CS 3190: Foundations of Data Analysis **Due: Wednesday 9.16 at 1:10pm**

Homework 1: Probability and Bayes' Rule

Instructions: Your answers are due at 1:10, before the beginning of class on the due date. You must turn in a pdf through canvas. I recommend using latex (`http://www.cs.utah.edu/~jeffp/teaching/latex/`, see also `http://overleaf.com`) for producing the assignment answers. If the answers are too hard to read you will lose points, entire questions may be given a 0 (e.g. **sloppy pictures with your phone's camera are not ok, but very careful ones are**)

Please make sure your name appears at the top of the page.

You may discuss the concepts with your classmates, but write up the answers entirely on your own. **Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.**

---

1. [**20 points**] Using the probability table below for the random variables X and Y, derive the following values

   (a) $\textbf{Pr}(Y \neq 1)$

   (b) $\textbf{Pr}(X = 1 \cap Y = 0)$

   (c) $\textbf{Pr}(X = 1 \mid Y = 0)$

   (d) Are X and Y independent? and explain why.

	$X = 0$	$X = 1$
$Y = 0$	0.4	0.1
$Y = 1$	0.4	0.1

2. [**25 points**] Natasha is the Data University football team field goal kicker. There are 12 games in the season. In game i, for $i = 1, \ldots, 12$ the number of field attempts taken by Natasha is described by the random variable A_i with distribution

$A_i = 1$	$A_i = 2$
0.5	0.5

   Natasha is an excellent early season kicker, but she struggles at the end of the season when the pressure is on to make the playoffs. For $i = 1, \ldots, 12$ and $j = 1, ..., A_i$ let M_{ji} be 1 if Natasha makes her jth field goal attempt in game i, and 0 otherwise. For $i = 1, \ldots, 8$ and $j = 1, \ldots, A_i$ M_{ji} is 1 with probability 0.7 . For $i = 9, \ldots, 12$ and $j = 1, \ldots, A_i$ M_{ji} is 1 with probability 0.6. Assume that within a given game, the outcome of field goal attempts are independent events.

   Finally, let C_i represent the total number of field goals Natasha makes in game i, for $i = 1, \ldots, 12$

   (a) What is $E(C_i)$ in the early season, where $i \in \{1, 2, 3, 4, 5, 6, 7, 8\}$?

(b) Using (a), find the expected number of field goals Natasha makes in total from playing in games 1 through 8

(c) What is $E(C_i)$ at the end of the season, where $i \in \{9, 10, 11, 12\}$?

(d) Using (c), find the expected number of field goals Natasha makes in total from playing in games 9 through 12

(e) What is the expected number of field goals made by Natasha this season?

(f) In football, a field goal is worth 3 points. What is the expected number of points Natasha scores this season?

3. **[35 points]** Isidor has been gifted a coin from his grandfather. Let p be the probability that on any given flip of the coin, it lands with heads up. Before conducting any flips of the coin, Isidor has no clue whatsoever as to the value of p

   (a) What is a reasonable prior distribution Isidor can use to model his current cluelessness about p? *[Hint: Keep it simple and uninformative.]*

   (b) Isidor now conducts 10 procedurally identical flips of the coin. On each flip, the coin has probability p of landing heads up. And when it does, Isidor records a 1. Otherwise he records a 0. The result is $\{1, 0, 1, 1, 1, 1, 1, 0, 1, 1\}$. Let the ith flip be denoted X_i. Given this setup, a reasonable likelihood model for the flips is: $X_1, \ldots, X_{10}$ are independent (given p) and identically distributed under a *Bernoulli distribution with parameter p*, specifically where $\mathbf{Pr}(X_i = 1 \mid p) = p$ and $\mathbf{Pr}(X_i = 0 \mid p) = 1 - p$. Using this likelihood, and the prior in (a), what value of p gives the most likely model?

4. **[20 points]** Use python to plot the pdf and cdf of the Rayleigh distribution ($f(x) = x \exp(-x^2/2)$) for values of x in the range $[-2, 4]$. The function `scipy.stats.rayleigh` may be useful.