MATH3091: Statistical Modelling II

Problem Sheet 4 (Solution)

1. Suppose $Y \sim \text{Exponential}(\lambda)$, with p.d.f.

$$f_Y(y;\lambda) = \lambda e^{-\lambda y}\,,$$

for $y > 0$ and $\lambda > 0$. Show that the exponential distribution is a member of the exponential family, and hence find $\mathbb{E}(Y)$, $\text{Var}(Y)$, and the variance function $V(\mu)$.

Solution: We have

$$f_Y(y;\lambda) = \lambda e^{-\lambda y} = \exp\left(-\lambda y + \log \lambda\right)\,.$$

This is of exponential family form, with $\theta = -\lambda$, $b(\theta) = -\log\lambda = -\log(-\theta)$, $a(\phi) = 1$ and $c(y,\phi) = 0$.

So

$$\mathbb{E}(Y) = b'(\theta) = \frac{d}{d\theta}\left(-\log(-\theta)\right) = -\frac{1}{\theta} = \frac{1}{\lambda} = \mu\,,$$

$$\text{Var}(Y) = a(\phi)b''(\theta) = 1 \times \frac{d}{d\theta}\left(-\theta^{-1}\right) = \theta^{-2} = \lambda^{-2}\,,$$

and the variance function is

$$V(\mu) = \theta^{-2} = [-\mu^{-1}]^2 = \mu^2\,.$$

2. Suppose $Y \sim \text{Geometric}(p)$, with p.d.f.

$$f_Y(y;p) = p(1-p)^{y-1}\,,$$

for $y \in \{1,2,3,\cdots\}$, $p \in (0,1)$. Show that the geometric distribution is a member of the exponential family, and hence find $\mathbb{E}(Y)$, $\text{Var}(Y)$, and the variance function $V(\mu)$.

Solution: We have

$$
\begin{aligned}
f_Y(y;p) &= p(1-p)^{y-1} \\
&= \exp\left(\log p + (y-1)\log(1-p)\right) \\
&= \exp\left(y\log(1-p) + \log\frac{p}{1-p}\right).
\end{aligned}
$$

This is of exponential family form, with $\theta = \log(1-p)$ (so $p = 1 - e^\theta$),

$$
b(\theta) = -\log\left(\frac{p}{1-p}\right) = -\log\left(\frac{1-e^\theta}{e^\theta}\right) = -\log(e^{-\theta}-1)\,,
$$

$a(\phi) = 1$ and $c(y,\phi) = 0$.

So

$$
\mathbb{E}(Y) = b'(\theta) = \frac{d}{d\theta}\left(-\log(e^{-\theta}-1)\right) = \frac{e^{-\theta}}{e^{-\theta}-1} = \frac{1}{1-e^\theta} = \frac{1}{p} = \mu\,,
$$

$$
\mathrm{Var}(Y) = a(\phi)b''(\theta) = 1 \times \frac{d}{d\theta}\left(\frac{1}{1-e^\theta}\right) = \frac{e^\theta}{(1-e^\theta)^2} = \frac{1-p}{p^2}\,,
$$

and the variance function is

$$
V(\mu) = \frac{e^\theta}{(1-e^\theta)^2} = \frac{1-\mu^{-1}}{(\mu^{-1})^2} = \mu(\mu-1)\,.
$$

3. What is the canonical link function for the exponential distribution (Q1)? What about for the geometric distribution (Q2)?

Solution:

For the **exponential distribution**, we have $b'(\theta) = -\theta^{-1} = \mu$, so the canonical link is

$$
g(\mu) = b'^{-1}(\mu) = -\mu^{-1}\,.
$$

For the **geometric distribution**, we have $b'(\theta) = (1-e^\theta)^{-1} = \mu$. Solving for θ, we find

$$
1 - e^\theta = \mu^{-1}\,.
$$

So

$$
\theta = \log(1-\mu^{-1})\,,
$$

and the canonical link is

$$
g(\mu) = b'^{-1}(\mu) = \log(1-\mu^{-1})\,.
$$

4. Suppose that Y_i, $i = 1, \cdots, n$, are independent Poisson(μ_i) random variables, that x_i is an explanatory variable, and that

$$\log \mu_i = \beta_1 + \beta_2 x_i .$$

   a. Write down the log-likelihood function of β_1 and β_2 explicitly. Hence, derive a pair of simultaneous equations, the solution of which are the maximum likelihood estimates for $\beta = (\beta_1, \beta_2)^\top$.

   b. Express the above model in terms of $\eta = \mathbf{X}\beta$, where $\mathbf{X}$ is the appropriate $n \times 2$ matrix.

Solution:

a. We have

$$\mu_i = \mu_i(\beta) = \exp(\beta_1 + \beta_2 x_i) ,$$

and the p.d.f. for the i-th observation is

$$f_{Y_i}(y_i; \mu_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} .$$

The likelihood is

$$\mathcal{L}(\beta) = \prod_{i=1}^{n} \frac{\mu_i(\beta)^{y_i} e^{-\mu_i(\beta)}}{y_i!} .$$

The log-likelihood is

$$\ell(\beta) = \log \prod_{i=1}^{n} \frac{\mu_i(\beta)^{y_i} e^{-\mu_i(\beta)}}{y_i!} = \sum_{i=1}^{n} y_i \log \mu_i(\beta) - \sum_{i=1}^{n} \mu_i(\beta) - \sum_{i=1}^{n} \log y_i! .$$

Differentiating with respect to β_j, $j = 1, 2$, gives

$$\frac{\partial}{\partial \beta_j} \ell(\beta) = \sum_{i=1}^{n} \frac{y_i}{\mu_i(\beta)} \frac{\partial \mu_i}{\partial \beta_j} - \sum_{i=1}^{n} \frac{\partial \mu_i}{\partial \beta_j} ,$$

where

$$\frac{\partial \mu_i}{\partial \beta_1} = \exp(\beta_1 + \beta_2 x_i) = \mu_i(\beta) ,$$

and

$$\frac{\partial \mu_i}{\partial \beta_2} = x_i \exp(\beta_1 + \beta_2 x_i) = x_i \mu_i(\beta) .$$

So

$$\frac{\partial}{\partial \beta_1} \ell(\beta) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \mu_i(\beta) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \exp(\beta_1 + \beta_2 x_i) ,$$

and

$$\frac{\partial}{\partial \beta_2}\ell(\beta) = \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \mu_i(\beta) = \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \exp(\beta_1 + \beta_2 x_i)\,.$$

So the MLE $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)^\top$ satisfies

$$\sum_{i=1}^{n} \exp(\hat{\beta}_1 + \hat{\beta}_2 x_i) = \sum_{i=1}^{n} y_i$$

and

$$\sum_{i=1}^{n} x_i \exp(\hat{\beta}_1 + \hat{\beta}_2 x_i) = \sum_{i=1}^{n} x_i y_i\,.$$

b. Let $\eta = \mathbf{X}\beta$, where

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix},$$

then $Y_i \sim \text{Poisson}(\mu_i)$, where $\mu_i = \exp(\eta_i)$.