- These exercises are designed for practice and for you to think through the key probability concepts we have seen in the course.
- No solutions will be provided for these exercises.

Discrete Probability

1. **Non-uniform Random Selection**: You have a box containing balls numbered 1 through 10, where the probability of selecting ball i is proportional to i^2.
   (a) Find the probability of selecting each ball.
   (b) What is the probability of selecting an even-numbered ball?
2. **Birthday Problem Generalization**: In a room with n people, what is the probability that exactly two people share a birthday? Extend this to find the probability that there is at least one shared birthday among n people.

Conditional Probability

1. **Expected Conditional Value**: A six-sided die is rolled, and you win an amount equal to the number shown on the die if it is greater than 3. Otherwise, you win nothing. Let X be the amount won and A be the event $X > 0$. Find $\mathbb{E}[X \mid A]$.
2. **Conditional Probability in Random Selection**: A bag contains 4 white balls, 3 black balls, and 3 red balls. You draw two balls at random without replacement. Let A be the event that the first ball drawn is white and B be the event that the second ball drawn is red. Find $\Pr[B \mid A]$.

Independence of Random Variables

1. **Independence of Events**: Suppose two events A and B satisfy $\Pr[A] = 0.4$ and $\Pr[B] = 0.5$. What are the possible values for $\Pr[A \cap B]$ if A and B are independent?

2. **Independence in a Dice Roll**: Roll two fair six-sided dice. Let X be the outcome of the first die, and Y the outcome of the second die. Define A as the event "X is even" and B as the event "Y is greater than 3." Are A and B independent? Justify your answer.

3. **Independent Random Variables and Expectations**: Let X and Y be two independent random variables with means $\mathbb{E}[X] = 3$ and $\mathbb{E}[Y] = 4$. Calculate $\mathbb{E}[XY]$ and show that independence implies $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

Union Bound

1. **Union Bound on Dice Rolls**: Roll a fair six-sided die 10 times. Let A_i be the event that the i-th roll is a 6. Use the union bound to find an upper bound on the probability that at least one of the rolls results in a 6.

2. **Union Bound for Failure Probability**: A random process fails with probability 0.05 each time it runs. Using the union bound, calculate the probability that at least one failure occurs over 100 independent runs.

3. **Probability of intersection**: Let A and B be two events such that $\Pr[A] = 1 - \epsilon$ and $\Pr[B] = 1 - \delta$.

   - What is the probability of $A \cap B$ if A and B are independent?
   - Suppose that $\epsilon, \delta \leq 0.01$ and A and B are arbitrary events not necessarily independent. Show that $\Pr[A \cap B] \geq 0.98$.

Linearity of Expectation

1. **Basic Definitions**: State the law of total expectation and the linearity of expectation. What are the required assumptions for these to apply (assuming we have a discrete and finite probability space)?

2. **Expected Value of Distinct Dice Sums**: You roll two six-sided dice n times. Let X be the number of unique sums (from 2 to 12) observed across the n rolls. Calculate $\mathbb{E}[X]$ using the linearity of expectation and indicator random variables.

Pokemon Collection Problem

1. ~~**Generalized Pokemon Collection**: Recall the pokemon collection problem from the lectures: each time we go to the pokemon store, we get a uniformly random pokemon and our goals if to collect all n pokemons. Suppose there are m "rare" pokemons that each appear with probability ten times smaller than the other $n-m$ "common" pokemons. Each rare pokemon and each common pokemon is equally likely within its own category. Find the expected number of draws needed to collect all n pokemons at least once.~~ *This problem is significantly harder than it looks. Moral: Don't ask ChatGPT for practice problems!*

Simulating Biased Coins and Random Variables

1. **Simulating a Biased Coin Using a Fair Coin**: If you only have a fair coin, describe a method to simulate a coin flip with probability $p = 1/3$. Then, generalize your method for any $p \in (0,1)$.

2. **Generate a Random Variable with a Geometric Distribution Using a Biased Coin**: Suppose you have a biased coin that comes up heads with probability p. Use this coin to simulate a random variable X with a geometric distribution, where $P(X = k) = (1-p)^{k-1}p$. Calculate the expected number of coin flips needed to simulate X.

Pairwise Independence

1. **Constructing Pairwise Independent Random Variables**: Construct three random variables X, Y, Z such that any pair of them is independent, but X, Y, Z are not mutually independent.

2. **Expectation of Pairwise Independent Sums**: Let $X_1, X_2, \ldots, X_n$ be pairwise independent random variables, each taking values in $\{0,1\}$ with the probability of 1 being p. Calculate $\mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right]$ and determine whether the result differs from the case where the variables are fully independent.

Treaps and QuickSort

1. **Expected Depth in Treap**: Suppose you are building a treap with n elements. Show that the expected depth of any node in the treap is $O(\log n)$.

2. Consider the following algorithm **quicksortRestart**. Prove the following propositions assuming the analysis of randomized quicksort we saw in the lectures.

quicksortRestart
Input: array A

```
while TRUE do
    Run A' = quicksort(A) for 2 · Θ(n log n) steps,
                 where Θ(n log n) is the expected runtime for quicksort.
    if quicksort finishes then
        return A'
    end if
end while
```

Proposition 1 *quicksortRestart* runs in $\Theta(kn\log n)$ steps with probability $1-\frac{1}{2^k}$.

Proposition 2 *quicksortRestart* runs in $\Theta(c \cdot n\log^2 n)$ steps with probability $1-\frac{1}{n^c}$.

Tail Bounds

1. **Tail Bounds on Sum of Indicators**: Let $X_1, X_2, \dots, X_n$ be independent random variables with $P(X_i = 1) = p$ and $P(X_i = 0) = 1-p$. Define $S = \sum_{i=1}^{n} X_i$. Use Markov's inequality, Chebyshev's inequality and Chernoff bounds to upper bound $P(S \geq (1+\delta)np)$ for $\delta = 0.1$. Compare the estimates and review the conditions for applying each of these tail inequalities.

2. **Chernoff Bound on Tail of Binomial Distribution**: Let $X_1, X_2, \dots, X_n$ be independent random variables with $P(X_i = 1) = p$ and $P(X_i = 0) = 1-p$ and $X = \sum_i X_i$. Use the Chernoff bound to derive a tail bound for $P(X \geq k)$ where $k > np$.