

MATH 523 - MIDTERM 2011: SOLUTIONS

(a) We have, if $\mathcal{Y} = \{0, 1, 2, \dots\}$,

$$f_Y(y; \pi) = \mathbb{1}_{\mathcal{Y}}(y) \exp\{y \log \pi + \log(1 - \pi)\}$$

so that the canonical parameter is $\theta = \log \pi$, and in the usual notation

$$b(\theta) = -\log(1 - e^\theta)$$

3 MARKS

(b) We have

$$\dot{b}(\theta) = \frac{e^\theta}{1 - e^\theta} \quad \ddot{b}(\theta) = \frac{e^\theta}{(1 - e^\theta)^2}$$

so therefore

$$\mathbb{E}[Y] = \frac{e^\theta}{1 - e^\theta} = \frac{\pi}{1 - \pi} \quad \text{Var}[Y] = \frac{e^\theta}{(1 - e^\theta)^2} = \frac{\pi}{(1 - \pi)^2}$$

and denoting $\mathbb{E}[Y] = \mu$, we have $\pi = \mu/(1 + \mu)$, and therefore

$$\text{Var}[Y] = \frac{\pi}{(1 - \pi)} \frac{1}{(1 - \pi)} = \mu(1 + \mu) = \phi V(\mu)$$

with $\phi = 1$.

6 MARKS

(c) The canonical link sets

$$\eta = \theta = \log \pi = \log \left(\frac{\mu}{1 + \mu} \right)$$

so $g(t) = \log(t/(1 + t))$.

5 MARKS

(d) As $0 < \pi < 1$, the canonical link is a mapping onto $\mathbb{R}^-$, rather than $\mathbb{R}$. This might be inconvenient, as the linear predictor could potentially take positive values. The given link function, however, maps onto $\mathbb{R}$.

Note: this link function is not available in $\mathbb{R}$.

4 MARKS

(e) The log likelihood in terms of π (common across all data) is

$$l_n(\pi) = \sum_{i=1}^n [y_i \log \pi + \log(1 - \pi)] = s \log \pi + n \log(1 - \pi)$$

where $s = \sum_{i=1}^n y_i$. By elementary calculations, we have

$$\hat{\pi} = \frac{s}{s + n} \quad \therefore \quad \hat{\mu}_0 = \frac{\hat{\pi}}{1 - \hat{\pi}} = \frac{s}{n}$$

by invariance.

3 MARKS

(f) Under the log link,

$$\eta = \beta_0 + \beta_1 x = \log \mu = \log \left(\frac{\pi}{1 - \pi} \right) \quad \therefore \quad \pi_i = \frac{\exp\{\beta_0 + \beta_1 x_i\}}{1 + \exp\{\beta_0 + \beta_1 x_i\}}.$$

and so, after some algebra

$$\dot{\pi}_{ij} = \frac{\partial \pi_i}{\partial \beta_j} = \begin{cases} \pi_i(1 - \pi_i) & j = 0 \\ x_i \pi_i(1 - \pi_i) & j = 1 \end{cases} = x_i^j \pi_i(1 - \pi_i).$$

We have

$$l_n(\beta_0, \beta_1) = \sum_{i=1}^n [y_i \log \pi(\beta_0, \beta_1, x_i) + \log(1 - \pi(\beta_0, \beta_1, x_i))] = \sum_{i=1}^n [y_i \log \pi_i + \log(1 - \pi_i)]$$

Therefore, for $j = 0, 1$,

$$\frac{\partial l_n}{\partial \beta_j} = \sum_{i=1}^n \dot{\pi}_{ij} \left[\frac{y_i}{\pi_i} - \frac{1}{1 - \pi_i} \right] = \sum_{i=1}^n x_i^j [y_i(1 - \pi_i) - \pi_i] = \sum_{i=1}^n \frac{x_i^j (y_i - \mu_i)}{1 + \mu_i}$$

as $\pi_i = \mu_i / (1 + \mu_i)$. Equating to zero yields the estimating equations.

5 MARKS

(g) Under the full model, $\hat{\mu}_i = y_i$, so that $\hat{\pi}_i = y_i / (1 + y_i)$. Therefore the deviance is

$$D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = 2 \sum_{i=1}^n \left[y_i \left\{ \log \left(\frac{y_i}{1 + y_i} \right) - \left(\frac{\hat{\mu}_i}{1 + \hat{\mu}_i} \right) \right\} + \left\{ \log \left(\frac{1}{1 + y_i} \right) - \left(\frac{1}{1 + \hat{\mu}_i} \right) \right\} \right]$$

4 MARKS

(h) This is the model described in (f), with a single covariate, and log link (the geometric distribution is a special case of the negative binomial). The estimates and standard errors indicate that the coefficient, β_1 is significantly different from zero; that is X influences failure, and as X increases, the expected lifetime **shortens** (the sign of $\hat{\beta}_1$ is negative)

4 MARKS

The change in deviance between the fitted model and the null model (corresponding to the fit given in (e)) is

$$84.730 - 61.232 = 23.498$$

which should be compared against the χ_1^2 density; clearly 23.498 is an atypical value from this density, so that allows us to conclude that the null model is not an adequate simplification of the fitted model.

4 MARKS

In this model, the dispersion parameter is $\phi = 1$. The estimate of the dispersion parameter under the fitted model is $\hat{\phi} = 61.232/48 = 1.276$ which is a little larger than 1, but not excessively so. Therefore there is no reason to question the adequacy of the model overall.

2 MARKS