Exercise 2: Probabilities

You can earn up to 10 points on this exercise.
You may work as a group of up to 3 people, but please submit your own version.
You may use any programming language you wish, but any submission that we cannot run on our computers without installing things must be presented to the class.

Please email your solution to `mittul.singh@lsv.uni-saarland.de` *or submit before the tutorial by* **10 am, November 11, 2015**.

The following tasks have been blatantly copied and inspired from Statistical Machine Translation book available here: `http://www.statmt.org/book/`

TASK 1

If we flip a coin 10 times, we might get the outcome *HTTHTHTHTT* (*H for heads, T for tails*).

a. Estimate a distribution by maximum likelihood estimation for this event. Estimating a distribution in this case would mean reporting probabilty of two disjoint events; first p(observing a heads) and second p(observing a tails). (3 points)

b. We want to test the quality of the estimation. We flip the coin five times and get *HHTTH*. What is the probability of this outcome according to

   i. the estimated distribution, and (1 point)

   ii. the uniform distribution or said another way assume that the coin is unbiased ? (1 point)

c. What is the entropy of a coin toss where the coin has a head on each side (fake coin) ? (2 points)

TASK 2

Computations with probability distributions:

a. Prove that $p(y|x) = p(y)$ if X and Y are independent. Do explain each of your steps. (3 points)

   OR

b. Derive the Bayes rule. Do explain each of your steps. (3 points)

BONUS

Given that the entropy of a random variable X is measured as:

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

where $0 \log_2 0 = 0$. Plot the entropy of an unfair coin showing heads ($X = H$) against probability of showing heads ($0 < p(X = H) < 1$) (3 points)