Assignment #8

Due: 12:00 noon November 16, 2020

Upload at: https://www.gradescope.com/courses/135869/assignments/816677

Assignments in COS 302 should be done individually. See the course syllabus for the collaboration policy.

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 (20pts)

Consider two random variables X and Y with a joint probability density function $p(x, y)$. Show that

$$\mathbb{E}_X[x] = \mathbb{E}_Y[\mathbb{E}_X[x \mid Y = y]]$$

where the notation $\mathbb{E}_X[x \mid Y = y]$ denotes the expectation of X under the conditional distribution $P(X \mid Y = y)$.

Problem 2 (18pts)

A coin is weighted so that its probability of landing on heads is 0.2, suppose the coin is flipped 20 times.

(A) Compute the probability that the coin lands on heads at least 16 times.

(B) Use Markov's inequality to bound the event that the coin lands on heads at least 16 times.

(C) Use Chebyshev's inequality to bound the event that the coin lands on heads at least 16 times.

Problem 3 (20pts)

The covariance between two random variables X and Y can be computed as:

$$\mathrm{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

If X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ which implies that $\mathrm{cov}(X, Y) = 0$. However, the converse is not true: if the covariance between two random variables is zero, this *does not* imply that these variables are independent. Let's construct a counterexample to show this is the case in a Colab notebook. Be sure to append your PDF and insert your link as usual.

(A) Let X be a random variable that can take on only the values -1 and 1 and $P(X = -1) = 0.5 = P(X = 1)$. Generate 1,000,000 samples of X.

(B) Let Y be random variable such that $Y = 0$ if $X = -1$, and Y is randomly either -1 or 1 with probability 0.5 if $X = 1$. Construct 1,000,000 samples of Y based on the samples of X you generated in part (A).

(C) Use numpy to numerically compute the covariance between X and Y.

Problem 4 (40pts)

In 1777, Georges-Louis Leclerc, Comte de Buffon introduced a simple Monte Carlo method for estimation, now called the Buffon's needle problem. It's a fun way to estimate π via physical simulation. In this problem, we'll look at the following variant: you have a piece of paper on a tabletop and drawn on it is a square with sides of length w. Centered and entirely contained in the square is a circle with radius r. (So $2r \leq w$.) You drop n points (say, grains of sand) onto the piece of paper in a spatially uniform way. You count the grains and find that the ratio of grains inside the circle (n_c) to the total number grains inside the square is α:

$$\alpha = \frac{\text{\# of grains inside circle}}{\text{\# of grains inside square}} = \frac{n_c}{n}.$$

(A) Write an estimate for π in terms of w, r, n, and n_c.

(B) What kind of probability distribution will n_c follow? What is the standard deviation of n_c as a function of n and α?

(C) Assume that the radius of the circle is 1. What box width w minimizes the standard deviation of your estimate of π when you form this Monte Carlo estimate? (Hint: use the standard deviation worked out in (B), and reason about the relationship between w and α. Do you want w as small as possible, as big as possible, or somewhere in between?)

(D) Perform a real-life version of this experiment and estimate π with physical simulation. You could do it with paper and sand or something, as described above. Or you can be more creative! Some ideas:

- Grate a known amount of cheese onto a square pan that has a pizza in the middle. Weigh the cheese that doesn't fall onto the pizza.
- Put a circular cup into a box and drop coins or peas or something into the box and count the fraction that fall into the cup.
- Make a batch of cookies that are all perfect circles. Spread sprinkles over the entire baking pan and count (or weigh) how many don't make it onto cookies. You'll need to work out the compensation for having more than one cookie.
- Put a round bucket inside a square bucket and stick it out in the rain. Measure the difference in water collected.

You'll need to measure the dimensions of the square region and the radius of the circle. Draw a couple of dozen samples and report your estimate. Explain what you did and include a photograph of your setup! (Use the command `\includegraphics[width=3in]{myimage.jpg}` to include an image file within LaTeX)

Problem 5 (2pts)

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

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

Changelog

- 3 November 2020 – Updated for Fall, 2020.
- 7 April 2020 – Initial version.