Abhishek Chatterjee
1004820615
chatt114

Introduction to Machine Learning and Data Mining Assignment 3

1. License plates

a. From lecture, we know the Bayes rule is defined as

$$\therefore P(L|X_{1:N}) = \frac{\prod_{i=1}^{N} P(X_i|L) \times P(L)}{P(X_{1:N})}$$
$$= \frac{\prod_{i=1}^{N} f(X_i; 1, L) \times f(L; 1, M)}{\sum_{j=1}^{M}\left(\prod_{k=1}^{N} P(X_k|L=j) \times f(L=j; 1, M)\right)} \quad using\ sum\ and\ product\ rules$$
$$= \frac{\prod_{i=1}^{N} f(X_i; 1, L) \times f(L; 1, M)}{\sum_{j=1}^{M}\left(\prod_{k=1}^{N} P(X_k; 1, L=j) \times f(L=j; 1, M)\right)}$$

The numerator works out to be the PDF of the given distribution while the denominator works out to be the CDF of the distribution.

b. Put another way, the question asks us to determine when $P(L|X_{1:N}) \neq 0$

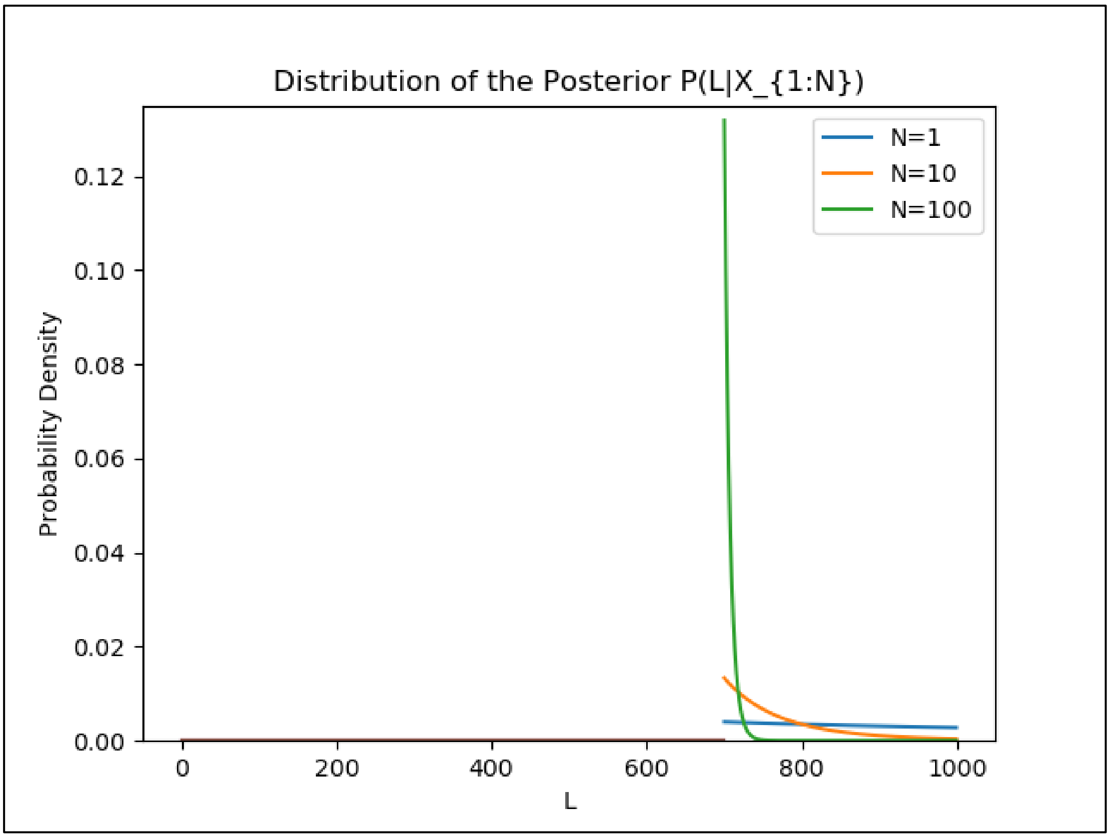
$$\Rightarrow \prod_{i=1}^{N} f(X_i; 1, L) \times f(L; 1, M) \neq 0$$
$$\Rightarrow \prod_{i=1}^{N} f(X_i; 1, L) \neq 0\ or\ f(L; 1, M) \neq 0$$
$$\Rightarrow 1 \leq X_i \leq L\ \forall X_i\ and\ 1 \leq L \leq M$$
$$\Rightarrow X_{max} \leq L \leq M$$

Thus, the posterior is nonzero under these conditions

c. We know that $X_{max} \leq L \leq M$. Therefore, simplifying our expression from part a, we get

$$\therefore P(L|X_{1:N}) = \frac{\left(\frac{1}{L}\right)^N \times \frac{1}{M}}{\sum_{j=X_{max}}^{M}\left(\left(\frac{1}{j}\right)^N \times \frac{1}{M}\right)}$$
$$= \frac{\left(\frac{1}{L}\right)^N \times \frac{1}{M}}{\frac{1}{M}\left(\sum_{j=X_{max}}^{M}\left(\left(\frac{1}{j}\right)^N\right)\right)}$$
$$= \frac{\left(\frac{1}{L}\right)^N}{\sum_{j=X_{max}}^{M}\left(\left(\frac{1}{j}\right)^N\right)}$$

Thus, under the assumption $X_{max} \leq L \leq M$, the posterior can be written as a function of L, M, N, X_{max} and observations d.



e. The maximum a posteriori (MAP) estimate is the value that maximises the posterior distribution. In this case, it would be X_{max}, as it is the positive value closest to 0. This holds true as we know that X_{max} is the closest to the actual value of L, which is the value that we observe

f. Plugging in the formula and values, we obtain the following results

<u>N = 1 case</u>

$$\therefore L_{mean} = \sum_{i=1}^{M} iP(L = i|X_1)$$

$$= \sum_{i=700}^{999} i \times \frac{\left(\frac{1}{i}\right)^1}{\sum_{j=700}^{999}\left(\left(\frac{1}{j}\right)^1\right)}$$

$$= 840.597$$

<u>N = 10 case</u>

$$\therefore L_{mean} = \sum_{i=700}^{999} i \times \frac{\left(\frac{1}{i}\right)^{10}}{\sum_{j=700}^{999}\left(\left(\frac{1}{j}\right)^{10}\right)}$$

$$= 772.775$$

N = 100 case

$$\therefore L_{mean} = \sum_{i=700}^{999} i \times \frac{\left(\frac{1}{i}\right)^{100}}{\sum_{j=700}^{999}\left(\left(\frac{1}{j}\right)^{1}\right)}$$

$$= 706.650$$

The general trend that we see is that the our estimate decreases as we sample more frequently. This is in line with our observations, as the more vehicles we see, the more likely it is that L approaches X_{max}

g. $\therefore P(X_{N+1}|X_{1:N}) = \sum_L P(X_{N+1}, L|X_{1:N})$ *by sum rule*

$$= \sum_L P(X_{N+1}|L, X_{1:N}) \times P(L|X_{1:N}) \quad \textit{by product rule}$$

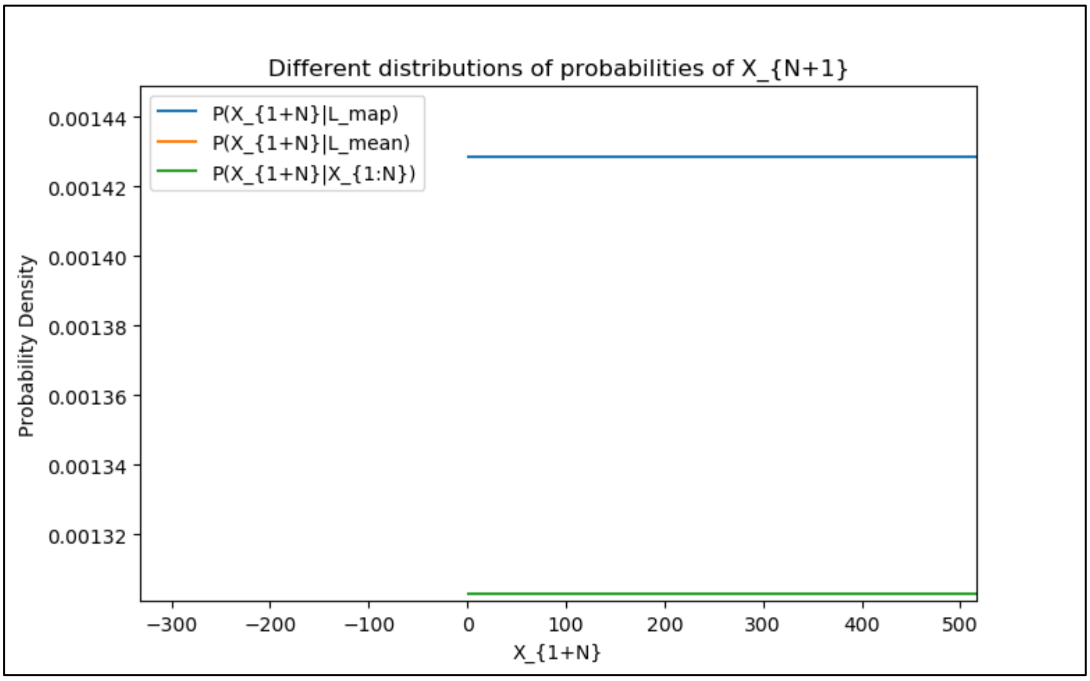
$$= \sum_L P(X_{N+1}|L) \times P(L|X_{1:N}) \quad \textit{by independence}$$

$$= \sum_{i=X_{max}}^{M} f(X_{N+1}; 1, l) \times P(L = l|X_{1:N})$$

$$= \sum_{i=X_{max}}^{M} \frac{1}{l} \times \frac{\left(\frac{1}{L}\right)^{N}}{\sum_{j=X_{max}}^{M}\left(\left(\frac{1}{j}\right)^{N}\right)}$$

Note: $f(X_{N+1}; 1, l) = likelihood\ of\ single\ license\ plate\ number$

h.



i. $P(X_{N+1}, L|X_{1:N})$ is most consistent with our intuition as it is the Bayes estimate, which takes all possible values of L into consideration. The probability of finding a car in the range is $P(X_{N+1}|X_{1:N}) \times (750 - 1) \approx 0.976$ which is almost guaranteed so don't take the bet