Thursday Jan. 30

Maximum Likelihood Estimation Continued

Model: $X_1, \dots X_n \overset{ind}{\sim} p(x|\theta)$ or $f(x|\theta)$

Likelihood function

$$L(\theta) = \begin{cases} \prod_{i=1}^{n} p(x_i|\theta) & \text{if discrete} \\ \prod_{i=1}^{n} f(x_i|\theta) & \text{if continuous} \end{cases}$$

Probability of getting the observed data, as a function of θ

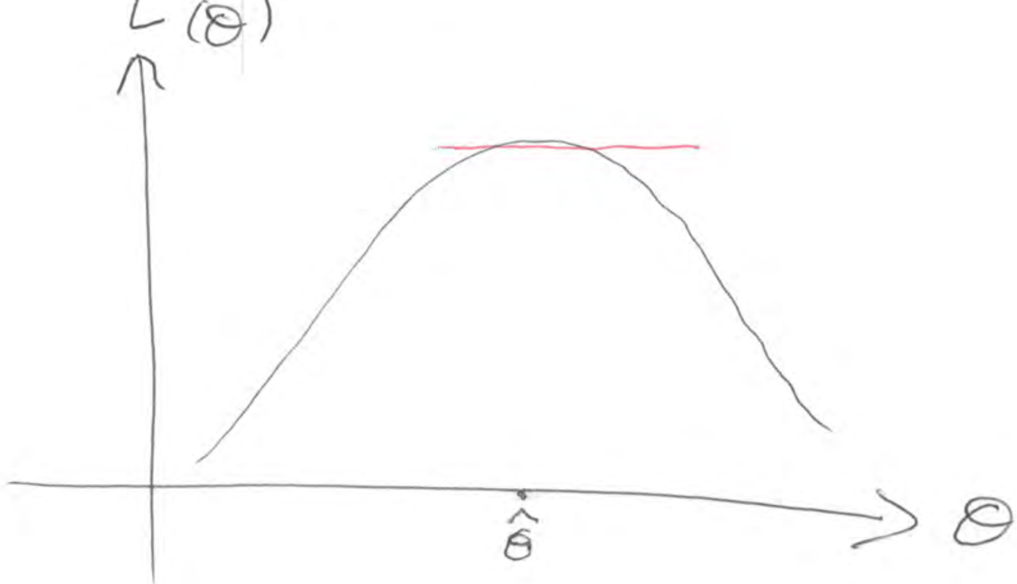
Book: "$L(\theta)$ establishes a preference ordering on $\theta \in \Omega$"

If $L(\theta_1) > L(\theta_2)$ for given data we prefer θ_1 over θ_2

Both Likelihoods (probabilities) could be very small. What matters is RELATIVE likelihood

$$\frac{L(\theta_1)}{L(\theta_2)} \quad \text{(Likelihood ratio)}$$

Method of Maximum Likelihood estimation is to estimate θ with with the value that makes $L(\theta)$ the greatest.



Ex $X_1, \ldots, X_n \overset{iid}{\sim}$ Bernoulli(θ)

$$L(\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$= \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{\sum_{i=1}^{n}(1-x_i)} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

Could differentiate w/resp to θ, but easier is Log likelihood

$\ell(\theta) = \ln[L(\theta)]$ Set $\ell'(\theta) = 0$

Why OKAY? Bec $\ln \uparrow$ $a > b$, $\ln(a) > \ln(b)$

4

$$\ell(\theta) = \ln[L(\theta)] = \ln\left(\theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}\right)$$

$$= \sum_{i=1}^{n} x_i \ln\theta + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-\theta)$$

$$\frac{d\ell}{d\theta} = \frac{\sum_{i=1}^{n} x_i}{\theta} + \frac{(n-\Sigma x_i)}{1-\theta}(-1) \overset{set}{=} 0$$

$$\Longrightarrow \frac{\sum_{i=1}^{n} x_i}{\theta} = \frac{n-\Sigma x_i}{1-\theta}$$

$$\Longrightarrow \sum_{i=1}^{n} x_i(1-\theta) = \theta\left(n - \sum_{i=1}^{n} x_i\right)$$

$$\Longrightarrow \sum_{i=1}^{n} x_i - \cancel{\theta\sum_{i=1}^{n} x_i} = n\theta - \cancel{\theta\sum_{i=1}^{n} x_i}$$

$$\Longrightarrow \theta = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}_n$$

MAX ?

$$\frac{\partial^2 \ell}{\partial\theta^2} = \frac{\partial}{\partial\theta}\left[\left(\sum_{i=1}^{n} x_i\right)\theta^{-1} + (n-\Sigma x_i)(1-\theta)^{-1}\right]$$

$$= \frac{\sum_{i=1}^{n} x_i}{\theta^2}(-1) + \frac{(n-\Sigma x_i)}{(1-\theta)^2}(-1) < 0$$

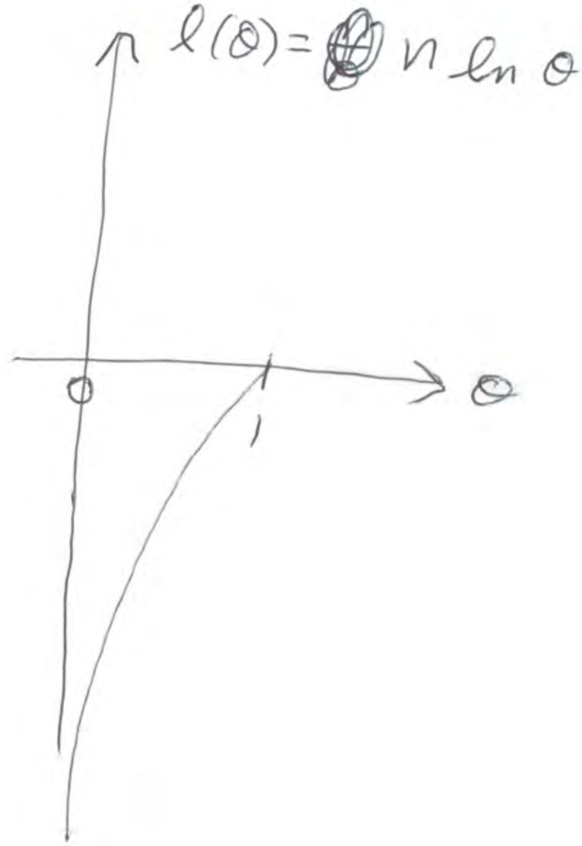
$\hat{\theta} = \bar{x}_n$

What happens if $\sum_{i=1}^{n} x_i = n$

$$L(\theta) = \theta^{\Sigma x}(1-\theta)^{n-\Sigma x} = \theta^{n}$$

$$l(\theta) = n \ln\theta, \quad l'(\theta) = \frac{n}{\theta} \stackrel{set}{=} 0$$

$$\Longrightarrow \cancel{n = 0}$$

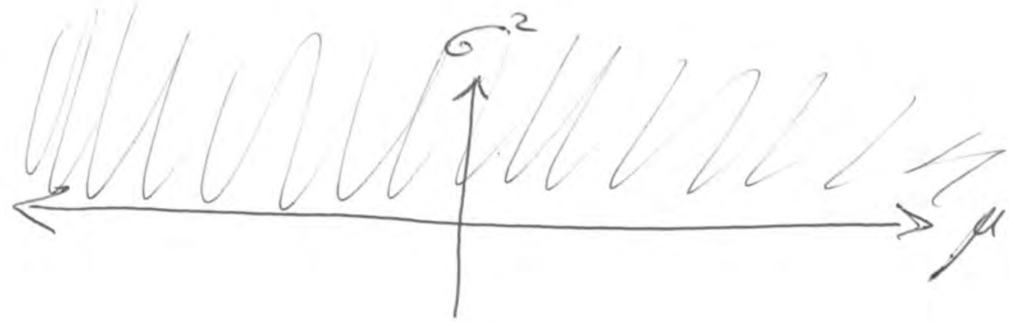


$$\frac{n}{\theta} > 0 \quad l \uparrow$$

MAX at $\theta = 1$

Ex $X_1 \cdots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Find MLE of $\theta = (\mu, \sigma^2)$

$\Omega =$ 

$$\ell(\theta) = \ln \prod_{i=1}^{n} f(x_i \mid \mu, \sigma^2)$$

$$= \ln \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

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

$$= \ln \left[(\sigma^2)^{-\frac{n}{2}} (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2} \right]$$

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

Strategy $\frac{\partial \ell}{\partial \mu} \overset{set}{=} 0$ & $\frac{\partial \ell}{\partial \sigma^2} \overset{set}{=} 0$

$$\frac{d\ell}{d\mu} = 0 + 0 - \frac{1}{2\sigma^2}\frac{\partial}{\partial\mu}\sum_{i=1}^{n}(x_i-\mu)^2$$ 7

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

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

$$\Longrightarrow \sum x_i = n\mu \quad (1)$$