Suggested reading: Meester, Chapter 2, §1.

Problems:

1. Consider the experiment of rolling a tetrahedral (four-sided) die twice. Find the probability mass functions for each of the following random variables (a table is fine—you don’t need to produce a formula!).

   (a) \( X = \) the sum of the rolls.
   (b) \( Y = \) the difference of the rolls.
   (c) \( M = \) the larger of the rolls.

2. A hand of five cards contains two aces and three kings. The five cards are shuffled, and then dealt face up one at a time until an ace appears.

   (a) Display in a table the probability mass function (PMF) for the number of cards dealt.
   (b) Suppose the dealing is continued until the second ace appears. Display the PMF for the total number of cards dealt.
   (c) Explain why the probabilities in the second table are just those in the first but in a different order. (Hint: Think about dealing from the bottom of the deck!)

3. The Benford pmf is a model for the leading digits of “naturally occurring” numbers.

   \[
   p_D(d) = P(D = d) = \log_{10} \left( 1 + \frac{1}{d} \right), \quad d = 1, 2, \ldots, 9.
   \]

   (a) Show that this is a valid pmf.
   (b) Show that the cumulative distribution function (CDF) is given by

   \[
   F_D(d) = P(D \leq d) = \log_{10}(d + 1), \quad d = 1, 2, \ldots, 9.
   \]

4. Ten dice are rolled. Five dice are red and five are green. Write unsimplified expressions for the probabilities of the events below.

   (a) Exactly four of the ten dice are sixes.
   (b) At least four of the ten dice are sixes.
   (c) Exactly two of the red dice are sixes and exactly three of the green dice are even numbers.
   (d) There are the same number of sixes among the red dice as among the green dice.
   (e) There are strictly more sixes among the red dice than among the green dice. (Hint: use (d) and symmetry.)
5. Now use R to find numerical values for the probabilities in the last problem.

6. Suppose that $X \sim \text{Geom}(p)$. Show that for all $m, k \geq 1$ we have

$$P(X = m + k \mid X > m) = P(X = k)$$

Explain why this property is called the lack of memory property of the geometric distribution.

7. In the World Series teams $A$ and $B$ play consecutive games, and the first to win four games wins the Series. Suppose that for each game the chance that team $A$ wins is $p$, independent of any other game.

(a) For $g = 4$ through 7, find a formula in terms of $p$ and $q = 1 - p$ for the probability that team $A$ wins the series in $g$ games.

(b) What is the probability that team $A$ wins the World Series.

(c) Let $X \sim \text{Binom}(7, p)$. Explain why $P(X \geq 4)$ is also equal to the probability that team $A$ wins the World Series.

(d) Let $G$ be the number of games played in the Series. What is the PMF of $G$?

8. In 1693 Samuel Pepys asked Isaac Newton to determine which was more likely: getting at least 1 six when 6 dice are rolled, at least 2 sixes when 12 dice are rolled, or at least 3 sixes when 18 dice are rolled.

Pepys originally thought the third case was the most likely, but Newton convinced him that the first had the best chance (leading Pepys to renege on a bet!).

(a) Use R to compare the 3 probabilities in question.

(b) Write a general analytical formula for the probability of $n$ or more sixes when $6n$ dice are rolled.

(c) Now write a function in R

```r
pepys <- function(n) {
    # your code here
}
```

that implements your formula. Use it to gather numerical evidence and make a conjecture about the limit as $n \to \infty$.

9. Fix $r$ and let $p(k), k = r, r + 1, \ldots$ denote the pmf of the negative binomial distribution with parameters $r$ and $p$; thus, $p(k)$ is the probability that the $r$th success in a sequence of Bernoulli($p$) trials occurs on the $k$th trial.

(a) Show that for $k \geq r + 1$,

$$\frac{p(k)}{p(k-1)} = \frac{(k - 1)}{(k - r)}q.$$

(b) Find a formula for the mode of the negative binomial distribution.

(c) For what values of $r$ and $p$ is there a double maximum? Which values of $k$ attain it.

(d) For what values of $r$ and $p$ does the mode occur at $k = r$?