COS302/SML305- Princeton University-Spring 2022
Assignment #7

Due: March 28, 2022 at 11:59 pm

Upload at: https://www.gradescope.com/courses/355863/assignments

---

Remember to append your Colab PDF as explained in the first homework, with all outputs visible.
When you print to PDF it may be helpful to scale at 95% or so to get everything on the page.

Problem 1 (18 pts)

Consider the following cumulative distribution function for a random variable X that takes values in $\mathbb{R}$:

$$F(x) = P(X \leq x) = \frac{1}{1 + e^{-x}}$$

(A) What is the probability density function for this random variable?

(B) Find the inverse distribution (quantile) function $F^{-1}(u)$ that maps from $(0, 1)$ to $\mathbb{R}$.

(C) In a Colab notebook, implement inversion sampling and draw 1000 samples from this distribution. Make a histogram of your results.

Problem 2 (18 pts)

You enter a chess tournament where your probability of winning a game is 0.2 against half of the players (call them type 1), 0.6 against a quarter of the players (call them type 2), and 0.5 against the remaining quarter of the players (call them type 3).

(A) What is the probability of winning against a randomly chosen opponent?

(B) Suppose that you win. What is the probability that you had an opponent of type 1?

Problem 3 (20 pts)

You're playing a game at a carnival in which there are three cups face down and if you choose the one with a ball under it, you win a prize. This is sometimes called a *shell game*. At this carnival, there is a twist to the game: after you pick a cup, but before you're shown what is beneath it, the game operator reveals to you that one of the other cups (one of the two you did not choose) is empty. The operator now gives you the opportunity to switch your selection to the other unrevealed cup.

To clarify with an example: imagine there are cups A, B, and C. It is equally probable that the ball is beneath any of the three. You choose B. Before you see what is under B, the operator lifts A and shows you there is nothing under it. You are now presented with the option to keep your selection of B, or switch to the still-unrevealed cup C.

(A) Is it better, worse, or the same to switch to the other cup? Explain your reasoning in terms of probabilities.

(B) In a Colab notebook, simulate this game. Run 1,000 games with the *stay* strategy and 1,000 games with the *switch* strategy. Report the win rate of each strategy and explain which one empirically seems better.

Problem 4 (22 pts)

Consider two random variables X and Y. X is distributed according to the exponential distribution with parameter λ, which is often written as follows:

$$X \sim \text{Exponential}(\lambda),$$

where the notation $\sim$ means "drawn according to the distribution ...".

Y is a random variable such that $Y \mid X = x$ is a Poisson distribution which can be written as follows:

$$Y \mid X = x \sim \text{Poisson}(X = x).$$

(A) Write the joint probability density function $p(x, y)$ for these random variables.

(B) Compute the marginal distribution for Y. Here are the rough steps you should follow to do this, which are reflective of the kinds of computations one often does in probabilistic machine learning:

1. Write down the definite integral that you solve in order to get $p(y)$ from $p(x, y)$.
2. Pull terms that don't depend on x outside the integral.
3. Inside the integral, collect exponents.
4. See if the pieces inside the integral have a form that you recognize. If they look like an unnormalized probability density with known form, then you can sometimes solve the integral by inspection using that PDF's normalization constant. This kind of thing comes up a lot. (Hint: for this problem you want to have a look at the gamma distribution.)
5. Simplify. In this problem you'll probably want to take advantage of the relationship between the factorial and the gamma function.

Problem 5 (20 pts)

Recall that two events X and Y are independent if and only if $P(X, Y) = P(X)P(Y)$. Two events X and Y are **conditionally** independent given a third event Z if $P(X, Y|Z) = P(X|Z)P(Y|Z)$. The following problem explores the relationship between independence and conditional independence, specifically whether independence implies conditional independence or vice versa.

(A) Imagine that you have two coins: one regular fair coin (P(heads) = 0.5) and one fake two-headed coin (P(heads) = 1). Consider the following experiment: choose a coin at random and toss it twice. Define the following events:

- A is the event that the first coin toss results in a heads.
- B is the event that the second coin toss results in a heads.
- C is the event that the regular fair coin has been selected.

Are the events A and B independent? Give a qualitative answer, i.e., either yes because... or no because... Are they conditionally independent given C? Show your work, i.e., explicitly show the definition holds as given in the problem statement.

(B) Roll a single six-sided die and consider the following events:

- E is that you roll 1 or 2, i.e., $E = \{1, 2\}$.
- F is that you roll an even number, i.e., $F = \{2, 4, 6\}$.
- G is that you roll 1 or 4, i.e., $G = \{1, 4\}$.

Are the events E and F independent? Are they conditionally independent given G? For both of these questions show the definitions hold as in the problem statement.

Problem 6 (2pts)

Approximately how many hours did this assignment take you to complete?

My notebook URL: `https://colab.research.google.com/XXXXXXXXXXXXXXXXXXXXXXX`