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. When $2 \le n \le 365$,

   $$p = 1 - P(\text{all birthday are distinct}) = 1 - \frac{\binom{365}{n} \times n!}{365^n}.$$

   And $p = 1$ when $n \ge 365$, $p = 0$ when $n = 1$.

   (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 $365 \times \left(1 - \left(\frac{364}{365}\right)^n\right)$.

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.

   Sol. First, we show that they are pairwise Independent. The number of event that satisfies A_{ij} are those with ith and jth being the same, which has 6, and the rest one roll with any result, so times another 6. Hence, the probability is

   $$P(A_{ij}) = \frac{6 \cdot 6}{6^3} = \frac{1}{6}$$

   On the other hand, we take $A_{12} \cap A_{23}$ as example. It means we roll the same outcome each times, which is six in total. The probability of $A_{12} \cap A_{23}$ is

   $$P(A_{12} \cap A_{23}) = \frac{6}{6^3} = \frac{1}{36} = P(A_{12})P(A_{23})$$

   The same goes with the other two combination $A_{12} \cap A_{13}$ and $A_{23} \cap A_{13}$. Therefore, we conclude that A_{12}, A_{23}, A_{13} are pairwise independent.

   Yet, when we consider $A_{12} \cap A_{23} \cap A_{13}$, it is also considering the event where all three rolls have the same outcome. The corresponding probability is

$$P(A_{12} \cap A_{23} \cap A_{13}) = \frac{6}{6^3} = \frac{1}{36} \neq P(A_{12})P(A_{23})P(A_{13})$$

Hence, they are not Independent.

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. Knowing $P(\text{you win in a round}) = \frac{1}{3}$, we have $P(\text{win}) = \sum_{n=0}^{\infty} \frac{1}{3^n} \times \frac{1}{3} = \frac{1}{2}$. In fact, 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.
Sol. See the solution of Quiz 2.

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)

$$\begin{aligned} E[XY] &= \sum_{x\in S_x, y\in S_y} xyP(X=x, Y=y) \\ &= \sum_{x\in S_x, y\in S_y} xyP(X=x)P(Y=y) \quad \text{(Independent)} \\ &= \sum_{x\in S_x} xP(X=x) \sum_{y\in S_y} yP(Y=y) \\ &= E[X]E[Y] \end{aligned}$$

$$\begin{aligned} \text{Var}(X+Y) &= E[(X+Y)^2] - E[(X+Y)]^2 \\ &\qquad \text{(Def. of Variance)} \\ &= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\ &\qquad \text{(Expand the terms and linearity of expectation)} \\ &= \underbrace{(E[X^2] - E[X]^2)}_{\text{Var}(X)} + \underbrace{(E[Y^2] - E[Y]^2)}_{\text{Var}(Y)} + 2\underbrace{(E[XY] - E[X]E[Y])}_{\text{0. By (a)}} \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

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

$$\begin{aligned}
\text{Cov}(1_{A_i}, 1_{A_j}) &= E[(1_{A_i} - E[1_{A_i}])(1_{A_j} - E[1_{A_j}])] \\
&= E[1_{A_i} \cap 1_{A_j}] - E[1_{A_i}]E[1_{A_j}] \\
&= P(A_i \cap A_j) - P(A_i)P(A_j)
\end{aligned}$$

$$\begin{aligned}
\text{Var}(X) &= \sum_{i,j\in\{1,2,\dots,n\}} \text{Cov}(1_{A_i}, 1_{A_j}) \\
&= \sum_{i,j\in\{1,2,\dots,n\}} P(A_i \cap A_j) - P(A_i)P(A_j)
\end{aligned}$$

7. **(15 points)** Let $(X_i)_{1\le i\le n}$ 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 $1 \le k \le n$ as the k-th partial sum.
(a) Compute $E(S_k^2)$ for any integer $k \ge 1$. (5 points)
Sol. For $i.i.d.$ sum, $Var(X_1 + \cdots + X_k) = kVar(X_1) = k$.

(b) Let N be a random variable taking values from $\{1, \cdots, n\}$ with equal probability, independent to $(X_i)_{1\le i\le n}$. 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

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

Sol. First, we calculate $E[S_N]$. The feeling of uncomfortable may arise because the index N now is not a fix number, but a random variable. That is why we use indicator to partition S_N to a series of S_k, where $k's$ are fixed numnber. Using the hint we have

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

Then, the problem state that N is independent to (X_i), so using 6.(b) we have

$E[S_k 1_{N=k}] = E[S_k]E[1_{N=k}]$, for each $k = 1, \ldots, n$. Consequently,

$$\begin{aligned} E\big(S_N\big) &= \sum_{k=1}^{n} E\big(S_k 1_{\{N=k\}}\big) \\ &= \sum_{k=1}^{n} (\underbrace{E[S_k]}_{0} \cdot \underbrace{E[1_{N=k}]}_{P(N=k)=\frac{1}{n}}) \\ &= 0 \end{aligned}$$

Here, using linearity of expectation we get $E[S_k] = E[\sum_{l=1}^{k} X_l] = \sum_{l=1}^{k} E[X_l] = 0$.

Next, we deal with $\text{Var}(S_N)$. Since $E[S_N] = 0$, Variance of S_N is just $E[S_N^2]$. Again, we use the hint to express $E[S_N^2]$ as

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

Still, S_k^2 is independent to N, so by 6.(b) we have $E[S_k 1_{N=k}] = E[S_k]E[1_{N=k}]$ for each $k = 1, ..., n$. Finally,

$$\begin{aligned} E\big(S_N^2\big) &= \sum_{k=1}^{n} E\big(S_k^2 1_{\{N=k\}}\big) \\ &= \sum_{k=1}^{n} (\underbrace{E[S_k^2]}_{k.\ \text{By}7.(a)} \cdot \underbrace{E[1_{N=k}]}_{P(N=k)=\frac{1}{n}}) \\ &= \sum_{k=1}^{n} k\frac{1}{n} \\ &= \frac{n+1}{2} \end{aligned}$$