Roman Shaikh

Student number: 18300989

CS7DS3-Applied Statistical Modelling

Short-Assignment-1

# Question 1:

## Given:

Let C -> Cohen brothers, E -> Eve

1. To show that 60% of the time I watch movies with Eve i.e. P (Eve) = 0.6

Assuming P (E) to be X which gives us

Therefore,

Therefor i.e. I watch movies with Eve 60% of the time.

1. Probability of Eve watching Coen Brother movie with you i.e. P (Eve | Coen brothers)?

By Bays theorem we know that,

Therefore, the probability of Eve watching Coen Brother movie with you is 30%

# Question 2:

## Given:

1. Let . be some random samples from exp(θ),

Then we can give the likelihood function by

=>



i.e. is the MLE for

1. We are given that

i.e.

to find the posterior distribution we have,

This equation again follows gamma distribution with parameters . Here we can note that the posterior parameters depend upon the prior once

1. Given:

; A is viewed 30 times

; B is viewed 20 times

for worst layout minimum engagement time is 45 mins i.e.

for optimistic layout maximum engagement time is more than 10 mins i.e.

We know that, and (because )…………………………………………………{1}

Now, Given and

Therefore, and

………………………………………………………………………………………………………{2}

From {1} and {2} we can say that,

Further,

Therefore,

Which gives

And there for

Finally, we get,

It is possible to specify a non-informative prior in this case non-informative is nothing but uniform distribution the prior given has gamma distribution. Thus, gamma distribution is converted to uniform distribution with . The information distribution will be of the form

1. Using Monte Carlo methods in the following R code we estimate the probability that the new layout B has increased user engagement compared to the existing Layout A i.e.
2. 95% interval for the difference in engagement rates

* 95% interval is [-0.03833525, 0.0614802]

1. The probability P ( = 0.5658, which is approximately 50 %. This suggests us that the we are unsure of the probability that is greater than .

Also, the density plot suggests that the difference in engagement rate is concentrated near zero, which supports our uncertainty about the probability.

So, we conclude that more information is required to support the proposed profess of Layout B being superior to Layout A.



1. Given:

y – time a user spends on a webpage

g - denotes the value (in euros) of this time to the company

Then is value of user engagement.

Using Monte Carlo methods in the following R code, we find the expected value of the company website when visiting layout A.

* The expected value comes out to be E2.532182
* The short comings of this projections can be
* Alternative approach.



# Question 3:

## Given:

where and U denotes a uniform distribution

1. since we know that u\* ranges from 0 to 1 can only be positive, the H-M algorithm will resemble an increasing graph.

Like

Optimal performing algorithm will look something like:

In this case of optimally performing algorithm, the kernel holds the following properties:

* Irreducible: cannot be reduced to separate smaller states
* Aperiodic: The system does not return to same state after.
* Recurrent: The expected number of steps for returning to the same state is finite.

1. By the property of Reversibility, we can deduce that if is taken from then can also be drawn from

* Therefor the condition will not be true in that case.
* Because the M-H algorithm can only be positive values, in this case and hence it won’t be able to trace back the path.
* Therefore cannot be concluded from even if we can draw from
* Visualizing the trace path of the from , it looks like below(somewhat)

1. A good alternative of proposal distribution for this would be

* This will help in controlling the ever-increasing value of the M-H algorithm to run optimally.
* Also, a normal distribution can also help with the problem.