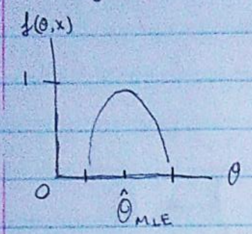
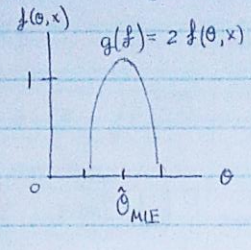
Lecture 2 – Frequentist model

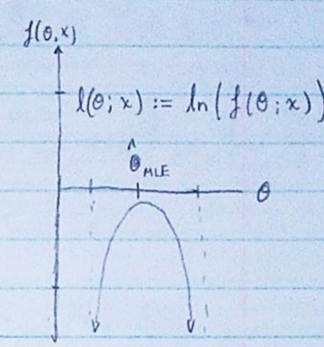
$$\hat{\theta}_{MLE} := \underset{\theta \in \Theta}{\operatorname{argmax}} \{ f(\theta; X) \} = \underset{\theta \in \Theta}{\operatorname{argmax}} \{ g(f) \} = \underset{\theta \in \Theta}{\operatorname{argmax}} \{ l(\theta; x) \} = \operatorname{argmax} \{ \ln(f(\theta; x)) \}$$

$$\forall_{x,y} \quad x > y \implies g(x) > g(y)$$



$\leftarrow \theta \rightarrow$





sample avg

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

$$\sum x_i = n\bar{x}$$

ex: $X = \langle 0,0,1,0,1,0 \rangle \qquad \left[X_1, \ldots, X_n \overset{iid}{\sim} Bern(\theta) \right]$

$$f(\theta; x_1 \ldots x_6) = P(x_1 \ldots x_6; \theta) = \prod_{i=1}^{6} P(x_i; \theta) = \prod_{i=1}^{6} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$l(\theta; x_1 \ldots x_6) = \ln\left(\prod_{i=1}^{6} \theta^{x_i} (1-\theta)^{1-x_i} \right)$$

$$= \sum_{i=1}^{6} \ln\left(\theta^{x_i} (1-\theta)^{1-x_i} \right)$$

$$= \sum_{1}^{6} (x_i) \ln\theta + (1-x_i)\ln(1-\theta)$$

Note

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

$$\overset{*}{=} (6\bar{x})\ln(\theta) + (6-6\bar{x})\ln(1-\theta)$$

$$6\left(\bar{x}\ln(\theta) + (1-\bar{x})\ln(1-\theta) \right)$$

to find Max $\hat{\theta}_{MLE}$

$\frac{d}{dx}[\ ] = 0$

set equal

find max

$$l'(\theta; X) = 6\left(\frac{\bar{x}}{\theta} - \frac{1-\bar{x}}{1-\theta} \right) = 0$$

$X_n : n = 6 \qquad = \hat{p} = \bar{x}$

$\bar{x} = 2/6 \qquad \hat{\theta}_{MLE} = 1/3$

$\theta =$

for $X \overset{iid}{\sim} Bern(\theta)$

Generally –

$$\left[\frac{\bar{x}}{\theta} - \frac{1-\bar{x}}{1-\theta} = 0 \right] \implies \left[\frac{\bar{x}}{\theta} = \frac{1-\bar{x}}{1-\theta} \right]$$

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

so

$$\left[\hat{\theta}_{MLE} = \bar{x} \right]$$

<u>Note</u> X is a r.v. like a function $\qquad$ $\overline{X}$ r.v. of multiple r.v., an average.
x : $x \in \text{supp}[X]$ $\qquad$ $\bar{x}$ (average)

- Properties of $\hat{\theta}_{MLE}$

① $\hat{\theta}_{MLE} \xrightarrow{P} \theta$

$$\forall \varepsilon > 0 : \lim_{n \to \infty} P\left(|\hat{\theta}_{MLE} - \theta| \geq \epsilon \right) = 0$$

for $\{X := \langle x_1, x_2, \ldots, x_n \rangle\}$ as n gets large $\hat{\theta}_{MLE}$ resembles true θ

② Asymptotic Normality

$$\hat{\theta}_{MLE} \xrightarrow{d} N\left(\theta, \underset{\text{standard error}}{SE}[\hat{\theta}_{MLE}]^2\right)$$

③ Efficiency
$SE[\hat{\theta}_{MLE}]$ is theoretically lowest

Recall!
$$** \; SE[X] = \sqrt{Var(\overline{X})} = \sqrt{Var\left(\frac{\Sigma x_i}{n}\right)}$$
$$= \sqrt{\frac{1}{n^2} Var(\Sigma x_i)} = \sqrt{\frac{1}{n^2} \Sigma Var(x_i)}$$
$$= \sqrt{\frac{n}{n^2} Var(x_i)} = \sqrt{\frac{\theta(1-\theta)}{n}}$$

find max likelihood est. for:

<u>ex:</u> $\left[X_1, \ldots, X_n \overset{iid^*}{\sim} Geom(\theta) \right]$ $\qquad$ $f(\theta; X) = (1-\theta)^{x_i}\theta$
$X : x_1, x_2, \ldots, x_i$
$\theta \in \Theta : (0, 1)$

so
$$f(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i; \theta)$$

and
$$f(\theta, x_i) = \prod_{i=1}^{n} (1-\theta)^{x_i}\theta \Rightarrow l(\theta; x_1 \ldots x_n) = \sum_{i=1}^{n} \ln\left((1-\theta)^{x_i}\theta\right)$$
$$= \sum_{i=1}^{n} x_i \ln(1-\theta) + \ln\theta \overset{*}{=} n\left(\bar{x}\ln(1-\theta) + \ln\theta\right)$$

$\frac{d}{dx}[\;] = 0$
$$l'(\theta; X) = n\left(\frac{1}{\theta} - \frac{\bar{x}}{1-\theta}\right) = 0$$
$$\frac{1}{\theta} = \frac{\bar{x}}{1-\theta} \Rightarrow \frac{1}{\theta} - 1 = \bar{x}$$
$$\frac{1}{\theta} = \bar{x} + 1 \Rightarrow \theta = \frac{1}{\bar{x}+1}$$

for $X \overset{iid}{\sim} Geom(\theta)$
$$\left[\hat{\theta}_{MLE} = \frac{1}{\bar{x}+1} \right]$$