

# ECE-210-A HW8

Instructor: Jonathan Lam

Spring 2022

In this problem set you'll be empirically simulating some probability experiments and comparing the results to the theoretical values.

Try your best to avoid loops when possible; vectorize and use logical indexing as much as you can to keep your code neat. There are many ways to do each problem – have fun with this!

1. Sample from the Poisson distribution with  $\lambda = 5$ . For each sample, sample again from the binomial distribution with  $p = 0.4$  and  $n$  equal to the Poisson sample. If the Poisson sample was equal to 0, don't sample from the binomial and just return 0 for that sample. In subplots, plot histograms of the p.m.f. from the result of this simulation and a different sampling from the Poisson distribution with  $\lambda' = \lambda p$ .

(If you have background in probability, think of the example where people enter a room at rate  $\lambda$  and each person is given a prize with probability 0.4. Then this distribution is the number of prize winners per second.)

2. You're going to simulate a problem that becomes relevant in cryptographic hashing algorithms (see the "birthday attack"). Simulate the following with a few values of  $n$ , and an appropriate sample size (what is a "sample" here?). Report your results clearly using plots or a table (see [help table](#)) as appropriate.

Given a classroom of  $n$  students with birthdays uniformly distributed throughout the year (and assume no leap years/days):

- (a) What is the probability that at least two students share a birthday? Empirically find the value of  $n$  for which the probability is roughly 0.5. (The "birthday problem.")
- (b) Given an arbitrary day of the year  $d$ , what is the probability that at least one student's birthday falls on that day?

Do these match the theoretical results? (You can derive or look up formulas to answer these questions.)

3. Go to [https://www.varsitytutors.com/gre\\_math-help/data-analysis/probability](https://www.varsitytutors.com/gre_math-help/data-analysis/probability). Simulate answers to at least 4 questions. Make sure to clearly label which questions you are answering, and include comments and/or plots as necessary to explain your thinking.