4 q_t)$ möglichst eng mit a_t korreliert ist.

- Um die optimalen Werte der Parameter π_1 bis π_4 zu finden, wird eine KQ-Schätzung der Gleichung (23.29) durchgeführt.
- **Frage:** Warum ist das Ergebnis dieser KQ-Schätzung bereits bekannt?
Antwort:

- Auf der zweiten Stufe der ZSKQ-Schätzung erfolgt eine KQ-Schätzung der Gleichung

$$w_t = \gamma + \delta_1 \hat{a}_t + \delta_2 q_t + v_t', \quad (23.30)$$

wobei v_t' die Störgröße bezeichnet.

- Diese KQ-Schätzung liefert die gesuchten ZSKQ-Schätzer $\hat{\gamma}^{ZSKQ}$, $\hat{\delta}_1^{ZSKQ}$ und $\hat{\delta}_2^{ZSKQ}$.

23.4.2 ZSKQ-Schätzung im Überblick

- Die ZSKQ-Schätzung einer Gleichung besteht aus zwei Stufen:
 - ① Aus der reduzierten Form werden die geschätzten Werte derjenigen Variable ermittelt, die in der betrachteten strukturellen Gleichung mit der Störgröße dieser Gleichung kontemporär korreliert ist.
 - ② In der betrachteten strukturellen Gleichung wird die korrelierte Variable durch ihr geschätztes Pendant ersetzt und eine KQ-Schätzung dieser modifizierten Gleichung vorgenommen. Man erhält daraus die ZSKQ-Schätzer.

- Das beschriebene Verfahren ist ebenso für genau identifizierte Gleichungen einsetzbar.
- Die Ergebnisse decken sich dann mit denjenigen der IKQ-Methode.
- Die ZSKQ-Schätzer – das heißt, die KQ-Schätzer der modifizierten Gleichung – sind konsistent.

- Würde man auch die Varianz der ZSKQ-Schätzer auf Basis der für KQ-Schätzer üblichen Berechnungsformeln ermitteln, so ergäbe sich allerdings eine systematische Verzerrung.
- Es ist aber möglich, die Ergebnisse der KQ-Schätzung manuell mit einem geeigneten Korrekturfaktor zu versehen und so zu konsistenten Ergebnissen zu kommen.

Numerische Illustration 23.5

Tabelle 23.3 zeigt die Ergebnisse einer KQ-Schätzung für Gleichung (23.25) der reduzierten Form. Setzt man die Schätzer $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$ und $\hat{\pi}_4$ in Gleichung (23.25) ein, so erhält man für jede Beobachtung t den geschätzten Wert \hat{a}_t .

(Forts.)

Tabelle 23.3: Schätzergebnisse für Gleichung (23.25) der reduzierten Form.

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	505,886	70,465	7,179	<0,001
Wirkstoffpreis	-1,194	0,130	-9,215	<0,001
Trend	-0,674	1,669	-0,404	0,690
Anzeigenpreis	0,586	2,559	0,229	0,821

Man ersetzt in (23.24) a_t durch \hat{a}_t und erhält Gleichung (23.30). Eine KQ-Schätzung dieser modifizierten Gleichung liefert die in Tabelle 23.4 abgedruckten ZSKQ-Schätzer.

(Forts.)

Tabelle 23.4: Schätzergebnisse für die modifizierte Gleichung (23.30).

Variable	Koeff.	$\widehat{sd}(\hat{\cdot})$	t-Wert	p-Wert
Konstante	174,131	40,827	4,265	<0,001
Geschätzter Absatz	0,269	0,089	3,038	0,006
Anzeigenpreis	-2,739	1,119	-2,448	0,023

Benutzt man ökonometrische Software, dann erhält man direkt die ZSKQ-Schätzer der Tabelle 23.4. Auch die berechneten Schätzwerte für die Standardabweichung dieser Schätzer wurden auf Basis konsistenter Schätzformeln ermittelt.

23.5 Weitere Beispiele interdependenter Gleichungssysteme

23.5.1 Gleichungssysteme mit Lag-Variablen

- Es wird ein erweitertes Pharma-Beispiel betrachtet:

$$a_t = \alpha + \beta_1 w_t + \beta_2 p_t + \beta_3 w_{t-1} + u_t \quad (23.31)$$

$$w_t = \gamma + \delta_1 a_t + \delta_2 q_t + \delta_3 a_{t-1} + v_t . \quad (23.32)$$

- Lag-Variablen (w_{t-1} und a_{t-1}) werden immer als system-exogen erachtet.
- **Frage:** Welche Resultate liefert das Ordnungskriterium für die Gleichungen (23.31) und (23.32)?

Antwort:

23.5.2 Keynesianisches Makromodell

- Die Volkswirtschaft eines Landes sei durch folgendes makroökonomisches Gleichungssystem beschrieben:

$$c_t = \alpha + \beta y_t + u_t \quad (23.33)$$

$$y_t = c_t + i_t . \quad (23.34)$$

- Dabei bezeichnet c_t den Konsum, y_t das Sozialprodukt bzw. das Einkommen und i_t die Investitionen einer Volkswirtschaft.
- Eine separate KQ-Schätzung der Gleichung (23.33) würde zu nicht-konsistenten Schätzergebnissen führen, denn u_t und y_t sind kontemporär korreliert.

- **Frage:** Welche der Gleichungen des betrachteten Makromodells sind genau identifiziert?

Antwort:

23.5.3 Partielles Marktgleichgewichtsmodell

- Es wird folgendes Angebots- und Nachfragesystem für ein Gut g betrachtet:

$$\text{Angebot} : g_t^A = \alpha + \beta p_t + u_t \quad (23.35)$$

$$\text{Nachfrage} : g_t^N = \gamma + \delta_1 p_t + \delta_2 y_t + v_t . \quad (23.36)$$

- Dabei bezeichnet g_t^A die angebotene Menge des Gutes, g_t^N die nachgefragte Menge, p_t den Preis des Gutes und y_t das verfügbare Einkommen.
- Wenn der Preis p_t flexibel ist, dann kann unterstellt werden, dass er Nachfrage g_t^N und Angebot g_t^A zum Ausgleich bringt:

$$\text{Marktgleichgewicht} : g_t^A = g_t^N = g_t ,$$

wobei g_t die Gleichgewichtsmenge ist.

- Wenn man Gleichung (23.35) umstellt, ergibt sich das folgende Gleichungssystem:

$$\text{Angebot} : p_t = \alpha^* + \beta^* g_t + u_t^* , \quad (23.37)$$

$$\text{Nachfrage} : g_t = \gamma + \delta_1 p_t + \delta_2 y_t + v_t , \quad (23.38)$$

wobei

$$\alpha^* = -\frac{\alpha}{\beta} , \quad \beta^* = \frac{1}{\beta} , \quad u_t^* = -\frac{1}{\beta} u_t .$$

- Im Gleichungssystem (23.37) und (23.38) herrscht in beiden Gleichungen kontemporäre Korrelation.

- **Frage:** Welches Resultat liefert das Ordnungskriterium für die Gleichungen (23.37) und (23.38)?

Antwort: