Jun.-Prof. Dr. Zakhar Kabluchko WS 2010/2011
Sebastian Lück 01.12.2010

Risikotheorie II
Übungsblatt 7

Aufgabe 1
Im verallgemeinerten linearen Modell $(g(\mathbb{E}Y_1), \ldots, g(\mathbb{E}Y_n))^T = X\beta$ wird der Zufallsvektor $(U_1(\beta), \ldots, U_m(\beta))^T$ mit

$$U_j(\beta) = \frac{\partial}{\partial \beta_j} \log L(Y, \beta)$$

Scorevektor und die Matrix $\left(I_{jk}(\beta)\right)_{j,k=1}^n = \left(\mathbb{E}(U_j(\beta)U_k(\beta))\right)_{j,k=1}^n$ Fisher-Informationsmatrix genannt. Sei g zweimal stetig differenzierbar mit $g'(x) \neq 0$.

(a) Zeige die Gültigkeit der Formeln

$$U_j(\beta) = \sum_{i=1}^n x_{ij}(Y_i - \mu_i) \frac{1}{\sigma_i^2(\beta) g'\left(g^{-1}(x_i^T \beta)\right)} \text{ und } I_{jk}(\beta) = \sum_{i=1}^n x_{ij} x_{ik} \frac{1}{\sigma_i^2(\beta)} \left(\frac{1}{g'\left(g^{-1}(x_i^T \beta)\right)} \right)^2.$$

(b) Zeige, dass sich obige Formeln im Falle der natürlichen Linkfunktion $g = (b')^{-1}$ wie folgt vereinfachen:

$$U_j(\beta) = \frac{1}{\tau^2} \sum_{i=1}^n x_{ij}(Y_i - \mu_i(\beta)) \text{ bzw. } U(\beta) = \frac{1}{\tau^2} X^T (Y - \mu(\beta)) \text{ und}$$

$$I_{jk}(\beta) = \frac{1}{\tau^4} \sum_{i=1}^n x_{ij} x_{ik} \sigma_i^2(\beta) \text{ bzw. } I(\beta) = \frac{1}{\tau^4} X^T V(\beta) X,$$

wobei $V(\beta) = \operatorname{diag}\left(\sigma_i^2(\beta)\right)$.

(c) Zeige, dass für die natürlich Linkfunktion gilt, dass

$$W(\beta) = \left(\frac{\partial^2}{\partial \beta_j \partial \beta_k} \log L(Y, \beta) \right)_{j,k=1}^n = -I(\beta).$$

Aufgabe 2
Gegeben sei ein verallgemeinertes lineares Modell mit natürlicher Linkfunktion.

(a) Gib die Matrizen $U(\beta)$ und $W(\beta)$ an, falls
 - die Stichprobenvariablen $Y_1, \ldots, Y_n$ gegeben sind mit $Y_i \sim N(\mu_i, \sigma^2), i = 1, \ldots, n$,
 - die Stichprobenvariablen $Y_1, \ldots, Y_n$ gegeben sind mit $Y_i \sim \mathrm{Bin}(1, \pi_i), i = 1, \ldots, n$,
 - die Stichprobenvariablen $Y_1, \ldots, Y_n$ gegeben sind mit $Y_i \sim \mathrm{Poi}(\lambda_i), i = 1, \ldots, n$.

(b) Leite den ML-Schätzer $\widehat{\boldsymbol{\beta}}$ für $\boldsymbol{\beta}$ her, falls $n = m$ und $rg(X) = m$.

Aufgabe 3

Eine Versicherung untersucht zur Tarifkalkulation in der KfZ-Sparte die erwartete Schadenanzahl in Abhängigkeit von der Altersgruppe des Versicherungsnehmers und dem Zulassungsort des KfZ (Stadt vs. Land). Unter der Annahme Poisson-verteilter Schäden mit orts- und altersabhängigem Parameter λ_i sollen die erwarteten Schadenzahlen über ein verallgemeinertes lineares Modell regressiert werden, um insbesondere einen altersunabhängigen Faktor für den Unterschied zwischen Stadt und Land zu ermitteln. Die beobachteten Schadenzahlen pro 1000 KfZ-Policen sind wie folgt gegeben:

Alter	<30	30-65	66-80	>80
Stadt	100	55	49	110
Land	80	35	37	90

(a) Formuliere ein geeignetes verallgemeinertes lineares Modell, so dass die Designmatrix X vollen Rang hat.

(b) Schreibe ein Programm, dass den ML-Schätzer $\hat{\beta}$ für β mit Hilfe des Newton-Verfahrens approximiert, d.h. nach Wahl eines geeigneten Startwerts $\hat{\beta}_0$ setzt man

$$\hat{\beta}_{k+1} = \hat{\beta}_k - W^{-1}(\hat{\beta}_k)U(\hat{\beta}_k),\ k \geq 1.$$

(c) Berechne den ML-Schätzer für den gegebenen Datensatz unter Wahl eines geeigneten Startwerts $\hat{\beta}_0$ und prüfe die Konvergenz des Verfahrens durch einen Plot des Iterationsverlaufs.

(d) Interpretiere die Rolle der Einflussfaktoren.