Prof. Dr. Wolf-Dieter Heller
Dr. Ruzana Davoian
Lehrstuhl für Ökonometrie und Statistik

Volkswirtschaftslehre III
Einführung in die Ökonometrie
Sommersemester 2013

Übungsaufgaben zur Vorlesung
VWL III - Einführung in die Ökonometrie

Vorlesung: Die Vorlesung VWL III wird als Blockveranstaltung angeboten.

Klausur: Der Termin der Klausur wird in Absprache mit den Teilnehmern der Vorlesung festgelegt. Rechtzeitig vorher findet eine Fragestunde statt. Die Klausur wird ohne Hilfsmittel geschrieben. Zugelassen ist ein Taschenrechner mit gelöschtem Programmspeicher und ohne Grafikfunktionen. Evtl. Hilfestellungen und Tabellen werden zuvor bekannt gegeben. Die Ergebnisse der Klausur werden nur auf der Homepage des Lehrstuhls veröffentlicht. Eine mögliche Nachklausur für Teilnehmer, die die Hauptklausur nicht bestanden haben, findet voraussichtlich mündlich statt.

Sonstige Informationen: Weitere Informationen werden in der Vorlesung sowie auf der Homepage des Lehrstuhls bekannt gegeben.

Übung 1 - bedingte Erwartungswerte

Gegeben sind die relativen Häufigkeiten f_{jk} diskreter Zahlenpaare (s_j, x_k) für Sparraten s_j und Einkommen x_k:

Table 1.1 Joint frequency distribution of X = income and S = savings rate.

s_j $(j=1...9)$	x_k $(k=1...10)$: 0.5	1.5	2.5	3.5	4.5	5.5	6.7	8.8	12.5	17.5
.50	.001	.011	.007	.006	.005	.005	.008	.009	.014	.004
.40	.001	.002	.006	.007	.010	.007	.008	.009	.008	.007
.25	.002	.006	.004	.007	.010	.011	.020	.019	.013	.006
.15	.002	.009	.009	.012	.016	.020	.042	.054	.024	.020
.05	.010	.023	.033	.031	.041	.029	.047	.039	.042	.007
0	.013	.013	.000	.002	.001	.000	.000	.000	.000	.000
−.05	.001	.012	.011	.005	.012	.016	.017	.014	.004	.003
−.18	.002	.008	.013	.006	.009	.008	.008	.008	.006	.002
−.25	.009	.009	.010	.006	.009	.007	.005	.003	.002	.003
f_k	.041	.093	.093	.082	.113	.103	.155	.155	.113	.052

Abbildung 1: Tabelle 1.1 (aus: Arthur S. Goldberger: A Course in Econometrics, Harvard University Press, Cambridge, Massachusetts, London, England, 1997)

- Transformieren Sie exemplarisch die ersten 2 Spalten $(f_{jk})_{j=1,...9}$ relativer Häufigkeiten f_{jk} in Spalten bedingter relativer Häufigkeiten $f_{j|k} = \frac{f_{jk}}{f_k}$.
- Berechnen Sie daraus die bedingten Mittelwerte $\bar{s}_{|k} = \sum_{j=1}^{9} s_j f_{j|k}$ Anmerkung: Die „conditional mean function" ist $g(k) := \bar{s}_{|k} (k = 1, ..., 10)$ (Vgl. Goldberger, Page 7, Figure 1.1).
- Die tatsächlichen Ersparnisse z_{jk} ergeben sich rechnerisch als Produkt von Einkommen und Sparrate ($z_{jk} = s_j \cdot x_k$). Stimmt ihr Mittelwert $\bar{z} = \sum_j \sum_k z_{jk} f_{jk}$ mit dem Produkt $\bar{s} \cdot \bar{x}$ der Randmittelwerte überein? Wann gilt dieser Fall?

Zusatzaufgabe - Das Ziegenproblem

Bei einer Spielshow im Fernsehen kann der Kandidat zwischen drei verschlossenen Türen wählen. Hinter einer Tür befindet sich der Gewinn, hinter den beiden anderen Türen jeweils eine meckernde Ziege.

Zuerst entscheidet sich der Kandidat für eine Tür, anschließend öffnet der Moderator der Sendung mit den Worten „Ich zeige Ihnen mal was" eine der beiden anderen Türen, jedoch wählt er die Tür immer so, dass eine Ziege den Kandidaten anschaut.

Es sind nun also noch zwei Türen verschlossen und der Kandidat erhält die Möglichkeit, seine Tür zu wechseln. Wozu würden Sie ihm raten?

Übung 2 - Wiederholung der linearen Regression

Aus Statistik 1 kennen Sie das Modell $y = ax + b$ der linearen Regression aus einer Grundgesamtheit:

- Stellen Sie die Normalgleichungen aus dem Kleinste-Quadrate-Ansatz auf und leiten sie daraus die Parameter a und b her.

- Zeigen Sie weiterhin, dass die Regressiongerade auf dem Schwerpunkt (d. h. Schnittpunkt von $\bar{x}$ und $\bar{y}$) liegt.
- Was sind Residuen? Berechnen Sie Summe und Mittelwert der Residuen.

Übung 3 - Markov-Ungleichung

Beweisen Sie die Markov-Ungleichung: Ist P eine Wahrscheinlichkeit auf $\mathbb{R}_+$ mit Dichte $\varphi(x)$, so gilt für beliebige Zahlen $k \in \mathbb{R}$:

$$P([k, \infty]) \leq \frac{1}{k} E(X)$$

Setzen Sie weiterhin $Y = (X - \mu)^2$. Welche Ihnen aus Statistik 2 bekannte Abschätzung folgt daraus?

Übung 4 - Kovarianz

Berechnen Sie für zwei reelle Zufallsvariablen X, Y:

$$Cov(a_1 X + a_2 Y + a_3, b_1 X + b_2 Y + b_3)$$

Übung 5 - Dachverteilung

Die gemeinsame Dichte (Vgl. Goldberger, Page 36, Example) ist $\varphi(x, y) = x + y$ auf $[0, 1] \times [0, 1]$. Bestimmen Sie:

- die Randdichten $\varphi(x)$ und $\varphi(y)$,
- die bedingten Dichten $\varphi_1(x|y)$ und $\varphi_2(y|x)$ für $x = 0, 0.5$ und 1,
- die Kovarianz $Cov(x, y)$.
- Skizzieren Sie $\varphi(x, y)$.
- Berechnen Sie weiterhin den bedingten Erwartungswert $E(Y|X)$ und die beste lineare Vorhersage $H(X)(= E^*(Y|X)) = \alpha X + \beta$ und stellen Sie diese in einer Grafik gegenüber.

Zusatzaufgabe - geschwungene Dachverteilung

Die gemeinsame Dichte (Vgl. Goldberger, Page 41 f., Exercises 4) ist $\varphi(x, y) = \frac{3}{11}(x^2 + y)$ auf $[0, 2] \times [0, 1]$. Bestimmen Sie:

- die Randdichten $\varphi(x)$ und $\varphi(y)$,
- die bedingten Dichten $\varphi_1(x|y)$ und $\varphi_2(y|x)$ für $x = 0, 1$ und 2,
- die Kovarianz $Cov(x, y)$,
- $P([0, 0.5]\,|x)$ für $x = 0, 1$ und 2.
- Skizzieren Sie $\varphi(x, y)$.
- Berechnen Sie weiterhin den bedingten Erwartungswert $E(Y|X)$ und die beste lineare Vorhersage $H(X)(= E^*(Y|X)) = \alpha X + \beta$ und stellen Sie diese in einer Grafik gegenüber.

Übung 6 - Inner- und Zwischenklassenvarianz

Übertragen Sie die Ihnen aus der Vorlesung bekannte Varianzzerlegungsformel

$$\sigma_y^2 = E_{(x)}(\sigma_{y|x}^2) + Var_{(x)}(\mu_{y|x})$$

auf folgende Stichprobensituation $(x_i, y_i)_{i=1,\ldots,n}$, in der x_i nur die Werte 1 und 2 annimmt und die Werte y_i zum Wert $x_i = 1$ im Intervall $I_1 = [0, c)$ und die Werte y_i zum Wert $x_i = 2$ im Intervall $I_1 = [c, 2c]$ liegen. (Dies entspricht einer Klassierung aller Beobachtungen y_i in 2 disjunkte Klassen I_1 und I_2.)

Beurteilen Sie danach die übliche Methode, aus einem Histogramm die approximative Varianz $\bar{s}^2$ wie in Statistik I zu bestimmen.

Verdeutlichen Sie sich das Ergebnis anhand des folgenden Beispiels: An der Klausur „VWL IV“ nehmen 10 Studierende, 4 aus dem 4. Semester und 6 aus dem 6. Semester, teil. Folgende Punktzahlen werden erzielt:

- aus dem 4. Semester: 0, 5, 25, 30
- aus dem 6. Semester: 1, 10, 25, 25, 35, 40

Berechnen Sie:

- den Mittelwert der Punktzahlen
  - aller Klausurteilnehmer
  - der Teilnehmer aus dem 4. Semester
  - der Teilnehmer aus dem 6. Semester
- die Varianz der Punktzahlen
- die Innerklassenvarianzen, deren Gewichte und die Zwischenklassenvarianz

Übung 7 - Transformation von Erwartungswert und Varianz

Die Verteilung der Vektoren $x \in \mathbb{R}^3$ habe den Erwartungswert $\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und die Varianz $\Sigma = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 4 \end{pmatrix}$.

Nach der Transformation $y = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 2 & 1 \end{pmatrix} x + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ wird gesucht:

$E(y)$, $Var(y)$, $E(yy^T)$, $E(y^Ty)$ und $Cov(x, y)$.

Übung 8 - Bedingter Median

Zeigen Sie: Ersetzt man zur Bestimmung der Vorhersagefunktion $h(x)$ für y das Problem $min\ E(y - h(x))^2$ durch $min\ E\,|y - h(x)|$, so resultiert der bedingte Median $Q_{0,5|x}$ als Lösung.

Übung 9 - Multivariate Normalverteilung und zugehörige Stichprobenverteilungen

- Wiederholen Sie die wichtigsten Eigenschaften der Normalverteilung $\mathcal{N}(\mu, \sigma^2)$ und der zugehörigen Stichprobenverteilungen χ_n^2 und t_n.
- Diskutieren Sie die bivariate Normalverteilung $\mathcal{N}(\mu, \Sigma)$:
  - Wie verändern sich graphisch die Höhenlinien der Dichte bei Veränderung der Parameter σ_x^2, σ_y^2 und ϱ?
  - Bestimmen Sie den Erwartungswert und die Varianz der bedingten Dichte $\varphi(y|x)$. Was fällt Ihnen auf?
  - Warum stimmen die Graphen der linearen Vorhersagen $H(X)$ und $H(Y)$ nicht überein? Wie lässt sich dieses Problem lösen? Zur Anschauung sei $x \in \mathbb{R}^2$ bivariat normalverteilt mit Parametern $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $\Sigma = \begin{pmatrix} \frac{5}{4} & \frac{-\sqrt{3}}{4} \\ \frac{-\sqrt{3}}{4} & \frac{7}{4} \end{pmatrix}$. Führen Sie eine Hauptachsentransformation durch.
- Diskutieren Sie die allgemeine n-dimensionale Normalverteilung $\mathcal{N}(\mu, \Sigma)$ und die F-Verteilung.

Übung 10 - Vorhersagefehler und lineare Vorhersage im bivariaten Fall

Bestimmen Sie für den Vorhersagefehler $\varepsilon^* = y - (\alpha x + \beta)$ der bivariaten linearen Vorhersage den Erwartungswert $E(\varepsilon^*)$ und die Kovarianz $Cov(\varepsilon^*, \alpha x + \beta)$ sowie die Zerlegung von $Var(y)$ nach Pythagoras.

Übung 11 - KQ-Schätzer

Wiederholen Sie die Herleitung des Schätzers $a \in \mathbb{R}^k$ für α aus dem KQ-Ansatz. Spezialisieren Sie nun diesen Schätzer im klassischen multiplen Regressionsmodell auf den bivariaten Fall der einfachen linearen Regression.

Übung 12 - Klassische lineare Regression

Für eine lineare Regression $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3$ sind aus 5 Beobachtungen folgende Daten bekannt:

$$X = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 1 \\ 3 & 4 & 1 \\ 5 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}, \; y = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 3 \\ 2 \end{pmatrix}$$

Gesucht sind: a, $\hat{y}$, e, P, $spur(P)$, $spur(I-p)$, $(I-P)\cdot P$, $\hat{y}^T e$, $1^T e$, s^2 (erwartungstreuer Schätzer für σ^2), $Var(a)$ (unter Verwendung von s^2), $Var(a_1)$, $Var(a_2)$, $Var(a_3)$ und R^2.

Übung 13 - Erwartungstreue des Kovarianzschätzers

Aus einer Stichprobe $\{(x_i, y_i)\}_{i=1...n}$ unabhängiger Beobachtungen schätzt M_{11} die Kovarianz σ_{xy} der zugrundeliegenden Verteilung auf $\mathbb{R} \times \mathbb{R}$. Ist M_{11} erwartungstreu für σ_{xy}? Sind $\bar{x}$ und $\bar{y}$ asymptotisch unkorreliert?

Übung 14 - Erwartungstreue von $a^T a$

Im klassischen Regressionsmodell habe sich bei einer Stichprobe vom Umfang $n = 40$ die Matrix $X^T X = \begin{pmatrix} 40 & 10 \\ 10 & 5 \end{pmatrix}$ ergeben. Bekannt sind weiter $\sigma^2 = 40$ und $\alpha = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ Bestimmen Sie $E(a^T a)$. Ist $a^T a$ ein erwartungstreuer Schätzer für $\alpha^T \alpha$?

Übung 15 - Mittlerer linearer Prognosefehler

Diskutieren Sie den „mittleren linearen Prognosefehler“ bei der bivariaten Schätzung von $\alpha x + \beta$ anhand einer KQ-besten Stichprobengeraden $\hat{y} = ax + b$. Was bezeichnet man als Interpolationsproblem, was als Extrapolationsproblem?

Übung 16 - Varianz des Prognosefehlers

Bestimmen Sie im Regressionsmodell von Übung 12 den Schätzwert für die Varianz des Prognosefehlers δ, falls $x_{n+1}^T = (3, 2, 1)$ bzw. $(-5, -6, 1)$ „neue Beobachtungen“ darstellen.

Übung 17 - Aufspaltung von Regressionen

1. Vergleichen Sie anhand der exogenen Datenmatrix $X = (x^1, x^2, 1)$ aus Übung 12 die zwei kurzen Regressionen von y auf $X_1 = (x^1, 1)$ und auf $X_2 = (x^2, 1)$ miteinander.

2. Durch die Aufspaltung $X = (X_1 | X_2)$ kann man allgemein zwei „kurze“, komplementäre Regressionen von y auf die $n \times k_1$-Matrix X_1 und auf die $n \times k_2$-Matrix X_2 berechnen ($k_1 + k_2 = k$). Die zugehörigen Regressionskoeffizienten bilden die Vektoren $a^1 \in \mathbb{R}^{k_1}$ und $a^2 \in \mathbb{R}^{k_2}$. Andererseits liefert die Regression von y auf X den Vektor $a \in \mathbb{R}^k$, aufteilbar in $a = \begin{pmatrix} \tilde{a}^1 \\ \tilde{a}^2 \end{pmatrix}$. Unter welcher Bedingung ist $a^1 \equiv \tilde{a}^1$ und $a^2 \equiv \tilde{a}^2$?

Übung 18 - Konfidenzintervallschätzung für den Parametervektor

1. Formulieren Sie für den Anstieg a der besten Geraden durch ein Stichprobenstreudiagramm $\{(x_i, y_i)\}_{i=1\dots n}$ den asymptotischen Intervallschätzer zu gegebenem Konfidenzniveau.

2. Bestimmen Sie Konfidenzintervalle zu den Niveaus 0.9, 0.95 und 0.975 für den Parametervektor α der linearen Regression

   - bei Kenntnis von $X^T X = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$, $\sigma^2 = 7$, $a = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$,
   - bei Kenntnis von $X^T X = \begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix}$, $\sigma^2 = 2$, $a = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$,
   - aus den Daten von Übung 12.

Übung 19 - Tests für den Parametervektor

1. Formulieren Sie für den Anstieg α der besten Geraden durch ein Stichprobenstreudiagramm $\{(x_i, y_i)\}_{i=1\dots n}$ asymptotische $\mathcal{N}$-Tests zu einseitigen und einfachen Nullhypothesen. Welches Interesse hat man z. B. an der einfachen Nullhypothese H_0: $\alpha = 0$?

2. Geben Sie für die einfache Nullhypothese H_0: $\alpha = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ zu den Daten $X^T X = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$, $\sigma^2 = 2$ den kritischen Bereich zum Niveau 0.05 an. Wird H_0 durch die Beobachtung $a = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ abgelehnt?

3. Formulieren Sie zu den Daten $X^T X = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$, $\sigma^2 = 7$ die einfache Nullhypothese H_0: $\alpha_1 = 1$ zum Niveau 0.1 . Wird H_0 durch die Beobachtung $a_1 = 3$ abgelehnt?

4. Testen Sie für den Daten aus Übung 12 auf H_0: $\alpha = \vec{0}$ zum Niveau 0.1.

Übung 20 - Vergleich von kurzer und langer Regression

Vergleichen Sie für die Daten aus Übung 12 die kurze mit der langen Regression, d. h. $X_1 = (x^1, 1)$ mit $X = (x^1, x^2, 1)$. Prüfen Sie die Nullhypothese H_0: $\alpha_2 = 0$ zum Niveau 0.99 .

Übung 21 - Verallgemeinertes klassisches Regressionsmodell

Zeigen Sie: Setzt man im verallgemeinerten klassischen Regressionsmodell speziell $\Sigma = \sigma^2 I_n$, so geht der Aitken-Schätzer a^* über in den klassischen KQ-Schätzer a.

Übung 22 - SUR und FGLS-Schätzer

Bestimmen Sie für die Merkmale $y^{(1)}$ und $y^{(2)}$ entsprechend des Regressionsansatzes $y^{(1)} = \alpha_1 \cdot x^{(1)} + \alpha_2 \cdot x^{(2)} + \alpha_3$ und $y^{(2)} = \beta_2 \cdot x^{(2)} + \beta_3 \cdot x^{(3)} + \beta_4$ die Parameter für unten gegebene Beobachtungen und bestimmen Sie anschließend den „feasible GLS-Schätzer“ für den Parametervektor. Was ist der „iterative GLS-Schätzer“?

$$X = (x^{(1)}, x^{(2)}, 1) \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 1 \\ 3 & 4 & 1 \\ 5 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}, \; y^{(1)} = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 3 \\ 2 \end{pmatrix}$$

$$X = (x^{(2)}, x^{(3)}, 1) \begin{pmatrix} 1 & 2 & 1 \\ 3 & 3 & 1 \\ 4 & 5 & 1 \\ 2 & 6 & 1 \\ 1 & 8 & 1 \end{pmatrix}, \; y^{(2)} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 3 \end{pmatrix}$$

Übung 23 - Kovarianz im Irrfahrt-Prozess

Bestimmen Sie $Cov(X_s, X_t)$ für den Irrfahrt-Prozess. Wie ist $X_t = \sum_{s=1}^{t} Z_s$ für Gaußsches weißes Rauschen, d. h. für $Z_s \sim \mathcal{N}(\mu, \sigma^2)$ verteilt?

Übung 24 - Kovarianzstationarität von MA- und AR-Prozessen

Sind MA- und AR-Prozesse kovarianzstationär? Wie ist jeweils X_t für $Z_s \sim \mathcal{N}(0, \sigma^2)$ verteilt? Ist der MA-Prozess für $p = 1$ aus der Kenntnis von $\xi(h)$ identifizierbar? Verdeutlichen Sie sich dies mittels eines Zahlenbeispiels.

Übung 25 - Koeffizienten von AR-Prozessen

Für die AR-Prozesse 1. und 2. Ordnung sind die Koeffizienten β_j aus einer Trajektorie $(x_1, ..., x_n)$ nach dem KQ-Ansatz zu berechnen. Verdeutlichen Sie sich dies mittels eines Zahlenbeispiels.

Übung 26 - Durbin-Watson-Test

Berechnen Sie im AR(1)-Störprozess $\varepsilon_t = \beta \cdot \varepsilon_{t-1} + u_t$ des verallgemeinerten klassischen Regressionsmodells $y_t = \alpha^T \cdot x_t + \varepsilon_t$ mit $x_t \in \mathbb{R}^4$ $(t = 1, ..., 15)$ aus dem beobachteten Residuenvektor

$$e^T = (-1, -1, -1, 1, 2, 1, 1, 1, 1, 1, -1, -2, -2, 0, 0)$$

den Wert der Teststatistik $d(e)$. Bestätigt e die Nullhypothese „keine Autokorrelation“ zum Niveau $0,05$? Wie wäre das Ergebnis für $x_t \in \mathbb{R}^6$ (6 exogene Variablen inklusive Konstante)?

Übung 27 - ARMA-Prozess

Berechnen Sie für den ARMA (1,1) die Umkehrungen:

- $X_t = P^{-1}(L)Q(L)Z_t$
- $Z_t = Q^{-1}(L)P(L)X_t$

sowie $\varrho(h)$ und $\xi(h)$. Ist der ARMA (1,1) kovarianzstationär? Ist der Irrfahrt-Prozess vom ARMA-Typ?