Q2) **application of bayesian techniques to the taxicab problem :**

Answer to Part (a) :

$$p(\theta\,|\,D) = Par(\theta\,|\,N + K, \text{max}(m, b)) = Par(\theta\,|\,1 + 0, \text{max}(0,100)) = Par(\theta\,|\,1,100)$$

Here as we have only one data point namely the first taxi number we observe the value of N = 1 while K = 0 ; the posterior is trying to convey to us the following statement " the distribution that the bound theta is greater than or equal to 100)

Answer to Part (b) :

To find the mode as we have only a single data point, the mode will be equal to 100 as it is the most observed/ dense data point (by definition of the mode) .

To find the mean of the data we have to find the expectation value of θ given D so that will be :

$$E[\theta\,|\,D] = \int \theta\, P(\theta\,|\,D)\, d\theta = \int \theta\, Par(\theta\,|\,K', M)\, d\theta = \int \theta\, Par(\theta\,|\,1,100)\, d\theta = \int \theta\, \frac{M}{\theta^2}\, d\theta$$

The following integral doesn't converge of the domain $[M, \infty)$ hence the mean doesn't exist .

To find the median we use the following result : median$(\theta\,|\,D) = M(2)^{\frac{1}{K}}$; which gives us the median for the distribution to be 200.

Answer to Part (c) :

$$p(\theta\,|\,D) = Par(\theta\,|\,1 + 0,\ \text{max}(m,\ 0)) = Par(\theta\,|\,1,\ m) \text{ .... (1)}$$

For the D' we can use the posterior for the first part and use it as a prior. For the predictive distribution the value of K' = 1 the value of b' = m and the number of samples N' = 1 also the maximum m' = max{D'} = max{[x]} = x

$$p(D' = (x)\,|\,D, [N' + K', m]) = \frac{K'}{(N' + K')b^{'N'}}\ \mathbb{I}(\,x \le m) + \frac{K'b^{'K'}}{(N' + K')m^{'N'+K'}}\ \mathbb{I}(\,x > m) = \frac{1}{2m}\mathbb{I}(x \le m) + \frac{m}{2x^2}\mathbb{I}(x > m)$$

Answer to Part (d) :

The answer to this part can be found out using the result from the previous part by substituting the values for [x] (D'), all that we have to do is to substitute the value of x in the equation as follows :

$$p(x\,=\,100\,|\,D, [N' + K', m']) = \frac{1}{2m}\mathbb{I}(100 \le m) + \frac{m}{2(100)^2}\mathbb{I}(100 > m)$$

$$p(x\,=\,50\,|\,D, [N' + K', m']) = \frac{1}{2m}\mathbb{I}(50 \le m) + \frac{m}{2(50)^2}\mathbb{I}(50 > m)$$

$$p(x\,=\,150\,|\,D, [N' + K', m']) = \frac{1}{2m}\mathbb{I}(150 \le m) + \frac{m}{2(150)^2}\mathbb{I}(150 > m)$$

Response to Part (e) :

Here in the problem we are trying to use a continuous variable, but in reality the distribution that should be used is a discrete one as x can't take all real values. We have been using a non-informative prior which provides very little (absolutely none here !) context, so instead of that we could use an informative prior to use that prior knowledge to improve the model accuracy.