Homework Assignment 3
Harshita Agarwala

November 13, 2017

1 Problem 1

Let the function be denoted by g(θ)

$$g(\theta) = \theta^t(1-\theta)^h$$

Differentitating w.r.t to θ we get the first derivative

$$g'(\theta) = t\theta^{\text{t-1}}(1-\theta)^h - h\theta^t(1-\theta)^{\text{h-1}}$$

Differentitating w.r.t to θ again, we get the second derivate:

$$g''(\theta) = t(t-1)\theta^{\text{t-2}}(1-\theta)^h - 2th\theta^{\text{t-1}}(1-\theta)^{\text{h-1}} + h(h-1)\theta^t(1-\theta)^{\text{h-2}}$$

Now we let f(θ)=log g(θ) and find the 1st and 2nd derivatives

$$f(\theta) = log(g(\theta)) = \text{t log}(\theta) + \text{h log(1-}\theta)$$

$$f'(\theta) = \frac{t}{\theta} - \frac{h}{1-\theta}$$

$$\text{and } f''(\theta) = \frac{h}{(1-\theta)^2} - \frac{t}{\theta^2}$$

2 Problem 2

We first find critical points on f(θ) = log g(θ) by equating f'(θ) to 0

$$\Rightarrow \frac{t}{\theta} - \frac{h}{1-\theta} = 0$$

$$\Rightarrow \frac{t-(t+h)\theta}{\theta(1-\theta} = 0$$

$$\Rightarrow t - (t+h)\theta = 0$$

$$\Rightarrow \theta = \frac{t}{t+h} \tag{1}$$

We substitute the value of θ in f"(θ) with that in equation (1) and get

$$f''(\theta) = \frac{(t+h)^2}{h} - \frac{(t+h)^2}{t} = \frac{t-h}{th}(t+h)^2 < 0 \text{ if h > t}$$

Again replacing the value of θ in g'(θ) with that in equation (1) we get

$$g'(\theta) = t\left[\frac{t}{t+h}\right]^{\text{t-1}}\left[1-\frac{t}{t+h}\right]^h - h\left[\frac{t}{t+h}\right]^{\text{t}}\left[1-\frac{t}{t+h}\right]^{\text{h-1}}$$

$$g'(\theta) = \frac{t^t h^h}{(t+h)^{\text{t+h-1}}} - \frac{t^t h^h}{(t+h)^{\text{t+h-1}}} = 0 \tag{2}$$

Hence, the critical points are same.
Now we find if it is a maximum or not.

Now g"((θ) is given by:

$$g''(\theta) = t(t-1)\theta^{\text{t-2}}(1-\theta)^h - 2th\theta^{\text{t-1}}(1-\theta)^{\text{h-1}} + h(h-1)\theta^t(1-\theta)^{\text{h-2}}$$

$$g''\left(\frac{t}{t+h}\right) = t(t-1)\left(\frac{t}{t+h}\right)^{\text{t-2}}\left(1-\frac{t}{t+h}\right)^h - 2th\left(\frac{t}{t+h}\right)^{\text{t-1}}\left(1-\frac{t}{t+h}\right)^{\text{h-1}} + h(h-1)\left(\frac{t}{t+h}\right)^t\left(1-\frac{t}{t+h}\right)^{\text{h-2}}$$

On solving further we get,

$$g''\left(\frac{t}{t+h}\right) = -\frac{t^{\text{t-1}}h^{\text{h-1}}}{(t+h)^{\text{t+h-1}}} < 0 \text{ as t,h} \in \mathbb{N} \tag{3}$$

Therefore, log function retains the critical points of the main function

3 Problem 3

We know that θ_{MAP} is the maximum of the p(θ = x | D) and is given by':

$$\theta_{\text{MAP}} = \frac{N_T+a-1}{N+a+b-1} \text{ and } \theta_{\text{MLE}} = \frac{N_T}{N}$$

If $\theta_{\text{MAP}} = \theta_{\text{MLE}}$ then a=b=1
This means that the prior distribution is uniform i.e there exists a prior p(θ) such that the result holds. Also, we know that such a prior will always exist.
Hence, it is true that θ_{MLE} is a special case of θ_{MAP}

4 Problem 4

$$\text{Now } \theta_{\text{MLE}} = \frac{m}{m+l}$$

$$\text{and } E_{\text{PR}}[\theta|a,b] = \frac{a}{a+b}$$

The consequent posterior distribution is given by the Beta distribution Beta(x | a+m, b+l) and the mean of this is:

$$E_{\text{PS}}[X] = \frac{a+m}{a+b+m+l} = \frac{a}{a+b+m+l} + \frac{m}{a+b+m+l}$$

$$\text{Let } 0 \le \lambda \le 1 \text{ be } \frac{a+b}{a+b+m+l}$$

Then we get,

$$\lambda E_{\text{PR}} = \left[\frac{a+b}{a+b+m+l}\right]\frac{a}{a+b} = \frac{a}{a+b+m+l} \tag{4}$$

$$\text{and } (1-\lambda)\theta_{\text{MLE}} = \left[\frac{m+l}{a+b+m+l}\right]\frac{m}{m+l} = \frac{m}{a+b+m+l} \tag{5}$$

On adding equation (4) and (5) we have,

$$\lambda E_{\text{PR}} + (1-\lambda)\theta_{\text{MLE}} = E_{\text{PS}}$$

5 Problem 5

The Poisson Distribution of X is given by

$$P(X = k|\lambda) = \frac{\lambda^k e^{\text{-}\lambda}}{k!}$$

Now, for n i.i.d samples for X the probability is given by:

$$P(D|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{k_i} e^{\text{-}\lambda}}{k_i!}$$

$$\Rightarrow P(D|\lambda) = e^{\text{-n}\lambda} \prod_{i=1}^{n} \frac{\lambda^{k_i}}{k_i!} \tag{6}$$

Taking log of equation (6) and denoting it by f(λ) we have,

$$f(\lambda) = -n\lambda + \Big(\sum_{i=1}^{n} k_i\Big) \log\lambda + \log\Big(\sum_{i=1}^{n} k_i\Big)$$

Differentiating w.r.t to λ and equating to 0 we get

$$f'(\lambda) = -n + \frac{\sum_{i=1}^{n} k_i}{\lambda} = 0$$

$$\Rightarrow \lambda = \frac{\sum_{i=1}^{n} k_i}{n}$$

Therefore we have $\theta_{\text{MLE}} = \frac{\sum_{i=1}^{n} k_i}{n}$

Now we let the prior have a Gamma distribution with constants α and β

$$P(\lambda|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma\alpha} \lambda^{\alpha \text{ -1}} e^{-\lambda\beta}$$

Now the posterior distribution is given by:

$$P(\lambda|D) = \frac{P(D|\lambda).P(\lambda)}{P(D)}$$

On replacing values we get,

$$P(\lambda|D) = \frac{1}{P(D)}.\frac{e^{\text{-n}\lambda}\lambda^{\sum k_i}}{\prod k_i}.\frac{\beta^\alpha}{\Gamma\alpha}\lambda^{\alpha \text{ -1}} e^{-\lambda\beta}$$

$$\Rightarrow P(\lambda|D) = c\Big[e^{(\text{-n + }\beta)\lambda}.\lambda^{(\alpha \text{ -1 + } \sum k_i)}\Big] \text{ where c is a constant}$$

Again taking log of the above function we get,

$$(-n+\beta)\lambda + (\alpha - 1 + \sum k_i) \log\lambda$$

On finding the derivative and equating it to 0 we get,

$$\lambda = \frac{\alpha - 1 + \sum k_i}{(-n+\beta)} = \theta_{\text{MAP}}$$