7-2-2020

D empirical process 2nd edition

full ~~R~~ likelihood

E-step: $$E\left[\cancel{Dg}\log f(Y;\theta) \mid Y_{obs}, \theta^{(k)}\right]$$

$$\downarrow$$

$$E\left[\frac{\partial}{\partial\theta}\log f(Y;\theta) \mid Y_{obs}, \theta^{(k)}\right] = S_{(k)}(\theta)$$

$$\downarrow$$

$$E\left[\left(\frac{\partial}{\partial\theta}\right)^{\otimes 2}\log f(Y;\theta) \,\Big|\, Y_{obs}, \theta^{(k)}\right] \cancel{\neq I_{(k)}(\theta)}$$
$$= -I_{(k)}(\theta)$$

M-step: Compute
$$\theta^{(k+1)} = \arg\max_{\theta} E\left[\log f(Y;\theta) \mid Y_{obs}, \theta^{(k)}\right]$$

$$\downarrow$$

$$\theta^{(k+1)} = \text{zero of } E\left[\frac{\partial}{\partial\theta}\log f(Y;\theta) \mid Y_{obs}, \theta^{(k)}\right]$$

$$\downarrow \text{ (approx by)}$$

$$\theta^{(k+1)} = \theta^{(k)} + \left[I_{(k)}(\theta^{(k)})\right]^{-1} S_{(k)}(\theta^{(k)})$$

Theorem 5.5. $\log f(Y_{obs};\theta^{(k+1)}) \geq \log f(Y_{obs};\theta^{(k)})$, with equality ~~only holding if~~ holding if and only if $\theta^{(k+1)} = \theta^{(k)}$.

Example 5.2. Suppose $Y \sim P_\theta$, where

$$p_\theta(y) = \{p\lambda e^{-\lambda y} + (1-p)\mu e^{-\mu y}\} I\{y \geq 0\}.$$

91. 2-2-2020

This is a mixture model, and we now introduce a complete model that leads to easier likelihood estimation: $X = (Y, \Delta) \sim P_\theta(X)$, where

$$P_\theta(X) = (p\lambda e^{-\lambda Y})^{\delta}\left((1-p)\mu e^{-\mu Y}\right)^{1-\delta} \quad \left(= f(X;\theta)\right)$$

$$\log P_\theta(X) = \Delta\left(\log p + \log\lambda - \lambda Y\right) + (1-\Delta)\left(\log(1-p) + \log\mu - \mu Y\right)$$

$$E\left[\log P_\theta(X) \,\middle|\, Y, \theta^{(k)}\right]$$

$$= E\left[\Delta \,\middle|\, Y, \theta^{(k)}\right]\left(\log p + \log\lambda - \lambda Y\right)$$

$$+ \left(1 - E\left[\Delta \,\middle|\, Y, \theta^{(k)}\right]\right)\left(\log(1-p) + \log\mu - \mu Y\right).$$

$$E\left[\Delta \,\middle|\, Y, \theta^{(k)}\right] = Pr\left(\Delta = 1 \,\middle|\, Y, \theta^{(k)}\right)$$

$$= \cancel{p\lambda e^{-\lambda Y}}$$

$$= \cancel{p_{(k)}\lambda e^{-\lambda_{(k)} Y}}$$

$$= \left.\frac{p\lambda e^{-\lambda Y}}{p\lambda e^{-\lambda Y} + (1-p)\mu e^{-\mu Y}}\right|_{\theta = \theta^{(k)}}$$

$$= \frac{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y}}{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y} + (1-p_{(k)})\mu_{(k)} e^{-\mu_{(k)} Y}}$$

 7-2-2020

Now, the full ~~se~~ likelihood score for p

is $\frac{\Delta}{p} - \frac{1-\Delta}{1-p}$,

the score for λ is $\Delta\left(\frac{1}{\lambda} - Y\right)$

and for μ is: $(1-\Delta)\left(\frac{1}{\mu} - Y\right)$.

~~This leads to~~ $\cancel{p_{(k+1)} = n^{-1}\sum_{i=1}^{n} \frac{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y_i}}{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y_i} + (1-p_{(k)})\mu_{(k)} e^{-\mu_{(k)} Y_i}}}$

Let $$\Delta_i(k) = \frac{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y_i}}{p_{(k)}\lambda_{(k)} e^{-\lambda_{(k)} Y_i} + (1-p_{(k)})\mu_{(k)} e^{-\mu_{(k)} Y_i}}.$$

This leads to:

$$p_{(k+1)} = n^{-1}\sum_{i=1}^{n} \Delta_i(k),$$

$$\lambda_{(k+1)} = \frac{\sum_{i=1}^{n} \Delta_i(k)}{\sum_{i=1}^{n} Y_i \Delta_i(k)},$$

$$\mu_{(k)} = \frac{\sum_{i=1}^{n} (1-\Delta_i(k))}{\sum_{i=1}^{n} Y_i (1-\Delta_i(k))}$$

7-2-2020

Let $f(\theta)$ be $K \geq 0$ times continuously differentiable, then

$$f(\theta+\Delta) = \sum_{j=0}^{K} \frac{f^{(j)}(\theta)\Delta^{\otimes j}}{j!} + o(\|\Delta\|^K)$$

$$\Rightarrow f(\theta+\Delta) - f(\theta) = \sum_{j=1}^{K} \frac{f^{(j)}(\theta)\Delta^{\otimes j}}{j!} + o(\|\Delta\|^K). \quad (*1)$$

Suppose $f(\theta)$ is continuous differentiable at θ, with $\dot{f}(\theta) \neq 0$, and suppose

$$r_n(\hat{\theta}_n - \theta) \rightsquigarrow Z \quad \text{(bdd in probability)},$$

for some sequence $r_n \to \infty$. Then

$$r_n(f(\hat{\theta}_n) - f(\theta)) \rightsquigarrow \dot{f}(\theta) Z.$$

Proof. By $(*1)$,

$$r_n(f(\hat{\theta}_n) - f(\theta)) = r_n\left(\dot{f}(\theta)(\hat{\theta}_n - \theta) + o_p(\|\hat{\theta}_n - \theta\|)\right)$$

$$= \underbrace{\dot{f}(\theta)\, r_n(\hat{\theta}_n - \theta)}_{\rightsquigarrow\ \dot{f}(\theta) Z} + \underbrace{o_p(r_n\|\hat{\theta}_n - \theta\|)}_{\rightsquigarrow\ 0}$$

Proof completed.

7-2-2020

Example. Let $X_1, \dots, X_n$ be iid mean 0 and variance 1. Then

$$\sqrt{n}(\bar{X}_n) \rightsquigarrow N(0,1) = Z.$$

What about convergence of $\cos \bar{X}_n$?

$$\cos \bar{X}_n - 1 = \sin(0)\,\bar{X}_n - \frac{\cos(0)}{2}\,\bar{X}_n^2 + o(|\bar{X}_n|^2)$$

$$= -\frac{1}{2}\bar{X}_n^2 + o(|\bar{X}_n|^2)$$

$$\Rightarrow \quad n(\cos \bar{X}_n - 1) \rightsquigarrow -\frac{1}{2} Z^2.$$