MATH 523 - ASSIGNMENT 1

To be handed in not later than 5pm, 27th January 2011.
Please hand in during lectures, to Burnside 1225, or to the Mathematics Office Burnside 1005

1 The constrained Poisson model for random variable Y and parameter λ has mass function

$$f_Y(y;\lambda) = \frac{e^{-\lambda}y^{\lambda}}{(1-e^{-\lambda})y!} \qquad y = 1, 2, \ldots$$

for parameter $\lambda > 0$. Find the mean and variance of this distribution. Is it an Exponential (or Exponential-Dispersion) family distribution ? Justify your answer.

6 MARKS

2 Suppose that data $y_1, \ldots, y_n$ are independently generated from a Poisson generalized linear model, where

$$\mathbb{E}[Y_i] = \mu(x_i\beta) \qquad i = 1, 2, \ldots, n$$

where $x_i = (x_{i1}, \ldots, x_{ip})$ are a vector of predictors, and $\beta = (\beta_1, \ldots, \beta_p)^{\mathsf{T}}$ are real-valued parameters, and

$$x_i\beta = \sum_{j=1}^{p} x_{ij}\beta_j$$

(a) Derive the maximum likelihood estimating equations for estimating the components of β. Leave your answer in terms of the partial derivatives

$$\dot{\mu}_j(x_i\beta) = \frac{\partial\mu(x_i\beta)}{\partial\beta_j} \qquad j = 1, \ldots, p.$$

4 MARKS

(b) The canonical parameter for the Poisson model is $\theta = \log\mu$, and therefore the canonical link function is determined by

$$\eta = \log\mu$$

where $\eta = x\beta$ is the linear predictor. Using this result, simplify the estimating equations.

4 MARKS

(c) Using the `glm` function in `R` (or another suitable computing package), fit the models

(i) Straight-line: $\eta_i = \beta_0 + \beta_1 x_i$

(ii) Quadratic: $\eta_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$

to the following data and report the parameter estimates and standard errors in each case. The response data y record the number of reports of self-harm attempts reported at a crisis centre over an 18 week period, where x records week number.

Week	x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Count	y	15	11	14	17	5	11	10	4	8	10	7	9	11	3	6	1	1	4

6 MARKS

R COMMANDS FOR FITTING GENERALIZED LINEAR MODELS.

The function `glm` is the basic fitting model for generalized linear models in `R`. If the data are contained in vectors `Y`, `X1` and `X2`, then the command

```
glm(Y∼X1+X2,family=poisson)
```

fits the Poisson GLM

$$1 + X_1 + X_2$$

with canonical link $g(t) = \log t$, that is, it specifies that $Y_i \sim Poisson(\lambda_i)$, with

$$\log \lambda_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$$

- To remove the intercept:

  ```
  summary(glm(Y∼-1+X1+X2,family=poisson))
  ```

- The parameter estimates and standard errors etc. can be obtained using the `summary` function

  ```
  summary(glm(Y∼X1+X2,family=poisson))
  ```

- To obtain the parameter estimates alone, use `coef`:

  ```
  coef(glm(Y∼X1+X2,family=poisson))
  ```

- To obtain the fitted values, use `fitted`:

  ```
  fitted(glm(Y∼X1+X2,family=poisson))
  ```

- The glm function can be assigned to an object in `R`:

  ```
  model1<-glm(Y∼X1+X2,family=poisson))
  summary(model1)
  coef(model1)
  fitted(model1)
  ```

- If the data are in an `R` data frame, called `data.set` say, then `glm` can be used on the data frame

  ```
  glm(Y∼X1+X2,family=poisson,data=data.set)
  ```