1 Double-Check Your Intuition

(a) (i)

Since $X \in \text{Bin}(5, \frac{1}{4})$, so using the formula given, we have that $\mathbb{P}[X = i] = \binom{5}{i} \cdot (\frac{1}{4})^i (1 - \frac{1}{4})^{5-i}$ for $i \in \{0, 1, 2, 3, 4, 5\}$, i.e. $\mathbb{P}[X = i] = f(i)$ where $f = \binom{5}{i}(\frac{1}{4})^i(1-\frac{1}{4})^{5-i}$ and its domain is $i \in \{0, 1, 2, 3, 4, 5\}$

Then, since we define $Y = 5 - X$, so $\mathbb{P}[Y = 5 - j] = \binom{5}{j} \cdot (\frac{1}{4})^j (1 - \frac{1}{4})^{5-j}$ for $j = 0, 1, 2, 3, 4, 5$, so we have that $\mathbb{P}[Y = i] = \binom{5}{5-i} \cdot (\frac{1}{4})^{5-i}(1 - \frac{1}{4})^i$ for $i \in \{0, 1, 2, 3, 4, 5\}$, i.e. $\mathbb{P}[Y = i] = g(i)$ where $g = \binom{5}{5-i}(\frac{1}{4})^{5-i}(1 - \frac{1}{4})^i$ and its domain is also $i \in \{0, 1, 2, 3, 4, 5\}$

(ii) $\mathbb{E}[Z^2] = \frac{91}{6}$

Using given information, we know that $Z \in \{1, 2, 3, 4, 5, 6\}$, and also $\mathbb{P}[Z = 1] = \mathbb{P}[Z = 2] = \mathbb{P}[Z = 3] = \mathbb{P}[Z = 4] = \mathbb{P}[Z = 5] = \mathbb{P}[Z = 6] = \frac{1}{6}$. Thus, we can conclude that $Z^2 \in \{1, 4, 9, 16, 25, 36\}$, and so $\mathscr{A} = \{1, 4, 9, 16, 25, 36\}$, with $\mathbb{P}[Z = 1] = \mathbb{P}[Z = 4] = \mathbb{P}[Z = 9] = \mathbb{P}[Z = 16] = \mathbb{P}[Z = 25] = \mathbb{P}[Z = 36] = \frac{1}{6}$.

Thus, $\mathbb{E}[Z^2] = \sum_{a \in \mathscr{A}} a \cdot \mathbb{P}[Z^2 = a] = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} = \frac{91}{6}$

(b) True

We proceed by a direct proof. Since $\sum_{i \in \mathbb{Z}} \mathbb{P}[A = i] = 1$ by definition of probability and we also have that for any $i \in \mathbb{Z}$, $\mathbb{P}[A = i] \geq 0$, which implies that there exists some $k \in \mathbb{Z}$ such that $\mathbb{P}[A = k] > 0$

Then, since $A = B$ is equivalent to $A = i$ and $B = i$ for all $i \in \mathbb{Z}$, so $\mathbb{P}[A = B] = \sum_{i \in \mathbb{Z}} \mathbb{P}[A = i] \cdot \mathbb{P}[B = i] = \sum_{i \in \mathbb{Z}} (\mathbb{P}[A = i])^2$ using given information, and thus $\mathbb{P}[A = B] = \sum_{i \in \mathbb{Z}} (\mathbb{P}[A = i])^2 \geq (\mathbb{P}[A = k])^2 > 0$, as desired. Q.E.D.

(c) False

We proceed by providing a counterexample. Let C be a random variable denoting the result of a die roll (so $1 \leq C \leq 6$uniformly at random). Using our result from part (a.ii), so $\mathbb{E}[C^2] = \frac{91}{6}$.

Now, using the example of a single die incidence provided in Note 15, so $\mathbb{E}[C] = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2}$, which means that $\mathbb{E}[C]^2 = (\frac{7}{2})^2 = \frac{49}{4} \neq \frac{91}{6} = \mathbb{E}[C]$, which gives the counterexample.

(d) False

We proceed by providing a counterexample. Let X be a random variable on a sample space $\{1, 10^6\}$ such that $\mathbb{P}[X = 1] = 99.9\%, \mathbb{P}[X = 10^6] = 0.1\%$; then let Y be a random variable on a sample space $\{2\}$ such that $\mathbb{P}[Y = 2] = 1$.

Now, $\mathbb{E}[X] = 1 \cdot 99.9\% + 10^6 \cdot 0.1\% = 1000.999$ and $\mathbb{E}[Y] = 2 \cdot 1 = 2$, so it satisfies the condition that $\mathbb{E}[X] > 100\,\mathbb{E}[Y]$. But, consider the probability of $X > Y$. Since $Y = 2$ is constant, so only when $X = 10^6$ does $X > Y$ hold, which means that $\mathbb{P}(X > Y) = \mathbb{P}[X = 10^6] = 0.1\% < 1/100$, which gives the countradiction and counterexample.

(e) False

We proceed by providing a counterexample. Let X be a random variable (taking positive values) on a sample space $\{1, 2\}$ such that $\mathbb{P}[X = 1] = \mathbb{P}[X = 2] = 0.5$; then let Y be a random variable (taking positive values) on a sample space $\{1\}$ such that $\mathbb{P}[Y = 1] = 1$.

Thus, we can easily determine that $\mathbb{P}[\frac{X}{X+Y} = \frac{1}{2}] = \mathbb{P}[\frac{X}{X+Y} = \frac{2}{3}] = 0.5$, which gives us that $\mathbb{E}[\frac{X}{X+Y}] = \frac{1}{2} \cdot 0.5 + \frac{2}{3} \cdot 0.5 = \frac{7}{12}$. Then, we also have that $\mathbb{P}[X + Y = 2] = \mathbb{P}[X + Y = 3] = 0.5 = \mathbb{P}[X = 1] = \mathbb{P}[X = 2]$, so we have $\mathbb{E}[X + Y] = 2 \cdot 0.5 + 3 \cdot 0.5 = 2.5$ and $\mathbb{E}[X] = 1.5$.

Thus, $\frac{\mathbb{E}[X]}{\mathbb{E}[X+Y]} = \frac{1.5}{2.5} = \frac{3}{5} \neq \frac{7}{12} = \mathbb{E}[\frac{X}{X+Y}]$, which gives the counterexample.

(f) True

Since A, B, C are events such that $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, then by Definition 14.4, so they're mutually independent. Q.E.D.

(g) False (No, an event A is not never independent with itself.)

Consider the sample space be flipping a fair coin, and the event A be that the coin lands and stays $45°$ from the ground (assuming a universe where a coin always lands on either heads or tails and doesn't behave like this), then we have that $\mathbb{P}(A) = 0$. Now, we can easily determine that $\mathbb{P}(A \cap A) = 0 = \mathbb{P}(A) \times \mathbb{P}(A)$, and since they're in the same probability space, so by Definition 14.3, the events A, A are indeed independent.

(h) True

Using given information, we have that $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A), \mathbb{P}(\overline{B}) = 1 - \mathbb{P}(B)$, and that $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$. Thus, $\mathbb{P}(\overline{A}) \times \mathbb{P}(\overline{B}) = (1 - \mathbb{P}(A)) \cdot (1 - \mathbb{P}(B)) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B)$.

Then, using the Theorem 14.2 (Inclusion-Exclusion), so we have that $\mathbb{P}(\overline{A}) \times \mathbb{P}(\overline{B}) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) = 1 - (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)) = 1 - \mathbb{P}(A \cup B) = \mathbb{P}(\overline{A} \cap \overline{B})$ where the last equality results from the definition of union and intersections of sets. Since $\mathbb{P}(\overline{A} \cap \overline{B}) = \mathbb{P}(\overline{A}) \times \mathbb{P}(\overline{B})$, so by Definition 14.3, we have that $\overline{A}, \overline{B}$ are also independent. Q.E.D.

2 Airport Revisited

(a) $\frac{n}{4}$

Let X_n denote the number of empty airports after all planes have landed. Then, we can first write

$$X_n = I_1 + I_2 + \cdots + I_n$$

where $I_i = 1$ if neither of the planes from airports $i-1, i+1$ landed at airport i (i.e. both chose the other direction); and $I_i = 0$ otherwise.

Then, specifically, $\mathbb{E}[I_i] = 0 \cdot \mathbb{P}[I_i = 0] + 1 \cdot \mathbb{P}[I_i = 1] = \mathbb{P}[I_i = 1] = \mathbb{P}[$both planes next to airport i chose the other direction$] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Thus, using Theorem 15.1, we have that $\mathbb{E}[X_n] = \mathbb{E}[I_1] + \mathbb{E}[I_2] + \cdots + \mathbb{E}[I_n] = \frac{1}{4} \cdot n = \frac{n}{4}$

(b) $\sum\limits_{i=1}^{n} \prod\limits_{a \in N(i)} \frac{\deg(a)-1}{\deg(a)}$

We proceed with a similar logic as part (a). Let X_n denote the number of empty airports after all planes have landed. Then, we can first write

$$X_n = I_1 + I_2 + \cdots + I_n$$

where $I_i = 1$ if none of the planes from $N(i)$ landed at airport i (i.e. all chose the other direction); and $I_i = 0$ otherwise.

Then, specifically, $\mathbb{E}[I_i] = 0 \cdot \mathbb{P}[I_i = 0] + 1 \cdot \mathbb{P}[I_i = 1] = \mathbb{P}[I_i = 1] = \mathbb{P}[$all planes from $N(i)$ chose another neighbor of theirs$] = \prod\limits_{a \in N(i)} \frac{\deg(a)-1}{\deg(a)}$

Thus, using Theorem 15.1, we have that $\mathbb{E}[X_n] = \mathbb{E}[I_1] + \mathbb{E}[I_2] + \cdots + \mathbb{E}[I_n] =$

$$\sum_{i=1}^{n} \prod_{a \in N(i)} \frac{\deg(a)-1}{\deg(a)}$$

3 Fizzbuzz

(a) $\frac{8}{15}n$

If $15 \mid n$, then there will be exactly $\frac{n}{3}$ multiples of 3 from 1 to n, $\frac{n}{5}$ multiples of 5 from 1 to n, and $\frac{n}{15}$ multiples of 15 from 1 to n.

Let A_1 be the event that picking an integer between 1 and n is a multiple of 3, and let A_2 be the event that picking an integer between 1 and n is a multiple of 5, so $A_1 \cup A_2$ is the event that picking an integer between 1 and n that is a multiple of 15. Thus, we have $\mathbb{P}[A_1] = \frac{\frac{n}{3}}{n} = \frac{1}{3}, \mathbb{P}[A_2] = \frac{\frac{n}{5}}{n} = \frac{1}{5}$, and $\mathbb{P}[A_1 \cup A_2] = \frac{\frac{n}{15}}{n} = \frac{1}{15}$.

Thus, the probability of randomly choosing an integer that printed words (i.e. multiple of 3 or 5) is $\mathbb{P}[\text{word}] = \mathbb{P}[U_{i=1}^{2} A_i] = \mathbb{P}[A_1] + \mathbb{P}[A_2] - \mathbb{P}[A_1 \cup A_2] = \frac{1}{3} + \frac{1}{5} - \frac{1}{15} = \frac{7}{15}$, which means that $\mathbb{P}[\text{integer}] = \mathbb{P}[\overline{\text{word}}] = 1 - \mathbb{P}[\text{word}] = \frac{8}{15}$.

Since the size of the sample space $|\Omega| = n$, so the size of the sample space, where the event is that the printed line contains integer, is $|\omega| = |\Omega| \cdot \mathbb{P}[\text{integer}] = \frac{8}{15}n$.

Thus, if n is a multiple of 15, then $\frac{8}{15}n$-many printed lines will contain an integer.

(b) Direct Proof

We proceed by a direct proof. Since the only prime factors of n are $p_1, p_2, \ldots, p_k$, and they're distinct, so we could eliminate the prime factors with a similar procedure/idea from part (a).

Using the Principle of Inclusion-Exclusion (Theorem 14.2), so the probability of randomly picking a line that contains words (not coprime with n) is that

$$\mathbb{P}[\cup_{j=1}^{k} A_j] = \sum_{a_1=1}^{k} \mathbb{P}[A_{a_1}] - \sum_{a_1<a_2} \mathbb{P}[A_{a_1} \cap A_{a_2}] + \sum_{a_1<a_2<a_3} \mathbb{P}[A_{a_1} \cap A_{a_2} \cap A_{a_3}] - \ldots + (-1)^{n-1} \mathbb{P}[A_1 \cap A_2 \cap \cdots \cap A_k]$$

$$= \sum_{a_1=1}^{k} \frac{1}{p_{a_1}} - \sum_{a_1<a_2} \frac{1}{p_{a_1} p_{a_2}} + \sum_{a_1<a_2<a_3} \frac{1}{p_{a_1} p_{a_2} p_{a_3}} - \ldots + (-1)^{n-1} \frac{1}{p_1 p_2 \ldots p_k}$$

Thus, the probability of randomly picking a line that contains an integer (i.e. picking a number that's coprime with n) is:

$$\mathbb{P}[integer] = 1 - \mathbb{P}[\cup_{j=1}^{k} A_j] = \prod_{j=1}^{k} (1 - \frac{1}{p_j})$$

.

Therefore, $\mathbb{P}[integer]$ is the same as the probability of randomly picking a number that's coprime with n, i.e. $\frac{\phi(n)}{n}$, which means that $\frac{\phi(n)}{n} = \prod_{j=1}^{k} (1 - \frac{1}{p_j})$, as desired.

Q.E.D.

4 Cliques in Random Graphs

(a) $2^{\frac{n(n-1)}{2}}$

The size of the sample space is that for each potential edge, we have two possibilities (heads for edge, or tails for no edge). Since there are a total of $\frac{n(n-1)}{2}$ possible edges, so the size of the sample space is

$$|\Omega| = 2^{\frac{n(n-1)}{2}}$$

(b) $2^{-\frac{k(k-1)}{2}}$

For a particular set of k vertices to form a k-clique, all the $|E| = \frac{k(k-1)}{2}$ possible edges should be connected by definition. Then, since the probability of each edge being connected is $\mathbb{P}[edge] = \frac{1}{2}$, so the total probability is $\mathbb{P} = \prod_{e \in E} \mathbb{P}[e] = (\frac{1}{2})^{|E|} = \frac{1}{2^{\frac{k(k-1)}{2}}} = 2^{-\frac{k(k-1)}{2}}$

(c) Direct Proof

Proof 1: We proceed by a direct combinatorial proof. $\binom{n}{k}$ means that we're choosing k elements from a set of n elements, sampling without replacement, while n^k is equivalent to choosing k elements from a set of n elements, sampling with replacement. Since we are choosing the same number of elements from the same set, and that we definitely have more options sampling with replacement (compared to without replacement), so $\binom{n}{k} \leq n^k$. Q.E.D.

Proof 2: Alternatively, we also provide an algebraic proof. $\binom{n}{k} = n \cdot (n-1) \cdot (n-2) \dots (n-k+1)$, and $n^k = n \cdot n \dots n$. Since they have the same number of elements (both have k terms), and that $0 \leq n - i \leq n$ for $i \in [0, k-1]$, so we have that each term of $\binom{n}{k}$ is less than or equal to the corresponding term in n^k, and that each term is non-negative, so we have $\binom{n}{k} \leq n^k$. Q.E.D.

(d) Direct Proof

We proceed by a direct proof. As we proved in part (b), the probability of a particular set of k vertices to form a k-clique is $\mathbb{P} = 2^{-\frac{k(k-1)}{2}}$. Now, there are $\binom{n}{k}$ total different sets of k particular vertices in a graph with n vertices, and for each set of k vertices, their probability of forming a k-clique is $\mathbb{P}[A] = 2^{-\frac{k(k-1)}{2}}$. Thus, the probability that a graph contains a k-clique is $\mathbb{P}[k\text{-clique}] = \mathbb{P}[\cup_{i=1}^{\binom{n}{k}} A_i] \leq \sum_{i=1}^{\binom{n}{k}} \mathbb{P}[A_i] = \binom{n}{k} \cdot 2^{-\frac{k(k-1)}{2}}$.

Given that $k \geq 4 \log_2 n + 1$, so $k - 1 \geq 4 \log_2 n$, so we have that $2^{-\frac{k(k-1)}{2}} \leq 2^{-\frac{k \cdot 4 \log_2 n}{2}} \leq 2^{-2k \log_2 n} = (2^{\log_2 n})^{-2k} = n^{-2k}$; and since we proved in part (c) that $\binom{n}{k} \leq n^k$, so combining these two gives us that:

$$\mathbb{P}[k\text{-clique}] \leq \binom{n}{k} \cdot 2^{-\frac{k(k-1)}{2}} \leq n^k \cdot n^{-2k} = n^{-k} \leq n^{-1} = \frac{1}{n}$$

Thus, we have proved that the probability that the graph contains a k-clique, for $k \geq 4 \log n + 1$, is at most $\frac{1}{n}$, as desired.

Q.E.D.

5 Balls and Bins, All Day Every Day

(a) $\binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$

The probability is $\binom{n}{k} \cdot (\frac{1}{n})^k \cdot (\frac{n-1}{n})^{n-k}$ as we first choose k balls from the n balls where order does not matter and then consider the probability of each ball getting into desired position (k balls into the first bin, and $(n-k)$ balls into any of the other $(n-1)$ bins).

(b) $p = \sum_{k=\frac{n}{2}}^{n} \binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$

The probability p that at least half of the balls land in the first bin is the sum of the probability of exactly k balls landing in the first bin, where $k \in [\frac{n}{2}, n]$ as n is a positive even number, which means that

$$p = \sum_{k=\frac{n}{2}}^{n} \binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$$

(c) Union bound: np

Let $\mathbb{P}[A_i]$ denote that there are at least half the balls in the i^{th} bin. Since there are a total of n bins, each of which has the same probability of containing at least half of the balls, which is the result of part (b). Thus, using the union bound, we have that the probability that some bin contains at least half of the balls, $\mathbb{P}[half] = \mathbb{P}[U_{i=1}^{n} Ai] \leq \sum_{i=1}^{n} p = np$

(d) $2p - \binom{n}{\frac{n}{2}} \cdot (\frac{1}{n})^n$

Let $\mathbb{P}[A_i]$ denote that there are at least half the balls in the i^{th} bin. So, using Theorem 14.2, the probability we want to know is $\mathbb{P} = \mathbb{P}[\cup_{i=1}^{2} A_i] = \mathbb{P}[A_1] + \mathbb{P}[A_1] - \mathbb{P}[A_1 \cap A_2]$ where $\mathbb{P}[A_i] = p$, and since $A_1 \cap A_2$ represents the probability of bins 1 and 2 both contain at least half the balls, which is equivalent to bins 1 and 2 both contain exactly half of the balls, so $\mathbb{P}[A_1 \cap A_2] = \binom{n}{\frac{n}{2}} \cdot (\frac{1}{n})^{\frac{n}{2}} \cdot (\frac{1}{n})^{\frac{n}{2}} = \binom{n}{\frac{n}{2}} \cdot (\frac{1}{n})^n$.

Thus, $\mathbb{P} = 2p - \binom{n}{\frac{n}{2}} \cdot (\frac{1}{n})^n$

(e) $\sum_{k=1}^{n} \frac{1}{k} \cdot \binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$

For a bin with exactly k balls, the probability that the random ball I picked up is the first ball I threw is $\frac{1}{k}$. Then, using a similar logic from part (a), so the probability of exactly k balls landing in the bin I threw my first ball into is: $\binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$, which gives us that the probability of picking up the first ball I threw in the bin where it contains exactly k balls is $\frac{1}{k} \cdot \binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$.

Thus, since the bin where the first ball landed contains at least one ball, so the total probability of this event is:

$$\sum_{k=1}^{n} \frac{1}{k} \cdot \binom{n}{k} \cdot (\frac{1}{n})^k (\frac{n-1}{n})^{n-k}$$