More Maximum Likelihood Estimation!

Special Cases in this video:

- Multiple parameters!
- Parameters in Indicators!!

Example 1:

$$X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

The pdf for one of them is

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

The joint pdf for all of them is

$$f(\vec{x}; \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2)$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$f(\vec{x};\mu,\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

The parameter space:

$$-\infty < \mu < \infty, \quad \sigma^2 > 0$$

$$L(\mu,\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\ell(\mu,\sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial}{\partial \mu}\ell(\mu, \sigma^2) \overset{\text{set}}{=} 0$$

$$\frac{\partial}{\partial \sigma^2}\ell(\mu, \sigma^2) \overset{\text{set}}{=} 0$$

Solve for μ and σ^2 simultaneously.

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial}{\partial\mu}\ell(\mu, \sigma^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}2(x_i - \mu)(-1)$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) \overset{\text{set}}{=} 0$$

$$\Rightarrow \sum_{i=1}^{n}(x_i - \mu) = 0 \quad \Rightarrow \sum_{i=1}^{n}x_i - n\mu = 0$$

$$\Rightarrow \mu = \sum_{i=1}^{n}x_i/n = \bar{x} \Rightarrow \quad \hat{\mu} = \bar{X}$$

Uppercase!

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial}{\partial\sigma^2}\ell(\mu, \sigma^2) = -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 \overset{\text{set}}{=} 0$$

Multiply both sides by $2(\sigma^2)^2$ and solve for σ^2 to get:

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

The maximum likelihood estimators for μ and σ^2 are:

Uppercase!

$$\hat{\mu} = \overline{X}$$

$$\widehat{\sigma^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}$$

Note: $\mathrm{Var}[X] = E[(X - \mu)^2]$

A "natural" estimator is the sample variance.

$$S_1^2 := \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}$$

However,

$$E[S_1^2] = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}\right] = \cdots = \frac{n-1}{n}\sigma^2$$

So an unbiased version of the sample variance is

$$S^2 := \frac{n}{n-1}S_1^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

Unless otherwise specified, we will be using S^2.

Back to the Normal Example:

$$\hat{\mu} = \overline{X}$$

$$\widehat{\sigma^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}$$

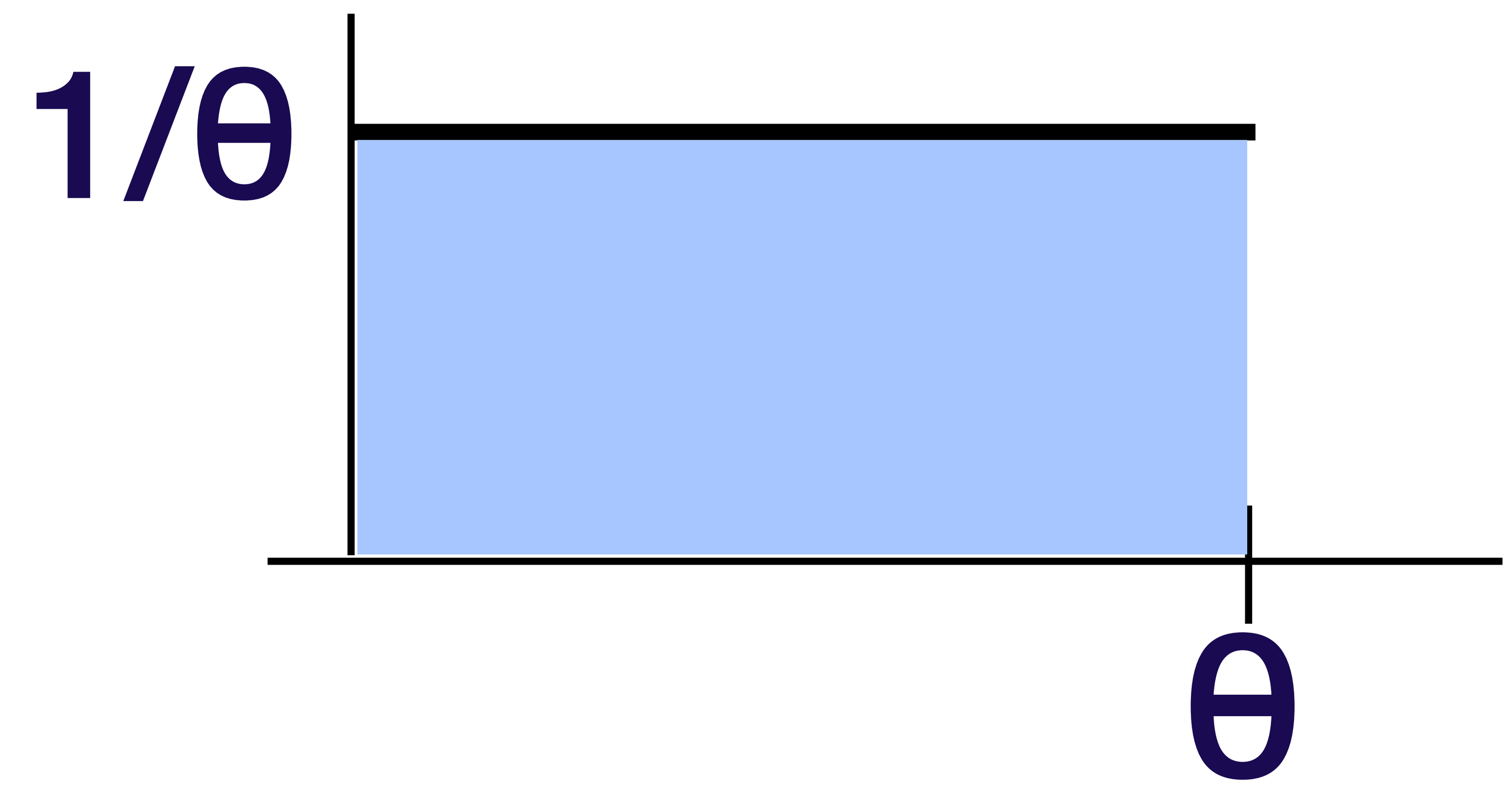
Another biased MLE! 😳

Example 2:

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} unif(0,\theta)$$

The pdf for one of them is

$$f(x;\theta) = \frac{1}{\theta} I_{(0,\theta)}(x)$$



The joint pdf for all of them is

$$f(\vec{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$= \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

$$f(\vec{x};\theta) = \frac{1}{\theta^n} \underbrace{\prod_{i=1}^{n} I_{(0,\theta)}(x_i)}$$

- Can't delete. It is part of the likelihood.
- However...

$$f(x) = \begin{cases} x^2 & , \quad 0 < x \le 1 \\ x + 1 & , \quad x > 1 \end{cases}$$

$$f'(x) = \begin{cases} 2x & , \quad 0 < x \le 1 \\ 1 & , \quad x > 1 \end{cases}$$

You don't take the derivative of this part!!!

$$f(\vec{x};\theta) = \frac{1}{\theta^n} \underbrace{\prod_{i=1}^{n} I_{(0,\theta)}(x_i)}$$

- Can't delete. It is part of the likelihood.
- However...

$$f(x) = \begin{cases} x^2 & , \quad 0 < x \leq 1 \\ x + 1 & , \quad x > 1 \end{cases}$$

$$\ln f(x) = \begin{cases} \ln x^2 & , 0 < x \leq 1 \\ \ln(x+1) & , \quad x > 1 \end{cases}$$

You don't take the log of this part!!!

$$f(\vec{x};\theta)=\frac{1}{\theta^n}\prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

Can’t delete but won’t be taking derivatives and logs . Let’s just put it aside but remember that it’s there!

maximize!

$$L(\theta)=\frac{1}{\theta^n}\prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

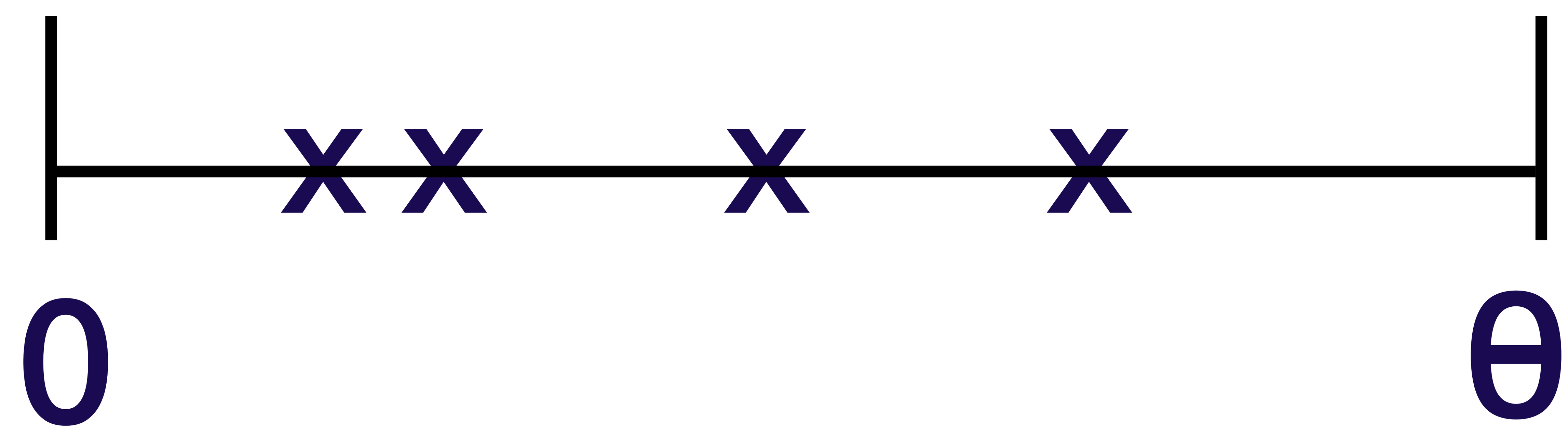
$$\ell(\theta)=-n\ln\theta$$

$$\frac{\partial}{\partial\theta}\ell(\theta)=-\frac{n}{\theta}\overset{\text{set}}{=}0 \quad ??$$

$$L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

Decreasing function of θ.

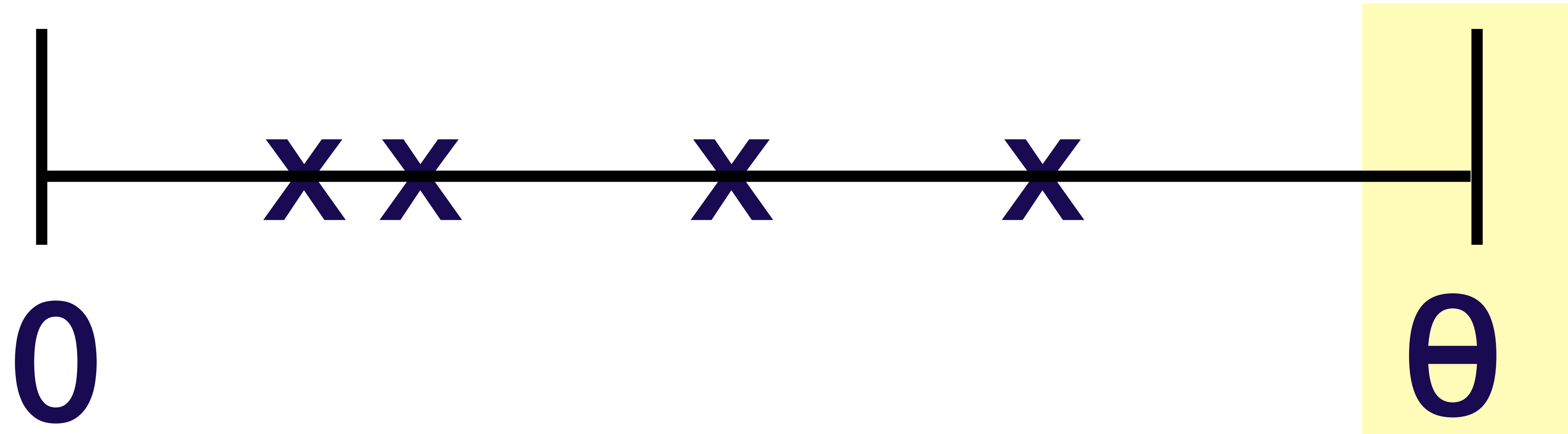
To maximize $L(\theta)$, take θ as small as possible.



$$L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

Decreasing function of θ.

To maximize $L(\theta)$, take θ as small as possible.



The smallest possible θ is the largest possible X!

The maximum likelihood estimator of θ is

$$\hat{\theta} = \max(X_1, X_2, \ldots, X_n)$$