1 Planetary Party

(a) Suppose we are at party on a planet where every year is 2849 days. If 30 people attend this party, what is the exact probability that two people will share the same birthday? You may leave your answer as an unevaluated expression.

(b) From lecture, we know that given n bins and m balls, $\mathbb{P}\left[\text{no collision}\right] \approx \exp(-m^2/(2n))$. Using this, give an approximation for the probability in part (a).

(c) What is the minimum number of people that need to attend this party to ensure that the probability that any two people share a birthday is at least 0.5? You can use the approximation you used in the previous part.

(d) Now suppose that 70 people attend this party. What the is probability that none of these 70 individuals have the same birthday? You can use the approximation you used in the previous parts.

Solution:

(a) Let's compute the probability that no two partygoers have the same birthday. We know the second person at the party cannot share a birthday with the first person, the third person at the party cannot share a birthday with the first two, etc. Thus

$$\mathbb{P}\left[\text{no collision}\right] = \left(1 - \frac{1}{2849}\right)\left(1 - \frac{2}{2849}\right)\left(1 - \frac{3}{2849}\right)\cdots\left(1 - \frac{29}{2849}\right)$$

Thus $\mathbb{P}[\text{collision}] = 1 - \mathbb{P}[\text{no collision}] = 1 - \left(1 - \frac{1}{2849}\right)\left(1 - \frac{2}{2849}\right)\left(1 - \frac{3}{2849}\right)\cdots\left(1 - \frac{29}{2849}\right)$.

(b) From lecture, we know that given n bins and m balls, $\mathbb{P}\left[\text{no collision}\right] \approx \exp(-m^2/(2n))$. Therefore in this case, if we want to find the probability of collision, we must find $1 - \mathbb{P}\left[\text{no collision}\right]$.

$$\mathbb{P}\left[\text{no collision}\right] = \exp\left(-\frac{30^2}{2 \cdot 2849}\right) = 0.854$$

This means that there is a 0.146 chance that two people share the same birthday in the group of 30.

(c) Rephrasing the question in terms of balls and bins, we want to find the minimum number of balls (m) such that there is at least 0.5 probability of collision when we have $n = 2849$ bins, which is the same as at **most** 0.5 probability of **no** collisions.

$$
\begin{aligned}
\mathbb{P}[\text{no collisions}] \approx \exp\left(\frac{-m^2}{n}\right) &\leq 0.5 \\
\implies \frac{-m^2}{n} &\leq \ln 0.5 \\
\implies m &\geq \sqrt{(-2\ln 0.5)n} \\
&= 62.845
\end{aligned}
$$

Since m must be an integer which is at least 62.485, we need at least $\boxed{63}$ people at the party.

(d) Once again we need to find $\mathbb{P}\left[\text{no collisions}\right]$ given that $m = 70$.

$$\mathbb{P}\left[\text{no collision}\right] = \exp\left(-\frac{70^2}{2\cdot 2849}\right) = 0.423$$

There is about a 42% chance that 70 people don't share the same birthday.

2 Faulty Lightbulbs

Box 1 contains 1000 lightbulbs of which 10% are defective. Box 2 contains 2000 lightbulbs of which 5% are defective.

(a) Suppose a box is given to you at random and you randomly select a lightbulb from the box. If that lightbulb is defective, what is the probability you chose Box 1?

(b) Suppose now that a box is given to you at random and you randomly select two light- bulbs from the box. If both lightbulbs are defective, what is the probability that you chose from Box 1?

Solution:

(a) Let:

- D denote the event that the lightbulb we selected is defective.
- B_i denote the event that the lightbulb we selected is from Box i.

We wish to compute $\mathbb{P}[B_1 \mid D]$. Using Bayes' Rule we get:

$$
\begin{aligned}
\mathbb{P}[B_1 \mid D] &= \frac{\mathbb{P}[D \mid B_1]\cdot\mathbb{P}[B_1]}{\mathbb{P}[B_1]\cdot\mathbb{P}[D \mid B_1] + \mathbb{P}[B_2]\cdot\mathbb{P}[D \mid B_2]} \\
&= \frac{0.1\cdot 0.5}{0.5\cdot 0.1 + 0.5\cdot 0.05} \\
&= \frac{2}{3}
\end{aligned}
$$

(b) Let:

- D' denote the event that both the lightbulbs we selected are defective.
- B_i denote the event that the lightbulb we selected is from Box i.

We wish to compute $\mathbb{P}[B_1 \mid D']$. Using Bayes' Rule we get:

$$\begin{aligned}
\mathbb{P}[B_1 \mid D'] &= \frac{\mathbb{P}[D' \mid B_1] \cdot \mathbb{P}[B_1]}{\mathbb{P}[B_1] \cdot \mathbb{P}[D' \mid B_1] + \mathbb{P}[B_2] \cdot \mathbb{P}[D' \mid B_2]} \\
&= \frac{\frac{100}{1000} \cdot \frac{99}{999} \cdot 0.5}{0.5 \cdot \frac{100}{1000} \cdot \frac{99}{999} + 0.5 \cdot \frac{100}{2000} \cdot \frac{99}{1999}} \\
&\approx 0.8
\end{aligned}$$

3 (Un)conditional (In)equalities

Let us consider a sample space $\Omega = \{\omega_1, \ldots, \omega_N\}$ of size $N > 2$, and two probability functions $\mathbb{P}_1$ and $\mathbb{P}_2$ on it. That is, we have two probability spaces: $(\Omega, \mathbb{P}_1)$ and $(\Omega, \mathbb{P}_2)$.

(a) If for every subset $A \subset \Omega$ of size $|A| = 2$ and every outcome $\omega \in \Omega$ it is true that $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A)$, then is it necessarily true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$? That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ are equal conditional on events of size 2, are they equal unconditionally? (*Hint*: Remember that probabilities must add up to 1.)

(b) If for every subset $A \subset \Omega$ of size $|A| = k$, where k is some fixed element in $\{2, \ldots, N\}$, and every outcome $\omega \in \Omega$ it is true that $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A)$, then is it necessarily true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$?

For the following two parts, assume that $\Omega = \left\{(a_1, \ldots, a_k) \mid \sum_{j=1}^{k} a_j = n\right\}$ is the set of configurations of n balls into k labeled bins, and let $\mathbb{P}_1$ be the probabilities assigned to these configurations by throwing the balls independently one after another into the bins, and let $\mathbb{P}_2$ be the probabilities assigned to these configurations by uniformly sampling one of these configurations.

(c) Let A be the event that all n balls land in exactly one bin. What are $\mathbb{P}_1(\omega \mid A)$ and $\mathbb{P}_2(\omega \mid A)$ for any $\omega \in A$? How about $\omega \in \Omega \setminus A$? Is it true that $\mathbb{P}_1(\omega) = \mathbb{P}_2(\omega)$ for all $\omega \in \Omega$?

(d) For the special case of $n = 9$ and $k = 3$, please give two outcomes B and C, so that $\mathbb{P}_1(B) < \mathbb{P}_2(B)$ and $\mathbb{P}_1(C) > \mathbb{P}_2(C)$.

Solution:

(a) Yes, this is indeed true. To see why, let's take the subset $A = \{\omega_i, \omega_j\}$ for some $i, j \in \{1, \ldots, N\}$ and compute: For any $k \in \{1, 2\}$, we have $\mathbb{P}_k(\omega_i \mid A) = \frac{\mathbb{P}_k(\omega_i)}{\mathbb{P}_k(A)}$. Since this expression (by assumption) is the same for $k = 1$ and $k = 2$, we conclude that $\frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_2(\omega_i)} = \frac{\mathbb{P}_1(A)}{\mathbb{P}_2(A)}$. Repeating the reasoning for ω_j, we similarly find that $\frac{\mathbb{P}_1(\omega_j)}{\mathbb{P}_2(\omega_j)} = \frac{\mathbb{P}_1(A)}{\mathbb{P}_2(A)}$, and whence $\frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_1(\omega_j)} = \frac{\mathbb{P}_2(\omega_i)}{\mathbb{P}_2(\omega_j)}$. Since this is true for any $i, j \in \{1, \ldots, N\}$, we can sum over i to get

$$\frac{1}{\mathbb{P}_1(\omega_j)} = \sum_{i=1}^{N} \frac{\mathbb{P}_1(\omega_i)}{\mathbb{P}_1(\omega_j)} = \sum_{i=1}^{N} \frac{\mathbb{P}_2(\omega_i)}{\mathbb{P}_2(\omega_j)} = \frac{1}{\mathbb{P}_2(\omega_j)},$$

which shows that $\mathbb{P}_1(\omega_j) = \mathbb{P}_2(\omega_j)$ for all $j \in \{1, \ldots, N\}$.

(b) Yes, it indeed would. There are two ways of verifying this. The first one is to observe that if $A' \subset A$ and $\omega \in A'$, then $\mathbb{P}_1(\omega \mid A') = \mathbb{P}_1(\omega \mid A' \cap A) = \frac{\mathbb{P}_1(\omega \mid A)}{\mathbb{P}_1(A' \mid A)} = \frac{\mathbb{P}_2(\omega \mid A)}{\mathbb{P}_2(A' \mid A)} = \mathbb{P}_2(\omega \mid A')$, where the second equality follows from the product rule (Theorem 13.1): $\mathbb{P}_1(A) \cdot \mathbb{P}_1(A' \mid A) \cdot \mathbb{P}_1(\omega \mid A \cap A') = \mathbb{P}_1(\{\omega\} \cap A \cap A') = \mathbb{P}_1(\omega) = \mathbb{P}_1(A)\mathbb{P}_1(\omega \mid A)$. That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ coincide conditional on some event A, they also coincide conditional on any smaller event A'. In particular, if they coincide on all events of size k, they also coincide on all events of size 2, which we have already dealt with in part (a).

The second way to convince ourselves that part (b) is true, is to observe that none of the arguments used in part (a) really relied on A having size 2, and so the very same reasoning carries through for A of size k.

(c) There are exactly k outcomes in A (namely, $(n, 0, \ldots, 0), (0, n, 0, \ldots), \ldots, (0, \ldots, 0, n)$; i.e. each bin could be the full one), and all of them are equally likely under either $\mathbb{P}_1$ or $\mathbb{P}_2$. That is, if $\omega \in A$, then $\mathbb{P}_1(\omega) = \left(\frac{1}{k}\right)^n$, and $\mathbb{P}_2(\omega) = \left[\binom{n+k-1}{k-1}\right]^{-1}$. Consequently, for $\omega \in A$,

$$\mathbb{P}_1(\omega \mid A) = \frac{k^{-n}}{k \cdot k^{-n}} = \frac{1}{k} \qquad \mathbb{P}_2(\omega \mid A) = \frac{\binom{n+k-1}{k-1}^{-1}}{k \cdot \binom{n+k-1}{k-1}^{-1}} = \frac{1}{k}.$$

If $\omega \notin A$, then $\mathbb{P}_1(\omega \mid A) = \mathbb{P}_2(\omega \mid A) = 0$, and so $\mathbb{P}_1(\omega \mid A)$ and $\mathbb{P}_2(\omega \mid A)$ coincide for all $\omega \in \Omega$. This, however, does *not* imply that $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same! Indeed, when computing the probability of $\omega \in A$ above, we saw that $\mathbb{P}_1(\omega) \neq \mathbb{P}_2(\omega)$ (remember that the assumption of part (b) was that the conditional probabilities coincide for *all* events of size k, here we have only shown equality conditional on *one* such event).

(d) Intuitively, throwing balls independently one after another makes it much less likely that all balls stack up in one bin as opposed to spreading out more evenly. This suggests taking, e.g., $A = \{\text{all balls land in bin } 1\}$, whose probability we already computed in part (c). That is, to show that $\mathbb{P}_1(A) < \mathbb{P}_2(A)$, we need to show that $k^{-n} < \binom{n+k-1}{k-1}^{-1}$. Plugging in $k = 3$ and $n = 9$, we have

$$k^{-n} = 3^{-9} = 3^{-2} \cdot 3^{-3} \cdot 3^{-4} = \frac{1}{9 \cdot 27 \cdot 3^4} < \frac{1}{5 \cdot 11} = \frac{2}{11 \cdot 10} = \binom{11}{2}^{-1} = \binom{n+k-1}{k-1}^{-1},$$

as desired.

Conversely, the same reasoning suggests that evenly distributed balls are much more likely under $\mathbb{P}_1$ than under $\mathbb{P}_2$. And indeed, letting $B = \{\text{each bin has exactly three balls}\}$, we have

$$\mathbb{P}_1(B) = \binom{9}{3}\binom{6}{3}\binom{3}{3}\left(\frac{1}{3}\right)^9 = \frac{9!}{(3!)^3 \cdot 3^9} = \frac{7!}{3^{10}} > \frac{7\cdot 6\cdot 5\cdot 4\cdot 3}{3^6}\cdot\frac{2}{11\cdot 10} = \frac{7}{3}\cdot\frac{6\cdot 5}{3^3}\cdot\frac{4}{3}\mathbb{P}_2(B),$$

and since the last factor is bigger than 1, we have $\mathbb{P}_1(B) > \mathbb{P}_2(B)$ as promised (of course, we could have also just plugged all these powers and factorials into a calculator to get the same result).

4 Max/Min Dice Rolls

Yining rolls three fair six-sided dice.

(a) Let X denote the maximum of the three values rolled. What is the distribution of X (that is, $\mathbb{P}[X = x]$ for $x = 1,2,3,4,5,6$)? You can leave your final answer in terms of "x". [*Hint*: Try to first compute $\mathbb{P}[X \le x]$ for $x = 1,2,3,4,5,6$]. If you want to check your answer, you can solve this problem using counting and make sure it matches with the formula you derived.

(b) Let Y denote the minimum of the three values rolled. What is the distribution of Y?

Solution:

(a) Let X denote the maximum of the three values rolled. We are interested in $\mathbb{P}(X = x)$, where $x = 1,2,3,4,5,6$. First, define X_1, X_2, X_3 to be the values rolled by the first, second, and third dice. These random variables are i.i.d. and uniformly distributed between 1 and 6 inclusive.

Following the hint we first compute $\mathbb{P}[X \le x]$ for $x = 1,2,3,4,5,6$:

$$\mathbb{P}(X \le x) = \mathbb{P}(X_1 \le x)\mathbb{P}(X_2 \le x)\mathbb{P}(X_3 \le x) = \left(\frac{x}{6}\right)\left(\frac{x}{6}\right)\left(\frac{x}{6}\right) = \left(\frac{x}{6}\right)^3$$

Then, observing that $\mathbb{P}(X = x) = \mathbb{P}(X \le x) - \mathbb{P}(X \le x-1)$:

$$\mathbb{P}(X = x) = \left(\frac{x}{6}\right)^3 - \left(\frac{x-1}{6}\right)^3 = \frac{3x^2 - 3x + 1}{216} = \begin{cases} \frac{1}{216}, & x = 1 \\ \frac{7}{216}, & x = 2 \\ \frac{19}{216}, & x = 3 \\ \frac{37}{216}, & x = 4 \\ \frac{61}{216}, & x = 5 \\ \frac{91}{216}, & x = 6 \end{cases}$$

One can confirm that $\sum_{x=1}^{6} \mathbb{P}(X = x) = 1$.

(b) Similarly to the previous part, we first compute $\mathbb{P}[Y \geq y]$.

$$\mathbb{P}(Y \geq y) = \mathbb{P}(X_1 \geq y)\mathbb{P}(X_2 \geq y)\mathbb{P}(X_3 \geq y) = \left(\frac{6-(y-1)}{6}\right)\left(\frac{6-(y-1)}{6}\right)\left(\frac{6-(y-1)}{6}\right) = \left(\frac{7-y}{6}\right)^3 .$$

Then, observing that $\mathbb{P}(Y = y) = \mathbb{P}(Y \geq y) - \mathbb{P}(Y \geq y-1)$:

$$\mathbb{P}[Y = y] = \left(\frac{7-y}{6}\right)^3 - \left(\frac{6-y}{6}\right)^3 .$$

5 Balls and Bins, All Day Every Day

Suppose n balls are thrown into n labeled bins one at a time, where n is a positive *even* integer.

(a) What is the probability that exactly k balls land in the first bin, where k is an integer $0 \leq k \leq n$?

(b) What is the probability p that at least half of the balls land in the first bin? (You may leave your answer as a summation.)

(c) Using the union bound, give a simple upper bound, in terms of p, on the probability that some bin contains at least half of the balls.

(d) What is the probability, in terms of p, that at least half of the balls land in the first bin, or at least half of the balls land in the second bin?

(e) After you throw the balls into the bins, you walk over to the bin which contains the first ball you threw, and you randomly pick a ball from this bin. What is the probability that you pick up the first ball you threw? (Again, leave your answer as a summation.)

Solution:

(a) The probability that a particular ball lands in the first bin is $1/n$. We need exactly k balls to land in the first bin, which occurs with probability $(1/n)^k$, and we need exactly $n-k$ balls to land in a different bin, which occurs with probability $(1-1/n)^{n-k}$, and there are $\binom{n}{k}$ ways to choose which of the k balls land in first bin. Thus, the probability is $\binom{n}{k}(1/n)^k(1-1/n)^{n-k}$.

(b) This is the summation over $k = n/2, \ldots, n$ of the probabilities computed in the first part, i.e., $\sum_{k=n/2}^{n} \binom{n}{k}(1/n)^k(1-1/n)^{n-k}$.

(c) The event that some bin has at least half of the balls is the union of the events A_k, $k = 1, \ldots, n$, where A_k is the event that bin k has at least half of the balls. By the union bound, $\mathbb{P}(\bigcup_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} \mathbb{P}(A_k) = np$.

(d) The probability that the first bin has at least half of the balls is p; similarly, the probability that the second bin has at least half of the balls is also p. There is overlap between these two events, however: the first bin has half of the balls and the second bin has the second half of the balls. The probability of this event is $\binom{n}{n/2}n^{-n}$: there are n^n total possible configurations for the n balls to land in the bins, but if we require exactly $n/2$ of the balls to land in the first bin and the remaining balls to land in the second bin, there are $\binom{n}{n/2}$ ways to choose which balls land in the first bin. By the principle of inclusion-exclusion, our desired probability is $p+p-\binom{n}{n/2}n^{-n} = 2p-\binom{n}{n/2}n^{-n}$.

(e) Condition on the number of balls in the bin. First we calculate the probability $\mathbb{P}(A_k)$, where A_k is the event that, in addition to the first ball you threw, an additional $k-1$ of the other $n-1$ balls landed in this bin, which by the reasoning in Part (a) has probability

$$\mathbb{P}(A_k) = \binom{n-1}{k-1}(1/n)^{k-1}(1-1/n)^{n-k} \ .$$

If we let B be the event that we pick up the first ball we threw, then

$$\mathbb{P}(B \mid A_k) = 1/k$$

since we are equally likely to pick any of the k balls in the bin. Thus the overall probability we are looking for is, by an application of the law of total probability,

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}(A_k \cap B) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}(B \mid A_k) = \sum_{k=1}^{n} \frac{1}{k}\binom{n-1}{k-1}\left(\frac{1}{n}\right)^{k-1}\left(1-\frac{1}{n}\right)^{n-k}.$$

6 Cookie Jars

You have two jars of cookies, each of which starts with n cookies initially. Every day, when you come home, you pick one of the two jars randomly (each jar is chosen with probability $1/2$) and eat one cookie from that jar. One day, you come home and reach inside one of the jars of cookies, but you find that is empty! Let X be the random variable representing the number of remaining cookies in non-empty jar at that time. What is the distribution of X?

Solution: Assume that you found jar 1 empty. The probability that $X = k$ and you found jar 1 empty is computed as follows. In order for there to be k cookies remaining, you must have eaten a cookie for $2n-k$ days, and then you must have chosen jar 1 (to discover that it is empty). Within those $2n-k$ days, exactly n of those days you chose jar 1. The probability of this is $\binom{2n-k}{n}2^{-(2n-k)}$. Furthermore, the probability that you then discover jar 1 is empty the day after is $1/2$. So, the probability that $X = k$ and you discover jar 1 empty is $\binom{2n-k}{n}2^{-(2n-k+1)}$. However, we assumed that we discovered jar 1 to be empty; the probability that $X = k$ and jar 2 is empty is the same by symmetry, so the overall probability that $X = k$ is:

$$\mathbb{P}(X = k) = \binom{2n-k}{n}\frac{1}{2^{2n-k}}, \qquad k \in \{0,\dots,n\}.$$

7 Testing Model Planes

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 any 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 $\mathbb{P}[X_2 = k]$? Name the distribution of X_2 and what its parameters are.

(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) Since Amin is performing n trials (flying a plane), each with an independent probability of "success" (not crashing), we have $X_1 \sim \text{Binom}(n, p)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$, for $0 \leq k \leq n$.

(b) Each model plane independently has probability p^2 of surviving both days. Whether a model plane survives both days is still independent from whether any other model plane survives both days, so we can say $X_2 \sim \text{Binom}(n, p^2)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{2k} (1-p^2)^{n-k}$, for $0 \leq k \leq n$.

(c) By extending the previous part we see each model plane has probability p^t of surviving t days, so $X_t \sim \text{Binom}(n, p^t)$, or $\mathbb{P}[X = k] = \binom{n}{k} p^{tk} (1-p^t)^{n-k}$, for $0 \leq k \leq n$.

(d) We consider the complement, the probability that no model planes remain after t days. By the previous part we know this to be $\mathbb{P}[X_t = 0] = \binom{n}{0} p^{t(0)} (1-p^t)^{n-0} = (1-p^t)^n$. So the probability of at least model plane remaining after t days is $1-(1-p^t)^n$.

(e) No. $\mathbb{P}[A_1 \cap B_1]$ is the probability that the first three model planes crash, which is $(1-p)^3$. But $\mathbb{P}[A_1]\mathbb{P}[B_1] = (1-p)^2(1-p)^2 = (1-p)^4$. So $\mathbb{P}[A_1 \cap B_1] \neq \mathbb{P}[A_1]\mathbb{P}[B_1]$ and A_1 and B_1 are not independent.

(f) $\mathbb{P}[A_1 \cap B_1]$ is the probability that only the first model plane and second model plane crash, which is $(1-p)^2 p^{n-2}$. $\mathbb{P}[A_1]$ is the probability that the first model plane crashes, and exactly one of the remaining $n-1$ model planes crashes, so $\mathbb{P}[A_2] = (1-p) \cdot \binom{n-1}{1}(1-p)p^{n-1-1} = (n-1)(1-p)^2 p^{n-2}$. Trivially, we have $\mathbb{P}[B_2] = 1-p$, so $\mathbb{P}[A_2]\mathbb{P}[B_2] = (n-1)(1-p)^3 p^{n-2}$ which is equal to $\mathbb{P}[A_2 \cap B_2] = (1-p)^2 p^{n-2}$ only when $(n-1)(1-p) = 1$, or when $n = \frac{1}{1-p} + 1$.

(g) No. Let $k_1 = 0$ and $k_2 = 1$. Then $\mathbb{P}[X_i = k_1 \cap X_j = k_2] = 0$ because you can't have 1 plane at the end of day 2 if there are no planes left at the end of day 1. But $\mathbb{P}[X_i = k_1] > 0$ and $\mathbb{P}[X_j = k_2] > 0$ so $\mathbb{P}[X_i = k_1]\mathbb{P}[X_j = k_2] > 0$. Since $\mathbb{P}[X_i = k_1]\mathbb{P}[X_j = k_2] \neq \mathbb{P}[X_i = k_1 \cap X_j = k_2]$, they are not independent.

8 Indicator Variables

(a) After throwing n balls into m bins at random, what is the expected number of bins that contains exactly k balls?

(b) Alice and Bob each draw k cards out of a deck of 52 distinct cards with replacement. Find k such that the expected number of common cards that both Alice and Bob draw is at least 1. You can use a calculator.

(c) A nefarious delivery guy in some company is out delivering n packages to n customers, where $n \in \mathbb{N}$, $n > 1$. Not only does he hand a random package to each customer, he opens the package before delivering it with probability $1/2$. Let X be the number of customers who receive their own packages unopened. Compute the expectation $\mathbb{E}(X)$

(d) Now, compute the variance of the previous part $\text{Var}(X)$.

Solution:

(a) Let $X_i = 1$ if bin i contains exactly k balls and $X_i = 0$ otherwise.

$$\mathbb{E}[X_i] = \binom{n}{k} \left(\frac{1}{m}\right)^k \left(\frac{m-1}{m}\right)^{n-k} = \binom{n}{k} \frac{(m-1)^{n-k}}{m^n}$$
$$\mathbb{E}[X] = \sum_{i=1}^{m} \binom{n}{k} \frac{(m-1)^{n-k}}{m^n} = \binom{n}{k} \frac{(m-1)^{n-k}}{m^{n-1}}$$

(b) Let $X_i = 1$ if card i is chosen by both Alice and Bob and $X_i = 0$ otherwise.
After drawing k cards, the probability that any given card appears at least once $1-(51/52)^k$ so

$$\mathbb{E}[X_i] = \left(1-\left(\frac{51}{52}\right)^k\right)\cdot\left(1-\left(\frac{51}{52}\right)^k\right)$$

$$\mathbb{E}[X] = \sum_{i=1}^{52}\left(1-\left(\frac{51}{52}\right)^k\right)^2 = 52\cdot\left(1-\left(\frac{51}{52}\right)^k\right)^2.$$

Setting $\mathbb{E}[X]=1$, we have $k = 7.69 \approx 8$.

(c) Define

$$X_i = \begin{cases} 1, & \text{if the } i\text{-th customer gets his/her package unopened,} \\ 0, & \text{otherwise.} \end{cases}$$

By linearity of expectation, $\mathbb{E}(X) = \mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}(X_i)$. We have

$$\mathbb{E}(X_i) = \mathbb{P}[X_i = 1] = \frac{1}{2n},$$

since the ith customer will get his/her own package with probability $1/n$ and it will be unopened with probability $1/2$ and the delivery guy opens the packages independently.
Hence,

$$\mathbb{E}(X) = n\cdot\frac{1}{2n} = \boxed{\frac{1}{2}}.$$

(d) To calculate $\mathrm{Var}(X)$, we need to know $\mathbb{E}(X^2)$.
By linearity of expectation:

$$\mathbb{E}(X^2) = \mathbb{E}\big((X_1+X_2+\cdots+X_n)^2\big) = \mathbb{E}\Big(\sum_{i,j} X_iX_j\Big) = \sum_{i,j}\mathbb{E}(X_iX_j).$$

Then we consider two cases, either $i=j$ or $i\neq j$.
Hence $\sum_{i,j}\mathbb{E}(X_iX_j) = \sum_i \mathbb{E}(X_i^2) + \sum_{i\neq j}\mathbb{E}(X_iX_j)$.

$$\mathbb{E}(X_i^2) = \mathbb{E}(X_i) = \frac{1}{2n}$$

for all i. To find $\mathbb{E}(X_iX_j)$, we need to calculate $\mathbb{P}[X_iX_j = 1]$.

$$\mathbb{P}[X_iX_j = 1] = \mathbb{P}[X_i = 1]\mathbb{P}[X_j = 1 \mid X_i = 1] = \frac{1}{2n}\cdot\frac{1}{2(n-1)}$$

since if customer i has received his/her own package, customer j has $n-1$ choices left.
Hence,

$$\mathbb{E}(X^2) = n\cdot\frac{1}{2n} + n\cdot(n-1)\cdot\frac{1}{2n}\cdot\frac{1}{2(n-1)} = \frac{3}{4},$$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{3}{4}-\frac{1}{4} = \boxed{\frac{1}{2}}.$$

9 Homework Process and Study Group

You must describe your homework process and study group in order to receive credit for this question.