1. **(24pts total).** Suppose a random variable X has the following probability distribution function

$$p_0 = 0.1, \qquad p_1 = 0.3, \qquad p_2 = 0.2, \qquad p_3 = 0.4,$$

where $p_i = P(X = i)$, $i = 0, 1, 2, 3$.

(a). **(8pts).** Find the mean and variance of X.

$$E[X] = 0.1 \times 0 + 0.3 \times 1 + 0.2 \times 2 + 0.4 \times 3 = 1.9$$

$$E[X^2] = 0.1 \times 0^2 + 0.3 \times 1^2 + 0.2 \times 2^2 + 0.4 \times 3^2 = 4.7$$

$$V[X] = EX^2 - (EX)^2 = 1.09$$

(b). **(8pts).** Suppose we conduct independent trials of the X in (a) until we observe the first 0 or 1. Then what is the expected number of trials we need to conduct?

Let the # of trials be $\cancel{X}$ Y

Then $Y \cancel{X} - 1 \sim \text{Geometric}(0.4)$.

$$\Rightarrow E\cancel{X}Y = 1 + \frac{0.6}{0.4} = 2.5.$$

(c). **(8pts).** Use Chebyshev's inequality to estimate the probability that the number of trials in (b) is less than 5.

$$EY = 2.5$$

$$V[Y] = \frac{0.6}{0.4^2} = 3.75$$

By Chebyshev,

$$P[\,|Y - EY| \geq 2.5] \leq \frac{3.75}{(2.5)^2} = 0.6 \quad (*)$$

Note Y is non-negative.

$$\Rightarrow P[Y \geq 5] \leq 0.6$$

$$\Rightarrow P[Y < 5] \geq 0.4.$$

So the estimate is 0.4.

P.S: Students may use other number instead of 2.5 in $(*)$. Please Grade accordingly. Partial marks can be given according to completeness and logic of arguments.

2. **(23pts total).**

Use the five axioms of expectation to prove the following claims:
Note: The five axioms should be the only assumptions in your proofs. All other claims need to be proved using the five axioms. Arguments of the proofs should be stated clearly in order to receive full marks.

(a). **(8pts).** For any two events A and B, show that $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$.

We first prove $I\{A \cup B\} = I(A) + I(B) - I(A \cap B)$. (*)

First note L.H.S & R.H.S. of (*) can only be 0 or 1.

If L.H.S. of (*) $= 1 \Rightarrow \omega \in A \cup B \Rightarrow$ $\begin{cases} \omega \in A \ \omega \in B & (1) \\ \text{or } \omega \in A \ \omega \notin B & (2) \\ \text{or } \omega \notin A \ \omega \in B. & (3) \end{cases}$

~~If (1)~~. In either (1), (2) or (3). R.H.S. $= 1$.

If L.H.S of (*) $= 0 \Rightarrow \omega \notin A$ & $\omega \notin B$
$\Rightarrow$ R.H.S. $= 0$ $\Rightarrow$ (*) holds.

Take expectation of (*) $\Rightarrow p(A \cup B) = p(A) + p(B) - p(A \cap B)$

P.S. Students can use Venn Diagram to show this. But arguments should be based on the five axioms.

(b). **(7pts).** Show that if A and B are events such that $P(A) = P(B) = 1$, then $P(A \bigcap B) = 1$.

Proof: By (a).

$p(A \cap B) = p(A) + p(B) - p(A \cup B)$. (*)

Note $A \cup B \subseteq \Omega \Rightarrow p(A \cup B) \leq p(\Omega) = 1$

$\overset{(*)}{\Rightarrow} p(A \cap B) \geq 1 + 1 - 1 = 1$.

On the other hand. $p(A \cap B) \leq p(\Omega) = 1$

$\Rightarrow p(A \cap B) = 1$.

Another method. $p(A \cap B) = 1 - p(\bar{A} \cup \bar{B})$.

Argue that Because $p(\bar{A}) = p(\bar{B}) = 0 \Rightarrow p(\bar{A} \cup \bar{B}) = 0$

$\Rightarrow p(A \cap B) = 1$

(c). **(8pts).** Suppose that X is a random variable such that $|X(\omega)| \leq 1$ for all $\omega \in \Omega$, where Ω is the sample space. Show that

$$E[X^2] \leq E[|X|] \text{ ① and } E[X^2] \geq (E[X])^2. \text{ ②}$$

Note that $\because |X| \leq 1 \Rightarrow$

$$|X| - X^2 = |X|(1-|X|) \geq 0$$

$\Rightarrow E[|X|] \geq EX^2$. ① holds.

To prove ②, we use Cauchy's inequality

$$(E[X])^2 = (E[X \cdot 1])^2 \leq E(X^2) \times E(1^2)$$
$$= E(X^2).$$

3. **(23pts total).**

Prove the following results. Arguments of the proofs should be stated clearly in order to receive full marks.

(a). **(8pts).** We say that two events A and B are independent if the two indicator random variables X_A and X_B are independent, where $X_A = I(A, \omega)$ and $X_B = I(B, \omega)$. Remember that $I(A, \omega) = 1$ if $\omega \in A$ and $I(A, \omega) = 0$ otherwise. $I(B, \omega)$ is defined similarly. Based on this definition, prove that A and B are independent if and only if $P(AB) = P(A)P(B)$.

We first prove $P(AB) = P(A)P(B) \Rightarrow A, B$ indep.

Since $I\{A\}$ & $I\{B\}$ can only be 0 or 1. and according to the definition of independence. we only need to check the four cases:

① $P\{I\{A\}=0, I\{B\}=0\} = P(I(A)=0)P(I(B)=0)$
② $P\{I\{A\}=0 \quad I\{B\}=1\} = P(I(A)=0)P(I(B)=1)$
③ 1 0 1 0
④ 1 1 1 1

I will only check ②. Students can obtain full mark as long as they correctly checked one of the four. But they must be aware that there are four cases to receive full mark.

to check ②.

$P(I(A)=0, I(B)=1)$
$= P(\omega \in B, \omega \notin A)$
$= P(B \cap \bar{A})$
$= P(A) - P(AB)$
$= P(A) - P(A)P(B)$
$= P(A)(1-P(B))$
$= P(A)P(\bar{B})$
~~$= P(I(A$ $E(I(A))$~~
$= P(I(A)=1)P(I(B)=0)$.

Now we show A, B independent $\Rightarrow P(AB) = P(A)P(B)$

To back of page 6.

(b). **(8pts).** Suppose that X is a Binomial(n, p) random variable and Y is a Bernoulli random variable with success probability p. Further assume that X and Y are independent. Show that $Z = X + Y$ is a Binomial random variable.

Think of tossing a coin of prob. p landing heads $(n+1)$ times.

Let X be the # of heads in the first n trials & $Y = \begin{cases} 1 & \text{if } (n+1)\text{th trial is head} \\ 0 & \text{o.w.} \end{cases}$

Then $X \sim \text{Binomial}(n, p)$
$Y \sim \text{Bernoulli}(p)$ & $X \perp\!\!\!\perp Y$

Observe that $X+Y$ is the # of heads in the n+1 trials

$\Rightarrow X+Y \sim \text{Binomial}(n+1, p)$.

P.S. students may use direct calculations (which is computationally intensive) Please grade accordingly.

(c). **(7pts).** Suppose that X is a random variable and c is a constant and $E|X-c| = 0$. Show that $P(X=c)=1$.

Proof:

Let events $A_k = \{ w : |X-c| \leq \frac{1}{k} \}$.

$k = 1, 2, \cdots$

Observe that A_k is a sequence of decreasing events.

& $A_\infty = \bigcap_{k=1}^{\infty} A_k = \{ w : |X-c| = 0 \} = \{ w : X = c \}$.

$\therefore P(X=c) = P(\bigcap_{k=1}^{\infty} A_k) = \lim_{k\to\infty} P(A_k)$ $\circledast$

On the other hand. according to Markov's Inequality

$$P(|X-c| > \frac{1}{k}) \leq \frac{E|X-c|}{\frac{1}{k}} = 0$$

$\Rightarrow P(|X-c| > \frac{1}{k}) = 0 \Rightarrow P(A_k) = 1$, $k = 1, 2, \cdots$

$\overset{\circledast}{\Rightarrow} P(X=c) = 1$.

4. **(30pts total).**

(a). **(8pts).** An absent minded secretary prepared five letters and envelopes to send to five different people. Then he randomly placed letters in the envelopes. A match occurs if the letter and its envelope are addressed to the same person. What is the probability that at least one of the five letters and envelopes match?

We use Inclusion - Exclusion Formula.

Let $A_i = \{ \text{ith letter \& envelope match} \}$ $i=1,2,3,4,5$.

$\Rightarrow P(A_i) = \frac{1}{5}$. $P(A_iA_j) = \frac{1}{20}$ $i<j$

$P(A_iA_jA_k) = \frac{1}{60}$ $i<j<k$ $P(A_iA_jA_kA_r) = \frac{1}{120}$ $i<j<k<r$

$P(A_1A_2A_3A_4A_5) = \frac{1}{120}$ $\Rightarrow P(\bigcup_{i=1}^{5} A_i) = 5\times\frac{1}{5} - \binom{5}{2}\cdot\frac{1}{20} + \binom{5}{3}\frac{1}{60}$

$- \binom{5}{4}\times\frac{1}{120} + \binom{5}{5}\times\frac{1}{120} = \frac{19}{30} = 0.6333$

(b). **(8pts).** The number of defect items that a factory produces, X, follows a Poisson distribution with mean 5 per day. The daily repair costs for the defect items is given by $C = X^2 + 3X$. Find the expectation of the daily repair cost and $P(C > 18)$.

Note $E[X] = V[X] = 5$.

~~V[X] =~~ $EX^2 = V[X] + (EX)^2 = 30$.

$\Rightarrow EC = E(X^2) + 3E(X) = 30 + 3\times5 = 45$.

Note $X \geq 0 \Rightarrow C > 18 \Leftrightarrow X > 3$

$\therefore P(C>18) = P(X>3) = 1 - P(X=0) - P(X=1) - P(X=2)$
$- P(X=3)$

$= 1 - [1 + \frac{5^1}{1!} + \frac{5^2}{2!} + \frac{5^3}{3!}]e^{-5}$

8 $= 0.735$.

(c). **(7pts).** Suppose that one's college GPA, X, is related to his/her IQ Y. Assume that $E(X) = 3.0$, $E(Y) = 100$, $V[X] = 0.25$, $V[Y] = 100$ and the covariance $\text{Cov}(X, Y) = 3$. If someone's IQ is 120. Find the best forecast of his/her college GPA.

$$\hat{X} = a_0 + a_1 Y.$$

$$a_1 = V_{YY}^{-1} V_{XY} = \frac{1}{100} \times 3 = 3\%$$

$$a_0 = EX - a_1 EY = 3 - 3\% \times 100 = 0.$$

$$\therefore \hat{X} = 3\% \times 120 = 3.6.$$

(d). **(7pts).** Tom wins a certain game with probability 0.5. He plays a series of this game and we assume that the outcome on any one game is independent of that on any other game. Suppose that Tom stops playing if he wins three games. However, if he plays a total of seven games and wins less than three in those seven games, he will also stop. Let X be the total number of games that Tom plays. Find $E(X)$.

Let Y be the # of games he plays until the third win.

Then $Y-3 \sim$ Negative Binomial $(3, 0.5)$

$$p(X=3) = p(Y=3) = \binom{0+3-1}{3-1} 0.5^3 \cdot 0.5^0 = \frac{1}{8}.$$

$$p(X=4) = p(Y=4) = \binom{1+2}{2} 0.5^3 \, 0.5^1 = \frac{3}{16}.$$

$$p(X=5) = p(Y=5) = \binom{2+2}{2} 0.5^3 \cdot 0.5^2 = \frac{3}{16}.$$

$$p(X=6) = p(Y=6) = \binom{3+2}{2} 0.5^3 \, 0.5^3 = \frac{5}{32}.$$

$$p(X=7) = p(Y \geq 7) = \left(1 - \frac{1}{8} - \frac{3}{16} - \frac{3}{16} - \frac{5}{32}\right) = \frac{11}{32}$$

9

$$\therefore E[X] = 3 \times \frac{1}{8} + 4 \times \frac{3}{16} + 5 \times \frac{3}{16} + 6 \times \frac{5}{32} + 7 \times \frac{11}{32} = 5.41$$

Selected list of formulas that may be useful for this test:

1. If X follows Poisson(λ), then $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,\cdots,$.

2. If X follows Binomial(n,p) distribution, then $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$, $k=0,1,\cdots,n$, where $\binom{n}{k}=n!/(k!(n-k)!)$.

3. If X follows a negative binomial distribution with parameters r and p, then

$$P(X=x)=\binom{x+r-1}{r-1}p^r(1-p)^x, \quad x=0,1,2,\cdots.$$

4. If X follows a geometric distribution with parameter p, then

$$P(X=x)=p(1-p)^x, \quad x=0,1,2,\cdots.$$

5. Inclusion-Exclusion Formula:

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= \sum_i P(A_i)-\sum_{i_1<i_2} P(A_{i_1}A_{i_2})+\cdots+(-1)^{r+1}\sum_{i_1<i_2<\cdots<i_r} P(A_{i_1}A_{i_2}\cdots A_{i_r}) \\ &+ \cdots+(-1)^{n+1}P(A_1A_2\cdots A_n). \end{aligned}$$

6. Least Squares Estimation:

The best linear estimate of X given $Y_1,Y_2,\cdots,Y_m$ is given by

$$\hat{X}=a_0+a_1Y_1+\cdots+a_mY_m,$$

where $\vec{a}=(a_1,\cdots,a_m)^T$ satisfies

$$V_{YY}\vec{a}=V_{XY}$$

and $a_0=E(X)-\sum_{j=1}^{m}a_jE(Y_j)$. Note that T in the definition of $\vec{a}$ denotes transpose.

In the above equations, $V_{YY}=\text{Cov}[(Y_1,Y_2,\cdots,Y_m)]$ is the covariance matrix of $(Y_1,Y_2,\cdots,Y_m)$ and $V_{XY}=\text{Cov}[X,(Y_1,Y_2,\cdots,Y_m)]$ is the covariance between X and $(Y_1,Y_2,\cdots,Y_m)$.

7. Chebyshev's Inequality:

For any random variable X and a real number $a>0$, we have

$$P(|X-E(X)|>a)\le V[X]/a^2.$$

8. Cauchy's Inequality: For any random variables X_1 and X_2, we have

$$[E(X_1X_2)]^2\le E(X_1^2)E(X_2^2)$$

with equality if and only if $P(c_1X_1+c_2X_2=0)=1$ holds for some constants c_1 and c_2 such that $c_1^2+c_2^2\neq 0$.

9. Markov's Inequality:

For any nonnegative random variable X and a real number $a>0$, we have

$$P(X>a)\le E[X]/a.$$