

MATH 523 - MIDTERM 2011: SOLUTIONS

- (a) We have by elementary results for expectation and variance $\mathbb{E}[Y] = \pi$, $\text{Var}[Y] = \pi(1 - \pi)/m$ and for $y \in \{0, 1/m, 2/m, \dots, (m-1)/m, 1\}$,

$$f_Y(y; \pi) = \exp \left\{ my \log \left(\frac{\pi}{1 - \pi} \right) + m \log(1 - \pi) + \log \left(\frac{m}{my} \right) \right\} = \exp \left\{ \frac{y\theta - \log(1 + e^\theta)}{1/m} + \log \left(\frac{m}{my} \right) \right\}$$

where $\theta = \log(\pi/(1 - \pi)) = \log(\mu/(1 - \mu))$ is the canonical parameter. Hence $\text{Var}[Y] = \mu(1 - \mu)/m$.

5 MARKS

- (b) The chosen link function g connects the linear predictor $\eta = x\beta$ to the conditional expected value of response Y , μ , that is $\eta = g(\mu)$. The canonical link function is the specific choice such that $\eta = g(\mu) = \theta$ for canonical parameter θ ; for (a) the canonical link is $g(t) = \log(t/(1 - t))$. Other plausible link functions include

- Probit: $g(t) = \Phi^{-1}(t)$, where $\Phi(\cdot)$ is the standard normal CDF.
- Complementary log log: $g(t) = \log(-\log(1 - t))$.

5 MARKS

- (c) Log likelihood is

$$l_n(\beta) = \sum_{i=1}^n \left[w_i \left(y_i \theta_i - \log(1 + e^{\theta_i}) \right) + \log \left(\frac{m_i}{m_i y_i} \right) \right]$$

where $\theta_i = h(x_i \beta)$ for some h related to the link function, $w_i = m_i$ is a fixed weight; as indicated in lectures, the convention is to take $\phi = 1$ for the binomial family (this is what R does), and to use the m_i as weights.

3 MARKS

- (d) Under the canonical link from (b), $\theta_i = x_i \beta$, so

$$l_n(\beta) = \sum_{i=1}^n \left[w_i \left(y_i x_i \beta - \log(1 + e^{x_i \beta}) \right) \right] + \text{const.}$$

hence, in a vector calculation, as

$$\frac{\partial \{y_i x_i \beta\}}{\partial \beta} = y_i x_i^\top \quad \frac{\partial \{e^{x_i \beta}\}}{\partial \beta} = x_i^\top e^{x_i \beta}$$

we have

$$\frac{\partial l_n(\beta)}{\partial \beta} = \sum_{i=1}^n w_i x_i^\top \left(y_i - \frac{e^{x_i \beta}}{1 + e^{x_i \beta}} \right) = \sum_{i=1}^n w_i x_i^\top (y_i - \pi_i)$$

3 MARKS

- (e) For a two level factor predictor $X \in \{0, 1\}$ say, we can derive the MLEs for the success probabilities for each level of X directly, and then use invariance of MLEs to obtain the parameter estimates for parameters in the linear predictor. To this end, let π_x denote the probability of success when $X = x$. we have that

$$\hat{\pi}_x = \frac{\sum_{i=1}^n z_i \mathbb{I}\{X_i = x\}}{\sum_{i=1}^n m_i \mathbb{I}\{X_i = x\}} \quad x = 0, 1.$$

that is, we aggregate the z and m values over all data for which the factor predictor X takes the value x . Moving to the GLM with canonical link, the standard linear predictor assumes

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 \mathbb{I}\{x_i = 1\}$$

which assumes a baseline group ($X = 0$) with parameter β_0 , and the second group ($X = 1$) with parameter $\beta_0 + \beta_1$. By invariance, we have

$$\hat{\beta}_0 = \log\left(\frac{\hat{\pi}_0}{1-\hat{\pi}_0}\right) \quad \hat{\beta}_1 = \log\left(\frac{\hat{\pi}_1}{1-\hat{\pi}_1}\right) - \log\left(\frac{\hat{\pi}_0}{1-\hat{\pi}_0}\right) = \log\left(\frac{\hat{\pi}_1/(1-\hat{\pi}_1)}{\hat{\pi}_0/(1-\hat{\pi}_0)}\right)$$

- β_0 is the **log-odds** on cure for the baseline group.
- β_1 is the **log-odds ratio** on cure comparing the second group to the baseline group.

6 MARKS

(f) The analysis in OUTPUT 1 demonstrates the following:

- By inspection of the parameter estimates table, we see that `Dose` is highly significant, with a z value of 7.321. The sign of the coefficient is positive, indicating an increasing (log) odds on cure as dose level increases.
- In the **standard** deviance test, we compare the change in deviance when moving from null model to the model which includes dose. We have in this case:

$$\text{Change in deviance : } 147.3237 - 3.8835 = 143.4402 \quad \text{Change in rdf : } 7 - 6 = 1$$

so compare 143.4402 with the right tail of a χ^2_1 distribution. This result is hugely significant.

- In the **heuristic** version of deviance test, we look at the ratio of the change in deviance to the change in rdf, that is

$$\frac{\text{Change in deviance}}{\text{Change in rdf}} = 143.4402/1 = 143.4402$$

Under the deviance properties, if this number is far from 1, we can (heuristically) reject the simpler model as an adequate simplification.

The logic of this heuristic test is that if $Z \sim \chi^2_\nu$, then $\mathbb{E}[Z] = \nu$, so $\mathbb{E}[Z/\nu] = 1$, hence if you do not have tables available, you can still carry out a broad significance assessment.

- For overall adequacy of the model including dose level, we can inspect the estimated precision $\hat{\phi} = \text{Deviance}/\text{rdf} = 3.885/6 = 0.6475$. Again, heuristically, we compare this value to 1, and conclude that the model assumptions are adequate, although there may be mild evidence that the dispersion is less than 1.

6 MARKS

(g) For a model comparison between model 1 (main effects) and model 2 (main effects plus interaction), we use the standard deviance test as given in the question:

$$\text{Change in deviance : } 6.0427 - 3.3273 = 2.7154 \quad \text{Change in rdf : } 13 - 12 = 1$$

Comparing 2.7154 with a χ^2_1 distribution, we see that the relevant quantile is at 3.841, so this test is not significant at the $\alpha = 0.05$ level. Thus the effect of `Sex` on the dose effect does not appear to be significant. Examining the output for model 1, we see that both main effects coefficients are

significant, indicating that a unit increase in dose has the same effect in both females and males on the log-odds scale, but that Sex is important in the model (that is, the cure rates are different for females and males at the same dose level, but the log odds ratio on being cured for females versus males is constant across doses. Specifically, at dose d , the log odds from model 1 are

$$\text{Females: } -2.6451 + 0.0229d \qquad \text{Males: } -2.6451 + (-0.8488) + 0.0229d$$

yielding a log odds ratio at this dose of

$$(-2.6451 + 0.0229d) - (-2.6451 + (-0.8488) + 0.0229d) = 0.8488$$

which does not depend on d .

For overall adequacy of the model including main effects only, we can inspect the estimated precision $\hat{\phi} = \text{Deviance}/\text{rdf} = 6.0427/13 = 0.4648231$. We compare this value to 1, and conclude that the model assumptions are adequate, although again there may be mild evidence that the dispersion is less than 1.

8 MARKS