Midterm exam

1. **(15 points)** Suppose there are n students in a class, and each has birthday equally likely to be 1 of 365 days (no leap year).
   (a) Write down the expression of probability that there exists at least a pair of student that share the same birthday. (5 points)
   Sol.
   $$p = 1 - P(\text{all birthday are distinct}) = 1 - \frac{\binom{365}{n}}{365^n}.$$

   (b) What is the expectation of number of distinct birthday ? (10 points)
   Sol. Let A_i be the event that day i is someone's birthday. Then $\sum_{i=1}^{n} \mathbb{1}_{A_i}$ is the number of distinct birthday. Also,
   $$\mathbb{E}(\mathbb{1}_{A_i}) = P(A_i) = 1 - P(\text{day } i \text{ is no one's birthday}) = 1 - \frac{364^n}{365^n}.$$
   So the average number of the distinct birthday is $n\big(1 - \big(\frac{364}{365}\big)^n\big)$.

2. **(10 points)** We roll a die three times. Let A_{ij} be the event that the ith and jth rolls produce the same number. Show that the events A_{12}, A_{23}, A_{13} are pairwise independent but not independent events.

3. **(15 points)** In your pocket there is a random number N of coins, where N has the Poisson distribution with parameter λ. You toss each coin once, with heads showing with probability p each time.
   (a) Compute $\mathbb{P}(H = h \mid N = n)$, where H is the total number of heads. (5 points)
   Sol.
   $$P(H = h \text{ given } N = n) = \binom{n}{h} p^h (1-p)^{n-h}.$$

   (b) Show that the total number of heads has the Poisson distribution with parameter λp. (10 points)
   Sol.
   $$\begin{aligned}
   P(H = h) &= \sum_{n=h}^{\infty} P(H = h \mid N = n) P(N = n) \\
   &= \sum_{n=h}^{\infty} \binom{n}{h} p^h (1-p)^{n-h} e^{-\lambda} \frac{\lambda^n}{n!} \\
   &= e^{-\lambda} (\lambda p)^h \sum_{k=0}^{\infty} \binom{k+h}{h} (1-p)^k \frac{\lambda^k}{(k+h)!} \\
   &= e^{-\lambda} \frac{(\lambda p)^h}{h!} \sum_{k=0}^{\infty} (1-p)^k \frac{\lambda^k}{k!} \\
   &= e^{-\lambda} \frac{(\lambda p)^h}{h!} e^{\lambda(1-p)} = e^{-\lambda p} \frac{(\lambda p)^h}{h!}.
   \end{aligned}$$
   So H has $Pois(\lambda p)$ distribution.

4. **(15 points.)** You and your opponent both roll a fair die. If one get a greater number than the other one, and that number > 3, then the game ends and whoever rolls the larger number wins. Otherwise, we repeat the game.

(a) Let N be the number of rounds in this game. Write down the p.m.f. of N. (5 points)

Sol. Let p be the probability of a round ends. Then $1 - p$ is the probability of getting the same number (this probability is $6/36 = 1/6$) or getting different ones but the larger one ≤ 3 (i.e. getting one of $\{1, 3\}$, $\{2, 3\}$, $\{1, 2\}$ as outcome. So the probability of this consequence is $6/36 = 1/6$). Hence $1 - p = 2 \times 1/6 = 1/3$, i.e. $p = 2/3$. So the p.m.f. of N is

$$P(N = n) = (1 - p)^{n-1} p = \frac{2}{3^n}.$$

(b) What is $P(\text{you win})$? (10 points)

Sol. As long as the probability of winning and losing are the same, $P(\text{win}) = P(\text{lose}) = 1/2$.

5. **(10 points.)** Consider a sequence of tosses of a p-coin. Let Y be the number of toss required to get the first head and Z be the number of tosses required to get the second head after getting the first head. Prove that Y and Z are independent and have the same probability mass functions.

6. **(20 points.)** (a) Let X and Y be two independent discrete random variables. Prove that $E(XY) = E(X)E(Y)$ and $Var(X + Y) = Var(X) + Var(Y)$. (10 points)

(b) Let $X = 1_{A_1} + \cdots + 1_{A_n}$. Compute $Cov(1_{A_i}, 1_{A_j})$ and then $Var(X)$. (10 points)

7. **(15 points)** Let $(X_i)_{1 \leq i \leq n}$ be a sequence n *i.i.d.* random variables with

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}.$$

Define $S_k = X_1 + X_2 + \cdots + X_k$ for $1 \leq k \leq n$ as the k-th partial sum.

(a) Compute $E(S_k^2)$ for any integer $k \geq 1$. (5 points)

(b) Let N be a random variable taking values from $\{1, \cdots, n\}$ with equal probability, independent to $(X_i)_{1 \leq i \leq n}$. What is the mean and variance of the random sum S_N? (10 points)

Hint: Note that $S_N = S_N \mathbb{1}_{\{N=1\}} + \cdots + S_N \mathbb{1}_{\{N=n\}}$, then by linearity of expectation,

$$\mathbb{E}\big(S_N\big) = \sum_{k=1}^{n} \mathbb{E}\big(S_N \mathbb{1}_{\{N=k\}}\big) = \sum_{k=1}^{n} \mathbb{E}\big(S_k \mathbb{1}_{\{N=k\}}\big)$$

and

$$\mathbb{E}\big(S_N^2\big) = \sum_{k=1}^{n} \mathbb{E}\big(S_N^2 \mathbb{1}_{\{N=k\}}\big) = \sum_{k=1}^{n} \mathbb{E}\big(S_k^2 \mathbb{1}_{\{N=k\}}\big)$$