# Stat 461-561: Quiz 1

Monday 29th January 2007

- **Exercise 1**. Suppose $X_1, X_2, ..., X_n$ are independent identically distributed from an exponential distribution $f(x|\theta^*)$, that is $\theta^*$ is the true parameter. The exponential distribution admits the following density

$$f(x|\theta) = \theta^{-1} \exp\left(-\theta^{-1}x\right) 1_{[0,\infty)}(x)$$

where $\theta \in (0, \infty)$.

**[461: 2 points, 561: 1 point]** Find the Maximum Likelihood Estimate (MLE) $\widehat{\theta}_n$ of $\theta$. Is the MLE biased?

We have

$$\log L(\theta|\mathbf{x}) = -n \log\theta - \left(\sum_{i=1}^n x_i\right)\theta^{-1}$$

so

$$\frac{\partial \log L(\theta|\mathbf{x})}{\partial\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\left(\sum_{i=1}^n x_i\right)$$

thus

$$\widehat{\theta} = \frac{\sum_{i=1}^n x_i}{n}.$$

This estimate is clearly unbiased as

$$\begin{aligned}
\mathbb{E}_\theta(X) &= \int_0^\infty x\theta^{-1}\exp\left(-\theta^{-1}x\right)dx \\
&= \theta\int_0^\infty u\exp(-u)\,du \\
&= \theta\left(\left[\frac{u^2}{2}\exp(-u)\right]_0^\infty + \int_0^\infty \exp(-u)\,du\right) \\
&= \theta.
\end{aligned}$$

**[461: 2 points, 561: 2 points]** Compute the Fisher information $I(\theta)$.

We have

$$\frac{\partial^2 \log f(x|\theta)}{\partial\theta^2} = \frac{1}{\theta^2} - \frac{2x}{\theta^3}$$

so

$$-\mathbb{E}_\theta\left(\frac{\partial^2 \log f(X|\theta)}{\partial\theta^2}\right) = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \frac{1}{\theta^2}.$$

**[461: 1 points, 561: 1 point]** What is the limiting distribution of $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$? How could you estimate $I(\theta^*)$ from the data?

We have

$$\sqrt{n}\left(\widehat{\theta}_n-\theta\right)\Rightarrow\mathcal{N}\left(0,I\left(\theta^*\right)^{-1}\right)$$

and

$$I\left(\theta^*\right)\approx-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2\log f\left(\left.x_i\right|\theta\right)}{\partial\theta^2}.$$

• **Exercise 2**. Suppose $Y_1,Y_2,...,Y_n$ are independent identically distributed from the following so-called zero-inflated Poisson distribution $g\left(\left.y\right|\theta^*\right)$, that is for $\theta=(\pi,\lambda)\in(0,1)\times(0,\infty)$

$$g\left(\left.y=0\right|\theta\right)=\pi+(1-\pi)\exp\left(-\lambda\right)$$

and for $i=1,2,3,...$

$$g\left(\left.y=i\right|\theta\right)=(1-\pi)\frac{\exp\left(-\lambda\right)\lambda^i}{i!}.$$

[**461: 2 points, 561: 1point**] Show that $g\left(\left.y\right|\theta\right)$ can be rewritten as a mixture, that is

$$g\left(\left.y\right|\theta\right)=\pi\delta_0\left(y\right)+(1-\pi)\,h\left(\left.y\right|\lambda\right)$$

where $\delta_0\left(y\right)=1$ if $y=0$ and $\delta_0\left(y\right)=0$ otherwise and $h\left(\left.y\right|\lambda\right)$ is a Poisson distribution of parameter $\lambda$.

This is trivial by the definition.

[**461: 3 points, 561: 2 points**] Derive the Expectation-Maximization algorithm to maximize the likelihood of the observations

$$L\left(\left.\theta\right|\mathbf{y}\right)=L\left(\left.\theta\right|y_1,...,y_n\right)=\sum_{k=1}^{n}\log g\left(\left.y_k\right|\theta\right).$$

We introduce $X_i\in\{1,2\}$ such that

$$f\left(\left.x_i,y_i\right|\theta\right)=f\left(\left.x_i\right|\theta\right)f\left(\left.y_i\right|\theta,x_i\right)=\begin{cases}\pi\delta_0\left(y_i\right) & \text{if } x_i=1,\\ (1-\pi)\,h\left(\left.y_i\right|\lambda\right) & \text{if } x_i=2.\end{cases}$$

So we have

$$Q\left(\theta,\widehat{\theta}_j\right)=\sum_{i=1}^{n}\left[\log\pi.\ f\left(\left.x_i=1\right|y_i,\widehat{\theta}_j\right)+\left(\log\left(1-\pi\right)-\lambda+y_i\log\lambda-y_i!\right)f\left(\left.x_i=2\right|y_i,\widehat{\theta}_j\right)\right]$$

and

$$\frac{\partial Q\left(\theta,\widehat{\theta}_j\right)}{\partial\pi}=\frac{\sum_{i=1}^{n}f\left(\left.x_i=1\right|y_i,\widehat{\theta}_j\right)}{\pi}-\frac{\sum_{i=1}^{n}f\left(\left.x_i=2\right|y_i,\widehat{\theta}_j\right)}{1-\pi}=0$$

thus

$$\widehat{\pi}_{j+1}=\frac{\sum_{i=1}^{n}f\left(\left.x_i=1\right|y_i,\widehat{\theta}_j\right)}{n}$$

where $f\left(x_i = 1 \middle| y_i, \widehat{\theta}_j\right) = 0$ if $y_j \neq 0$ and

$$f\left(x_i = 1 \middle| y_i = 0, \widehat{\theta}_j\right) = \frac{\widehat{\pi}_j}{\widehat{\pi}_j + (1 - \widehat{\pi}_j) \exp\left(-\widehat{\lambda}_j\right)}.$$

Now

$$\frac{\partial Q\left(\theta, \widehat{\theta}_j\right)}{\partial \lambda} = -\sum_{i=1}^{n} f\left(x_i = 2 \middle| y_i, \widehat{\theta}_j\right) + \frac{\sum_{i=1}^{n} y_i}{\lambda} = 0$$

so

$$\widehat{\lambda}_{j+1} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} f\left(x_i = 2 \middle| y_i, \widehat{\theta}_j\right)}$$

where

$$f\left(x_i = 2 \middle| y_i, \widehat{\theta}_j\right) = \begin{cases} \frac{(1-\widehat{\pi}_j)\exp\left(-\widehat{\lambda}_j\right)}{\widehat{\pi}_j + (1-\widehat{\pi}_j)\exp\left(-\widehat{\lambda}_j\right)} & \text{if } y_i = 0 \\ 1 & \text{otherwise} \end{cases}.$$

**[461: Optional 2 points, 561: 2 points]** Use Fisher's identity to compute

$$\frac{\partial \log L(\theta|\, \mathbf{y})}{\partial \theta} = \begin{pmatrix} \frac{\partial \log L(\theta|\mathbf{y})}{\partial \pi} \\ \frac{\partial \log L(\theta|\mathbf{y})}{\partial \lambda} \end{pmatrix}.$$

We have

$$\begin{aligned} \frac{\partial \log f(x_i, y_i|\, \theta)}{\partial \pi} &= \begin{cases} 1 & \text{if } x_i = 1 \text{ and } y_i = 0, \\ \frac{-1}{(1-\pi)} & \text{if } x_i = 2. \end{cases}, \\ \frac{\partial \log f(x_i, y_i|\, \theta)}{\partial \lambda} &= \begin{cases} 0 & \text{if } x_i = 1 \text{ and } y_i = 0, \\ -1 + \frac{y_i}{\lambda} & \text{if } x_i = 2. \end{cases} \end{aligned}$$

so

$$\begin{aligned} \frac{\partial \log L(\theta|\, \mathbf{y})}{\partial \pi} &= \sum_{i=1}^{n} f(x_i = 1|\, y_i, \theta) + \frac{\sum_{i=1}^{n} f(x_i = 2|\, y_i, \theta)}{(1-\pi)}, \\ \frac{\partial \log L(\theta|\, \mathbf{y})}{\partial \pi} &= \sum_{i=1}^{n} f(x_i = 2|\, y_i, \theta)\left(-1 + \frac{y_i}{\lambda}\right). \end{aligned}$$

• **Optional Exercise 3.** Suppose $X^1 = \left(X_1^1, X_2^1\right)$, $X^2 = \left(X_1^2, X_2^2\right)$,..., $X^n = \left(X_1^n, X_2^n\right)$ are independent identically distributed from

$$f(x|\, \theta^*) = f(x_1, x_2|\, \theta^*).$$

**[461: For the glory 0 point, 561: 2 points]** Suppose your estimator is based on the maximization of the pseudo-log-likelihood function

$$PL\left(\theta \middle| x^1, ..., x^n\right) = \sum_{i=1}^{n} \left\{\log f\left(x_1^i \middle| \theta, x_2^i\right) + \log f\left(x_2^i \middle| \theta, x_1^i\right)\right\}.$$

Establish the expression of

$$PL(\theta) = \lim_{n\to\infty} \frac{1}{n} PL\left(\theta | x^1, ..., x^n\right).$$

Is $\theta^*$ a maximum of $PL(\theta)$?

We have by the law of large numbers

$$PL(\theta) = \int (\log f(x_1|\theta, x_2) + \log f(x_2|\theta, x_1)) f(x_1, x_2|\theta^*) dx_1 dx_2$$

and it is equivalent to maximize

$$\begin{aligned} M(\theta) &= PL(\theta) - \int (\log f(x_1|\theta^*, x_2) + \log f(x_2|\theta^*, x_1)) f(x_1, x_2|\theta^*) dx_1 dx_2 \\ &= \int \left( \log \left( \frac{f(x_1|\theta, x_2)}{f(x_1|\theta^*, x_2)} \right) + \log \left( \frac{f(x_2|\theta, x_1)}{f(x_2|\theta^*, x_1)} \right) \right) f(x_1, x_2|\theta^*) dx_1 dx_2 \\ &= \int \left( \int \log \left( \frac{f(x_1|\theta, x_2)}{f(x_1|\theta^*, x_2)} \right) f(x_1|\theta^*, x_2) dx_1 \right) f(x_2|\theta^*) dx_2 \\ &\quad + \int \left( \int \log \left( \frac{f(x_2|\theta, x_1)}{f(x_2|\theta^*, x_1)} \right) f(x_2|\theta^*, x_1) dx_2 \right) f(x_1|\theta^*) dx_1 \end{aligned}$$

We have $M(\theta^*) = 0$ and $M(\theta^*) \leq 0$ for $\theta \neq \theta^*$ as

$$\int \log \left( \frac{f(x_1|\theta, x_2)}{f(x_1|\theta^*, x_2)} \right) f(x_1|\theta^*, x_2) dx_1 \leq 0,$$

$$\int \log \left( \frac{f(x_2|\theta, x_1)}{f(x_2|\theta^*, x_1)} \right) f(x_2|\theta^*, x_1) dx_2 \leq 0.$$