Lecture 6 : More MLEs

$$L(\eta) = \left(\frac{\eta}{1+\eta}\right)^{N\bar{x}} \left(1 - \eta/_{1+\eta}\right)^{N - N\bar{x}}$$

$$\overbrace{\quad}^{1/_{1+\eta}}$$

$$\ell(\eta) = N\bar{x} \log\left(\eta/_{1+\eta}\right) + (N - N\bar{x}) \log\left(1 - \eta/_{1+\eta}\right)$$

recall: $\log(a/_b) = \log a - \log b$

$$= N\bar{x}\left(\log\eta - \log(1+\eta)\right) + (N - N\bar{x})\left(-\log(1+\eta)\right)$$

$$= N\bar{x}\log\eta - \cancel{N\bar{x}\log(1+\eta)} - N\log(1+\eta) + \cancel{N\bar{x}\log(1+\eta)}$$

$$= N\bar{x}\log\eta - N\log(1+\eta)$$

$$\frac{\partial \ell}{\partial \eta} = \frac{N\bar{x}}{\eta} - \frac{N}{1+\eta} = 0$$

$$\Rightarrow N\bar{x}(1+\eta) - N\eta = 0$$

$$\Rightarrow \cancel{N}\bar{x} + \cancel{N}\bar{x}\eta - \cancel{N}\eta = 0$$

$$\Rightarrow \bar{x} = \eta(1 - \bar{x})$$

$$\Rightarrow \boxed{\hat{\eta}_{MLE} = \frac{\bar{x}}{1 - \bar{x}}}$$

$$= \frac{\hat{p}}{1 - \hat{p}}$$

Recall: $\hat{p} = \bar{x}$

$$\eta = \frac{p}{1-p}$$

Theorem: Transformation of MLEs

If $\hat{\theta}$ is the MLE for θ then the MLE for $g(\theta)$ is $g(\hat{\theta})$.

Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$

Let's get the MLE.

no λ so I can ignore

① $L(\lambda) = f_\lambda(\underline{x}) = \prod_{n=1}^{N} \frac{e^{-\lambda} \lambda^{x_n}}{x_n!} \mathbb{1}(x_n \in \mathbb{N}_0) = \frac{e^{-N\lambda} \lambda^{\sum x_n}}{\prod_n x_n!}$

$\sum x_n = N\bar{x}$

$\ell(\lambda) = \log L(\lambda)$

$= -N\lambda + N\bar{x}\log\lambda - \log \prod_n x_n!$

② $\frac{\partial \ell}{\partial \lambda} = -\cancel{N} + \frac{\cancel{N}\bar{x}}{\lambda} = 0$

$\Rightarrow \frac{\bar{x}}{\lambda} = 1$

$\Rightarrow \boxed{\hat{\lambda} = \bar{x}}$

Ex. $X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$

Find the MLE for λ.

① $L(\lambda) = f_\lambda(\underset{\sim}{x}) = \prod_n \lambda e^{-\lambda x_n}$

$$= \lambda^N e^{-\lambda N \bar{x}}$$

$$\ell(\lambda) = N \log \lambda - \lambda N \bar{x}$$

② $\dfrac{\partial \ell}{\partial \lambda} = \dfrac{\cancel{N}}{\lambda} - \cancel{N}\bar{x} = 0$

$\Rightarrow \dfrac{1}{\lambda} = \bar{x}$

recall: $\mathbb{E} X_n = 1/\lambda$

$\Rightarrow \boxed{\hat{\lambda} = 1/\bar{x}}$

Could also consider $X_n \overset{iid}{\sim} \mathrm{Exp}(\text{mean} = \beta)$

so that $\mathbb{E} X_n = \beta$

i.e. $\beta = 1/\lambda$

What's the MLE of β?

Apply the transf. theorem, $\hat{\lambda} = 1/\bar{x}$

so $\hat{\beta} = 1/\hat{\lambda} = \bar{x}$.

Ex. $X_n \overset{iid}{\sim} U(0,\theta)$ $\theta > 0$ — $f(x) = \frac{1}{\theta}\mathbb{1}(0<x<\theta)$

whats the MLE of θ?

① $L(\theta) = f_\theta(\underline{x}) = \prod_n \frac{1}{\theta}\mathbb{1}(0<x_n<\theta)$

$$= \theta^{-N} \underbrace{\prod_n \mathbb{1}(x_n>0)}_{\mathbb{1}(x_{(1)}>0)} \underbrace{\prod_n \mathbb{1}(x_n<\theta)}_{\mathbb{1}(x_{(N)}<\theta)}$$

$$= \theta^{-N}$$

not differentiable

Since $L(\theta) = 0$ for $\theta < x_{(N)}$

$L(\theta)$ is decreasing for $\theta \geq x_{(N)}$

then the max is achieved at $\boxed{\hat{\theta} = x_{(N)}}$

Ex. $X_n \overset{iid}{\sim} N(\theta, 1)$ where $\theta \geq 0$ — $\Theta = [0,\infty)$

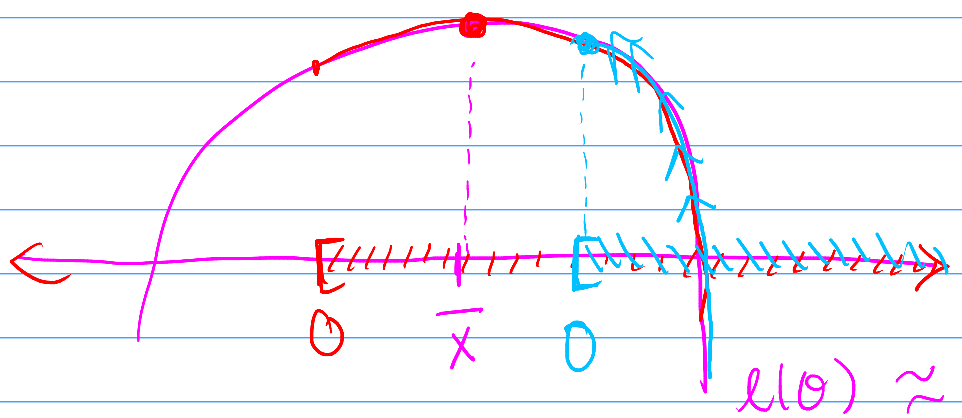
Remember: $\hat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}}\ L(\theta)$

must be in $[0,\infty)$

(1) $$L(\theta) = \prod_n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_n-\theta)^2\right)$$

$$= (2\pi)^{-N/2} \exp\left(-\frac{1}{2}\sum_n (x_n-\theta)^2\right)$$

$$\ell(\theta) = -\frac{N}{2}\lg 2\pi - \frac{1}{2}\sum_n (x_n-\theta)^2$$ looks like $-a\theta^2 + \cdots$



Case 1: $\bar{X} \geq 0$ so $\hat{\theta} = \bar{X}$

Case 2: $\bar{X} < 0$ so $\hat{\theta} = 0$

all together: $\hat{\theta} = \max(\{0, \bar{X}\})$

Ex. $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ $\mu \in \mathbb{R}$, $\sigma^2 > 0$

Let's get the MLE for both μ and σ^2.

(1) $$L(\mu, \sigma^2) = \prod_n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_n-\mu)^2\right)$$

$$= (2\pi)^{-N/2} (\sigma^2)^{-N/2} \exp\left(-\frac{1}{2\sigma^2}\sum_n (x_n-\mu)^2\right)$$

$$\ell(\mu,\sigma^2) = -\frac{N}{2}\log 2\pi - \frac{N}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_n (x_n-\mu)^2$$

(2) $\frac{\partial \ell}{\partial \mu} = 0$; $\frac{\partial \ell}{\partial \sigma^2} = 0$

$\tau = \sigma^2$

$$\ell(\mu,\tau) = -\frac{N}{2}\log 2\pi - \frac{N}{2}\log \tau - \frac{1}{2\tau}\sum_n (x_n-\mu)^2$$

$\frac{\partial \ell}{\partial \mu} = 0$ and $\frac{\partial \ell}{\partial \tau} = 0$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\tau}\sum_n(-2)(x_n-\mu) = \frac{1}{\tau}\sum_n(x_n-\mu) = 0$$

$$\Rightarrow \sum_n x_n - N\mu = 0$$

$$\Rightarrow \boxed{\hat{\mu} = \frac{1}{N}\sum_n x_n = \bar{x}}$$

$$\frac{\partial \ell}{\partial \tau} = -\frac{N}{2\tau} + \frac{1}{2\tau^2}\sum_n (x_n-\mu)^2 = 0$$

$$\Rightarrow -\frac{N}{2}\tau + \frac{1}{2}\sum_n (x_n-\mu)^2 = 0$$

$$\Rightarrow N\tau = \sum_n (x_n-\mu)^2$$

$$\Rightarrow \tau = \frac{1}{N}\sum_n (x_n-\mu)^2$$

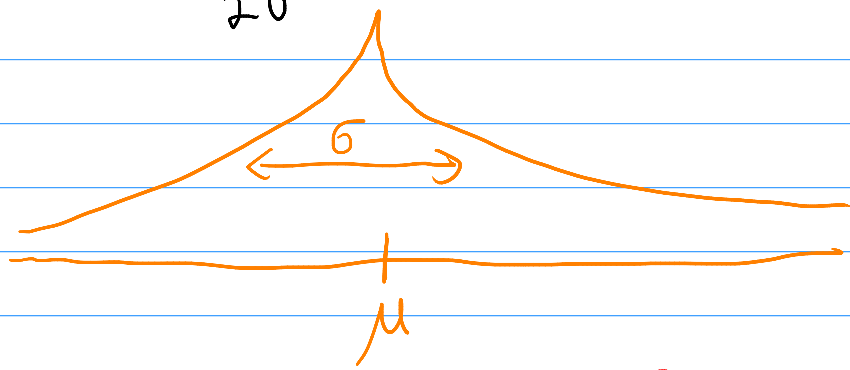
looks a lot like $S^2_{N-1} = \frac{1}{N-1}\sum_n (x_n-\bar{x})^2$

by $\frac{\partial \ell}{\partial \mu} = 0 \Rightarrow \mu = \bar{x}$ so

$$\boxed{\hat{\sigma}^2 = \hat{\tau} = \frac{1}{N}\sum_n (x_n-\bar{x})^2}$$

Ex. $X_n \overset{iid}{\sim} \text{Laplace}(\mu, \sigma)$

$$f(x) = \frac{1}{2\sigma} \exp\left(-\frac{1}{\sigma}|x-\mu|\right)$$



what's the MLE? μ and σ?

① $$L(\mu, \sigma) = \prod_n \frac{1}{2\sigma} \exp\left(-\frac{1}{\sigma}|x_n - \mu|\right)$$
$$= 2^{-N} \sigma^{-N} \exp\left(-\frac{1}{\sigma} \sum_n |x_n - \mu|\right)$$

$$\ell(\mu, \sigma) = -N \log 2 - N \log \sigma - \frac{1}{\sigma} \sum_n |x_n - \mu|$$

A problem, not diff'able wrt μ
is " " σ

$$\frac{\partial \ell}{\partial \sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^2} \sum_n |x_n - \mu| = 0$$

$$\Rightarrow N\sigma = \sum_n |x_n - \mu|$$

$$\Rightarrow \boxed{\hat{\sigma} = \frac{1}{N} \sum_n |x_n - \hat{\mu}|}$$

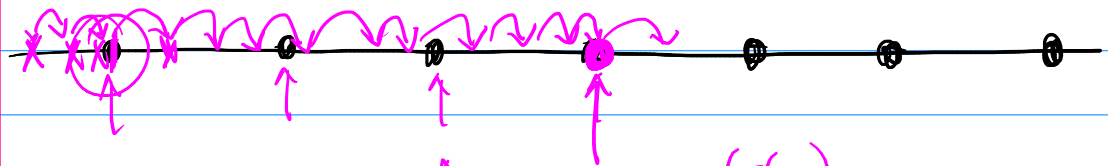
$$L(\mu,\sigma) \propto \exp\Big(-\frac{1}{\sigma}\sum_n |x_n-\mu|\Big)$$

decreases as this term increases

i.e. to maximize L

I should make $\sum_n |x_n-\mu|$ as small as possible.

$T = \sum_n |x_n-\mu|$ = total dist. of x_ns to μ



$\hat{\mu} = \text{median}(\underset{\sim}{x})$