WS 09/10

Multivariate Analysis Übung 3

K. Utikal

1. für das regressions modell

$$y_i \sim \mathcal{N}(\mu_i,\ \sigma^2)$$
$$\mu_i = \mu(\mathbf{x}) = \beta_0 + \sum_j^p \beta_j x_{ij}$$

teste mit LR-methode die hypo $H_0:\ \beta_1 = \cdots = \beta_p = 0$

Lösung:

$$L(\beta,\sigma^2 \big|\ \mathbf{y}, X) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^n} \exp\left(-\frac{1}{2}\frac{(\mathbf{y}-X\beta)'(\mathbf{y}-X\beta)}{\sigma^2}\right)$$

log likelihood

$$l(\beta,\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\frac{(\mathbf{y}-X\beta)'(\mathbf{y}-X\beta)}{\sigma^2}$$

$\Longrightarrow$ mle:

$$\widehat{\beta} = (X'X)^{-1} X'\mathbf{y}$$
$$\widehat{\sigma}^2 = \frac{1}{n}\left(\mathbf{y}-X\widehat{\beta}\right)'\left(\mathbf{y}-X\widehat{\beta}\right)$$

$\Longrightarrow$ max.log.likelihood

$$l(\widehat{\beta},\widehat{\sigma}^2) = -\frac{n}{2}\log\widehat{\sigma}^2 - \frac{n}{2}$$

null hyp

$$H_0:\ \beta_1 = \cdots = \beta_p = 0$$

mle

$$\widehat{\beta}_0 = \overline{y},\ \widehat{\sigma}_0^2 = \widehat{s}^2$$

max.log.likelihood

$$l(\widehat{\beta}_0,\ \widehat{\sigma}_0^2) = -\frac{n}{2}\log\widehat{\sigma}_0^2 - \frac{n}{2}$$

LR-test

$$-2\log \mathrm{LR} = \log\left(\widehat{\sigma}_0^2/\,\widehat{\sigma}^2\right) \sim \chi^2(p)$$

2. für das anova modell

$$\begin{aligned} y_{ij} &\sim \mathcal{N}(\mu_{ij},\ \sigma^2) \\ \mu_{ij} &= \mu + \alpha_j \\ \textstyle\sum_{j=1}^{J} \alpha_j &= 0 \end{aligned}$$

finde die ML schätzer von $\mu_{ij},\ \mu,\ \alpha_j,\ j = 1, ..., J$. Das gleiche unter der bedingung: $\alpha_1 = 0$

Lösung: with design matrix

$$X^* = \begin{bmatrix} 1 & 1 & & \\ \vdots & \vdots & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ \vdots & & & \vdots \\ & & & 1 \\ & -1 & \cdots & -1 \\ \vdots & \vdots & & \vdots \\ 1 & -1 & \cdots & -1 \end{bmatrix} \quad \text{or } X^* = \begin{bmatrix} 1 & & & \\ \vdots & & & \\ & 1 & & \\ & \vdots & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ \vdots & & & \vdots \\ 1 & & & 1 \end{bmatrix}$$

mle

$$\begin{aligned} \widehat{\beta} &= \left(X^{*\prime} X^*\right)^{-1} X^{*\prime} \mathbf{y} \\ \widehat{\sigma}^2 &= \frac{1}{n} \left(\mathbf{y} - X\widehat{\beta}\right)' \left(\mathbf{y} - X\widehat{\beta}\right) \end{aligned}$$

coordinate wise solution:

$$\begin{aligned} l(\mu, \alpha) &= \sum_{j=1}^{J} \sum_{i}^{n_j} \left(y_{ij} - \mu - \alpha_j\right)^2 \\ \frac{\partial l}{\partial \alpha_j} &= \sum_{i}^{n_j} \left(y_{ij} - \mu - \alpha_j\right) = 0 \end{aligned}$$

$$\left.\begin{array}{r} \sum_i^{n_j} (y_{ij} - \mu - \alpha_j) = 0 \implies y_{\cdot j} - n_j\mu - n_j\alpha_j = 0 \implies \alpha_j = \overline{y}_{\cdot j} - \mu \\ \sum_j \alpha_j = 0 \implies \widehat{\mu} = \frac{1}{J} \sum_j \overline{y}_{\cdot j} \end{array}\right\}$$

$$\implies \left\{ \begin{array}{l} \widehat{\mu} = \frac{1}{J} \sum_j \overline{y}_{\cdot j} \\ \widehat{\alpha}_j = \overline{y}_{\cdot j} - \frac{1}{J} \sum_j \overline{y}_{\cdot j} \end{array} \right.$$

alternatively:

$$\left.\begin{array}{r}\sum_i^{n_j}(y_{ij}-\mu-\alpha_j)=0\\ \alpha_1=0\end{array}\right\}\implies\sum_i^{n_1}(y_{i1}-\mu)=0\implies\widehat{\mu}=\overline{y}_{\cdot 1}$$

$$\left.\begin{array}{r}\sum_i^{n_j}(y_{ij}-\mu-\alpha_j)=0,\quad j>1\\ \widehat{\mu}=\overline{y}_{\cdot 1}\end{array}\right\}\implies\widehat{\alpha}_j=\frac{1}{n_j}\sum_i^{n_j}(y_{ij}-\overline{y}_{\cdot 1})=\overline{y}_{\cdot j}-\overline{y}_{\cdot 1}$$

3. give an example of a 2x2x2 table for three binary variables X, Y, Z with conditional independence of the X, Y on Z but not joint independence of X, Y from Z and vice versa

solution:

$$\begin{aligned}\log\mu_{ijk}=\mu_0&+\sum\nolimits_{i>1}^{I}\alpha_i x_{1i}+\sum\nolimits_{j>1}^{J}\beta_j x_{2j}+\sum\nolimits_{k>1}^{K}\gamma_k x_{3k}\\ &+\sum\nolimits_{i>1,j>1}^{I,J}(\alpha\beta)_{ij}\,x_{1i}x_{2j}+\sum\nolimits_{i>1,k>1}^{I,K}(\alpha\gamma)_{ik}\,x_{1i}x_{3k}+\sum\nolimits_{j>1,k>1}^{J,K}(\beta\gamma)_{jk}\,x_{2j}x_{3k}\\ &+\sum\nolimits_{i>1,j>1,k>1}^{I,J,K}(\alpha\beta\gamma)_{ijk}\,x_{1i}x_{2j}x_{3k}\end{aligned}$$

in anov-notation

$$\log\mu_{ijk}=\mu_0+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}+(\alpha\beta\gamma)_{ijk}$$

cases

condl indep	$\alpha\perp\beta\mid\gamma$:	$\log\mu_{ijk}=\mu_0+\alpha_i+\beta_j+\gamma_k+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}$
joint indep	$\alpha\beta\perp\ \gamma$:	$\log\mu_{ijk}=\mu_0+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\gamma_k$

important note: *the model for joint indep is a special case of the model of condl independence in the case*

condl indep	$\alpha\perp\gamma\mid\beta$:	$\log\mu_{ijk}=\mu_0+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\beta\gamma)_{jk}$
joint indep	$\alpha\beta\perp\ \gamma$:	$\log\mu_{ijk}=\mu_0+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\gamma_k$

implies: $\alpha\perp\gamma\mid\beta$ and $\beta\perp\gamma\Leftrightarrow\alpha\beta\perp\ \gamma$
but $\alpha\beta\perp\ \gamma\ \not\Leftrightarrow\ \alpha\perp\beta\mid\gamma$

examples: ((death $\perp$ gender) | smoking), (smoking $\perp$ gender) $\Leftrightarrow$ (death,smoking) $\perp$ gender
but (death,smoking) $\perp$ gender $\not\Leftrightarrow$ (death $\perp$ gender) | smoking

4. finden sie die den ML-schätzer definierenden gleichungen für die binäre logistische regression

$$P(Y=1|\ x)=1-P(Y=0|\ x)=\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}$$

und konstruieren sie das newton verfahren

solution:

$$P\left(y|\ x\right)=\binom{n}{y}p^{y}\left(x\right)\left(1-p\left(x\right)\right)^{n-y}$$
$$p\left(x\right)=\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}$$

likelihood (one obs)

$$L(\alpha,\beta|\ x)=p(x)^{y}\left(1-p\left(x\right)\right)^{1-y}=\left(\frac{p\left(x\right)}{1-p\left(x\right)}\right)^{y}\left(1-p\left(x\right)\right)$$
$$=\left(\exp\left(\alpha+\beta x\right)\right)^{y}\left(\frac{1}{1+\exp\left(\alpha+\beta x\right)}\right)$$

log.likelihood

$$l(\alpha,\beta|\ x)=y(\alpha+\beta x)-\log\left(1+\exp\left(\alpha+\beta x\right)\right)$$

likelihood eqns

$$\frac{\partial l(\alpha,\beta|\ x)}{\partial\beta}=yx-\frac{x\exp\left(\alpha+\beta x\right)}{1+\exp\left(\alpha+\beta x\right)}=x\left(y-p\left(x\right)\right)$$
$$\frac{\partial l(\alpha,\beta|\ x)}{\partial\alpha}=y-\frac{\exp\left(\alpha+\beta x\right)}{1+\exp\left(\alpha+\beta x\right)}=\left(y-p\left(x\right)\right)$$

information:

$$\frac{\partial^{2}l(\alpha,\beta|\ x)}{\partial\beta^{2}}=-\frac{x^{2}\exp\left(\alpha+\beta x\right)\left(1+\exp\left(\alpha+\beta x\right)\right)-\left(x\exp\left(\alpha+\beta x\right)\right)^{2}}{\left(1+\exp\left(\alpha+\beta x\right)\right)^{2}}$$
$$=-\frac{x^{2}\exp\left(\alpha+\beta x\right)}{\left(1+\exp\left(\alpha+\beta x\right)\right)^{2}}=-x^{2}p(x)$$
$$\frac{\partial^{2}l(\alpha,\beta|\ x)}{\partial\alpha^{2}}=-\frac{\exp\left(\alpha+\beta x\right)}{\left(1+\exp\left(\alpha+\beta x\right)\right)^{2}}=-p(x)$$
$$\frac{\partial^{2}l(\alpha,\beta|\ x)}{\partial\alpha\partial\beta}=-\frac{x\exp\left(\alpha+\beta x\right)}{\left(1+\exp\left(\alpha+\beta x\right)\right)^{2}}=-xp(x)$$

note: information independent of y hence expected and observed information equal

Newton method: solution of

$$f\left(x\right)=0$$

as limit of the sequence

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(n)})}$$

solve likelihood eqn

$$f(\theta) = \frac{\partial l\left(\theta\right)}{\partial \theta}$$

as limit of sequence

$$\theta^{(n+1)} = \theta^{(n)} - \frac{\partial l\left(\theta^{(n)}\right)}{\partial \theta} \Bigg/ \frac{\partial l^2\left(\theta^{(n)}\right)}{\partial \theta^2}$$

relation to Fisher information

$$\mathbb{I}(\theta) = -\mathbb{E}\frac{\partial l^2\left(\theta\right)}{\partial \theta^2}$$

Newton method:

$$\theta^{(n+1)} = \theta^{(n)} + \frac{\partial l\left(\theta^{(n)}\right)}{\partial \theta} \Bigg/ \widehat{\mathbb{I}}(\theta^{(n)})$$

where

$$\widehat{\mathbb{I}}(\theta) = -\frac{\partial l^2\left(\theta\right)}{\partial \theta^2}$$

is the observed information matrix
using expected information matrix $\Longrightarrow$ Fisher scoring

$$\theta^{(n+1)} = \theta^{(n)} + \frac{\partial l\left(\theta^{(n)}\right)}{\partial \theta} \Bigg/ \mathbb{I}(\theta^{(n)})$$

sometimes better convergence
matrix form

$$\begin{aligned}
\theta^{(n+1)} &= \theta^{(n)} - \left[\frac{\partial l^2\left(\theta^{(n)}\right)}{\partial \theta_i \partial \theta_j}\right]^{-1} \nabla l\left(\theta^{(n)}\right) \\
&= \theta^{(n)} + \widehat{\mathbb{I}}^{-1}\left(\theta^{(n)}\right) \nabla l\left(\theta^{(n)}\right) \\
\widehat{\mathbb{I}}\left(\theta\right) &= -\left[\frac{\partial l^2\left(\theta\right)}{\partial \theta_i \partial \theta_j}\right]
\end{aligned}$$

for binary logit

$$\begin{aligned}
\left[\frac{\partial l\left(\theta\right)}{\partial \theta_i}\right] &= \begin{bmatrix} x\left(y - p\left(x\right)\right) \\ \left(y - p\left(x\right)\right) \end{bmatrix} \\
-\left[\frac{\partial l^2\left(\theta\right)}{\partial \theta_i \partial \theta_j}\right] &= \begin{bmatrix} x^2 p\left(x\right) & x p\left(x\right) \\ x p\left(x\right) & p\left(x\right) \end{bmatrix}
\end{aligned}$$

(in this case: fisher information=observed information) hence

$$\begin{bmatrix} \beta^{(n+1)} \\ \alpha^{(n+1)} \end{bmatrix} = \begin{bmatrix} \beta^{(n)} \\ \alpha^{(n)} \end{bmatrix} +$$

$$\begin{bmatrix} \sum_i x_i^2 p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right) & \sum_i x_i p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right) \\ \sum_i x_i p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right) & \sum_i p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right) \end{bmatrix}^{-1} \begin{bmatrix} \sum_i \left(x_i y_i - p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right)\right) \\ \sum_i \left(x_i - p\left(x_i \middle|\ \alpha^{(n)}, \beta^{(n)}\right)\right) \end{bmatrix}$$

with

$$p\left(x \middle|\ \alpha, \beta\right) = \frac{\exp\left(\alpha + \beta x\right)}{1 + \exp\left(\alpha + \beta x\right)}$$

5. finden sie die likelihood eqns für das multinomiale regressionsmodell

multinomial:

$$\begin{aligned} p(y_{1t}, \ldots, y_{It}) &= \binom{n}{y_{1t} \cdots y_{It}} p_1^{y_{1t}}(x_t) \ldots p_1^{y_{It}}(x_t) \\ p_i(x) &= \frac{\exp\left(\alpha_i + \beta_i x\right)}{\sum \exp\left(\alpha_i + \beta_i x\right)} \\ \alpha_1 &= \beta_1 = 0 \end{aligned}$$

solution:

$$\left.\begin{array}{r} l(\mathbf{p}|\ \mathbf{y}) = \sum_i^I y_i \log p_i \\ \log p_i = \alpha_i + \beta_i x - \log\left(1 + \sum_{i>1} \exp\left(\alpha_i + \beta_i x\right)\right) \\ (\text{for } i > 1) \end{array}\right\} \Longrightarrow$$

$$\begin{aligned} l(\alpha, \beta|\ \mathbf{y}) &= \sum_t^T \sum_i^I y_{it} \log \frac{\exp\left(\alpha_i + \beta_i x_t\right)}{1 + \sum_{j>1} \exp\left(\alpha_j + \beta_j x_t\right)} = \\ &= \sum_t^T \sum_i^I \left(y_{it}\left(\alpha_i + \beta_i x_t\right) - y_{it} \log\left(1 + \sum_{j>1} \exp\left(\alpha_j + \beta_j x_t\right)\right)\right) \end{aligned}$$

scores

$$\begin{aligned} \frac{\partial}{\partial \beta_{i_0}} l_{Mnom}(\alpha, \beta|\ \mathbf{y}) &= \sum_t^T \left(y_{i_0 t} x_t - \sum_i^I y_{it} x_t \frac{\exp\left(\alpha_{i_0} + \beta_{i_0} x_t\right)}{1 + \sum_{j>1} \exp\left(\alpha_j + \beta_j x_t\right)} \right) \\ &= \sum_t^T x_t \left(y_{i_0 t} - p_{i_0}(x_t) \sum_i^I y_{it} \right) = \sum_t^T x_t \left(y_{i_0 t} - n p_{i_0}(x_t)\right) \end{aligned}$$

conjecture: similarly (see binary regression)

$$\frac{\partial}{\partial \alpha_{i_0}} l_{Mnom}(\alpha, \beta| \ \mathbf{y}) = \sum_t^T (y_{i_0 t} - np_{i_0}(x_t))$$

and

$$\frac{\partial^2}{\partial \beta_{i_0}^2} l_{Mnom}(\alpha, \beta| \ \mathbf{y}) = -\sum_t^T x_t^2 np_{i_0}(x_t)$$
$$\frac{\partial^2}{\partial \alpha_{i_0} \partial \beta_{i_0}} l_{Mnom}(\alpha, \beta| \ \mathbf{y}) = -\sum_t^T x_t np_{i_0}(x_t)$$
$$\frac{\partial^2}{\partial \alpha_{i_0}^2} l_{Mnom}(\alpha, \beta| \ \mathbf{y}) = -\sum_t^T np_{i_0}(x_t)$$

6. condl independence and joint independence: show

$$\left.\begin{array}{l} X,Z \text{ indep condl on } Y \\ Y,Z \text{ marginal indep} \end{array}\right\} \implies X,Y \text{ joint indep from } Z$$

$$\left.\begin{array}{l} X,Y \text{ joint indep from } Z \\ Y,Z \text{ marginal indep} \end{array}\right\} \implies X,Z \text{ indep condl on } Y$$

Proof: $\implies$
$P(x/z,y) = P(x/y)$
$P(x,y,z)/P(y,z) = P(x,y)/P(y)$
$P(x,y,z)P(y) = P(x,y)P(y,z)$
$P(x,y,z)P(y) = P(x,y)P(y/z)P(z) = P(x,y)P(y)P(z)$ (hypothesis)
$P(x,y,z)/P(z) = P(x,y)$
$\impliedby$ same argument

7. show by example, that without this condition (Y,Z marginal indep) both concepts in 5. are unrelated

solution X : babies, Y : families, Z : nests
$P(x/z,y) = P(x/y)$ but $P(x,y/z = \text{hi}) \neq P(x,y/z = \text{lo})$
also: $P(x,y/z) = P(x,y)$ does not imply condl independence
$P(x,y) = P(x,y/z) \implies P(x) = P(x/z), \ P(y) = P(y/z) \implies$ joint indep from Z would imply marginal independence.