Ex. 1 page 146 (Ch. 9)

$X_1, \dots, X_N \overset{i.i.d}{\sim} \Gamma(\alpha,\beta)$ Gamma distr.

$\alpha, \beta > 0$

$$f(x_i \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-x_i \beta} \mathbb{1}_{]0,+\infty[}(x_i)$$

Special case:

$$\alpha = 1 \Rightarrow f(x_i \mid \beta) = \frac{\beta}{\Gamma(1)} e^{-x_i \beta} \mathbb{1}_{]0,+\infty[}(x_i)$$

$\Rightarrow$ exponential R.V.

Recall: i) $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

Gamma function

$\Gamma(1) = 1$

Since we estimate (α, β) we need two empirical moments:

i) $\mathbb{E}(X_i) = \int_0^{+\infty} u \frac{\beta^{\alpha}}{\Gamma(\alpha)} u^{\alpha-1} e^{-u\beta} \, du =$

$$= \underbrace{\int_0^{+\infty} \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} u^{\alpha} e^{-u\beta}\, du}_{=1} \cdot \frac{\alpha}{\beta} = \frac{\alpha}{\beta}$$

ii) $$\mathbb{E}(X_i^2) = \int_0^{+\infty} u^2 \frac{\beta^{\alpha}}{\Gamma(\alpha)} u^{\alpha-1} e^{-u\beta}\, du$$

$$= \int_0^{+\infty} \frac{\alpha^2+\alpha}{\beta^2} \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} u^{\alpha+1} e^{-u\beta}\, du = \frac{\alpha^2+\alpha}{\beta^2}$$

Since

$$\Gamma(\alpha+2) = (\alpha+1)\Gamma(\alpha+1) = (\alpha+1)\alpha\Gamma(\alpha)$$

$$\Rightarrow \Gamma(\alpha) = \frac{\Gamma(\alpha+2)}{\alpha^2+\alpha}$$

Empirical moments:

i) $\overline{X}$

ii) $\frac{1}{n}\sum_{i=1}^{n} X_i^2$

$$\begin{cases} \frac{\alpha}{\beta} = \overline{X} \\ \frac{\alpha^2+\alpha}{\beta^2} = \frac{1}{n}\sum_i X_i^2 \end{cases}$$

The solution is the MM estimator $(\hat{\alpha}, \hat{\beta})$

$$\begin{cases} \alpha = \bar{x}\beta \\[2ex] \dfrac{\bar{x}^2\beta^2 + \bar{x}\beta}{\beta^2} = \dfrac{1}{n}\sum_i x_i^2 \end{cases}$$

from Eq. 2 : $\beta^{\cancel{2}}\left[\frac{1}{n}\sum_i x_i^2 - \bar{x}^2\right] = \bar{x}\cancel{\beta}$

$$\Rightarrow \quad \hat{\beta} = \frac{\bar{x}}{\frac{1}{n}\sum_i x_i^2 - \bar{x}^2}$$

$$\hat{\alpha} = \frac{\bar{x}^2}{\frac{1}{n}\sum_i x_i^2 - \bar{x}^2}$$

Ex. 2

$X_1, \dots, X_N \sim \mathcal{U}[a,b]$ $\quad a < b$

i) Find MM estimates of (a,b)

Theoretical moments

i) $\mathbb{E}(X_i) = \int_a^b u \frac{1}{b-a} du = \frac{1}{b-a} \frac{1}{2} u^2 \Big|_a^b$

$= \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}}$

$= \frac{a+b}{2}$

ii) $\mathbb{E}(X_i^2) = \frac{a^2+b^2+ab}{3}$

Empirical moments : i) $\bar{X}$

ii) $S = \frac{1}{N} \sum_i X_i^2$

$$\begin{cases} \frac{a+b}{2} = \bar{X} \\ \frac{a^2+b^2+ab}{3} = S \end{cases}$$

$$\begin{cases} a = 2\bar{x} - b \\ 4\bar{x}^2 + b^2 - 4\bar{x}b + \cancel{b^2} + 2\bar{x}b - \cancel{b^2} = 3S \end{cases}$$

Eq. 2: $4\bar{x}^2 - 2\bar{x}b + b^2 - 3S = 0$

$$b^2 - 2\bar{x}b + 4\bar{x}^2 - 3S = 0$$

$$b = \bar{x} \pm \sqrt{\bar{x}^2 - 4\bar{x}^2 + 3S} = \bar{x} \pm \sqrt{3(S - \bar{x}^2)}$$

Eq. 1: $a = \bar{x} \mp \sqrt{3(S - \bar{x}^2)}$

Since $a < b \Rightarrow$

$$\hat{a} = \bar{x} - \sqrt{3(S - \bar{x}^2)}$$

$$\hat{b} = \bar{x} + \sqrt{3(S - \bar{x}^2)}$$

Recall: $ax^2 + bx + c = 0$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} =$$

$$= \frac{-\frac{b}{2} \pm \sqrt{\left(\frac{b}{2}\right)^2 - ac}}{a}$$

2)
$$\hat{a}_{ML} = \min\{x_1, \dots, x_N\}$$
$$\hat{b}_{ML} = \max\{x_1, \dots, x_N\}$$

c) $\tau = \int x f(x) dx = E(X_i)$

The plug-in estimator of τ, say $\hat{\tau}_{PEN} = \overline{X}$

Recall that $E(X_i) = \frac{a+b}{2}$

Equivariance of ML $\Rightarrow$ $\hat{\tau}_{ML} = \frac{\hat{a}_{ML} + \hat{b}_{ML}}{2}$

d) Since $\hat{\tau}_{PEN} = \overline{X} \Rightarrow \mathbb{E}(\overline{X}) = \frac{1}{n}\sum_i \underbrace{E(X_i)}_{\frac{a+b}{2}}$

$$= \frac{a+b}{2} = \tau$$

Recall:
$$MSE(\hat{\tau}) = Var(\hat{\tau}) + bias^2(\hat{\tau})$$

and $bias(\hat{\tau}_{PEN}) = 0$

$$Var(\hat{\tau}_{PEN}) \stackrel{\perp}{=} \frac{1}{N^2}\sum_i Var(X_i) = \frac{Var(X_i)}{N}$$

$$\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - [\mathbb{E}(X_i)]^2 =$$

$$= \frac{a^2+b^2+ab}{3} - \frac{a^2+b^2+2ab}{4}$$

$$= \frac{4a^2+4b^2+4ab-3a^2-3b^2-6ab}{12}$$

$$= \frac{a^2+b^2-2ab}{12} = \frac{(a-b)^2}{12}$$

$$\Rightarrow \mathrm{Var}(\hat{\tau}_{\mathrm{DiN}}) = \frac{(a-b)^2}{12N} = \mathrm{MSE}(\hat{\tau}_{\mathrm{DiN}})$$

ASSUME $a = 1, b = 3 \quad n = 10$

Properties of ML estimators

Consistency:

Kullback-Leibler $\rightarrow$ $KL(f\|g) = \mathbb{E}_f \log \frac{f(x)}{g(x)}$

$$= \int f(x) \log \frac{f(x)}{g(x)}\, dx$$

Facts:

$KL(f\|g) \geq 0$

and $KL(f\|g) = 0$ iff. $g \overset{ae}{=} f$

Given $x_1, \ldots, x_N \overset{i.i.d}{\sim} f(\theta^*)$ $\quad \theta^* \in \mathbb{R}$ ← true value

We maximize $\ell(\theta) = \underbrace{\frac{1}{N} \sum_i \log f(x_i | \theta)}$

i) $M_N(\theta) \triangleq \frac{1}{N} \sum_{i=1}^{N} \log \frac{f(x_i|\theta)}{f(x_i|\theta^*)} \xrightarrow{\mathbb{P}_{\theta^*}} KL(f_{\theta^*} \| f_\theta)$

$$\sup_{\theta \in \mathbb{R}} |M_N(\theta) - KL(f_{\theta^*} \| f_\theta)| \xrightarrow{\mathbb{P}_{\theta^*}} 0$$

ii) There exists a unique max for $KL(f_{\theta^*} \| f_\theta)$

$$\hat{\theta}_{ML} \longrightarrow \theta^*$$

ii) Equivariance: $\tau = g(\theta)$, with g smooth then $\hat{\tau}_{ML} = g(\hat{\theta}_{ML})$

iii) Asymptotic normality:

$\theta \in \mathbb{R}$ $\qquad \lim_{n \to +\infty} se(\hat{\theta}_N) = \left[\frac{1}{I_n(\theta)}\right]^{1/2}$

where $I_n(\theta)$ is the Fisher information

$I_n(\theta) \overset{i.i.d.}{=} n I(\theta)$ with

$$I(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log f(x_i|\theta)\right]$$

It means that $\hat{se}(\hat{\theta}_N) = I_N(\hat{\theta}_{ML})$

$$\hat{\theta}_{ML} - \theta \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{I_n(\theta)}\right)$$

$$\Rightarrow \left[\hat{\theta}_{ML} - z_{\alpha/2}\sqrt{\frac{1}{I_n(\hat{\theta}_{ML})}}, \hat{\theta}_{ML} + z_{\frac{\alpha}{2}}\sqrt{\frac{1}{I_n(\hat{\theta}_{ML})}}\right]$$

is an asymptotic CI of level $1-\alpha$.

Ex: $X_1, \dots, X_N \sim \mathcal{N}(\mu, \sigma^2)$ with μ known

AIM: provide <u>ML estimate for σ</u> and an approx. 95% CI for σ

$$\ell(\sigma) = \sum_{i=1}^{N} \log \underbrace{f(x_i \mid \mu, \sigma^2)}_{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}} =$$

$$= \sum_i \left[-\log\sigma - \frac{1}{2\sigma^2}(x_i-\mu)^2 - cst \right]$$

$$= -N\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2 + cst$$

$$\frac{d}{d\sigma}\ell(\sigma) = -\frac{N}{\sigma} + \frac{\cancel{2}\sigma^{-3}}{\cancel{2}}\sum_{i=1}^{N}(x_i-\mu)^2 = 0$$

$$\iff \frac{1}{\sigma^{\cancel{3}\,2}}\sum_{i=1}^{N}(x_i-\mu)^2 = \frac{N}{\cancel{\sigma}}$$

$$\iff \hat{\sigma}_{ML} = \sqrt{\frac{1}{N}\sum_i (x_i-\mu)^2}$$

$$\hat{se}(\hat{\sigma}_{ML}) = \sqrt{\frac{1}{I_n(\hat{\sigma}_{ML})}}$$

with $I_n(\theta) \stackrel{!}{=} n I(\theta)$ and

$$I_\theta = -\mathbb{E}\left[\frac{d^2}{d\sigma^2} \underbrace{\log f(x_i|\mu,\sigma)}\right]$$

$$\log \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} =$$

$$= -\log\sigma - \frac{1}{2\sigma^2}(x_i-\mu)^2 + cst$$

$$\frac{d}{d\sigma}\log f(x_i|\mu,\sigma) = -\frac{1}{\sigma} + \frac{1}{\sigma^3}(x_i-\mu)^2$$

$$\frac{d^2}{d\sigma^2}\log f(x_i|\mu,\sigma) = \frac{1}{\sigma^2} - 3\sigma^{-4}(x_i-\mu)^2$$

$$-\mathbb{E}\,\frac{d}{d\sigma^2}\log f(x_i|\mu,\sigma) = \mathbb{E}\left[\frac{1}{\sigma^2} - \frac{3}{\sigma^4}(x_i-\mu)^2\right]$$

$$= \frac{1}{\sigma^2} - \frac{3}{\sigma^4}\underbrace{\mathbb{E}(x_i-\mu)^2}_{\sigma^2} = + \frac{2}{\sigma^2}$$

$$\Rightarrow I_n(\sigma) = \frac{2n}{\sigma^2} \quad \text{and finally}$$

$$\hat{se}(\hat{\sigma}_{ML}) = \sqrt{\frac{\hat{\sigma}^2_{ML}}{2n}} = \frac{\hat{\sigma}_{ML}}{\sqrt{2n}}$$

Finally,

$$CI_{1-\alpha} = \left[\hat{\sigma}_{ML} \pm z_{\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_{ML}}{\sqrt{2n}} \right]$$