You will have 55 minutes to complete this test.

Each of the problems is worth 25% of your exam grade.

No calculators, notes, or books are permitted.

If a question calls for a numerical answer, you don’t need to multiply everything out. (For example, it is fine to write something like \( (0.9)^7/\binom{7}{3} \) as your answer.)

Don’t forget to write your name on the top of every page.

Please show your work and explain your answers we will not award full credit for the correct numerical answer without proper explanation. Good luck!
Problem 1 (35 points) A fair coin is tossed infinitely many times.

1. (3 points) What is the probability that the outcome of the first $2N$ tosses is $N$ heads followed by $N$ tails.

2. (3 points) What is the probability of having exactly $k$ heads among the $N$ first tosses?

3. (4 points) Compute the probability that the first head appears at the $n^{th}$ toss.

4. (10 points) Let $T$ be the first toss when a head appears. Compute $E[T]$ and $Var(T)$.

5. (15 points) Let $Q_n$ denote the probability that no run of 3 consecutive heads appears in $n$ tosses of a fair coin. Show that

$$Q_n = \frac{1}{2}Q_{n-1} + \frac{1}{4}Q_{n-2} + \frac{1}{8}Q_{n-3}$$

by conditioning on the first tail. Compute $Q_6$.

Proof. 1. Let $H_i$ and $T_i$ be the events that $i$-th toss resulted in heads and tails respectively. Then the probability is

$$P(H_1, ..., H_N, T_{N+1}, ..., T_{2N}) = P(H_1)P(H_2)...P(H_N)P(T_{N+1})...P(T_{2N}) = (1/2)^{2N}.$$ 

2. Let $A$ be the event of having exactly $k$ heads among the first $N$ tosses. Let $C_k$ be the number of ways to positions $k$ heads in first $N$ tosses. Then we have

$$P(A) = \frac{1}{2^N} C_k,$$

where $C_k = \binom{N}{k}$ if $k \leq N$ and 0 otherwise.

3. The above is precisely

$$P(T_1, ..., T_{n-1}, H_n) = P(T_1)...P(T_{n-1})P(H_n) = (1/2)^n.$$

4. The random variable $T$ has the geometric distribution with probability $p = 1/2$. Thus one concludes

$$E[T] = 2 \quad Var(T) = \frac{1 - 1/2}{(1/2)^2} = 2.$$ 

See Examples 8b and 8c from Chapter 4 in Ross for details.

5. Let $E^n_i$ be the event there are no 3 consecutive heads in tosses $i + 1, ..., n$. Then one has

$$Q_n = P(E^n_0) = P(E^n_0 | T_1)P(T_1) + P(E^n_0 | H_1, T_2)P(H_1, T_2) + P(E^n_0 | H_1, H_2, T_2)P(H_1, H_2, T_2) +$$

$$P(E^n_0 | H_1, H_2, H_3)P(H_1, H_2, H_3) =$$

$$= \frac{1}{2}P(E^n_1) + \frac{1}{4}P(E^n_2) + \frac{1}{8}P(E^n_3) = \frac{1}{2}Q_{n-1} + \frac{1}{4}Q_{n-2} + \frac{1}{8}Q_{n-3}$$
Then we have $Q_3 = 7/8$, $Q_2 = Q_1 = 1$, which implies consecutively

\[
Q_4 = (1/2)(7/8) + (1/4) + (1/8) = 13/16,
\]

\[
Q_5 = (1/2)(13/16) + (1/4)(7/8) + (1/8) = 24/32 = 3/4,
\]

\[
Q_6 = (1/2)(3/4) + (1/4)(13/16) + (1/8)(7/8) = \frac{24 + 13 + 7}{64} = \frac{44}{64} = \frac{11}{16}.
\]
Problem 2 (10 points) One tosses a fair coin 3 times and denote $X_i$ the random variable which is equal to 1 if the $i^{th}$ toss is head, and zero otherwise.

- Compute $P(X_1 + X_2 + X_3 = 1|X_1 - X_2 = 0)$
- Are the events $\{X_1 = X_2\}$, $\{X_2 = X_3\}$, $\{X_3 = X_1\}$ pairwise independent? independent? Explain.

Proof. One observes the equality of events

$$E = \{X_1 + X_2 + X_3 = 1, X_1 - X_2 = 0\} = \{X_1 = 0, X_2 = 0, X_3 = 1\} \text{ and}$$

$$F = \{X_1 - X_2 = 0\} = \{X_1 = 1, X_2 = 1\} \cup \{X_1 = 0, X_2 = 0\}$$

Thus

$$P(X_1 + X_2 + X_3 = 1|X_1 - X_2 = 0) = \frac{P(E)}{P(F)} = \frac{1/8}{1/2} = \frac{1}{4}.$$ 

Let $A_1 = \{X_2 = X_3\}$, $A_2 = \{X_1 = X_3\}$, $A_3 = \{X_1 = X_2\}$ Then

$$P(A_i) = P(F) = \frac{1}{2} \text{ and } P(A_i \cap A_j) = P(X_1 = X_2 = X_3) = \frac{1}{4},$$

which proves pairwise independence.

We observe

$$P(A_1 \cap A_2 \cap A_3) = P(X_1 = X_2 = X_3) = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3),$$

which proves the three events are not independent. This is fairly clear, since if we know $A_1$ and $A_2$ occurred then we know that $A_3$ occurred as well, so it is not independent from them. 

$\square$
**Problem 3** (10 points) Compute how many quintuple \((x_1, x_2, \ldots, x_5)\) of non-negative integer numbers so that
\[ x_1 + x_2 + \cdots + x_5 = 30 \]

*Proof.* This is essentially Proposition 6.2 in Chapter 1 of Ross’s “A First Course in Probability” with \(n = 30\) and \(r = 5\). So the answer is
\[
\binom{n + r - 1}{r - 1} = \binom{34}{4} = \frac{34 \cdot 33 \cdot 32 \cdot 31}{1 \cdot 2 \cdot 3 \cdot 4} = 46376.
\]

Here is one (combinatorial) proof of the above: We consider a sequence of 34 boxes.

\[
\square \quad \square \quad \square \quad \cdots \quad \square
\]

Then we color 4 of the boxes in black. This leaves 30 white boxes and splits the group of 30 boxes into 5 subgroups of non-negative size \(x_1, x_2, x_3, x_4, x_5\). For example

\[
\square \quad \blacksquare \quad \square \quad \square \quad \blacksquare \quad \square \quad \square \cdots \quad \square
\]

corresponds to \(x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 2, x_5 = 25\), and

\[
\blacksquare \quad \square \quad \blacksquare \quad \blacksquare \quad \square \quad \square \quad \square \quad \cdots \quad \square
\]

corresponds to \(x_1 = 0, x_2 = 1, x_3 = 0, x_4 = 2, x_5 = 27\).

More generally, let the boxes be placed on positions 1, 2, 3, \ldots, 34 and let the colored boxes be in positions \(i_1, i_2, i_3, i_4\). Then we have \(x_1 = i_1 - 1; x_2 = i_2 - i_1; x_3 = i_3 - i_2; x_4 = i_4 - i_3; x_5 = 34 - i_4\). Then one observes that each choice of 4 boxes to color gives in the above way a distinct 5-tuple \((x_1, x_2, \ldots, x_5)\), satisfying our problem and in fact this is a bijective correspondence. This means that the number of 5-tuples \((x_1, x_2, \ldots, x_5)\) is exactly the number of ways to choose 4 boxes from 34, i.e. \(\binom{34}{4}\). \(\blacksquare\)
Problem 4 (15 points) Take 100 men and assume their birthdate is independent, equally distributed on \{1, 2, \ldots, 365\}. Let \( X \) be the number of men with birthdate on April 1. Compute \( E[X] \) by exact computation and then approximate the probability that \( \{X \leq 3\} \) by using the Poisson approximation.

Proof. Number the men from 1 to 100. We know \( X = \sum_{i=1}^{100} X_i \), where \( X_i = 1 \) if man \( i \) was born on April 1, and 0 otherwise. We observe \( X_i \) are independent Bernoulli random variables with probability \( p = 1/365 \). Thus

\[
E[X] = \sum_{i=1}^{100} E[X_i] = \sum_{i=1}^{100} \frac{1}{365} = \frac{100}{365} = \frac{20}{73}.
\]

We observe that the number of people born on April 1 has the binomial distribution with parameters \((p, n) = (\frac{1}{365}, 100)\). Thus we may approximate the distribution by \( \text{Poisson}(np) = \text{Poisson}(\frac{100}{365}) \). Then

\[
P(\{X \leq 3\}) \approx e^{-\lambda}(1 + \lambda + \lambda^2/2 + \lambda^3/6),
\]
where \( \lambda = 20/73 \).
Problem 5 (25 points) Roll 6 dice independently.

1. (10 points) Find the probability that there are at least 4 sixes given that there is at least one.

2. (15 points) Let $X$ be the number of ways of couple of dice that have the same value (if they all have the same value, $X = 6 \times 5/2$). Compute $E[X]$ and $Var(X)$. Hint: decompose $X$ by considering the indicator functions of the sets $E_i^j = \{i^{th} \text{ die equals } j^{th}\}$.

Proof. (1)Let $E_i$ be the event there are $i$ sixes. Then we have the desired probability is

$$P(\cup_{i=4}^6 E_i | \cup_{i=1}^6 E_i) = \frac{P(\cup_{i=4}^6 E_i, \cup_{i=1}^6 E_i)}{P(\cup_{i=1}^6 E_i)} = \frac{P(\cup_{i=4}^6 E_i)}{P(\cup_{i=1}^6 E_i)}$$

We have

$$P(\cup_{i=1}^6 E_i) = 1 - P(E_0) = 1 - (5/6)^6,$$

$$P(\cup_{i=4}^6 E_i) = P(E_4) + P(E_5) + P(E_6) = 15 \times \frac{5^2}{6^6} + 6 \times \frac{5}{6^6} + 1 \times \frac{1}{6^6} = \frac{375 + 30 + 1}{6^6}.$$

Thus we conclude

$$P(\cup_{i=4}^6 E_i | \cup_{i=1}^6 E_i) = \frac{406}{6^6 - 5^6}.$$

(2) We will use the hint. Let us number our dice from 1 to 6. Then let $E_i^j$ be random variables that equal 1 if the value of die $i$ is the same as that of die $j$. Then we have

$$X = \sum_{i=1}^6 \sum_{j<i}^6 E_i^j.$$

We then observe $E_i^j$ are Bernoulli random variables with parameter 1/6 (if $j < i$). Thus $E[E_i^j] = \frac{1}{6}$ and we get

$$E[X] = \sum_{j<i} E[E_i^j] = \frac{1}{6} \times \binom{6}{2} = 2.5$$

We have

$$Var(X) = E[X^2] - E[X]^2,$$

so it suffices to compute $E[X^2]$. We have

$$E[X^2] = E[\sum_{j<i} E_i^j \sum_{l<k} E_l^k] = \sum_{j<i} \sum_{l<k} E[E_i^j E_l^k].$$

We now have that $E[E_i^j E_l^k] = 1/6$ if $i = k$ and $j = l$ and 1/36 otherwise. There are $\binom{6}{2}$ summands of the first type and $\binom{6}{2}^2 - \binom{6}{2}$ of the second so we get

$$E[X^2] = \frac{1}{6} \binom{6}{2} + \frac{1}{36} \left( \binom{6}{2}^2 - \binom{6}{2} \right) = \frac{15}{6} + \frac{210}{36} = \frac{300}{36} = \frac{25}{3}.$$

We conclude that

$$Var(X) = \frac{25}{3} - \frac{25}{4} = \frac{25}{12}. \qed$$