A4

MASD 2021
Department of Computer Science
University of Copenhagen

Mikkel/pwz854 | partners = mfh144 & qfh987

20/12/2021

Contents

Exercise 1

a) Start by writing out the likelihood function and take the log(the natural logarithm) of it, so we have

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

Then we want to take the derivative in relation to θ since we are interested in theta

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

The + between the two fractions becomes a - due to the chain rule being applied to $log((1-\theta))$. Now we want to get an expression for θ so we equate this to 0 and manipulate it

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

Add the sum $\sum_{i=1}^{n}\frac{x_i-1}{1-\theta}$ on both sides so we get

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

Then multiply with θ on both sides

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

Now multiply by $1-\theta$ on both sides

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

$$\iff$$

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

$$\iff$$

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

$$\iff$$

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

We then notice that we have $-n\theta$ no both sides and therefor we add $n\theta$ on both sides

$$n = \theta \sum_{i=1}^{n} (x_i)$$

Now all we have left is to divide by $\sum_{i=1}^{n} (x_i)$ on both sides

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

Above we see that the given expression was indeed the likelihood estimator for θ and now we need to show that it is the maximum likelihood estimator. This can be shown by taking the second derivative and see if it is less than 0

$$\frac{\partial^2 l}{\partial\theta\partial\theta} = \left(\sum_{i=1}^{n} \left(\frac{1}{\theta} - \frac{x_i - 1}{1 - \theta} \right) \right)'$$

Using the quotient rule and $(f - g)' = f' - g'$ we get

$$\frac{1}{\theta}' = -\frac{1}{\theta^2}$$

$$\left(\frac{x_i - 1}{1 - \theta} \right)' = \frac{(1 - \theta) \cdot 0 - (x_i - 1) \cdot -1}{(1 - \theta)^2} = \frac{x_i - 1}{(1 - \theta)^2}$$

writing $f' - g'$ out we get

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

Knowing that $0 < \theta \leq 1$ we can conclude that the $-\frac{1}{\theta^2}$ expression can only result in negative values. Now for the $\frac{1 - x_i}{(1 - \theta)^2}$ expression we see that the denominator will always be a value between 0 and 1, since $0 < \theta \leq 1$. We also notice that we have to make the statement that $\theta \neq 1$ since this will result in the denominator being 0. Now knowing that the denominator will always be between 1 and 0, but never be 1 or 0 we can look at the numerator. We know that x_i can be all positive integers, so for a start lets look at the case where $x_i = 1$, then the numerator will become 0 and the $-\frac{1}{\theta}^2$ becomes the determining factor. Now for $x_i > 1$ the numerator will result in a negative number and therefore the $\frac{1 - x_i}{(1-\theta)^2}$ expression will become negative. We now know that both expressions in the summation are negative and can therefor conclude that the summation will just sum up negative numbers resulting in a new negative number, hence the MLE for the parameter θ given in the assignment is actually a maximum.

Exercise 2

a) The properties of a probability function states that integrating over the range of $f(x)$ is equal to 1. Therefor c can by found by double integrating over the range of $f_\theta(x.y)$

$$\int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} f_\theta(x,y)\ dxdy = \int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} c\ dxdy = 1$$

Start by integrating over x

$$\int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} c\ dxdy = \int_{y_{min}}^{y_{max}} cx_{max} - cx_{min}\ dy$$

Then integrate by y

$$\int_{y_{min}}^{y_{max}} cx_{max} - cx_{min}\ dy = [cx_{max}y - cx_{min}y]_{y_{min}}^{y_{max}}$$

$$= (cx_{max}y_{max} - cx_{min}y_{max}) - (cx_{max}y_{min} - cx_{min}y_{min})$$

$$= cx_{max}y_{max} - cx_{min}y_{max} - cx_{max}y_{min} + cx_{min}y_{min} = 1$$

now isolate c in this expression, starting by factorising it so we have

$$c(x_{max} - x_{min})(y_{max} - y_{min}) = 1$$

now then divide both sides by $(x_{max} - x_{min})(y_{max} - y_{min})$

$$c = \frac{1}{(x_{max} - x_{min})(y_{max} - y_{min})}$$

b) To compute the likelihood we start by computing c

$$c = \frac{1}{(4-(-1))(3-(-1))} = \frac{1}{20}$$

Now to find the likelihood we have to compute

$$\prod_{i=1}^{4} f_{\theta_1}(x,y) = \prod_{i=1}^{4} c = \prod_{i=1}^{4} \frac{1}{20} = \frac{1}{160.000} = 6.25 \cdot 10^{-6}$$

We want the likelihood of 4 stars and therefore n=4. Then we repeat this for θ_2

$$c = \frac{1}{63}$$

$$\prod_{i=1}^{4} f_{\theta_2}(x,y) = \prod_{i=1}^{4} \frac{1}{63} = \frac{1}{15.752.961} \approx 6.35 \cdot 10^{-8}$$

c) Knowing that the PDF is uniform we know that the likelihood of seeing a star is given by the expression for c or else it is 0. So rather than considering when the likelihood > 0, we will consider when it is not bigger than 0, which is only when the placement of the star is outside the borders of the window. By this we know that as long as the star(s) are within the window then likelihood > 0 is true. Therefor the MLE $\hat{\theta}^{ML} = (0, 2, 0, 2)$ since the is the smallest we can make the window while still being able to see the stars.

Exercise 3

compute the posterior distribution p(r | yN) for three different priors:

a) Bayes' Rule $p(r|y_N) = \dfrac{p(y_N|r)p(r)}{p(y_N)}$
Because the posterior is also a beta distribution, we can omit $p(y_N)$ from the above equation so

$$p(r|y_N) \propto p(y_N|r)p(r)$$

Now we have a proportional constant $c = \dfrac{1}{p(y_N)}$

Now if we multiply the prior from the assignment with the binomial likelihood we have

$$p(r|y_N) \propto \binom{N}{y_N} r^{y_N}(1-r)^{N-y_N} \cdot 1 \propto \binom{N}{y_N} r^{y_N}(1-r)^{N-y_N}$$

Now we notice that this proportionality is close to the general form of a beta density: $p(r) = Kr^{\delta-1}(1-r)^{\gamma-1}$. On the right hand side of $\binom{N}{y_N}$ we actually have something similar to the right hand side of K. If we let $\delta = y_N + 1$ and $\gamma = N - y_N + 1$ then we actually have something exactly similar on those right hand sides. Meaning that if we put $\binom{N}{y_N}$ into the proportional constant c so $c = \dfrac{\binom{N}{y_N}}{p(y_N)}$, then this must be equal to K. Therefore our posterior beta density $p(r|y_N) = \dfrac{\binom{N}{y_N}}{p(y_N)} r^{y_N}(1-r)^{N-y_N}$ where $\delta = y_N + 1$ and $\gamma = N - y_N + 1$.

b) Repeating the same method as in task a) we have

$$p(r|y_N) \propto p(y_N|r)p(r) \propto \binom{N}{y_N} r^{y_N}(1-r)^{N-y_N} \cdot 2r$$

The case where $p(r) = 0$ is disregarded since this will result in the whole expression being 0. Now manipulating the above expression a bit we have

$$\binom{N}{y_N} r^{y_N}(1-r)^{N-y_N} \cdot 2r = 2\binom{N}{y_N} r^{y_N} r(1-r)^{N-y_N} = 2\binom{N}{y_N} r^{y_N+1}(1-r)^{N-y_N}$$

Now just as in task a) we can determine $\delta = y_N + 2$ and $\gamma = N - y_N + 1$ and then the proportionality constant $c = \dfrac{2\binom{N}{y_N}}{p(y_N)}$, which again must be equal to K. So we have the posterior beta density $p(r|y_N) = \frac{2\binom{N}{y_N}}{p(y_N)} r^{y_N+1}(1-r)^{N-y_N}$

For the values of α and β look below. The values has been found using the expression $p(r) = K \cdot r^{\alpha-1}(1-x)^{\beta-1} = 2r$

$$\alpha = 2$$

$$\beta = 1$$

c) Repeating what we did in task b) we have

$$p(r|y_N) \propto p(y_N|r)p(r) \propto 3\binom{N}{y_N} r^{y_N+2}(1-r)^{N-y_N}$$

we determine that $\delta = y_N + 3$ and $\gamma = N - y_N + 1$ and $K = \frac{3\binom{N}{y_N}}{p(y_N)}$

so we have the posterior beta density $p(r|y_N) = \frac{3\binom{N}{y_N}}{p(y_N)} r^{y_N+2}(1-r)^{N-y_N}$

For the values of α and β look below. The values has been found using the expression $p(r) = K \cdot r^{\alpha-1}(1-x)^{\beta-1} = 3r^2$

$$\alpha = 3$$

$$\beta = 1$$

Exercise 4

a)

Appendix