Homework 6 (due Saturday, December 7, 2024)

1. (2.5 points) **Hitting time of Gambler's Ruin**

   Recall that in gambler's ruin, the player starts with k gold coins. For each turn, the player wins 1 gold coin with probability p, stays the same with probability r, and loses 1 gold coin with probability q, where $p + q + r = 1$. The game ends when the number of gold coins hit either a or b for the first time, $a < k < b$.

   Let H_k be the average duration of play (i.e. number of turns from start until the game ends) starting with k gold coins. H_k can be computed by the boundary conditions, $H_a = 0$ and $H_b = 0$ and the recurrence relation,

   $$H_k = 1 + pH_{k+1} + rH_k + qH_{k-1}.$$

   For the following exercises, we fix $p = q = r = \dfrac{1}{3}$.

   a) For the givens a, b, write the program to compute H_k for $k = a, a+1, \ldots, b-1, b$.
   b) Show your result of part a) with $a = 5$ and $b = 15$.
   c) Use program `GuessPol` to help you to guess the formula of H_k as a function k, a, b.

2. (4 points) **Coupon Collector**

   a) Generalize the program to simulate `Coupon2(n)` the number of boxes you need to buy to complete **TWO** sets of toys (say, one set for you and the other for your brother). What is the expectation? Discuss the results.
   b) Is the result in (a) same as the program `Coupon(2n)`? Discuss.
   c) Generalize the program to simulate `CouponK(n)` the number of boxes to buy to complete **K** sets of toys. What is the expectation? Discuss.

3. `(total 2.5 points + extra credit 0.5 point)` Recall that in Lecture 18 we did simulations of symmetric random walks for a number of steps and counted the number of points which lie outside the parabola region. Let us denote the random walk by $\{S_n\}_{n=0}^{\infty}$, where $S_0 = 0$, and

   $$S_n = X_1 + X_2 + \cdots + X_n,$$

   where $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with

   $$X_i = \begin{cases} +1 & \text{with probability } 1/2, \\ -1 & \text{with probability } 1/2. \end{cases}$$

   For each nonnegative n, let

   $$W_n := \frac{1}{n+1} \cdot \# \left\{ k \in \{0, 1, \ldots, n\} \; : \; |S_k| > 1.96 \cdot \sqrt{k} \right\}.$$

   That is, W_n is the fraction of points *outside* the parabola region.

   Feel free to write a computer program to help solve the following problems.

a) What is $\mathbb{E}(W_3)$? (You should be able to compute this explicitly.)

b) What is $\mathbb{E}(W_4)$? (You should be able to compute this explicitly.)

c) Prove that

$$\mathbb{E}(W_{10}) = \frac{17}{512} \approx 3.3203\,\%.$$

d) Show that

$$\mathbb{E}(W_{100}) \approx 4.8391\,\%.$$

e) Show that

$$\mathbb{E}(W_{1000}) \approx 4.9787\,\%.$$

f) `(extra credit + 0.5 point)` Determine, with proof, the limit

$$\lim_{n\to\infty} \mathbb{E}(W_n).$$

Hint 1. The answer is *not* 5.00000 %. Perhaps you might provide the answer in terms of a value of some function we have seen in this class.

Hint 2. Stolz–Cesàro theorem.

4. `(extra credit + 2.5 points)` Let $p \in (0,1)$. Consider the random walk where you start at 0 and in each step there is a probability p of going one step to the right and a probability $1-p$ of going one step to the left. Let us denote by $\alpha(p)$ the probability that we visit some negative integer at least once. The goal of this problem is to analyze $\alpha(p)$.

We model this random walk as follows. Let $X_1, X_2, \ldots$ be i.i.d. random variables where

$$X_i = \begin{cases} +1 & \text{with probability } p, \\ -1 & \text{with probability } 1-p, \end{cases}$$

and for each n, the location after n steps is

$$S_n := X_1 + X_2 + \cdots + X_n.$$

By convention, define $S_0 = 0$. Using these notation, we can write $\alpha(p)$ as

$$\alpha(p) := \mathbb{P}\{\exists\, n \geq 0 \,:\, S_n < 0\}\,.$$

a) If $p < 1/2$, argue why $\alpha(p) = 1$.

b) If $p = 1/2$, argue why $\alpha(p) = 1$.

Now let us assume for the rest of this problem that $p > 1/2$. Let us examine the value of $\alpha(p)$ using first-step analysis.

To that end, for each $m \in \mathbb{Z}$, define a_m to be the probability that we visit some negative integer at least once *if* the random walk starts at m instead. (So if $m = 0$, the probability $a_m = a_0$ is the same as $\alpha(p)$.)

If $m < 0$ is a negative integer, then the random walk starts at a negative number, so $a_m = 1$.

c) For each nonnegative integer m, argue why

$$a_m = p \cdot a_{m+1} + (1-p) \cdot a_{m-1}.$$

Now for each nonnegative integer m, let us write

$$\Delta_m := a_m - a_{m-1}.$$

From part c), we find that for every nonnegative integer m,

$$(1-p) \cdot \Delta_m = p \cdot \Delta_{m+1}.$$

This means

$$\Delta_{m+1} = r \cdot \Delta_m,$$

where $r := \frac{1-p}{p} \in (0,1)$. (Recall that we are assuming $p > 1/2$.) Hence, for every nonnegative integer m, we have

$$\Delta_m = r^m \cdot \Delta_0 = r^m \cdot (a_0 - a_{-1}) = r^m \cdot (a_0 - 1).$$

d) Argue why for every nonnegative integer m, we have

$$a_m = \Delta_m + \Delta_{m-1} + \cdots + \Delta_1 + \Delta_0 + 1,$$

and therefore,

$$a_m = \frac{1 - r^{m+1}}{1-r} \cdot (a_0 - 1) + 1.$$

Now let us assume we know that

$$\lim_{m \to \infty} a_m = 0.$$

From the formula above, we obtain

$$\lim_{m \to \infty} a_m = \lim_{m \to \infty} \frac{1 - r^{m+1}}{1-r} \cdot (a_0 - 1) + 1.$$

Since $0 < r < 1$, the above equation becomes

$$0 = \frac{a_0 - 1}{1-r} + 1.$$

e) Conclude that

$$\alpha(p) = a_0 = r = \frac{1-p}{p}$$

The argument above uses the assumption that $\lim_{m \to \infty} a_m = 0$ for which we did not provide a proof. This might be unsatisfactory. Let us describe a different approach to prove that

$$\alpha(p) = \frac{1-p}{p},$$

now using a gambler's ruin type argument.

For each positive integer m, let us denote by τ_m the first time where the random walk is at either -1 or m. That is,

$$\tau_m := \min\{t \geq 0 \,:\, S_t \in \{-1, m\}\}.$$

We continue to assume $p > 1/2$, and let $r = (1-p)/p$ be as above.

f) Prove that
$$\mathbb{P}\{X_{\tau_m} = m\} = \frac{1-r}{1-r^{m+1}}.$$

Let E_m denote the event $\{X_{\tau_m} = m\}$ from part f).

g) Argue why we have the following descending chain of events:
$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots.$$

h) Argue why the event
$$\bigcap_{i=1}^{\infty} E_i$$
is the equivalent to the event that the random walk never visits any negative number.

Now we invoke a lemma from measure theory.[1] Suppose that

$$H_1 \supseteq H_2 \supseteq H_3 \supseteq \cdots$$

is a descending chain of events. Then the probability that *all events* in the chain occur is
$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} H_i\right) = \lim_{i\to\infty} \mathbb{P}(H_i).$$

i) Conclude that the probability that the random walk never visits any negative number is $1-r$.

j) Conclude that the probability that the random walk visits a negative number at least once is
$$r = \frac{1-p}{p},$$
whence $\alpha(p) = \frac{1-p}{p}$.

Yay! We are done.

[1]which is actually not too hard to prove, but people seem to be afraid of measure theory at this point.