Lecture 5 - MLEs

Thursday, September 12, 2024 11:01 AM

Maximum Likelihood Estimation (MLE)

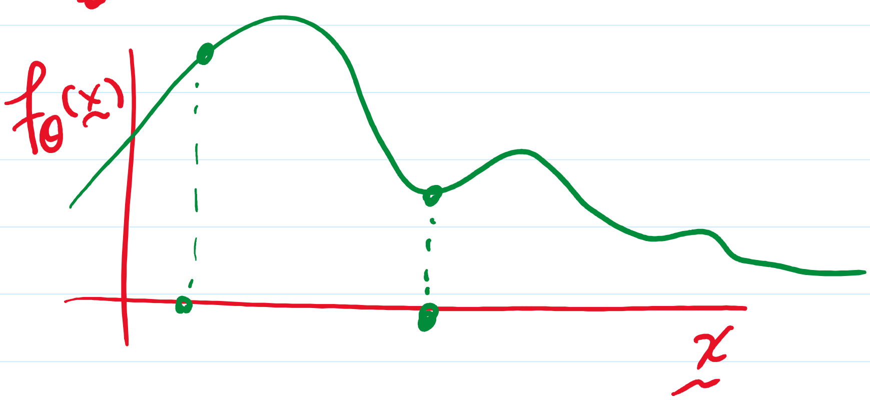
Assume $X_n \overset{iid}{\sim} f_\theta$ where $\theta \in \Theta$

recall: joint

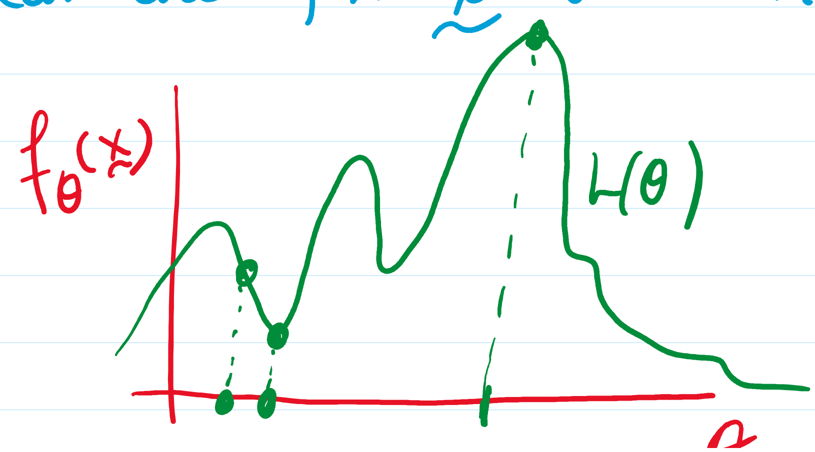
$$f_\theta(\underset{\sim}{x}) = \prod_{n=1}^{N} f_\theta(x_n)$$

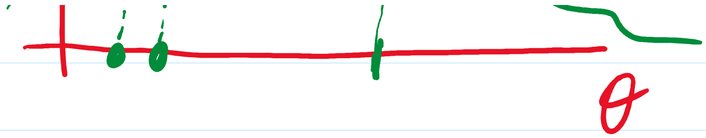
typ. look at as a function of $\underset{\sim}{x}$

$f_\theta : \mathbb{R}^N \to \mathbb{R}$



Can also fix $\underset{\sim}{x}$ and think of as a fn of θ





Called the likelihood function

$$L : \Theta \to \mathbb{R}$$

defn:

$$L(\theta) = f_\theta(\underline{x})$$

MLE says choose $\hat{\theta}_{MLE}$ as val. of θ that max. L.

MLE:

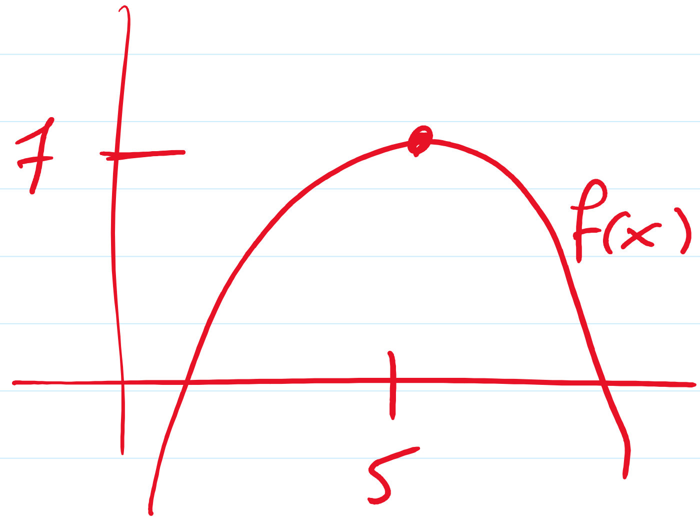
$$\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\arg\max}\, L(\theta)$$

---

argmax

$$\max_x f(x) = 7$$

$$\underset{x}{\operatorname{argmax}} f(x) = 5$$



Often, we work with log-likelihood

$$\ell(\theta) = \log L(\theta).$$

Alt. defn of MLE

$$\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\operatorname{argmax}} \ell(\theta)$$

This is equiv. b/c log is inc. fn.

Ex, $X_n \overset{iid}{\sim} N(\theta, 1)$ where $\theta \in \mathbb{R}$.

What is the MLE?

What is the MLE?

① Let's get $\ell(\theta)$

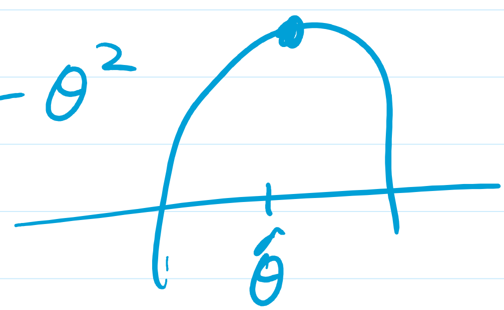
$$L(\theta) = f_\theta(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n)$$

$$= \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_n-\theta)^2\right)$$

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

$$\ell(\theta) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{n=1}^{N}(x_n-\theta)^2$$

$\propto -\theta^2$



② Let's get $\frac{\partial \ell}{\partial \theta}$

$$\frac{\partial \ell}{\partial \theta} = 0 - \frac{1}{2}\sum_{n=1}^{N} 2(x_n-\theta)(-1)$$

$$\sum^{N}(x_n - \theta)$$

$$= \sum_{n=1}^{N} (x_n - \theta)$$

③ Set $\frac{\partial \ell}{\partial \theta} = 0$ and solve for θ

$$\frac{\partial \ell}{\partial \theta} = \sum_{n=1}^{N} x_n - N\theta = 0$$

$$\text{So } \boxed{\hat{\theta}_{MLE} = \frac{1}{N} \sum_{n=1}^{N} x_n = \overline{X}}$$

Technically, need to check $\frac{\partial^2 \ell}{\partial \theta^2} < 0$ and consider $\lim_{\theta \to \pm\infty} L(\theta) = 0$

Theorem: MLEs are functions of SS

$$\hat{\theta}_{MLE} = \text{function}(T)$$

$\leftarrow$ T suff.

$\hat{}_{MLE}$ $\uparrow$ T suff. for θ.

Pf: Factorization thm says

$$L(\theta) = f_\theta(\underset{\sim}{x}) = g(T, \theta) h(\underset{\sim}{x})$$

So

$$\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\operatorname{argmax}} L(\theta)$$

$$= \underset{\theta \in \Theta}{\operatorname{argmax}}\, g(T, \theta) h(\underset{\sim}{x})$$

$$= \underset{\theta \in \Theta}{\operatorname{argmax}}\, g(T, \theta)$$

$\uparrow$ fn of T

$$\ell(\theta) = \log h(\underset{\sim}{x}) + \log g(T, \theta)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{\partial}{\partial \theta} \log g(T, \theta)$$

Ex. $X_n \overset{iid}{\sim} \text{Bern}(p)$, $p \in [0,1]$

What's my MLE for p?

1 Get $L(p)$ and $\ell(p)$

$$L(p) = f_p(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n)$$

$$= \prod_{n=1}^{N} p^{x_n}(1-p)^{1-x_n}\mathbb{1}(x_n = 0 \text{ or } 1)$$

$$\boxed{\begin{aligned} \bar{X} &= \frac{1}{N}\sum_n X_n \\ N\bar{X} &= \sum_n X_n \end{aligned}}$$

$$= p^{\sum_n x_n}(1-p)^{\sum_n (1-x_n)} \prod_n \mathbb{1}(X_n = 0 \text{ or } 1)$$

$$= p^{N\bar{X}}(1-p)^{N(1-\bar{X})} \prod_n \mathbb{1}(X_n = 0 \text{ or } 1)$$

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

$$= N\bar{X}\cdot\log(p) + N(1-\bar{X})\log(1-p)$$

$$+ \log \prod_n \mathbb{1}(X_n = 0 \text{ or } 1)$$

$$+ \underbrace{\log \prod_n \mathbb{1}(x_n = 0 \text{ or } 1)}_{0}$$

② Take deriv, set eq to zero.

$$\frac{\partial \ell}{\partial p} = \frac{N\bar{x}}{p} - \frac{N(1-\bar{x})}{1-p} = 0$$

$$\Rightarrow \frac{N\bar{x}}{p} = \frac{N(1-\bar{x})}{1-p}$$

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

$$\Rightarrow N\bar{x} - \cancel{N\bar{x}}p = Np - \cancel{N\bar{x}}p$$

$$\Rightarrow N\bar{x} = Np$$

$$\Rightarrow \boxed{\hat{p}_{MLE} = \bar{x}} = \text{pct of data that are 1s.}$$

1-1 fn w/ D

Consider

$$\eta = \frac{p}{1-p} = \text{odds} \qquad p = \frac{\eta}{(1+\eta)}$$

1-1 fn w/ p

Might want $\hat{\eta}_{MLE}$.

Get likelihood in terms of η ...

$$L(p) = p^{N\bar{x}}(1-p)^{N(1-\bar{x})} \cdots$$

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

$$\hat{\eta} = \underset{\eta}{\operatorname{argmax}}\, L(\eta)$$

(1) Get $\frac{\partial \ell}{\partial \eta}$.

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

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

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

$$\frac{\partial \ell}{\partial \eta} = N\bar{x}\left[\frac{1}{\eta} - \frac{1}{1+\eta}\right]$$
$$- N(1-\bar{x})\left(\frac{1}{1+\eta}\right)$$

$$= \frac{N\bar{x}}{\eta} - \cancel{\frac{N\bar{x}}{1+\eta}} - \frac{N}{1+\eta} + \cancel{\frac{N\bar{x}}{1+\eta}}$$

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

(2) Set $\frac{\partial \ell}{\partial \eta} = 0$ solve for η.

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

$$\Rightarrow \frac{\bar{x}}{\eta} = \frac{1}{1+\eta}$$

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

$\Rightarrow \bar{x} + \bar{x}\eta = \eta$

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

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

Recall: $\eta = \dfrac{p}{1-p}$.

Theorem: Transf. of MLEs

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

Ex. Let $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$, $\lambda > 0$.

Let's get MLE for λ.

(1) $L(\lambda) = f_\lambda(\underset{\sim}{x}) = \prod_{n=1}^{N} f(x_n)$

ignore

① $L(\lambda) = f_\lambda(\tilde{x}) \quad \prod_{n=1}$

$$= \prod_{n=1}^{N} \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} \mathbb{1}(x_n \in \mathbb{N}_0)$$

ignore

no λ

$$= \frac{\lambda^{\sum_n x_n} e^{-N\lambda}}{\prod_n x_n!}$$

$$\ell(\lambda) = \left(\sum_n x_n\right) \log(\lambda) - N\lambda - \log\left(\prod_n x_n!\right)$$

② $\dfrac{\partial \ell}{\partial \lambda} = 0$

$$\frac{\partial \ell}{\partial \lambda} = \left(\sum_n x_n\right) \frac{1}{\lambda} - N - 0 = 0$$

$$\Rightarrow \sum_n x_n / \lambda - N = 0$$

$$\Rightarrow \boxed{\hat{\lambda} = \frac{1}{N} \sum x_n = \bar{x}}$$

$$\Rightarrow \boxed{\hat{\lambda} = \frac{1}{N}\sum_n x_n = \overline{x}}$$