SC4000/CZ4041/CE4041: Machine Learning
Lecture 3 Tutorial Question Sets

Question 1: Consider the data set shown in Table 1.

Table 1: Data set for Question 1.

Record	A	B	C	Class
1	0	0	0	-
2	0	0	1	-
3	0	1	1	+
4	0	1	1	+
5	0	0	1	-
6	1	0	1	+
7	1	1	0	-
8	1	0	0	-
9	1	0	1	+
10	0	1	0	-

1. Estimate the conditional probabilities for $P(A = 1|+)$, $P(B = 1|+)$, $P(C = 1|+)$, $P(A = 1|-)$, $P(B = 1|-)$, and $P(C = 1|-)$.

2. Use the estimate of conditional probabilities given in the previous question to predict the class label for a test example $(A = 1, B = 1, C = 1)$ using the naïve Bayes approach.

Question 2: On Page 28 of the lecture notes "Lecture 3", recalculate the likelihoods using m-estimate. Compare the m-estimate method with the original method shown on Page 25 for estimating probabilities. Which method is better and why?