CSE 512: Homework 1 Due Wednesday September 11

1. **(2pts)** Alice, Bob, Carlos, Dima, and Estella are all elementary school kids, and the sit in the classroom in a circle. Alice is closest to the teacher, Mrs. Novid, and next to her is Bob, then Carlos, then Dima, then Estella.

   The students are somewhat well-spaced apart, but, you know kids. They like to cough on each other, touch each other's stuff, there's no privacy or social distancing. So disease transmission is indeed an issue.

   Transmission probabilities

   - If Mrs. Novid has Covid, then Alice has a 50% chance of getting it from Mrs. Novid.
   - If Alice has Covid, then Bob has 50% chance of getting Covid from Alice.
   - If Bob has Covid, then Carlos has 50% chance of getting Covid from Bob.
   - If Carlos has Covid, then Dima has 50% chance of getting Covid from Carlos.
   - If Dima has Covid, then Estella has 50% chance of getting Covid from Dima.

   However, even if no one else in the class has Covid, each child still has a 25% chance of getting Covid, from an outside source.

   (a) Draw a graphical model in which each node stands for a random variable, which represents the probability that a person has Covid. (e.g. node A may represent the event that Alice has Covid, node Bob that Bob has covid, etc...)

   (b) We find out one day that Mrs. Novid is home sick, and has had Covid for the past few days. What is the probability that Estella has Covid? Give your answer to the nearest 0.001

   Hint: The question is a little easier if you first try to find the probability that each person *doesn't* have Covid.

2. **(2 pts)** *Probability and statistics.* I have 4 children, Alexa, Siri, Googs, and Zuckie. Every morning I tell them to put on their socks.

   - Alexa only listens to me on Mondays and Thursdays and puts on her socks. The rest of the days, she puts on her socks only half of the time. She either puts on both her socks or none of her socks.
   - Siri always runs and gets her socks, but only puts one sock on.
   - Googs tells me all this random trivia about socks, but never puts on his socks.
   - Zuckie wears both his socks 4/7 of the time and sells the rest of them to CambridgeAnalytica.

   Assume the children all act independently. Round all answers to at least 3 significant digits.

   (a) **(0.5 pts)** What are the chances that either Alexa or Zuckie is wearing a sock?

   (b) **(0.5 pt)** On a random day, a girl is wearing a sock. What are the chances that it's Alexa?

   (c) **(0.5 pts)** What is the expected number of socks being worn by each child?

   (d) **(0.5 pts)** What is the variance in the number of socks being work by each child?

3. **(2 pts) Exponential distribution.** Wait time is often modeled as an exponential distribution, e.g.

$$\mathbf{Pr}(\text{I wait less than } x \text{ hours at the DMV}) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0, \end{cases}$$

and this cumulative density function is parametrized by some constant $\lambda > 0$. A random variable X distributed according to this CDF is denoted as $X \sim \exp[\lambda]$.

(a) **(0.5 pts)** In terms of λ, give the probability distribution function for the exponential distribution.

(b) **(0.5 pt)** Show that if $X \sim \exp(\lambda)$, then the mean of X is $1/\lambda$ and the variance is $1/\lambda^2$.
(You may use a symbolic integration tool such as Wolfram Alpha. If you do wish to do the integral by hand, my hint is to review integration by parts.)

(c) **(0.5 pt)** Now suppose I run a huge server farm, and I am monitoring the server's ability to respond to web requests. I have m observations of delay times, $x_1, ..., x_m$, which I assume are i.i.d., distributed according to $\exp[\lambda]$ for some λ. Given these m observations, what is the maximum likelihood estimate $\hat{\lambda}$ of λ?

(d) **(0.5 pt)** Given the estimate of $\hat{\lambda}$ in your previous question, is $1/\hat{\lambda}$ an unbiased estimate of the mean wait time? Is $1/\hat{\lambda}^2$ an unbiased estimate of the variance in wait time?

4. **(2 pts)** *Independent or not independent.* Variables A and B are random variables for two distributions. Decide if A and B are independent. Justify your answer.

(a) **(0.5 pts)** A and B are discrete random variables and have the following p.m.f.s

$$p_A(a) = \begin{cases} 0.25, & a = \text{red} \\ 0.25, & a = \text{blue} \\ 0.5, & a = \text{green} \end{cases}, \qquad p_B(b) = \begin{cases} 0.3, & b = \text{hat} \\ 0.3, & b = \text{T-shirt} \\ 0.2, & b = \text{skirt} \\ 0.2, & b = \text{shoes} \end{cases}$$

and $p_{A,B}(a, b)$ are defined by the table below

	a = red	a = blue	a = green
b = hat	0.075	0.075	0.15
b = T-shirt	0.075	0.075	0.15
b = skirt	0.05	0.05	0.1
b = shoes	0.05	0.05	0.1

(b) **(1 pt)** A and B are uniform distributions, where

$$f_A(a) = \begin{cases} 1 & -1 \leq a \leq 0 \\ 0 & \text{else,} \end{cases} \qquad f_B(b) = \begin{cases} 1 & 0 \leq b \leq 1 \\ 0 & \text{else,} \end{cases}, \qquad f_{A,B}(a, b) = \begin{cases} 4/3 & |a + b| \leq 1/2 \\ 0 & \text{else,} \end{cases}$$

(c) **(0.5 pts)** A and B are Gaussian distributions, with the following properties:

$$\mathbb{E}[A] = 0, \quad \mathbb{E}[B] = 1, \quad \mathbb{E}[A^2] = 1, \quad \mathbb{E}[(B-1)^2] = 1/2, \quad \mathbb{E}[A(B-1)] = -1.$$

Writing in terms of the usual Gaussian distribution form, if we form a random vector as $X = \begin{bmatrix} A \\ B \end{bmatrix}$, then

$$\mu = \begin{bmatrix} \mathbb{E}[A] \\ \mathbb{E}[B] \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \mathbb{E}[(A - \mathbb{E}[A])^2] & \mathbb{E}[(A - \mathbb{E}[A])(B - \mathbb{E}[B])] \\ \mathbb{E}[(A - \mathbb{E}[A])(B - \mathbb{E}[B])] & \mathbb{E}[(B - \mathbb{E}[B])^2] \end{bmatrix}$$

5. **Naive Bayes. (2pts)** Consider the following dataset 1

Feature A	Feature B	Decision (D)	Inference, Naive Bayes	Inference, full Bayes
1	1	1		
1	-1	-1		
-1	1	-1		
-1	-1	1		
1	1	1		
-1	-1	1		
1	-1	-1		
-1	-1	1		
-1	1	-1		
-1	-1	1		

(a) **(0.5 pts)** Using Naive Bayes, fill in the table below

$$\mathbf{Pr}(D=1|A=1,B=1) = \qquad \mathbf{Pr}(D=-1|A=1,B=1) =$$

$$\mathbf{Pr}(D=1|A=1,B=-1) = \qquad \mathbf{Pr}(D=-1|A=1,B=-1) =$$

$$\mathbf{Pr}(D=1|A=-1,B=1) = \qquad \mathbf{Pr}(D=-1|A=-1,B=1) =$$

$$\mathbf{Pr}(D=1|A=-1,B=-1) = \qquad \mathbf{Pr}(D=-1|A=-1,B=-1) =$$

(b) **(0.5 pts)** Using the probabilities you have computed above, compute the Naive Bayes inference of your training data, by filling in the 4th column. Give the training set error rate using this prediction.

(c) **(0.5 pts)** Using full Bayes, fill in the table below

$$\mathbf{Pr}(D=1|A=1,B=1) = \qquad \mathbf{Pr}(D=-1|A=1,B=1) =$$

$$\mathbf{Pr}(D=1|A=1,B=-1) = \qquad \mathbf{Pr}(D=-1|A=1,B=-1) =$$

$$\mathbf{Pr}(D=1|A=-1,B=1) = \qquad \mathbf{Pr}(D=-1|A=-1,B=1) =$$

$$\mathbf{Pr}(D=1|A=-1,B=-1) = \qquad \mathbf{Pr}(D=-1|A=-1,B=-1) =$$

(d) **(0.5 pts)** Using the probabilities you have computed above, compute the full Bayes inference of your training data, by filling in the 4th column. Give the training set error rate using this prediction.