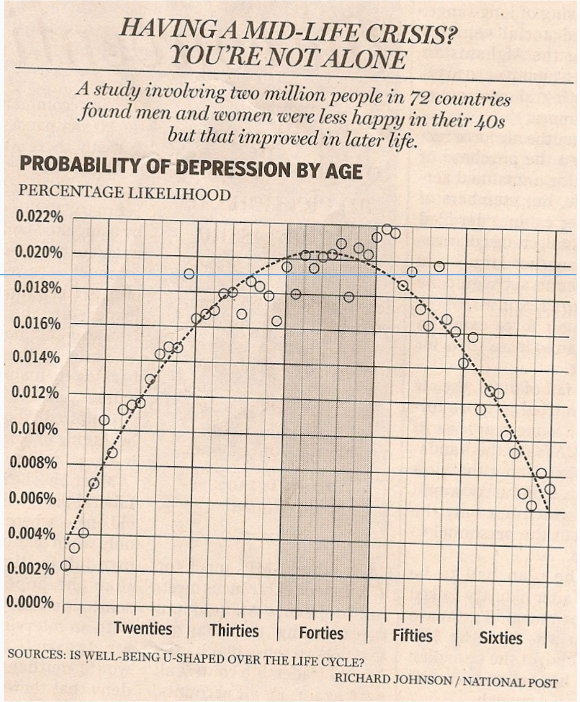
Mathematical Statistics I

STA2212H S LEC9101

Week 1

January 15 2021



1. Course Space – Amar
2. Non-regular distributions
3. HW3 from MS I

Davison

Likelihood and Estimation in SM: §4.1 – 4.4; 7.1.1

Likelihood Inference in EH: §4.1,2; 5.1,2,3,5

Office Hours: Monday 7 pm; Thursday, Friday 11 am

Bbcollaborate

Not all families are smooth

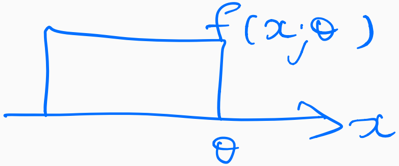
$$I_n^{1/2}(\theta)(\hat\theta-\theta) \xrightarrow{d} N(0,1) \Longrightarrow \underline{\hat\theta \,\dot\sim\, N\left(\theta, I^{-1}(\theta)\right)}$$

$\cancel{I_n^{1/2}(\hat\theta)}$

C.I.

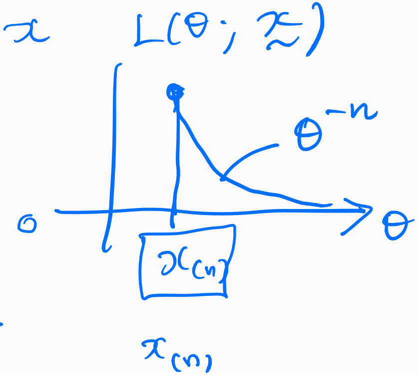
$\hat\theta \pm \ldots$

$X_1,\ldots,X_n$ iid $U(0,\theta)$



$$\prod_{i=1}^{n} f(x_i;\theta) = \begin{cases} \left(\frac{1}{\theta}\right)^n, & 0<x_i<\theta \\ 0 \end{cases} = \theta^{-n}$$

$$L(\theta; x_1,\ldots,x_n) = \begin{cases} \left(\frac{1}{\theta}\right)^n, & 0<x_1,\ldots x_n<\theta \\ 0 \end{cases}$$



$$\ell(\theta;\underset{\sim}{x}) = -n\log\theta, \quad 0 < x_{(n)} \le \theta$$

$\underline{\theta \geq x_{(n)}}$

$x_{(n)} = \hat\theta$

Not all families are smooth

$\boxed{U(0,\theta)}$ $\quad x_{(n)} = \hat{\theta}_n \quad$ X's ind't $\quad$ $\boxed{\text{not } \ell'(\hat{\theta})=0}$

$$P_\theta(X_{(n)} \le z) = P_\theta(X_1 \le z) P_\theta(X_2 \le z) \cdots P_\theta(X_n \le z)$$

$$= [F_\theta(z;\theta)]^n$$

$$f(X_{(n)};\theta) = n F(z;\theta)^{n-1} f(z;\theta)$$

$$= n\left(\frac{z}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} = \boxed{\frac{n}{\theta^n} z^{n-1}, \quad 0<z<\theta}$$

$$\frac{\partial}{\partial\theta} E(X_{(n)};\theta) = \frac{\partial}{\partial\theta}\int_0^{\theta} z \cdot f(z;\theta)\,dz \neq \int \frac{\partial}{\partial\theta} \cdots dz$$

Calculating maximum likelihood estimators

Newton-Raphson ← HW M.S 1

EM-algorithm

$$\ell'(\hat\theta) = 0 \doteq \underbrace{\ell'(\theta) + (\hat\theta - \theta)\,\ell''(\theta)}$$

$$\hat\theta \doteq \theta + \frac{\ell'(\theta)}{E[-\ell''(\theta)]}$$

Fisher scoring

in R, can use mle(loglik, ...)

wrapper optim

many methods

$$\hat\theta^{(0)} : \quad \hat\theta^{(1)} = \hat\theta^{(0)} + \frac{\ell'(\hat\theta^{(0)})}{-\ell''(\hat\theta^{(0)})}$$

$$\theta^{(t)} = \hat\theta^{(t-1)} + \frac{\ell'(\hat\theta^{(t-1)})}{-\ell''(\hat\theta^{(t-1)})}$$

$\ell'' < 0$

$\boxed{\text{glm's : smooth}}$

Newton-Raphson

$\boxed{\text{EM-algorithm}}$

AoS p. 144

E - step compute an exp^r

M - step maximize

if $L(\theta; \underset{\sim}{x})$ is hard to maximize

but latent variables so that $\tilde{L}(\theta; \underset{\sim}{x}, \underset{\sim}{z})$ is easy to maximize

mixture models $pf_1 + (1-p)f_2$

$z=1; f_1$

Problem 4 - EXTRA CREDIT

Let $X_1, \ldots, X_n \sim \text{Pareto}(\theta, \nu)$ be iid. The Pareto distribution is commonly used in social and actuarial science. The density is given by

$$\boxed{f(x|\theta,\nu) = \frac{\theta\nu^\theta}{x^{\theta+1}}, x \geq \nu, \theta > 0, \nu > 0}$$

where ν is the scale parameter and θ is the shape parameter. We will assume that ν is fixed throughout the problem.

(a) 5 pts

Find the method of moments estimator $\hat{\theta}_{MOM}$ for θ. Note that for $X \sim \text{Pareto}(\theta, \nu)$,

$$E(X) = \frac{\theta \times \nu}{\theta - 1}.$$

$X_1, \ldots, X_n$ iid Pareto

$\bar{X} = \frac{\theta \cdot \nu}{\theta - 1}$ defines $\hat{\theta}_{mom}$

ν is assumed known

(b) 5 pts

Find the MLE $\hat{\theta}_{MLE}$ for θ.

$$\hat{\theta}_{mom} = \frac{\bar{X}/\nu}{1-\bar{X}/\nu}$$

$$L(\theta) = \prod f(x_i;\theta)$$

$$= \begin{cases} \dfrac{\theta^n \nu^{n\theta}}{\prod x_i^{\theta+1}}, & x_1,\dots,x_n \geq \nu \\ 0 \end{cases}$$

$$\ell(\theta; x) = n\ln\theta + n\theta\log\nu - (\theta+1)\sum\log x_i, \quad x_{(1)} > \nu$$

suff. stat. $(x_{(1)}, \sum\log x_i)$

$$\hat{\theta}_{mle}: \quad \frac{n}{\theta} + n\log\nu - \sum\log x_i = 0 \quad \dots$$

$$\underline{\underline{\hat{\theta}_n}} = \frac{1}{\frac{1}{n}\sum\log x_i - \log(\nu)}$$

Suppose ν is unknown. Now $\hat{\theta} = \hat{\theta}_\nu$ constrained mle

$$\ell(\hat{\theta}_\nu, \nu; \underset{\sim}{x}) = n \log \hat{\theta}_{n,\nu} + n\, \hat{\theta}_{n,\nu} \log \nu - \hat{\theta}_\nu \sum \log x_i$$

$$= \ell_p(\nu; \underset{\sim}{x}) \qquad p \text{ for profile}$$

$\ell_p'(\hat{\nu}) = 0$ determine $\hat{\nu}$: $\hat{\theta}_{\hat{\nu}} = $ mle of θ

mle of ν

$$\boxed{\frac{\partial \ell(\theta, \nu; \underset{\sim}{x})}{\partial \theta} \overset{①}{=} 0 \quad \text{and} \quad \frac{\partial \ell(\theta, \nu; \underset{\sim}{x})}{\partial \nu} \overset{②}{=} 0}$$

?? $\quad \nu \geq x_{(1)}$

$$u(\hat\theta) = 0 = u(\theta) + (\hat\theta - \theta) u'(\theta) + R$$

$$= \ell'(\theta) + (\hat\theta - \theta)\ell''(\theta) + R \qquad \underbrace{(\hat\theta-\theta)^2 \ell'''(\tilde\theta)}_{\xrightarrow{p} 0}$$

$$\hat\theta - \theta = \frac{\ell'(\theta)}{-\ell''(\theta)} + R'$$

$$\sqrt{n}(\hat\theta - \theta) = \frac{\frac{1}{\sqrt n}\ell'(\theta) \xrightarrow{d} N(0, I_1(\theta))}{-\frac{1}{n}\ell''(\theta) \xrightarrow{p} E(-\ell''(\theta)) = I_1(\theta)} \qquad X_n \xrightarrow{d} \quad Y_n \xrightarrow{p}$$

$$\underline{\hat\theta \xrightarrow{p} \theta} \qquad |\tilde\theta - \theta| \le |\hat\theta - \theta| \qquad \hat\theta \xrightarrow{p} \theta \Rightarrow \tilde\theta \xrightarrow{p} \theta$$

$$E\hat\theta \simeq \theta \qquad \underline{E_\theta \hat\theta = \theta + \frac{1}{n}(\text{bias}) + O(n^{-2})}$$