

## CS544 Module 2 Assignment (Self-Graded)

### Part1) Probability - 30 points

a) From the Bayes' rule example given in Section 3.10, compute the probabilities that a randomly selected non-smoker i) has lung disease and ii) does not have lung disease. Show the calculations without using R. Then, verify with the *bayes* function provided in the samples.

b) Suppose that in a particular state, among the registered voters, 40% are democrats, 50 % are republicans, and the rest are independents. Suppose that a ballot question is whether to impose sales tax on internet purchases or not. Suppose that 70% of democrats, 40% of republicans, and 20% of independents favor the sales tax. If a person is chosen at random that favors the sales tax, what is the probability that the person is i) a democrat? ii) a republican, iii) an independent. Show the solutions with the calculations without using R. Then, verify with the *bayes* function provided in the samples.

### Part2) Random Variables - 30 points

a) Consider the experiment of rolling a pair of dice. Using R, show how would you define a random variable for the absolute value of the difference of the two rolls, using a user-defined function.

b) Using the above result, what is the probability that the two rolls differ by exactly 2? What is the probability that the two rolls differ by at most 2? What is the probability that the two rolls differ by at least 3? Use the **Prob** function of R.

c) Show the marginal distribution of the above random variable (using R).

d) Using R, add another random variable to the above probability space using a user defined function. The random variable is TRUE if the sum of the two rolls is even, and FALSE otherwise. What is the probability that the sum of the two rolls is even? Show also the marginal distribution for this random variable.

### Part3) Functions - 40 points

Write your own R function, **all.primes(n)**, that returns a vector of all the prime numbers up to n (inclusive). Use the following rules in the function:

- i) Initialize the variable, **source**, to a vector sequence from 2 to n.
- ii) Initialize the variable, **result**, to the empty vector, or NULL.

In a loop, do the following as long as there are numbers remaining in the variable, **source**.

- iii) take the first element of **source**.
- iv) concatenate this element to **result**.
- v) Modify **source** by eliminating all numbers that are multiples of the element taken in step iii). This has to be done in a single statement without using any loop.

After the loop terminates, return the **result**.

Sample output:

```
> all.primes(15)
[1] 2 3 5 7 11 13
>
> all.primes(100)
[1] 2 3 5 7 11 13 17 19 23 29 31 37 41 43 47 53 59 61 67 71 73 79 83 89 97
```