STAT 543

HW 9 Solution

Spring 2016

#1. (a)

$$L(p,\lambda) = f(\underline{x}\mid p,\lambda) = \prod_{i=1}^{n}\left[pe^{-\lambda}+1-p\right]^{I[x_i=0]}\left[p\,\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}\right]^{I[x_i\neq 0]}$$

$$= \left(pe^{-\lambda}+1-p\right)^{\sum_1^n I[x_i=0]}\left[p^{\sum_1^n I[x_i\neq 0]}\, e^{-\lambda\cdot\sum_1^n I[x_i\neq 0]}\cdot\frac{\lambda^{\sum_1^n x_i}}{\prod_{i=1}^n x_i!}\right]$$

since $\lambda^{\sum_1^n x_i I[x_i\neq 0]} = \lambda^{\sum_1^n x_i}$ and $\prod_{i=1}^{n}\left(\frac{1}{x_i!}\right)^{I[x_i\neq 0]} = \frac{1}{\prod_{i=1}^n x_i!}$

$$= \left(pe^{-\lambda}+1-p\right)^{n_0}\left[p^{\,n-n_0}\, e^{-\lambda(n-n_0)}\cdot\frac{\lambda^{\sum_1^n x_i}}{\prod_1^n x_i!}\right] \quad \text{where} \quad n_0 = \sum_1^n I[x_i=0]$$

$$\Rightarrow\ l(p,\lambda) \equiv \log L(p,\lambda) = n_0\log\left(pe^{-\lambda}+1-p\right) + (n-n_0)\log p - \lambda(n-n_0)$$

$$+ \sum_1^n x_i\log\lambda - \sum_1^n \log x_i!$$

$$\Rightarrow\ \begin{cases} \dfrac{\partial l}{\partial p} = \dfrac{n_0\left(e^{-\lambda}-1\right)}{pe^{-\lambda}+1-p} + \dfrac{n-n_0}{p} \overset{\text{set}}{=} 0 \\[2ex] \dfrac{\partial l}{\partial \lambda} = \dfrac{-n_0 p e^{-\lambda}}{pe^{-\lambda}+1-p} - (n-n_0) + \dfrac{\sum_1^n x_i}{\lambda} \overset{\text{set}}{=} 0 \end{cases}$$

//

(b)

i) $$\mu_1 \equiv EX = 0 + \sum_{x=1}^{\infty} p\frac{e^{-\lambda}\lambda^{x}}{x!}\cdot x = p\cdot\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\lambda = p\lambda \overset{\text{set}}{=} \frac{1}{n}\sum_1^n x_i \equiv \hat{\mu}_1$$

$$\mu_2 \equiv EX^2 = 0 + \sum_{x=1}^{\infty} p\frac{e^{-\lambda}\lambda^{x}}{x!}\cdot x^2 = \sum_{x=1}^{\infty} p\cdot\frac{e^{-\lambda}\lambda^{x}}{x!}\,x(x-1) + \sum_{x=1}^{\infty} p\frac{e^{-\lambda}\lambda^{x}}{x!}\,x$$

$$= p\left[\sum_{x=2}^{\infty} \frac{e^{-\lambda}\cdot\lambda^{x-2}}{(x-2)!}\cdot\lambda^2 + \sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\cdot\lambda\right] = p\lambda(\lambda+1) \overset{\text{set}}{=} \frac{1}{n}\sum_{1}^{n} x_i^2 \equiv \hat{\mu}_2 .$$

$$\Rightarrow \begin{cases} \tilde{p} = \dfrac{\hat{\mu}_1^2}{\hat{\mu}_2 - \hat{\mu}_1} \\ \\ \tilde{\lambda} = \dfrac{\hat{\mu}_2}{\hat{\mu}_1} - 1 \end{cases}$$

$\Rightarrow$ By the multivariate CLT,
$$\sqrt{n}\begin{bmatrix} \hat{\mu}_1 - p\lambda \\ \hat{\mu}_2 - p\lambda(\lambda+1)\end{bmatrix} \xrightarrow{d} N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \Sigma\right)$$

where
$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1, X_1^2) \\ Cov(X_1, X_1^2) & Var(X_1^2)\end{bmatrix} = p\lambda\begin{bmatrix} 1+(1-p)\lambda & (1-p)\lambda^2+(3-p)\lambda+1 \\ (1-p)\lambda^2+(3-p)\lambda+1 & (1-p)\lambda^3+(6-2p)\lambda^2+(7-p)\lambda+1\end{bmatrix}$$

$\Rightarrow$ Letting
$$\underset{\sim}{g}(x,y) = \begin{bmatrix} \overbrace{\dfrac{x^2}{y-x}}^{\equiv g_1} \\ \underbrace{\dfrac{y}{x}-1}_{\equiv g_2}\end{bmatrix}, \quad g(\hat{\mu}_1, \hat{\mu}_2) = \begin{bmatrix}\tilde{p}\\ \tilde{\lambda}\end{bmatrix}, \quad g(\mu_1,\mu_2) = \begin{bmatrix}p\\ \lambda\end{bmatrix} \quad \text{and}$$

$$\nabla \underset{\sim}{g} = \begin{bmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y}\end{bmatrix} = \begin{bmatrix} \dfrac{x(2y-x)}{(y-x)^2} & -\dfrac{x^2}{(y-x)^2} \\ -\dfrac{y}{x^2} & \dfrac{1}{x}\end{bmatrix} \quad \cdots (*)$$

By the delta method,

$$\sqrt{n}\begin{bmatrix}\tilde{p}-p \\ \tilde{\lambda}-\lambda\end{bmatrix} = \sqrt{n}\left(g(\hat{\mu}_1,\hat{\mu}_2) - g(\mu_1,\mu_2)\right) \xrightarrow{d} N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \nabla \underset{\sim}{g}\, \Sigma\, \nabla \underset{\sim}{g}'\right)$$

where
$$\nabla \underset{\sim}{g} = \begin{bmatrix} \dfrac{2\lambda+1}{\lambda^2} & -\dfrac{1}{\lambda^2} \\ -\dfrac{\lambda+1}{p\lambda} & \dfrac{1}{p\lambda}\end{bmatrix} \quad \text{by } (*) \text{ and } \mu_1 = p\lambda,\ \mu_2 = p\lambda(\lambda+1)$$

$\therefore$ when $p=0.7$ and $\lambda=3$,
$$\sqrt{n}\begin{bmatrix}\tilde{p}-0.7\\ \tilde{\lambda}-3\end{bmatrix} \xrightarrow{d} N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0.3656 & -0.6667 \\ -0.6667 & 7.1429\end{bmatrix}\right)$$ //

(c)

$$\text{I)}\quad \frac{\partial^2 l}{\partial p^2} = \frac{-n_0(e^{-\lambda}-1)^2}{(pe^{-\lambda}+1-p)^2} - \frac{n-n_0}{p^2}$$

$$\frac{\partial^2 l}{\partial p \partial \lambda} = \frac{-n_0 e^{-\lambda}(pe^{-\lambda}+1-p) + n_0(e^{-\lambda}-1)pe^{-\lambda}}{(pe^{-\lambda}+1-p)^2} = \frac{-n_0 e^{-\lambda}}{(pe^{-\lambda}+1-p)^2}$$

$$\frac{\partial^2 l}{\partial \lambda^2} = \frac{n_0 pe^{-\lambda}(pe^{-\lambda}+1-p) - n_0(pe^{-\lambda})^2}{(pe^{-\lambda}+1-p)^2} - \frac{\sum_1^n x_i}{\lambda^2} = \frac{n_0 pe^{-\lambda}(1-p)}{(pe^{-\lambda}+1-p)^2} - \frac{\sum_1^n x_i}{\lambda^2}$$

Note that $I_{[x_i=0]} \overset{iid}{\sim} Ber(pe^{-\lambda}+1-p) \Rightarrow E n_0 = n.EI_{[x_i=0]} = n.(pe^{-\lambda}+1-p)$

and $E\sum_1^n x_i = nEX_1 = np\lambda$

$$\Rightarrow \quad -E\left[\frac{\partial^2 l}{\partial p^2}\right] = \frac{n(pe^{-\lambda}+1-p)(e^{-\lambda}-1)^2}{(pe^{-\lambda}+1-p)^2} + \frac{np(1-e^{-\lambda})}{p^2} = \frac{n(e^{-\lambda}-1)^2}{pe^{-\lambda}+1-p} + \frac{n(1-e^{-\lambda})}{p}$$

$$-E\left[\frac{\partial^2 l}{\partial p \partial \lambda}\right] = \frac{ne^{-\lambda}}{pe^{-\lambda}+1-p}$$

$$-E\left[\frac{\partial^2 l}{\partial \lambda^2}\right] = -\frac{npe^{-\lambda}(1-p)}{pe^{-\lambda}+1-p} + \frac{np}{\lambda}$$

$$\Rightarrow \quad I_1^{-1}(p,\lambda) = \frac{1}{n}\begin{bmatrix} \frac{(e^{-\lambda}-1)^2}{pe^{-\lambda}+1-p} + \frac{1-e^{-\lambda}}{p} & \frac{e^{-\lambda}}{pe^{-\lambda}+1-p} \\ \frac{e^{-\lambda}}{pe^{-\lambda}+1-p} & -\frac{pe^{-\lambda}(1-p)}{pe^{-\lambda}+1-p} + \frac{p}{\lambda} \end{bmatrix}^{-1}$$

$$= \frac{1}{n}\begin{bmatrix} 0.2535 & -0.1865 \\ -0.1865 & 5.0850 \end{bmatrix}$$

Thus,

$$\sqrt{n}\begin{bmatrix} \hat{p}_{MLE} - 0.7 \\ \hat{\lambda}_{MLE} - 3 \end{bmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.2535 & -0.1865 \\ -0.1865 & 5.0850 \end{bmatrix}\right)$$

Compared with the variance of the MOM est. in (b),

the MLEs have smaller variance. //

(d) For given data, $\tilde{p} = 0.8229$ and $\tilde{\lambda} = 2.9167$.

$$l(p,\lambda) = 6\log(pe^{-\lambda}+1-p) + 14\log p - 14\lambda + 48\log\lambda + C, \quad C \text{ is a const.}$$

$$\Rightarrow \quad l' = \begin{bmatrix} \dfrac{6(e^{-\lambda}-1)}{pe^{-\lambda}+1-p} + \dfrac{14}{p} \\ \dfrac{-6pe^{-\lambda}}{pe^{-\lambda}+1-p} - 14 + \dfrac{48}{\lambda} \end{bmatrix}$$

and

$$l'' = \begin{bmatrix} \dfrac{-6(e^{-\lambda}-1)^2}{(pe^{-\lambda}+1-p)^2} - \dfrac{14}{p^2} & \dfrac{-6e^{-\lambda}}{(pe^{-\lambda}+1-p)^2} \\ \dfrac{-6e^{-\lambda}}{(pe^{-\lambda}+1-p)^2} & \dfrac{6pe^{-\lambda}(1-p)}{(pe^{-\lambda}+1-p)^2} - \dfrac{48}{\lambda^2} \end{bmatrix}$$

Then, a one-step Newton improvement on $(\tilde{p}, \tilde{\lambda})$ is

$$\begin{bmatrix} \tilde{p}_1 \\ \tilde{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \tilde{p} \\ \tilde{\lambda} \end{bmatrix} - l''^{-1} l' \Big|_{\binom{p}{\lambda} = \binom{\tilde{p}}{\tilde{\lambda}}} = \begin{bmatrix} 0.7370 \\ 3.3055 \end{bmatrix}$$

//

(e)

For $(\hat{p}, \hat{\lambda}) = (0.72675, 3.30242)$, the large sample 90% C.I. for p and λ

are

$$\hat{p}_{MLE} \pm z_{0.95} \sqrt{\{-l''|^{-1}_{(\hat{p}_{MLE}, \hat{\lambda}_{MLE})}\}_{(1,1)}} = (0.5502, 0.9032)$$

and

$$\hat{\lambda}_{MLE} \pm z_{0.95} \sqrt{\{-l''|^{-1}_{(\hat{p}_{MLE}, \hat{\lambda}_{MLE})}\}_{(2,2)}} = (2.4636, 4.1412) \quad //$$

#2. (a)

$$E_\alpha X = \int [\alpha f_1(x) + (1-\alpha) f_0(x)] x \, dx = \alpha E_1 X + (1-\alpha) E_0 X = \alpha \cdot 1 + (1-\alpha) \cdot 0 = \alpha$$

$$E_\alpha X^2 = \int [\alpha f_1(x) + (1-\alpha) f_0(x)] x^2 \, dx = \alpha E_1 X^2 + (1-\alpha) E_0 X^2 = \alpha [1+1] + (1-\alpha)[0+1]$$

$$= 1 + \alpha$$

$$\therefore \quad Var_\alpha X = E_\alpha X^2 - (E_\alpha X)^2 = 1 + \alpha - \alpha^2 \quad //$$

(b)

$$L(\alpha) = f(x|\alpha) = \alpha f_1(x) + (1-\alpha) f_0(x) = \alpha [f_1(x) - f_0(x)] + f_0(x)$$

$$\Rightarrow l(\alpha) \equiv \log L(\alpha) = \log\left(\alpha [f_1(x) - f_0(x)] + f_0(x)\right)$$

$$\Rightarrow l'(\alpha) = \frac{f_1(x) - f_0(x)}{\alpha [f_1(x) - f_0(x)] + f_0(x)}, \quad \text{where} \quad \alpha \in [0, 1].$$

$$\Rightarrow l'(\alpha) = \begin{cases} > 0 & \text{if} \quad f_1(x) > f_0(x) \\ \le 0 & \text{if} \quad f_1(x) \le f_0(x) \end{cases}$$

Since $f_1(x) > f_0(x) \iff \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} > \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \iff x > 0.5$.

when $x > 0.5$, $l'(\alpha) > 0$ which means that $l(\alpha)$ is increasing in α

and when $x \le 0.5$, $l'(\alpha) \le 0$ which means that $l(\alpha)$ is decreasing in α.

$$\therefore \hat{\alpha} = \begin{cases} 1 & \text{if } x > 0.5 \\ 0 & \text{o.w.} \end{cases}$$

Then, $E\hat{\alpha} = P_\alpha(X > 0.5) = \int_{0.5}^{\infty} [\alpha f_1(x) + (1-\alpha) f_0(x)]\,dx$

$$= \alpha P_1(X > 0.5) + (1-\alpha) P_0(X > 0.5)$$

$$= \alpha(1 - \Phi(-0.5)) + (1-\alpha)(1 - \Phi(0.5))$$

$$= 0.6915\alpha + (1-\alpha)\,0.3085 = 0.3085 + 0.3830\alpha.$$

and

$$\mathrm{Var}\,\hat{\alpha} = E\hat{\alpha}^2 - (E\hat{\alpha})^2 = E\hat{\alpha}(1 - E\hat{\alpha}) = (0.3085 + 0.3830\alpha)(0.6915 - 0.3830\alpha)$$

$\Rightarrow$ $\hat{\alpha}$ is not unbiased for α, except for $\alpha = \frac{0.3085}{0.6170}$

(c) We know that for any $p \in [0,1]$, $p(1-p) \le \frac{1}{4}$ ... $(*)$

$$\Rightarrow MSE(\hat{\alpha}) = \mathrm{Var}(\hat{\alpha}) + \mathrm{Bias}(\hat{\alpha})^2 = P_\alpha(X > 0.5)[1 - P_\alpha(X > 0.5)] + [E\hat{\alpha} - \alpha]^2$$

$$\overset{(*)}{\le} \frac{1}{4} + (0.3085 + 0.6915\alpha)^2 \le 0.25 + 0.3085^2 \quad \text{as } 0 \le \alpha \le 1$$

$$MSE(X) = \mathrm{Var}(X) + \mathrm{Bias}(X)^2 = 1 + \alpha - \alpha^2 = 1 + \alpha(1-\alpha) \ge 1 > 0.25 + 0.3085^2 \ge MSE(\hat{\alpha})$$

$\therefore$ $\hat{\alpha}$ is biased but has smaller MSE than X. //

(d)

$$\ell'(\alpha) = \frac{f_1(x) - f_0(x)}{f(x|\alpha)}$$

$$\Rightarrow \quad I(\alpha) = E_\alpha\left[\left(\frac{f_1(x) - f_0(x)}{f(x|\alpha)}\right)^2\right] = \int_{-\infty}^{\infty} \frac{[f_1(x) - f_0(x)]^2}{f(x|\alpha)}\, dx. \quad //$$

(e)

By CLT, $\bar{X}_n \xrightarrow{d} N(\alpha, \frac{1}{n}(1+\alpha+\alpha^2))$ as $n \to \infty$

By large sample theory of MLE, $\hat{\alpha}_n \xrightarrow{d} N(\alpha, \frac{1}{nI(\alpha)})$ as $n \to \infty$

Figure 1 shows that $\frac{1}{I(\alpha)} < 1+\alpha-\alpha^2$

$\therefore$ $\hat{\alpha}_n$ is better than $\bar{X}_n$ under large sample consideration. //

(f) $I(\hat{\alpha}_n) = \frac{1}{1.2}$, $\hat{\alpha}_n = 0.4$

$\Rightarrow$ 90% C.I. is $\hat{\alpha}_n \pm z_{0.95}\sqrt{\frac{1}{nI(\hat{\alpha}_n)}} = 0.4 \pm 1.645 \cdot \sqrt{\frac{1.2}{20}}$

$\therefore$ $(-0.0029, \ 0.8029)$ //