170E Week 4 Discussion Notes

Colin Ni

October 19, 2025

This week for discussion section we will do the following four problems which, in my opinion, would make for a reasonable (maybe somewhat difficult) midterm. As promised, all but one of the problems will be extremely similar to homework problems.

Problem 1. Find c so that

$$f_X(x) = \frac{c}{x(x+1)} \qquad \text{for } x = 5, 6, \ldots$$

defines a pdf for a random variable X. Determine $\mathbb{P}(X \geq 20 \mid X \geq 10)$.

Solution. The constant c is determined by the requirement

$$1 = \sum_{x=5}^{\infty} \frac{c}{x(x+1)} = c \sum_{x=5}^{\infty} \frac{(x+1) - x}{x(x+1)} = c \sum_{x=5}^{\infty} \left(\frac{1}{x} - \frac{1}{x+1} \right) = \frac{c}{5},$$

so $c = 5$. Moreover

$$\mathbb{P}(X \geq 20 \mid X \geq 10) = \frac{\mathbb{P}(X \geq 20 \text{ and } X \geq 10)}{\mathbb{P}(X \geq 10)} = \frac{\mathbb{P}(X \geq 20)}{\mathbb{P}(X \geq 10)} = \frac{1}{2}.$$

since

$$\mathbb{P}(X \geq k) = \sum_{x=k}^{\infty} \frac{5}{x(x+1)} = \frac{5}{k}$$

by the same argument as before. □

Problem 2. Let X be the number of flips of a fair coin until we see the same face on two consecutive flips. Find the pdf, mgf, mean, and variance of X. (You may freely use that $\sum_{k=1}^{\infty} k(\frac{1}{2})^{k-1} = 4$ and $\sum_{k=1}^{\infty} k^2(\frac{1}{2})^{k-1} = 12$.)

Solution. It takes exactly k flips if and only if the faces alternate for the first $k-1$ flips and the last two flips are the same, which happens with probability $(1/2)^{k-1}$. In other words, the pdf of X is

$$f_X(k) = \left(\frac{1}{2} \right)^{k-1} \qquad \text{for } k = 2, 3, \ldots.$$

The mgf is therefore

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{k=2}^{\infty} e^{tk} \left(\frac{1}{2}\right)^{k-1}.$$

To compute the mean and variance, we directly compute

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=2}^{\infty} k \left(\frac{1}{2}\right)^{k-1} \\
&= \sum_{k=1}^{\infty} k \left(\frac{1}{2}\right)^{k-1} - 1 \\
&= 3.
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{k=2}^{\infty} k^2 \left(\frac{1}{2}\right)^{k-1} \\
&= \sum_{k=1}^{\infty} k^2 \left(\frac{1}{2}\right)^{k-1} - 1 \\
&= 11,
\end{aligned}$$

so $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 2$. □

Problem 3. Suppose we select n random drivers from LA and define the random variables $X_1, \ldots, X_n$ as

$$X_i = \begin{cases} 1 & \text{the } i\text{th person is a good driver} \\ 0 & \text{otherwise.} \end{cases}$$

If exactly 10% of drivers in LA are good drivers, determine the distribution, mean, and variance of

$$S = X_1 + \cdots + X_n.$$

Solution. Each X_i is Bernoulli:

$$X_i \sim \mathrm{Bern}\left(\frac{1}{10}\right).$$

Since the X_i are mutually independent, their sum is binomial:

$$S \sim \mathrm{Binom}\left(n, \frac{1}{10}\right).$$

The mean and variance of S are therefore

$$\mathbb{E}(S) = \frac{n}{10} \qquad \text{and} \qquad \mathrm{Var}(X) = \frac{9n}{100}.$$ □

Problem 4. Suppose there are n types of coupons and that every cereal box contains a single coupon of a uniformly random type.

(i) Find the expected number of boxes that we need to open in order to collect a coupon of every type.

(ii) Given that we have opened k boxes, find the expected number of distinct coupons that we have collected.

Solution. For (i), supposing we already have r coupon types, let X_r be the number of boxes we open to find a new coupon type. We are looking for $\mathbb{E}(X_0 + X_1 + \cdots X_{n-1})$.

Note that if we already have r coupon types, the probability a new box has a new coupon type is $\frac{n-r}{n}$. Thus $X_r \sim \text{Geom}(\frac{n-r}{n})$, and

$$\begin{aligned}\mathbb{E}(X_0 + X_1 + \cdots + X_{n-1}) &= \mathbb{E}(X_0) + \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_{n-1}) \\ &= \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1} \\ &= n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right).\end{aligned}$$

For (ii), let $C_1, \ldots, C_n$ be the indicator random variables where C_i is 1 if one of our k boxes contains the ith coupon type and 0 otherwise. We are looking for $\mathbb{E}(C_1 + \cdots + C_n)$.

Since the probability that none of our k boxes contains a particular coupon type is $\left(\frac{n-1}{n}\right)^k$, we have that

$$C_i \sim \text{Bern}\left(1 - \left(\frac{n-1}{n}\right)^k\right).$$

Thus

$$\mathbb{E}(C_1 + \cdots + C_n) = \mathbb{E}(C_1) + \cdots + \mathbb{E}(C_n) = n\left(1 - \left(\frac{n-1}{n}\right)^k\right). \qquad \square$$