Probability Theory: Homework 1

姓名： 陈小川　　**学号：** 231240058

评分： ______　　**评阅：** ______

2025 年 2 月 28 日

1 Principle of Inclusion and Exclusion

题目 1 (Union bound)
Prove

$$\mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbf{Pr}\left(A_i\right)$$

using the definition of probability space.

解答：
We use induction on n to prove the statement.
For the base step, we have $n = 1$, the statement is trivial.
For the induction step, we assume that the statement holds for $n = k$, we need to prove it for $n = k + 1$. Recall that we have already proved $\mathbf{Pr}\left(A \cup B\right) = \mathbf{Pr}\left(A\right) + \mathbf{Pr}\left(B\right) - \mathbf{Pr}\left(A \cap B\right)$ on the textbook, then we have

$$\begin{aligned}
\mathbf{Pr}\left(\bigcup_{i=1}^{k+1} A_i\right) &= \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}\right) \\
&= \mathbf{Pr}\left(\bigcup_{i=1}^{k} A_i\right) + \mathbf{Pr}\left(A_{k+1}\right) - \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right) \\
&\leq \sum_{i=1}^{k} \mathbf{Pr}\left(A_i\right) + \mathbf{Pr}\left(A_{k+1}\right) - \mathbf{Pr}\left(\bigcup_{i=1}^{k} \left(A_i \cap A_{k+1}\right)\right) \\
&\leq \sum_{i=1}^{k+1} \mathbf{Pr}\left(A_i\right)
\end{aligned}$$

The last inequality holds because $\mathbf{Pr}\left(\bigcup_{i=1}^{k} \left(A_i \cap A_{k+1}\right)\right) \geq 0$. Thus, by induction, the statement holds for all $n \in \mathbb{N}$.

题目 2 (Principle of Inclusion and Exclusion (PIE))
Prove that

$$\mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right)$$

where $[n] = \{1, 2, \ldots, n\}$.

解答：
We prove the statement by induction on n.
For the base step, we have $n = 1$, the statement is trivial.
For the induction step, we assume that the statement holds for $n = k$, we need to

prove it for $n = k+1$. We have

$$\begin{aligned}
\mathbf{Pr}\left(\bigcup_{i=1}^{k+1} A_i\right) &= \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}\right) \\
&= \mathbf{Pr}\left(\bigcup_{i=1}^{k} A_i\right) + \mathbf{Pr}\left(A_{k+1}\right) - \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right) \\
&= \mathbf{Pr}\left(\bigcup_{i=1}^{k} A_i\right) + \mathbf{Pr}\left(A_{k+1}\right) - \mathbf{Pr}\left(\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right)
\end{aligned}$$

By induction hypothesis, we have

$$\mathbf{Pr}\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right)$$

and

$$\begin{aligned}
&\mathbf{Pr}\left(A_{k+1}\right) - \mathbf{Pr}\left(\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right) \\
=&\mathbf{Pr}\left(A_{k+1}\right) - \sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} (A_i \cap A_{k+1})\right) \\
=&\mathbf{Pr}\left(A_{k+1}\right) - \sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|-1} \mathbf{Pr}\left(\left(\bigcap_{i \in S} A_i\right) \cap A_{k+1}\right) \\
=&\mathbf{Pr}\left(A_{k+1}\right) - \sum_{\{k+1\} \subsetneqq S \subseteq [k+1]} (-1)^{|S|} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) \\
=& \sum_{\{k+1\} \subseteq S \subseteq [k+1]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right)
\end{aligned}$$

Summing up the above two equations, we have

$$\begin{aligned}
\mathbf{Pr}\left(\bigcup_{i=1}^{k+1} A_i\right) &= \sum_{\emptyset \neq S \subseteq [k]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) + \sum_{\{k+1\} \subseteq S \subseteq [k+1]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) \\
&= \sum_{\emptyset \neq S \subseteq [k+1]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right)
\end{aligned}$$

Thus, by induction, the statement holds for all $n \in \mathbb{N}$.

题目 3 (Surjection)

For positive integers $m \geq n$, prove that the probability of a uniform random function $f : [m] \to [n]$ to be surjective (满射) is $\displaystyle\sum_{k=1}^{n} (-1)^{n-k} \binom{n}{k} \left(\frac{k}{n}\right)^m$.

解答:

Let A_i be the event that the i-th element in $[n]$ is not in the range of f. Then the probability of f not being surjective is $\mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right)$. By the Principle of Inclusion and Exclusion, we have

$$\begin{aligned}
\mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) \\
&= \sum_{k=1}^{n} \sum_{S \subseteq [n], |S| = k} (-1)^{k-1} \mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) \\
&= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \left(\frac{n-k}{n}\right)^m
\end{aligned}$$

The last equality holds because $\mathbf{Pr}\left(f(i) \notin S\right) = \frac{n-k}{n}$ for all $i \in [m]$ and

$$\mathbf{Pr}\left(\bigcap_{i\in S} A_i\right) = \mathbf{Pr}\left(\forall i \in [m], f(i) \notin S\right) = \prod_{i=1}^{m} \mathbf{Pr}\left(f(i) \notin S\right) = \left(\frac{n-k}{n}\right)^m$$

Thus, the probability of f being surjective is

$$\begin{aligned} 1 - \mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) &= 1 - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \left(\frac{n-k}{n}\right)^m \\ &= \sum_{k=0}^{n} (-1)^{k} \binom{n}{k} \left(\frac{n-k}{n}\right)^m \\ &= \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{n-k} \left(\frac{k}{n}\right)^m \\ &= \sum_{k=1}^{n} (-1)^{n-k} \binom{n}{k} \left(\frac{k}{n}\right)^m \end{aligned}$$

题目 4 (Bonferroni's inequality and Kounias' inequality)
Prove that

$$\sum_{i=1}^{n} \mathbf{Pr}(A_i) - \sum_{1\le i<j\le n} \mathbf{Pr}(A_i \cap A_j) \le \mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} \mathbf{Pr}\left(A_i\right) - \sum_{i=2}^{n} \mathbf{Pr}(A_1 \cap A_i)$$

Hint: This is sometimes called Kounias' inequality which is weaker than the Bonferroni's inequality. You can try using Venn diagram to understand these inequalities.

解答:
We first show that

$$\sum_{i=1}^{n} \mathbf{Pr}(A_i) - \sum_{1\le i<j\le n} \mathbf{Pr}(A_i \cap A_j) \le \mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right)$$

To prove this, we consider induction on n.
For the base step, we have $n = 2$, the statement is trivial since $\mathbf{Pr}(A_1) + \mathbf{Pr}(A_2) - \mathbf{Pr}(A_1 \cap A_2) = \mathbf{Pr}(A_1 \cup A_2)$.
For the induction step, we assume that the statement holds for $n = k$, we need to prove it for $n = k + 1$. We have

$$\begin{aligned} &\sum_{i=1}^{k+1} \mathbf{Pr}(A_i) - \sum_{1\le i<j\le k+1} \mathbf{Pr}(A_i \cap A_j) \\ =&\sum_{i=1}^{k} \mathbf{Pr}(A_i) - \sum_{1\le i<j\le k} \mathbf{Pr}(A_i \cap A_j) + \mathbf{Pr}(A_{k+1}) - \sum_{i=1}^{k} \mathbf{Pr}(A_i \cap A_{k+1}) \\ \le&\mathbf{Pr}\left(\bigcup_{i=1}^{k} A_i\right) + \mathbf{Pr}(A_{k+1}) - \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right) \\ \le&\mathbf{Pr}\left(\bigcup_{i=1}^{k+1} A_i\right) \end{aligned}$$

The second to last inequality holds because

$$\sum_{i=1}^{k} \mathbf{Pr}(A_i \cap A_{k+1}) \le \mathbf{Pr}\left(\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right) = \mathbf{Pr}\left(\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right)$$

Thus, by induction, the left half of the inequality holds for all $n \in \mathbb{N}$.
Then we show that

$$\mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} \mathbf{Pr}\left(A_i\right) - \sum_{i=2}^{n} \mathbf{Pr}(A_1 \cap A_i)$$

Just note that

$$\begin{aligned}
\sum_{i=1}^{n} \mathbf{Pr}(A_i) - \sum_{i=2}^{n} \mathbf{Pr}(A_1 \cap A_i) &= \mathbf{Pr}(A_1) + \sum_{i=2}^{n} [\mathbf{Pr}(A_i) - \mathbf{Pr}(A_1 \cap A_i)] \\
&= \mathbf{Pr}(A_1) + \sum_{i=2}^{n} \mathbf{Pr}(A_i \setminus A_1) \\
&\ge \mathbf{Pr}(A_1) + \mathbf{Pr}\left(\bigcup_{i=2}^{n} (A_i \setminus A_1)\right) \\
&= \mathbf{Pr}(A_1) + \mathbf{Pr}\left(\left(\bigcup_{i=2}^{n} A_i\right) \setminus A_1\right) \\
&\ge \mathbf{Pr}\left(A_1 \cup \left(\left(\bigcup_{i=2}^{n} A_i\right) \setminus A_1\right)\right) \\
&= \mathbf{Pr}\left(\bigcup_{i=1}^{n} A_i\right)
\end{aligned}$$

Thus, the right half of the inequality also holds for all $n \in \mathbb{N}$.

---

2 Probability space

题目 1 (Nonexistence of probability space)
Prove that it is impossible to define a uniform probability law on natural numbers $\mathbb{N}$. More precisely, prove that there does not exist a probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbf{Pr})$ such that $\mathbf{Pr}(\{i\}) = \mathbf{Pr}(\{j\})$ for all $i, j \in \mathbb{N}$. Please explain why the same argument fails to prove that there is no uniform probability law on the real interval $[0, 1]$, that is, there is no such probability space $([0, 1], \mathcal{F}, \mathbf{Pr})$ that for any interval $(l, r] \subseteq [0, 1]$, it holds that $(l, r] \in \mathcal{F}$ and $\mathbf{Pr}((l, r]) = r - l$. (Actually, such probability measure does exist and is called the Lebesgue measure on $[0, 1]$).

解答:
Suppose to the contrary that there exists a probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbf{Pr})$ such that $\mathbf{Pr}(\{i\}) = p$ for all $i, j \in \mathbb{N}$.
If $p = 0$, then

$$\mathbf{Pr}(\mathbb{N}) = \mathbf{Pr}(\bigcup_{i=1}^{\infty}\{i\}) = \sum_{i=1}^{\infty} \mathbf{Pr}(\{i\}) = \sum_{i=0}^{\infty} 0 = 0$$

which contradicts the fact that $\mathbf{Pr}(\mathbb{N}) = 1$.
If $p > 0$, then there exists a positive integer N such that $N > \frac{1}{p}$. Then

$$\mathbf{Pr}(\bigcup_{i=1}^{N}\{i\}) = \sum_{i=1}^{N} \mathbf{Pr}(\{i\}) = Np > 1$$

which contradicts the fact that probability measure $\mathbf{Pr}$ is a function from $2^{\mathbb{N}}$ to $[0, 1]$.
Thus, there does not exist a probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbf{Pr})$ such that $\mathbf{Pr}(\{i\}) = \mathbf{Pr}(\{j\})$ for all $i, j \in \mathbb{N}$.
However, the same argument fails to prove that there is no uniform probability law on the real interval $[0, 1]$. The reason is that the real interval $[0, 1]$ is uncountable,

while the natural numbers $\mathbb{N}$ is countable. We cannot write $\mathbf{Pr}\left(\bigcup_{x\in[0,1]}\{x\}\right) = \sum_{x\in[0,1]}\mathbf{Pr}(\{x\})$, since the definition of probability measure requires that there can be only countable disjoint sets in the equation. So it's possible that $\mathbf{Pr}(\{x\}) = 0$ for all $x \in [0,1]$.

题目 2 (Smallest σ-field (I))
For any subset $S \subseteq 2^\Omega$, prove that the smallest σ-field containing S is given by

$$\sigma(S) := \bigcap_{\substack{S\subseteq\mathcal{F}\subseteq 2^\Omega \\ \mathcal{F}\text{ is a }\sigma\text{-field}}} \mathcal{F}$$

Hint: You should show that it is indeed a σ-field and also it is the smallest one containing S.

解答:
We first show that $\sigma(S)$ is a σ-field.
(1) $\emptyset \in \sigma(S)$. For any σ-field $\mathcal{F}$ containing S, we have $\emptyset \in \mathcal{F}$, thus $\emptyset \in \sigma(S)$.
(2) $A \in \sigma(S) \Rightarrow A^c \in \sigma(S)$. We have

$$A \in \sigma(S) \Rightarrow \forall\sigma\text{-field } \mathcal{F} \supseteq S, A \in \mathcal{F} \Rightarrow \forall\sigma\text{-field } \mathcal{F} \supseteq S, A^c \in \mathcal{F} \Rightarrow A^c \in \sigma(S)$$

(3) $\forall i \in \mathbb{N}, A_i \in \sigma(S) \Rightarrow \bigcup_{i=1}^\infty A_i \in \sigma(S)$. We have

$$\begin{aligned}
&\forall i \in \mathbb{N}, A_i \in \sigma(S) \\
\Rightarrow &\forall i \in \mathbb{N}, \forall\sigma\text{-field } \mathcal{F} \supseteq S, A_i \in \mathcal{F} \\
\Rightarrow &\forall\sigma\text{-field } \mathcal{F} \supseteq S, \forall i \in \mathbb{N}, A_i^c \in \mathcal{F} \\
\Rightarrow &\forall\sigma\text{-field } \mathcal{F} \supseteq S, \bigcup_{i=1}^\infty A_i \in \mathcal{F} \\
\Rightarrow &\bigcup_{i=1}^\infty A_i \in \sigma(S)
\end{aligned}$$

Thus, $\sigma(S)$ is a σ-field.
Then we show that $\sigma(S)$ is the smallest σ-field containing S.
Note that for all σ-field $\mathcal{F}'$ containing S, we have $\sigma(S) \subseteq \mathcal{F}'$ since $\mathcal{F}'$ is among the $\mathcal{F}$ in the intersection. Thus, $\sigma(S)$ is the smallest σ-field containing S.

题目 3 (Smallest σ-field (II))
Let $S, T \subseteq 2^\Omega$. Show that $\sigma(S) = \sigma(T)$ if and only if $S \subseteq \sigma(T)$ and $T \subseteq \sigma(S)$.

解答:
On one hand, if $\sigma(S) = \sigma(T)$, then $S \subseteq \sigma(T)$ and $T \subseteq \sigma(S)$ since $S \subseteq \sigma(S)$ and $T \subseteq \sigma(T)$.
On the other hand, if $S \subseteq \sigma(T)$ and $T \subseteq \sigma(S)$, then $\sigma(T)$ is a σ-field containing S, and $\sigma(S)$ is a σ-field containing T. Note that $\sigma(T) = \bigcap_{\substack{T\subseteq\mathcal{F}\subseteq 2^\Omega \\ \mathcal{F}\text{ is a }\sigma\text{-field}}} \mathcal{F}$, so we have $\sigma(T) \subseteq \sigma(S)$. Similarly, we have $\sigma(S) \subseteq \sigma(T)$. Thus, $\sigma(S) = \sigma(T)$.

题目 4 (Union of σ-field)
Let $\mathcal{F}$ and $\mathcal{G}$ be σ-fields of subsets of Ω. Show that $\mathcal{F} \cup \mathcal{G}$ is not necessarily a σ-field.
Suppose $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \dots$ is a sequence of σ-fields. Is $\bigcup_{i=1}^{+\infty} \mathcal{F}_i$ a σ-field?

解答:
To show that $\mathcal{F} \cup \mathcal{G}$ is not necessarily a σ-field, we provide a counterexample.
Let $\Omega = \{1, 2, 3\}$, then $\mathcal{F} = \{\emptyset, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$ and $\mathcal{G} = \{\emptyset, \{2\}, \{1, 3\}, \{1, 2, 3\}\}$ are both σ-field. However, $\mathcal{F} \cup \mathcal{G} = \{\emptyset, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ is not a σ-field since $\{1\} \cup \{2\} = \{1, 2\} \notin \mathcal{F} \cup \mathcal{G}$. So $\mathcal{F} \cup \mathcal{G}$ is not necessarily a σ-field.
Now suppose $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \dots$ is a sequence of σ-fields, if $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{F}_3 = \dots$, then $\bigcup_{i=1}^{+\infty} \mathcal{F}_i = \mathcal{F}_1$ is indeed a σ-field. However, we claim that $\bigcup_{i=1}^{+\infty} \mathcal{F}_i$ is not necessarily a σ-field. Here's a counterexample.
Let $\Omega = \mathbb{N}$ and $\mathcal{F}_i = \{X \subseteq \mathbb{N} : X \subseteq [i] \text{ or } X^c \subseteq [i]\}$, we have

1. For all $i \in \mathbb{N}$, $\mathcal{F}_i$ is a σ-field.
   (1) $\emptyset \in \mathcal{F}_i$ since $\emptyset \subseteq [i]$.
   (2) $A \in \mathcal{F}_i \Rightarrow A^c \in \mathcal{F}_i$. If $A \subseteq [i]$, then $A^c \subseteq [i]$, thus $A^c \in \mathcal{F}_i$. If $A^c \subseteq [i]$, then $A \subseteq [i]$, thus $A^c \in \mathcal{F}_i$.
   (3) $\forall i \in \mathbb{N}, A_i \in \mathcal{F}_i \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_i$. If $A_i \subseteq [i]$ for all $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} A_i \subseteq [i]$, thus $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_i$. If there exists $i \in \mathbb{N}$ such that $A_i^c \subseteq [i]$, then $(\bigcup_{i=1}^{\infty} A_i)^c \subseteq A_i^c \subseteq [i]$, thus $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_i$.
2. But $\mathcal{F} = \bigcup_{i=1}^{+\infty} \mathcal{F}_i$ is not a σ-field.
   Let $A_i = \{2i\}$, then $\bigcup_{i=1}^{+\infty} A_i = \{2, 4, 6, \dots\} \notin \mathcal{F}$ since for all $i \in \mathbb{N}$, $\{2, 4, 6, \dots\} \notin \mathcal{F}_i$. Therefore, $\bigcup_{i=1}^{+\infty} \mathcal{F}_i$ is not necessarily a σ-field.

题目 5 (Projection)
Let $\mathcal{F}$ be a σ-field of subsets of Ω and $T \subseteq \Omega$ be a subset. Show that $\{S \cap T \mid S \in \mathcal{F}\}$ is a σ-field.

解答:
Let $\mathcal{F}_T = \{S \cap T \mid S \in \mathcal{F}\}$, then $\mathcal{F}_T \subseteq 2^T$. We have
(1) $\emptyset \in \mathcal{F}_T$. Since $\emptyset \in \mathcal{F}$, we have $\emptyset \cap T = \emptyset \in \mathcal{F}_T$.
(2) $A \in \mathcal{F}_T \Rightarrow T \setminus A \in \mathcal{F}_T$. For any $A \in \mathcal{F}_T$, there exists $S \in \mathcal{F}$ such that $A = S \cap T$. Then $T \setminus A = T \setminus (S \cap T) = (S^c) \cap T \in \mathcal{F}_T$, where $S^c = \Omega \setminus S \in \mathcal{F}$ since $\mathcal{F}$ is a σ-field.
(3) $\forall i \in \mathbb{N}, A_i \in \mathcal{F}_T \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_T$. For all $A_i \in \mathcal{F}_T$, there exists $S_i \in \mathcal{F}$ such that $A_i = S_i \cap T$. Then $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} (S_i \cap T) = (\bigcup_{i=1}^{\infty} S_i) \cap T \in \mathcal{F}_T$.
Thus, $\mathcal{F}_T$ is a σ-field.

题目 6 (Probability space?)
Let $\Omega = \mathbb{R}$, $\mathcal{F}$ is the set of all subsets $A \subseteq \Omega$ so that A or $\overline{A}$ (complement of A) is countable, $P(A) = 0$ in the first case and $P(A) = 1$ in the second. Is $(\Omega, \mathcal{F}, P)$ a probability space? Please explain your answer.

解答:
We claim that $(\Omega, \mathcal{F}, P)$ is indeed a probability space.
We first show that $\mathcal{F}$ is a σ-field. Note that
(1) $\emptyset \in \mathcal{F}$. Since $\emptyset$ is countable, we have $\emptyset \in \mathcal{F}$.
(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$. If A is countable, then $\overline{A^c} = A$ is countable, thus $A^c \in \mathcal{F}$. If A is uncountable, then A^c is countable since $A \in \mathcal{F}$, thus $A^c \in \mathcal{F}$.
(3) $\forall i \in \mathbb{N}, A_i \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. If there exists A_i that is uncountable, then A_i^c is countable, thus $(\bigcup_{i=1}^{\infty} A_i)^c = \bigcap_{i=1}^{\infty} A_i^c$ is countable, and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. Otherwise, if all A_i are countable, then $\bigcup_{i=1}^{\infty} A_i$ is countable ①, thus $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
Therefore, $\mathcal{F}$ is a σ-field.
Then we show that P is a probability measure.

① Actually, since A_i are countable for all $i \in \mathbb{N}$, there exists a bijection $f_i : A_i \to \mathbb{N}$ for all $i \in \mathbb{N}$. Then $f : \mathbb{N}^2 \to \bigcup_{i=1}^{\infty} A_i$ defined by

$$f(j, k) = f_j^{-1}(k)$$

is onto, thus $|\mathbb{N}| \leq |\bigcup_{i=1}^{\infty} A_i| \leq |\mathbb{N}^2|$. Since $|\mathbb{N}| = |\mathbb{N}^2| = \aleph_0$, $|\bigcup_{i=1}^{\infty} A_i| = \aleph_0$

(1) $P(\Omega) = 1$. Since $\Omega = \mathbb{R}$ is uncountable, we have $P(\Omega) = 1$.
(2) If $A_1, A_2, \ldots$ is a collection of disjoint members of $\mathcal{F}$, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

- If there exists A_i that is uncountable, then $(\bigcup_{i=1}^{\infty} A_i)^c$ is countable, thus $P(\bigcup_{i=1}^{\infty} A_i) = 1$. For all $j \neq i$, we have $A_j \subseteq A_i^c$ since they are disjoint. So A_j is countable since A_i^c is countable. That said, $P(A_j) = 0$ for all $j \neq i$, and $\sum_{i=1}^{\infty} P(A_i) = P(A_i) = 1$.
- Otherwise, if all A_i are countable, then $\bigcup_{i=1}^{\infty} A_i$ is countable, thus $P(\bigcup_{i=1}^{\infty} A_i) = 0 = \sum_{i=1}^{\infty} P(A_i)$.

In both cases, $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
Therefore, $(\Omega, \mathcal{F}, P)$ is a probability space.

---

3 Birthday paradox

题目 1
Please design a **randomized algorithm using the birthday paradox** that solves the following problem in $\mathrm{poly}(n) \cdot 2^{n/2}$ time with high probability (for example, 0.99 when n is sufficiently large). Please provide a detailed error analysis as well. (**WARNING:** You will **NOT** receive any points if you solve this task using Gaussian elimination)

- Given an integer sequence $a_1, a_2, \ldots, a_{100n}$ of length $100n$ satisfying $0 \leq a_i < 2^n$ for all $1 \leq i \leq 100n$. Please find out a non-empty subset $S \subseteq \{1, 2, \ldots, 100n\}$ satisfying $\bigoplus_{i \in S} a_i = 0$, i.e. the exclusive or of the elements whose indices are in S equals to 0.

解答:
Here is a randomized algorithm using the birthday paradox to solve the problem.

Algorithm 1 Randomized Algorithm using Birthday Paradox

```
1: Partition {1, 2, ..., 100n} into 200n sets, i.e. S_i = {i·n/2+1, i·n/2+2, ..., (i+1)·n/2}
   for 0 ≤ i < 200.
2: for each set S_i do
3:     for each subset T ⊆ S_i do
4:         if ⊕_{j∈T} a_j = 0 then
5:             return the subset T
6:         end if
7:     end for
8: end for
```

Error Analysis:
We assume that each a_i is chosen uniformly at random from $\{0, 1, \ldots, 2^n - 1\}$.
Let A be the event that all the non-empty subsets of S_1 do not satisfy the condition, and A_i be the event that all the non-empty subsets of $\{1, 2, \ldots, i\}$ do not satisfy the condition. Then we have

$$\mathbf{Pr}(A) = \mathbf{Pr}(\bigcap_{i=1}^{n/2} A_i) = \prod_{i=1}^{n/2} \mathbf{Pr}(A_i | \bigcap_{j<i} A_j) = \prod_{i=1}^{n/2} \mathbf{Pr}(A_i | A_{i-1}) = \prod_{i=1}^{n/2} \left(1 - \frac{2^{i-1}}{2^n}\right),$$

where A_0 is certain event since $\emptyset$ has no non-empty subset.

$$\prod_{i=1}^{n/2} \left(1 - \frac{2^{i-1}}{2^n}\right) = \prod_{i=1}^{n/2} (1 - 2^{i-n-1}) = \prod_{i=1}^{n/2} e^{-2^{i-n-1}}$$

---

4 Conditional probability

题目 1 (Positive correlation)
We say that events B gives *positive information* about event A if $\mathbf{Pr}(A|B) > \mathbf{Pr}(A)$, that is, the occurrence of B makes the occurrence of A more likely. Now suppose that B gives positive information about A.

1. Does A give positive information about B?
2. Does $\overline{B}$ give negative information about A, that is, is it true that $\mathbf{Pr}(A|\overline{B}) < \mathbf{Pr}(A)$?
3. Does $\overline{B}$ give positive information or negative information about $\overline{A}$?

解答:

1. Since B gives positive information about A, we have $\mathbf{Pr}(A|B) > \mathbf{Pr}(A)$. Then by Bayes' theorem, we have
$$\mathbf{Pr}(B|A) = \frac{\mathbf{Pr}(A|B)\cdot\mathbf{Pr}(B)}{\mathbf{Pr}(A)} > \mathbf{Pr}(B)$$
Thus, A gives positive information about B.

2. By law of total probability, we have
$$\begin{aligned}\mathbf{Pr}(A|\overline{B}) &= \frac{1}{\mathbf{Pr}(\overline{B})}\left(\mathbf{Pr}(A) - \mathbf{Pr}(A|B)\mathbf{Pr}(B)\right)\\ &< \frac{1}{\mathbf{Pr}(\overline{B})}\left(\mathbf{Pr}(A) - \mathbf{Pr}(A)\mathbf{Pr}(B)\right)\\ &= \frac{1}{\mathbf{Pr}(\overline{B})}\mathbf{Pr}(A)(1-\mathbf{Pr}(B))\\ &= \frac{1}{\mathbf{Pr}(\overline{B})}\mathbf{Pr}(A)\mathbf{Pr}(\overline{B})\\ &= \mathbf{Pr}(A)\end{aligned}$$
Thus, $\overline{B}$ does give negative information about A.

3. Note that $\mathbf{Pr}(A|\overline{B}) = 1 - \mathbf{Pr}(\overline{A}|\overline{B})$ and $\mathbf{Pr}(A) = 1 - \mathbf{Pr}(\overline{A})$. Then we have
$$\mathbf{Pr}(\overline{A}|\overline{B}) = 1 - \mathbf{Pr}(A|\overline{B}) > 1 - \mathbf{Pr}(A) = \mathbf{Pr}(\overline{A})$$
Thus, $\overline{B}$ gives positive information about $\overline{A}$.

题目 2 (Balls in urns (I))
There are n urns of which the r-th contains $r-1$ white balls and $n-r$ black balls. You pick an urn uniformly at random (here, "uniformly" means that each urn has equal probability of being chosen) and remove two balls from that urn, uniformly at random without replacement (which means that each of the $\binom{n-1}{2}$ pairs of balls are chosen to be removed with equal probability). Find the following probabilities:

1. the second ball is black;
2. the second ball is black, given the first is black.

解答:

题目 3 (Balls in urns (II))
Suppose that an urn contains w white balls and b black balls. The balls are drawn from the urn one by one, each time uniformly and independently at random, without replacement (which means we do not put the chosen ball back after each drawing). Find the probabilities of the events:

1. the first white ball drawn is the $(k+1)$th ball;
2. the last ball drawn is white.

解答:

---

5 Independence

Let's consider a series of n outputs $(X_1, X_2, \cdots, X_n) \in \{0,1\}^n$ of n independent Bernoulli trials, where each trial succeeds with the same probability $0 < p < 1$.

题目 1 (Limited independence)
Construct three events A, B and C out of n Bernoulli trials such that A, B and C are pairwise independent but are not (mutually) independent. You need to prove that the constructed events A, B and C satisfy this. (Hint: Consider the case where $n = 2$ and $p = 1/2$.)

解答:

题目 2 (Product distribution)
Suppose someone has observed the output of the n trials, and she told you that precisely k out of n trials succeeded for some $0 < k < n$. Now you want to predict the output of the $(n+1)$-th trial while the parameter p of the Bernoulli trial is unknown. One way to estimate p is to find such $\hat{p}$ that makes the observed outcomes most probable, namely you need to solve

$$\arg\max_{\hat{p}\in(0,1)} \mathbf{Pr}_{\hat{p}}[k \text{ out of } n \text{ trials succeed}].$$

1. Estimate p by solving the above optimization problem.
2. If someone tells you exactly which k trials succeed (in addition to just telling you the number of successful trials, which is k), would it help you to estimate p more accurately? Why?

解答:

---

6 Probabilistic method

题目 1 (Tournaments)
A tournament can be interpreted as the outcome of a round-robin tournament in which every player faces every other player exactly once, and no draws occur. Given two players (vertices) x and y, we draw an arrow from x to y if x beats y (and vice

versa). We say a tournament has property S_k if for every k players, there exists another player v who defeats all of them. For example, a triangle $x \to y \to z \to x$ has property S_1 but not S_2. Prove that if

$$\binom{n}{k}(1-2^{-k})^{n-k} < 1,$$

then there is a tournament on n vertices that has the property S_k.

解答:

Let G be a random tournament on n vertices, where each of the $\binom{n}{2}$ edges is chosen independently with probability $1/2$, and E be the event that G has property S_k. To prove that there exists a tournament on n vertices that has the property S_k, it suffices to show that $\mathbf{Pr}(E) > 0$.

Let E_S be the event that the k players in S are defeated by another player and F_S be the event that k players in S are defeated by a specific player not in S.

We claim that $\mathbf{Pr}(F_S) = 2^{-k}$, this is because there are exactly k edges from the specific player to the k players in S, and each edge is chosen independently with probability $1/2$.

Then we have

$$\mathbf{Pr}(\overline{E_S}) = \left[\mathbf{Pr}(\overline{F_S})\right]^{n-k} = \left(1-2^{-k}\right)^{n-k}$$

Since there are $\binom{n}{k}$ ways to choose k players from n players, we have

$$\mathbf{Pr}(\overline{E}) = \mathbf{Pr}(\bigcup_{S\subseteq V(G),|S|=k} \overline{E_S}) \leq \sum_{S\subseteq V(G),|S|=k} \mathbf{Pr}(\overline{E_S}) = \binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1$$

Therefore, $\mathbf{Pr}(E) = 1 - \mathbf{Pr}(\overline{E}) > 0$, and there must exists a tournament on n vertices that has the property S_k.
