

Klausur / Exam
Methoden der Ökonometrie / Econometric Methods

am / at 27.02.2008

Dauer/Examination Duration 150 Min.

Lesezeit/Reading Time 10 Min.

Nachname/Last Name:

Vorname/First Name:

Immatrikulations–Nr/Enrolment No.:

Erlaubte Hilfsmittel: Formelsammlung (ausgeteilt), Taschenrechner (nicht programmierbar)
Allowed Resources: Formulary (provided), Pocket Calculator (nonintelligent)

Bitte bestätigen Sie durch Ihre Unterschrift, dass Ihnen bekannt ist, dass jeder Täuschungsversuch zum Abbruch der Arbeit mit der Bewertung "ungenügend" (5.0) führt.

Please confirm by signing that you are aware that any attempt to deceive results in a grade "failed" (5.0).

.....
Unterschrift/Signature

(nur für die Prüfer/only for examiner)

Punkte/Points	Problem 1	Problem 2	Problem 3	Problem 4	Gesamt/Total
mögliche / possible	35	32	33	26	126
erreichte / attained					

Note / Grade:

.....
1. Prüfer/1st examiner

.....
2. Prüfer/2nd examiner

Problem 1 (35 Points)

1. Consider the linear regression model with stochastic regressors and linear restriction/s $R\beta = r$:

$$y = X\beta + \varepsilon, \quad \varepsilon|X \sim N(0, \sigma^2 I_n).$$

The $(n \times k)$ -matrix X has $\text{rank}(X) = k$ with probability 1. The deterministic $(m \times k)$ -matrix R has full row rank, i. e. $\text{rank}(R) = m$.

- (a) **(5 Points)** Show that the unrestricted ordinary least-squares estimator (OLSE), $\hat{\beta}_U$, and the restricted OLSE, $\hat{\beta}_R = \hat{\beta}_U + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_U - r)$, are unbiased under the restriction $R\beta = r$.
- (b) **(4 Points)** Derive the unconditional covariance matrix of $\hat{\beta}_U$ if $R\beta = r$.
Hint: For two random vectors Z_1 and Z_2 it holds : $\mathbb{V}[Z_1] = \mathbb{E}[\mathbb{V}[Z_1|Z_2]] + \mathbb{V}[\mathbb{E}[Z_1|Z_2]]$.
- (c) **(5 Points)** If the restriction $R\beta = r$ holds, prove that $\mathbb{V}[\hat{\beta}_R] = \sigma^2 \mathbb{E}[M^*(X'X)^{-1}]$, where $M^* = I_k - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R$.
Hint 1: Note that $M^*(X'X)^{-1}M^{*'} = M^*(X'X)^{-1}$.
Hint 2: You can use the fact that $\hat{\beta}_R = \beta + M^*(X'X)^{-1}X'\varepsilon$ if $R\beta = r$. (You do not have to prove the hints.)
- (d) **(2 Points)** Under the restriction $R\beta = r$, it can be shown that $\mathbb{V}[\hat{\beta}_U] - \mathbb{V}[\hat{\beta}_R]$ is a positive semidefinite matrix. How do you interpret this result?
- (e) Assume now that the restriction/s does/do not hold, i.e. $R\beta \neq r$.
 - i. **(3 Points)** Derive the expectations $\mathbb{E}[\hat{\beta}_U]$ and $\mathbb{E}[\hat{\beta}_R]$.
 - ii. **(1 Point)** It can be shown that the formulae for the covariance matrices $\mathbb{V}[\hat{\beta}_U]$ and $\mathbb{V}[\hat{\beta}_R]$ under $R\beta = r$ as well as under $R\beta \neq r$ correspond to those from part 1(b) and part 1(c). What can you say about the relative efficiency of $\hat{\beta}_U$ and $\hat{\beta}_R$?
 - iii. **(1 Point)** In how far can the MSE criterion be used to compare $\hat{\beta}_U$ and $\hat{\beta}_R$?

2. Consider now the linear regression model from above with a constant and two explanatory variables, i.e. $k = 3$ and $\beta = (\beta_1, \beta_2, \beta_3)'$. Suppose an economic theory implies that $\beta_2 + \beta_3 = 0$. Moreover, for a sample of size $n = 20$ let the following data be given:

$$(X'X)^{-1} = \begin{bmatrix} 4.000 & -0.277 & -0.425 \\ -0.277 & 0.045 & 0.010 \\ -0.425 & 0.010 & 0.065 \end{bmatrix}, \quad X'y = \begin{bmatrix} 11 \\ 57 \\ 59 \end{bmatrix}, \quad e'e = 51.$$

- (a) **(12 Points)** Use the F-test for linear restrictions to test $H_0 : \beta_2 + \beta_3 = 0$ against $H_1 : \beta_2 + \beta_3 \neq 0$ using a significance level of $\alpha = 0.05$. The critical value is given by $F_{(1, 17, 95\%)} = 4.45$. Based on the test decision which estimator would you use, $\hat{\beta}_U$ or $\hat{\beta}_R$?
- (b) **(2 Points)** Suppose you rejected H_0 in part 2(a). Which type of error would you make if the true values of β_2 and β_3 were 0.5 and -0.5 respectively? How large is the probability of such an error?

Problem 2 (32 Points)

1. Consider the linear regression model:

$$y_i = x_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where x_i and β are $(k \times 1)$ -vectors. Moreover, it is assumed that:

- A1) $\{(y_i, x_i')'\}$ is a sequence of i.i.d. $(k+1 \times 1)$ -random vectors,
- A2) $\mathbb{E}[x_i x_i'] = \Omega_{XX}$ is finite and strictly positive definite,
- A3) 4th moments of regressors exist,
- A4) $\mathbb{E}[\varepsilon_i x_i] = 0$,
- A5) $\mathbb{E}[\varepsilon_i^2 x_i x_i'] = \Omega_{\varepsilon XX}$ is finite and strictly positive definite.

Denote the $(n \times k)$ -matrix $X = [x_1, \dots, x_n]'$ and the $(n \times 1)$ -vector $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$. Under assumptions A1), A2) and A3), it can be shown that $(X'X/n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \Omega_{XX}$.

- (a) **(7 Points)** Which weak law of large numbers holds for $X'\varepsilon/n$? Justify briefly and give the probability limit. Which central limit theorem holds for $X'\varepsilon/\sqrt{n}$? Justify briefly and give the limiting distribution.
- (b) **(3 Points)** Show that the OLSE $\hat{\beta}_n$ in model (1) is consistent.
- (c) **(3 Points)** Show that $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow[n \rightarrow \infty]{\mathbb{D}} N(0, \Omega_{XX}^{-1} \Omega_{\varepsilon XX} \Omega_{XX}^{-1})$.
- (d) **(2 Points)** Approximate the distribution of $\hat{\beta}_n$ using the asymptotic result from part 1(c). Give a consistent estimator for the approximate covariance matrix of $\hat{\beta}_n$.
- (e) **(3 Points)** Explain briefly why it is necessary to approximate the distribution of $\hat{\beta}_n$. Comment also on the usefulness of the transformation with \sqrt{n} in part 1(c).

2. Now, consider the linear regression model:

$$y_i = x_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (2)$$

where x_i and β are $(k \times 1)$ -vectors with $k \geq 2$. Again, assume A1), A2), A3) and replace assumptions A4) and A5) by:

- AA3) $\mathbb{E}[x_i x_i' x_{i2}^2] = \Omega_{XXX}$ is finite and strictly positive definite,
- AA4) $\mathbb{E}[\varepsilon_i | x_i] = 0$,
- AA5) $\mathbb{E}[\varepsilon_i^2 | x_i] = \sigma^2 x_{i2}^2$.

- (a) **(2 Points)** How do you interpret assumptions AA4) and AA5)? Explain briefly.
- (b) **(4 Points)** Show that assumptions AA3), AA4) and AA5) imply assumptions A4) and A5) from model (1). Which model is more general, model (1) or (2)?
- (c) **(1 Point)** Give the limiting distribution of $\sqrt{n}(\hat{\beta}_n - \beta)$.
- (d) **(4 Points)** Show that model (2) is a linear regression model with unconditional homoskedastic errors terms.

3. **(3 Points)** What is wrong with the following text? Explain briefly.

Any consistent estimator converges in mean square and is asymptotically normally distributed. Moreover, an asymptotically unbiased estimator is consistent.

Problem 3 (33 Points)

Consider $y_i, i = 1, \dots, n$, i.i.d. variables that follow a Poisson distribution with $\mathbb{E}[y_i] = \lambda$ and $\mathbb{V}[y_i] = \lambda$. Assume that λ is finite.

1. **(4 Points)** Show that the log likelihood function is given by

$$l(\lambda|y) = -n\lambda + \ln \lambda \sum_{i=1}^n y_i - \sum_{i=1}^n \ln(y_i!),$$

with $y = (y_1, \dots, y_n)'$.

2. **(2 Points)** Derive the score function of the log likelihood function. Interpret this function.
3. **(2 Points)** Show that the maximum likelihood estimator (MLE) is given by $\hat{\lambda}_{ML} = (1/n) \sum_{i=1}^n y_i$.
4. **(2 Points)** Show that the MLE $\hat{\lambda}_{ML}$ is unbiased.
5. **(2 Points)** Derive the formula for the variance of $\hat{\lambda}_{ML}$.
6. **(4 Points)** Determine the second derivative of $l(\lambda|y)$ and the (scalar) Fisher Information $I(\lambda)$. Show that $\hat{\lambda}_{ML}$ is efficient. Which theorem do you use to show that $\hat{\lambda}_{ML}$ is efficient?
7. **(2 Points)** Is the MLE $\hat{\lambda}_{ML}$ consistent if the distribution is unknown, but it still holds that the $y_i, i = 1, \dots, n$, are i.i.d. with $\mathbb{E}[y_i] = \lambda, \mathbb{V}[y_i] = \lambda$? Justify your answer briefly.
8. **(8 Points)** Suppose you observe a random sample of $y_i, i = 1, \dots, n$, with $\sum_{i=1}^{50} y_i = 450$. Perform a Wald test to test the null hypothesis of $\lambda = 8$ against the alternative $\lambda \neq 8$. The critical value is given by $\chi^2_{(1, 95\%)} = 3.842$. Proceed as follows:
- (a) Calculate $\hat{\lambda}_{ML}$.
 - (b) Formulate your null hypothesis in terms of $r(\lambda)$ and derive $R(\hat{\lambda}_{ML})$ as defined in the formulary.
 - (c) Estimate the Fisher Information by $I(\hat{\lambda}_{ML})$.
 - (d) Calculate the value for the Wald test statistic.
 - (e) Make your test decision.
9. **(3 Points)** Given the pair of hypotheses from part 8, what does it mean to make a type II error in this test? Can you reduce the probability of making a type II error by changing the significance level α ? In this context, what does $\alpha = 1$ imply? Explain briefly.
10. **(2 Points)** Is it really necessary to use an asymptotic test in part 8? Explain your answer.
11. **(2 Points)** What is wrong in the following text? Explain briefly.
The Wald, LM and LR tests are always exact tests of hypotheses about the parameters of a statistical model. In finite samples, the LM test statistic is always larger than the Wald test statistic.

Problem 4 (26 Points)

1. Consider the following two linear regression equations

$$y_i^{(m)} = x_i^{(m)'} \beta^{(m)} + \varepsilon_i^{(m)}, \quad \varepsilon_i^{(m)} \sim N(0, \sigma^2), \quad i = 1, \dots, n_m \quad (3)$$

$$y_i^{(f)} = x_i^{(f)'} \beta^{(f)} + \varepsilon_i^{(f)}, \quad \varepsilon_i^{(f)} \sim N(0, \sigma^2), \quad i = n_m + 1, \dots, n, \quad (4)$$

where $\varepsilon^{(m)}$ and $\varepsilon^{(f)}$ are independent of each other. Moreover,

$y_i^{(g)}$ denotes the log wage of person i with gender g (either male (m) or female (f)),

$x_i^{(g)'}$ is a stochastic (1×2) -regressor vector of person i including a constant and years of work experience,

$\beta^{(g)}$ is a (2×1) parameter vector and

$n_f = n - n_m$ is the number of female persons.

Take the following data as given: $n_m = n_f = 25$ and

$$(X^{(m)'} X^{(m)})^{-1} = \begin{bmatrix} 0.1511 & -0.0226 \\ -0.0226 & 0.0046 \end{bmatrix}, \quad X^{(m)'} y^{(m)} = \begin{bmatrix} 134.97 \\ 816.64 \end{bmatrix}, \quad \frac{\sum_{i=1}^{n_m} x_{i2}^{(m)}}{n_m} = 4.92,$$

$$(X^{(f)'} X^{(f)})^{-1} = \begin{bmatrix} 0.1104 & -0.0195 \\ -0.0195 & 0.0054 \end{bmatrix}, \quad X^{(f)'} y^{(f)} = \begin{bmatrix} 61.16 \\ 273.78 \end{bmatrix}, \quad \frac{\sum_{i=n_m+1}^n x_{i2}^{(f)}}{n_f} = 3.60.$$

(a) (8 Points) Show that $\Delta = E[y_i^{(f)} - y_i^{(m)}]$ is given by

$$\Delta = \left(E[x_i^{(f)}] - E[x_i^{(m)}] \right)' \beta^{(f)} + E[x_i^{(m)}]' (\beta^{(f)} - \beta^{(m)}).$$

Interpret the two right-hand-side terms as different determinants of log wage differences among genders. Which term measures wage discrimination and what is the meaning of the other term? Does $\Delta < 0$ necessarily imply wage discrimination?

(b) (5 Points) Estimate $\beta^{(m)}$ and $\beta^{(f)}$ of models (3) and (4) separately using the OLSE and the data given above. Calculate an estimate for $\hat{\Delta}$.

2. Consider the linear regression model

$$y_i = D_i x_i' \beta^{(f)} + (1 - D_i) x_i' \beta^{(m)} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \dots, n, \quad (5)$$

$$\text{where } y_i = \begin{cases} y_i^{(m)}, & \text{for } i = 1, \dots, n_m \\ y_i^{(f)}, & \text{for } i = n_m + 1, \dots, n \end{cases}, \quad x_i' = \begin{cases} x_i^{(m)'}, & \text{for } i = 1, \dots, n_m \\ x_i^{(f)'}, & \text{for } i = n_m + 1, \dots, n \end{cases},$$

$$\text{and } \varepsilon_i = \begin{cases} \varepsilon_i^{(m)}, & \text{for } i = 1, \dots, n_m \\ \varepsilon_i^{(f)}, & \text{for } i = n_m + 1, \dots, n \end{cases}, \quad D_i = \begin{cases} 0, & \text{for } i = 1, \dots, n_m \\ 1, & \text{for } i = n_m + 1, \dots, n. \end{cases}$$

(a) (2 Points) Show that model (5) stacks the observations of equations (3) and (4), i.e. rewrite equation (5) for $i = 1, \dots, n_m$ and $i = n_m + 1, \dots, n$ in terms of $y_i^{(g)}$, $x_i^{(g)'}$, $\beta^{(g)}$ and $\varepsilon_i^{(g)}$ with $g \in \{m, f\}$.

(b) **(5 Points)** Show that model (5) can also be written as

$$y_i = D_i x_i' \delta + x_i' \beta^{(m)} + \varepsilon_i, \quad \delta = \beta^{(f)} - \beta^{(m)}, \quad i = 1, \dots, n. \quad (6)$$

Rewrite equation (6) in matrix form

$$y = X\gamma + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n), \quad (7)$$

where X is an $(n \times 4)$ -matrix and γ is a (4×1) -vector. Give the detailed representation of the X -matrix and the γ -vector, i.e. write down the X -matrix and the γ -vector in terms of $x_i^{(g)}$, $i = 1, \dots, n$, and $\beta^{(g)}$ with $g \in \{m, f\}$.

(c) **(4 Points)** Consider the estimation output in Figure 1 for model (6). Unfortunately, some entries have been deleted. Test the two null hypotheses that δ_1 and δ_2 are equal to zero. Use the critical value $t_{(44, 97.5\%)} = 2.015$ and the significance level $\alpha = 0.05$. What would a rejection of the null hypotheses indicate with respect to a possible discrimination effect discussed in part 1(a)?

Note, in the regression output in Figure 1, D denotes the dummy variable and $WorkExp$ is the number of years of work experience. $WorkExpD$ denotes the elementwise product of years of work experience and the dummy variable D .

Dependent Variable: LN WAGE Method: Least Squares Date: 02/22/08 Time: 22:52 Sample: 1 50 Included observations: 50 LN WAGE=C(1)*D+C(2)*WorkExpD+C(3)+C(4)*WorkExp				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.554	0.575	*****	*****
C(2)	-0.409	0.113	*****	*****
C(3)	1.952	0.437	*****	*****
C(4)	0.701	0.076	*****	*****
R-squared	0.799	Mean dependent var		3.923
Adjusted R-squared	0.786	S.D. dependent var		2.431
S.E. of regression	1.125	Akaike info criterion		3.149
Sum squared resid	58.18	Schwarz criterion		3.302
Log likelihood	-74.73	Durbin-Watson stat		1.924

Figure 1: Regression output of model (6)

3. **(2 Points)** Which advantage has the approach in part 2 compared to the approach in part 1?

Aufgabe 1 (35 Punkte)

1. Betrachten Sie das lineare Regressionsmodell mit stochastischen Regressoren und der/den linearen Restriktion/en $R\beta = r$:

$$y = X\beta + \varepsilon, \quad \varepsilon|X \sim N(0, \sigma^2 I_n).$$

Für die $(n \times k)$ -Matrix X gilt $\text{Rang}(X) = k$ mit Wahrscheinlichkeit 1. Die deterministische $(m \times k)$ -Matrix R hat vollen Zeilenrang, d. h. $\text{Rang}(R) = m$.

- (a) **(5 Punkte)** Zeigen Sie, dass der unrestringierte gewöhnliche Kleinste-Quadrate Schätzer (KQS), $\hat{\beta}_U$, und der restringierte KQS, $\hat{\beta}_R = \hat{\beta}_U + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_U - r)$, unverzerrt sind, wenn die Restriktion $R\beta = r$ gilt.
- (b) **(4 Punkte)** Leiten Sie die unbedingte Kovarianzmatrix von $\hat{\beta}_U$ her, wenn $R\beta = r$ gilt.
Hinweis: Für zwei Zufallsvektoren Z_1 und Z_2 gilt: $\mathbb{V}[Z_1] = \mathbb{E}[\mathbb{V}[Z_1|Z_2]] + \mathbb{V}[\mathbb{E}[Z_1|Z_2]]$.
- (c) **(5 Punkte)** Wenn die Restriktion $R\beta = r$ gilt, beweisen Sie, dass $\mathbb{V}[\hat{\beta}_R] = \sigma^2 \mathbb{E}[M^*(X'X)^{-1}]$, wobei $M^* = I_k - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}R$ ist.
Hinweis 1: Beachten Sie, dass $M^*(X'X)^{-1}M^{*'} = M^*(X'X)^{-1}$.
Hinweis 2: Unter der Restriktion $R\beta = r$ können Sie die Darstellung $\hat{\beta}_R = \beta + M^*(X'X)^{-1}X'\varepsilon$ verwenden. (Sie müssen die Hinweise nicht beweisen.)
- (d) **(2 Punkte)** Wenn die Restriktion $R\beta = r$ gilt, kann gezeigt werden, dass $\mathbb{V}[\hat{\beta}_U] - \mathbb{V}[\hat{\beta}_R]$ eine positiv semidefinite Matrix ist. Wie interpretieren Sie dieses Resultat?
- (e) Nehmen Sie nun an, dass die Restriktion/en nicht gilt/gelten, d.h. $R\beta \neq r$.
- (3 Punkte)** Leiten Sie die Erwartungswerte $\mathbb{E}[\hat{\beta}_U]$ und $\mathbb{E}[\hat{\beta}_R]$ her.
 - (1 Punkt)** Es kann gezeigt werden, dass die Formeln für die Kovarianzmatrizen $\mathbb{V}[\hat{\beta}_U]$ bzw. $\mathbb{V}[\hat{\beta}_R]$ sowohl unter $R\beta = r$ als auch unter $R\beta \neq r$ denen aus den Teilaufgaben 1(b) und 1(c) entsprechen. Was können Sie über die relative Effizienz der Schätzer $\hat{\beta}_U$ und $\hat{\beta}_R$ sagen?
 - (1 Punkt)** Inwiefern kann das MSE Kriterium zum Vergleich von $\hat{\beta}_U$ und $\hat{\beta}_R$ verwendet werden?

2. Betrachten Sie nun das lineare Regressionsmodell aus Teilaufgabe 1 mit einer Konstanten und zwei erklärenden Variablen, d.h. $k = 3$ und $\beta = (\beta_1, \beta_2, \beta_3)'$. Nehmen Sie an, eine ökonomische Theorie impliziert, dass $\beta_2 + \beta_3 = 0$. Darüber hinaus verfügen Sie über eine Stichprobe der Größe $n = 20$ mit den folgenden Daten:

$$(X'X)^{-1} = \begin{bmatrix} 4.000 & -0.277 & -0.425 \\ -0.277 & 0.045 & 0.010 \\ -0.425 & 0.010 & 0.065 \end{bmatrix}, \quad X'y = \begin{bmatrix} 11 \\ 57 \\ 59 \end{bmatrix}, \quad e'e = 51.$$

- (a) **(12 Punkte)** Testen Sie $H_0 : \beta_2 + \beta_3 = 0$ gegen $H_1 : \beta_2 + \beta_3 \neq 0$. Führen Sie dazu einen F-Test zur Überprüfung linearer Restriktionen auf einem Signifikanzniveau von $\alpha = 0.05$ durch. Der kritische Wert ist $F_{(1, 17, 95\%)} = 4.45$. Basierend auf Ihrer Testentscheidung, welchen Schätzer würden Sie verwenden, $\hat{\beta}_U$ oder $\hat{\beta}_R$?
- (b) **(2 Punkte)** Angenommen Sie haben H_0 in Teilaufgabe 2(a) verworfen. Welche Art von Fehler würden Sie begehen, wenn die wahren Parameterwerte von β_2 und β_3 durch 0.5 bzw. -0.5 gegeben wären? Wie groß ist die Wahrscheinlichkeit einen solchen Fehler zu begehen?

Aufgabe 2 (32 Punkte)

1. Betrachten Sie das lineare Regressionsmodell:

$$y_i = x_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (8)$$

wobei x_i und β zwei $(k \times 1)$ -Vektoren sind. Darüber hinaus wird folgendes angenommen:

- A1) $\{(y_i, x_i')'\}$ ist eine Folge von identisch und unabhängig verteilten $(k+1 \times 1)$ -Zufallsvektoren,
- A2) $\mathbb{E}[x_i x_i'] = \Omega_{XX}$ ist endlich und strikt positiv definit,
- A3) 4. Momente der Regressoren existieren,
- A4) $\mathbb{E}[\varepsilon_i x_i] = 0$,
- A5) $\mathbb{E}[\varepsilon_i^2 x_i x_i'] = \Omega_{\varepsilon XX}$ ist endlich und strikt positiv definit.

$X = [x_1, \dots, x_n]'$ bezeichne die $(n \times k)$ -Regressormatrix und $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ den $(n \times 1)$ -Fehlervektor. Mit den Annahmen A1) bis A3) kann man zeigen, dass $(X'X/n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \Omega_{XX}$.

- (a) **(7 Punkte)** Welches schwache Gesetz der großen Zahlen gilt für $X'\varepsilon/n$. Begründen Sie kurz. Geben Sie den Grenzwert in Wahrscheinlichkeit an. Welcher Zentraler Grenzwertsatz gilt für $X'\varepsilon/\sqrt{n}$? Begründen Sie kurz. Geben Sie die Grenzverteilung an.
- (b) **(3 Punkte)** Zeigen Sie, dass der gewöhnliche Kleinste-Quadrate Schätzer (KQS) $\hat{\beta}_n$ in Modell (8) konsistent ist.
- (c) **(3 Punkte)** Zeigen Sie, dass $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow[n \rightarrow \infty]{\mathbb{D}} N(0, \Omega_{XX}^{-1} \Omega_{\varepsilon XX} \Omega_{XX}^{-1})$.
- (d) **(2 Punkte)** Approximieren Sie die Verteilung von $\hat{\beta}_n$ mit Hilfe des asymptotischen Resultats aus Teilaufgabe 1(c). Geben Sie einen konsistenten Schätzer für die approximative Kovarianzmatrix von $\hat{\beta}_n$ an.
- (e) **(3 Punkte)** Erläutern Sie kurz, warum die Verteilung von $\hat{\beta}_n$ approximiert werden muss. Kommentieren Sie auch die Zweckmäßigkeit der Transformation mit \sqrt{n} in Teilaufgabe 1(c).

2. Betrachten Sie nun das lineare Regressionsmodell:

$$y_i = x_i' \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (9)$$

wobei x_i und β zwei $(k \times 1)$ -Vektoren sind und $k \geq 2$ ist. Nehmen Sie an, dass A1), A2), A3) gelten und ersetzen Sie die Annahmen A4) und A5) durch:

- AA3) $\mathbb{E}[x_i x_i' x_{i2}^2] = \Omega_{XXX}$ ist endlich und strikt positiv definit,
- AA4) $\mathbb{E}[\varepsilon_i | x_i] = 0$,
- AA5) $\mathbb{E}[\varepsilon_i^2 | x_i] = \sigma^2 x_{i2}^2$.

- (a) **(2 Punkte)** Wie interpretieren Sie die Annahmen AA4) und AA5)? Erklären Sie kurz.
- (b) **(4 Punkte)** Zeigen Sie, dass die Annahmen AA3) bis AA5) die Annahmen A4) und A5) aus Modell (8) implizieren. Welches Modell ist allgemeiner, Modell (8) oder (9) ?
- (c) **(1 Punkt)** Geben Sie die Grenzverteilung von $\sqrt{n}(\hat{\beta}_n - \beta)$ an.
- (d) **(4 Punkte)** Zeigen Sie, dass Modell (9) ein lineares Regressionsmodell mit unbedingt homoskedastischen Fehlern ist.

3. **(3 Punkte)** Was ist falsch an folgendem Text? Erklären Sie kurz.

Jeder konsistente Schätzer konvergiert im quadratischen Mittel und ist asymptotisch normalverteilt. Ausserdem ist ein asymptotisch unverzerrter Schätzer konsistent.

Aufgabe 3 (33 Punkte)

Betrachten Sie die unabhängig und identisch Poisson verteilten Zufallsvariablen y_i , $i = 1, \dots, n$, mit Erwartungswert $\mathbb{E}[y_i] = \lambda$ und Varianz $\mathbb{V}[y_i] = \lambda$. Der Parameter λ sei endlich.

1. (4 Punkte) Zeigen Sie, dass die logarithmierte Likelihood Funktion die folgende Form hat

$$l(\lambda|y) = -n\lambda + \ln \lambda \sum_{i=1}^n y_i - \sum_{i=1}^n \ln(y_i!),$$

wobei $y = (y_1, \dots, y_n)'$ ist.

2. (2 Punkte) Geben Sie die Score Funktion der logarithmierten Likelihood Funktion an. Interpretieren Sie diese.
3. (2 Punkte) Zeigen Sie, dass der Maximum Likelihood (ML) Schätzer durch $\hat{\lambda}_{ML} = (1/n) \sum_{i=1}^n y_i$ gegeben ist.
4. (2 Punkte) Zeigen Sie die Unverzerrtheit des ML Schätzers $\hat{\lambda}_{ML}$.
5. (2 Punkte) Leiten Sie die Varianzformel für $\hat{\lambda}_{ML}$ her.
6. (4 Punkte) Bestimmen Sie die zweite Ableitung von $l(\lambda|y)$ und die (skalare) Fisher Information $I(\lambda)$. Zeigen Sie, dass $\hat{\lambda}_{ML}$ ein effizienter Schätzer ist. Welches Theorem benutzen Sie, um auf Effizienz von $\hat{\lambda}_{ML}$ zu schließen?
7. (2 Punkte) Ist der ML Schätzer $\hat{\lambda}_{ML}$ konsistent, wenn die Verteilungsfunktion der Zufallsvariablen zwar unbekannt ist, aber dennoch gilt, dass y_i , $i = 1, \dots, n$, i.i.d. sind und dass $\mathbb{E}[y_i] = \lambda$, $\mathbb{V}[y_i] = \lambda$ gegeben ist? Begründen Sie Ihre Antwort kurz.
8. (8 Punkte) Es liegen Beobachtungen für y_i , $i = 1, \dots, n$, aus einer Zufallsstichprobe mit $\sum_{i=1}^{50} y_i = 450$ vor. Führen Sie zur Überprüfung der Nullhypothese $\lambda = 8$ gegen die Alternative $\lambda \neq 8$ einen Wald Test durch. Der kritische Wert ist $\chi^2_{(1, 95\%)} = 3.842$. Gehen Sie wie folgt vor:
- Berechnen Sie $\hat{\lambda}_{ML}$.
 - Formulieren Sie die Nullhypothese unter Verwendung von $r(\lambda)$ und bestimmen Sie $R(\hat{\lambda}_{ML})$ (wie in der Formelsammlung definiert).
 - Schätzen Sie die Fisher Information mittels $I(\hat{\lambda}_{ML})$.
 - Berechnen Sie den Wert der Wald Teststatistik.
 - Treffen Sie die Testentscheidung.
9. (3 Punkte) Erläutern Sie kurz anhand des Hypothesenpaares in Teilaufgabe 8 die Bedeutung eines Fehlers 2. Art. Können Sie die Wahrscheinlichkeit eines Fehlers 2. Art reduzieren, indem Sie das Signifikanzniveau α ändern? Was bedeutet es demnach, wenn $\alpha = 1$ ist? Erklären Sie kurz.
10. (2 Punkte) Ist ein asymptotischer Test in Teilaufgabe 8 wirklich notwendig? Erklären Sie Ihre Antwort kurz.
11. (2 Punkte) Was ist im folgenden Text falsch? Erklären Sie kurz.
Der Wald Test, der LM Test und der LR Test sind immer exakte Tests von Hypothesen bezüglich der Parameter eines statistischen Modells. In endlichen Stichproben ist der Wert der LM Teststatistik immer größer als der Wert der Wald Statistik.

Aufgabe 4 (26 Punkte)

1. Betrachten Sie die folgenden zwei linearen Regressionsgleichungen

$$y_i^{(m)} = x_i^{(m)'} \beta^{(m)} + \varepsilon_i^{(m)}, \quad \varepsilon_i^{(m)} \sim N(0, \sigma^2), \quad i = 1, \dots, n_m \quad (10)$$

$$y_i^{(f)} = x_i^{(f)'} \beta^{(f)} + \varepsilon_i^{(f)}, \quad \varepsilon_i^{(f)} \sim N(0, \sigma^2), \quad i = n_m + 1, \dots, n. \quad (11)$$

$e^{(m)}$ und $e^{(f)}$ sind voneinander unabhängig.

$y_i^{(g)}$ bezeichnet den logarithmierten Lohn der Person i des Geschlechts g (entweder männlich (m) oder weiblich (f)),

$x_i^{(g)'}$ ist ein stochastischer (1×2) -Vektor der Regressoren bestehend aus einer Konstanten und der Anzahl der Jahre Berufserfahrung für Person i ,

$\beta^{(g)}$ ist ein (2×1) -Koeffizientenvektor und

$n_f = n - n_m$ bezeichnet die Anzahl der weiblichen Personen.

Gegeben sind: $n_m = n_f = 25$ und

$$(X^{(m)'} X^{(m)})^{-1} = \begin{bmatrix} 0.1511 & -0.0226 \\ -0.0226 & 0.0046 \end{bmatrix}, \quad X^{(m)'} y^{(m)} = \begin{bmatrix} 134.97 \\ 816.64 \end{bmatrix}, \quad \frac{\sum_{i=1}^{n_m} x_{i2}^{(m)}}{n_m} = 4.92,$$

$$(X^{(f)'} X^{(f)})^{-1} = \begin{bmatrix} 0.1104 & -0.0195 \\ -0.0195 & 0.0054 \end{bmatrix}, \quad X^{(f)'} y^{(f)} = \begin{bmatrix} 61.16 \\ 273.78 \end{bmatrix}, \quad \frac{\sum_{i=n_m+1}^n x_{i2}^{(f)}}{n_f} = 3.60.$$

(a) (8 Punkte) Zeigen Sie, dass $\Delta = E[y_i^{(f)} - y_i^{(m)}]$ umgeformt werden kann zu

$$\Delta = \left(E[x_i^{(f)}] - E[x_i^{(m)}] \right)' \beta^{(f)} + E[x_i^{(m)}]' (\beta^{(f)} - \beta^{(m)}).$$

Interpretieren Sie die Terme der rechten Seite der Gleichung. Welcher Term sagt etwas über Lohndiskriminierung aus und was misst der andere Term? Lässt sich für $\Delta < 0$ notwendigerweise auf Lohndiskriminierung schließen?

(b) (5 Punkte) Schätzen Sie $\beta^{(m)}$ und $\beta^{(f)}$ der Modelle (10) und (11) einzeln unter Verwendung der KQ-Methode und der gegebenen Daten. Berechnen Sie einen Schätzwert für $\hat{\Delta}$.

2. Betrachten Sie nun das folgende lineare Regressionsmodell

$$y_i = D_i x_i' \beta^{(f)} + (1 - D_i) x_i' \beta^{(m)} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \dots, n, \quad (12)$$

$$\text{wobei} \quad y_i = \begin{cases} y_i^{(m)}, & \text{für } i = 1, \dots, n_m \\ y_i^{(f)}, & \text{für } i = n_m + 1, \dots, n \end{cases}, \quad x_i' = \begin{cases} x_i^{(m)'}, & \text{für } i = 1, \dots, n_m \\ x_i^{(f)'}, & \text{für } i = n_m + 1, \dots, n \end{cases},$$

$$\text{und} \quad \varepsilon_i = \begin{cases} \varepsilon_i^{(m)}, & \text{für } i = 1, \dots, n_m \\ \varepsilon_i^{(f)}, & \text{für } i = n_m + 1, \dots, n \end{cases}, \quad D_i = \begin{cases} 0, & \text{für } i = 1, \dots, n_m \\ 1, & \text{für } i = n_m + 1, \dots, n \end{cases}.$$

(a) (2 Punkte) Zeigen Sie, dass Modell (12) die Beobachtungen der Gleichungen (10) und (11) untereinander aufreicht, d.h. schreiben Sie Gleichung (12) für $i = 1, \dots, n_m$ und $i = n_m + 1, \dots, n$ mit Hilfe von $y_i^{(g)}$, $x_i^{(g)'}$, $\beta^{(g)}$ und $\varepsilon_i^{(g)}$ um (mit $g \in \{m, f\}$).

(b) (5 Punkte) Zeigen Sie, dass Modell (12) wie folgt dargestellt werden kann

$$y_i = D_i x_i' \delta + x_i' \beta^{(m)} + \varepsilon_i, \quad \delta = \beta^{(f)} - \beta^{(m)}, \quad i = 1, \dots, n. \quad (13)$$

Schreiben Sie Gleichung (13) in die folgende Matrixform um

$$y = X\gamma + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n), \quad (14)$$

wobei X eine $(n \times 4)$ -Matrix und γ ein (4×1) -Vektor ist. Geben Sie die detaillierte Schreibweise der X -Matrix und des γ -Vektors an, d.h. notieren Sie die X -Matrix und den γ -Vektor mit Hilfe von $x_i^{(g)}$, $i = 1, \dots, n$, und $\beta^{(g)}$, wobei $g \in \{m, f\}$ ist.

- (c) **(4 Punkte)** Betrachten Sie den Regressionsoutput in Abbildung 2 (Figure 2) für das Modell (13). Leider sind einige Einträge gelöscht worden. Testen Sie die Nullhypothesen dass δ_1 und δ_2 null sind. Benutzen Sie den kritischen Wert $t_{(44, 97.5\%)} = 2.015$ und das Signifikanzniveau $\alpha = 0.05$). Was würde eine Ablehnung beider Nullhypothesen im Hinblick auf einen in Teilaufgabe 1(a) diskutierten möglichen Diskriminierungseffekt bedeuten?

Beachten Sie die Notation. In Abbildung 2 bezeichnet LN Wage den logarithmierten Lohn, D die Dummy Variable, WorkExp steht für die Jahre Berufserfahrung. WorkExpD bezeichnet den Vektor der elementweise mit der Dummy Variablen D multiplizierten Jahre Berufserfahrung.

Dependent Variable: LN WAGE Method: Least Squares Date: 02/22/08 Time: 22:52 Sample: 1 50 Included observations: 50 LN WAGE=C(1)*D+C(2)*WorkExpD+C(3)+C(4)*WorkExp				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.554	0.575	*****	*****
C(2)	-0.409	0.113	*****	*****
C(3)	1.952	0.437	*****	*****
C(4)	0.701	0.076	*****	*****
R-squared	0.799	Mean dependent var		3.923
Adjusted R-squared	0.786	S.D. dependent var		2.431
S.E. of regression	1.125	Akaike info criterion		3.149
Sum squared resid	58.18	Schwarz criterion		3.302
Log likelihood	-74.73	Durbin-Watson stat		1.924

Figure 2: Regressionsoutput des Modells (13)

3. **(2 Punkte)** Welchen Vorteil hat der Ansatz in Teilaufgabe 2 gegenüber dem Ansatz in Teilaufgabe 1?