Question 1

Given
$$X(\omega) = \omega, \qquad Y(\omega) = (-1)^\omega, \qquad P(\{\omega\}) = 2^{-\omega}.$$
Since $E[X \mid Y]$ is a function of the random variable Y, we can write
$$E[X \mid Y](\omega) = g(Y(\omega))$$
for some function $g : \{-1, 1\} \to \mathbb{R}$. That is, all ω with $Y(\omega) = -1$ (odd ω) will map to the same value $g(-1)$, and all ω with $Y(\omega) = 1$ (even ω) will map to the same value $g(1)$.

Next, we compute $g(-1) = E[X \mid Y = -1]$ using the formula
$$E[X \mid Y = -1] = \frac{\sum_{\omega \text{ odd}} \omega P(\{\omega\})}{P(Y = -1)}.$$
The probability of odd ω is
$$P(Y = -1) = \sum_{j=0}^{\infty} 2^{-(2j+1)} = \frac{2}{3},$$
and the numerator is
$$\sum_{j=0}^{\infty} (2j+1) 2^{-(2j+1)} = \frac{10}{9}.$$
Thus,
$$g(-1) = \frac{10/9}{2/3} = \frac{5}{3}.$$
Similarly, we compute $g(1) = E[X \mid Y = 1]$ as
$$E[X \mid Y = 1] = \frac{\sum_{\omega \text{ even}} \omega P(\{\omega\})}{P(Y = 1)}.$$
Here,
$$P(Y = 1) = \sum_{j=1}^{\infty} 2^{-2j} = \frac{1}{3}, \qquad \sum_{j=1}^{\infty} (2j) 2^{-2j} = \frac{8}{9}.$$
Hence,
$$g(1) = \frac{8/9}{1/3} = \frac{8}{3}.$$
Therefore, the conditional expectation as a function of ω is
$$E[X \mid Y](\omega) = g(Y(\omega)) = \begin{cases} \frac{5}{3}, & Y(\omega) = -1 \ (\omega \text{ odd}), \\ \frac{8}{3}, & Y(\omega) = 1 \ (\omega \text{ even}). \end{cases}$$

Question 2

Part (a): $\mathbb{E}[X|X] = X$

We are asked to show that the conditional expectation of a random variable X given itself is equal to the variable itself.

Proof. By definition, the conditional expectation $\mathbb{E}[X|X = x]$ is the expected value of X with respect to its conditional distribution, given that $X = x$.

Given that we know the value of the random variable X is exactly x, the conditional probability distribution of X is a point mass at x. That is, $P(X = x|X = x) = 1$.

Therefore, the expected value is simply the value itself:

$$\mathbb{E}[X|X = x] = \sum_k k \cdot P(X = k|X = x) = x \cdot P(X = x|X = x) = x \cdot 1 = x$$

Since this holds for any specific value x, we can write this in terms of the random variable:

$$\mathbb{E}[X|X] = X$$

□

Part (b): $\mathbb{E}[Xg(Y)|Y] = g(Y)\mathbb{E}[X|Y]$ and $\mathbb{E}[g(Y)|Y] = g(Y)$

Proof of $\mathbb{E}[Xg(Y)|Y] = g(Y)\mathbb{E}[X|Y]$

We aim to prove that if $g(Y)$ is a function of the conditioning variable Y, it can be treated as a constant and pulled out of the conditional expectation.

Proof. We will use the definition of conditional expectation for discrete random variables.

$$\mathbb{E}[Xg(Y)|Y = y] = \sum_x xg(y)P(X = x|Y = y)$$

Since we are conditioning on $Y = y$, the value $g(y)$ is a fixed constant with respect to the summation over x. We can therefore pull it out:

$$= g(y)\sum_x xP(X = x|Y = y)$$

By the definition of conditional expectation, the summation term is exactly $\mathbb{E}[X|Y = y]$.

$$= g(y)\mathbb{E}[X|Y = y]$$

Since this equality holds for any possible value y, we can replace y with the random variable Y to get the desired result:

$$\mathbb{E}[Xg(Y)|Y] = g(Y)\mathbb{E}[X|Y]$$

□

Deduction of $\mathbb{E}[g(Y)|Y] = g(Y)$

We can deduce the second part of the question by using the result from the first part.

Proof. Let's set $X = 1$ in the previous result, $\mathbb{E}[Xg(Y)|Y] = g(Y)\mathbb{E}[X|Y]$. Note that $X = 1$ is a constant, and can be considered a valid (albeit trivial) random variable.

$$\mathbb{E}[1 \cdot g(Y)|Y] = g(Y)\mathbb{E}[1|Y]$$

The conditional expectation of a constant is the constant itself. Therefore, $\mathbb{E}[1|Y] = 1$.

$$\mathbb{E}[g(Y)|Y] = g(Y) \cdot 1$$
$$\mathbb{E}[g(Y)|Y] = g(Y)$$

□

Part (c): $\mathbb{E}[\mathbb{E}[X|Y,Z]|Y] = \mathbb{E}[X|Y]$

Proof. By the definition of conditional expectation, for discrete random variables, the left-hand side is:

$$\mathbb{E}[\mathbb{E}[X|Y,Z]|Y = y] = \sum_z \mathbb{E}[X|Y = y, Z = z]P(Z = z|Y = y)$$

Now, let's substitute the definition of the inner conditional expectation:

$$\mathbb{E}[X|Y = y, Z = z] = \sum_x xP(X = x|Y = y, Z = z)$$

So, the full expression becomes:

$$\mathbb{E}[\mathbb{E}[X|Y,Z]|Y = y] = \sum_z \left(\sum_x xP(X = x|Y = y, Z = z) \right) P(Z = z|Y = y)$$

We can rewrite the conditional probabilities using the product rule: $P(A|B) = \frac{P(A,B)}{P(B)}$

$$= \sum_z \left(\sum_x x \frac{P(X = x, Y = y, Z = z)}{P(Y = y, Z = z)} \right) \frac{P(Y = y, Z = z)}{P(Y = y)}$$

The terms $P(Y = y, Z = z)$ cancel out:

$$= \sum_z \sum_x x \frac{P(X = x, Y = y, Z = z)}{P(Y = y)}$$

We can switch the order of summation:

$$= \frac{1}{P(Y = y)} \sum_x x \sum_z P(X = x, Y = y, Z = z)$$

The inner summation over z is the marginal probability of the joint event $(X = x, Y = y)$:

$$\sum_z P(X = x, Y = y, Z = z) = P(X = x, Y = y)$$

Substituting this back, we get:

$$= \sum_x x \frac{P(X = x, Y = y)}{P(Y = y)}$$

The expression inside the summation is the definition of conditional probability $P(X = x|Y = y)$.

$$= \sum_x x P(X = x|Y = y)$$

This is the definition of the conditional expectation of X given $Y = y$.

$$= \mathbb{E}[X|Y = y]$$

Since this holds for any value of y, we can write this in terms of the random variable:

$$\mathbb{E}[\mathbb{E}[X|Y, Z]|Y] = \mathbb{E}[X|Y]$$

$\square$

Solution Key for Q3

Problem data:

- Y is integer–valued with CDF

$$F_Y(y) = 1 - \frac{2}{(y+1)(y+2)}, \qquad y = 0, 1, 2, \ldots$$

- Conditional on $Y = y$, the random variable $Z \mid Y = y$ is uniform on $\{1, 2, \ldots, y^2\}$:

$$P(Z = z \mid Y = y) = \frac{1}{y^2}, \qquad z = 1, 2, \ldots, y^2.$$

- Goal: compute $\mathbb{E}[Z]$.

Step 1: PMF of Y

For integers $y \geq 1$,

$$\begin{aligned} P(Y = y) &= F_Y(y) - F_Y(y-1) \\ &= \Big(1 - \frac{2}{(y+1)(y+2)}\Big) - \Big(1 - \frac{2}{y(y+1)}\Big) \\ &= \frac{4}{y(y+1)(y+2)}. \end{aligned}$$

Step 2: Conditional mean of Z

Given $Y = y$, Z is uniform on $\{1, 2, \ldots, y^2\}$. Therefore,

$$\mathbb{E}[Z \mid Y = y] = \frac{1}{y^2} \sum_{z=1}^{y^2} z = \frac{1}{y^2} \cdot \frac{y^2(y^2+1)}{2} = \frac{y^2+1}{2}.$$

Step 3: Law of total expectation

$$\begin{aligned}\mathbb{E}[Z] &= \sum_{y=1}^{\infty} \mathbb{E}[Z \mid Y = y]\; P(Y = y) \\ &= \sum_{y=1}^{\infty} \frac{y^2+1}{2} \cdot \frac{4}{y(y+1)(y+2)} \\ &= 2\sum_{y=1}^{\infty} \frac{y^2+1}{y(y+1)(y+2)}.\end{aligned}$$

Let

$$S := \sum_{y=1}^{\infty} \frac{y^2+1}{y(y+1)(y+2)}.$$

Step 4: Partial fraction expansion

We decompose:

$$\frac{y^2+1}{y(y+1)(y+2)} = \frac{1}{2y} - \frac{2}{y+1} + \frac{5}{2(y+2)}.$$

$$\frac{y^2+1}{y(y+1)(y+2)} = \frac{A}{y} + \frac{B}{y+1} + \frac{C}{y+2}.$$

$$y^2 + 1 = A(y+1)(y+2) + B\,y(y+2) + C\,y(y+1).$$

$$= A(y^2 + 3y + 2) + B(y^2 + 2y) + C(y^2 + y).$$

$$= (A + B + C)\,y^2 + (3A + 2B + C)\,y + 2A.$$

$$\begin{cases} A + B + C = 1, \\ 3A + 2B + C = 0, \\ 2A = 1. \end{cases}$$

$$A = \tfrac{1}{2}, \quad B = -2, \quad C = \tfrac{5}{2}.$$

$$\therefore \quad \frac{y^2+1}{y(y+1)(y+2)} = \frac{1}{2y} - \frac{2}{y+1} + \frac{5}{2(y+2)}.$$

Why the sum diverges:

We consider

$$\sum_{y=1}^{\infty} \left(\frac{1}{2y} - \frac{2}{y+1} + \frac{5}{2(y+2)} \right).$$

The first term is

$$\frac{1}{2y}.$$

If we only summed this, we would get

$$\sum_{y=1}^{\infty} \frac{1}{2y} = \tfrac{1}{2} \sum_{y=1}^{\infty} \frac{1}{y},$$

which already diverges.

For large y,

$$-\frac{2}{y+1} \approx -\frac{2}{y}, \qquad \frac{5}{2(y+2)} \approx \frac{2.5}{y}.$$

$-2/y + 2.5/y = 0.5/y$.

Thus the whole expression is approximately

$$\frac{1}{2y} + \frac{0.5}{y} = \frac{1}{y}.$$

Therefore the infinite sum behaves like

$$\sum_{y=1}^{\infty} \frac{1}{y},$$

which is the harmonic series and diverges.

$$\boxed{\text{Hence, the series diverges to infinity.}}$$

Final Answer

$$\boxed{\mathbb{E}[Z] = \infty}$$

The expectation does not exist as a finite number.

Q.4.

Given: $\mathbb{E}[X] = \mathbb{E}[Y] = 0, \quad \mathrm{Var}(X) = \mathrm{Var}(Y) = 1, \quad \mathrm{Cov}(X,Y) = \rho$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \quad\Rightarrow\quad \mathbb{E}[X^2] = 1$$
$$\text{Similarly,}\quad \mathbb{E}[Y^2] = 1$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] = \rho \quad \text{(as the means are 0).}$$

Now, using the identity:

$$\max\{x^2, y^2\} = \frac{x^2 + y^2}{2} + \frac{|x^2 - y^2|}{2}$$

$$\begin{aligned}
\Rightarrow \mathbb{E}[\max\{X^2, Y^2\}] &= \tfrac{1}{2}\mathbb{E}[X^2 + Y^2] + \tfrac{1}{2}\mathbb{E}[\,|X^2 - Y^2|\,] \\
&= \tfrac{1}{2}(\mathbb{E}[X^2] + \mathbb{E}[Y^2]) + \tfrac{1}{2}\mathbb{E}[\,|X^2 - Y^2|\,] \\
&= 1 + \tfrac{1}{2}\mathbb{E}[\,|X^2 - Y^2|\,].
\end{aligned}$$

Using Cauchy–Schwarz inequality:

$$\mathbb{E}[\,|X^2 - Y^2|\,] \le \sqrt{\mathbb{E}[(X - Y)^2]\;\mathbb{E}[(X + Y)^2]}.$$

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] - 2\mathbb{E}[XY] = 1 + 1 - 2\rho = 2(1 - \rho)$$

$$\mathbb{E}[(X + Y)^2] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] = 1 + 1 + 2\rho = 2(1 + \rho)$$

$$\Rightarrow \mathbb{E}[\,|X^2 - Y^2|\,] \le \sqrt{2(1 - \rho)\cdot 2(1 + \rho)} = 2\sqrt{1 - \rho^2}.$$

$$\therefore \mathbb{E}[\max\{X^2, Y^2\}] = 1 + \tfrac{1}{2}\mathbb{E}[\,|X^2 - Y^2|\,] \le 1 + \sqrt{1 - \rho^2}.$$

Additional note- For $\rho = 0$:

$$\mathbb{E}[\max\{X^2, Y^2\}] \le 1 + 1 = 2,$$

eg- if X and Y are independent (like symmetric bernoulli).
For $\rho = 1$:

$$X = Y \quad\Rightarrow\quad \max\{X^2, Y^2\} = X^2, \quad \mathbb{E}[X^2] = 1.$$

Similarly for $\rho = -1$.

Hence, the bound is sharp.

Q5

Solution (Indicator variables)

Let π be a permutation chosen uniformly at random from all permutations of $[1:n]$. For each $i \in \{1, \ldots, n\}$ define the indicator random variable

$$I_i = \mathbf{1}\{\pi(i) = i\},$$

that is, $I_i = 1$ if i is a fixed point of π and $I_i = 0$ otherwise. The number of fixed points is

$$X = \sum_{i=1}^{n} I_i.$$

Expectation

By linearity of expectation,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[I_i].$$

For each i, since all permutations are equally likely,

$$\mathbb{P}(\pi(i) = i) = \frac{(n-1)!}{n!} = \frac{1}{n},$$

so $\mathbb{E}[I_i] = \frac{1}{n}$. Hence

$$\mathbb{E}[X] = n \cdot \frac{1}{n} = 1.$$

Variance

We use

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Now,

$$X^2 = \left(\sum_{i=1}^{n} I_i \right)^2 = \sum_{i=1}^{n} I_i^2 + 2 \sum_{1 \le i < j \le n} I_i I_j.$$

Since $I_i^2 = I_i$,

$$\mathbb{E}[X^2] = \sum_{i=1}^{n} \mathbb{E}[I_i] + 2 \sum_{1 \le i < j \le n} \mathbb{E}[I_i I_j].$$

We already know $\mathbb{E}[I_i] = 1/n$, so

$$\sum_{i=1}^{n} \mathbb{E}[I_i] = 1.$$

For $i \neq j$, we need $\mathbb{E}[I_i I_j] = \mathbb{P}(\pi(i) = i, \pi(j) = j)$. Fixing both i and j reduces the permutation count to $(n-2)!$, so

$$\mathbb{E}[I_i I_j] = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

Thus,

$$2 \sum_{1 \leq i < j \leq n} \mathbb{E}[I_i I_j] = 2 \cdot \binom{n}{2} \cdot \frac{1}{n(n-1)} = 1.$$

Therefore,

$$\mathbb{E}[X^2] = 1 + 1 = 2.$$

Finally,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 2 - 1^2 = 1.$$

Conclusion

$$\boxed{\mathbb{E}[X] = 1, \qquad \mathrm{Var}(X) = 1.}$$

Problem 6. Let $X_1, X_2, \dots, X_n$ be independent discrete random variables and let $X = X_1 + X_2 + \cdots + X_n$. Suppose that each X_i is a geometric random variable with parameter p_i, and that $p_1, p_2, \dots, p_n$ are chosen so that the mean of X is a given $\mu > 0$. Show that the variance of X is minimized if the p_i values are chosen to be all equal to $\frac{n}{\mu}$.

Solution: For a geometric random variable with parameter p_i, we have:

$$E[X_i] = \frac{1}{p_i}, \quad \mathrm{Var}(X_i) = \frac{1-p_i}{p_i^2} = \frac{1}{p_i^2} - \frac{1}{p_i} \tag{1}$$

Since $X = X_1 + X_2 + \cdots + X_n$ and the X_i are independent, linearity of variance applies:

$$E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{p_i} = \mu \tag{2}$$

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} \left(\frac{1}{p_i^2} - \frac{1}{p_i} \right) = \sum_{i=1}^{n} \frac{1}{p_i^2} - \mu \tag{3}$$

To minimize the variance, we need to minimize $\sum_{i=1}^{n} \frac{1}{p_i^2}$ subject to $\sum_{i=1}^{n} \frac{1}{p_i} = \mu$.

We use the Cauchy-Schwarz Inequality: for any real numbers $a_1, \dots, a_n$ and $b_1, \dots, b_n$,

$$\left(\sum_{i=1}^{n} a_i b_i \right)^2 \leq \left(\sum_{i=1}^{n} a_i^2 \right) \left(\sum_{i=1}^{n} b_i^2 \right)$$

Let $a_i = \frac{1}{p_i}$ and $b_i = 1$ for all i. Then:

$$\left(\sum_{i=1}^{n} \frac{1}{p_i} \right)^2 \leq \left(\sum_{i=1}^{n} \frac{1}{p_i^2} \right) \cdot n$$

Using $\sum_{i=1}^{n} \frac{1}{p_i} = \mu$, we get:

$$\mu^2 \leq n \sum_{i=1}^{n} \frac{1}{p_i^2} \quad \Rightarrow \quad \sum_{i=1}^{n} \frac{1}{p_i^2} \geq \frac{\mu^2}{n}$$

Equality in Cauchy-Schwarz holds if and only if the vectors are proportional, i.e., $\frac{a_i}{b_i} = \frac{1/p_i}{1} = \frac{1}{p_i}$ is constant for all i. This means $p_1 = p_2 = \cdots = p_n = p$ for some constant p.

From the constraint $\sum_{i=1}^{n} \frac{1}{p_i} = \mu$, we have $n \cdot \frac{1}{p} = \mu$, so $p = \frac{n}{\mu}$.

Therefore, the minimum value of $\sum_{i=1}^{n} \frac{1}{p_i^2}$ is $\frac{\mu^2}{n}$, achieved when $p_i = \frac{n}{\mu}$ for all i.

Problem 7. Let X_1, X_2, X_3 be independent random variables taking values in the positive integers and having PMFs

$$P_{X_i}(k) = (1-p_i)p_i^{k-1}, \qquad k = 1, 2, \ldots, \ i = 1, 2, 3.$$

Compute $P(X_1 < X_2 < X_3)$ and $P(X_1 \leq X_2 \leq X_3)$.

Solution. Each X_i is a geometric random variable with parameter $1-p_i$.

—

1. Computing $P(X_1 \leq X_2 \leq X_3)$:
Condition on $X_2 = m$. Then

$$P(X_1 \leq X_2 \leq X_3) = \sum_{m=1}^{\infty} P(X_2 = m)\, P(X_1 \leq m)\, P(X_3 \geq m).$$

Now,

$$P(X_2 = m) = (1-p_2)p_2^{m-1}, \qquad P(X_1 \leq m) = 1-p_1^m, \qquad P(X_3 \geq m) = p_3^{m-1}.$$

Thus,

$$P(X_1 \leq X_2 \leq X_3) = \sum_{m=1}^{\infty} (1-p_2)p_2^{m-1}\bigl(1-p_1^m\bigr)p_3^{m-1}.$$

Let $b = p_2p_3$ and $c = p_1p_2p_3$. Then

$$P(X_1 \leq X_2 \leq X_3) = (1-p_2)\left(\sum_{m=1}^{\infty} b^{m-1} - \sum_{m=1}^{\infty} p_1^m b^{m-1}\right).$$

The series are geometric:

$$\sum_{m=1}^{\infty} b^{m-1} = \frac{1}{1-b}, \qquad \sum_{m=1}^{\infty} p_1^m b^{m-1} = \frac{p_1}{1-c}.$$

Hence,

$$P(X_1 \leq X_2 \leq X_3) = (1-p_2)\left(\frac{1}{1-b} - \frac{p_1}{1-c}\right).$$

Since $c = p_1 b$, simplifying yields

$$\boxed{P(X_1 \leq X_2 \leq X_3) = \frac{(1-p_1)(1-p_2)}{(1-p_2p_3)(1-p_1p_2p_3)}}.$$

—

2. Computing $P(X_1 < X_2 < X_3)$:
Again, condition on $X_2 = m$:

$$P(X_1 < X_2 < X_3) = \sum_{m=1}^{\infty} P(X_2 = m)\, P(X_1 \le m-1)\, P(X_3 \ge m+1).$$

We have

$$P(X_1 \le m-1) = 1 - p_1^{m-1}, \qquad P(X_3 \ge m+1) = p_3^m.$$

Thus,

$$P(X_1 < X_2 < X_3) = \sum_{m=1}^{\infty} (1-p_2)p_2^{m-1}\big(1 - p_1^{m-1}\big)p_3^m.$$

This equals

$$(1-p_2)\left(p_3 \sum_{m=1}^{\infty} b^{m-1} - p_3 \sum_{m=1}^{\infty} c^{m-1}\right).$$

So,

$$P(X_1 < X_2 < X_3) = (1-p_2)p_3 \left(\frac{1}{1-b} - \frac{1}{1-c}\right).$$

Simplifying,

$$P(X_1 < X_2 < X_3) = \frac{(1-p_1)(1-p_2)\, p_2\, p_3^2}{(1-p_2p_3)(1-p_1p_2p_3)}.$$

$$\boxed{P(X_1 < X_2 < X_3) = \frac{(1-p_1)(1-p_2)\, p_2\, p_3^2}{(1-p_2p_3)(1-p_1p_2p_3)}}$$

—

Final Results:

$$P(X_1 \le X_2 \le X_3) = \frac{(1-p_1)(1-p_2)}{(1-p_2p_3)(1-p_1p_2p_3)}, \qquad P(X_1 < X_2 < X_3) = \frac{(1-p_1)(1-p_2)\, p_2\, p_3^2}{(1-p_2p_3)(1-p_1p_2p_3)}.$$