Homework 5

蒋翌坤 20307100013

6.6

The model is $\text{logit}(\pi(x)) = \alpha$, we score 1-5 from worse to marked improvement. From software, we have $\hat{\alpha} = -1.02$ and $G^2 = 7.28$ with $df = 4$, p-value$= 0.12$. The standard pearson residual is $-1.47, -1.54, 0.22, 1.52, 1.24$, since the residual's absolute value are all less than 2, the model fits well and there is no lack of fit.

6.22

(a)

We know $G^2(M_j|M_k) + G^2(M_k|M_s) = G^2(M_j|M_s)$, $G^2(M_k|M_s) \geq 0$, so $G^2(M_j|M_k) \leq G^2(M_j|M_s)$

(b)

Since M_k also holds, model M_j fit good under α confidence level. We also know as $n \to \infty$, G^2 converges to χ^2_{df}, i.e., $P(G^2(M_j|M_k) > \chi^2_{df}(\alpha)) \leq \alpha$, df is the difference of parameters between M_j and M_k. v is the difference of parameters between M_1 and M_s, so $v \geq df$, $\chi^2_v(\alpha) \geq \chi^2_{df}(\alpha)$, so $P(G^2(M_j|M_k) > \chi^2_{df}(\alpha)) \leq P(G^2(M_j|M_k) > \chi^2_v(\alpha)) \leq \alpha$

(c)

$P(G^2(M_j|M_k) \leq \chi^2_v(\alpha)) = 1 - P(G^2(M_j|M_k) > \chi^2_v) \geq 1 - \alpha$, so the procedure has type I error no greater than α.

6.24(a)

$\pi_{ik} = \frac{\exp(\alpha+\beta x_i+\beta_k^Z)}{1+\exp(\alpha+\beta x_i+\beta_k^Z)}$, the likelihood function is $L = \prod_{i=1}^2 \prod_{k=1}^K \pi_{ik}^{n_{i1k}}(1-\pi_{ik})^{n_{i2k}}$, so the log-likelihood function is

$$
\begin{aligned}
l &= \sum_{i=1}^2 \sum_{k=1}^K \Big(n_{i1k} \log \pi_{ik} + n_{i2k} \log(1-\pi_{ik}) \Big) \\
&= \sum_{i=1}^2 \sum_{k=1}^K n_{i1k}(\alpha + \beta x_i + \beta_k^Z) - \sum_{i=1}^2 \sum_{k=1}^K n_{i+k} \log(1 + \exp(\alpha + \beta x_i + \beta_k^Z))
\end{aligned}
$$

For α, the sufficient statistic is $\sum_{i=1}^2 \sum_{k=1}^K n_{i1k}$; For β, the sufficient statistic is $\sum_{k=1}^K n_{11k}$; For β_k^Z, the sufficient statistic is $\sum_{i=1}^2 n_{i1k}$.

6.32

(a)

$\pi_i = \Phi(\sum_j \beta_j x_{ij})$, so the log-likelihood is

$$l(\beta) = \sum_i y_i \log \pi_i + (1-y_i)\log(1-\pi_i) = \sum_i y_i \log \Phi(\sum_j \beta_j x_{ij}) + (1-y_i)\log(1-\Phi(\sum_j \beta_j x_{ij}))$$

(b)

$$\frac{\partial l}{\partial \beta_j} = \sum_i \frac{y_i \phi(\sum_j \beta_j)}{\Phi(\sum_j \beta_j x_{ij})} x_{ij} - \frac{(1-y_i)\phi(\sum_j \beta_j x_{ij})}{1-\Phi(\sum_j \beta_j x_{ij})} x_{ij} = \sum_i \frac{(y_i - \pi_i)x_{ij}}{\pi_i(1-\pi_i)} \phi(\sum_j \beta_j x_{ij})$$

So the likelihood equation is $\sum_i \frac{(y_i-\hat{\pi}_i)x_{ij}}{\hat{\pi}_i(1-\hat{\pi}_i)}\phi(\sum_j \hat{\beta}_j x_{ij}) = 0$, we can see that $z_i = \frac{\phi(\sum_j \hat{\beta}_j x_{ij})}{\hat{\pi}_i(1-\hat{\pi}_i)}$.

For logistic model, $\phi(x) = \frac{e^x}{(1+e^x)^2}$, so $z_i = \frac{e^{\sum_j \hat{\beta}_j x_{ij}}}{(1+e^{\sum_j \hat{\beta}_j x_{ij}})^2} \Big/ \left(\frac{e^{\sum_j \hat{\beta}_j x_{ij}}}{1+e^{\sum_j \hat{\beta}_j x_{ij}}} \frac{1}{1+e^{\sum_j \hat{\beta}_j x_{ij}}}\right) = 1$