Stat 461-561: Solutions Exercises 2

January 24, 2007

Exercise 7.20

• Let

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}.$$

So we have

$$\mathbb{E}\left(\widehat{\beta}_1\right) = \frac{\sum_{i=1}^n \mathbb{E}[Y_i]}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n \beta x_i}{\sum_{i=1}^n x_i} = \beta.$$

• We have

$$var\left[\widehat{\beta}_1\right] = \left(\sum_{i=1}^n x_i\right)^{-2} \sum_{i=1}^n var[Y_i]$$

where

$$var[Y_i] = \sigma^2$$

so

$$var\left[\widehat{\beta}_1\right] = \frac{n\sigma^2}{\left(\sum_{i=1}^n x_i\right)^2}.$$

For the MLE (see solutions Exercises 1), we have

$$var\left[\widehat{\beta}_{MLE}\right] = \frac{\sigma^2}{\left(\sum_{i=1}^n x_i^2\right)}.$$

It follows that

$$var\left[\widehat{\beta}_{MLE}\right] \leq var\left[\widehat{\beta}_1\right]$$

as

$$\left(\sum_{i=1}^n x_i\right)^2 = \left(\sum_{i=1}^n x_i.1\right)^2 \leq \sum_{i=1}^n x_i^2.\sum_{i=1}^n 1 = n\sum_{i=1}^n x_i^2$$

by Cauchy-Schwartz inequality.

Exercise 7.21

• Let

$$\widehat{\beta}_2 = \frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}$$

then

$$\begin{aligned}\mathbb{E}\left(\widehat{\beta}_2\right) &= \frac{1}{n}\sum_{i=1}^n \frac{\mathbb{E}[Y_i]}{x_i} = \frac{1}{n}\sum_{i=1}^n \frac{\beta x_i}{x_i} \\ &= \beta.\end{aligned}$$

• We have

$$var\left(\widehat{\beta}_2\right) = \frac{1}{n^2}\sum_{i=1}^{n}\frac{var\left[Y_i\right]}{x_i^2} = \frac{\sigma^2}{n^2}\sum_{i=1}^{n}\frac{1}{x_i^2}.$$

To show that

$$var\left(\widehat{\beta}_2\right) \geq var\left(\widehat{\beta}_{MLE}\right)$$

is easy as

$$n^2 = \left(\sum_{i=1}^{n}\frac{x_i^2}{x_i^2}\right)^2 \leq \left(\sum_{i=1}^{n}\frac{1}{x_i^2}\right)\left(\sum_{i=1}^{n}x_i^2\right)$$

by Cauchy-Schwartz. So

$$\begin{aligned} var\left(\widehat{\beta}_2\right) &= \frac{\sigma^2}{n^2}\sum_{i=1}^{n}\frac{1}{x_i^2} \\ &\geq \frac{\sigma^2}{\left(\sum_{i=1}^{n}\frac{1}{x_i^2}\right)\left(\sum_{i=1}^{n}x_i^2\right)}\sum_{i=1}^{n}\frac{1}{x_i^2} \\ &\geq \frac{\sigma^2}{\left(\sum_{i=1}^{n}x_i^2\right)} \\ &= var\left(\widehat{\beta}_{MLE}\right). \end{aligned}$$

We cannot show that

$$var\left(\widehat{\beta}_2\right) \geq var\left(\widehat{\beta}_1\right) \text{ or } var\left(\widehat{\beta}_2\right) \leq var\left(\widehat{\beta}_1\right)$$

as $var\left(\widehat{\beta}_2\right) = \infty$ and $var\left(\widehat{\beta}_1\right) < \infty$ if there exists i such that $x_i = 0$ and $\sum_{i=1}^{n} x_i \neq 0$ whereas $var\left(\widehat{\beta}_2\right) < \infty$ and $var\left(\widehat{\beta}_1\right) = \infty$ if $\sum_{i=1}^{n} x_i = 0$ but $x_i^2 > 0$ for $i = 1, ..., n$.

Exercise 7.27

• We have

$$\begin{aligned} &\log f\left(x_1, y_1, ..., x_n, y_n \middle| \beta, \tau_1, ..., \tau_n\right) \\ = &-\left(\beta+1\right)\sum_{i=1}^{n}\tau_i + \left(\sum_{i=1}^{n}y_i\right)\log\beta + \left(\sum_{i=1}^{n}\left(y_i+x_i\right)\log\left(\tau_i\right)\right) - \sum_{i=1}^{n}\log\left(x_i!y_i!\right) \end{aligned}$$

so by differentiating the log-likelihood with respect to β and τ_j

$$\begin{aligned} -\sum_{i=1}^{n}\tau_i + \frac{\left(\sum_{i=1}^{n}y_i\right)}{\beta} &= 0, \\ -\left(\beta+1\right) + \frac{\left(y_i+x_i\right)}{\tau_i} &= 0. \end{aligned}$$

So we obtain from the second equation

$$\widehat{\tau}_i = \frac{\left(y_i+x_i\right)}{\left(\widehat{\beta}+1\right)}.$$

Plugging this equation in the first equation, we obtain

$$-\sum_{i=1}^{n}\frac{(y_i+x_i)}{\left(\widehat{\beta}+1\right)}+\frac{\left(\sum_{i=1}^{n}y_i\right)}{\widehat{\beta}}=0$$

and the result follows

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

• In the limit, we have the three fixed point equations

$$\begin{aligned}\widehat{\beta} &= \frac{\sum_{i=1}^{n}y_i}{\widehat{\tau}_1+\sum_{i=2}^{n}x_i},\\ \widehat{\tau}_1 &= \frac{\widehat{\tau}_1+y_1}{\widehat{\beta}+1},\\ \widehat{\tau}_j &= \frac{x_j+y_j}{\widehat{\beta}+1}\text{ for } j=2,...,n.\end{aligned}$$

We can solve the 2nd equation, this yields

$$\widehat{\tau}_1\widehat{\beta}=y_1.$$

The 3rd equation yields

$$x_j+y_j=\widehat{\tau}_j\left(\widehat{\beta}+1\right)$$

and the first follows directly.

• Direct as calculations are similar to the first question, except that x_1 is omitted.

Exercise 7.29

• Follows directly from the definition of Poisson and multinomial distributions, i.e. we have

$$f\left(\mathbf{x},\mathbf{y}|\,\beta,\boldsymbol{\tau}\right)=\left[\prod_{i=1}^{n}\frac{e^{-m\beta\tau_i}\left(m\beta\tau_i\right)^{y_i}}{y_i!}\right]\frac{m!}{\prod_{i=1}^{n}x_i!}\prod_{i=1}^{n}\tau_i^{x_i}$$

(there is a typo in C&B p. 360 as $m!$ is not inside the product).

• We have

$$\begin{aligned}\log f\left(\mathbf{x},\mathbf{y}|\,\beta,\boldsymbol{\tau}\right) &= -m\beta\sum_{i=1}^{n}\tau_i+\left(\sum_{i=1}^{n}y_i\right)\log m\beta+\sum_{i=1}^{n}\left(y_i+x_i\right)\log\left(\tau_i\right)\\ &\quad -\sum_{i=1}^{n}\log\left(x_i!y_i!\right)-\log m!\end{aligned}$$

so by differentiating with respect to β and τ_i

$$\begin{aligned}-m\sum_{i=1}^{n}\tau_i+\frac{\left(\sum_{i=1}^{n}y_i\right)}{\beta} &= 0,\\ -m\beta+\frac{y_i+x_i}{\tau_i} &= 0.\end{aligned}$$

However, do not forget that we have the constraint $\sum_{i=1}^{n}\tau_i = 1$ so by plugging this into the first equation

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i},$$

and the result follows for $\widehat{\tau_i}$. An alternative, maybe "cleaner", way to do it consists of using a Lagrange multiplier technique as discussed during the lecture.

• We reuse a technique similar to Example 7.2.19 in C&B and the result follows.

Exercise 7.30

• We have

$$\begin{aligned} f(\mathbf{x}, \mathbf{z}|\, p) &= \prod_{i=1}^{n} f(x_i, z_i|\, p) \\ &= \prod_{i=1}^{n} f(z_i|\, p)\, f(x_i|\, p, z_i) \\ &= \prod_{i=1}^{n} [pf(x_i)]^{z_i} [(1-p)\, g(x_i)]^{1-z_i}. \end{aligned}$$

• We have

$$f(\mathbf{z}|\, p, \mathbf{x}) = \prod_{i=1}^{n} f(z_i|\, p, x_i)$$

where

$$f(z_i|\, p, x_i) = \frac{f(z_i, x_i|\, p)}{f(x_i|\, p)} = \frac{[pf(x_i)]^{z_i} [(1-p)\, g(x_i)]^{1-z_i}}{pf(x_i) + (1-p)\, g(x_i)}$$

so $Z_i|\, x_i, p$ is a Bernoulli with success probability $\frac{pf(x_i)}{pf(x_i)+(1-p)g(x_i)}$.

• We have

$$\begin{aligned} Q\left(p, p^{(r)}\right) &= \sum_{z_{1:n}} \log\left(\prod_{i=1}^{n} [pf(x_i)]^{z_i} [(1-p)\, g(x_i)]^{1-z_i}\right) \prod_{i=1}^{n} f(z_i|\, p, x_i) \\ &= \log(p) \sum_{i=1}^{n} f\left(z_i = 1|\, p^{(r)}, x_i\right) + \sum_{i=1}^{n} f\left(z_i = 1|\, p^{(r)}, x_i\right) \log f(x_i) \\ &\quad + \log(1-p) \sum_{i=1}^{n} f\left(z_i = 0|\, p^{(r)}, x_i\right) + \sum_{i=1}^{n} f\left(z_i = 0|\, p^{(r)}, x_i\right) \log g(x_i) \end{aligned}$$

By differentiating with respect to p

$$\frac{\sum_{i=1}^{n} f\left(z_i = 1|\, p^{(r)}, x_i\right)}{p} = \frac{\sum_{i=1}^{n} f\left(z_i = 0|\, p^{(r)}, x_i\right)}{1-p}$$

so

$$\begin{aligned} p^{(r+1)} &= \frac{\sum_{i=1}^{n} f\left(z_i = 1 \middle| p^{(r)}, x_i\right)}{\sum_{i=1}^{n} f\left(z_i = 1 \middle| p^{(r)}, x_i\right) + \sum_{i=1}^{n} f\left(z_i = 0 \middle| p^{(r)}, x_i\right)} \\ &= \frac{1}{n} \sum_{i=1}^{n} f\left(z_i = 1 \middle| p^{(r)}, x_i\right) \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{p^{(r)} f\left(x_i\right)}{p^{(r)} f\left(x_i\right) + \left(1 - p^{(r)}\right) g\left(x_i\right)}. \end{aligned}$$