STAT 410

Practice Problems 19

Summer 2015
A. Stepanov

1. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the distribution with probability density function

$$f_X(x) = f_X(x;\theta) = (\theta-1)^2 \cdot \frac{\ln x}{x^\theta}, \qquad x > 1, \qquad \theta > 1.$$

Recall: If $\theta > 2$, the method of moments estimator of θ is $\tilde{\theta} = \dfrac{2\sqrt{\overline{X}} - 1}{\sqrt{\overline{X}} - 1}$.

The maximum likelihood estimator of θ is $\hat{\theta} = 1 + \dfrac{2n}{\sum_{i=1}^{n} \ln X_i}$.

$Y = \sum_{i=1}^{n} \ln X_i$ has Gamma($\alpha = 2n$, "usual θ" $= \dfrac{1}{\theta - 1}$) distribution.

a) Is $\hat{\theta}$ a consistent estimator for θ?

b) Show that $\hat{\theta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

c) Suppose $\theta > 3$. Is $\tilde{\theta}$ a consistent estimator for θ?

d) Suppose $\theta > 3$. Show that $\tilde{\theta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

e) Suggest a $(1-\alpha)\,100\,\%$ confidence interval for θ based on $\sum_{i=1}^{n} \ln X_i$.

f) Suppose $n = 5$, and

$$x_1 = 5, \qquad x_2 = 1.2, \qquad x_3 = 2, \qquad x_4 = 12, \qquad x_5 = 1.5.$$

Construct a 95% confidence interval for θ.

g) Find the sufficient statistic $Y = u(X_1, X_2, \ldots, X_n)$ for θ.

h) Determine the Fisher information $I(\theta)$.

i) Recall: $\hat{\hat{\theta}} = 1 + \dfrac{2n-1}{\sum_{i=1}^{n} \ln X_i}$ is an unbiased estimator for θ.

Is $\hat{\hat{\theta}}$ an efficient estimator of θ? If $\hat{\hat{\theta}}$ is not an efficient estimator of θ, find its efficiency.

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability density function

$$f(x) = 4\theta x^3 e^{-\theta x^4} \qquad x > 0 \qquad \theta > 0.$$

Recall: $Y = \sum_{i=1}^{n} X_i^4$ has Gamma($\alpha = n$, "usual θ" $= \dfrac{1}{\theta}$) distribution.

a) Suggest a confidence interval for θ with $(1-\alpha)\,100\,\%$ confidence level.

b) Find the sufficient statistic $Y = u(X_1, X_2, \ldots, X_n)$ for θ.

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability mass function

$$P(X_i = 1) = \frac{\theta}{3+\theta}, \quad P(X_i = 2) = \frac{2}{3+\theta}, \quad P(X_i = 3) = \frac{1}{3+\theta}, \qquad \theta > 0.$$

a) Find a sufficient statistic for θ.

b) Obtain the method of moments estimator $\tilde{\theta}$ of θ.

c) Obtain the maximum likelihood estimator $\hat{\theta}$ of θ.

4. Let $X_1, X_2, \ldots, X_n, \ldots$ be independent random variables, each with the probability density function

$$f(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

For each of the following sequences, state whether the sequence convergences in probability or/and in distribution, and identify the limits (limiting distributions).

a) $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$;

b) $T_n = \frac{1}{n}\left(X_1^2 + X_2^2 + \ldots + X_n^2\right)$;

c) $U_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right)$;

d) $V_n = n\left(\overline{X}_n - \frac{1}{3}\right)^2$;

e) $W_n = \sqrt{n}\left(\overline{X}_n^2 - \frac{1}{9}\right)$.

Answers:

1. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the distribution with probability density function

$$f_X(x) = f_X(x;\theta) = (\theta-1)^2 \cdot \frac{\ln x}{x^{\theta}}, \qquad x > 1, \qquad \theta > 1.$$

Recall: If $\theta > 2$, the method of moments estimator of θ is $\tilde{\theta} = \dfrac{2\sqrt{\overline{X}} - 1}{\sqrt{\overline{X}} - 1}$.

The maximum likelihood estimator of θ is $\hat{\theta} = 1 + \dfrac{2n}{\sum_{i=1}^{n} \ln X_i}$.

$Y = \sum_{i=1}^{n} \ln X_i$ has Gamma($\alpha = 2n$, "usual θ" $= \dfrac{1}{\theta - 1}$) distribution.

a) Is $\hat{\theta}$ a consistent estimator for θ?

By WLLN, $\dfrac{1}{n}\sum_{i=1}^{n} \ln X_i \xrightarrow{P} E(\ln X) = \dfrac{2}{(\theta-1)}$.

$g(x) = 1 + \dfrac{2}{x}$ is continuous at $\dfrac{2}{(\theta-1)}$.

$g\left(\dfrac{1}{n}\sum_{i=1}^{n} \ln X_i\right) = \hat{\theta}$. $\qquad g\left(\dfrac{2}{(\theta-1)}\right) = \theta$. $\qquad \Rightarrow \qquad \hat{\theta} \xrightarrow{P} \theta$.

b) Show that $\hat{\theta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

$$\operatorname{Var}(\ln X) = \frac{2}{(\theta-1)^2}.$$

By CLT, $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\ln X_i - \frac{2}{(\theta-1)}\right) \xrightarrow{D} N(0, \frac{2}{(\theta-1)^2})$.

$g(x) = 1+\frac{2}{x}$. $\qquad g'(x) = -\frac{2}{x^2}$.

$g(\frac{1}{n}\sum_{i=1}^{n}\ln X_i) = \hat{\theta}$. $\qquad g(\frac{2}{(\theta-1)}) = \theta$. $\qquad g'(\frac{2}{(\theta-1)}) = -\frac{(\theta-1)^2}{2}$.

$$\sqrt{n}\left(g\left(\frac{1}{n}\sum_{i=1}^{n}\ln X_i\right) - g\left(\frac{2}{(\theta-1)}\right)\right) = \sqrt{n}\left(\tilde{\theta}-\theta\right) \xrightarrow{D} N(0, \frac{(\theta-1)^2}{2}).$$

For large n, $\quad \tilde{\theta} \sim N(\theta, \frac{(\theta-1)^2}{2n})$.

c) Suppose $\theta > 3$. Is $\tilde{\theta}$ a consistent estimator for θ?

By WLLN, $\quad \overline{X} \xrightarrow{P} \mu = E(X) = \frac{(\theta-1)^2}{(\theta-2)^2}$.

$g(x) = \frac{2\sqrt{x}-1}{\sqrt{x}-1}$ is continuous at $\frac{(\theta-1)^2}{(\theta-2)^2}$.

$g(\overline{X}) = \tilde{\theta}$. $\qquad g(\frac{(\theta-1)^2}{(\theta-2)^2}) = \theta$. $\qquad \Rightarrow \qquad \tilde{\theta} \xrightarrow{P} \theta$.

d) Suppose $\theta > 3$. Show that $\tilde{\theta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

By CLT, $\quad \sqrt{n}\left(\overline{X}-\mu\right) \xrightarrow{D} N(0, \sigma^2)$.

$$\sigma^2 = \frac{(\theta-1)^2}{(\theta-3)^2} - \left[\frac{(\theta-1)^2}{(\theta-2)^2}\right]^2 = \frac{(\theta-1)^2\left(2\theta^2-8\theta+7\right)}{(\theta-2)^4(\theta-3)^2}.$$

$g(x) = \frac{2\sqrt{x} - 1}{\sqrt{x} - 1}$. $\qquad g'(x) = \frac{1}{2\sqrt{x}\left(\sqrt{x} - 1\right)^2}$.

$g(\overline{\mathrm{X}}) = \tilde{\theta}$ $\qquad g\left(\frac{(\theta - 1)^2}{(\theta - 2)^2}\right) = \theta$. $\qquad g'\left(\frac{(\theta - 1)^2}{(\theta - 2)^2}\right) = \frac{(\theta - 2)^3}{2(\theta - 1)}$.

$$\sqrt{n}\left(g\left(\overline{\mathrm{X}}\right) - g(\mu)\right) = \sqrt{n}\left(\tilde{\theta} - \theta\right) \xrightarrow{D} N\left(0, \frac{(\theta - 2)^2\left(2\theta^2 - 8\theta + 7\right)}{4(\theta - 3)^2}\right).$$

For large n, $\quad \tilde{\theta} \sim N\left(\theta, \frac{(\theta - 2)^2\left(2\theta^2 - 8\theta + 7\right)}{4n(\theta - 3)^2}\right)$.

e) Suggest a $(1 - \alpha)\,100\,\%$ confidence interval for θ based on $\sum_{i=1}^{n} \ln \mathrm{X}_i$.

$\sum_{i=1}^{n} \ln \mathrm{X}_i$ has Gamma($\alpha = 2n$, "usual θ" $= \frac{1}{\theta - 1}$) distribution.

Then $2(\theta - 1)\sum_{i=1}^{n} \ln \mathrm{X}_i$ has a $\chi^2(2\alpha = 4n)$ distribution.

$$\Rightarrow \qquad \mathrm{P}\left(\chi^2_{1-\alpha/2}(4n) < 2(\theta - 1)\sum_{i=1}^{n} \ln \mathrm{X}_i < \chi^2_{\alpha/2}(4n)\right) = 1 - \alpha.$$

$$\Rightarrow \qquad \mathrm{P}\left(1 + \frac{\chi^2_{1-\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln \mathrm{X}_i} < \theta < 1 + \frac{\chi^2_{\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln \mathrm{X}_i}\right) = 1 - \alpha.$$

A $(1 - \alpha)\,100\,\%$ confidence interval for θ

$$\left(1 + \frac{\chi^2_{1-\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln x_i},\ 1 + \frac{\chi^2_{\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln x_i}\right).$$

f) Suppose $n = 5$, and

$$x_1 = 5, \qquad x_2 = 1.2, \qquad x_3 = 2, \qquad x_4 = 12, \qquad x_5 = 1.5.$$

Construct a 95% confidence interval for θ.

$$\sum_{i=1}^{5} \ln x_i = \ln 216 \approx 5.3753. \qquad \chi^2_{0.975}(20) = 9.591. \qquad \chi^2_{0.025}(20) = 34.17.$$

A $(1-\alpha)\,100\,\%$ confidence interval for θ

$$\left(1 + \frac{\chi^2_{1-\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln x_i},\ 1 + \frac{\chi^2_{\alpha/2}(4n)}{2\sum_{i=1}^{n} \ln x_i} \right) = \left(1 + \frac{9.591}{2 \cdot 5.3753},\ 1 + \frac{34.17}{2 \cdot 5.3753} \right)$$

$$= (\,1.89, 4.18\,)$$

g) Find the sufficient statistic $\mathrm{Y} = u(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n)$ for θ.

$$f(x_1;\theta)\, f(x_2;\theta)\ \ldots\ f(x_n;\theta) = \prod_{i=1}^{n} (\theta-1)^2 \cdot \frac{\ln x_i}{x_i^{\theta}}$$

$$= (\theta-1)^{2n} \cdot \left(\prod_{i=1}^{n} x_i \right)^{-\theta} \cdot \prod_{i=1}^{n} \ln x_i\,.$$

$\Rightarrow$ $\mathrm{Y}_1 = \prod_{i=1}^{n} \mathrm{X}_i$ is a sufficient statistic for θ.

$\Rightarrow$ $\mathrm{Y}_2 = \ln \prod_{i=1}^{n} \mathrm{X}_i = \sum_{i=1}^{n} \ln \mathrm{X}_i$ is also a sufficient statistic for θ.

OR

$$f_{\mathrm{X}}(x;\theta) = (\theta-1)^2 \cdot \frac{\ln x}{x^{\theta}} = \exp\{-\theta \ln x + \ln \ln x + 2\ln(\theta-1)\}$$

$\Rightarrow$ $\mathrm{K}(x) = \ln x$.

$\Rightarrow \qquad Y_2 = \sum_{i=1}^{n} \ln X_i$ is a sufficient statistic for θ.

$\Rightarrow \qquad Y_1 = e^{Y_2} = \prod_{i=1}^{n} X_i$ is also a sufficient statistic for θ.

h) Determine the Fisher information $I(\theta)$.

$$\ln f(x;\theta) = 2 \cdot \ln(\theta - 1) + \ln \ln x - \theta \cdot \ln x$$

$$\frac{\partial}{\partial \theta} \ln f(x;\theta) = \frac{2}{\theta - 1} - \ln x$$

$$\frac{\partial^2}{\partial \theta^2} \ln f(x;\theta) = -\frac{2}{(\theta - 1)^2}$$

$$I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \ln f(X;\theta) \right] = \frac{2}{(\theta - 1)^2}$$

i) Recall: $\hat{\hat{\theta}} = 1 + \dfrac{2n-1}{\sum_{i=1}^{n} \ln X_i}$ is an unbiased estimator for θ.

Is $\hat{\hat{\theta}}$ an efficient estimator of θ? If $\hat{\hat{\theta}}$ is not an efficient estimator of θ, find its efficiency.

Recall: $\operatorname{Var}(\hat{\hat{\theta}}) = \dfrac{(2n-1)^2}{(2n)^2} \operatorname{Var}(\hat{\theta}) = \dfrac{(\theta-1)^2}{(2n-2)}$.

Rao-Cramer lower bound $= \dfrac{1}{n \cdot I(\theta)} = \dfrac{(\theta-1)^2}{2n}$.

$\Rightarrow \qquad \hat{\hat{\theta}}$ is NOT an efficient estimator of θ,

its efficiency $= \dfrac{2n-2}{2n} = \dfrac{n-1}{n} \to 1$ as $n \to \infty$.

2. Let $X_1, X_2, \dots, X_n$ be a random sample from the distribution with probability density function

$$f(x) = 4\theta\, x^3\, e^{-\theta x^4} \qquad x > 0 \qquad \theta > 0.$$

Recall: $\quad Y = \sum_{i=1}^{n} X_i^4$ has Gamma($\alpha = n$, "usual θ" $= \frac{1}{\theta}$) distribution.

a) Suggest a confidence interval for θ with $(1-\alpha)\,100\,\%$ confidence level.

$$2Y/\text{"usual }\theta\text{"} = 2\theta \sum_{i=1}^{n} X_i^4 \text{ has a chi-square distribution with } r = 2\alpha = 2n \text{ d.f.}$$

$$\Rightarrow \quad P\left(\chi^2_{1-\alpha/2}(2n) < 2\theta \sum_{i=1}^{n} X_i^4 < \chi^2_{\alpha/2}(2n) \right) = 1-\alpha.$$

$$\Rightarrow \quad P\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} X_i^4} < \theta < \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} X_i^4} \right) = 1-\alpha.$$

A $(1-\alpha)\,100\,\%$ confidence interval for θ: $\quad \left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} X_i^4},\ \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} X_i^4} \right).$

b) Find the sufficient statistic $\mathrm{Y} = u(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n)$ for θ.

$$f(x_1, x_2, \ldots x_n; \theta) = f(x_1; \theta)\, f(x_2; \theta)\, \ldots\, f(x_n; \theta)$$

$$= \left[4^n \theta^n e^{-\theta \sum_{i=1}^{n} x_i^4} \right] \left(\prod_{i=1}^{n} x_i^3 \right).$$

By Factorization Theorem, $\mathrm{Y} = \sum\limits_{i=1}^{n} \mathrm{X}_i^4$ is a sufficient statistic for θ.

OR

$$f(x; \theta) = \exp\left\{ -\theta \cdot x^4 + \ln\theta + \ln 4 + 3\ln x \right\}.$$

$\Rightarrow$ $\mathrm{K}(x) = x^4.$

$\Rightarrow$ $\mathrm{Y} = \sum\limits_{i=1}^{n} \mathrm{K}(\mathrm{X}_i) = \sum\limits_{i=1}^{n} \mathrm{X}_i^4$ is a sufficient statistic for θ.

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability mass function

$$P(X_i = 1) = \frac{\theta}{3+\theta}, \quad P(X_i = 2) = \frac{2}{3+\theta}, \quad P(X_i = 3) = \frac{1}{3+\theta}, \qquad \theta > 0.$$

a) Find a sufficient statistic for θ.

$$f(x_1;\theta)\, f(x_2;\theta)\, \ldots\, f(x_n;\theta) = \frac{1}{(3+\theta)^n} \cdot \theta^{(\#\text{ of 1's})} \cdot 2^{(\#\text{ of 2's})} \cdot 1^{(\#\text{ of 3's})}.$$

$\Rightarrow$ $Y = (\#\text{ of 1's})$ is a sufficient statistic for θ.

b) Obtain the method of moments estimator $\tilde{\theta}$ of θ.

$$E(X) = 1 \times \frac{\theta}{3+\theta} + 2 \times \frac{2}{3+\theta} + 3 \times \frac{1}{3+\theta} = \frac{\theta+7}{3+\theta}.$$

$$\frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \bar{x} = \frac{\tilde{\theta}+7}{3+\tilde{\theta}}. \qquad 3\bar{x} + \tilde{\theta}\,\bar{x} = \tilde{\theta} + 7.$$

$$\Rightarrow \qquad \tilde{\theta} = \frac{7 - 3\bar{x}}{\bar{x} - 1}.$$

c) Obtain the maximum likelihood estimator $\hat{\theta}$ of θ.

$$L(\theta) = \frac{1}{(3+\theta)^n} \cdot \theta^{(\#\text{ of 1's})} \cdot 2^{(\#\text{ of 2's})} \cdot 1^{(\#\text{ of 3's})}.$$

$$\ln L(\theta) = -n \ln(3+\theta) + (\#\text{ of 1's}) \ln(\theta) + (\#\text{ of 2's}) \ln(2) + (\#\text{ of 3's}) \ln(1).$$

$$(\ln L(\theta))' = -\frac{n}{3+\theta} + \frac{(\#\text{ of 1's})}{\theta} = 0 \qquad \Rightarrow \qquad \hat{\theta} = \frac{3 \cdot (\#\text{ of 1's})}{n - (\#\text{ of 1's})}.$$

4. Let $X_1, X_2, \dots, X_n, \dots$ be independent random variables, each with the probability density function

$$f(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

For each of the following sequences, state whether the sequence convergences in probability or/and in distribution, and identify the limits (limiting distributions).

a) $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$;

$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$;

By WLLN, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{P} \mu = E(X) = \frac{1}{3}$.

b) $T_n = \frac{1}{n}\left(X_1^2 + X_2^2 + \dots + X_n^2\right)$;

$T_n = \frac{1}{n}\left(X_1^2 + X_2^2 + \dots + X_n^2\right)$;

By WLLN, $T_n = \frac{1}{n}\left(X_1^2 + X_2^2 + \dots + X_n^2\right) \xrightarrow{P} E(X^2) = \frac{1}{6}$.

c) $U_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right)$;

$U_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right)$;

By CLT, $U_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right) \xrightarrow{D} N(0, \sigma^2) = N(0, \frac{1}{18})$.

d) $V_n = n\left(\overline{X}_n - \frac{1}{3}\right)^2;$

$V_n = n\left(\overline{X}_n - \frac{1}{3}\right)^2;$

From part (c), $\sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right) \xrightarrow{D} \frac{1}{\sqrt{18}} N(0, 1).$

Since $g(x) = x^2$ is continuous,

$$g\left(\sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right)\right) = n\left(\overline{X}_n - \frac{1}{3}\right)^2 \xrightarrow{D} \left[\frac{1}{\sqrt{18}} N(0, 1)\right]^2 = \frac{1}{18}\chi^2(1).$$

e) $W_n = \sqrt{n}\left(\overline{X}_n^2 - \frac{1}{9}\right).$

$W_n = \sqrt{n}\left(\overline{X}_n^2 - \frac{1}{9}\right).$

From part (c), $\sqrt{n}\left(\overline{X}_n - \frac{1}{3}\right) \xrightarrow{D} N(0, \frac{1}{18}).$

Since $g(x) = x^2$ is differentiable and $g'(\frac{1}{3}) = \frac{2}{3} \neq 0,$

$$\sqrt{n}\left(g(\overline{X}_n) - g\left(\frac{1}{3}\right)\right) = \sqrt{n}\left(\overline{X}_n^2 - \frac{1}{9}\right) \xrightarrow{D} N(0, \left(\frac{2}{3}\right)^2 \times \frac{1}{18}) = N(0, \frac{2}{81}).$$