Lecture 4: Maximum likelihood estimation

Reminders

- HW 1 due Friday 11am on Canvas
- You are always welcome to work together on assignments, just make sure to write up your own solutions
- You have a bank of 5 extension days; you can use 1 or 2 on any assignment or exam

Recap: maximum likelihood estimation

Definition: Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a sample of n observations, and let $f(\mathbf{y}|\theta)$ denote the joint pdf or pmf of $\mathbf{Y}$, with parameter(s) θ. The *likelihood function* is

function of θ $\rightarrow$

$$L(\theta|\mathbf{Y}) = f(\mathbf{Y}|\theta)$$

Definition: Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a sample of n observations. The *maximum likelihood estimator* (MLE) is

$$\hat{\theta} = \operatorname{argmax}_\theta L(\theta|\mathbf{Y})$$

Common approach:

- write down likelihood, take log
- differentiate, set $= 0$
- solve

Example: $N(\theta, 1)$

$\theta \in (-\infty, \infty)$

$Y_1, \ldots, Y_n \overset{iid}{\sim} N(\theta, 1)$

$$f(y|\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\theta)^2}$$

$$L(\theta|Y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(Y_i - \theta)^2\right\}$$

$$= (2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_i (Y_i - \theta)^2\right\}$$

$$\ell(\theta|Y) = \log L(\theta|Y) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(Y_i - \theta)^2$$

$$\frac{\partial}{\partial\theta}\ell(\theta|Y) = -\frac{1}{2}\sum_{i=1}^{n} 2(Y_i - \theta)(-1) = \sum_{i=1}^{n}(Y_i - \theta) \overset{\text{set}}{=} 0$$

$$\Rightarrow \sum_{i=1}^{n} Y_i = n\theta \Rightarrow \theta = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}$$

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta|Y) = \frac{\partial}{\partial\theta}\sum_{i=1}^{n}(Y_i - \theta) = -n < 0 \checkmark$$

$\Rightarrow$ unique maximum $(\hat{\theta} = \overline{Y})$

Example: $\mathrm{Uniform}(0, \theta)$

Let $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \mathrm{Uniform}(0, \theta)$, where $\theta > 0$. We want the maximum likelihood estimator of θ.

Discuss with your neighbors what the MLE of θ might be.

Hint: focus on finding and sketching the likelihood function $L(\mathbf{Y}|\theta)$

$Y_1, \dots, Y_n \overset{iid}{\sim} \text{Uniform}(0, \theta)$

$$f(y|\theta) = \begin{cases} \frac{1}{\theta} & 0 \le y \le \theta \\ 0 & \text{else} \end{cases}$$

$$= \frac{1}{\theta} \underbrace{\mathbb{1}\{0 \le y \le \theta\}}_{\substack{=1 \text{ if } 0 \le y \le \theta \\ =0 \text{ else}}}$$

$$L(\theta|Y) = \prod_{i=1}^{n} \frac{1}{\theta} \mathbb{1}\{0 \le Y_i \le \theta\}$$

$$= \frac{1}{\theta^n} \underbrace{\prod_{i=1}^{n} \mathbb{1}\{0 \le Y_i \le \theta\}}_{\substack{=1 \text{ if } 0 \le Y_i \le \theta \ \forall i \\ =0 \text{ else}}}$$

$$= \mathbb{1}\{0 \le Y_1, Y_2, \dots, Y_n \le \theta\}$$

$$= \mathbb{1}\{0 \le Y_{(n)} \le \theta\}$$

$Y_{(1)} = \min\{Y_1, \dots, Y_n\}$

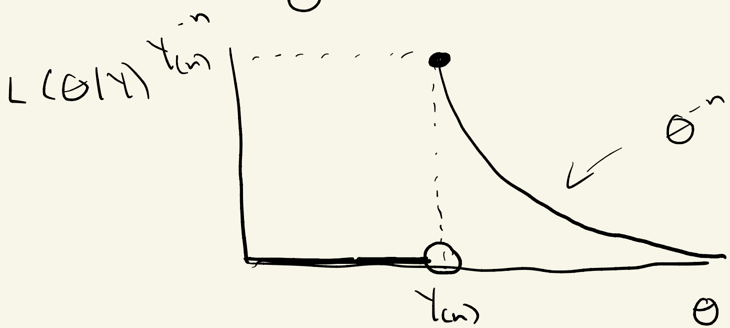
$Y_{(n)} = \max\{Y_1, \dots, Y_n\}$

$$\Rightarrow L(\theta|Y) = \frac{1}{\theta^n} \mathbb{1}\{0 \le Y_{(n)} \le \theta\}$$

Intuition: suppose $Y_i = 0.5$

can $\theta = 0.4$? No!

$\Rightarrow \theta \ge$ all Y_i



$L(\theta|Y) = 0 \quad \forall \theta < Y_{(n)}$

$L(\theta|Y) \downarrow$ in θ for $\theta \ge Y_{(n)}$

$$\Rightarrow \arg\max_{\theta} L(\theta|Y) = Y_{(n)}$$

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

Example: $N(\mu, \sigma^2)$

$$Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$$

$$\Theta = (\mu, \sigma^2)$$
$$\mu \in (-\infty, \infty)$$
$$\sigma^2 > 0$$

$$f(y \mid \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2}(y-\mu)^2 \right\}$$

$$\Rightarrow \quad L(\theta \mid Y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - \mu)^2 \right\}$$

$$\Rightarrow \quad \ell(\theta \mid Y) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mu)^2$$

Plan: maximize wrt μ, then maximize wrt σ^2

$$\frac{\partial}{\partial \mu} \ell(\theta \mid Y) = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - \mu)(2)(-1) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \mu) \overset{set}{=} 0$$

$$\Rightarrow \quad \sum_{i=1}^{n}(Y_i - \mu) = 0 \quad \Rightarrow \quad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}$$

For any value of σ^2

$$L(\theta|Y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\mu)^2\right\}$$

$$\leq (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right\}$$

Want to maximize

$$\ell^*(\sigma^2|Y) = -\frac{n}{2}\underbrace{\log(2\pi\sigma^2)}_{\log(2\pi)+\log(\sigma^2)} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\bar{Y})^2$$

$$\frac{d\ell^*}{d\sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i-\bar{Y})^2 \overset{\text{set}}{=} 0$$

$$\Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i-\bar{Y})^2}$$

N.B. compare to $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2$ ...

Logistic regression

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{i,1} + \cdots + \beta_k X_{i,k}$$

Suppose we observe independent samples $(X_1, Y_1), \ldots, (X_n, Y_n)$. Write down the likelihood function

$$L(\beta|\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n} f(Y_i|\beta, X_i)$$

for the logistic regression problem.