Section A.
Answer each of the following questions [Marks per questions as shown; 25% total]

1. [5 points]

   In what follows, suppose that X, Y, Z are random variables with possible outcomes $\{0, 1\}$. For each of the following statements, write whether they are True (**T**) or False (**F**). In each case, briefly justify your reasoning. (Answers that omit reasoning will *not* be awarded points.)

   i. It is always true that
   $$p(X = 0 \mid Y = 1) + p(X = 1 \mid Y = 1) = 1.$$

   ii. It is always true that
   $$p(X = 1 \mid Y = 1) + p(X = 1 \mid Y = 0) = 1.$$

   iii. If
   $$p(X = 1, Y = 1, Z = 1) = p(X = 1)p(Y = 1)p(Z = 1)$$
   then X, Y, Z must be statistically independent.

   iv. Assuming none of the probabilities involved are zero, it is always true that
   $$\frac{p(X = 0, Y = 0)}{p(X = 0, Y = 1)} = \frac{p(Y = 0 \mid X = 0)}{p(Y = 1 \mid X = 0)}.$$

   v. Assuming none of the probabilities involved are zero, it is always true that
   $$p(X = 0, Y = 0) + p(X = 1, Y = 0) = \frac{p(X = 0, Y = 0)}{p(X = 0 \mid Y = 0)}.$$

2. [14 points]

   Steve, a cricket captain, has been very successful: his team has won 90% of the games that Steve has played. Steve is also superstitious: he believes that when he keeps a red handkerchief with him, his team has a much better chance of winning. This is because of all the games where Steve played *with* the red handkerchief, his team won 95% of games; on the other hand, of all the games where Steve played *without* the red handkerchief, his team won 85% of games.

   i. What fraction of games must Steve have played *with* the red handkerchief?

   ii. What is the probability of Steve playing with the red handkerchief *given* that his team wins? (You may leave your answer as a fraction, but simplify this as much as possible.)

   iii. Andy, an internet commentator, thinks Steve's faith in the handkerchief is misguided. He claims: "We must keep in mind that of all the games where Steve has played with his red handkerchief, 90% of them have been against India. And the probability of him winning given that he plays with the red handkerchief *and* against India is *twice* that of him winning given that he plays with the red handkerchief *and* does *not* play against India! So, I think it is his playing against India that is the cause of his success."

   What is the probability of Steve winning given he has the handkerchief *and* does *not* play against India?

3. [6 points]

Jorge has been given a gold coin for his birthday. However, he notices it has a slight defect; as a result, he suspects that it will land "heads" more often than "tails". In order to test this, he flips it ten times. The outcomes he observes are

$$\mathcal{D} = \{\mathtt{h}, \mathtt{h}, \mathtt{h}, \mathtt{t}, \mathtt{h}, \mathtt{t}, \mathtt{t}, \mathtt{h}, \mathtt{h}, \mathtt{t}\},$$

where $\mathtt{h}$ denotes the coin landing "heads", and $\mathtt{t}$ denotes the coin landing "tails".

  i. Let $\theta_\mathtt{h}$ denote the unknown bias of the coin. Given the observations $\mathcal{D}$ above, what is the maximum likelihood estimate for $\theta_\mathtt{h}$?

  ii. Suppose Jorge were to conduct ten *thousand* flips of his gold coin, and compute the maximum likelihood estimate for $\theta_\mathtt{h}$. By referencing an appropriate theorem, explain why he can be confident in his estimate being close to the true bias of the coin.

 iii. Zoe, Jorge's impatient friend, takes the coin from him one day and performs two flips. She observes:
$$\mathcal{D}' = \{\mathtt{t}, \mathtt{t}\}.$$

Based on this, she concludes the coin will always land "tails". Jorge is skeptical of Zoe's conclusion, but she adamantly says: *"It has to be completely biased in favour of tails – it's the only possibility that agrees perfectly with the observations!"*

Suppose Jorge is able to express his *prior* belief that the coin is more likely to land "heads" via some distribution $p(\theta)$. Explain in one or two sentences how Jorge could use this to construct a different estimate for $\theta_\mathtt{h}$, and reach a different conclusion to Zoe.

Section B.
Answer each of the following questions [Marks per questions as shown; 25% total]

1. [12 points]

    Ray and Naks are playing a chess game. At a crucial moment, Ray has four possible moves he can play, labelled $\mathcal{X} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$. Let X denote a random variable representing Ray's move, with possible outcomes $\mathcal{X}$. The probability of Ray playing each of the moves are

    $$p(X = \mathtt{a}) = \frac{1}{2} \qquad\qquad p(X = \mathtt{c}) = \frac{1}{12}$$
    $$p(X = \mathtt{b}) = \frac{1}{3} \qquad\qquad p(X = \mathtt{d}) = \frac{1}{12}$$

    Depending on the move Ray plays, Naks will have a certain probability of winning the game. Let Y denote a random variable representing the winner of the game, with possible outcomes $\{\mathtt{Ray}, \mathtt{Naks}\}$. The probabilities of Naks winning given each of Ray's moves are

    $$p(Y = \mathtt{Naks} \mid X = \mathtt{a}) = \frac{1}{6} \qquad\qquad p(Y = \mathtt{Naks} \mid X = \mathtt{c}) = \frac{1}{2}$$
    $$p(Y = \mathtt{Naks} \mid X = \mathtt{b}) = \frac{1}{4} \qquad\qquad p(Y = \mathtt{Naks} \mid X = \mathtt{d}) = \frac{1}{2}$$

    i. Compute $H(X)$. (You may use the approximation $\log_2(3) \approx 1.58$.)

    ii. Compute $H(Y)$. (You may use the approximation $\log_2(3) \approx 1.58$.)

    iii. Give an example of outcomes $x, x' \in \mathcal{X}$ such that

    $$H(Y \mid X = x) < H(Y) < H(Y \mid X = x').$$

    (If necessary, you may use the approximation $\log_2(5) \approx 2.32$.)

    iv. Explain in one or two sentences why, intuitively, the value of $H(Y \mid X)$ cannot be larger than $H(Y)$. (You do *not* need to explicitly compute the value of $H(Y \mid X)$ for this question.)

    v. Watson, a computer program, is analysing the game as it happens. As soon as Ray plays his move, the program outputs its decision as to whether that move was good, or not. Let Z be a random variable denoting Watson's decision, with possible outcomes $\{\mathtt{good}, \mathtt{bad}\}$. It is known that $Z = \mathtt{bad}$ with certainty if $X = \mathtt{d}$, and $Z = \mathtt{good}$ otherwise; that is, the output of the algorithm is *deterministic* given Ray's move.
    Compute $H(Z \mid X)$. Explain your answer intuitively.

2. [9 points]

    Recall that the 3-tuple of random variables (X, Y, Z) form a Markov chain if and only if

    $$p(X, Y, Z) = p(X)p(Y \mid X)p(Z \mid Y).$$

    i. Show that if (X, Y, Z) form a Markov chain, then (Z, Y, X) form a Markov chain.

    ii. Using (i), and the data-processing inequality, show that if (X, Y, Z) form a Markov chain, then
    $$I(X; Z) \leq \min(I(X; Y), I(Y; Z)).$$

    In one or two sentences, explain the above formula intuitively.

iii. Give an example of random variables (X, Y, Z) that form a Markov chain, and for which $I(X; Y) < I(Y; Z)$.

3. [4 points]

Let X be a random variable denoting the number of cars passing through Northbourne Avenue during the morning rush hour on a given weekday. From historical records, the expected value of X is known to be 5000.

Similarly, let Y be a random variable denoting the number of cars passing through London Circuit during the morning rush hour on a given weekday. From historical records, the expected value of Y is known to be 3000. The random variables X and Y are known to be strongly dependent.

i. Let $Z = X + Y$ denote the total number of cars passing through Northbourne Avenue and London Circuit *combined*. Compute $\mathbb{E}[Z]$.

ii. Using Markov's inequality, given an upper bound on the probability of Z being 20,000 or more vehicles.

iii. In or two sentences, explain why the probability in (ii) may be difficult to compute exactly.