Probability Theory 1 – Proposed solution of moed aleph exam 2021

1. **(a)** Since $Z \sim Bin(4, 1/2)$, the support of the distribution of Z is $\{0, 1, 2, 3, 4\}$. Moreover, the value of Z uniquely determines the value of the pair (X, Y). It holds that

$$\mathbb{P}(X=0, Y=2) = \mathbb{P}(Z=0) = \binom{4}{0}(1/2)^0(1-1/2)^{4-0} = 1/16;$$

$$\mathbb{P}(X=1, Y=2) = \mathbb{P}(Z=1) = \binom{4}{1}(1/2)^4 = 1/4;$$

$$\mathbb{P}(X=2, Y=2) = \mathbb{P}(Z=2) = \binom{4}{2}(1/2)^4 = 3/8;$$

$$\mathbb{P}(X=2, Y=3) = \mathbb{P}(Z=3) = \binom{4}{3}(1/2)^4 = 1/4;$$

$$\mathbb{P}(X=2, Y=4) = \mathbb{P}(Z=4) = \binom{4}{4}(1/2)^4 = 1/16.$$

The results are summarised in the table below.

	$X=0$	$X=1$	$X=2$
$Y=2$	1/16	1/4	3/8
$Y=3$	0	0	1/4
$Y=4$	0	0	1/16

(b) Using the table from (a) we conclude that

$$\begin{aligned} P(X=0) &= P(X=0, Y=2) + P(X=0, Y=3) + P(X=0, Y=4) \\ &= 1/16 + 0 + 0 = 1/16; \end{aligned}$$

$$\begin{aligned} P(X=1) &= P(X=1, Y=2) + P(X=1, Y=3) + P(X=1, Y=4) \\ &= 1/4 + 0 + 0 = 1/4; \end{aligned}$$

and

$$\begin{aligned} P(X=2) &= P(X=2, Y=2) + P(X=2, Y=3) + P(X=2, Y=4) \\ &= 3/8 + 1/4 + 1/16 = 11/16. \end{aligned}$$

Similarly

$$\begin{aligned} P(Y=2) &= P(X=0, Y=2) + P(X=1, Y=2) + P(X=2, Y=2) \\ &= 1/16 + 1/4 + 3/8 = 11/16; \end{aligned}$$

$$\begin{aligned}P(Y=3) &= P(X=0,Y=3)+P(X=1,Y=3)+P(X=2,Y=3)\\ &= 0+0+1/4=1/4;\end{aligned}$$

and

$$\begin{aligned}P(Y=4) &= P(X=0,Y=4)+P(X=1,Y=4)+P(X=2,Y=4)\\ &= 0+0+1/16=1/16.\end{aligned}$$

(c) One can calculate $\mathbb{E}(X+Y)$ using the linearity of expectation and the marginal distributions of X and Y calculated in Part (b) of this question. However, it is easier to observe that $X+Y=Z+2$ (this inequality is immediate from the definition of X,Y and Z). Hence

$$\mathbb{E}(X+Y)=\mathbb{E}(Z+2)=\mathbb{E}(Z)+2=4\cdot 1/2+2=4,$$

where the second equality holds by the linearity of expectation and the third equality holds since $Z\sim Bin(4,1/2)$.

2. **(a)** It follows from the description of the experiment that $X\sim Geom(1/7)$. Hence

$$\begin{aligned}\mathbb{P}(X\geq 10) &= \sum_{k=10}^{\infty}\mathbb{P}(X=k)=\sum_{k=10}^{\infty}1/7\cdot(1-1/7)^{k-1}=\frac{1}{7}\left(\frac{6}{7}\right)^{9}\cdot\sum_{i=0}^{\infty}(6/7)^{i}\\ &= \frac{1}{7}\left(\frac{6}{7}\right)^{9}\frac{1}{1-6/7}=\left(\frac{6}{7}\right)^{9}.\end{aligned}$$

We conclude that $\mathbb{P}(X<10)=1-\mathbb{P}(X\geq 10)=1-(6/7)^9$.

(b) $Y=0$ if and only if we drew from the urn k black balls for some $k\in\mathbb{N}\cup\{0\}$ and then one white ball. It thus follows by the law of total probability that

$$\begin{aligned}\mathbb{P}(Y=0) &= \sum_{k=1}^{\infty}\mathbb{P}(Y=0|X=k)\mathbb{P}(X=k)=\sum_{k=1}^{\infty}(4/6)^{k-1}\cdot 1/7\cdot(6/7)^{k-1}\\ &= 1/7\cdot\sum_{k=1}^{\infty}(4/7)^{k-1}=1/7\cdot\frac{1}{1-4/7}=1/3.\end{aligned}$$

(c) $X=Y+1$ if and only if we drew from the urn k red balls for some $k\in\mathbb{N}\cup\{0\}$ and then one white ball. It thus follows by the law of total probability that

$$\begin{aligned}\mathbb{P}(X=Y+1) &= \sum_{k=1}^{\infty}\mathbb{P}(X=k,Y=k-1)=\sum_{k=1}^{\infty}\mathbb{P}(Y=k-1|X=k)\mathbb{P}(X=k)\\ &= \sum_{k=1}^{\infty}(2/6)^{k-1}\cdot 1/7\cdot(6/7)^{k-1}=1/7\cdot\sum_{k=1}^{\infty}(2/7)^{k-1}\\ &= 1/7\cdot\frac{1}{1-2/7}=1/5.\end{aligned}$$

3. **(a)** This statement is true. Assume that $\mathbb{P}(X \geq a) \geq \mathbb{P}(Y \geq a)$ for every $a \in \mathbb{N}$. Note that

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{\infty} k\mathbb{P}(X=k) = \sum_{k=1}^{\infty}\sum_{a=1}^{k}\mathbb{P}(X=k) \\ &= \sum_{a=1}^{\infty}\sum_{k=a}^{\infty}\mathbb{P}(X=k) = \sum_{a=1}^{\infty}\mathbb{P}(X \geq a),\end{aligned}$$

where the third equality follows by changing the order of summation. Similarly, $\mathbb{E}(Y) = \sum_{a=1}^{\infty} \mathbb{P}(Y \geq a)$. Therefore

$$\mathbb{E}(X) = \sum_{a=1}^{\infty}\mathbb{P}(X \geq a) \geq \sum_{a=1}^{\infty}\mathbb{P}(Y \geq a) = \mathbb{E}(Y),$$

where the above inequality holds by the assumption that $\mathbb{P}(X \geq a) \geq \mathbb{P}(Y \geq a)$ holds for every $a \in \mathbb{N}$.

(b) This statement is false. Consider random variables $X \sim U(\{1,5\})$ and $Y = 2$. Clearly $Supp(X) \subseteq \mathbb{N}$ and $Supp(Y) \subseteq \mathbb{N}$. Moreover $\mathbb{E}(X) = \frac{1+5}{2} = 3 > 2 = \mathbb{E}(Y)$, but $\mathbb{P}(X \geq 2) = \mathbb{P}(X = 5) = 1/2 < 1 = \mathbb{P}(Y = 2) = \mathbb{P}(Y \geq 2)$.

(c) This statement is false. Consider random variables $X \sim U(\{1,2\})$ and $Y = 3 - X$. Clearly $Supp(X) \subseteq \mathbb{N}$ and $Supp(Y) \subseteq \mathbb{N}$. However

$$\rho(X,Y) = \rho(X, 3-X) = \rho(X,-X) = -1,$$

where the second and third equalities hold by properties of ρ that were discussed in the lectures.

4. **(a)** For every $1 \leq i \leq n-2$, let X_i be the indicator random variable for the event that $Y_{i+1} = 2Y_i$ and $Y_{i+2} = Y_i$. It is evident that the only possible values for such a triple (Y_i, Y_{i+1}, Y_{i+2}) are $(1,2,1), (2,4,2)$ and $(3,6,3)$. It follows that $\mathbb{E}(X_i) = \mathbb{P}(X_i = 1) = 3/6^3 = 1/72$ for every $1 \leq i \leq n-2$. Since $X = \sum_{i=1}^{n-2} X_i$, it follows by the linearity of expectation that $\mathbb{E}(X) = \sum_{i=1}^{n-2}\mathbb{E}(X_i) = (n-2)/72$.

(b) For every $1 \leq i \leq n-2$ it holds that $Var(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \mathbb{E}(X_i) - (\mathbb{E}(X_i))^2 = 1/72 - 1/72^2$, where the second equality holds since X_i is an indicator random variable and thus $X_i^2 = X_i$.

Next, fix some $1 \leq i < j \leq n-2$. If $j > i+2$, then $\{i, i+1, i+2\} \cap \{j, j+1, j+2\} = \emptyset$, implying that X_i and X_j are independent. In particular, $Cov(X_i, X_j) = 0$. Suppose then that $j = i+1$. We then have

$$\begin{aligned}Cov(X_i, X_j) &= \mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \mathbb{P}(X_i = 1, X_{i+1} = 1) - 1/72^2 \\ &= \mathbb{P}(Y_{i+1} = 2Y_i = 2Y_{i+2} = 4Y_{i+1}) - 1/72^2 = 0 - 1/72^2 = -1/72^2.\end{aligned}$$

Finally, suppose that $j = i+2$. We then have

$$\begin{aligned}Cov(X_i, X_j) &= \mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \mathbb{P}(X_i = 1, X_{i+2} = 1) - 1/72^2 \\ &= 3/6^5 - 1/72^2 = 1/72^2,\end{aligned}$$

where $\mathbb{P}(X_i = 1, X_{i+2} = 1) = 3/6^5$ since $X_i = X_{i+2} = 1$ if and only if $(Y_i, Y_{i+1}, Y_{i+2}, Y_{i+3}, Y_{i+4}) \in \{(1,2,1,2,1), (2,4,2,4,2), (3,6,3,6,3)\}$.
We conclude that

$$\begin{aligned}
Var(X) &= \sum_{i=1}^{n-2} Var(X_i) + 2 \sum_{1 \leq i < j \leq n-2} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n-2} Var(X_i) + 2 \sum_{i=1}^{n-3} Cov(X_i, X_{i+1}) + 2 \sum_{i=1}^{n-4} Cov(X_i, X_{i+2}) \\
&= (n-2)/72 - (n-2)/72^2 - 2(n-3)/72^2 + 2(n-4)/72^2 \\
&= (n-2)/72 - n/72^2.
\end{aligned}$$

(c) Note that

$$\begin{aligned}
\mathbb{P}(X \geq n/2) \leq \mathbb{P}(X \geq \mathbb{E}(X) + n/3) &\leq \mathbb{P}(|X - \mathbb{E}(X)| \geq n/3) \\
&\leq \frac{Var(X)}{(n/3)^2} \leq 1/n,
\end{aligned}$$

where in the first inequality we used the value of $\mathbb{E}(X)$ as calculated in Part (a), the third inequality is Chebyshev's inequality, and in the last inequality we used the value of $Var(X)$ as calculated in Part (b). We conclude that

$$\lim_{n \to \infty} \mathbb{P}(X \geq n/2) = 0$$

as claimed.