Homework-1

T. sai varsha
EE18BTECH11042

1.1)

i) $\underline{X} = x_1, x_2, x_3 \ldots, x_n$ $\{1, 2, \ldots m\}$

Take $\theta_1, \theta_2, \ldots \theta_m$ are estimated ML parameters.

$\cancel{P(x_i/\theta)}$ $\mathbb{I}(x_i = j) = \begin{cases} 1 & \text{if } x_i = j \\ 0 & \text{else} \end{cases}$

$$P(x_i|\theta) = \prod_{j=1}^{m} \theta_j^{\mathbb{I}(x_i = j)}$$

$$L(\theta_1, \theta_2, \ldots, \theta_m) = \underset{\substack{\theta_i \\ \sum \theta_i = 1}}{\arg\max} \; P(\underline{X}|\theta)$$

$$= \prod_{i=1}^{n} P(x_i|\theta)$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{m} \theta_j^{\mathbb{I}(x_i = j)}$$

$$= \prod_{j=1}^{m} \prod_{i=1}^{n} \theta_j^{\mathbb{I}(x_i = j)}$$

$$= \prod_{j=1}^{m} \theta_j^{\sum_{i=1}^{n} \mathbb{I}(x_i = j)}$$

$$L(\theta_1, \theta_2, --- \theta_n) = \prod_{j=1}^{m} \theta_j^{\sum_{i=1}^{n} I(x_i = j)}$$

Take,

$$n_j(\underline{x}) \sum_{i=1}^{n} I(x_i = j)$$ = no. of data points in $\underline{x}$ for which $x_i = j$

$$= n_j(\underline{x})$$ = no. of occurances of j

$$\Rightarrow L(\theta_1, \theta_2, --- \theta_n) = \prod_{j=1}^{m} \theta_j^{n_j(\underline{x})}$$

$$\Rightarrow \log L(\theta_1, \theta_2, -- \theta_n) = \sum_{j=1}^{m} n_j(\underline{x}) \log \theta_j$$

$$\Rightarrow \frac{1}{n} \log L(\theta_1, \theta_2, --- \theta_n) = \sum_{j=1}^{m} \frac{n_j(\underline{x})}{n} \log \theta_j$$

$$\Rightarrow \boxed{\frac{n_j(\underline{x})}{n} \text{ is empirical pmf}} \rightarrow \text{Take, } q_j = \frac{n_j(\bar{x})}{n}$$

$$\Rightarrow \frac{1}{n} \log L(\theta_1, \theta_2, --- \theta_n) = \sum_{j=1}^{m} q_i \log \theta_j$$

$$= \sum_{j=1}^{m} q_j \log \theta_j - q_j \log q_j + q_j \log q_j$$

$$= -\sum_{j=1}^{m} q_j \log \frac{q_j}{\theta_j} + \sum_{j=1}^{m} q_j \log q_j$$

$$= -D(q \| \theta) - H(q)$$

we know,

$$D(q \| \theta) \geq 0$$

To maximize log likelihood, $D(q\|\theta)$ should be 0.

so,

$$D(q\|\theta) = 0$$

$$\Rightarrow \quad q = \theta$$

so,

$$\Rightarrow \quad \boxed{\theta_i = q_i = \frac{n_i(\underline{x})}{n}}$$

$\Rightarrow$ ML estimator $\hat{\theta}$ equal to empirical pmf.

$$\Rightarrow E[\theta_j] = E\left[\frac{n_j(\underline{x})}{n}\right]$$

We know,

$$n_j(\underline{x}) = \sum_{i=1}^{n} I(x_i = j)$$

$$I(x_i = j) = \begin{cases} 1 & \text{if } x_i = j \\ 0 & \text{else} \end{cases}$$

$$E[I(x_i = j)] = 0 \times P[I_a(x_i = j) = 0] + 1 \times P[I_a(x_i = j) = 1]$$

$$= P[x_i = j] = P_x(j)$$

$$E[\theta_j] = \cancel{E[nx]} \frac{E[n_j(\underline{x})]}{n}$$

$$= \frac{E\left[\sum_{i=1}^{n} I(\cancel{x_i}\, x_i = j)\right]}{n}$$

$$= \frac{\sum_{i=1}^{n} E[I(x_i = j)]}{n}$$

$$= \frac{\sum_{i=1}^{n} P_X(j)}{n}$$

$$= \frac{\cancel{n} P_X(j)}{\cancel{n}}$$

so,

$$\Rightarrow \boxed{E[\theta_j] = P_X(j)}$$

so, it is unbiased estimator.

ii) $m_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i = E[x^i]$

Take,

$$m_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i$$

$$= \frac{1}{n}\sum_{a=1}^{m} n_a(\underline{x})\, a^i$$

$$m_i = \sum_{a=1}^{m} \frac{n_a(\underline{x})}{n}\, a^i$$

$$m_1 = \sum_{a=1}^{m} \frac{n_a(\underline{x})}{n}\, a$$

$$m_2 = \sum_{a=1}^{m} \frac{n_a(\underline{x})}{n}\, a^2$$

$$\vdots$$

$$m_m = \sum_{a=1}^{m} \frac{n_a(\underline{x})}{n}\, a^m$$

We can write,

$$\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_m \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & \cdots & m \\ 1^2 & 2^2 & 3^2 & \cdots & m^2 \\ \vdots \\ 1^m & 2^m & 3^m & \cdots & m^m \end{bmatrix} \begin{bmatrix} \frac{n_1(\underline{x})}{n} \\ \frac{n_2(\underline{x})}{n} \\ \vdots \\ \frac{n_m(\underline{x})}{n} \end{bmatrix}$$

Eq 1

$$\cancel{E[x^i] = \sum_{a=1}^{m} a^m p_X(a)}$$

$$\cancel{E[X] = \sum_{a=1}^{m} a^m}$$

$$E[x^i] = \sum_{a=1}^{m} a^i p_X(a)$$

so,

$$E[X_1] = \sum_{a=1}^{m} a\, p_X(a)$$

$$E[X_2] = \sum_{a=1}^{m} a^2 p_X(a)$$

$$\vdots$$

$$E[X_m] = \sum_{a=1}^{m} a^m p_X(a)$$

we can write this as,

$$\begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_m] \end{bmatrix} = \begin{bmatrix} 1 & 2 & \cdots & m \\ 1^2 & 2^2 & \cdots & m^2 \\ \vdots & & & \\ 1^m & 2^m & \cdots & m^m \end{bmatrix} \begin{bmatrix} p_X(1) \\ p_X(2) \\ \vdots \\ p_X(m) \end{bmatrix}$$

Eq (2)

we have equate,

$$\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_m \end{bmatrix} = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_m] \end{bmatrix}$$

So)

$$\begin{bmatrix} 1 & 2 & \cdots & m \\ 1^2 & 2^2 & \cdots & m^2 \\ \vdots & & & \\ 1^m & 2^m & \cdots & m^m \end{bmatrix} \begin{bmatrix} \frac{n_1(\underline{X})}{n} \\ \frac{n_2(\underline{X})}{n} \\ \vdots \\ \frac{n_m(\underline{X})}{n} \end{bmatrix} = \begin{bmatrix} 1 & 2 & \cdots & m \\ 1^2 & 2^2 & \cdots & m^2 \\ \vdots & & & \\ 1^m & 2^m & \cdots & m^m \end{bmatrix} \begin{bmatrix} P_X(1) \\ P_X(2) \\ \vdots \\ P_X(m) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \frac{n_1(\underline{X})}{n} \\ \frac{n_2(\underline{X})}{n} \\ \vdots \\ \frac{n_m(\underline{X})}{n} \end{bmatrix} = \begin{bmatrix} P_X(1) \\ P_X(2) \\ \vdots \\ P_X(m) \end{bmatrix}$$

$$\Rightarrow \boxed{P_X(a) = \frac{n_a(\underline{X})}{n}}$$

estimated pmf using moment matching

So, here estimator is equal to empirical pmf's.

We proved in 1.1) a) that this estimator is unbiased estimator

Moment matching & ML estimator gives us the same estimator. Both performances is same .

iii)

$$\hat{P_x}(a) = \frac{n_a(X)+1}{n+m}$$

$$E[\hat{P_x}(a)] = \frac{E[n_a(X)+1]}{n+m}$$

$$= \frac{E[n_a(X)]+1}{n+m}$$

We proved,

$E[n_a(X)] = nP_x(a)$ in 1.1) a. so,

$$E[\hat{P_x}(a)] = \frac{nP_x(a)+1}{n+m}$$

$$= \frac{P_x(a)+\frac{1}{n}}{1+\frac{m}{n}}$$

Here,

$E[\hat{P_x}(a)] \neq P_x(a)$. so, it is a biased estimator.

As $\underset{n\to\infty}{Lt}$ $E[\hat{P_x}(a)] = \underset{n\to\infty}{Lt} \frac{P_x(a)+\frac{1}{n}}{1+\frac{m}{n}}$ $\left[\because \text{As } n\to\infty, \frac{1}{n}\to 0\right]$

$$= \frac{P_x(a)+0}{1+0}$$

$$= P_x(a).$$

Asymptotically,

$$\lim_{n \to \infty} E[\hat{P_x}(a)] = P_x(a)$$

So, asymptotically ~~t~~ add-1 estimator is a unbiased estimator.

1.3) We have to estimate h_0, h_1 from y_i & x_i

$$y_i = \begin{cases} h_0 x_0 + h_1 x_{n-1} + z_0 & \text{if } i=0 \\ h_0 x_i + h_1 x_{i-1} + z_i & \text{if } i = 1,2,3,\dots n-1 \end{cases}$$

$$z_i \sim N(0, \sigma^2)$$

$$L(h_0, h_1) = $$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - h_0 x_i - h_1 x_{i-1})^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N e^{-\sum_{i=1}^{n} \frac{(y_i - h_0 x_i - h_1 x_{i-1})^2}{2\sigma^2}}$$

$$\log L(h_0, h_1) = -\frac{N}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n} \frac{(y_i - h_0 x_i - h_1 x_{i-1})^2}{2\sigma^2}$$

$$\Rightarrow \frac{d}{dh_0} \log(L(h_0, h_1)) = 0$$

$$\Rightarrow \sum (y_i - h_0 x_i - h_1 x_{i-1}) x_i = 0$$

$$\Rightarrow \boxed{\sum x_i y_i - h_0 \sum x_i^2 - h_1 \sum x_i x_{i-1} = 0} \quad (1)$$

$$\Rightarrow \frac{d}{dh_1}[\log(L(h_0, h_1))] = 0$$

$$\Rightarrow \quad \sum(y_i - h_0 x_i - h_1 x_{i-1}) x_{i-1} = 0$$

$$\Rightarrow \boxed{\sum x_{i-1} y_i - h_0 \sum x_i x_{i-1} - h_1 \sum x_{i-1}^2 = 0} \longrightarrow (2)$$

Eq (1) & (2) can be ~~re~~written as,

$$\Rightarrow \begin{bmatrix} \sum x_i y_i \\ \sum x_{i-1} y_i \end{bmatrix} = \begin{bmatrix} \sum x_i^2 & \sum x_i x_{i-1} \\ \sum x_i x_{i-1} & \sum x_{i-1}^2 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix}$$

& $\boxed{\sum x_i^2 = \sum x_{i-1}^2}$. so,

$$\Rightarrow \begin{bmatrix} \sum x_i y_i \\ \sum x_{i-1} y_i \end{bmatrix} = \begin{bmatrix} \sum x_i^2 & \sum x_i x_{i-1} \\ \sum x_i x_{i-1} & \sum x_i^2 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix}$$

$$\Rightarrow \boxed{\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \sum x_i^2 & \sum x_i x_{i-1} \\ \sum x_i x_{i-1} & \sum x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum x_i y_i \\ \sum x_{i-1} y_i \end{bmatrix}}$$