Chapter 9 - Parametric Inference

Q_1: $X_1, \dots, X_n \sim \text{Gamma}(\alpha, \beta)$

method of moments estimator for α, β ?

$\alpha_1 = E(X) = \alpha\beta \quad , \quad \hat{\alpha}_1 = n^{-1} \sum X_i = \bar{X}$

$\alpha_2 = E(X^2) = \alpha\beta^2 + \alpha^2\beta^2 \quad , \quad \hat{\alpha}_2 = n^{-1} \sum X_i^2 = \overline{X^2}$

$$\Rightarrow \begin{cases} \alpha\beta = \bar{X} \\ \alpha\beta(\alpha\beta + \beta) = \overline{X^2} \end{cases} \Rightarrow \begin{cases} \alpha = \bar{X}^2 / (\overline{X^2} - \bar{x}^2) \\ \beta = (\overline{x^2} - \bar{x}^2) / \bar{x} \end{cases}$$

Q_2: $X_1, \dots, X_n \sim \text{Unif}(a, b) \quad a < b$ unknown

a) method of moment estimator for a & b ?

$\alpha_1 = E(X) = \frac{a+b}{2} = \hat{\alpha}_1 = \bar{x}$

$\alpha_2 = E(X^2) = \frac{(b-a)^2}{12} + \frac{(a+b)^2}{4} = \hat{\alpha}_2 = \overline{x^2}$

define $\hat{\sigma}^2 = \overline{x^2} - \bar{x}^2 \quad , \quad c = b - a$

$$\Rightarrow \begin{cases} (a+b)/2 = \bar{x} \\ c^2/12 = \hat{\sigma}^2 \end{cases} \Rightarrow \begin{cases} a+b = 2\bar{x} \\ b-a = 2\sqrt{3}\,\hat{\sigma} \end{cases}$$

$\Rightarrow b = \bar{x} + \sqrt{3}\,\hat{\sigma}$, $a = \bar{x} - \sqrt{3}\,\hat{\sigma}$

b) find MLE $\hat{a}, \hat{b}$

$L_n(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = (b-a)^{-n} \prod_{i=1}^{n} I_{(a,b)}(x_i)$

maximize $L_n(\theta)$ by $\hat{a} = \min(x_i)$, $\hat{b} = \max(x_i)$

note: if $\hat{a} \neq \min(x_i)$ or $\hat{b} \neq \max(x_i) \Rightarrow L_n(\theta) = 0$

c) Let $\tau = \int x \, dF(x)$ find MLE for τ

$\tau = (a+b)/2 = g(\theta)$

by equivariance $\hat{\tau} = g(\hat{\theta}) = (\hat{a} + \hat{b})/2$

d) $\tau = \int x \, dF(x)$, $\hat{\tau}$: MLE $\tilde{\tau}$: non param plug in estimator

find MSE of $\hat{\tau}$ by simulation , find MSE of $\tilde{\tau}$ analytically

$a=1$, $b=3$, $n=10$

MSE of $\hat{\tau}$ by simulation:

$\hat{\tau} = (\hat{a} + \hat{b})/2$ $\Rightarrow$ results in Jupyter notebook

$\hat{a} = \min(x_i)$, $\hat{b} = \max(x_i)$

MSE of $\tilde{\tau}$ analytically:

$\tilde{\tau} = E_{\hat{F}} X = n^{-1} \sum X_i$ unbiased estimator

$MSE = \cancel{bias^2(\tilde{\tau})} + Var(\tilde{\tau}) = Var(\tilde{\tau}) = \frac{(b-a)^2}{12\,n} = \frac{1}{30} = 0.033$

$Q_3:$ $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ $P(X < \tau) = 0.95$

a) MLE of τ ?

$P(X < \tau) = P\left(\frac{X-\mu}{\sigma} < \frac{\tau-\mu}{\sigma}\right) = F_Z\left(\frac{\tau-\mu}{\sigma}\right) = 0.95$

Z: Standard normal

$\Rightarrow \tau = F_Z^{-1}(0.95)\,\sigma + \mu = g(\mu, \sigma)$

$\Rightarrow$ by equivariance : $\hat{\tau} = g(\hat{\mu}, \hat{\sigma})$

where $\hat{\mu} = \bar{x}$, $\hat{\sigma}^2 = n^{-1} \sum (X_i - \bar{X})^2 = \overline{X^2} - \bar{x}^2$

b) find $1-\alpha$ confidence interval for τ

$\tau = g(\mu, \sigma) = F_Z^{-1}(0.95)\,\sigma + \mu$

$\hat{se}(\hat{\tau}) = \sqrt{(\hat{\nabla} g)^T \hat{J}_n \, \hat{\nabla} g}$

$J_n = I_n^{-1}(\theta)$, $\hat{J}_n = J_n(\hat{\theta})$

$\nabla g^T = \left[\frac{\partial g}{\partial \mu}, \frac{\partial g}{\partial \sigma}\right]$, $\hat{\nabla} g = \nabla g(\hat{\theta})$

$$L_n(\mu, \sigma) = \sigma^{-n} \exp\left\{ \frac{-1}{2\sigma^2} \sum_i (X_i - \mu)^2 \right\}$$

$$\Rightarrow l_n = -n \log \sigma - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2$$

$$\partial l_n / \partial \mu = \sigma^{-2} \sum (X_i - \mu)$$

$$\Rightarrow H_{11} = \partial^2 l_n / \partial \mu^2 = -n \sigma^{-2}$$

$$H_{12} = H_{21} = \partial^2 l_n / \partial \sigma \partial \mu = -2\sigma^{-3} \sum (X_i - \mu)$$

$$\partial l_n / \partial \sigma = -n\sigma^{-1} + \sigma^{-3} \sum (X_i - \mu)^2$$

$$\Rightarrow H_{22} = \partial^2 l_n / \partial \sigma^2 = n\sigma^{-2} - 3\sigma^{-4} \sum (X_i - \mu)^2$$

$$\Rightarrow I_n(\mu, \sigma) = - \begin{bmatrix} E_\theta(H_{11}) & E_\theta(H_{12}) \\ E_\theta(H_{21}) & E_\theta(H_{22}) \end{bmatrix} = \begin{bmatrix} n\sigma^{-2} & 0 \\ 0 & 2n\sigma^{-2} \end{bmatrix}$$

$$\Rightarrow J_n = I_n^{-1} = \frac{1}{n} \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2/2 \end{bmatrix}$$

$$\nabla g^T = \left[\frac{\partial g}{\partial \mu} \; \frac{\partial g}{\partial \sigma} \right] = [1, c] \qquad c = F_Z^{-1}(0.95)$$

$$\Rightarrow \hat{se}(\hat{\tau}) = \left([1, c] \frac{1}{n} \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2/2 \end{bmatrix} \begin{bmatrix} 1 \\ c \end{bmatrix} \right)^{0.5}$$

$$= \sigma \sqrt{\frac{1}{n} + \frac{c^2}{2n}}$$

$\Rightarrow$ $1-\alpha$ Conf. Interval : $\hat{\tau} \pm z_{\alpha/2} \hat{se}(\hat{\tau})$

C) Simulation

Jupyter notebook.

Q4: $X_1, \dots, X_n \sim Unif(0,\theta)$ show MLE is consistent.

hint $y = \max\{X_1, \dots, X_n\}$ $\forall c$ $P(y<c) = P(X_1<c) \cdots P(X_n<c)$

using results from question 2b: $\hat{\theta} = \max(X_i) = y$

$$P(|y-\theta|>\varepsilon) = 1-P(\theta-\varepsilon \leq y \leq \theta+\varepsilon) = 1-F_y(\theta+\varepsilon)+F_y(\theta-\varepsilon)$$

$$F_y(y) = P(X_1<y) \cdots P(X_n<y) = F_X^n(y)$$

$$F_X(y) = P(X<y) = \begin{cases} \frac{y}{\theta} & 0 \leq y \leq \theta \\ 1 & y \geq \theta \end{cases}$$

$$\Rightarrow F_y(y) = \begin{cases} (y/\theta)^n & 0 \leq y \leq \theta \\ 1 & y \geq \theta \end{cases} \Rightarrow \begin{cases} F_y(\theta-\varepsilon) = \left(\frac{\theta-\varepsilon}{\theta}\right)^n \\ F_y(\theta+\varepsilon) = 1 \end{cases}$$

$$\Rightarrow P(|y-\theta|>\varepsilon) = 1-F_y(\theta+\varepsilon)-F_y(\theta-\varepsilon) = \left(1-\frac{\varepsilon}{\theta}\right)^n \to 0$$

$$\lim_{n\to\infty} \left(1-\frac{\varepsilon}{\theta}\right)^n = 0 \Rightarrow y \xrightarrow{P} \theta$$ consistent.

Q_5: $X_1, \dots, X_n \sim \text{Poisson}(\lambda)$

find moment estimator, MLE, Fisher information $I(\lambda)$

- MME $\quad \alpha_1 = E_\lambda(X) = \lambda \quad \hat{\alpha}_1 = \bar{X} \quad \Rightarrow \quad \hat{\lambda} = \bar{X}$

- MLE $\quad L_n(\lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-\lambda n} \lambda^{\sum x_i} / \prod x_i!$

$\ell_n(\lambda) = -n\lambda + \log\lambda \sum x_i - \sum \log x_i!$

$\partial \ell_n / \partial \lambda = -n + \lambda^{-1} \sum x_i = 0 \Rightarrow \hat{\lambda} = n^{-1} \sum x_i = \bar{x}$

- Fisher information $\quad I(\lambda) = -E_\lambda\left(\partial^2 \log f(x,\lambda) / \partial \lambda^2\right)$

$\log f = -\lambda + x \log \lambda - \log x! \Rightarrow \frac{\partial \log f}{\partial \lambda} = -1 + \frac{x}{\lambda}$

$\Rightarrow \frac{\partial^2 \log f}{\partial \lambda^2} = -x\lambda^{-2} \quad \Rightarrow I(\lambda) = -E_\lambda(-x\lambda^{-2}) = \frac{1}{\lambda}$

Q_6: $X_1, \dots, X_n \sim N(\theta, 1) \quad Y_i = \begin{cases} 1 & x_i > 0 \\ 0 & x_i \leq 0 \end{cases} \quad \psi = P(Y_1 = 1)$

a) find MLE $\hat{\psi}$

$\psi = P(X > 0) = 1 - P(X < 0) = 1 - P(X - \theta < -\theta) = 1 - F_Z(-\theta) = F_Z(\theta)$

by equivariance $\hat{\psi} = F_Z(\hat{\theta}) \quad , \hat{\theta} = \bar{X}$

b) find 95% C.I. for ψ

$\psi = g(\theta) = F_Z(\theta) \Rightarrow g' = f_Z(\theta)$ f_Z is normal dist. function

by delta method: $\hat{se}(\hat{\psi}) = |g'(\hat{\theta})|\, \hat{se}(\hat{\theta}) = f_Z(\hat{\theta}) \times \frac{1}{\sqrt{n}}$

95% C.I. $= \hat{\psi} \pm 2\,\hat{se}(\hat{\psi})$

C) define $\tilde{\psi} = n^{-1}\sum Y_i$ show $\tilde{\psi}$ is consistent est. for ψ

Using LLN: $\tilde{\psi}$ converges in prob. to $E(Y_i) = \psi \Rightarrow$ consistent.

d) compute Assymptotic Relative Efficiency of $\tilde{\psi}$ to $\hat{\psi}$

$se^2(\hat{\psi}) = f_Z^2(\bar{X})/n$ derived in part b

$se^2(\tilde{\psi}) = Var(y)/n$, $Var(y_i) = E(y^2) - E^2(y) = \psi(1-\psi)$

$\Rightarrow ARE(\tilde{\psi}, \hat{\psi}) = \frac{f_Z^2(\bar{X})}{\psi(1-\psi)} < 1$ optimality of MLE

e) Suppose data are not normal. Show $\hat{\psi}$ is not consistent.

does $\hat{\psi}$ converge to anything?

$\hat{\theta} = \bar{X} \xrightarrow{P} \mu_x \Rightarrow \hat{\psi} \xrightarrow{P} F_Z(\mu_x) \neq \psi = P(X>0) = 1 - F_x(0)$

if X is not normal $\Rightarrow 1 - F_X(0) \neq F_Z(\mu_x)$

Q_7: $X_1 \sim Bin(n_1, p_1)$ $X_2 \sim Bin(n_2, p_2)$ $\psi = p_1 - p_2$

a) MLE $\hat{\psi}$

$$f(x,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\Rightarrow \log f(x,p) = \log \binom{n}{x} + x \log p + (n-x)\log(1-p)$$

taking derivative: $\frac{x}{p} + \frac{x-n}{1-p} = 0 \Rightarrow \hat{p} = \frac{x}{n}$

by equivariance $\hat{\psi} = \hat{p}_1 - \hat{p}_2 = \frac{x_1}{n_1} - \frac{x_2}{n_2}$

b) Fisher info matrix $I(p_1, p_2)$

$$\ell_n = \sum_{i=1}^{2} \log f(x_i) = \sum \log \binom{n_i}{x_i} + x_i \log p_i + (n_i - x_i)\log(1-p_i)$$

$$\partial \ell_n / \partial p_1 = \frac{x_1}{p_1} + \frac{x_1 - n_1}{1-p_1} \qquad H_{11} = \frac{\partial^2 \ell_n}{\partial p_1^2} = \frac{-x_1}{p_1^2} + \frac{x_1 - n_1}{(1-p_1)^2}$$

$$H_{12} = H_{21} = \partial^2 \ell_n / \partial p_1 \partial p_2 = 0$$

$$E_\theta(H_{11}) = -\frac{n_1 p_1}{p_1^2} - \frac{n_1(1-p_1)}{(1-p_1)^2} = \frac{-n_1}{p_1(1-p_1)} \Rightarrow I(\theta) = \begin{bmatrix} \frac{n_1}{p_1(1-p_1)} & 0 \\ 0 & \frac{n_2}{p_2(1-p_2)} \end{bmatrix}$$

c) use delta method to find $\hat{se}(\hat{\psi})$

$$J(\theta) = I^{-1}(\theta) = diag\left(p_1(1-p_1)/n_1,\ p_2(1-p_2)/n_2\right)$$

$$\psi = g(p_1, p_2) = p_1 - p_2 \Rightarrow \nabla g = [1, -1]^T$$

$$\hat{se}(\hat{\psi}) = \left(\hat{\nabla g}^T \hat{J}\, \hat{\nabla g}\right)^{1/2} = \left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}\right)^{1/2}$$

d) $n_1 = n_2 = 200$ $\quad X_1 = 160$ $\quad X_2 = 148$ $\quad \hat{\psi} = ?$

90% C.I. using delta method & Parametric bootstrap

Param bootstrap in Jupyter notebook

$\hat{p}_1 = 0.8 \quad \hat{p}_2 = 0.74 \quad \hat{\psi} = \hat{p}_1 - \hat{p}_2 = 0.06$

$$\hat{se}(\hat{\psi}) = \left(\frac{0.8 \times 0.2 + 0.74 \times 0.26}{200}\right)^{1/2} =$$

$\alpha = 0.1 \Rightarrow Z_{\alpha/2} = 1.645 \Rightarrow 90\%$ C.I. $= (-0.009, 0.129)$

Q8: Fisher info matrix for Eample 9.29

$X_1, \ldots, X_n \sim N(\mu, \sigma^2) \quad \tau = g(\mu, \sigma) = \sigma/\mu$

$$I_n(\mu, \sigma) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}$$

$$\ell_n = -n \log \sigma - \frac{1}{2\sigma^2} \sum (X_i - \mu)^2$$

$$\partial \ell_n / \partial \mu = \sigma^{-2} \sum (X_i - \mu)$$

$$\Rightarrow H_{11} = \partial^2 \ell_n / \partial \mu^2 = -n\sigma^{-2}$$

$$H_{12} = H_{21} = \partial^2 \ell_n / \partial \sigma \partial \mu = -2\sigma^{-3} \sum (X_i - \mu)$$

$\partial \ell_n / \partial \sigma = -n\sigma^{-1} + \sigma^{-3} \sum (X_i - \mu)^2$

$\Rightarrow H_{22} = \partial^2 \ell_n / \partial \sigma^2 = n\sigma^{-2} - 3\sigma^{-4} \sum (X_i - \mu)^2$

$$I_n(\mu, \sigma) = - \begin{bmatrix} E_\theta(H_{11}) & E_\theta(H_{12}) \\ E_\theta(H_{21}) & E_\theta(H_{22}) \end{bmatrix} = \begin{bmatrix} n\sigma^{-2} & 0 \\ 0 & 2n\sigma^{-2} \end{bmatrix}$$

Q9: $X_1, \cdots, X_n \sim N(\mu, 1)$, $\theta = e^{\mu}$, $MLE = \hat{\theta} = e^{\bar{X}}$

$\mu = 5 \quad n = 100$

a) $\hat{se}$ and 95% C.I. using delta method, param./nonparam. boot.

$\theta = g(\mu) = e^{\mu} \Rightarrow \hat{se}(\hat{\theta}) = |g'(\hat{\mu})| \hat{se}(\hat{\mu})$

$\hat{\mu} = \bar{x} \Rightarrow \hat{se}(\hat{\mu}) = 1/\sqrt{n} \Rightarrow \hat{se}(\hat{\theta}) = e^{\bar{X}}/\sqrt{n}$

95% C.I. $= e^{\bar{X}} \pm 1.96\, e^{\bar{X}}/\sqrt{100}$

param./nonparam bootstrap in Jupyter notebook.

b) Simulation

Jupyter notebook.

Q10: $X_1, \dots, X_n \sim \text{Unif}(0,\theta)$ $\quad MLE = \hat{\theta} = \max\{X_1, \dots, X_n\}$

$n = 50$, $\theta = 1$

a) find dist. of $\hat{\theta}$ analytically, compare it to param. and nonparam bootstrap histograms.

$\hat{\theta} = \max\{X_i\} = Y \implies$ CDF $F_Y(y) = \left(\frac{y}{\theta}\right)^n$ $\quad f_Y(y) = \frac{n}{\theta^n} y^{n-1}$

comp. plots in Jupyter notebook.

b) param. bootstrap $P(\hat{\theta}^* = \hat{\theta}) = 0$

nonparam bootstrap $P(\hat{\theta}^* = \hat{\theta}) = 0.632$

param bootstrap: cont. dist. $\implies P(\hat{\theta}^* = \hat{\theta}) = 0$

non param bootstrap: refer to Chapter 8, question 7.