

Aufgabensammlung Mikroökonomie

2015S

Ulrike Schneider

Aufgaben zu Kapitel 1

1. Sei $\hat{\theta}$ ein Schätzer für den Parameter $\theta \in \mathbb{R}$. Dann ist der *mean square error* MSE definiert als $\text{MSE}_\theta(\hat{\theta}) := E[(\hat{\theta} - \theta)^2]$, also als die erwartete quadratische Abweichung des Schätzers vom wahren Wert.

(a) Zeigen Sie, dass $\text{MSE}_\theta(\hat{\theta}) = \text{Bias}_\theta^2(\hat{\theta}) + \text{Var}_\theta(\hat{\theta})$ gilt.

(b) Es sei $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ mit $\mu \in \mathbb{R}$ und $\sigma^2 > 0$ unbekannt. Geschätzt werden soll der Parameter σ^2 . Geben Sie das statistische Modell, den Parameterraum und den Raum des zu schätzenden Parameters an. Berechnen Sie die Werte $\text{MSE}_\theta(\hat{\sigma}^2)$ und $\text{MSE}_\theta(s^2)$. Kommentieren Sie das Ergebnis.

HINWEIS: Sie können die Tatsache verwenden, dass $(n-1)s^2/\sigma^2 \sim \chi_{n-1}^2$ gilt.

2. Die Funktion $g: \mathbb{R}^k \rightarrow \mathbb{R}$ ist *konkav*, wenn für alle $x, y \in \mathbb{R}^k$ und für alle $t \in (0, 1)$

$$g(tx + (1-t)y) \geq tg(x) + (1-t)g(y)$$

gilt. Die Funktion ist *strikt konkav*, wenn obige Ungleichung strikt ist.

(a) Nehmen Sie an, dass g auf \mathbb{R}^k zweimal stetig differenzierbar ist und zeigen Sie, dass die Funktion strikt konkav ist, falls die Hessematrix negativ definit auf ganz \mathbb{R}^k ist.

(b) Zeigen Sie, dass in dem Fall ein stationärer Punkt $x_0 \in \mathbb{R}^k$ (falls er existiert) ein globales Maximum darstellt.

3. Zeigen Sie

$$E_{\theta_0} \left(\frac{\partial}{\partial \theta} \ln f(Z, \theta_0) \frac{\partial}{\partial \theta'} \ln f(Z, \theta_0) \right) = -E_{\theta_0} \left(\frac{\partial^2}{\partial \theta \partial \theta'} \ln f(Z, \theta_0) \right).$$

Dabei dürfen Sie die Reihenfolge von Differentiation und Integration vertauschen.

4. Überlegen Sie, in welchem Zusammenhang der LM-Test mit Lagrange-Multiplier stehen könnte.

5. Zeigen Sie, dass im klassischen linearen Regressionsmodell mit normalverteilten Fehlern der F -Test zur Hypothese $R\beta = r$ auch als Wald Test aufgefasst werden kann. Wie passt die asymptotische χ^2 -Verteilung ins Bild?

6. Für Folgen von Zufallsvariablen $\{X_n\}$ und $\{Y_n\}$ ($n \in \mathbb{N}$), die auf demselben W-Raum definiert sind, wiederholen Sie die Definitionen von *stochastisch beschränkt*, $X_n = O_P(Y_n)$ und $X_n = o_p(Y_n)$. Was heißt insbesondere $X_n = O_P(1)$, $X_n = O_P(\frac{1}{\sqrt{n}})$ und $X_n = o_p(1)$?

Aufgaben zu Kapitel 2

7. Wir betrachten die logistische (Standard-)Verteilung mit cdf $F(x) = \Lambda(x) = 1/(1 + e^{-x})$.

- (a) Berechnen Sie die Dichtefunktion $\lambda(x)$ und zeigen Sie, dass diese symmetrisch um den Ursprung ist.
- (b) Berechnen Sie Erwartungswert und Varianz.

HINWEIS: Die Berechnung der Varianz ist zB über die mgf $M_X(t)$ möglich. Dazu können folgende Identitäten hilfreich sein:

$$B(x, y) = \int_0^\infty \frac{u^{x-1}}{(1+u)^{x+y}} \text{ für } x, y > 0$$

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

$$\frac{\pi^2}{6} = \Gamma''(1) - \Gamma'(1)^2$$

- (c) Zeigen Sie, dass $\lambda(x) = \Lambda(x)(1 - \Lambda(x))$ gilt.
- (d) Plotten Sie die Dichtefunktion der logistischen Verteilung zusammen mit der Dichtefunktion einer Standardnormalverteilung.

8. Zeigen Sie, dass für unabhängige Gumbel-verteilte Zufallsvariablen u und v die Differenz $u - v$ einer logistischen Verteilung folgt.

HINWEIS: Die cdf einer Gumbel- (oder auch log-Weibull-) Verteilung ist durch $G(x) = \exp(-\exp(-x))$ gegeben.

9. Laden Sie den Datensatz *SwissLabor* aus dem R-package *AER*. Details über die Daten und die Verwendung der *glm*-Funktion können Sie dem dem entsprechenden Helpfile entnehmen). Regressieren Sie die Variable *participation* auf die restlichen Variablen des Datensatzes, sowie auf *age*². Verwenden Sie dabei sowohl Logit als auch Probit Modell.

10. Zeigen Sie, dass der *marginal effect* $\partial P(y_i = 1)/\partial x_{ij}$ im logit-Modell „sinnvoll“ durch $0.25\hat{\beta}_{ML,j}$ nach oben abgeschätzt werden kann.

11. Zeigen Sie, dass die Log-Likelihood Funktion $l_n(\beta)$ für das Probit Modell strikt konkav in β ist, falls die Regressormatrix X vollen Spaltenrank k hat.

HINWEIS: Eine Möglichkeit ist es zu zeigen, dass die Funktionen $\ln(\Phi(x))$ und $\ln(1 - \Phi(x))$ auf ganz \mathbb{R} strikt konkav sind.

12. Berechnen Sie im Binary-Response Modell den restringierten ML-Schätzer unter der Restriktion $\beta_2 = \dots = \beta_k = 0$, sowie den Wert der Log-Likelihood Funktion an dieser Stelle.

13. Berechnen Sie für den Datensatz aus Beispiel 9 jeweils für Logit und Probit Modell das Maß R_{RG}^2 .

14. Rechnen Sie für den Datensatz aus Beispiel 9 jeweils für Logit und Probit Modell nach, wie in R die Standardfehler und *p*-Werte für die Koeffizienten gebildet werden. Tun Sie dies nur mithilfe der Daten und dem entsprechenden ML-Schätzers.

Aufgaben zu Kapitel 3

15. Überlegen Sie ein Beispiel für ein Multinomial Logit Modell.
16. Geben Sie einen Schätzer für die VC-Matrix des ML-Schätzers im Conditional Logit Modell an. Stellen Sie die Verbindung zur heuristischen Herleitung der asymptotischen Verteilung des ML-Schätzers aus Kapitel 1 her.
17. Zeigen Sie, dass der Erwartungswert der Gumbelverteilung durch die Euler-Mascheroni Konstante γ gegeben ist und berechnen Sie auch die Varianz.
HINWEIS: $\gamma = -\int_0^\infty \ln xe^{-x} dx$, Beispiel 8.
18. Leiten Sie die *marginal effects* im CLM und im MLM her.
19. Berechnen Sie den Korrelationskoeffizienten der bivariaten Gumbelverteilung mit cdf $F(x, y) = \exp(-[\exp(-x/\rho) + \exp(-y/\rho)]^\rho)$ mit $\rho \in (0, 1]$.
HINWEIS: Achtung schwer.
20. Verwenden Sie den Datensatz *Fishing* aus dem R-package *mlogit* (siehe Datei fishing.pdf für eine Beschreibung). Fitten Sie ein CLM, z.B. mithilfe der Funktion *logit* wo Sie die Variable *mode* auf die Variablen *price* und *catch* regressieren. Das Modell enthält keine Konstante. Beachten Sie, dass die Regressoren “alternative-varying” sind, nicht aber die zu schätzenden Parameter. Schätzen Sie die *marginal effects*. Kommentieren Sie Ihre Ergebnisse.
HINWEIS: In der Funktion *mlogit* sollte die Variable *shape* auf den Wert *wide* gesetzt werden (bezieht sich auf die Struktur des Datensatzes), mit der Variable *varying* werden diejenigen Regressoren beschrieben, die “alternative-varying” sind.
21. Fitten Sie für den Datensatz *BankWages* (siehe entsprechendes Helpfile) aus dem R-package *AER* ein MLM indem Sie die Variable *job* auf *education* und *minority* regressieren. Führen Sie die Regression jeweils für ein subset des Datensatzes durch, einmal für *gender = male*, einmal für *gender = female*. Sie können für den fit auch die R-Funktion *multinom* verwenden. Schätzen Sie die *marginal effects*.
22. Überlegen Sie ein Beispiel für ein Nested Logit Modell.
23. Zeigen Sie, dass das Verhältnis der Wahrscheinlichkeiten der Wahl zweier verschiedener Alternativen innerhalb eines Nests der IIA-Eigenschaft gehorcht, nicht aber wenn die Alternativen in verschiedenen Nestern liegen. Das heißt, zeigen Sie dass

$$\frac{p_{ijl}}{p_{irs}}$$

für $j = r$ nicht von den anderen Alternativen abhängt, sehr wohl aber für $j \neq r$.

24. (a) Fitten Sie für den Datensatz *HC* (siehe entsprechendes Helpfile) aus dem R-package *mlogit* ein NLM indem Sie die Variable *depuar* auf die übrigen Variablen (ohne *income*) und alternativ-abhängige Konstante regressieren. Dabei sollten Sie 2 Nester, jeweils mit den Variablen *gcc*, *ecc*, *erc*, *hpc* (cooling) und *gc*, *ec*, *er* (nocooling) bilden. Da die Variable *icca* und *occa* nur für das Nest *cooling* Sinn ergeben, sollten die Werte für die restlichen Variablen auf Null gesetzt werden. (Achtung: das Datenset muss vorher mit der Funktion *mlogit.data* in die richtige Form gebracht werden.) Diskutieren Sie den output.

- (b) Geben Sie das Modell für v_{ijl} an, das dabei verwendet wird. Wieviele Parameter sind enthalten?

[Alternativ können Sie ein NLM für einen Datensatz Ihrer Wahl fitten!!]

Aufgaben zu Kapitel 4

25. Zeigen Sie, dass für eine gestutzte (*truncated*) normalverteilte Zufallsvariable $Y|Y > a$ mit $Y \sim N(\mu, \sigma^2)$ gilt:

(a) $E[Y|Y > a] = \mu + \sigma\lambda(\alpha)$, und

(b) $\text{Var}(Y|Y > a) = \sigma^2[1 + \alpha\lambda(\alpha) - \lambda^2(\alpha)] = \sigma^2[1 - \delta(\alpha)]$,

wobei $\lambda(\alpha) = \phi(\alpha)/(1 - \Phi(\alpha))$, $\delta(\alpha) = \lambda'(\alpha) = \lambda(\alpha)[\lambda(\alpha) - \alpha]$ und $\alpha = \frac{a-\mu}{\sigma}$.

26. Zeigen Sie, dass der *inverse Mill's ratio* $\lambda(a) = \phi(a)/(1 - \Phi(a))$ die Ungleichung

$$\lambda(a) > a$$

erfüllt und somit auch $\delta(a) > 0$ gilt.

HINWEIS: Beispiel 25(a).

27. Fitten Sie ein Standard-Tobit Modell für Tobin's Originaldatensatz. [Z.B. Datensatz *tobin* und Funktion *tobit* im package *AER*.]

28. Zeigen Sie, dass für eine *left-censored* normalverteilte Zufallsvariable Y , d.h. für

$$Y = \begin{cases} Y^* & \text{für } Y^* > a \\ a & \text{für } Y^* \leq a \end{cases}$$

mit $Y^* \sim N(\mu, \sigma^2)$ gilt:

(a) $E(Y) = \mu + \sigma[\Phi(\alpha)\alpha + \phi(\alpha)]$

(b) (*optional*) $\text{Var}(Y) = \sigma^2(1 - \Phi(\alpha))[1 - \delta(\alpha) + (\lambda(\alpha) - \alpha)^2\Phi(\alpha)]$,

wobei wieder $\alpha = \frac{a-\mu}{\sigma}$ gilt.

29. Berechnen Sie die *marginal effects* im Standard Tobit Modell

$$\frac{\partial E(y_i)}{\partial x_i}$$

30. Fitten Sie ein Standard-Tobit Modell für Tobin's Originaldatensatz. [Z.B. Datensatz *tobin* und Funktion *tobit* im package *AER*.] Vergleichen Sie die Ergebnisse sowohl mit dem Probit-Schätzer $\hat{\alpha}_{\text{PROB}}$ alleine, als auch mit dem Heckit-Schätzer $(\hat{\beta}'_{\text{HECK}}, \hat{\sigma}^2_{\text{HECK}})$.

31. Fitten Sie ein Tobit-II-Modell für den den *Mroz87* Datensatz (*labor supply of married women*) aus dem R-package *sampleSelection*. Die abhängige Variable ist das entsprechende Einkommen (*wage*). Die *participation equation* soll durch

$$lfp = \beta_1^{(1)} + age \beta_2^{(1)} + age^2 \beta_3^{(1)} + faminc \beta_4^{(1)} + kids \beta_5^{(1)} + educ \beta_6^{(1)} + u_1$$

beschrieben werden, wobei *kids* ein dummy dafür ist, ob es Kinder gibt oder nicht (die Variable muss im Datensatz erst definiert werden). Die Gleichung für das Einkommen soll dann durch

$$wage = \beta_1^{(2)} + \mathit{exper}\beta_2^{(2)} + \mathit{exper}^2\beta_3^{(2)} + \mathit{educ}\beta_4^{(2)} + \mathit{city}\beta_5^{(2)} + u_2$$

gegeben sein. (Für genauere Informationen über die Variablen aus dem Datensatz schauen Sie bitte ins entsprechende helpfile.) Geben Sie sowohl den ML-Schätzer als auch den Heckitschätzer an (beide können mit der R-Funktion *selection* berechnet werden).

32. Es sei $(X, Y)' \sim N(\mu, \Sigma)$ mit $\mu = (\mu_1, \mu_2)'$ und $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$.

- (a) $\xi := Y - \mu_2 - \frac{\sigma_{12}}{\sigma_1^2}(X - \mu_1)$ ist unabhängig von X und erfüllt $\xi \sim N(0, \sigma_2^2 - \sigma_{12}^2/\sigma_1^2)$.
 (b) Berechnen Sie $E(Y|X > a)$.
 (c) Berechnen Sie $\text{Var}(Y|X > a)$.

HINWEIS: (b) und (c): (a) und Beispiel 25.

33. Berechnen Sie die *marginal effects* im Tobit-V Modell.

HINWEIS: Erweitern Sie Beispiel 32, indem Sie auf das Ereignis $X \leq 0$ bedingen. Dazu können Sie verwenden, dass für $Z \sim N(\mu, \sigma^2)$ gilt, dass

$$\begin{aligned} E(Z|Z \leq 0) &= \mu - \sigma\lambda(\mu/\sigma) \\ \text{Var}(Z|Z \leq 0) &= \sigma^2(1 - \delta(\mu/\sigma)) \end{aligned}$$

34. Beschreiben Sie den Heckit-Schätzer im Tobit-V Modell. Welcher Parameter kann nicht geschätzt werden?

Aufgaben zu Kapitel 5

35. Zeigen Sie, dass für eine stetige nicht-negative Zufallsvariable T

$$E[T] = \int_0^\infty S(t)dt$$

gilt, wobei $S(t)$ die *survival function* bezeichnet.

36. Zeigen Sie, dass die Likelihoodfunktion für diskrete rechts-zensierte Daten durch

$$l_n(\lambda_1, \dots, \lambda_p) = \prod_{j=1}^p \lambda_j^{d_j} (1 - \lambda_j)^{r_j - d_j}$$

gegeben ist, wobei die Bezeichnungen aus der Vorlesung gelten. Berechnen Sie auch die entsprechenden ML-Schätzer $\hat{\lambda}_{\text{ML},j}$ und argumentieren Sie, warum es keine Korrelation zwischen diesen gibt.

37. Leiten Sie folgenden Varianzschätzer her

$$\widehat{\text{Var}}(\hat{\lambda}_{\text{ML},j}) = \frac{d_j(r_j - d_j)}{r_j^3}$$

38. Berechnen (plotten) Sie den Kaplan-Meier Schätzer für den Datensatz *aml* aus dem R-package *survival*, einmal für Patienten mit erhaltender Chemotherapie und einmal ohne (für Details über den Datensatz siehe entsprechendes helpfile). Geben Sie in Ihren Plots auch Konfidenzbänder mithilfe der Greenwood-Formel an ("*plain*" für Option *conf.type* der Funktion *survfit*).
39. Basierend auf dem Konfidenzintervall für $\ln(-\ln S(t))$, leiten Sie folgendes Konfidenzintervall für $S(t)$ her:
- $$\left(\hat{S}_{\text{KM}}(t)^{\exp\{z_{1-\alpha/2}\hat{\sigma}(t)\}}, \hat{S}_{\text{KM}}(t)^{\exp\{-z_{1-\alpha/2}\hat{\sigma}(t)\}} \right),$$
- wobei die Bezeichnungen aus der Vorlesung gelten.
40. Plotten Sie für verschiedene Werte der Parameter die Ausfallsrate des *Weibull*, des *generalized Weibull* und des *log-logistic models*.
41. Zeigen Sie, dass auch das *exponential* und das *log-logistic model* als AFT Modell aufgefasst werden können.
42. Simulieren Sie Daten $T_i \sim \text{Exp}(\gamma_i)$ mit $\gamma_i = e^{-x_i^\beta}$ ($\neq \text{VO!}$) und schätzen Sie die Parameter ein AFT-Modell. Korrigieren Sie den Bias des entsprechenden Schätzers und vergleichen Sie diesen mit dem wahren Wert. (Sie können vereinfacht auch $k = 1$ und $x_i = 1$ wählen.)
43. Zeigen Sie, dass das *exponential* und das *Weibull model* als *proportional hazard model* betrachtet werden können (verwenden Sie dabei das Weibull Modell in der Parametrisierung $F(x) = 1 - e^{-\gamma t^\alpha}$).
44. Fitten Sie ein Cox Proportional Hazard Modell für den Datensatz *ema1996.dta*. Details finden Sie im entsprechenden Kapitel von Cameron & Trivedi.

Aufgaben zu Kapitel 5

45. Berechnen Sie Erwartungswert und Varianz einer Poissonverteilung.
46. Fitten Sie ein Poisson-Regressionsmodell für den Datensatz *DoctorVisits* im R-package *AER*. Verwenden Sie zusätzlich die erklärende Variable *age*². Führen Sie auch einen Test auf *overdispersion* durch.
47. Berechnen Sie die Varianz einer negativen Binomialverteilung.