1. **(Sanity Check!)** Define a random variable X to be the result of rolling two standard dice and summing the results.

   (a) Directly compute the expectation of X.

   Answer: The expectation of X is

$$\begin{aligned}\mathrm{E}[X] &= 2\times\Pr[X=2]+3\times\Pr[X=3]+\cdots+12\times\Pr[X=12]\\ &= \frac{2\times1+3\times2+4\times3+5\times4+6\times5+7\times6+8\times5+9\times4+10\times3+11\times2+12\times1}{36}\\ &= 7.\end{aligned}$$

   (b) Compute the expectation of X using linearity of expectation.

   Answer: Let Y be the random variable corresponding to the result of the first die and Z the result of the second die. Clearly,

$$\mathrm{E}[Y]=\mathrm{E}[Z]=\frac{1}{6}(1+2+3+4+5+6)=\frac{21}{6}=3.5$$

   Hence, $\mathrm{E}[X]=\mathrm{E}[Y+Z]=\mathrm{E}[Y]+\mathrm{E}[Z]=7$.

2. **(True or False?)** If X,Y are random variables, and $E[XY]=E[X]E[Y]$, then X and Y are independent.

   Answer: The converse is true, but the original statement is not true in general. Consider an example from note. Let X be a fair coin toss that we consider as taking values $+1$ and -1 equally likely. Suppose Y is an independent fair coin toss that takes values $+1$ and $+2$. The random variables X and Y are independent by construction.

   Let's consider a new random variable $Z=XY$. Is Z independent of Y? Obviously not. Z takes on four possible values $-2,-1,+1,+2$ and the magnitude of Z reveals exactly what Y is. We also know that $E[Y]=\frac{1}{2}(1+2)=\frac{3}{2}$, and $E[Z]=\frac{1}{4}(-2-1+1+2)=0$, so $E[Y]E[Z]=0$. However,

$$\begin{aligned}E[YZ] &= \sum_y\sum_z yzPr[Y=y,Z=z]\\ &= \frac{1}{4}(1(-1)+1(1)+2(-2)+2(2))\\ &= \frac{1}{4}(-1+1-4+4)\\ &= 0,\end{aligned}$$

   which means $E[YZ]=E[Y]E[Z]=0$, but Y and Z are not independent. Hence, this is a counterexample to the original claim, and the statement must be false.

3. **(Geometric Distribution)** Coco is trying to collect four coupons, C_1, C_2, C_3, C_4. She does not know which coupon(s) she will get before paying. There are two stores selling coupons:

   - In Store A, she pays 3 dollars to get a coupon, and the probability of getting each coupon is always $\frac{1}{4}$.
   - In Store B, C_1 and C_2 are always packaged together, and C_3 and C_4 are also always packaged together. She pays 8 dollars to get one package (with two coupons), and the probability of getting each pair (either (C_1, C_2) or (C_3, C_4)) is always $\frac{1}{2}$.

   Which store should she go, or it does not matter?

   Answer: Store B.

   Let $A = A_1 + A_2 + A_3 + A_4$ where A_i is the number of tries until Coco gets the i-th new coupon. Similarly, $B = B_1 + B_2$. In this case, there are only 2 distinct packages.

$$\mathbb{E}(A) = \sum_{i=1}^{4} \mathbb{E}(A_i) = \frac{4}{4} + \frac{4}{3} + \frac{4}{2} + \frac{4}{1} = \frac{25}{3};$$

$$\mathbb{E}(B) = \sum_{i=1}^{2} \mathbb{E}(B_i) = \frac{2}{2} + \frac{2}{1} = 3.$$

   Coco should go to Store B because $3 \cdot \frac{25}{3} > 8 \cdot 3$.

4. **(Conditional Expectation)** Suppose the number of children in a family is a random variable X with mean μ, and given $X = n$ for $n \geq 1$, each of the n children in the family is a girl with probability p and a boy with probability $1 - p$. What is the expected number of girls in a family?

   Answer: Intuitively, the answer should be $p\mu$. To show this is correct, let G be the random number of girls in a family. Given $X = n$, G is the sum of n indicators of events with probability p, so

$$\mathrm{E}[G|X = n] = np$$

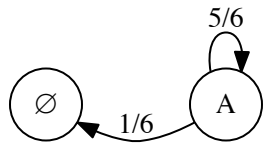
   Note that this is correct even for $n = 0$. By conditioning on X,

$$\begin{aligned}
\mathrm{E}[G] &= \sum_n \mathrm{E}[G|X = n]\mathrm{P}[X = n] \\
&= p \sum_n n \cdot \mathrm{P}[X = n] \\
&= p\mu.
\end{aligned}$$

5. **(James Bond)** James Bond is playing a game that involves repeatedly rolling a fair standard 6-sided die.

   (a) What is the expected number of rolls until he gets a 5?

   Answer: Let X be the random variable denoting the number of rolls until 007 gets a 5, so $E[X]$ is the expected number of rolls. Notice that he must roll at least once. On his first try, he gets a 5 with probability $\frac{1}{6}$, and fails with probability $\frac{5}{6}$, which means he has to start all over again (but with one additional roll under his belt).



The question can be represented by the following Markov Chain (where state A represents "last roll not a 5", and state $\emptyset$ represents game termination): Starting from state A, we want to find the expected rolls until he hits $\emptyset$.

By the reasoning above, we can express $E[X]$ in a recursive fashion and solve for it.

$$\begin{aligned} E[X] &= \frac{1}{6} \cdot 1 + \frac{5}{6}(E[X]+1) \\ &= \frac{1}{6} + \frac{5}{6} + \frac{5}{6}E[X] \\ \frac{1}{6}E[X] &= 1 \\ E[X] &= 6 \end{aligned}$$

which means the expected number of rolls until 007 gets a 5 is $\boxed{6}$. Notice that it doesn't matter what the roll value he wants is (any other number between 1 and 6 that is not 5), it is expected that Bond tries 6 times before he gets the desired number.

You could solve this question by noting that the number of rolls is a geometric random variable with parameter $p = \frac{1}{6}$, and since we know that the expected value of a geometric random variable is $\frac{1}{p}$, you will also get the same correct answer of 6.

You could also solve this question by writing out the sum of a geometric series and calculating that sum, but this method would take the longest amount of time, given that this is only a 3-point question.

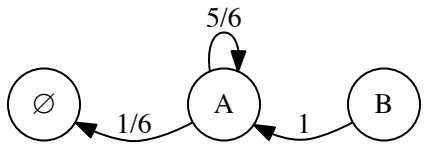
The recursive approach will be the most useful way to think about this problem, as we will see in the next two parts.

(b) What is the expected number of rolls until the last two rolls sum to 7?

Answer: Intuitively, the answer to this question is the same as part (a), plus one. Why? First, notice that no matter what your "first" roll is, there is only a $\frac{1}{6}$ chance of making the sum of that roll and the roll immediately after it equal 7. This is because every number between 1 and 6 has a unique "partner" that sums to 7, e.g. 1 will be paired with 6, 2 with 5, etc. Second, notice that 007 cannot have a pair of rolls until he has rolled at least twice, so we need to always roll once first, and then we can calculate the desired expected number of rolls starting from the second roll.

Again, this can be modeled by the following Markov chain, where state A represents at least one previous roll has taken place, and state B represents the start of the game (no previous rolls). We want the expected rolls to hit $\emptyset$, starting from state B.

Mathematically, let β be the expected number of rolls, and α be the expected number of rolls after the first roll. This implies that $\beta = 1 + \alpha$, as explained above. To calculate α, note that



there are two cases. The second roll of the pair either "matches" the first one (so their sum is 7) with probability $\frac{1}{6}$ or it doesn't (with probability $\frac{5}{6}$). Either way it costs 007 an additional roll to find out. In the first case, the game ends, so the additional expected number of rolls is 0. In the second case, the additional expected number of rolls is just α, since this so-called "failed" second roll is now the first roll and basically waits for the next roll to match it.

In other words, $\alpha = 1 + \frac{5}{6}\alpha$. Solving for α gives 6 as expected. However, recall that we are interested in the number of rolls, and we must always roll once to actually start the game (since the event that we're interested in is the sum of the last two rolls). Therefore, the expected number of rolls until the last two rolls sum to 7 is $\beta = 1 + \alpha = 1 + 6 = \boxed{7}$.