SUGGESTED SOLUTION

Exam TMA 4295 Statistical Inference H2015

Sat 19.12.2015 0900-1300

Ref. Henning Omre / 05.01.2016

Problem 1

Pareto cdf

$$\underline{X} \sim f(x/\alpha,\beta) = Par(\alpha,\beta) = \begin{cases} \alpha^{\beta}\cdot\beta\cdot x^{-(\beta+1)} & ; \alpha \leq x \\ 0 & ; x < \alpha \end{cases}$$

$$\underline{X} \sim F(x/\alpha,\beta) = CPar(\alpha,\beta) = \begin{cases} 1-\alpha^{\beta}x^{-\beta} & ; \alpha \leq x \\ 0 & ; x < \alpha \end{cases}$$

with $\alpha > 0$, $\beta > 2$.

$$\mu = E\{\underline{X}\} = [\beta-1]^{-1}\beta\alpha$$

$$\sigma^2 = Var\{\underline{X}\} = [\beta-1]^{-2}[\beta-2]\beta\alpha^2$$

Random sample:

$\underline{\mathbb{X}}_n : \underline{X}_1, \dots, \underline{X}_n$ iid $Par(\alpha,\beta)$

$\mathbb{x}_n \quad x_1 \quad x_n$

Initially in a), b) & c):

(α,β) - unknown

a) Factorization Theorem - Theo 6.2.6

Pdf of random sample:

$$f(\mathbf{x}_n|\alpha,\beta)=\prod_{i=1}^{n} f(x_i|\alpha,\beta)$$

$$=\alpha^{n\beta}\beta^n\prod_{i=1}^{n} x_i^{-(\beta+1)}\prod_{i=1}^{n} I(\alpha\le x_i)$$

$$=\underbrace{\alpha^{n\beta}\beta^n}_{h(\alpha,\beta)}\underbrace{\left[\prod_{i=1}^{n} x_i\right]^{-(\beta+1)}}_{g_1(T_1(\mathbf{x})|\beta)}\underbrace{I(\alpha\le\min\{\mathbf{x}_n\})}_{g_2(T_2(\mathbf{x})|\alpha)}$$

Sufficient statistics:

$$T_1(\mathbf{x})=\prod_{i=1}^{n} x_i \xrightarrow[\text{transf}]{\text{monotone}} \sum_{i=1}^{n}\log x_i$$

$$T_2(\mathbf{x})=\min\{\mathbf{x}_n\}=x_{(1)}$$ - order statistic

b) Likelihood:

$$L(\alpha,\beta|\mathbf{x}_n)=f(\mathbf{x}_n|\alpha,\beta)=\alpha^{n\beta}\beta^n\left[\prod_{i=1}^{n} x_i\right]^{-(\beta+1)} I(\alpha\le x_{(1)})$$

Log-likelihood:

$$\log\text{-}L(\alpha,\beta|\mathbf{x}_n)=$$

$$n\beta\log\alpha+n\log\beta-(\beta+1)\sum_{i=1}^{n}\log x_i+\log I(\alpha<x_{(1)})$$

Note:

Max: α - β fixed:

$$\log I(\alpha<x_{(1)})=\begin{cases}0 & \alpha<x_{(1)}\\ -\infty & x_{(1)}<\alpha\end{cases}$$

$n\beta\log\alpha$ - increase with α

Hence: Log-L max at $x_{(1)}$

MLE: $\hat{\alpha}=X_{(1)}=\min\{\mathbf{X}_n\}$

Max: β - α fixed:

$$\frac{\partial}{\partial\beta} \log\text{-}L(\alpha,\beta|\mathbb{x}_n)$$

$$= n\log\alpha + n\frac{1}{\beta} - \sum_i \log x_i = 0$$

$$\hat\beta = \left[\frac{1}{n}\sum_i \log x_i - \log\alpha\right]^{-1}$$

MLE: $\hat\beta = \left[\frac{1}{n}\sum_{i=1}^{n} \log X_i - \log X_{(1)}\right]^{-1}$

MLE for (μ, σ^2):

$$\left.\begin{aligned}\hat\mu &= [\hat\beta-1]^{-1}\hat\beta\hat\alpha\\ \hat\sigma^2 &= [\hat\beta-1]^{-2}[\hat\beta-2]^{-1}\hat\beta\hat\alpha^2\end{aligned}\right\}$$ Invariance property of MLE Theo 7.2.10

c) Hypothesis:

$H_0: \beta=\beta_0$ versus $H_1: \beta\neq\beta_0$ $\quad \beta_0 > 2$

Likelihood ratio:

$\alpha = \hat\alpha \quad \beta=\beta_0$

$$\lambda(\mathbb{x}_n) = \frac{\sup_{\alpha>0,\beta\in H_0} L(\alpha,\beta|\mathbb{x}_n)}{\sup_{\alpha>0,\beta\in H_1} L(\alpha,\beta|\mathbb{x}_n)}$$

$\alpha=\hat\alpha \quad \beta=\hat\beta$

$$= \frac{\hat\alpha^{n\beta_0}\beta_0^n\left[\prod_i x_i\right]^{-(\beta_0+1)} I(\hat\alpha\le x_{(1)})}{\hat\alpha^{n\hat\beta}\hat\beta^n\left[\prod_i x_i\right]^{-(\hat\beta+1)} I(\hat\alpha\le x_{(1)})}$$

$$= x_{(1)}^{n(\beta_0-\hat\beta)}\beta_0^n\hat\beta^{-n}\left[\prod_i x_i\right]^{\hat\beta-\beta_0}$$

Rejection region R_c:

$$R_c : \{ \mathbb{X}_n \mid \lambda(\mathbb{X}_n) < c \}$$

$$= \left\{ \mathbb{X}_n \mid x_{(1)}^{n(\beta_0 - \hat{\beta})} \beta_0^n \hat{\beta}^{-n} \left[\prod_i x_i \right]^{\hat{\beta} - \beta_0} < c \right\}$$

NOTE: $-2 \ln \lambda(\mathbb{X}_n)$ is NOT asymptotically Chi-square distributed, since the regularity conditions in $f(x \mid \alpha, \beta)$ is not fullfilled.

Assume: fixed $\alpha_0 > 0$ - $\beta > 2$ unknown

Sufficient statistic: $T_1(\mathbb{X}_n) = \sum_{i=1}^{n} \log X_i$

d) Note:

$$X \Rightarrow f(x|\beta) = Par(\beta) = \begin{cases} \alpha_0^\beta \beta x^{-(\beta+1)} & ; \alpha_0 < x \\ 0 & ; \text{else} \end{cases}$$

$$f(x|\beta) = \underbrace{\alpha_0^\beta \beta}_{c(\beta)} \cdot \exp\{\underbrace{-(\beta+1)}_{\omega(\beta)} \underbrace{\log x}_{t(x)}\}$$

Exponential family of pdf

From Theo 6.2.25 $T(\mathbb{X}_n) = \sum_{i=1}^{n} t(X_i) = \sum_{i=1}^{n} \log X_i$ is a complete statistic if $\omega(\beta) = -(\beta+1)$ is an open set - $\beta > 2 \Rightarrow \omega(\beta) < 3 \Rightarrow$ open.

Hence $T_1(\mathbb{X}_n)$ is a complete sufficient statistics

The MLE for β is

$$\beta^* = \left[\frac{1}{n} \sum_{i=1}^{n} \log X_i - \log \alpha_0\right]^{-1}$$

Can be seen from:

$$\beta^* = \underset{\beta}{\text{argmax}} \{\log\text{-}L(\alpha_0, \beta | x_n)\} \text{ in point b)}$$

e) The estimator β^* is consistent for β – since all MLE are consistent for exponential family pdfs – Theo 10.1.6

Consider:

$$\sqrt{n}\left[\beta^* - \beta\right] \xrightarrow[n\to\infty]{dist} N(0, v(\beta))$$

with $v(\beta)$ being Cramér-Rao Lower Bound

Theo 10.1.12

$$v(\beta) = \left[-\frac{d^2}{d\beta^2} \log\text{-}L(\beta|x)\right]^{-1}$$

$$= \left[\frac{1}{\beta^2}\right]^{-1} = \beta^2$$

The estimator β^* is asymptotically efficient for β – since β^* is MLE for β – Theo 10.1.12.

f) Parameter inference in Bayesian setting.

Prior model:

$$\beta \Rightarrow f(\beta|\lambda, k) = Gam(\lambda, k)$$

$$= \begin{cases} \lambda^k [k!]^{-1} \beta^{k-1} \exp\{-\lambda\beta\} & ; \beta \geq 0 \\ 0 & ; \beta < 0 \end{cases}$$

for known $\lambda \in \mathbb{R}_+$ and $k \in \mathbb{N}_+$.

$$f(\beta | \alpha_0, \lambda, k, x_n)$$

$$= const_1 \times f(x_n|\beta, \alpha_0) \times f(\beta|\lambda, k)$$

$$= const_1 \times \alpha_0^{n\beta} \cdot \beta^n \prod_i x_i^{-(\beta+1)} \cdot \beta^{k-1} \exp\{-\lambda\beta\}$$

$$= const_1 \times \beta^{k+n-1} \cdot \exp\{n\beta \ln\alpha_0\} \exp\{-(\beta+1)\sum_i \ln x_i\} \exp\{-\lambda\beta\}$$

$$= const_2 \times \beta^{k+n-1} \times \exp\{-(\lambda + \sum_i \ln x_i - n\ln\alpha_0)\beta\}$$

$$= Gam(\overbrace{\lambda + \sum_i \ln x_i - n\ln\alpha_0}^{\lambda|x_n}, \overbrace{k+n}^{k|x_n})$$

Hence the Gamma pdf is a conjugate prior model for iid samples from the Pareto pdf with given α_0.

$$E(\beta|\alpha_0, \lambda, k, x_n) = \frac{k|x_n}{\lambda|x_n} = \frac{k+n}{\lambda + \sum_i \ln x_i - n\ln\alpha_0}$$

from the Gamma pdf.

Assume: fixed $\beta_0 > 2$ − $\alpha > 0$ unknown

Sufficient statistic: $T_2(\mathbb{X}_n) = \min\{\mathbb{X}_n\} = X_{(1)}$

g) $X_1 \sim Par(\alpha, \beta_0)$ $\quad E\{X_1\} = [\beta_0 - 1]^{-1}\beta_0 \alpha$

$Var\{X_1\} = [\beta_0 - 1]^{-2}[\beta_0 - 2]^{-1}\beta_0 \alpha^2$

$$\alpha^+ = \beta_0^{-1}[\beta_0 - 1] X_1$$

$$E\{\alpha^+\} = \beta_0^{-1}[\beta_0 - 1] E\{X_1\} = \alpha$$

Use Rao-Blackwell construction Theo 7.3.17:

$$\tilde{\alpha} = E\{\beta_0^{-1}[\beta_0 - 1] X_1 \mid X_{(1)} = t\}$$

$$= \beta_0^{-1}[\beta_0 - 1] E\{X_1 / X_{(1)} = t\}$$

$$= \beta_0^{-1}[\beta_0 - 1]\Big[\overset{(A)}{E\{X_1 / X_{(1)} = t, (1) = 1\} p((1) = 1 / X_{(1)} = t)}$$

$$+ \overset{(B)}{E\{X_1 / X_{(1)} = t, (1) \neq 1\} p((1) \neq 1 / X_{(1)} = t)}\Big]$$

(A): $E\{X_1 / X_1 = t\}\frac{1}{n} = \frac{1}{n} t$

(B): $E\{X_1 / X_1 > t\}\frac{n-1}{n} = [1 - F(t)]^{-1}\int_t^\infty u f(u)\,du \cdot \frac{n-1}{n}$

$$= \frac{n-1}{n} \cdot \alpha^{-\beta_0} t^{\beta_0} \int_t^\infty u \beta_0 \alpha^{\beta_0} u^{-(\beta_0+1)}\,du$$

$$= \frac{n-1}{n} \alpha^{-\beta_0} t^{\beta_0} \beta_0 \alpha^{\beta_0} \int_t^\infty u^{-\beta_0}\,du \longrightarrow \Big[\frac{1}{-(\beta_0+1)} u^{-(\beta_0-1)}\Big]_t^\infty$$

$$\downarrow$$

$$\frac{1}{(\beta_0 - 1)} t^{-(\beta_0-1)}$$

$$= \frac{n-1}{n} t^{\beta_0} \beta_0 [\beta_0 - 1]^{-1} t^{-\beta_0 + 1}$$

$$= \frac{n-1}{n} \beta_0 [\beta_0 - 1]^{-1} t$$

$$\tilde{\alpha} = \beta_0^{-1}[\beta_0 - 1]\left[\frac{1}{n}t + \frac{n-1}{n}\beta_0[\beta_0 - 1]^{-1}t\right]$$

$$= \left[1 - [n\beta_0]^{-1}\right]t$$

Hence the estimator is:

$$\tilde{\alpha} = \left[1 - [n\beta_0]^{-1}\right]X_{(1)}$$

h) consider:

$$\alpha^* = X_{(1)}$$

Order stat cdf:

$$X_{(1)} \rightsquigarrow F_{X_{(1)}}(x/\alpha, \beta_0) = [1 - F(x)]^n \Longrightarrow$$

$$X_{(1)} \rightsquigarrow f_{X_{(1)}}(x/\alpha, \beta_0) = n f(x)[1 - F(x)]^{n-1}$$

$$= n\beta_0 \alpha^{\beta_0} x^{-(\beta_0+1)}[\alpha^{\beta_0} x^{-\beta_0}]^{n-1} I(\alpha \geq x)$$

$$= n\beta_0 \alpha^{n\beta_0} x^{-(n\beta_0+1)} I(\alpha \geq x)$$

$$= Par(\alpha, n\beta_0) \longrightarrow \begin{cases} E\{X_{(1)}\} = [n\beta_0 - 1]^{-1} n\beta_0 \alpha \\ Var\{X_{(1)}\} = [n\beta_0 - 1]^{-2}[n\beta_0 - 2]^{-1} n\beta_0 \alpha^2 \end{cases}$$

$$E\{X_{(1)}\} = \left[1 - [n\beta_0]^{-1}\right]^{-1}\alpha \xrightarrow[n\to\infty]{} \alpha \quad - \text{NOTE: } \tilde{\alpha} \;!!$$

$$Var\{X_{(1)}\} = [n\beta_0 - 1]^{-2}\left[1 - 2\cdot[n\beta_0]^{-1}\right]^{-1}\cdot\alpha \xrightarrow[n\to\infty]{} 0$$

α^* is consistent since:

$$E\{\alpha^*\} \xrightarrow[n\to\infty]{} \alpha, \quad Var\{\alpha^*\} \xrightarrow[n\to\infty]{} 0 \quad \text{Theo } 10.1.3$$