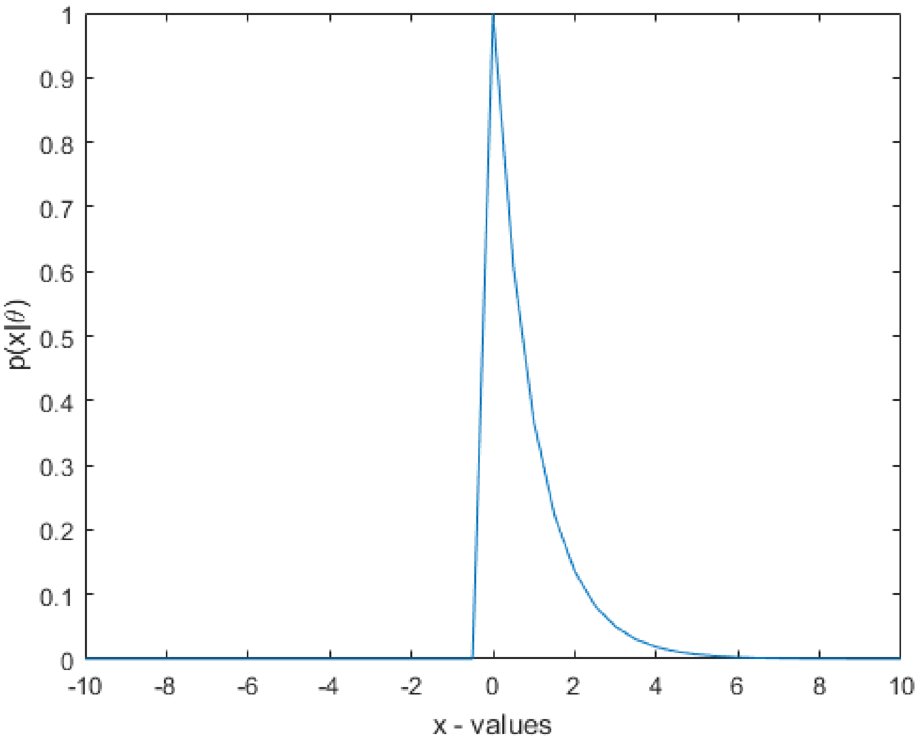
Fundamentals of Statistical Learning & Pattern Recognition – CSE 569

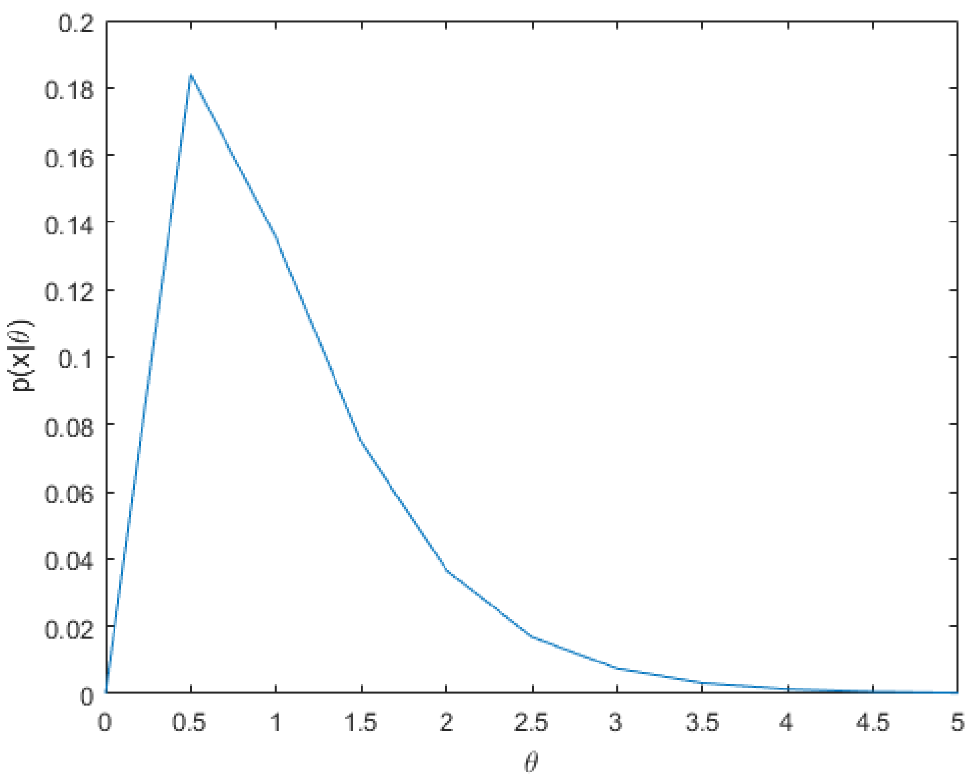
Name: Dhananjayan Santhanakrishnan ASU ID # 1211181423

Problem # 1

(a) Plot $p(x|\theta)$ versus x for θ =1 (i.e., $p(x|\theta)$ is viewed as a function of x).



(b) Plot $p(x|\theta)$ versus θ , $0 \leq \theta \leq 5$, for x =2 ((i.e., $p(x|\theta)$ is viewed as a function of θ).



(c) Given a training set of n samples D ={x1, x2, …, xn}(i.i.d. samples drawn from the above distribution), find the MLE for θ.

1. (c)

$$p(x|\theta) = \begin{cases} \theta e^{-\theta x} & \text{if } x>0 \\ 0 & \text{otherwise} \end{cases}$$

set of 'n' samples $D = \{x_1, x_2, x_3, \dots, x_n\}$

find MLE for θ

$$l(\theta) = \log p(D|\theta) = \sum_{k=1}^{n} \log p(x_k|\theta)$$

$$= \sum_{1}^{n} \log \theta e^{-\theta x_k} = n \log\theta + \sum_{1}^{n} (-\theta x_k)$$

$$l(\theta) = n \log \theta - \theta \sum_{k=1}^{n} x_k$$

$$\nabla l(\theta) = 0$$

$$\Rightarrow \frac{n}{\theta} - \sum_{k=1}^{n} x_k = 0 \Rightarrow \frac{n}{\theta} = \sum_{1}^{n} x_k$$

$$\Rightarrow \boxed{\theta = \frac{n}{\sum_{1}^{n} x_K}}$$

Problem #2 (a) find value of 'c'?

$$p(x|\mu) = ce^{-|2x-\mu|}$$

To find value of c integrating the above eq$\underline{n}$ and equating to '1'.

$$\int_{-\infty}^{\infty} p(x|\mu)\,dx = 1 \Rightarrow \int_{-\infty}^{\infty} c e^{-|2x-\mu|}\,dx = 1$$

breaking up integral

$$\Rightarrow c\int_{-\infty}^{0} e^{-(-(2x-\mu))}\,dx + c\int_{0}^{\infty} e^{-(2x-\mu)}\,dx = 1$$

$$\Rightarrow c\left[\lim_{a\to-\infty}\int_{a}^{0} e^{2x-\mu}\,dx + \lim_{b\to\infty}\int_{0}^{b} e^{-2x+\mu}\,dx\right] = 1$$

$$\Rightarrow c\left[\lim_{a\to-\infty} \frac{e^{2x-\mu}}{2}\bigg|_{a}^{0} + \left\{-\frac{e^{-2x+\mu}}{2}\right\}_{0}^{b}\right] = 1$$

$$\Rightarrow c\left[\left\{\frac{e^{-\mu}}{2} - \frac{e^{2a-\mu}}{2}\right\}_{a\to-\infty} + \left\{-\frac{e^{-2b+\mu}}{2} + \frac{e^{\mu}}{2}\right\}_{b\to\infty}\right] = 1$$

$$\Rightarrow c\left[\frac{e^{-\mu}}{2}+\frac{e^{\mu}}{2}\right]=1$$

remaining terms equates to zero
Since $e^{-\infty}=0$

$$\Rightarrow \boxed{c=\frac{2}{e^{-\mu}+e^{\mu}}}$$

(b) Assume we have training data set $D=\{x_1\ x_2\ x_3\}$ find MLE of μ given D

$$p(x)=\frac{2}{e^{-\mu}+e^{\mu}}\, e^{-|2x-\mu|}$$

$$=\frac{2e^{\mu}}{1+e^{2\mu}}\, e^{-|2x-\mu|}$$

$$l(\mu)=2\ln e^{\mu}-\ln(1+e^{2\mu})+\{-|2x-\mu|\}$$

$$= 2\mu - \ln(1+e^{2\mu}) - \sum_{i=1}^{3} |2x_i - \mu|$$

$$\nabla l(\mu) = 0$$

$$\Rightarrow \quad 2 - \frac{2e^{2\mu}}{e^{2\mu}+1} - \sum_{i=1}^{3} \operatorname{sgn}(2x_i - \mu) = 0$$

$$\Rightarrow \quad \frac{2e^{2\mu}}{e^{2\mu}+1} + \sum_{i=1}^{3} \operatorname{sgn}(2x_i - \mu) = 2$$

In summary $\mu = \text{median}(x_1, x_2, x_3)$ is the maximum likelihood estimator.

Problem # 3

$$p(X = n \mid \lambda) = \frac{\lambda^n}{n!} e^{-\lambda} \qquad D = \{x_1\}$$

$$l(\theta) = \log p(D \mid \lambda) = \log \left(\frac{\lambda^{x_1}}{(x_1)!} e^{-\lambda} \right)$$

$$= \log \frac{\lambda^{x_1}}{(x_1)!} + \log \frac{e^{-\lambda}}{(x_1)!} = \frac{1}{(x_1)!} \left[\log \lambda^{x_1} - \lambda \right]$$

$$\nabla l(\theta) = 0 \Rightarrow \frac{d}{d\lambda} l(\theta) = 0$$

$$\frac{1}{(x_1)!} \left[\frac{1}{\lambda} (x_1) - 1 \right] = 0$$

$$\Rightarrow \frac{1}{\lambda} (x_1) = 1 \Rightarrow \lambda = x_1$$

So, Max Likelihood Estimate for

$$\lambda = x_1$$

(b) $f(\lambda) = \begin{cases} e^{-\lambda} & \text{if } \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$

$$p(\lambda | D) = \frac{p(D|\lambda)\, p(\lambda)}{\int p(D|\theta)\, p(\theta)\, d\theta}$$

Ignoring denominator as it is constant

$$p(D|\lambda) = \frac{\lambda^{x_1}}{(x_1)!} e^{-\lambda} \quad \because D = \{x_1\}$$

$$p(\lambda) = e^{-\lambda} \quad \lambda > 0.$$

$$p(\lambda | D) = \frac{\lambda^{x_1}}{(x_1)!} e^{-2\lambda}$$

To find estimate for λ, we can find mean of posterior density

$\hat{\theta} = E(\theta | D)$, ||| ly $\hat{\lambda} = E(\lambda | D)$

$$= \int_0^\infty \lambda\, p(\lambda|D) = \int_0^\infty \lambda \left(\frac{\lambda^{x_1}}{(x_1)!} e^{-2\lambda} \right) d\lambda$$

$$\hat{\theta} = \int_0^\infty \lambda \left(\frac{\lambda^{x_1}}{(x_1)!} e^{-2\lambda} \right) d\lambda$$

$$= \int_0^\infty \frac{\lambda^{x_1+1}}{(x_1)!} e^{-2\lambda} d\lambda = \frac{e^{-2}}{(x_1)!} \int_0^\infty \lambda^{(x_1+2)-1} e^{-\lambda} d\lambda$$

$$= \frac{e^{-2}}{(x_1)!} \Gamma(x_1+2)$$ According to gamma function definition

$$= \frac{e^{-2}}{(x_1)!} (x_1+1)! = \frac{e^{-2}}{(x_1)!} (x_1+1)(x_1)!$$

$$= e^{-2}(x_1+1)$$

$$\boxed{\hat{\lambda} = e^{-2}(x_1+1)}$$

Problem # 4 HMM Training Model

Code:

```
obs = importdata('Observations.txt');
obsArr = [];

% converting the read values to 1 and 2 representing Heads and Tails
for i=1:size(obs,1)
    str = obs(i);
    str = char(str);
    numArr = uint8(str);
    numArr(2:2:end,:)=[];
    numArr(:,2:2:end)=[];
    obsArr = [obsArr;numArr];
end
for i=1:size(obsArr,1)
    for j=1:size(obsArr,2)
        if(obsArr(i,j)==72)
            obsArr(i,j) = 1;
        end
        if(obsArr(i,j)==84)
            obsArr(i,j) = 2;
        end
    end
end

states = importdata('States.txt');
statesArr = [];

% reading states from file
for i=1:size(states,1)
    str = states(i,:);
    numArr = uint8(str);
    statesArr = [statesArr;numArr];
end

% estimating transision and emission parameters
[trans,emis] = hmmestimate(obsArr,statesArr);

transNew = [0.5 0.5;0.5 0.5];
sumH = 0;
sumT = 0;
for i=1:size(obsArr,1)
    for j=1:size(obsArr,2)
        if (obsArr(i,j)==1)
            sumH = sumH+1;
        end
        if(obsArr(i,j)==2)
            sumT = sumT+1;
        end
    end
end
emisNew = [sumH/600 sumT/600;sumH/600 sumT/600];

[updatedTrans,updatedEmis] = hmmtrain(obsArr,transNew,emisNew);
states = [];
states = hmmviterbi(obsArr,trans,emis);
```

4(a). Output Emission Matrix:

Variables - emis

emis

2x2 double

	1	2	3
1	0.4500	0.5500	
2	0.2000	0.8000	
3			

Transition Matrix:

Variables - trans

trans

2x2 double

	1	2	3
1	0.7368	0.2632	
2	0.6000	0.4000	
3			
4			