CSCI 2244 – Homework 4

Out: Friday, September 20, 2019
Due: Friday, September 27, 2019, 11:59pm

This homework consists of **only written** exercises. You *must* type your solutions. See the "Assignments" section in the syllabus for advice about doing this. You should submit your homework via Canvas. In particular, you should upload a PDF file called:

`FirstName_LastName_Homework4.pdf`

Please use your full first name and last name, as they appear in official university records. The reason for doing so is that the TAs and I must match up these names with the entries in the gradebook.

1 Written Exercises

1.1 Inequalities

Task 1.1 (2 pts). A professor wants students to run a Monte Carlo simulation to estimate the expected value of a random variable X. The professor knows that X is a Bernoulli random variable and $\mathbb{E}[X] = .41$. She wants to have her students run enough trials to ensure that if a student's code is correct, the probability that their Monte Carlo estimate of $\mathbb{E}[X]$ will differ from .41 by .01 or more is less than 1/1000. According to Chebyshev's inequality, what is the fewest number of trials she can ask the students to run to guarantee this? (Hint: review the proof of the Weak Law of Large Numbers).

Task 1.2 (2 pts). In the textbook, Theorem 5.4.2 (Chernoff bound) shows that if $X_1, \ldots, X_n$ are independent Bernoulli random variables with expected values $p_1, \ldots, p_n$, and X is the sum of the X_i, then for $0 < \delta \leq 1$, we have:

$$P(X \geq p(1+\delta)) \leq e^{-\frac{p\delta^2}{3}}$$

and

$$P(X \leq p(1-\delta)) \leq e^{-\frac{p\delta^2}{3}}$$

(**Note**: the textbook states this in terms of strict inequalities, using $>$ and $<$, and I wrote it this way in class, but the form above is preferable. Also, as I said in class, there are many related inequalities that are all called "Chernoff bounds", so I don't expect you to memorize this.) Using the union bound, we can rearrange the above to get:

$$P(|X - p| \geq p\delta) \leq 2e^{-\frac{p\delta^2}{3}} \tag{1}$$

Concretely, suppose $n = 100$, and all of the $p_i = .5$. Then Equation 1 becomes:

$$P(|X - 50| \geq 50\delta) \leq 2e^{-\frac{50\delta^2}{3}} \tag{2}$$

We also know that it's possible to apply Chebyshev's bound to X to get an inequality of the form:

$$P(|X - 50| \geq k\sqrt{\mathsf{Var}[X]}) \leq \frac{1}{k^2} \tag{3}$$

Use both Chebyshev and Chernoff's inequality to bound $P(|X - 50| > 10)$, $P(|X - 50| > 20)$, $P(|X - 50| > 30)$, and $P(|X - 50| > 40)$. List the decimal values you get for the bounds from applying the two theorems to each of these probabilities.

1.2 Conditional Probabilities

Recall that $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. This is only well defined when $P(B) > 0$. That is, we have to condition on an event that occurs with non-zero probability. In the below, you can assume that this is always the case for all of the mentioned events.

Task 1.3 (1 pts). You roll two 6-sided dice. Let A be the event that the sum of the die values is ≥ 9, and let B be the event that the first die has value 4. What is $P(B \mid A)$?

Task 1.4 (1 pts). You have 10 coins. 9 are normal fair coins, but the last coin has heads on both sides. You pick one of the coins at random, each equally likely, and then flip that coin 3 times. Let A be the event that all 3 flips resulted in heads, and let B be the event that the coin you picked was the one with two heads. What is $P(B \mid A)$?

Task 1.5 (2 pts). You have two jars. Jar "A" has 2 blue marbles and 15 red marbles. Jar "B" has 12 blue marbles and 5 red marbles. Imagine you close your eyes, randomly stick your hand into one of the two jars (each equally likely), and pull out a marble. Let A be the event that the marble you pulled out came from Jar "A", B be the event that the marble came from Jar "B", and R be the event that the marble is red.

Compute $P(A \mid R)$ and $P(B \mid R)$.

Task 1.6 (2 pts). In class, the conditional probability $P(A \mid B)$ was defined as $\frac{P(A \cap B)}{P(B)}$. However, even though we called this a "probability", we never *proved* in class that that this number actually lies in the range $[0, 1]$. Prove it.

Task 1.7 (2 pts). Let A and B be two events such that $A \subseteq B$. Show that $P(B \mid A) \geq P(B)$.

Task 1.8 (2 pts). Show that if $P(A \mid B) = P(A \mid \overline{B})$ then A and B are independent.

Task 1.9 (2 pts). In class we discussed the Law of Total Probability for conditional probabilities. There is also a Law of Total Expectation for conditional expectations. It says that, given events

$B_1, \ldots, B_n$ such that $B_1 \cup \cdots \cup B_n = S$ and $B_i \cap B_j = \emptyset$ for $i \neq j$, then for any random variable X on the sample space S,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}\left[X \mid B_i\right] P(B_i)$$

Suppose you play the following gambling game: you flip a coin. If the coin is tails, you roll a fair 6-sided die and whatever the resulting value is, you win that many dollars. If the coin is heads, you roll *two* 6-sided dice and you win the sum of their values in dollars. Let X be the random variable giving your winnings. What is $\mathbb{E}[X]$? (Use the Law of Total Expectation.)