Solution proposal finals STK3100/4100-f15

Problem 1

a) The frequency function of a binomially distributed variable is

$$f(y;\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y} = \binom{n}{y}\exp(y\log(\pi/(1-\pi)) + n\log(1-\pi))$$

Thus $\theta = \log(\pi/(1-\pi))$, $a(\theta) = -n\log(1-\pi)$, $\phi = 1$ and $c(y,\phi) = \log\binom{n}{y}$.

The parameter θ is called the canonical parameter. The connection between the canonical parameter and the expectation is $E(y) = a'(\theta)$. If $\eta = x\beta'$ is the predictor, the link function defines the connection between the predictor and the expectation. Hence the canonical parameter can be expressed by the coefficients in the predictor, β.

b) The likelihood in a generalized linear model is $L(\theta) = \prod_{i=1}^n c(y_i,\phi)\exp(\frac{\theta_i y_i - a(\theta_i)}{\phi})$. Hence if $\check{\theta}$ and $\hat{\theta}$ are the fitted parameters in a saturated and another model the deviance Δ is -2 log likelihood ratio:

$$\Delta = 2\sum_{i=1}^n[(\check{\theta}_i - \hat{\theta}_i)y_i - a(\check{\theta}_i) + a(\hat{\theta}_i)]$$

For the binomial distribution $\check{\theta}_i = \log(y_i/(n_i - y_i))$, $\hat{\theta}_i = \log(\hat{\mu}_i/(n_i - \hat{\mu}_i))$, $a(\check{\theta}_i) = -n_i\log(1 - y_i/n_i)$ and $a(\hat{\theta}_i) = -n_i\log(1-\hat{\mu}_i/n_i)$, so

$$\Delta = 2\sum_{i=1}^n[y_i\log(y_i/\hat{\mu}_i) + (n_i - y_i)\log((n_i - y_i)/(n_i - \hat{\mu}_i))]$$

The most common use of the deviance is for comparing two nested models. Then the χ^2-distribution can be a good approximation. For use of the deviance as a goodness-of-fit measure the situation is more complicated and the χ^2 approximation can be bad.

Problem 2

a) Within the same hospital $e^{\hat{\beta}_1} = 1.67$ represents the predicted proportional increase of the odds of survival of having a benign tumor (level 2) with respect to having a malign tumor.

The predicted odds for survival within country j with benign tumor is

$$\frac{\hat{\pi}_{bj}}{1-\hat{\pi}_{bj}} = \begin{cases} e^{\hat{\beta}_0+\hat{\beta}_1} & \text{if } j = 1 \\ e^{\hat{\beta}_0+\hat{\beta}_1+\hat{\beta}_2} & \text{if } j = 2 \\ e^{\hat{\beta}_0+\hat{\beta}_1+\hat{\beta}_3} & \text{if } j = 3 \end{cases}$$

The predicted odds for survival within country j with malign tumor is

$$\frac{\hat{\pi}_{mj}}{1-\hat{\pi}_{mj}} = \begin{cases} e^{\hat{\beta}_0} & \text{if } j=1 \\ e^{\hat{\beta}_0+\hat{\beta}_2} & \text{if } j=2 \\ e^{\hat{\beta}_0+\hat{\beta}_3} & \text{if } j=3 \end{cases}$$

Thus, the odds ratios $\text{OR} = \frac{\hat{\pi}_{bj}}{1-\hat{\pi}_{bj}} / \frac{\hat{\pi}_{mj}}{1-\hat{\pi}_{mj}} = e^{\hat{\beta}_1}$ for all three countries $j = 1, 2, 3$ or $\hat{\beta}_1 = \log \text{OR}$.

b) The output below is a deviance table from fitting various binomial models. Fill out the positions indicated by a question mark.

```
Analysis of Deviance Table

Model 1: cbind(surv, nsurv) ~ fapp + fage + fcountry
Model 2: cbind(surv, nsurv) ~ fapp + fage + finfl + fcountry
Model 3: cbind(surv, nsurv) ~ fapp + finfl + fage * fcountry
Model 4: cbind(surv, nsurv) ~ fapp * finfl + fage * fcountry
Model 5: cbind(surv, nsurv) ~ fapp * finfl + fapp * fage + fage * fcountry
Model 6: cbind(surv, nsurv) ~ fapp * finfl * fage * fcountry
  Resid. Df Resid. Dev Df Deviance
1        30     33.198
2        29     33.197  1   0.0009
3        25     25.718  4   7.4790
4        24     25.511  1   0.2079
5        22     22.059  2   3.4519
6         0      0.000 22  22.0587
```

b) Use the formula that if factor A has a levels and factor B has b levels A*B means intercept +(a-1) main effects parameters of A + (b-1) main effects parameters of B and (a-1)(b-1) interactions. Hence, remembering that the intercept and the main effects of a factor can only be counted once in a model specification:

(i) model 2 has p= 1+ 1+ 2 +1+2=7 parameters so n-p=36-7=29
(ii) model 3 has p=1+1+1+2+2+4= 11 parameters. Hence $p_{mod3} - p_{mod2} = 11 - 7 = 4$
(iii) $25.718 - 25.511 = 0.207 \approx 0.0.2079$
(iv) model 6 has 36 parameters and model 5 has 1+ 1+1+1+ +2+2 +2+4=14 parameters so $p_{mod6} - p_{mod5} = 36 - 14 = 22$.

In the remaining parts of this problem consider the hypothesis

$$H_0 : \beta_2 + \beta_3 = -1 \text{ versus } H_a : \beta_2 + \beta_3 \neq -1$$

c) $\hat{\beta}_2 + \hat{\beta}_3 + 1 = -0.6616 - 0.4946 + 1 = -0.1562$
$Var(\hat{\beta}_2 + \hat{\beta}_3 + 1) = Var(\hat{\beta}_2) + Var(\hat{\beta}_3) + 2Cov(\hat{\beta}_2, \hat{\beta}_2) = 0.040 + 0.043 + 2 \times 0.021 = 0.125$ so $st.err_{\hat{\beta}_2+\hat{\beta}_3+1} = \sqrt{0.125} = 0.354$ and the Wald statistic is $-0.156/0.354 = -0.441$ which has a p-value $2P(Z \leq -0.441)) = 0.66$ for $Z \sim N(0,1)$, so the hypothesis is not rejected.

d) fcountry2 corresponds to a dummy variable, dum2, which is equal to 1 when the level of country is 2, i.e. hospital is in US, and 0 for all combinations, fcountry3 corresponds to a dummy variable, dum3, which is equal to 1 when the level of country is 3, i.e. hospital is in UK, and 0 for all combinations. Thus the model from part a) corresponds to a model $\beta_0+\beta_1 fapp+\beta_2 dum2+\beta 3dum3$. Using that $\beta_2+\beta_3 = 1$ the model under H_0 becomes $\beta_0+\beta_1 fapp+\beta_2 dum2+(-1-\beta_2)dum3 = \beta_0+\beta_1 fapp + \beta_2(dum2 - dum3) - dum3$. This can be fitted by specifying a model of the form $offset(-dum3) + \beta_1 fapp + \beta_2(dum2 - dum3)$. Here dum2-dum3 is a variable which is 0 for treatments which takes place in Japan, 1 for treatments in US and -1 for treatments in UX. The test now consists of comparing the two deviances, and using a χ^2_1 distribution as reference.

Problem 3

a)

$$y_i = X_i\beta + Z_i b_i + \varepsilon_i, \ i = 1, \ldots, 54$$

where

$$X_i = \begin{pmatrix} 1 & 1 & I_{[AVED\in\{7,8,9\}]} & I_{[AVED\in\{10,11,\ldots\}]} \\ 1 & 2 & I_{[AVED\in\{7,8,9\}]} & I_{[AVED\in\{10,11,\ldots\}]} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 6 & I_{[AVED\in\{7,8,9\}]} & I_{[AVED\in\{10,11,\ldots\}]} \end{pmatrix}$$

$$Z_i = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 6 \end{pmatrix}$$

of dimensions 6×4 and 6×2 respectively. The indicator function is denoted as $I_{[\cdot]}$ The fixed effects parameters are collected in the 4×1 vector $\beta = (\beta_0, \beta_1, \beta_2, \beta_1)'$. The random effect are the elements of the 2×1 vectors $b_i = (b1_i, b2_i)', i = 1, \ldots, 54$ which is binormally distributed with expectation $(0,0)'$ and covariance matrix D and are independent of the errors $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{i6})'$ where all the elements are independent $N(0, \sigma^2)$ distributed.

b) $(\hat{\beta}_1 \beta_1)/\widehat{std.err}_{\hat{\beta}_1}$ is approximately $N(0,1)$ distributed which implies that an approximately 95% confidence interval has boundaries $706.00 \pm 1.9639.55$.

c) A model not containing the random effect `YEAR` is a simplifivcation of the covariance structure. This can be performed by fitting models containing `YEAR` and not containing `YEAR` by REML and comparing the values of -2 log LR. But the approximating distribution is a linear combination of χ^2-distributions, in this case $\frac{1}{1}\chi_1^2 + \frac{1}{1}\chi_2^2$.

d) The covariance matrix of y_i is $Cov(Z_i b_i + \varepsilon_i = Z_i Cov(b_i) Z_i' + \sigma^2 I_6 = Z_i D Z_i' + \sigma^2 I_6$ which equals

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 6 \end{pmatrix} \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 6 \end{pmatrix}$$

$$= \begin{pmatrix} d_{11} + 2d_{12} + d_{22} & \cdots & d_{11} + 7d_{12} + 6d_{22} \\ \vdots & & \vdots \\ d_{11} + 7d_{12} + 6d_{22} & \cdots & d_{11} + 42d_{12} + 36d_{22} \end{pmatrix}$$

e) The hypothesis implies a simplification of the fixed effect structure. This can be performed by fitting the model from part a) by maximum likelihood, and also the simplified model

$$y_{ij} = \beta_0 + \beta_1 \times j + \beta_3(AVETD_2 + 2AVETD_2) + b1_i + j \times b2_i + \varepsilon_{ij}, j = 1, \dots, 6\ , i = 1, \dots, 54$$

also by maximum likelihood. Then one compares the values of -2 log LR. The approximating distribution a χ_1^2-distribution, since the hypothesis represents one restriction.

Also a Wald test along the lines described in part 1 c) can be used. The estimate of the covariance matrix of the estimators is listed in the output.