Homework 6

● Graded

Student

NATHAN LEUNG

Total Points

25 / 25 pts

Question 1

1 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 2

2 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 3

3 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 4

4 — **5** / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

Question 5

5 5 / 5 pts

✔ **– 0 pts** Correct

– 1 pt Minor issue(s)

– 3 pts Major issue(s)

– 5 pts Completely incorrect

1 This is a nice way to do it since it shows exactly what is causing $\tilde{Z}$ to be larger than Z

Question assigned to the following page: 1

MATH 170A Homework 6

Nathan Leung, UID 005835316

16 February 2024

Problem 1

Each of n people (whom we label $1, 2, ..., n$) are randomly and independently assigned a number from the set $\{1, 2, 3, ..., 365\}$ according to the uniform distribution. We will call this number their birthday. Let j and k be distinct labels (between 1 and n) and let A_{jk} denote the event that the corresponding people share a birthday. Let X_{jk} denote the indicator random variable associated to A_{jk}.

Part (a)

Tabulate the joint PMF for X_{12} and X_{13}. Compute the PMF for the product $X_{12}X_{13}$.

First, we define our sample space $\Omega = \{1, 2, 3, ..., 365\}^n$ where the first coordinate of $\omega = (x_1, x_2, ..., x_n) \in \Omega$ corresponds to the first person's birthday, the second corresponds to the second person's birthday, etc.

Next, we expand the given definitions. By the definition of an indicator random variable, we know that for an elementary event $\omega \in \Omega$, we have

$$X_{jk}(\omega) = \begin{cases} 1 & \text{if } \omega \in A_{jk} \\ 0 & \text{if } \omega \notin A_{jk} \end{cases}$$

with $A_{jk} = \{\text{person } j \text{ and person } k \text{ have the same birthday}\}$, $j \neq k$ and $1 \leq j, k \leq n$.

Since we are told that birthdays are randomly and independently assigned, we can assume that each elementary event $\omega \in \omega$ is equiprobable. There are 365^n total elementary events, since each of the n people can have one of 365 different birthdays. Once we fix a person j and person k with $j \neq k$, we claim that there are 365^{n-1} different elementary events $\omega = (x_1, ..., x_n) \in \Omega$ such that $x_j = x_k$ (i.e. j and k have the same birthday).

Indeed, forcing x_j and x_k to have the same birthday is equivalent to removing one degree of freedom from our assignment of birthdays — once we pick x_j's birthday or x_k's birthday (of which there are 365 choices), the other's birthday is also fixed. The remaining $n-2$ choices can be made arbitrarily, with 365 choices for each. Thus, the total number of elementary events $\omega \in \Omega$ satisfying $x_j = x_k$ is $365 \cdot 365^{n-2} = 365^{n-1}$ which is what we claimed. Hence, the probability is $\frac{365^{n-1}}{365^n} = \frac{1}{365}$. Since person j and person k either have the same birthday or have a different birthday (these cases are exhaustive), the probability that they have different birthdays is thus $1 - \frac{1}{365} = \frac{364}{365}$.

Translating this into the language of indicator random variables, we get

Question assigned to the following page: 1

$$
\begin{aligned}
p_{X_{jk}}(\ell) &= \mathbb{P}(X_{jk} = \ell) \\
&= \begin{cases} \mathbb{P}(X_{jk} = 0) & \text{if } \ell = 0 \\ \mathbb{P}(X_{jk} = 1) & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases} && \text{Since } X_{jk} \text{ is only ever 0 or 1} \\
&= \begin{cases} \mathbb{P}(\{\omega \in \Omega : X_{jk}(\omega) = 0\}) & \text{if } \ell = 0 \\ \mathbb{P}(\{\omega \in \Omega : X_{jk}(\omega) = 1\}) & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathbb{P}(\{\omega \in \Omega : \omega \notin A_{jk}\}) & \text{if } \ell = 0 \\ \mathbb{P}(\{\omega \in \Omega : \omega \in A_{jk}\}) & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases} && \text{By definition of } X_{jk} \\
&= \begin{cases} \mathbb{P}(A_{jk}^C) & \text{if } \ell = 0 \\ \mathbb{P}(A_{jk}) & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{364}{365} & \text{if } \ell = 0 \\ \frac{1}{365} & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases} && \text{From our calculations above}
\end{aligned}
$$

We will use this general fact later on.
For now, to compute the joint pmf, we apply the definition:

$$
\begin{aligned}
p_{X_{12}, X_{13}}(\ell, m) &= \begin{cases} \mathbb{P}(X_{12} = 1 \text{ and } X_{13} = 1) & \text{if } \ell = m = 1 \\ \mathbb{P}(X_{12} = 1 \text{ and } X_{13} = 0) & \text{if } \ell = 1 \text{ and } m = 0 \\ \mathbb{P}(X_{12} = 0 \text{ and } X_{13} = 1) & \text{if } \ell = 0 \text{ and } m = 1 \\ \mathbb{P}(X_{12} = 0 \text{ and } X_{13} = 0) & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathbb{P}(\{\omega \in \Omega : X_{12}(\omega) = 1 \text{ and } X_{13}(\omega) = 1\}) & \text{if } \ell = m = 1 \\ \mathbb{P}(\{\omega \in \Omega : X_{12}(\omega) = 1 \text{ and } X_{13}(\omega) = 0\}) & \text{if } \ell = 1 \text{ and } m = 0 \\ \mathbb{P}(\{\omega \in \Omega : X_{12}(\omega) = 0 \text{ and } X_{13}(\omega) = 1\}) & \text{if } \ell = 0 \text{ and } m = 1 \\ \mathbb{P}(\{\omega \in \Omega : X_{12}(\omega) = 0 \text{ and } X_{13}(\omega) = 0\}) & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathbb{P}(\{\omega \in \Omega : \omega \in A_{12} \cap A_{13}\}) & \text{if } \ell = m = 1 \\ \mathbb{P}(\{\omega \in \Omega : \omega \in A_{12} \cap A_{13}^C\}) & \text{if } \ell = 1 \text{ and } m = 0 \\ \mathbb{P}(\{\omega \in \Omega : \omega \in A_{12}^C \cap A_{13}\}) & \text{if } \ell = 0 \text{ and } m = 1 \\ \mathbb{P}(\{\omega \in \Omega : \omega \in A_{12}^C \cap A_{13}^C\}) & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Question assigned to the following page: 1

Nathan Leung, UID 005835316
MATH 170A Homework 6
16 February 2024

$$= \begin{cases} \mathbb{P}(A_{12} \cap A_{13}) & \text{if } \ell = m = 1 \\ \mathbb{P}(A_{12} \cap A_{13}^C) & \text{if } \ell = 1 \text{ and } m = 0 \\ \mathbb{P}(A_{12}^C \cap A_{13}) & \text{if } \ell = 0 \text{ and } m = 1 \\ \mathbb{P}(A_{12}^C \cap A_{13}^C) & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}$$

From here, we can manually compute the probabilities, again assuming each elementary event $\omega \in \Omega$ is equiprobable. We translate each set into words, then rewrite in terms of the sample space:

$$\begin{aligned} A_{12} \cap A_{13} &= \{\text{persons 1 and 2 and 2 and 3 share birthdays}\} \\ &= \{\text{persons 1, 2, and 3 all have the same birthday }\} \\ &= \{(x_1, ..., x_n) \in \Omega : x_1 = x_2 = x_3\} \\ A_{12} \cap A_{13}^C &= \{\text{persons 1 and 2 share a birthday; person 3's is different from person 1's}\} \\ &= \{(x_1, ..., x_n) \in \Omega : x_1 = x_2, x_1 \neq x_3\} \\ A_{12}^C \cap A_{13} &= \{\text{persons 1 and 3 share a birthday; person 2's is different from person 1's}\} \\ &= \{(x_1, ..., x_n) \in \Omega : x_1 = x_3, x_1 \neq x_2\} \\ A_{12}^C \cap A_{13}^C &= \{\text{persons 1 and 2 don't share birthdays; persons 1 and 3 don't share birthdays}\} \\ &= \{(x_1, ..., x_n) \in \Omega : x_1 \neq x_2, x_1 \neq x_3\} \end{aligned}$$

Using similar logic from above, we can count the number of elements in each set. For $A_{12} \cap A_{13}$, we are essentially taking away two degrees of freedom from our original choice. Once we pick a birthday for persons 1, 2, and 3 to share (of which there are 365 choices), there are 365^{n-3} different choices for the remaining people's birthdays (which can be assigned arbitrarily), for a total of $365 \cdot 365^{n-3} = 365^{n-2}$ total elementary events in this set.

For $A_{12} \cap A_{13}^C$, we pick a birthday for persons 1 and 2 to share (of which there are 365 choices), then we pick a different birthday for person 3 (of which there are 364 choices), then we can assign the remaining birthdays arbitrarily (365^{n-3} choices). In total, this gives $365 \cdot 364 \cdot 365^{n-3} = 364 \cdot 365^{n-2}$ total elementary events for this set. A symmetrical argument applies for counting $A_{12}^C \cap A_{13}$.

Finally, for $A_{12}^C \cap A_{12}^C$, we pick a birthday for person 1 (365 choices), then we pick a different birthday for person 2 (364 choices). For person 3, we still have 364 choices since we only need to pick a birthday different from person 1's, not necessarily person 2's. The remaining birthdays can be assigned arbitrarily (365^{n-3} choices). Altogether, we have $365 \cdot 364^2 \cdot 365^{n-3} = 364^2 \cdot 365^{n-2}$ total elementary events for this set.

Since every elementary event is equiprobable, we simply divide these counts by the total cardinality of the sample space, 365^n, to get the probability of each event. Plugging these probabilities into the joint PMF, we have

Question assigned to the following page: 1

$$
\begin{aligned}
p_{X_{12},X_{13}}(\ell,m) &= \begin{cases} \mathbb{P}(A_{12}\cap A_{13}) & \text{if } \ell=m=1 \\ \mathbb{P}(A_{12}\cap A_{13}^C) & \text{if } \ell=1 \text{ and } m=0 \\ \mathbb{P}(A_{12}^C\cap A_{13}) & \text{if } \ell=0 \text{ and } m=1 \\ \mathbb{P}(A_{12}^C\cap A_{13}^C) & \text{if } \ell=0 \text{ and } m=0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{365^{n-2}}{365^n} & \text{if } \ell=m=1 \\ \frac{364\cdot 365^{n-2}}{365^n} & \text{if } \ell=1 \text{ and } m=0 \\ \frac{364\cdot 365^{n-2}}{365^n} & \text{if } \ell=0 \text{ and } m=1 \\ \frac{364^2\cdot 365^{n-2}}{365^n} & \text{if } \ell=0 \text{ and } m=0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Simplifying, we get a final joint PMF of

$$
\boxed{p_{X_{12},X_{13}}(\ell,m) = \begin{cases} \frac{1}{365^2} & \text{if } \ell=m=1 \\ \frac{364}{365^2} & \text{if } \ell=1 \text{ and } m=0 \\ \frac{364}{365^2} & \text{if } \ell=0 \text{ and } m=1 \\ \frac{364^2}{365^2} & \text{if } \ell=0 \text{ and } m=0 \\ 0 & \text{otherwise} \end{cases}}
$$

We are also asked to find the PMF for the product $X_{12}X_{13}$. By the definition of a PMF (and from the fact that $X_{jk}=0$ or $X_{jk}=1$ only), we have

$$
\begin{aligned}
p_{X_{12}X_{13}}(k) &= \mathbb{P}(X_{12}X_{13}=k) \\
&= \begin{cases} \mathbb{P}(X_{12}=1 \text{ and } X_{13}=1) & \text{if } k=1 \\ \mathbb{P}(X_{12}=1 \text{ and } X_{13}=0) & \text{if } k=0 \\ \quad +\mathbb{P}(X_{12}=0 \text{ and } X_{13}=1) & \\ \quad +\mathbb{P}(X_{12}=0 \text{ and } X_{13}=0) & \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} p_{X_{12},X_{13}}(1,1) & \text{if } k=1 \\ p_{X_{12},X_{13}}(1,0)+p_{X_{12},X_{13}}(0,1)+p_{X_{12},X_{13}}(0,0) & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Plugging in the values we computed for the joint PMF above, we have

$$
p_{X_{12}X_{13}}(k) = \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{364}{365^2}+\frac{364}{365^2}+\frac{364^2}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}
$$

Question assigned to the following page: 1

$$= \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{364+364+364^2}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{364(1+1+364)}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{364(366)}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{(365-1)(365+1)}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ \frac{365^2-1^2}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}$$

which simplifies to

$$\boxed{p_{X_{12}X_{13}}(k) = \begin{cases} \frac{1}{365^2} & \text{if } k=1 \\ 1-\frac{1}{365^2} & \text{if } k=0 \\ 0 & \text{otherwise} \end{cases}}$$

Part (b)

Tabulate the joint PMF for X_{12} and X_{34}. Compute the PMF for the product $X_{12}X_{34}$.

We can follow a similar procedure as for **Part (a)** to do this part. By the definition of a joint PMF, we have

$$p_{X_{12},X_{34}}(\ell,m) = \begin{cases} \mathbb{P}(X_{12}=1 \text{ and } X_{34}=1) & \text{if } \ell=m=1 \\ \mathbb{P}(X_{12}=1 \text{ and } X_{34}=0) & \text{if } \ell=1 \text{ and } m=0 \\ \mathbb{P}(X_{12}=0 \text{ and } X_{34}=1) & \text{if } \ell=0 \text{ and } m=1 \\ \mathbb{P}(X_{12}=0 \text{ and } X_{34}=0) & \text{if } \ell=0 \text{ and } m=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \mathbb{P}(A_{12}\cap A_{34}) & \text{if } \ell=m=1 \\ \mathbb{P}(A_{12}\cap A_{34}^C) & \text{if } \ell=1 \text{ and } m=0 \\ \mathbb{P}(A_{12}^C\cap A_{34}) & \text{if } \ell=0 \text{ and } m=1 \\ \mathbb{P}(A_{12}^C\cap A_{34}^C) & \text{if } \ell=0 \text{ and } m=0 \\ 0 & \text{otherwise} \end{cases} \qquad \text{Same logic as } \textbf{Part (a)}$$

Question assigned to the following page: 1

The set expressions correspond to the following events:

$$\begin{aligned}
A_{12} \cap A_{34} &= \{\text{persons 1 and 2 share a birthday, persons 3 and 4 share a birthday}\} \\
&= \{(x_1, ..., x_n) \in \Omega : x_1 = x_2, x_3 = x_4\} \\
A_{12} \cap A_{34}^C &= \{\text{persons 1 and 2 share a birthday, persons 3 and 4 do not share a birthday}\} \\
&= \{(x_1, ..., x_n) \in \Omega : x_1 = x_2, x_3 \neq x_4\} \\
A_{12}^C \cap A_{34} &= \{\text{persons 1 and 2 do not share a birthday, persons 3 and 4 share a birthday}\} \\
&= \{(x_1, ..., x_n) \in \Omega : x_1 \neq x_2, x_3 = x_4\} \\
A_{12}^C \cap A_{34}^C &= \{\text{persons 1 and 2 do not share a birthday, persons 3 and 4 do not share a birthday}\} \\
&= \{(x_1, ..., x_n) \in \Omega : x_1 \neq x_2, x_3 \neq x_4\}
\end{aligned}$$

For each event, we count the number of elementary events corresponding to the event. For $A_{12} \cap A_{34}$, there are 365 choices for the shared birthday of persons 1 and 2, 365 choices for the shared birthday of persons 3 and 4, and 365^{n-4} choices for the birthdays of the remaining $(n-4)$ people for a total of 365^{n-2} elementary events.

In the case of $A_{12} \cap A_{34}^C$, there are 365 choices for the shared birthday of persons 1 and 2, 365 choices for the birthday of person 3, 364 choices for the birthday of person 4 (since it must be different from person 3's birthday), and 365^{n-4} choices for the birthdays of the remaining people for a total of $365 \cdot 365 \cdot 364 \cdot 365^{n-4} = 364 \cdot 365^{n-2}$ elementary events. A symmetrical argument applies for $A_{12}^C \cap A_{34}$.

Finally, for $A_{12}^C \cap A_{34}^C$, there are 365 choices for the birthday of person 1, 364 choices for the birthday of person 2 (different from person 1's), and similarly 365 choices for person 3's birthday and 364 choices for person 4's. There are 365^{n-4} ways to pick birthdays for the remaining people. Thus, there are a total of $365 \cdot 364 \cdot 365 \cdot 364 \cdot 365^{n-4} = 364^2 \cdot 365^{n-2}$ total elementary events corresponding to this event.

Again, since each elementary event is equiprobable, we divide these counts by the total number of elementary events in the sample space, 365^n, to get the absolute probability of each event. The joint PMF thus becomes

$$\begin{aligned}
p_{X_{12}, X_{34}}(\ell, m) &= \begin{cases} \mathbb{P}(A_{12} \cap A_{34}) & \text{if } \ell = m = 1 \\ \mathbb{P}(A_{12} \cap A_{34}^C) & \text{if } \ell = 1 \text{ and } m = 0 \\ \mathbb{P}(A_{12}^C \cap A_{34}) & \text{if } \ell = 0 \text{ and } m = 1 \\ \mathbb{P}(A_{12}^C \cap A_{34}^C) & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{365^{n-2}}{365^n} & \text{if } \ell = m = 1 \\ \frac{364 \cdot 365^{n-2}}{365^n} & \text{if } \ell = 1 \text{ and } m = 0 \\ \frac{364 \cdot 365^{n-2}}{365^n} & \text{if } \ell = 0 \text{ and } m = 1 \\ \frac{364^2 \cdot 365^{n-2}}{365^n} & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Question assigned to the following page: 1

We simplify to get a final joint PMF of

$$\boxed{p_{X_{12},X_{34}}(\ell, m) = \begin{cases} \frac{1}{365^2} & \text{if } \ell = m = 1 \\ \frac{364}{365^2} & \text{if } \ell = 1 \text{ and } m = 0 \\ \frac{364}{365^2} & \text{if } \ell = 0 \text{ and } m = 1 \\ \frac{364^2}{365^2} & \text{if } \ell = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}}$$

We also want to compute the PMF for the product $X_{12}X_{34}$. Since $X_{jk} = 0$ or $X_{jk} = 1$ only, the only two possible products are 0 or 1. The product PMF is thus

$$\begin{aligned} p_{X_{12}X_{34}}(k) &= \mathbb{P}(X_{12}X_{34} = k) \\ &= \begin{cases} \mathbb{P}(X_{12} = 1 \text{ and } X_{34} = 1) & \text{if } k = 1 \\ \mathbb{P}(X_{12} = 1 \text{ and } X_{34} = 0) & \text{if } k = 0 \\ \quad +\mathbb{P}(X_{12} = 0 \text{ and } X_{34} = 1) & \\ \quad +\mathbb{P}(X_{12} = 0 \text{ and } X_{34} = 0) & \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} p_{X_{12},X_{34}}(1,1) & \text{if } k = 1 \\ p_{X_{12},X_{34}}(1,0) + p_{X_{12},X_{13}}(0,1) + p_{X_{12},X_{34}}(0,0) & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

We can plug in the values we computed for the joint PMF above.

$$p_{X_{12}X_{34}}(k) = \begin{cases} \frac{1}{365^2} & \text{if } k = 1 \\ \frac{364}{365^2} + \frac{364}{365^2} + \frac{364^2}{365^2} & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

Note that the joint PMF of X_{12} and X_{34} is actually the same as that of X_{12} and X_{13} from **Part (a)** above, so our expressions above simplify to the same product PMF as before:

$$\boxed{p_{X_{12}X_{34}}(k) = \begin{cases} \frac{1}{365^2} & \text{if } k = 1 \\ 1 - \frac{1}{365^2} & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}}$$

Part (c)

Are A_{12} and A_{34} independent? Are they independent conditioned on A_{13}?

We claim that A_{12} and A_{34} are independent. It suffices to show that

$$\mathbb{P}(A_{12} \cap A_{34}) = \mathbb{P}(A_{12}) \cdot \mathbb{P}(A_{34})$$

We can write each term above in terms of indicator random variables as follows:

Question assigned to the following page: 1

$$\begin{aligned}
\mathbb{P}(A_{12}) &= \mathbb{P}(X_{12} = 1) \\
\mathbb{P}(A_{34}) &= \mathbb{P}(X_{34} = 1) \\
\mathbb{P}(A_{12} \cap A_{34}) &= \mathbb{P}(X_{12}X_{34} = 1)
\end{aligned}$$

We can compute these probabilities using the PMFs we found in **Part (a)** and **Part (b)**.

$$\begin{aligned}
\mathbb{P}(X_{12} = 1) &= p_{X_{12}}(1) = \frac{1}{365} \\
\mathbb{P}(X_{34} = 1) &= p_{X_{34}}(1) = \frac{1}{365} \\
\mathbb{P}(X_{12}X_{34} = 1) &= p_{X_{12}X_{34}}(1) = \frac{1}{365^2}
\end{aligned}$$

Thus we have

$$\begin{aligned}
\mathbb{P}(A_{12} \cap A_{34}) &= \mathbb{P}(X_{12}X_{34} = 1) \\
&= \frac{1}{365^2} \\
\mathbb{P}(A_{12}) \cdot \mathbb{P}(A_{34}) &= \mathbb{P}(X_{12} = 1) \cdot \mathbb{P}(X_{34} = 1) \\
&= \frac{1}{365} \cdot \frac{1}{365} \\
&= \frac{1}{365^2}
\end{aligned}$$

and hence $\boxed{A_{12} \text{ and } A_{34} \text{ are indeed independent}}$.

We also want to check that A_{12} and A_{34} are independent conditioned on A_{13}. To demonstrate independence conditioned on A_{13}, it suffices to show that

$$\mathbb{P}(A_{12} \cap A_{34} \mid A_{13}) = \mathbb{P}(A_{12} \mid A_{13}) \cdot \mathbb{P}(A_{34} \mid A_{13})$$

Expanding out the definition of conditional probability, we have

$$\begin{aligned}
\mathbb{P}(A_{12} \cap A_{34} \mid A_{13}) &= \frac{\mathbb{P}(A_{12} \cap A_{34} \cap A_{13})}{\mathbb{P}(A_{13})} \\
\mathbb{P}(A_{12} \mid A_{13}) &= \frac{\mathbb{P}(A_{12} \cap A_{13})}{\mathbb{P}(A_{13})} \\
\mathbb{P}(A_{34} \mid A_{13}) &= \frac{\mathbb{P}(A_{34} \cap A_{13})}{\mathbb{P}(A_{13})}
\end{aligned}$$

Hence, we equivalently want to show that

Question assigned to the following page: 1

$$\frac{\mathbb{P}(A_{12} \cap A_{34} \cap A_{13})}{\mathbb{P}(A_{13})} = \frac{\mathbb{P}(A_{12} \cap A_{13})}{\mathbb{P}(A_{13})} \cdot \frac{\mathbb{P}(A_{34} \cap A_{13})}{\mathbb{P}(A_{13})}$$

We can write each term above in terms of indicator random variables as follows:

$$\begin{aligned}
\mathbb{P}(A_{13}) &= \mathbb{P}(X_{13} = 1) \\
\mathbb{P}(A_{12} \cap A_{13}) &= \mathbb{P}(X_{12}X_{13} = 1) \\
\mathbb{P}(A_{34} \cap A_{13}) &= \mathbb{P}(X_{34}X_{13} = 1) \\
\mathbb{P}(A_{12} \cap A_{34} \cap A_{13}) &= \mathbb{P}(X_{12}X_{34}X_{13} = 1)
\end{aligned}$$

The first three terms above can be computed using the PMFs we found in **Part (a)**:

$$\begin{aligned}
\mathbb{P}(X_{13} = 1) &= p_{X_{13}}(1) = \frac{1}{365} \\
\mathbb{P}(X_{12}X_{13} = 1) &= p_{X_{12}X_{13}}(1) = \frac{1}{365^2} \\
\mathbb{P}(X_{34}X_{13} = 1) &= p_{X_{34}X_{13}}(1) = \frac{1}{365^2}
\end{aligned}$$

where we get the second product by relabeling $1 \mapsto 3$, $2 \mapsto 4$, $3 \mapsto 1$ and $4 \mapsto 2$ (i.e. X_{12} becomes X_{34} and X_{13} becomes X_{31}), noting that $X_{31} = X_{13}$ (if persons 1 and 3 have the same birthday, then persons 3 and 1 have the same birthday), and then applying **Part (a)**.

To find $\mathbb{P}(A_{12} \cap A_{34} \cap A_{13})$, note that the event $A_{12} \cap A_{34} \cap A_{13}$ occurs when persons 1, 2, 3 and 4 all share the same birthday (by the transitive property of equality: person 1's is the same as 2's and 3's, and 3's is the same as 4's). There are 365 choices for this shared birthday, and 365^{n-4} ways to pick birthdays for the remaining people, giving a total of $365 \cdot 365^{n-4} = 365^{n-3}$ elementary events. Since each elementary event is equiprobable, we get $\mathbb{P}(A_{12} \cap A_{34} \cap A_{13}) = \frac{365^{n-3}}{365^n} = \frac{1}{365^3}$.

Thus we have

$$\begin{aligned}
\mathbb{P}(A_{12} \cap A_{34} \mid A_{13}) = \frac{\mathbb{P}(A_{12} \cap A_{34} \cap A_{13})}{\mathbb{P}(A_{13})} &= \frac{\frac{1}{365^3}}{\frac{1}{365}} \\
&= \frac{365}{365^3} \\
&= \frac{1}{365^2} \\
\mathbb{P}(A_{12} \mid A_{13}) = \frac{\mathbb{P}(A_{12} \cap A_{13})}{\mathbb{P}(A_{13})} &= \frac{\frac{1}{365^2}}{\frac{1}{365}} \\
&= \frac{365}{365^2} \\
&= \frac{1}{365} \\
\mathbb{P}(A_{34} \mid A_{13}) = \frac{\mathbb{P}(A_{34} \cap A_{13})}{\mathbb{P}(A_{13})} &= \frac{\frac{1}{365^2}}{\frac{1}{365}}
\end{aligned}$$

Question assigned to the following page: 1

$$
\begin{aligned}
&= \frac{365}{365^2} \\
&= \frac{1}{365}
\end{aligned}
$$

and so

$$
\begin{aligned}
\mathbb{P}(A_{12} \mid A_{13}) \cdot \mathbb{P}(A_{34} \mid A_{13}) &= \frac{1}{365} \cdot \frac{1}{365} \\
&= \frac{1}{365^2} \\
&= \mathbb{P}(A_{12} \cap A_{34} \mid A_{13})
\end{aligned}
$$

and so $\boxed{A_{12} \text{ and } A_{34} \text{ are also independent when conditioned on } A_{13}}$.

Part (d)

Are A_{12} and A_{13} independent? Are they independent conditioned on A_{23}?

We claim that A_{12} and A_{13} are independent. It suffices to show that

$$\mathbb{P}(A_{12} \cap A_{13}) = \mathbb{P}(A_{12}) \cdot \mathbb{P}(A_{13})$$

We can write each term above in terms of indicator random variables as follows:

$$
\begin{aligned}
\mathbb{P}(A_{12}) &= \mathbb{P}(X_{12} = 1) \\
\mathbb{P}(A_{13}) &= \mathbb{P}(X_{13} = 1) \\
\mathbb{P}(A_{12} \cap A_{13}) &= \mathbb{P}(X_{12}X_{13} = 1)
\end{aligned}
$$

We can compute these probabilities using the PMFs we found in **Part (a)**.

$$
\begin{aligned}
\mathbb{P}(X_{12} = 1) = p_{X_{12}}(1) &= \frac{1}{365} \\
\mathbb{P}(X_{13} = 1) = p_{X_{13}}(1) &= \frac{1}{365} \\
\mathbb{P}(X_{12}X_{13} = 1) = p_{X_{12}X_{13}}(1) &= \frac{1}{365^2}
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\mathbb{P}(A_{12} \cap A_{13}) &= \mathbb{P}(X_{12}X_{13} = 1) \\
&= \frac{1}{365^2} \\
\mathbb{P}(A_{12}) \cdot \mathbb{P}(A_{13}) &= \mathbb{P}(X_{12} = 1) \cdot \mathbb{P}(X_{13} = 1) \\
&= \frac{1}{365} \cdot \frac{1}{365}
\end{aligned}
$$

Question assigned to the following page: 1

$$= \frac{1}{365^2}$$

and hence $\boxed{A_{12} \text{ and } A_{13} \text{ are indeed independent}}$.

We also want to check that A_{12} and A_{13} are independent conditioned on A_{23}. To demonstrate independence conditioned on A_{23}, it suffices to show that

$$\mathbb{P}(A_{12} \cap A_{13} \mid A_{23}) = \mathbb{P}(A_{12} \mid A_{23}) \cdot \mathbb{P}(A_{13} \mid A_{23})$$

Expanding out the definition of conditional probability, we have

$$\begin{aligned}
\mathbb{P}(A_{12} \cap A_{13} \mid A_{23}) &= \frac{\mathbb{P}(A_{12} \cap A_{13} \cap A_{23})}{\mathbb{P}(A_{23})} \\
\mathbb{P}(A_{12} \mid A_{23}) &= \frac{\mathbb{P}(A_{12} \cap A_{23})}{\mathbb{P}(A_{23})} \\
\mathbb{P}(A_{13} \mid A_{23}) &= \frac{\mathbb{P}(A_{13} \cap A_{23})}{\mathbb{P}(A_{23})}
\end{aligned}$$

Hence, we equivalently want to show that

$$\frac{\mathbb{P}(A_{12} \cap A_{13} \cap A_{23})}{\mathbb{P}(A_{23})} = \frac{\mathbb{P}(A_{12} \cap A_{23})}{\mathbb{P}(A_{23})} \cdot \frac{\mathbb{P}(A_{13} \cap A_{23})}{\mathbb{P}(A_{23})}$$

We can write each term above in terms of indicator random variables as follows:

$$\begin{aligned}
\mathbb{P}(A_{23}) &= \mathbb{P}(X_{23} = 1) \\
\mathbb{P}(A_{12} \cap A_{23}) &= \mathbb{P}(X_{12}X_{23} = 1) \\
\mathbb{P}(A_{13} \cap A_{23}) &= \mathbb{P}(X_{13}X_{23} = 1) \\
\mathbb{P}(A_{12} \cap A_{13} \cap A_{23}) &= \mathbb{P}(X_{12}X_{13}X_{23} = 1)
\end{aligned}$$

The first three terms above can be computed using the PMFs we found in **Part (a)**:

$$\begin{aligned}
\mathbb{P}(X_{23} = 1) &= p_{X_{23}}(1) = \frac{1}{365} \\
\mathbb{P}(X_{12}X_{23} = 1) &= p_{X_{12}X_{23}}(1) = \frac{1}{365^2} \\
\mathbb{P}(X_{13}X_{23} = 1) &= p_{X_{13}X_{23}}(1) = \frac{1}{365^2}
\end{aligned}$$

where the second term was calculated by reindexing by $1 \mapsto 2$, $2 \mapsto 3$, and $3 \mapsto 1$ (so X_{12} becomes X_{23} and X_{13} becomes $X_{21} = X_{12}$), then applying **Part (a)**, and similarly for the third term we reindex by $1 \mapsto 3$, $2 \mapsto 2$, and $3 \mapsto 1$ (so X_{12} becomes $X_{32} = X_{23}$ and X_{13} becomes $X_{31} = X_{13}$).

Question assigned to the following page: 1

To find $\mathbb{P}(A_{12} \cap A_{13} \cap A_{23})$, note that the event $A_{12} \cap A_{13} \cap A_{23}$ occurs when persons 1, 2 and 3 all share the same birthday (by the transitive property of equality: person 1's is the same as 2's and 3's, so 2's and 3's must be the same). There are 365 choices for this shared birthday, and 365^{n-3} ways to pick birthdays for the remaining people, giving a total of $365 \cdot 365^{n-3} = 365^{n-2}$ elementary events. Since each elementary event is equiprobable, we get $\mathbb{P}(A_{12} \cap A_{13} \cap A_{23}) = \frac{365^{n-2}}{365^n} = \frac{1}{365^2}$.

Thus we have

$$\begin{aligned}
\mathbb{P}(A_{12} \cap A_{13} \mid A_{23}) = \frac{\mathbb{P}(A_{12} \cap A_{34} \cap A_{23})}{\mathbb{P}(A_{23})} &= \frac{\frac{1}{365^2}}{\frac{1}{365}} \\
&= \frac{365}{365^2} \\
&= \frac{1}{365} \\
\mathbb{P}(A_{12} \mid A_{23}) = \frac{\mathbb{P}(A_{12} \cap A_{23})}{\mathbb{P}(A_{23})} &= \frac{\frac{1}{365^2}}{\frac{1}{365}} \\
&= \frac{365}{365^2} \\
&= \frac{1}{365} \\
\mathbb{P}(A_{13} \mid A_{23}) = \frac{\mathbb{P}(A_{13} \cap A_{23})}{\mathbb{P}(A_{23})} &= \frac{\frac{1}{365^2}}{\frac{1}{365}} \\
&= \frac{365}{365^2} \\
&= \frac{1}{365}
\end{aligned}$$

and so

$$\begin{aligned}
\mathbb{P}(A_{12} \mid A_{23}) \cdot \mathbb{P}(A_{13} \mid A_{23}) &= \frac{1}{365} \cdot \frac{1}{365} \\
&= \frac{1}{365^2} \\
&\neq \frac{1}{365} = \mathbb{P}(A_{12} \cap A_{13} \mid A_{23})
\end{aligned}$$

and so $\boxed{A_{12} \text{ and } A_{13} \text{ are in fact not independent when conditioned on } A_{23}}$.

Part (e)

Compute the expected number of pairs of people who share a birthday (hint: write this number as a sum of X_{jk}s).

Let Z be the random variable corresponding to the number of pairs of people who share a birthday. We want to compute $\mathbb{E}[Z]$. We can write Z in terms of X_{jk} as

Question assigned to the following page: 1

$$Z = \sum_{1 \le j < k \le n} X_{jk}$$

where X_{jk} is the indicator random variable associated with A_{jk} (i.e. the random variable that is 1 when persons j and k share a birthday and 0 otherwise). Note that the constraint $1 \le j < k \le n$ in the summation ensures that each unique pair of people (j, k) is only counted once. To calculate the expectation of Z, we can invoke linearity of expectation:

$$\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}\left[\sum_{1 \le j < k \le n} X_{jk}\right] \\
&= \sum_{1 \le j < k \le n} \mathbb{E}\left[X_{jk}\right] \\
&= \sum_{1 \le j < k \le n} \mathbb{P}(X_{jk} = 1) && \text{Definition of indicator random variable, } \mathbb{E} \\
&= \sum_{1 \le j < k \le n} \frac{1}{365} && \text{Calculated in } \textbf{Part (a)} \\
&= \binom{n}{2} \cdot \frac{1}{365} && \text{There are } \binom{n}{2} \text{ unique pairs}
\end{aligned}$$

Thus, the expected number of pairs of people who share a birthday is $\mathbb{E}[Z] = \boxed{\binom{n}{2} \cdot \frac{1}{365}}$.

Part (f)

Compute the second moment and variance of the number of pairs of people who share a birthday.

With Z as the random variable corresponding to the number of pairs of people who share a birthday as in **Part (e)**, we want to compute the second moment $\mathbb{E}[Z^2]$ and the variance $\mathrm{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$. First, we compute the second moment. We have

$$\begin{aligned}
\mathbb{E}[Z^2] &= \mathbb{E}\left[\left(\sum_{1 \le j < k \le n} X_{jk}\right)^2\right] \\
&= \mathbb{E}\left[\underbrace{\left(\sum_{1 \le j < k \le n} X_{jk}^2\right)}_{\text{product of same } X_{jk}} + 2 \cdot \underbrace{\left(\sum_{\substack{1 \le j < k \le n \\ 1 \le \ell < m < n \\ \text{if } j = \ell, \text{ then } k < m \\ \text{otherwise } j < \ell}} X_{jk} X_{\ell m}\right)}_{\text{product of distinct } X_{jk}}\right]
\end{aligned}$$

Question assigned to the following page: 1

where the condition under the second sum above ensures that each product of two different X_{jk}s is counted exactly once. Applying linearity of expectation yields

$$\begin{aligned}
\mathbb{E}[Z^2] &= \mathbb{E}\left[\left(\sum_{1\leq j<k\leq n} X_{jk}^2\right) + 2\cdot\left(\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} X_{jk}X_{\ell m}\right)\right]\\
&= \mathbb{E}\left[\sum_{1\leq j<k\leq n} X_{jk}^2\right] + 2\cdot\mathbb{E}\left[\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} X_{jk}X_{\ell m}\right]
\end{aligned}$$

Note that $X_{jk}^2 = X_{jk}$, since X_{jk} is only ever 1 or 0. Thus, we can write

$$\begin{aligned}
\mathbb{E}[Z^2] &= \mathbb{E}\left[\sum_{1\leq j<k\leq n} X_{jk}^2\right] + 2\cdot\mathbb{E}\left[\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} X_{jk}X_{\ell m}\right]\\
&= \mathbb{E}\left[\sum_{1\leq j<k\leq n} X_{jk}\right] + 2\cdot\mathbb{E}\left[\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} X_{jk}X_{\ell m}\right]\\
&= \binom{n}{2}\cdot\frac{1}{365} + 2\cdot\mathbb{E}\left[\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} X_{jk}X_{\ell m}\right] && \text{From } \textbf{Part (e)}\\
&= \binom{n}{2}\cdot\frac{1}{365} + 2\cdot\left(\sum_{\substack{1\leq j<k\leq n\\ 1\leq \ell<m<n\\ \text{if } j=\ell,\text{ then } k<m\\ \text{otherwise } j<\ell}} \mathbb{E}\left[X_{jk}X_{\ell m}\right]\right) && \text{Linearity of expectation}
\end{aligned}$$

Question assigned to the following page: 1

$$= \binom{n}{2} \cdot \frac{1}{365} + 2 \cdot \left(\sum_{\substack{1 \le j < k \le n \\ 1 \le \ell < m < n \\ \text{if } j = \ell, \text{ then } k < m \\ \text{otherwise } j < \ell}} \mathbb{P}\left(X_{jk} X_{\ell m} = 1\right) \right) \quad \text{Definition of indicator r.v. and } \mathbb{E}$$

$$= \binom{n}{2} \cdot \frac{1}{365} + 2 \cdot \left(\sum_{\substack{1 \le j < k \le n \\ 1 \le \ell < m < n \\ \text{if } j = \ell, \text{ then } k < m \\ \text{otherwise } j < \ell}} \frac{1}{365^2} \right) \quad \text{From } \textbf{Part (a)} \text{ and } \textbf{Part (b)}$$

The last equality above comes from relabeling X_{jk} and $X_{\ell m}$ as we have done in previous parts to X_{12} and X_{13} or X_{12} and X_{34}; more precisely, the pairs (j, k) and (ℓ, m) will share at most one common component, and we saw in both **Part (a)** and **Part (b)** that the product PMF is the same regardless.

Note that there are $\binom{n}{2} \cdot \left(\binom{n}{2} - 1\right) \cdot \frac{1}{2}$ different pairs of (j, k) and (ℓ, m) satisfying the constraints of the last summation above: we pick a first pair of people, then we pick a second, distinct pair of people, then we divide by two since we are counting both $((j, k), (\ell, m))$ and $((\ell, m), (j, k))$ (we need to take into account the ordering condition of the sum as well). Thus we get

$$\begin{aligned} \mathbb{E}[Z^2] &= \binom{n}{2} \cdot \frac{1}{365} + 2 \cdot \left(\sum_{\substack{1 \le j < k \le n \\ 1 \le \ell < m < n \\ \text{if } j = \ell, \text{ then } k < m \\ \text{otherwise } j < \ell}} \frac{1}{365^2} \right) \\ &= \binom{n}{2} \cdot \frac{1}{365} + 2 \cdot \left(\binom{n}{2} \cdot \left(\binom{n}{2} - 1 \right) \cdot \frac{1}{2} \right) \cdot \frac{1}{365^2} \end{aligned}$$

which simplifies to

$$\boxed{\mathbb{E}[Z^2] = \binom{n}{2} \cdot \frac{1}{365} + \binom{n}{2} \cdot \left(\binom{n}{2} - 1 \right) \cdot \frac{1}{365^2}}$$

For the expectation, we can plug the second moment into the formula

$$\begin{aligned} \operatorname{Var}(Z) &= \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 \\ &= \left(\binom{n}{2} \cdot \frac{1}{365} + \binom{n}{2} \cdot \left(\binom{n}{2} - 1 \right) \cdot \frac{1}{365^2} \right) - \left(\binom{n}{2} \cdot \frac{1}{365} \right)^2 \end{aligned}$$

Questions assigned to the following page: 1 and 2

$$
\begin{aligned}
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(\left(1 + \left(\binom{n}{2} - 1\right) \cdot \frac{1}{365}\right) - \left(\binom{n}{2} \cdot \frac{1}{365}\right)\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(1 + \left(\binom{n}{2} - 1\right) \cdot \frac{1}{365} - \binom{n}{2} \cdot \frac{1}{365}\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(1 + \frac{1}{365} \cdot \left(\left(\binom{n}{2} - 1\right) - \binom{n}{2}\right)\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(1 + \frac{1}{365} \cdot \left(\binom{n}{2} - 1 - \binom{n}{2}\right)\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(1 + \frac{1}{365} \cdot (-1)\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \left(1 - \frac{1}{365}\right) \\
&= \left(\binom{n}{2} \cdot \frac{1}{365}\right) \cdot \frac{364}{365}
\end{aligned}
$$

which simplifies to

$$\boxed{\mathrm{Var}(Z) = \binom{n}{2} \cdot \frac{364}{365^2}}$$

Problem 2

An infant repeatedly attempts to build a stack of three blocks. Her probability of balancing any particular block is p (including the first block of the stack), independent of any other attempt. Failure to balance any particular block results in the whole stack collapsing.

Part (a)

What is the probability she successfully builds the stack on any one attempt?

In order to build the stack on one attempt, she needs to balance three blocks in a row. Since the probability of balancing each block is independent, the probability of balancing three blocks in a row is $p \cdot p \cdot p = \boxed{p^3}$.

Part (b)

What is the PMF for the number of attempts needed to first successfully complete the stack?

Let X be the random variable corresponding to the number of attempts needed for the first successful completion of the stack. We define a sample space

$$\Omega = \{0, 1, 2, 3\}^{\mathbb{N}}$$

with elementary outcomes $\omega = (x_1, x_2, x_3, ...)$ where x_i denotes the number of successfully balanced blocks on attempt i. The stack is successfully built on attempt i if $x_i = 3$. For a

Question assigned to the following page: 2

specific example, we might have $\omega = (1, 2, 0, 2, 3, ...)$. In this case, $X(\omega) = 5$ since the stack is first successfully completed on attempt 5.

From **Part (a)**, we know that the probability of successfully building the stack on any one attempt is p^3. Thus, the probability of failing to build the stack on any one attempt is $1-p^3$. Since the probability of succeeding on any attempt is independent of the previous attempts (this follows from the independence of probabilities of balancing any single block), the number of attempts X needed for the first successful completion of the stack can in fact be modeled as a Geometric(p^3) variable. Thus, we have the PMF

$$\boxed{p_X(k) = \begin{cases} (1-p^3)^{k-1}p^3 & \text{if } k \geq 1 \\ 0 & \text{otherwise} \end{cases}}$$

for the number of attempts needed to first successfully complete the stack.

Part (c)

What is the (conditional) PMF for the number of blocks successfully balanced in any particular attempt to build the stack, given that the attempt fails?

Fix an attempt $i \in \mathbb{N}$. Let Y_i denote the random variable corresponding to the number of blocks successfully balanced on attempt i, given that attempt i fails. Given that the attempt fails, we have three cases:

1. exactly 0 blocks balanced
2. exactly 1 block balanced
3. exactly 2 blocks balanced

Since the attempt ends when the first failure to balance occurs, and the probability of balancing blocks is independent each time, we can calculate the absolute probabilities by taking products:

1. $\mathbb{P}(\{\text{ exactly 0 blocks balanced on attempt } i \}) = (1-p)$ (i.e. fail on first block)
2. $\mathbb{P}(\{\text{ exactly 1 block balanced on attempt } i \}) = p(1-p)$ (i.e. fail on second block)
3. $\mathbb{P}(\{\text{ exactly 2 blocks balanced on attempt } i \}) = p^2(1-p)$ (i.e. fail on third block)

By summing the probabilities above, we can also get the absolute probability that attempt i fails (since the attempt only fails if exactly 0, 1, or 2 blocks are balanced):

$$\begin{aligned}
&\mathbb{P}(\{\text{ attempt } i \text{ fails }\}) \\
&\qquad = \mathbb{P}(\{\text{ exactly 0 blocks balanced on attempt } i \}) \\
&\qquad\quad + \mathbb{P}(\{\text{ exactly 1 block balanced on attempt } i \}) \\
&\qquad\quad + \mathbb{P}(\{\text{ exactly 2 blocks balanced on attempt } i \}) \\
&\qquad = (1-p) + p(1-p) + p^2(1-p) \\
&\qquad = (1+p+p^2)(1-p)
\end{aligned}$$

Question assigned to the following page: 2

$$= 1 - p^3$$

By the definition of conditional probability, we have

$$\begin{aligned}\mathbb{P}(Y_i = k) &= \mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i \mid \text{attempt } i \text{ fails}\}) \\ &= \frac{\mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i \cap \text{ attempt } i \text{ fails}\})}{\mathbb{P}(\{\text{attempt } i \text{ fails}\})}\end{aligned}$$

Note that

$$\begin{aligned}&\mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i \cap \text{ attempt } i \text{ fails}\}) \\ &\qquad = \mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i\})\end{aligned}$$

for $k = 0, 1, 2$ since success only occurs when exactly 3 blocks are balanced. In terms of the random variable Y_i, then, for $k = 0, 1, 2$, we get

$$\begin{aligned}\mathbb{P}(Y_i = k) &= \frac{\mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i\})}{\mathbb{P}(\{\text{attempt } i \text{ fails}\})} \\ &= \frac{\mathbb{P}(\{\text{exactly } k \text{ blocks successfully balanced on attempt } i\})}{1 - p^3}\end{aligned}$$

Thus, our conditional PMF is

$$\boxed{p_{Y_i}(k) = \begin{cases} \frac{1-p}{1-p^3} & \text{if } k = 0 \\ \frac{p(1-p)}{1-p^3} & \text{if } k = 1 \\ \frac{p^2(1-p)}{1-p^3} & \text{if } k = 2 \\ 0 & \text{otherwise} \end{cases}}$$

Part (d)

What is the expected number of blocks balanced successfully in a failed attempt to build the stack?

We fix an attempt $i \in \mathbb{N}$. We want to calculate $\mathbb{E}[Y_i]$. By the definition of expectation, we have

$$\begin{aligned}\mathbb{E}[Y_i] &= \sum_{k=0}^{2} k \cdot \mathbb{P}(Y_i = k) \\ &= \sum_{k=1}^{2} k \cdot \mathbb{P}(Y_i = k) \\ &= \mathbb{P}(Y_i = 1) + 2 \cdot \mathbb{P}(Y_i = 2) \\ &= p_{Y_i}(1) + 2 \cdot p_{Y_i}(2) \\ &= \frac{p(1-p)}{1-p^3} + 2 \cdot \frac{p^2(1-p)}{1-p^3}\end{aligned}$$

Question assigned to the following page: 2

$$
\begin{aligned}
&= \frac{p(1-p) + 2p^2(1-p)}{1-p^3} \\
&= \frac{(p+2p^2)(1-p)}{1-p^3} \\
&= \frac{(p+2p^2)(1-p)}{(1-p)(1+p+p^2)}
\end{aligned}
$$

which simplifies to

$$\boxed{\mathbb{E}[Y_i] = \frac{p+2p^2}{1+p+p^2}}$$

Part (e)

Fix $\ell \in \{1, 2, 3, ...\}$. Given that the first successful completion of the stack happened on the ℓth attempt, what is the expected value of the total number of blocks balanced successfully over the course of these ℓ attempts?

Let Z_ℓ be the random variable corresponding to the total number of blocks balanced successfully over the course of ℓ attempts given that the ℓth attempt is the first success. We can write

$$Z_\ell = \left(\sum_{i=1}^{\ell-1} Y_i\right) + 3$$

since attempts $1 \leq i \leq \ell - 1$ are all failures (since the ℓth attempt is the first successful completion of the stack), and Y_i is defined to be the number of blocks successfully balanced on attempt i, given that attempt i fails (from **Part (c)**). We add 3 to the end because on the ℓth attempt, since it was a success, the infant must have balanced 3 blocks successfully. We want to calculate $\mathbb{E}[Z_\ell]$. We have

$$
\begin{aligned}
\mathbb{E}[Z_\ell] &= \mathbb{E}\left[\left(\sum_{i=1}^{\ell-1} Y_i\right) + 3\right] \\
&= \left(\sum_{i=1}^{\ell-1} \mathbb{E}\left[Y_i\right]\right) + \mathbb{E}[3] && \text{Linearity of } \mathbb{E} \\
&= \left(\sum_{i=1}^{\ell-1} \mathbb{E}\left[Y_i\right]\right) + 3 \\
&= \left(\sum_{i=1}^{\ell-1} \frac{p+2p^2}{1+p+p^2}\right) + 3 && \text{From } \textbf{Part (d)}
\end{aligned}
$$

which simplifies to

Question assigned to the following page: 2

$$\boxed{\mathbb{E}[Z_\ell] = (\ell - 1) \cdot \left(\frac{p + 2p^2}{1 + p + p^2}\right) + 3}$$

Part (f)

What is the expected total number of blocks balanced successfully up to and including the first successful completion of the whole stack?

Let W be the random variable corresponding to the total number of blocks balanced successfully up to the first success. We want to calculate $\mathbb{E}[W]$. Note that we can partition Ω based on the attempt on which the first success occurs: the first success is either on the first attempt, on the second attempt, on the third attempt, etc., and the first success cantnot occur on more than one attempt. We can apply the partition theorem to the expectation to get

$$\begin{aligned}
\mathbb{E}[W] &= \sum_{\ell=1}^{\infty} \mathbb{E}[W \mid \{\text{first success on } \ell\text{th attempt}\}] \cdot \mathbb{P}(\{\text{first success on } \ell\text{th attempt}\}) \\
&= \sum_{\ell=1}^{\infty} \mathbb{E}[Z_\ell] \cdot \mathbb{P}(\{\text{first success on } \ell\text{th attempt}\}) \\
&= \sum_{\ell=1}^{\infty} \mathbb{E}[Z_\ell] \cdot \mathbb{P}(X = \ell)
\end{aligned}$$

Plugging in the expressions we got in **Part (b)** and **Part (e)**, we get

$$\begin{aligned}
\mathbb{E}[W] &= \sum_{\ell=1}^{\infty} \left((\ell - 1) \cdot \left(\frac{p + 2p^2}{1 + p + p^2}\right) + 3\right) \cdot \left((1 - p^3)^{\ell-1} p^3\right) \\
&= \left(\left(\frac{p + 2p^2}{1 + p + p^2}\right) \cdot \sum_{\ell=1}^{\infty} (\ell - 1) \cdot (1 - p^3)^{\ell-1} p^3\right) + \left(3 \cdot \sum_{\ell=1}^{\infty} (1 - p^3)^{\ell-1} p^3\right)
\end{aligned}$$

Let $V \sim \text{Geometric}(p^3)$. Then by the definition of $\mathbb{P}$ and $\mathbb{E}$, we have

$$\begin{aligned}
1 &= \sum_{\ell=1}^{\infty} \mathbb{P}(V = \ell) \\
&= \sum_{\ell=1}^{\infty} (1 - p^3)^{\ell-1} p^3 \\
\mathbb{E}[V] &= \sum_{\ell=1}^{\infty} \ell \cdot \mathbb{P}(V = \ell)
\end{aligned}$$

Question assigned to the following page: 2

$$= \sum_{\ell=1}^{\infty} \ell \cdot (1-p^3)^{\ell-1} p^3$$

If we let $g(x) = x - 1$, we get

$$\begin{aligned}
\mathbb{E}[V] - 1 &= \mathbb{E}[V] - \mathbb{E}[1] \\
&= \mathbb{E}[V-1] && \text{Linearity of expectation} \\
&= \mathbb{E}[g(V)] \\
&= \sum_{\ell=1}^{\infty} g(\ell) \cdot \mathbb{P}(V = \ell) \\
&= \sum_{\ell=1}^{\infty} g(\ell) \cdot (1-p^3)^{\ell-1} p^3 \\
&= \sum_{\ell=1}^{\infty} (\ell - 1) \cdot (1-p^3)^{\ell-1} p^3
\end{aligned}$$

From **Homework 5, Problem 2, Part (a)**, we know that

$$\mathbb{E}[V] = \frac{1}{p^3}$$

so we get

$$\begin{aligned}
\sum_{\ell=1}^{\infty} (\ell - 1) \cdot (1-p^3)^{\ell-1} p^3 &= \mathbb{E}[V] - 1 \\
&= \frac{1}{p^3} - 1
\end{aligned}$$

Returning to our original expression for $\mathbb{E}[W]$, we had

$$\begin{aligned}
\mathbb{E}[W] &= \left(\left(\frac{p + 2p^2}{1 + p + p^2} \right) \cdot \sum_{\ell=1}^{\infty} (\ell - 1) \cdot (1-p^3)^{\ell-1} p^3 \right) \\
&\quad + \left(3 \cdot \sum_{\ell=1}^{\infty} (1-p^3)^{\ell-1} p^3 \right) \\
&= \left(\left(\frac{p + 2p^2}{1 + p + p^2} \right) \cdot \sum_{\ell=1}^{\infty} (\ell - 1) \cdot (1-p^3)^{\ell-1} p^3 \right) + 3(1) && \text{Sum over PMF of } V \\
&= \left(\frac{p + 2p^2}{1 + p + p^2} \right) \cdot \left(\frac{1}{p^3} - 1 \right) + 3 && \text{Substitute in } \mathbb{E}[V-1] \text{ above} \\
&= \left(\frac{(p + 2p^2)(1-p)}{1 - p^3} \right) \cdot \left(\frac{1 - p^3}{p^3} \right) + 3
\end{aligned}$$

Questions assigned to the following page: 2 and 3

$$
\begin{aligned}
&= (p + 2p^2)(1 - p) \cdot \left(\frac{1}{p^3}\right) + 3 \\
&= \left(\frac{(p + 2p^2)(1 - p)}{p^3}\right) + 3 \\
&= \left(\frac{p(1 + 2p)(1 - p)}{p^3}\right) + 3 \\
&= \left(\frac{(1 + 2p)(1 - p)}{p^2}\right) + 3 \\
&= \left(\frac{1 + 2p - p - 2p^2}{p^2}\right) + 3
\end{aligned}
$$

which simplifies to

$$\boxed{\mathbb{E}[W] = \left(\frac{1 + p - 2p^2}{p^2}\right) + 3}$$

Problem 3

Suppose $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ are statistically independent. Show that $X + Y \sim \text{Poisson}(\mu + \lambda)$. Hint: Use the binomial theorem.

We want to show that $Z = X + Y \sim \text{Poisson}(\mu + \lambda)$. It suffices to show that the PMF of Z is equal to the PMF of a $\text{Poisson}(\mu + \lambda)$ random variable. We have, for $k \geq 0$:

$$
\begin{aligned}
p_Z(k) &= \mathbb{P}(Z = k) \\
&= \sum_{i=0}^{k} \mathbb{P}(X = i \,\cap\, Y = k - i) && X \text{ and } Y \text{ are } \mathbb{N}\text{-valued} \\
&= \sum_{i=0}^{k} \mathbb{P}(X = i) \cdot \mathbb{P}(Y = k - i) && \text{Independence of } X \text{ and } Y \\
&= \sum_{i=0}^{k} \left(\frac{\lambda^i}{i!} e^{-\lambda}\right) \cdot \left(\frac{\mu^{k-i}}{(k-i)!} e^{-\mu}\right) && \text{Definition of Poisson distribution} \\
&= e^{-\lambda} e^{-\mu} \cdot \sum_{i=0}^{k} \left(\frac{\lambda^i}{i!}\right) \cdot \left(\frac{\mu^{k-i}}{(k-i)!}\right) && \text{Factor out terms which don't depend on } i \\
&= e^{-\lambda} e^{-\mu} \cdot \sum_{i=0}^{k} \left(\lambda^i \mu^{k-i}\right) \cdot \left(\frac{1}{i!(k-i)!}\right) \\
&= e^{-\lambda} e^{-\mu} \cdot \sum_{i=0}^{k} \left(\lambda^i \mu^{k-i}\right) \cdot \left(\frac{k!}{i!(k-i)!} \cdot \frac{1}{k!}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-\lambda}e^{-\mu} \cdot \sum_{i=0}^{k} \left(\lambda^i \mu^{k-i}\right) \cdot \binom{k}{i} \cdot \frac{1}{k!} \\
&= \frac{e^{-\lambda}e^{-\mu}}{k!} \cdot \sum_{i=0}^{k} \binom{k}{i} \cdot \left(\lambda^i \mu^{k-i}\right) \\
&= \frac{e^{-\lambda}e^{-\mu}}{k!} \cdot (\lambda + \mu)^k && \text{Binomial theorem} \\
&= \frac{(\lambda + \mu)^k}{k!} \cdot e^{-(\lambda+\mu)}
\end{aligned}
$$

Note that the final expression above is exactly the PMF for a Poisson($\lambda + \mu$) random variable for $k \geq 0$, so hence $Z = X + Y$ is indeed Poisson($\lambda + \mu$) distributed, which is what we wanted to show.

Problem 4

We have a physical process that produces independent identically distributed random variables X_j whose distribution is unknown to us. We wish to find the true mean μ via experimentation. We may assume that $\mathbb{E}(X_i^2) < \infty$.

Part (a)

Whose inequality shows $\mathbb{E}(|X_i|) < \infty$? Elaborate.

It is $\boxed{\text{Jensen's inequality}}$ which shows $\mathbb{E}(|X_i|) < \infty$. More precisely, we are told that $\mathbb{E}[X_i^2] < \infty$. Since squares of real numbers are nonnegative, $X_i^2 = |X_i|^2$ and we can write $\mathbb{E}[X_i^2] = \mathbb{E}[|X_i|^2] < \infty$ also. Since $g(x) = x^2$ is a convex function, we can then apply Jensen's inequality to conclude that

$$
\begin{gathered}
g(\mathbb{E}[|X_i|]) \leq \mathbb{E}[g(|X_i|)] \\
\mathbb{E}[|X_i|]^2 \leq \mathbb{E}[|X_i|^2]
\end{gathered}
$$

Thus we have $\mathbb{E}[|X_i|]^2 \leq \mathbb{E}[|X_i|^2] < \infty$, and so $\mathbb{E}[|X_i|]^2 < \infty$, i.e. $\mathbb{E}[|X_i|]^2$ is finite. Thus, its square root exists (and is finite), and $\mathbb{E}[|X_i|] < \infty$ also. Since the absolute expectation of X_i exists, by the definition of expectation, the raw expectation $\mathbb{E}[X_i]$ must also exist, i.e. $\mathbb{E}[X_i] < \infty$.

Part (b)

Show that the sample mean

$$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$$

is an unbiased estimate of the true mean.

To show that the sample mean $\bar{X}$ is an unbiased estimate of the true mean, we need to show that the expected value of the sample mean is equal to the true mean, i.e.

Question assigned to the following page: 4

$$\mathbb{E}[\bar{X}] = \mu = \mathbb{E}[X_i]$$

Expanding out the definition of $\bar{X}$, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n}(X_1 + \cdots + X_n)\right] &= \frac{1}{n}\mathbb{E}[X_1 + \cdots + X_n] && \text{Linearity of expectation} \\
&= \frac{1}{n}(\mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}\mathbb{E}[X_i]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}\mathbb{E}[X_1]\right) && \text{Variables are identically distributed} \\
&= \frac{1}{n}(n\mathbb{E}[X_1]) \\
&= \mathbb{E}[X_1] \\
&= \mu && \text{Definition of true mean}
\end{aligned}$$

Since $\mathbb{E}[\bar{X}] = \mu$, the sample mean $\bar{X}$ is indeed an unbiased estimate of the true mean.

Part (c)

Determine $\mathrm{Var}(\bar{X})$ in terms of $\mathrm{Var}(X_i)$.

We proceed by expanding out the definitions:

$$\begin{aligned}
\mathrm{Var}(\bar{X}) &= \mathbb{E}[(\bar{X} - \mathbb{E}[\bar{X}])^2] \\
&= \mathbb{E}[\bar{X}^2] - \mathbb{E}[\bar{X}]^2 && \text{Variance formula in terms of } \mathbb{E} \\
&= \mathbb{E}[\bar{X}^2] - \mu^2 && \text{From } \textbf{Part (b)} \\
&= \mathbb{E}\left[\left(\frac{1}{n}\cdot(X_1 + \cdots + X_n)\right)^2\right] - \mu^2 && \text{Definition of } \bar{X} \\
&= \mathbb{E}\left[\frac{1}{n^2}\cdot(X_1 + \cdots + X_n)^2\right] - \mu^2 \\
&= \frac{1}{n^2}\cdot\mathbb{E}\left[(X_1 + \cdots + X_n)^2\right] - \mu^2 \\
&= \frac{1}{n^2}\cdot\mathbb{E}\left[\left(\sum_{i=1}^{n}X_i^2\right) + \left(2\cdot\sum_{1\le j<k\le n}X_jX_k\right)\right] - \mu^2 && \text{Expand multinomial} \\
&= \frac{1}{n^2}\cdot\mathbb{E}\left[\sum_{i=1}^{n}X_i^2\right] + \frac{2}{n^2}\cdot\mathbb{E}\left[\sum_{1\le j<k\le n}X_jX_k\right] - \mu^2 && \text{Linearity of expectation}
\end{aligned}$$

Question assigned to the following page: 4

$$
\begin{aligned}
&= \left(\frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right]\right) + \left(\frac{2}{n^2} \cdot \sum_{1 \le j < k \le n} \mathbb{E}\left[X_j X_k\right]\right) - \mu^2 \\
&= \left(\frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right]\right) + \left(\frac{2}{n^2} \cdot \sum_{1 \le j < k \le n} \mathbb{E}\left[X_j\right] \mathbb{E}\left[X_k\right]\right) - \mu^2 && \text{Variables are independent} \\
&= \left(\frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{2}{n^2} \cdot \sum_{1 \le j < k \le n} \mathbb{E}\left[X_1\right] \mathbb{E}\left[X_1\right]\right) - \mu^2 && \text{Variables are identically distributed} \\
&= \left(\frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{2}{n^2} \cdot \sum_{1 \le j < k \le n} \mathbb{E}\left[X_1\right]^2\right) - \mu^2 \\
&= \left(\frac{1}{n^2} \cdot n \cdot \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{2}{n^2} \cdot \binom{n}{2} \cdot \mathbb{E}\left[X_1\right]^2\right) - \mu^2 \\
&= \left(\frac{1}{n} \cdot \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{2}{n^2} \cdot \frac{n(n-1)}{2} \cdot \mathbb{E}\left[X_1\right]^2\right) - \mu^2 \\
&= \left(\frac{1}{n} \cdot \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{n-1}{n} \cdot \mathbb{E}\left[X_1\right]^2\right) - \mu^2
\end{aligned}
$$

Note that we also know that

$$
\begin{aligned}
\mathrm{Var}(X_i) &= \mathbb{E}\left[(X_i - \mathbb{E}[X_i])^2\right] \\
&= \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2
\end{aligned}
$$

for all $1 \le i \le n$. Rearranging, we get

$$\mathrm{Var}(X_i) + \mathbb{E}[X_i]^2 = \mathbb{E}[X_i^2]$$

and more specifically,

$$\mathbb{E}[X_1^2] = \mathrm{Var}(X_1) + \mathbb{E}[X_1]^2$$

Substituting this expression for $\mathbb{E}[X_1^2]$ in our first computation, we get

$$
\begin{aligned}
\mathrm{Var}(\bar{X}) &= \left(\frac{1}{n} \cdot \mathbb{E}\left[X_1^2\right]\right) + \left(\frac{n-1}{n} \cdot \mathbb{E}\left[X_1\right]^2\right) - \mu^2 \\
&= \frac{1}{n} \cdot \left(\mathrm{Var}(X_1) + \mathbb{E}[X_1]^2\right) + \left(\frac{n-1}{n} \cdot \mathbb{E}\left[X_1\right]^2\right) - \mu^2 \\
&= \frac{1}{n} \cdot \left(\mathrm{Var}(X_1) + \mu^2\right) + \left(\frac{n-1}{n} \cdot \mu^2\right) - \mu^2 \\
&= \frac{1}{n} \cdot \left(\mathrm{Var}(X_1) + \mu^2\right) + \left(\frac{(n-1)\mu^2}{n}\right) - \mu^2
\end{aligned}
$$

$$\begin{aligned}
&= \frac{\text{Var}(X_1) + \mu^2}{n} + \frac{(n-1)\mu^2}{n} - \frac{n\mu^2}{n} \\
&= \frac{\text{Var}(X_1) + \mu^2}{n} + \frac{n\mu^2 - \mu^2}{n} - \frac{n\mu^2}{n} \\
&= \frac{\text{Var}(X_1) + \mu^2 + n\mu^2 - \mu^2 - n\mu^2}{n} \\
&= \frac{\text{Var}(X_1)}{n}
\end{aligned}$$

Since the variables X_i are all identically distributed, we can write

$$\boxed{\text{Var}(\bar{X}) = \frac{\text{Var}(X_i)}{n}}$$

Part (d)

Show that if the mystery distribution is Poisson(λ), albeit with λ unknown, then $\bar{X}$ is the maximum likelihood estimator for $\mu = \lambda$.

To show that $\bar{X}$ is the maximum likelihood estimator for $\mu = \lambda$ (the unknown parameter of our Poisson distribution), we need to show that $\lambda = \bar{X}$ maximizes the likelihood of our observations of $X_1, ..., X_n$.

Formally, let $\omega = (x_1, ..., x_n)$ be an arbitrary observation of X_1, ..., X_n respectively. We want to show that a Poisson parameter of $\lambda = \bar{x} = \bar{X}(\omega)$ maximizes the likelihood of this observation. Let $\mathbb{P}_\lambda(\cdot)$ denote the probability law of a Poisson(λ) distributed random variable with parameter λ. Given our observations, we have a likelihood function

$$L(\lambda) = \mathbb{P}_\lambda(x_1) \cdots \mathbb{P}_\lambda(x_n)$$

where we are able to calculate the joint probability of all the observations $(x_1, ..., x_n)$ as a product since the variables are independently distributed. Expanding the definition of the Poisson(λ) probability law, we have

$$\begin{aligned}
L(\lambda) &= \mathbb{P}_\lambda(x_1) \cdots \mathbb{P}_\lambda(x_n) \\
&= \left(\frac{\lambda^{x_1}}{x_1!} e^{-\lambda}\right) \cdots \left(\frac{\lambda^{x_n}}{x_n!} e^{-\lambda}\right) \\
&= \frac{\lambda^{x_1} \cdots \lambda^{x_n}}{x_1! \cdots x_n!} \cdot (e^{-\lambda})^n \\
&= \frac{\lambda^{x_1 + \cdots + x_n}}{x_1! \cdots x_n!} \cdot e^{-n\lambda} \\
&= \frac{1}{x_1! \cdots x_n!} \cdot \lambda^{x_1 + \cdots + x_n} e^{-n\lambda}
\end{aligned}$$

We want to find the λ that maximizes our likelihood function. To do so, we can take the derivative and solve for the values of λ which make the derivative zero. We have

Question assigned to the following page: 4

$$\begin{aligned}
\frac{\partial}{\partial\lambda}L(\lambda) &= \frac{\partial}{\partial\lambda}\left(\frac{1}{x_1!\cdots x_n!}\cdot\lambda^{x_1+\cdots+x_n}e^{-n\lambda}\right)\\
&= \frac{1}{x_1!\cdots x_n!}\cdot\frac{\partial}{\partial\lambda}\left(\lambda^{x_1+\cdots+x_n}e^{-n\lambda}\right)\\
&= \frac{1}{x_1!\cdots x_n!}\cdot\left(\lambda^{x_1+\cdots+x_n}\cdot -ne^{-n\lambda}+e^{-n\lambda}\cdot(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1}\right) \quad \text{Product rule}
\end{aligned}$$

Solving for λ when the derivative is zero, we have

$$\begin{aligned}
\frac{1}{x_1!\cdots x_n!}\cdot\left(\lambda^{x_1+\cdots+x_n}\cdot -ne^{-n\lambda}+e^{-n\lambda}\cdot(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1}\right) &= 0\\
\lambda^{x_1+\cdots+x_n}\cdot -ne^{-n\lambda}+e^{-n\lambda}\cdot(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1} &= 0\\
e^{-n\lambda}(-n\lambda^{x_1+\cdots+x_n}+(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1}) &= 0\\
-n\lambda^{x_1+\cdots+x_n}+(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1} &= 0\\
\lambda^{x_1+\cdots+x_n}(-n+(x_1+\cdots+x_n)\lambda^{-1}) &= 0\\
-n\lambda^{x_1+\cdots+x_n}+(x_1+\cdots+x_n)\lambda^{x_1+\cdots+x_n-1} &= 0\\
-n+(x_1+\cdots+x_n)\lambda^{-1} &= 0
\end{aligned}$$

Note that $\lambda \neq 0$ by the definition of a Poisson distribution, so we can safely divide both sides by λ. Continuing to solve for λ, we have

$$\begin{aligned}
-n+(x_1+\cdots+x_n)\lambda^{-1} &= 0(x_1+\cdots+x_n)\lambda^{-1} & &= n\\
x_1+\cdots+x_n &= n\lambda\\
\frac{x_1+\cdots+x_n}{n} &= \lambda\\
\bar{x} &= \lambda
\end{aligned}$$

That is, $\bar{x}$ is a maxima or minima of λ. To determine what type of critical point $\bar{x}$ is, we can investigate what occurs when we perturb λ to be slightly larger or smaller. We have

$$\begin{aligned}
L(\lambda) &= \frac{1}{x_1!\cdots x_n!}\cdot\lambda^{x_1+\cdots+x_n}e^{-n\lambda}\\
&= \frac{1}{x_1!\cdots x_n!}\cdot(e^{\log(\lambda)})^{x_1+\cdots+x_n}e^{-n\lambda}\\
&= \frac{1}{x_1!\cdots x_n!}\cdot e^{\log(\lambda)(x_1+\cdots+x_n)}e^{-n\lambda}\\
&= \frac{1}{x_1!\cdots x_n!}\cdot e^{\log(\lambda)(x_1+\cdots+x_n)-n\lambda}
\end{aligned}$$

Since e^x is monotonically increasing, this reduces to determining what happens to the exponent $\log(\lambda)(x_1+\cdots+x_n)-n\lambda$ when we perturb λ. If the expression gets smaller for $\pm\epsilon$, then $\bar{x}$ is a maximum. It suffices to run the second-derivative test on this exponent

Question assigned to the following page: 4

$$M(\lambda) = \log(\lambda)(x_1 + \cdots + x_n) - n\lambda$$

We have

$$\begin{aligned}
\frac{\partial}{\partial\lambda} M(\lambda) &= \frac{\partial}{\partial\lambda}\left(\log(\lambda)(x_1 + \cdots + x_n) - n\lambda\right) \\
&= \frac{\partial}{\partial\lambda}\left(\log(\lambda)(x_1 + \cdots + x_n)\right) - \frac{\partial}{\partial\lambda}(n\lambda) \\
&= (x_1 + \cdots + x_n)\frac{\partial}{\partial\lambda}\log(\lambda) - n\frac{\partial}{\partial\lambda}\lambda \\
&= (x_1 + \cdots + x_n)\cdot\frac{1}{\lambda} - n
\end{aligned}$$

with second derivative

$$\begin{aligned}
\frac{\partial^2}{\partial^2\lambda} M(\lambda) &= \frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}M(\lambda)\right) \\
&= \frac{\partial}{\partial\lambda}\left((x_1 + \cdots + x_n)\cdot\frac{1}{\lambda} - n\right) \\
&= \frac{\partial}{\partial\lambda}\left((x_1 + \cdots + x_n)\cdot\frac{1}{\lambda}\right) - \frac{\partial}{\partial\lambda}(n) \\
&= (x_1 + \cdots + x_n)\cdot\frac{\partial}{\partial\lambda}\left(\frac{1}{\lambda}\right) \\
&= (x_1 + \cdots + x_n)\cdot\left(-\frac{1}{\lambda^2}\right)
\end{aligned}$$

Plugging in $\lambda = \bar{x} = \frac{x_1 + \cdots x_n}{n}$, we get

$$\begin{aligned}
\frac{\partial^2}{\partial^2\lambda} M(\lambda)\mid_{\lambda=\bar{x}} &= (x_1 + \cdots + x_n)\cdot\left(-\frac{1}{\bar{x}^2}\right) \\
&= (x_1 + \cdots + x_n)\cdot -\left(\frac{1}{\bar{x}}\right)^2 \\
&= (x_1 + \cdots + x_n)\cdot -\left(\frac{n}{x_1 + \cdots x_n}\right)^2 \\
&= (x_1 + \cdots + x_n)\cdot\frac{-n^2}{(x_1 + \cdots x_n)^2} \\
&= \frac{-n^2}{(x_1 + \cdots x_n)}
\end{aligned}$$

Since the x_i must always be positive (by the definition of a Poisson distribution — the probability of observing a negative number of events is 0), and $-n^2$ is always negative, the second derivative

Questions assigned to the following page: 4 and 5

will be negative at $\bar{x}$. Since $\bar{x}$ is a critical point of our likelihood function $L(\lambda)$ with negative second derivative, it must be a maximum. Hence, the function $\bar{X}$ is indeed a maximum-likelihood estimator for λ.

Problem 5

(Continued from previous problem.) We may also wish to know the (common) variance of our random variables X_i, which we may estimate using

$$Z = \frac{1}{n}\left([X_1 - \bar{X}]^2 + \cdots + [X_n - \bar{X}]^2\right)$$

or, if by some miracle we knew the true mean μ of the distribution

$$\tilde{Z} = \frac{1}{n}\left([X_1 - \mu]^2 + \cdots + [X_n - \mu]^2\right)$$

Part (a)

Prove that $Z \leq \tilde{Z}$ for every elementary outcome.

It suffices to show that $\tilde{Z} - Z \geq 0$. We have

$$\begin{aligned}
\tilde{Z} - Z &= \frac{1}{n}\left([X_1 - \mu]^2 + \cdots + [X_n - \mu]^2\right) - \frac{1}{n}\left([X_1 - \bar{X}]^2 + \cdots + [X_n - \bar{X}]^2\right) \\
&= \frac{1}{n}\left([X_1 - \mu]^2 + \cdots + [X_n - \mu]^2 - [X_1 - \bar{X}]^2 - \cdots - [X_n - \bar{X}]^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i - \mu]^2 - \sum_{i=1}^{n}[X_i - \bar{X}]^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i^2 - 2X_i\mu + \mu^2] - \sum_{i=1}^{n}[X_i^2 - 2X_i\bar{X} + \bar{X}^2]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[-2X_i\mu + \mu^2] - \sum_{i=1}^{n}[-2X_i\bar{X} + \bar{X}^2]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[-2X_i\mu + 2X_i\bar{X}] + \sum_{i=1}^{n}[\mu^2 - \bar{X}^2]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[2X_i\bar{X} - 2X_i\mu] + \sum_{i=1}^{n}[\mu^2 - \bar{X}^2]\right) \\
&= \frac{1}{n}\left(2 \cdot \sum_{i=1}^{n}[X_i(\bar{X} - \mu)] + \sum_{i=1}^{n}[\mu^2 - \bar{X}^2]\right) \\
&= \frac{1}{n}\left(2(\bar{X} - \mu) \cdot \sum_{i=1}^{n}[X_i] + \sum_{i=1}^{n}[\mu^2 - \bar{X}^2]\right)
\end{aligned}$$

Question assigned to the following page: 5

$$
\begin{aligned}
&= \frac{1}{n}\left(2(\bar{X}-\mu)\cdot\sum_{i=1}^{n}[X_i] + n[\mu^2-\bar{X}^2]\right) \\
&= 2(\bar{X}-\mu)\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}[X_i] + [\mu^2-\bar{X}^2] \\
&= 2(\bar{X}-\mu)\cdot\bar{X} + [\mu^2-\bar{X}^2] \\
&= 2\bar{X}^2 - 2\mu\bar{X} + \mu^2 - \bar{X}^2 \\
&= \bar{X}^2 - 2\mu\bar{X} + \mu^2 \\
&= (\bar{X}-\mu)^2 \\
&\geq 0 && \text{Squares are nonnegative}
\end{aligned}
$$

Since $\tilde{Z} - Z \geq 0$, we have $\tilde{Z} \geq Z$, which is what we wanted to show.

Part (b)(1)

Prove that

$$Z = \left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right) - \bar{X}^2$$

By definition, we have

$$
\begin{aligned}
Z &= \frac{1}{n}\left([X_1-\bar{X}]^2 + \cdots + [X_n-\bar{X}]^2\right) \\
&= \frac{1}{n}\left([X_1^2 - 2X_1\bar{X} + \bar{X}^2] + \cdots + [X_n^2 - 2X_n\bar{X} + \bar{X}^2]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i^2] - 2\bar{X}\sum_{i=1}^{n}[X_i] + \sum_{i=1}^{n}[\bar{X}^2]\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i^2] - 2\bar{X}(n\bar{X}) + \sum_{i=1}^{n}[\bar{X}^2]\right) && \text{Definition of } \bar{X} \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i^2] - 2n\bar{X}^2 + n\bar{X}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n}[X_i^2] - n\bar{X}^2\right) \\
&= \left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right) - \bar{X}^2
\end{aligned}
$$

which is what we wanted to show.

Question assigned to the following page: 5

Nathan Leung, UID 005835316
MATH 170A Homework 6
16 February 2024

Part (b)(2)

Find $\mathbb{E}[Z]$ and $\mathbb{E}[\tilde{Z}]$ in terms of $\mathrm{Var}(X_i)$.

From **Part (b)**, we know that

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - \bar{X}^2\right] \\ &= \frac{1}{n}\left(\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right]\right) - \mathbb{E}[\bar{X}^2] && \text{Linearity of expectation} \end{aligned}$$

By the definition of variance, we know that

$$\begin{aligned} \mathbb{E}[X_i^2] &= \mathrm{Var}(X_i) + \mathbb{E}[X_i]^2 \\ \mathbb{E}[\bar{X}^2] &= \mathrm{Var}(\bar{X}) + \mathbb{E}[\bar{X}]^2 \end{aligned}$$

From **Problem 4, Part (c)**, we also know that

$$\mathrm{Var}(\bar{X}) = \frac{\mathrm{Var}(X_i)}{n}$$

so we also have

$$\begin{aligned} \mathbb{E}[\bar{X}^2] &= \frac{\mathrm{Var}(X_i)}{n} + \mathbb{E}[\bar{X}]^2 \\ &= \frac{\mathrm{Var}(X_i)}{n} + \mu^2 \end{aligned}$$

Making these substitutions into the expression we have above, we get

$$\begin{aligned} \mathbb{E}[Z] &= \frac{1}{n}\left(\sum_{j=1}^{n} \mathbb{E}\left[X_j^2\right]\right) - \mathbb{E}[\bar{X}^2] \\ &= \frac{1}{n}\left(\sum_{j=1}^{n} \left[\mathrm{Var}(X_j) + \mathbb{E}[X_j]^2\right]\right) - [\frac{\mathrm{Var}(X_i)}{n} + \mu^2] \\ &= \frac{1}{n}\left(\sum_{j=1}^{n} [\mathrm{Var}(X_1) + \mathbb{E}[X_1]^2]\right) - \left[\frac{\mathrm{Var}(X_i)}{n} + \mu^2\right] && \text{Identically distributed} \\ &= \frac{1}{n} \cdot n \cdot [\mathrm{Var}(X_1) + \mathbb{E}[X_1]^2] - \left[\frac{\mathrm{Var}(X_i)}{n} + \mu^2\right] \\ &= [\mathrm{Var}(X_1) + \mathbb{E}[X_1]^2] - \left[\frac{\mathrm{Var}(X_i)}{n} + \mu^2\right] \end{aligned}$$

Question assigned to the following page: 5

$$
\begin{aligned}
&= \mathrm{Var}(X_i) + \mathbb{E}[X_i]^2 - \frac{\mathrm{Var}(X_i)}{n} - \mu^2 && \text{Indentically distributed} \\
&= \mathrm{Var}(X_i) + \mu^2 - \frac{\mathrm{Var}(X_i)}{n} - \mu^2 \\
&= \mathrm{Var}(X_i) - \frac{\mathrm{Var}(X_i)}{n} \\
&= \frac{n\mathrm{Var}(X_i)}{n} - \frac{\mathrm{Var}(X_i)}{n} \\
&= \frac{(n-1)\mathrm{Var}(X_i)}{n}
\end{aligned}
$$

which simplifies to

$$\boxed{\mathbb{E}[Z] = \frac{n-1}{n} \cdot \mathrm{Var}(X_i)}$$

For $\mathbb{E}[\tilde{Z}]$, we have

$$
\begin{aligned}
\mathbb{E}[\tilde{Z}] &= \mathbb{E}\left[\frac{1}{n} \cdot ((X_1 - \mu)^2 + \cdots + (X_n - \mu)^2)\right] \\
&= \frac{1}{n} \cdot \mathbb{E}\left[(X_1^2 - 2X_1\mu + \mu^2) + \cdots + (X_n^2 - 2X_\mu + \mu^2)\right] \\
&= \frac{1}{n} \cdot \mathbb{E}\left[\sum_{j=1}^{n}[X_j^2] - 2\mu\sum_{j=1}^{n}[X_j] + n\mu^2\right] \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\mathbb{E}\left[X_j^2\right] - 2\mu\sum_{j=1}^{n}\mathbb{E}[X_j] + n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}[\mathrm{Var}(X_j) + \mathbb{E}[X_j]^2] - 2\mu\sum_{j=1}^{n}\mathbb{E}[X_j] + n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}[\mathrm{Var}(X_j) + \mu^2] - 2\mu\sum_{j=1}^{n}\mathbb{E}[X_j] + n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\mathrm{Var}(X_j) + n\mu^2 - 2\mu\sum_{j=1}^{n}\mathbb{E}[X_j] + n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\mathrm{Var}(X_j) + n\mu^2 - 2\mu\sum_{j=1}^{n}\mu + n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{j=1}^{n}\mathrm{Var}(X_j) + n\mu^2 - 2n\mu^2 + n\mu^2\right)
\end{aligned}
$$

Question assigned to the following page: 5

$$
\begin{aligned}
&= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Var}(X_j) \\
&= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Var}(X_1) && \text{Identically distributed} \\
&= \frac{1}{n}\cdot n\mathrm{Var}(X_1) \\
&= \mathrm{Var}(X_1)
\end{aligned}
$$

Since the variables are identically distributed, we can conclude that

$$\boxed{\mathbb{E}[\tilde{Z}] = \mathrm{Var}(X_i)}$$

Part (c)

The idea of using $\frac{n}{n-1}Z$ as an estimator for the true variance is known as the Bessel correction. Explain how this leads to bias (systematic underestimation) in determining the standard deviation.

In **Part (b)(2)**, we saw that using $\frac{n}{n-1}Z$ is unbiased with respect to the true variance, because

$$
\begin{aligned}
\mathbb{E}\left[\frac{n}{n-1}Z\right] &= \frac{n}{n-1}\mathbb{E}[Z] \\
&= \frac{n}{n-1}\left(\frac{n-1}{n}\cdot \mathrm{Var}(X_i)\right) && \text{From } \textbf{Part (b)(2)} \\
&= \mathrm{Var}(X_i)
\end{aligned}
$$

However, we run into problems if we use

$$\sqrt{\frac{n}{n-1}Z}$$

to estimate the standard deviation. First, note that $g(x) = -\sqrt{x}$ is a convex function over positive numbers. By the second derivative test, we have

$$
\begin{aligned}
\frac{d^2}{dx^2}\sqrt{x} &= \frac{d^2}{dx^2}x^{\frac{1}{2}} \\
&= \frac{d}{dx}\left(\frac{d}{dx}x^{\frac{1}{2}}\right) \\
&= \frac{d}{dx}\left(\frac{1}{2}x^{-\frac{1}{2}}\right) \\
&= -\frac{1}{4}x^{-\frac{3}{2}} \\
&= -\frac{1}{4x^{\frac{3}{2}}}
\end{aligned}
$$

---

which is always negative for positive x; hence, $g(x)$ is convex. Thus, by Jensen's inequality, we have

$$\begin{aligned} g\left(\mathbb{E}\left[\frac{n}{n-1}Z\right]\right) &\leq \mathbb{E}\left[g\left(\frac{n}{n-1}Z\right)\right] \\ -\sqrt{\mathbb{E}\left[\frac{n}{n-1}Z\right]} &\leq \mathbb{E}\left[-\sqrt{\frac{n}{n-1}}Z\right] \\ -\sqrt{\mathrm{Var}(X_i)} &\leq -\mathbb{E}\left[\sqrt{\frac{n}{n-1}}Z\right] \end{aligned}$$

Multiplying both sides and flipping the direction of the inequality, we get

$$\sqrt{\mathrm{Var}(X_i)} \geq \mathbb{E}\left[\sqrt{\frac{n}{n-1}}Z\right]$$

i.e. if we use $\sqrt{\frac{n}{n-1}Z}$ to estimate the true standard deviation $\sqrt{\mathrm{Var}(X_i)}$, the value we get will, on average, be less than or equal to the true value — never greater — and hence we will systematically underestimate it.