CSE 312: Foundations of Computing II
Quiz Section #4: Conditional Probability, Random Variables, Naive Bayes

Review: Main Theorems and Concepts

Conditional Probability: $\mathbb{P}(A|B) =$ ______________________________

Independence: Events E and F are independent iff $\mathbb{P}(E \cap F) =$ ______________________________, or equivalently $\mathbb{P}(F) =$ ______________________________, or equivalently $\mathbb{P}(E) =$ ______________________________

Conditional Independence: Let E, F, G be events, and $\mathbb{P}(G) > 0$. E and F are conditionally independent given G iff $\mathbb{P}(E \cap F \mid G) =$ ______________________________, or equivalently $\mathbb{P}(F|E, G) =$ ______________________________, or equivalently $\mathbb{P}(E|F, G) =$ ______________________________

Bayes Theorem: $\mathbb{P}(A|B) =$ ______________________________

Partition: Nonempty events $E_1, \ldots, E_n$ partition the sample space Ω iff

- $E_1, \ldots, E_n$ are exhaustive: $E_1 \cup E_2 \cup \cdots \cup E_n = \bigcup_{i=1}^n E_i = \Omega$, and
- $E_1, \ldots, E_n$ are pairwise mutually exclusive: $\forall i \neq j, E_i \cap E_j = \emptyset$
  - Note that for any event A (with $A \neq \emptyset, A \neq \Omega$): A and A^C partition Ω

Law of Total Probability (LTP): Suppose $A_1, \ldots, A_n$ partition Ω and let B be any event. Then

$\mathbb{P}(B) =$ ______________________________

Bayes Theorem with LTP: Suppose $A_1, \ldots, A_n$ partition Ω and let A, B be any events with $\mathbb{P}(A), \mathbb{P}(B) > 0$.

Then $\mathbb{P}(A_1|B) =$ ____. In particular,

$\mathbb{P}(A|B) =$ ____

Chain Rule: Suppose $A_1, \ldots, A_n$ are events. Then

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_}$$

Random Variable (rv): A numeric function $X : \Omega \to \mathbb{R}$ of the outcome.

Range/Support: The support/range of a random variable X, denoted Ω_X, is the set of all possible values that X can take on.

Discrete Random Variable (drv): A random variable taking on a ____________ (either finite or countably infinite) number of possible values.

Probability Mass Function (pmf) for a discrete random variable X: a function $p_X : \Omega_X \to [0, 1]$ with $p_X(x) = \mathbb{P}(X = x)$ that maps possible values of a discrete random variable to the probability of that value happening, such that $\sum_x p_X(x) =$ ____________.

Exercises

1. Suppose there are three possible teachers for CSE 312: Martin Tompa, Anna Karlin, and Anup Rao. Suppose the ratio of grades $A : B : C : D : F$ for Martin's class is $1 : 2 : 3 : 4 : 5$, for Anna's class is $3 : 4 : 5 : 1 : 2$, and for Anup's class is $5 : 4 : 3 : 2 : 1$. Suppose you are assigned a grade randomly according to the given ratios when you take a class from one of these professors, irrespective of your performance. Furthermore, suppose Martin teaches your class with probability $\frac{1}{2}$ and Anna and Anup have an equal chance of teaching if Martin isn't. What is the probability you had Martin, given that you received an A? Compare this to the unconditional probability that you had Martin.

2. Suppose we have a coin with probability p of heads. Suppose we flip this coin n times independently. Let X be the number of heads that we observe. What is $\mathbb{P}(X = k)$, for $k = 0, \dots n$? Verify that $\sum_{k=0}^{n} \mathbb{P}(X = k) = 1$, as it should.

3. Suppose we have a coin with probability p of heads. Suppose we flip this coin until we flip a head for the first time. Let X be the number of times we flip the coin *up to and including* the first head. What is $\mathbb{P}(X = k)$, for $k \in \mathbb{N}$? Verify that $\sum_{k=1}^{\infty} \mathbb{P}(X = k) = 1$, as it should.

4. A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability p_1, to the left with probability p_2, and doesn't move with probability p_3, where $p_1 + p_2 + p_3 = 1$. After 2 seconds, let X be the location of the frog. Find $p_X(k)$, the probability mass function for X. Find $p_Y(k)$, the probability mass function for $Y = |X|$.

5. Corrupted by their power, the judges running the popular game show America's Next Top Mathematician have been taking bribes from many of the contestants. During each of two episodes, a given contestant is either allowed to stay on the show or is kicked off. If the contestant has been bribing the judges, she will be allowed to stay with probability 1. If the contestant has not been bribing the judges, she will be allowed to stay with probability 1/3, independent of what happens in earlier episodes. Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds.

   (a) If you pick a random contestant, what is the probability that she is allowed to stay during the first episode?

   (b) If you pick a random contestant, what is the probability that she is allowed to stay during both episodes?

   (c) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she gets kicked off during the second episode?

   (d) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she was bribing the judges?

6. Questions about the Naive Bayes Classifier:

   (a) Naive Bayes assumes conditional independence of words in an email, given that we know the label (ham or spam) of the email. Why is that assumption necessary to make Naive Bayes work?

   (b) Is the conditional independence assumption actually true in the real world? That is, are the occurrences of words in an email independent of each other, if we know the label of the email? Explain.

   (c) Do you expect the Naive Bayes Classifier to correctly classify all emails in a test set? Explain why or why not.

   (d) If you were a spammer and you knew we used Naive Bayes to filter spam, how would you change your emails to try to get past the filter?