ENEE324, Home assignment 4. Date due October 25, 2026, 11:59pm EDT.

Instructor: Alexander Barg

Please upload your work as a **single PDF file** to ELMS (under the "Assignments" tab)

- Submissions on paper or by email will not be accepted.
- Please do not submit your solutions as multiple separate files (pictures of individual pages). Such submissions are difficult to grade and will not be accepted.
- Justification of solutions is required.
- Each problem is worth 10 points unless noted otherwise.

Problem 1. Let X be a discrete random variable taking values in $\{1,2,3,4,5\}$ with

$$P(X=k)=\frac{6-k}{15}, \qquad k=1,2,3,4,5.$$

(1) Verify that this defines a probability distribution.
(2) Compute $\mathbb{E}[X]$.
(3) Compute $\mathbb{E}[X^2]$ and $\mathrm{Var}(X)$.

Solution. *(1) The values* $6-k$ *for* $k=1,\dots,5$ *are* $5,4,3,2,1$*, summing to* 15*, so* $\sum_k P(X=k)=15/15=1$ *and each term is nonnegative; it is a valid pmf.*

(2) $\mathbb{E}[X]=\dfrac{1}{15}\displaystyle\sum_{k=1}^{5}k(6-k)=\frac{5+8+9+8+5}{15}=\frac{35}{15}=\frac{7}{3}\approx 2.333.$

(3) $\mathbb{E}[X^2]=\dfrac{1}{15}\displaystyle\sum_{k=1}^{5}k^2(6-k)=\frac{5+16+27+32+25}{15}=\frac{105}{15}=7$*. Hence* $\mathrm{Var}(X)=7-\left(\frac{7}{3}\right)^2=7-\frac{49}{9}=\frac{14}{9}\approx 1.556.$

Problem 2. Let $X\sim \mathrm{FS}(p)$ and $Y\sim \mathrm{FS}(p)$ be independent First Success ($=$ geometric $+1$) random variables with pms's

$$P(X=k)=P(Y=k)=p(1-p)^{k-1}, \qquad k=1,2,\dots.$$

Define $U=\min(X,Y)$ and $W=\max(X,Y)$.

(1) Compute $\mathbb{E}[X]$.
(2) Compute $P(X=Y)$.
(3) Using symmetry, compute $P(X>Y)$.
(4) Show that $U=\min(X,Y)$ is again a First Success random variable and find its parameter.
(5) Are $U=\min(X,Y)$ and $W=\max(X,Y)$ independent? Justify your answer.

Solution. *(1)* $\mathbb{E}[X]=\sum_{k\ge 1}k\,p(1-p)^{k-1}=\dfrac{1}{p}$ *(the mean of a First Success variable).*

(2) $P(X=Y)=\displaystyle\sum_{k\ge 1}\left(p(1-p)^{k-1}\right)^2=p^2\sum_{k\ge 1}(1-p)^{2(k-1)}=\frac{p^2}{1-(1-p)^2}=\frac{p^2}{2p-p^2}=\frac{p}{2-p}.$

(3) By symmetry $P(X > Y) = P(Y > X)$, and these two events together with $\{X = Y\}$ partition the sample space, so

$$P(X > Y) = \frac{1 - P(X = Y)}{2} = \frac{1}{2}\left(1 - \frac{p}{2-p}\right) = \frac{1-p}{2-p}.$$

(4) For any $k \geq 1$, $P(X \geq k) = (1-p)^{k-1}$, and by independence

$$P(U \geq k) = P(X \geq k)P(Y \geq k) = (1-p)^{2(k-1)} = \big((1-p)^2\big)^{k-1}.$$

This is the tail of a First Success variable with parameter q satisfying $1-q = (1-p)^2$, i.e. $q = 1-(1-p)^2 = p(2-p)$. Thus $U \sim \mathrm{FS}\big(p(2-p)\big)$.

(5) No. For instance $P(U = 2, W = 1) = 0$ because $U \leq W$ always, whereas $P(U = 2) > 0$ and $P(W = 1) = P(X = Y = 1) = p^2 > 0$, so $P(U = 2, W = 1) \neq P(U = 2)P(W = 1)$. The constraint $U \leq W$ makes them dependent.

Problem 3. At a coat check, n guests leave their coats and later the coats are returned in a uniformly random order (a uniformly random permutation of the guests). Let X be the number of guests who receive their own coat.

(1) For $i = 1, \ldots, n$ let I_i be the indicator of the event that guest i receives their own coat. Express X in terms of $I_1, \ldots, I_n$ and compute $\mathbb{E}[I_i]$.
(2) Use linearity of expectation to compute $\mathbb{E}[X]$. Does the answer depend on n?
(3) Compute $\mathbb{E}[X(X-1)]$ by considering $\mathbb{E}[I_i I_j]$ for $i \neq j$, and deduce $\mathrm{Var}(X)$.

Solution. *(1) $X = \sum_{i=1}^{n} I_i$. Guest i receives their own coat with probability $\mathbb{E}[I_i] = P(I_i = 1) = \dfrac{1}{n}$ (each coat is equally likely to go to guest i).*

(2) By linearity, $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[I_i] = n \cdot \frac{1}{n} = 1$. The expected number of matches is 1 for every $n \geq 1$.

(3) For $i \neq j$, $\mathbb{E}[I_i I_j] = P(I_i = 1, I_j = 1) = \dfrac{1}{n} \cdot \dfrac{1}{n-1}$. There are $n(n-1)$ ordered pairs, so

$$\mathbb{E}[X(X-1)] = \sum_{i \neq j} \mathbb{E}[I_i I_j] = n(n-1) \cdot \frac{1}{n(n-1)} = 1.$$

Then $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = 1 + 1 = 2$, so for $n \geq 2$,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 2 - 1 = 1.$$

Problem 4. Let X be a discrete random variable with

$$P(X = k) = \frac{c}{k(k+1)}, \qquad k = 1, 2, 3, \ldots.$$

(1) Find the constant c so that this is a probability distribution.
(2) Determine whether $\mathbb{E}[X]$ is finite.
(3) Let $Y = \dfrac{1}{X}$. Compute $\mathbb{E}[Y]$ in closed form.
(4) Let $Z = \mathbf{1}_{\{X \text{ is even}\}}$. Compute $\mathbb{E}[Z]$ in closed form.

Solution. *(1) Since $\dfrac{1}{k(k+1)} = \dfrac{1}{k} - \dfrac{1}{k+1}$, the sum telescopes: $\sum_{k \geq 1} \dfrac{1}{k(k+1)} = 1$. Hence $c = 1$.*

(2) $\mathbb{E}[X] = \displaystyle\sum_{k \geq 1} \frac{k}{k(k+1)} = \sum_{k \geq 1} \frac{1}{k+1} = \infty$*; the expectation is* not *finite (harmonic series).*

(3) $\mathbb{E}[1/X] = \sum_{k\geq 1} \frac{1}{k} \cdot \frac{1}{k(k+1)} = \sum_{k\geq 1} \frac{1}{k^2(k+1)}$. *Using* $\frac{1}{k^2(k+1)} = \frac{1}{k^2} - \frac{1}{k(k+1)}$,

$$\mathbb{E}[1/X] = \sum_{k\geq 1} \frac{1}{k^2} - \sum_{k\geq 1} \frac{1}{k(k+1)} = \frac{\pi^2}{6} - 1 \approx 0.6449.$$

(4) $P(X \text{ even}) = \sum_{j\geq 1} \frac{1}{(2j)(2j+1)} = \sum_{j\geq 1} \left(\frac{1}{2j} - \frac{1}{2j+1}\right) = 1 - \ln 2 \approx 0.3069$, *using the power series for* $\ln(x)$, *namely,* $\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \cdots$. *Hence* $\mathbb{E}[Z] = 1 - \ln 2$.

Problem 5. n balls are thrown independently and uniformly at random into m bins. Let X be the number of bins that contain *exactly one* ball.

(1) For $i = 1, \ldots, m$, define an indicator variable I_i for the event that bin i contains exactly one ball, and express X in terms of the I_i.
(2) Compute $\mathbb{E}[I_i]$.
(3) Use linearity of expectation to compute $\mathbb{E}[X]$.

Solution. *(1)* $X = \sum_{i=1}^{m} I_i$, *where* $I_i = 1$ *iff exactly one of the* n *balls lands in bin* i.

(2) The number of balls in bin i *is* $\text{Binomial}(n, 1/m)$, *so*

$$\mathbb{E}[I_i] = P(\text{exactly one ball in bin } i) = \binom{n}{1} \frac{1}{m} \left(1 - \frac{1}{m}\right)^{n-1} = \frac{n}{m} \left(1 - \frac{1}{m}\right)^{n-1}.$$

(3) By linearity,

$$\mathbb{E}[X] = m \cdot \frac{n}{m} \left(1 - \frac{1}{m}\right)^{n-1} = n \left(1 - \frac{1}{m}\right)^{n-1}.$$

Problem 6. A fair die is rolled once. Let X be the number shown, and let $Y = X \bmod 3$ be the remainder of X upon division by 3 (so $Y \in \{0, 1, 2\}$).

(1) Find the distributions of X and of Y.
(2) Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.
(3) Are X and Y independent? Justify your answer.

Solution. *(1)* X *is uniform on* $\{1, \ldots, 6\}$, *each value with probability* $1/6$. *For* Y: $Y = 0$ *when* $X \in \{3, 6\}$, $Y = 1$ *when* $X \in \{1, 4\}$, $Y = 2$ *when* $X \in \{2, 5\}$; *each of these has probability* $2/6 = 1/3$, *so* Y *is uniform on* $\{0, 1, 2\}$.

(2) $\mathbb{E}[X] = \frac{1+2+\cdots+6}{6} = \frac{21}{6} = 3.5$; $\mathbb{E}[Y] = \frac{0+1+2}{3} = 1$.

(3) No. Y *is a deterministic function of* X, *so they cannot be independent (unless* Y *were constant). Concretely,* $P(X = 1, Y = 0) = 0$ *while* $P(X = 1)P(Y = 0) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \neq 0$.

Problem 7. (The coupon collector)

Each cereal box contains one of $m = 10$ equally likely coupon types, independently across boxes. Let T be the number of boxes you must buy until you have collected all m types.

(1) Suppose you have already collected exactly $i - 1$ distinct types. Let G_i be the number of additional boxes needed to obtain a type you do not yet have. Explain why G_i is a First Success random variable, and give its parameter p_i. Write T in terms of $G_1, \ldots, G_m$.
(2) Using linearity of expectation and $\mathbb{E}[G_i] = 1/p_i$, show that

$$\mathbb{E}[T] = m \sum_{i=1}^{m} \frac{1}{i} = m\, H_m,$$

and evaluate it for $m = 10$.

(3) Simulate the coupon-collector process by computer to confirm your answer for $m = 10$. What is the smallest number of trials you needed to match the exact value of $\mathbb{E}[T]$ to within 10^{-1}? To within 10^{-2}? You may use any software package; include your code in your write-up.

Solution. *(1) Once $i-1$ types are held, each new box shows an unseen type with probability $p_i = \dfrac{m-(i-1)}{m}$, independently of previous boxes, so the wait G_i for the next new type is a First Success variable, $G_i \sim \mathrm{FS}(p_i)$. Collecting all types is done stage by stage, so $T = G_1 + G_2 + \cdots + G_m$.*

(2) By linearity, $\mathbb{E}[T] = \sum_{i=1}^{m} \mathbb{E}[G_i] = \sum_{i=1}^{m} \dfrac{1}{p_i} = \sum_{i=1}^{m} \dfrac{m}{m-i+1} = m \sum_{j=1}^{m} \dfrac{1}{j} = m\,H_m$. For $m = 10$, $H_{10} = \dfrac{7381}{2520} \approx 2.9290$, so $\mathbb{E}[T] = \dfrac{7381}{252} \approx 29.29$.

(3) A Monte Carlo simulation (repeatedly drawing coupons until all 10 types appear, then averaging T) yields a sample mean converging to ≈ 29.29. Since the estimate is an average of independent trials, its typical error decreases like $N^{-1/2}$; matching the exact value to within 10^{-1} takes on the order of 10^4 trials, and to within 10^{-2} roughly two orders of magnitude more (on the order of 10^6).