Empirische Wirtschaftsforschung
Regressionsanalyse

Univ.-Prof. Dr. Joachim Merz

Skriptum zur Vorlesung

Fünfzehnte Auflage 2016

Leuphana Universität Lüneburg

Fakultät Wirtschaftswissenschaften

Forschungsinstitut Freie Berufe

Professur 'Statistik und Freie Berufe'

www.leuphana.de/ffb

Fünfzehnte Auflage 2016

Univ.-Prof. Dr. Joachim Merz, LEUPHANA Universität Lüneburg, Fakultät Wirtschaft, Forschungsinstitut Freie Berufe (FFB), Professur 'Statistik und Freie Berufe', Campus, Geb. 5, Scharnhorststr. 1, 21335 Lüneburg, Tel.: 04131/677-2051, Fax: 04131/677-2059, e-mail: merz@uni.leuphana.de, url: www.leuphana.de/ffb.

Generelle Informationen und Veranstaltungen zur empirischen Wirtschaftsforschung finden Sie auf unserer FFB-Homepage www.leuphana.de/ffb.

.

Empirische Wirtschaftsforschung – Regressionsanalyse

Skriptum zur Vorlesung

Univ.-Prof. Dr. Joachim Merz

THEMENBEREICHE

Empirische Wirtschaftsforschung
Regressionsanalyse

VII HYPOTHESENTESTS IM CLR 105

VIII BEISPIELE CLR-ANALYSE 120

IX TESTS DER ANNAHMEN DES KLASSISCHEN LINEAREN REGRESSIONSMODELLS (CLR) 131

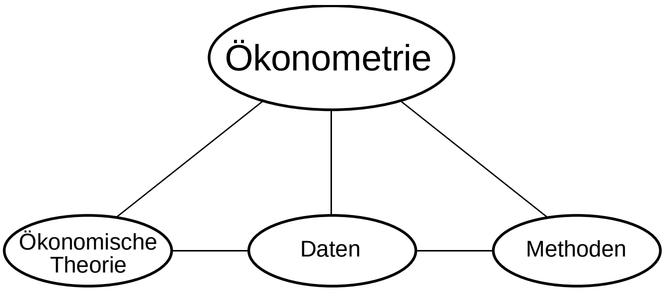
I EINLEITUNG: EMPIRISCHE WIRTSCHAFTSFORSCHUNG UND ÖKONOMETRIE – ZIELE UND AUFGABEN

Was ist Empirische Wirtschaftsforschung (EWF)? Welche Ziele werden damit verfolgt? Wie ist das Verhältnis zwischen Empirischer Wirtschaftsforschung und Ökonometrie zu sehen? Wie ist die Vorgehensweise einer theoretisch/empirisch fundierten Analyse im Rahmen der Empirischen Wirtschaftsforschung? Diese Fragen wollen wir zu Beginn klären und anhand der Keynes'schen Konsumtheorie acht zentrale Schritte der Vorgehensweise der Empirischen Wirtschaftsforschung/Ökonometrie illustrieren.

1 EMPIRISCHE WIRTSCHAFTSFORSCHUNG (EWF) UND ÖKONOMETRIE - ABGRENZUNG

Empirische Wirtschaftsforschung:

Die beobachtete Realität mit wirtschaftstheoretischen Modellen zu erklären, zu verstehen und vorauszusagen (nach Hansen 1997, Hübler 2005).



Ökonometrie:

Griech. oikonomia (Verwaltung, Wirtschaft); metron (Maß, Messung)

Econometrics:

Literally interpreted, econometrics means "econometric measurement". Although measurement is an important part of econometrics, the scope of econometrics is much broader. As can be seen from the following quotations:

- Econometrics, the result of a certain outlook on the role of economics, consists of the application of mathematical statistics to economic data to lend empirical support to the models constructed by mathematical economics and to obtain numerical results (Tintner 1968, S.74).
- …econometrics may be defined as the social science in which the tools of economic theory, mathematics, and statistical inference are applied to the analysis of economic phenomena (Goldberger 1964, S.1).

- …econometrics may be defined as the quantitative analysis of actual economic phenomena based on the concurrent development of the theory and observation, related by appropriate methods of inference (Samuelson/Koopmans/Stone 1954, S. 141).
- Econometrics is concerned with the empirical determination of economic laws (Theil 1971, S.1).
- The art of the econometrician consists in finding the set of assumptions that are both sufficiently specific and sufficiently realistic to allow him to take the best possible advantage of the data available to him (Malinvaud, 1966, S.514).
- Econometricians are a positive help in trying to dispel the poor public image of economics (quantitative or otherwise) as a subject in which empty boxes are opened by assuming the existence of can-openers to reveal contents which any ten economists will interpret in eleven ways (Darnell, Evans, 1990, S.54).
- The method of econometric research aims, essentially, at the conjunction of economic theory and actual measurements, using the theory and technique of statistical inference as a bridge pier (Haavelmo, 1944, S.iii).
- In the first issue of Econometrica, the Econometric Society stated that ... its main object shall be to promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate the natural sciences. But there are several aspects of the quantitative approach to economics, and no single one of these aspects, taken by itself, should be confounded with econometrics. Thus, econometrics is by no means the same as economic statistics. Nor is it identical with what we call general economic theory, although a considerable portion of this theory has a definitely quantitative character. Nor should economics be taken as synonymous [sic] with the application of mathematics to economics. Experience has shown that each of these three view-points [sic], that of **statistics, economic theory, and mathematics**, is a necessary, but not by itself a sufficient condition for a real understanding of the quantitative relations in modern economic life. **It is the unification of all three that is powerful.** And it is this unification that constitutes econometrics (Frisch 1933).
- Econometrics consists of an application of statistical methods to economic data (Maddala 1977, S. 3).
- Models play a major role in all econometric studies, whether theoretical or applied. Indeed, defining econometrics as the branch of economics concerned with the empirical estimation of economic relationships, models, together with data, represent the basic ingredients of any economic study. Typically, the theory of phenomena under investigation is developed into a model which is further refined into an econometric model (Intrilligator 1983, S. 182).

EWF kann weiter gefasst werden → vermehrte Betonung auf:

- Datengewinnung

- Empirische Sozialforschung: Untersuchungsplanung, Datenerhebung, Datenauswertung (multivariate Verfahren) (Diekmann 1995)
- Politikanalyse mit Mikro-, Makrosimulation

4-M-Strategie der Mikro-/ Makrosimulation (Merz 1993):

*M*ikro-/ Makro – Theorie
*M*ikro-/ Makro – Daten
*M*ikro-/ Makro – Methoden
*M*ikro-/ Makro – Simulation

2 Ökonometrie - Methodologie und Vorgehensweise

2.1 Acht Schritte der ökonometrischen Modellierung

1. Ökonomische Theorie
2. Mathematisches Modell
3. Ökonometrisches Modell
4. Daten
5. Schätzen des ökonometrischen Modells
6. Hypothesentests
7. Vorhersage
8. Politikanalyse, -kontrolle, -simulation

2.2 Ökonometrische Modellbildung am Beispiel der Keynesschen Konsumtheorie

1. Ökonomische Theorie: Keynes

(A) "The fundamental psychological law ... is that men [women] are disposed, as a rule and on average, to increase their consumption as their income increases, but not as much as the increase of their income".

(B): "But, apart from the short-period in the level of income, it is also obvious that a higher absolute level of income will tend as a rule to widen the gap between income and consumption ... . These reasons will lead, as a rule, to a greater proportion of income being saved as real income increases." (Keynes 1936, p. 96)

2. Spezifikation des mathematischen Modells

(A) Konsum C ist abhängig vom Einkommen y: $C = f(y)$

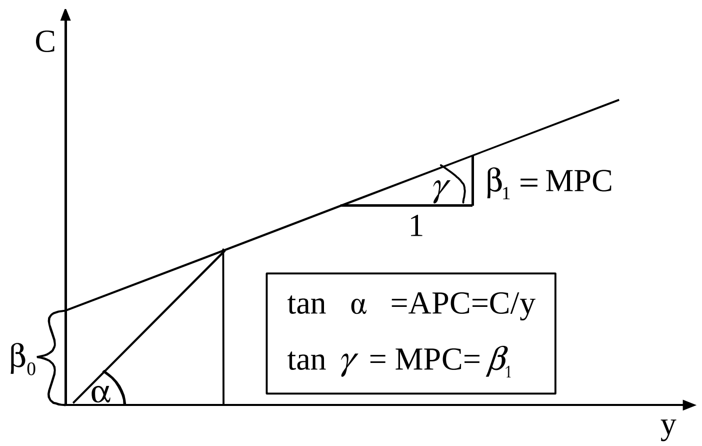
Marginal Propensity to Consume (MPC): $0 < MPC = \frac{dC}{dy} < 1$

(B) Average Propensity to Consume fällt, wenn das Einkommen steigt :

$$\frac{d(APC)}{dy} = \frac{d(C/y)}{dy} = \frac{(MPC - APC)}{y} < 0 \quad \Rightarrow \quad MPC < APC$$

$$APC = \tan\alpha = C/y$$

Mathematisches Modell: $C = \beta_0 + \beta_1 y$ wobei APC > MPC



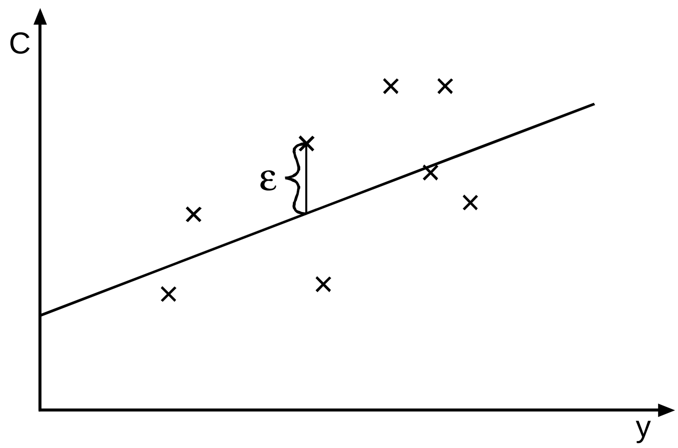
APC > MPC, tanα > tanγ, d.h. $\alpha > \gamma$ wenn β_0 positiv ist.

3. Spezifikation des ökonometrischen Modells

Nicht nur deterministische, auch stochastische Einflüsse:

$$C = \beta_0 + \beta_1 y + \varepsilon$$

ε = error term, Zufallsvariable



4. Daten

Mikro- oder Makrodaten, z.B. aggregierte Daten

C = personal consumption expenditure

y = gross domestic product (GDP, BIP)

in bestimmten Maßeinheiten etc. sachlich, räumlich, zeitlich bestimmt.

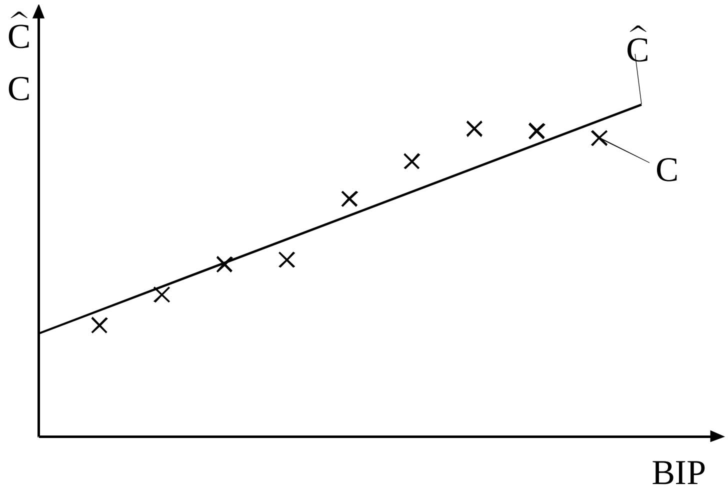
5. Schätzen des ökonometrischen Modells

Schätzen der Parameter β_0, β_1 → Regressionsanalyse

$$\hat{C} = 39894{,}71 + 0{,}54 \text{ BIP}$$

$\hat{C}$ = privater Verbrauch (geschätzt)

BIP = Bruttoinlandsprodukt in der BRD (West) 1975 - 1988 [Mio. DM]



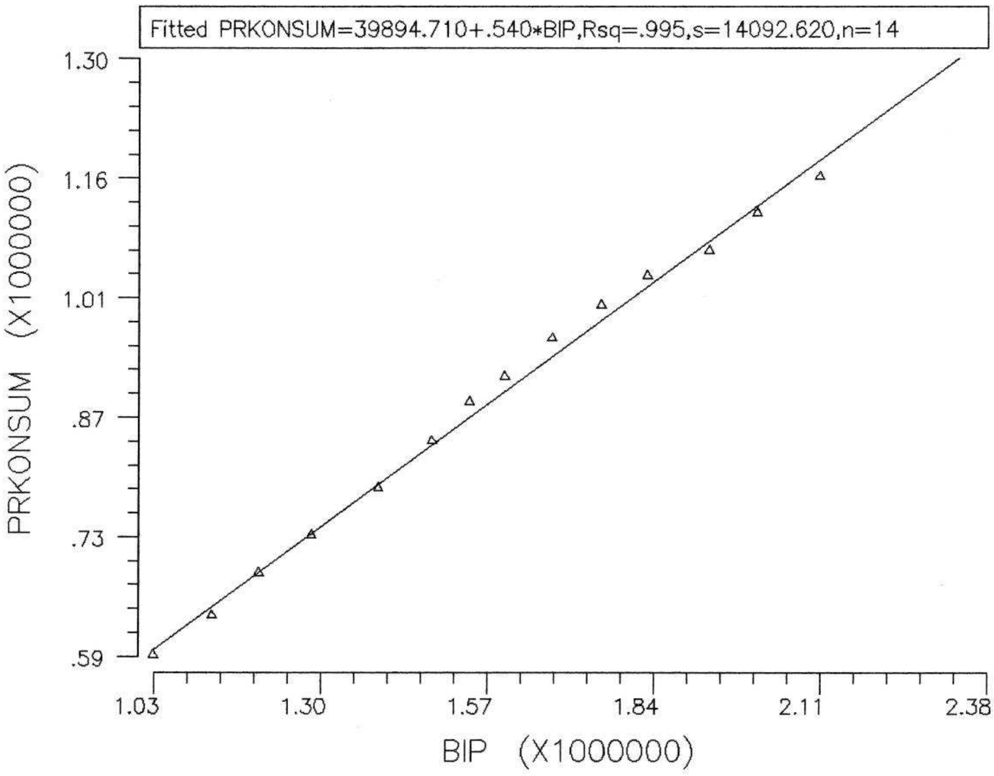
Steigung = MPC = β_1 = 0,54

Interpretation der Steigung:

Im Durchschnitt erhöht eine zusätzliche Einheit des BIP den privaten Verbrauch um 0,54 Einheiten, wenn alle anderen Einflussgrößen konstant gehalten werden.

Keynes-Beispiel:

```
PRKONSUM: Privater Verbrauch BRD (West) 1975-1988 (C)
BIP:      Bruttoinlandsprodukt (nominal) BRD (West) 1975-1988 (Y)
          in Mio. DM

Row  Jahr        PRKONSUM    BIP
  1 1975       .58533E+06  .10269E+07
  2 1976       .63251E+06  .11217E+07
  3 1977       .68316E+06  .11978E+07
  4 1978       .72890E+06  .12833E+07
  5 1979       .78502E+06  .13923E+07
  6 1980       .84078E+06  .14789E+07
  7 1981       .88785E+06  .15409E+07
  8 1982       .91805E+06  .15979E+07
  9 1983       .96416E+06  .16748E+07
 10 1984       .10036E+07  .17558E+07
 11 1985       .10383E+07  .18305E+07
 12 1986       .10686E+07  .19312E+07
 13 1987       .11138E+07  .20091E+07
 14 1988       .11572E+07  .21109E+07

ET Regressionsergebnis
===========================================================================
Ordinary    Least Squares
Dependent Variable        PRKONSUM     Number of Observations            14
Mean of Dep. Variable  886235.7143     Std. Dev. of Dep. Var. 183070.365569
Durbin Watson statistic      .4943     Estimated Autocorrelation     .75284
Std. Error of Regr.     14089.9110     Sum of Squared Residuals .238231E+10
Total Variation         .43569E+12     Regression Variation      .43331E+12
Regression degrees of freedom =  1     Residual degrees of freedom =     12
R - squared                 .99453     Adjusted R - squared          .99408
F( 1,  12)               2182.6383     Prob. Value for F             .00000
Akaike Information        19.23799     Amemiya Prediction   226886390.43136
===========================================================================
Variable Coefficient Std. Error t-ratio Prob|t|>x  Mean of X   Std.Dev.of X
---------------------------------------------------------------------------
Constant   39892.9   .1850E+05   2.156   .05208
BIP        .539753   .1155E-01  46.719   .00000    1568019.2857 338245.9044
```



6. Hypothesentests

Test der Keynes´schen Theorie: 0 < MPC < 1?

MPC = β_1 = 0,54 < 1? ok

signifikant < 1? mittels statistischer Inferenz wäre noch zu testen, ob sich der geschätzte Parmeter signifikant von 1 unterscheidet. Dies entspricht einem t-Test, in dem der H_0 Wert nicht 0 sondern 1 ist (siehe Kapitel VII).

APC = $\tan\alpha = \beta_0 > 0$? ok

7. Vorhersage, Prognose

Wie hoch ist der vorhergesagte Konsum bei einem BIP von 2,5 Billionen DM (unter der Annahme: ähnliches Verhalten)?

$$\hat{C} = 39894{,}71 + 0{,}54 \times (2\,500\,000[\text{Mio.}])$$

$$\hat{C} = 1{,}39 \text{ Billionen DM}$$

Politikanalyse: Welche Wirkung hat eine Investitionsänderung auf die Wirtschaft?

Einkommensmultiplikator

$$M = \frac{1}{1 - MPC} = \frac{1}{1 - 0{,}54} = 2{,}17$$

Ceteris paribus, d.h. unter sonst gleichen Bedingungen gilt:

Eine Erhöhung (Senkung) einer Geldeinheit der Investitionen wird (in etwa) das Einkommen (längerfristig) um mehr als das Doppelte des Eingesetzten erhöhen (senken).

8. Politikanalyse, -kontrolle und –simulation

Optimal Control: In welcher Höhe müssen sog. Kontrollvariablen sein, damit bestimmte Ziele erreicht werden?

Basis: umfangreiches makroökonimisches Modell

Makrosimulation: Mit veränderten Staatsausgaben/ Parametern ergeben sich bestimmte Werte von Zielvariablen (Szenarienanalyse).

Beispiele:

- Optimal control mit linearen Modellen in: Stöppler 1980
- Optimal control mit nichtlinearen Modellen: Galler 1975
- ‚Einfache' Ermittlung eines Kontrollvariablenwertes:

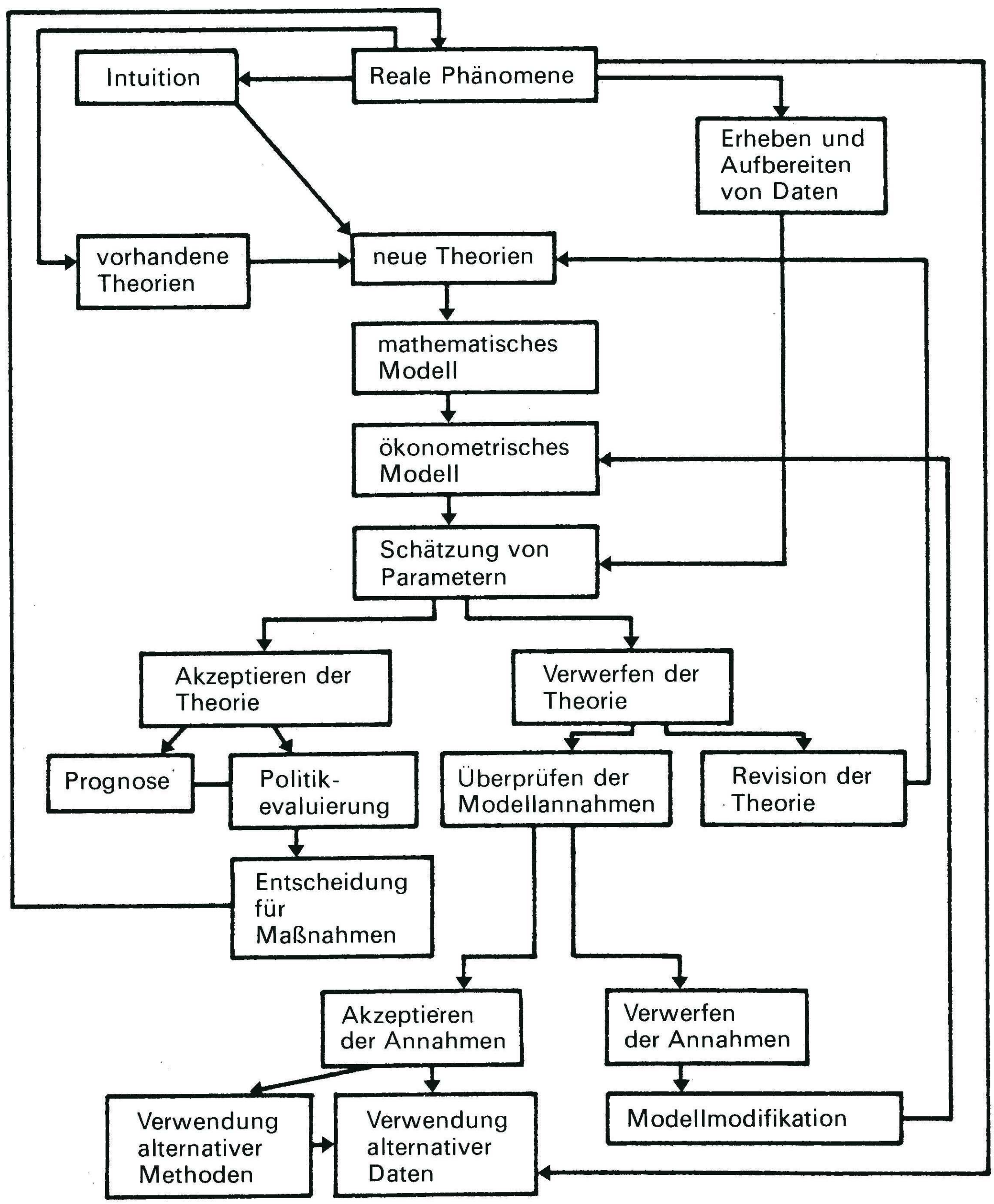
Wie hoch müssen die Staatsausgaben (z.B. Infrastrukturmaßnahmen) sein, damit eine Arbeitslosenquote von 8% (die mit einem Konsumlevel von 1,5 Billionen DM verbunden sei) erreicht würde?

$$1500000 = 39894,71 + 0,54 \times \text{BIP}$$
$$\text{BIP} = 2,7\ [\text{Bio. DM}]$$

Interpretation:

Steigerung der Staatsausgaben auf ein BIP von 2,7 Billionen DM wird - bei einer marginalen Konsumquote von 54% - Konsumausgaben von 1,5 Billionen DM und damit verbundene 8% Arbeitslose „produzieren".

Ablaufschema ökonometrischen Arbeitens



Quelle: Hübler (1989)

3 REGRESSIONSANALYSE - EINFÜHRENDE BEISPIELE

- Gesamtkosten = f(Produktionsmenge, ...)
- Ausgaben = f(Preisen, Haushaltseinkommen, sozioökonomischen Variablen)
- Ernteertrag = f(Niederschlagsmenge)
- Elektrizitätsnachfrage / Strukturansatz / Zeitreihenansatz
- Ausgaben privater Haushalte (Makro)

- Messung von Qualitätsänderung / Hedonic Price Models
- Portfolio-Analyse / Capital Asset Pricing Model (CAPM)
- Lohnsatz - Determinanten / Lohn-Diskriminierung
- Aggregierte Investitionen / Determinante / Politikanalyse
- Arbeitsangebot von Frauen und Männern
- …

Einführende Beispiele

Weitere Beispielprobleme, die mit der (klassischen und verallgemeinerten) Regressionsanalyse untersucht werden können (z.B. Gruber 1997, Fahrmeir, Kneib und Lang 2009).

Beispiel 1

Preise: Der Preis eines kurzfristig nicht produzierbaren Gutes, den der Verbraucher in Periode t zu zahlen hat, hängt vor allem ab (ist eine Funktion) von der in Periode t angebotenen Menge dieses Gutes und von den Preisen von konkurrierenden Gütern in Periode t.

Wie groß ist die Änderung des Preises dieses Gutes, die von Änderungen der angebotenen Menge oder der Preise konkurrierender Güter „verursacht“ wird?

Beispiel 2

Konsumausgaben: Welchen Einfluss haben der jeweilige Güterpreis und die Preise anderer Güter, das verfügbare Einkommen, das Humankapital und andere sozio-ökonomische Größen auf die privaten Konsumausgaben?

Beispiel 3

Produktionskosten: Die durchschnittlichen Produktionskosten in einem Betrieb in Periode t hängen im wesentlichen ab von den Preisen und den eingesetzten Mengen der Produktionsfunktion. Wie stark ist diese funktionale Abhängigkeit der durchschnittlichen Produktionskosten?

Beispiel 4

Absatz: Der Absatz eines Unternehmens (Y) hänge positiv ab von den Werbeaufwendungen (X2) und dem Index der Nettoproduktion für das produzierende Gewerbe (X3), der ein für das betreffende Unternehmen wichtiger „Umweltindikator“ sein soll.

Die funktionale Abhängigkeit der „Zielgröße“ Y von den Regressoren X2 und X3 ist zu bestimmen.

Beispiel 5

Return on investment: Unternehmen einer Branche, die in ein und derselben Volkswirtschaft und damit unter zumindest ähnlichen Rahmenbedingungen arbeiten, erzielen unterschiedliche Wirtschaftsergebnisse.

Wie hängt die Rentabilität (return on investment, ROI) funktional ab von den Bestimmungsfaktoren Marktanteil, Produktqualität, Marketingausgaben, Ausgaben für Forschung und Entwicklung, Investitionsausgaben?

Beispiel 6

Ernteertrag: Der Ernteertrag von Tomaten hängt bei gegebenen Boden- und Klimaverhältnissen und gegebenen sonstigen ertragsbestimmenden Faktoren ab von der „richtigen" Kombination der Reinnährstoffe Stickstoff, Phosphorsäure und Kalium sowie verschiedener Spurenelemente, die als Handelsdünger gegeben werden.

Von welcher Form ist die funktionale Abhängigkeit des Tomatenertrags von den genannten Reinnährstoffen?

Beispiel 7

Arzneiwirksamkeit: Die Wirksamkeit einer blutdrucksteigernden Arznei ist zu untersuchen. Ein Ziel der Untersuchung ist es, für Patienten mit unterschiedlichem (aber in jedem Fall zu niedrigem) Ausgangsblutdruck die im Durchschnitt vieler Patienten optimale Dosierung dieser Arznei, die in Tropfen verabreicht wird, zu ermitteln.

Wie ist die funktionale Abhängigkeit der blutdrucksteigernden Wirkung von der verabreichten Tropfenzahl und dem Ausgangsblutdruck der Patienten?

Beispiel 8

Wahlen: Die Unterschiede zwischen Wahlkreisen im Stimmenanteil einer Partei hängen vor allem ab von der Altersstruktur der Bevölkerung, dem Katholikenanteil, dem Frauenanteil, der Bildungsstruktur der Bevölkerung sowie von der Attraktivität der Programme und Kandidaten der konkurrierenden Parteien.

Von welcher Form ist die funktionale Abhängigkeit? Ist die Abhängigkeit in der Zeit konstant?

Beispiel 9

Entwicklungsökonomie: Analyse sozio-ökonomischer Determinanten der Unterernährung neugeborener Kinder in Entwicklungsländern.

Beispiel 10

Credit Scoring: Analyse der Kreditwürdigkeit von privaten Bankkunden

Beispiel 11

Marktforschung: Wie wirken Verkaufsförderungsmaßnahmen**?**

Beispiel 12

Versicherungsprämien: Analyse von Schadenshäufigkeit und Schadenshöhe und weiterer Merkmale in ihrem Einfluss auf die Versicherungsprämien?.

Beispiel 13

Ökologie: Wirkungsanalyse von Einflussfaktoren auf den Waldzustand

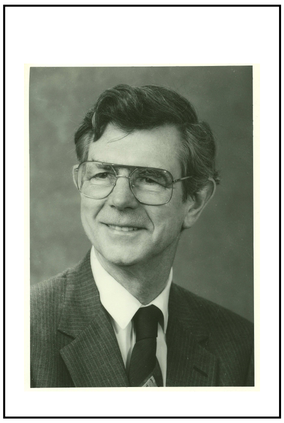
Beispiel 14

Einkommen: Welche Faktoren bestimmen das persönliche Einkommen?

…

4 GUY ORCUTT: A PIONEER IN ECONOMETRIC COMPUTATION

Born in a Detroit suburb in 1917, Guy Orcutt spent much of his spare time as a high school student experimenting with electricity and chemistry, even accidentally creating an explosion that redecorated part of his bedroom ceiling. After graduating from the University of Michigan with majors in physics and mathematics, Orcutt spent the summer of 1939 at a Quaker work camp in a severely depressed coal mining region of West Virginia. This experience kindled an interest in economics and led him to enrol as a graduate student in economics at Michigan, even though he had not taken a single undergraduate economics course. Mentored by Arthur Smithies, as a graduate student Orcutt was encouraged to expand his interests in using electronic devices for undertaking complex mathematical computations. This led to Orcutt's design of a "regression analyser", the theory and circuit diagrams of which are prominently displayed in his Ph.D. dissertation.



Since the actual construction of such a machine required considerable time, funds for materials and help from machinists (support that was not available to Orcutt as a graduate student), it was not until late 1945 that a prototype of Orcutt's regression analyser was successfully constructed. This took place at the Massachusetts Institute of Technology, where Orcutt was hired in 1944 to teach beginning economics courses to engineering undergraduates but was simultaneously involved with Professor George Wadsworth of the MIT Mathematics Department in an effort to improve weather forecasting by relating the weather at a given location to the weather somewhat earlier at other locations.

Orcutt's regression analyser was successfully demonstrated at the annual meetings of the American Economic Association in Cleveland in 1945, and his interest in auto correlated time series attracted the attention of Richard Stone, who invited Orcutt to Cambridge University beginning in the fall of 1946. There, in the newly created Department of Applied Economics under the direction of Stone, Orcutt refurbished his regression analyser, availing himself of facilities at the Cavendish Laboratory. He then used the regression analyser in Monte Carlo studies that focused on developing tests of significance of apparent relationships between time series having auto correlated characteristics similar to actual time series being used in econometric modelling of national economies. These efforts resulted in two papers, one in the 1948 volume of the Journal of the Royal Statistical Society, Series A, and the other with S.F: James in the December 1948 issue of Biometrika.

Together with a Cambridge graduate student named Donald Cochrane, Orcutt also developed the now widely used Cochrane-Orcutt estimator for models with first-order autocorrelated disturbances, which was described in the March and September 1949 issues of the Journal of the American Statistical Association. In addition to his pioneering work on electronic computations and the estimation of autocorrelated time series models, Orcutt is widely known for research on the measurement of price elasticities in international trade and the development of microanalytic models of socio-economic behaviour.

Guy Orcutt retired from Yale University in 1988 and now lives with his wife Gail (and his personal computer) in Grantham, New Hampshire. Quelle: Berndt (1991).

Guy Orcutt was a pioneer of microsimulation with his seminal paper on 'A New Type of Socio Economic Systems (Orcutt 1957). Some further information about microsimulation: Merz 1991; Orcutt, Merz and Quinke 1986). Guy Orcutt, born July 5, 1917 died in March 2006.

II DATEN UND STRUKTUR ÖKONOMETRISCHER MODELLE, ÖKONOMETRISCHE SOFTWARE

Bevor wir auf die eigentliche Modellformulierung eingehen, werden zentrale Bausteine und Tools behandelt. Das Kapitel bietet einen Überblick über Anforderungen und unterschiedliche Klassifizierungen der Daten und mögliche Datenquellen. Zum anderen wird die Struktur ökonometrischer Modelle, Datenbanken als Quellen für die Analyse sowie ökonometrische Software vorgestellt, mit der die Daten dann weiter verarbeitet werden können.

1 DATEN

1.1 Anforderungen und Messniveaus

Anforderungen an die Datenqualität:

- Objektivität: Eindeutigkeit; frei von subjektiver Wahrnehmung
- Reliabilität: Zuverlässigkeit; unter gleichen Bedingungen jeweils das gleiche erfassen
- Validität: Gültigkeit; theoretisches Konstrukt empirisch erfassbar operationalisieren

Probleme: z.B. Messung von Wohlstand → Volkseinkommen pro Kopf

Messung von technischem Fortschritt → Forschungsausgaben

Messniveaus:

Adäquationsproblem (Übereinstimmung von theoretischer Frage und empirischer Operationalisierung); Frankfurter Schule: Blind, Flasskämper, Grohmann

Begriffe:

Messung: Merkmalszuordnung
Daten: gemessene Merkmale (Variablen)
Messniveaus: Skalierung

Skalenniveaus:

- Nominalskala:
  Verschiedene Ausprägungen eines Merkmals drücken nur Gleichheit oder Ungleichheit aus, keine Wertung oder Reihung. Die Ausprägungen werden durch Zahlen codiert. Beispiele: Postleitzahlen, Gliederungsziffern, Farben, Religion
- Ordinalskala:
  Merkmalsausprägungen können in Rangreihenfolge gebracht werden. Keine Informationen über die Abstände der Ausprägungen. Beispiele: Schulnoten, Pferderennen: 1,2,3 (Dreier-, Zweierwette)
- Kardinal- oder metrische Skala:
  Merkmalsausprägungen können in eine Rangfolge gebracht werden und:

- *Intervallskala:* Zusätzliche Definition des Abstandes zwischen zwei verschiedenen Merkmalsausprägungen; Bsp.: Thermometer (in °C)
- *Verhältnisskala* (stetig oder diskret): Es kommt noch ein natürlicher Nullpunkt hinzu; Bsp.: Körpergröße, Einkommen, Alter
- *Absolutskala:* Es kommt noch eine natürliche Einheit dazu; Bsp.: Stückzahl

Qualitative, sozioökonomische Größen:
Erfassung durch Dummy-Variablen (Nominalskala), z.B. - Haushaltstyp, Familienstand, Geschlecht ("the unfortunate econometrician who has to take a dummy as a proxy for sex")

Referenz bei sich erschöpfenden Benennungen: ausgelassene Kategorien (im Bezug auf ...)

1.2 Querschnitts-, Längsschnitt- und Paneldaten

- Querschnittsdaten:
  Daten werden zu einem bestimmten Zeitpunkt erhoben (i =1,2,...,N), z.B. Umfragedaten (cross section):
  - ALLBUS
  - Sfb3-Nebenerwerbstätigkeitsumfrage
  - Einkommens und Verbrauchsstichprobe (EVS)
  - Mikrozensus
- Längsschnittdaten:
  Daten werden zu verschiedenen Zeitpunkten erhoben (t=1,2,...,T), z.B. Zeitreihendaten (time series)
  - volkswirtschaftliche Gesamtrechnung mehrerer Perioden, C, Y, L ...
  - Betriebskennzahlen mehrerer Perioden (Beschäftigung, Bilanzen, etc.)
- Kombination von Zeitreihen- und Querschnittsdaten: Datenmerge (Mikro- und Makrodaten)
  - Mikrozensus und Aggregate
  - Family Expenditure Survey (FES) und Aggregate
- Paneldaten:
  Gleiche Beobachtungsträger werden zu verschiedenen Zeitpunkten befragt
  - Sozioökonomisches Panel von Sfb3/ DIW (Deutsches Institut für Wirtschaftsforschung)

2 STRUKTUR ÖKONOMETRISCHER MODELLE

2.1 Variablen, Parameter, Störgrößen und Modellgleichungen verbinden

Begriffe:

Variablen: Merkmal
Parameter: konstante Größen, Koeffizienten
Störgröße: Fehlerterm, Residuum

Beispiel:

$$C_t = 1{,}5 + 0{,}8y_t \left(+\varepsilon_t\right)$$

ε_t = Störgröße

y_t = Volkseinkommen (Variable)

0,8 = geschätzter Koeffizient

$$0{,}8 = \frac{\partial C_t}{\partial y_t} = \text{marginale Konsumquote (MPC)}$$

Unterscheidung:

- institutionelle Relationen: $y_{netto} = (1-\tau)y_{brutto} = 0{,}67 \times y_{brutto}$
  (τ =0,33 = beispielhafter Steuersatz)
- technologische Relationen: Energieverbleib = 0,7*Energieinput
  (30% Umwandlungsverlust)
- Verhaltensrelationen: $y_i = a_0 + a_1$ Schuljahre

Variablen:

endogene: werden innerhalb des Modells erklärt
exogene: werden von außen vorgegeben
verzögert: x_{t-1}
unverzögert: x_t
prädeterminierte: alle Variablen der rechten Seite einer Gleichung

(→ simultane Gleichungssysteme)

2.2 Strukturelle und reduzierte Form

Strukturform: inhaltlich interessante, eigentliche Menge von Modellgleichungen
Reduzierte Form: abgeleitete Form mit prädeterminierten Variablen auf der rechten Seite

3 Datenbanken und Ökonometriesoftware

3.1 Datenbanken

- Statistisches Bundesamt:
  http://www.destatis.de

- Forschungsdatenzentren des Bundes und der Länder:
  Mit vielen Mikrodaten, Campusfiles
  http://www.forschungsdatenzentrum.de

- Bundesagentur für Arbeit:
  Mit vielen Arbeitsmarkdaten
  http://www.statistik.arbeitsagentur.de

- DSI (Data Service & Information) Macro Data International:
  Makrozeitreihen aus Bundesbankberichten, OECD, IMF, Eurostat Disclosure Spectrum Database: Jährliche Berichte von mehr als 15.000 amerikanischen Unternehmen

- USA:
  Bureau of Economic Analysis (BEA) of the Department of Commerce: "Statistical Abstract of the United States" (jährliche Daten), alle Daten elektronisch verfügbar im Economic Bulletin Board (EBB) CITIBASE (Citicorp Database System mit über 5600 ökonomischen Zeitreihen); Micro Tsp Nutzer haben Zugriff auf die EBB-Daten

- US-Finanzdaten: CRSP Files, ISL, Compustat (vgl. z.B. Gujarati 1995, S.29ff.)

- CIA World Factbook:
  https://www.cia.gov/library/publications/the-world-factbook/

3.2 Ökonometrie/Statistik Programme

Statistics and Econometrics Software for PC and Compatibles and for Apple Macintosh: A Partial List of Products and Vendors

BMDP/PC. BMDP Statistical Software Inc., 1440 Sepulveda Blvd., Suite 316, Los Angeles, CA 90025. (213)-479-7799.

DATA-FIT. Oxford Electronic Publishing, Oxford University Press, Walton Street, Oxford OX2 6DP, UK.

ECONOMIST WORKSTATION. Data Resources/McGraw-Hill, 24 Hartwell Avenue, Lexington, MA 02173, (617)-863-5100.

ESP. Economic Software Package. 76 Bedford St., Suite 33, Lexington, MA 02173, (617)-861-8852.

ET. William H. Greene, Stern Graduate School of Business, New York University, 100 Trinity Place, New York, NY 10006.

EVIEWS, 4521 Campus Drive, #336, Irvine, CA 92612-2621, USA, www.eviews.com

GAUSS. Aptech Systems Inc., 26250 196th Place SE, Kent, WA 98042, (206)-361-6679, www.aptech.com .

LIMDEP. William H. Greene, Stern Graduate School of Business, New York University, 100 Trinity Place, New York, NY 10006, www.limdep.com .

MATLAB. MathWorks Inc., 20 N. Main St., Sherborn,. MA 01770, (617)-653-1415.

MICRO TSP. Quantitative Micro Software, 4521 Campus drive, Suite 336, Irvine, CA 92715, (714)-856-3368.

MINITAB. Minitab, 3081 Enterprise Dr., State College, PA 16801, (814)-238-3280.

PC-GIVE. University of Oxford, Institute of Economics and Statistics, St. Cross Building, Manor rd., Oxford OX1 3UL, UK, (0865)-249631.

PC-TSP. TSP International, P.O. Box 61015, Palo Alto, CA 94306, (415)-326-1927.

R, Günter Faes, Koniferenstrasse 82, 41542 Dormagen, www.r-statistik.de.

RATS. VAR Econometrics, P.O. Box 1818, Evanston, IL 60204-1818, (312)-864-8772, www.estima.com .

SAS/STAT. SAS Institute Inc., P.O. Box 8000, SAS Circle, Cary, NC 27511-8000, (919)-467-8000, www.sas.com .

SHAZAM. Kenneth J. White, Department of Economics, University of British Columbia, Vancouver, BC V6T 1Y2 Canada, (604)-228-5062, shazam.econ.ubc.ca .

SORITEC. The Sorites Group Inc., P.O. Box 2939, 8136 Old Keene Mill Road, Springfield, VA 22152, (703)-569-1400.

SPSS/PC+. SPSS Inc. 444 N. Michigan ave., Chicago, IL. 60611, (312)-329-3600.

SST. Dublin/Rivers Research, 1510 Ontario Ave., Passadena, CA 91103, (818)-577-8361.

STATA. Computing Resource Center, 10801 National Blvd, 3rd Floor, Los Angeles, CA 90064, (800)-STATAPC, in California (213)-470-4341, www.stata.com .

STATGRAPHICS. STSC Inc., 2115 E. Jefferson St, Rockville, MD 20892, (800)-592-0050, in Maryland (301)-984-5123.

STATPRO. Penton Software Inc., 420 Lexington Avenue, Suite 2846, New York, NY 10017. (800)-221-3414, in New York (212)-878-9600.

SYSTAT. Systat Inc. 1800 Sherman Ave., Evaston, IL 60201, (312)-864-5670.

TSP, TSP International, Stanford, Calif., www.tspint1.com .

III EMPIRISCHE REGRESSION: EINFACHREGRESSION

Der zentrale Modellansatz der Empirischen Wirtschaftsforschung/Ökonometrie ist die Regression, die Untersuchung des zielgerichteten Zusammenhangs zwischen erklärenden und zu erklärender Größe eines Problems. Die Regressionsanalyse ist Grundlage weiterer Modelle und wird deshalb in der Vorlesung und in diesem Skript ausführlich behandelt.

Die empirische Regression, also die Regressionsanalyse allein auf der Basis und für die beobachteten Werte einer Teilerhebung/Stichprobe, wird in Kapitel II für die Einfachregression (nur eine erklärende Variable) und in Kapitel IV für die Mehrfachregression (viele erklärende Größen) entwickelt. Nach einführenden, typischen Fragestellungen, die sich mit Hilfe einfacher und mehrfacher Regressionsanalyse untersuchen lassen, werden die Begriffe Regression und Korrelation erläutert. Neben einigen anderen Verfahren eine Ausgleichsgerade zu finden, wird dann für die Einfachregression die Methode der kleinsten Quadrate (MKQ) - oder Ordinary Least Squares (OLS) - vorgestellt, die im Weiteren die zentrale Rolle spielen wird. Zum Schluss wird dann auf die Möglichkeit der Variabelentransformation für eine vereinfachte Berechnung eingegangen.

Im Literaturverzeichnis sind vielfältige Quellen gegeben. An neuerer englischsprachiger ökonometrischer Literatur sind Greene (2008), Wooldridge (2006, 2002) und Studenmund (2006), aus betrieblicher Sicht Anderson et al. (2010) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

Mit dem **FFB e-learning Modul**: Lineare Regression - Deskriptives Modell (www.leuphana.de/ffb) (Merz und Stolze 2010a), eine Einführung in die empirische Regression, und mit dem Modul: Lineare Regression – Stochastisches Modell (Merz und Stolze 2010b) wird jeweils eine audio/visuelles internetbasierte Einführung in das klassische lineare Regressionsmodel (CLR) angeboten.

1 REGRESSION UND KORRELATION

Korrelation:

Aussagen über die Stärke des Zusammenhangs zwischen den Variablen; keine Unterscheidung zwischen abhängigen (endogenen) und unabhängigen (exogenen) Variablen.

Regression:

Aussagen über die funktionale Form der Abhängigkeit zwischen den Variablen; gerichtete, kausale Analyse zwischen abhängigen und unabhängigen Variablen.

$y = f(x)$

y = abhängige Variable, Regressand, erklärte oder endogene Variable

x = unabhängige Variable, Regressor, erklärende oder exogene Variable

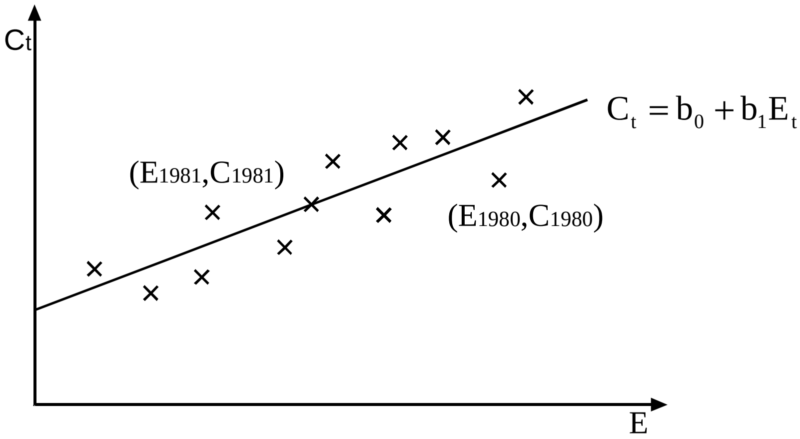
Grundsätzliche Unterscheidung:

Einfachregression: $y = f(x)$

Multiple Regression: $y = f(x_1, x_2, ..., x_K)$

Beispiel: Einfachregression mit Makrodaten (Zeitreihendaten)

$C_t = f(E_t)$ C = Konsum; E = Einkommen; t = Zeitindex



Achtung:

Zeitliche Zuordnung nicht ersichtlich, nicht relevant, relevant ist der inhaltliche Zusammenhang zwischen Konsum und Einkommen.

Ziel:

Grundtendenz bestimmen über die zu schätzenden Parameter b_0, b_1 aus

$\hat{C}_t = b_0 + b_1 E_t$ b_0 = Konstante, absolutes Glied, y - Achsenabschnitt

linearer Ansatz b_1 = Steigung der Geraden

Beispiel: Einfachregression mit Mikrodaten (Querschnittsdaten)

Das Engel'sche Gesetz:

Mit steigendem Einkommen wachsen zwar die Ausgaben für Nahrungsmittel im Durchschnitt, relativ zum Einkommen nehmen die Ausgaben jedoch ab.

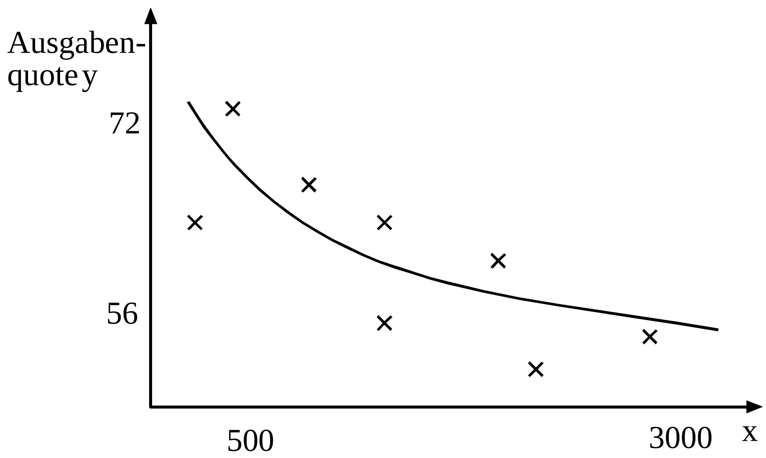
Belgische Arbeiterfamilien (1853)

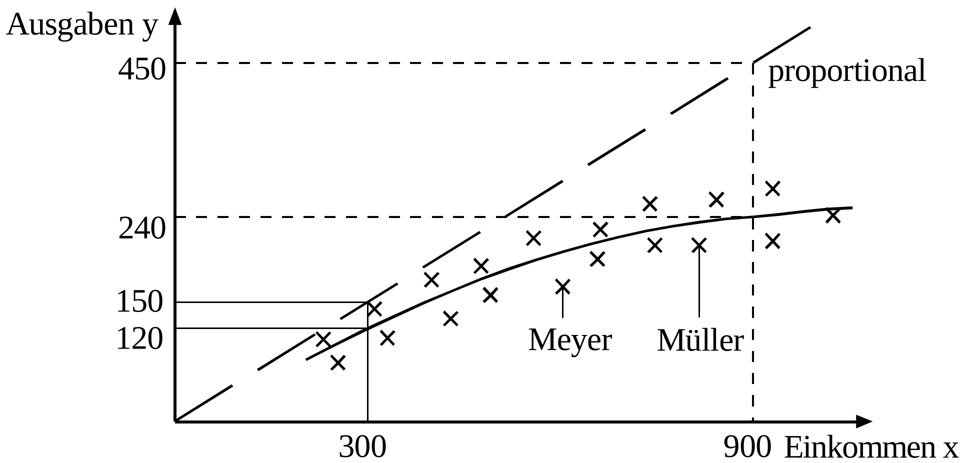
Ausgaben für Nahrungsmittel$_i$ = f(Einkommen)

i=(1,2,...,n) Haushalte

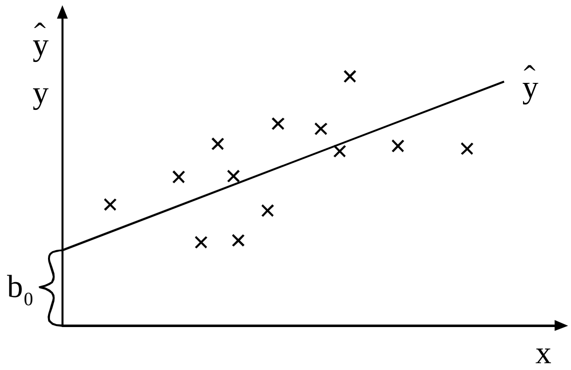
Fiktive Querschnittsdaten:

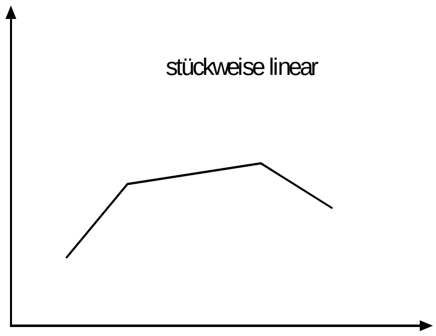
Einkommen x	**∅ Ausgaben in %** y
200	72,9
300	71,4
...	...
3000	56,9





2 DIE METHODE DER KLEINSTEN QUADRATE (MKQ/ OLS) – EINFACHREGRESSION





$$\boxed{\hat{y} = b_0 + b_1 x}$$

bezüglich Einzelwerte $(x_i; y_i)$ (i=1,...,n)

$$\boxed{\hat{y}_i = b_0 + b_1 x_i}$$

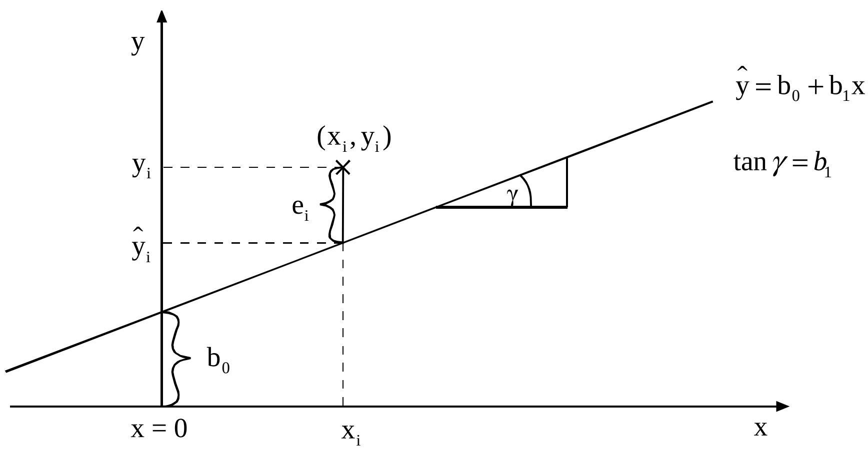
b_0, b_1: Regressionskoeffizienten

b_0: y-Achsenabschnitt, absolutes Glied, intercept, Ordinatenabschnitt

b_1: Steigung (tan α zwischen Gerade und x-Achse)

y_i: beobachtete/empirische Werte

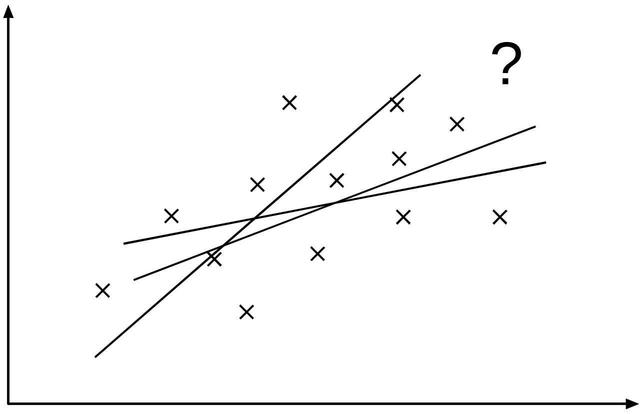
$\hat{y}_i$: Schätzwerte für y_i



Residuen: Abweichungen e_i zwischen beobachteten y_i und geschätzten Werten $\hat{y}_i$

$$e_i = y_i - \hat{y}_i \qquad \text{(i=1,...,n)}$$

2.1 Abweichungskriterien: LAD und MKQ/ OLS



1. Möglichkeit: „Minimum" der Summe der Abweichungen

Summe der einfachen vertikalen Abstände = 0

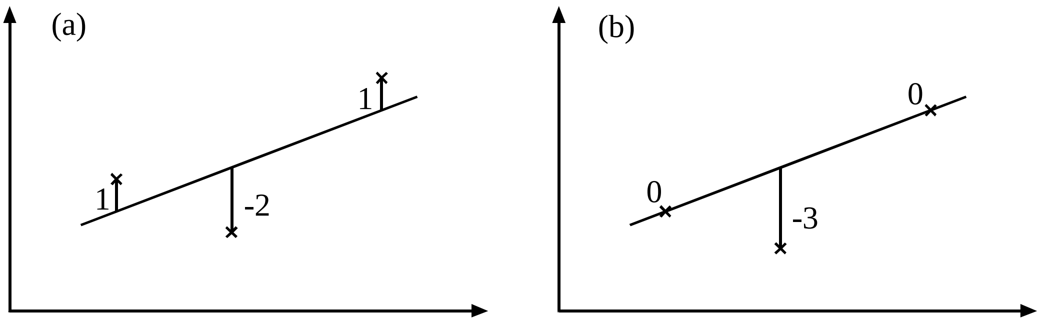
$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right) = \min! = 0 \qquad \text{n = Anzahl der Beobachtungen}$$

Keine eindeutige Regressionsgerade, da positive und negative Abstände sich (bei bestimmten Eigenschaften von $\hat{y}$) gegenseitig aufheben.

2. Möglichkeit: Minimum der Summe der absoluten Abweichungen - LAD (Least Absolute Deviation)

$$\sum_{i=1}^{n} |e_i| = \sum_{i=1}^{n} \left|y_i - \hat{y}_i\right| \rightarrow \text{minimieren} \qquad y_i = \text{beob. Wert;}\ \hat{y}_i = \text{geschätzter Wert}$$

Nicht ideal, da extreme Abweichungen nicht ausreichend berücksichtigt werden.



$$\sum |e_i| = 4 \qquad\qquad \sum |e_i| = 3$$

(a) ist ‚ausgleichsgünstiger', obwohl (b) eine kleinere Summe aufweist ist

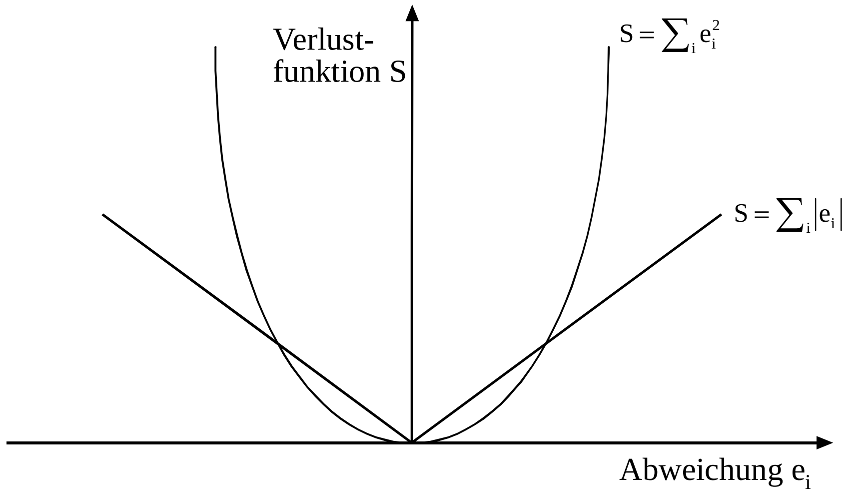
3. Möglichkeit: Minimum der Summe der quadrierten Abstände - MKQ (Methode der kleinsten Quadrate) = OLS (Ordinary Least Squares)

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \rightarrow \text{minimieren}$$

$$(a): \sum_{i=1}^{n} e_i^2 = 1^2 + 2^2 + 1^2 = 6$$

$$(b): \sum_{i=1}^{n} e_i^2 = 0^2 + 3^2 + 0^2 = 9$$

Vorteil: Größere Abweichungen werden bei MKQ/ OLS stärker berücksichtigt als bei LAD.



MKQ/ OLS ist die zentrale Methode in der Ökonometrie

2.2 Parameterschätzung nach der Methode der kleinsten Quadrate

Wooldridge 2006, Greene 2003, Gujarati 1995, Hübler 2005, 1989, Frohn 1980

I	$\hat{y}_i$	$= b_0 + b_1 x_i$	Schätzwerte $(i = 1, 2, ..., n)$
II	y_i	$= b_0 + b_1 x_i + e_i$	beobachtete tatsächliche Werte
II – I	$y_i - \hat{y}_i$	$= e_i$	
	$y_i - b_0 - b_1 x_i$	$= e_i$	

Gesucht:

b_0, b_1 so, das

$$S = \sum_{i=1}^{n} e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - b_0 - b_1 x_i)^2 = \min!$$

Lösung des Extremwertproblem durch MKQ, OLS:

Extremamöglichkeit aus den nullgesetzten ersten Ableitungen (Notwendige Bedingungen):

$$\frac{\partial S}{\partial b_0} = \frac{\partial S}{\partial b_1} = 0$$

$$(1) \quad \frac{\partial S}{\partial b_0} = \sum_{i=1}^{n} \left[2(y_i - b_0 - b_1 x_i)^{2-1}(-1) \right] = -2 \sum (y_i - b_0 - b_1 x_i)$$

$$(2) \quad \frac{\partial S}{\partial b_1} = \sum_{i=1}^{n} \left[2(y_i - b_0 - b_1 x_i)^{2-1}(-x_i) \right] = -2 \sum x_i (y_i - b_0 - b_1 x_i)$$

Nullsetzen von (1) und (2):

$$(1)\quad \frac{\partial S}{\partial b_0} = -2\sum (y_i - b_0 - b_1 x_i) = 0$$
$$\Leftrightarrow \sum y_i - nb_0 - \sum b_1 x_i = 0$$
$$\Leftrightarrow nb_0 + \sum b_1 x_i = \sum y_i$$

$$(2)\quad \frac{\partial S}{\partial b_1} = -2\sum x_i (y_i - b_0 - b_1 x_i) = 0$$
$$\Leftrightarrow \sum x_i y_i - \sum b_0 x_i - \sum b_1 x_i^2 = 0$$
$$\Leftrightarrow \sum b_0 x_i + \sum b_1 x_i^2 = \sum x_i y_i$$

$$\boxed{\left.\begin{array}{ll} (3) & nb_0 + b_1 \sum x_i = \sum y_i \\ (4) & b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i \end{array}\right\} \text{Normalgleichungen}}$$

$$\boxed{(5)(3)\quad b_0 = \frac{1}{n}\sum y_i - b_1 \frac{1}{n}\sum x_i = \bar{y} - b_1 \bar{x}}$$

Einsetzen in (4):

$$\left(\frac{1}{n}\sum y_i - b_1 \frac{1}{n}\sum x_i\right)\sum x_i + b_1 \sum x_i^2 = \sum x_i y_i$$
$$\frac{1}{n}\sum x_i \sum y_i - b_1 \frac{1}{n}\left(\sum x_i\right)^2 + b_1 \sum x_i^2 = \sum x_i y_i$$
$$b_1\left(\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right) = \sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i$$

$$\boxed{(6)\quad b_1 = \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\mathrm{Cov}(x,y)}{\mathrm{Var}(x)}}$$

Erst b_1 berechnen, dann b_1 in (3) einsetzen und b_0 berechnen.

$$b_0 = \frac{1}{n}\sum y_i - \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} \cdot \frac{1}{n}\sum x_i$$
$$= \frac{1}{n}\left[\frac{\sum y_i \cdot \left(n\sum x_i^2 - \left(\sum x_i\right)^2\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2} - \frac{n\sum x_i \sum x_i y_i - \sum x_i \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}\right]$$
$$= \frac{1}{n}\left[\frac{n\sum x_i^2 \sum y_i - \left(\sum x_i\right)^2 \sum y_i - n\sum x_i \sum x_i y_i + \left(\sum x_i\right)^2 \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}\right]$$

$$\boxed{(7)\quad b_0 = \frac{1}{n}\left[\frac{n\sum x_i^2 \sum y_i - n\sum x_i \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}\right] = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}}$$

Dass b_0, b_1 tatsächlich $S(b_0,b_1)$ minimiert kann anhand der höheren Ableitungen gezeigt werden (vgl. Kap. 4.2 für den allgemeinen K-Variablen-Fall).

Beispiel: Einkommensregression, Minimierung der Fehlerquadratsumme

Eine zu einem bestimmten Zeitpunkt durchgeführte Umfrage (Querschnitt) liefere folgende Mikrodaten mit monatlichem Nettoeinkommen (income) (y_i) und Alter der Personen (age) (x_i) in Jahren (n = 6).

Tab. IV.5: Mikrodaten zu Nettoeinkommen und Alter der Personen

i	Daten y_i income	x_i age	Arbeitstabelle x_i^2	$x_i \cdot y_i$	y_i^2
1	1200	22	484	26400	1440000
2	1700	24	576	40800	2890000
3	3500	28	784	98000	12250000
4	4200	27	729	113400	17640000
5	1600	23	529	36800	2560000
6	5200	36	1296	187200	27040000
Σ	17400	160	4398	502600	63820000

Gesucht: income = b0 + b1*age

$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 26,66$ Jahre $\qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 2900(\,mtl.)$

$$b_1 = \frac{\frac{1}{n}\sum y_i \cdot x_i - \overline{y}\cdot\overline{x}}{\frac{1}{n}\sum x_i^2 - \overline{x}^2} = \frac{\frac{1}{6}\cdot 502600 - 2900\cdot 26,66}{\frac{1}{6}\cdot 4398 - 26,66^2} = \frac{6433,3353}{21,889} = 293,91$$

$b_0 = \overline{y} - b_1 \cdot \overline{x} = 2900 - 293,91 \cdot 26,666 = -4937,40406$ (mit gerundeten Werten)

Regressionsgerade: $\hat{y}_i = -4937,40 + 293,91 x_i$

Interpretation: Bei einer Zunahme des Alters x um 1 Jahr steigt das Einkommen im Durchschnitt um 293,91 Euro (im Wertebereich 20-44 Jahre).

```
============================================================================
Ordinary    Least Squares
Dependent Variable           INCOME     Number of Observations          6
Mean of Dep. Variable     2900.0000     Std. Dev. of Dep. Var.   1634.625339
Std. Error of Regr.        709.7758     Sum of Squared Residuals  201513E+07
R - squared                  .84917     Adjusted R - squared           .81146
F( 1,   4)                  22.5194     Prob. Value for F              .00900
============================================================================
Variable Coefficient Std. Error t-ratio Prob|t|>x    Mean of X Std.Dev.of X
----------------------------------------------------------------------------
Constant   -4937.56      1677.  -2.945   .04220
AGE          293.909      61.93   4.745   .00900       26.66667      5.12510
```

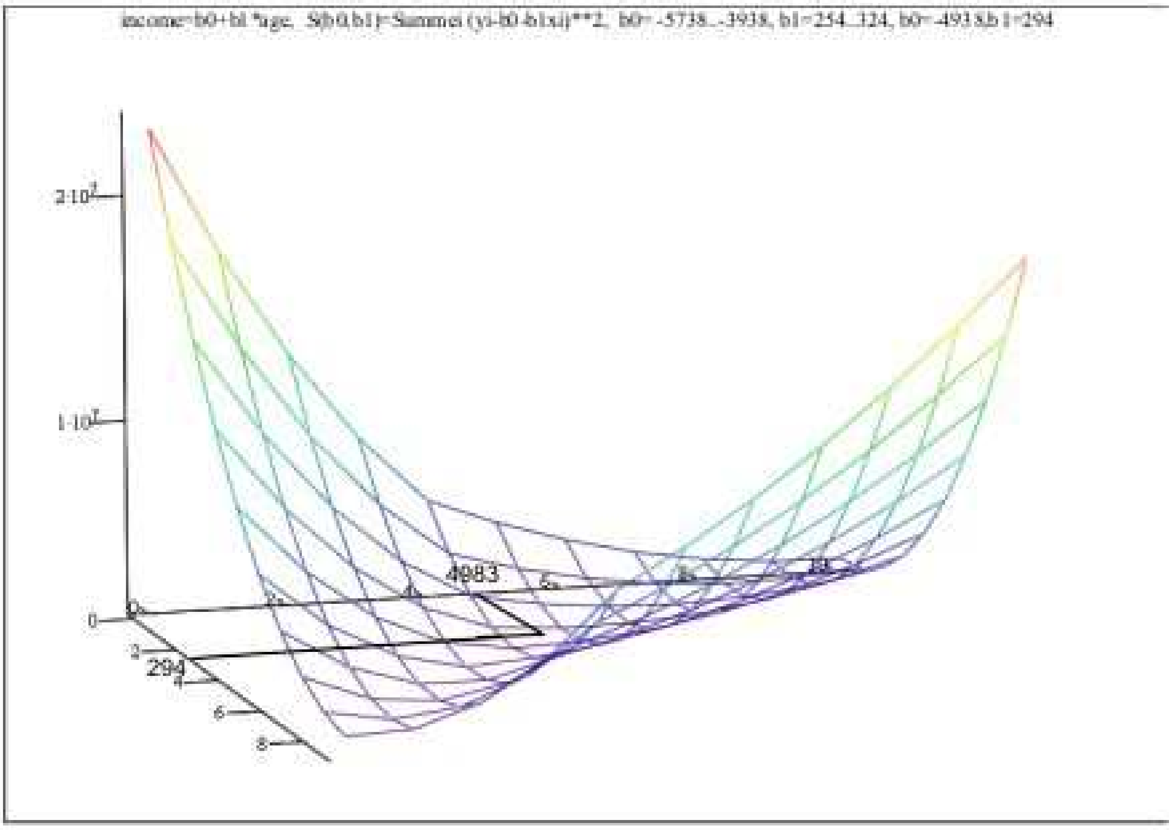
Regressionsgerade: $\hat{y}_i = -4937{,}56 + 293{,}91x_i$

Alternativ: Direkte Minimierung der Fehlerquadratsumme

MKQ/OLS-Aufgabe: $S(b_0,b_1)=\sum_i (y_i - b_0 - b_1 x_i)^2 = \min!$

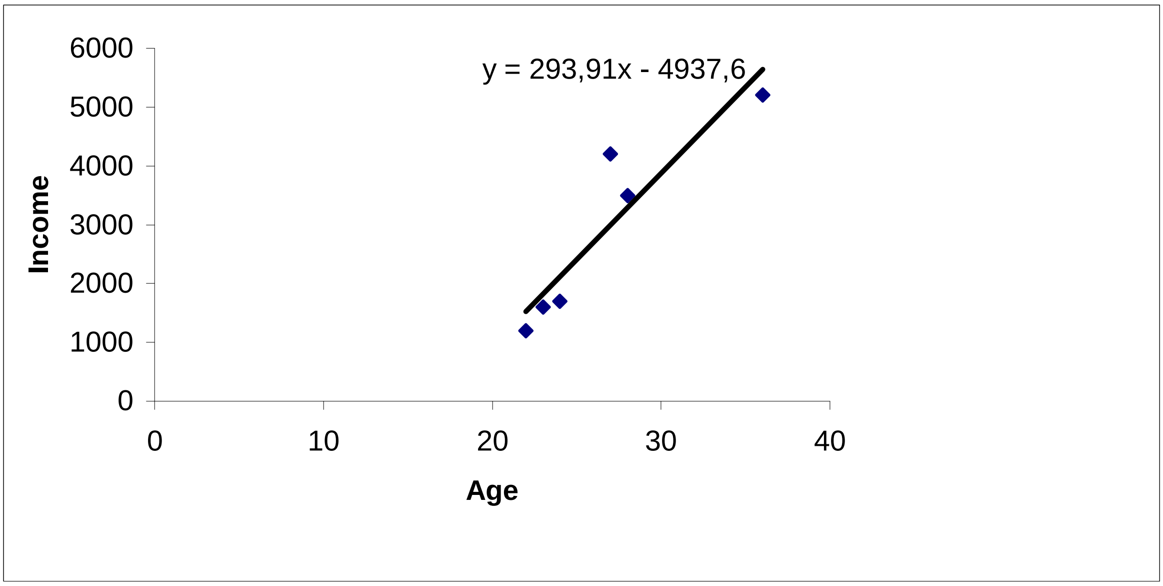
Minimierung von $S(b_0,b_1) = f(b_0 = -5738 \ldots -3938, \quad b_1 = 254 \ldots 354)$

$$
\begin{array}{cc}
 & \begin{array}{ccccccccccc} 254 & & & & \overset{b_1}{294} & & & & & 344 & 354 \end{array} \\
\text{smp} := & \left(\begin{array}{ccccccccccc}
23086192 & 17454232 & 12701872 & 8829112 & 5835952 & 4669222 & 3722392 & 2995462 & 2488432 & 2201302 & 2134072 \\
18850992 & 13859032 & 9746672 & 6513912 & 4160752 & 3314022 & 2687192 & 2280262 & 2093232 & 2126102 & 2378872 \\
15095792 & 10743832 & 7271472 & 4678712 & 2965552 & 2438822 & 2131992 & 2045062 & 2178032 & 2530902 & 3103672 \\
11820592 & 8108632 & 5276272 & 3323512 & 2250352 & 2043622 & 2056792 & 2289862 & 2742832 & 3415702 & 4308472 \\
9025392 & 5953432 & 3761072 & 2448312 & \boxed{2015152} & 2128422 & 2461592 & 3014662 & 3787632 & 4780502 & 5993272 \\
6710192 & 4278232 & 2725872 & 2053112 & 2259952 & 2693222 & 3346392 & 4219462 & 5312432 & 6625302 & 8158072 \\
4874992 & 3083032 & 2170672 & 2137912 & 2984752 & 3738022 & 4711192 & 5904262 & 7317232 & 8950102 & 10802872 \\
3519792 & 2367832 & 2095472 & 2702712 & 4189552 & 5262822 & 6555992 & 8069062 & 9802032 & 11754902 & 13927672 \\
2644592 & 2132632 & 2500272 & 3747512 & 5874352 & 7267622 & 8880792 & 10713862 & 12766832 & 15039702 & 17532472 \\
2249392 & 2377432 & 3385072 & 5272312 & 8039152 & 9752422 & 11685592 & 13838662 & 16211632 & 18804502 & 21617272
\end{array}\right)
\end{array}
\begin{array}{l} -5738 \\ \\ \\ \\ -4938 \;\; b_0 \\ \\ \\ \\ \\ -3938 \end{array}
$$



Minimum der Fehlerquadratsumme also bei b_0=-4938 und b_1=294 und damit

Income = -4938 + 294 age



Beispiel Einfachregression: Jahresumsatz und Ladenverkaufsfläche

Ein Filialunternehmen will den Zusammenhang zwischen Jahresumsatz und Ladenverkaufsfläche untersuchen. In einem bestimmten Jahr liefern die n = 12 Filialen folgende Daten:

Tabelle 1: Jahresumsatz und Verkaufsfläche von 12 Filialen

Filiale [i]	Verkaufsfläche (in Tsd. qm) [x_i]	Jahresumsatz (in Mio. €) [y_i]
1	0,31	2,93
2	0,98	5,27
3	1,21	6,85
4	1,29	7,01
5	1,12	7,02
6	1,49	8,35
7	0,78	4,33
8	0,94	5,77
9	1,29	7,68
10	0,48	3,16
11	0,24	1,52
12	0,55	3,15

Trägt man die Beobachtungspunkte in ein xy-Koordinatensystem ein, so erkennt man, dass hier der Zusammenhang zwischen den beiden Variablen recht gut durch eine lineare Funktion wiedergegeben werden kann (siehe Abbildung).

Die zur Berechnung der Regressionskoeffizienten notwendigen Summen lassen sich am besten anhand der folgenden Arbeitstabelle Tabelle 2 bestimmen.

Tabelle 2: Arbeitstabelle: Jahresumsatz und Verkaufsfläche von 12 Filialen

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	0,31	2,93	0,0961	8,5849	0,9083
2	0,98	5,27	0,9604	27,7729	5,1646
3	1,21	6,85	1,4641	46,9225	8,2885
4	1,29	7,01	1,6641	49,1404	9,0429
5	1,12	7,02	1,2544	49,2804	7,8624
6	1,49	8,35	2,2201	69,7225	12,4415
7	0,78	4,33	0,6084	18,7489	3,3774
8	0,94	5,77	0,8836	33,2929	5,4238
9	1,29	7,68	1,6641	58,9824	9,9072
10	0,48	3,16	0,2304	9,9856	1,5168
11	0,24	1,52	0,0576	2,3104	0,3648
12	0,55	3,15	0,3025	9,9225	1,7325
Σ	10,68	63,04	11,4058	384,6660	66,0307

Für unser Beispiel erhalten wir also:

$$N = 12$$
$$\sum x_i = 10,68; \quad \sum x_i^2 = 11,4058; \quad \sum y_i = 63,04$$
$$\sum y_i^2 = 384,660; \quad \sum x_i y_i = 66,0307$$

Damit findet man die Kleinste-Quadrate-Regressionskoeffizienten:

$$\begin{aligned} b_0 &= \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} \\ &= \frac{11,4058 \cdot 63,04 - 10,68 \cdot 66,0307}{12 \cdot 11,4058 - 10,68 \cdot 10,68} \\ &= \frac{13,813756}{22,8072} \\ &= 0,605675 \approx 0,6057 \end{aligned}$$

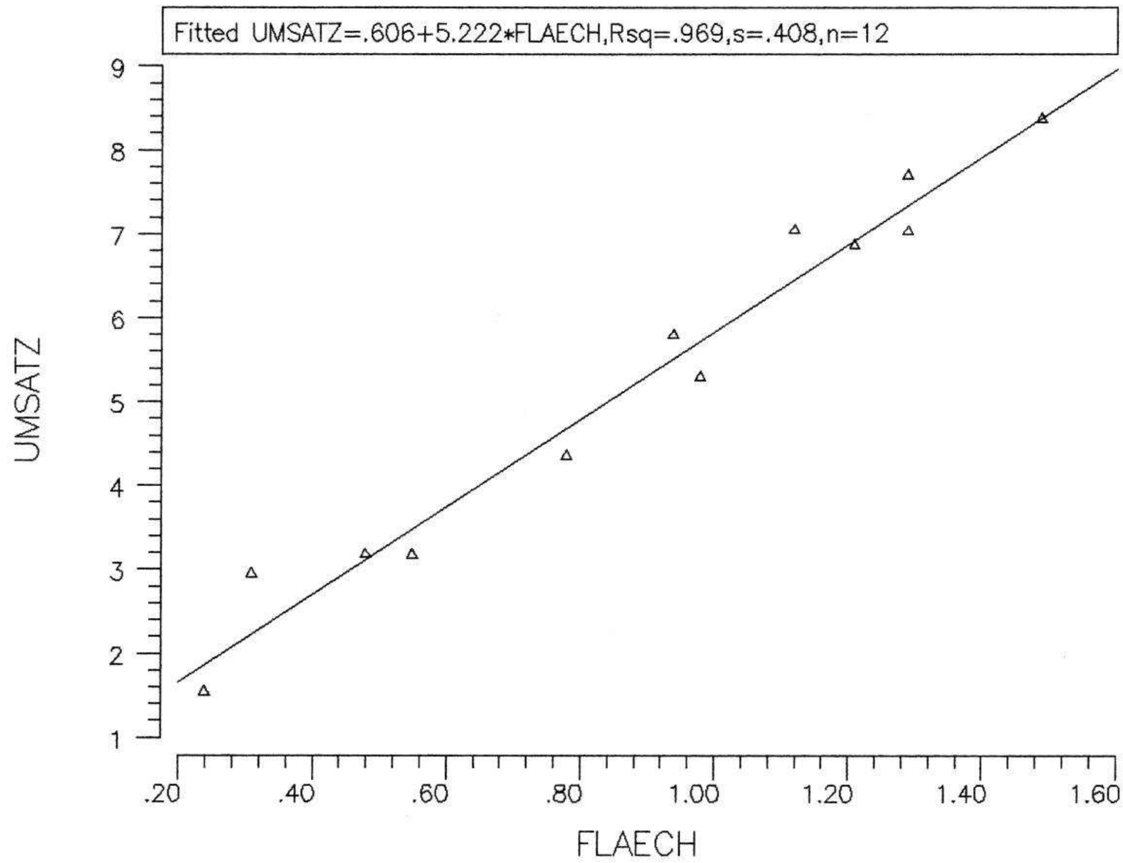
und

$$\begin{aligned} b_1 &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \\ &= \frac{12 \cdot 66,0307 - 10,68 \cdot 63,04}{12 \cdot 11,4058 - 10,68 \cdot 10,68} \\ &= \frac{119,1012}{22,8072} \\ &= 5,222088 \approx 5,2221 \end{aligned}$$

Die Kleiste-Quadrate-Regressionsfunktion lautet demnach

$$\hat{y} = 0,6057 + 5,2221 x_i$$

Die Regressionsfunktion ist in der folgenden Abbildung dargestellt:



2.3 Vereinfachungen durch Koordinatentransformation

Für eine Vereinfachung der Berechnungen oder wenn bereits mittelwerttransformierte Daten vorliegen, können OLS-Ergebnisse durch einfache bzw. zweifache Koordinatentransformation gefunden werden.

Einfache Koordinatentransformation

Verschieben des Koordinatensystems um das arithmetische Mittel $\bar{x} = \frac{1}{n}\sum x_i$

$$\text{sei } x_i^* = x_i - \bar{x} \Rightarrow x_i = x_i^* + \bar{x}$$

Regressionsfunktion:

$$\hat{y}_i = b_0 + b_1 x_i$$

$$\hat{y}_i = b_0 + b_1(x_i^* + \bar{x})$$

$$\hat{y}_i = \underbrace{b_0 + b_1\bar{x}}_{b_0^*} + b_1 x_i^*$$

$$\hat{y}_i = b_0^* + b_1 x_i^* \qquad \text{im } (x^*, y) - \text{Koordinatensystem}$$

Einsetzen der "Sternchenwerte" in die 1. Normalgleichung:

$$(3^*) \quad nb_0^* + b_1 \underbrace{\sum x_i^*}_{0} = \sum y_i$$

$$nb_0^* = \sum y_i$$

$$\boxed{b_0^* = \tfrac{1}{n}\sum y_i = \overline{y}}$$

Nulleigenschaft des arithm. Mittels:

$$\sum x_i^* = \sum(x_i - \overline{x}) = \sum x_i - \sum \overline{x}$$

$$= n\overline{x} - n\overline{x} = 0$$

$$(4^*) \quad b_0^* \underbrace{\sum x_i^*}_{0} + b_1 \sum x_i^{*2} = \sum x_i^* y_i$$

$$\boxed{b_1 = \frac{\sum x_i^* y_i}{\sum x_i^{*2}}}$$

Rücktransformation für b_0:

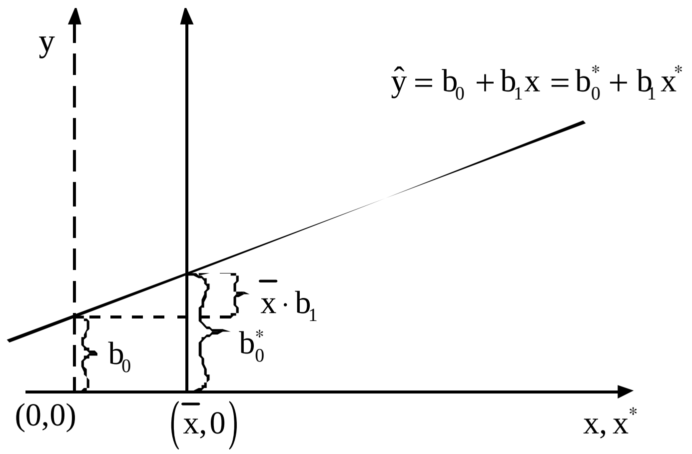
$$b_0^* = b_0 + b_1\overline{x}$$

$$\Leftrightarrow \quad b_0 = b_0^* - b_1\overline{x} = \overline{y} - b_1\overline{x}$$

Graphische Darstellung im (x*, y) – Koordinatensystem:

$$\hat{y} = b_0 + b_1 x = b_0^* + b_1 x^*$$

gleiche Steigung



4. Zweifache Koordinatentransformation

Verschieben des Koordinatensystems um die arithmetischen Mittel: $\overline{x} = \frac{1}{n}\sum x_i$; $\overline{y} = \frac{1}{n}\sum y_i$

sei $\quad x_i^* = x_i - \overline{x} \Rightarrow x_i = x_i^* + \overline{x}$

$\quad\quad y_i^* = y_i - \overline{y} \Rightarrow y_i = y_i^* + \overline{y}$

Regressionsfunktion:

$$\hat{y}_i = b_0 + b_1 x_i$$

$$\hat{y}_i^* + \overline{y} = b_0 + b_1(x_i^* + \overline{x})$$

$$\hat{y}_i^* = \underbrace{b_0 + b_1\overline{x}}_{b_0^*} - \overline{y} + b_1 x_i^* = \underbrace{b_0^* - \overline{y}}_{0\,[\text{vgl.}(3^*)]} + b_1 x_i^*$$

$$\hat{y}_i^* = b_1 x_i^*$$

im (x^*, y^*) – Koordinatensystem

Einsetzen der "Sternchenwerte" in die 1. Normalgleichung:

$$(4^*) \quad b_0^* \sum x_i^* + b_1 \sum x_i^{*2} = \sum x_i^* y_i^* \qquad \Big| \sum x_i^* = 0$$

$$\boxed{b_1 = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}}}$$

Rücktransformation für b_0:

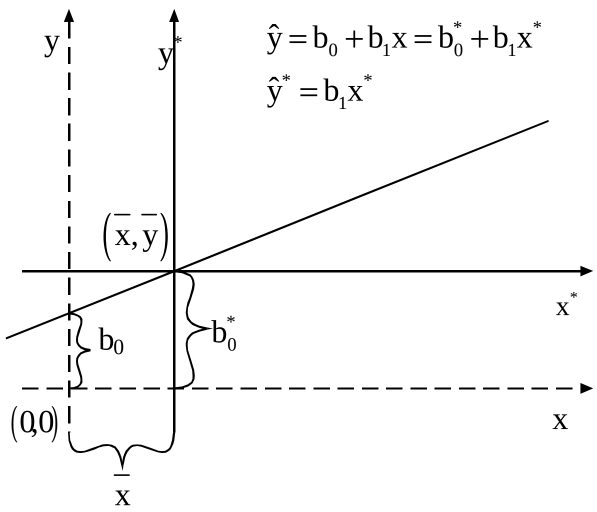
$$b_0^* = b_0 + b_1 \bar{x}$$

$$\Leftrightarrow \quad b_0 = b_0^* - b_1 \bar{x} = \bar{y} - b_1 \bar{x}$$

Graphische Darstellung im (x*, y*)-Koordinatensystem

$$\hat{y} = b_0 + b_1 x = b_0^* + b_1 x^*$$

$\hat{y}^* = b_1 x^*$ gleiche Steigung



2.4 Zusammenfassung: Empirische Regression – Einfachregression

Allgemein:

$$\hat{y}_i = b_0 + b_1 x_i$$

$$\boxed{\begin{aligned} b_0 &= \overline{y} - b_1\overline{x} \\ b_1 &= \frac{\sum x_i y_i - n\overline{x}\overline{y}}{\sum x_i^2 - n\overline{x}^2} = \frac{\mathrm{Cov}(x,y)}{\mathrm{Var}(x)} \end{aligned}}$$

Einfache Transformation:

$$\hat{y}_i = b_0^* + b_1 x_i^*$$

$$\boxed{\begin{aligned} b_0^* &= \overline{y} \\ b_1 &= \frac{\sum x_i^* y_i}{\sum x_i^{*2}} \end{aligned}}$$

$$b_0 = \overline{y} - b_1\overline{x}$$

Rücktransformation:

$$b_0^* = b_0 + b_1\overline{x}$$

$$\Leftrightarrow \quad b_0 = b_0^* - b_1\overline{x} = \overline{y} - b_1\overline{x}$$

Zweifache Transformation:

$$x_i^* = x_i - \overline{x} \qquad y_i^* = y_i - \overline{y}$$

$$\hat{y}_i^* = b_1 x_i^*$$

$$\boxed{b_1 = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}}}$$

$$b_0 = \overline{y} - b_1\overline{x}$$

2.5 Beispiele BWL/VWL Einfachregression und Variablentransformation

Beispiel aus der BWL / Variablentransformation

Die Montage der Produktion, die in großen Serien gefertigt werden, etwa Motoren, findet an einem Montageband statt. Werden bei der Produktion Typenänderungen am produzierten Modell vorgenommen, so entstehen zunächst Zeitverluste, weil der neue Typ von den Arbeitern nicht so schnell montiert werden kann. Es wird eine geraume „**Anpassungszeit**“ verstreichen, bis die Motoren wieder im üblichen Tempo produziert werden können. Die Dauer der „Anpassungszeit „ kann nun durch verschiedene Faktoren beeinflusst werden, etwa von der Komplexität des neuen Typs, von seiner „Fabrikationsreife“ und möglicherweise von der Taktzeit, d.h. der Zeit, die den Arbeiter unter normalen Verhältnissen für jeden Arbeitsgang vorgegeben wird. Es soll nun untersucht werden, inwieweit die Taktzeit einen Einfluss auf die „Anpassungszeit“ ausübt. Um dieses zu untersuchen, müssen Daten gesammelt werden. Folgende Daten seien beobachtet worden:

Taktzeit (in Minuten)	x_i	1,0	1,2	1,7	1,8	2,1	2,4	2,6	2,8	3,0
Anpassungszeit (in Wochen)	y_i	2	5	4	7	6	8	6	7	10

Es sollen nun die Regressionsparameter für das Beispiel der Abhängigkeit der Anpassungszeit von der Taktzeit numerisch berechnet werden. Dazu werden die Beobachtungswerte zweckmäßig in einer Tabelle der folgenden Form angeordnet.

Rechentabelle zur Bestimmung der Regressionskoeffizienten für die Abhängigkeit der „Anpassungszeit“ y von der Taktzeit x

	x_i	x_i^*	$(x_i^*)^2$	y_i	y_i^*	$(y_i^*)^2$	$x_i^* y_i^*$
	1,0	-1,07	1,145	2	-4,1	16,81	4,387
	1,2	-0,87	0,757	5	-1,1	1,21	0,957
	1,7	-0,37	0,137	4	-2,1	4,41	0,777
	1,8	-0,27	0,073	7	0,9	0,81	-0,243
	2,1	0,03	0,001	6	-0,1	0,01	-0,003
	2,4	0,33	0,109	8	1,9	3,61	0,627
	2,6	0,53	0,281	6	-0,1	0,01	-0,053
	2,8	0,73	0,533	7	0,9	0,81	0,657
	3,0	0,93	0,865	10	3,9	15,21	3,627
Σ	18,6	-0,03	3,901	55	0,1	42,89	10,733

$\overline{x} = 2,07$ $\qquad \overline{y} = 6,11$

$x_i^* = x_i - \overline{x}$ $\qquad y_i^* = y_i - \overline{y}$

$\overline{x} = 2,07 \qquad x_i^* = x_i - \overline{x}$

Regressionsansatz einfache Variablentransformation:

$$\hat{y}_i = b_0^* + b_1 x_i^*$$

$$\boxed{\begin{aligned} b_0^* &= \overline{y} \\ b_1 &= \frac{\sum x_i^* y_i}{\sum x_i^{*2}} \end{aligned}}$$

$$b_0 = \overline{y} - b_1 \overline{x}$$

Man erhält mit zweifacher Variablentransformation:

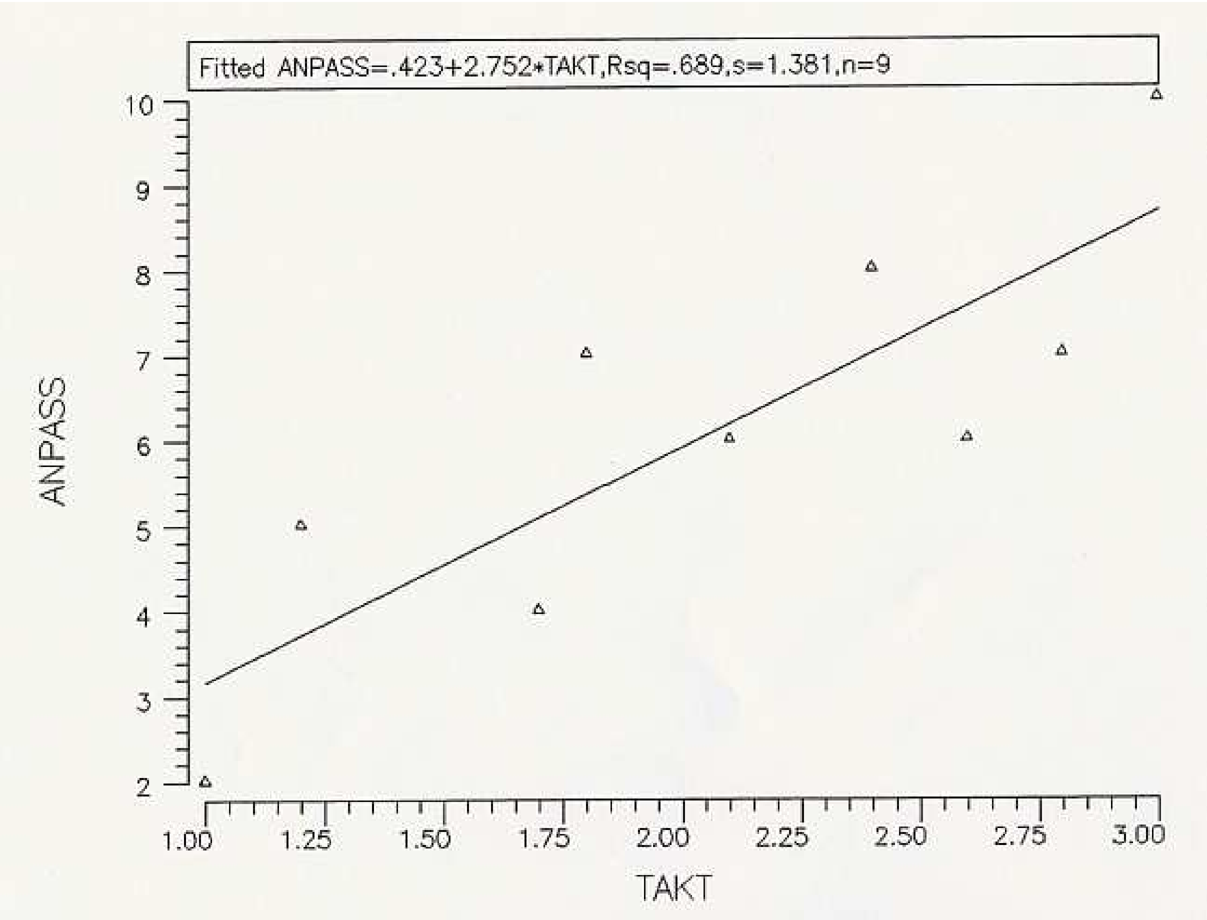
$$b_1 = \frac{\sum x_i^* y_i^*}{\sum \left(x_i^*\right)^2} = \frac{10,733}{3,901} = 2,75$$

$$b_0 = \overline{y} - b_1 \overline{x} = 6,11 - 2,75 \cdot 2,07 = 6,11 - 5,69 = 0,42$$

Damit lautet die gesuchte Regressionsgerade

$\hat{y} = 0,42 + 2,75x$ im (x,y)-Koordinatensystem der Beobachtungswerte, und

$\hat{y}^* = 2,75x^*$ im (x^*, y^*)- Koordinatensystem der zentrierten Werte.



Beispiel aus der VWL / Variablentransformation

Anhand von Daten für die Bundesrepublik Deutschland aus den Jahren 1961 bis 1972 soll eine Regression für den **privaten Verbrauch** (in Preisen von 1962) in Abhängigkeit vom verfügbaren Einkommen der privaten Haushalte bestimmt werden.

Die Daten aus der folgenden Tabelle wurden dem Jahresgutachten 1973 des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung entnommen (Drucksache 7/1273 des Deutschen Bundestages, 7. Wahlperiode, Tabelle 20 und 22)

Rechentabelle zur Bestimmung der Regressionskoeffizienten für die Abhängigkeit des privaten Konsums C von dem verfügbaren privatem Einkommen Y_{pr}^{V} (in Mrd. DM)

	Y_{pr}^{V}	Y^*	$(Y^*)^2$	C	C^*	$(C^*)^2$	C*Y*
1961	219,99	-125,88	15845,15	194,70	-59,35	3522,52	7470,93
1962	236,07	-109,80	12055,49	205,20	-48,85	2386,40	5363,70
1963	251,41	-94,46	8922,22	212,53	-41,52	1723,98	3921,95
1964	236,16	-109,71	12035,74	223,51	-30,54	932,74	3350,56
1965	307,70	-38,17	1456,76	238,50	-15,55	241,83	593,54
1966	325,57	-20,30	411,99	247,39	-6,66	44,37	135,20
1967	330,97	-14,90	221,94	249,55	-4,50	20,26	67,05
1968	399,73	53,86	2901,17	258,98	4,93	24,30	265,50
1969	390,73	44,86	2012,64	279,00	24,95	622,46	1119,28
1970	443,08	97,21	9450,27	298,45	44,40	1971,29	4316,15
1971	481,68	135,81	18445,04	315,05	61,00	3720,90	8284,45
1972	527,32	181,45	32925,01	325,75	71,70	5140,77	13009,99
Σ	4150,4	0,00	116683,40	3048,61	0,00	20351,81	47898,31

$\overline{Y} = 345,87 \qquad \overline{C} = 254,05$

Regressionsansatz zweifache Variablentransformation:

$$\hat{y}_i^* = b_1 x_i^*$$

$$\boxed{b_1 = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}}}$$

$$b_0 = \overline{y} - b_1 \overline{x}$$

Man erhält:

$$\overline{x} = 345,87 \qquad \overline{y} = 254,05$$

$$x_i^* = x_i - \overline{x} \qquad y_i^* = y_i - \overline{y}$$

$$b_1 = \frac{\sum x_i^* y_i^*}{\sum \left(x_i^*\right)^2} = \frac{47898,31}{116683,40} = 0,41$$

$$b_0 = \overline{y} - b_1 \overline{x} = 254,05 - 0,41 \quad 345,87 = 112,24$$

Damit lautet die gesuchte Regressionsgerade mit

$$b_0^* = \overline{y} = 254,05$$

$$b_1 = \quad 0,41$$

$$b_0 = 112,24$$

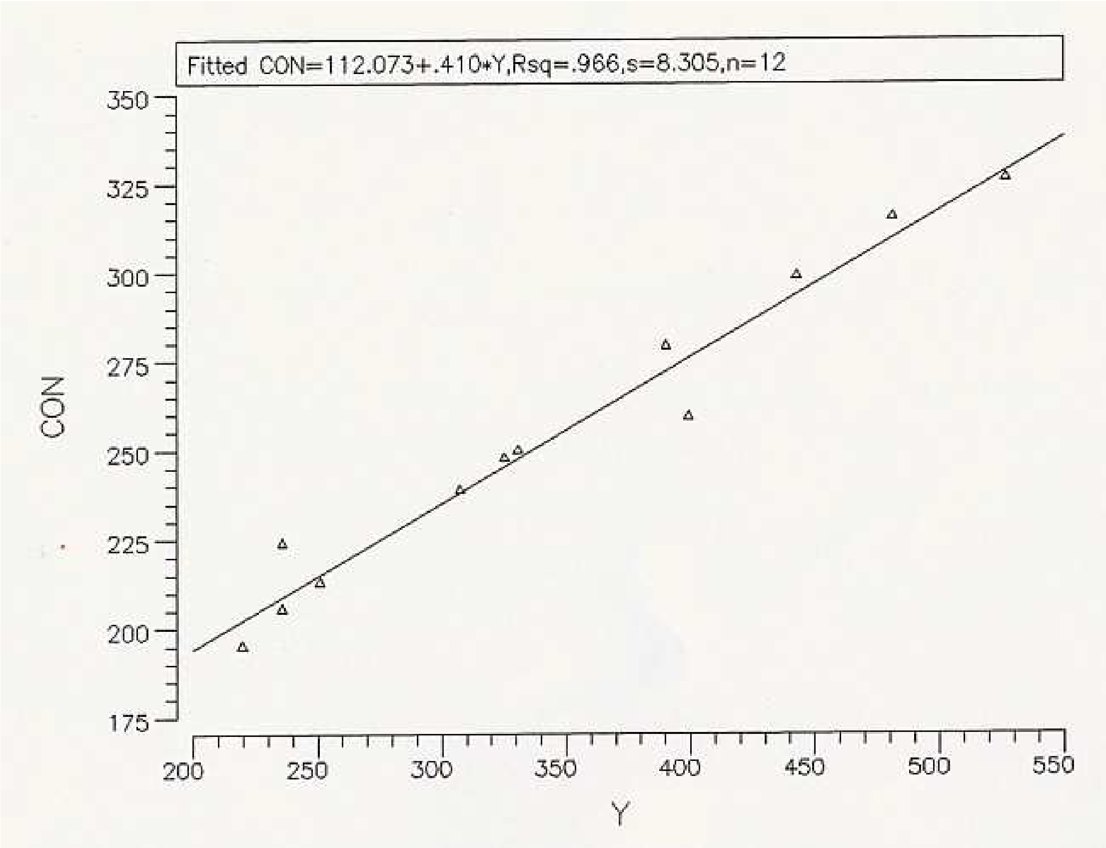
C = 112,24 + 0,41 Y_{pr}^V im (Y,C)- Koordinatensystem der Beobachtungswerte.

C= 112,24 + 0,41 Y_{pr}^V im (x,y)-Koordinatensystem der Beobachtungswerte, und

C^*= 0,41 Y^* im (x^*, y^*)- Koordinatensystem der zentrierten Werte.

Der Wert b_1=0,41 gibt die Steigung der Regressionsgeraden an, er kann wie folgt gedeutet werden: Wenn das private verfügbare Einkommen um eine Einheit (= 1 Mrd. DM) erhöht (bzw. gesenkt wird) wird, hätte dies zur Folge, dass sich die Ausgaben für den privaten Konsum im Durchschnitt um 0,41 Einheiten (= 410 Mio. DM) erhöhen (bzw. senken) würden. Ebenfalls könnte man den Wert für b_0 interpretieren, falls der zugehörige Abszissenwert Y_{pr}^V = 0 zum Wertebereich des Beobachtungsmaterials gehören würde. Dann wären 112,24 Mrd. DM Ausgaben für den privaten Konsum zu verzeichnen, obwohl kein privates Einkommen existieren würde. Da aber der Wert Y_{pr}^V = 0 nicht zum Wertebereich des Beobachtungsmaterials gehört, ist eine sinnvolle Interpretation der Größe b_0 in diesem Beispiel nicht möglich.

Allgemein muss man feststellen, dass die Regressionsgerade nur für einen bestimmten, von der jeweiligen Problemstellung festgelegten Bereich der Abszisse aussagefähig ist.



2.6 Übungsaufgaben / Exercises

Schädlingsbekämpfung

Sie sind von einem Bauern aus der Nachbarschaft beauftragt, den Zusammenhang zwischen dem Einsatz von Schädlingsbekämpfungsmittel (Faktoreinsatzmenge) und Menge am geerntetem Weizen (Outputmenge) zu bestimmen. Er stellt Ihnen zu diesem Zweck produktionstechnische Daten der vergangenen Jahre zur Verfügung, die Sie in Tabelle 3: finden:

Tabelle 3: Mengen ausgebrachten Schädlingsbekämpfungsmittels x [in l] und Weizenproduktion y [in hl]

Input x	Output y
0.58	130
1.10	151
1.20	162
1.30	193
1.95	143
2.55	202
2.60	176
2.90	198
3.45	239
3.50	231
3.60	256
4.10	278

4.35	301
4.40	289
4.50	276

Quelle: Hill, Griffith, Judge (1997)

a) Nehmen Sie an, die Produktionsfunktion könne durch das einfache Regressionsmodell $y = \beta_1 + \beta_2 x + e$ beschrieben werden. Schätzen Sie anhand dieses Modells die Parameter b_1 und b_2 nach dem Prinzip der kleinsten Quadrate.
b) Geben Sie eine ökonomische Interpretation der geschätzten Parameter.
c) Zeichnen Sie den Graphen der geschätzten Produktionsfunktion unter Verwendung der Ergebnisse aus a).
d) Der Preis eines Kilogramms Düngemittel betrage 6 Euro. Bestimmen Sie die Kosten- und Grenzkostenfunktion.
e) Welche Kritik üben Sie am oben geschätzten Modell?

Capital Asset Pricing Model (CAPM)

Das Capital Asset Pricing Model (CAPM) erklärt Preis und interne Verzinsung von Wertpapieren als Funktion der Verzinsung eines Portfolios mit allen öffentlich gehandelten Wertpapieren, dem so genannten Marktportfolio. Allgemein wird die interne Verzinsung jeder Investition relativ zu ihren Opportunitätskosten gemessen, der Verzinsung einer risikolosen Anlage. Die daraus resultierende Differenz nennt sich Risikoprämie und stellt einen Auf- oder Abschlag für das übernommene Risiko der Wertentwicklung dar. Das CAPM unterstellt, dass die Risikoprämie eines einzelnen Wertpapiers proportional zur Risikoprämie des Marktportfolios ist:

$$r_j - r_f \quad = \beta_j \, (r_m - r_f)$$

Dabei sind r_j und r_f die Returns auf einen Anteilsschein j und der risikolose Zins, r_m bezeichnet die Verzinsung des Marktportfolios. β_j ist der „Beta-Wert“ von Wertpapier j. Er misst die Preissensitivität dieses Papiers gegenüber Schwankungen in der Notation des Marktportfolios. Ein Beta kleiner als 1 deutet auf einen eher defensiven Titel hin, ein höherer Wert ist dagegen stärkeren Schwankungen in Preis und Rendite ausgesetzt. Die ökonometrische Version des Modells von oben enthält einen Störterm und eine Konstante (obwohl diese in der Theorie nicht vorgesehen ist):

$$r_j - r_f \quad = \alpha_j + \beta_j \, (r_m - r_f) + \varepsilon$$

In der folgenden Tabelle sind die Renditen der MobilOil- Aktie (Mobil Oil) an der NYSE, die Verzinsung eines Marktportfolios(Markt) und einer risikolosen Anlage in Form des 30-Tage- US-Treasury- Bills (Sicher) der Jahre 1971 bis 1990 aufgeführt (Berndt 1991):

a) Bestimmen Sie mit Hilfe der Methode der kleinsten Quadrate den Beta-Wert für Mobil-Oil in der genannten Zeitspanne.
b) Handelt es sich bei dem Wert um eine eher konservative oder spekulative Anlage?
($\beta > 1$ spekulativ; $\beta < 1$ konservativ)
c) Kommentieren Sie die Aussage des CAPM, nach der der Ordinatenabschnitt gleich Null ist.

Jahr	MobilOil	Markt	Sicher
1971	-0.046	-0.045	0.00487
1972	-0.017	0.010	0.00494
1973	0.049	0.050	0.00526
1974	0.077	0.063	0.00491
1975	-0.011	0.067	0.00513
1976	-0.043	0.007	0.00527
1977	0.028	0.071	0.00528
1978	0.056	0.079	0.00607
1979	0.064	0.002	0.00645
1980	-0.069	-0.189	0.00685
1981	0.037	0.084	0.00719
1982	0.041	0.015	0.00690
1983	0.061	0.058	0.00761
1984	-0.002	0.011	0.00761
1985	0.029	0.123	0.00769
1986	0.079	0.026	0.00764
1987	-0.086	0.014	0.00772
1988	0.088	0.075	0.00715
1989	0.018	-0.013	0.00728
1990	0.111	0.095	0.00789

$$\hat{y} = b_0 + b_1 x \qquad b_0 = 0{,}0047966, \quad b_1 = 0{,}5036174$$
$$\hat{y} = bx \qquad b = 0{,}52764$$

Exercise 1

Assume that you are in charge of the central monetary authority in a mystical country. You are given the following historical data on the quantity of money and national income (both in millions of dollars):

Year	Quantity of money	National income	Year	Quantity of money	National income
1973	2,0	5,0	1978	4,0	7,7
1974	2,5	5,5	1979	4,2	8,4
1975	3,2	6,0	1980	4,6	9,0
1976	3,6	7,0	1981	4,8	9,7
1977	3,3	7,2	1982	5,0	10,0

a) Plot these points on a scatter diagram. Then estimate the regression of national income Y on the quantity of money X and plot the line on the scatter diagram.
b) How do you interpret the intercept and slope of the regression line?
c) If you had sole control over the money supply and wished to achieve a level of national income of 120 in 1983, at what level would you set the money supply? Explain.

Exercise 2

Calculate the regression of income on grade-point average and compare it with the regression of grade-point average income. Why are the two results different?

Y [grade point average]	**X [Income of parents in $1000]**	$x^* = x_i - \bar{x}$	$y^* = y_i - \bar{y}$	$x^* y^*$	x^{*2}
4,0	21,0	7,5	1,0	7,5	56,25
3,0	15,0	1,5	0	0	2,25
3,5	15,0	1,5	0,5	0,75	2,25
2,0	9,0	-4,5	-1,0	4,5	20,25
3,0	12,0	-1,5	0	0	2,25
3,5	18,0	4,5	0,5	2,25	20,25
2,5	6,0	-7,5	-0,5	3,75	56,25
2,5	12,0	-1,5	-0,5	0,75	2,25

Exercise 3

a) Assume that least-squares estimates are obtained for the relationship Y = a + bX. After the work is completed, it is decided to multiply the units of the X variable by a factor of 10. What will happen to the resulting least-squares slope and intercept?

b) Generalise the results of part a) by evaluating the effects on the regression of changing the units of X and Y in the following manner: $Y^* = c_1 + c_2Y$ $X^* = d_1 + d_2X$

What can you conclude?

Exercise 4

What happens to the least-squares intercept and slope estimate when all observations on the independent variable are identical? Can you explain intuitively why this occurs?

Exercise 5

Prove that the estimated regression line passes through the point of means $(\bar{x}, \bar{y})$. Hint: Show that $\bar{x}$ and $\bar{y}$ satisfy the equation Y = a +bX, where a and b are defined in Equations (2 and 3).

Exercise 6

How would you interpret the – 1,78 value of the intercept in the regression equation of example 1.2? Explain why the value of the intercept is not likely to be of much practical interest.

Exercise 7

To test for sensitivity of least-squares estimates of intercept and slope to the presence of outliers, perform the following calculations:

a) Reestimate the slope and intercept in Example 1.1 under the assumption that the first observation was (21.0, 1.0) rather than (21.0, 4.0)

b) Reestimate the slope and intercept dropping the first observation from the sample.
   - Describe how the slope and intercept estimates in 1 and 2 compare with those given in the example. A graph of both straight lines would be helpful. Why are least-squares estimates so sensitive to individual data points?
   - Having graphed the least-squares line in case 1, would you conclude that the first data point is an outlier? Discuss.

Answers:

1.a $\hat{y} = 1{,}16 + 1{,}72x$

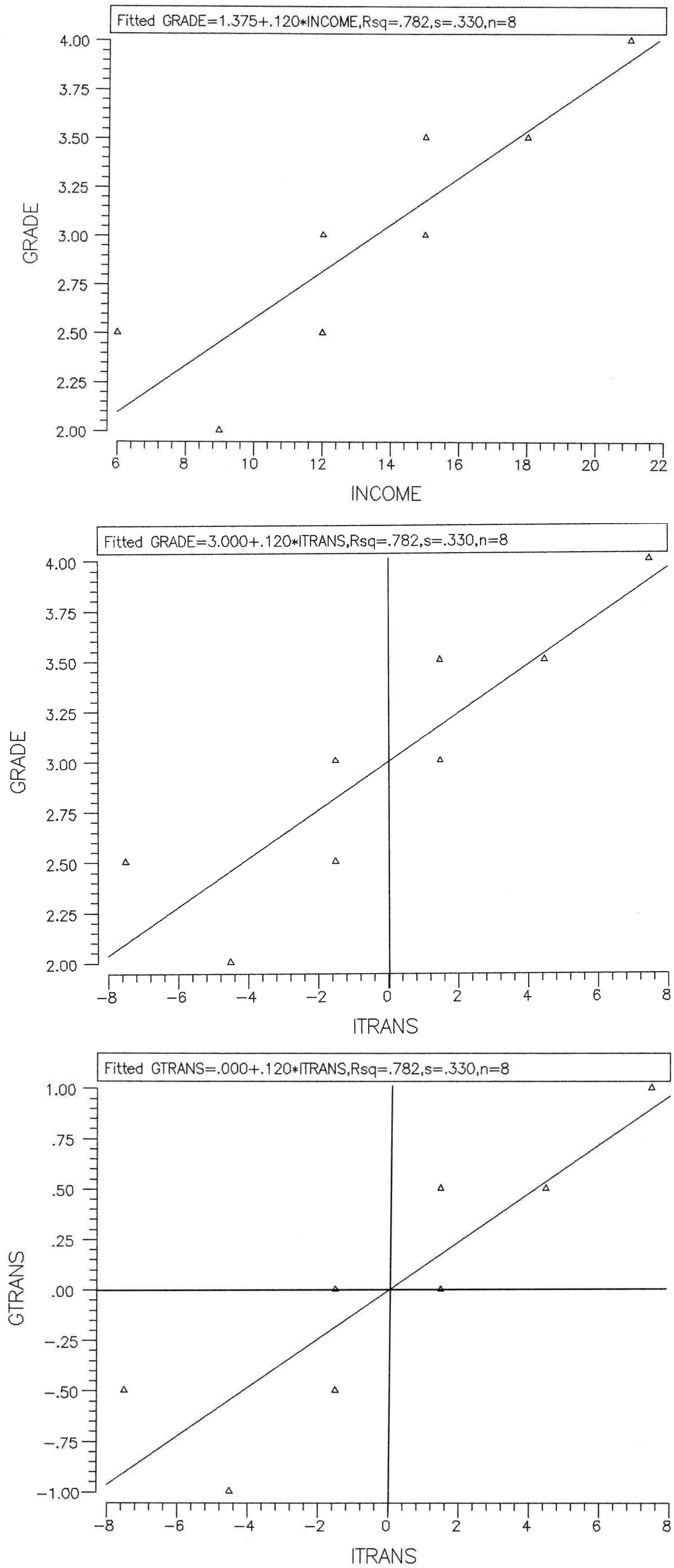
1.c y = 12; x = 6,3

3.a $\tilde{b} = \frac{1}{10} b_1 \quad \tilde{b}_0 = b_0$

3.b $b_1^* = \frac{c_2}{d_2} b_1 \quad b_0^* = c_1 + c_2 b_0 - \frac{c_2 d_1}{d_2} b_1$

Quelle: Pindyck und Rubinfeld (1981, S. 17f.

Grade-point Regression (Exercise 2): Variablentransformation



IV EMPIRISCHE REGRESSION: MEHRFACHREGRESSION

Das Thema „Mehrfachregression" wird methodisch eingeleitet mit einer Formulierung der Problemstellung in Matrixschreibweise. Die Matrixschreibweise erlaubt eine relativ einfache, umfassende Darstellung sowohl des funktionalen Ansatzes einer Untersuchung als auch die Einbeziehung aller Beobachtungen in die ökonometrische Formulierung. Dabei werden wir einige Grundregeln der Matrix Algebra wiederholen. Danach werden wichtige statistische und geometrische Eigenschaften des KQ-Schätzers vorgestellt und für eine Beurteilung der Modell-/Schätzgüte das Bestimmtheitsmaß aus dem Korrelationskoeffizienten entwickelt.

Im Literaturverzeichnis sind – wie schon erwähnt – für die Einfach- wie Mehrfachregression vielfältige Quellen gegeben. An neuerer englischsprachiger ökonometrischer Literatur sind Greene (2008), Wooldridge (2006, 2002) und Studenmund (2006), aus betrieblicher Sicht Anderson et al. (2010) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

Mit dem **FFB e-learning Modul**: **Lineare Regression - Deskriptives Modell** (www.leuphana.de/ffb) (Merz und Stolze 2010) wird eine Einführung in die empirische Regression gegeben. Das klassische lineare Regressionsmodel (CLR) ist dort in dem Modul: Lineare Regression – Stochastisches Modell (Merz und Stolze 2010) zu finden.

1 DIE METHODE DER KLEINSTEN QUADRATE (MKQ/OLS) – MEHRFACHREGRESSION

1.1 Lineare Mehrfachregression - Formulierung des Problems in Matrixnotation

Lineare multiple Regression:

lineare Beziehung zwischen y (Regressand) und (x_1, x_2, ..., x_K) Variablen (Regressoren)

n Beobachtungen $(y_i, x_{i1}, x_{i2}, \ldots, x_{iK})$ $\quad (i = 1, \ldots, n)$; n = Anzahl der Fälle

→ n Beobachtungsgleichungen

$$y_i = \underbrace{b_0(1 = x_{i0}) + b_1 x_{i1} + b_2 x_{i2} + \ldots b_K x_{iK}}_{\text{Ausgleichsebene}} + \underbrace{e_i}_{\text{Abweichung}}$$

Einfachregression: Ausgleichsgerade im $\mathbb{R}^2$

Mehrfachregression: Ausgleichshyperebene im $\mathbb{R}^{K+1}$

$$y_i = \sum_{k=0}^{K} b_k x_{ik} + e_i \qquad (i = 1, \ldots, n = \text{Anzahl der Fälle})$$

$$x_{i0} = 1 \text{ (absolutes Glied)} \qquad (k = 1, \ldots, K = \text{Anzahl der Variablen})$$

→ Lineares Gleichungssystem (LGS)

$$
\begin{array}{l}
y_1 = b_0 + b_1 x_{11} + b_2 x_{12} + \ldots + b_K x_{1K} + e_1 \\
y_2 = b_0 + b_1 x_{21} + b_2 x_{22} + \ldots + b_K x_{2K} + e_2 \\
\vdots \quad \vdots \quad \vdots \quad \vdots \quad \ddots \quad \vdots \quad \vdots \\
y_n = b_0 + b_1 x_{n1} + b_2 x_{n2} + \ldots + b_K x_{nK} + e_n
\end{array}
$$

Kompakt in Matrixnotation

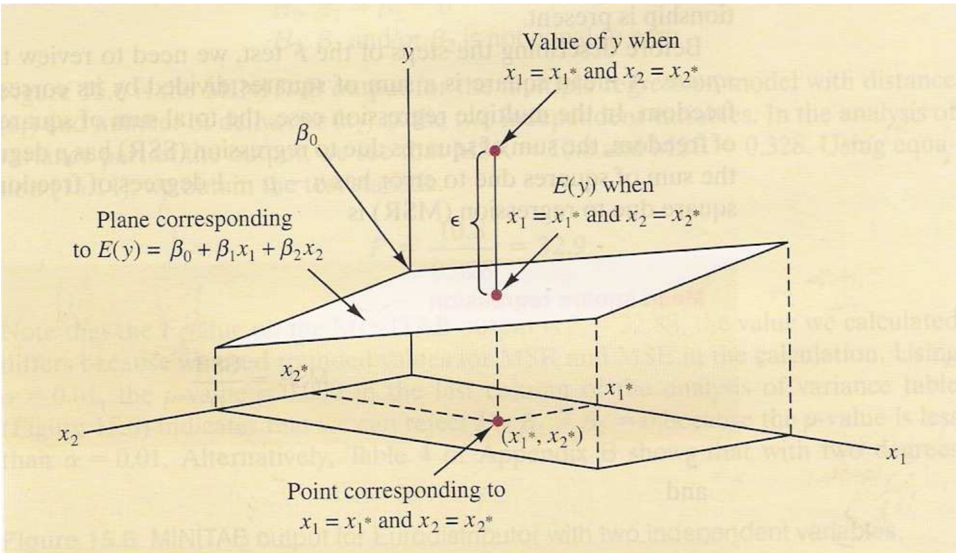
(Großbuchstaben stehen für Matrizen, Kleinbuchstaben für Vektoren)

$$
\begin{array}{ccccc}
y & = & Xb & + & e \\
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} & = & \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1K} \\ 1 & x_{21} & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nK} \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_K \end{pmatrix} & + & \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \\
y_{n\times 1} & = & X_{n\times(K+1)} \quad b_{(K+1)\times 1} & + & e_{n\times 1}
\end{array}
$$

Gesucht:

Parametervektor **b**, der die optimale Lage der Ausgleichshyperebene durch die Punktewolke bestimmt.

Regression Hyperebene im R3



Quelle: Anderson et al. 2010, p. 573

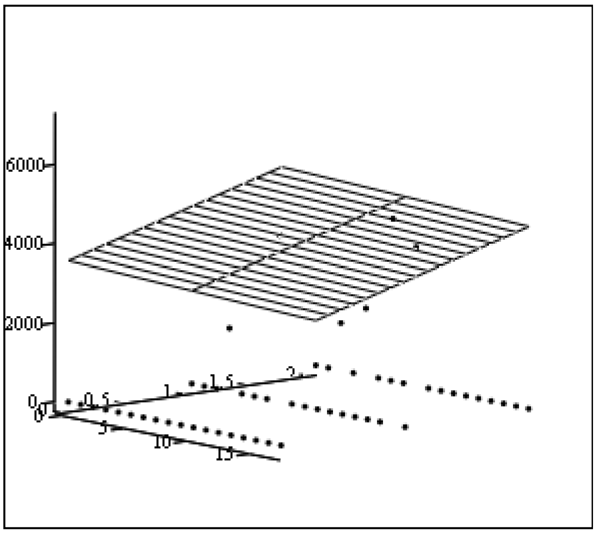
Beispiel: Einkommensregression (K=2)

income= b_0 + b_1*age + b_2*sex

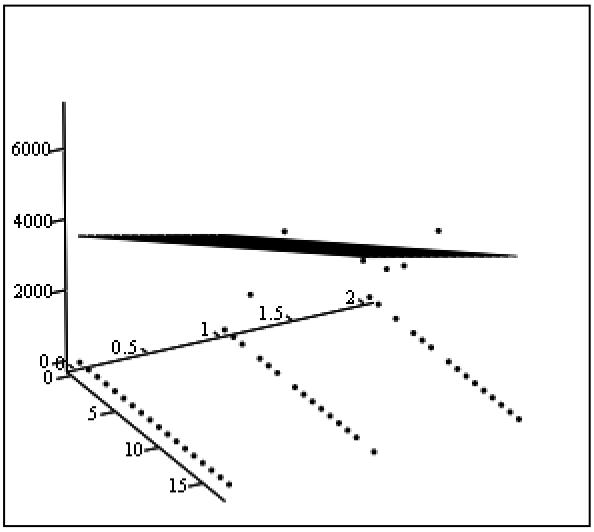
Daten:

Observation	AGE	SEX	INCOME
1	22.000	1.0000	1200.0
2	24.000	1.0000	1700.0
3	28.000	1.0000	3500.0
4	27.000	.00000	4200.0
5	23.000	1.0000	1600.0
6	36.000	.00000	5200.0

MATHCAD: MKQ/OLS-Hyperebene
income = -1728.57 + 204.081 age -1220.41 sex



coul, B



coul, B

Rechnerische Lösung: siehe 1.3 Beispiel 3

1.2 Parameterschätzung bei K Variablen

OLS-Schätzung für b

Minimierung der Summe der Residuenquadrate. Residuum: $e_i = y_i - \hat{y}_i$

In Summenschreibweise:

$$S = \sum_{i=1}^{n} e_i^2 = e_1^2 + e_2^2 + \ldots + e_n^2 = \min!$$

In Vektorschreibweise:

$$S = \mathbf{e'e} = (e_1, e_2, \ldots, e_n)\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \sum_{i=1}^{n} e_i^2 = e_1^2 + e_2^2 + \ldots + e_n^2$$

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \sum_{k=0}^{K} b_k x_{ik})^2 = \sum_{i=1}^{n} (y_i - \mathbf{b}'\mathbf{x}_i)^2 = \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) = \text{min!}$$

Vorgehensweise bei der Bestimmung der OLS-Parameter:

- $S = \mathbf{e'e}$ umformen
- Nullsetzen der 1. Ableitung (notwendige Bedingung) → Lösung $\hat{\mathbf{b}}_{OLS}$
- Überprüfen ob $S(\hat{\mathbf{b}}) = \text{min!}$ (hinreichende Bedingung)

Schritt 1: S = e'e umformen

$$\begin{aligned} S = \mathbf{e'e} &= (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) = (\mathbf{y}' - \mathbf{b}'\mathbf{X}')(\mathbf{y} - \mathbf{Xb}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{Xb} - \mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{Xb} \\ &= \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{Xb} \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{Xb} \end{aligned}$$

Anmerkung: $\mathbf{y}'\mathbf{Xb} = y'_{1\times n}\mathbf{X}_{n\times(K+1)}\mathbf{b}_{(K+1)\times 1} = \mathbf{y}'_{1\times n}(\mathbf{Xb})_{n\times 1} = (\mathbf{y}'\mathbf{Xb})_{1\times 1}$ → Skalar (Zahlenwert)

Für ein Skalar c gilt: $c' = c$

Folglich gilt auch: $\mathbf{y}'\mathbf{Xb} = (\mathbf{y}'\mathbf{Xb})' = (\mathbf{Xb})'\mathbf{y} = \mathbf{b}'\mathbf{X}'\mathbf{y}$

Schritt 2: Extremamöglichkeit

Notwendige Bedingung: Nullsetzen der 1. Ableitung

Exkurs: Differentiation von Matrizen I

(1) $\dfrac{\partial \mathbf{x}'\mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}$ $\quad \mathbf{x}_{m\times 1}$-Variablenvektor

$\mathbf{a}_{m\times 1}$-Konstantenvektor

$\dfrac{\partial \mathbf{b}'\mathbf{X}'\mathbf{y}}{\partial \mathbf{b}} = \mathbf{X}'\mathbf{y}$ $\quad \mathbf{x}_{m\times 1} \triangleq \mathbf{b}_{(K+1)\times 1}$-Variablenvektor

$\mathbf{a}_{m\times 1} \triangleq \mathbf{X}'_{(K+1)\times n}\mathbf{y}_{n\times 1}$-Konstantenvektor

(2) $\dfrac{\partial \mathbf{x}'\mathbf{Ax}}{\partial \mathbf{x}} = 2\mathbf{Ax}$ $\quad \mathbf{A}_{m\times m}$-Konstantenmatrix

$\dfrac{\partial \mathbf{b}'\mathbf{X}'\mathbf{Xb}}{\partial \mathbf{b}} = 2\mathbf{X}'\mathbf{Xb}$ $\quad \mathbf{A}_{m\times m} \triangleq \mathbf{X}'\mathbf{X}_{(K+1)\times(K+1)}$-Konstantenmatrix

(3) $\dfrac{\partial \mathbf{Ax}}{\partial \mathbf{x}} = \mathbf{A}'$ $\qquad$ $\mathbf{x}_{n\times 1}$-Variablenvektor

$\dfrac{\partial \mathbf{X'Xb}}{\partial \mathbf{b}} = (\mathbf{X'X})' = (\mathbf{X'X''}) = (\mathbf{X'X})$ $\qquad$ $\mathbf{A}_{n\times n}$-Konstantenmatrix

Ist $2\mathbf{X'X}$ positiv definit, dann minimiert $\hat{\mathbf{b}}$ den Ausdruck $S(\mathbf{b})$.

Positive Definitheit: $\mathbf{x'Ax} > 0 \ \forall\ \mathbf{x} \neq \mathbf{0}$

also:

$$S = \mathbf{e'e} = \mathbf{y'y} - 2\mathbf{b}'\underbrace{\mathbf{X'y}}_{\mathbf{a}} + \mathbf{b}'\underbrace{\mathbf{X'X}}_{\mathbf{A}}\mathbf{b}$$

$$\frac{\partial S(\mathbf{b})}{\partial \mathbf{b}} = \frac{\partial \mathbf{e'e}}{\partial \mathbf{b}} = -2\mathbf{X'y} + 2\mathbf{X'Xb}$$

$$\frac{\partial S(\hat{\mathbf{b}})}{\partial \hat{\mathbf{b}}} = \mathbf{0}$$ $\hat{\mathbf{b}}$ = minimierendes $\mathbf{b}$, minimiert $S(\mathbf{b})$

$$\boxed{\mathbf{X'X}\hat{\mathbf{b}} = \mathbf{X'y}}$$ (Normalgleichungssystem)

$$(\mathbf{X'X})^{-1}(\mathbf{X'X})\hat{\mathbf{b}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$$

(Bildung der Inverse $(\mathbf{X'X})^{-1}$ nur möglich wenn $\text{Det}(\mathbf{X'X}) \neq 0$)

$$\boxed{\hat{\mathbf{b}} = \mathbf{b}_{OLS} = (\mathbf{X'X})^{-1}\mathbf{X'y}}$$

Schritt 3: Extremasicherheit (hinreichende Bedingung)

$\hat{\mathbf{b}}$ minimiert $S(\hat{\mathbf{b}})$, wenn $\dfrac{\partial^2 S(\hat{\mathbf{b}})}{\partial \hat{\mathbf{b}}^2}$ positiv definit ist.

$$\frac{\partial^2 S(\hat{\mathbf{b}})}{\partial \hat{\mathbf{b}}^2} = 2(\mathbf{X'X})' = 2(\mathbf{X'X}) > 0$$

Satz:

Eine symmetrische Matrix **A** ist positiv definit, dann und nur dann, wenn **c'Ac** >0 für alle $\mathbf{c} \neq \mathbf{0}$

Hier:

$A = X'X$

$c'(X'X)c = (c'X')(Xc) = (Xc)'(Xc)$ (ergibt Skalar)

Diese Summe der Quadrate ist immer größer Null für alle $c \neq 0$

Alternativ:

Satz:

Ist **P** eine $n \times m$-Matrix mit Rang(**P**) = m, dann ist **P´P** positiv definit (z. B. Goldberger, A. Econometric Theory, 1966, S.36)

Annahme:

Rang(**X**) = (K+1), d.h. die erklärenden Variablen sind unabhängig voneinander. Dann ist **X'X** positiv definit und $\hat{\mathbf{b}} = \hat{\mathbf{b}}_{OLS}$ minimiert S(**b**) q.e.d.

Damit minimiert **b** = (**X'X**)$^{-1}$**X'y** tatsächlich die Fehlerquadratsumme.

Zur Berechnung des Schätzvektors b

Zur Bestimmung von b_{ols} sind drei zentrale Bausteine zu berechnen: **X'X, X'y** und die Inverse **(X'X)**$^{-1}$

X'X: Momentenmatrix mit Produktmomenten um Null 2. Ordnung

$$\mathbf{X'X} = \overbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1K} & x_{2K} & \cdots & x_{nK} \end{pmatrix}_{(K+1)\times n}}^{X'} \overbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1K} \\ 1 & x_{21} & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nK} \end{pmatrix}_{n\times(K+1)}}^{X}$$

$$= \begin{pmatrix} \sum_i 1 = n & \sum_i x_{i1} & \sum_i x_{i2} & \cdots & \sum_i x_{iK} \\ \sum_i x_{i1} & \sum_i x_{i1}^2 & \sum_i x_{i1}x_{i2} & \cdots & \sum_i x_{i1}x_{iK} \\ \sum_i x_{i2} & \sum_i x_{i2}x_{i1} & \sum_i x_{i2}^2 & \cdots & \sum_i x_{i2}x_{iK} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_i x_{iK} & \sum_i x_{iK}x_{i1} & \sum_i x_{iK}x_{i2} & \cdots & \sum_i x_{iK}^2 \end{pmatrix}_{(K+1)\times(K+1)\text{ symmetrisch}}$$

Momentenmatrix **X'X**

- symmetrisch
- Element "links oben": $\sum_i 1 = n$ Anzahl der Beobachtungen
- 1. Zeile: $\sum_i x_{ik}$ Summe der Beobachtungen von Variable k
  $(k = 1, 2, \ldots, K)$ K = Anzahl der erklärenden Variablen
- Diagonalelemente: $\sum_i x_{ik}^2$ (x- Quadratsummen)
- Nebendiagonalelemente: $\sum_i x_{ik}x_{il}$ (Summe der x- Kreuzprodukte)

X'y , Kreuzprodukte x_k und y

$$\mathbf{X'y} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1K} & x_{2K} & \dots & x_{nK} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_{i1} y_i \\ \sum_i x_{i2} y_i \\ \vdots \\ \sum_i x_{iK} y_i \end{pmatrix}_{(K+1)\times 1}$$

- 1. Wert: Summe der y-Werte
- andere Werte: Kreuzprodukte x_k und y

$(X'X)^{-1}$, Berechnung der Inversen: siehe Anhang

1.3 Beispiele: Regression in Matrixnotation

Beispiel 1: Bruttosozialprodukt und Geldmenge (in Mrd. DM) - Einfachregression

Jahr	Geldmenge	BSP	**Jahr**	Geldmenge	BSP
1973	2,0	5,0	1978	4,0	7,7
1974	2,5	5,5	1979	4,2	8,4
1975	3,2	6,0	1980	4,6	9,0
1976	3,6	7,0	1981	4,8	9,7
1977	3,3	7,2	1982	5,0	10,0

Gleichungssystem:

$$\begin{pmatrix} y_i & = & b_0 1 & + & b_1 x_{i1} & + & e_i & = & \sum_{k=0}^{1} b_k x_{ik} + e_i \\ BSP_i & = & b_0 1 & + & b_1 \text{Geldmenge}_i & + & e_i \end{pmatrix} \quad (i = 1, \dots, 10); K = 1$$

$$\text{gesucht}: b_0 \text{ und } b_1, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$$

MKQ/ OLS: $\mathbf{b}_{\mathrm{OLS}} = \hat{\mathbf{b}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$

$$\mathbf{X} = \begin{pmatrix} 1 & 2,0 \\ 1 & 2,5 \\ 1 & 3,2 \\ 1 & 3,6 \\ 1 & 3,3 \\ 1 & 4,0 \\ 1 & 4,2 \\ 1 & 4,6 \\ 1 & 4,8 \\ 1 & 5,0 \end{pmatrix} \qquad \mathbf{X'} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2,0 & 2,5 & 3,2 & 3,6 & 3,3 & 4,0 & 4,2 & 4,6 & 4,8 & 5,0 \end{pmatrix}$$

$$(\mathbf{X'X}) = \begin{pmatrix} n & \sum_i x_{i1} \\ \sum_i x_{i1} & \sum_i x_{i1}^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2,0 & 2,5 & \cdots & 5,0 \end{pmatrix} \begin{pmatrix} 1 & 2,0 \\ 1 & 2,5 \\ \vdots & \vdots \\ 1 & 5,0 \end{pmatrix} = \begin{pmatrix} 10 & 37,2 \\ 37,2 & 147,18 \end{pmatrix}$$

(X'X)$^{-1}$ über Adjunkte

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \qquad |\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21} = 10 \cdot 147,18 - 37,2 \cdot 37,2 = 87,96$$

$$(\mathbf{X'X})^{-1} = \frac{1}{87,96}\begin{pmatrix} 147,18 & -37,2 \\ -37,2 & 10 \end{pmatrix} = \begin{pmatrix} 1,6733 & -0,4229 \\ -0,4229 & 0,1137 \end{pmatrix}$$

$$\mathbf{X'y} = \begin{pmatrix} \sum_i x_{i0}y_i \\ \sum_i x_{i1}y_i \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2,0 & 2,5 & \cdots & 5,0 \end{pmatrix} \begin{pmatrix} 5,0 \\ 5,5 \\ 6,0 \\ 7,0 \\ 7,2 \\ 7,7 \\ 8,4 \\ 9,0 \\ 9,7 \\ 10,0 \end{pmatrix} = \begin{pmatrix} 75,5 \\ 295,95 \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_{i1}y_i \end{pmatrix}$$

$$\hat{\mathbf{b}} = \left((\mathbf{X'X})^{-1}\right)\mathbf{X'y} = \begin{pmatrix} 1,6733 & -0,4229 \\ -0,4229 & 0,1137 \end{pmatrix}\begin{pmatrix} 75,5 \\ 295,95 \end{pmatrix} = \begin{pmatrix} 1,168 \\ 1,715 \end{pmatrix} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \end{pmatrix}$$

$$\Rightarrow \widehat{\mathrm{BSP}}_{i=t} = 1,168 + 1,715 \cdot \mathrm{Geldmenge}_{i=t} \qquad (t = 1, \ldots, T)$$

Interpretation:

Erhöht sich die Geldmenge um 1 Mrd. DM dann erhöht sich das Bruttosozialprodukt (im Mittel) um 1,715 Mrd. DM unter sonst gleichen Bedingungen.

Beispiel 2: MacProd – Cobb-Douglas-Produktionsfunktion - Mehrfachregression

Die Firma MacProd möchte ihren Produktionsprozess über eine Cobb-Douglas-Produktionsfunktion beschreiben.

(1) $\mathrm{Y} = \gamma \mathrm{K}^{\alpha} \mathrm{L}^{\beta}$ Y = Output, K = Kapitalinput, L = Arbeit (Labor)

Gesucht sind die Elastizitäten α, β sowie der Faktor γ

Beobachtete Daten:

i	y	ln Y	K	ln K	L	ln L
1	20,086	3	20,086	3	148,413	5
2	2,718	1	2,718	1	54,598	4
3	2980,958	8	148,413	5	403,429	6
4	20,086	3	7,386	2	54,598	4
5	148,413	5	54,598	4	403,429	6

Logarithmieren von (1) → lineare Gleichung

$$\ln \mathrm{Y} = \ln \gamma + \alpha \ln \mathrm{K} + \beta \ln \mathrm{L}$$

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} \qquad (i = 1, \ldots, 5),\ K = 2$$

$$\mathbf{y} = \begin{pmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{pmatrix}_{5\times1} \qquad \mathbf{X} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{pmatrix}_{5\times3} \qquad \mathbf{X'X} = \begin{pmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{pmatrix}_{3\times3} \qquad \mathbf{X'y} = \begin{pmatrix} 20 \\ 76 \\ 109 \end{pmatrix}_{3\times1}$$

Normalgleichungen: $(\mathbf{X'X})\mathbf{b} = \mathbf{X'y}$

$$\begin{pmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 20 \\ 76 \\ 109 \end{pmatrix}$$

Lösungsmethode 1: Gauß'sches Eliminationsverfahren nach Triangulation ("Dreiecksmatrix")

Triangulation durch elementare Zeilentransformationen

1. $Z_2 = Z_2 - 3Z_1$ dann 2. $Z_3 = Z_3 - \frac{6}{10} Z_2$
   $Z_3 = Z_3 - 5Z_1$

$$\begin{pmatrix} 5 & 15 & 25 \\ 0 & 10 & 6 \\ 0 & 0 & 0,4 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 20 \\ 16 \\ -0,6 \end{pmatrix}$$

aus 3. Gleichung folgt:	$0{,}4b_2 = -0,6$	$\rightarrow b_2 = -1,5$
b_2 in Gleichung 2:	$10b_1 + 6(-1,5) = 16$	$\rightarrow b_1 = 2,5$
1. Gleichung:	$5b_0 + 15(2,5) + 25(-1,5) = 20$	$\rightarrow b_0 = 4$

Regressionsgleichung

$$y_i = 4 + 2,5x_{i1} - 1,5x_{i2}$$

bzw. Cobb-Douglas

$$y_i = 54,6K^{2,5}L^{-1,5} \qquad (\ln \gamma = 4 \Leftrightarrow \gamma = e^4 = 54,6)$$

- Niveaukonstante: 54,6
- Kapitalelastzitität: 2,5
- Arbeitselastizität: -1,5

(Achtung Vorzeicheninterpretation! Fiktives Beispiel)

Elastizität:

Eine1%-ige Erhöhung von x führt zu einer b%-igen Änderung von y. Die Interpretation der Koeffizienten im logarithmierten Modell als Elastizitäten (also als prozentualer Änderung der modellendogenen Variablen bei Änderung der exogenen Größe zum ein Prozent) lässt sich wie folgt verdeutlichen:

Die Regression bezieht sich auf logarithmisch transformierte Größen der eigentlichen Variablen:

$$\ln y = b_0 + b_1 \ln x_1 + b_2 \ln x_2$$

Eine partielle Ableitung nach einer der exogenen Variablen z.B. x_1 ergibt:

$$\frac{\partial(\ln y)}{\partial x_1} = \frac{\partial(b_0 + b_1 \ln x_1 + b_2 \ln x_2)}{\partial x_1} = \frac{b_1}{x_1}$$

Zusammen mit der Gleichung

$$\frac{\partial(\ln y)}{\partial y} = \frac{1}{y} \Leftrightarrow \partial(\ln y) = \frac{\partial y}{y}$$

lässt sich der einzelne Koeffizient direkt als Elastizität darstellen:

$$\frac{\frac{\partial y}{y}}{\partial x_1} = \frac{b_1}{x_1} \Leftrightarrow b_1 = \frac{\frac{\partial y}{y}}{\frac{\partial x_1}{x_1}}$$

Lösungsmethode 2: Über Inverse nach Gauß

$$(\mathbf{A}, \mathbf{I}) \xrightarrow{\text{Gauß}} (\mathbf{I}, \mathbf{A}^{-1})$$

$$((\mathbf{X'X}), \mathbf{I}) \xrightarrow{\text{Gauß}} (\mathbf{I}, (\mathbf{X'X})^{-1})$$

$\mathbf{X'X}$			I			
5	15	25	1	0	0	$Z_1^* = \frac{1}{5} Z_1$
15	55	81	0	1	0	$Z_2^* = Z_2 - 3Z_1$
25	81	129	0	0	1	$Z_3^* = Z_3 - 5Z_1$
1	3	5	0,2	0	0	$Z_1^* = Z_1 - \frac{3}{10} Z_2$
0	10	6	-3	1	0	$Z_2^* = \frac{1}{10} Z_2$
0	6	4	-5	0	1	$Z_3^* = Z_3 - \frac{6}{10} Z_2$
1	0	3,2	1,1	-0,3	0	$Z_1^* = Z_1 - \frac{3,2}{0,4} Z_3$
0	1	0,6	-0,3	0,1	0	$Z_2^* = Z_2 - \frac{0,6}{0,4} Z_3$
0	0	0,4	-3,2	-0,6	1	$Z_3^* = Z_3 / 0,4$
1	0	0	26,7	4,5	-8	* im nächstens Tableau
0	1	0	4,5	1	-1,5	
0	0	1	-8	-1,5	2,5	
I			$(\mathbf{X'X})^{-1}$			

$$\hat{\mathbf{b}} = (\mathbf{X'X})^{-1} \mathbf{X'y} = \begin{pmatrix} 26,7 & 4,5 & -8 \\ 4,5 & 1 & -1,5 \\ -8 & -1,5 & 2,5 \end{pmatrix} \begin{pmatrix} 20 \\ 76 \\ 109 \end{pmatrix} = \begin{pmatrix} 4 \\ 2,5 \\ -1,5 \end{pmatrix}$$

$$\ln \mathrm{Y} = \ln \gamma + \alpha \ln \mathrm{K} + \beta \ln \mathrm{L}$$

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2}$$

$$y_i = 4 + 2,5 x_{i1} - 1,5 x_{i2}$$

Beispiel 3: income = b0 + b1*age + b2*sex (ET)

```
? ---------------------------------
? ET income: ols by matrix algebra
? ---------------------------------
?
```

```
? read data
?      y : income
?      x0: one
?      x1: age
?      x2: sex (0=male,1=female)
?
read; file=income.dat;nvar=4;nobs=6;names=1$
list; income, age, sex, asex$
?
? create X'X (=XSX) and X'y (XSy)
? ------------------------------
namelist; X=one,age,sex$
matrix; XSX=xdot(one,age,sex)
      ; XSX=xdot(X)$
matrix; XSy=xdot(X,income)$
?
? create invers of (X'X)
? ------------------------------
matrix; XSXINV=ginv(XSX)
      ; XSXINV=xpxi(one,age,sex)$
?
? compute ols
? ------------------------------
matrix; bols=XSXINV | XSy
      ; bet=xlsq(X,income)$
?
? regression by et
? ------------------------------
regres; dep=income; ind=one,age,sex$
calc; y0fit=-1728.57 +204.082*22 -1220.41*1$
list; age,sex,income,yfit$
```

```
Ergebnisse: ET: income = b0 + b1*age + b2*sex
===========================================
DATA LISTING (Current sample)
Observation      INCOME      AGE         SEX
      1         1200.0      22.000      1.0000
      2         1700.0      24.000      1.0000
      3         3500.0      28.000      1.0000
      4         4200.0      27.000      .00000
      5         1600.0      23.000      1.0000
      6         5200.0      36.000      .00000



  1. Matrix -> XSX=XDOT(X)

      <<<< XSX      >>>>   COLUMN
                1             2             3
 ROW   1   6.00000       160.000       4.00000
 ROW   2   160.000       4398.00       97.0000
 ROW   3   4.00000       97.0000       4.00000

     1. Matrix -> XSY=XDOT(X,INCOME)

      <<<< XSY      >>>>   COLUMN
                1
 ROW   1   17400.0
 ROW   2   502600.
 ROW   3   8000.00
```

```
 1. Matrix -> XSXINV=GINV(XSX)

     <<<< XSXINV   >>>>   COLUMN
              1              2             3
ROW   1   16.7000      -.514286      -4.22857
ROW   2  -.514286       .163265E-01   .118367
ROW   3  -4.22857       .118367       1.60816

    1. Matrix -> BOLS=XSXINV|XSY

     <<<< BOLS     >>>>   COLUMN
              1
ROW   1  -1728.57
ROW   2   204.081
ROW   3  -1220.41

===========================================================================
Ordinary    Least Squares
Dependent Variable        INCOME       Number of Observations            6
Mean of Dep. Variable    2900.0000     Std. Dev. of Dep. Var.   1634.625339
Std. Error of Regr.       602.4892     Sum of Squared Residuals .108898E+07
R - squared                 .91849     Adjusted R - squared          .86415
F( 2,   3)                 16.9025     Prob. Value for F             .02327
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x   Mean of X  .Dev.of X
---------------------------------------------------------------------------
Constant  -1728.57       2462.       -.702   .53320
AGE      204.082     76.98       2.651   .07693     26.66667       5.12510
SEX     -1220.41     764.0      -1.597   .20848       .66667        .51640

>>> Y0FIT    =     1540.824     <<<
```

Beispiel 4: Umsatz (KPMT) = f(Werbeausgaben, Mitarbeiterstand) (ET)

Frage: Wird der Umsatz der Firma KPMT wesentlich bestimmt durch die Werbeausgaben (x_1) und Mitarbeiterstand (x_2)?

Folgende Beobachtungen (für die Jahre 1995 bis 2000) seien gegeben:

y	2	7	6	8	9	4
$\mathbf{x_1}$	6	1	2	4	3	8
$\mathbf{x_2}$	4	2	4	5	2	3

i	t	y	$x_0=1$	x_1	x_2
1995	1	2	1	6	4
1996	2	7	1	1	2
1997	3	6	1	2	4
1998	4	8	1	4	5
1999	5	9	1	3	2
2000	6	4	1	8	3

$$\rightarrow \mathbf{X} = \begin{pmatrix} 1 & 6 & 4 \\ 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 4 & 5 \\ 1 & 3 & 2 \\ 1 & 8 & 3 \end{pmatrix}_{6\times 3}$$

$$\mathbf{X'X} = \overbrace{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 6 & 1 & 2 & 4 & 3 & 8 \\ 4 & 2 & 4 & 5 & 2 & 3 \end{pmatrix}_{3\times 6}}^{\mathbf{x}'} \overbrace{\begin{pmatrix} 1 & 6 & 4 \\ 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & 4 & 5 \\ 1 & 3 & 2 \\ 1 & 8 & 3 \end{pmatrix}_{6\times 3}}^{\mathbf{x}} = \begin{pmatrix} 6 & 24 & 20 \\ 24 & 130 & 84 \\ 20 & 84 & 74 \end{pmatrix}_{3\times 3} \quad \text{symmetrisch}$$

Anzahl der Beobachtungen: $\sum_i x_{i0}^2 = \sum_i 1 = n = 6$ Jahre

Summe der Beobachtungen: $\sum_i 1 \cdot x_{i1} = \sum_i x_{i1} = 24$ alle Werbeausgaben

Summe der Kreuzprodukte: $\sum_i x_{i1} \cdot x_{i2} = 84$ Summe Werbeausgaben × Mitarbeiter

Summe der quadrierten x_{i2}: $\sum_i x_{i2}^2 = 74$ Summe Mitarbeiterquadrate

$$(\mathbf{X'y}) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 6 & 1 & 2 & 4 & 3 & 8 \\ 4 & 2 & 4 & 5 & 2 & 3 \end{pmatrix}_{3\times 6} \begin{pmatrix} 2 \\ 7 \\ 6 \\ 8 \\ 9 \\ 4 \end{pmatrix}_{6\times 1} = \begin{pmatrix} 36 \\ 122 \\ 116 \end{pmatrix}_{3\times 1}$$

Mit **X'X** und **X'y** sind zwei zentrale Bausteine für $\mathbf{b}_{OLS} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ gegeben. Notwendig zur Berechnung ist dann noch die Inverse der Momentenmatrix $(\mathbf{X'X})^{-1}$, beispielsweise mit dem Statistikpaket „Econometric Toolkit“ (ET) mit dem Befehl GINV(matrix):

```
 1. Matrix -> XSXINV=GINV(XSX)
     <<<< XSXINV   >>>>   COLUMN
             1             2             3
 ROW   1   1.83143      -.685714E-01  -.417143
 ROW   2  -.685714E-01   .314286E-01  -.171429E-01
 ROW   3  -.417143      -.171429E-01   .145714

     1. Matrix -> BOLS=XSXINV|XSY
      <<<< BOLS     >>>>   COLUMN
             1
 ROW   1   9.17714
 ROW   2  -.622857
 ROW   3  -.205715
```

```
===========================================================================
Ordinary    Least Squares
Dependent Variable        UMSATZ       Number of Observations            6
Mean of Dep. Variable       6.0000     Std. Dev. of Dep. Var.     2.607681
Std. Error of Regr.         2.5478     Sum of Squared Residuals    19.4743
R - squared                 .42723     Adjusted R - squared         .04538
F( 2,   3)                  1.1188     Prob. Value for F            .43348
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant   9.17714      3.448       2.662   .07623
WERBE     -.622857      .4517      -1.379   .26173   4.00000       2.60768
MITARB    -.205714      .9726       -.212   .84604   3.33333       1.21106
```

2 Eigenschaften der MKQ/ OLS und geometrische Interpretation

Gegeben:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \text{Vektor aller beobachteten y-Werte}$$

$$\hat{\mathbf{y}} = \mathbf{Xb} \qquad \text{Ausgleichshyperebene; } \mathbf{X} = \text{Matrix aller beobachteten x-Werte}$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e}$$

$$\rightarrow \mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{Xb} \quad \text{Residuen}$$

2.1 Statistische Eigenschaften der MKQ/ OLS

a) Die Summe der Kreuzprodukte eines jeden Regressors und der Residuen ist Null

Bedeutung: Unabhängigkeit zwischen x_{ik} und e_i

Summenschreibweise:

$$\sum_{i=1}^{n} x_{ik} e_i = 0 \qquad \text{für alle Regressoren } k = 0,1,\ldots,K$$

Matrixschreibweise:

$$\mathbf{X'e} = 0 \qquad \begin{pmatrix} x_{10} & x_{20} & \cdots & x_{n0} \\ x_{11} & x_{21} & \cdots & x_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1K} & x_{2K} & \cdots & x_{nK} \end{pmatrix}_{(K+1)\times n} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}_{n\times 1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(K+1)\times 1}$$

Beweis:

$$\begin{aligned} \mathbf{X'e} &= \mathbf{X'}(\mathbf{y} - \hat{\mathbf{y}}) = \mathbf{X'y} - \mathbf{X'}\hat{\mathbf{y}} = \mathbf{X'y} - \mathbf{X'Xb} \\ &= \mathbf{X'y} - \mathbf{X'X}\left((\mathbf{X'X})^{-1}\mathbf{X'y}\right) = \mathbf{X'y} - \underbrace{\mathbf{X'X}(\mathbf{X'X})^{-1}}_{\mathrm{I}}\mathbf{X'y} \\ &= \mathbf{X'y} - \mathbf{IX'y} = \mathbf{X'y} - \mathbf{X'y} = \mathbf{0} \qquad \text{q.e.d.} \end{aligned}$$

b) Die Summe der Kreuzprodukte der berechneten y-Werte und der Residuen ist Null

Bedeutung: Unabhängigkeit zwischen $\hat{y}_i$ und e_i

Summenschreibweise:

$$\sum_{i=1}^{n} \hat{y}_i e_i = 0$$

Matrixschreibweise:

$$\hat{\mathbf{y}}'\mathbf{e} = 0 \Leftrightarrow \begin{pmatrix} \hat{y}_1 & \cdots & \hat{y}_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = 0$$

Beweis:

$$\begin{aligned} \hat{\mathbf{y}}'\mathbf{e} &= \hat{\mathbf{y}}'(\mathbf{y} - \hat{\mathbf{y}}) = \hat{\mathbf{y}}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{y}} & \qquad \hat{\mathbf{y}} &= \mathbf{Xb} \\ &= \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{Xb} = \mathbf{b}'(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{Xb}) & \qquad \mathbf{b} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{b}'(\mathbf{X}'\mathbf{y} - \underbrace{\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}}_{\mathbf{I}}\mathbf{X}'\mathbf{y}) \\ &= \mathbf{b}'(\mathbf{X}'\mathbf{y} - \mathbf{IX}'\mathbf{y}) = \mathbf{b}'(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{y}) = 0 & \qquad &\text{q.e.d.} \end{aligned}$$

c) Die Summe der Residuen ist Null

Bedeutung: Ausgleichseigenschaft

Fehlerangleichende Hyperebene: $\sum$ der positiven = $\sum$ der negativen Abweichungen

Summenschreibweise:

$$\sum_{i=1}^{n} e_i = 0$$

Matrixschreibweise:

$$\boldsymbol{\iota}'\mathbf{e} = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \qquad \boldsymbol{\iota} = \text{Einheitsvektor}$$

Beweis:

$$\begin{aligned} \boldsymbol{\iota}'\mathbf{e} &= \boldsymbol{\iota}'(\mathbf{y} - \hat{\mathbf{y}}) \\ &= \boldsymbol{\iota}'(\mathbf{y} - \mathbf{Xb}) \\ &= \boldsymbol{\iota}'\mathbf{y} - \boldsymbol{\iota}'\mathbf{Xb} \end{aligned}$$

Nebenrechnung :

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \Leftrightarrow (\mathbf{X}'\mathbf{X})\mathbf{b} = (\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$\Leftrightarrow (\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y} \qquad \text{Normalgleichungen}$$

$$\underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ x_{11} & \cdots & x_{n1} \\ \vdots & \ddots & \vdots \\ x_{1K} & \cdots & x_{nK} \end{pmatrix}}_{X'} \underbrace{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1K} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nK} \end{pmatrix}}_{X} \begin{pmatrix} b_0 \\ \vdots \\ b_K \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ x_{11} & \cdots & x_{n1} \\ \vdots & \ddots & \vdots \\ x_{1K} & \cdots & x_{nK} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \begin{pmatrix} n & \sum_i x_{i1} & \cdots & \sum_i x_{iK} \\ \sum_i x_{i1} & \sum_i x_{i1}^2 & \cdots & \sum_i x_{i1}x_{iK} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_i x_{iK} & \sum_i x_{iK}x_{i1} & \cdots & \sum_i x_{iK}^2 \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_K \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_{i1}y_i \\ \vdots \\ \sum_i x_{iK}y_i \end{pmatrix}$$

Aus der 1. Normalgleichung folgt:

$$\sum_i x_{i0}b_0 + \sum_i x_{i1}b_1 + \ldots + \sum_i x_{iK}b_K = \sum_i y_i$$

$$\begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & x_{11} & \cdots & x_{1K} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nK} \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_K \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\boldsymbol{\iota}'\mathbf{Xb} = \boldsymbol{\iota}'\mathbf{y}$$

$= \boldsymbol{\iota}'\mathbf{y} - \boldsymbol{\iota}'\mathbf{Xb}$

$= \boldsymbol{\iota}'\mathbf{Xb} - \boldsymbol{\iota}'\mathbf{Xb}$

$= 0$ q.e.d.

d) Das arithmetische Mittel $\overline{y}$ der beobachteten y_i-Werte ist gleich dem arithmetischen Mittel $\overline{\hat{y}}_i$ der geschätzten Werte.

Summenschreibweise:

$$\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \frac{1}{n}\sum_{i=1}^{n} y_i \Leftrightarrow \overline{\hat{y}} = \overline{y}$$

Matrixschreibweise:

$$\frac{1}{n}(\boldsymbol{\iota}'\hat{\mathbf{y}}) = \frac{1}{n}(\boldsymbol{\iota}'\mathbf{y})$$

$$\boldsymbol{\iota}'\hat{\mathbf{y}} = \boldsymbol{\iota}'\mathbf{y}$$

Beweis:

$\hat{\mathbf{y}} = \mathbf{Xb}$ (Ausgleichshyperebene)

Aus 3.: $\boldsymbol{\iota}'\mathbf{y} = \boldsymbol{\iota}'\mathbf{Xb}$

$\Rightarrow \boldsymbol{\iota}'\hat{\mathbf{y}} = \boldsymbol{\iota}'\mathbf{Xb} = \boldsymbol{\iota}'\mathbf{y}$ q.e.d.

Anschauliche Bedeutung: Regressionsgerade verläuft durch $\overline{\hat{y}}, \overline{y}$

Zusammenfassung der statistischen Eigenschaften der MKQ/ OLS:

1. $\mathbf{X}'\mathbf{e} = 0$	$\sum_i x_{ik} e_i = 0$
2. $\hat{\mathbf{y}}'\mathbf{e} = 0$	$\sum_i \hat{y}_i e_i = 0$
3. $\boldsymbol{\iota}'\mathbf{e} = 0$	$\sum_{i=1}^{n} e_i = 0$
4. $\boldsymbol{\iota}'\hat{\mathbf{y}} = \boldsymbol{\iota}'\mathbf{y}$	$\overline{\hat{y}} = \overline{y}$

2.2 Geometrische Interpretation der MKQ/ OLS

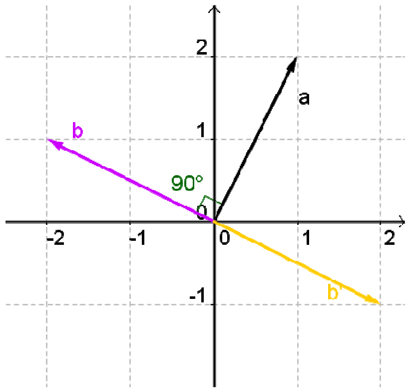
Orthogonalität: Zwei Vektoren **a** und **b** sind orthogonal ($\mathbf{a} \perp \mathbf{b}$, rechtwinklig) dann und nur dann, wenn ihr Skalarprodukt gleich Null ist: $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} = 0$

Beispiel Orthogonalität:

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$\mathbf{a}'\mathbf{b} = 0 \Leftrightarrow \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = 0$$

$$\mathbf{a}'\mathbf{b} = 1 \cdot (-2) + 2 \cdot (1) = 0$$



OLS-Aufgabe:

b ist so zu wählen, dass die Länge von **e** (Vektornorm: $\|\mathbf{e}\| = \sqrt{\mathbf{e}'\mathbf{e}}$) minimiert wird. Die Lösung ist derjenige Vektor b, der **e** orthogonal zu **Xb** macht, also für den $(\mathbf{Xb})'\mathbf{e} = 0$ gilt:

Orthogonalität führt zu den bekannten Normalgleichungen:

$$(\mathbf{Xb})'\mathbf{e} = 0$$

$$(\mathbf{Xb})'\mathbf{e} = (\mathbf{Xb})'(\mathbf{y} - \hat{\mathbf{y}}) = (\mathbf{Xb})'(\mathbf{y} - \mathbf{Xb})$$

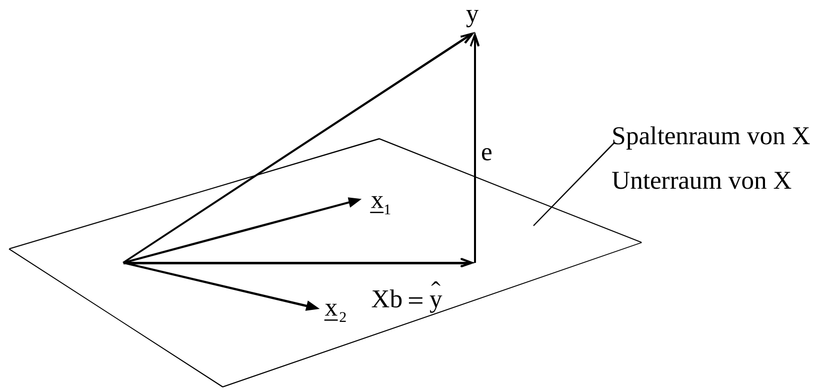
$$= (\mathbf{Xb})'\mathbf{y} - (\mathbf{Xb})'\mathbf{Xb} = \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{Xb}$$

$$= \mathbf{b}'(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{Xb}) = 0$$

Gleichung ist nur Null ($\mathbf{b} \neq 0$) wenn $\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{Xb}$ (= Normalgleichungen OLS)

Damit ist die Orthogonalität die bestimmende OLS-Eigenschaft.

OLS-Eigenschaft und Orthogonalität graphisch:



$$\mathbf{Xb} = \begin{pmatrix} x_1, & x_2, & \dots & x_K \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_K \end{pmatrix} = b_0 x_1 + b_1 x_2 + \ldots + b_K x_K$$

x_k ist der Vektor der beobachteten Werte von der Variablen k. **Xb** ist ein Vektor aus der Linear-kombination $\sum_k b_k x_{ik}$ die den Spaltenraum von $\mathbf{X}_{n\times(K+1)}$ aufspannt. Der Vektor **b**, der **Xb** auf die Länge bringt, bei der **e** orthogonal, also rechtwinklig, ist, minimiert die Fehlerquadratsumme $\mathbf{e'e}$ und damit auch die Länge von **e** ($\|\mathbf{e}\| = \sqrt{\mathbf{e'e}}$).

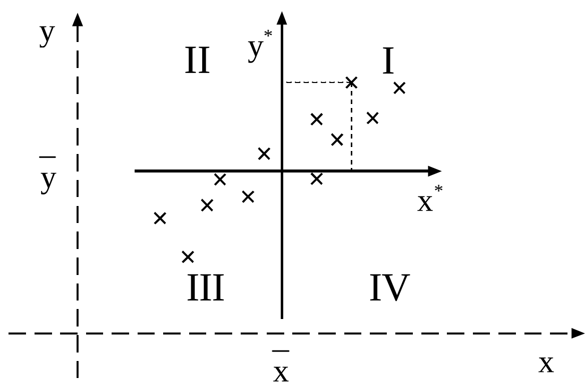
Zur Vertiefung vgl. z.B. Johnston (1985) S.111ff. oder Greene (2003), S.25ff.

3 KORRELATION, VARIANZANALYSE UND "GOODNESS OF FIT"

Gesucht:
Maß für die Güte der Anpassung der Regressionsgeraden an die Beobachtungswerte → Korrelationskoeffizient: Je enger die Punkte sich um die Regressionsgerade scharen, umso strammer ist der Zusammenhang zwischen den beobachteten Werten.

3.1 Korrelationskoeffizient und partielle Korrelation

$$\text{cov}\,(x,y) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}) = \frac{1}{n}\sum_{i=1}^{n} y_i^* \cdot x_i^*$$



(a) positive Korrelation: Punkte größtenteils in I, III; je größer x, desto größer auch y
(b) negative Korrelation: Punkte größtenteils in II, IV; je größer x, desto kleiner y

Produktsumme $\sum_i x_i^* y_i^*$ $\left(x_i^* = x_i - \overline{x}\right)$ umso größer, je enger die Punkte an der Regressionsgeraden liegen.

Produktsumme $\sum_i x_i^* y_i^*$ geht bei einer willkürlichen Verteilung gegen 0.

Problem:

Produktsumme verändert sich, wenn sich Maßeinheit oder Anzahl der Werte verändern → Normierung notwendig in den beiden Punkten:

- Maßeinheit

$$\tilde{x}_i^* = \frac{x_i^*}{s_x} \quad s_x^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} x^{*2} = \mathrm{Var}(x)$$

$$\tilde{y}_i^* = \frac{y_i^*}{s_y} \quad s_y^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} y^{*2} = \mathrm{Var}(y)$$

mit Stichprobenumfang n

Korrelationskoeffizient nach Bravais-Pearson

$$\boxed{\begin{aligned} r &= \frac{1}{n}\sum_i \tilde{x}_i^* \tilde{y}_i^* \\ r &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})\cdot(y_i - \overline{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}} \\ &= \frac{\sum_{i=1}^{n} x_i^* y_i^*}{\sqrt{\sum_{i=1}^{n}(x_i^*)^2}\sqrt{\sum_{i=1}^{n}(y_i^*)^2}} \\ &= \frac{\mathrm{cov}(x,y)}{s_x s_y} = \frac{s_{xy}}{s_x s_y} \end{aligned}}$$

r ist ein Maß für die Stärke des (linearen!) Zusammenhangs: $\boxed{-1 \le r \le +1}$

r nahe bei -1 (+1) → sehr enger negativer (positiver) Zusammenhang

r nahe bei 0 → kein (linearer) Zusammenhang

Partielle Korrelation

Einfache Korrelation:

Linearer Zusammenhang zwischen zwei Variablen

z.B. r_{12} für Korrelation zwischen y und x_1
r_{13} " " y und x_2
r_{23} " " x_1 und x_2

Partielle Korrelation:

Will man die Korrelation direkt nur zwischen zwei Größen unabhängig vom Einfluss einer dritten Größe analysieren, dann verwendet man die partielle Korrelation (z.B. Gujarati 1995, S.212f.).

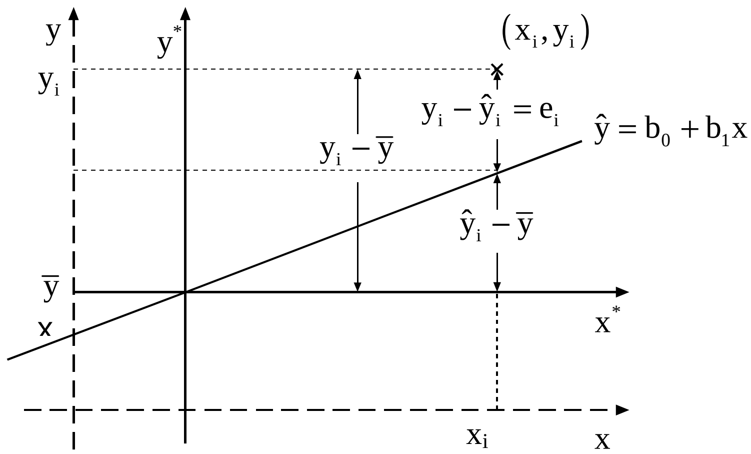
Beispiel: $r_{12.3}$ = partieller Korrelationskoeffizient zwischen y und x_1, wenn x_3 konstant gehalten wird.

$$r_{12.3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}; \qquad r_{13.2} = \frac{r_{13} - r_{12} r_{23}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}; \qquad r_{23.1} = \frac{r_{23} - r_{12} r_{13}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}$$

3.2 ANOVA-Varianzanalyse

Streuungszerlegung

Ziel:
Variation (Streuung) der abhängigen Variablen erklären aus der Variation der unabhängigen Variabeln.



$$\underbrace{(y_i - \bar{y})}_{\substack{\text{zu erklärende Abweichung} \\ \text{(Gesamte Abweichung)}}} = \underbrace{(\hat{y}_i - \bar{y})}_{\substack{\text{erklärte Abweichung} \\ \text{(Regression)}}} + \underbrace{(y_i - \hat{y}_i)}_{\text{nicht erklärte Abweichung}} \qquad (i = 1, \ldots, n)$$

Analog zur Bestimmung der Regressionsfunktion → Abweichungsquadrate:

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

Quadrieren auf beiden Seiten und summieren über alle Beobachtungen:

$$\sum_{i=1}^{n}(y_i-\overline{y})^2 = \sum_{i=1}^{n}\left[(\hat{y}_i-\overline{y})^2-2(\hat{y}_i-\overline{y})(y_i-\hat{y}_i)+(y_i-\hat{y}_i)^2\right]$$

$$\underbrace{\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2}_{A}-\underbrace{2\sum_{i=1}^{n}(\hat{y}_i-\overline{y})(y_i-\hat{y}_i)}_{2B}+\underbrace{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}_{C}$$

$$B=\sum_{i=1}^{n}(\hat{y}_i-\overline{y})(y_i-\hat{y}_i)=\sum_{i=1}^{n}(\hat{y}_i-\overline{y})e_i \qquad \left|e_i=(y_i-\hat{y}_i)\right.$$

$$=\underbrace{\sum_{i=1}^{n}e_i\hat{y}_i}_{=0}-\overline{y}\underbrace{\sum_{i=1}^{n}e_i}_{=0}$$

$\left|\text{Eigenschaft 2}: \sum_i e_i\hat{y}_i = 0 \text{ und}\right.$

$\left|\text{Eigenschaft 3}: \sum_i e_i = 0\right.$

→ für eine lineare KQ-Regressionsfunktion (Verwendung der Eigenschaften 1 und 3) gilt die einfache Zerlegung der Streuung

$$\underbrace{\sum_{i=1}^{n}(y_i-\overline{y})^2}_{\substack{\text{Quadratsumme der zu}\\ \text{erklärenden Abweichung}}}=\underbrace{\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2}_{\substack{\text{Quadratsumme der}\\ \text{erklärten Abweichung}}}+\underbrace{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}_{\substack{\text{Quadratsumme der nicht}\\ \text{erklärten Abweichungen,}\\ \text{(der Residuen } e_i\text{)}}}$$

$$\text{SQT} = \text{SQE} + \text{SQR}$$

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=b_1^2\sum_{i=1}^{n}(x_i^*)^2+\sum_{i=1}^{n}e_i^2$$

Anmerkung: Für andere lineare Regressionsfunktionen, die nicht nach OLS ermittelt wurden, gilt diese einfache Zerlegung nicht!

Alternative Notation:

SQT		SQE		SQR
SST	=	SSE	+	SSR
TSS		ESS		RSS
Total Sum of Squares		Explained Sum of Squares		Residual Sum of Squares

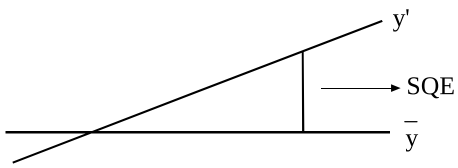
3.3 Goodness of Fit: R^2 (Bestimmheitsmaß) und korrigiertes R^2

R^2 ist ein Maß dafür, wieviel Variation der abhängigen Variablen durch die Variation der unabhängigen Variablen erklärt wird.

Lineares, einfaches Bestimmtheitsmaß

$$R^2=B_{y,x}=\frac{\text{SQE}}{\text{SQT}}=\frac{\sum(\hat{y}-\overline{y})^2}{\sum(y-\overline{y})^2}=\frac{\text{erklärte}}{\text{gesamte}}\text{Abweichungsquadratsumme}$$

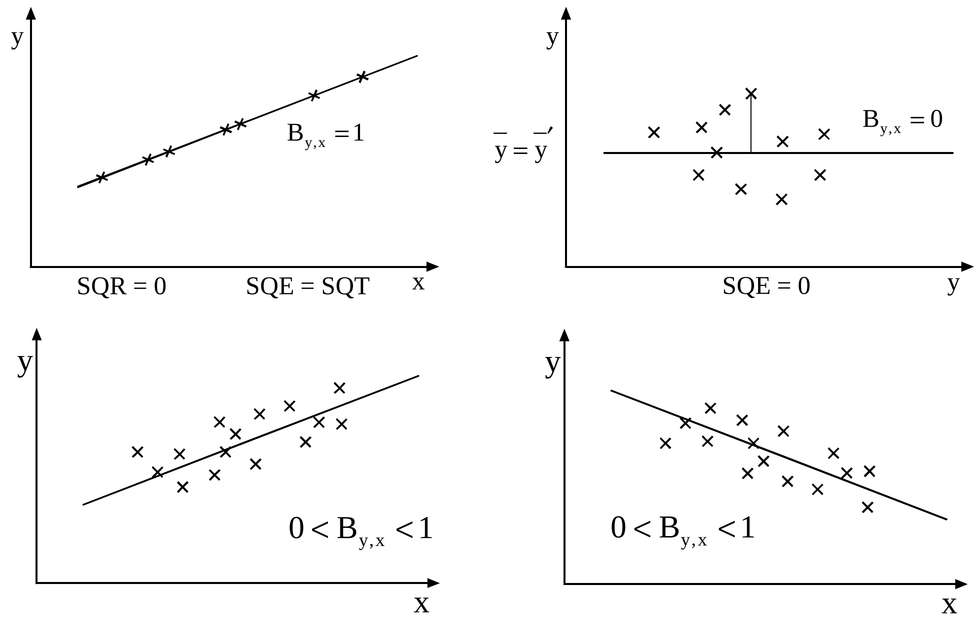
$$0\le B\le 1$$



Je näher $B_{y,x}$ an 1, desto

- besser passt sich die Regression den Beobachtungen an
- größer ist die durch die Regression erklärte Gesamtstreuung
- geringer ist die Streuung der Punkte um die Regressionsgerade

$B_{y,x}$ nahe bei 0: Kein Beitrag zur Erklärung der Gesamtstreuung



Das Bestimmtheitsmaß ist also ein Maß für die Anpassungsgüte der Gesamtregression ("Goodness of Fit").

Bestimmtheitsmaß B und Korrelationskoeffizient R

Behauptung:

$$B = r^2$$

Beweis:

$$B = \frac{s_{\hat{y}}^2}{s_y^2} = \frac{\frac{1}{n}\sum_i \left(\hat{y} - \bar{y}\right)^2}{\frac{1}{n}\sum_i \left(y - \bar{y}\right)^2} = \frac{SQE}{SQT} = \frac{\sum_i \left(\hat{y}_i^*\right)^2}{\sum_i \left(y_i^*\right)^2}$$

$$\hat{y}_i^* = b_1 x_i^* \text{ und } b_1 = \frac{\sum_i x_i^* y_i^*}{\sum_i \left(x_i^*\right)^2}$$

$$B = \frac{\sum_i \left(b_1 x_i^*\right)^2}{\sum_i \left(y_i^*\right)^2} = \frac{b_1^2 \sum_i \left(x_i^*\right)^2}{\sum_i \left(y_i^*\right)^2} = \underbrace{\frac{\left(\sum_i x_i^* y_i^*\right)^2}{\left(\sum_i \left(x_i^*\right)^2\right)^2}}_{b_1^2} \frac{\sum_i \left(x_i^*\right)^2}{\sum_i \left(y_i^*\right)^2}$$

$$B = \frac{\sum_i x_i^* y_i^*}{\sum_i \left(x_i^*\right)^2 \sum_i \left(y_i^*\right)^2} = r^2 \quad \text{q.e.d.}$$

$$B = r^2 = \frac{SQT - SQR}{SQT} = 1 - \frac{SQR}{SQT} = 1 - \frac{\sum_i e_i^2}{\sum_i \left(y_i - \bar{y}\right)^2}$$

B ist bedeutungsvoller in der Regressionsrechnung als r (r als 2. Wurzel von $B = r^2$ hat kein eindeutiges Vorzeichen), da B den Anteil der erklärten Variation angibt.

Praktische Berechnung von B:

$$\begin{aligned}
&\sum e_i^2 = \sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i \\
&\sum \left(y_i - \bar{y}\right)^2 = \sum \left(y_i^2 - 2 y_i \bar{y} + \bar{y}^2\right) = \sum y_i^2 - 2 \sum y_i \bar{y} + \sum \bar{y}^2 \\
&= \sum y_i^2 - 2n\bar{y}\bar{y} + n\bar{y}^2 = \sum y_i^2 - n\bar{y}\bar{y} = \sum y_i^2 - n\left(\tfrac{1}{n} \sum y_i\right)^2 \\
&= \sum y_i^2 - \tfrac{1}{n}\left(\sum y_i\right)^2 = \sum y_i^2 - n\bar{y}^2 \\
&B = R^2 = f\left(b_0, b_1, \sum y_i, \sum y_i^2, \sum x_i y_i\right)
\end{aligned}$$

→ Rechenersparnis: Statt zweimaligem Stichprobendurchlauf erst für $\bar{y}$, dann für $\left(y_i - \bar{y}\right)^2$ nur einmaliger Stichprobendurchlauf für $\left(y_i^2, \sum y_i\right)$.

Bei multipler Regression kennzeichnet $B = r^2_{y(y\,\text{geschätzt})}$, also die Korrelation zwischen den beobachteten und geschätzten y-Werten.

Korrigiertes Bestimmtheitsmaß $\bar{R}^2$ oder R^2_{adj}

Fügt man zu einem Regressionsmodell eine oder mehrer erklärende Variablen hinzu, so steigt R^2 oder bleibt zumindest konstant.

R^2 nützt also wenig, wenn man beurteilen will, ob die Hereinnahme einer weiteren Variablen eine Bedeutung bei der Erklärung der abhängigen Variable hat. Für diese Korrektur hat sich in der Praxis das korrigierte Bestimmtheitsmaß $\bar{R}^2$ oder R^2_{adj} ("adj" für "adjusted") durchgesetzt:

$$\begin{aligned}
\bar{R}^2 = R^2_{adj} &= 1 - \frac{n-1}{n-K-1}\left(1 - R^2\right) \qquad K = \text{Anzahl der Regressoren} \\
&= R^2 - \frac{K}{n-K-1}\left(1 - R^2\right)
\end{aligned}$$

$\bar{R}^2$ ist genauso zu interpretieren wie das Bestimmtheitsmaß R^2. $\bar{R}^2$ liegt zwischen 0 und 1, je näher der 1 desto besser wird die Varianz der y- Werte durch die Regression erklärt.

Weitere Informationen: Goldberger (1964), S.217; Hübler (1989), S.56.

Beispiel: R^2, Bestimmtheitsmaß

BSP-/ Geldmengenbeispiel:

$$\widehat{BSP}_t = 1{,}168 + 1{,}715 \cdot \text{Geldmenge}_t \qquad (t = 1, \ldots, T)$$
$$\hat{y}_i = b_0 + b_1 \cdot x_t$$

Bestimmtheitsmaß:

$$B_{y,x} = \frac{SQE}{SQT} = \frac{\sum_i (\hat{y}_i - \overline{y})^2}{\sum_i (y_i - \overline{y})^2} = r^2 = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \overline{y})^2}$$

$$\sum_i e_i^2 = \sum_i y_i^2 - b_0 \sum_i y_i - b_1 \sum_i x_i y_i$$

$$\sum_i (y_i - \overline{y})^2 = \sum_i y_i^2 - \frac{1}{n}\left(\sum_i y_i\right)^2$$

bekannt:

$$(X'X) = \begin{pmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{pmatrix} = \begin{pmatrix} 10 & 37{,}2 \\ 37{,}2 & 147{,}18 \end{pmatrix}$$

$$X'y = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{pmatrix} = \begin{pmatrix} 75{,}5 \\ 295{,}95 \end{pmatrix}$$

fehlt $\sum_i y_i^2$; $\quad \sum_i y_i^2 = 597{,}03$

Berechnung:

$$\sum_i e_i^2 = 597{,}03 - 1{,}168 \cdot 75{,}5 - 1{,}715 \cdot 295{,}95 = 1{,}2918$$

$$\sum_i (y_i - \overline{y})^2 = 597{,}03 - \frac{1}{10} 75{,}5^2 = 27{,}005$$

$$R^2 = B = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \overline{y})^2} = 1 - \frac{1{,}2918}{27{,}005} = 0{,}9522$$

$$\overline{R}^2 = R^2_{adj} = 1 - \frac{n-1}{n-K-1}\left(1 - R^2\right) = 1 - \frac{9}{8}(1 - 0{,}9522) = 0{,}9462$$

D.h. 95,2 bzw. 94,6% der Gesamtvarianz des BSP wird durch die Regressionsfunktion erklärt.

3.4 Goodness of Fit: Akaike Information, Amemiya Prediction und Schwarz Information Kriterium

Das Akaike Information Criterion (AIC), das Amemiya Prediction Criterion (APC) und das Schwarz Information Criterion (SIC) sind dem korrigierten R^2 ähnlich; sie basieren auf der Summe der quadrierten Residuen aber bestrafen sozusagen den Verlust von Freiheitsgraden wenn mehrere Variablen zur Erklärung herangezogen werden.

$APC = s^2[1 + (K\text{-}1)/n]$, wobei $s^2 = \mathbf{e}'\mathbf{e}/(n\text{-}K)$.

$AIC = \ln(\mathbf{e}'\mathbf{e}/n) + 2(K\text{-}1)/n$

$SIC = \ln(\mathbf{e}'\mathbf{e}/n) + (K\text{-}1)/n \ln n$.

AIC und SIC sind sehr ähnlich, SIC 'bestraft' etwas schärfer. Kleine Werte beider Maße sind wünschenswert. Weitere Information finden sich in Auer (2003, S. 255 ff), Greene (2008, S. 142 ff).

3.5 Übungsaufgaben Mehrfachregression

Schädlingsbekämpfung: Mehrfachregression

Nachdem Sie den Ernteertrag des benachbarten Landwirtes als Funktion des ausgebrachten Schädlingsbekämpfungsmittels geschätzt haben, fällt Ihnen auf, dass auch das Wetter einen Einfluss gehabt haben könnte. Zufällig liegen hierzu auch Daten vor.

Zahl der Sonnenstunden x_1, Mengen ausgebrachten Schädlingsbekämpfungsmittels x_2 [in l] und Weizenproduktion y [in hl]:

x_1	x_2	y
1508	0,93	154
1721	1,1	164
1476	1,2	175
1630	1,3	193
1278	1,95	174
1321	2,55	188
1176	2,6	176
1123	2,9	198
1374	3,45	239
1546	3,5	256
1666	3,6	253
1475	4,1	278
1765	4,35	301
1232	4,83	278

a) Gehen Sie wiederum davon aus, dass die Produktionsfunktion in den Argumenten x_1 und x_2 linear sei. Schätzen Sie nun im erweiterten Modell die Parameter b_0, b_1 und b_2 nach der KQ-Methode.
b) Geben Sie eine ökonomische Interpretation der geschätzten Parameter.
c) Vergleichen Sie das Bestimmtheitsmaß dieser Regression mit dem aus der Einfachregression.

Einkommen: Einfach- und Mehrfachregression

Das Einkommen soll als Funktion des Alters berechnet werden:

$$\text{income} = b_0 + b_1 * \text{age}$$

Berechnen Sie anhand der Daten aus Beispiel 3 (ET: $\text{income} = b_0 + b_1 * \text{age} + b_2 * \text{sex}$)

- mit der Summenformel aus der Einfachregression sowie
- mit der Matrizenrechnung die OLS-Koeffizienten
- Interpretieren Sie Ihr Ergebnis.

V Statistische Grundlagen: Wahrscheinlichkeit, Schätzeigenschaften und statistische Inferenz

Dieses Kapitel fasst grundlegende, für das stochastische Modell notwendige Begriffe und Konzepte der Wahrscheinlichkeitsrechung und der schließenden Statistik zusammen.

Im Literaturverzeichnis sind mit fast allen Ökonometrielehrbüchern auch Grundlagen der Hypothesentests der statistischen Inferenz sowie wünschenswerte Eigenschaften von Schätzern gegeben.

1 Zufallsvariablen und Wahrscheinlichkeitsverteilung

Beobachtung bestimmter Aspekte von Wirtschaft und Gesellschaft ist das Ergebnis eines zufälligen Prozesses.

Begriffe:

Zufallsvariable X:	große Buchstaben	"random variable"
Ausprägung x:	kleine Buchstaben	
Prob (X=x):	Wahrscheinlichkeit, dass X die Ausprägung x annimmt	
Diskrete Zufallsvariable:	endliche, zählbare Menge der Ergebnisse	
Kontinuierliche Zufallsvariable:	Menge der Ergebnisse ist unendlich, teilbar, nicht zählbar	

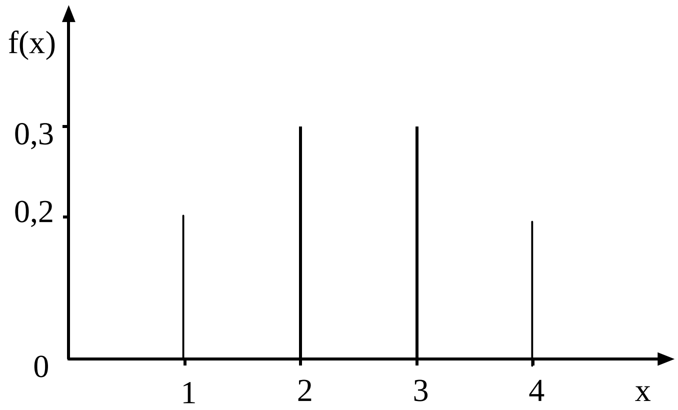
Der Prozess, der die Ausprägungen der Zufallsvariablen generiert, kann durch seine Wahrscheinlichkeitsfunktion beschrieben werden. Die **Wahrscheinlichkeitsfunktion ("probability function") f(x)** listet alle Ergebnisse x, die die Zufallsvariable X annimmt, mit ihren Wahrscheinlichkeiten des Eintritts auf.

Diskrete Zufallsvariable:

$$f(x_i) = \Pr ob(X = x_i) = P(X = x_i) = p_i$$

P ist die Wahrscheinlichkeit, mit der X die Ausprägung x_i annimmt.

1. $0 \le f(x_i) \le 1$
2. $f(x_i) \ge 0$
3. $\sum_i f(x_i) = 1$

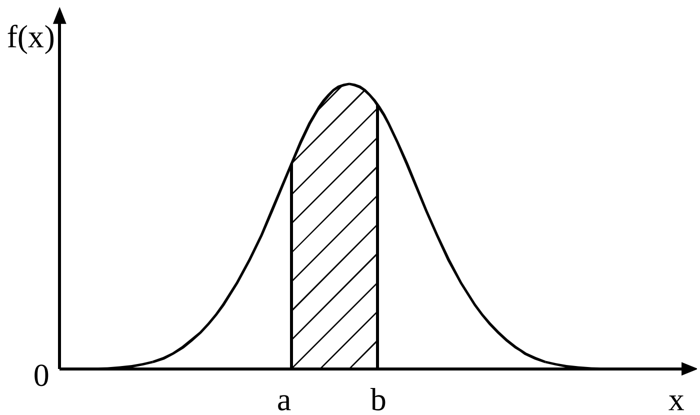


Kontinuierliche Zufallsvariable (nur für Intervalle möglich):

$$\Pr ob(a \leq x \leq b) = \int_a^b f(x)\,dx \geq 0$$

Fläche unter der Dichtefunktion in den Grenzen a und b gibt die Wahrscheinlichkeit an. Wahrscheinlichkeitsdichte (Dichtefunktion) = probability distribution function (pdf)

1. $f(x) \geq 0$
2. $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

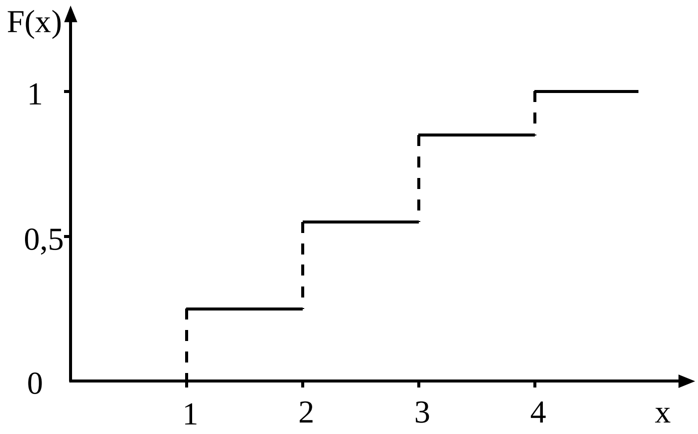


Die **Verteilungsfunktion F(x)** (cumulative distribution function, cdf) gibt die Wahrscheinlichkeit an, dass die Zufallsvariable X höchstens den Wert x annimmt.

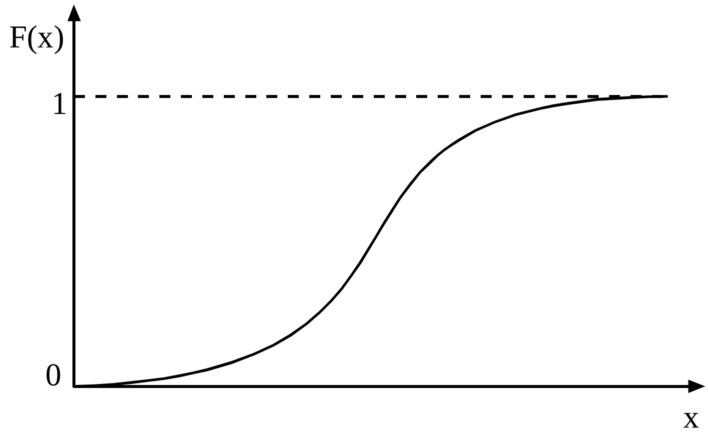
Diskrete Zufallsvariable:

$$F(x) = P(X \leq x) = \sum_{X \leq x_i} f(x_i)$$

$$f(x_i) = F(x_i) - F(x_{i-1})$$



Kontinuierliche Zufallsvariable



$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

Eigenschaften

$$0 \leq F(x) \leq 1$$

$$x > y,\ F(x) \geq F(y)$$

$$F(+\infty) = 1$$

$$F(-\infty) = 0$$

$$\text{Prob}(a \leq x \leq b) = F(b) - F(a)$$

2 Erwartungswert und Eigenschaften des Erwartungswertoperators

Beschreibung der Wahrscheinlichkeitsfunktionen oft durch Mittelwert (mean) und Varianz (variance) über den Erwartungswertoperator E einer Zufallsvariablen X.

Erwartungswert, Mittelwert

$$\mathrm{Var}(x) = \sigma_x^2 = \sum_i p_i \left(x_i - E(x)\right)^2$$

$$E(X) = \mu_x = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_N f(x_N) = \sum_i f(x_i) x_i$$

x_i mögliche Ergebinsse von X

$$f(x_i) = \mathrm{Prob}(X = x_i) = p_i$$

$$\sum_i f(x_i) = 1$$

$$E(X) = \begin{cases} \sum_i x_i \; f(x_i) & x_i \text{ diskret} \\ \int x \; f(x) dx & x \text{ kontinuierlich} \end{cases}$$

Varianz

$$\mathrm{Var}(X) = \begin{cases} \sum_i (x_i - \mu)^2 f(x_i) = \sum_i x_i^2 f(x_i) - \mu^2 & x_i \text{ diskret} \\ \int (x - \mu)^2 f(x) dx = \int x^2 f(x) dx - \mu^2 & x \text{ stetig} \end{cases}$$

Die Varianz ist also ein gewichtetes Mittel der Abweichungsquadrate der Ergebnisse von X von ihrem erwarteten Wert (Gewichte = Eintrittswahrscheinlichkeiten)

$$\mathrm{Var}(x) = \sigma_x^2 = E\left(X - E(X)\right)^2 = E\left(X^2\right) - \mu^2 \qquad \text{(Verschiebungssatz)}$$

Eigenschaften des Erwartungswertoperators:

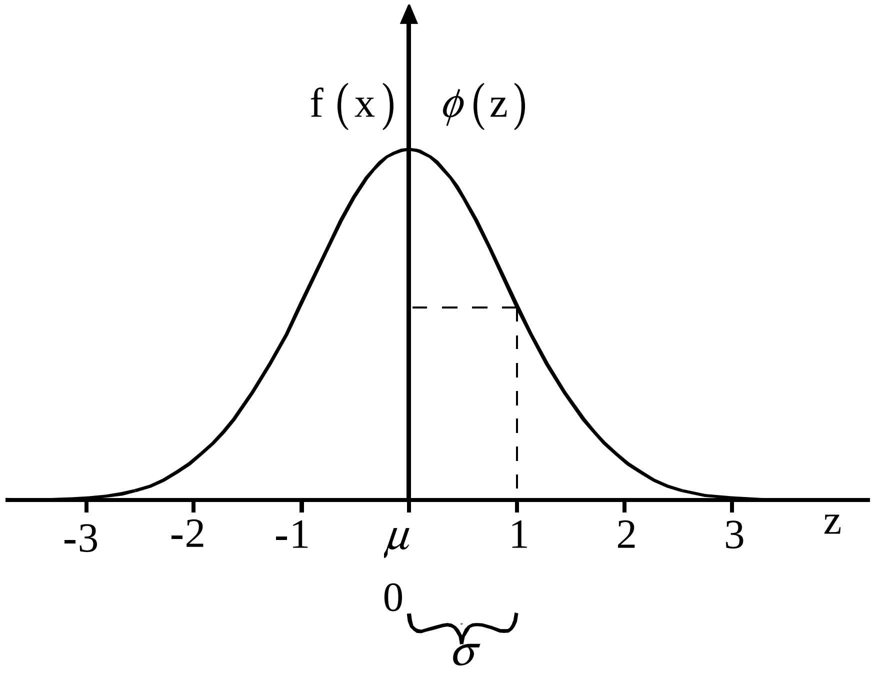
1. $E(aX + b) = aE(X) + b$ a, b Konstanten
2. $E\left[(aX)^2\right] = a^2 E\left(X^2\right)$
3. $\mathrm{Var}(aX+b) = a^2 \mathrm{Var}(X)$
4. $E(X+Y) = E(X) + E(Y)$ X, Y Zufallsvariblen
5. $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(XY)$

Bei Unabhängigkeit zweier Zuvallsvariabler X und Y

6. $E(XY) = E(X)E(Y)$
7. $\mathrm{Cov}(XY) = 0$

3 Spezifische Verteilungen von Zufallsvariablen

3.1 Normalverteilung



$$f_N\left(x \middle| \mu, \sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \mu = \text{Mittelwert}, \ \sigma^2 = \text{Varianz der Grundgesamtheit}$$

X heißt normalverteilt mit den Parametern μ und σ^2: $X \sim N\left(\mu, \sigma^2\right)$

wichtige Transformationseigenschaft: wenn $X \sim N\left(\mu, \sigma^2\right)$, dann $a + bX \sim N\left(a + b\mu, b^2\sigma^2\right)$

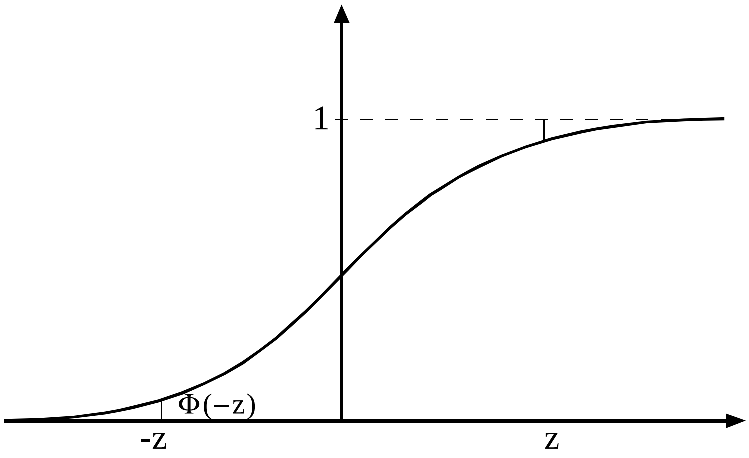
Standardnormalverteilung

$$Z = \frac{X - \mu}{\sigma} \qquad \left(a = \frac{-\mu}{\sigma}; b = \frac{1}{\sigma}\right)$$

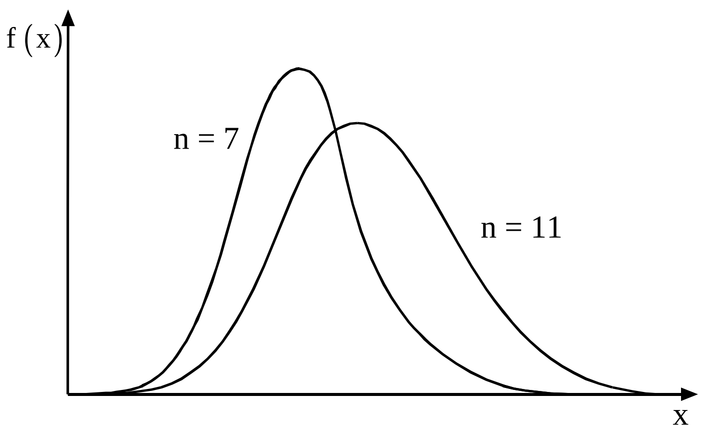
$$\phi(z) = f_N(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

Symmetrische Verteilung um 0: $\phi(z) = \phi(-z)$

cdf: $\Phi(-z) = 1 - \Phi(z)$



3.2 Chi-Quadrat Verteilung (χ^2)



Abgeleitet aus der Normalverteilung. Verwendung: Varianztests

Sind (Z_1,Z_2,...,Z_n) unabhängig standardnormalverteilte Zufallsvariablen, so hat die Quadratsumme $\sum_i Z_i^2$ eine $\chi^2(n)-$ Verteilung (n = Freiheitsgrade, DF: Degrees of Freedom).

3.3 t-Verteilung (Student-Verteilung)

Abgeleitet aus der Normalverteilung. Verwendung für Signifikanztests der Parameter in ökonometrischen Modellen (wenn die Varianz von X unbekannt ist).

Sind zwei Zufallsvariablen Z und Y voneinander unabhängig verteilt mit $Z \sim N(0,1)$ und $Y \sim \chi^2(\nu)$, so ist die Zufallsvariable $T = \dfrac{Z}{\sqrt{Y/\nu}} \sim t(\nu)$ (t-verteilt mit ν Freiheitsgraden).

3.4 F-Verteilung

Abgeleitet aus Normalverteilung. Verwendung: Test auf Gleichheit zweier Varianzen.

Sind X_1 und X_2 zwei unabhängige χ^2-verteilte Variablen mit ν_1, bzw. ν_2 Freiheitsgraden, dann ist der Quotient $F = \frac{X_1/\nu_1}{X_2/\nu_2} \sim F(\nu_1, \nu_2)$ (F-verteilt mit ν_1, bzw. ν_2 Freiheitsgraden).

weitere Verteilungen: Lognormal, Gamma, Beta, ...

Gemeinsame Verteilungen (joint distribution) für zwei und mehr Variablen (bivariate und multivariate Verteilungen).

4 SCHÄTZUNGEN UND WÜNSCHENSWERTE SCHÄTZEIGENSCHAFTEN

Die Parameter der Grundgesamtheit sind meist unbekannt, daher versucht man sie auf Basis der Stichprobenparameter zu schätzen.

4.1 Punktschätzung

Geg.:
Stichprobenwerte x_1, x_2, ..., x_n, die als Realisationen von Zufallsvariablen X_1, X_2, ..., X_n aufgefasst werden.

Gesucht:
Wahrer, aber unbekannter Parameter der Verteilung der Grundgesamtheit θ (Theta).

Der Schätzwert (Näherungswert, Punktschätzung) für θ ist eine Funktion der Stichprobenergebnisse (Stichprobenfunktion):

$$\hat{\theta} = f(x_1, x_2, \ldots x_n)$$

Dieser Schätzwert kann als Realisation einer speziellen Stichprobenfunktion, der Schätzfunktion, interpretiert werden:

$$\theta = f(X_1, X_2, \ldots X_n)$$

Allgemeine Prinzipien zur Festlegung dieser Schätzfunktion

→ Schätzprinzipien und wünschenswerte Schätzeigenschaften

Schätzprinzipien:

- Methode der Momente (empirische Momente als Schätzer für die Momente der Zufallsvariablen)
- Maximum Likelihood (ML)
- MKQ/ OLS
- Bayes
- Minimum - Chi-Quadrat
- Minimum – Distanz

4.2 Wünschenswerte Schätzeigenschaften für kleine Stichproben: Erwartungstreue, Effizienz und BLUE

Problem:
Güte einer Schätzfunktion? Wahl der besten Schätzfunktion, wenn verschiedene Schätzfunktionen gegeben sind.

a) Erwartungstreue

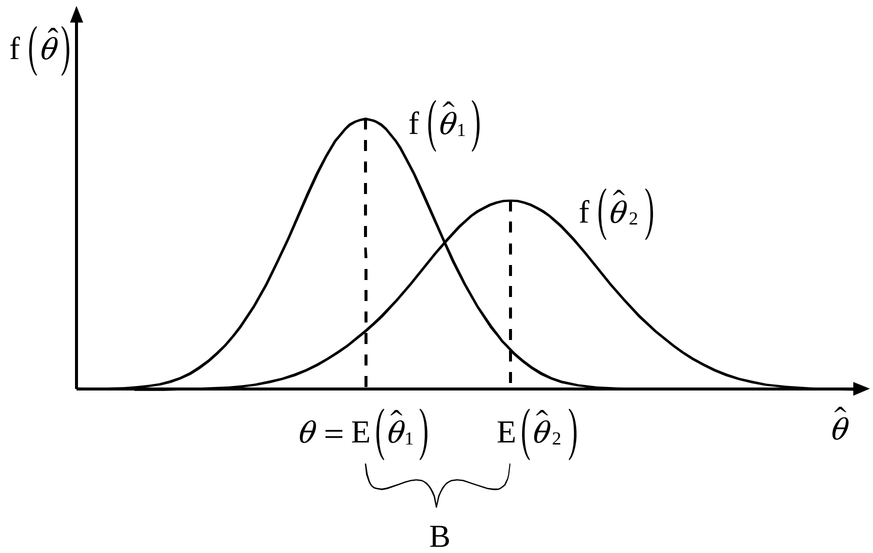
Schätzfunktion $\hat{\theta}$ für unbekannten Parameter θ heißt erwartungstreu (unverzerrt), wenn gilt:

$$\boxed{E\left(\hat{\theta}\right) = \theta} \qquad \text{z.B. } E\left(b_{OLS}\right) = b$$

Erwartungstreue (Unverzerrtheit):

Ein Schätzer ist erwartungstreu, wenn die Abweichung der Schätzwerte $\hat{\theta}$ vom wahren Wert θ im Mittel aller möglichen Schätzwerte gleich Null ist.

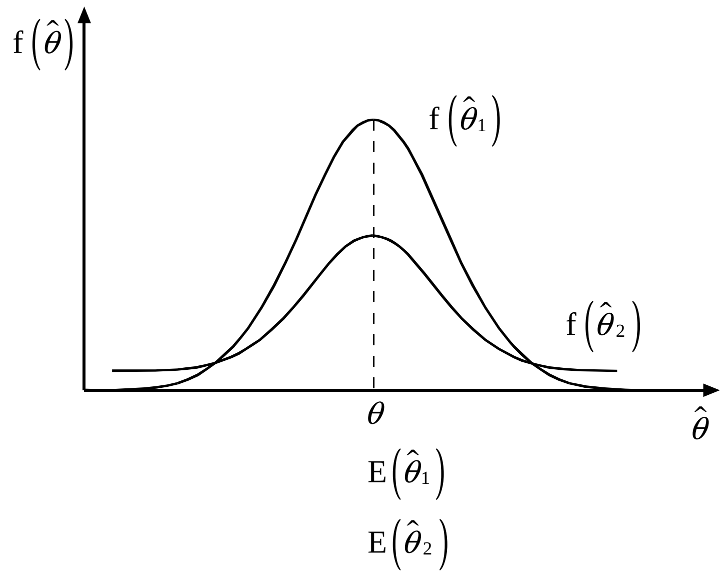
Verzerrung, Bias $B = E\left(\hat{\theta}\right) - \theta =$ 0, d.h. erwartungstreu.



Schätzungen von θ aus verschiedenen Stichproben; hier nur $\hat{\theta}_1$ erwartungstreu mit $E(\hat{\theta}_1) = \theta$

b) Effizienz (minimale Varianz)

Es kann mehrere erwartungstreue Schätzfunktionen $E(\hat{\theta}) = \theta$ geben, die sich in ihrer Varianz unterscheiden. Das heißt, dass die Wahrscheinlichkeiten für Abweichungen vom Mittelwert für die gleichen Abweichungen verschieden groß sein können.



$E(\hat{\theta}_1 - \theta)^2 \leq E(\hat{\theta}_2 - \theta)^2$ $\qquad$ $\sigma^2_{\hat{\theta}_1} \leq \sigma^2_{\hat{\theta}_2}$ $\qquad$ Vor.: $E(\hat{\theta}_1) = \theta$

$E(\hat{\theta}_2) = \theta$

relative Effizienz:

Die erwartungstreue Schätzfunktion $\hat{\theta}_1$ für den unbekannten Parameter θ ist relativ effizient, wenn ihre Streuung (Varianz) kleiner ist als die Streuung eines anderen erwartungstreuen Schätzers $\hat{\theta}_2$.

absolute Effizienz:

... Streuung kleiner als die Streuung aller anderen ...

BLUE-Eigenschaft:

"best linear unbiased estimator"

Wünschenswert ist eine Schätzfunktion für den Parameter θ mit der BLUE-Eigenschaft

- best linear: aus der Klasse der linearen Schätzer den Schätzer mit der kleinsten Varianz wählen
- unbiased: unverzerrt/ erwartungstreu

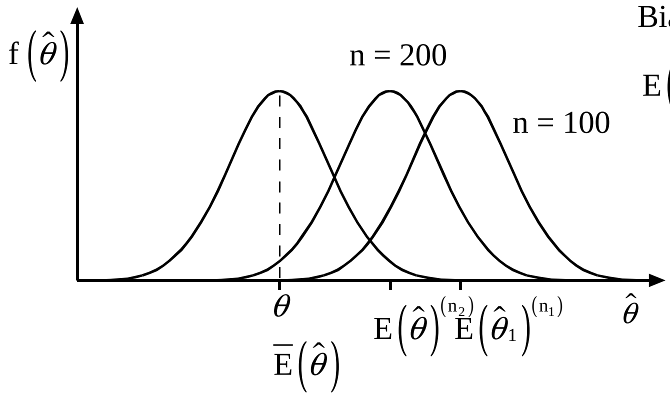
4.3 Wünschenswerte Schätzeigenschaften für große Stichproben: Asymptotische Erwartungstreue und Konsistenz

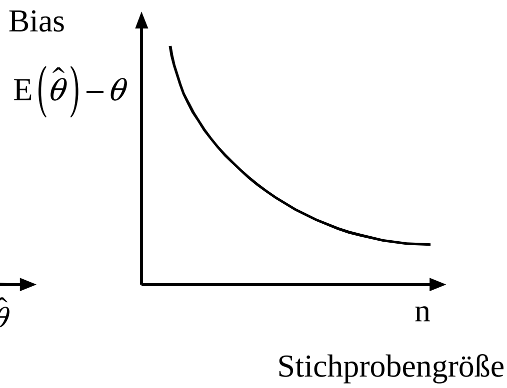
c) Asymptotische Erwartungstreue

Erwartungswerte von immer größer werdenden Stichproben weichen immer weniger von θ ab.

$$\boxed{\overline{\mathrm{E}}\left(\hat{\theta}\right) = \lim_{n\to\infty} \mathrm{E}\left(\hat{\theta}^{(n)}\right) = \theta} \qquad \text{n = Stichprobengröße}$$

Eine asymptotisch erwartungstreue Schätzfunktion ist nicht notwendigerweise erwartungstreu, aber eine erwartungstreue Schätzfunktion ist auch asymptotisch erwartungstreu.





d) Konsistenz

Eine Schätzfunktion $\hat{\theta}$ für θ heißt konsistent, wenn Verzerrung und Varianz mit wachsendem Stichprobenumfang gegen 0 streben

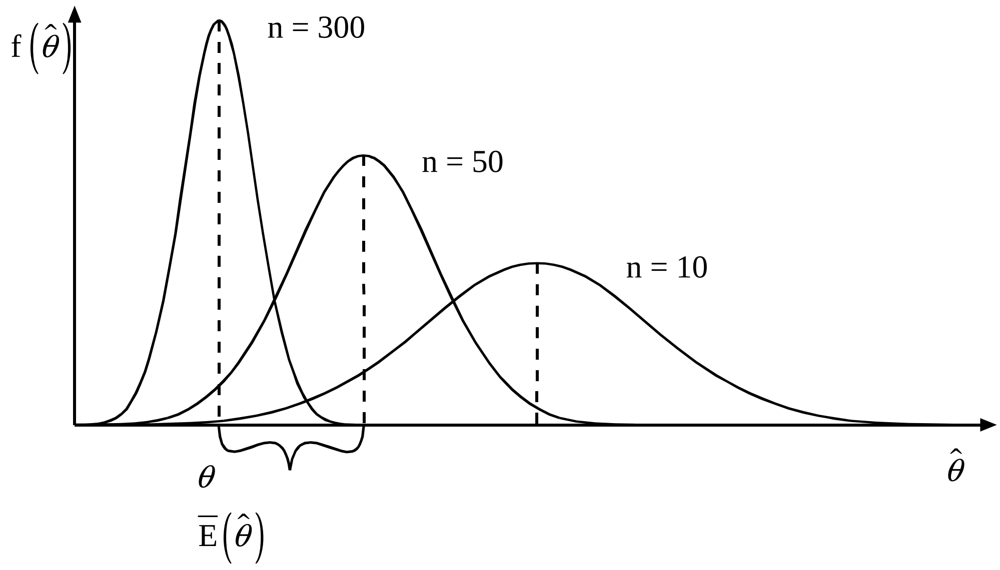
$$\boxed{\mathrm{p}\lim\left(\hat{\theta}\right) = \theta}$$

$$\lim_{n\to\infty} \text{ Prob } \left\{\left|\hat{\theta}^{(n)} - \theta\right| \geq \delta\right\} = 0$$

$$\lim_{n\to\infty} \mathrm{B} = \lim_{n\to\infty} \mathrm{E}\left(\hat{\theta} - \theta\right) = 0$$

$$\text{und } \lim_{n\to\infty} \mathrm{var}\left(\hat{\theta}\right) = \lim_{n\to\infty} \mathrm{E}\left(\hat{\theta} - \theta\right)^2 = 0$$

Verteilung bricht auf einen Punkt zusammen:



Eine konsistente Schätzfunktion ist auch asymptotisch erwartungstreu, aber eine asymptotisch erwartungstreue Schätzfunktion ist nicht notwendigerweise konsistent.

VI Das Klassische Lineare Regressionsmodell (CLR)

Nach der zunächst nur beschreibenden, empirischen Regression der letzten Kapitel behandelt das klassische lineare Regressionsmodell (Classical Linear Regression, CLR) die Übertragung der Stichprobenergebnisse auf eine dahinter liegende Grundgesamtheit unter stochastischen Gesichtspunkten. Erst dieses stochastische Modell erlaubt dann Aussagen über die Signifikanz des Ansatzes hinsichtlich der ‚wahren' Relationen aus der Grundgesamtheit. Anschließend werden die Schätzeigenschaften des KQ-Schätzers, insbesondere die ‚Best Linear Unbiased Estimator' (BLUE)-Eigenschaft dargestellt. Den Abschluss des Kapitels bilden die Schätzung der Parameter nach dem Maximum-Likelihood-Prinzip und Prognoseeigenschaften der KQ-Schätzung.

Das klassische lineare Regressionsmodell ist die zentrale Grundlagen der Ökonometrie, das dann zahlreiche Erweiterungen (bspw. mit diskreten/kategorialen Variablen als zu erklärende Seite) erfahren hat

Im Literaturverzeichnis sind für das klassische lineare Regressionsmodell (stochastisches Modell) wie auch schon für die empirische Einfach- wie Mehrfachregression (deskriptives Modell) vielfältige Quellen gegeben. An neuerer englischsprachiger Literatur sind Greene (2008), Wooldridge (2006, 2001) und Studenmund (2006) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

Mit dem **FFB e-learning Modul**: **Lineare Regression – Stochastisches Modell** (www.leuphana.de/ffb) (Merz und Stolze 2010) wird eine Einführung in das klassische lineare Regressionsmodell gegeben (siehe auch das FFB e-learning Modul: Parametertests).

1 Formulierung des stochastischen Modells

Ziel: Aussagen über die Grundgesamtheit auf Basis der Stichprobe

Stochastisches (Probabilistisches) Modell:

Für eine bestimmte Beobachtung x (unabhängige, erklärende Variable) gibt es viele mögliche Beobachtungen y (abhängige Variable)

$c = f(x)$ x = Einkommen

z.B. Umfrage für ein bestimmtes Einkommen (z.B. x = 47.000 € im Jahr) gibt es viele Personen/ Hauhalte die in unterschiedlicher Höhe Konsumausgaben für Nahrungsmittel haben (y = Ausgaben für Essen und Trinken).

Mögliche Gründe: Unterschiedliche Regionale Bedingungen, Präferenzen, etc.

Solange kein zusätzliches Wissen vorhanden ist (z.B. über sozioökonomische Hintergründe), wird angenommen, dass für jede Beobachtung x (Einkommen), die Beobachtungen y (Konsumausgaben) sich nur zufällig unterscheiden.

→ "Mittlerer" Zusammenhang; Störungen, nicht erfasste Einflüsse, Fehlerterm

Additiver Fehlerterm ("error term") ε (oder μ)

ε = Zufallsvariable (griechische Buchstaben stehen für Zufallsgrößen)

Stochastischer Zusammenhang:

$$y_i = f(x_i) + \varepsilon_i$$

Der Fehlerterm ε fängt die Einflüsse auf, die ich nicht erklären kann.

Stichprobenmodell:

x = gegeben, nicht stochastisch, kann man messen

y = zufällige Auswahl über ε

Urnenmodell:

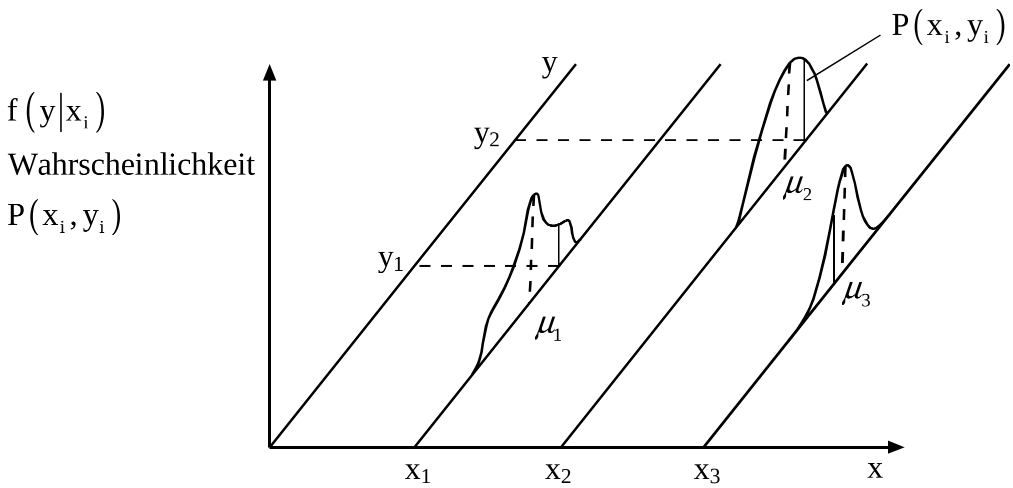
Es gibt n Urnen (i = 1, ..., n).

Jede Urne entspricht einem fest vorgegebenen x_i.

Stichprobenziehung:

	1. Urne	2. Urne	...	n. Urne
Ziehung 1	x_1 y_{11}	x_2 y_{21}	...	x_n y_{n1}
Ziehung 2	x_1 y_{12}	x_2 y_{22}	…	x_n y_{n2}
…				

→ Realisationen y_{ij} der n Zufallsvariablen Y_1, ..., Y_n



$P(x_i, y_i)$ = Wahrscheinlichkeit, relative Häufigkeit für (x_i, y_i)

Wiederholte Beobachtungen in einer Stichprobe

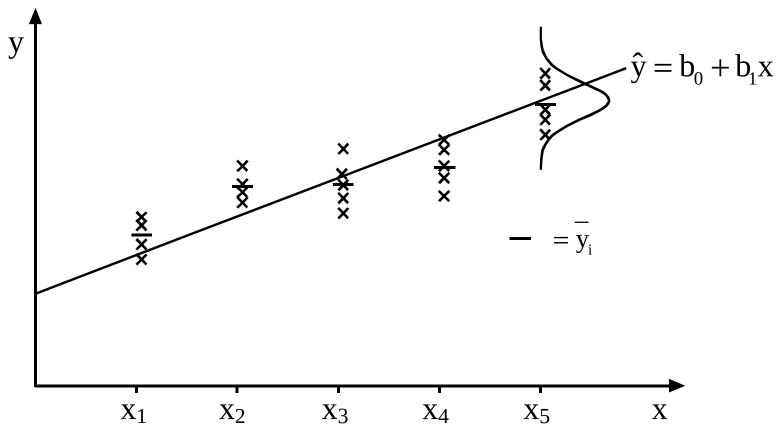
x	y				Mittelwert $\overline{y}_i$
x_1	y_{11}	y_{12}	$\cdots$	y_{1m}	$\overline{y}_1$
x_2	y_{21}	y_{22}	$\cdots$	y_{2m}	$\overline{y}_2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	y_{i1}	y_{i2}	$\cdots$	y_{im}	$\overline{y}_i$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_p	y_{p1}	y_{p2}	$\cdots$	y_{pm}	$\overline{y}_p$

p verschiedene x-Werte

m Beobachtungen zu jedem x - Wert (Vereinfachung zu m_i)

$\Rightarrow$ Anzahl der Beobachtungen: $n = m \cdot p$

Ziel: Regressionsgerade durch Mittelwerte $\overline{y}_i$ = Regressionsgerade durch y_i ($m_i = m$)



$n = m \cdot p \qquad 24 = 6 \cdot 4$

$$y_i = \mu_i + \varepsilon_i = \underbrace{\underbrace{f(x_i)}_{\text{systematische}} + \underbrace{\varepsilon_i}_{\text{zufällige}}}_{\text{Komponente}}$$

Beispiel: Kindergröße

Länge $\text{Kind}_i = \beta_0 + \beta_1 \cdot$ Länge Mutter $+ \varepsilon_i$

Aus einem Sample von Kindern gibt es mehrere Kinder unterschiedlicher Länge bei gleich langen Müttern

Gründe für den stochastischen Ansatz:

- nicht alle Einflüsse sind durch das Modell erfasst (Rest über ε); manche Einflüsse sind gar nicht erfassbar (Beispiel: Unsicherheit über zukünftige Entwicklung)
- Messfehler und nicht "richtige" Erfassung der Variablen ("Adäquationsproblem")
- der "wahre" funktionale Zusammenhang ist durch die Wahl der Funktionsform (Linearität, ...) nicht erfasst

- Schwankungen im menschlichen Verhalten (in der Natur, etc.)

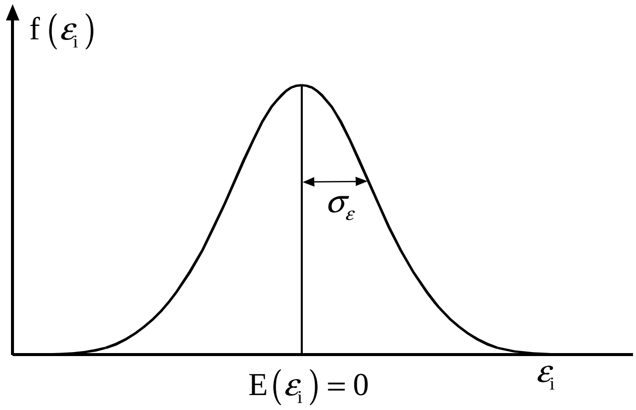
Modellformulierung CLR (Classical Linear Regression)

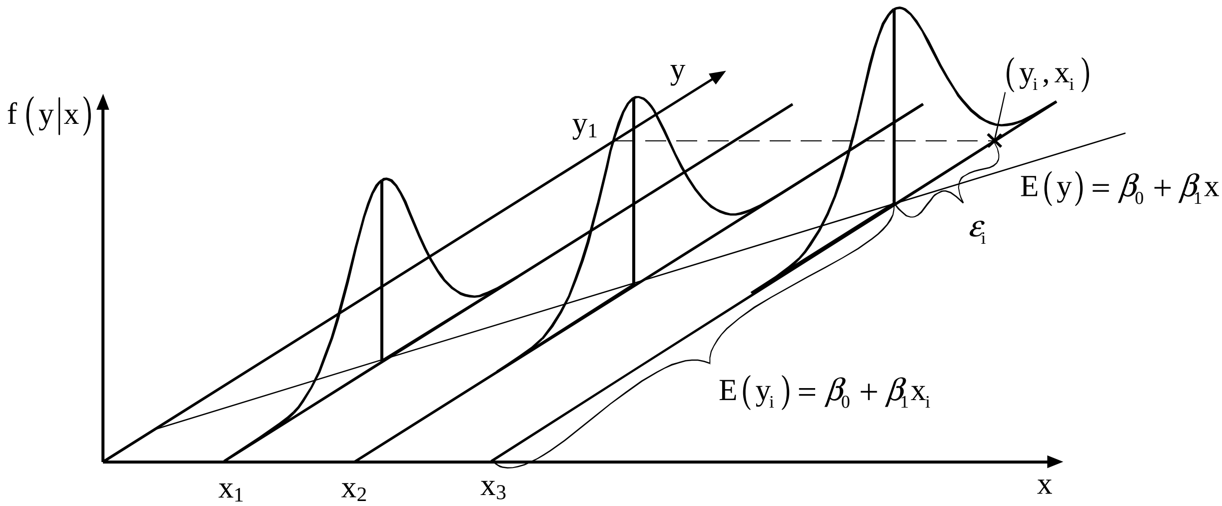
- n Beobachtungen als Stichprobe einer Grundgesamtheit
- bedingte Verteilung der abhängigen Variablen y bei gegebenen unabhängigen Variablen $x_1, \ldots, x_K$
- Mittel der bedingten Zufallsverteilung $\bar{y}\left(\mu, E(Y)\right)$ ist eine lineare Funktion der Variablen $x_1, \ldots, x_K$

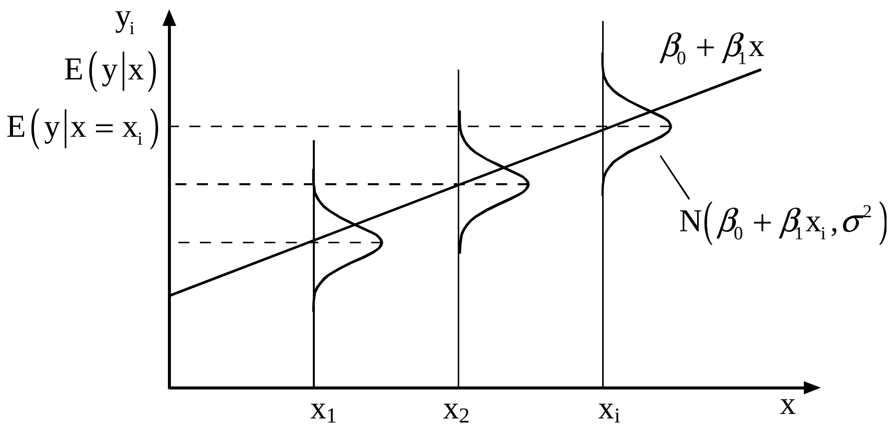
  d.h. $E(y_i|\mathbf{x_i}) = \mathbf{x_i'\beta}$ $\qquad \mathbf{x}_i' = (x_{i0}, x_{i1}, x_{i2}, \ldots, x_{iK})$ bzw.

  $E(\mathbf{y}|\mathbf{X}) = \mathbf{X\beta}$
- Varianz von y ist konstant; y ist (über ε) stochastisch
- Die durch wiederholte Ziehung gewonnenen y-Werte seien nicht korreliert; die unabhängigen Variablen x_K sind nicht stochastisch und bei jeder Stichprobe gleich ("fixed in repeated samples").
- gleiche (konstante) Varianz σ_ε^2

Abb. Regressionsfunktion der Grundgesamtheit und bedingte Verteilungen der (stetigen) Störvariablen ε_i, bzw. der abhängigen Variablen Y_i







2 Modellannahmen des klassischen linearen Regressionsmodells (CLR)

A1: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$	**Funktionale Form**

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad \mathbf{x}_i' = (x_{i0}, x_{i1}, x_{i2}, \ldots, x_{iK})$$

Spezifikationsannahme: Linearität in den Parametern

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{10} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{n0} & \cdots & x_{nK} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_K \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

(n = Anzahl der Ziehungen aus K verschiedenen Urnen)

y_i	$= \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i$
$\mathbf{y}_{n\times 1}$	Vektor der abhängigen, stochastischen Regressanden
$\mathbf{X}_{n\times(K+1)}$	Matrix der unabhängigen, nicht stochastischen (K+1) Regressoren
$\boldsymbol{\beta}_{(K+1)\times 1}$	Vektor der Parameter der Grundgesamtheit ("wahre" Werte)
$\boldsymbol{\varepsilon}_{n\times 1}$	Vektor der Zufallsterme bei n Beobachtungen

Linearisierung nichtlinearer Relationen:

$$f(y_i) = \beta_0 + \beta_1 g_1(x_1) + \beta_2 g_2(x_2) + \ldots + \beta_K g_K(x_K) + \varepsilon_i$$

Beispiele für g(x):

$\ln x$	$y = ax^{\beta} + \varepsilon_i \rightarrow \ln y = \beta_0 + \beta_1 \ln x$
$1/x$	$y = a + b(1/x)$
e^x	$y = a + b(e^x)$

Implikationen:

- Es ist aus der ökonomischen Theorie bekannt, welche Variable die abhängige ist.
- Alle notwendigen Regressoren zur systematischen Erklärung von y sind aufgeführt, d.h. keine Fehlspezifikation (Idealmodell).
- y_i ist eine lineare Funktion der x_{ik} (k = 0,1,...,K); y_i ist über ε_i eine stochastische Funktion.
- Ist die Verteilung von ε gegeben, dann ist auch die Verteilung von y gegeben.
- y und X enthalten keine Beobachtungsfehler; Fehler schlagen sich im Störterm ε nieder.
- β und ε sind nicht bekannt, da sie aus der Grundgesamtheit stammen.

A2:	$\mathbf{E(\boldsymbol{\varepsilon}) = 0}$	**Erwartungswert der Fehlerterme ist 0**

Für die erklärenden Größen ergeben sich Abweichungen nach oben und nach unten, die in der Erwartung jedoch 0 sein sollten.

$$\mathrm{E}(\varepsilon_i) = 0 \quad (i = 1, \dots, n) \qquad \begin{pmatrix} \mathrm{E}(\varepsilon_1) \\ \vdots \\ \mathrm{E}(\varepsilon_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ε hat keinen systematischen Einfluss auf y, der Fehler (latente Variable) ist zufällig.

Erwartungswert von **y** ist $\mathbf{X\beta}$:

$$\begin{aligned} \mathrm{E}(\mathbf{y}) &= \mathrm{E}(\mathbf{X\beta} + \boldsymbol{\varepsilon}) && \text{E ist linearer Operator} \\ &= \mathrm{E}(\mathbf{X\beta}) + \underbrace{\mathrm{E}(\boldsymbol{\varepsilon})}_{0} && \mathbf{X} \text{ ist nicht stochastisch (Annahme 4)} \\ &= \mathbf{X}\mathrm{E}(\boldsymbol{\beta}) = \mathbf{X\beta} \end{aligned}$$

A3:	$\mathbf{E(\boldsymbol{\varepsilon\varepsilon}') = \boldsymbol{\sigma}^2 I}$	**Konstante Varianz:**	$\mathbf{Var(\boldsymbol{\varepsilon}_i) = \boldsymbol{\sigma}^2}$	$\forall \mathbf{i}$
		Null Kovarianz:	$\mathbf{Cov(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j) = 0}$	$\mathbf{i \neq j}$

Homoskedastizität

bedeutet eine konstante Varianz der Fehlerterme unabhängig vom Wert der unabhängigen Variablen. Die Varianz der Fehlerterme sei gleich für alle Kombinationen der Werte der unabhängigen Variablen. Zu jeder Beobachtung, und damit für alle Kombinationen der Werte der unabhängigen Variablen, ist also die Varianz der Fehlerterme gleich groß. Das Gegenteil zur Homoskedastizität ist die Heteroskedastizität.

$\mathrm{E}(\varepsilon - \mathrm{E}(\varepsilon))(\varepsilon - \mathrm{E}(\varepsilon))' = \mathrm{E}(\varepsilon\varepsilon')$ $\quad (\mathrm{E}(\varepsilon) = \mathbf{0}$ (A2)) ist die Varianz-Kovarianzmatrix der Fehler.

$$\mathrm{E}\left[\begin{pmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{pmatrix}\begin{pmatrix}\varepsilon_1 & \dots & \varepsilon_n\end{pmatrix}\right] = \mathrm{E}\begin{pmatrix}\varepsilon_1^2 & \varepsilon_1\varepsilon_2 & \cdots & \varepsilon_1\varepsilon_n\\ \varepsilon_2\varepsilon_1 & \varepsilon_2^2 & \cdots & \varepsilon_2\varepsilon_n\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & \cdots & \varepsilon_n^2\end{pmatrix}$$

$$= \begin{pmatrix}\mathrm{Var}(\varepsilon_1) & \mathrm{Cov}(\varepsilon_1\varepsilon_2) & \cdots & \mathrm{Cov}(\varepsilon_1\varepsilon_n)\\ \mathrm{Cov}(\varepsilon_2\varepsilon_1) & \mathrm{Var}(\varepsilon_2) & \cdots & \mathrm{Cov}(\varepsilon_2\varepsilon_n)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{Cov}(\varepsilon_n\varepsilon_1) & \mathrm{Cov}(\varepsilon_n\varepsilon_2) & \cdots & \mathrm{Var}(\varepsilon_n)\end{pmatrix}$$

$$= \begin{pmatrix}\sigma^2 & 0 & \cdots & 0\\ 0 & \sigma^2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma^2\end{pmatrix} = \sigma^2\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix} = \sigma^2\mathbf{I}$$

Implikationen

- Varianzen aller ε_i (i = 1, ..., n) sind immer gleich
  Homoskedastizität $\mathrm{E}(\varepsilon_i^2) = \sigma^2 \quad \forall \mathrm{i}\ (\mathrm{i} = 1, \dots, \mathrm{n})$
- Störterme sind nicht miteinander korreliert
  $\mathrm{E}(\varepsilon_i\varepsilon_j) = 0 \quad \mathrm{i} \neq \mathrm{j} \quad (\mathrm{i}, \mathrm{j} = 1, \dots, \mathrm{n})$ d.h. in Zeitreihenanalysen keine Autokorrelation

(Autokorrelation 1. Ordnung: Dann wäre $\mathrm{E}(\varepsilon_i\varepsilon_{i-1}) \neq 0$)

A4: X ist nicht stochastisch

In wiederholten Stichproben ist **X** immer gleich, **y** aber möglicherweise verändert.

Implikationen

- **X** kommt von außerhalb des Modells; **X** ist fix
- **X** hat in wiederholten Stichproben den gleichen Wert
- **X** und **ε** sind unabhängig, nicht korreliert (systematischer und nicht systematischer (stochastischer) Teil sind voneinander unabhängig), d.h.

  $\mathrm{E}(\boldsymbol{\varepsilon}|\mathbf{x}_i) = \mathrm{E}\boldsymbol{\varepsilon} = 0$ oder

  $\mathrm{E}(\mathbf{X}'\boldsymbol{\varepsilon}) = \mathbf{X}'\mathrm{E}\boldsymbol{\varepsilon} = 0$

  $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}) = \mathrm{E}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' = \sigma^2\mathbf{I}$ und

  $\mathrm{E}(\mathbf{y}|\mathbf{x}) = \mathbf{X}\boldsymbol{\beta} + \mathrm{E}(\boldsymbol{\varepsilon}|\mathbf{x}) = \mathbf{X}\boldsymbol{\beta} \quad (\mathrm{E}\boldsymbol{\varepsilon} = \mathbf{0})$

bedingte Erwartung von y ist eine lineare Funktion der x_k.

Endogenitätsproblem

Endogenität beschreibt das Vorhandensein einer endogen erklärenden Variablen. Eine endogen erklärende Variable ist eine erklärende Variable in einem multiplen Regressionsmodell, die mit dem Fehlerterm korreliert ist, sei es durch eine ausgelassene Variable (‚omitted variable'), durch Messfehler bedingt oder wegen eines simultanen Ansatzes (Wooldridge 2009, S. 838).

Endogenität führt zu einem verzerrten und inkonsistenten Schätzer. Betrachten wir dazu eine einfache Regression mit mittelwertzentrierten y- und x-Werten (die jeweiligen Mittelwerte sind von den Ausgangsdaten abgezogen). Dann ist für die Regression

$$\mathbf{y} = \beta_1 \mathrm{x} + \varepsilon$$

der OLS-Schätzer $\hat{\beta}_1$ von β_1

$$\hat{\beta}_1 = \frac{\mathrm{cov}(x,y)}{\mathrm{var}(x)} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \qquad (\text{mittelwertbereinigte x und y}) \qquad .$$

Setzen wir y in den OLS-Schätzer ein, so ergibt sich

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i(\beta_1 x_i + \varepsilon_i)}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n}(\beta_1 x_i^2 + x_i\varepsilon_i)}{\sum_{i=1}^{n} x_i^2} = \frac{\beta_1 \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_i\varepsilon_i}{\sum_{i=1}^{n} x_i^2} = \beta_1 + \frac{\sum_{i=1}^{n} x_i\varepsilon_i}{\sum_{i=1}^{n} x_i^2}.$$

Wenn x und ε korreliert sind, dann ist der Erwartungswert des zweiten Terms ungleich Null und der Zähler wird nicht zu Null konvergieren, wenn die Stichprobengröße wächst. Damit ist die Schätzung von β_1 um den zweiten Term verzerrt (biased) und die Verzerrung wird auch durch größer werdende Stichproben nicht aufgehoben (siehe auch Kapitel IX, Spezifikationstests).

A5: $\quad \mathbf{Rang}(\mathbf{X}) = \mathbf{K}+\mathbf{1} \le \mathbf{n}$

Implikationen

- Sicherstellung der Regularität von $\mathbf{X'X}$ ($\mathbf{X'X}$ ist invertierbar)
- Lineare Unabhängigkeit der Vektoren $\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_K$. Zwischen den Regressoren ist kein perfekter linearer Zusammenhang erlaubt (dies wäre der Fall wenn z.B. eine erklärende Größe der Summe anderer erklärender Größen entspricht). Bei Vorliegen eines linearen Zusammenhangs zwischen den Regressoren spricht man von **Multikollinearität.**
  $\underline{y} = \beta_0 \underline{x}_0 + \beta_1 \underline{x}_1 + \ldots \beta_K \underline{x}_K + \varepsilon)$
- Wäre (K+1) = n $\rightarrow$ exakt bestimmtes lineares GS $(\hat{y} = y; \hat{\varepsilon} = 0)$

A6: ε ist (multivariat) normalverteilt

Da der Fehlerterm die Summe vieler unterschiedlicher unbeobachteter Faktoren ist der y beeinflusst, kann der zentrale Grenzwertsatz (alle Verteilungen nähern sich der Normalverteilung mit wachsender Anzahl von Beobachtungen an) herangezogen werden um e als approximativ normalverteilt anzunehmen.

Implikationen

- Damit ist die Verteilung von $\boldsymbol{\varepsilon}$ und damit von $\mathbf{y}$ vollständig spezifiziert.
- Elemente von $\boldsymbol{\varepsilon}$ sind stochastisch unabhängig.

Kombination der Annahmen A2, A3, und A6: $\boldsymbol{\varepsilon} \sim \mathrm{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$

Die Normalverteilung der Störterme ist für die BLUE-Eigenschaft (Best Linear Unbiased Estimator) der OLS-Schätzung nicht notwendig. Sie ist jedoch für Tests (t-Test, F-Test) erforderlich.

Zusammenfassung der Annahmen

A1	$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
A2	$\mathrm{E}\boldsymbol{\varepsilon} = \mathbf{0}$
A3	$\mathrm{E}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' = \sigma^2 \mathbf{I}$
A4	$\mathbf{X}$ ist nichtstochastisch
A5	$\mathrm{Rang}(\mathbf{X}) = \mathrm{K} + 1 \leq \mathrm{n}$
A6	$\boldsymbol{\varepsilon}$ normalverteilt

3 SCHÄTZUNG DER MODELLPARAMETER

3.1 MKQ/ OLS-Schätzung des Parametervektors β; Varianz von b_{OLS}

CLR-Annahme: Daten sind generiert durch die stochastische Relation

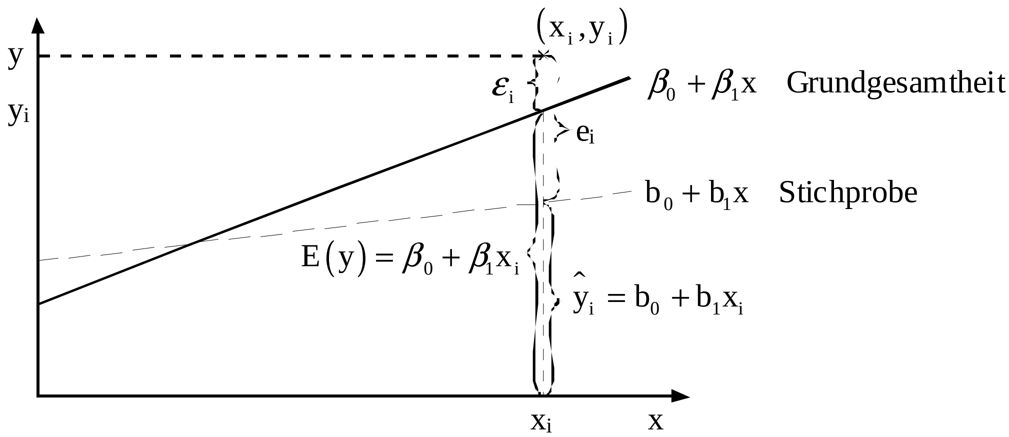
$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{iK} + \varepsilon_i = E\left(y_i \middle| x_{i0}, \ldots, x_{iK}\right) + \varepsilon_i$$

Werte der Grundgesamtheit ("wahre Werte"): $\boldsymbol{\beta}$, $\boldsymbol{\varepsilon}$

Schätzung des Parametervektors: $\mathbf{b}$

Grundgesamtheitsregression: $E(\mathbf{y}) = \mathbf{X\beta}$

Schätzung von $E(\mathbf{y})$: $E(\hat{\mathbf{y}}) = \hat{\mathbf{y}} = \mathbf{Xb}$



$$\Rightarrow \varepsilon_i = y_i - \overbrace{(\beta_0 + \beta_1 x_i)}^{\mu_i = X\beta} \qquad \text{bzw.} \quad \boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X\beta}$$

$$e_i = y_i - \underbrace{(b_0 + b_1 x_i)}_{\hat{y} = Xb} \qquad \text{bzw.} \quad \mathbf{e} = \mathbf{y} - \mathbf{Xb} = \mathbf{y} - \hat{\mathbf{y}}$$

aus der Definition

$$y = \begin{cases} \mathbf{X\beta} + \boldsymbol{\varepsilon} & \text{Grundgesamtheit} \\ \mathbf{Xb} + \mathbf{e} & \text{Stichprobe} \end{cases}$$

Nahe liegend: Stichprobe für die Schätzung der Grundgesamtheit verwenden!

$\hat{\mathbf{y}} = \mathbf{Xb}$ als Schätzer für E$\mathbf{y}$ $\Rightarrow$ Ausgleichslinie (-ebene) der Stichprobe als Schätzer für die Ausgleichslinie (-ebene) der Grundgesamtheit.

"**Fitting Criterion**: Least Squares": Minimierung der Fehlerquadratsumme

OLS:

$$\sum_i \tilde{\varepsilon}_i^2 = \sum_i \left(y_i - \widetilde{\boldsymbol{\beta}'\mathbf{x}_i}\right)^2$$

$\mathbf{x}_i$ ist (K+1) Zeilenvektor

$\tilde{\boldsymbol{\beta}}$ ist beliebiger (K+1) Spaltenvektor

$$\begin{aligned} \operatorname{Min} S\left(\tilde{\boldsymbol{\beta}}\right) &= \tilde{\boldsymbol{\varepsilon}}'\tilde{\boldsymbol{\varepsilon}} = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) = \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\tilde{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \tilde{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}} \\ &= \mathbf{y}'\mathbf{y} - 2\tilde{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \tilde{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}} \end{aligned}$$

siehe Abschnitt III

Notwendige Minimumbedingung

$$\frac{\partial S(\tilde{\boldsymbol{\beta}})}{\partial \tilde{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}} = \mathbf{0}$$

$\tilde{\boldsymbol{\beta}} \to \mathbf{b}$ löst die Bedingung und erfüllt die Normalgleichung

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y} \quad \Leftrightarrow \quad \mathbf{b} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$$

$$\frac{\partial^2 S(\mathbf{b})}{\partial \mathbf{b}\partial \mathbf{b}'} = 2(\mathbf{X}'\mathbf{X})' = 2(\mathbf{X}'\mathbf{X}) \text{ ist positiv definit (Abschnitt II)}$$

$$\Rightarrow \mathbf{b} \text{ minimiert } S\left(\tilde{\boldsymbol{\beta}}\right)$$

$$\boxed{\text{Schätzer für } \boldsymbol{\beta}\text{:} \quad \hat{\boldsymbol{\beta}} = \mathbf{b} = \mathbf{b}_{OLS} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}}$$

Varianz-Kovarianzmatrix von b

Da $\hat{\boldsymbol{\beta}} = \mathbf{b} = \mathbf{b}_{OLS} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$ eine Funktion von $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ist und damit über $\boldsymbol{\varepsilon}$ stochastisch ist, ist auch der Vektor **b** stochastisch. Mit der Annahme der Normalverteilung für $\boldsymbol{\varepsilon}$ ist über die lineare Verknüpfung mit **y** und **b** auch der Vektor **b** eine normalverteilte Zufallsvariable mit einem Erwartungswert – wie wir noch sehen werden $\boldsymbol{\beta}$ – und einer Varianz, bzw. eine (K+1)(K+1) Varianz-Kovarianzmatrix **Var**(**b**) aller Komponenten des (K+1) **b** Vektors.

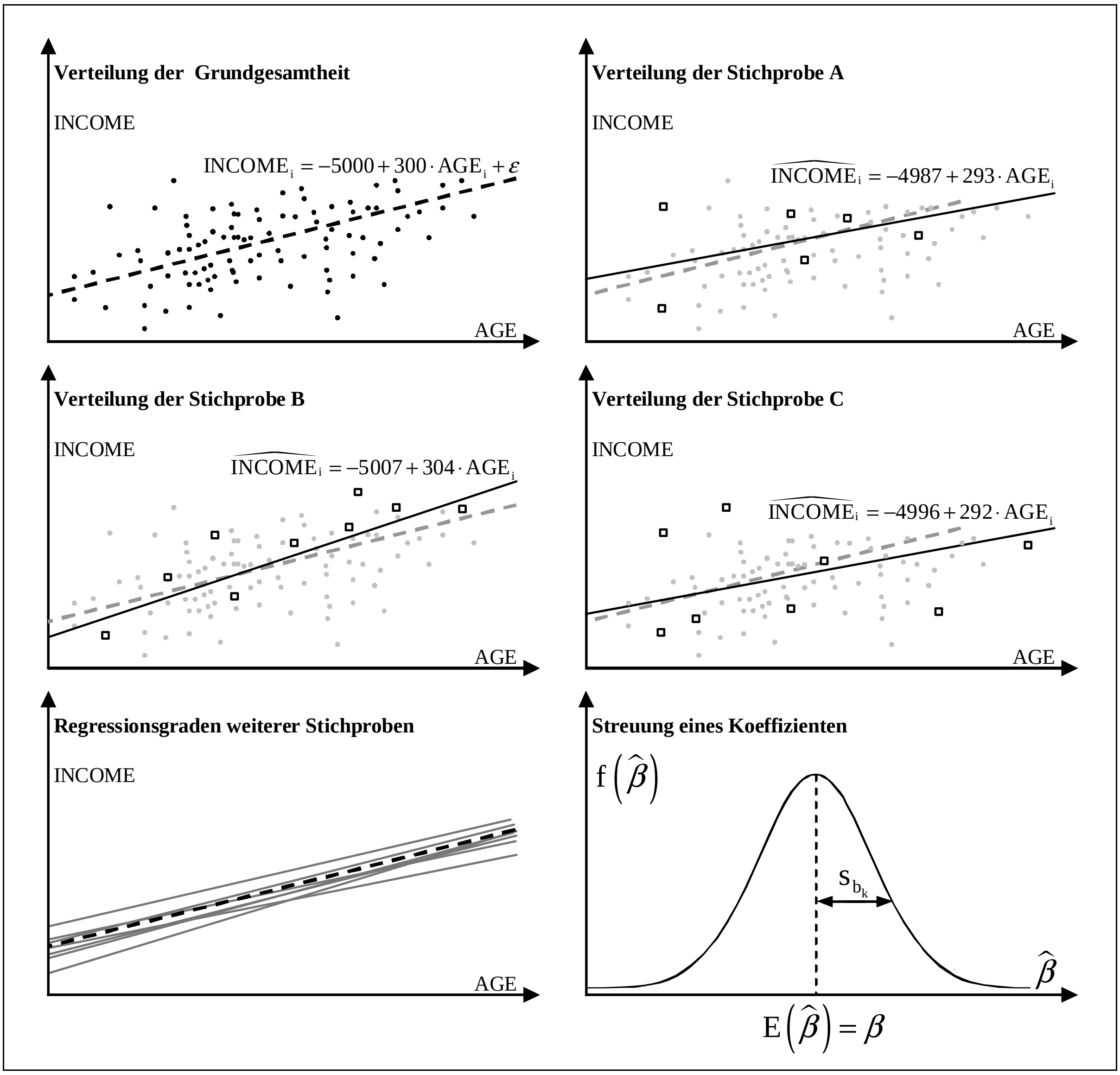
Es gibt also eine Varianz von **b**, weil der Schätzer eine Funktion der Stichprobe ist, die selbst wieder eine von vielen möglichen Stichproben ist.

Wollen wir Hypothesen mit den Parametern $\boldsymbol{\beta}$ testen und damit Konfidenzintervalle zu $\boldsymbol{\beta}$ bilden, dann brauchen wir eine entsprechende Streuung, die wir mit der Stichprobenschätzung der Varianz-Kovarianzmatrix **Var**(**b**) formulieren können und die wir nun ermitteln wollen.

Entwicklungsschritte: Dazu formulieren wir zunächst die Varianz-Kovarianzmatrix **Var**(**b**) abhängig von den Stichprobenwerten **X** und **y** mit dem Ergebnis $\mathbf{Var}(\mathbf{b}) = \sigma^2\mathbf{I}\left(\mathbf{X}'\mathbf{X}\right)^{-1}$. Da σ^2 aber aus der Grundgesamtheit stammt, also unbekannt ist, benötigen wir noch einen Schätzer für σ^2, der als erwartungstreuer Schätzer $\sigma^2 = \frac{1}{(n-K-1)}E\left[\mathbf{e}'\mathbf{e}\right]$ aus der Stichprobe abgeleitet werden kann.

Beispiel Zur Streuung des Parametervektors b

Die Streuung des Vektors **b** aus mehreren Stichprobenziehungen aus einer Grundgesamtheit sei exemplarisch für das Beispiel einer Einfachregression dargestellt mit:



Damit kann schließlich die gesuchte Varianz-Kovarianzmatrix von **b** mit den Daten einer einzigen Stichprobe geschätzt werden als

$$\mathbf{Var}(\mathbf{b}) = \frac{\mathbf{e}'\mathbf{e}}{n-K-1}(\mathbf{X}'\mathbf{X})^{-1}.$$

Entwicklung im Einzelnen:

Schätzvektor $\hat{\boldsymbol{\beta}} = \mathbf{b} = \mathbf{b}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ als Funktion der Grundgesamtheitsparameter $\boldsymbol{\beta}$ und $\boldsymbol{\varepsilon}$

$$\left|\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})\right.$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$
$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

$$\mathrm{Var}(\mathbf{b}) = \boldsymbol{\Sigma}_{\mathrm{bb}} = \mathrm{E}\left[\left(\mathbf{b} - \underbrace{\mathrm{E}(\mathbf{b})}_{\beta}\right)(\mathbf{b} - \mathrm{E}(\mathbf{b}))'\right]$$

$$= \mathrm{E}\left[(\mathbf{b}-\boldsymbol{\beta})(\mathbf{b}-\boldsymbol{\beta})'\right] \qquad \left|\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\right.$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$
$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

$$= \mathrm{E}\left\{\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right]\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\right]'\right\} \qquad \left|\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]' = \mathbf{X}\left[(\mathbf{X}'\mathbf{X})^{-1}\right]'\right.$$
$$\left| = \mathbf{X}\left[(\mathbf{X}'\mathbf{X})'\right]^{-1} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right.$$

$$\mathrm{Var}(\mathbf{b}) = \mathrm{E}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right]$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$
$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\sigma^2\mathbf{I})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \qquad |\mathrm{A3}$$
$$= (\mathbf{X}'\mathbf{X})^{-1}(\sigma^2\mathbf{I})$$

Varianz von b auf der Hauptdiagonalen von $\sum_{bb.}$

Kovarianzen b_k, b_{k*} (k, k* = 0, ..., K) übrige Elemente.

Zur Berechnung von $\sum_{bb}$ ist σ^2 notwendig, aber – da aus der Grundgesamtheit – unbekannt.

3.2 Schätzung von σ^2

Für die Varianzberechnung Var**(b)** benötigen wir noch eine Schätzung der unbekannten Grundgesamtheitsvarianz σ^2.

Denkbar wäre eine direkte Schätzung von σ^2 mit der Stichprobenvarianz

$$\hat{\sigma}^2 = \frac{1}{n}\sum(e_i - 0)^2 = \frac{1}{n}\sum e_i^2 = \frac{1}{n}\mathbf{e}'\mathbf{e}.$$

Allerdings sind die OLS Residuen nur ungenügende Schätzer für die Grundgesamtheitsresiduen:

$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} - \mathbf{X}\mathbf{b} = \boldsymbol{\varepsilon} - \mathbf{X}(\mathbf{b} - \boldsymbol{\beta})$, da $\boldsymbol{\beta}$ nicht direkt beobachtbar ist. Dennoch können die Stichprobenfehler wie folgt mit den Grundgesamtheitsfehlern verknüpft und dann weiter als Schätzer verwendet.

Über $\mathbf{e'e}$ und $E(\mathbf{e'e})$ werden wir nun zeigen, dass $\hat{\sigma}^2 = \frac{1}{n}\mathbf{e'e}$ nur ein verzerrter Schätzer ist. Im nächsten Abschnitt dient uns dieses Ergebnis als Vorlage für einen dann erwartungstreuen Schätzer für σ^2.

Transformationen 1: Verknüpfung von e mit ε

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{Xb} = \mathbf{y} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'y} = \left[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right]\mathbf{y} = \mathbf{My}$$

Eigenschaften von M:

- symmetrisch: $\mathbf{M} = \mathbf{M'}$ (Durch Ausführen der Operation leicht zu zeigen)
- idempotent: $\mathbf{M} = \mathbf{M}^2$

$$\begin{aligned}
\mathbf{e} &= \mathbf{My} = \mathbf{M}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\
&= \left[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right]\mathbf{X}\boldsymbol{\beta} + \mathbf{M}\boldsymbol{\varepsilon} \\
&= \mathbf{X}\boldsymbol{\beta} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'X}\boldsymbol{\beta} + \mathbf{M}\boldsymbol{\varepsilon} \\
&= \mathbf{M}\boldsymbol{\varepsilon} && |\text{also } \mathbf{MX} = 0
\end{aligned}$$

$$\begin{aligned}
\mathbf{e'e} &= (\mathbf{M}\boldsymbol{\varepsilon})'\mathbf{M}\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'\mathbf{M'M}\boldsymbol{\varepsilon} \\
&= \boldsymbol{\varepsilon}'\mathbf{M'M}\boldsymbol{\varepsilon} && |\text{Symmetrie} \\
&= \boldsymbol{\varepsilon}'\mathbf{M}^2\boldsymbol{\varepsilon} && |\text{Idempotenz} \\
&= \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}
\end{aligned}$$

Damit ist eine Verknüpfung von **e** mit **ε** gefunden.

Die $(n \times n)$ Matrix **M** ist fundamental und kann zudem in unterschiedlicher Weise verwendet werden:

$$\mathbf{e'e} = \mathbf{y'M'My} = \mathbf{y'My} = \mathbf{y'e} = \mathbf{e'y}$$

oder

$$\mathbf{e'e} = \mathbf{y'y} - \mathbf{b'X'Xb}$$

$$\mathbf{e'e} = \mathbf{y'y} - \mathbf{b'X'y} = \mathbf{y'y} - \mathbf{y'Xb}$$

Transformationen 2: $E[\mathbf{e'e}]$

Wie sieht $E[\mathbf{e'e}]$ nach der Verknüpfung der **e** mit **ε** aus?

$$E[\mathbf{e'e}] = E[\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}] = E[\boldsymbol{\varepsilon}'_{1n}\mathbf{M}_{nn}\boldsymbol{\varepsilon}_{n1}] \qquad |\text{Skalar!}$$

Spur (trace) einer Matrix A (tr(A)) ist die Summe der Diagonalelemente → tr (Skalar) = Skalar

$$E\left[\operatorname{tr}(\varepsilon'\mathbf{M}\varepsilon)\right] = E\left[\operatorname{tr}(\mathbf{M}\varepsilon\varepsilon')\right] \qquad \left|\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})\right.$$

$$= \operatorname{tr}\left(\mathbf{M}E[\varepsilon\varepsilon']\right) = \operatorname{tr}\left(\mathbf{M}\sigma^2\mathbf{I}\right)$$

| A3 $\quad$ **M** nicht stochastisch
| $\quad\quad$ tr lineare Funktion

$$= \sigma^2 \operatorname{tr}(\mathbf{M})$$

Nun ist

$$\operatorname{tr}(\mathbf{M}) = \operatorname{tr}\left[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] = \operatorname{tr}[\mathbf{I}_n] - \operatorname{tr}\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\right]$$
$$= \operatorname{tr}[\mathbf{I}_n] - \operatorname{tr}[\mathbf{I}_{K+1}] = n - (K+1)$$

Damit erhalten wir

$$E[\mathbf{e}'\mathbf{e}] = \sigma^2(n-K-1)$$

$$\boxed{\sigma^2 = \frac{E[\mathbf{e}'\mathbf{e}]}{(n-K-1)} = \frac{1}{(n-K-1)}E[\mathbf{e}'\mathbf{e}]}$$

und ist somit nicht identisch mit dem zunächst denkbaren Schätzer σ^2 mit

$$E(\hat{\sigma}^2) = E[(1/n)\mathbf{e}'\mathbf{e}] = \frac{1}{n}E[\mathbf{e}'\mathbf{e}]$$

Ein erwartungstreuer Schätzer von σ^2 der dieses Ergebnis direkt verwendet, wird im nächsten Abschnitt entwickelt.

4 SCHÄTZEIGENSCHAFTEN DES MKQ/ OLS-SCHÄTZERS

4.1 Erwartungstreuer Schätzer des Parametervektors $\beta = b_{OLS}$

$$\text{CLR}: \quad \mathbf{y} = \mathbf{X}\beta + \varepsilon$$
$$\mathbf{b}_{OLS} = \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$
$$\Rightarrow \quad \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \varepsilon)$$

$$E[\mathbf{b}] = E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon\right]$$
$$= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\varepsilon] \qquad |\text{A2: } E\varepsilon = 0;\ \text{A4: } \mathbf{X} \text{ nicht stochatisch}$$
$$= \beta$$

$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ ist erwartungstreuer Schätzer von β (unabhängig von der Verteilung von ε).

Alternative Beweisführung:

Allgemein: Sei $\hat{\beta}$ ein linearer Schätzer von β $\qquad$ $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

$$\begin{aligned} \hat{\beta} &= \mathbf{C}'\mathbf{y} = \mathbf{C}'(\mathbf{X}\beta + \varepsilon) && \Big| \mathbf{C}'_{(K+1)\times n} \text{ nicht stochastisch} \\ &= \mathbf{C}'\mathbf{X}\beta + \mathbf{C}'\varepsilon \\ E\hat{\beta} &= \mathbf{C}'\mathbf{X}\beta && \Big| A2: E\varepsilon = 0 \end{aligned}$$

$\hat{\beta}$ ist erwartungstreu, wenn $E\hat{\beta} = \beta$

d.h. wenn $\mathbf{C'X}=\mathbf{I}$ $\quad\Rightarrow\quad$ $E\hat{\beta} = \mathbf{C}'\mathbf{X}\beta = \mathbf{I}\beta = \beta$

Speziell: b ist linearer Schätzer mit

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \hat{\mathbf{C}}'\mathbf{y}$$

b ist erwartungstreuer Schätzer wenn $\hat{\mathbf{C}}'\mathbf{X} = \mathbf{I}$

$$\hat{\mathbf{C}}'\mathbf{X} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}$$

$\rightarrow$ **b** ist erwartungstreuer Schätzer von β

4.2 Erwartungstreuer Schätzer von σ^2

Ergebnis aus 3.2: Die Stichprobenvarianz ist nicht unmittelbar ein Schätzer der Grundgesamtheitsvarianz (σ^2)

$$E[\mathbf{e}'\mathbf{e}] = \sigma^2(n-K-1).$$

Sei

$$s^2 = \frac{\mathbf{e}'\mathbf{e}}{n-K-1} = \frac{SSE}{n-K-1}$$

ein Schätzer für σ^2. Sein Erwartungswert ist

$$E\left[s^2\right] = E\left[\frac{\mathbf{e}'\mathbf{e}}{n-K-1}\right] = \frac{\sigma^2(n-K-1)}{(n-K-1)} = \sigma^2.$$

Im CLR gilt also:

$$\boxed{s^2 = \frac{\mathbf{e}'\mathbf{e}}{n-K-1} \text{ ist erwartungstreuer Schätzer von } \sigma^2}$$

4.3 Erwartungstreue Schätzung der Varianz-Kovarianz-Matrix Var(b) = $\sum_{bb}$ des OLS-Schätzers b

Nach Abschnitt 3.1 wird ein erwartungstreuer Schätzer für

$$\mathrm{Var}[\mathbf{b}] = \Sigma_{bb} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

gesucht.

$(\mathbf{X}'\mathbf{X})^{-1}$ ist nicht stochastisch, s^2 ist erwartungstreuer Schätzer für σ^2

$$S_{bb} = s^2 (\mathbf{X'X})^{-1}$$
$$E[S_{bb}] = E\left[s^2 (\mathbf{X'X})^{-1}\right] = E\left[s^2\right](\mathbf{X'X})^{-1} = \sigma^2 (\mathbf{X'X})^{-1}$$

Damit ist ein erwartungstreuer Schätzer für $\mathrm{Var}[\mathbf{b}] = \Sigma_{bb}$ gefunden:

$$\boxed{S_{bb} = s^2 (\mathbf{X'X})^{-1} = \frac{\mathbf{e'e}}{n-K-1}(\mathbf{X'X})^{-1}}$$

4.4 Minimale Varianz (Effizienz) und BLUE-Eigenschaft von $\mathbf{b_{OLS}}$ – Gauss-Markov Theorem

In der Klasse aller linearen, erwartungstreuen Schätzer ist derjenige mit der geringsten Varianz auch ein effizienter Schätzer.

Allgemein: Abweichung der Schätzwerte

$$\begin{aligned}
\hat{\beta} - \beta &= \hat{\beta} - E\hat{\beta} \\
&= \mathbf{C'y} - E[\mathbf{C'y}] \\
&= C'(\mathbf{X\beta} + \varepsilon) - E\left[\mathbf{C'}(\mathbf{X\beta} + \varepsilon)\right] \\
&= \cancel{\mathbf{C'X\beta}} + \mathbf{C'}\varepsilon - \cancel{\mathbf{C'X\beta}} - \mathbf{C'}E[\varepsilon] && \big|E\varepsilon = \mathbf{0};\ \mathbf{C}, \mathbf{X} \text{ nicht stochastisch} \\
&= \mathbf{C'}\varepsilon && \text{lineare Funktion der Störvariablen } \varepsilon
\end{aligned}$$

Varianz-Kovarianz-Matrix

$$\begin{aligned}
\Sigma_{\hat{\beta}\hat{\beta}} &= E\left[\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)'\right] \\
&= E\left[\mathbf{C'}\varepsilon(\mathbf{C'}\varepsilon)'\right] = E[\mathbf{C'}\varepsilon\varepsilon'\mathbf{C}] \\
&= \mathbf{C'}E[\varepsilon\varepsilon']\mathbf{C} = \mathbf{C'}\sigma^2\mathbf{IC} && \big|A3 \\
&= \sigma^2\mathbf{C'C}
\end{aligned}$$

Zerlegung von $\mathbf{C'}$ des allgemeinen Schätzers $\hat{\beta}$:

$$\mathbf{C'} = \mathbf{\hat{C}'} + \mathbf{D'} = \underbrace{(\mathbf{X'X})^{-1}\mathbf{X'}}_{\text{aus OLS}} + \underbrace{\mathbf{D'}_{(k+1)\times n}}_{\text{nicht stochastisch}}$$

$$\begin{aligned}
\Rightarrow \Sigma_{\hat{\beta}\hat{\beta}} &= \sigma^2\mathbf{C'C} \\
&= \sigma^2\left(\mathbf{\hat{C}'} + \mathbf{D'}\right)\left(\mathbf{\hat{C}} + \mathbf{D}\right) = \sigma^2\left(\mathbf{\hat{C}'\hat{C}} + \mathbf{\hat{C}'D} + \mathbf{D'\hat{C}} + \mathbf{D'D}\right) \\
&= \sigma^2\left[\underbrace{(\mathbf{X'X})^{-1}\mathbf{X'X}}_{I}(\mathbf{X'X})^{-1} + (\mathbf{X'X})^{-1}\mathbf{X'D} + \mathbf{D'X}(\mathbf{X'X})^{-1} + \mathbf{D'D}\right] \\
&= \sigma^2\left[(\mathbf{X'X})^{-1} + (\mathbf{X'X})^{-1}\mathbf{X'D} + (\mathbf{X'X})^{-1}\mathbf{X'D} + \mathbf{D'D}\right]
\end{aligned}$$

$\hat{\beta}$ ist erwartungstreu, wenn

$$\mathbf{C'X} = \mathbf{I}$$
$$\Rightarrow \left(\hat{\mathbf{C}}' + \mathbf{D}'\right)\mathbf{X} = \hat{\mathbf{C}}'\mathbf{X} + \mathbf{D}'\mathbf{X}$$
$$= \left(\mathbf{X'X}\right)^{-1}\mathbf{X'X} + \mathbf{D'X}$$
$$= \mathbf{I} + \mathbf{D'X}$$

$\mathbf{C'X}$ ist aber nur dann gleich $\mathbf{I}$ wenn $\mathbf{D'X} = \left(\mathbf{X'D}\right)' = \mathbf{X'D} = \mathbf{0}$

$$= \sigma^2\left[\left(\mathbf{X'X}\right)^{-1} + \mathbf{D'D}\right]$$
$$= \sigma^2\left(\mathbf{X'X}\right)^{-1} + \sigma^2\mathbf{D'D}$$
$$= \Sigma_{\mathbf{bb}} + \sigma^2\mathbf{D'D}$$

Die Varianz-Kovarianzmatrix einer linearen erwartungstreuen Schätzung $\hat{\boldsymbol{\beta}}$ setzt sich aus der $\boldsymbol{\Sigma}_{\mathrm{bb}}$ -Matrix des OLS-Schätzers $\mathbf{b}$ und aus $\sigma^2\mathbf{D'D}$ zusammen.

Da nun $\mathbf{D'D}$ nicht negativ definit ist (Satz: $\mathbf{D}$ beliebig $\Rightarrow \mathbf{D'D}$ n.n.d.), und somit auch $\sigma^2\mathbf{D'D}\left(\sigma^2 \geq 0\right)$, so kann es allgemein keine kleinere Varianz-Kovarianz-Matrix geben als $\boldsymbol{\Sigma}_{\mathrm{bb}}$ des OLS-Schätzers $\mathbf{b} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y}$.

Somit gilt das

Gauss-Markov-Theorem

Im klassischen linearen Regressionsmodell ist der beste (minimale Varianz) lineare erwartungstreue Schätzer von β (BLUE = best linear unbiased Estimator) der OLS-Schätzer

$$\mathbf{b} = \mathbf{b}_{\mathrm{OLS}} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y}$$

mit der Varianz-Kovarianz-Matrix

$$\mathrm{Var}\left[\mathbf{b}\right] = \Sigma_{\mathbf{bb}} = \sigma^2\left(\mathbf{X'X}\right)^{-1}$$

und seiner erwartungstreuen Schätzung

$$\mathrm{S_{bb}} = \mathrm{s}^2\left(\mathbf{X'X}\right)^{-1} = \frac{\mathbf{e'e}}{\mathrm{n-K-1}}\left(\mathbf{X'X}\right)^{-1}$$

Goodness of fit, alternative Berechnung des Bestimmtheitsmaßes:

$$0 \leq \mathrm{R}^2 \leq 1$$

$$B = r^2 = R^2 = \frac{SQE}{SQT} = \frac{\sum(\hat{y}-\overline{y})^2}{\sum(y-\overline{y})^2} = \frac{SQT - SQR}{SQT} = 1 - \frac{SQR}{SQT} = 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{y} - n\overline{y}^2}$$

$$= 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \overline{y})^2} = 1 - \frac{SQR}{SQT} = 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{M}^\circ\mathbf{y}} \qquad \left| \ \mathbf{M}^\circ \mathrel{\hat{=}} \mathbf{M} \left\langle \begin{array}{l} \text{mit Variablen} \\ \text{als Differenz} \\ \text{zu Mittelwert} \end{array} \right. \right.$$

$$\overline{R}^2 = 1 - \frac{n-1}{n-K-1}\left(1-R^2\right) \qquad \left(\text{korrigiertes } R^2\right)$$

5 Maximum Likelihood (ML) und ML-Schätzung der Modellparameter

Der Maximum Likelihood Ansatz ist ein genereller Schätzansatz für nichtlineare Modell. Als ein Spezialfall führt er zur gleichen Schätzprozedur wie die Methode der kleinsten Quadrate (OLS).

Im Literaturverzeichnis finden sich mit allen Ökonometrielehrbüchern auch entsprechende Kapitel zum ML-Ansatz.

5.1 Maximum Likelihood (ML) Ansatz

ML-Prinzip
Unter allen möglichen Regressionsgeraden in der Grundgesamtheit werde diejenige ausgewählt, die - unterstellt man, sie wäre die 'wahre' Regressionsgerade - die Wahrscheinlichkeit maximiert, gerade die Werte der Stichprobe zu beobachten (R.A. Fisher 1890 - 1962)

Schätzansatz über die Residuen (CLR Annahme 6, Normalverteilung der Störterme)
Störvariable ε sei normalverteilt

$N\left(0, \sigma^2\right)$ 0 = Mittelwert, Erwartungswert σ^2 = Varianz

Gesucht

Schätzfunktion, für die unbekannten Parameter $\beta_0, \ldots, \beta_K$ und σ_ε^2 derart, dass die Wahrscheinlichkeit dafür, dass die beobachtete Stichprobe aus der angenommenen Grundgesamtheit stammt, maximiert wird.

Wahrscheinlichkeitsdichte von ε_i bei Normalverteilung

$$f\left(\varepsilon_i\right) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \exp\left[-\frac{\varepsilon_i^2}{2\sigma^2}\right] \qquad \text{Normalverteilung}$$

CLR-Modell: Gemeinsame Wahrscheinlichkeitsdichte der Stichprobe

Für das CLR-Modell sind mit der Annahme A3: E($\boldsymbol{\varepsilon\varepsilon}'$) = $\sigma^2\mathbf{I}$ die einzelnen ε_i sind nicht korreliert, also stochastisch unabhängig, und besitzen eine konstante Varianz. Mit der stochastischen Unabhängigkeit wird nach dem Multiplikationssatz der Wahrscheinlichkeitsrechnung die gemeinsame Dichte, also die Wahrscheinlichkeit für die beobachtete Stichprobe, aus dem Produkt der Einzelwahrscheinlichkeiten gebildet (z.B.: die Wahrscheinlichkeit bei zwei unabhängigen, sukzessiven Münzwürfen, zwei Mal 'Zahl' zu erhalten, entspricht dem Produkt der Einzelwahrscheinlichkeiten: $0,5 \cdot 0,5 = 0,25 \;\; -> L(\cdot) = f(\varepsilon_1) \cdot f(\varepsilon_2)$).

Gemeinsame Wahrscheinlichkeitsdichte aller normalverteilten Residuen der Stichprobe:

$$\begin{aligned} L(\cdot) &= f(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_v) \\ &= \prod_{i=1}^{n} f(e_i) = f(\varepsilon_1) \cdot f(\varepsilon_2) \cdot \ldots \cdot f(\varepsilon_n) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[-\tfrac{1}{2\sigma^2} \sum_{i=1}^{n} \varepsilon_i^2\right] \end{aligned}$$

Die gesuchten Parameter $\boldsymbol{\beta}$ lassen sich durch das Einsetzen des Schätzansatzes in die Fehlerterme ε_i mit $\varepsilon_i = y_i - \mathbf{x}_i'\boldsymbol{\beta}$ und $\sum_i \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ finden.

ML-Schätzung der CLR-Parameter: Vorgehensweise

1. Bilde die Wahrscheinlichkeit für die Stichprobe mit der

Likelihood-Funktion:

$$L = \prod_{i=1}^{n} f(\varepsilon_i) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[-\tfrac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mathbf{x}_i'\boldsymbol{\beta})^2\right]$$

2. Vereinfache durch logarithmieren (Produkt -> Summe)
3. Lösung der Maximierungsaufgabe $L(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2) = \max!$ über die ersten zu Null gesetzten Ableitungen nach $\boldsymbol{\beta}$ und σ^2 mit $\hat{\boldsymbol{\beta}}$ und $\hat{\sigma}^2$.

5.2 ML-Schätzung der CLR-Modellparameter

Rechentechnische Vereinfachung:

Da der natürliche Logarithmus eine streng monotone Transformation ist, ist max L(·) = max ln L(·) an der gleichen Stelle $(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2)$.

$$\ln L(\boldsymbol{\beta}, \sigma^2) = -\tfrac{n}{2} \ln(2\pi) - \tfrac{n}{2} \ln(\sigma^2) - \tfrac{1}{2\sigma^2} \underbrace{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'}_{\boldsymbol{\varepsilon}'} \underbrace{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}_{\boldsymbol{\varepsilon}}$$

Maximierung von ln L(·) durch Nullsetzen der ersten partiellen Ableitungen

(1) $$\frac{\partial \ln L}{\partial \boldsymbol{\beta}} = -\tfrac{1}{2\hat{\sigma}^2}\left(-2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}\right) = \tfrac{1}{\hat{\sigma}^2}\left(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}\right) = \mathbf{0}$$

$$\hat{=} \text{ OLS Ableitung mit } S = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = \ldots = \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

$$\frac{\partial S}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

(2) $$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right) = 0$$

Wird (1) mit $\hat{\sigma}^2$ multipliziert, so entspricht dies genau den OLS-Normalgleichungen $\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$

ML-Lösung: Der ML-Schätzer von $\boldsymbol{\beta}$ identisch ist mit dem OLS-Schätzer von $\boldsymbol{\beta}$

ML-Lösung:

$$\hat{\boldsymbol{\beta}}_{ML} = \mathbf{b}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$\sigma^2_{ML} = \frac{1}{n}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right) = \frac{\mathbf{e}'\mathbf{e}}{n}$$

Eigenschaften des ML-Schätzers

- ML-Schätzer wie OLS-Schätzer von β ist unverzerrt (unbiased, erwartungstreu)
- ML-Schätzer von σ^2_ε ist nicht erwartungstreu $\left(\tilde{s}^2 = \dfrac{\mathbf{e}'\mathbf{e}}{n-K-1} \text{ ist dagegen erwartungstreu}\right)$,

  der ML-Schätzer ist dagegen ‚nach Null' verzerrt:

  $$E(\hat{\sigma}_{ML}) = \frac{n-K}{n}\sigma^2 = \left(1-\frac{K}{n}\right)\sigma^2 < \sigma^2.$$

  Die Verzerrung verschwindet allerdings bei wachsendem Stichprobenumfang $\left(\frac{K}{n} \to 0\right)$.

ML-Schätzer hat alle wünschenswerten asymptotischen Eigenschaften:

- ML ist konsistent
- ML ist asymptotisch normal verteilt
- ML ist asymptotisch effizient

5.3 Beispiel: ML-Schätzung auf der Basis einer Binomialverteilung

Beispiel: Schätzung fehlerhafter Produktion

Geg.: Fertigungslos von Mikroprozessoren

daraus Stichprobe mit n = 5 Stück (mit Zurücklegen); untauglich 2 Stück

Stichprobenfehleranteil: $p = \frac{x}{n} = \frac{2}{5} = 0,4$

Falls Fehleranteil θ der Grundgesamtheit bekannt ist, kann die Wahrscheinlichkeit für p=0,4 mit Hilfe der Binominalverteilung bestimmt werden.

Stichproben aus den Grundgesamtheiten Nr. 1, 2, 3, ... mit unterschiedlichen Fehleranteilen

$$\theta_1 = 0$$
$$\theta_2 = 0,1$$
$$\theta_3 = 0,2$$
$$\dots$$

Für jede dieser Grundgesamtheiten ist die Wahrscheinlichkeit dafür zu berechnen, dass eine Stichprobe mit p = 0,4 auftritt.

ML-Schätzer für ein unbekanntes θ: Das θ bei dem das beobachtete Stichprobenergebnis die größte Wahrscheinlichkeit (Maximum Likelihood) besitzt.

$$W(X = x) = f_b(x/n;\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \quad \text{oder}$$

$$W(X = 2) = \binom{5}{2}\theta^2(1-\theta)^3$$

θ	$L(\theta) = W(X = 2\mid\theta)$
0,0	0,0
0,1	0,0729
0,2	0,2048
0,3	0,3087
0,4	**0,3456**
0,5	0,3125
0,6	0,2304
0,7	0,1323
0,8	0,0512
0,9	0,0081
1,0	0,0

Max. Wahrscheinlichkeit $W(X = 2\mid\theta)$ ist bei θ = 0,4 gegeben. Der ML-Schätzer $\hat{\theta}$ für den unbekannten Anteilswert in der Grundgesamtheit ist also gleich dem Stichprobenanteilswert:

$$\hat{\theta} = p = \frac{x}{n} = 0,4$$

Formal

$$L(\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$\ln L(\theta) = \ln \binom{n}{x} + x \ln \theta + (n-x) \ln(1-\theta)$$

Maximierung durch Nullsetzen der 1. Ableitung nach θ

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{x}{\tilde{\theta}} - \frac{n-x}{1-\tilde{\theta}} = 0$$

$$x - x\tilde{\theta} = n\tilde{\theta} - x\tilde{\theta}$$

$$\boxed{\tilde{\theta} = \frac{x}{n} = p} \qquad \left(\text{hier mit } p = \frac{x}{n} = 0,4\right)$$

ML-Schätzer für das arithmetische Mittel μ (Normalverteilung):

arithmetisches Mittel der Stichprobe $\hat{\mu} = \overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

6 PROGNOSE UND PROGNOSEINTERVALLE

Prognose von y° mit dem Regressorvektor x°:

$$y^{\circ} = \boldsymbol{\beta}' \mathbf{x}^{\mathbf{o}} + \varepsilon^{\circ}$$

Gauss-Markov-Theorem: $\hat{y}^{\circ} = \boldsymbol{\beta}' \mathbf{x}^{\mathbf{o}}$ ist BLUE - Schätzer von $E\left(y^{\circ}\right)$.

Prognosefehler

Greene 1993, Kap. 6.8

$$e^{\circ} = y^{\circ} - \hat{y}^{\circ} = (\boldsymbol{\beta} - \mathbf{b})' \mathbf{x}^{\mathbf{o}} + \varepsilon^{\circ}$$

Prognosevarianz

$$\operatorname{var}\left(e^{\circ}\right) = \sigma^2 + \operatorname{var}\left[(\boldsymbol{\beta} - \mathbf{b})' \mathbf{x}^{\circ}\right]$$

$$= \sigma^2 + \mathbf{x}^{\circ\prime} \operatorname{var}\left[\sigma^2 (\mathbf{X}'\mathbf{X})^{-1}\right] \mathbf{x}^{\circ}$$

bei einer Regression mit einer Konstanten

$$\operatorname{var}\left(e^{\circ}\right) = \sigma^2 + \frac{\sigma^2}{n} + \sigma^2 \left\{ \sum_{i=2}^{K+1} \sum_{j=2}^{K+1} \left(x_i^{\circ} - \overline{x}_i\right)\left(x_j^{\circ} - \overline{x}_j\right)\left(\underline{\mathbf{X}}' \mathbf{M}^{\circ} \underline{\mathbf{X}}\right)^{ij} \right\}$$

wobei $\underline{\mathbf{X}}$ die K Spalten von **X** ohne die Einsen enthält.

σ^2 durch s^2 geschätzt → Konfidenzintervall

$$\hat{y}^{\circ} \pm t_{\alpha/2} s\left(\hat{y}^{\circ}\right)$$

Prognoseintervall: K = 1-Variablen Fall

$$\hat{y}^{\circ} = b_0 + b_1 x^{\circ}$$

Prognoseintervall

$$\hat{y}^{\circ} \pm t_{\alpha}^{*} s_f \qquad t^{*} = t - \text{Wert}$$

$$s_f = s\left[1 + \frac{1}{n} + \frac{\left(x^{\circ} - \bar{x}\right)}{\sum_i \left(x_i - \bar{x}\right)}\right]^{\frac{1}{2}} \quad \text{Prognosefehler (Greene 1990, ET, S.69)}$$

Beispiele: Prognose der realen Investitionen

$$\text{Real investment} = f\left(\text{GNP}\left(1{,}6\text{T}\right), \text{discount rate}\left(10\%\right), \text{inflation}\left(4\%\right)\right)$$

$$\xrightarrow{\text{Example 6.5}} \left(1{,}16{,}1.5{,}10{,}4\right)\underbrace{\left(-0.509, -0.017, 0.67, -0{,}02, -0.001\right)}_{b} = 0{,}20764$$

Varianz: $$S^2\left(1 + \mathbf{X}^{\circ\prime}\right)\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}^{\circ}$$

Prognoseintervall: $$0.20764 \pm \underbrace{2.228\left(0.01022\right)}_{0{,}0228\ (=11\%\ \text{des Prognosewertes})} = \left\langle 0.1849, 0.2304 \right\rangle$$

Vereinfachung durch eine erweiterte Regression (Salkever 1976) vgl. Greene 1993 S: 196f.

(Greene 1993, S.195 Example 6.11)

7 ÜBUNGSAUFGABEN

- Verdeutlichen Sie sich die Bedeutung der folgenden Eigenschaften:
  - Linearer Schätzer
  - Unverzerrter Schätzer
  - Bester, linearer, unverzerrter Schätzer.
- Was besagt das Gauss-Markov-Theorem? Welche Annahmen liegen der Aussage zugrunde? Welche Konsequenzen hat das Theorem für Ihr weiteres Vorgehen bei der Schätzung?
- Welche zwei Bedingungen sind identisch mit der Eigenschaft der Konsistenz? Formulieren Sie diese in Gleichungsform.
- Bestimmen Sie den unverzerrten Schätzer der Varianz des Störterms.
- Bestimmen Sie den unverzerrten Schätzer der Varianz des Schätzers $b_{k.}$

VII HYPOTHESENTESTS IM CLR

Im Folgenden werden – aufbauend auf dem klassischen linearen Regressionsmodell – Tests herangezogen, die eine Hypothese zu überprüfen helfen. Testverfahren für Einzelparameter (t-Test) und später für verbundene Hypothesen - über Teilvektoren oder den gesamten Koeffizientenvektor (F-Test) – werden vorgestellt. Anschließend wird auf Verfahren eingegangen, mit denen Verletzungen der Annahmen aus Kapitel VI festgestellt werden können. Insbesondere wird auf Probleme der Autokorrelation, der Heteroskedastizität, fehlender Strukturkonstanz oder Normalverteilung der Residuen eingegangen.

Im Literaturverzeichnis werden mit allen Ökonometrielehrbüchern auch Hypothesentests behandelt. An neuerer englischsprachiger Literatur sind wieder Greene (2003), Wooldridge (2006, 2002) und Studenmund (2006), aus betrieblicher Sicht Anderson et al. (2010) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

Mit dem **FFB e-learning Modul Parametertests** wird eine entsprechende einfache Einführung gegeben (www.leuphana.de/ffb).

1 HYPOTHESEN UND HYPOTHESENTESTS (EINSEITIGE UND ZWEISEITIGE TESTS)

Wesentliches Anliegen der Ökonometrie:

Die Parameterschätzwerte eines (linearen) Modells mit sozioökonomischen Theorien zu konfrontieren.

Voraussetzung:

Theorien so formuliert, dass sie sich als Hypothese einem statistischen Test unterziehen lassen können.

Prüf- oder Nullhypothese $H_0 \leftrightarrow$ Alternativhypothese H_1

Hypothesen können sich auf einzelne, einige oder alle Parameter $\beta_0, \ldots, \beta_k, \sigma^2$ beziehen.

Skalarhypothesen:

- einfache Hypothesen: $H_0: \beta_k = \beta_k^0$ oder $\beta_1 \leq 0$ oder $\sigma^2 = 0$
- Linearkombination aus $\boldsymbol{\beta}$: $H_0: \boldsymbol{\alpha}'\boldsymbol{\beta} = \alpha^0$ oder $\boldsymbol{\alpha}'\boldsymbol{\beta} = 0$

  $\boldsymbol{\alpha}$ vorgegebener Vektor, α^0 Skalar postuliert

Vektorhypothesen:

- über den gesamten Koeffizientenvektor: H_0: $\boldsymbol{\beta} = \boldsymbol{\beta_0}$ bzw. $\boldsymbol{\beta} = \mathbf{0}$
- über Teile des Koeffizientenvektors: H_0: $\boldsymbol{\beta}^* = \boldsymbol{\beta}^{*0}$ bzw. $\boldsymbol{\beta}^* = \mathbf{0}$

- über verschiedene Linearkombinationen: H_0: $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}^0$ bzw. $\mathbf{R}\boldsymbol{\beta} = \mathbf{0}$

(**R**, $\mathbf{r}^0$ erfasst die Restriktionen → siehe auch restringiertes Modell RLS, Kapitel VIII)

Beispiel: Einfaches Modell zur Erklärung von Investitionen:

$$\text{Investitionen} = \beta_0 + \beta_1 \cdot \text{Zins} + \beta_2 \cdot \text{Inflation} + \varepsilon$$

Frage:

Beeinflussen Zins und/oder Inflation die Investition(shöhe)? Gilt damit $\beta_1 \neq 0$, $\beta_2 \neq 0$? D.h. ist die Investition (y) abhängig vom Zins und/ oder Inflation (x_2)?

→ Skalarhypothese: H_0: $\beta_1 = \beta_1^0 = 0$ H_1: $\beta_1 \neq 0$

Vektorhypothese: H_0: $\underline{\beta} = \underline{\beta}^0 = \underline{0}$ H_1: $\underline{\beta}_1 \neq \underline{0}$

Die geschätzten Parameter $\hat{\boldsymbol{\beta}} = \mathbf{b}_{OLS}$ oder $\mathbf{b}_{ML}$, $\hat{\sigma}^2$ sind zufällige, stochastische Größen, da über eine Stichprobe (x_i, y_i) ermittelt:

$$\left(y_i \middle| x_i\right) \text{zufällig} \Rightarrow \hat{\beta} \text{ zufällig} \Rightarrow \hat{\beta} \text{ hat Verteilung}$$

$$\text{z.B. } y_i \sim N(0,1) \Rightarrow \hat{\beta} \sim N\left(b, \sigma_b^2\right)$$

Prinzipielle Testvorgehensweise:

Konstruktion einer Prüfvariablen (Zufallsvariablen) aus den geschätzten Parametern, von der die Verteilung bekannt ist. Bei Vorgabe einer Irrtumswahrscheinlichkeit α wird dann festgestellt, ob die Realisation der Prüfvariablen in den Annahme- oder Ablehnungsbereich fällt.

1. Fall: $\sigma^2_{\hat{\beta}_k}$ bekannt

$$\Rightarrow \text{ Prüfgröße } z = \frac{\hat{\beta}_k - \beta_k^0}{\sigma_{\hat{\beta}_k}} \sim N(0,1) \qquad \text{standardnormalverteilt}$$

$\sigma_{\hat{\beta}_k}$ bekannt aus anderen Untersuchungen

2. Fall: $\sigma^2_{\hat{\beta}_k}$ unbekannt

$$\Rightarrow \text{ Prüfgröße } t_{beob} = \frac{\hat{\beta}_k - \beta_k^0}{s_{\hat{\beta}_k}} \sim t_{n-K-1} \qquad t-\text{verteilt mit } (n-K-1) \text{ Freiheitsgraden}$$

Zweiseitiger Test: $H_0: \beta_k = 0$

$H_1: \beta_k \neq 0$

Einseitiger Test: $H_0: \beta_k = 0$ oder $H_0: \beta_k = 0$

$H_1: \beta_k > 0$ $H_1: \beta_k < 0$

2 KONFIDENZINTERVALLE

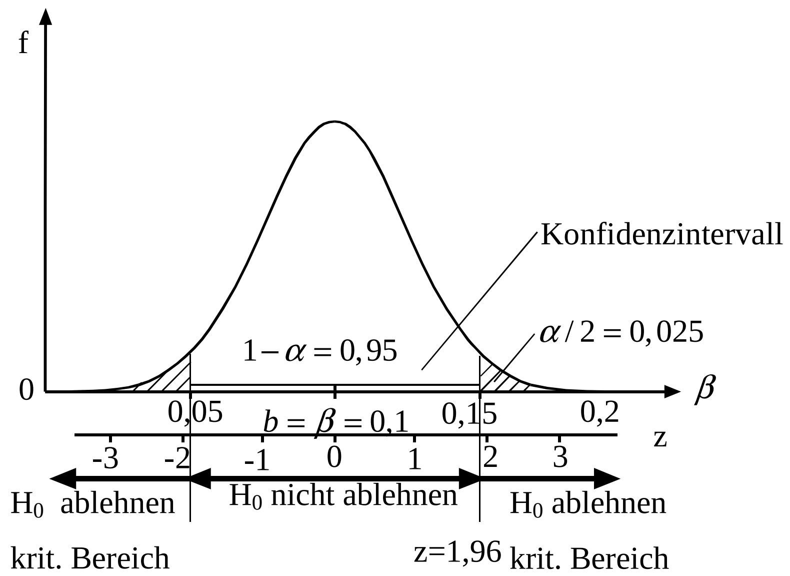
Die gegebene spezielle Stichprobe **y**, **X** ergibt eine Punktschätzung $\hat{y} = Xb$. Erwartungstreuer Schätzer liefert E(**b**) = **β**. Andere Stichproben der gleichen Grundgesamtheit liefern analoge $\mathbf{b}^*$ (zwar auch E($\mathbf{b}^*$)=β), aber mit Abweichungen vom wahren Wert β.

→ **Intervallschätzung:** Bereich um $\hat{\boldsymbol{\beta}}\left(= \mathbf{b} = \mathbf{b}_{OLS}\right)$, innerhalb dessen die 'wahren' Werte vermutet werden können.

Konfidenzintervall:

Intervall, das mit einer vorgegebenen Sicherheitswahrscheinlichkeit 1-α den wahren Wert der Grundgesamtheit überdeckt.

$$\Pr\text{ob}\left\{-\text{Prüfwert}_{\alpha/2} \leq \text{Prüfwert }(z,t) \leq \text{Prüfwert}_{\alpha/2}\right\} = 1-\alpha$$



Stichproben- und Testverteilung sowie kritischer Bereich

Konfidenzintervall für β

OLS-Modell: **ε** ist multivariat normalverteilt und $\boldsymbol{\varepsilon} \sim N\left(0, \sigma^2 \mathbf{I}\right)$; X nicht stochastisch

$$\mathbf{b} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y} \text{ und } \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ folgt:}$$

$$\mathbf{b} = \boldsymbol{\beta} + \left(\mathbf{X'X}\right)^{-1}\mathbf{X'}\boldsymbol{\varepsilon}$$

→ wegen X nicht stochastisch und ε multivariat ist auch b multivariat normal mit

$$b \sim N\left(\beta, \sigma^2 \left(X'X\right)^{-1}\right)$$

für die k-te Variable (inklusive der Konstanten)

$$b_k \sim N\left(\beta, \sigma^2 \left(X'X\right)^{-1}_{kk}\right) \qquad \left(X'X\right)^{-1}_{kk} = \text{k-tes Diagonalelement}$$

Prüfvariable/ Standardisierung:

$$z = \frac{b_k - \beta_k}{\sqrt{\sigma^2 \left(X'X\right)^{-1}_{kk}}} \sim N\left(0,1\right)$$

σ^2 unbekannt (Grundgesamtheit); da $\frac{\left(n-K-1\right)s^2}{\sigma^2} \sim \chi^2\left(n-K-1\right)$ (Johnston 1985, S.18) kann eine Prüfgröße t gebildet werden mit

$$t = \frac{b_k - \beta_k}{\sigma\sqrt{\left(X'X\right)^{-1}_{kk}}} \frac{\sigma\sqrt{n-K-1}}{s\sqrt{n-K-1}}$$

$\left| s = s_\varepsilon \right.$ = erwartungstreuer Schätzer von $\sigma = \sigma_\varepsilon = \mathbf{e}'\mathbf{e}/\left(n-K-1\right)$

$$t = \frac{b_k - \beta_k}{\sqrt{\sigma^2 \left(X'X\right)^{-1}_{kk}}} = \frac{b_k - \beta_k}{s_{b_k}} \sim t\left(n-K-1\right)$$

Konfidenzintervall:

$$\Pr ob\left\{-t^{n-K-1}_{\alpha/2} \le t \le t^{n-K-1}_{\alpha/2}\right\} = 1-\alpha \qquad |\text{umformen}$$

$$\Pr ob\left\{b_k - t^{n-K-1}_{\alpha/2} s_{b_k} \le \beta_k \le b_k + t^{n-K-1}_{\alpha/2} s_{b_k}\right\} = 1-\alpha$$

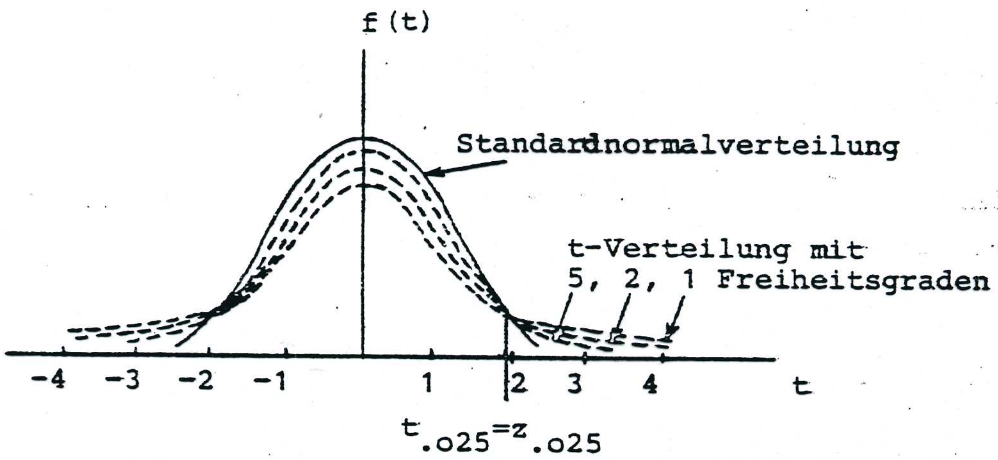
$$\boxed{\text{Konfidenzintervall: } \beta_k = b_k \pm t^{n-K-1}_{\alpha/2} s_{b_k}} \qquad \left| s^2_{b_k} = s^2 \left(\mathbf{X}'\mathbf{X}\right)^{-1}_{kk} \right. \qquad s^2 = s^2_\varepsilon = \frac{\mathbf{e}'\mathbf{e}}{n-K-1}$$

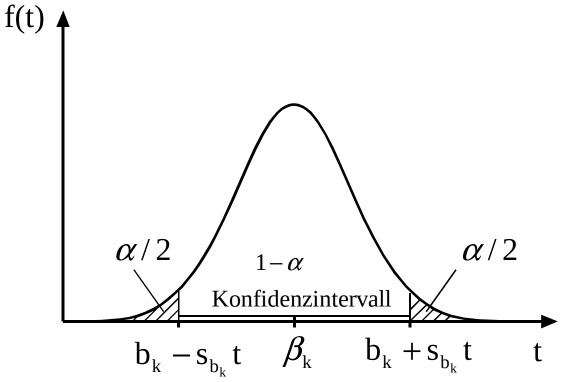
Interpretation:

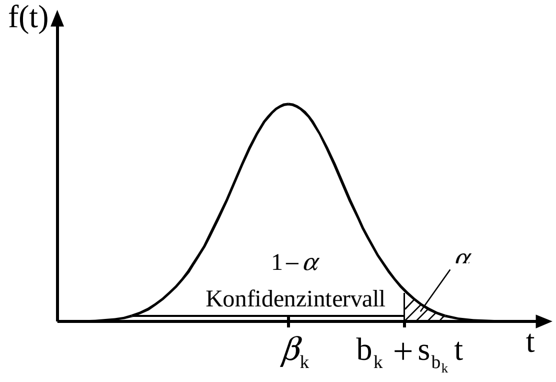
1-α= 0,95 α= Irrtumswahrscheinlichkeit

100 Stichproben → in 95 Fällen wird das Konfidenzintervall den wahren Wert des Regressionskoeffizienten einschließen.

Abb.: Vergleich der Standardnormalverteilung mit der t-Verteilung







3 TEST DER GESAMTERKLÄRUNGSGÜTE R^2 UND F-TEST

Hypothese: Alle Parameter außer der Konstanten (β_0) sind Null, der Gesamtansatz ist nicht signifikant..

$$H_0:\ \beta_1 = \beta_2 = \ldots = \beta_K = 0$$
$$H_1:\ \beta_1, \beta_{2,} \ldots, \beta_K \neq 0$$

b_{OLS} minimiert $e'e$, da

$$R^2 = r^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT} \text{ maximiert } b_{OLS}\ R^2\left(b = 0 \Rightarrow R^2 = 0\right).$$

F-Statistik heranziehen:

$$F_{beob} = \frac{R^2/K}{\left(1-R^2\right)/\left(n-K-1\right)} = \frac{SQE/K}{SQR/\left(n-K-1\right)} \sim F_{\alpha}^{(K,n-K-1)}$$

F- verteilt mit K-Freiheitsgraden (Zähler) und n-K-1- Freiheitsgraden (Nenner)

Wenn $F_{beob} > F_{krit} = F_{\alpha}^{(K,n-K-1)}$ dann wird H_0 verworfen

Zusammenfassung F-Test für R^2 (Bestimmtheitsmaß)

$H_0:\ \beta_1 = \beta_2 = \ldots = \beta_K = 0$
$H_1:\ \beta_1, \beta_{2,} \ldots, \beta_K \neq 0$

1. Signifikanzniveau α festlegen α 'klein'
2. Berechne $F_{beob} = \dfrac{R^2/K}{\left(1-R^2\right)/\left(n-K-1\right)}$
3. $F_{krit} = F_{\alpha}^{(K,n-K-1)}$ Tabelle
4. Für $F_{beob} > F_{krit}$ oder p-value < α H_0 wird verworfen
   Für $F_{beob} \leq F_{krit}$ oder p-value > α H_0 wird nicht verworfen

Beispiel F-Test für R^2:

Mac Prod Produktionsfunktion (EViews)

Dependent Variable: LNY
Method: Least Squares
Date: 09/20/06 Time: 16:03
Sample: 1 5
Included observations: 5

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.000000	4.474930	0.893869	0.4657
LNK	2.500000	0.866025	2.886751	0.1020
LNL	-1.500000	1.369306	-1.095445	0.3876

R-squared	0.946429	Mean dependent var	4.000000
Adjusted R-squared	0.892857	S.D. dependent var	2.645751
S.E. of regression	0.866025	Akaike info criterion	2.833904
Sum squared resid	1.500000	Schwarz criterion	2.599567
Log likelihood	-4.084760	F-statistic	17.66667
Durbin-Watson stat	1.666667	Prob(F-statistic)	0.053571

H_0: $\beta_1^0 = \beta_2^0 = 0$

H_1: $\beta_1^0, \beta_2^0 \neq 0$

1. **Signifikanzniveau**

   $\alpha = 0{,}01$

2. **Berechnung von F_{beob}**

$$F_{beob} = \frac{R^2/2}{(1-R^2)/2}$$

$$R^2 = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \bar{y})^2}$$

$$\sum_i e_i^2 = \mathbf{e'e} = 1{,}5 \qquad \text{(siehe Kapitel IV.1.3)}$$

$$\sum_i (y_i - \bar{y})^2 = \sum_i y_i^2 - \frac{1}{n}\left(\sum_i \bar{y}\right)^2 = 108 - \tfrac{1}{5}20^2 = 28$$

$$R^2 = 1 - \frac{1{,}5}{28} = 0{,}9464$$

$$F_{beob} = \frac{0{,}9464/2}{(1-0{,}9464)/2}$$

$$F_{beob} = 17{,}6567$$

3. **F_{krit}**

$$F_{krit} = F_{0,01}^{(2,2)} = 99{,}00 \qquad \text{Tabelle}$$

$$p-\text{value}(F_{beob}) = ?$$

$F_{beob} = 17{,}6567$ liegt zwischen $F_{0,05}^{(2,2)} = 19{,}00$ und $F_{0,10}^{(2,2)} = 9{,}00$ aus Tabelle

lineare Interpolation:

$$p-\text{value} = 10 - 5 * 8{,}6567/10 = 5{,}67\%$$

4. **H_0 ablehnen?**

   H_0 kann nicht abgelehnt werden, da $F_{beob} \leq F_{krit}$ bzw. p-value=0,054 > α =0,01 'schwache' Gesamterklärungsgüte (fiktives Beispiel, wenige Beobachtungen, n = 5)

4 TEST DER REGRESSIONSPARAMETER: T – TEST

$$H_0:\ \beta_k = 0 \qquad \Rightarrow \qquad t_{beob} = \frac{b_k - 0}{s_{b_k}}$$
$$H_1:\ \beta_k \neq 0$$

Ablehnung von H_0: $\beta_k = \beta_k^0 = 0$, wenn b_1 'stark' (in Relation zur Streuung) von 0 abweicht.

$$\beta_k = b_k \pm t_{\alpha/2}^{n-K-1} s_{b_k}$$

$$s_{b_k}^2 = s^2 \left(\mathbf{X'X}\right)_{kk}^{-1} \qquad s^2 = s_\varepsilon^2 = \frac{\mathbf{e'e}}{n-K-1}$$

$\left(\mathbf{X'X}\right)_{kk}^{-1}$ = k-tes Diagonalelement (inklusive Konstante)

$$\Pr ob\left\{-t_{krit} \leq \underbrace{t_{beob}}_{\text{Prüfgröße}} \leq t_{krit}\right\} = 1-\alpha$$

Idee: Einen Fehler in der '$\beta_k = 0$ ist richtig' - Entscheidung in Kauf zu nehmen

$\rightarrow t_{krit}$ = f(Risikobereitschaft)

Prob. H_0 abzulehnen, obwohl richtig $\alpha = 1\% \rightarrow t_{krit} \approx 3^*$

Prob. H_0 abzulehnen, obwohl richtig $\alpha = 5\% \rightarrow t_{krit} \approx 2$

α = Irrtumswahrscheinlichkeit = f(Freiheitsgraden)

Ist $t_{beob} > t_{krit} \rightarrow H_0$: $\beta_k = 0$ kann nicht aufrechterhalten werden, H_0 verwerfen. Das heißt x_k hat einen signifikanten Einfluss auf y.

Zusammenfassung t-Test $\beta_k = \beta_k^0 (= 0)$

H_0: $\beta_k = 0$

H_1: $\beta_k \neq 0$

1. Signifikanzniveau α festlegen α 'klein'
2. Berechne

$$t_{beob} = \frac{b_k - 0}{s_{b_k}}$$

$$s_{b_k}^2 = s^2 \left(\mathbf{X}'\mathbf{X}\right)_{kk}^{-1} \qquad s^2 = s_\varepsilon^2 = \frac{\mathbf{e}'\mathbf{e}}{n-K-1}$$

$$\left(\mathbf{X}'\mathbf{X}\right)_{kk}^{-1} = \text{k-tes Diagonalelement (inklusive Konstante)}$$

3. Aus der Tabelle der t-Verteilung entnehmen

$$t_{krit} = t_{\alpha/2}^{n-K-1} \qquad (\alpha/2 - \text{zweiseitig}; \alpha - \text{einseitig})$$

4. Für

$\left|t_{beob}\right| > t_{krit}$ oder p-value < α H_0 verwerfen

b_k ist statistisch gegen $\beta_k^0\ (=0)$ gesichert bei Signifikanzniveau α

$\left|t_{beob}\right| \le t_{krit}$ oder p-value > α H_0 nicht verwerfen

b_k ist nicht gegen $\beta_k^0\ (=0)$ gesichert

Beispiel t-Test Mac Prod Produktionsfunktion:

Mac Prod Produktionsfunktion

Dependent Variable: LNY
Method: Least Squares
Date: 09/20/06 Time: 16:03
Sample: 1 5
Included observations: 5

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.000000	4.474930	0.893869	0.4657
LNK	2.500000	0.866025	2.886751	0.1020
LNL	-1.500000	1.369306	-1.095445	0.3876

R-squared	0.946429	Mean dependent var	4.000000
Adjusted R-squared	0.892857	S.D. dependent var	2.645751
S.E. of regression	0.866025	Akaike info criterion	2.833904
Sum squared resid	1.500000	Schwarz criterion	2.599567
Log likelihood	-4.084760	F-statistic	17.66667
Durbin-Watson stat	1.666667	Prob(F-statistic)	0.053571

Signifikanztest von x_2 = Arbeitsinput

$$H_0: \beta_2^0 = 0$$

$$H_1: \beta_2^0 \neq 0$$

1. Signifikanzniveau

$\alpha = 0{,}01$ gewählt

2. Berechnung von t_{beob}

$$t_{beob} = \frac{b_k - \beta_k^0}{s_{b_k}} \qquad k = 2 \quad b_k = -1{,}5 \qquad \beta_2^0 = 0$$

$$s_{b_k}^2 = s_\varepsilon^2 \left(\mathbf{X}'\mathbf{X}\right)_{kk}^{-1} == \frac{\mathbf{e}'\mathbf{e}}{n-K-1}\left(\mathbf{X}'\mathbf{X}\right)_{kk}^{-1}$$

Erinnerung (Kapitel VI.3.2):

$$\begin{aligned}\mathbf{e} &= \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{Xb} \\ &= \mathbf{y} - \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} \\ &= \left[\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right]\mathbf{y} = \mathbf{My}\end{aligned}$$

Es gibt mehrer Wege, $e'e$ auszurechnen:

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{M}'\mathbf{My} = \mathbf{y}'\mathbf{My} = \mathbf{y}'\mathbf{e} = \mathbf{e}'\mathbf{y} \qquad (\mathbf{M} \text{ idempotent})$$

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{Xb}$$

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{Xb}$$

y_i	$\hat{\mathbf{y}}_i = \mathbf{x}_i'\mathbf{b}$	$e_i = y_i - \hat{y}_i$
3	4	-1
1	0,5	0,5
8	7,5	0,5
3	3	0
5	5	0

```
    <<<< XSXINV   >>>>  COLUMN
              1            2            3
ROW  1  26.7000      4.50000     -8.00000
ROW  2  4.50000      1.00000     -1.50000
ROW  3 -8.00000     -1.50000      2.50000
```

$$\mathbf{e}'\mathbf{e} = \sum_i e_i^2 = 1 + 0,25 + 0,25 = 1,5 \qquad \left(\mathbf{X}'\mathbf{X}\right)_{33}^{-1} = 2,5$$

$$s_{b_2}^2 = \frac{1,5}{5-2-1} \cdot 2,5 = 1,875 \qquad \left|t_{beob}\right| = \frac{-1,5}{\sqrt{1,875}} = 1,0954$$

3. t_{krit}

$$t_{krit} = t_{\alpha/2}^{n-K-1} = t_{0,01/2}^{2} = 9,925 \qquad \text{zweiseitig Tabellenwert } (H_1 \neq 0)$$

4. H_0 ablehnen?

$$\left|t_{beob}\right| \underset{\leq}{\overset{>}{}} t_{krit} \;??$$

1,0954 < 9,925 → H_0 kann nicht verworfen werden

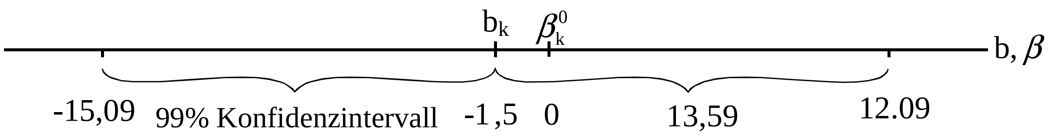
b_2 ist also nicht signifikant von Null verschieden

x_2 = Arbeit hat keinen signifikanten Produktionseinfluss in diesem fiktiven Beispiel.

Konfidenzintervall:

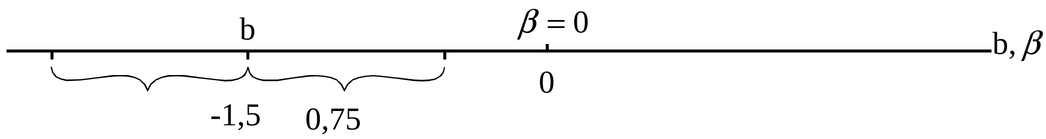
$\boxed{\alpha = 0{,}01 \rightarrow 99\%\ \text{Konfidenzintervall:}\ b_k \pm s_{b_k} t_{krit}}$

$-1{,}5 \pm 1{,}3693 \cdot 9{,}925{:}\ \ -1{,}5 \pm 13{,}49$



Da β_k^0 im Konfidenzintervall liegt, ist eben $b_k = -1{,}5$ nicht signifikant von Null verschieden.

Wäre z.B.



dann läge β_k^0 außerhalb des Vertrauensbereiches und $b_k = -1{,}5$ wäre signifikant von Null verschieden. $H_0{:}\ \beta_k^0 = 0$ würde zugunsten $H_1{:}\ \beta_k^0 \neq 0$ abgelehnt.

Beispiel F-Test und t-Test Arbeitsangebot:

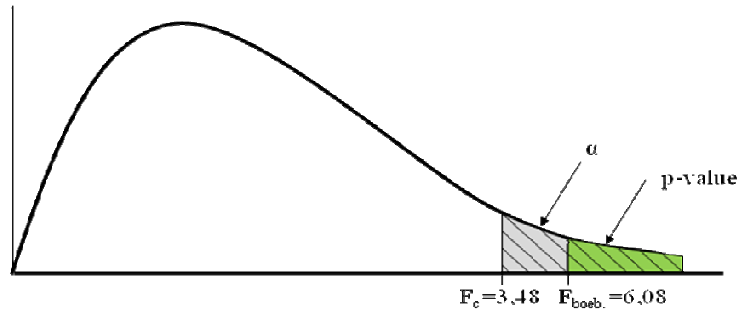
working hours = f(wage, sex, age, hhsize) (fiktive Daten)

CLR Schätzung (ET/LIMDEP)

```
===========================================================================
Ordinary    Least Squares
Dependent Variable        HOURS      Number of Observations            15
Mean of Dep. Variable      35.5333   Std. Dev. of Dep. Var.       9.738485
Std. Error of Regr.         6.2195   Sum of Squared Residuals      386.825
R - squared                 .70866   Adjusted R - squared           .59212
F( 4,  10)                  6.0810   Prob. Value for F              .00954
===========================================================================
Variable  Coefficient Std. Error  t-ratio Prob|t|>x Mean of X  Std.Dev.of X
---------------------------------------------------------------------------
Constant  -.844767      8.059       -.105   .91859
WAGE        .733063      .2350      3.120   .01088   25.53333       8.02555
SEX       -4.29397      3.630      -1.183   .26417     .60000        .50709
AGE         .710602      .2856      2.488   .03208   33.00000       7.40656
HHSIZE    -1.14748      1.593       -.720   .48780    2.80000       1.20712
```

F-Test (einseitiger Test)

Für den Test der Gesamterklärungsgüte mittels eines F-Tests gehört in diesem Arbeitsangebotsbeispiel zu dem Prüfwert dieser Stichprobe $F_{beob.} = F(4{,}10) = 6{,}081$ der p-value von 0,00954.



Testentscheidung:

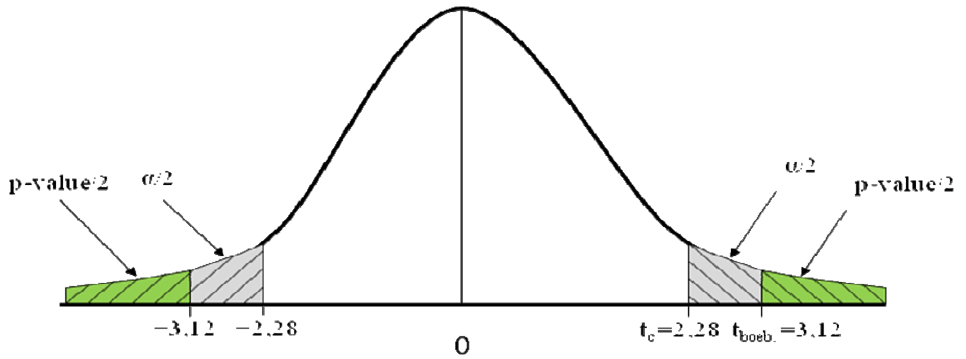
$F_{beob.}$ = F(4,10) = 6,081 > Fc = 3,478 (α=0,05)

p-value = 0,00954 < α=0,05

H_0 = „kein signifikanter Gesamterklärungsansatz" wird zugunsten einer signifikanten Gesamterklärungskraft abgelehnt

t-Test (zweiseitiger Test):

In diesem Arbeitsangebotsbeispiel ist für den geschätzten WAGE-Koeffizienten der Prüfwert t_c = 3,120 der entsprechende p-value 0,01088.



Testentscheidung:

$t_{beob.}$ = 3,120 > t_c = 2,2281 (ν= n-K-1=10, α=0,05)

p-value = 0,01088 < α=0,05 bzw.

p-value/2 = 0,00544 < α/2=0,025

H_0 = „kein signifikanter Einfluss des Lohnsatzes" wird zugunsten eines signifikanten Einflusses abgelehnt.

5 ALLGEMEINER F-TEST FÜR DEN KOEFFIZIENTENVEKTOR UND TEILVEKTOREN

Mit einem t-Test kann die Signifikanz eines einzelnen Regressors überprüft werden. Sollen mehrere Regressoren zusammen auf ihre Signifikanz getestet werden, empfiehlt sich ein F-Test auf multiple lineare Restriktionen (test on common exclusion restrictions, Wooldridge 2006, 150-157). Ein ähnlicher F-Test ist der F-Test auf die Gesamterklärungsgüte.

Ist bspw. eine Einkommensregression spezifiziert als

$$\log(\text{income}) = \beta_0 + \beta_1 \text{female} + \beta_2 \text{age} + \beta_3 \text{age}^2 + \beta_4 \text{university} + \varepsilon$$

mit den dummy.-Variablen female und university, dann interessiert bspw. ob das Alter signifikant Einkommen mit erklärt. Damit sind zwei Variablen gemeinsam gegen Null zu testen.

Dafür sind zwei Regressionen zu schätzen, einmal die gesamte Regression (nicht restringiertes Modell) und einmal die Regression ohne age und age^2 (restringiertes Modell).

Der F-Test ist definiert als

$$F_{\text{beob}} = \frac{(SSR_r - SSR_{ur})/c}{SSR_{ur}/(n-K-1)} \sim F(c, n-K-1)$$

wobei SSR die Summe der quadrierten Residuen (sum of squared residuals) des restringierten (r) bzw. unrestringierten (ur) Modells und c die Anzahl der ausgeschlossenen Variablen ist. In unserem log(income) Beispiel ist c=2 (für age und age2), K ist die Anzahl aller Regressoren (K=4) und n ist die Anzahl der Beobachtungen.

Kritische Werte können aus der entsprechenden F-Tabelle entnommen werden mit (c, n-K-1) Freiheitsgraden (df, degrees of freedom). Ist $F_{\text{beob}} > F_{\text{Krit}}$ dann ist die Nullhypothese, die betrachteten Variablen sind nicht signifikant von Null verschieden, zugunsten der Signifikanz abzulehnen.

Allgemeiner Fall multipler linearer Restriktionen

Ganz allgemein lässt sich der F-Test für den Koeffizientenvektor und Teilvektoren mit allgemein linearen Restriktionen formulieren als:

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$
$$H_1 : \mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$$

(Greene 1993, S.190ff., Hübler 1988, S.65ff.)

$\mathbf{R}$ = bekannte $c \times (K+1) - \text{Matrix}$

(mit vollem Rang, d.h. es gibt keine lineare Abhängigkeit zwischen den Hypothesen)

$\mathbf{r}$ = bekannter c-Vektor

Zur Illustration:

1. Ein Koeffizient ist Null: $\beta_k = 0$

   $\mathbf{R} = [0, \ldots, 0, 1, 0, \ldots, 0]$ und $\mathbf{r} = 0$

   für eine bestimmten anderen Wert β_k = a: r = a

2. Zwei Koeffizienten sind gleich: $\beta_k = \beta_l$
   $\mathbf{R} = [0,\dots\ ,\ 1,\dots,\ -1,\ \dots 0]$ und $\mathbf{r} = 0$
3. Eine Anzahl von Koeffizienten summiert sich auf 1: $\beta_2 + \beta_3 + \beta_4 = 1$
   $\mathbf{R} = [0,1,1,1,0\dots]$ und $\mathbf{r} = 1$
4. Eine Anzahl von Koeffizienten ist Null: $\beta_1 = \beta_2 = \beta_3 = \dots = 0$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \end{pmatrix} \text{und}\, \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

bzw. $[\mathbf{I} : 0]\beta = 0$ — $c = K + 1$ wenn alle $\beta_k = 0$ (siehe Abschnitt F-Test)

5. Mehrere Restriktionen gelten simultan: $\beta_2 + \beta_3 = 1,\ \beta_4 + \beta_6 = 0$ und $\beta_5 + \beta_6 = 0$

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_6 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$\rightarrow$ F-Statistik als allgemeiner Likelihood- Ratio- Test (Hübler 1985, S.65ff.)

$$F_{beob} = \frac{(\mathbf{Rb} - \mathbf{r})' \left[\mathbf{R} (\mathbf{X'X})^{-1} \mathbf{R}' \right]^{-1} (\mathbf{Rb} - \mathbf{r}) / c}{\mathbf{e'e} / (n - K - 1)} \sim F(c, n - K - 1)$$

$c, n - K - 1$: Freiheitsgrade (c= Anzahl der 'Joint' Hypothesen=Anzahl der Restriktionen)

Ist $F_{beob} > F_{\alpha}^{(1,n-k-1)}$: die verbundenen Hypothesen H_0: $R\beta = \mathbf{r}$ werden abgelehnt.

Beispiel: Restriktionen in Matrixschreibweise (Greene 1993, S.193)

$$\text{Investment} = \text{Constant} + \beta_1 \text{Time} + \beta_2 \text{Real GNP} + \beta_3 \text{Interest} + \beta_4 \text{Inflation} + \varepsilon$$

Hypothesen:

	$\beta_1 = 0$	kein Zeittrend
	$\beta_2 = 1$	Grenzrate der Investition = 1
	$\beta_3 + \beta_4 = 0$	Investoren betrachten nur realen Zinssatz
	$n - K - 1 = 10$	

$$\Rightarrow \mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \text{und}\ \mathbf{r} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$F_{beob} = 1409{,}59 > F_{0,05}^{(3,10)}$$

$\rightarrow$ Daten und Theorie stimmen nicht überein

6 LIKELIHOOD RATIO (LR)-, LAGRANGE MULTIPLIER (LM)- UND WALD TEST

Zentrale Tests im linearen Regressionsmodell sind t-, F- und χ^2-Tests.

Restriktionen auf den Parametern (z.B. Test ob $\beta_K \neq 0$) in linearen – und vor allem in nicht-linearen Modellen - werden auch durch die zusammengehörenden Likelihood Ratio (LR)-, Lagrange Multiplier (LM) und Wald Test getestet.

Die Nullhypothese H_0 ist dabei das restringierte Modell $H_0 : c(\beta) = 0$

Die Alternativhypothese H_1 ist das unrestringierte Modell $H_1 : c(\beta) \neq 0$, also ein CLR Schätzergebnis $y = \mathbf{x}'\boldsymbol{\beta} + \varepsilon \ \textit{mit} \ \hat{y} = \mathbf{x}'\mathbf{b}$.

z.B. Test auf Gesamterklärungsgüte mit $H_0 : \beta = 0 \ \textit{und} \ H_1 : \beta \neq 0$

Likelihood Ratio Test (LR)

verwendet H_0- und H_1- Schätzungen

Teststatistik:

$$LR = -2\ln\frac{L_r}{L_u} = -2\left(\ln L_r - \ln L_u\right) = 2\left(\ln L_u - \ln L_r\right)$$

mit L_r der Likelihood Wert im Optimum (Schätzung) des restringierten (z.B. y=f(const.)) und L_u der Likelihood Wert im Optimum des unrestringierten Modells (z.B. y=f(const., x_1, …, x_k)).

LR ist asymptotisch $\chi^2(\nu)$ verteilt mit $\nu = n - K - 1$ Freiheitsgraden (df)

Ist der LR-Wert größer als der $\chi^2_{krit} - Wert$, hat sich also die Schätzung durch unser eigentliches Modell $y=\mathbf{x}'\boldsymbol{\beta}+\varepsilon$ deutlich verbessert $\left(L_u > L_r\right)$, dann wird H_0 zugunsten H_1 verworfen.

Beispiel:

$\alpha = 0,05 \quad \chi^2_{krit}(\nu = 2) = 5,99$

Sei LR = 88,7,

dann ist $LR > \chi^2_{krit} \Rightarrow H_0 : c(\beta) = 0 \ \textit{wird verworfen} \ H_1 : c(\beta) \neq 0$

Lagrange Multiplier Test (LM)

verwendet H_0-Schätzungen (H_1 evtl. zu komplex)

$$LM = (\mathbf{Rb} - \mathbf{r})' \left[\mathbf{R}s^2 * \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{R}'\right]^{-1}(\mathbf{Rb} - \mathbf{r})$$

mit s^2* basierend auf den restringierten Koeffizienten ist $\mathbf{e}'^{*}\mathbf{e}^{*}/n$.

Der LM-Test ist rechentechnisch von Vorteil, wenn H_0 nicht abgedeckt wird. LM ist ebenfalls asymptotisch χ^2-verteilt mit df = Anzahl der Restriktionen = Zeilen von **R** mit der gleichen Testprozedur wie beim LR-Test.

Wald Test

verwendet H_1-Schätzungen

$$W = \left[(\mathbf{Rb} - \mathbf{r})' \left(\mathbf{R}\, Est.Var\, \mathbf{b} \right) \mathbf{R}' \right]^{-1} (\mathbf{Rb} - \mathbf{r})$$

W ist ebenfalls asymptotisch χ^2-verteilt (mit df = Anzahl der Restriktionen).

Es gilt W > LR > LM

Asymptotisch sind alle drei Tests äquivalent.

Für ‚finite sample' führen diese Tests im linearen (!) Regressionsmodell allerdings zu keiner zusätzlichen Testinformation gegenüber dem bekannten F-Test.

VIII BEISPIELE CLR-ANALYSE

Die folgenden Beispiele illustrieren mit unterschiedlichen Programmpaketen verschiedene CLR-Anwendungen sowohl auf der Makro- als auch auf der Mikroebene. Zunächst werden an den fiktiven Einkommensbeispielen aus dem bisherigen Text (income=f(age,sex)) die Zusammenhänge der Parametertests verdeutlicht und damit auch gezeigt, wie eine multiple Regressionsanalyse mit EXCEL durchgeführt wird. ‚Gasoline Sales in the US-Market' ist eine Makroanwendung mit Resultaten der Programmpakete ET/LIMDEP, EViews und SPSS.

Mikroanwendungen schätzen mit ‚Return on Human Capital (Stata)' Lohnsatz-Determinanten mit über 20.000 Zeittagebuchdaten der deutschen Zeitbudgeterhebung. Diese Zeitdaten sind auch die Basis für eine Schätzung mit vielen sozio-ökonomischen Variablen der täglichen Arbeitszeit mit ‚Daily Working Hour Arrangements (LIMDEP)'. Die bisher genanten CLR-Anwendungen illustrieren die Ergebnisse mit dem jeweiligen Programmpaketausdruck. Das letzte Beispiel zu ‚Happiness (Ferrer-i-Carbonell und Frijters 2004)' zeigt die Ergebnispräsentation in einer wissenschaftlichen Zeitschrift.

1 INCOME = F(AGE) (ET/LIMDEP)

Beispiel Income = f(age) (ET/LIMDEP)

Observation	INCOME	AGE	SEX
1	1200.0	22.000	1.0000
2	1700.0	24.000	1.0000
3	3500.0	28.000	1.0000
4	4200.0	27.000	.00000
5	1600.0	23.000	1.0000
6	5200.0	36.000	.00000

XSX=XDOT(ONE,AGE)

<<<< **XSX** >>>> **COLUMN**

		1	2
ROW	1	6.00000	160.000
ROW	2	160.000	4398.00

XSY=XDOT(X,INCOME)

<<<< **XSY** >>>> **COLUMN**

		1
ROW	1	17400.0
ROW	2	502600.

XSXINV=GINV(XSX)

<<<< **XSXINV** >>>> **COLUMN**

		1	2
ROW	1	5.58122	-.203046
ROW	2	-.203046	.761421E-02

Inverse of the 2*2 matrix $(\mathbf{X}'\mathbf{X})^{-1}$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{Inverse:} \quad A^{-1} = \frac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

with determinant $|A| = a_{11}a_{22} - a_{12}a_{21}$

$$A = (X'X) = \begin{pmatrix} 6 & 160 \\ 160 & 4398 \end{pmatrix}$$

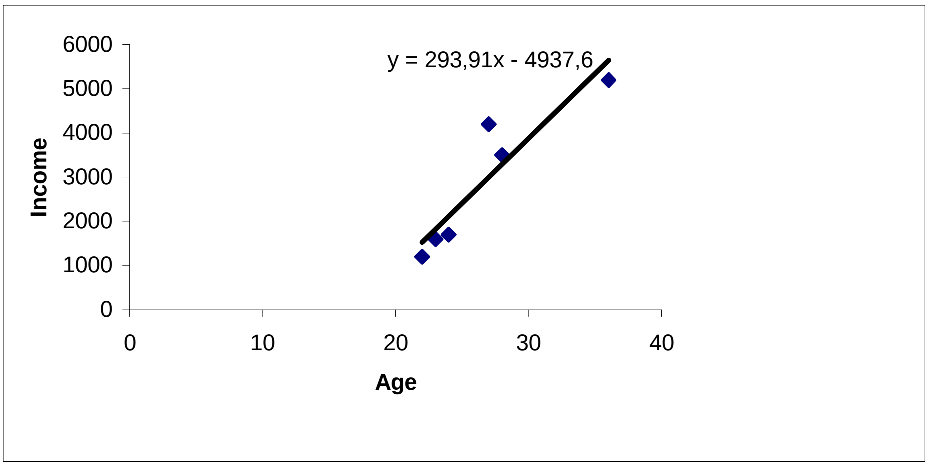
$$A^{-1} = \frac{1}{6 \cdot 4398 - 160 \cdot 160}\begin{pmatrix} 4398 & -160 \\ -160 & 6 \end{pmatrix}$$

$$= \begin{pmatrix} 5{,}5812 & -0{,}2030 \\ -0{,}2030 & 0{,}7614 \cdot 10^{-2} \end{pmatrix}$$

```
      BOLS=XSXINV|XSY

      <<<< BOLS      >>>>   COLUMN
                  1
 ROW   1  -4937.56
 ROW   2   293.909


      BET=XLSQ(X,INCOME)

      <<<< BET       >>>>   COLUMN
                  1
 ROW   1  -4937.56
 ROW   2   293.909

      regres; dep=income; ind=one,age$
```

```
Ordinary    Least Squares
Dependent Variable          INCOME       Number of Observations           6
Mean of Dep. Variable    2900.0000       Std. Dev. of Dep. Var.     1634.625339
Std. Error of Regr.       709.7758       Sum of Squared Residuals   .201513E+07
R - squared                 .84917       Adjusted R - squared            .81146
F( 1,   4)                 22.5194       Prob. Value for F               .00900
```

Variable	Coefficient	Std. Error	t-ratio	Prob|t|>x	Mean of X	Std.Dev.of X
Constant	-4937.56	1677.	-2.945	.04220		
AGE	293.909	61.93	4.745	.00900	26.66667	5.12510



Exkurs: Standard error and standard error of regression

A standard error is an *estimate of the standard deviation* of an estimator (mean, variance, regression coefficient, etc.). It is the standard deviation of the sampling distribution of a statistic. For example, in the case of estimating the variability of a sample mean, the standard error of the mean is the standard deviation of the sample means over all possible samples drawn from the population.

The **standard error of the data** is the estimate of the standard deviation of the data itselves

$$se(X) = \sqrt{\operatorname{var}(X)} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

with X as a random variable of a sample. The correction (using n-1 instead of n, Bessel's correction) provides with se^2 an unbiased estimator for the variance σ^2 of the underlying population. Additionally, if n = 1, then there is no indication of deviation from the mean, and standard deviation should therefore be undefined. The term sample standard deviation is used for the corrected estimator (using n-1). Note, the population standard deviation is $\sigma = \sqrt{\operatorname{var}(x)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ **.**

The **standard error of the mean** is the estimate of the standard deviation of the mean

$$se(\bar{X}) = se(X)/\sqrt{n}$$.

The **standard error of regression** is the estimate of the variability of the data around the regression line.

$$se(e) = \sqrt{\frac{1}{n-K-1}\sum_{i=1}^{n}e_i^2} = \sqrt{\frac{\mathbf{e'e}}{n-K-1}} .$$

Imagine a point on the regression line for a particular value of x. The standard error of regression is the estimate of the spread of the data above and below the fitted point.

```
ET:
Std. Error of Regr.    = sqrt(1/(n-K-1) Sum of Squared Residuals)
Std. Dev. of Dep. Var. = sqrt(1/(n-1) Sum(Y-mean of Y)²)
```

$\mathbf{e'e}$ = Sum of Squared Residuals = .201513E+07
n=6
K=1

Std. Error of Regr. = s(e) = sqrt(e'e/4) = sqrt(503782,5) = 709,7758

$$\mathrm{S_{bb}} = \mathrm{s}^2\left(\mathbf{X'X}\right)^{-1} = \frac{\mathbf{e'e}}{\mathrm{n-K-1}}\left(\mathbf{X'X}\right)^{-1}$$

$$\mathrm{S_{bb}} = \mathrm{s}^2\left(\mathbf{X'X}\right)^{-1} = 503782,5\left(\mathbf{X'X}\right)^{-1}$$

$$S_{bb} = 503782,5\begin{pmatrix} 5,58122 & -0,203046 \\ -0,203046 & 0,761421E-02 \end{pmatrix}_{(2\times 2)}$$

$$= \begin{pmatrix} 2811720,965 & -102291,02 \\ -102291,02 & 3835,91 \end{pmatrix}$$

$$s_{b_0} = \sqrt{2811720{,}965} = 1676{,}819$$
$$s_{b_1} = \sqrt{3835{,}91} = 61{,}934$$

which is exact the above result of Std. Error (standard error of b):

Variable	Coefficient	Std. Error	t-ratio	Prob|t|>x
Constant	-4937.56	*1677.*	-2.945	.04220
AGE	293.909	*61.93*	4.745	.00900

Std. Error of coefficient = $\sqrt{\mathrm{var}(b_k)} = \sqrt{Hauptdiagonalelement\, von\, S_{bb}}$

2 INCOME = F(AGE, SEX) (EXCEL)

Beispiel Income = f(age,sex) (EXCEL)

Regression module in EXCEL:
Activate in EXTRAS Add-Ins 'Analyse Funktionen', then click in EXTRAS 'Analyse Funktionen' the submodul 'Regression'

Income=b0+b1age+b2sex

i	income	age	sex
1	1200	22	1
2	1700	24	1
3	3500	28	1
4	4200	27	0
5	1600	23	1
6	5200	36	0

sex=0 male, 1 female

AUSGABE: ZUSAMMENFASSUNG

Regressions-Statistik	
Multipler Korrelat	0,958378606
Bestimmtheitsma	0,918489552
Adjustiertes Best	0,864149253
Standardfehler	602,4891678
Beobachtungen	6

ANOVA

	Freiheitsgrade (df	*adratsummen (S*	*Quadratsumme*	*Prüfgröße (F)*	*F krit*
Regression	2	12271020,41	6135510,204	16,90254873	0,023271261
Residue	3	1088979,592	362993,1973		
Gesamt	5	13360000			

	Koeffizienten	*Standardfehler*	*t-Statistik*	*P-Wert*	*Untere 95%*	*Obere 95%*	*Untere 95,0%*	*Obere 95,0%*
Schnittpunkt	-1728,571429	2462,110151	-0,702069088	0,533204051	-9564,112132	6106,969275	-9564,112132	6106,969275
X Variable 1	204,0816327	76,98324199	2,650987765	0,076928807	-40,91363122	449,0768965	-40,91363122	449,0768965
X Variable 2	-1220,408163	764,0368613	-1,597315817	0,208480104	-3651,916731	1211,100404	-3651,916731	1211,100404

ACHTUNG: EXCEL FKRIT ist der p-value für den F-Test

3 GASOLINE SALES IN THE US-MARKET (ET/LIMDEP, EVIEWS, SPSS)

Beispiel Gasoline Sales in the US-Market (ET)

Berechnung mit Econometric Toolkit (ET, Computer Package by W. Greene) oder LIMDEP.

The table below lists data that characterize sales of gasoline in the U.S. market from 1960 to 1986.

```
Gasoline sales in the US market from 1960 – 1982.
      G     = gasoline sales, in billions of gallons
      PG    = price index for gasoline
      Y     = per capita income
      PNC   = price index for new cars
      PUC   = price index for used cars
      PPT   = price index for public transportation
      PD    = price index for consumer durables
      PN    = price index for nondurables
      PS    = price index for services

DATA LISTING (Current sample)
Year   G       PG      Y    PNC    PUC    PPT    PD     PN     PS
1960 129.7   .925  6036  1.045   .836   .810   .444   .331   .302
1961 131.3   .914  6113  1.045   .869   .846   .448   .335   .307
1962 137.1   .919  6271  1.041   .948   .874   .457   .338   .314
1963 141.6   .918  6378  1.035   .960   .885   .463   .343   .320
1964 148.8   .914  6727  1.032  1.001   .901   .470   .347   .325
1965 155.9   .949  7027  1.009   .994   .919   .471   .353   .332
1966 164.9   .970  7280   .991   .970   .952   .475   .366   .342
1967 171.0  1.000  7513  1.000  1.000  1.000   .483   .375   .353
1968 183.4  1.014  7728  1.028  1.028  1.046   .501   .390   .368
1969 195.8  1.047  7891  1.044  1.031  1.127   .514   .409   .386
1970 207.4  1.056  8134  1.076  1.043  1.285   .527   .427   .407
1971 218.3  1.063  8322  1.120  1.102  1.377   .547   .442   .431
1972 226.8  1.076  8562  1.110  1.105  1.434   .555   .458   .451
1973 237.9  1.181  9042  1.111  1.176  1.448   .566   .497   .474
1974 225.8  1.599  8867  1.175  1.226  1.480   .604   .572   .513
1975 232.4  1.708  8944  1.276  1.464  1.586   .659   .615   .556
1976 241.7  1.779  9175  1.357  1.679  1.742   .695   .638   .598
1977 249.2  1.882  9381  1.429  1.828  1.824   .727   .671   .648
1978 261.3  1.963  9735  1.538  1.865  1.878   .769   .719   .698
1979 248.9  2.656  9829  1.660  2.010  2.003   .821   .800   .756
1980 226.8  3.691  9722  1.793  2.081  2.516   .892   .894   .839
1981 225.6  4.109  9769  1.902  2.569  3.120   .957   .969   .926
1982 228.8  3.894  9725  1.976  2.964  3.460  1.000  1.000  1.000
1983 239.6  3.764  9930  2.026  3.297  3.626  1.041  1.021  1.062
1984 244.7  3.707 10421  2.085  3.757  3.852  1.038  1.050  1.117
1985 245.8  3.738 10563  2.152  3.797  4.028  1.045  1.075  1.173
1986 269.4  2.921 10780  2.240  3.632  4.264  1.053  1.069  1.224

regres; dep=g; ind=ONE,YEAR,PG,Y,PNC,PUC,PPT$
===========================================================================
Ordinary    Least Squares
Dependent Variable       G           Number of Observations           27
Mean of Dep. Variable    207.0333    Std. Dev. of Dep. Var.    43.798927
Std. Error of Regr.        5.4157    Sum of Squared Residuals    586.606
R - squared                .98824    Adjusted R - squared         .98471
```

```
F( 6,  20)                280.0881     Prob. Value for F             .00000
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant  -10209.5      4644.       -2.198   .03987
YEAR       5.21657      2.401        2.172   .04203   1973.00000    7.93725
PG        -18.4347      3.695       -4.989   .00007      1.90211    1.16791
Y          .187961E-01  .9968E-02   1.886    .07396   8513.51852 1455.62903
PNC        27.4422      18.65        1.471   .15680      1.38133     .42797
PUC       -14.4631      8.007       -1.806   .08593      1.71230     .97474
PPT       -7.44493      7.654        -.973   .34235      1.86233    1.10831
===========================================================================

Estimated V-C Matrix for Parameters

               1-ONE        2-YEAR       3-PG         4-Y          5-PNC
   1-ONE       .215710E+08
   2-YEAR     -11153.3      5.76690
   3-PG        8666.22     -4.46827      13.6542
   4-Y         45.7814     -.236774E-01  .184537E-01  .993618E-04
   5-PNC       5180.41     -2.78086     -37.5415      .728855E-02  347.929
   6-PUC       746.702     -.375470      7.81011      .425357E-02 -65.9089
   7-PPT       10645.2     -5.48129      2.72227      .191899E-01 -28.0169

               6-PUC        7-PPT
   6-PUC       64.1078
   7-PPT      -40.6463      58.5889


P r e d i c t e d   V a l u e s
(* => not in estimating sample)
Observation     Observed Y   Predicted Y   Residual   95% Forecast Interval
       1          129.70       121.94        7.7582      108.659   135.225
       2          131.30       128.06        3.2368      115.567   140.560
       3          137.10       134.70        2.4035      122.448   146.945
       4          141.60       141.52         .0774      128.933   154.112
       5          148.80       152.58       -3.7783      140.582   164.574
       6          155.90       162.12       -6.2246      150.287   173.962
       7          164.90       171.32       -6.4169      159.481   183.153
       8          171.00       179.82       -8.8156      167.956   191.675
       9          183.40       188.84       -5.4362      177.078   200.594
      10          195.80       196.30        -.5008      184.545   208.057
      11          207.40       205.45        1.9527      193.483   217.411
      12          218.30       213.74        4.5624      201.691   225.784
      13          226.80       222.48        4.3166      210.171   234.796
      14          237.90       233.68        4.2172      220.194   247.172
      15          225.80       228.70       -2.8993      216.558   240.841
      16          232.40       231.89         .5059      219.617   244.171
      17          241.70       238.10        3.6045      225.979   250.212
      18          249.20       244.50        4.7044      232.121   256.870
      19          261.30       256.93        4.3731      243.896   269.958
      20          248.90       251.46       -2.5552      237.990   264.920
      21          226.80       234.38       -7.5843      220.907   247.861
      22          225.60       224.22        1.3849      210.780   237.651
      23          228.80       226.35        2.4454      213.436   239.273
      24          239.60       233.14        6.4590      220.406   245.876
      25          244.70       241.92        2.7793      228.078   255.763
      26          245.80       249.18       -3.3846      236.138   262.231
      27          269.40       276.59       -7.1854      261.302   291.869
```

(ET, Greene(1997, S.327f.)

Beispiel Gasoline Sales in the US-Market (EViews)

EViews: US Gasoline Market

Dependent Variable: G
Method: Least Squares
Date: 11/12/02 Time: 18:14
Sample: 1960 1986
Included observations: 27

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y	0.029679	0.004543	6.533306	0.0000
PG	-7.601784	8.500123	-0.894315	0.3808
PUC	-45.07951	15.34882	-2.937001	0.0076
PD	-181.5067	101.8276	-1.782490	0.0885
PS	281.3091	59.43764	4.732845	0.0001

R-squared	0.953216	Mean dependent var	207.0333
Adjusted R-squared	0.944710	S.D. dependent var	43.79893
S.E. of regression	10.29879	Akaike info criterion	7.667505
Sum squared resid	2333.430	Schwarz criterion	7.907475
Log likelihood	-98.51132	F-statistic	112.0624
Durbin-Watson stat	0.628661	Prob(F-statistic)	0.000000

Beispiel Gasoline Sales in the US-Market (SPSS)

Aufgenommene/Entfernte Variablen[b]

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	PS, Y, PG, PUC, PD[a]	,	Eingeben

a. Alle gewünschten Variablen wurden aufgenommen.

b. Abhängige Variable: G

Modellzusammenfassung[b]

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Durbin-Watson-Statistik
1	,997[a]	,993	,991	4,03858	,637

a. Einflußvariablen : (Konstante), PS, Y, PG, PUC, PD

b. Abhängige Variable: G

ANOVA[b]

Modell		Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
1	Regression	49534,487	5	9906,897	607,407	,000[a]
	Residuen	342,513	21	16,310		
	Gesamt	49877,000	26			

a. Einflußvariablen : (Konstante), PS, Y, PG, PUC, PD

b. Abhängige Variable: G

Koeffizienten[a]

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	-146,273	13,239		-11,048	,000
	Y	3,462E-02	,002	1,151	18,850	,000
	PG	-29,124	3,861	-,777	-7,544	,000
	PUC	-14,311	6,632	-,318	-2,158	,043
	PD	305,844	59,500	1,558	5,140	,000
	PS	-113,096	42,634	-,781	-2,653	,015

a. Abhängige Variable: G

4 RETURN ON HUMAN CAPITAL (STATA)

Beispiel Return on Human Capital (Stata)

Datenbasis: Zeitbudgeterhebung des Statistischen Bundesamtes 2001/02, Zeittagebücher

(Merz: . use "X:\timeuse\Wha\zbe2001-02_eIJTUR wha 2009\wha_mnl_cols\zbe0102_merz stata.dta")

Deskription

```
. sum wage lnwage experien woman if wage>0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        wage |     19470    9.987384    6.254399   .3501401   96.72619
      lnwage |     19470    2.129235    .6176006  -1.049422   4.571884
    experien |     24885    25.67077      12.493          0         49
       woman |     26802    .5335423     .498883          0          1
```

Log wage Schätzung, Männer und Frauen getrennt

```
. regress lnwage experien exper2 if (wage>0) & (woman==0)

      Source |       SS       df       MS              Number of obs =    9538
-------------+------------------------------           F(  2,  9535) = 1524.75
       Model |   889.30034     2   444.65017           Prob > F      =  0.0000
    Residual |  2780.61851  9535  .291622287           R-squared     =  0.2423
-------------+------------------------------           Adj R-squared =  0.2422
       Total |  3669.91885  9537  .384808519           Root MSE      =  .54002
```

```
------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    experien |   .0805268   .0016766    48.03   0.000     .0772403    .0838133
      exper2 |  -.0013267   .0000346   -38.35   0.000    -.0013945   -.0012589
       _cons |   1.269845   .0190835    66.54   0.000     1.232438    1.307253
------------------------------------------------------------------------------

. regress lnwage experien exper2 if (wage>0) & (woman==1)

      Source |       SS       df       MS              Number of obs =    9082
-------------+------------------------------           F(  2,  9079) =  677.25
       Model |  378.295661     2  189.14783            Prob > F      =  0.0000
    Residual |  2535.65226  9079  .279287615           R-squared     =  0.1298
-------------+------------------------------           Adj R-squared =  0.1296
       Total |  2913.94792  9081  .320884035           Root MSE      =  .52848

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    experien |   .0570792   .0017064    33.45   0.000     .0537342    .0604242
      exper2 |  -.0010003   .0000364   -27.46   0.000    -.0010717   -.0009289
       _cons |   1.341976   .0190472    70.46   0.000     1.304639    1.379313
```

Log wage Schätzung mit dummy woman

```
. regress lnwage woman experien exper2 if wage>0

      Source |       SS       df       MS              Number of obs =   18620
-------------+------------------------------           F(  3, 18616) = 1766.59
       Model |  1529.36323     3  509.787743           Prob > F      =  0.0000
    Residual |  5372.05879 18616  .288572131           R-squared     =  0.2216
-------------+------------------------------           Adj R-squared =  0.2215
       Total |  6901.42203 18619  .370665558           Root MSE      =  .53719

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       woman |  -.2635038   .0078969   -33.37   0.000    -.2789825   -.2480251
    experien |   .0686598   .0012007    57.18   0.000     .0663063    .0710132
      exper2 |  -.0011562   .0000252   -45.95   0.000    -.0012055   -.0011069
       _cons |   1.436838   .0140526   102.25   0.000     1.409293    1.464382
```

Log wage Schätzung mit dummy woman und Berücksichtigung von mehreren Tagebüchern pro Person (cluster)

```
. regress lnwage woman experien exper2 if wage>0, vce(cluster idpers)

Linear regression                                      Number of obs =   18620
                                                       F(  3,     6) =  890.91
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.2216
                                                       Root MSE      =  .53719

                                  (Std. Err. adjusted for 7 clusters in idpers)
------------------------------------------------------------------------------
             |               Robust
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       woman |  -.2635038   .0391315    -6.73   0.001     -.359255   -.1677525
    experien |   .0686598   .0076634     8.96   0.000     .0499081    .0874115
      exper2 |  -.0011562   .0001194    -9.68   0.000    -.0014483    -.000864
       _cons |   1.436838   .1089093    13.19   0.000     1.170346    1.703329
------------------------------------------------------------------------------
```

5 Daily Working Hour Arrangements (LIMDEP)

Beispiel Daily Working Hour Arrangements (LIMDEP)

Merz, Joachim, Böhm, Paul and Derik Burgert (2009), Timing and Fragmentation of Daily Working Hours Arrangements and Income Inequality – An Earnings Treatment Effects Approach with German Time Use Diary Data, in: electronic International Journal of Time Use Research, 6/2, 200-239

Datenbasis: Zeitbudgeterhebung des Statistischen Bundesamtes 2001/02, Zeittagebücher

CLR-Beispiel für Tägliche Arbeitszeitarrangements (Kernarbeitszeit am Tage/nicht fragmentiert) (Schätzungen in Merz et al. 2009 mit einem komplexeren Modell als CLR)

```
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = HWORK    Mean=   7.534279684    , S.D.=   2.629386890     |
| Model size: Observations =    4896, Parameters =  33, Deg.Fr.=   4863 |
| Residuals:  Sum of squares= 19358.74146    , Std.Dev.=        1.99520 |
| Fit:        R-squared=  .427974, Adjusted R-squared =          .42421 |
| Model test: F[ 32,   4863] =  113.70,    Prob value =          .00000 |
| Diagnostic: Log-L = -10312.4512, Restricted(b=0) Log-L =  -11679.8345 |
|             LogAmemiyaPrCrt.=    1.388, Akaike Info. Crt.=      4.226 |
| Autocorrel: Durbin-Watson Statistic =   1.69951,   Rho =       .15024 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     3.567611749       .44055231    8.098   .0000
 AGE           .1535422414   .20946408E-01    7.330   .0000     41.287786
 AGE2100     -.1785802420   .25470817E-01   -7.011   .0000     18.078854
 WOMAN       -.1881477754   .75960614E-01   -2.477   .0133      .45894608
 MARRIED     -.3137882175   .94533900E-01   -3.319   .0009      .66768791
 ELEMENTA  .4004220803E-01       .12289913     .326   .7446      .25776144
 INTERMED  .1388616499E-01       .11407012     .122   .9031      .40257353
 SUPPER   -.1653603737E-01       .12186310    -.136   .8921      .37520425
 SPECVC        .1248644419   .66693555E-01    1.872   .0612      .36254085
 UNIVERSI -.5648683006E-01       .10486537    -.539   .5901      .15706699
 FREIBERU    -.3620200726       .18035780   -2.007   .0447  .29207516E-01
 ENTREPRE     .4617177965       .16502991    2.798   .0051  .34313725E-01
 SZEITHH  -.1222957281E-01   .31428768E-03  -38.912   .0000     121.94444
 ZEITKIND -.1024598376E-01   .89770811E-03  -11.413   .0000     15.998775
 PHACTHH  -.5562468704E-02   .44656589E-02   -1.246   .2129     3.3926879
 PHACTCHI -.9916299012E-02   .56620624E-02   -1.751   .0799      .71531863
 PHACTCAR  .5756283161E-02   .20042920E-01     .287   .7740      .22824755
 WORKDAY      2.823572997       .10441312   27.042   .0000      .91462418
 WORKDIST  .2030814589E-02   .13455994E-02    1.509   .1312     25.428922
 SECONJOB -.7316559212E-01   .84749055E-01    -.863   .3880      .13582516
 WAGE     -.2817667992E-01   .12971176E-01   -2.172   .0298     10.389687
 WAGE2100  .3722019539E-01   .27623139E-01    1.347   .1778     1.4203866
 INDUSTRY     .3236532474   .97534220E-01    3.318   .0009      .27920752
 SERVICES     .1294761324   .85788969E-01    1.509   .1312      .57843137
 PARTNORW  .1655870401E-02   .20153711E-02     .822   .4113     18.745347
 PERSHH    .6766849025E-01   .30586785E-01    2.212   .0269     3.1921977
 YOUNGKID     .1717078335       .15310445    1.122   .2621      .13786765
 YKIDPWOR  .5089898187E-01       .16927709     .301   .7637  .90482026E-01
 OWNHOUSE  .2779368117E-01   .64883264E-01     .428   .6684      .63378268
 RESINCOM -.7282643633E-04   .16133670E-04   -4.514   .0000     2781.6883
 HHPASHH  -.2581328861E-02   .45603134E-02    -.566   .5714     1.9031454
 HHPASCHI  .3701933395E-02   .52305136E-02     .708   .4791     1.6376430
 OST          .6321685326   .77991805E-01    8.106   .0000      .21997549
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)
```

6 HAPPINESS (FERRER-I-CARBONELL UND FRIJTERS 2004)

Beispiel Happiness (Ferrer-i-Carbonell und Frijters 2004)

Die bisherigen Beispiele haben Regressions-Ergebnisprotokolle von verschiedenen Ökonometrieprogrammen gezeigt. Für eine Veröffentlichung werden schließlich nur zentrale Regressionsergebnisse (wie R2, Schätzkoeffizienten, t-Werte, n) ähnlich wie im folgenden Beispiel von Ferrer-i-Carbonell und Frijters 2004 publiziert.

Das Paper untersucht, inwiefern die Wahl der Analysemethode (OLS; Fixed-Effects-Schätzung) die Ergebnisse von Happiness/Satisfaction (hier General Satisfaction, GS) Regressionen beeinflusst.

Es zeigen sich nur geringe Unterschiede für die OLS mit und ohne Kontrollvariablen (1). Die Berücksichtigung individueller zeitinvarianter fixer Effekte hingegen verändert die Ergebnisse relativ stark hinsichtlich der Zufriedenheitsdifferenzen (2).

$$GS_{it} = x_{it}\beta + \varepsilon_{it}. \tag{1}$$

$$GS_{it} - GS_{it-1} = \Delta x_{it}\beta + \Delta\varepsilon_{it}, \tag{2}$$

Table 1

The Determinants of Cardinal General Satisfaction for West German Workers in the GSOEP

	Model (1)				Model (2)	
	OLS on GS		OLS on GS		Fixed-eff. OLS	
	Estimate	t-val	Estimate	t-val	Estimate	t-val
Age	−0.03	5.8	−0.05	10.0		
Age × age	0.0005	7.5	0.0007	11.3	−0.0006	6.5
ln(household income)	0.34	18.7	0.38	18.6	0.11	4.3
Number of children	−0.07	5.5	−0.05	5.2	0.01	0.9
Steady partner (1 = yes)	0.13	4.8	0.23	12.3	0.07	2.4
Subjective health	0.54	93.8	0.39	97.3	0.32	44.1
Controls	No		Yes		No	
Number of individuals	7,806		7,806		6,664	
R^2	0.25		0.26		0.09	
Number of cases	30,569		30,569		21,104	

Time-dummies were present in all estimates but are not shown. The number of individuals is lower for the fixed-effects because they require at least 2 observations per individual.
The controls for the OLS on GS contains the following variables: education, working hours, gender, and the number of adults in the household.

Quelle:

Ferrer-i-Carbonell A. and P. Frijters (2004), How important is Methodology for the estimates of determinants of happiness? in: The Economic Journal, Vol. 114, No. 497, 641–659.

IX TESTS DER ANNAHMEN DES KLASSISCHEN LINEAREN REGRESSIONSMODELLS (CLR)

Wenn ein Regressionsmodell nicht "passt", also die Realität nicht "gut" erklären kann und/oder die CLR-Modellannahmen nicht erfüllt sind, dann verbleiben strukturelle Spuren in der Verteilung des stochastischen Teils, den Residuen.

Die im Folgenden behandelten Tests der CLR-Modellannahmen werden deshalb direkt oder indirekt etwas mit dem stochastischen Teil - dem Fehlerterm - und seinem Verhältnis zum deterministischen Teil des Modells zu tun haben, seien es die Themen Spezifikation des Modells, Heteroskedastizität, Autokorrelation, Strukturkonstanz, verzögerte Variablen und dynamische Modelle oder Fehler in den Variablen.

Die Verletzungen der Modellannahmen können mit entsprechenden Tests diagnostiziert werden. Anschließend kann eine neue Modellierung vorgenommen werden und/oder Modifikationen mit dem Generalisierten linearen Regressionsmodell (GLR) berücksichtigt werden (vgl. dazu Abschnitt IX).

Im Folgenden werden deshalb Testverfahren zur Überprüfung der Annahmen des klassischen linearen Modells (Spezifikation und Modellselektion, Autokorrelation, Homoskedastizität, Multikollinearität (Rang(X)=k+1), Normalverteilung der Residuen) vorgestellt.

Im Literaturverzeichnis wird mit allen Ökonometrielehrbüchern auch das Vorgehen bei Verletzung der Annahmen des klassischen linearen Regressionsmodells behandelt. An neuerer englischsprachiger Literatur sind Greene (2003), Wooldridge (2006) und Studenmund (2006) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

Zum Test der Annahmen im klassischen linearen Regressionsmodell (CLR)

Generell kann eine Hypothese mit einer Teststatistik T getestet werden, die eine Funktion der Stichprobe ist. Mit ihr und einer bestimmten Irrtumswahrscheinlichkeit α kann entschieden werden, ob die Hypothese gültig ist bzw. ob die Daten der Stichprobe durch die hypothetische Grundgesamtheit generiert sind.

Dafür wird ein kritischer Wert c herangezogen, der angibt, ab welchem Wert die Nullhypothese H_0 zugunsten der Alternativhypothese H_A abgelehnt wird (Neyman-Pearson Ansatz). Dieser kritische Wert hängt vom Signifikanzniveau α ab und wird durch die Verteilung der Teststatistik unter der Annahme bestimmt, dass H_0 die wahre Hypothese ist; bei einem rechtsseitigen Test gilt beispielsweise: $P(t>c \mid H_0)=\alpha$. Liegt der Testwert t oberhalb des kritischen Wertes – oder was identisch ist, ist die zum Testwert t gehörende Wahrscheinlichkeit (p-value) kleiner als α – , dann ist die Nullhypothese abzulehnen. Der p-value ist also – unabhängig von α – die Wahrscheinlichkeit dafür, dass die Nullhypothese fälschlicherweise abgelehnt wird. Je geringer die gewählte Irrtumswahrscheinlichkeit α (Fehler erster Art, α-Fehler), desto schwieriger ist es, die Nullhypothese abzulehnen.

Die Nullhypothese wird im Allgemeinen so formuliert, dass bei Ablehnung (zugunsten der Alternativhypothese) eine möglichst große Sicherheit besteht. Das heißt, dass die Irrtumswahrscheinlichkeit α klein und damit über die Annahme der H_0 als wahre Hypothese berechenbar ist (der Fehler zweiter Art, der β-Fehler, ist nicht direkt berechenbar). Die eigentlich zu bestätigende Hypothese wird somit im Normalfall als Alternativhypothese formuliert.

In den nun folgenden Tests der Annahmen des klassischen Regressionsmodells (CLR) wird allerdings die zu bestätigende Hypothese, der eigentlich interessierende Sachverhalt, nicht als Alternativhypothese sondern als Nullhypothese formuliert. So wird beispielsweise die Normalverteilungsannahme der Residuen als H_0 formuliert und jede andere Verteilung als H_A. Wäre nämlich jede andere Verteilung die H_0, dann könnte der α-Fehler nicht berechnet werden.

Die Wahl eines sehr kleinen α-Fehlers, um relativ sichere Aussagen bezüglich der Alternativhypothese treffen zu können, ist hier also unpassend (da sie in unserem Beispiel ja nur jede beliebige Verteilung relativ sicher macht).

Höhere α-Fehler dagegen belassen ein kleineres Konfidenzintervall und stehen daher für eine kritische Haltung gegenüber einer CLR-Annahme (Normalverteilung als Beispiel); die Gültigkeit einer CLR-Annahme ist dann leichter zu verwerfen.

1 Spezifikation und Modellselektion

Zentrale Fragen:

- Ist das Modell richtig spezifiziert? Sind also die zentralen erklärenden Variablen vorhanden und „passt“ die funktionale Form (deterministischer Teil)?
- Gilt der vermutete Zusammenhang für den gesamten Bereich der Daten oder nur für Abschnitte daraus?

Die erste Frage spricht ausgelassene (omitted) Variablen und Konsequenzen davon an. Bestimmte Arten von Fehlspezifikationen des Regressionsmodells können mit dem RESET- Test überprüft werden. Die zweite Frage der Strukturkonstanz lässt sich mit einem F-Test (Chow-Test) testen. Die damit verbundene Frage der Instabilität der Regressionskoeffizienten ist zudem mit dem CUSUM- und CUSUMSQ-Test überprüfbar.

1.1 Verzerrung durch ausgelassene Variablen (Omitted Variable Bias)

Werden wichtige Variablen, die einen Einfluss auf die erklärte Variable haben und mit anderen erklärenden Variablen korrelieren, in einem Regressionsmodell nicht berücksichtigt, dann wird die CLR-Annahme, dass es keine Korrelation zwischen einer der erklärenden Variablen und dem Fehlerterm ε gibt, verletzt. Die geschätzten Koeffizienten geben dann nicht mehr alleine den kausalen Effekt einer erklärenden Variablen x auf y an. Die geschätzten Koeffizienten sind vielmehr eine Mischung des kausalen Effekts von x und den mit dieser Variablen korrelierten Faktoren, die im Fehlerterm vorhanden sind. Die geschätzten Koeffizienten sind verzerrt, im Extremfall ist sogar ein Vorzeichenwechsel durch das Auslassen wichtiger Variablen möglich.

Zur Illustration:

Es sei das wahre Modell

$$\mathbf{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon ,$$

für das aber x_2 im Datensatz nicht vorhanden ist, oder eventuell nicht berücksichtigt wurde. Geschätzt wird

$$\mathbf{y} = \beta_0 + \beta_1 x_1 + \upsilon .$$

Die nicht beobachtete Variable ist nun Bestandteil des Fehlerterms $\upsilon = \beta_2 x_2 + \varepsilon$.

Die Konsequenz ausgelassener (omitted) Variablen ist eine verzerrte Schätzung der Regressionskoeffizienten. Mit Greene (2003, 148-149) sei allgemein das korrekt spezifizierte Modell

$$\mathbf{y} = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

wobei die zwei Teile von X K_1 und K_2 erklärende Variablen haben. Die Regression ohne den zweiten Teil ergibt den OLS-Schätzer

$$\begin{aligned} b_1 = \hat{\beta} &= (X_1'X_1)^{-1}X_1'y = (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \varepsilon) \\ &= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\varepsilon) \end{aligned}.$$

Der erwartete Wert von b_1 ist um die beiden Zusatzterme verzerrt, es sei denn $X_1'X_2 = 0$ oder $\beta_2 = 0$

$$E[E[b_1 \mid X] = \beta_1 + P_{12}\beta_2 \qquad \text{mit} \qquad P_{12} = (X_1'X_1)^{-1}X_1'X_2 \quad \text{und} \qquad E(\varepsilon) = 0.$$

Jede Spalte der $K_1 * K_2$ Matrix P_{12} ist der Spaltenvektor der Steigungen der OLS-Regression der entsprechenden Spalte von X_2 auf die Spalte von X_1.

Die Richtung der Verzerrung hängt sowohl von β_2 als auch von P_{12} ab.

	$Corr(x_1, x_2) > 0$	$Corr(x_1, x_2) = 0$	$Corr(x_1, x_2) < 0$
$\beta_2 > 0$	positive Verzerrung	keine Verzerrung	negative Verzerrung
$\beta_2 = 0$	keine Verzerrung	keine Verzerrung	keine Verzerrung
$\beta_2 < 0$	negative Verzerrung	keine Verzerrung	positive Verzerrung

So ist zum Beispiel $\hat{\beta}_1$ positiv verzerrt, d.h. größer als der „wahre" zu schätzende Parameter β_1, wenn $\beta_2 < 0$ gilt (x_2 hat einen negativen Einfluss auf y) und x_1 und x_2 negativ korrelieren (x_1 und x_2 seien hier Variablen).

Im 'omitted variable' Fall ist der OLS-Schätzer von β nicht nur verzerrt (biased) sondern auch inkonsistent, da mit wachsendem Stichprobenumfang die Zusatzterme nicht verschwinden.

Zur Verzerrung durch unbeobachtete Variablen siehe Wooldridge 2006, S. 89-94, der auch das ‚omitted variable problem' ausführlicher behandelt.

Beispiel $\text{wage} = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{abil} + \varepsilon$

Angenommen wir wollen den kausalen Effekt der Bildungsjahre (educ) auf den Stundenlohn (wage) bestimmen. Für das „wahre“ Modell gilt

$$\text{wage} = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{abil} + \varepsilon .$$

Weil wir keine Informationen über die Begabung/Intelligenz (abil(ity)) der Personen in unserem Datensatz haben oder weil wir nicht wussten, dass auch die Begabung/Intelligenz (abil) einen Einfluss auf den Stundenlohn hat, schätzen wir lediglich

$$\text{wage} = \beta_0 + \beta_1 \text{educ} + \upsilon \qquad \text{mit } \upsilon = \beta_2 \text{abil} + \varepsilon$$

und erhalten zum Beispiel (Hourly Wage Equation, Wooldridge 2006, S. 89-94)

$$\widetilde{\text{wage}} = 3,374 + 1,232\text{educ}$$

mit $n = 533$ und $R^2 = 0,224$.

Wenn wir eine mögliche Verzerrung ignorieren, würden wir sagen, dass ein zusätzliches Bildungsjahr (educ) unseren späteren Stundenlohn (wage) um durchschnittlich 1,23 Euro erhöht. Wenn wir allerdings davon ausgehen, dass die Begabung/Intelligenz (abil) einen positiven Einfluss auf den Stundenlohn (wage) hat und dass die Begabung/Intelligenz (abil) positiv mit den Bildungsjahren (educ) korreliert, dann müssten wir annehmen, dass der ermittelte Koeffizient positiv verzerrt ist bzw. der „wahre“ Effekt der Bildungsjahre (educ) auf den Stundenlohn (wage) unterhalb der ermittelten 1,23 Euro liegt (siehe Wooldridge 2006, S. 89-94).

Ausgelassene (omitted) Variablen: Was kann getan werden?

Was kann getan werden, wenn eine für wichtig erachtete Variable nicht im Datensatz vorliegt und deshalb nicht bei der Schätzung berücksichtigt werden kann?

- Die Schätzung wird ohne die Variable durchgeführt, eine Verzerrung durch unbeobachtete Variablen wird in Kauf genommen. Bei der Interpretation der Schätzergebnisse wird dann auf eine mögliche Verzerrung hingewiesen und es wird diskutiert, ob durch Annahmen zur Richtung einer Korrelation mit im Model vorhandenen Variablen, eine Richtung der Verzerrung abgeleitet werden kann (siehe obiges Beispiel); x_2 hat keinen Einfluss auf y , d.h. $\beta_2 = 0$.
- Es wird nach einer Proxy-Variablen im Datensatz gesucht, die stellvertretend für die fehlende Variable in das Regressionsmodell aufgenommen wird (siehe Wooldridge 2006, S. 306 ff.).
- Verwendung von Instrumentvariablen-Schätzung (siehe hierzu Wooldridge 2006, S. 506 ff.).
- Liegen Paneldaten vor, kann unter Verwendung von panelökonometrischen Methoden die unbeobachtete Heterogenität kontrolliert werden.

Der nachfolgende RESET-Test testet die Spezifikation des Modells, dabei können bestimmte Arten von Fehlspezifikationen identifiziert werden. Da jedoch nicht in jedem Fall das Problem von ausgelassenen Variablen erkannt wird, sollte eine Modellspezifikation auf der Grundlage detaillierter theoretischer Überlegungen gefunden werden.

Literatur: z.B. Bauer, Fertig und Schmidt 2009, Kapitel 8, Wooldridge 2006, S. 89 ff oder Studenmund 2006, Kapitel 6.1.

1.2 Spezifikation, Funktionale Form und RESET-Test

Mit dem RESET-Test kann allgemein die Spezifikation der funktionalen Form überprüft werden. Zwar kann durch Variablentransformation eine Vielzahl nichtlinearer Ansätze zu linearen Ansätzen transformiert werden, und somit die funktionale Form der ‚Punktwolke' weitgehend angepasst werden, es bleibt aber die Frage, ob auch ein transformierter Ansatz "passt". Zu testen wäre dann die lineare Form nach Transformation. Der RESET-Test ist ein allgemeiner Test, der allerdings keinen Hinweis auf die Ursache der Fehlspezifikation gibt.

Transformationsbeispiele:

$y = b_0 + b_1 x^2$. Mit $z = x^2$ als Transformation ergibt sich linear $y = b_0 + b_1 z$

Modell	Strukturform	lineare Form	Steigung dy/dx	Elastizität* (dy/dx)x/y
Linear	$y = b_0 + b_1 x$	$y = b_0 + b_1 x$	b_1	$b_1 (x/y)$
Quadratisch	$y = b_0 + b_1 x^2$	$y = b_0 + b_1 (x^2)$	$b_1 (2x)$	$b_1 (2x^2/y)$
Reziprok	$y = b_0 + b_1 1/x$	$y = b_0 + b_1 (1/x)$	$-b_1 (1/x^2)$	$-b_1 [1/(xy)]$
Log-linear	$y = \exp(b_0 + b_1 \ln x)$	$\ln y = b_0 + b_1 \ln x$	$b_1 (y/x)$	b_1
Log-lin	$y = \exp(b_0 + b_1 x)$	$\ln y = b_0 + b_1 x$	$b_1 (y)$	$b_1 (x)$
Lin-log	$\exp(y) = \exp(b_0 + b_1 \ln x)$	$y = b_0 + b_1 \ln x$	$b_1 (1/x)$	$b_1 (1/y)$

* Elastizitäten meist bei den Mittelwerten von x und / oder y berechnet

Ramsey`s RESET-Test

Der RESET-Test ist ein genereller Test auf die Modellspezifikation. Dieser Test wird angewendet, wenn überprüft werden soll, ob im gewählten Ansatz wichtige Erklärungsvariablen fehlen (ausgelassene Variablen, omitted variables). Er testet die funktionale Form, indem zusätzlich das Quadrat der geschätzten y-Werte $\hat{y}_i$ in die Regressionsgleichung aufgenommen und auf Signifikanz getestet wird. Die Idee dahinter: quadratischer Ansatz als Taylor-Entwicklung höherer Ordnung.

RESET-Testprozedur

1. Schätze
   $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
   $\rightarrow \hat{\mathbf{y}} = \mathbf{Xb}$
2. Schätze
   $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{y}}^{*}\alpha + \boldsymbol{\varepsilon}$
   $\rightarrow \hat{\mathbf{y}} = \mathbf{Xb} + \hat{\mathbf{y}}^{2}a_2 (+\hat{\mathbf{y}}^{3}a_3 + \hat{\mathbf{y}}^{4}a_4 + ... + \hat{\mathbf{y}}^{p}a_p)$
   wobei $\hat{\mathbf{y}}^{*}$ die quadrierten $\hat{y}_i$ bzw. auch Terme höherer Ordnungen enthält.
3. Teste a auf Signifikanz (funktionale Form: quadratisch), z.B. mit t-Test.
   Teste, ob sich R^2 signifikant verbessert hat gegenüber dem Ansatz aus 1 mit
   $$F = \frac{(R_2^2 - R_1^2)/(K_2 - K_1)}{(1-R_2^2)/(n-K_2-1)}$$
   Freiheitsgrade $\upsilon_1 = K_2 - K_1 = p-1,\ \upsilon_2 = n - K_2 - 1$ mit K_i als Anzahl der erklärenden Variablen (ohne Konstante) in Modell i.
   Ist $F(\upsilon_1;\upsilon_2) > F_{krit}(\alpha)$, dann wird H_0: keine Verbesserung (no omitteld variables), verworfen; es liegt dann ein Spezifikationsfehler vor (z.B. es gibt ausgelassene Variablen, die abeer nicht im Modell berücksichtigt sind).

Für einen Ansatz mit einer höheren Ordnung (p) als der quadratische (p = 2) hat der t-Test $\upsilon = T - K - (p-1)$, der F-Test $\upsilon_1 = K_2 - K_1 = p-1,\ \upsilon_2 = T - K_2 - (p-1)$ und der χ^2-Test $\upsilon = p-1$ Freiheitsgrade.

Beispiel RESET Test: US-Gasoline Sales

```
ET: Gasoline Sales in the US Market

Gasoline sales in the US market from 1960 – 1982.

      G     = gasoline sales, in billions of gallons
      PG    = price index for gasoline
      Y     = per capita income
      PNC   = price index for new cars
      PUC   = price index for used cars
      PPT   = price index for public transportation
      PD    = price index for consumer durables
      PN    = price index for nondurables
      PS    = price index for services

DATA LISTING (Current sample)
Year  G       PG     Y    PNC    PUC    PPT    PD     PN     PS
1960 129.7   .925  6036  1.045   .836   .810   .444   .331   .302
1961 131.3   .914  6113  1.045   .869   .846   .448   .335   .307
1962 137.1   .919  6271  1.041   .948   .874   .457   .338   .314
1963 141.6   .918  6378  1.035   .960   .885   .463   .343   .320
1964 148.8   .914  6727  1.032  1.001   .901   .470   .347   .325
1965 155.9   .949  7027  1.009   .994   .919   .471   .353   .332
1966 164.9   .970  7280   .991   .970   .952   .475   .366   .342
1967 171.0  1.000  7513  1.000  1.000  1.000   .483   .375   .353
1968 183.4  1.014  7728  1.028  1.028  1.046   .501   .390   .368
1969 195.8  1.047  7891  1.044  1.031  1.127   .514   .409   .386
1970 207.4  1.056  8134  1.076  1.043  1.285   .527   .427   .407
```

```
1971 218.3  1.063  8322  1.120  1.102  1.377   .547   .442   .431
1972 226.8  1.076  8562  1.110  1.105  1.434   .555   .458   .451
1973 237.9  1.181  9042  1.111  1.176  1.448   .566   .497   .474
1974 225.8  1.599  8867  1.175  1.226  1.480   .604   .572   .513
1975 232.4  1.708  8944  1.276  1.464  1.586   .659   .615   .556
1976 241.7  1.779  9175  1.357  1.679  1.742   .695   .638   .598
1977 249.2  1.882  9381  1.429  1.828  1.824   .727   .671   .648
1978 261.3  1.963  9735  1.538  1.865  1.878   .769   .719   .698
1979 248.9  2.656  9829  1.660  2.010  2.003   .821   .800   .756
1980 226.8  3.691  9722  1.793  2.081  2.516   .892   .894   .839
1981 225.6  4.109  9769  1.902  2.569  3.120   .957   .969   .926
1982 228.8  3.894  9725  1.976  2.964  3.460  1.000  1.000  1.000
1983 239.6  3.764  9930  2.026  3.297  3.626  1.041  1.021  1.062
1984 244.7  3.707 10421  2.085  3.757  3.852  1.038  1.050  1.117
1985 245.8  3.738 10563  2.152  3.797  4.028  1.045  1.075  1.173
1986 269.4  2.921 10780  2.240  3.632  4.264  1.053  1.069  1.224
```

1. Schritt

```
===========================================================================
Ordinary    Least Squares
Dependent Variable        G              Number of Observations           27
Mean of Dep. Variable     207.0333       Std. Dev. of Dep. Var.    43.798927
Durbin Watson statistic      .7023       Estimated Autocorrelation    .64884
Std. Error of Regr.         5.4157       Sum of Squared Residuals   586.606
Total Variation         49877.           Regression Variation       49290.
Regression degrees of freedom =  6       Residual degrees of freedom =    20
R - squared                 .98824       Adjusted R - squared         .98471
F( 6,  20)                280.0881       Prob. Value for F            .00000
Akaike Information         3.59703       Amemiya Prediction         36.93444
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant  -10209.5      4644.       -2.198   .03987
YEAR       5.21657      2.401        2.172   .04203   1973.00000    7.93725
PG        -18.4347      3.695       -4.989   .00007      1.90211    1.16791
Y          .187961E-01   .9968E-02   1.886   .07396   8513.51852 1455.62903
PNC        27.4422      18.65        1.471   .15680      1.38133     .42797
PUC       -14.4631      8.007       -1.806   .08593      1.71230     .97474
PPT       -7.44493      7.654        -.973   .34235      1.86233    1.10831
```

2. Schritt:

ET-Command: CREATE; YFIT2=YFIT*YFIT$

```
===========================================================================
Ordinary    Least Squares
Dependent Variable        G              Number of Observations           27
Mean of Dep. Variable     207.0333       Std. Dev. of Dep. Var.    43.798927
Durbin Watson statistic      .8744       Estimated Autocorrelation    .56280
Std. Error of Regr.         5.1161       Sum of Squared Residuals   497.305
Total Variation         49877.           Regression Variation       49380.
Regression degrees of freedom =  7       Residual degrees of freedom =    19
R - squared                 .99003       Adjusted R - squared         .98636
F( 7,  19)                269.5136       Prob. Value for F            .00000
Akaike Information         3.50596       Amemiya Prediction         33.92922
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant  -2326.50      6121.        -.380   .70808
YEAR       1.27988      3.113         .411   .68553   1973.00000    7.93725
```

PG	10.8722	16.25	.669	.51139	1.90211	1.16791
Y	-.812493E-02	.1735E-01	-.468	.64493	8513.51852	1455.62903
PNC	-69.0069	55.11	-1.252	.22569	1.38133	.42797
PUC	6.67999	13.72	.487	.63190	1.71230	.97474
PPT	5.76646	10.17	.567	.57737	1.86233	1.10831
YFIT2	.290817E-02	.1574E-02	1.847	.08036	44688.37095	7012.11792

3. Schritt:

F-Test auf Veränderung der Gesamterklärungsgüte (R^2)

$R_1^2 = .98824$, $R_2^2 = .99003$; $K_1 = 6$; $K_2 = 7$

$$F = \frac{(R_2^2 - R_1^2)/(K_2 - K_1)}{(1 - R_2^2)/(n - K_2 - 1)} = 3{,}4112$$

Da $F_{krit}(1,19)(\alpha=0.05)=4{,}381 > F=3{,}4112$ ist, kann H_0: keine Änderung im R^2, nicht verworfen werden:

Der zweite Ansatz ist nicht signifikant besser.

t-Test auf Signifikanz von YFIT2:

Da der p-value (Prob|t|>x)von b(YFIT2)=.08036 > α=.05, kann H_0: YFIT2 ist nicht signifikant also nicht verworfen werden. Damit hat der quadratische Term keinen Einfluss. Es liegt keine Fehlspezifikation mit dem ursprünglichen Ansatz vor.

1.3 Test auf Strukturkonstanz 1: F-Test (Chow-Test)

Strukturbruch

Ein Strukturbruch liegt vor, wenn ein oder mehrere Schätzkoeffizienten von einem Indexbereich $I_A = (i_1, ..., i_m)$ zum anderen Bereich $I_B = (i_{m+1}, ..., i_n)$ ihre Werte verändern. Damit sind nicht nur Strukturbrüche in der zeitlichen Entwicklung angesprochen, sondern auch allgemein hinsichtlich anderer sozio-ökonomischer Variablen (z.B. Unterschiede zwischen Mann und Frau).

H_0: Es liegt Strukturkonstanz vor

H_A: Strukturkonstanz liegt nicht vor

Wenn der Bruchpunkt bekannt ist, dann kann ein F-Test, bekannt als Chow-Test, angewendet werden. Wenn aber weder die Anzahl noch der zeitliche oder strukturelle Bruchpunkt (shift-point) bekannt ist, dann empfiehlt sich der CUSUM Test.

Beispiel: Daten 1960 - 1990 Vermuteter Strukturbruch 1972/1973 Ölpreisschock

Frage:

Unterscheiden sich – aufgrund eines vermuteten Strukturbruchs 1972/1973 verursacht durch den damaligen Ölpreisschock – die Koeffizienten einer Regression 1960 - 1972 von den Koeffizienten der Regression 1973 - 1990?

Strukturbruch durch Dummy - Variablen berücksichtigen:

$$d_{A_i} = \begin{cases} 1 & i \in I_A \\ 0 & i \in I_B \end{cases} \qquad d_{B_i} = \begin{cases} 0 & i \in I_A \\ 1 & i \in I_B \end{cases}$$

Testidee ähnlich wie F-Test zur Homoskedastizität

Gruppe A z.B. 1960 - 1972 Gruppe B z.B. 1973 - 1990

Bilde gruppenspezifische Regressionen:

$$\hat{\hat{y}}_i = b_0^A d_{A_i} + b_1^A x_i d_{A_i} + b_0^B d_{B_i} + b_1^B x_i d_{B_i}$$

$$\hat{y}_i = b_0 + b_1 x_i \quad \left(i \in I_A \cup I_B\right)$$

$$H_0: b^A = b^B \qquad H_1: b^A \neq b^B$$

$$y_i = \hat{y}_i + e_i$$

$$y_i = \hat{\hat{y}}_i + \hat{e}_i \Rightarrow e_i = \hat{e}_i + \left(\hat{\hat{y}}_i - \hat{y}_i\right)$$

$$\sum_{i=1}^{n} e_i^2 = \sum\left[\hat{e}_i + \left(\hat{\hat{y}}_i - \hat{y}_i\right)\right]^2 = \sum \hat{e}_i^2 + 2\underbrace{\sum \hat{e}_i\left(\hat{\hat{y}}_i - \hat{y}_i\right)}_{0} + \sum\left(\hat{\hat{y}}_i - \hat{y}_i\right)^2$$

$$\sum\left(\hat{\hat{y}}_i - \hat{y}_i\right)^2 = \sum_{i=1}^{n} e_i^2 - \left[\sum_{i=1}^{n_A} \hat{e}_{Ai}^2 + \sum_{i=n_A+1}^{n} \hat{e}_{Bi}^2\right]$$

Nun sind

$$\frac{\sum e_i^2}{\sigma^2}, \frac{\sum \hat{e}_{Ai}^2}{\sigma^2}, \frac{\sum \hat{e}_{Bi}^2}{\sigma^2} \quad \text{jeweils } \chi^2_{n-K-1} - \text{verteilt}$$

$$\Rightarrow z_1 = \frac{\sum \hat{e}_{Ai}^2}{\sigma^2} + \frac{\sum \hat{e}_{Bi}^2}{\sigma^2} \sim \chi^2_{N_A-K-1}$$

$$z_2 = \frac{\sum\left(\hat{\hat{y}}_i - \hat{y}_i\right)}{\sigma^2} \sim \chi^2_{N_B-k-1}$$

$$\Rightarrow F_{beob} = \frac{z_2/K+1}{z_1/N-2K-2} = \frac{\sum\left(\hat{\hat{y}}_i - \hat{y}_i\right)}{\sum \hat{e}_{Ai}^2 + \sum \hat{e}_{Bi}^2} \cdot \frac{N-2K-2}{K+1} \sim F\left(K+1, N-2K-2\right) - \text{verteilt}$$

Überprüfen der Nullhypothese H_0: Strukturkonstanz

- Irrtumswahrscheinlichkeit $\alpha \rightarrow F_{krit}$
- Ist $F_{beob} > F_{krit} \rightarrow H_0$ ablehnen, Strukturbruch signifikant

Zusammenfassung F-Test auf Strukturkonstanz (Chow Test)

1. Signifikanzniveau von α festlegen
2. Bilde $I_A: 1 \leq i \leq i_m \qquad I_B: i > i_m$

   Regression: $\hat{\hat{y}}_i = b_0^A d_{Ai} + b_1^A x_i d_{Ai} + b_0^B d_{Bi} + b_1^B x_i d_{Bi}$

   $\hat{y}_i = b_0 + b_1 x_i \quad (\forall\, i = 1, \ldots, n)$

   Berechne: $\sum_{i \in I_A} \hat{e}_{A_i}^2, \qquad \sum_{i \in I_B} \hat{e}_{B_i}^2 \qquad \Big|\, y_i = \hat{\hat{y}}_i + \hat{e}_i, \qquad y_i = \hat{y}_i + e_i$

   $\sum_{i=1}^{n} \left(\hat{\hat{y}}_i - \hat{y}_i\right)^2$

   $$F_{beob} = \frac{\sum_{i=1}^{n} \left(\hat{\hat{y}}_i - \hat{y}_i\right)^2}{\sum_{i \in I_A} \hat{e}_{A_i}^2 + \sum_{i \in I_B} \hat{e}_{B_i}^2} \frac{N - 2K - 2}{K + 1}$$
3. Aus der Tabelle entnehmen

   $F_{krit} = F_\alpha$
4. Für $F_{beob} > F_{krit}$ H_0 (Strukturkonstanz verwerfen)

   $F_{beob} \leq F_{krit} \qquad H_0$ nicht verwerfen

Beispiel F-Test auf Strukturkonstanz

$BSP = b_0 + b_1 \text{Geldmenge} \quad$ (wie oben)

Vermuteter Strukturbruch 1976/ 1977

$\Rightarrow I_A = 1973, \ldots, 1976\ (i = 1, 2, 3, 4) \qquad I_B = 1977, \ldots, 1982\ (i = 5, 6, 7, 8, 9, 10)$

1. Signifikanzniveau

$\alpha = 5\%$ gewählt

2. OLS

$\hat{\hat{y}}_i = \widehat{\widehat{BSP}} = 2{,}62 d_{Ai} + 1{,}15 (\text{Geldmenge} \cdot d_{Ai}) + 1{,}2 d_{Bi} + 1{,}73 (\text{Geldmenge} \cdot d_{Bi})$

$\hat{y}_i = \widehat{BSP} = 1{,}17 + 1{,}72 \text{Geldmenge}$

$\sum_{i \in I_A} \hat{e}_{Ai}^2 = 0{,}1538462 \qquad \sum_{i \in I_B} \hat{e}_{Bi}^2 = 0{,}3513052$

$\sum_{i=1}^{n} \left(\hat{\hat{y}}_i - \hat{y}\right)^2 = 0{,}6121608$

$$F_{beob} = \frac{\sum \left(\hat{\hat{y}}_i - \hat{y}_i\right)^2}{\sum \hat{e}_{Ai}^2 + \sum \hat{e}_{Bi}^2} \cdot \frac{n - 4}{2} = \frac{0{,}6121608}{0{,}1538462 + 0{,}3513052} \cdot 3 = 3{,}6355$$

3. F_{krit}

$F_{krit} = F_{0{,}05}^{(2,6)} = 5{,}14$

4. H_0 ablehnen?

Da $F_{beob} = 3{,}64 < F_{krit} = 5{,}14$ kann H_0: $b^A = b^B$, also Strukturkonstanz nicht verworfen werden (bei $\alpha = 5\%$).

Berechnung mit Econometric Toolkit (ET) oder LIMDEP

```
? -------------------------------
? ET F-Test auf Strukturbruch
? -------------------------------
?
? read data
?      y : bsp
?      x0: one
?      x1: money
?      x2: da
?      x3: db
read; file=struktur.dat;nvar=5;nobs=10;names=1$
list; bsp,money,da,db$

create; xda=money*da
      ;  xdb=money*db$

regress; dep=bsp; ind=da,xda,db,xdb$
create; yfit2=yfit$
calculate; ssr=SUMSQDEV$

regress; dep= bsp; ind= one,money$
create; yfit1=yfit
      ; help=(yfit2-yfit1)^2$
calculate; Fbeob=Sum(help)/ssr*3$
```

ET-Ergebnis: F-Test auf Strukturkonstanz

```
DATA LISTING (Current sample)
Observation      BSP        MONEY       DA          DB
      1        5.0000      2.0000      1.0000      .00000
      2        5.5000      2.5000      1.0000      .00000
      3        6.0000      3.2000      1.0000      .00000
      4        7.0000      3.6000      1.0000      .00000
      5        7.2000      3.3000      .00000      1.0000
      6        7.7000      4.0000      .00000      1.0000
      7        8.4000      4.2000      .00000      1.0000
      8        9.0000      4.6000      .00000      1.0000
      9        9.7000      4.8000      .00000      1.0000
     10        10.000      5.0000      .00000      1.0000
 1. Command: -> XDA=MONEY*DA
 2. Command: -> XDB=MONEY*DB
===========================================================================
Ordinary    Least Squares
Dependent Variable        BSP        Number of Observations           10
Mean of Dep. Variable      7.5500     Std. Dev. of Dep. Var.     1.732211
Std. Error of Regr.         .2902     Sum of Squared Residuals     .505151
R - squared                .98129     Adjusted R - squared          .97194
F( 3,   6)               104.9185     Prob. Value for F            .00001
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
DA          2.61538       .6789       3.852   .00844       .40000      .51640
XDA         1.15385       .2348       4.915   .00267      1.13000     1.51588
DB          1.20484       .9097       1.324   .23358       .60000      .51640
XDB         1.72861       .2090       8.273   .00017      2.59000     2.27667
```

```
 1. Command: -> YFIT2=YFIT

>>> SSR      =     .5051512     <<<

===========================================================================
Ordinary    Least Squares
Dependent Variable         BSP          Number of Observations          10
Mean of Dep. Variable       7.5500      Std. Dev. of Dep. Var.     1.732211
Std. Error of Regr.          .3737      Sum of Squared Residuals    1.11731
R - squared                 .95863      Adjusted R - squared         .95345
F( 1,   8)                185.3569      Prob. Value for F            .00000
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant   1.16814       .4834       2.416   .04208
MONEY      1.71555       .1260      13.615   .00000      3.72000     .98860

 1. Command: -> YFIT1=YFIT
 2. Command: -> HELP=(YFIT2-YFIT1)^2

>>> FBEOB    =     3.635511     <<<
```

F-Test (Chow-Test) auf Strukturbruch, alternative Vorgehensweise (Greene 2003, chapt. 7.4)

Die ersten Beobachtungen I_A von y und X $\rightarrow \mathbf{y}_A, \mathbf{X}_A$

Regression: $\begin{pmatrix} \mathbf{y}_A \\ \mathbf{y}_B \end{pmatrix} = \begin{pmatrix} \mathbf{X}_A & 0 \\ 0 & \mathbf{X}_B \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_A \\ \boldsymbol{\beta}_B \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_A \\ \boldsymbol{\varepsilon}_B \end{pmatrix}$

$y = X\beta + \varepsilon$

OLS: $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{pmatrix} \mathbf{X}_A{}'\mathbf{X}_A & 0 \\ 0 & \mathbf{X}_B{}'\mathbf{X}_B \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$

Der restringierte Koeffizientenvektor kann auf zwei Wegen ermittelt werden:

1. über die Restriktion: $\beta_A = \beta_B$ über $\beta = (\beta_A, \beta_B)$

$\mathbf{R}\boldsymbol{\beta} = \mathbf{q}$ mit $\mathbf{R} = [\mathbf{I} : \mathbf{-I}]$ und $\mathbf{q} = \mathbf{0}$

also $\mathbf{R}\boldsymbol{\beta} = \mathbf{q}$ $\quad [\mathbf{I}\ \mathbf{-I}] \begin{pmatrix} \boldsymbol{\beta}_A \\ \boldsymbol{\beta}_B \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$

hier: $\begin{aligned} 1\beta_{Ai} + 0\beta_{A2} - 1\beta_{B1} - 0\beta_{B2} = 0 \\ 0\beta_{Ai} + 1\beta_{A2} - 0\beta_{B1} - 1\beta_{B2} = 0 \end{aligned}$

2. in dem die Restriktion $\beta_A = \beta_B$ über $\beta = (\beta_A, \beta_B)$ direkt in das Modell eingebaut wird mit

$$\begin{pmatrix} \mathbf{y}_A \\ \mathbf{y}_B \end{pmatrix} = \begin{pmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{pmatrix} \beta + \begin{pmatrix} \boldsymbol{\varepsilon}_A \\ \boldsymbol{\varepsilon}_B \end{pmatrix}$$

(stacking the data). Die Residuensumme ist dann die Testbasis für einen F-test (Greene 2003, S. 131 und S.97)

1.4 Test auf Strukturkonstanz 2: CUSUM- und CUSUMSQ-Test

Die bisherigen Tests auf Strukturkonstanz setzen einen bestimmten vermuteten Strukturbruch voraus. Sind weder dieser Zeitpunkt (oder allgemeiner dieser Wert) noch die Anzahl der Strukturbrüche vorhanden, dann helfen der CUSUM- (cumulative sum, Brown, Durbin und Evans 1975) und CUSUMSQ- (cumulative sum of squares, Eckey, Kosfeld und Dreger 2001, 218 ff.) Test überprüfen. Allerdings kann eine fehlende Konstanz der Regressionskoeffizienten auch auf eine Fehlspezifikation des Modells zurückzuführen sein.

Falls die Regressionskoeffizienten bis zu einem bestimmten Zeitpunkt r=k+1,...,T(n) konstant sind und danach variieren, haben die Residuen bis zum Zeitpunkt des Strukturbruchs einen Erwartungswert von Null und ab der (r+1)-ten Beobachtung einen Erwartungswert von ungleich Null.

Zur Durchführung der Tests benötigt man die rekursiven Residuen. Die rekursiven Residuen werden folgendermaßen berechnet:

$$\hat{\varepsilon}_t = y_t - x_t b_{t-1}$$

$\mathbf{b}_{t-1}$ ist der Vektor der OLS-Schätzer, bei dem zur Schätzung der Koeffizienten nur die ersten (t-1) Beobachtungen herangezogen wurden.

Die Nullhypothese bei beiden Tests (CUSUM und CUSUMSQ) unterstellt die Stabilität der Parameter.

CUSUM-Test:

Unter der Nullhypothese (Stabilität der Parameter) hat die kumulierte Summe der Residuen (C_r) ein arithmetisches Mittel von Null und eine Varianz die in etwa der Anzahl der summierten Residuen entspricht.

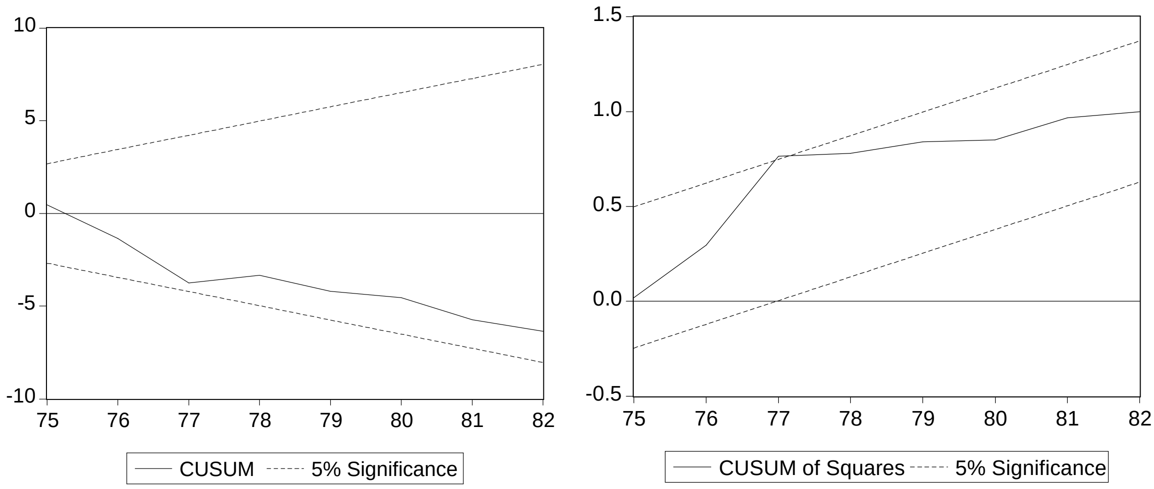
Aus der Entwicklung der kumulierten Residuen (C_r) in einem (r, C_r)-Diagramm kann man einen möglichen Strukturbruch ersehen.

$$C_r = \frac{1}{\hat{\sigma}} \sum_{j=k+1}^{r} \hat{\varepsilon}_j$$

In der nachfolgenden Grafik (vgl. Geldmengenbeispiel S. 123) ist C_r gegen r abgetragen. Die gestrichelten Linien geben das 5 % Konfidenzintervall an. Die Nullhypothese wird verworfen, wenn C_r außerhalb dieser Grenzen liegt – sich also signifikant von 0 unterscheidet. Ein Strukturbruch ist dann vor dem Überschreiten der Signifikanzlinie gegeben.

Für das Geldmengenbeispiel ist vor 1977 ein gewisser Strukturbruch sichtbar, der aber ‚knapp' signifikant ist.

CUSUM und CUSUMSQ Test für das Beispiel BSP=f(Geldmenge)



CUSUMSQ-Test:

Der CUSUMSQ(uared) Test wird preferiert, wenn die Parameterinstabilität weniger systematisch sondern eher zufällig ist. Der CUSUMSQ-Test basiert auf dem Vergleich der quadrierten Residuen mit

$$\mathrm{CSQ}_r = \sum\nolimits_{j=k+1}^{r} \hat{\varepsilon}_j^2 \Big/ \sum\nolimits_{j=k+1}^{n} \hat{\varepsilon}_j^2 = C_r / C_n \; .$$

Die CSQ folgen einer Beta-Verteilung (approximativ einer Chi^2-Verteilung). Weichen die CSQ im (r,CSQ$_r$)-Diagramm vom Signifikanzpfad – Linienpaar parallel zum Erwartungswert mit $\pm c + (r-k)/(n-k)$, der Parameter c ist eine Funktion des Signifikanzniveaus (z.B. α=0.05 dann ist c=1.115) – ab, ist dies ein Indiz für die Instabilität.

Für das Geldmengenbeispiel ist auch mit dem CUSUMSQ-Test vor 1977 ein gewisser Strukturbruch sichtbar, der aber gerade auf der Signifikanzgrenze liegt. Eine neue Spezifikation wäre deshalb zu überlegen.

1.5 Misspezifikation: Was kann getan werden?

Zentrale Fragen:

- Ist das Modell richtig spezifiziert? Sind also die zentralen erklärenden Variablen vorhanden (deterministischer Teil)?
- Gilt der vermutete Zusammenhang für den gesamten Bereich der Daten oder nur für Abschnitte daraus.

Die erste Frage wurde mit dem Reset-Test überprüft, die Frage der Strukturkonstanz testet ein dafür konzipierter F-Test (Chow-Test). Beide Tests brauchen einen vermuteten Zeitpunkt des Strukturbruchs. Ist dieser Zeitpunkt unbekannt, dann kann mit dem CUSUM- und CUSUMSQ-Test der Zeitpunktt und die Signifikanz eines Strukturbruchs ermittelt werden.

Was tun?

Gibt es über die obigen Diagnostiktests Hinweise auf eine Fehlspezifikation, also auf eine noch ungenügende Modellierung mit dem deterministischen Teil des Regressionsansatzes, dann muss die Modellformulierung eben verbessert werden durch Überdenken der Operationalisierung des theoretischen Ansatzes (Adäquationsproblem) und oder durch einen anderen funktionalen Zusammenhang.

Weitere Literatur: Bauer, Fertig und Schmidt 2009, Kapitel 6, Studenmund 2006, Kapitel 6, 7 oder Wooldridge 2006, Kapitel 9.

2 HOMOSKEDASTIZITÄT/HETEROSKEDASTIZITÄT

Eine weitere im CLR unterstellte Eigenschaft des Modells ist die der Homoskedastizität. Wird die Annahme der Homoskedastizität, also die Annahme gleicher Fehlervarianzen, verletzt, dann hat der OLS-Schätzer **b** keine Minimum-Varianz-Eigenschaft, die geschätzten Koeffizienten sind nicht mehr effizient (no longer Best, but ‚LUE'). Resultat: die gewöhnlichen Konfidenzintervalle und Hypothesentests, wie F- oder t-Test sind nicht zuverlässig mit der Gefahr, falsche Signifikanzschlüsse zu ziehen.

Zur Überprüfung dieser Annahme wurde eine Reihe von Tests entwickelt, die im Folgenden vorgestellt werden sollen. Zunächst sollen einige Beispiele die Homoskedastizität veranschaulichen.

2.1 Charakteristika und Beispiele

Heteroskedastizität der Residuen liegt dann vor, wenn $\mathrm{var}(\varepsilon_\mathrm{i}) \neq \mathrm{var}(\varepsilon_\mathrm{j})\ \mathrm{i} \neq \mathrm{j}$ ist, wenn also die Homoskedastizitätsannahme des CLR-Modells mit $\mathrm{var}(\varepsilon_i) = \mathrm{var}(\varepsilon_j)$ nicht erfüllt ist.

Die Kovarianzmatrix der Residuen hat bei Homoskedastizität die Form

$$\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \begin{pmatrix} \mathrm{E}\varepsilon_1^2 & 0 \\ & \ldots & \\ 0 & & \mathrm{E}\varepsilon_n^2 \end{pmatrix}$$

Heteroskedastizität wird vor allem in Querschnittsanalysen beobachtet, weniger in Zeitreihenanalysen.

Mögliches Muster: Varianz von ε steigt mit wachsendem x.

Beispiele Heteroskedastizität

1. **Querschnittsregression: Sparen = f(Einkommen)**

Personen/Haushalte mit höherem Einkommen (x) werden im Modell mehr sparen (y) als Personen/Haushalte mit geringerem Einkommen und mit größerer Variabilität/Streuung.

2. **Querschnittsregression: Forschung und Entwicklung = f(Firmenprofit,x)**

Ausgaben für Forschung und Entwicklung (FuE)

$$\mathrm{FuE}_\mathrm{i} = \beta_0 + \beta_1 \mathrm{profit} + \beta_2 \mathrm{x} + \varepsilon_\mathrm{i}$$

Auch hier dürften die FuE-Ausgaben – bei gegebenem weiteren x-Wert – größer und mit stärkerer Streuung bei Firmen mit größerem Gewinn (y) sein als bei Firmen mit kleinerem Gewinn.

3. Humankapital: Schooling, Experience and Earnings

Hourly earnings (y) = f(years of school (S), occupational experience (x), socio-economic variables (**z**))

Mincer's (1974) Schätzgleichung

$$\ln y_i = \beta_0 + \beta_1 x_i - \beta_2 x_i^2 + \beta_3 S_i + \boldsymbol{\gamma}'\mathbf{z} + \varepsilon_i \,,$$

wobei $\boldsymbol{\gamma}$ der Parametervektor der sozioökonomischen Variablen ist.

Warum könnte die Varianz der Einkommen in Abhängigkeit von der beruflichen Erfahrung sich ändern?

Berufsbeginn: Start als Trainee mit geringerer Bezahlung; kein großer Unterschied in der Entlohnung.

Später im Beruf: zunehmende Berufserfahrung führt zu steigender Entlohnung und zunehmender Spreizung der Einkommen.

2.2 Tests auf Homoskedastizität

2.2.1 F-Test auf Homoskedastizität

Betrifft Modellannahme des CLR A3: $E\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' = \sigma^2\mathbf{I}$

$$E\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' = \sigma^2\mathbf{I} = \begin{pmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{pmatrix}$$

Homoskedastizität:

Varianz der Fehlerterme ist für alle Beobachtungen i=1,...,n gleich; $\sigma_{ii}^2 = \sigma^2$

Heteroskedastizität:

$$\sigma_{ii}^2 \neq \sigma_{jj}^2 \qquad (i, j = 1, \ldots, n)$$

Im Generalisierten Linearen Regressionsmodell GLS ist $E\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' = \sigma^2\boldsymbol{\Omega}$

($\boldsymbol{\Omega}$ ist positiv definite Matrix $\rightarrow$ Kapitel IX)

Verwendung der aus der Schätzung hervorgegangenen $e_i = \hat{\varepsilon}_i$

Testprinzip:

1. Bilde zwei Gruppen von Beobachtungen

   A: z.B. kleinere x_k-Werte $s_A^2 = \frac{1}{n_A}\sum_{i=1}^{n_A} e_i^2$

   B: z.B. größere x_k-Werte $s_B^2 = \frac{1}{n_B}\sum_{i=1}^{n_B} e_i^2$

   andere Gruppen:
   SEX: Männer, Frauen
   LOCATION: Nord, Süd, etc.

2. Test: H_0: $\sigma_A^2 = \sigma_B^2 = \sigma_e^2$

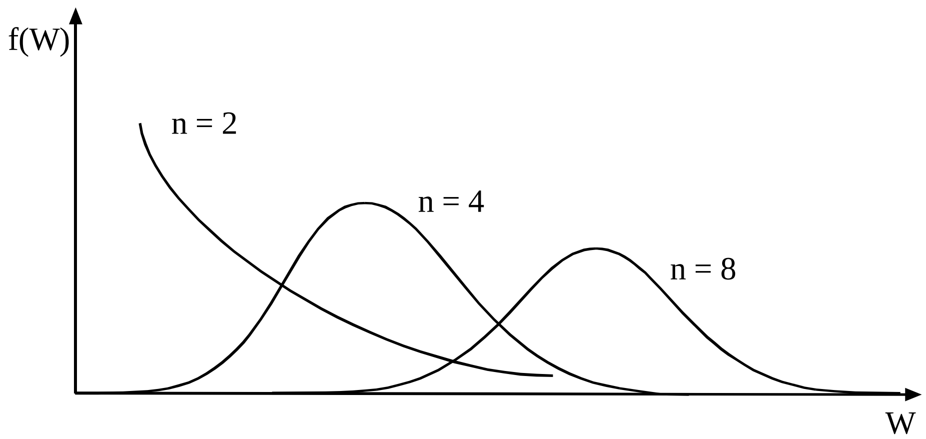
Notwendig:

Verteilungsaussage über s_A^2 und s_B^2, dafür Verteilung von $\sum_{i=1}^{n} e_i^2$ nötig.

$\rightarrow W = \sum_{i=1}^{n} Z_i^2$ ist Zufallsvariable, die χ_n^2 - verteilt ist.

Vor.: Z_i sind voneinander unabhängige N(0,1) - verteilte Zufallsvariablen.

Wahrscheinlichkeitsdichte der χ_n^2 - Verteilung ist von n abhängig.



Satz von Cochran:

Die Summe unabhängiger Chi-Quadrat-verteilter Zufallsvariablen ist wieder Chi-Quadrat-verteilt.

Zahl der Freiheitsgrade (degrees of freedom: DF) ist Summe der Einzelfreiheitsgrade

$\rightarrow C = \sum_{i=1}^{n} e_i^2$ ist χ^2 - verteilt. Freiheitsgrade: $\upsilon = (n - K - 1)$

Ist $C = \sum_{i=1}^{n} e_i^2$ χ^2 - verteilt, dann ist auch $\frac{\sum e_i^2}{\sigma_e^2}$ χ^2 - verteilt.

Vorgehensweise: Bilde Regression getrennt für jede Gruppe: OLS

A: $y_i = \mathbf{b}_A{}' \mathbf{x}_i + e_i \qquad i \in \text{Gruppe A}$

B: $y_j = \mathbf{b}_B{}' \mathbf{x}_j + e_j \qquad j \in \text{Gruppe B}$

$$\rightarrow \sum \frac{e_{A_i}^2}{\sigma_A^2} \sim \chi^2_{n_A - K - 1} \qquad \sum \frac{e_{B_i}^2}{\sigma_B^2} \sim \chi^2_{n_B - K - 1}$$

$F = \dfrac{W_1 / \upsilon_1}{W_2 / \upsilon_2}$ ist Zufallsvariable, die $F(\upsilon_1, \upsilon_2)$ - verteilt ist mit W_i = χ^2- verteilte Zufallsvariable und υ_i = Freiheitsgrade (i = 1,2).

F-Test auf Homoskedastizität:

$$F = \frac{\sum_i e_{A_i}^2}{n_A - K - 1} \Bigg/ \frac{\sum_i e_{B_i}^2}{n_B - K - 1} = \frac{n_B - K - 1}{n_A - K - 1} \frac{\sum_i e_{A_i}^2}{\sum_i e_{B_i}^2}$$

ist $F(n_A - K - 1, n_B - K - 1)$ - verteilt (ohne σ_e^2 berechenbar!)

Achtung: Größere Streuung steht immer im Zähler.

Je größer F_{beob} desto eher wird man H_0 (Homoskedastizität) ablehnen.

F_{krit} aus Tabelle nach Festlegung der Irrtumswahrscheinlichkeit.

Zusammenfassung F-Test auf Homoskedastizität

H_0: Homoskedastizität

H_1: Heteroskedastizität

1. Signifikanzniveau von α festlegen
2. Bilde A: z.B. kleinere x_k- Werte B: z.B. größere x_k- Werte

   Berechne OLS

   $$b_A \rightarrow \sum e_{A_i}^2 \qquad e_i = y_i - \mathbf{b}' \mathbf{x}_i$$

   $$b_B \rightarrow \sum e_{B_i}^2$$

   $$F_{beob} = \frac{\sum_i e_{A_i}^2}{\sum_i e_{B_i}^2} \frac{n_B - K - 1}{n_A - K - 1} \qquad \left| \begin{array}{l} \upsilon_1 = n_A - K - 1 \\ \upsilon_2 = n_B - K - 1 \end{array} \right.$$
3. aus der Tabelle entnehmen

   $$F_{krit} = F_\alpha^{(\upsilon_1, \upsilon_2)} \qquad \alpha = \Pr ob(F > F_{krit}) = 1 - \Pr ob(F \le F_{krit}) = 0{,}05$$
4. Für $F_{beob} > F_{krit}$ H_0 (Homoskedastizität) verwerfen

   $F_{beob} \le F_{krit}$ H_0: $\sigma_A^2 = \sigma_B^2 = \sigma_e^2$

Beispiel F-Test auf Homoskedastizität

Frage: Ist die Fehlervarianz bei dem Beispiel der Einkommenserklärung zwischen Jüngeren und Älteren verschieden oder nicht?

$$\text{income} = b_0 + b_1\text{age}$$

i	income	age
1	1200	22
2	1700	24
3	3500	28
4	4200	27
5	1600	23
6	5200	36

Gruppierung nach dem Alter:: Jüngere: Gruppe A: 1,2,5 $n_A = 3$
Ältere: Gruppe B: 3,4,6 $n_B = 3$

1. Signifikanzniveau

$\alpha = 5\%$ gewählt

2. OLS zweier Gruppen

Gruppe A: income= $\text{income} = -4250 + 250\text{age}$

Gruppe B: $\text{income} = -250 + 150\text{age}$

$e_{A_i} = y_i - \mathbf{b}'\mathbf{x}_i$ $\qquad e_{B_i} = y_i - \mathbf{b}'\mathbf{x}_i$

$\sum e_{A_i}^2 = 15000,0$ $\qquad \sum e_{B_i}^2 = 365000,0$

Achtung: Da $\sum e_{B_i}^2 > \sum e_{A_i}^2$ kommt e_{B_i} in den Zähler!

$$F_{beob} = \frac{365000}{15000} \cdot \frac{3-1-1}{3-1-1} = 24,33$$

3. F_{krit}

$$F_{krit} = F_{0,05}^{(1,1)} = 161,4$$

4. H_0 ablehnen?

Da $F_{beob} = 24,33 < F_{krit} = 161,4$ wird H_0: $\sigma_A^2 = \sigma_B^2 = \sigma_e^2$ (also Homoskedastizität) nicht verworfen (bei $\alpha = 5\%$); die Fehlervarianzen der Jüngeren und Älteren unterscheiden sich damit nicht signifikant.

Berechnung mit Econometrics Toolkit (ET; Computer Package by W. Greene) oder LIMDEP

```
? ---------------------------------
? ET F-Test auf Homoskedastizität
? ---------------------------------
?
? read data
?      y : bsp
?      x0: one
?      x1: money
? -----------------------------------------
? ET F-Test auf Homoskedastizit income=f(age)
? -----------------------------------------
?
? read data
?      y : income
?      x0: one
```

```
?      x1: age
?      x2: sex (0=male,1=female)
?
read; file=income.dat;nvar=3;nobs=6;names=1$
list; income, age, sex$

regress; dep=income; ind=one,age$

sample; 1,2,5$
regress; dep=income; ind=one,age$
calculate; ssra=SUMSQDEV
         ; dfa=DEGFRDM$

sample; 3,4,6$
regress; dep=income; ind=one,age$
calculate; ssrb=SUMSQDEV
         ; dfb=DEGFRDM$
calculate; Fbeob=ssrb/ssra*dfa/dfb$
```

```
ET Ergebnisse: F-Test of Homoskedastizität
------------------------------------------

DATA LISTING (Current sample)
Observation      INCOME      AGE         SEX

      1        1200.0      22.000      1.0000
      2        1700.0      24.000      1.0000
      3        3500.0      28.000      1.0000
      4        4200.0      27.000      .00000
      5        1600.0      23.000      1.0000
      6        5200.0      36.000      .00000

==============================================================================
Ordinary    Least Squares
Dependent Variable        INCOME      Number of Observations             6
Mean of Dep. Variable    2900.0000     Std. Dev. of Dep. Var.    1634.625339
Std. Error of Regr.       709.7758     Sum of Squared Residuals   .201513E+07
R - squared                 .84917     Adjusted R - squared           .81146
F( 1,   4)                 22.5194     Prob. Value for F              .00900
==============================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x   Mean of X  Std.Dev.of X
------------------------------------------------------------------------------
Constant  -4937.56       1677.      -2.945   .04220
AGE        293.909       61.93       4.745   .00900     26.66667       5.12510


Sample set to -> 1,2,5

==============================================================================
Ordinary    Least Squares
Dependent Variable        INCOME      Number of Observations             3
Mean of Dep. Variable    1500.0000     Std. Dev. of Dep. Var.     264.575131
Std. Error of Regr.       122.4745     Sum of Squared Residuals   15000.0
R - squared                 .89286     Adjusted R - squared           .78571
F( 1,   1)                  8.3333     Prob. Value for F              .21230
==============================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x   Mean of X  Std.Dev.of X
------------------------------------------------------------------------------
Constant  -4250.00       1993.      -2.132   .27917
AGE        250.000       86.60       2.887   .21230     23.00000       1.00000


>>> SSRA     =     15000.00     <<<

>>> DFA      =     1.000000     <<<


Sample set to -> 3,4,6
```

```
==============================================================================
Ordinary    Least Squares
Dependent Variable        INCOME       Number of Observations             3
Mean of Dep. Variable    4300.0000     Std. Dev. of Dep. Var.      854.400375
Std. Error of Regr.       604.1523     Sum of Squared Residuals   365000.
R - squared                 .75000     Adjusted R - squared            .50000
F( 1,   1)                 3.0000     Prob. Value for F               .33333
==============================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x   Mean of X  Std.Dev.of X
------------------------------------------------------------------------------
Constant  -250.000       2650.       -.094   .94012
AGE        150.000       86.60       1.732   .33333     30.33333       4.93288


>>> SSRB     =     365000.0     <<<

>>> DFB      =     1.000000     <<<

>>> FBEOB    =     24.33333     <<<
```

Da FBEOB=24,33 < FKRIt=161,4 wird Ho: Homoskedastizität nicht verworfen, d.h. die Fehlervarianzen sind zwischen Jüngeren und Älteren in diesem fiktiven Beispiel bei einem Signifikanzniveau von 5% gleich.

2.2.2 Weitere Tests auf Heteroskedastizität: White, Goldfeld-Quandt, Breusch-Pagan/Godfrey- und Glejser Test

Die folgenden Tests beginnen mit einem allgemeinen Ansatz (White Test) und werden dann spezifischer (Goldfeld-Quandt Test, Breusch-Pagan/Godfrey Test, Glejser Test). Dabei gilt: je allgemeiner diese Tests, desto geringer ist die Macht (Power) der Tests. Mehr dazu: Wooldridge 2006, S. 271-276.

White's genereller Heteroskedastizität Test

$H_0 : \sigma_i^2 = \sigma^2$ für alle i (Homoskedastizität)

H_1 : nicht H_0, keine Homoskedastizität, also Heteroskedastizität

White's Test (1980) sei an folgendem Beispiel illustriert:

(1) $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_{1i}$

Testschritte:

1. Schätze (1) mit OLS, um die Residuen ε_i zu erhalten
2. Schätze die Residuenquadrate durch die quadratische Form
   (2) $\varepsilon_i^2 = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \alpha_3 x_{1i}^2 + \alpha_4 x_{2i}^2 + \alpha_5 x_{1i} x_{2i} + \varepsilon_{2i}$
   also inklusive aller quadrierten X-Werte und ihrer Kreuzprodukte (allgemein $\mathbf{x} \otimes \mathbf{x}$).
3. Berechne den Bestimmtheitskoeffizienten R^2 aus (2)
   Gäbe es keine Heteroskedastizität, müssten alle α_K Null sein.
4. Teststatistik: nR^2 ist $\chi^2(\upsilon = K(\alpha))$ verteilt mit der Anzahl der Freiheitsgrade υ gleich K aus (2) ohne die Konstante

5. Ist $\chi^2 > \chi^2_{\text{krit.}}$ bei gewähltem α - oder der p- (bzw. PROB-) Wert ist kleiner als α - dann wird H_0: keine Heteroskedastizität, verworfen.

Goldfeld-Quandt-Test

Goldfeld und Quandt (1965) unterscheiden zwei Gruppen von Beobachtungen in der Weise, dass unter der Homoskedastizitäts-Hypothese die Varianz der Residuen in beiden Gruppen gleich wäre. Unter der Alternativhypothese (Heteroskedastizität) differieren die Varianzen der Residuen systematisch.

Eigentlich wäre die Gruppe mit geringerer Residuenvarianz zu vergleichen mit der Gruppe mit höherer Residuenvarianz. Damit müssten die Varianzen von jedem x-Wert zu bestimmen sein, dann die Beobachtungen aufsteigend nach der Größe der Varianzen zu ordnen und dann zwei Gruppen zu bilden sein. Dafür werden aber für jeden x-Wert zu wenige Residuen e vorhanden sein.

Eine Gruppierung kann aber nach der Größe der x-Werte erfolgen, so wie von Goldfeld und Quandt (1965) beschrieben. Damit werden die Residuenvarianzen als Funktion sich ändernder x-Werte spezifiziert.

Beispiele: Gruppenweise Heteroskedastizität oder $\sigma_i^2 = \sigma^2 x_i^2$.

Testschritte:

1. Ordnen (‚Ranking') der Beobachtungen mit x_i von hohen bis zu niedrigen x_i-Werten (eigentlich Varianzen der Residuen)
   Gruppe 1: n_1 Gruppe 2: n_2
2. Berechne zwei Regressionen (für Gruppe 1 und Gruppe 2)
3. Teststatistik

$$F(n_1 - K, n_2 - K) = \frac{\mathbf{e}_1'\mathbf{e}_1/(n_1 - K - 1)}{\mathbf{e}_2'\mathbf{e}_2/(n_2 - K - 1)}$$

   mit größeren Residuenvarianzen in der ersten Gruppe
   Ist $F > F_{\text{krit}}$ bei gewähltem α, dann wird H_0: keine Heteroskedastizität, verworfen.

Breusch-Pagan/Godfrey-Test

Der Test von Breusch und Pagan (1979) gruppiert wie im Goldfeld-Quandt Test, allerdings wird zugelassen, dass die Varianz der Residuen sich in Abhängigkeit eines Sets von Regressoren ändert:

$$\sigma_i^2 = \sigma^2 f(\alpha_0 + \boldsymbol{\alpha}'\mathbf{z}_i) \text{ mit } \mathbf{z}_i \text{ als Vektor unabhängiger Variablen.}$$

Homoskedastizität: $\boldsymbol{\alpha} = \mathbf{0}$.

Testschritte:

H_0: Homoskedastizität

H_1: Heteroskedastizität

1. Ordnen der Beobachtungen wie bei Goldfeld-Quandt.
2. Berechne die Regression

   $y_i = e_i^2 = \boldsymbol{\alpha}'\mathbf{z}_i + \varepsilon$ bzw. $y_i = e_i / (e'e/n) = \boldsymbol{\alpha}'\mathbf{z}_i + \varepsilon$ (Greene 2000, S. 510)
3. Teststatistik

   Lagrange Multiplier Test

   $\text{LM} = 1/2 \sum_i \left(y_i - \overline{y}\right)^2$
4. Mit H_0: Homoskedastizität, ist LM asymptotisch χ^2 verteilt mit der Anzahl der Freiheitsgraden υ gleich der Anzahl der Variablen in z.

   Ist $\chi^2 > \chi^2_{krit}$ bei gewähltem Signifikanzniveau, dann wird H_0: keine Heteroskedastizität, verworfen.

Glejser-Test

Testidee: Im Falle von Heteroskedastizität kann die Varianz systematisch verbunden sein mit einer oder mehreren erklärenden Variablen. Glejser (1969) testet spezifische Formen der Residuen Varianz:

$$\operatorname{var}(\varepsilon_i) = \sigma^2 \left(\boldsymbol{\alpha}'\mathbf{z}_i\right),\ \operatorname{var}(\varepsilon_i) = \sigma^2 \left(\boldsymbol{\alpha}'\mathbf{z}_i\right)^2 \text{ und } \operatorname{var}(\varepsilon_i) = \sigma^2 \exp\left(\boldsymbol{\alpha}'\mathbf{z}_i\right)$$

Testschritte:

1. Berechne die Ausgangsregression, um e_i zu erhalten.
2. Schätze

   $|e_i| = \alpha_0 + \alpha_1 x_i + \varepsilon_{2i}$ oder

   $|e_i| = \alpha_0 + \alpha_2 \sqrt{x_i} + \varepsilon_{2i}$ oder

   $|e_i| = \alpha_0 + \alpha_1 \left(\frac{1}{x_i}\right) + \varepsilon_{2i}$

   Statt $|e_i|$ wird auch e_i^2 oder ln $|e_i|$ einbezogen. Der Park Test (1966) verwendet hier e_i^2 oder ln $|e_i|$ als Regressand.
3. Teststatistik

   z. B. t-Test (Glejser: Wald Test) auf Signifikanz von α_2.
4. Ist $t > t_{krit}$ bei gewähltem Signifikanzniveau, dann wird H_0: keine Heteroskedastizität, verworfen.

2.3 Heteroskedastizität: Was kann getan werden?

Konsequenzen der Heteroskedastizität für das CLR-Modell mit OLS-Schätzung

- OLS-Schätzer sind weiterhin linear

- OLS-Schätzer sind erwartungstreu (unbiased)
- OLS-Schätzer sind konsistent und asymptotisch normalverteilt
- *Aber bezüglich Var(b): OLS-Schätzer haben keine Minimum-Varianz-Eigenschaft, sie sind nicht mehr effizient (no longer Best, but ‚LUE').*

Resultat: die gewöhnlichen Konfidenzintervalle und Hypothesentests, wie F- oder t-Test sind nicht zuverlässig mit der Gefahr, falsche Signifikanzschlüsse zu ziehen.

Was kann getan werden? Eine mögliche funktionale Beziehung zur Erklärung der Residuen kann direkt bei der verallgemeinerten Methode der kleinsten Quadrate (GLS) direkt mit korrekter Varianzmatrix der geschätzten Parameter berücksichtigt werden (vgl. Abschnitt IX).

White's robuster Schätzer

Wie kann eine mögliche Heteroskedastizität direkt bei der Schätzung aufgefangen werden? Mit White's robustem Schätzer.

Der White (1980) Schätzer der asymptotischen Kovarianzmatrix des OLS-Schätzers von b ist robust gegenüber den meisten Arten von Heteroskedastizität:

$$\mathrm{var}(b) = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\left(\sum_i e_i^2 \mathbf{x}_i \mathbf{x}_i'\right)\left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

Damit werden korrekte Schätzungen der Streuung von b – ob nun Heteroskedastizität vorliegt oder nicht – gegeben; entsprechende Tests (t-Test für die Koeffizienten b) sind dann wieder zutreffend.

3 AUTOKORRELATION

Annahme A3 fordert neben der Homoskedastizität auch die Freiheit von Autokorrelation. Insbesondere der Durbin-Watson-Test wird herangezogen, um zu überprüfen, inwieweit die Störgrößen der einzelnen Beobachtungen untereinander korrelieren.

Die Frage der Autokorrelation stellt sich vor allem bei Zeitreihenanalysen. Bei Querschnitten kann im Allgemeinen eher von der Unabhängigkeit der Beobachtungen ausgegangen werden.

Autokorrelation generell:

Korrelation einer Beobachtungsreihe mit sich selbst (AK)

AK1: bedeutet, dass die Beobachtungen mit den vorangegangenen Beobachtungen (t-1) korreliert

AK2: bedeutet, dass die Beobachtungen mit den Beobachtungen (t-2) korrelieren…

Verletzung der CLR-Annahme:

CLR- Annahme A3 (Kapitel VI.2): Störgrößen (Fehlerterme) ε sind unkorreliert.

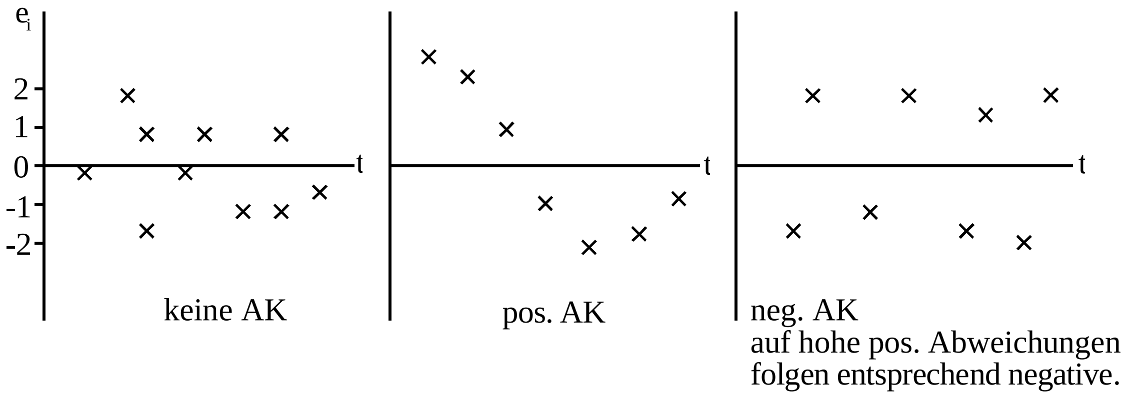
Keine Autokorrelation: $E(\varepsilon_i \varepsilon_j) = 0 \quad i \neq j \quad (i, j = 1, \ldots, n)$, alle Nebendiagonalelement der Varianz-Kovarianzmatrix der Fehler $E(\varepsilon - E(\varepsilon))(\varepsilon - E(\varepsilon))' = E(\varepsilon\varepsilon')$ sind Null

Autokorrelation liegt z.B. bei fehlerhafter Spezifikation vor, es fehlen wichtige Erklärungsvariablen

z.B. $C_i = \beta_0 + \beta_1 x_i + \beta_2 C_{i-1} + \varepsilon_i \rightarrow \varepsilon_i$ und ε_{i-1} sind über die Lagstruktur korreliert

Mögliche Folgen:

- OLS Schätzer von **b** ist nicht mehr BLUE
- Tests fehlerhaft
- Unterschätzung (bei positiver Autokorrelation)/Überschätzung (bei negativer Autokorrelation) der tatsächlichen Streuung der b_k



Angenommen ε_i und ε_{i-1} seien korreliert

$$\varepsilon_i = \rho \varepsilon_{i-1} + \eta_i$$

AK ersten Grades; Störterme η_i und η_{i-1} sind unkorreliert; (ρ = Rho)

ρ = 0: fehlende AK
ρ = 1: vollständige positive AK
ρ = -1: vollständige negative AK

ε ist nicht beobachtbar, lediglich $e_i = \hat{\varepsilon}_i$

$$e_i = \rho e_{i-1} + \eta_i$$

OLS-Schätzer von ρ ist dann

$$\hat{\rho} = \frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sum_{i=2}^{n} e_{i-1}^2} \quad \triangleq \quad \frac{\text{Kovarianz}}{\text{Varianz}} \qquad |e_i = e_i - \bar{e} = e_i - 0$$

Gewünscht: Fehlende Autokorrelation

$H_0: \rho = 0$

$H_1: \rho \neq 0$

Die zu bestätigende Hypothese ist also nicht die Alterrnativhypothese.

Ein Test mittels

$$\hat{\rho} = \frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sum_{i=2}^{n} e_{i-1}^2}$$

aus OLS-Schätzung ist nicht zulässig, da die Verteilung der e_i und damit auch die von $\hat{\rho}$ von den x-Werten abhängt.

3.1 Durbin-Watson-Test

Die ökonometrische Praxis verwendet den Durbin-Watson-Test zur Autokorrelationsanalyse

Prüfvariable:

$$\boxed{d_{beob} = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}}$$

Vor.: Regressionsgerade mit absolutem Glied

Verteilung unabhängig von x.

Da Verteilung von ρ unbekannt $\rightarrow$ Hilfsgröße d

Zusammenhang zwischen d und ρ

$$\sum_{i=2}^{n} (e_i - e_{i-1})^2 = \sum e_i^2 - 2\sum e_i e_{i-1} + \sum e_{i-1}^2 \qquad \left| \sum_{i=1}^{n} e_i^2 \approx \sum_{i=2}^{n} e_i^2 \right.$$

$$\approx 2\sum e_i^2 - 2\sum e_i e_{i-1}$$

$$= 2\left(\sum e_i^2 - \sum e_i e_{i-1}\right)$$

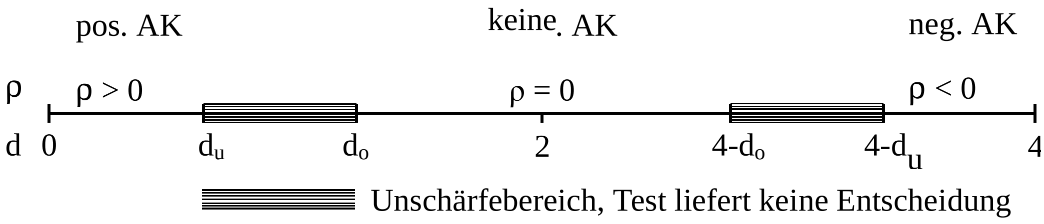
$$\Rightarrow d \approx \frac{\sum e_i^2 - \sum e_i e_{i-1}}{\sum e_i^2} = 2 \cdot (1 - \hat{\rho})$$

$\hat{\rho} = 0$: $d = 2$ keine Autokorrelation

$\hat{\rho} = 1$: $d = 0$ positive Autokorrelation

$\hat{\rho} = -1$: $d = 4$ negative Autokorrelation

Approximation durch d führt zu Unschärfebereichen.



H_0: $\rho = 0$ keine Autokorrelation

$H_{1:}$ $\rho > 0$ (einseitiger Test)

H_0: $\rho = 0$ keine Autokorrelation

$H_{1:}$ $\rho < 0$

Zusammenfassung Durbin-Watson-Test auf Autokorrelation

1. Signifikanzniveau von α festlegen
2. Berechne

$$d_{beob} = \frac{\sum_{i=2}^{n}\left(e_i - e_{i-1}\right)^2}{\sum_{i=1}^{n} e_i^2}$$

einer Regression mit Konstante.

3. Aus der Tabelle entnehmen
   d_u, d_o in Abhängigkeit von K = Anzahl der unabhängigen Variablen x
4. einseitiger Test

$$\left.\begin{array}{l} d_{beob} < 2 \quad \text{und} \quad d_{beob} > d_o \\ d_{beob} > 2 \quad \text{und} \quad d_{beob} < 4 - d_o \end{array}\right\} H_0\text{: Aufeinanderfolgende Beobachtungen von n sind}$$

voneinander unabhängig; H_0 kann nicht verworfen werden.

Testinteresse: Entdeckung der Autokorrelation der Residuen

Danach ist H_0: keine Autokorrelation

H_1: Autokorrelation

Beispiel: Durbin-Watson-Test (DW-Test) auf Autokorrelation

MacProd Produktionsfunktion (Kap. VI.3.2)

$Y = \gamma K^{\alpha} L^{\beta}$ Y = Output, K = Kapitalinput, L = Arbeit (Labor)

Nach der CLR-Schätzung ergeben sich die n=5 Residuen $e' = \left(-1;\ 0,5;\ 0,5;\ 0;\ 0\right)$

1. Signifikanzniveau
   $\alpha = 0{,}05$ gewählt

2. Berechnung von d_{beob}

$$d_{beob} = \frac{\sum_{i=2}^{5}\left(e_i - e_{i-1}\right)^2}{\sum_{i=1}^{5} e_i^2} = \frac{1,5^2 + 0^2 + (-0,5)^2 + 0^2}{1,5} = \frac{2,5}{1,5} = 1,667$$

3. d_u, d_o aus Tabelle

$$\left.\begin{array}{l} d_u = 0,467 \\ d_o = 1,896 \end{array}\right\rangle d_{beob} = 1,667$$

   K = 2 = Anzahl der Variablen ohne die Konstante

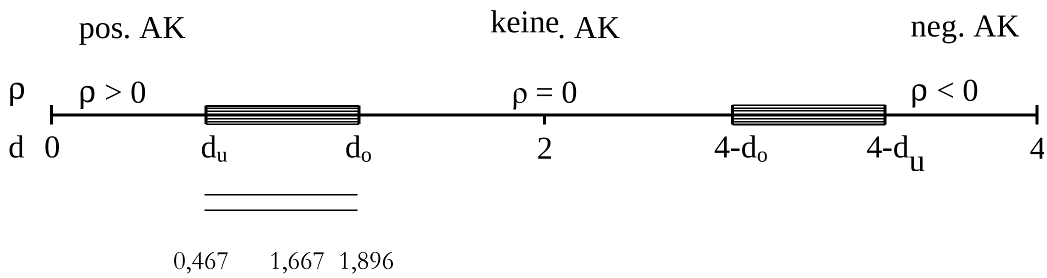
   Achtung: Tabellenwerte beziehen sich aber auf K inklusive der Konstante also hier K = 3

   n = 5 eigentlich zu wenig Beobachtungen

   Sei n = 7

4. H_0 ablehnen?

   d_{beob} liegt im Unschärfebereich (zur positiven Autokorrelation)



→ keine Aussage über AK bei $\alpha = 0,05$

Wäre d_{beob} = 1,92, dann könnte H_0: keine Autokorrelation nicht verworfen werden.

3.2 Autokorrelation: Was kann getan werden?

Die Annahmen des CLR-Modells mit dem Gauss-Markov Theorem verlangen Homoskedatiszität und keine Autokorrelation, also seriell unkorrelierte Störgrößen. Sind diese Annahmen verletzt, hat das Konsequenzen auf die Ergebnisse insgesamt und Testergebnisse insbeondere.

Konsequenzen der Autokorrelation:

- OLS-Schätzer sind nicht mehr BLUE.
- OLS β-Streuungen sind verzerrt und Teststatistiken (t-Test, F-Test, LM-Test …) sind nicht mehr gültig, sogar asymptotisch nicht mehr gültig..
- Die relevanten Diagonalelemente der Streuung von b_{OLS}, also $\sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$, unterschätzen (bei positiver Autokorrelation) die wahre Stichprobenvarianz von b_{OLS}. Die Standardfehler von β sind dann nach unten verzerrt. Bei negativer Autokorrelation wird die wahre Streuung von b überschätzt. Es wird eine größere Genauigkeit der Schätzung vorgespielt – t-Werte sind zu groß (bei positiver Autokorrelation) – als dies tatsächlich ist. Die Konfidenzintervalle sind

kleiner und eine Hypothese wird früher als zulässig abgelehnt. Weitere Diskussion z.B. Wooldridge (2009, S. 409ff), Hübler (1989, S. 176 f).

Was kann getan werden?

Bei statistisch gesicherter Autokorrelation kann Folgendes für die Spezifikation/Modellbildung getan werden:

- Inhaltlich: zu überprüfendes Modell überdenken.
- Neue unverzögerte und verzögerte exogene Variablen gegebenenfalls hinzufügen.
- Änderung des Funktionstyps

Also: Neue Spezifikation ist angebracht.

Für die Berücksichtigung der Autokorrelation für effiziente Schätzer im generalisierten Regressionsmodell → GLS Kap. IX.

4 MULTIKOLLINEARITÄT

Multikollinearität betrifft den Zusammenhang zwischen den erklärenden Variablen. Multikollinearität bedeutet, dass es eine lineare Abhängigkeit zwischen den Regressoren gibt und (A5) rg(**X**)=K+1 verletzt ist. Liegt Multikollinearität vor, dann gibt es entweder keine algebraische Lösung der Normalgleichungen (exakte Multikollinearität) oder als Folge einer erhöhten Parameterschätzer-Varianz werden die errechneten t-Werte für die Signifikanz der Parameter ‚verzerrt', sie werden zu klein (Quasi Multikollinearität).

4.1 Exakte Multikollinearität

Bei exakter Multikollinearität ist ein Regressor direkt linear abhängig von einem anderen Regressor (z.B. $ALTER2 = 2 * ALTER$; nicht aber $ALTER2 = ALTER^2$, was ein zulässiger nichtlinearer Zusammenhang wäre). Exakte Multikollinearität kann sich durch Zufall erklären oder Folge einer definitorischen Beziehung der erklärenden Variablen sein. Inhaltlich bedeutet dies, dass damit keine neue Erklärung gefunden werden kann.

Beispiel: Exakte Multikollinearität Konsumfunktion

Beispiel (Frohn 1980): Konsumfunktion

C_t=Konsum; Y_t=Einkommen; W_t= Lohn- und Gehaltseinkommen; G_t=Profite; Tr_t= Transfereinkommen.

$$C_t = \beta_1 + \beta_2 Y_t + \beta_3 W_t + \beta_4 G_t + \beta_5 Tr_t + u_t$$

Für die Regressoren gilt folgende Beziehung:

$$Y_t = W_t + G_t + Tr_t .$$

Damit ist Y_t (exakt) linear abhängig von den anderen Variablen.

Exakte Multikollinearität lässt sich meist durch sorgfältige Auswahl der Regressoren vermeiden.

Formal: Lineare Abhängigkeit zwischen den Regressoren führt dazu, dass der geforderte Rang der Matrix **X** zu klein ist (nicht mehr K+1 ist). Damit kann die Inverse von (**X'X**) nicht mehr gebildet werden, die Normalgleichungen können damit nicht mehr algebraisch gelöst werden.

Beispiel: Exakte Multikollinearität im K=2 Fall

$$x_{t2} = ax_{t1} \qquad K = 2$$

$$(A5)\;\; r(\mathbf{X}) = 1 + K$$
$$\text{hier}:\;\; r(\mathbf{X}) < K + 1$$
$$\Rightarrow \mathbf{X'X} \text{ ist singulär d.h. } |\mathbf{X'X}| = 0$$
$$\Rightarrow \;\text{es gibt keine } (\mathbf{X'X})^{-1}$$

Beweis für K=2:

$$(\mathbf{X'X}) = \begin{pmatrix} x_{t0} & \dots & x_{T0} \\ x_{t1} & \dots & x_{T1} \\ x_{t2} & \dots & x_{T2} \end{pmatrix}\begin{pmatrix} x_{t0} & x_{t1} & x_{t2} \\ \dots & \dots & \dots \\ x_{T0} & x_{T1} & x_{T2} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ x_{t1} & \dots & x_{T1} \\ x_{t2} & \dots & x_{T2} \end{pmatrix}\begin{pmatrix} 1 & x_{t1} & x_{t2} \\ \dots & \dots & \dots \\ 1 & x_{T1} & x_{T2} \end{pmatrix}$$

$$= \begin{bmatrix} T & \sum x_{t1} & \sum x_{t2} \\ \sum x_{t1} & \sum x^2_{t1} & \sum x_{t2}x_{t1} \\ \sum x_{t2} & \sum x_{t2}x_{t1} & \sum x^2_{t2} \end{bmatrix} = \begin{bmatrix} T & \sum x_1 & a\sum x_1 \\ \sum x_1 & \sum x^2_1 & a\sum x^2_1 \\ a\sum x_1 & a\sum x^2_1 & a^2\sum x^2_1 \end{bmatrix}$$

$$\begin{aligned} \det(\mathbf{X'X}) &= a^2T\sum x^2_1\sum x^2_1 + a^2\sum x_1\sum x_1\sum x^2_1 + a^2\sum x_1\sum x_1\sum x^2_1 \\ &-(a^2T\sum x^2_1\sum x^2_1 + a^2\sum x_1\sum x_1\sum x^2_1 + a^2\sum x_1\sum x_1\sum x^2_1) = 0 \end{aligned} \qquad \text{q.e.d.}$$

Die Determinante ist also Null; eine Berechnung der Inversen $(\mathbf{X'X})^{-1}$ ist nicht möglich (keine Lösung des Problems).

4.2 Quasi Multikollinearität

Quasi-Multikollinearität liegt vor, wenn Regressoren stark korreliert sind. In diesem Fall kann zwar die Inverse von (**X'X**) gebildet werden, die aber sehr groß werden kann.

Für den k+1=2 Fall wäre die dazu notwendige Determinante $|\mathbf{X'X}|$ sehr klein $\rightarrow (\mathbf{X'X})^{-1}$ sehr groß von (**X'X**) aus

$$(\mathbf{X'X})^{-1} = \frac{1}{|\mathbf{X'X}|}\text{adj.}(\mathbf{X'X})$$

$$|\mathbf{X'X}| \text{ sehr klein} \Rightarrow (\mathbf{X'X})^{-1} \text{ sehr groß}$$

Das hat zur Folge, dass mit

$$S_{bb} = s^2 \left(\mathbf{X'X}\right)^{-1} \qquad \Rightarrow \text{ auch } S_{bb} \text{ sehr groß wird}$$

$$t_b^* > t_b = \frac{b_K}{s_{bk}} \rightarrow \frac{b_K}{\rightarrow \infty} \rightarrow 0 \quad t_{b_K} - \text{Werte werden schließlich sehr klein}$$

Mit zu kleinen t-Werten der Parameter werden somit bei Quasi-Multikollinearität signifikante Parameter unter Umständen nicht erkannt.

4.3 Wie lässt sich Quasi-Multikollinearität feststellen?

Korrelationsmatrix

Hohe Korrelationskoeffizienten zwischen zwei Variablen deuten in der Korrelationsmatrix auf einen starken Zusammenhang hin.

Hilfsregressionen

Eine Möglichkeit Multikollinearität festzustellen ist, Hilfsregressionen zu berechnen, in denen Regressoren durch andere Regressoren erklärt werden. Das Bestimmtheitsmaß dieser Hilfsregressionen kann Auskunft über die lineare Abhängigkeit der Regressoren geben.

Beispiel: Multikollinearität und Hilfsregression

Im folgenden Ansatz wird Multikollinearität vermutet:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \varepsilon_t$$

Als Hilfsregressionen werden nun folgende Ansätze geschätzt:

$$x_{1t} = \gamma + \mu_2 x_{2t} + \eta_t$$

$$x_{1t} = \gamma + \mu_3 x_{3t} + \eta_t \quad \text{(und weiter für alle K Regressoren)}$$

$$x_{1t} = \gamma + \mu_2 x_{2t} + \mu_3 x_{3t} + \eta_t$$

$$x_{2t} = \gamma + \mu_1 x_{1t} + \mu_3 x_{3t} + \eta_t$$

$$x_{3t} = \gamma + \mu_1 x_{1t} + \mu_2 x_{2t} + \eta_t$$

Nun betrachtet man das Bestimmtheitsmaß dieser Hilfsregressionen. Ergibt sich für die Hilfsregressionen ein niedriges Bestimmheitsmaß, so liegt mit hoher Wahrscheinlichkeit keine Multikollinearität vor. Nehmen die sich ergebenden Korrelationsmaße $R^2_{1\bullet 2}, R^2_{1\bullet 3}, R^2_{2\bullet 3}, R^2_{1\bullet 23}, R^2_{2\bullet 13}, R^2_{3\bullet 12}$ (teilweise) hohe Werte an, so deutet das auf Multikollinearität (bestimmter Regressoren) hin.

Variance Inflation Factors (VIFs), Studenmund 2001

Ein ähnlicher Ansatz um Multikollinearität zu erkennen sind Variance Inflation Factors (VIFs). Auch hier werden Hilfsregressionen geschätzt und das daraus resultierende Bestimmtheitsmaß als Maß für Multikollinearität verwendet. Um folgenden Ansatz auf das Vorliegen von Multikollinearität zu untersuchen, wird für jeden der K Regressoren ein zugehöriger VIF berechnet.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_K x_K + \varepsilon$$

Dabei ist folgendermaßen vorzugehen:

1. OLS-Schätzung der Ansätze
   $x_1 = \alpha_1 + \alpha_2 x_2 + \alpha_3 x_3 + ... + \alpha_k x_k + \mu$ (μ ist Fehlerterm)
   analog für jedes x_k (k=1,…,K)
2. Berechnung der VIFs
   Aus den K Bestimmtheitsmaßen R_k^2 der K OLS-Schätzungen werden K VIFs berechnet
   $$\mathrm{VIF}(\hat{\beta}_k) = \frac{1}{(1-R_k^2)}$$
3. Beurteilung der Multikollinearität anhand der VIF$(\hat{\beta}_i)$

Je größer der VIF-Wert ist, umso stärker ist die vorliegende Multikollinearität. Das bedeutet auch je höher der VIF einer Variablen, umso höher ist die Varianz des geschätzten Koeffizienten der Variablen.

Wertebereich VIF: Für perfekte Multikollinearität nimmt R_k^2 den Wert 1 an. In diesem Fall ist der Term nicht definiert. Liegt R_k^2 sehr Nahe an 1, so geht der VIF gegen Unendlich ($\lim_{R_k^2 \to 1} \mathrm{VIF} = +\infty$). Nimmt R_i^2 den Wert Null an, d.h. es liegt keine Multikollinearität vor, ergibt sich ein VIF von 1.

Das bedeutet: je höher der VIF-Wert, umso stärker ausgeprägt ist die vorliegende Multikollinearität.

Faustregel: Nimmt VIF Werte größer 5 an, so liegt erhebliche Multikollinearität vor. Probleme: Es existieren keine festen Kriterien zur Entscheidung ob Multikollinearität vorliegt. Des Weiteren ist es möglich, dass starke Multikollinearität trotz relativ geringer VIF-Werte vorliegt.

4.4 Was kann bei Multikollinearität getan werden?

Es gibt mehrere Möglichkeiten, mit Multikollinearität umzugehen:

Weglassen von Variablen:

Variablen mit niedrigen t-Werten werden als insignifikant aus dem Ansatz gestrichen. Problem: Geringe t-Werte können auf Multikollinearität zurückzuführen sein. Werden aber relevante Variablen gestrichen, so führt das zu verzerrten Parameterschätzern. Liegen also Anhaltspunkte für Multikollinearität von Variablen mit niedrigem t-Wert vor, so sollten die betroffenen Variablen nicht ausgelassen werden.

Differenzenbildung:

Multikollinearität lässt sich häufig durch Bildung der ersten Differenzen eliminieren. Schätzung des Ansatzes:

$$\Delta y_t = \beta_1 \Delta x_{1t} + \beta_2 \Delta x_{2t} + \ldots + \beta_k \Delta x_{kt} + \Delta u_t$$
$$\text{mit } \Delta y_t = y_t - y_{t-1}, \Delta x_{1t} = x_{1t} - x_{1t-1} \ldots \Delta x_{kt} = x_{kt} - x_{kt-1};\ \Delta u_t = u_t - u_{t-1}$$

Problem: Δu_t und Δu_{t-1} sind über u_{t-1} korreliert. Schätzung der Parametervarianz und t-Werte sind verzerrt.

Verwendung zusätzlicher Daten:

Problem: Weitere Datenerhebung unter Umständen nicht möglich oder zu teuer. Des Weiteren ist zu erwarten, dass der Informationsgehalt zusätzlicher Datensätze unter demselben Multikollinearitätsproblem leidet.

Verwendung externer Informationen:

Restringierte-Kleinste-Quadrate Schätzung, RKQ:

Liegen für bestimmte Parameter des Modells - aus externer Quelle - Informationen über einen linearen Zusammenhang vor, so lassen sich diese für die Schätzung des Modells nutzen.

Für den Ansatz $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + e_t$ sei bekannt, dass $\beta_2 = 0$ und $\beta_3 = -2\beta_1$. Integriert man dieses Wissen in den ursprünglichen Ansatz so ergibt sich

$$\begin{aligned} y_t &= \beta_0 + \beta_1 x_{1t} + 0x_{2t} - 2\beta_1 x_{3t} + \beta_4 x_{4t} + e_t \\ &= \beta_0 + \beta_1 x^*_{1t} + \beta_4 x_{4t} + e_t \end{aligned}$$
$$(x^*_{1t} = x_{1t} - 2x_{3t})$$

Auf diese Weise wird die Zahl der zu schätzenden Parameter bei unverändertem n (T) reduziert. Die Zahl der Freiheitsgrade erhöht sich und die verbleibenden Parameter können mit geringerer Varianz geschätzt werden. Der Schätzer für β_3 lässt sich dann aus der Beziehung $\beta_3 = -2\beta_1$ ermitteln.

Problem: Die RKQ-Schätzung ist nur dann unverzerrt, wenn die verwendeten Informationen exakt stimmen. Weicht die unterstellte Beziehung von der Realität ab, so resultieren verzerrte Schätzer.

Neue Spezifikation?

Eine mögliche Lösung wäre den Ansatz neu zu spezifizieren. Dies macht dann Sinn, wenn der neue Ansatz nicht mehr von Multikollinearität betroffen ist. Zu bedenken ist dabei, dass der Ausschluss von relevanten Variablen aus dem Modell zu einer Verzerrung der Schätzergebnisse führen kann (siehe ‚omitted variable bias'). Da man sich mit allen Reaktionen auf Multikollinearität neue Probleme einhandelt, ist es in vielen Fällen am besten, den Ansatz trotz Multikollinearität beizubehalten. Dies gilt insbesondere dann, wenn die Parameterschätzer trotz vorliegender Multikollinearität signifikant sind und eine geringe Varianz aufweisen. Die Varianz der Parameterschätzer lässt sich folgendermaßen darstellen:

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{s_{11}(1 - R^2_{12})}$$

Hier lässt sich leicht erkennen durch welche Faktoren die Varianz der Parameterschätzer bestimmt wird. Sie hängt vom Ausmaß der Korrelation der Regressoren, der Variation innerhalb der Regressoren und der Varianz des Fehlerterms ab. Eine hohe Varianz für $\hat{\beta}_1$ ergibt sich also wenn

- σ^2 hoch ist
- s_{11} klein ist
- R_{12}^2 hoch ist

Ein hohes R_{12}^2 kann somit durch eine hohe Variation der Regressorenwerte und ein niedriges σ^2 kompensiert werden.

5 NORMALVERTEILUNG DER STÖRGRÖßEN

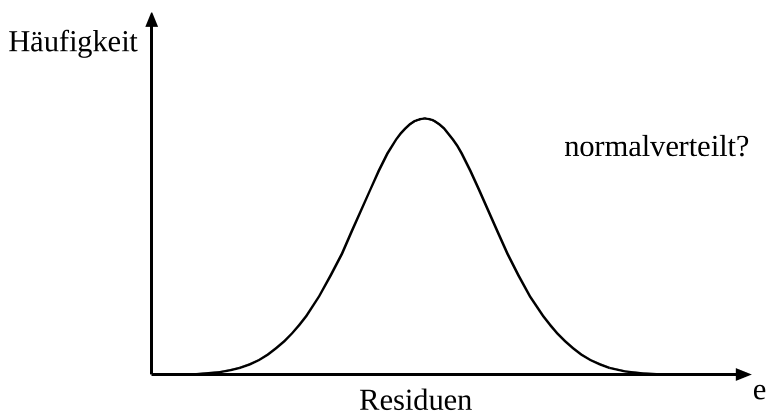
Die letzte Annahme im Referenzmodell besagte, dass die stochastischen Störterme einer Normalverteilung unterliegen. Zwar bleiben die OLS Schätzungen auch bei Nicht-Normalität der Residuen erwartungstreu, Aussagen zur Signifikanz sind allerdings nicht mehr zutreffend. Zur Überprüfung dieser Annahme zieht man den Jarque-Bera-Test oder einen χ^2 Test zu Rate.

5.1 Normalität: Histogramm, Normal Probability Plot

Zur Diagnose werden die Residuen zunächst graphisch auf Normalität untersucht - mittels eines Histogramms sowie dem Normal Probability Plot.

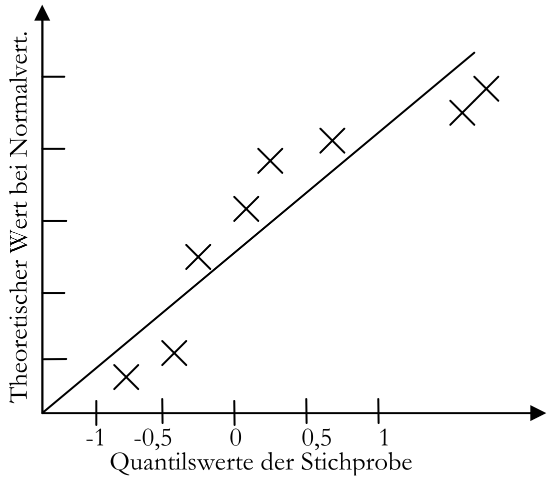
Histogramm

Graphische Inspektion durch ein Histogramm der Residuen $e_i \left(i = 1, 2, \ldots, n\right)$



Normal Probability Plot

x-Achse: Quantilswerte der vorliegenden Stichproben e_i
y-Achse: erwartete Quantilswerte von e_i wenn diese normalverteilt wären



Ist e_i exakt standardnormalverteilt $N(0, \sigma^2 = 1)$, dann liegen alle Punkte auf einer Geraden.

Test: Anderson-Darling-Normalitätstest: A^2-Statistik

5.2 Jarque-Bera-Test (JB) auf Normalverteilung

Teststatistik:

$$JB = \frac{n}{6}\left[SKEW^2 + \frac{(KURT-3)^2}{4}\right]$$

wobei

SKEW = Skewness: Schiefemaß/ Maß für die Asymmetrie

$$SKEW = \left(\frac{1}{n}\sum_i (e_i - \overline{e})^3\right) \Big/ s^3$$

KURT = Kurtosis: Wölbungsmaß

$$KURT = \left(\frac{1}{n}\sum_i (e_i - \overline{e})^4\right) \Big/ s^4 \text{ und } s = \sqrt{1/n \sum_i (e_i - \overline{e})^2}$$

JB ist asymptotisch χ^2-verteilt mit $\nu = 2$ Freiheitsgraden

H_0: ε_i sind normalverteilt

H_1: ε_i sind nicht normalverteilt

H_0 wird verworfen, wenn $JB > \chi^2(\upsilon)$ bei einem ausgewählten Signifikanzniveau α ist.

5.3 χ^2-Verteilungstest auf Normalität

Gruppierte Residuen: Beobachtete Gruppenhäufigkeit (n_j) wird mit theoretischer Gruppenhäufigkeit (np_j) verglichen. Dabei ist p_i aus den Gruppenintervallen der Normalverteilung heranzuziehen (vgl. Merz 2010a).

Teststatistik:

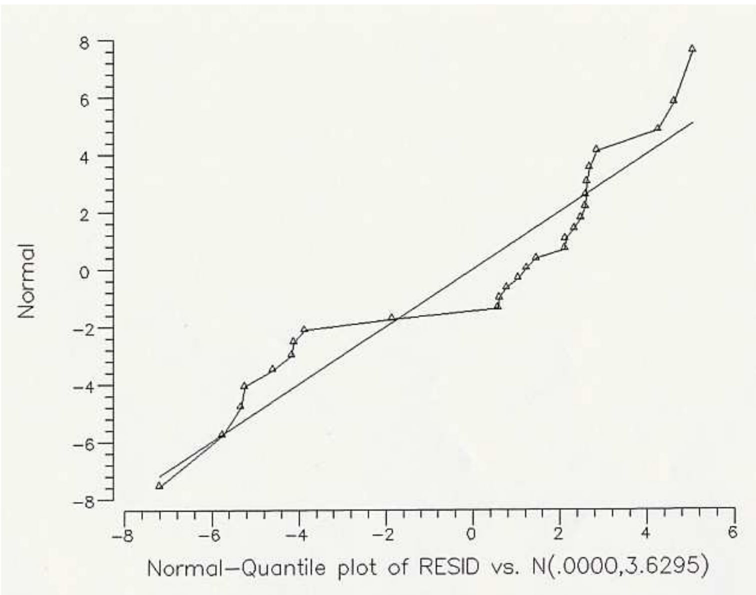
$$\chi^2 = \sum_{i=1}^{K} \frac{\left(n_i - np_i\right)^2}{np_i} \qquad \text{mit } \nu = K - 1 - m \text{ Freiheitsgraden}$$

(m = 2 bei Normalverteilung)

H_0: ε_i sind normalverteilt, wird verworfen wenn $\chi^2 > \chi^2_{krit}(\nu)$ bei gewähltem Signifikanzniveau von α ist.

Beispiel: Normalitätstest: US-Gasoline Sales

```
Descriptives, Histogramm, Normal Probability Plot und Jarque-Bera Test
ET Residuen: RESID (=e)


Descriptive Statistics
Variable     Mean    Std. Dev.      Skew.   Kurt.   Minimum   Maximum Cases
---------------------------------------------------------------------------
RESID      .11620E-05  4.7499       -.351   1.834   -8.816    7.758      27

Histogram for RESID    computed using   27 observations
Individual data      Mean=     .000, std.dev.=    4.750

                                         Frequency         Cumulative
      Lower Limit    Upper Limit    Total   Relative   Total   Relative
   0       -8.816         -7.158        3      .1111       3      .1111
   1       -7.158         -5.501        2      .0741       5      .1852
   2       -5.501         -3.843        1      .0370       6      .2222
   3       -3.843         -2.186        4      .1481      10      .3704
   4       -2.186          -.529        0      .0000      10      .3704
   5        -.529          1.129        3      .1111      13      .4815
   6        1.129          2.786        5      .1852      18      .6667
   7        2.786          4.443        5      .1852      23      .8519
   8        4.443          6.101        2      .0741      25      .9259
   9        6.101          7.758        2      .0741      27     1.0000
```



Bei vollkommener Normalverteilung der Residuen müssten alle Punkte auf der 45 Grad-Linie liegen.

```
Jarque-Bera Test: Chi-squared Test
Variable  Chi-squared  P-Value  Kolmogorov-Smirnov  P-Value
RESID          2.0822   .35306               .5000   .00000
```

H_0: ε_i sind normalverteilt
H_1: ε_i sind nicht normalverteilt

Da JB=2.0822 < Chi2(2)=5.99147 (Tabellenwert), bzw. der p-value=.35306 > α=.05 kann H_0 nicht verworfen werden: Die Verteilung der Residuen weicht also nicht signifikant von einer Normalverteilung ab. Allerdings weist der Kolmogorov-Smirnov χ^2-Test auf nicht normalverteilte Fehlerterme hin.

5.4 Fehlende Normalverteilung: Was kann getan werden?

Konsequenzen fehlender Normalverteilung

- OLS-Schätzungen für β bleiben unverzerrt, also erwartungstreu, da $\sum_{i=1}^{n} \varepsilon_i = 0$.
- Inferenzstatistische Aussagen über den geschätzten Koeffizientenvektor lassen sich nicht mehr treffen, da die Verteilungseigenschaften der Schätzung nicht mehr bekannt sind.

Was kann getan werden?

Überprüfen der Spezifikation und funktionalen Form z.B. mit Ramsay's RESET Test und dann Neuspezifikation.

6 Regression und einflußreiche Beobachtungen (Ausreißer?)

Die Daten sind die Basis jeder empirischen Regression. Vor jeder Regressionsanalyse ist es daher besonders wichtig, die Daten zu inspizieren und auf etwaige Ausreißer zu untersuchen. Allerdings, ob ein Wert wirklich ein Ausreißer ist, muss letztlich inhaltlich geklärt werden. Um solche einflußreichen Beobachtungen, die einen besonderen Effekt auf das Regressionsergenis haben, erkennen zu können, sind mittlerweile mehrere Ansätze der Regression Diagnostic vorhanden, die im Folgenden skizziert werden.

Neben rein deskriptiven Indikatoren (Begrenzung auf das zweifache der jeweiligen Standardabweichung) und Illustrationen (wie z.B. Box and Whisker Plot) sollen zwei Verfahren genannt werden, die bspw. auch in Stata zur Verfügung stehen.

Literatur: Kohler und Kreuter 2008, Kapitel 8.3 Regressionsdiagnostik; Atkinson und Riani 2000, Kapitel 2 Regression and Forward Search

6.1 Einfluss systematisch ausgelassener Beobachtungen: DFBETA

Vorgehensweise der Analyse: Nach multipler Regressionsrechnung wird eine Beobachtung herausgenommen und erneut eine Regression ohne diese Beobachtung durchgeführt. Dies geschieht systematisch nacheinander mit allen Beobachtungen. Danach wird für jeden geschätzten Koeffizienten die Differenzen mit und ohne heraus genommener Beobachtung getestet mit

$$DFBETA_{ik} = \frac{b_k - b_{k(i)}}{s_{e(i)} / \sqrt{RSS_k}}$$

Wobei $b_{k(i)}$ der Koeffizient der Schätzung ohne die ite Beobachtung ist. Alle Beobachtungen, die ein

$$DFBETA_{ik} > 2/\text{Wurzel}(n)$$

haben, erfordern besondere Aufmerksamkeit, da besonders einflußreich.

6.2 Cook's distance: D

Dieses Maß schätzt den Einfluss einer Beobachtung simultan für alle Regressionskoeffizienten. Dabei setzt sich der Einfluss einer Beobachtung aus einer außergewöhnlichen Kombination der unabhängigen Variablen und einem ungewöhlichen Wert der abhängigen Variable zusammen.

$$D_i = \frac{e_i^2 h_i}{(k+1)s^2(1-h_i)^2} = \underbrace{\frac{h_i}{1-h_i}}_{\substack{\text{außergewöhnliche}\\ \text{Kombination der}\\ \text{unabhängigen Variablen}\\ (\text{Leverage})}} \cdot \underbrace{\frac{r_i^2}{k+1}}_{\substack{\text{ungewöhlicher Wert}\\ \text{der abhängigen Variable}\\ (\text{Diskrepanz})}}$$

wobei $h_i = x_i(X'X)^{-1}x_i$ und $r_i = \frac{e_i}{s\sqrt{1-h_i}}$ auch *studentized residuals* genannt werden.

Beobachtungen mit einem Wert

$$D_i > 4/n$$

erfordern besondere Aufmerksamkeit.

6.3 Einflußreiche Beobachtungen (Ausreißer?): Was kann getan werden?

Konsequenzen einflußreicher Beobachtungen (Ausreißer?)

Die eigentliche Wirkung der Einflußgröße, die durch alle anderen Variablenwerte getragen wird, wird durch den veränderten geschätzten Regressionskoeffizienten „verschleiert".

Was kann getan werden?

Wenn einflußreiche Variablen gefunden werden, können Sie – falls Messfehler – aus der Analyse ausgeschlossen werden. Sollte es sich nicht um Mess- oder Editierungsfehler handeln, sollten die Auswirkungen mit zwei Regressionen (mit und ohne entsprechende Variablenwerte) analysiert und/oder weitere Variablen zu einer Neuspezifikation des Modells verwendet werden.

X BEISPIELE ANALYSE DER CLR-ANNAHMEN

Mit den folgenden zwei Beispielen, US-Gasoline Sales (Makro) und Return on Human Capital (Mikro), sollen die Diagnostikmöglichkeiten der zwei Programmpakete LIMDEP und Stata hnsichtlich der CLR-Annahmen illustriert und die Annahmen jeweils des klassischen linearen Regressionsmodell untersucht und diskutiert werden.

1 US-GASOLINE SALES

ET: Gasoline Sales in the US Market

Gasoline sales in the US market from 1960 – 1982.

```
      G     = gasoline sales, in billions of gallons
      PG    = price index for gasoline
      Y     = per capita income
      PNC   = price index for new cars
      PUC   = price index for used cars
      PPT   = price index for public transportation
      PD    = price index for consumer durables
      PN    = price index for nondurables
      PS    = price index for services
```

DATA LISTING (Current sample)

Year	G	PG	Y	PNC	PUC	PPT	PD	PN	PS
1960	129.7	.925	6036	1.045	.836	.810	.444	.331	.302
1961	131.3	.914	6113	1.045	.869	.846	.448	.335	.307
1962	137.1	.919	6271	1.041	.948	.874	.457	.338	.314
1963	141.6	.918	6378	1.035	.960	.885	.463	.343	.320
1964	148.8	.914	6727	1.032	1.001	.901	.470	.347	.325
1965	155.9	.949	7027	1.009	.994	.919	.471	.353	.332
1966	164.9	.970	7280	.991	.970	.952	.475	.366	.342
1967	171.0	1.000	7513	1.000	1.000	1.000	.483	.375	.353
1968	183.4	1.014	7728	1.028	1.028	1.046	.501	.390	.368
1969	195.8	1.047	7891	1.044	1.031	1.127	.514	.409	.386
1970	207.4	1.056	8134	1.076	1.043	1.285	.527	.427	.407
1971	218.3	1.063	8322	1.120	1.102	1.377	.547	.442	.431
1972	226.8	1.076	8562	1.110	1.105	1.434	.555	.458	.451
1973	237.9	1.181	9042	1.111	1.176	1.448	.566	.497	.474
1974	225.8	1.599	8867	1.175	1.226	1.480	.604	.572	.513
1975	232.4	1.708	8944	1.276	1.464	1.586	.659	.615	.556
1976	241.7	1.779	9175	1.357	1.679	1.742	.695	.638	.598
1977	249.2	1.882	9381	1.429	1.828	1.824	.727	.671	.648
1978	261.3	1.963	9735	1.538	1.865	1.878	.769	.719	.698
1979	248.9	2.656	9829	1.660	2.010	2.003	.821	.800	.756
1980	226.8	3.691	9722	1.793	2.081	2.516	.892	.894	.839
1981	225.6	4.109	9769	1.902	2.569	3.120	.957	.969	.926
1982	228.8	3.894	9725	1.976	2.964	3.460	1.000	1.000	1.000
1983	239.6	3.764	9930	2.026	3.297	3.626	1.041	1.021	1.062
1984	244.7	3.707	10421	2.085	3.757	3.852	1.038	1.050	1.117
1985	245.8	3.738	10563	2.152	3.797	4.028	1.045	1.075	1.173
1986	269.4	2.921	10780	2.240	3.632	4.264	1.053	1.069	1.224

Descriptive Statistics

Variable	Mean	Std. Dev.	Skew.	Kurt.	Minimum	Maximum	Cases
Y	8513.5	1455.6	-.272	1.800	6036.	.1078E+05	27

```
PG          1.9021      1.1679      .814   1.963  .9140    4.109       27
PUC         1.7123      .97474     1.080   2.691  .8360    3.797       27
PD          .67489      .22317      .619   1.763  .4440    1.053       27
PS          .60081      .30261      .791   2.158  .3020    1.224       27

Correlation Matrix

                  1-Y          2-PG          3-PUC         4-PD        5-PS
   1-Y         1.000000
   2-PG         .812162     1.000000
   3-PUC        .831816      .910825     1.000000
   4-PD         .898680      .963304      .969624     1.000000
   5-PS         .891549      .937920      .985491      .992958   1.000000

                             Order Statistics for Variables
Percentile     Y            PG            PUC           PD            PS
Min.           6036.0       .91400        .83600        .44400        .30200
10th           6271.0       .91800        .94800        .45700        .31400
20th           7027.0       .94900        .99400        .47100        .33200
25th           7280.0       .97000        1.0000        .47500        .34200
30th           7728.0       1.0140        1.0280        .50100        .36800
40th           8134.0       1.0560        1.0430        .52700        .40700
Med.           8867.0       1.1810        1.1760        .56600        .47400
60th           9175.0       1.7790        1.6790        .69500        .59800
70th           9722.0       1.9630        1.8650        .76900        .69800
75th           9735.0       2.9210        2.0810        .89200        .83900
80th           9769.0       3.6910        2.5690        .95700        .92600
90th           10421.       3.7640        3.6320        1.0410        1.1170
Max.           10780.       4.1090        3.7970        1.0530        1.2240

Partition of range Min to Max
Range of X     Y            PG            PUC           PD            PS
Minimum        6036.0       .91400        .83600        .44400        .30200
1st.Qrtl       7222.0       1.7128        1.5763        .59625        .53250
Midpoint       8408.0       2.5115        2.3165        .74850        .76300
3rd.Qrtl       9594.0       3.3103        3.0567        .90075        .99350
Maximum        10780.       4.1090        3.7970        1.0530        1.2240


1. Matrix -> XSX=XDOT(X)

      <<<< XSX       >>>>   COLUMN
                1          2          3          4          5          6
 ROW   1   27.0000    229865.    51.3570    46.2320    18.2220    16.2220
 ROW   2   229865.  .201205E+10  473127.    424283.    162724.    148317.
 ROW   3   51.3570    473127.    133.151    114.898    41.1883    39.4746
 ROW   4   46.2320    424283.    114.898    103.866    36.6855    35.3348
 ROW   5   18.2220    162724.    41.1883    36.6855    13.5928    12.6916
 ROW   6   16.2220    148317.    39.4746    35.3348    12.6916    12.1274

     1. Matrix -> XSY=XDOT(X,G)

      <<<< XSY       >>>>   COLUMN
                1
 ROW   1   5589.90
 ROW   2   .491790E+08
 ROW   3   11478.6
 ROW   4   10317.8
 ROW   5   3965.61
 ROW   6   3616.81

     1. Matrix -> XSXINV=GINV(XSX)
```

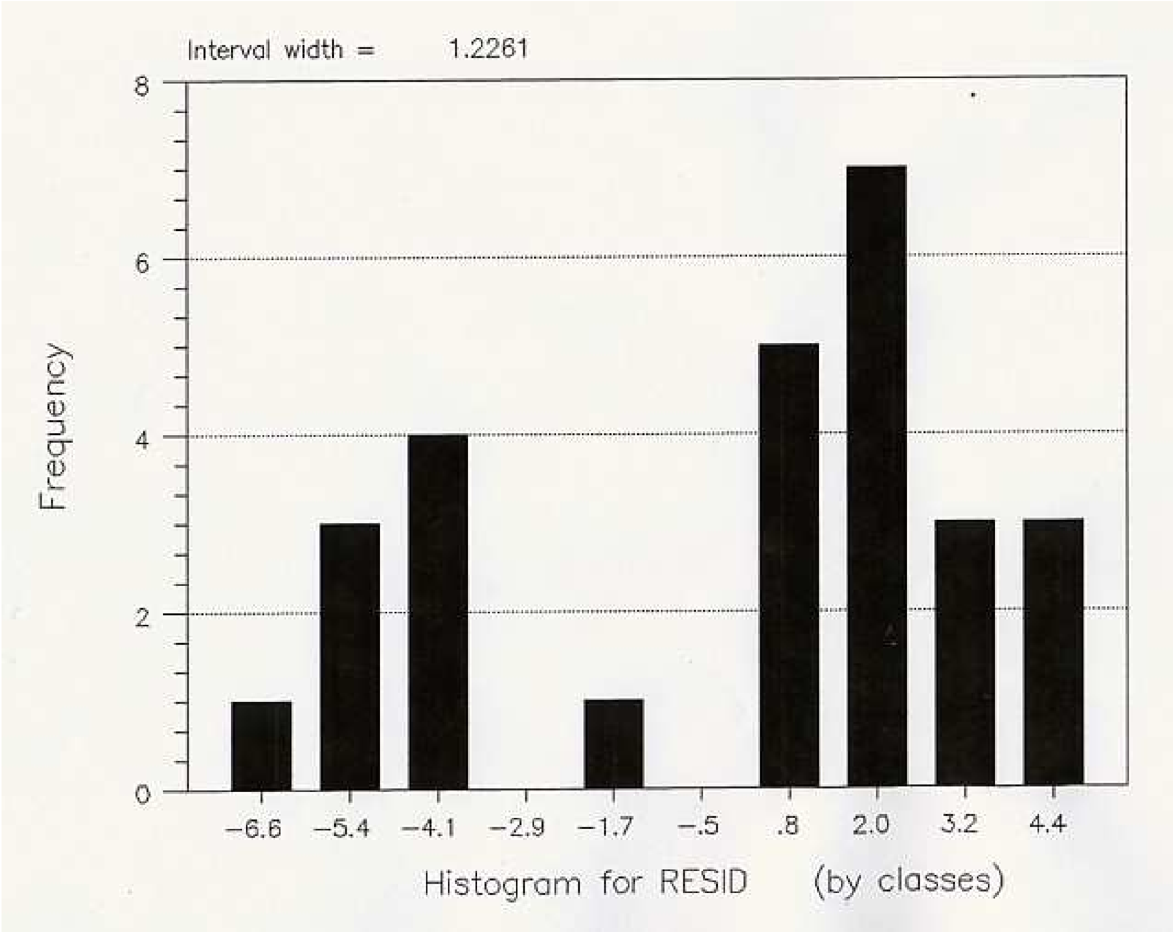
```
     <<<< XSXINV   >>>>   COLUMN
           1         2          3          4          5           6
 ROW 1  10.7477  -.363085E-03  1.58149    -2.26064     -35.8107    28.9797
 ROW 2-.36308E-03 .20683E-06 .24585E-03 .42368E-03  -.30446E-02 -.89231E-03
 ROW 3  1.58149   .245853E-03  .913961     .282591     -11.6497    3.27109
 ROW 4 -2.26064   .423683E-03  .282591     2.69674      1.48108   -12.4849
 ROW 5 -35.8107  -.304468E-02 -11.6497     1.48108      217.085   -108.443
 ROW 6  28.9797  -.892317E-03  3.27109    -12.4849     -108.443    111.448


     1. Matrix -> BOLS=XSXINV|XSY

     <<<< BOLS     >>>>   COLUMN
              1
 ROW   1  -146.289
 ROW   2   .346198E-01
 ROW   3  -29.1291
 ROW   4  -14.3086
 ROW   5   305.885
 ROW   6  -113.157

===========================================================================
Ordinary    Least Squares
Dependent Variable       G            Number of Observations           27
Mean of Dep. Variable     207.0333     Std. Dev. of Dep. Var.    43.798927
Std. Error of Regr.        43.7989     Sum of Squared Residuals    49877.0
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant   207.033       8.429      24.562   .00000


===========================================================================
Ordinary    Least Squares
Dependent Variable       G            Number of Observations           27
Mean of Dep. Variable     207.0333     Std. Dev. of Dep. Var.    43.798927
Durbin Watson statistic      .6367     Estimated Autocorrelation     .68164
Std. Error of Regr.         4.0386     Sum of Squared Residuals   342.514
Total Variation          49877.        Regression Variation       49534.
Regression degrees of freedom =  5     Residual degrees of freedom =     21
R - squared                 .99313     Adjusted R - squared          .99150
F( 5,  21)                607.4059     Prob. Value for F             .00000
Akaike Information         2.98492     Amemiya Prediction          19.93466
LM test for heteroskedasticity         Chi-squared[ 5] =    2.872  P= .7197
LM test for autocorrelation            Chi-squared[ 2] =   16.245  P= .0000
===========================================================================
Variable  Coefficient  Std. Error  t-ratio Prob|t|>x Mean of X Std.Dev.of X
---------------------------------------------------------------------------
Constant  -146.273       13.24     -11.048   .00000
Y           .346210E-01   .1837E-02  18.850   .00000   8513.51852 1455.62903
PG        -29.1239       3.861      -7.544   .00000      1.90211    1.16791
PUC       -14.3108       6.632      -2.158   .04267      1.71230     .97474
PD         305.844       59.50       5.140   .00004       .67489     .22317
PS        -113.096       42.63      -2.653   .01489       .60081     .30261

Estimated V-C Matrix for Parameters

                1-ONE        2-Y          3-PG         4-PUC        5-PD
   1-ONE      175.282
   2-Y       -.592271E-02 .337338E-05
   3-PG       25.7903     .400945E-02  14.9053
   4-PUC     -36.8706     .690996E-02  4.60857  43.9831
```

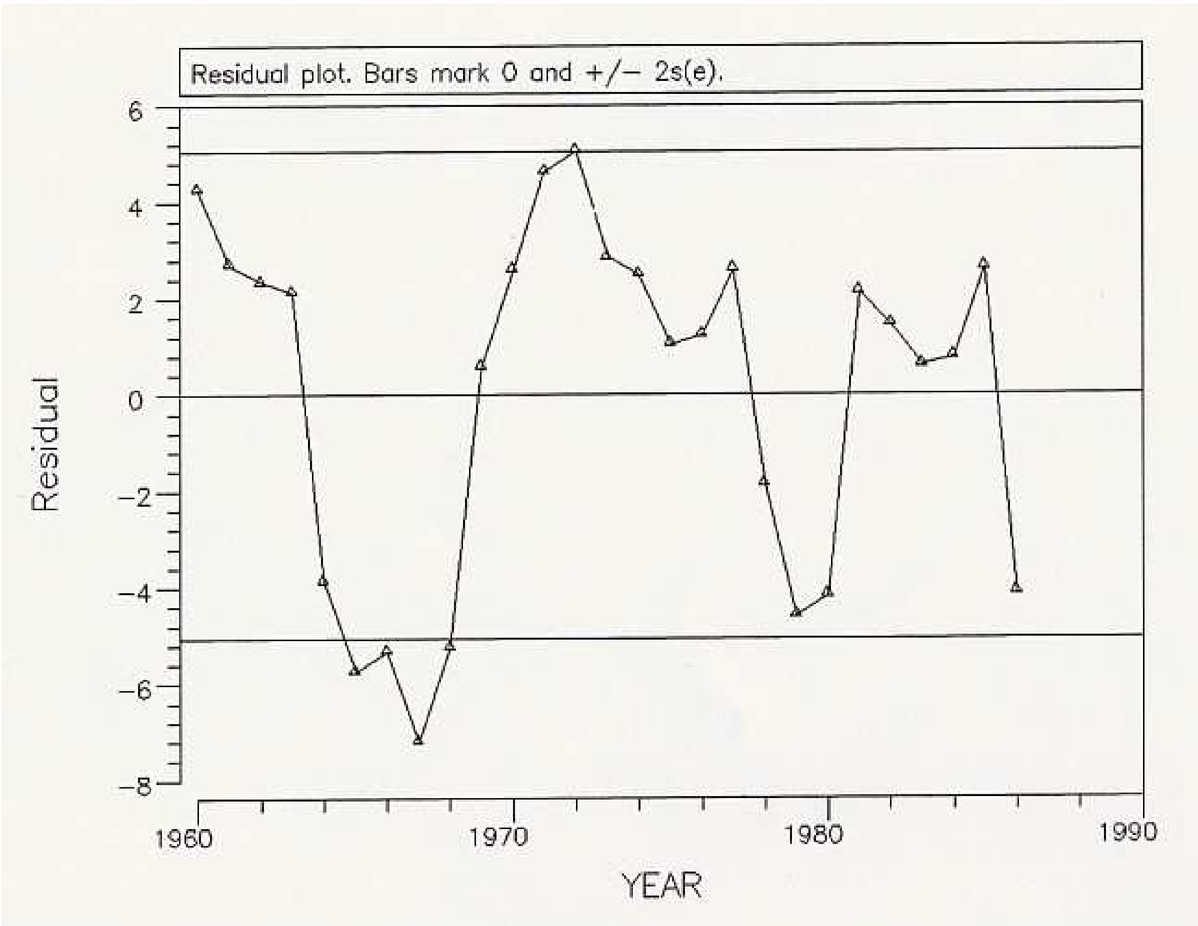
```
   5-PD      -584.000    -.496526E-01 -189.981   24.1609  3540.24
   6-PS       472.622    -.145551E-01  53.3419  -203.626 -1768.52 1817.62

P r e d i c t e d   V a l u e s
(* => not in estimating sample)
Observation Observed Y   Predicted Y   Residual   95% Forecast Interval
  1         129.70        125.44         4.2644      116.103    134.768
  2         131.30        128.61         2.6926      119.368    137.847
  3         137.10        134.76         2.3377      125.688    143.837
  4         141.60        139.48         2.1194      130.451    148.510
  5         148.80        152.67        -3.8685      143.774    161.563
  6         155.90        161.65        -5.7499      152.853    170.446
  7         164.90        170.23        -5.3332      161.460    179.007
  8         171.00        178.20        -7.1996      169.366    187.033
  9         183.40        188.64        -5.2434      179.892    197.395
 10         195.80        195.22          .5771      186.527    203.919
 11         207.40        204.80         2.5971      196.080    213.526
 12         218.30        213.67         4.6339      204.928    222.404
 13         226.80        221.74         5.0616      212.869    230.608
 14         237.90        235.05         2.8545      225.762    244.329
 15         225.80        223.31         2.4912      214.158    232.460
 16         232.40        231.35         1.0475      222.581    240.124
 17         241.70        240.47         1.2343      231.393    249.538
 18         249.20        246.60         2.6022      237.552    255.644
 19         261.30        263.16        -1.8557      253.695    272.616
 20         248.90        253.50        -4.5965      244.529    262.464
 21         226.80        230.96        -4.1608      220.830    241.091
 22         225.60        223.47         2.1290      213.659    233.283
 23         228.80        227.34         1.4612      218.096    236.581
 24         239.60        238.98          .6157      229.327    248.642
 25         244.70        243.92          .7775      233.681    254.164
 26         245.80        243.17         2.6290      233.009    253.333
 27         269.40        273.52        -4.1181      262.113    284.923

LM-Test for heteroskedasticity
==============================
Breusch-Pagan Lagrange Multiplier Test
H0: Homoskedatizität

 1. Command: -> ELM=NREG*RESID^2/SUMSQDEV
      calc; lm = .5 * Xss(X,elm)
    ; chi(lm,6)$

>>> LM       =     2.871945     <<<
>>> Result   =     .1752610     <<<    (1 – p-value)

Da LM=ch2=2.871945 < chi2(critical, df=6, alpha=0.05)=12.592
oder
da 1-p-value=0.1752610  bzw. p-value=0.824739 > alpha=0.05
wird H0:Homoskedastizität NICHT abgelehnt
```

```
Exkurs: Stata LM-Test for heteroskedasticity

. regress  y  pg puc pd ps

      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  4,    22) =   57.17
       Model |  50255283.8     4    12563821           Prob > F      =  0.0000
    Residual |   4834968.9    22  219771.314           R-squared     =  0.9122
-------------+------------------------------           Adj R-squared =  0.8963
       Total |  55090252.7    26  2118855.87           Root MSE      =   468.8

------------------------------------------------------------------------------
```

```
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          pg |  -1188.556   369.6332    -3.22   0.004    -1955.129   -421.9839
         puc |  -2048.381   633.9785    -3.23   0.004    -3363.172   -733.5902
          pd |   14718.97   6152.661     2.39   0.026     1959.134    27478.81
          ps |    4314.69   4862.642     0.89   0.385    -5769.812    14399.19
       _cons |    1755.72   1490.542     1.18   0.251    -1335.476    4846.916
------------------------------------------------------------------------------
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of y

         chi2(1)      =     0.16
         Prob > chi2  =   0.6892


Da pvalue(LM=0.16)= 0.6892 > alpha=0.05 wird H0:Homoskedastizität NICHT abgelehnt.

Exkurs Stata Ende
```

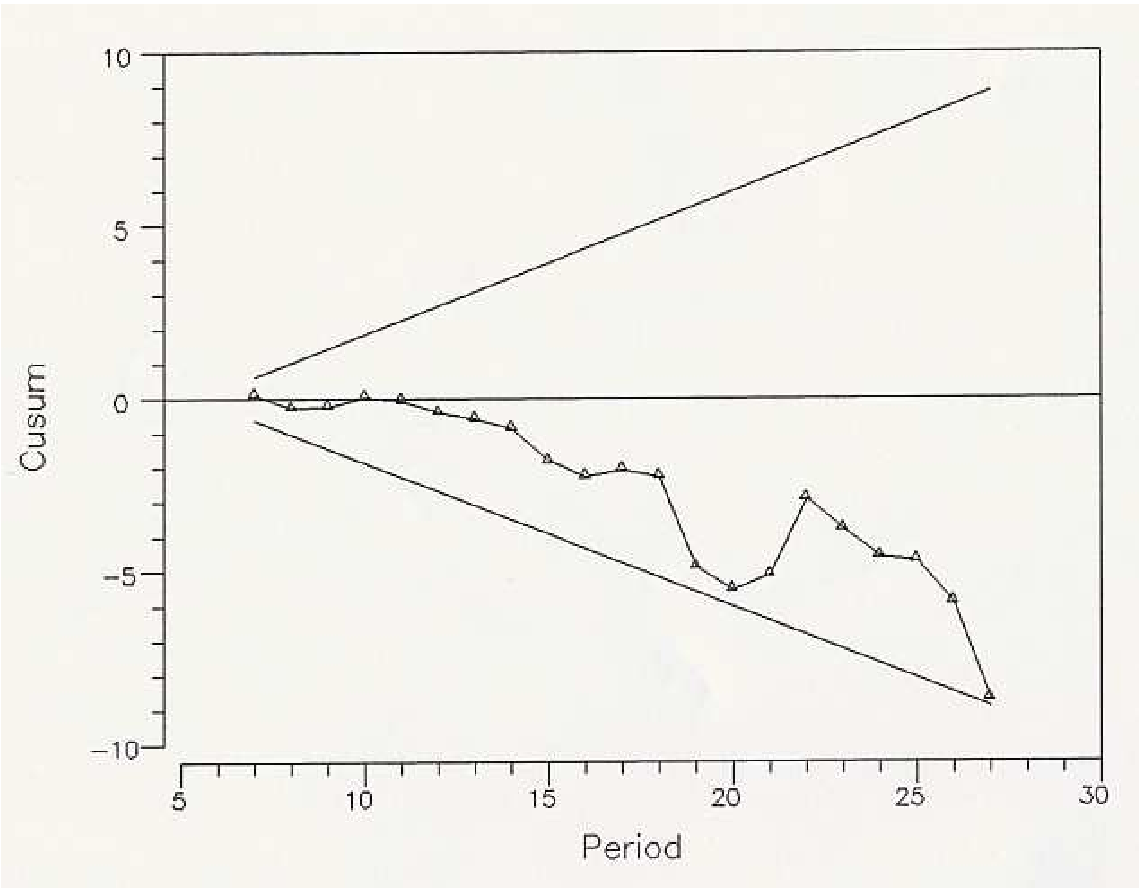
```
DATA LISTING (Current sample)
Observation      RESID
      1        4.2644
      2        2.6926
      3        2.3377
      4        2.1194
      5       -3.8685
      6       -5.7499
      7       -5.3332
      8       -7.1996
      9       -5.2434
     10        .57710
     11        2.5971
     12        4.6339
     13        5.0616
     14        2.8545
     15        2.4912
     16        1.0475
     17        1.2343
     18        2.6022
     19       -1.8557
     20       -4.5965
     21       -4.1608
     22        2.1290
     23        1.4612
     24        .61565
     25        .77746
     26        2.6290
     27       -4.1181

       Histogram for RESID    computed using   27 observations
       Obs. out of range: too low=   0, too high=   0
       Individual data      Mean=     .000, std.dev.=    3.630
                                       Frequency          Cumulative
      Lower Limit    Upper Limit    Total   Relative   Total   Relative
   0      -7.200         -5.973       1       .0370       1      .0370
   1      -5.973         -4.747       3       .1111       4      .1481
   2      -4.747         -3.521       4       .1481       8      .2963
   3      -3.521         -2.295       0       .0000       8      .2963
   4      -2.295         -1.069       1       .0370       9      .3333
   5      -1.069           .157       0       .0000       9      .3333
   6        .157          1.383       5       .1852      14      .5185
   7       1.383          2.609       7       .2593      21      .7778
   8       2.609          3.835       3       .1111      24      .8889
   9       3.835          5.062       3       .1111      27     1.0000
```

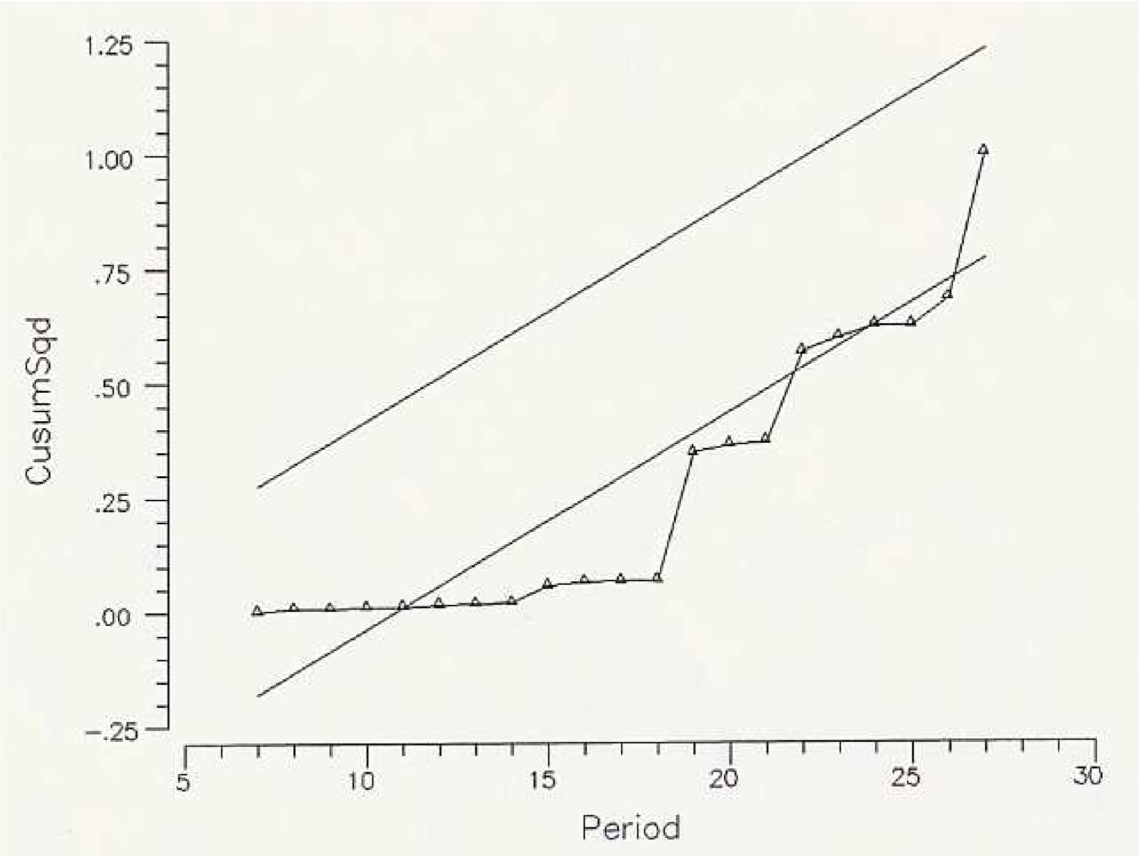

Interval width = 1.2261
Frequency
0
2
4
6
8
-6.6 -5.4 -4.1 -2.9 -1.7 -.5 .8 2.0 3.2 4.4
Histogram for RESID (by classes)


Residual plot. Bars mark 0 and +/- 2s(e).
Residual
6
4
2
0
-2
-4
-6
-8
1960
1970
1980
1990
YEAR

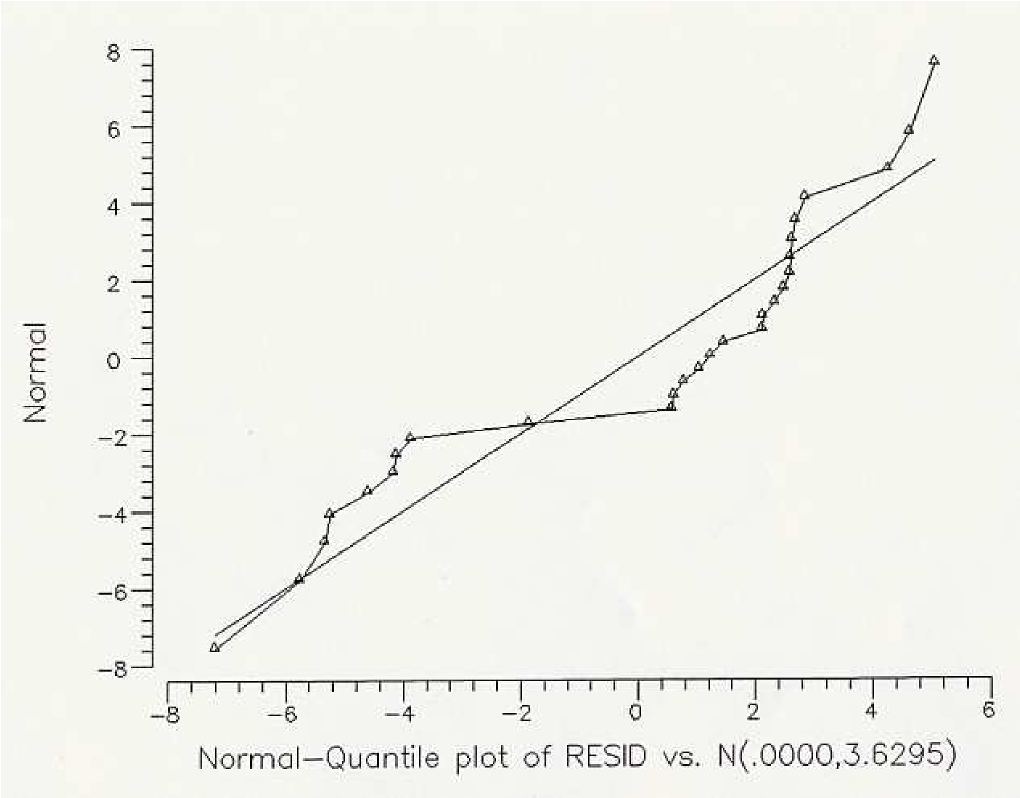
Pd.	Cusum	$Cusum^2$
7	.12	.001
8	-.27	.007
9	-.22	.007
10	.05	.010
11	-.05	.010
12	-.40	.015*
13	-.60	.017*
14	-.86	.020*
15	-1.81	.056*
16	-2.26	.064*
17	-2.05	.066*
18	-2.26	.068*
19	-4.88	.346*
20	-5.52	.363*
21	-5.10	.370*
22	-2.91	.567
23	-3.77	.597
24	-4.57	.623*
25	-4.70	.624*
26	-5.90	.683*
27	-8.69	1.000

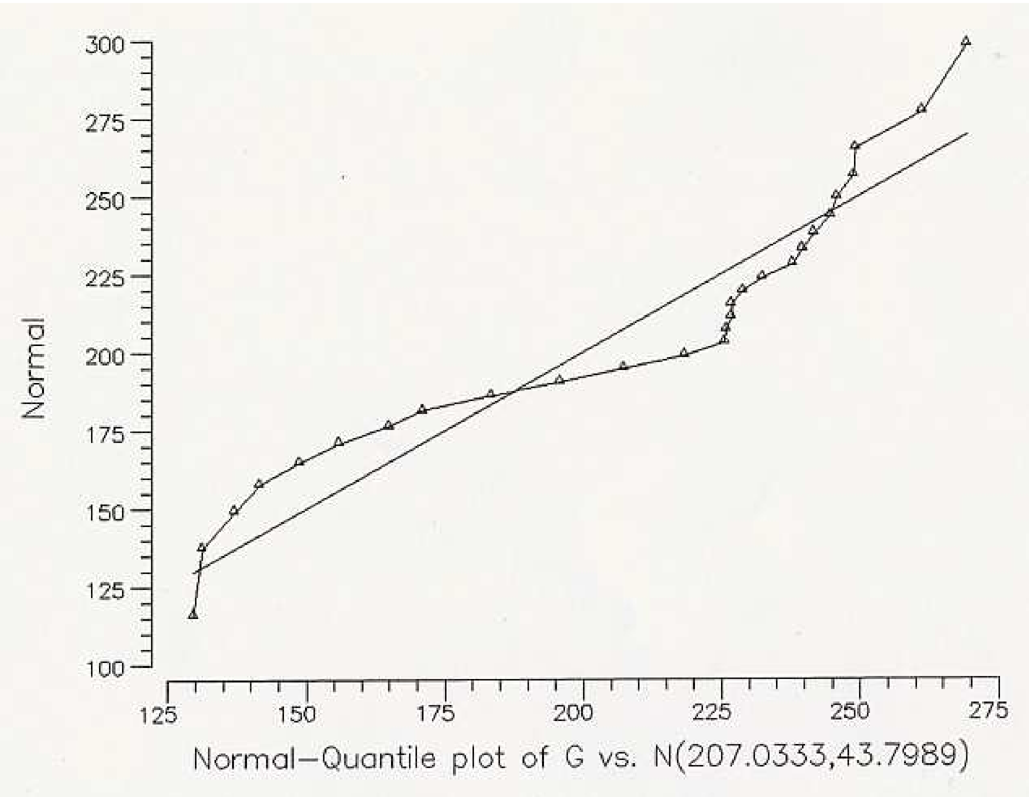
Die mit * markierten Werte liegen außerhalb des Vertrauensintervalls.






```
Normality Tests
Variable  Chi-squared  P-Value  Kolmogorov-Smirnov  P-Value
RESID          2.8839   .23646               .5000   .00000
```


Normal
Normal–Quantile plot of RESID vs. N(.0000,3.6295)


Normal
Normal–Quantile plot of G vs. N(207.0333,43.7989)

2 Return on Human Capital

Datenbasis: Zeitbudgeterhebung des Statistischen Bundesamtes 2001/02, Zeittagebücher

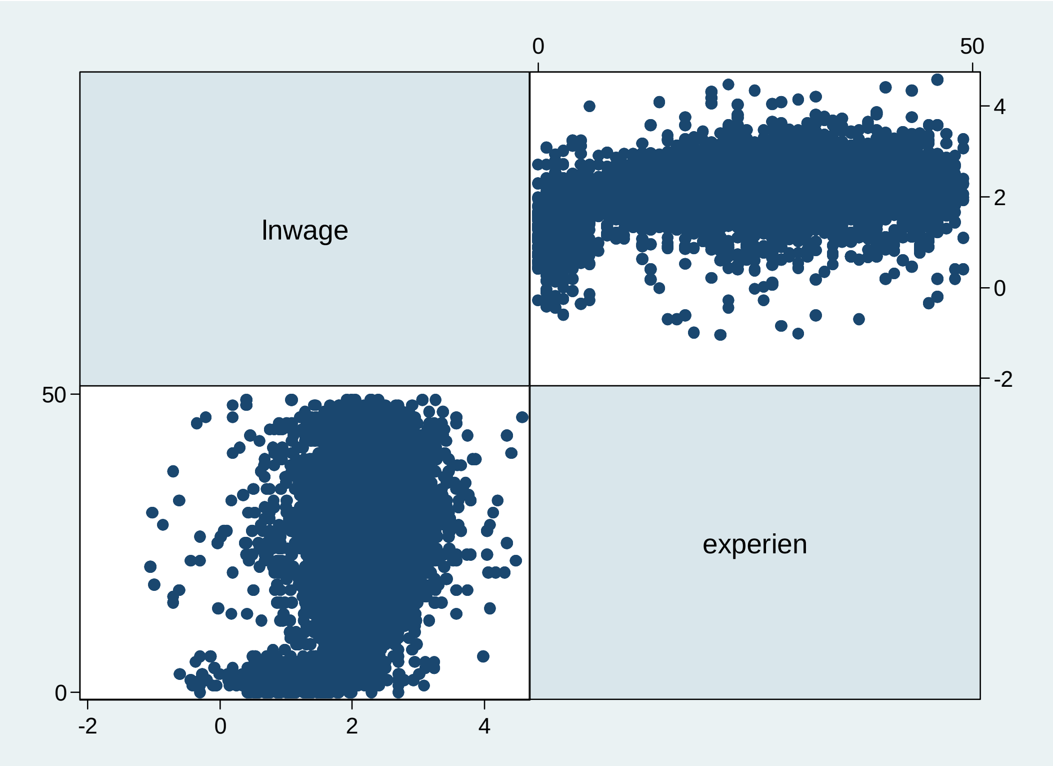
(Merz:. use "X:\EWF\Regression\Skript\Version 9 Regression noch nicht\zbe0102_merz_1.dta")

Deskription

```
. sum lnwage woman experien exper2 if wage>0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      lnwage |     19470    2.129235    .6176006  -1.049422   4.571884
       woman |     26802    .5335423     .498883          0          1
    experien |     24885    25.67077      12.493          0         49
      exper2 |     24885    815.0569    633.7039          0       2401
```

Spezifikation

```
. graph matrix lnwage experien
```



Graph deutet auf quadratischen Einfluss von experien

Regression

```
. regress lnwage woman experien exper2 if wage>0

      Source |       SS       df       MS              Number of obs =   18620
-------------+------------------------------           F(  3, 18616) = 1766.59
       Model |  1529.36323     3  509.787743           Prob > F      =  0.0000
    Residual |  5372.05879 18616  .288572131           R-squared     =  0.2216
-------------+------------------------------           Adj R-squared =  0.2215
       Total |  6901.42203 18619  .370665558           Root MSE      =  .53719

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       woman |  -.2635038   .0078969   -33.37   0.000    -.2789825   -.2480251
    experien |   .0686598   .0012007    57.18   0.000     .0663063    .0710132
      exper2 |  -.0011562   .0000252   -45.95   0.000    -.0012055   -.0011069
```

```
       _cons |   1.436838   .0140526   102.25   0.000     1.409293    1.464382
------------------------------------------------------------------------------


. estat sum

  Estimation sample regress             Number of obs =  18620

  -------------------------------------------------------------
      Variable |       Mean     Std. Dev.       Min        Max
  -------------+-----------------------------------------------
        lnwage |   2.146407     .6088231   -1.04942    4.57188
         woman |   .4877551     .4998635          0          1
      experien |   24.46149     11.37525          0         49
        exper2 |    727.754     543.2549          0       2401
  -------------------------------------------------------------

. estat vce

Covariance matrix of coefficients of regress model

        e(V) |      woman    experien      exper2       _cons
-------------+------------------------------------------------
       woman |  .00006236
    experien | -3.616e-07   1.442e-06
      exper2 |  1.082e-08  -2.892e-08   6.330e-10
       _cons | -.00002945  -.00001404   2.416e-07   .00019748
```

Damit können die t-Tests der Koeffizienten berechnet werden

Test der Annahmen des CLR

```
. predict residuals,r
(8248 missing values generated)
```

Goodness of fit

```
. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  18620   -17180.38   -14848.07      4     29704.14    29735.47
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Goodness of fit: Hohe Werte von AIC (Akaike Information Criterion) sprechen für guten fit (vgl. auch das für Querschnitte relativ gute R2 von 22%).

Spezifikation

```
. estat linktest

      Source |       SS       df       MS              Number of obs =   18620
-------------+------------------------------           F(  2, 18617) = 2653.14
       Model |  1530.76255     2  765.381273           Prob > F      =  0.0000
    Residual |  5370.65948 18617  .288481467           R-squared     =  0.2218
-------------+------------------------------           Adj R-squared =  0.2217
       Total |  6901.42203 18619  .370665558           Root MSE      =   .5371

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _hat |   1.310816   .1417901     9.24   0.000     1.032895    1.588738
      _hatsq |  -.0794456   .0360716    -2.20   0.028    -.1501492    -.008742
       _cons |  -.2946021    .137028    -2.15   0.032    -.5631894   -.0260148
------------------------------------------------------------------------------
```

linktest creates two new variables, the variable of prediction, **_hat**, and the variable of squared prediction, **_hatsq**. The model is then refit using these two variables as predictors. **_hat** should be significant since it is the predicted value. On the other hand, **_hatsq** shouldn't, because if our model is specified correctly, the squared predictions should not have much explanatory power. That is we wouldn't expect **_hatsq** to be a significant predictor if our model is specified correctly. So we will be looking at the p-value for **_hatsq**.

_hatsq ist significant (alpha=5%), d.h. die Ergebnisse deuten auf eine Fehlspezifikation hin.

Ausgelassene (omitted) Variablen

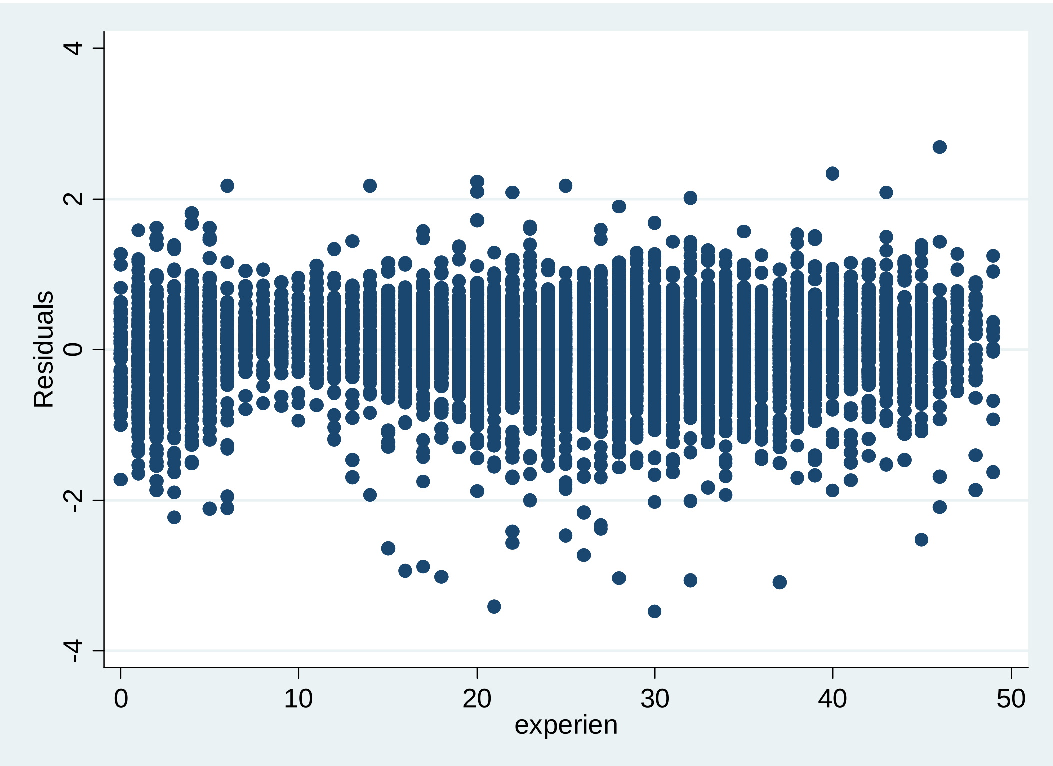
```
. estat ovtest

Ramsey RESET test using powers of the fitted values of lnwage
       Ho:  model has no omitted variables
               F(3, 18613) =      6.87
                  Prob > F =      0.0001
```

H0 wird abgelehnt (alpha=5%), d.h. die Ergebnisse deuten auf ausgelassene Variablen hin (siehe auch linktest, d.h. Spezifikation ist noch nicht optimal).

Homoskedastizität/Heteroskedastizität

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lnwage

         chi2(1)      =   118.79
         Prob > chi2  =   0.0000
```

H0 wird abgelehnt, d.h. Hinweis auf Heteroskedastizität.



Multikollinearität

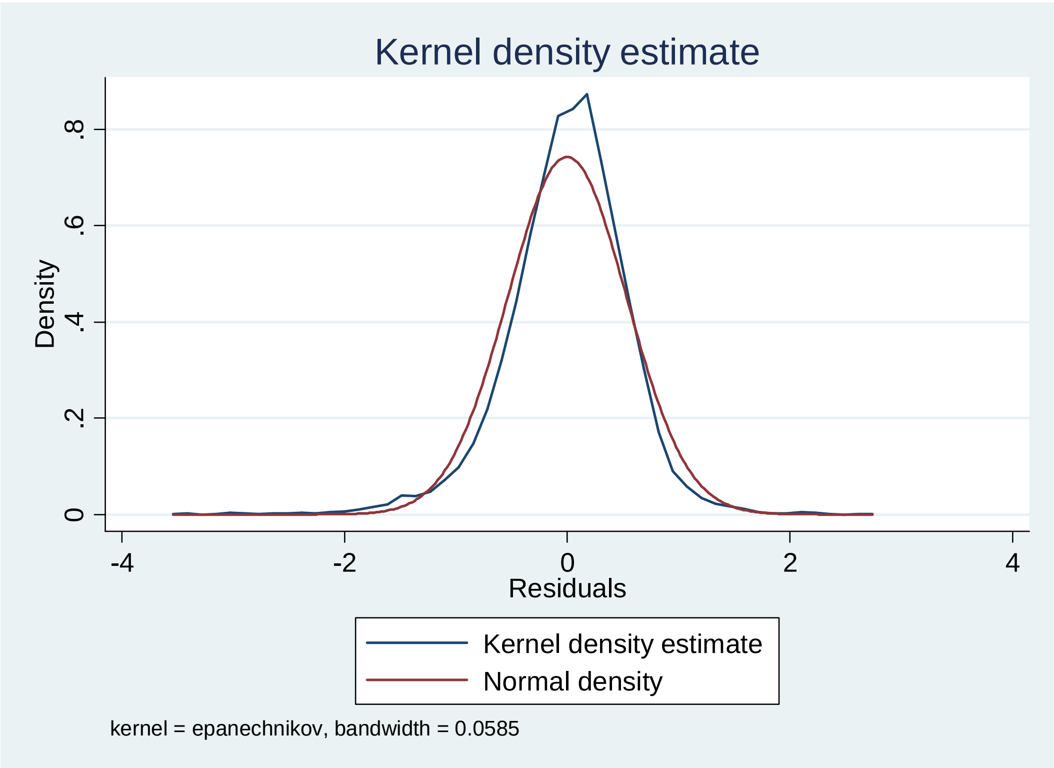
```
. vif

    Variable |       VIF       1/VIF
-------------+----------------------
      exper2 |     12.05    0.082957
    experien |     12.04    0.083083
       woman |      1.01    0.994672
-------------+----------------------
    Mean VIF |      8.37
```

vif stands for *variance inflation factor*. As a rule of thumb, a variable whose VIF values are greater than 10 may merit further investigation. Tolerance, defined as 1/VIF, is used by many researchers to check on the degree of collinearity. A tolerance value lower than 0.1 is comparable to a VIF of 10. It means that the variable could be considered as a linear combination of other independent variables.

Der Wert vif >10 deutet auf Multikollinearität zwischen den experience Variablen hin, nicht aber hinsichtlich woman (ist aber dummy Variable).

Normalität

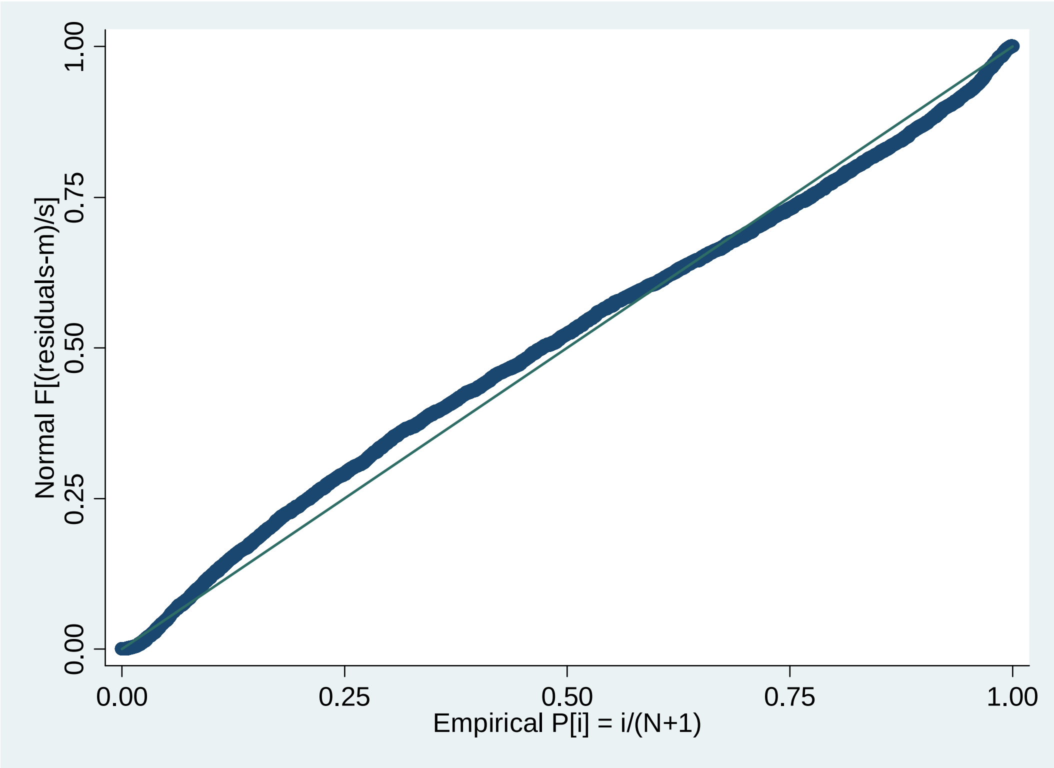
. kdensity residuals,normal



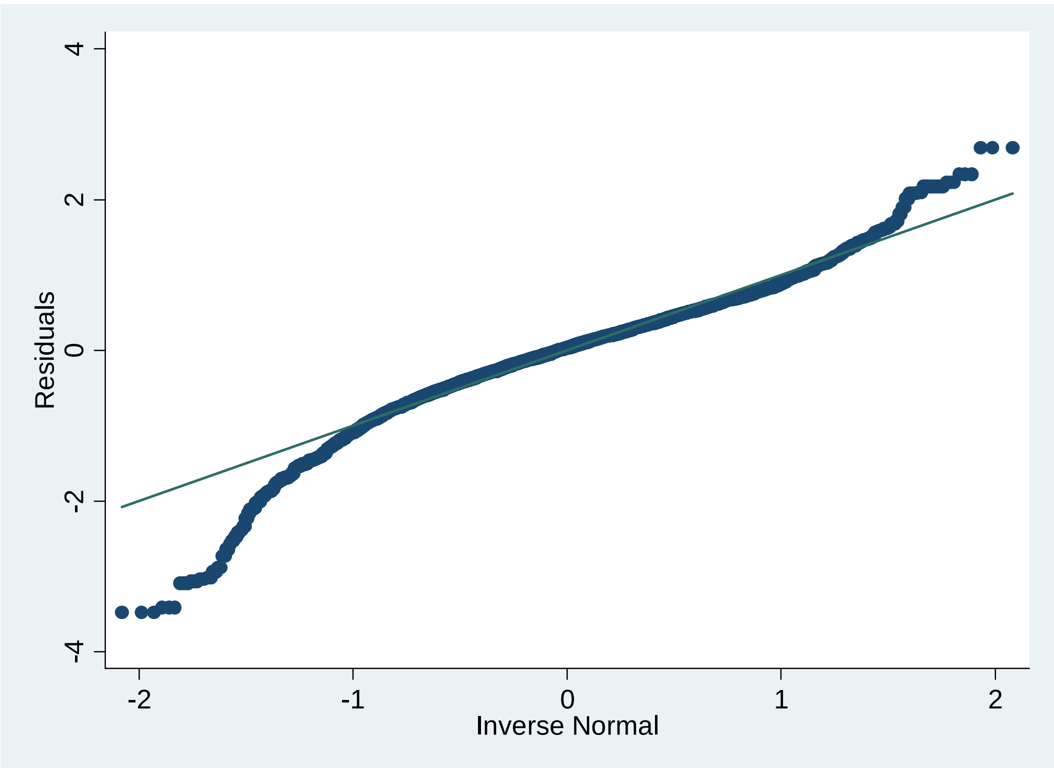
The **kdensity** command produce a kernel density plot with the **normal** option requesting that a normal density be overlaid on the plot. **kdensity** stands for kernel density estimate. It can be thought of as a histogram with narrow bins and moving average.

Die Übereinstimmung beider Verteilungen weist auf eine Normalverteilung der Residuen hin.

```
. pnorm residuals
```




```
. qnorm residuals
```



The **pnorm** command graphs a standardized normal probability (P-P) plot while **qnorm** plots the quantiles of a variable against the quantiles of a normal distribution. **pnorm** is sensitive to non-normality in the middle range of data and **qnorm** is sensitive to non-normality near the tails.

Kdensity und pnorm deuten auf Normalverteilung der Residuen hin. Allerdings gibt es Abweichungen im unteren und oberen Bereich der Verteilung der Residuen (qnorm).

```
. sktest residuals

                   Skewness/Kurtosis tests for Normality
                                                    ------- joint ------
    Variable |    Obs   Pr(Skewness)   Pr(Kurtosis)  adj chi2(2)    Prob>chi2
-------------+--------------------------------------------------------------
   residuals |   1.9e+04    0.000          0.000            .             .
```

Skewness und Kurtosis sind die zentralen Elemente des Jarque-Bera Tests (H0: Normalverteilung). Hier zeigen kleine p-values (pr(.) auf Ungleichheit der Schiefe und Wölbung der Regressionsresiduen mit der entsprechenden Normalverteilung

```
. *dfbeta: influential variables?
. dfbeta
(8248 missing values generated)
                       _dfbeta_1: dfbeta(woman)
(8248 missing values generated)
                       _dfbeta_2: dfbeta(experien)
(8248 missing values generated)
                       _dfbeta_3: dfbeta(exper2)

. *influential if dfbeta_*i > 2/sqrt(n)
. . display 2/(sqrt(18620))
.01465683

. sum _dfbeta_* if _dfbeta_1>0.01465683

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   _dfbeta_1 |       467    .0207707    .0067927   .0146683   .0484624
   _dfbeta_2 |       467   -.0009564    .0260667  -.0721212   .0702621
   _dfbeta_3 |       467    .0016756    .0254104  -.0804605    .096849
```

D.h. von den 18620 Beobachtungen der Regression sind 467 Beobachtungen besonders einflussreich

```
.
. *cooksd: influential variables together?
. predict cooksd if e(sample), cooksd
(8248 missing values generated)

. *influential if cooksdi>4/n
. . display 4/18620
.00021482

. sum cooksd if cooksd>0.00021482

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      cooksd |      1114    .0005203    .0004328   .0002149   .0043046
```

D.h. von den 18620 Beobachtungen der Regression sind über alle Variablen gemeinsam 1114 Beobachtungen besonders einflussreich.

Mit den jeweiligen Grenzwerten können dann diese Beobachtungen von der weiteren Regression ausgeschlossn werden:

```
. *Original regression
. regress lnwage woman experien exper2 if wage>0

      Source |       SS       df       MS              Number of obs =   18620
-------------+------------------------------           F(  3, 18616) = 1766.59
       Model |  1529.36323     3  509.787743           Prob > F      =  0.0000
    Residual |  5372.05879 18616  .288572131           R-squared     =  0.2216
-------------+------------------------------           Adj R-squared =  0.2215
       Total |  6901.42203 18619  .370665558           Root MSE      =  .53719

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       woman |  -.2635038   .0078969   -33.37   0.000    -.2789825   -.2480251
    experien |   .0686598   .0012007    57.18   0.000     .0663063    .0710132
      exper2 |  -.0011562   .0000252   -45.95   0.000    -.0012055   -.0011069
       _cons |   1.436838   .0140526   102.25   0.000     1.409293    1.464382
------------------------------------------------------------------------------

.
. *Regression without particular influential variables
. regress lnwage woman experien exper2 if wage>0 & cooksd <= 0.00021482

      Source |       SS       df       MS              Number of obs =   17506
-------------+------------------------------           F(  3, 17502) = 2160.63
       Model |  1186.59716     3  395.532388           Prob > F      =  0.0000
    Residual |  3203.97317 17502  .183063259           R-squared     =  0.2703
-------------+------------------------------           Adj R-squared =  0.2701
       Total |  4390.57033 17505  .250818071           Root MSE      =  .42786

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       woman |  -.2900815   .0064838   -44.74   0.000    -.3027904   -.2773726
    experien |   .0640905   .0010671    60.06   0.000      .061999    .0661821
      exper2 |  -.0010975   .0000221   -49.57   0.000    -.0011409   -.0010541
       _cons |   1.546821   .0126639   122.14   0.000     1.521999    1.571644
------------------------------------------------------------------------------
```

Resultat: Goodness of fit mit Adj R-squared ist besser geworden, alle Variablen haben an Signifikanz gewonnen. Trotzdem: Die ausgelassenen Variablen sollten weiter inspiziert warden.

```
. sum wage lnwage woman experien exper2 if wage>0 & cooksd > 0.00021482

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        wage |      1964    9.145597    12.56913   .3501401   96.72619
      lnwage |      1964    1.589771    1.105677  -1.049422   4.571884
       woman |      9296     .613167    .4870511          0          1
    experien |      7379    27.97235    15.55944          0         49
      exper2 |      7379    1024.516     805.294          0       2401
```

XI DAS GENERALISIERTE LINEARE REGRESSIONSMODELL (GLS)

In der ökonometrischen Praxis erweisen sich die Annahmen des Klassischen Linearen Modells (CLR) in der Regel als zu restriktiv. Oft lassen die im vorangegangenen Kapitel vorgestellten Tests darauf schließen, dass mindestens eine der Annahmen verletzt ist. Insbesondere A3, die Homoskedastizität und die Freiheit von Autokorrelation unterstellt, hält einer empirischen Untersuchung zugunsten des Klassischen Linearen Modells oft nicht stand. Damit verliert die im Gauss-Markov-Theorem bewiesene Effizienzeigenschaft der Methode der kleinsten Quadrate ihre Gültigkeit und die OLS- Schätzung damit ihre Legitimation. Die verallgemeinerte Methode der kleinsten Quadrate (Generalized Least Squares, GLS) ist hierfür eine Lösung.

Im Literaturverzeichnis sind wieder in den Ökonometrielehrbüchern ausführliche Unterlagen zu GLS zu finden. An neuerer englischsprachiger Literatur sind Greene (2003), Wooldridge (2006) und Studenmund (2006) und an neuerer deutscher Literatur Fahrmeir, Kneib und Lang (2009), Bauer, Fertig und Schmidt (2009), Hübler (2005) und von Auer (2003) besonders zu nennen.

1 GLS BEI BEKANNTER KOVARIANZMATRIX DER STÖRTERME

Statt der OLS- Methode lässt sich im Fall von Autokorrelation oder Heteroskedastizität eine Verallgemeinerung dieses Schätzansatzes anwenden: In beiden Fällen ist Annahme A3 verletzt, die vorsieht, dass

$$E(\boldsymbol{\varepsilon\varepsilon}') = \sigma^2 \mathbf{I}$$

Die Verallgemeinerte Methode der kleinsten Quadrate oder Generalized Least Squares (GLS) – oder auch Aitken-Schätzer genannt - hebt diese Annahme auf und geht allgemein davon aus, dass sich die Varianz-Kovarianz-Matrix der Störterme als

$$E(\boldsymbol{\varepsilon\varepsilon}') = \sigma^2 \mathbf{\Omega}$$

darstellen lässt.

Wenn $\mathbf{\Omega}$ symmetrisch und positiv definit ist, gilt dies auch für ihre Inverse $\mathbf{\Omega}^{-1}$, die sich dann auch schreiben lässt als

$$\mathbf{\Omega}^{-1} = \mathbf{T}'\mathbf{T}.$$

Man formt nun das möglicherweise fehlerbehaftete ökonometrische Modell um, indem die zunächst als bekannt unterstellte Matrix $\mathbf{T}$ von links anmultipliziert wird:

$$\mathbf{Ty} = \mathbf{TX}\boldsymbol{\beta} + \mathbf{T}\boldsymbol{\varepsilon}$$

Mit den Definitionen $\mathbf{Ty} = \mathbf{y}^*$, $\mathbf{TX} = \mathbf{X}^*$, $\mathbf{T}\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^*$ ergibt sich das transformierte, ökonometrische Modell:

$$\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$$

In diesem Modell erfüllen die Residuen wieder die klassischen Eigenschaften des weißen:

$$E(\varepsilon^*) = E(\mathbf{T}\varepsilon) = \mathbf{T}\, E(\varepsilon) = 0$$

und

$$\begin{aligned} E(\varepsilon^* \varepsilon^{*\prime}) &= E(\varepsilon^*) = E(\mathbf{T}\,\varepsilon\,(\mathbf{T}\,\varepsilon)') = E(\mathbf{T}\,\varepsilon\,\varepsilon'\mathbf{T}') = \mathbf{T}\,\mathbf{E}(\varepsilon\,\varepsilon')\mathbf{T}' \\ &= \mathbf{T}\sigma^2\mathbf{\Omega}\mathbf{T}' = \sigma^2\mathbf{T\Omega T}' \\ &= \sigma^2\,\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}' = \sigma^2\mathbf{T}\mathbf{T}^{-1}(\mathbf{T}')^{-1}\mathbf{T}' \\ &= \sigma^2\,\mathbf{I} \end{aligned}$$

Die Transformation mit der Matrix **T** mit $\mathbf{T\Omega T'=I}$ führt zu den gewünschten klassischen Eigenschaften und damit zur Effizienz der OLS- Schätzung.

Einer Schätzung steht nun nichts mehr im Wege:

$$\begin{aligned} \mathbf{b}_{GLS} &= (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{y}^* \\ &= (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\,\mathbf{\Omega}^{-1}\mathbf{y} \end{aligned}$$

Um zudem inferenzstatistische Aussagen über den so gewonnenen Koeffizientenvektor $\mathbf{b}_{GLS}$ zu treffen, muss die Varianz-Kovarianz-Matrix dieses Schätzers bekannt sein. Sie ergibt sich als:

$$\begin{aligned} Var(\mathbf{b}_{GLS}) &= E[(\mathbf{b}_{GLS} - E(\mathbf{b}_{GLS}))(\mathbf{b}_{GLS} - E(\mathbf{b}_{GLS}))'] = \sigma^2_{GLS}\,(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1} \\ &= \sigma^2_{GLS}\,(\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1} \end{aligned}$$

Eine erwartungstreue Schätzung der Varianz der transformierten Störgrößen ε^* ist wie im klassischen Modell:

$$\hat{\sigma}^2_{GLS} = \frac{\mathbf{e}^{*\prime}\mathbf{e}^*}{n - K - 1}$$

mit $e^* = y^* - X^{*\prime}b_{GLS}$ (GLS - Residuen) :

$$\hat{\sigma}^2_{GLS} = \frac{(\mathbf{y}^* - \mathbf{X}^{*\prime}\mathbf{b}_{GLS})'(\mathbf{y}^* - \mathbf{X}^{*\prime}\mathbf{b}_{GLS})}{n - K - 1}$$

$$= \frac{(\mathbf{Ty} - \mathbf{TX}'\mathbf{b}_{GLS})'(\mathbf{Ty} - \mathbf{TX}'\mathbf{b}_{GLS})}{n - K - 1}$$

$$= \frac{(\mathbf{y} - \mathbf{X}'\mathbf{b}_{GLS})'\mathbf{T}'\mathbf{T}(\mathbf{y} - \mathbf{X}'\mathbf{b}_{GLS})}{n - K - 1}$$

$$\hat{\sigma}^2_{GLS} = \frac{(\mathbf{y} - \mathbf{X}'\mathbf{b}_{GLS})'\mathbf{\Omega}^{-1}(\mathbf{y} - \mathbf{X}'\mathbf{b}_{GLS})}{n - K - 1}$$

Demnach ist weder für die Schätzung des Koeffizientenvektors, noch der Varianz-Kovarianz-Matrix Kenntnis der Transformationsmatrix T notwendig. Stattdessen ist die hier noch bekannte Kovarianzmatrix **Ω** des ursprünglichen, untransformierten Modells einbeziehbar.

2 GLS bei unbekannter Kovarianzmatrix der Störterme

Ist die Kovarianzstruktur des ursprünglichen Modells nicht bekannt, dies trifft in der Praxis auf fast alle empirischen Untersuchungen zu, muss diese geschätzt werden; $\mathbf{\Omega}$ ist also durch $\hat{\mathbf{\Omega}}$ zu schätzen[1]

$$\mathbf{b}_{\mathrm{GLS}} = (\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{y}$$

$$\widehat{\mathrm{Var}}\,(\mathbf{b}_{\mathrm{GLS}}) = \hat{\sigma}^2_{\mathrm{GLS}}\,(\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X})^{-1}$$

$$\hat{\sigma}^2_{\mathrm{GLS}} = \frac{(\mathbf{y} - \mathbf{X'b}_{\mathrm{GLS}})'\hat{\mathbf{\Omega}}^{-1}(\mathbf{y} - \mathbf{X'b}_{\mathrm{GLS}})}{\mathrm{n - K - 1}}$$

b_{GLS} ist in diesem Fall zwar nicht mehr erwartungstreu, aber weiterhin konsistent.

Das Problem der Schätzung der Varianz-Kovarianz-Matrix besteht darin, dass

$$\mathrm{E}(\boldsymbol{\varepsilon\varepsilon}') = \mathrm{E}\begin{bmatrix} \varepsilon_1\varepsilon_1 & \cdots & \varepsilon_1\varepsilon_N \\ \vdots & \ddots & \vdots \\ \varepsilon_1\varepsilon_N & \cdots & \varepsilon_N\varepsilon_N \end{bmatrix} = \sigma^2\,\mathbf{\Omega}$$

(aufgrund der Symmetrie) N(N-1)/2 verschiedene Elemente enthält, die mit den n zur Verfügung stehenden Beobachtungen prinzipiell nicht geschätzt werden können. Man ist in diesem ersten Schritt deshalb darauf angewiesen, Annahmen über eine mögliche Struktur der $\mathbf{\Omega}$- Matrix zu treffen, mit denen sich die Anzahl der zu schätzenden Parameter verringert. Solche Annahmen über die Heteroskedastizität als auch die Autokorrelation werden im Folgenden besprochen.

2.1 GLS bei Heteroskedastizität

Wurde im Ausgangsmodell ausschließlich Heteroskedastizität und damit keine Autokorrelation der Residuen festgestellt, stellt sich die Varianz-Kovarianz-Matrix allgemein als

$$\mathrm{E}(\boldsymbol{\varepsilon\varepsilon}') = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_N^2 \end{bmatrix}$$

dar. Damit sind allerdings weiterhin mehr Parameter unbestimmt als mit den N Beobachtungen geschätzt werden können.

Handelt es sich um Zeitreihendaten, so bietet es sich an, die Hauptdiagonale von $\sigma^2\,\mathbf{\Omega}$ in $m < N$ Gruppen mit identischen Elementen aufzuteilen, um die Anzahl der zu schätzenden Parameter zu verringern:

[1] In der ökonometrischen Literatur wird ein GLS- Schätzer, der sich auf die geschätzte $\mathbf{\Omega}$-Matrix stützt, auch FGLS- Schätzer (Feasible Generalized Least Squares, vgl. z. B. Greene (1999, S.511ff.) oder EGLS (Estimated Generalized Least Squares, vgl. Griffith et al. (1993, S.498ff) genannt.

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_1^2 & 0 & & \cdots & & 0 \\ 0 & \ddots & & & & \\ & & \sigma_1^2 & & & \vdots \\ \vdots & & & \sigma_m^2 & & \\ & & & & \ddots & 0 \\ 0 & & \cdots & & 0 & \sigma_m^2 \end{bmatrix}$$

Anlass für die Formulierung eines Modells mit m=2 bietet insbesondere ein vorangegangener Goldfeld- Quandt- Test, der auf eine solche Kovarianzstruktur hindeutet.

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_1^2 & 0 & & \cdots & & 0 \\ 0 & \ddots & & & & \\ & & \sigma_1^2 & & & \vdots \\ \vdots & & & \sigma_2^2 & & \\ & & & & \ddots & 0 \\ 0 & & \cdots & & 0 & \sigma_2^2 \end{bmatrix}$$

Eine andere Herangehensweise empfiehlt sich bei der Verwendung von Querschnittsdaten. In diesem Fall ist die Annahme, die Varianz der Störterme könne von der individuellen Beobachtung beeinflusst sein i.d.R. nicht gerechtfertigt. Heteroskedastizität drückt sich in diesem Fall als Abhängigkeit der Fehlerstreuung von Elementen der Regressormatrix aus. Die beobachtungsspezifische Varianz lässt sich dann darstellen als:

$$\sigma_i^2 = \sigma^2 z_i^\delta \qquad i = 1, \ldots N$$

mit z_i als Beobachtung eines Regressors oder einer Linearkombination aus mehreren Regressorvariablen der i- ten Beobachtung. Für den Parameter δ muss wegen der Nichtnegativität von σ_i^2 eine gerade Zahl festgelegt werden.

Einen Hinweis auf diese Form des Modelldefektes liefert der im vorangegangenen Kapitel vorgestellte Breusch-Pagan-Test: Liefert er als Ergebnis eine Abhängigkeit der Varianz von den Regressoren, so lässt sich die Varianz-Kovarianz-Matrix allgemein folgendermaßen darstellen:

$$E(\varepsilon\varepsilon') = \sigma^2\,\mathbf{\Omega} = \begin{bmatrix} \sigma^2 z_1^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 z_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma^2 z_N^2 \end{bmatrix} =$$

Die Ω- Matrix muss in diesem Fall nicht geschätzt werden, sie ist unter Annahme der oben genannten Struktur der Störterme einfach gegeben als:

$$\hat{\mathbf{\Omega}} = \begin{bmatrix} z_1^2 & 0 & \cdots & 0 \\ 0 & z_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & z_N^2 \end{bmatrix}, \text{ ihre Inverse als } \hat{\mathbf{\Omega}}^{-1} = \begin{bmatrix} 1/z_1^2 & 0 & \cdots & 0 \\ 0 & 1/z_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1/z_N^2 \end{bmatrix}.$$

Daraus lässt sich leicht die Transformationsmatrix **T** aus **T'T**= $\hat{\mathbf{\Omega}}^{-1}$ ableiten. Sie macht deutlich, warum diese Form der Verallgemeinerten KQ- Schätzung auch „**gewichtete KQ- Schätzung**" genannt wird:

$$\mathbf{T} = \begin{bmatrix} 1/z_1 & 0 & \cdots & 0 \\ 0 & 1/z_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1/z_N \end{bmatrix},$$

Die Transformationsmatrix **T** sorgt dafür, dass jede einzelne Beobachtung durch den Wert der Linearkombination z_i geteilt wird, und damit eine Gewichtung erhält. Je größer dieser Wert z_i, von dem die Varianz abhängt, desto geringer das Gewicht, mit dem dazugehörige Beobachtung in die Regression eingeht.

2.2 GLS bei Autokorrelation

Im Falle von autokorrelierten Störgrößen erster Ordnung und eines über die Beobachtungen konstanten Autokorrelationskoeffizienten ϱ, also bei Annahme folgender Beziehung zwischen den Störtermen

$$\varepsilon_i = \rho\varepsilon_{i-1} + \upsilon_i\,,$$

lässt sich folgende Varianz-Kovarianz-Matrix der Störgrößen ableiten[2]:

$$E(\varepsilon\varepsilon') = \sigma^2\,\mathbf{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & & & \vdots \\ \rho^2 & & \ddots & & \vdots \\ \vdots & & & 1 & \rho \\ \rho^{N-1} & \rho^{N-2} & \cdots & \rho & 1 \end{bmatrix} =$$

Die Matrix ist von dem einzigen Parameter ϱ abhängig, der zunächst unbekannt ist und deshalb geschätzt werden muss. Dies geschieht, indem im ersten Schritt das untransformierte, autokorrelierte Modell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

geschätzt und dann mittels der Residuen dieser Regression der Autokorrelationskoeffizient ϱ aus

$$e_i = \varrho\, e_{i-1} + v_i$$

ermittelt wird.

Im dritten Schritt wird schließlich der geschätzte Autokorrelationskoeffizient ϱ zur Bestimmung der Schätzung für die Kovarianzmatrix verwendet:

[2] Vgl. z.B. Eckey et al. (2001, S. 128 ff.)

$$\sigma^2\ \hat{\boldsymbol{\Omega}} = \sigma^2 \begin{bmatrix} 1 & \hat{\rho} & \hat{\rho}^2 & \cdots & \hat{\rho}^{N-1} \\ \hat{\rho} & 1 & & & \vdots \\ \hat{\rho}^2 & & \ddots & & \vdots \\ \vdots & & & 1 & \hat{\rho} \\ \hat{\rho}^{N-1} & \hat{\rho}^{N-2} & \cdots & \hat{\rho} & 1 \end{bmatrix}$$

Analog zum Fall von GLS bei Heteroskedastizität erhält man den Schätzer als:

$$\mathbf{b}_{\mathrm{GLS}} = (\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\ \hat{\boldsymbol{\Omega}}^{-1}\mathbf{y},$$

die (geschätzte) Varianz der geschätzten Koeffizienten als:

$$\widehat{\mathrm{Var}}\ (\mathbf{b}_{\mathrm{GLS}}) = \hat{\sigma}^2_{\mathrm{GLS}}\,(\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X})^{-1}$$

und den Schätzer für die Fehlerstreuung mit

$$\hat{\sigma}^2_{\mathrm{GLS}} = \frac{(\mathbf{y} - \mathbf{X}'\mathbf{b}_{\mathrm{GLS}})'\hat{\boldsymbol{\Omega}}^{-1}(\mathbf{y} - \mathbf{X}'\mathbf{b}_{\mathrm{GLS}})}{\mathrm{n - K - 1}}.$$

Neben der separaten Behandlung von Heteroskedastizität und Autokorrelation können natürlich auch beide Verletzungen der Grundannahmen in einer Ω-Matrix berücksichtigt werden.

2.3 Alternative GLS-Schätzung bei Autokorrelation: Cochran-Orcutt- und Hildreth-Lu-Ansatz

Abschließend soll noch eine etwas einfachere Herangehensweise vorgestellt werden, die bei Autokorrelation der Störgrößen statt der OLS- Methode empfohlen werden: Die Ansätze

- nach Cochran und Orcutt sowie
- nach Hildreth und Lu

kommen ohne die Bestimmung und Invertierung einer **Ω**-Matrix aus. Beide Ansätze gehen von dem untransformierten Modell

$$y_i = x_i\beta + \varepsilon_i$$

mit Autokorrelation erster Ordnung aus:

$$\varepsilon_i = \rho\varepsilon_{i-1} + \upsilon_i \qquad \text{und} \qquad \upsilon_i \sim \mathrm{N}(0,\sigma^2).$$

Durch Einsetzen erhält man

$$\begin{aligned} y_i &= x_i\beta + \rho\varepsilon_{i-1} + \upsilon_i. \\ &= x_i\beta + \rho(y_{i-1} - x_{i-1}\beta) + \upsilon_i. \end{aligned}$$

$$y_i - \rho\ y_{i-1} = \rho\ (x_i - x_{i-1})\ \beta + \upsilon_i$$

Da ϱ und β in nichtlinearer Beziehung stehen, lassen sich die beiden nicht nach MKQ bestimmen. An dieser Stelle setzen die beiden Lösungsvorschläge an:

Cochran/Orcutt gehen wie folgt vor: Sie schätzen zunächst das Ausgangsmodell und mit Hilfe der Residuen den Autokorrelationskoeffizienten. Eingesetzt in die letzte Gleichung ergibt sich ein schätzbares Modell mit dem gewünschten additiven weißen Rauschen im Störterm υ_i. Die Koeffizienten in β lassen sich mit OLS bestimmen.

Hildreth/Lu berechnen für alternative, zunächst hypothetische Werte von ρ [ρ = 0,1; 0,2;…;0,9] die Regressionskoeffizienten. Anschließend wählen sie diejenige ρ /β –Kombination als Schätzer, die die resultierende Residuenquadratsumme $\upsilon_i'\upsilon_i$ minimiert.

XII PANELDATENMODELLE

1 PANELMODELLE

Paneldaten werden aus wiederholten Beobachtungen der gleichen Einheiten über die Zeit gewonnen. In den Wirtschafts- und Sozialwissenschaften sind das in der Regel wiederholte Befragungen der gleichen Mikroeinheiten. In diesen Daten wird somit ein relativ großer Querschnitt mit den gleichen Mikroeinheiten (Personen, Firmen) über mehrere Perioden betrachtet. Beispiele sind das Sozio-ökonomische Panel (SOEP), begonnen durch den Sonderforschungsbereich 3, Mikroanalytische Grundlagen der Gesellschaftspolitik' der Universitäten Frankfurt und Mannheim und nun vom DIW in Berlin weitergeführt, mit nun ca. 8.500 Haushalten und über 20.000 Personen, das Hannoveraner Firmenpanel oder die Michigan Panel Study of Income Dynamics (PSID), das mit ca. 6.000 Familien und ca. 15.000 Personen bereits 1968 begonnen wurde.

Mit Schwerpunkt auf vielen Querschnittsbeobachtungen in den Wirtschafts- und Sozialwissenschaften spielt die Heterogenität über die Mikroeinheiten eine besondere Rolle. Mit Paneldaten kann zwischen Kohorten-, Alters- und Periodeneffekten unterschieden werden.

Für die explizite Berücksichtigung dieser Heterogenität sowie des zeitlichen Einflusses sind generell Fixed und Random Effects Modelle entwickelt worden. Greene (2000, chapt. 14) bspw. behandelt allgemein den Fall linearer Panelmodelle.

2 ERFASSUNG VON UNBEOBACHTETEN ZEITEINFLÜSSEN

Bei Panelbeobachtungen liegen häufig Beobachtungen einer großen Anzahl an Individuen i über mehrere (wenige) Perioden t vor. Dadurch lassen sich unbeobachtete Zeiteinflüsse durch Jahres-, Quartals- oder Monatsdummies erfassen.

$$D_t = \begin{cases} 1 & \text{wenn Beobachtung aus Periode t} \\ 0 & \text{sonst} \end{cases}$$

Eine Periode ist als Basisperiode zu definieren. Die Koeffizienten der Dummies messen den Unterschied der Periode t zum Einfluss der Basisperiode auf y. Sind die Koeffizientenschätzer nicht signifikant von 0 verschieden (t-Test), so liegt kein Zeiteffekt gegenüber der Basisperiode vor. Mit einem F-Test kann geprüft werden, ob überhaupt statistisch signifikante unbeobachtete Zeiteffekte vorliegen (Vergleich der Residuenquadratsumme des Modells mit Zeitdummies und der des Modells ohne Zeitdummies).

3 ZEITINVARIANTE UNBEOBACHTETE INDIVIDUELLE HETEROGENITÄT

3.1 Fixed-Effects-Modell

Bei den fixed effects Panelmodellen wird davon ausgegangen, dass die Differenzen zwischen den Mikroeinheiten mit Differenzen im konstanten Term eingefangen werden können. Bei Paneldaten können unbeobachtete individuelle Eigenschaften erfasst werden, wenn diese über t konstant sind. α_i erfasst dabei die gesamte Heterogenität von Individuum i. Wird zugelassen wird, dass

x_{ikt} und α_i korrelieren, d.h. es gibt einen individuenspezifischen Zusammenhang von α_i und allen x_{ikt} (t=1,…,T), dann spricht man vom Fixed-Effects-Modell (FEM).

Fixed Effects Modell (FEM):

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \qquad \varepsilon_{it} \text{ genügt den CLR Anforderungen.}$$

Eine Zurechnung von α_i zum Störterm und eine einfache OLS-Schätzung führen in diesem Fall zu inkonsistenten Schätzern. Denkbar wäre eine Schätzung, bei der α_i als individuenspezifisches Absolutglied definiert wird (vgl. oben, Erfassung von Zeiteffekten durch Dummyvariablen). Dafür müsste das Modell um n-1 Dummies ergänzt werden. Da Paneldaten üblicherweise sehr viele Individuen umfassen, müssten folglich auch sehr viele zusätzliche Variablen in den Ansatz mit aufgenommen werden (Problem der Freiheitsgrade).

Eine andere Möglichkeit den Individualeffekt zu eliminieren, ist die einfache Differenzenbildung.

$$\triangle y_{it} = y_{it} - y_{i,t-1} = (\mathbf{x}_{it} - \mathbf{x}_{i,t-1})'\boldsymbol{\beta} + (\varepsilon_{it} - \varepsilon_{i,t-1}) = \triangle\mathbf{x}'_{it}\boldsymbol{\beta} + \triangle\varepsilon_{it}$$

oder Differenzenbildung zwischen Individual- und Durchschnittswert:

$$y^W_{it} = y_{it} - \overline{y}_i = (\mathbf{x}_{it} - \overline{\mathbf{x}}_i)'\boldsymbol{\beta} + (\varepsilon_{it} - \overline{\varepsilon}_i) =: \mathbf{x}^{W\prime}_{it}\boldsymbol{\beta} + \varepsilon^W_{it}$$

(Within-Transformation, Differenz aus gepooltem und Between-Modell; Between-Modell siehe unten).

Beide Ansätze haben den Nachteil, dass durch die Differenzenbildung nicht nur α_i, sondern auch zeitinvariante Regressoren entfernt werden. Der OLS Schätzer der Within-Transformation heißt Within- oder Fixed-Effects-Schätzer ($\hat{\boldsymbol{\beta}}^{FE}$ oder $\hat{\boldsymbol{\beta}}^{W}$) und ist bei strikter Exogenität der Regressoren unverzerrt.

Die Schätzung von σ^2_ε erfolgt durch

$$\hat{\sigma}^2_\varepsilon = \frac{1}{nT - K - (n-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} (\hat{\varepsilon}^W_{it})^2$$

Die Individualeffekte können mit Hilfe der Fixed-Effects-Schätzer berechnet werden.

$$\hat{\alpha}_i = (\overline{y}_i - \overline{y}) - (\overline{x}_i - \overline{x})'\hat{\beta}^{FE}$$

Mit Hilfe eines F-Tests lässt sich prüfen, ob überhaupt Individualeffekte vorliegen:

$$F = \frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} - \hat{\boldsymbol{\varepsilon}}'_W\hat{\boldsymbol{\varepsilon}}_W}{\hat{\boldsymbol{\varepsilon}}'_W\hat{\boldsymbol{\varepsilon}}_W} \cdot \frac{nT - K - (n-1)}{n-1} \sim F^{n-1}_{nT-K-(n-1)}$$

3.2 Random-Effects-Modell

Die random effects Modelle zeichnen sich dadurch aus, dass die individuenspezifischen konstanten Terme zufällig (randomly) über die Mikroeinheiten verteilt sind. Wird angenommen, dass alle x_{ikt} und α_i unkorreliert sind, kann α_i über den Störterm erfasst werden; in diesem Fall spricht man von einem Random-Effects-Modell (REM).

Random Effects-Modell (REM):

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + u_{it} \qquad \text{mit } u_{it} = \alpha_i + \varepsilon_{it} \qquad \text{und } \alpha_i \sim N(0, \sigma_\alpha^2).$$

Auch hier ist keine OLS-Schätzung anzuwenden, da sie die serielle Korrelation des Störterms u_{it} vernachlässigen würde. Eine Rückführung auf einen klassischen Störterm kann durch folgende Transformation erreicht werden:

$$u_{it} - \delta \bar{u}_i$$

$$\text{mit } \delta = 1 - \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2}}$$

Das gesamte Modell wird dadurch folgendermaßen transformiert:

$$y_{it} - \delta \bar{y}_i = \beta_o(1-\delta) + \beta_1(x_{i1t} - \delta \bar{x}_{i1}) + \cdots + \beta_K(x_{iKt} - \delta \bar{x}_{iK}) + u_{it} - \delta \bar{u}_i$$

Da δ unbekannt ist, müssen $\hat{\sigma}_\varepsilon^2$ und $\hat{\sigma}_\alpha^2$ geschätzt werden. Wenn sich $\hat{\sigma}_\varepsilon^2$ bestimmen lässt kann über das „Between-Modell" $\bar{y}_i = \sum_{k=1}^{K} \bar{x}_{ik}\beta_k + \bar{u}_i$ durch OLS-Schätzung $\hat{\bar{u}}_i$ berechnet werden. Damit lässt sich nun σ_α^2 schätzen:

$$\hat{\sigma}_\alpha^2 = \frac{1}{n-K}\sum_{i=1}^{n} \hat{\bar{u}}_i^2 - \frac{1}{T}\hat{\sigma}_\varepsilon^2$$

Die OLS-Schätzungen des transformierten Modells heißen REM-Schätzungen ($\hat{\boldsymbol{\beta}}^{RE}$). Ob unbeobachtete Heterogenität vorliegt wird durch folgenden Test geprüft:

$$H_0: \sigma_\alpha^2 = 0 \qquad H_1: \sigma_\alpha^2 \neq 0$$

Die Teststatistik lautet:

$$TS = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} \hat{u}_{it}\hat{u}_{is}}{\left[\sum_{i=1}^{n}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} \hat{u}_{it}\hat{u}_{is}\right)^2\right]^{1/2}}$$

Wenn $|TS| > z_{1-\alpha/2} \rightarrow H_0$ ablehnen - das bedeutet, dass unbeobachtete individuelle Heterogenität vorliegt.

3.3 Fixed Effects- oder Random Effects-Schätzer?

Anhand eines Hausman-Tests kann überprüft werden, ob ein FE- oder RE Schätzer vorzuziehen ist. Die Nullhypothese lautet, dass x_i und α_i unkorreliert sind. Die Teststatistik lautet:

$$(\hat{\boldsymbol{\beta}}^{FE} - \hat{\boldsymbol{\beta}}^{RE})'[V(\hat{\boldsymbol{\beta}}^{FE}) - V(\hat{\boldsymbol{\beta}}^{RE})]^{-1}(\hat{\boldsymbol{\beta}}^{FE} - \hat{\boldsymbol{\beta}}^{RE}) \sim \chi_{K-1}^2$$

Übersteigt die Teststatistik den kritischen Wert der χ^2-Verteilung, ist H_0 abzulehnen und der FE-Schätzer vorzuziehen.

Literatur: Hübler (2005) S.268 ff., Hübler (1990), Greene (2003) Chapter 13, Hsiao (2003).

XIII Weitere Topics

1 Systeme von Regressionsgleichungen (SUR)

2 Systeme simultaner Gleichungen

3 Modelle mit diskret abhängigen Variablen - Logit, Probit

4 Modelle mit beschränkt abhängigen Variablen (LDV) – Tobit, Heckman

5 Zeitreihenanalyse, Kointegration

ANHANG 1: MATRIX-ALGEBRA: INVERSE UND DETERMINANTE

Bisheriges zentrales Ergebnis:

OLS-Schätzer: $\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ als Lösung des linearen Gleichungssystems

$(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}$ $\qquad (\mathbf{Ab} = \mathbf{c})$

$\hat{\mathbf{b}}_{(K+1)\times 1} = (\mathbf{X}'\mathbf{X})^{-1}_{(K+1)\times(K+1)}\mathbf{X}'\mathbf{y}_{(K+1)\times 1}$ $\qquad (\mathbf{b} = \mathbf{A}^{-1}\mathbf{c})$

Lösungsvoraussetzung: $\quad \mathrm{rg}(\mathbf{X}) = \mathrm{rg}(\mathbf{X}'\mathbf{X}) = K+1 \leq n$,

d.h. die Anzahl der linear unabhängigen Variablen muss kleiner gleich der Fallzahl sein.

Ist rg(X)=n, dann liegen alle Beobachtungen genau auf der Hyperebene.

Lösungsverfahren linearer Gleichungssysteme:

1. Gauß'sches Eliminationsverfahren ("Treppensystem")
2. Vollständige Elimination mit Gauß
3. Cramer'sche Regel
4. Über die Inverse

Im Folgenden werden nur die Ergebnisse angesprochen. Details mit Beispielen z.B. in Merz 1990 oder in der "Matrix-Algebra"-Sektion eines Ökonometriebuches.

Gegeben: lineares Gleichungssystem

$$\mathbf{A}_{m\times m}\mathbf{b}_{m\times 1} = \mathbf{c}_{m\times 1}$$

1. Gauß'sches Eliminationsverfahren nach Triangulation

Die Lösung b von Ab = c ergibt sich durch schrittweise Substitution aus dem durch Gauß (elementare Zeilentransformation) gebildeten Treppensystemen.

$$\mathbf{A}^*\mathbf{b} = \mathbf{c}^*$$

$$\begin{pmatrix} 1b_1 & + & a^*_{12}b_2 & + & \cdots & + & a^*_{1m}b_m & = & c^*_1 \\ & & 1b_2 & + & \cdots & + & a^*_{2m}b_m & = & c^*_2 \\ & & & & & & \vdots & \vdots & \vdots \\ & & & & & & 1b_m & = & c^*_m \end{pmatrix}$$

2. Vollständige Elimination mit Gauß

Lösung b von **Ab** = **c** durch vollständige Elimination

$$(\mathbf{A},\mathbf{c}) \xrightarrow{\text{Gauß}} (\mathbf{I},\mathbf{b}) \qquad \mathbf{I} = \text{Einheitsmatrix}$$

3. Cramer'sche Regel

Ersetzen der Spalte k von **A** durch **c**:

$$\mathbf{A}_k = (A_1 \ldots A_{\cdot(j-1)}\ c\ A_{\cdot(j+1)} \ldots A_m)$$

$$b_j = \frac{\det \mathbf{A}_j}{\det \mathbf{A}}$$

Berechnung einer Determinanten: $\det \mathbf{A} = |\mathbf{A}|$

- Sarrus'sche Regel (m≤3):

$$m = 1 \quad |\mathbf{A}| = a_{11}$$

$$m = 2 \quad |\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$$m = 3 \quad |\mathbf{A}| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

- Nach Jacoby-Cauchy:

$$\sum_{P(j)} (-1)^{I(j)} a_{1j1} \cdot a_{2j2} \cdot \ldots a_{mjm} \qquad \text{P = Permutation, I = Inversion}$$

- Nach dem Laplace'schen Entwicklungssatz

$$|\mathbf{A}| = \sum_{k}^{K} a_{ik} \underbrace{(-1)^{i+k} D_{ik}}_{\text{Kofaktor, Adjunkte}} \qquad (D_{ik} \text{ Minor, Unterdeterminante})$$

- Als Produkt der Hauptdiagonalen nach Triangulation (Dreiecksmatrix)

$$|\mathbf{A}| = \prod_{i=1}^{m} a_{ii} \qquad \text{Zeilentausch = Vorzeichenänderung}$$

4. Lösung über die Inverse

$$\mathbf{Ab} = \mathbf{c}$$

$$\underbrace{\mathbf{A}^{-1}\mathbf{A}}_{\mathbf{I}}\mathbf{b} = \mathbf{A}^{-1}\mathbf{c}$$

$$\mathbf{Ib} = \mathbf{A}^{-1}\mathbf{c} \qquad \mathbf{I} = \text{Einheitsmatrix}$$

$$\mathbf{b} = \mathbf{A}^{-1}\mathbf{c} \qquad \text{Voraussetzung:} |\mathbf{A}| \neq 0$$

Inversenbestimmung nach Gauß

$$(\mathbf{A}, \mathbf{I}) \xrightarrow{\text{Gauß}} (\mathbf{I}, \mathbf{A}^{-1})$$

Inversenbestimmung mit Hilfe der Adjunkten

- Berechnung der Determinanten: $|\mathbf{A}|$
- Berechung der Unterdeterminanten D_{ik} (aus der Matrix A ohne i-te Zeile und k-te Spalte)
- Die Adjunkte ergibt sich als: $A_{adj,ik} = (-1)^{i+k} D_{ik}$

$$\text{Beispiel für } \mathbf{A}_{3\times 3} : \mathbf{A}_{adj} = \begin{pmatrix} +D_{11} & -D_{21} & +D_{31} \\ -D_{12} & +D_{22} & -D_{32} \\ +D_{13} & -D_{23} & +D_{33} \end{pmatrix}$$

- Matrix $\mathbf{A}_{adj}$ transponieren (Vertauschen von Zeilen und Spalten)

Berechnung der Determinanten: $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{A}_{\text{adj}}$

Besondere Vereinfachung für m = 2:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \rightarrow \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

ANHANG 2: MATRIZENRECHNUNG

Typen von Matrizen:

Matrix: $\mathbf{A}_{m\,n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

Zeilenvektoren: $\mathbf{a}_i' = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$

Spaltenvektoren: $\mathbf{a}_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{bmatrix}$

Quadratische Matrix: $\mathbf{A}_{m\,n} \quad \forall\ m=n$

Symmetrische Matrix: $\mathbf{A} = \mathbf{A}' \Leftrightarrow a_{ik} = a_{ki} \quad \forall\ i, k$

Nullmatrix: $\mathbf{A} = \mathbf{0} \Leftrightarrow a_{ik} = 0 \quad \forall\ i, k$

Einheitsmatrix: $\mathbf{A} = \mathbf{I} \Leftrightarrow \mathbf{A}=\mathbf{A}'$ und $a_{ik} = \begin{cases} 1 & \forall\ i = k \\ 0 & \forall\ i \neq k \end{cases}$

Rechenoperationen mit Matrizen:

Gleichheit von Matrizen $\mathbf{A} = \mathbf{B} \Leftrightarrow a_{ik} = b_{ik} \quad \forall\ i, k$

I. Transponieren: $\mathbf{A}' = \mathbf{B} \Leftrightarrow a_{ik} = b_{ki} \quad \forall\ i, k$

II. Addition: $\mathbf{C} = \mathbf{A} + \mathbf{B} \Leftrightarrow c_{ik} = a_{ik} + b_{ik} \quad \forall\ i, k$

Bedingung: $\mathbf{A}_{m\,n1}, \mathbf{B}_{m2\,n2}\ m_1 = m_2$ und $n_1 = n_2$

Regeln: $\mathbf{A} + \mathbf{0} = \mathbf{A}$

$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$

$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

III: Multiplikation: $\mathbf{C} = \mathbf{AB} \Leftrightarrow c_{ik} = a_i'\, b_k = \sum_j^n a_{ij} b_{jk} \quad \forall\ i, k$

Bedingung: $\mathbf{A}_{m1\,n1}, \mathbf{B}_{m2\,n2}\ n_1 = m_2$

Regeln: $\mathbf{AI} = \mathbf{A}$

$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ [$(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$ usw.]

IV. Inversion $\mathbf{B} = \mathbf{A}^{-1} \quad \Leftrightarrow \quad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

Regel: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Eine Matrix, deren Inverse existiert, heißt nicht-singulär.

V. Rang — Der Rang ($\mathbf{A}$) ist das Minimum aus Zeilen- und Spaltenrang von A, wobei der Zeilenrang die Anzahl der unabhängigen Zeilenvektoren und der Spaltenrang die Anzahl der unabhängigen Spalten bezeichnet.

VI. Spur $\mathrm{tr}(\mathbf{A}) = \sum_i a_{ii}$

Regeln:

$\mathrm{tr}(c\mathbf{A}) = c\ \mathrm{tr}(\mathbf{A})$

$(\mathbf{A}') = \mathrm{tr}\ (\mathbf{A})$

$\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$

$\mathrm{tr}(\mathbf{I}_K) = \mathbf{K}$

$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

VII. Differentiation

Regeln:

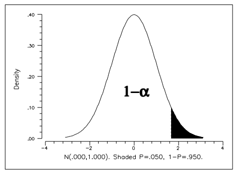
$$\frac{\partial(\mathbf{Ax})}{\partial\mathbf{x}} = \mathbf{A}'$$

$$\frac{\partial(\mathbf{xAx})}{\partial\mathbf{x}} = 2\mathbf{Ax}$$

$$\frac{\partial(\mathbf{x}'\mathbf{Ax})}{\partial\mathbf{A}} = \mathbf{xx}'$$

ANHANG 3: TABELLEN VERTEILUNGSFUNKTIONEN

VERTEILUNGSFUNKTION $\Phi(z)$ DER STANDARDNORMALVERTEILUNG



$\mathbf{N(0,1)}$ (Quelle: Hartung et al. 1982, S. 734)

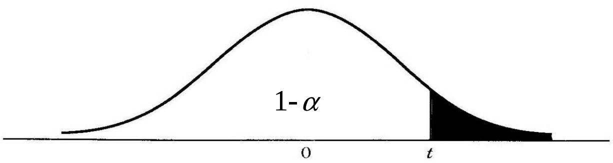
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Ablesebeispiel: $\Phi(1,56) = 0,9406$; $\Phi(z) = 1 - \alpha$

Erweiterung der Tafel: $\Phi(-z) = 1 - \Phi(z)$

VERTEILUNGSFUNKTION DER T-VERTEILUNG

Quelle: J. Schwarze, Grundlagen der Statistik II, 1993, S.308 – 309.

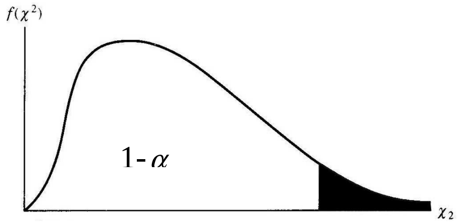


ν	$1-\alpha$: 0,6000	0,7000	0,8000	0,9000	0,9500	0,9750	0,9900	0,9950	0,9990	0,9995
1	0,3249	0,7265	1,3764	3,0777	6,3137	12,7062	31,8205	63,6568	318,3127	636,5894
2	0,2887	0,6172	1,0607	1,8856	2,9200	4,3027	6,9646	9,9248	22,3273	31,5983
3	0,2767	0,5844	0,9785	1,6377	2,3534	3,1825	4,5407	5,8409	10,2146	12,9238
4	0,2707	0,5686	0,9410	1,5332	2,1318	2,7763	3,7470	4,6041	7,1732	8,6102
5	0,2672	0,5594	0,9195	1,4759	2,0150	2,5706	3,3648	4,0322	5,8934	6,8688
6	0,2648	0,5534	0,9057	1,4398	1,9432	2,4469	3,1426	3,7074	5,2076	5,9588
7	0,2632	0,5491	0,8960	1,4149	1,8946	2,3646	2,9979	3,4995	4,7851	5,4079
8	0,2619	0,5459	0,8889	1,3968	1,8595	2,3060	2,8965	3,3554	4,5007	5,0411
9	0,2610	0,5435	0,8834	1,3830	1,8331	2,2622	2,8214	3,2498	4,2968	4,7808
10	0,2602	0,5415	0,8791	1,3722	1,8125	2,2281	2,7638	3,1693	4,1437	4,5868
11	0,2596	0,5399	0,8755	1,3634	1,7959	2,2010	2,7181	3,1058	4,0247	4,4369
12	0,2590	0,5386	0,8726	1,3562	1,7823	2,1788	2,6810	3,0545	3,9296	4,3178
13	0,2586	0,5375	0,8702	1,3502	1,7709	2,1604	2,6503	3,0123	3,8520	4,2208
14	0,2582	0,5366	0,8681	1,3450	1,7613	2,1448	2,6245	2,9768	3,7874	4,1404
15	0,2579	0,5357	0,8662	1,3406	1,7530	2,1314	2,6025	2,9467	3,7328	4,0727
16	0,2576	0,5350	0,8647	1,3368	1,7459	2,1199	2,5835	2,9208	3,6862	4,0150
17	0,2573	0,5344	0,8633	1,3334	1,7396	2,1098	2,5669	2,8982	3,6458	3,9651
18	0,2571	0,5338	0,8620	1,3304	1,7341	2,1009	2,5524	2,8784	3,6105	3,9216
19	0,2569	0,5333	0,8610	1,3277	1,7291	2,0930	2,5395	2,8609	3,5794	3,8834
20	0,2567	0,5329	0,8600	1,3253	1,7247	2,0860	2,5280	2,8453	3,5518	3,8495
21	0,2566	0,5325	0,8591	1,3232	1,7207	2,0796	2,5176	2,8314	3,5272	3,8193
22	0,2564	0,5321	0,8583	1,3212	1,7171	2,0739	2,5083	2,8188	3,5050	3,7921
23	0,2563	0,5317	0,8575	1,3195	1,7139	2,0687	2,4999	2,8073	3,4850	3,7676
24	0,2562	0,5314	0,8569	1,3178	1,7109	2,0639	2,4922	2,7969	3,4668	3,7454
25	0,2561	0,5312	0,8562	1,3163	1,7081	2,0595	2,4851	2,7874	3,4502	3,7251

ν	$1-\alpha$									
	0,6000	0,7000	0,8000	0,9000	0,9500	0,9750	0,9900	0,9950	0,9990	0,9995
26	0,2560	0,5309	0,8557	1,3150	1,7056	2,0555	2,4786	2,7787	3,4350	3,7066
27	0,2559	0,5306	0,8551	1,3137	1,7033	2,0518	2,4727	2,7707	3,4210	3,6896
28	0,2558	0,5304	0,8546	1,3125	1,7011	2,0484	2,4671	2,7633	3,4082	3,6739
29	0,2557	0,5302	0,8542	1,3114	1,6991	2,0452	2,4620	2,7564	3,3962	3,6594
30	0,2556	0,5300	0,8538	1,3104	1,6973	2,0423	2,4573	2,7500	3,3852	3,6459
31	0,2555	0,5298	0,8534	1,3095	1,6955	2,0395	2,4528	2,7440	3,3749	3,6334
32	0,2555	0,5297	0,8530	1,3086	1,6939	2,0369	2,4487	2,7385	3,3653	3,6218
33	0,2554	0,5295	0,8526	1,3077	1,6924	2,0345	2,4448	2,7333	3,3563	3,6109
34	0,2553	0,5294	0,8523	1,3070	1,6909	2,0322	2,4411	2,7284	3,3479	3,6007
35	0,2553	0,5292	0,8520	1,3062	1,6896	2,0301	2,4377	2,7238	3,3400	3,5911
36	0,2552	0,5291	0,8517	1,3055	1,6883	2,0281	2,4345	2,7195	3,3326	3,5821
37	0,2552	0,5289	0,8514	1,3049	1,6871	2,0262	2,4314	2,7154	3,3256	3,5737
38	0,2551	0,5288	0,8512	1,3042	1,6860	2,0244	2,4286	2,7116	3,3190	3,5657
39	0,2551	0,5287	0,8509	1,3036	1,6849	2,0227	2,4258	2,7079	3,3128	3,5581
40	0,2550	0,5286	0,8507	1,3031	1,6839	2,0211	2,4233	2,7045	3,3069	3,5509
41	0,2550	0,5285	0,8505	1,3025	1,6829	2,0195	2,4208	2,7012	3,3013	3,5442
42	0,2550	0,5284	0,8503	1,3020	1,6820	2,0181	2,4185	2,6981	3,2960	3,5377
43	0,2549	0,5283	0,8501	1,3016	1,6811	2,0167	2,4163	2,6951	3,2909	3,5316
44	0,2549	0,5282	0,8499	1,3011	1,6802	2,0154	2,4141	2,6923	3,2861	3,5258
45	0,2549	0,5281	0,8497	1,3006	1,6794	2,0141	2,4121	2,6896	3,2815	3,5202
46	0,2548	0,5281	0,8495	1,3002	1,6787	2,0129	2,4102	2,6870	3,2771	3,5149
47	0,2548	0,5280	0,8493	1,2998	1,6779	2,0117	2,4083	2,6846	3,2729	3,5099
48	0,2548	0,5279	0,8492	1,2994	1,6772	2,0106	2,4066	2,6822	3,2689	3,5051
49	0,2547	0,5278	0,8490	1,2991	1,6765	2,0096	2,4049	2,6800	3,2651	3,5004
50	0,2547	0,5278	0,8489	1,2987	1,6759	2,0086	2,4033	2,6778	3,2614	3,4960
60	0,2545	0,5272	0,8477	1,2958	1,6706	2,0003	2,3901	2,6603	3,2317	3,4602
70	0,2543	0,5268	0,8468	1,2938	1,6669	1,9944	2,3808	2,6479	3,2108	3,4350
80	0,2542	0,5265	0,8461	1,2922	1,6641	1,9901	2,3739	2,6387	3,1953	3,4163
90	0,2541	0,5263	0,8456	1,2910	1,6620	1,9867	2,3685	2,6316	3,1833	3,4019
100	0,2540	0,5261	0,8452	1,2901	1,6602	1,9840	2,3642	2,6259	3,1737	3,3905

VERTEILUNGSFUNKTION DER χ^2-VERTEILUNG

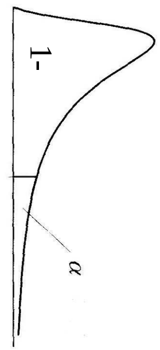
für $P(0 < \chi^2 \le \chi^2_\alpha) = 1 - \alpha$



α	0,5	0,25	0,20	0,1	0,05	0,025	0,01	0,005	0,001
Freiheitsgrade									
1	0,455	1,323	1,642	2,706	3,841	5,024	6,635	7,879	10,827
2	1,386	2,773	3,219	4,605	5,991	7,378	9,210	10,597	13,815
3	2,366	4,108	4,642	6,251	7,815	9,348	11,345	12,838	16,266
4	3,357	5,385	5,989	7,779	9,488	11,143	13,277	14,860	18,466
5	4,351	6,626	7,289	9,236	11,070	12,832	15,086	16,750	20,515
6	5,348	7,841	8,558	10,645	12,592	14,449	16,812	18,548	22,457
7	6,346	9,037	9,803	12,017	14,067	16,013	18,475	20,278	24,321
8	7,344	10,219	11,030	13,362	15,507	17,535	20,090	21,955	26,124
9	8,343	11,389	12,242	14,684	16,919	19,023	21,666	23,589	27,877
10	9,342	12,549	13,442	15,987	18,307	20,483	23,209	25,188	29,588
11	10,341	13,701	14,631	17,275	19,675	21,920	24,725	26,757	31,264
12	11,340	14,845	15,812	18,549	21,026	23,337	26,217	28,300	32,909
13	12,340	15,984	16,985	19,812	22,362	24,736	27,688	29,819	34,527
14	13,339	17,117	18,151	21,064	23,685	26,119	29,141	31,319	36,124
15	14,339	18,245	19,311	22,307	24,996	27,488	30,578	32,801	37,698
16	15,338	19,369	20,465	23,542	26,296	28,845	32,000	34,267	39,252
18	17,338	21,605	22,760	25,989	28,869	31,526	34,805	37,156	42,312
20	19,337	23,828	25,038	28,412	31,410	34,170	37,566	39,997	45,314
22	21,337	26,039	27,301	30,813	33,924	36,781	40,289	42,796	48,268
24	23,337	28,241	29,553	33,196	36,415	39,364	42,980	45,558	51,179
26	25,336	30,435	31,795	35,563	38,885	41,923	45,642	48,290	54,051
28	27,336	32,620	34,027	37,916	41,337	44,461	48,278	50,994	56,892
30	29,336	34,800	36,250	40,256	43,773	46,979	50,892	53,672	59,702
32	31,336	36,973	38,466	42,585	46,194	49,480	53,486	56,328	62,487
34	33,336	39,141	40,676	44,903	48,602	51,966	56,061	58,964	65,247
36	35,336	41,304	42,879	47,212	50,998	54,437	58,619	61,581	67,985
38	37,335	43,462	45,076	49,513	53,384	56,895	61,162	64,181	70,704
40	39,335	45,616	47,269	51,805	55,758	59,342	63,691	66,766	73,403
42	41,335	47,766	49,456	54,090	58,124	61,777	66,206	69,336	76,084
44	43,335	49,913	51,639	56,369	60,481	64,201	68,710	71,892	78,749
46	45,335	52,056	53,818	58,641	62,830	66,616	71,201	74,437	81,400
48	47,335	54,196	55,993	60,907	65,171	69,023	73,683	76,969	84,037
50	49,335	56,334	58,164	63,167	67,505	71,420	76,154	79,490	86,660
55	54,335	61,665	63,577	68,796	73,311	77,380	82,292	85,749	93,167
60	59,335	66,981	68,972	74,397	79,082	83,298	88,379	91,952	99,608
70	69,334	77,577	79,715	85,527	90,531	95,023	100,425	104,215	112,317
80	79,334	88,130	90,405	96,578	101,879	106,629	112,329	116,321	124,839
90	89,334	98,650	101,054	107,565	113,145	118,136	124,116	128,299	137,208
100	99,334	109,141	111,667	118,498	124,342	129,561	135,807	140,170	149,449
120	119,334	130,055	132,806	140,233	146,567	152,211	158,950	163,648	173,618
200	199,334	213,102	216,609	226,021	233,994	241,058	249,445	255,264	267,539

F-VERTEILUNGSTABELLE

Quelle: J. Schwarze, Grundlagen der Statistik II, 1993, S. 301 – 305.



$1-\alpha = 0{,}900$

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	39,86	49,50	53,59	55,83	57,24	58,20	59,14	59,69	60,12	60,46	60,98	61,36	61,64	61,86	62,04	62,58	62,85	63,01	63,33	63,33
2	8,526	9,000	9,162	9,244	9,293	9,326	9,350	9,367	9,381	9,392	9,409	9,421	9,430	9,437	9,442	9,459	9,467	9,472	9,482	9,491
3	5,538	5,462	5,390	5,343	5,309	5,285	5,266	5,251	5,240	5,230	5,215	5,204	5,196	5,189	5,184	5,167	5,159	5,154	5,143	5,134
4	4,545	4,325	4,191	4,107	4,051	4,010	3,979	3,955	3,936	3,920	3,895	3,877	3,864	3,853	3,844	3,817	3,803	3,795	3,777	3,761
5	4,060	3,780	3,619	3,520	3,453	3,404	3,368	3,339	3,316	3,297	3,268	3,247	3,230	3,217	3,206	3,174	3,157	3,147	3,126	3,105
6	3,776	3,463	3,289	3,181	3,107	3,055	3,014	2,983	2,958	2,937	2,905	2,881	2,863	2,848	2,836	2,800	2,781	2,770	2,746	2,722
7	3,589	3,257	3,074	2,961	2,883	2,827	2,785	2,752	2,725	2,703	2,668	2,643	2,623	2,607	2,595	2,555	2,535	2,523	2,497	2,471
8	3,458	3,113	2,924	2,806	2,726	2,668	2,624	2,589	2,561	2,538	2,502	2,475	2,454	2,438	2,425	2,383	2,361	2,348	2,321	2,293
9	3,360	3,006	2,813	2,693	2,611	2,551	2,505	2,469	2,440	2,416	2,379	2,351	2,329	2,312	2,298	2,255	2,232	2,218	2,189	2,159
10	3,285	2,924	2,728	2,605	2,522	2,461	2,414	2,377	2,347	2,323	2,284	2,255	2,233	2,215	2,201	2,155	2,132	2,117	2,087	2,055
11	3,225	2,859	2,660	2,536	2,451	2,389	2,342	2,304	2,274	2,248	2,209	2,179	2,156	2,138	2,123	2,076	2,052	2,036	2,005	1,972
12	3,177	2,807	2,606	2,480	2,394	2,331	2,283	2,245	2,214	2,188	2,147	2,117	2,094	2,075	2,060	2,011	1,986	1,970	1,938	1,904
13	3,136	2,763	2,560	2,434	2,347	2,283	2,234	2,195	2,164	2,138	2,097	2,066	2,042	2,023	2,007	1,958	1,931	1,915	1,882	1,846
14	3,102	2,726	2,522	2,395	2,307	2,243	2,193	2,154	2,122	2,095	2,054	2,022	1,998	1,978	1,962	1,912	1,885	1,869	1,834	1,797
15	3,073	2,695	2,490	2,361	2,273	2,208	2,158	2,119	2,086	2,059	2,017	1,985	1,961	1,941	1,924	1,873	1,845	1,828	1,793	1,755
16	3,048	2,668	2,462	2,333	2,244	2,178	2,128	2,088	2,055	2,028	1,985	1,953	1,928	1,908	1,891	1,839	1,811	1,793	1,757	1,718
17	3,026	2,645	2,437	2,308	2,218	2,152	2,102	2,061	2,028	2,001	1,958	1,925	1,900	1,879	1,862	1,809	1,781	1,763	1,726	1,686
18	3,007	2,624	2,416	2,286	2,196	2,130	2,079	2,038	2,005	1,977	1,933	1,900	1,875	1,854	1,837	1,783	1,754	1,736	1,698	1,657
19	2,990	2,605	2,397	2,266	2,176	2,109	2,058	2,017	1,984	1,956	1,912	1,878	1,852	1,831	1,814	1,759	1,730	1,711	1,673	1,631
20	2,975	2,589	2,380	2,249	2,158	2,091	2,040	1,999	1,965	1,937	1,892	1,859	1,833	1,811	1,794	1,738	1,708	1,690	1,650	1,607
21	2,961	2,574	2,365	2,233	2,142	2,075	2,023	1,982	1,948	1,920	1,875	1,841	1,815	1,793	1,776	1,719	1,689	1,670	1,630	1,586
22	2,949	2,561	2,351	2,219	2,128	2,060	2,008	1,967	1,933	1,904	1,859	1,825	1,798	1,777	1,759	1,702	1,671	1,652	1,611	1,567
23	2,937	2,549	2,339	2,207	2,115	2,047	1,995	1,953	1,919	1,890	1,845	1,811	1,784	1,762	1,744	1,686	1,655	1,636	1,594	1,549
24	2,927	2,538	2,327	2,195	2,103	2,035	1,983	1,941	1,906	1,877	1,832	1,797	1,770	1,748	1,730	1,672	1,641	1,621	1,579	1,533
25	2,918	2,528	2,317	2,184	2,092	2,024	1,971	1,929	1,895	1,866	1,820	1,785	1,758	1,736	1,718	1,659	1,627	1,607	1,565	1,518
26	2,909	2,519	2,307	2,174	2,082	2,014	1,961	1,919	1,884	1,855	1,809	1,774	1,747	1,724	1,706	1,647	1,615	1,594	1,551	1,504
27	2,901	2,511	2,299	2,165	2,073	2,005	1,951	1,909	1,874	1,845	1,799	1,764	1,736	1,714	1,695	1,636	1,603	1,583	1,539	1,491
28	2,894	2,503	2,291	2,157	2,064	1,996	1,943	1,900	1,865	1,836	1,790	1,754	1,726	1,704	1,685	1,625	1,592	1,572	1,528	1,478
29	2,887	2,495	2,283	2,149	2,057	1,988	1,934	1,892	1,857	1,827	1,781	1,745	1,717	1,695	1,676	1,616	1,583	1,562	1,517	1,467
30	2,881	2,489	2,276	2,142	2,049	1,980	1,927	1,884	1,849	1,819	1,773	1,737	1,709	1,686	1,667	1,606	1,573	1,552	1,507	1,456
40	2,835	2,440	2,226	2,091	1,997	1,927	1,872	1,829	1,793	1,763	1,715	1,678	1,649	1,625	1,605	1,541	1,506	1,483	1,434	1,377
50	2,809	2,412	2,197	2,061	1,966	1,895	1,840	1,796	1,760	1,729	1,680	1,643	1,613	1,588	1,568	1,502	1,465	1,441	1,388	1,327
60	2,791	2,393	2,177	2,041	1,946	1,875	1,819	1,775	1,738	1,707	1,657	1,619	1,589	1,564	1,543	1,476	1,437	1,413	1,358	1,291
70	2,779	2,380	2,164	2,027	1,931	1,860	1,804	1,760	1,723	1,691	1,641	1,603	1,572	1,547	1,526	1,457	1,418	1,392	1,335	1,265
80	2,769	2,370	2,153	2,016	1,921	1,849	1,793	1,748	1,711	1,680	1,629	1,590	1,559	1,534	1,513	1,443	1,403	1,377	1,318	1,245
90	2,762	2,362	2,146	2,008	1,912	1,841	1,785	1,739	1,702	1,670	1,620	1,581	1,550	1,524	1,503	1,432	1,391	1,365	1,304	1,228
100	2,756	2,356	2,139	2,002	1,906	1,834	1,778	1,732	1,695	1,663	1,612	1,573	1,542	1,516	1,494	1,423	1,382	1,355	1,293	1,214
∞	2,706	2,303	2,084	1,945	1,847	1,774	1,717	1,670	1,632	1,599	1,546	1,505	1,471	1,444	1,421	1,342	1,295	1,263	1,185	1,000

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
										$1-\alpha = 0{,}950$										
1	161,4	199,5	215,8	224,7	230,4	234,2	237,0	239,1	240,8	242,1	244,2	245,6	246,8	247,6	248,3	250,4	251,5	252,1	253,4	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,42	19,43	19,44	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,71	8,69	8,67	8,66	8,62	8,59	8,58	8,54	8,53
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,912	5,873	5,844	5,821	5,803	5,746	5,717	5,699	5,658	5,628
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,678	4,636	4,604	4,579	4,558	4,496	4,464	4,444	4,403	4,365
6	5,987	5,143	4,757	4,534	4,387	4,284	4,206	4,147	4,099	4,060	4,000	3,956	3,922	3,896	3,874	3,808	3,773	3,753	3,711	3,669
7	5,591	4,737	4,347	4,120	3,971	3,866	3,787	3,726	3,677	3,636	3,574	3,529	3,494	3,467	3,444	3,375	3,340	3,318	3,274	3,230
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,284	3,237	3,202	3,173	3,150	3,079	3,042	3,020	2,974	2,928
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,073	3,025	2,989	2,960	2,936	2,864	2,826	2,803	2,755	2,707
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,913	2,865	2,828	2,798	2,774	2,699	2,661	2,637	2,588	2,538
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,788	2,739	2,701	2,671	2,646	2,570	2,531	2,507	2,456	2,404
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,687	2,637	2,599	2,568	2,544	2,466	2,426	2,401	2,350	2,296
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,604	2,554	2,515	2,484	2,459	2,380	2,339	2,314	2,261	2,206
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646	2,602	2,534	2,484	2,445	2,413	2,388	2,308	2,266	2,240	2,187	2,131
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588	2,544	2,475	2,424	2,385	2,353	2,328	2,247	2,204	2,178	2,123	2,066
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538	2,494	2,425	2,373	2,333	2,302	2,276	2,194	2,151	2,124	2,068	2,010
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494	2,450	2,381	2,329	2,289	2,257	2,230	2,148	2,104	2,077	2,020	1,960
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456	2,412	2,342	2,290	2,250	2,217	2,191	2,107	2,063	2,035	1,978	1,917
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423	2,378	2,308	2,256	2,215	2,182	2,155	2,071	2,026	1,999	1,940	1,878
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393	2,348	2,278	2,225	2,184	2,151	2,124	2,039	1,994	1,966	1,907	1,843
21	4,325	3,467	3,072	2,840	2,685	2,573	2,488	2,420	2,366	2,321	2,250	2,197	2,156	2,123	2,096	2,010	1,965	1,936	1,876	1,812
22	4,301	3,443	3,049	2,817	2,661	2,549	2,464	2,396	2,342	2,297	2,226	2,173	2,131	2,098	2,071	1,984	1,938	1,909	1,849	1,783
23	4,279	3,422	3,028	2,796	2,640	2,528	2,442	2,375	2,320	2,275	2,204	2,150	2,109	2,075	2,048	1,961	1,914	1,885	1,823	1,757
24	4,260	3,403	3,009	2,776	2,621	2,508	2,423	2,355	2,300	2,255	2,183	2,130	2,088	2,054	2,027	1,939	1,892	1,863	1,800	1,733
25	4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,282	2,236	2,165	2,111	2,069	2,035	2,007	1,919	1,872	1,842	1,779	1,711
26	4,225	3,369	2,975	2,743	2,587	2,474	2,388	2,321	2,265	2,220	2,148	2,094	2,052	2,018	1,990	1,901	1,853	1,823	1,760	1,691
27	4,210	3,354	2,960	2,728	2,572	2,459	2,373	2,305	2,250	2,204	2,132	2,078	2,036	2,002	1,974	1,884	1,836	1,806	1,742	1,672
28	4,196	3,340	2,947	2,714	2,558	2,445	2,359	2,291	2,236	2,190	2,118	2,064	2,021	1,987	1,959	1,869	1,820	1,790	1,725	1,654
29	4,183	3,327	2,934	2,701	2,545	2,432	2,346	2,278	2,223	2,177	2,104	2,050	2,007	1,973	1,945	1,854	1,806	1,775	1,710	1,638
30	4,171	3,316	2,922	2,690	2,534	2,420	2,334	2,266	2,211	2,165	2,092	2,037	1,995	1,960	1,932	1,841	1,792	1,761	1,695	1,622
40	4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,124	2,077	2,003	1,948	1,904	1,868	1,839	1,744	1,693	1,660	1,589	1,509
50	4,034	3,183	2,790	2,557	2,400	2,286	2,199	2,130	2,073	2,026	1,952	1,895	1,850	1,814	1,784	1,687	1,634	1,599	1,525	1,438
60	4,001	3,150	2,758	2,525	2,368	2,254	2,166	2,097	2,040	1,993	1,917	1,860	1,815	1,778	1,748	1,649	1,594	1,559	1,481	1,389
70	3,978	3,127	2,735	2,503	2,346	2,231	2,143	2,074	2,017	1,969	1,893	1,836	1,790	1,753	1,722	1,622	1,566	1,530	1,450	1,353
80	3,960	3,110	2,719	2,486	2,329	2,214	2,126	2,056	1,999	1,951	1,875	1,817	1,772	1,734	1,703	1,602	1,545	1,508	1,426	1,325
90	3,947	3,097	2,706	2,473	2,316	2,201	2,113	2,043	1,986	1,938	1,861	1,803	1,757	1,720	1,688	1,586	1,528	1,491	1,407	1,302
100	3,936	3,087	2,695	2,463	2,305	2,191	2,102	2,032	1,975	1,927	1,850	1,792	1,746	1,708	1,676	1,573	1,515	1,477	1,392	1,283
∞	3,841	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,880	1,831	1,752	1,692	1,644	1,604	1,571	1,459	1,394	1,350	1,243	1,000

$1-\alpha = 0,975$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	647,5	799,5	864,3	899,8	922,0	937,3	948,5	956,9	963,5	968,9	977,0	982,8	987,2	990,6	993,4	1002	1006	1008	1014	1018
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,44	39,44	39,45	39,46	39,47	39,48	39,98	39,50
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,28	14,23	14,20	14,17	14,08	14,04	14,01	13,96	13,90
4	12,22	10,65	9,979	9,605	9,364	9,197	9,074	8,980	8,905	8,844	8,751	8,684	8,633	8,592	8,560	8,461	8,411	8,381	8,319	8,257
5	10,01	8,43	7,764	7,388	7,146	6,978	6,853	6,757	6,681	6,619	6,525	6,456	6,403	6,362	6,329	6,227	6,175	6,144	6,080	6,015
6	8,813	7,260	6,599	6,227	5,988	5,820	5,695	5,600	5,523	5,461	5,366	5,297	5,244	5,202	5,168	5,065	5,012	4,980	4,912	4,849
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823	4,761	4,666	4,596	4,543	4,501	4,467	4,362	4,309	4,276	4,208	4,142
8	7,571	6,059	5,416	5,053	4,817	4,651	4,528	4,433	4,357	4,295	4,200	4,130	4,076	4,034	3,999	3,894	3,839	3,807	3,738	3,670
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026	3,964	3,868	3,798	3,744	3,701	3,667	3,560	3,505	3,472	3,403	3,333
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779	3,717	3,621	3,550	3,496	3,453	3,418	3,311	3,255	3,221	3,151	3,080
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588	3,526	3,430	3,359	3,304	3,261	3,226	3,117	3,061	3,027	2,956	2,883
12	6,554	5,096	4,474	4,121	3,891	3,728	3,606	3,512	3,436	3,374	3,277	3,206	3,152	3,108	3,073	2,963	2,906	2,871	2,799	2,725
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312	3,250	3,153	3,082	3,027	2,983	2,948	2,837	2,780	2,744	2,671	2,595
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209	3,147	3,050	2,979	2,923	2,879	2,844	2,732	2,674	2,638	2,564	2,487
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123	3,060	2,963	2,891	2,836	2,792	2,756	2,644	2,585	2,549	2,474	2,395
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049	2,986	2,889	2,817	2,761	2,717	2,681	2,568	2,508	2,472	2,396	2,316
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985	2,922	2,825	2,753	2,697	2,652	2,616	2,502	2,442	2,405	2,328	2,247
18	5,978	4,560	3,954	3,608	3,382	3,221	3,100	3,005	2,929	2,866	2,769	2,696	2,640	2,596	2,559	2,444	2,384	2,347	2,269	2,187
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880	2,817	2,720	2,647	2,591	2,546	2,509	2,394	2,333	2,295	2,217	2,133
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837	2,774	2,676	2,603	2,547	2,501	2,464	2,349	2,287	2,249	2,170	2,085
21	5,827	4,420	3,819	3,475	3,250	3,089	2,969	2,874	2,798	2,735	2,637	2,564	2,507	2,462	2,425	2,308	2,246	2,208	2,128	2,042
22	5,786	4,383	3,783	3,440	3,215	3,055	2,934	2,839	2,763	2,700	2,602	2,528	2,472	2,426	2,389	2,272	2,210	2,171	2,090	2,003
23	5,750	4,349	3,750	3,408	3,183	3,023	2,902	2,808	2,731	2,668	2,570	2,497	2,440	2,394	2,357	2,239	2,176	2,137	2,056	1,968
24	5,717	4,319	3,721	3,379	3,155	2,995	2,874	2,779	2,703	2,640	2,541	2,468	2,411	2,365	2,327	2,209	2,146	2,107	2,024	1,935
25	5,686	4,291	3,694	3,353	3,129	2,968	2,848	2,753	2,677	2,613	2,515	2,441	2,384	2,338	2,300	2,182	2,118	2,079	1,996	1,906
26	5,659	4,265	3,670	3,329	3,105	2,945	2,824	2,729	2,653	2,590	2,491	2,417	2,360	2,314	2,276	2,157	2,093	2,053	1,969	1,878
27	5,633	4,242	3,647	3,307	3,083	2,923	2,802	2,707	2,631	2,568	2,469	2,395	2,337	2,291	2,253	2,133	2,069	2,029	1,945	1,853
28	5,610	4,221	3,626	3,286	3,063	2,903	2,782	2,687	2,611	2,547	2,448	2,374	2,317	2,270	2,232	2,112	2,048	2,007	1,922	1,829
29	5,588	4,201	3,607	3,267	3,044	2,884	2,763	2,669	2,592	2,529	2,430	2,355	2,298	2,251	2,213	2,092	2,028	1,987	1,901	1,807
30	5,568	4,182	3,589	3,250	3,026	2,867	2,746	2,651	2,575	2,511	2,412	2,338	2,280	2,233	2,195	2,074	2,009	1,968	1,882	1,787
40	5,424	4,051	3,463	3,126	2,904	2,744	2,624	2,529	2,452	2,388	2,288	2,213	2,154	2,107	2,068	1,943	1,875	1,832	1,741	1,637
50	5,340	3,975	3,390	3,054	2,833	2,673	2,553	2,458	2,381	2,317	2,216	2,140	2,081	2,033	1,993	1,866	1,796	1,752	1,656	1,545
60	5,286	3,925	3,342	3,008	2,786	2,627	2,507	2,412	2,334	2,270	2,169	2,093	2,033	1,985	1,944	1,815	1,744	1,699	1,599	1,482
70	5,247	3,890	3,309	2,975	2,754	2,595	2,474	2,379	2,302	2,237	2,136	2,059	1,999	1,950	1,910	1,779	1,707	1,660	1,558	1,436
80	5,218	3,864	3,284	2,950	2,730	2,571	2,450	2,355	2,277	2,213	2,111	2,035	1,974	1,925	1,884	1,752	1,679	1,632	1,527	1,400
90	5,196	3,844	3,265	2,932	2,711	2,552	2,431	2,336	2,259	2,194	2,092	2,015	1,955	1,905	1,864	1,731	1,657	1,610	1,503	1,371
100	5,179	3,828	3,249	2,917	2,696	2,537	2,417	2,321	2,244	2,179	2,077	2,000	1,939	1,890	1,849	1,715	1,640	1,592	1,483	1,347
∞	5,024	3,689	3,116	2,786	2,567	2,408	2,288	2,192	2,114	2,048	1,945	1,866	1,803	1,751	1,708	1,566	1,484	1,428	1,296	1,000

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
	$1-\alpha = 0{,}990$																			
1	4052	5000	5403	5625	5764	5859	5929	5981	6023	6056	6107	6143	6170	6192	6209	6261	6287	6303	6334	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,75	99,78	99,80	99,83	99,86	99,88	99,89	99,90	99,93	99,95	99,96	99,98	99,50
3	34,11	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,05	26,92	26,83	26,75	26,69	26,50	26,41	26,35	26,24	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,25	14,15	14,08	14,02	13,84	13,75	13,69	13,58	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,888	9,770	9,680	9,610	9,553	9,379	9,291	9,238	9,130	9,020
6	13,74	10,92	9,779	9,148	8,746	8,466	8,260	8,102	7,976	7,874	7,718	7,605	7,519	7,451	7,396	7,229	7,143	7,091	6,987	6,880
7	12,25	9,55	8,451	7,847	7,460	7,191	6,993	6,840	6,719	6,620	6,469	6,359	6,275	6,209	6,155	5,992	5,908	5,858	5,751	5,650
8	11,26	8,65	7,591	7,006	6,632	6,371	6,178	6,029	5,911	5,814	5,667	5,559	5,477	5,412	5,359	5,198	5,116	5,065	4,961	4,859
9	10,56	8,02	6,992	6,422	6,057	5,802	5,613	5,467	5,351	5,257	5,111	5,005	4,924	4,860	4,808	4,649	4,567	4,516	4,414	4,311
10	10,04	7,56	6,552	5,994	5,636	5,386	5,200	5,057	4,942	4,849	4,706	4,601	4,520	4,457	4,405	4,246	4,165	4,115	4,013	3,909
11	9,646	7,206	6,217	5,668	5,316	5,069	4,886	4,744	4,631	4,539	4,397	4,293	4,213	4,150	4,099	3,941	3,859	3,809	3,707	3,602
12	9,330	6,927	5,952	5,412	5,064	4,820	4,639	4,499	4,387	4,296	4,155	4,052	3,972	3,909	3,858	3,701	3,619	3,569	3,466	3,361
13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,191	4,100	3,960	3,857	3,778	3,715	3,664	3,507	3,425	3,375	3,272	3,165
14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	4,030	3,939	3,800	3,697	3,619	3,556	3,505	3,347	3,266	3,215	3,112	3,004
15	8,683	6,359	5,417	4,893	4,556	4,318	4,141	4,004	3,895	3,805	3,666	3,564	3,485	3,423	3,372	3,214	3,132	3,081	2,977	2,868
16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,780	3,691	3,553	3,451	3,372	3,310	3,259	3,101	3,018	2,967	2,863	2,753
17	8,400	6,112	5,185	4,669	4,336	4,101	3,927	3,791	3,682	3,593	3,455	3,353	3,275	3,212	3,161	3,003	2,920	2,869	2,764	2,653
18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,597	3,508	3,371	3,269	3,190	3,128	3,077	2,918	2,835	2,784	2,678	2,566
19	8,185	5,926	5,010	4,500	4,171	3,939	3,765	3,631	3,522	3,434	3,297	3,195	3,116	3,054	3,003	2,844	2,761	2,709	2,602	2,489
20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,457	3,368	3,231	3,130	3,051	2,989	2,938	2,778	2,695	2,643	2,535	2,421
21	8,016	5,780	4,874	4,369	4,042	3,812	3,640	3,506	3,398	3,310	3,173	3,071	2,993	2,931	2,880	2,720	2,636	2,584	2,475	2,360
22	7,945	5,719	4,817	4,313	3,988	3,758	3,587	3,453	3,346	3,258	3,121	3,019	2,941	2,879	2,827	2,667	2,583	2,531	2,422	2,305
23	7,881	5,664	4,765	4,264	3,939	3,710	3,539	3,406	3,299	3,211	3,074	2,973	2,894	2,832	2,780	2,620	2,535	2,483	2,373	2,256
24	7,823	5,614	4,718	4,218	3,895	3,667	3,496	3,363	3,256	3,168	3,032	2,930	2,852	2,789	2,738	2,577	2,492	2,439	2,329	2,211
25	7,770	5,568	4,675	4,177	3,855	3,627	3,457	3,324	3,217	3,129	2,993	2,892	2,813	2,751	2,699	2,538	2,453	2,400	2,289	2,169
26	7,721	5,526	4,636	4,140	3,818	3,591	3,421	3,288	3,182	3,094	2,958	2,857	2,778	2,715	2,664	2,503	2,417	2,364	2,252	2,131
27	7,677	5,488	4,601	4,106	3,785	3,558	3,388	3,256	3,149	3,062	2,926	2,824	2,746	2,683	2,632	2,470	2,384	2,330	2,218	2,097
28	7,636	5,453	4,568	4,074	3,754	3,528	3,358	3,226	3,120	3,032	2,896	2,795	2,716	2,653	2,602	2,440	2,354	2,300	2,187	2,064
29	7,598	5,420	4,538	4,045	3,725	3,499	3,330	3,198	3,092	3,005	2,868	2,767	2,689	2,626	2,574	2,412	2,325	2,271	2,158	2,034
30	7,562	5,390	4,510	4,018	3,699	3,473	3,304	3,173	3,067	2,979	2,843	2,742	2,663	2,600	2,549	2,386	2,299	2,245	2,131	2,006
40	7,314	5,179	4,312	3,828	3,514	3,291	3,124	2,993	2,888	2,801	2,665	2,563	2,484	2,421	2,369	2,203	2,114	2,058	1,938	1,805
50	7,170	5,056	4,199	3,720	3,408	3,186	3,020	2,890	2,785	2,698	2,563	2,461	2,382	2,318	2,265	2,098	2,007	1,949	1,825	1,683
60	7,077	4,977	4,126	3,649	3,339	3,119	2,953	2,823	2,718	2,632	2,496	2,394	2,315	2,251	2,198	2,028	1,936	1,877	1,749	1,601
70	7,011	4,922	4,074	3,600	3,291	3,071	2,906	2,777	2,672	2,585	2,450	2,348	2,268	2,204	2,150	1,980	1,886	1,826	1,695	1,540
80	6,962	4,881	4,036	3,563	3,255	3,036	2,871	2,742	2,637	2,551	2,415	2,313	2,233	2,169	2,115	1,944	1,849	1,788	1,655	1,494
90	6,925	4,849	4,007	3,535	3,228	3,009	2,844	2,715	2,611	2,524	2,389	2,286	2,206	2,142	2,088	1,916	1,820	1,759	1,623	1,457
100	6,895	4,824	3,983	3,513	3,206	2,988	2,823	2,694	2,590	2,503	2,368	2,265	2,185	2,120	2,067	1,893	1,797	1,735	1,598	1,427
∞	6,635	4,605	3,782	3,319	3,017	2,802	2,639	2,511	2,407	2,321	2,185	2,082	2,000	1,934	1,878	1,696	1,592	1,523	1,358	1,000

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
	$1-\alpha = 0,995$																			
1	16210	20000	21615	22500	23056	23437	23715	23926	24091	24225	24427	24572	24682	24768	24836	25044	25149	25211	25338	25465
2	198,5	199,0	199,3	199,5	199,6	199,7	199,7	199,8	199,8	199,8	199,8	199,9	199,9	199,9	199,9	199,9	200,0	200,0	200,0	199.5
3	55,55	49,80	47,46	46,19	45,39	44,84	44,43	44,12	43,88	43,68	43,38	43,17	43,00	42,88	42,77	42,46	42,30	42,21	42,02	41,83
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,51	20,37	20,26	20,17	19,89	19,75	19,67	19,50	19,32
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,21	13,09	12,98	12,90	12,66	12,53	12,45	12,30	12,14
6	18,635	14,544	12,916	12,028	11,464	11,073	10,786	10,566	10,391	10,250	10,034	9,877	9,758	9,664	9,589	9,358	9,241	9,170	9,026	8,879
7	16,236	12,404	10,882	10,050	9,522	9,155	8,885	8,678	8,514	8,380	8,176	8,028	7,915	7,826	7,754	7,534	7,422	7,354	7,210	7,076
8	14,688	11,042	9,596	8,805	8,302	7,952	7,694	7,496	7,339	7,211	7,015	6,872	6,763	6,678	6,608	6,396	6,288	6,222	6,084	5,951
9	13,614	10,107	8,717	7,956	7,471	7,134	6,885	6,693	6,541	6,417	6,227	6,089	5,983	5,899	5,832	5,625	5,519	5,452	5,321	5,188
10	12,826	9,427	8,081	7,343	6,872	6,545	6,302	6,116	5,968	5,847	5,661	5,526	5,422	5,340	5,274	5,071	4,965	4,901	4,771	4,639
11	12,226	8,912	7,600	6,881	6,422	6,102	5,865	5,682	5,537	5,418	5,236	5,103	5,001	4,921	4,855	4,654	4,550	4,487	4,358	4,226
12	11,754	8,510	7,226	6,521	6,071	5,757	5,525	5,345	5,202	5,085	4,906	4,774	4,674	4,594	4,530	4,331	4,228	4,165	4,036	3,904
13	11,373	8,186	6,926	6,233	5,791	5,482	5,253	5,076	4,935	4,820	4,643	4,513	4,413	4,334	4,270	4,073	3,970	3,908	3,779	3,647
14	11,060	7,922	6,680	5,998	5,562	5,257	5,031	4,856	4,717	4,603	4,428	4,299	4,200	4,122	4,058	3,862	3,760	3,697	3,569	3,436
15	10,798	7,701	6,476	5,803	5,372	5,071	4,847	4,674	4,536	4,423	4,250	4,122	4,024	3,946	3,882	3,687	3,585	3,522	3,394	3,260
16	10,575	7,514	6,303	5,638	5,212	4,913	4,692	4,521	4,384	4,272	4,099	3,972	3,875	3,797	3,734	3,539	3,437	3,375	3,246	3,112
17	10,384	7,354	6,155	5,497	5,075	4,779	4,559	4,389	4,253	4,142	3,971	3,844	3,747	3,670	3,607	3,412	3,311	3,248	3,119	2,984
18	10,218	7,215	6,028	5,375	4,956	4,663	4,445	4,276	4,141	4,030	3,860	3,734	3,637	3,560	3,498	3,303	3,201	3,139	3,009	2,873
19	10,073	7,093	5,916	5,268	4,853	4,561	4,345	4,177	4,043	3,933	3,763	3,638	3,541	3,464	3,402	3,208	3,106	3,043	2,913	2,776
20	9,944	6,986	5,818	5,174	4,762	4,472	4,257	4,090	3,956	3,847	3,678	3,553	3,457	3,380	3,318	3,123	3,021	2,959	2,828	2,690
21	9,830	6,891	5,730	5,091	4,681	4,393	4,179	4,013	3,880	3,771	3,602	3,478	3,382	3,305	3,243	3,049	2,947	2,884	2,753	2,614
22	9,727	6,806	5,652	5,017	4,609	4,322	4,109	3,944	3,812	3,703	3,535	3,411	3,315	3,239	3,176	2,982	2,880	2,817	2,685	2,546
23	9,635	6,730	5,582	4,950	4,544	4,259	4,047	3,882	3,750	3,642	3,475	3,351	3,255	3,179	3,116	2,922	2,820	2,756	2,624	2,484
24	9,551	6,661	5,519	4,890	4,486	4,202	3,990	3,826	3,695	3,587	3,420	3,296	3,201	3,125	3,062	2,868	2,765	2,702	2,569	2,428
25	9,475	6,598	5,461	4,835	4,433	4,150	3,939	3,776	3,645	3,537	3,370	3,247	3,151	3,075	3,013	2,819	2,716	2,652	2,519	2,377
26	9,406	6,541	5,409	4,785	4,384	4,103	3,893	3,730	3,599	3,492	3,325	3,202	3,107	3,031	2,968	2,774	2,671	2,607	2,473	2,330
27	9,342	6,488	5,361	4,740	4,340	4,059	3,850	3,687	3,557	3,450	3,284	3,161	3,066	2,990	2,928	2,733	2,630	2,565	2,431	2,287
28	9,284	6,440	5,317	4,698	4,300	4,020	3,811	3,649	3,519	3,412	3,246	3,123	3,028	2,952	2,890	2,695	2,592	2,527	2,392	2,247
29	9,230	6,396	5,276	4,659	4,262	3,983	3,775	3,613	3,483	3,377	3,211	3,088	2,993	2,917	2,855	2,660	2,557	2,492	2,357	2,210
30	9,180	6,355	5,239	4,623	4,228	3,949	3,742	3,580	3,450	3,344	3,179	3,056	2,961	2,885	2,823	2,628	2,524	2,459	2,323	2,176
40	8,828	6,066	4,976	4,374	3,986	3,713	3,509	3,350	3,222	3,117	2,953	2,831	2,737	2,661	2,598	2,401	2,296	2,230	2,088	1,932
50	8,626	5,902	4,826	4,232	3,848	3,578	3,376	3,219	3,092	2,988	2,825	2,703	2,609	2,533	2,470	2,272	2,164	2,097	1,951	1,786
60	8,494	5,795	4,729	4,140	3,760	3,492	3,291	3,134	3,008	2,904	2,742	2,620	2,526	2,450	2,387	2,187	2,079	2,010	1,861	1,689
70	8,402	5,720	4,661	4,076	3,698	3,431	3,232	3,075	2,950	2,846	2,684	2,563	2,468	2,392	2,329	2,128	2,019	1,949	1,797	1,618
80	8,334	5,665	4,611	4,029	3,652	3,387	3,188	3,032	2,907	2,803	2,641	2,520	2,425	2,349	2,286	2,084	1,974	1,903	1,748	1,563
90	8,282	5,623	4,573	3,992	3,617	3,352	3,154	2,999	2,874	2,770	2,609	2,487	2,393	2,316	2,253	2,051	1,939	1,868	1,711	1,520
100	8,240	5,589	4,542	3,963	3,589	3,325	3,127	2,972	2,847	2,744	2,583	2,461	2,367	2,290	2,227	2,024	1,912	1,840	1,681	1,485
∞	7,879	5,298	4,279	3,715	3,350	3,091	2,897	2,744	2,621	2,519	2,358	2,237	2,142	2,064	2,000	1,789	1,669	1,590	1,402	1,000

DURBIN-WATSON-TABELLE

Table 5 Critical Values for the Durbin–Watson Test: 5% Significance Level[a]

	$K=2$		$K=3$		$K=4$		$K=5$		$K=6$		$K=7$		$K=8$		$K=9$		$K=10$		$K=11$	
T	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U
6	0.610	1.400																		
7	0.700	1.356	0.467	1.896																
8	0.763	1.332	0.559	1.777	0.368	2.287														
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588												
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822										
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005								
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149						
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266				
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360		
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	2.071	0.798	2.188	0.723	2.309	0.650	2.431
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.682	2.396
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333

[a] K referes to the number of columns in X, including the constant term.

Table 5 (Continued)

	$K = 2$		$K = 3$		$K = 4$		$K = 5$		$K = 6$		$K = 7$		$K = 8$		$K = 9$		$K = 10$		$K = 11$	
T	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U	d^*_L	d^*_U
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.795	2.281
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	1.979	0.950	2.069	0.885	2.162	0.821	2.257
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.198
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.945	2.149
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874

Quelle: Hansen (1993) S. 436

Griechisches Alphabet

α	Alpha	ν	Ny
β	Beta	ξ	Xi
$\Gamma\,\gamma$	Gamma	$\Pi\,\pi$	Pi
$\Delta\,\delta$	Delta	ρ	Rho
ε	Epsilon	$\Sigma\,\sigma$	Sigma
ζ	Zeta	τ	Tau
η	Eta	χ	Chi
$\Theta\,\theta$	Theta	$\Phi\,\phi$	Phi
κ	Kappa	$\Psi\,\psi$	Psi
$\Lambda\,\lambda$	Lambda	$\Omega\,\omega$	Omega
μ	My		

LITERATURVERZEICHNIS

Amemiya, T. (1985), Advanced Econometrics, Oxford.

Anderson, Davi, R., Sweeney, Dennis, J., Williams, Thomas, A., Freeman, Jim and Essie Shoesmith (2010), Statistics for Business and Economics, Thomson Publisher, London, United Kingdom.

Atkinson, Anthony und Marco Riani (2000), Robust Diagnostic Regression Analysis, Springer Series in Statistics, Springer Verlag, New York Berlin Heidlberg.

Bauer, Th.K., Fertig, M. und Chr.M. Schmidt (2009), Empirische Wirtschaftsforschung – Eine Einführung, Springer-Verlag, berlin Heidelberg.

Berndt, E. R. (1991), The Practice of Econometrics - Classic and Contemporary, Reading (Mass.)/San Juan.

Bleymüller, J., Gehlert, G. und H. Gülicher (1983), Statistik für Wirtschaftswissenschaftler, 3. Auflage, München.

Brown, R.L., Durbin, J. und Evans, J.M. (1975), Techniques for Testing the Constancy of Regression Relationships over Time, Journal of the Royal Statistical Society, Series B, Vol. 37, 149-163.

Cramer, J. S. (1969), Empirical Econometrics, Amsterdam/London.

Cramer, J. S. (1989), Econometric Applications of Maximum Likelihood Methods, Cambridge/Sydney.

Czayka, L. (1978), Erkenntnisprobleme der Ökonometrie, Meisenheim am Glan.

Davidson, R. and J. G. Mac Kinnon (2004), Econometric Theory and Methods, Oxford University Press, Oxford.

Dhrymes, Ph. J. (1970), Econometrics - Statistical Foundations and Applications, New York/London.

Dhrymes, Ph. J. (1994), Topics in Advanced Econometrics, Volume II - Linear and Nonlinear Simultaneous Equations, New York/Budapest.

Diekmann, A. (2007), Empirische Sozialforschung - Grundlagen, Methoden, Anwendungen, Reinbek bei Hamburg.

Dixit, A. K. (1990), Optimization in Economic Theory, Second edition, Oxford.

Eckey, H.-Fr., Kosfeld, R. und Chr. Dreger (2001), Ökonometrie, Grundlagen – Methoden – Beispiele, 2. Auflage, Wiesbaden.

Fahrmeir, L., Kneib, Th. und St. Lang (2009), Regression – Modelle, Methoden und Anwendungen, 2. Auflage, Springer Heidelberg/New York.

Fair, R. (1978), A Theory of Extramarital Affairs, in Journal of Political Economy, 86, 45-61.

Ferrer-i-Carbonell A. and P. Frijters (2004), How important is Methodology for the estimates of determinants of happiness? in: The Economic Journal, Vol. 114, No. 497, 641–659

Fomby, Th. B., Hill, R.C. und St. R. Johnson (1988), Advanced Econometric Methods, New York/Tokyo.

Frohn, J. (1980), Grundausbildung in Ökonometrie, Walter de Gruyter, Berlin/New York.

Galler, H.-P. (1976), Optimale Wirtschaftspolitik mit nichtlinearen ökonometrischen Modellen, Frankfurt.

Godfrey, L. G. (1988), Misspecification Tests in Econometrics, Cambridge/Sydney.

Goldberger, A. S. (1964), Econometric Theory, New York/London/Sydney.

Goldberger, A. S. (1991), A Course in Econometrics, New York.

Goldfeld, S.M. and R.E. Quandt (1965), Some Tests for Homoskedasticity, in: Journal of the American Statistical Association, 60, 539–547

Gourieroux, Chr. und A. Monfort (1995), Statistics and Econometric Models, Vol. 1, General Concepts, Estimation, Prediction and Algorithms, Cambridge University Press, Cambridge.

Greene, W. (2008, 2003, 2000, 1997), Econometric Analysis, Sixth, Fifth, Fourth, Third Edition, New York/London.

Greene, W. (1992), ET - The Econometrics Toolkit, Version 3.0, User's Guide, New York.

Greene, W. (1995), LIMDEP, Version 7.0, New York.

Griffith, W., Hill, R.C. und G. G. Judge, Learning and Practicing Econometrics, New York.

Griliches, Z. und M. D. Intriligator (eds.) (1983), Handbook of Econometrics, Volume 1, Amsterdam/Tokyo

Griliches, Z. und M. D. Intriligator (eds.) (1984), Handbook of Econometrics, Volume 2, Amsterdam/Tokyo.

Griliches, Z. und M. D. Intriligator (eds.) (1986), Handbook of Econometrics, Volume 3, Amsterdam/Tokyo.

Gruber, J. (1997), Ökonometrie, Band 1, Einführung in die multiple Regression und Ökonometrie, WiSt Studienkurs, München.

Gruber, J. (1997), Ökonometrie, Band 2, Ökonometrische Prognose- und Optimierungsmodelle, WiSt Studienkurs, München.

Gujarati, D. (1999), Essentials of Econometrics, Second edition, McGraw-Hill, Boston/Toronto

Gujarati, D. N. (1995), Basic Econometrics, Third edition, New York/Toronto.

Hansen, G. (1993), Quantitative Wirtschaftsforschung, München.

Hill, C., Griffith, W. E. und G. Judge (1997), Undergraduate Econometrics, New York.

Hübler, O. (1989), Ökonometrie, UTB-Taschenbücher 1538, Stuttgart/New York.

Hübler, O. (1990), Lineare Paneldatenmodelle mit alternativer Störgrößenstruktur, in: Nakhae, Z.G. und K.H. Vollmer (Hg.), Neuere Entwicklungen in der Angewandten Ökonometrie, Heidelberg, 65-99.

Hübler, O. (2005), Einführung in die empirische Wirtschaftsforschung, Verlag Oldenbourg, Oldenbourg.

Hübler, O. und G. Tsertsvadze (2007), Übungsbuch zur empirischen Wirtschaftsforschung – Aufgaben, Wiederholungsfragen, Ergänzungen und Lösungen, Oldenbourg Wissenschaftsverlag, München.

Hujer, R. und R. Cremer (1978), Methoden der empirischen Wirtschaftsforschung, München.

Intrilligator, M. D. (1978), Econometric Models, Techniques, and Applications, Amsterdam/Oxford.

Intrilligator, M. D. (1983), Handbook of Econometrics, Amsterdam/Oxford.

Johnston, J. (1985), Econometric Methods, Third edition, Auckland/Tokyo.

Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H. und T.-C. Lee (1985), The Theory and Practice of Econometrics, Second edition, New York/Singapore.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lütkepohl, H. und T.-C. Lee (1988), Introduction to the Theory and Practice of Econometrics, Second edition, New York/Singapore.

Kennedy, P. (2003), A Guide to Econometrics, Fifth edition, Malden/USA.

Keynes, J. M. (1936), Allgemeine Theorie der Beschäftigung, des Zinses und des Geldes, München.

Kohler, Ulrich und Frauke Kreuter (2008), Datenanalyse mit Stata, Kapitel 8.3 Regressionsdiagnostik, Oldenbourg Veerlag, München Wien

Krämer, W., Sonnberger, H., Maurer, J. und P. Havlik (1985), Diagnostic Checking in Practice, in: The Review of Economics and Statistics, Vol. LXVII,/1, 118-123.

Maddala, G. S. (1977), Econometrics, Tokyo/Sydney.

Malinvaud, E. (1980), Statistical Methods in Econometrics, Third edition, American Elsevier, New York.

Merz, J (1980), Prognosegüte und Spektraleigenschaften ökonomischer Modelle, in: Stöppler, S. (Hg.), Dynamische ökonomische Systeme - Analyse und Steuerung, 31-66.

Merz, J. (1990), Mathematik II für Wirtschaftswissenschaftler (Lineare Algebra), Skriptum zur Vorlesung, 2. Auflage, Frankfurt/Lüneburg.

Merz, J. (1991), Microsimulation - A Survey of Principles, Developments and Applications, in: International Journal of Forecasting, 7, 77-104.

Merz, J. (1993), Microsimulation as an Instrument to Evaluate Economic and Social Programmes, FFB-Discussion Paper No. 5, Department of Economic and Social Sciences, University of Lüneburg, Lüneburg.

Merz, J. (2009a), Statistik I - Deskription, Skriptum zur Vorlesung, Lüneburg.

Merz, J. (2009b), Statistik I - Deskription, Übungs- und Klausuraufgaben mit Lösungen, Lüneburg.

Merz, J. (2010a), Statistik II - Wahrscheinlichkeitsrechnung und induktive Statistik, Skriptum zur Vorlesung, Lüneburg.

Merz, J. (2010b), Statistik II - Wahrscheinlichkeitsrechnung und induktive Statistik, Übungs- und Klausuraufgaben mit Lösungen, Lüneburg.

Merz, J. und H. Stolze (2010), FFB e-learning: Lineare Regression - Deskriptives Modell, Lüneburg (www.leuphana.de/ffb)

Merz, J. und H. Stolze (2010), FFB e-learning: Lineare Regression – Stochastisches Modell, Lüneburg (www.leuphana.de/ffb)

Merz, J. und H. Stolze (2010), FFB e-learning: Parametertests, Lüneburg (www.leuphana.de/ffb)

Merz, J., Böhm, P. und D. Burgert (2009), Timing and Fragmentation of Daily Working Hours Arrangements and Income Inequality – An Earnings Treatment Effects Approach with German Time Use Diary Data, in: electronic International Journal of Time Use Research, 6/2, 200-239.

Mittelhammer; R. C., Judge G.J. und J.M. Douglas (2000), Econometric Foundations, Cambridge.

Moosmüller, G. (2004), Methoden der empirischen Wirtschaftsforschung, Pearson Studium, München.

Morgan, M. S. (1992), The History of Econometric Ideas, Cambridge/Sydney.

Orcutt, G., Merz, J. and H. Quinke (Eds.) (1986), Microanalytic Simulation Models to Support Social and Financial Policy, Amsterdam: North Holland.

Orcutt, Guy H. (1957), A New Type of Socio Economic Systems, in: Review of Economics and Statistics, Vol. 58, 773-97.

Patterson, K. (2000), An Introduction to Applied Econometrics – A Time Series Approach, Palgrave, , New York

Peracchi, Fr. (2001), Econometrics, John Wiley & Sons, Chichester/Toronto

Pindyck, R. S. und D. L. Rubinfeld (1981), Econometric and Models - Economic Forecasts, Second edition, New York/Toronto.

Rinne, H. (1995), Taschenbuch der Statistik, Thun/Frankfurt am Main.

Schneeweiß, H. (1990), Ökonometrie, 4. Auflage, Würzburg/Wien.

Schönfeld, P. (1969), Methoden der Ökonometrie, Band I, Berlin/Frankfurt.

Schönfeld, P. (1971), Methoden der Ökonometrie, Band II, Berlin/Frankfurt.

Spahn, P. B., Galler, H. P., Kaiser, H., Kassella, Th. und J. Merz (1992), Mikrosimulation in der Steuerpolitik, Heidelberg.

Spanos, A. (1986), Statistical Foundations of Econometric Modelling, Cambridge/Sydney.

Stock, J. H. and M. W. Watson (2003), Introduction to Econometrics, Boston.

Stöppler, S. (Hg.) (1980), Dynamische ökonomische Systeme - Analyse und Steuerung, 2. Auflage, Wiesbaden.

Studenmund, A.H. (2006), Using Econometrics – A Practical Guide, Fifth Edition, Pearson/Addison Wesley, Boston – Montreal.

Theil, H. (1971), Principles of Econometrics, New York/Singapore.

Urban, D. (1982), Regressionstheorie und Regressionstechnik, Stuttgart.

Verbeek, M. (2004), A Guide to Modern Econometrics, 2. Auflage, John Wiley & Sons, Chichester.

Vinod, H. D. und A. Ullah (1981), Recent Advances in Regression Methods, New York/Basel.

Vogelvang, B. (2005), Econometrics – Theory and Applications with EViews, Pearson/Addison Wesley, Harlow.

von Auer, L. (2003), Ökonometrie – Eine Einführung, Springer-Verlag, Berlin Heidelberg.

Winker, P. (1997), Empirische Wirtschaftsforschung, Springer, Berlin.

Wooldridge, J.M. (2002), Econometric Analysis of Cross Section and Panel Data, The MIT Press, Cambridge, Mass.

Wooldridge, J.M. (2006, 2009), Introductory Econometrics, A Modern Approach, Third, Fourth Edition, Thomson, South-Western, Canada.