

Beispiele für Prüfungsfragen Ökonometrie I und II

Von den sieben Beispielen von Prüfungsfragen wären in einer konkreten einstündigen Prüfung etwa deren zwei zu beantworten.

Aufgaben 3, 4, 6 und 7 beziehen sich auf Ökonometrie II

Pro memoria:

F-Test:
$$F = \frac{(RRSS - URSS)/r}{URSS/(n - k')}$$

n = Anzahl Beobachtungen

r = Anzahl Restriktionen

k' = (Anzahl erklärende Variablen + 1)
bzw. Anzahl geschätzte Parameter der
unrestringierten Regression

Durbin-h-Test:
$$h = \hat{\rho} \sqrt{\frac{n}{1 - n s_{\beta}^2}} \quad \text{mit}$$

$\hat{\rho}$: geschätzter Autokorrelationskoeffizient erster Ordnung, der
approximativ als $\hat{\rho} = 1 - \frac{DW}{2}$ berechnet werden kann.

s_{β}^2 : geschätzte Varianz des Koeffizienten vor der verzögerten
endogenen Variablen.

n : Anzahl Beobachtungen.

AUFGABE 1

Sie schätzen eine logarithmische Konsumfunktion für die Schweiz. Der reale Konsum (CONSPR) soll mit dem real verfügbaren Haushaltseinkommen (YDISPR) erklärt werden (Quartalsdaten 1971-1995): $\text{Log}(\text{CONSPR}_t) = \alpha + \beta \text{Log}(\text{YDISPR}_t) + u_t$

Für diese Ausgangsgleichung erhalten Sie:

Sample: 1971:1 1995:4				
Included observations: 100				
LOG(CONSPR) = ALFA + BETA*LOG(YDISPR)				
	Coefficient	Std. Error	t-Statistic	
ALFA	1.797417	0.217328	8.270538	
BETA	0.817537	0.021046	38.84467	
R-squared		0.939113	Adjusted R-squared	0.938391
S.E. of regression		0.025618	Durbin-Watson stat	0.394284

a) Um wieviele Prozent steigt gemäss diesen Schätzergebnissen der Konsum, wenn sich das Einkommen um 1% erhöht? Testen Sie die Hypothese, dass die Einkommenselastizität der Konsumnachfrage gleich 1 ist!

Sie erweitern die Gleichung, indem Sie einen Einfluss der Arbeitslosigkeit (ARBLOS) auf die Konsumnachfrage in Betracht ziehen. Sie erhalten:

Sample: 1971:1 1995:4				
Included observations: 100				
LOG(CONSPR) = ALFA + BETA1*LOG(YDISPR) + BETA2*LOG(ARBLOS)				
	Coefficient	Std. Error	t-Statistic	
ALFA	1.705973	0.237628	7.179175	
BETA1	0.845376	0.026768	31.58159	
BETA2	-0.134531	0.088835	-1.514392	
R-squared		0.939000	Adjusted R-squared	0.938916
S.E. of regression		0.025426	Durbin-Watson stat	1.295285

b) Hat die Arbeitslosigkeit einen signifikanten Einfluss auf die Konsumnachfrage? Gibt es - unter Beachtung aller Teststatistiken - einen Grund, diese Gleichung der Ausgangsgleichung vorzuziehen?

c) Beim oben ausgewiesenen R-Quadrat (0.939000) muss ein Typfehler unterlaufen sein. Warum?

Eine andere Art der Modifikation der Ausgangsgleichung ist ein Anpassungsmodell der Form:

$$\text{Log}(\text{CONSPR}_t) = \lambda\alpha + \lambda\beta \text{Log}(\text{YDISPR}_t) + (1 - \lambda)\text{Log}(\text{CONSPR}_{t-1}) + u_t$$

Sie erhalten dafür die folgenden Schätzergebnisse:

Sample: 1971:1 1995:4			
Included observations: 100			
LOG(CONSPR)=C(1)+C(2)*LOG(YDISPR)+C(3)*LOG(CONSPR(-1))			
	Coefficient	Std. Error	t-Statistic
C(1)	0.260007	0.119256	2.180238
C(2)	0.082359	0.036514	2.255547
C(3)	0.891845	0.062910	12.23214
R-squared	0.988817	Adjusted R-squared	0.988586
S.E. of regression	0.011027	Durbin-Watson stat	1.640435

- d) *Wie gross sind gemäss dieser Schätzgleichung die kurzfristige und die langfristige Einkommenselastizität des Konsums?*
- e) *Testen Sie die Hypothese zeitlich unkorrelierter Störterme mit der Durbin-Watson-Statistik sowie dem Durbin-h-Test, wobei sie den Autokorrelationskoeffizienten 1. Ordnung approximativ aus der Durbin-Watson-Statistik ableiten (vgl. Deckblatt). Welcher der beiden Tests ist im vorliegenden Fall zuverlässiger?*

AUFGABE 2

Zur Schätzung eines Regressionsmodells $y = \alpha + \beta x + u$ verfügen Sie über die folgenden 3 Beobachtungen (was natürlich eine minimale Stichprobe ist):

$$x = 3, 3, 6$$

$$y = 2, 4, 6$$

a) Zeichnen Sie die 3 Beobachtungspunkte in ein x-y-Diagramm ein, überlegen Sie sich, wie die Regressionsgerade aussieht, und berechnen Sie TSS, RSS und R^2 !

b) Testen Sie die Hypothese, dass x keinen Einfluss auf y hat (F-Test, 5%-Signifikanzniveau)!

c) Stellen Sie sich vor, die Stichprobe würde 9 Beobachtungen umfassen, die gerade 3 mal der Stichprobe von oben entsprechen, d.h:

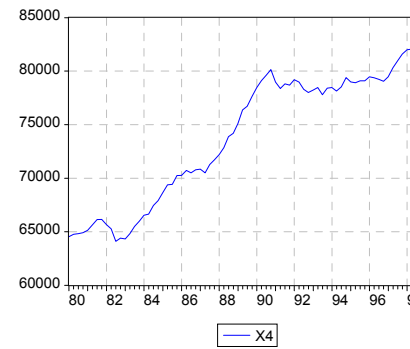
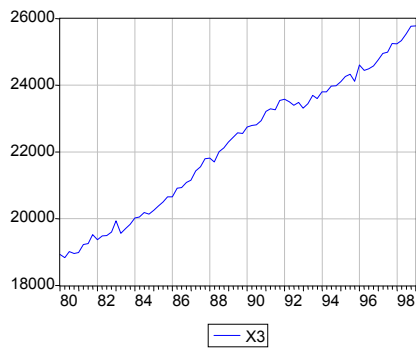
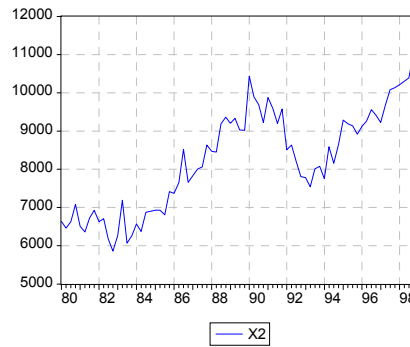
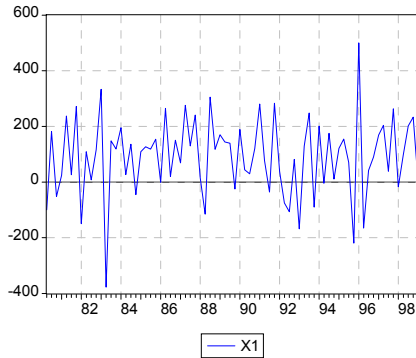
$$x = 3, 3, 6, 3, 3, 6, 3, 3, 6$$

$$y = 2, 4, 6, 2, 4, 6, 2, 4, 6$$

Wie fällt nun der F-Test für die in b) postulierte Hypothese aus?

AUFGABE 3

Von den in den folgenden Grafiken gezeigten vier Variablen ist eine stationär, eine trend-stationär und zwei sind nicht-stationär (Quartalswerte 1980-1998).



a) Geben Sie aufgrund des grafischen Eindrucks an, welche der Variablen X1, X2, X3 und X4 stationär bzw. trend-stationär ist und welche beide nicht-stationär sind!

b) Betrachten sie die folgenden Regressionsansätze:

- b1) $X3 = \alpha + \beta X1 + u$
- b2) $X3 = \alpha + \beta_1 X1 + \beta_2 \text{Zeittrend} + u$
- b3) $X1 = \alpha + \beta \Delta X4 + u$
- b4) $X2 = \alpha + \beta X4 + u$

Welche dieser Regressionsansätze erscheinen ihnen sinnvoll und welche nicht?

c) Für die berechneten Residuen der Regression b4 führen sie folgende Regression durch:

$$\Delta \hat{u}_t = \alpha + \theta \hat{u}_{t-1} + \varepsilon_t$$

Für θ erhalten Sie einen Schätzwert von -0.7 mit einem Standardfehler von 0.1 .

Was folgern Sie daraus bezüglich der Frage, ob X2 und X4 kointegriert sind?

Welche Art von Schätzansatz würden Sie vorschlagen, um den Einfluss von X4 auf X2 zu untersuchen ?

AUFGABE 4

Sie erhalten von einem Reiseveranstalter der Auftrag, die Preiselastizität der Nachfrage nach Auslandferien zu schätzen. Der Auftraggeber liefert Ihnen Quartalsdaten (1980:Q1 bis 2001:Q3) für die folgenden Variablen:

- FA Ausgaben der Schweizer Haushalte für Auslandferien
- HE Haushaltseinkommen
- PF Preisindex für Ferien im Ausland, 1990=1 (in Schweizer Franken)
- PC Landesindex der Konsumentenpreise, 1990 = 1.

Sie beginnen Ihre Analyse mit der folgenden Regression:

$$(A) \quad \text{LOG}(FA) = C(1) + C(2) * \text{LOG}(HE) + C(3) * \text{LOG}(PF/PC)$$

	Coefficient	Std. Error	t-Statistic
C(1)	-0.464031	0.241866	-1.918545
C(2)	0.778512	0.742423	1.048610
C(3)	-1.328983	1.385880	-0.958945

R-squared	0.961802	Adjusted R-squared	0.960960
S.E. of regression	0.031205	Durbin-Watson stat	1.610733

Die geschätzte Preiselastizität hat das erwartete negative Vorzeichen, eine präzise Aussage ist aber aufgrund des grossen Standardfehlers nicht möglich (was Sie z.B. durch Berechnung eines 95%-Vertrauensbereichs sehen). Sie lassen sich von zwei Kollegen beraten. Beide vermuten, dass der grossen Standardfehler von C(3) auf ein Multikollinearitätsproblem zurückzuführen ist (HE steigt über die Zeit ziemlich gleichmässig an, während der relative Preis PF/PC fällt. Die beiden erklärenden Variablen sind somit negativ miteinander korreliert.)

Da Sie letztlich an der Preiselastizität interessiert sind, schlägt Kollege B Ihnen vor, das Einkommen aus der Gleichung zu eliminieren:

$$(B) \quad \text{LOG}(FS) = C(1) + C(3) * \text{LOG}(PS/PF)$$

	Coefficient	Std. Error	t-Statistic
C(1)	1.594722	0.287422	5.548364
C(3)	-2.797661	0.158559	-17.64429

R-squared	0.961523	Adjusted R-squared	0.960958
S.E. of regression	0.031104	Durbin-Watson stat	1.628363

Kollege C hingegen schlägt vor, das Einkommen in den Nenner auf die rechte Gleichungsseite zu schreiben:

(C) $\text{LOG}(\text{FS}/\text{HE}) = \text{C}(1) + \text{C}(3) * \text{LOG}(\text{PS}/\text{PF})$

	Coefficient	Std. Error	t-Statistic
C(1)	-3.075947	0.187401	-16.41017
C(3)	-0.678497	0.138389	-4.902825

R-squared	0.961223	Adjusted R-squared	0.961126
S.E. of regression	0.031104	Durbin-Watson stat	1.612223

Fragen:

- Berechnen Sie für Gleichung A einen 95%-Vertrauensbereich für die Preiselastizität!*
- Welche Restriktion enthält Gleichung B bzw. Gleichung C im Vergleich zu Gleichung A? Testen Sie die Zulässigkeit dieser Restriktionen!*
- Wie ist es zu erklären, dass die Preiselastizität in Gleichung B einen viel stärker negativen Wert annimmt als in Gleichung C.*
- Auf welche Gleichung würden Sie ihren Bericht an den Reiseveranstalter abstützen?*

AUFGABE 5

Sie wollen die Behauptung überprüfen, dass Frauen bei gleicher Ausbildungsdauer und gleichem Alter weniger verdienen als Männer. Sie schätzen zu diesem Zweck Regressionen der Form

$$L_i = \alpha + \beta A_i + u_i,$$

worin L_i der Monatslohn in 1000 Franken und A_i die gesamte Ausbildungsdauer in Jahren (Schule, Lehre, Universität, Kurse usw.) der Person i darstellen. Leider steht Ihnen nur eine sehr kleine Stichprobe von je 4 fünfzigjährigen Frauen und Männern zur Verfügung:

Frauen:	i	L_i	A_i	Männer:	i	L_i	A_i
	1	6	10		1	8	10
	2	8	10		2	10	10
	3	11	20		3	19	20
	4	15	20		4	23	20

Fragen:

- Wie gross ist das R^2 der Regressionsgleichung für die Frauen?
- Testen Sie die Null-Hypothese, dass die Ausbildungsdauer keinen Einfluss auf den Lohn der Frauen hat, mit einem F-Test! Wie fällt der gleiche Test für die Männer aus?
- Testen Sie die Null-Hypothese, dass der Zusammenhang zwischen Ausbildungsdauer und Lohn für Frauen und Männern identisch ist, mit Hilfe eines F-Tests, indem Sie die separaten Regressionsgleichungen für Frauen und Männer mit einer gemeinsamen Regressionsgleichung für Frauen und Männer vergleichen.

Hinweis zum Vorgehen:

Zeichnen Sie (am besten auf kariertem Papier) die Beobachtungspunkte ein in x-y-Diagramm ein (A_i auf die x-Achse, L_i auf die y-Achse) und überlegen Sie sich anhand der Grafik (ohne viel zu rechnen), wie die drei Regressionsgeraden und die entsprechenden Residuen aussehen!

AUFGABE 6

Sie schätzen ein Marktmodell für ein landwirtschaftliches Produkt. Die Menge Q ist nachfrageseitig negativ vom Preis P und positiv vom Einkommen X der Haushalte abhängig:

$$\text{Nachfrage: } Q = \text{ALFA1} + \text{ALFA2} \cdot P + \text{ALFA3} \cdot X \quad \text{ALFA2} < 0, \text{ALFA3} > 0$$

Angebotsseitig ist die Menge Q positiv vom Preis P und negativ vom Lohnniveau Z in der Landwirtschaft abhängig:

$$\text{Angebot: } Q = \text{BETA1} + \text{BETA2} \cdot P + \text{BETA3} \cdot Z \quad \text{BETA2} > 0, \text{BETA3} < 0$$

Es ist angenommen, dass der Markt stets im Gleichgewicht ist, d.h. sich der Preis immer so einstellt, dass die beiden Gleichungen die gleiche Menge Q bestimmen. Die Variablen X und Z sind als exogen zu betrachten.

- a) *Warum ist bei einer Schätzung der Nachfragegleichung nach der Methode der kleinsten Quadrate (OLS) der Schätzwert für ALFA2 verzerrt?*
- b) *Diskutieren Sie die Identifikation der beiden Gleichungen in den folgenden Fällen:*
 - b1) *in der oben gegebenen Form,*
 - b2) *wenn die Variable X auch in der Angebotsgleichung erscheint,*
 - b3) *wenn die Angebotsgleichung (ausgehend von Fall b1) eine weitere exogene Variable W enthält.*
- c) *Skizzieren Sie das Vorgehen zur Schätzung der beiden Gleichungen im Fall b3 !*

AUFGABE 7

"Je länger eine Person arbeitslos ist, desto geringer ist die Wahrscheinlichkeit, dass sie wieder Arbeit findet."

Sie überprüfen diese Behauptung mit einer Stichprobe von 100 Personen, die im April 2008 für x_i Monate arbeitslos waren. 17 dieser Personen haben im Mai wieder eine Stelle gefunden ($y_i = 1$), 83 blieben arbeitslos ($y_i = 0$).

Das Probit-Modell lautet:

$$\text{Prob}(y_i = 1) = F(\alpha + \beta x_i) \quad F(\cdot): \text{kumulative Standard-Normalverteilung}$$

$y_i = 1$ im Mai Stelle gefunden $y_i = 0$ im Mai weiterhin arbeitslos

x_i Dauer der bisherigen Arbeitslosigkeit in Monaten
(x_i bewegt sich zwischen 1 und 18 Monaten)

Dependent Variable: Y

Method: ML - Binary Probit (Quadratic hill climbing)

Included observations: 100

	Coefficient	Std. Error	z-Statistic	Prob.
ALFA	2.767835	0.822296	3.365983	0.0008
BETA	-0.561418	0.127101	-4.417111	0.0000
Mean dependent var	0.170000	S.D. dependent var		0.368453
S.E. of regression	0.245927	Akaike info criterion		0.408034
Sum squared resid	5.927071	Schwarz criterion		0.460138
Log likelihood	-18.40171	Hannan-Quinn criter.		0.429121
Restr. log likelihood	-43.96699	Avg. log likelihood		-0.184017
LR statistic (1 df)	51.13056	McFadden R-squared		0.581465

Fragen:

a) Bei welcher Dauer der Arbeitslosigkeit beträgt die Wahrscheinlichkeit, im folgenden Monat eine Stelle zu finden, gemäss Modellschätzung exakt 50%? Schätzen Sie weiter ab, mit welcher Wahrscheinlichkeit eine für zwei Monate arbeitslose Person im nächsten Monat eine Stelle findet! (Für die Beantwortung dieser Frage müssen sie die Tabelle "Cumulative Normal Distribution" konsultieren).

b) Sie klassifizieren die 100 Personen nach der Regel

$$\text{Prob}(y_i = 1) > 0.5 \Leftrightarrow y_i = 1, \quad \text{Prob}(y_i = 1) \leq 0.5 \Leftrightarrow y_i = 0$$

Es zeigt sich, dass das Modell 92 Personen richtig und 8 fehlklassifiziert. Wieviel Personen werden vom restringierten Modell ($\beta = 0$, d.h. Dauer der Arbeitslosigkeit hat keinen Einfluss) richtig bzw. fehlklassifiziert?

c) Für eine Person, die gemäss Modell mit 50% Wahrscheinlichkeit eine Stelle findet, tritt dieser Fall nicht ein, sie bleibt also einen Monat länger arbeitslos. Wie hoch ist nun für diese Person die Wahrscheinlichkeit, im nachfolgenden Monat eine Stelle zu finden. (Hinweis: Die Dichtefunktion der Standard-Normalverteilung hat an der Stelle Null den Wert $1/\sqrt{2\pi} = 0.399$.)