Prof. Regina T. Riphahn, Ph.D. WS 04/05

Zwischenklausur zur Veranstaltung **"Einführung in die Ökonometrie"**
am 14. Dezember 2004, 9.00-10.00 Uhr

______________________ ______________________ ______________________

Name, Vorname Matr.-Nr. Unterschrift

Erlaubte Hilfsmittel: 2 DIN A4-Seiten eigene Notizen, Taschenrechner, Fremdwörterbuch.
Bearbeitungszeit: 60 Minuten
Die Klausur besteht aus 5 Aufgaben in denen insgesamt 60 Punkte erworben werden können. Die Punktzahl je Aufgabe, bzw. je Teilaufgabe, ist in eckigen Klammern angegeben und entspricht der für die Aufgabe aufzuwendenden Zeit in Minuten.

Das Team der Abteilung Statistik und Ökonometrie wünscht Ihnen viel Erfolg!

Klausurergebnis:

Aufgabe	1	2	3	4	5
Punkte:					

Aufgabe 1: **[30]**

Der folgende EViews-Output liefert Ihnen Informationen über den Zusammenhang zwischen der Anzahl der Verbrechen in einem Gebiet X auf der einen Seite und der Anzahl der Gefangenen, der Anzahl der Polizisten, der Arbeitslosenrate und dem nominalem Einkommen desselben Gebietes X auf der anderen Seite. Dabei wurden folgende Variablen verwendet:

crime:	Anzahl Verbrechen pro 100'000 Einwohner
log(crime):	Natürlicher Logarithmus von crime
pris:	Anzahl Gefangene pro 100'000 Einwohner
log(pris):	Natürlicher Logarithmus von pris
police:	Anzahl Polizisten pro 100'000 Einwohner
unem:	Arbeitslosenrate in Prozent
income:	Pro Kopf Einkommen, nominal
log(income):	Natürlicher Logarithmus von income

Der aus den USA stammende Datensatz beinhaltet 714 Beobachtungen aus 51 Bundesstaaten.

a) Bestimmen Sie den Standardfehler für b_4 und den Standardfehler der Regression. [3]
b) Interpretieren Sie die Koeffizienten b_1, b_2 und b_3, und beschreiben Sie die geschätzten Zusammenhänge zwischen den erklärenden Variablen und der abhängigen Variablen. Erläutern Sie, ob Ihnen die Ergebnisse plausibel erscheinen. [9]
c) Bestimmen Sie das 99% Konfidenzintervall für b_3 (geben Sie die für die Berechnung korrekten Freiheitsgrade an, unterstellen Sie jedoch für Ihre Berechnung $DF = \infty$). Interpretieren Sie Ihr Ergebnis. [5]
d) Nehmen Sie an, dass der im Anschluss einer linearen Regression durchgeführten Jarque Bera Test den Wert 19.459447 ergibt und der entsprechende p-Wert 0.003 ist. Was überprüft dieser Test typischerweise? Zu welchem Schluss kommen Sie aufgrund der genannten Ergebnisse? Erläutern Sie, ob und gegebenenfalls welche Konsequenzen Ihr Ergebnis für Ihre Lösung der Teilaufgabe c) hat. [6]
e) Bestimmen Sie das R^2, wenn Sie wissen, dass $\sum y_t^2 = 10380.2427$. Was sagt es aus? [7]

Dependent Variable: LOG(CRIME)				
Method: Least Squares				
Date: 14/12/04 Time: 08.59				
Sample: 1 714				
Included observations: 714				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C (b_1)	3.817823	0.357300	10.68519	0.0000
LOG(PRIS) (b_2)	0.102309	0.018234	5.610776	0.0000
POLICE (b_3)	0.001335	0.000116	11.53013	0.0000
UNEM (b_4)	-0.483089		-1.128397	0.2595
LOG(INCOME) (b_5)	-0.090287	0.039386	-2.292377	0.0222
R-squared		Mean dependent var		3.803649
Adjusted R-squared		S.D. dependent var		0.265534
S.E. of regression		Akaike info criterion		-0.153690
Sum squared resid	35.35041	Schwarz criterion		-0.121681
Log likelihood	59.86745	F-statistic		74.82040
Durbin-Watson stat	0.264170	Prob(F-statistic)		0.000000

Aufgabe 2: **[7]**

Ihnen liegt ein Datensatz mit Informationen über Pensionäre vor. Sie regressieren die jährlich ausgezahlten Pensionen (P) einer Pensionskasse (gemessen in CHF) auf die Anzahl der Jahre mit Beitragszahlung (BJ) der Versicherten, sowie auf das letzte Jahreseinkommen (JE) (gemessen in CHF), das die Pensionäre verdient haben:

$$P_t = \alpha_0 + \alpha_1 \, BJ_t + \alpha_2 \, JE_t + e_t$$

a. Wie interpretieren Sie den Koeffizienten α_1 wenn er beispielsweise den Wert 2'000.8 annimmt? [2]

b. Ihr Kollege behauptet, dass nicht die Anzahl der Jahre mit Beitragszahlung die relevante Grösse darstellt, sondern vielmehr die Quartale, in denen Beiträge bezahlt wurden. Wie ändert sich der Koeffizient α_1 wenn Sie diese Anpassung in der Variable BJ vornehmen? [2]

c. Sie verwerfen den Vorschlag des Kollegen und legen Ihre ursprünglichen Schätzergebnisse Ihrem Vorgesetzten vor, der daraufhin vorschlägt, die abhängige Variable statt mit ihrem jährlichen mit ihrem monatlichen Wert in der Schätzung zu benutzen. Welche Wirkung hat dies auf die drei geschätzten Koeffizienten? [3]

Aufgabe 3: **[9]**

Sie schätzen ein einfaches lineares Regressionsmodell und betrachten anschliessend die Residuen. Dabei stellen Sie fest, dass die Residuen bei den sehr niedrigen und den sehr hohen Werten der erklärenden Variable negative Werte aufweisen, während sie im mittleren Bereich meist positiv sind.

a. Wie interpretieren Sie diesen Befund? [2]

b. Entspricht dieser Befund den Grundannahmen des einfachen linearen Regressionsmodells bezüglich der Residuen? [3]

c. Welche Möglichkeit sehen Sie, Ihr Regressionsmodell zu verbessern und wie könnte der Erfolg dieser Massnahme gemessen werden? [4]

Aufgabe 4: **[8]**

Der Kleinstquadratschätzer des Modells $y_t = \beta_1 + \beta_2 x_t + e_t$ minimiert eine Funktion der Störterme e_t, die im X-Y Diagramm als *vertikale* Abweichungen von der Regressionsgeraden dargestellt werden. Wie würden Sie die Parameterschätzer für β_1 und β_2 ableiten, wenn Sie stattdessen die *horizontalen* Abweichungen von der Regressionsgerade nutzen sollten? Geben Sie die zu minimierende Zielfunktion S an.

Aufgabe 5: **[6]**

Wahr oder Falsch? Tragen Sie für zutreffende Aussagen den Buchstaben w (für wahr), für nicht zutreffende f (für falsch) ein.

(Für jede richtige Antwort gibt es 0,5 Punkte, für jede falsche Antwort werden 0,5 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.)

	Das einfache lineare Regressionsmodell wird unter der Annahme geschätzt, dass die erklärenden Variablen unkorreliert und normalverteilt sind.
	Bei einem einseitigen t-Test verwerfen wir die Nullhypothese, wenn der p-Wert kleiner als α ist.
	Das angepasste R^2 kann nicht als Anteil der erklärten Varianz interpretiert werden.
	Intervallschätzer von Parametern erlauben Rückschlüsse auf deren statistische Signifikanz.
	Für die Ableitung des Kleinstquadratschätzers ist die Annahme der Normalverteilung unerheblich.
	Wenn Schätzer erwartungstreu sind, stimmt der geschätzte Wert des Parameters mit dem wahren Wert überein.
	Ein Typ II Fehler liegt vor, wenn eine Nullhypothese verworfen wurde, obwohl sie zutrifft.
	Kollineare erklärende Variablen erhöhen die Präzision der Parameterschätzung im multiplen Regressionsmodell.
	Bei einseitigen Tests von H_0: β [r1]$\geq$ c liegt die Verwerfungsregion auf der linken Seite der t-Verteilung.
	Eine hohe Streuung der abhängigen Variable erhöht die Präzision, mit der Parameter des linearen Modells geschätzt werden können.
	Die Anzahl der Freiheitsgrade der Chi^2-Verteilung richtet sich nach der Anzahl der aufsummierten standardnormalverteilten und quadrierten Zufallsvariablen, auf deren Basis die Chi^2 verteilte Zufallsvariable bestimmt wurde.
	R^2 lässt sich im einfachen Regressionsmodell aus dem Stichprobenkorrelationskoeffizienten von y und $\hat{y}$ berechnen.