Recap : Maximum likelihood estimation

Example 8.5.2. As in Example 8.1.4, suppose that the data $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ are independently sampled from a $\mathrm{N}(\theta, 1)$ distribution. The likelihood $L(\theta|\mathbf{x})$ is

$$L(\theta|\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2\right).$$

$$f(x_i|\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \theta)^2\right)$$

$$L(\theta|x) = f(x|\theta)$$

$$= \prod_{i=1}^{n} f(x_i|\theta) \quad (X_i \text{ indep.})$$

$$= \prod_{i=1}^{n} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}(x_i - \theta)^2\right)$$

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

equiv.

$$L(\theta|x) = \exp\left(-1/2 \sum_{i=1}^{n}(x_i - \theta)^2\right)$$

$$\frac{d}{d\theta} L(\theta|x)$$

$$= \frac{d}{d\theta}\left[(2\pi)^{-n/2} \exp\left(-1/2 \sum_{i=1}^{n}(x_i - \theta)^2\right)\right.$$

$$= (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right) \cdot$$

$$\cdot\, 2\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)\right)(-1)$$

$$\frac{d}{d\theta} L(x|\theta) = 0$$

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

$$\Leftrightarrow \quad \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\theta = 0$$

$$\Leftrightarrow \quad n\bar{x} - n\theta = 0$$

$$\boxed{\begin{aligned} \bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ n\bar{x} &= \sum_{i=1}^{n} x_i \end{aligned}}$$

$$\Leftrightarrow \quad n\bar{x} = n\theta$$

$$\Leftrightarrow \quad \theta = \bar{x}$$

Max or min? Second derivative

$$\frac{d}{d\theta} L(\theta|x) = (2\pi)^{-n/2} \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right)$$

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

$$\frac{d^2}{d\theta^2} L(\theta|x)$$

$$= 2\pi^{-n/2}\Big[\exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right)\left(\sum_{i=1}^{n}(x_i-\theta)\right)^2$$

$$+ \underbrace{\exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right)}_{A}\sum_{i=1}^{n}(-1)\Big]$$

$$\frac{d^2}{d\theta^2} L(\theta|x)\Big|_{\theta=\bar{x}}$$

$$= 2\pi^{-n/2}\left[A\left(\sum_{i=1}^{n}(x_i-\theta)\right)\Big|_{\theta=\bar{x}} + A(-n)\right]$$

$$A\big|_{\theta=\bar{x}} = \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$= \exp\left(\frac{-(n-1)s^2}{2}\right) > 0$$

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

$$= n\bar{x} - n\bar{x} = 0$$

$$\frac{d^2}{d\theta^2} L(\theta|x) = (2\pi)^{-n/2}\left(A \cdot 0 + A(-n)\right)$$

$$= -n \underbrace{A\,(2\pi)^{-n/2}}_{>0} < 0$$

$\Rightarrow$ local maximum.

Need to check boundary of values for θ (domain of L)

$\theta \in \mathbb{R}$. boundaries: $\pm\infty$

$$L(\theta|x) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right)$$

$$\lim_{\theta\to\infty} L(\theta|x) \to \exp(-\infty) = 0$$

$$\lim_{\theta\to-\infty} L(\theta|x) = 0$$

$\Rightarrow \theta = \bar{x}$

$$L(\theta|x)\big|_{\theta=\bar{x}} = \underbrace{\frac{1}{(2\pi)^{n/2}} \exp\left(\frac{-(n-1)s^2}{2}\right)}_{>0}$$

$\Rightarrow \hat{\theta} = \bar{x}$ is the maximum likelihood estimate of θ

Maximum likelihood ESTIMATOR

$$\hat{\Theta} = \bar{X}$$

function of random variables

ML estimate is $f(x_1, x_2, \ldots, x_n)$

ML estimator is $f(X_1, X_2, \ldots, X_n)$

<u>Exercise 1.2.10</u>

$$\sum_{i=1}^{n} (x_i - a)^2 \geq \sum_{i=1}^{n} (x_i - \bar{x})^2$$

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

$$\exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-a)^2\right) \leq \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$L(\theta|x) = \frac{1}{(2\pi)^{n/2}}\exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right)$$

$$\leq \frac{1}{(2\pi)^{n/2}}\exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

globally.

$\Rightarrow$ $\bar{x}$ is MLE of θ.

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

Example 8.5.5

Suppose the independent random variables $X_1, X_2, \dots, X_n$ follow a $\text{Bern}(\theta)$

θ unknown, and we observe data $x_1, \dots, x_n$.

The likelihood is:

$$P(X=x) = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \end{cases}$$

$$= \theta^{x}(1-\theta)^{1-x} \quad \text{if } x \in \{0,1\}$$

$$\theta^{1}(1-\theta)^{1-1} = \theta$$

$$\theta^{0}(1-\theta)^{1-0} = 1-\theta$$

For each x_i :

$$P(X=x_i) = \theta^{x_i}(1-\theta)^{1-x_i}$$

$$\begin{aligned} L(\theta|x) &= f(x|\theta) \\ &= \prod_{i=1}^{n} f(x_i|\theta) \\ &= \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} \\ &= \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \\ &= \theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}} \end{aligned}$$

Let $y = n\bar{x}$

$$L(\theta|x) = \theta^{y}(1-\theta)^{n-y} \quad \theta \in [0,1]$$

Maximise log-likelihood

$$a < b$$
$$\log(a) < \log(b)$$

$$\log L(\theta|x) = y\log\theta + (n-y)\log(1-\theta)$$
$$\theta \in (0,1)$$

$$\frac{d}{d\theta}\log L(\theta|x) = \frac{y}{\theta} + \frac{(n-y)}{(1-\theta)}(-1)$$

$$= \frac{y}{\theta} - \frac{n-y}{1-\theta}$$

Set equal to 0:

$$\frac{y}{\theta} - \frac{n-y}{1-\theta} = 0$$

$$\frac{y}{\theta} = \frac{n-y}{1-\theta}$$

$$y(1-\theta) = (n-y)\theta$$

$$y - \cancel{y\theta} = n\theta - \cancel{y\theta}$$

$$\theta = \frac{y}{n} = \frac{\sum x_i}{n} = \bar{x}$$

- check θ on boundary

$\theta = 0$:

$$L(\theta|x)|_{\theta=0} = \theta^y (1-\theta)^{n-y} = 0$$

$\theta = 1$

$$L(\theta|x)|_{\theta=1} = 1^y (0)^y = 0$$

$$0 < y < n$$

what if $y = 0$ or $y = n$

Second derivative

$$\frac{d}{d\theta} \log L = \frac{y}{\theta} - \frac{n-y}{1-\theta}$$

$$(n-y)(1-\theta)^{-1}$$

$$\frac{d^2}{d\theta^2} \log L = -\frac{y}{\theta^2} - \left[-(n-y)(1-\theta)^{-2}(-1)\right]$$

$$= -\frac{y}{\theta^2} - \frac{1-y}{(1-\theta)^2} < 0$$

Example 8.5.6. Suppose the observations $x_1, x_2, \dots, x_n$ are independently sampled from a $\mathrm{N}(\mu, \sigma^2)$ distribution, with both $\theta_1 = \mu$ and $\theta_2 = \sigma^2$ unknown. Similarly to Example 8.1.4, the likelihood can be written as

$$L(\mu, \sigma^2|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right).$$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

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

Simpson's Paradox

Adult passengers and crew members				
Class	Survived	Did not survive	Total	Survival rate
Third	151	476	627	**24.08 %**
Crew	212	673	885	23.95 %

Table 5.1: The survival rate of adult 3rd-class passengers and crew members on the Titanic.

Men					Women				
Class	Survived	Did not survive	Total	Survival rate	Class	Survived	Did not survive	Total	Survival rate
Third	75	387	462	16.23%	Third	76	89	165	46.06%
Crew	192	670	862	**22.27%**	**Crew**	20	3	23	**86.96%**

Table 5.2: The survival rates of crew members and third-class passengers on the Titanic when the data is split into male and female subgroups.

All patients					
Treatment	Improved	Not Improved	Total	Percent Improved	
Standard	24	16	40	**60%**	a+b
New	20	20	40	50%	A+B

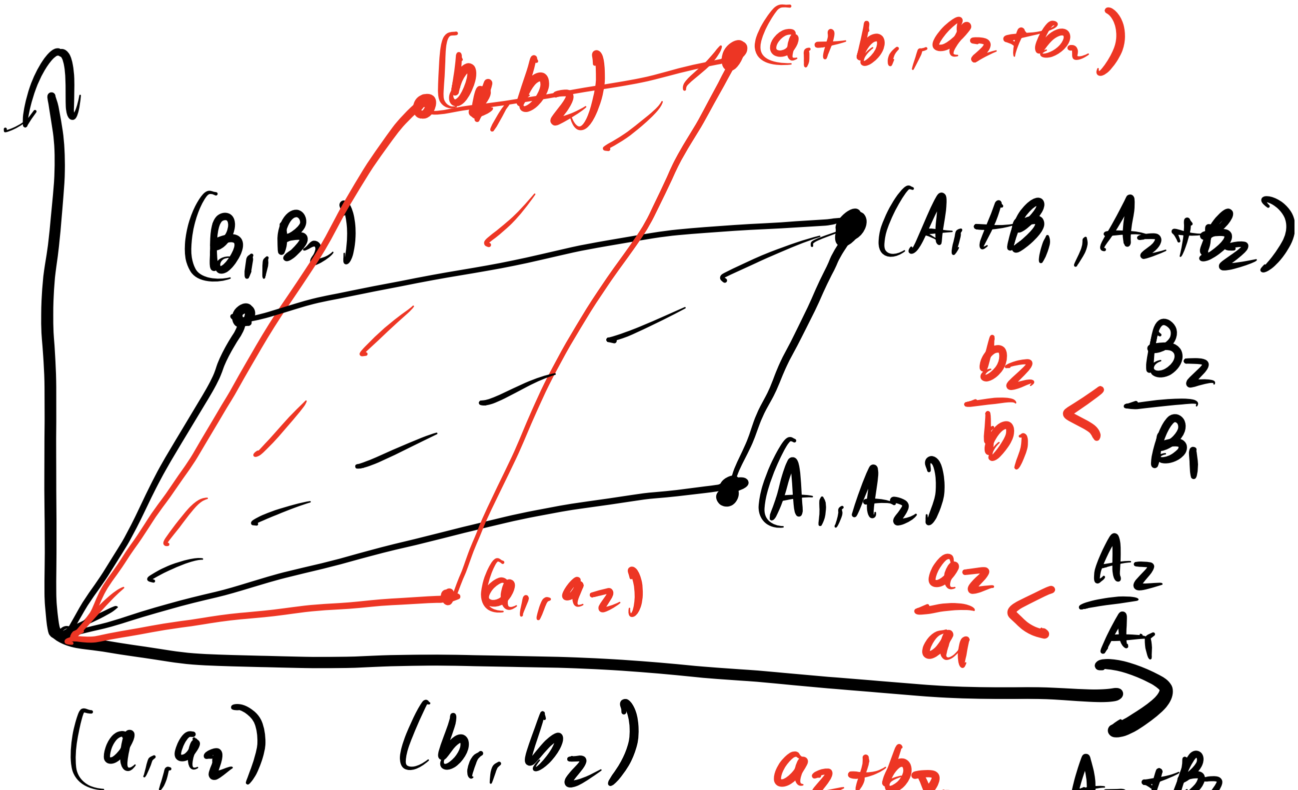
Table 5.3: Improvement percentage across patients for the standard and new treatments.

Group A					Group B				
Treatment	Improved	Not improved	Percent improved		Treatment	Improved	Not Improved	Percent improved	
Standard	3	7	30%	a_1	Standard	21	9	70%	b
New	12	18	**40%**	A	**New**	8	2	**80%**	B

Table 5.4: Improvement percentage across patients for the standard and new treatments with the data disaggregated in to subgroups Group A and Group B.

improved a_2 : improvement rate a_2/a_1

total : a_1



$$\frac{b_2}{b_1} < \frac{B_2}{B_1}$$

$$\frac{a_2}{a_1} < \frac{A_2}{A_1}$$

$$\frac{a_2+b_2}{a_1+b_1} > \frac{A_2+B_2}{A_1+B_1}$$

$<$