SUGGESTED SOLUTION

Exam TMA4295 Statistical Inference · H2013

Sat 14.12.2013 0900-1300

Ref: Henning Omre / 03.01.2014

Problem 1

$$X \Rightarrow f(x|p) = (1-p)^{x-1} p \quad ; \quad x = 1, 2, \ldots \quad p \in [0,1]$$

$$\mathbb{X}_n : X_1, \ldots, X_n \text{ iid } f(x|p)$$

a)

$$f(x|p) = \left[\frac{p}{1-p}\right](1-p)^x = \underbrace{\frac{p}{1-p}}_{c(p)} \cdot \exp\{\underbrace{\ln(1-p)}_{\omega_1(p)} \cdot \underbrace{x}_{t_1(x)}\}$$

Hence in the exponential family of pmfs

For an exponential family the sufficient statistic for p is:

$$T_1(\mathbb{X}_n) = \sum_{i=1}^{n} t_1(X_i) = \sum_{i=1}^{n} X_i$$

The statistic is also complete since $\{\ln(1-p); p \in [0,1]\}$ contain an open supset of $\mathbb{R}^1$.

The likelihood function is:

$$L(p|\mathbb{x}_n) = \prod_{i=1}^{n} (1-p)^{x_i - 1} p = (1-p)^{\Sigma x_i - n} p^n$$

$$\frac{d}{dp} \log L(p|\mathbb{x}_n) = (\Sigma x_i - n)\frac{-1}{1-p} + n\frac{1}{p} = 0 \Rightarrow$$

$$(\Sigma x_i - n) = n\frac{1-p}{p} \Rightarrow p = n/\Sigma x_i = \bar{x}^{-1}$$

hence

$$\hat{p} = [\bar{X}]^{-1}$$

b) Due to the invariance property of MLE one has

$$\hat{\mu}(p) = \mu(\hat{p}) = \frac{1}{\hat{p}} = \bar{X}$$

$$\hat{\sigma}^2(p) = \sigma^2(\hat{p}) = \frac{(1-\hat{p})}{\hat{p}^2} = \bar{X}^2 - \bar{X}$$

Expectations of MLE:

$$E_p(\hat{p}) = E\left[\frac{1}{\bar{X}}\right] \underset{\uparrow}{>} \frac{1}{E[\bar{X}]} = \frac{1}{\frac{1}{p}} = p \quad \text{– Not unbiased!}$$

Jensens uneq for $g(x) = \frac{1}{x}$

$$E_p(\hat{\mu}) = E[\bar{X}] = \frac{1}{p} = \mu \quad \text{– Unbiased!}$$

$$E_p(\hat{\sigma}^2) = E[\bar{X}^2] - E[\bar{X}] = Var[\bar{X}] + [E[\bar{X}]]^2 - E[\bar{X}]$$

$$= \frac{1-p}{np^2} + \frac{1}{p^2} - \frac{1}{p} = \frac{n-1}{n}\frac{1-p}{p^2}$$

– Not unbiased.

Define

$$\tilde{p} = I[\bar{X} = 1] = \begin{cases} 1 & \text{if } \bar{X} = 1 \\ 0 & \text{else} \end{cases}$$

$$E_p[\tilde{p}] = E[I[\bar{X}=1]] = 1 \cdot Prob_p\{\bar{X}=1\} + 0 \cdot [1 - Prob_p\{\bar{X}=1\}]$$

$$= Prob_p\{\bar{X}=1\} = (1-p)^{1-1} p = p \quad \text{– Unbiased!}$$

c) Unbiased

$$\tilde{p} = I(X_1 = 1) = \begin{cases} 1 & \text{if } X_1 = 1 \\ 0 & \text{else} \end{cases}$$

First $(t_n - 1)$th $\quad t_n$th

Bin$(n-1, p)$

$\underbrace{0,100100001}\ =1$

Sufficient statistic:

$$T(\mathbb{X}_n) = \sum_{i=1}^{n} X_i \longrightarrow f(t_n|p) = \binom{t_n - 1}{n-1} p^{n-1} (1-p)^{(t_n-1)-(n-1)} \cdot p$$

$$= \binom{t_n - 1}{n-1} p^{n} (1-p)^{t_n - n}$$

↑ Neg-Bin dist!

$t_n = n, n+1, \ldots$

From Rao-Blackwell:

$$E_p[\tilde{p} | t_n] = 1 \cdot \text{Prob}_p\{X_1 = 1 | T(\mathbb{X}_n) = t_n\} + 0 \cdot [1 - \text{Prob}_p\{X_1 = 1 | T(\mathbb{X}_n) = t_n\}]$$

$$= \text{Prob}\{X_1 = 1 | \sum_{i=1}^{n} X_i = t\}$$

We have:

$$\text{Prob}_p\{X_1 = 1 | \sum_{i=1}^{n} X_i = t\} = \frac{\text{Prob}\{X_1 = 1 \cap \sum_{i=1}^{n} X_i = t\}}{\text{Prob}\{\sum_{i=1}^{n} X_i = t\}}$$

$$= \frac{\text{Prob}\{X_1 = 1\}\text{Prob}\{\sum_{i=2}^{n} X_i = t-1\}}{\text{Prob}\{\sum_{i=1}^{n} X_i = t\}}$$

$$= \frac{p \cdot \binom{t-2}{n-2} p^{n-1} (1-p)^{t-1-(n-1)}}{\binom{t-1}{n-1} p^{n} (1-p)^{t-n}} = \frac{\binom{t-2}{n-2}}{\binom{t-1}{n-1}}$$

$$= \frac{\frac{(t-2)!}{(n-2)!(t-2-(n-2))!}}{\frac{(t-1)!}{(n-1)!(t-1-(n-1))!}} = \frac{n-1}{t-1}$$

hence

$$p^*_{RB} = E_p[\tilde{p} | T(\mathbb{X}_n)] = (n-1) \cdot \left[\sum_{i=1}^{n} X_i - 1\right]^{-1}$$

Since $T(\underline{X}_n) = \sum_{i=1}^{n} X_i$ is a complete, sufficient statistic for p and p^*_{RB} is unbiased for p and function of $T(\underline{X}_n)$ only $\Rightarrow$ p^*_{RB} is UMVUE!

d) Recall

$$E_p[\hat{\mu}] = E_p[\bar{X}] = \frac{1}{p}$$

$$Var_p[\hat{\mu}] = E_p[\bar{X}] = \frac{1}{n} Var_p[X] = \frac{1-p}{n\,p^2}$$

From Cramér-Rao Theorem, all estimators μ^* with $E_p[\mu^*] = \mu = \frac{1}{p}$ have:

$$Var_p[\mu^*] \geq \frac{\left[\frac{d}{dp} E_p[\mu^*]\right]^2}{-n E_p\left[\frac{d^2}{d^2p} \log f(X|p)\right]}$$

$$= \frac{\left[\frac{d}{dp}\frac{1}{p}\right]^2}{-n E_p\left[-(X-1)\frac{1}{(1-p)^2} - \frac{1}{p^2}\right]}$$

$$= \frac{1/p^4}{n/(1-p)p^2}$$

$$= \frac{(1-p)}{n\,p^2} = Var_p[\hat{\mu}]$$

Hence $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$ is UMVUE for μ.

[Alternative:

Reparametrize: $f(x|p) \longrightarrow f(x|\mu) = \left(1-\frac{1}{\mu}\right)^{x-1} \cdot \frac{1}{\mu}$

C&R:

$$Var_\mu[\mu^*] = \frac{\left[\frac{d}{d\mu} E_\mu[\mu^*]\right]^2}{-n E_\mu\left[\frac{d^2}{d\mu^2} \log f(X|\mu)\right]}$$

$$= \frac{1}{n/\mu^2(1-\frac{1}{\mu})} = \frac{\mu^2(1-\frac{1}{\mu})}{n} \longrightarrow \frac{(1-p)}{n\,p^2}$$]

c) The MLE's $\hat{p}, \hat{\mu}$ and $\hat{\sigma}^2$ are asymptotic efficient.

Note that for $f(x|p)$ the Fisher information is:

$$I(p) = -E_p\left[\frac{d^2}{dp^2} \log f(\underline{X}|p)\right] = \left[p^2(1-p)\right]^{-1}$$

The Delta method provides:

For $\hat{\beta}$ MLE for β with $\beta = \tau(p)$

$$\sqrt{n}\left[\hat{\beta}-\beta\right] \xrightarrow[n\to\infty]{dist} Norm\left[0, \left[\frac{d}{dp}\tau(p)\right]^2 \left[I(p)\right]^{-1}\right]$$

From this:

$\hat{p} = [\bar{X}]^{-1}$ MLE for p with $p = \tau(p) = p$

$$\Rightarrow \sqrt{n}\left[\hat{p}-p\right] \xrightarrow[n\to\infty]{dist} Norm\left[0, p^2(1-p)\right]$$

$\hat{\mu} = \bar{X}$ MLE for μ with $\mu = \tau(p) = \frac{1}{p}$

$$\Rightarrow \sqrt{n}\left[\hat{\mu}-\mu\right] \xrightarrow[n\to\infty]{dist} Norm\left[0, \underbrace{\frac{1}{p^4} p^2(1-p)}_{\frac{1-p}{p^2}}\right]$$

$\hat{\sigma}^2 = \bar{X}^2 - \bar{X}$ MLE for σ^2 with $\sigma^2 = \tau(p) = \frac{1-p}{p^2}$

$$\Rightarrow \sqrt{n}\left[\hat{\sigma}^2-\sigma^2\right] \xrightarrow[n\to\infty]{dist} Norm\left[0, \underbrace{\frac{(2-p)^2}{p^6} p^2(1-p)}_{\frac{(2-p)^2(1-p)}{p^4}}\right]$$

Since p is unknown, in practice we use:

- $I(\hat{p})$ and $\left[\frac{d}{dp}\tau(p)\right]^2\Big|_{p=\hat{p}}$

or sometimes even

- $I'(\hat{p}) = -\frac{1}{n}\sum_i\left[\frac{d^2}{dp^2}\log f(x_i|p)\right]\Big|_{p=\hat{p}}$

f) Hypothesis

$$H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$$

we have LRT statistics

$$\lambda(\mathbb{X}_n) = \frac{L_0(p_0 | \mathbb{X}_n)}{L(\hat{p} | \mathbb{X}_n)} = \frac{(1-p_0)^{\sum x_i - n} p_0^n}{(1-\hat{p})^{\sum x_i - n} \hat{p}^n}$$

$$= \left[\frac{1-p_0}{1-\hat{p}}\right]^{\sum x_i - n} \left[\frac{p_0}{\hat{p}}\right]^n = \left[\frac{1-p_0}{1-\hat{p}}\right]^{n\left[\frac{1-\hat{p}}{\hat{p}}\right]} \left[\frac{p_0}{\hat{p}}\right]^n$$

For the finite sample case we have rejection region:

$$R_{c_n} : \left\{ \mathbb{X}_n \,\middle|\, \left[\frac{1-p_0}{1-\hat{p}}\right]^{n\left[\frac{1-\hat{p}}{\hat{p}}\right]} \left[\frac{p_0}{\hat{p}}\right]^n \leq c_n \right\}$$

difficult to find for finite n!!

In asymptotic LRT we use:

$$-2 \log \lambda(\mathbb{X}_n) \xrightarrow[n \to \infty]{dist} Chi(1)$$

The level α asymptotic LRT has rejection region:

$$R : \left\{ \mathbb{X}_n \,\middle|\, -2 \log \lambda(\mathbb{X}_n) \geq \chi^2_{1,\alpha} \right\}$$

$$= \left\{ \mathbb{X}_n \,\middle|\, -2n\left[\left[\frac{1-\hat{p}}{\hat{p}}\right] \log \frac{1-p_0}{1-\hat{p}} + \log \frac{p_0}{\hat{p}}\right] \geq \chi^2_{1,\alpha} \right\}$$

The corresponding asymptotic $(1-\alpha)$ confidence region for p is:

$$C_{1-\alpha}(\mathbb{X}_n) : \left\{ p \,\middle|\, -2n\left[\left[\frac{1-\hat{p}}{\hat{p}}\right] \log \frac{1-p}{1-\hat{p}} + \log \frac{p}{\hat{p}}\right] \leq \chi^2_{1,\alpha} \right\}$$

hence the set of p such that H_0 is NOT rejected - ie. accepted.

Problem 2

$$X \rightsquigarrow f(x|\lambda) = Pois(\lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad ; \quad x = 0, 1, \ldots \quad \lambda > 0$$

$$\mathbb{X} = (X_1, X_2, X_3) = (x_1, x_2, x_3)$$

$$f(\mathbb{x}|\lambda) = \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \cdot \frac{(2\lambda)^{x_2}}{x_2!} e^{-2\lambda} \cdot \frac{(3\lambda)^{x_3}}{x_3!} e^{-3\lambda}$$

a) Factorization theorem:

$$f(\mathbb{x}/\lambda) = \underbrace{\frac{2^{x_2} 3^{x_3}}{x_1! x_2! x_3!}}_{h(\mathbb{x})} \cdot \underbrace{e^{-6\lambda} \cdot \lambda^{\Sigma x_i}}_{g(\Sigma x_i|\lambda)}$$

The sufficient statistic for λ is $T(\mathbb{X}) = \sum_{i=1}^{n} X_i$

The MLE for λ is:

$$\hat{\lambda} = \underset{\lambda}{\text{argmax}} \{\log f(\mathbb{x}/\lambda)\}$$

$$\downarrow$$

$$\frac{d}{d\lambda}\left[\log h(\mathbb{x}) - 6\lambda + \Sigma x_i \log \lambda\right] = 0$$

$$\Downarrow \quad -6 + \Sigma x_i \cdot \frac{1}{\lambda} = 0$$

$$\Downarrow \quad \underline{\hat{\lambda} = \frac{1}{6} \sum_i X_i}$$

The characteristics of $\hat{\lambda}$ is:

$$E(\hat{\lambda}) = \frac{1}{6}\left[E(X_1) + E(X_2) + E(X_3)\right] =$$

$$= \frac{1}{6}\left[\lambda + 2\lambda + 3\lambda\right] = \lambda$$

hence $\hat{\lambda}$ is unbiased.

$$Var(\hat{\lambda}) = \frac{1}{6^2}\left[Var(X_1) + Var(X_2) + Var(X_3)\right]$$

$$= \frac{1}{6^2}\left[\lambda + 2\lambda + 3\lambda\right] = \frac{\lambda}{6}$$

From the Cramér-Rao Theorem, a lower bound for unbiased estimators λ^* for λ is:

$$Var(\lambda^*) \geq \frac{1}{-E_\lambda\left[\frac{d^2}{d\lambda^2} \log f(\underline{X}|\lambda)\right]}$$

$$= \left[-E_\lambda \frac{d^2}{d\lambda^2}\left[\log h(x) - 6\lambda + \sum_i X_i \log\lambda\right]\right]^{-1}$$

$$= \left[-E_\lambda\left(\sum_i X_i\left(-\frac{1}{\lambda^2}\right)\right)\right]^{-1}$$

$$= \left[\frac{6\lambda}{\lambda^2}\right]^{-1} = \frac{\lambda}{6} = Var(\hat{\lambda})$$

Hence $\hat{\lambda}$ is UMVUE for λ.

b) Recall

$$f(\underset{\sim}{x}|\lambda) = \frac{1^{x_1} 2^{x_2} 3^{x_3}}{x_1!\, x_2!\, x_3!} e^{-6\lambda} \lambda^{\Sigma x_i}$$

Define prior model:

$$\pi(\lambda;\alpha,\beta) = [\Gamma(\alpha)]^{-1} \beta^{-\alpha} \lambda^{\alpha-1} e^{-\frac{1}{\beta}\lambda} \quad ; \quad \lambda \geq 0, \ \alpha,\beta > 0$$

The posterior model for λ is:

$$\begin{aligned}\pi(\lambda|\underset{\sim}{x};\alpha,\beta) &= const^1_{\underset{\sim}{x}}\, f(\underset{\sim}{x}/\lambda)\,\pi(\lambda) \\ &= const^2_{\underset{\sim}{x}} \cdot e^{-6\lambda} \lambda^{\Sigma x_i} \lambda^{\alpha-1} e^{-\frac{1}{\beta}\lambda} \\ &= const^2_{\underset{\sim}{x}} \cdot \lambda^{\alpha+\Sigma x_i - 1} e^{-(\frac{1}{\beta}+6)\lambda}\end{aligned}$$

Hence

$$[\lambda|\underset{\sim}{x}] \rightsquigarrow \pi(\lambda|\underset{\sim}{x};\alpha,\beta) = Gam(\overbrace{\alpha + \Sigma_i x_i}^{\alpha'}, \overbrace{[\tfrac{1}{\beta}+6]^{-1}}^{\beta'})$$

and from 'T&F in Stat':

$$E[\lambda|\underset{\sim}{x}] = \alpha'\beta' = (\alpha + \Sigma_i x_i)\left[\frac{1}{\beta}+6\right]^{-1}$$

$$Var[\lambda|\underset{\sim}{x}] = \alpha'\cdot\beta'^2 = (\alpha + \Sigma x_i)\left[\frac{1}{\beta}+6\right]^{-2}$$

Various asymptotic cases:

i) $\alpha \to 0$, $\beta \to \infty$ and $\alpha\beta = \mu$ $[\beta^{-1} \to 0]$

Prior model $\pi(\lambda)$:

$$E[\lambda] = \alpha \cdot \beta = \mu$$

$$Var[\lambda] = \alpha \cdot \beta^2 = (\alpha \cdot \beta)\beta = \mu\beta \longrightarrow \infty$$

Hence $\pi(\lambda) \longrightarrow$ "$Uni[0, \infty]$" - uniform pdf

Estimates:

$$E[\lambda | x] \longrightarrow \sum x_i \cdot \frac{1}{6} = \hat{\lambda}$$

$$Var[\lambda | x] \longrightarrow \sum x_i \cdot \frac{1}{6^2} = \frac{\hat{\lambda}}{6}$$

Hence, estimators are MLE for λ and $Var[\hat{\lambda}]$.

ii) $\alpha \to \infty$, $\beta \to 0$ and $\alpha\beta = \mu$ $[\alpha^{-1} \to 0]$

Prior model $\pi(\lambda)$:

$$E[\lambda] = \alpha \cdot \beta = \mu$$

$$Var[\lambda] = \alpha \cdot \beta^2 = (\alpha\beta)\beta = \mu\beta \longrightarrow 0$$

Hence $\pi(\lambda) \longrightarrow$ "$Dir[\mu]$" - dirac pdf

Estimates:

$$E[\lambda | x] = (\alpha + \sum x_i)\left[\frac{1}{\beta} + 6\right]^{-1}$$

$$= \frac{\alpha\beta}{1+6\beta} + \frac{\beta}{1+6\beta}\sum x_i \longrightarrow \mu$$

$$Var[\lambda | x] = (\alpha + \sum x_i)\left[\frac{1}{\beta} + 6\right]^{-2}$$

$$= \frac{(\alpha\beta)\beta}{1+6\beta} + \frac{\beta^2}{1+6\beta}\sum x_i \longrightarrow 0$$

Hence, estimators are constants corresponding to an estimate at μ without variance which coinside with the prior model $\pi(\lambda)$.