Probability Theory 1 – Proposed solution of moed aleph summer exam

1. **(a)** Every time we roll the dice, the probability that the blue die and the red die yield the same result is 1/3. Since different dice rolls are independent, it follows that $X \sim Geom(1/3)$.

   (b) If $X = 1$, then $Y \in \{2, 4, 6\}$ and each such value is obtained with probability 1/3. That is $(Y|X = 1) \sim U(\{2, 4, 6\})$. Hence, for every $k \in \{2, 4, 6\}$ we have

$$Pr(X = 1, Y = k) = Pr(X = 1) \cdot Pr(Y = k|X = 1) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

   For every other value of k, we have $Pr(X = 1, Y = k) = 0$.
   Next, for every integer $t > 1$ we have that if $X = t$, then $Y \in \{3, 4, 5\}$. Moreover

$$Pr(Y = 3|X = t) = \frac{|\{(2, 1), (1, 2)\}|}{|\{(2, 1), (1, 2), (3, 1), (1, 3), (3, 2), (2, 3)\}|} = \frac{1}{3}.$$

   where (x, y) indicates that the result of the red die was x and the result of the blue die was y.
   Similarly, $Pr(Y = 4|X = t) = Pr(Y = 5|X = t) = 1/3$, that is, $(Y|X = t) \sim U(\{3, 4, 5\})$. Hence, for every $k \in \{3, 4, 5\}$ we have

$$Pr(X = t, Y = k) = Pr(X = t) \cdot Pr(Y = k|X = t) = \frac{1}{3} \cdot \left(\frac{2}{3}\right)^{t-1} \cdot \frac{1}{3} = \frac{1}{6} \cdot \left(\frac{2}{3}\right)^{t}.$$

   For every other value of k, we have $Pr(X = t, Y = k) = 0$.

   (c) It follows by the law of total probability that

$$\begin{aligned} Pr(Y = 4) &= Pr(Y = 4, X = 1) + Pr(Y = 4, X \geq 2) = \frac{1}{9} + Pr(X \geq 2) \cdot Pr(Y = 4|X \geq 2) \\ &= \frac{1}{9} + \frac{2}{3} \cdot \frac{1}{3} = \frac{1}{3}. \end{aligned}$$

2. **(a)** For every $1 \leq i \leq n$, let X_i be the indicator random variable for the event "the members of married couple i sit next to each other". By viewing the members of married couple i as one element we see that

$$\mathbb{E}(X_i) = Pr(X_i = 1) = \frac{2 \cdot (2n-1)!}{(2n)!} = \frac{1}{n}.$$

   Therefore

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n \cdot 1/n = 1.$$

(b) For every $1 \le i \le 2n-2$, let Y_i be the indicator random variable for the event "seats $i, i+1$ and $i+2$ are occupied by men". A direct calculation then shows that

$$\mathbb{E}(Y_i) = Pr(Y_i = 1) = \frac{\binom{n}{3} \cdot 3! \cdot (2n-3)!}{(2n)!} = \frac{n(n-1)(n-2)}{2n(2n-1)(2n-2)} = \frac{n-2}{4(2n-1)}.$$

Therefore

$$\mathbb{E}(Y) = \mathbb{E}\left(\sum_{i=1}^{2n-2} Y_i\right) = \sum_{i=1}^{2n-2} \mathbb{E}(Y_i) = (2n-2) \cdot \frac{n-2}{4(2n-1)} = \frac{(n-1)(n-2)}{2(2n-1)}.$$

(c) X and Y are not independent. For example, if the members of every couple sit next to each other, then no three men sit consecutively. In particular, $Pr(X = n, Y = 1) = 0$. On the other hand, there are arrangements such that the members of every couple sit next to each other (for example, for every $1 \le i \le n$, the ith man sits in seat $2i-1$ and the ith woman sits in seat $2i$), and there are arrangements such that $Y_i = 1$ for exactly one $1 \le i \le 2n-2$ (for example, men sit in seats 1, 2, 3, women sit in seats 4, 5, 6, and for every $4 \le i \le n$, a man sits in seat $2i-1$ and a woman sits in seat $2i$). Therefore $Pr(X = n) > 0$ and $Pr(Y = 1) > 0$. We conclude that $Pr(X = n, Y = 1) \neq Pr(X = n) \cdot Pr(Y = 1)$.

(d) Fix an arbitrary $1 \le k \le n$. Then

$$Pr(X \ge k) \le \frac{\mathbb{E}(X)}{k} = \frac{1}{k},$$

where the inequality above holds by Markov's inequality which is applicable here as X is a non-negative random variable, and the equality holds by Part (a) of this question.

3. **(a)** Clearly $X \sim Bin(30, 1/2)$. Therefore $\mathbb{E}(X) = 30 \cdot 1/2 = 15$ and $Var(X) = 30 \cdot 1/2 \cdot (1 - 1/2) = 7.5$.

(b) Similarly to Part (a) of this question, $Y \sim Bin(30, 1/3)$ and $Z \sim Bin(30, 1/6)$. Therefore $Var(Y) = 30 \cdot 1/3 \cdot (1 - 1/3) = 20/3$ and $Var(Z) = 30 \cdot 1/6 \cdot (1 - 1/6) = 25/6$. We will now calculate $Cov(X, Y)$. For every $1 \le i \le 30$, let Y_i be the indicator random variable for the event "Ariel ate a peach on day i" and let Z_i be the indicator random variable for the event "Ariel ate a banana on day i". Clearly, $Y = \sum_{i=1}^{30} Y_i$ and $Z = \sum_{i=1}^{30} Z_i$. Therefore

$$Cov(Y, Z) = Cov\left(\sum_{i=1}^{30} Y_i, \sum_{i=1}^{30} Z_i\right) = \sum_{i=1}^{30}\sum_{j=1}^{30} Cov(Y_i, Z_j). \tag{1}$$

Since dice rolls on different days are independent, it follows that Y_i and Z_j are independent and thus $Cov(Y_i, Z_j) = 0$ for every $1 \le i \neq j \le 30$. On the other hand, for every $1 \le i \le 30$ we have

$$Cov(Y_i, Z_i) = \mathbb{E}(Y_i Z_i) - \mathbb{E}(Y_i) \cdot \mathbb{E}(Z_i) = Pr(Y_i = 1, Z_i = 1) - \frac{1}{3} \cdot \frac{1}{6} = 0 - 1/18 = -1/18.$$

Plugging this into (1) we obtain

$$Cov(Y, Z) = 30 \cdot (-1/18) = -5/3.$$

Finally

$$\rho(Y,Z) = \frac{Cov(Y,Z)}{\sqrt{Var(Y)} \cdot \sqrt{Var(Y)}} = -\frac{5/3}{\sqrt{20/3 \cdot 25/6}} = -\frac{5/3}{5/3 \cdot \sqrt{10}} = -\frac{1}{\sqrt{10}}.$$

Remark: This shows that Y and Z are negatively correlated, which makes sense as the more peaches Ariel eats, the less bananas he eats.

(c) We will use Chebyshev's inequality (an application of Markov's inequality yields a weaker result). Observe that $X + Z \geq 25$ if and only if $Y \leq 5$. Moreover, recall from Part (b) of this question that $\mathbb{E}(Y) = 30 \cdot 1/3 = 10$ and $Var(Y) = 20/3$. Hence

$$Pr(X + Z \geq 25) = Pr(Y \leq 5) \leq Pr(|Y - \mathbb{E}(Y)| \geq 5) \leq \frac{Var(Y)}{5^2} = \frac{20}{3 \cdot 25} = \frac{4}{15}.$$