MAD Assignment 4

Ask Jensen

19. december 2021

Indhold

1 Problem 1

1.1 (a)

Assessing the following:

$$\theta_n = \frac{n}{\sum_{i=1}^{n} x_i}$$

I'm working with a geometric distribution so I have the following PMF:

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

rewriting the expression using pi-notation (Product notation):

$$P_\theta(X_i = x_i) = P_\theta(x_i) = (1-\theta)^{x_i-1}\theta$$

$$P_\theta(x_1, ..., x_n; \theta) = \prod_{i=1}^{n}(1-\theta)^{x_i-1}\theta$$

Rewriting this new expression by taking the logarithm of the expression:

$$log(\prod_{i=1}^{n}(1-\theta)^{x_i-1}\theta) = \sum_{i=1}^{n}(log((1-\theta)^{x_i-1} + log(\theta)))$$
$$= l(\theta) = \sum_{i=1}^{n}(x_i-1)log(1-\theta) + log(\theta)$$

Differentiating the expression and setting it equal to zero and dividing the expression into three sums

$$\frac{\partial l}{\partial \theta} = \sum_{i=1}^{n}((x_i-1)(-1)\frac{1}{1-\theta}\frac{1}{\theta}) = 0$$

$$\implies \sum_{i=1}^{n}((\frac{-1}{1-\theta} + 1\frac{1}{1-\theta}) + \frac{1}{\theta}) = 0$$

$$\frac{-1}{1-\theta}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\frac{1}{1-\theta} + \sum_{i=1}^{n}\frac{1}{\theta} = 0$$

$$\implies \frac{-1}{1-\theta}\sum_{i=1}^{n} x_i + n\frac{1}{1-\theta} + n\frac{1}{\theta} = 0$$

multiplying with $(1-\theta)$ through the equation:

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

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

$$-\sum_{i=1}^{n} x_i + \frac{n}{\theta} = 0$$

$$\iff \frac{n}{\theta} = \sum_{i=1}^{n} x_i \Longrightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i}$$

Calculating the second derivative of the function, in order to check if the function is either convex or concave. Which tells whether it's a maximum or minimum I've found.

$$\frac{\partial^2 l}{\partial\theta\partial\theta} = \frac{-1}{1-\theta}\sum_{i=1}^{n} x_i + n\frac{1}{1-\theta} + n\frac{1}{\theta}$$

$$= -(-1)\frac{-1}{(1-\theta)^2}\sum_{i=1}^{n} x_i + n(-1)\frac{-1}{(1-\theta)^2} + n(-1)\frac{-1}{\theta^2}$$

$$= \frac{-1}{(1-\theta)^2}\sum_{i=1}^{n} x_i + n\frac{1}{(1-\theta)^2} - n\frac{1}{\theta^2}$$

$$= -\frac{\sum_{i=1}^{n} x_i}{(1-\theta)^2} + \frac{n}{(1-\theta)^2} - \frac{n}{\theta^2}$$

Thus since the function will always be negative, I can conclude that I'm dealing with a maximum

$$\sum_{i=1}^{n} x_i \geq n \quad \text{since} \quad x_i \exists Z_+$$

Which concludes the proof.

2 Problem 2

3 Problem 3

3.1 (a)

From comment 3.1 in the book[1] by Rogers and Girolami I can see that I'm dealing with a conjugate pair. I have a beta distribution and a binomial liklihood. This means that when calculating the result, using the Baysian method, I can discard the denominator. Furthermore since my prior is a beta distributed and I as mentioned, is working with a conjugate pair my posterior must be beta distributed aswell.

$$p(r|y_N) \propto \left[\binom{N}{y_N} r^{y_N}(1-r)^{N-y_N} \right] \times \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta-1} \right]$$

$$= \left[\binom{N}{y_N} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \right] \times \left[r^{y_N} r^{\alpha-1}(1-r)^{N-y_N}(1-r)^{\beta-1} \right]$$

$$\propto r^{y_N+\alpha-1}(1-r)^{N-y_N+\beta-1}$$

$$\propto r^{\delta-1}(1-r)^{\gamma-1}$$

The parameters for δ and γ of this particular distribution is:

$$\delta = y_N + \alpha$$
$$\gamma = N - y_N + \beta$$

With the posterior beta density with the following general form, where K is a constant.

$$p(r) = K r^{\delta-1}(1-r)^{\beta-1}$$

In my case, the posterior will look as the following:

$$p(r|y_N) = \frac{\Gamma(\delta+\gamma)}{\Gamma(\delta)\Gamma(\gamma)} r^{\delta-1}(1-r)^{\beta-1}$$

where $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ is now the constant.

[1] A first course in machine learning

3.2 (b)

In order to identify the prior i will make use of the beta density function

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}r^{\alpha-1}(1-r)^{\beta-1}$$

To obtain $2r$ I will need to find some values for α and β $2r$ can also be rewritten as $2*r^1$ so I will try to find the values for alpha and beta such that the fraction with the Gamma function will be equal to 2. Note that the "Gamma function"is the generalization of the factorial function.

$$\frac{\Gamma(2+1)}{\Gamma(2)\Gamma(1)}r^{2-1}(1-r)^{1-1}$$

$$\frac{\Gamma(3-1)!}{\Gamma(2-1)!\Gamma(1-1)!}r^{2-1}(1-r)^{1-1}$$

$$2r^1(1-r)^0$$
$$2r^1\cdot 1$$
$$=2r$$

Following the same steps as in (a) and plugging in the new values for α and β I have the following.

$$r^{y_N}r^{2-1}(1-r)^{N-y_N}(1-r)^{1-1}$$
$$r^{y_N+2-1}(1-r)^{N-y_N+1-1}$$

The parameters for δ and γ of this particular distribution is:

$$\delta = y_N + 2$$
$$\gamma = N - y_N + 1$$

3.3 (c)

I will not be showing all the calculations, since they are exactly the same steps as in (3b).

To get r^2 I set the alpha value to 3, since $r^3 - 1 = r^2$

Then I have to finde a beta value that makes it possible for me to end up with $3r^2$. As show in (3b) this is where the fraction with the Gamma function will be used. I will set $\beta = 1$

$$\frac{\Gamma(3+1)}{\Gamma(2)\Gamma(1)} r^{3-1}(1-r)^{1-1}$$

$$\frac{\Gamma(4-1)!}{\Gamma(3-1)!\Gamma(1-1)!} r^{3-1}(1-r)^{1-1}$$

$$3r^2(1-r)^0$$
$$3r^2 \cdot 1$$
$$=3r^2$$

Following the same steps as in (a) and plugging in the new values for α and β I have the following.

$$r^{y_N} r^{3-1}(1-r)^{N-y_N}(1-r)^{1-1}$$
$$r^{y_N+3-1}(1-r)^{N-y_N+1-1}$$

The parameters for δ and γ of this particular distribution is:

$$\delta = y_N + 3$$
$$\gamma = N - y_N + 1$$

4 Problem 4

4.1 (a)