CS 70 Discrete Mathematics and Probability Theory
Spring 2025 Rao

HW 11

Due: Saturday, 4/12, 4:00 PM
Grace period until Saturday, 4/12, 6:00 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

Zachary Brandt
zbrandt@berkeley.edu

1 Combined Head Count

Note 19

Suppose you flip a fair coin twice.

(a) What is the sample space Ω generated from these flips?

(b) Define a random variable X to be the number of heads. What is the distribution of X?

(c) Define a random variable Y to be 1 if you get a heads followed by a tails and 0 otherwise. What is the distribution of Y?

(d) Compute the conditional probabilities $\mathbb{P}[Y = i \mid X = j]$ for all i, j.

(e) Define a third random variable $Z = X + Y$. Use the conditional probabilities you computed in part (d) to find the distribution of Z.

Solution:

(a) The sample space Ω generated from two flips is $\{HH, HT, TH, TT\}$.

(b) The distribution of the random variable X is the collection of values $\{(a, \mathbb{P}[X = a]) : a \in \mathscr{A}\}$, which is displayed below:

outcomes ω	value of X (# heads)	probability of occurring
TT	0	$\frac{1}{4}$
HT, TH	1	$\frac{2}{4}$
HH	2	$\frac{1}{4}$

(c) The distribution of the random variable Y is displayed below:

outcomes ω	value of Y (heads, tails)	probability of occurring
TT, TH, HH	0	$\frac{3}{4}$
HT	1	$\frac{1}{4}$

(d) The conditional probabilities are displayed below:

$$\begin{aligned}
Y \text{ can only be 1 when HT present} \quad & \mathbb{P}[Y=0 \mid X=0] = 1, \\
\text{TH also present} \quad & \mathbb{P}[Y=0 \mid X=1] = 0.5, \\
Y \text{ can only be 1 when HT present} \quad & \mathbb{P}[Y=0 \mid X=2] = 1, \\
Y \text{ can only be 1 when HT present} \quad & \mathbb{P}[Y=1 \mid X=0] = 0, \\
\text{TH also present} \quad & \mathbb{P}[Y=1 \mid X=1] = 0.5, \\
Y \text{ can only be 1 when HT present} \quad & \mathbb{P}[Y=1 \mid X=2] = 0
\end{aligned}$$

(e) The distribution of the random variable Z is displayed below:

outcomes ω	value of Z (X+Y)	probability of occurring
TT	0	$\frac{1}{4}$
TH	1	$\frac{1}{4}$
HH, HT	2	$\frac{2}{4}$

2 Testing Model Planes

Note 15

Amin is testing model airplanes. He starts with n model planes which each independently have probability p of flying successfully each time they are flown, where $0 < p < 1$. Each day, he flies every single plane and keeps the ones that fly successfully (i.e. don't crash), throwing away all other models. He repeats this process for many days, where each "day" consists of Amin flying all remaining model planes and throwing away any that crash. Let X_i be the random variable representing how many model planes remain after i days. Note that $X_0 = n$. Justify your answers for each part.

(a) What is the distribution of X_1? That is, what is $\mathbb{P}[X_1 = k]$?

(b) What is the distribution of X_2? That is, what is $\mathbb{P}[X_2 = k]$? Recognize the distribution of X_2 as one of the famous ones and provide its name and parameters.

(c) Repeat the previous part for X_t for arbitrary $t \geq 1$.

(d) What is the probability that at least one model plane still remains (has not crashed yet) after t days? Do not have any summations in your answer.

(e) Considering only the first day of flights, is the event A_1 that the first and second model planes crash independent from the event B_1 that the second and third model planes crash? Recall that two events A and B are independent if $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$. Prove your answer using this definition.

(f) Considering only the first day of flights, let A_2 be the event that the first model plane crashes *and* exactly two model planes crash in total. Let B_2 be the event that the second plane crashes on the first day. What must n be equal to in terms of p such that A_2 is independent from B_2? Prove your answer using the definition of independence stated in the previous part.

(g) Are the random variables X_i and X_j, where $i < j$, independent? Recall that two random variables X and Y are independent if $\mathbb{P}[X = k_1 \cap Y = k_2] = \mathbb{P}[X = k_1]\,\mathbb{P}[Y = k_2]$ for all k_1 and k_2. Prove your answer using this definition.

Solution:

(a) The probability that the number of remaining planes on day 1 is some k, is $\mathbb{P}[X_1 = k] = \binom{n}{k} p^k (1-p)^{n-k}$.

(b) There need to be at least k planes left over from day 1 for there to be k planes that remain after day 2. So, from the Total Probability Law, the probability that there are 2 planes remaining on day 2 is the probability that 2 remain and k remained from day 1, or $k+1$ remained, and so until all n planes remained from day 1.

$$\begin{aligned}
\mathbb{P}[X_2 = k] &= \mathbb{P}[X_2 = k, X_1 = k] + \mathbb{P}[X_2 = k, X_1 = k+1] + \cdots + \mathbb{P}[X_2 = k, X_1 = n] \\
&= \sum_{i=k}^{n} \mathbb{P}[X_2 = k \mid X_1 = i]\,\mathbb{P}[X_1 = i] \\
&= \sum_{i=k}^{n} \mathbb{P}[X_2 = k \mid X_1 = i] \cdot \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=k}^{n} \binom{i}{k} p^k (1-p)^{i-k} \cdot \binom{n}{i} p^i (1-p)^{n-i} \\
&= p^k (1-p)^{n-k} \sum_{i=k}^{n} \binom{i}{k} \binom{n}{i} p^i \\
&= p^k (1-p)^{n-k} \sum_{i=k}^{n} \binom{n}{k} \binom{n-k}{i-k} p^i \\
(j = i-k) \qquad &= \binom{n}{k} p^k (1-p)^{n-k} \sum_{j=0}^{n-k} \binom{n-k}{j} p^{j+k} \\
&= \binom{n}{k} p^{2k} (1-p)^{n-k} \sum_{j=0}^{n-k} \binom{n-k}{j} p^j \cdot 1^{n-k-j} \\
&= \binom{n}{k} p^{2k} (1-p)^{n-k} (1+p)^{n-k} \\
&= \binom{n}{k} p^{2k} ((1-p)(1+p))^{n-k} \\
&= \binom{n}{k} p^{2k} (1-p^2)^{n-k} \\
X_2 &\sim \text{Binomial}(n, p^2)
\end{aligned}$$

The above is a binomial distribution on parameters n and p^2, which makes sense considering that every plane flight outcome is independent of another and each plane has to fly successfully twice, i.e. p then p, to remain. $1-p^2$ represents all those planes who did not make it twice.

(c) Using this aformentioned logic, for an arbitrary $t \geq 1$, the distribution of X_t is then $X_t \sim \text{Binomial}(n, p^t)$.

(d) The set of outcomes where at least one model plane still remains is exactly the complement of the set of outcomes where no model planes remain after t days. Therefore, the probability is $1-(1-p^t)^n$.

(e) If the planes are numbered like they are here, the probability that the first two crash is $(1-p)^2$, and the probability that the second two crash is also $(1-p)^2$. For the joint event $A \cap B$, the first three planes crash, since the second plane doesn't crash twice, which happens with probability $(1-p)^3$. However, $\mathbb{P}[A] \cdot \mathbb{P}[B] = (1-p)^2 \cdot (1-p)^2 = (1-p)^4 \neq (1-p)^3$, and therefore the events are not independent.

(f) For events A_2 and B_2 to be independent, n nust be equal to $p-1$. This is because the probability $\mathbb{P}[A_2]$ is equal to $\binom{n}{1}(1-p)^2$, since the first plane must crash, and then there are n other options for the other plane. The probability $\mathbb{P}[B_2]$ is equal to $(1-p)$ since just the second plane has to crash. If $n = p-1$, then

$$\binom{p-1}{1} \cdot (1-p)^2 = (1-p)^3 = \mathbb{P}[A_2 \cap B_2]$$
$$(1-p) \cdot (1-p)^2 = (1-p)^3 = \mathbb{P}[A_2] \cdot Pr[B_2].$$

(g) False, consider the case where there are more planes on a following day than there were on a previous day, which is impossible considering the problem setup. E.g. $\mathbb{P}[X_1 = 1 \mid X_2 = 2] = 0$ because there cannot be more planes than there were on the initial day, however $\mathbb{P}[X_1 = 1]$ and $\mathbb{P}[X_2 = 2]$ are both not equal to zero. Therefore, $\mathbb{P}[X_1 = 1 \mid X_2 = 2] \neq \mathbb{P}[X_1 = 1] \cdot \mathbb{P}[X_2 = 2]$, and the two events are not independent.

3 Fishy Computations

Note 19

Assume for each part that the random variable can be modelled by a Poisson distribution.

(a) Suppose that on average, a fisherman catches 20 salmon per week. What is the probability that he will catch exactly 7 salmon this week?

(b) Suppose that on average, you go to Fisherman's Wharf twice a year. What is the probability that you will go at most once in 2024?

(c) Suppose that in March, on average, there are 5.7 boats that sail in Laguna Beach per day. What is the probability there will be *at least* 3 boats sailing throughout the *next two days* in Laguna?

(d) Denote $X \sim \text{Pois}(\lambda)$. Prove that

$$\mathbb{E}[Xf(X)] = \lambda\, \mathbb{E}[f(X+1)]$$

for any function f.

Solution:

(a) The probability that a fisher catches exactly 7 salmon this week is $\mathbb{P}[X=7] = \frac{20^7 \cdot e^{-20}}{7!}$.

(b) The probability that you will go to the Fisherman's Whart at most once in 2024 is $\mathbb{P}[X \le 1] = \mathbb{P}[X=0] + \mathbb{P}[X=1] = \frac{2^0 \cdot e^{-2}}{0!} + \frac{2^1 \cdot e^{-2}}{1!} = \frac{3}{e^2}$.

(c) The combined random variable for the number of boats sailing throughout the next two days is $X_1 + X_2 \sim \text{Pois}(\lambda+\lambda)$. Then, the probability that there are at least 3 boats observed is the complement of the probability that there are less than 3 boats observed, i.e.,

$$\begin{aligned}
\mathbb{P}[X_1+X_2 \ge 3] &= 1 - \mathbb{P}[X_1+X_2 < 3] \\
&= 1 - \mathbb{P}[X_1+X_2=0] - \mathbb{P}[X_1+X_2=1] - \mathbb{P}[X_1+X_2=2] \\
&= 1 - \frac{(11.4)^0}{0!}e^{11.4} - \frac{(11.4)^1}{1!}e^{11.4} - \frac{(11.4)^2}{2!}e^{11.4}
\end{aligned}$$

(d) The proof is as follows:

$$\begin{aligned}
E[Xf(X)] &= \sum_x x f(x)\, \mathbb{P}[X=x] \\
&= \sum_x x f(x)\, \mathbb{P}[X=x] \\
&= \sum_x x f(x) \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_x f(x) \frac{\lambda^x}{(x-1)!} e^{-\lambda} \\
&= \sum_{i=0}^{\infty} f(i+1) \frac{\lambda^{i+1}}{i!} e^{-\lambda} \\
&= \lambda \sum_{i=0}^{\infty} f(i+1) \frac{\lambda^i}{i!} e^{-\lambda} \\
&= \lambda \sum_x f(x+1) \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \lambda \sum_x f(x+1)\, \mathbb{P}[X=x] \\
&= \lambda E[f(X+1)]
\end{aligned}$$

4 Such High Expectations

Note 19

Suppose X and Y are independently drawn from a Geometric distribution with parameter p. For each of the below subparts, your answer must be simplified (i.e. NOT left in terms of a summation).

(a) Compute $\mathbb{E}[\min(X,Y)]$.

(b) Compute $\mathbb{E}[\max(X,Y)]$.

(c) Compute $\mathbb{P}[X+Y \geq t]$

Solution:

(a) I will define a new random variable Z as $Z = \min(X,Y)$. Then, the computation is as follows:

$$
\begin{aligned}
\mathbb{E}[Z] &= (0 \times \mathbb{P}[Z=0]) + (1 \times \mathbb{P}[Z=1]) + (2 \times \mathbb{P}[Z=2]) + \dots \\
&= \mathbb{P}[Z=1] + (\mathbb{P}[Z=2] + \mathbb{P}[Z=2]) + (\mathbb{P}[Z=3] + \mathbb{P}[Z=3] + \mathbb{P}[Z=3]) \dots \\
&= (\mathbb{P}[Z=1] + \mathbb{P}[Z=2] + \mathbb{P}[Z=3] + \dots) + (\mathbb{P}[Z=2] + \mathbb{P}[Z=3] + \dots) + \dots \\
&= \mathbb{P}[Z \geq 1] + \mathbb{P}[Z \geq 2] + \mathbb{P}[Z \geq 3] + \dots \\
&= \sum_{i=1}^{\infty} \mathbb{P}[Z \geq i] \\
X \text{ \& } Y \text{ need to be } > \quad &= \sum_{i=1}^{\infty} \mathbb{P}[X \geq i]\,\mathbb{P}[Y \geq i] \\
&= \sum_{i=1}^{\infty} (1-p)^{i-1} \cdot (1-p)^{i-1} \\
&= \sum_{i=1}^{\infty} (1-p)^{2i-2} \\
&= (1-p)^{-2} \sum_{i=1}^{\infty} (1-p)^{2i} \\
&= \frac{1}{1-(1-p)^2}
\end{aligned}
$$

(b) From before we can use the fact that $\mathbb{E}[Z] = \sum_{i=1}^{\infty} \mathbb{P}[Z \geq i]$ to see that

$$
\begin{aligned}
\mathbb{E}[\max(X,Y)] &= \sum_{i=1}^{\infty} \mathbb{P}[\max(X,Y) \geq i] \\
&= \sum_{i=1}^{\infty} 1 - \mathbb{P}[\max(X,Y) < i] \\
&= \sum_{i=1}^{\infty} 1 - \mathbb{P}[X < i] \cdot \mathbb{P}[Y < i] \\
&= \sum_{i=1}^{\infty} 1 - (1 - \mathbb{P}[X \geq i]) \cdot (1 - \mathbb{P}[Y \geq i]) \\
&= \sum_{i=1}^{\infty} 1 - (1 - (1-p)^{i-1}) \cdot (1 - (1-p)^{i-1}) \\
&= \sum_{i=1}^{\infty} 2(1-p)^{i-1} - (1-p)^{2i-2}
\end{aligned}
$$

(c) For any value x of X, the value of y of Y must then satisfy the inequality $y \geq t - x$. Therefore, the computation is

$$\begin{aligned}
\mathbb{E}[X+Y \geq t] &= \sum_t \mathbb{P}[X+Y \geq t] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[X=i]\,\mathbb{P}[Y \geq t-i] \\
&= \sum_{i=1}^{\infty} \mathbb{P}[X=i]\,\mathbb{P}[Y \geq t-i] \\
&= \sum_{i=1}^{t-1} \mathbb{P}[X=i]\,\mathbb{P}[Y \geq t-i] + \sum_{i=t}^{\infty} \mathbb{P}[X=i]\,\mathbb{P}[Y \geq t-i] \\
\text{Y has no bearing once $X \geq t$} \quad &= \sum_{i=1}^{t-1} \mathbb{P}[X=i]\,\mathbb{P}[Y \geq t-i] + \sum_{i=t}^{\infty} \mathbb{P}[X=i] \\
&= \sum_{i=1}^{t-1} (1-p)^{i-1} p \cdot (1-p)^{t-i-1} + \mathbb{P}[X \geq t] \\
&= (1-p)^{t-2} p \sum_{i=1}^{t-1} 1 + \mathbb{P}[X \geq t] \\
&= (1-p)^{t-2} p(t-1) + (1-p)^{t-1}
\end{aligned}$$

5 Swaps and Cycles

Note 15

A permutation of n objects is a bijection from $(1,\ldots,n)$ to itself. For example, the permutation $\pi = (2,1,4,3)$ of 4 objects is the mapping $\pi(1) = 2$, $\pi(2) = 1$, $\pi(3) = 4$, and $\pi(4) = 3$. We'll say that a permutation $\pi = (\pi(1),\ldots,\pi(n))$ contains a *swap* if there exist $i, j \in \{1,\ldots,n\}$ so that $\pi(i) = j$ and $\pi(j) = i$, where $i \neq j$. The example above contains two swaps: $(1,2)$ and $(3,4)$.

(a) In terms of n, what is the expected number of swaps in a random permutation?

(b) In the same spirit as above, we'll say that π contains a *k-cycle* if there exist $i_1,\ldots,i_k \in \{1,\ldots,n\}$ with $\pi(i_1) = i_2, \pi(i_2) = i_3, \ldots, \pi(i_k) = i_1$. Compute the expectation of the number of k-cycles.

Solution:

(a) For a set of n objects, there are $n!$ permutations, which make sup the sample space. When two elements are swapped, there are $n-2$ remaining elements which can be in any order, i.e., $(n-2)!$ permutations with any two elements swapped. There are n choose 2 potential pairs out of the n objects. Since each pair has the same probability of being swapped, $\frac{(n-2)!}{n!}$, the expected value is $\binom{n}{2} \times \frac{(n-2)!}{n!} = \frac{n!}{(n-2)! \cdot 2!} \times \frac{(n-2)!}{n!} = \frac{1}{2!} = \frac{1}{2}$.

(b) For any subset of k elements, there are $k!$ possible permutations, but each permutation has k equivalents, i.e. (1, 2, 3) is equivalent to (2, 3, 1), so there are $k! \div k$ cycles for a k, and there are n choose k different k-lenght subsets for a total of $\binom{n}{k} \times \frac{k!}{k}$ cycles. Then, for any permutation of n objects, if there is a permutation with k elements, the remaining $n-k$ elements can be ordered however, so the probability of getting a cycle is $(n-k)! \div n!$. The expected value is then the number of such cycles times the probability, i.e.,

$$\binom{n}{k} \times \frac{k!}{k} \times \frac{(n-k)!}{n!} = \frac{n!}{(n-k)!k!} \times \frac{k!}{k} \times \frac{(n-k)!}{n!} = \frac{1}{k}.$$