(*): Assigned to weekly problem set.

Poisson Distribution and Paradigm

1. For each of the following situations, determine whether it would be appropriate to use the Poisson Paradigm to approximate the variable N. If it is appropriate, briefly explain why the conditions for the paradigm are satisfied. If it isn't appropriate, explain what condition isn't satisfied, as well as why it might seem tempting to apply the Poisson Paradigm.
   (a) Suppose $A_1, A_2, \ldots$ is an infinite sequence of independent events, each with probability $p = 10^{-1000}$ of occurring. Let N be the number of events that occur.
   (b) A particular data file is stored as a sequence of 10^6 binary digits. When the data file is copied, each term in the sequence has probability $p = 10^{-4}$ of having an error, independent of other terms. Let N be the number of errors.
   (c) Suppose two copies of each of 52 cards are throughly shuffled. Cards are drawn from the deck two at a time. Let N be the number of cards that are paired with their other copy.
   (d) Let $X \sim \text{Binomial}(100, \frac{1}{100})$. For each n in $0, 1, 2, \ldots, 100$ let $\mathbf{1}_n$ be the indicator for the event "$X = n$". Let $N = \mathbf{1}_0 + \mathbf{1}_1 + \cdots + \mathbf{1}_{100}$.
   (e) Each of 20 students in a statistics classroom reveals the last 3 digits of their phone number. Let N be the number of phone numbers that are repeated at least once.

2. (*) A survey is being conducted in a city with a million 10^6 people. A sample of size 1000 is collected by choosing people in the city at random, with replacement and with equal probabilities for everyone in the city. Find a simple, accurate approximation to the probability that at least one person will get chosen more than once). *Hint: Indicator r.v.s are useful here, but creating 1 indicator for each of the million people is not recommended since it leads to a messy calculation. Feel free to use the fact that* $999 \approx 1000$.

3. The number of fish in a certain lake is a Poisson(λ) random variable. Worried that there might be no fish at all, a statistician adds one fish to the lake. Let Y be the resulting number of fish (so Y is 1 plus a Poisson(λ) random variable).
   (a) Find $\mathbb{E}[Y^2]$.
   (b) Find $\mathbb{E}[\frac{1}{Y}]$.

4. (*) If a random variable X has a discrete distribution with PMF $p(x)$, then the value of $x = x^* \in S_X$ that maximizes $p(x)$ is called the *mode* of the distribution. If this same maximum $f(x^*)$ is attained at more than one value of x, then all such values of x are called the modes of the distribution. If $X \sim \text{Poisson}(\lambda)$, find the mode or modes of this distribution. *Hint 1: consider the ratio* $\frac{p_X(k+1)}{p_X(k)}$, *and try to understand why this ratio is useful for solving this problem. Hint 2: consider the cases when* λ *is an integer versus when it is not.*