Notes for Tutorial 12

6.042 - May 9 and May 12

There are two topics, each accompanied by a group problem.

- Two useful formulas.
- Analyzing random processes.

1 Two Useful Formulas

In lecture, we saw two important new formulas. First, we have an alternate way to compute expectated values:

Fact 1 *For a random variable X that takes on values in $\mathbb{N}$:*

$$Ex[X] \quad = \quad \sum_{k=0}^{\infty} \Pr\{X > k\}$$

Many experiments consist of a sequence of independent trials, where each trial either succeeds or fails.

Fact 2 *Suppose there is a sequence of independent trials and each trial succeeds with probability p. Then the expected number of trials needed to get one success is $1/p$.*

1.1 Short Problems

Problem 1: Suppose that we roll a fair, six-sided die until we get a five. What is the expected number of rolls? Each trial succeeds with probabilty $\frac{1}{6}$, so the expected number of trials to get one success is $1/(1/6) = 6$.

Problem 2: Suppose that we roll a fair, six-sided die until we get a total of three fives. What is the expected number of rolls? By linearity of expectation, this is the expected number of rolls to get the first five, plus the expected number of rolls to get the second, plus the expected number of rolls to get the third. That is, $6 + 6 + 6 = 18$.

Problem 3: What is the probability of actually getting three or more fives within 18 rolls?

$$1 - \binom{18}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{16} - \binom{18}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{17} - \binom{18}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{18} \quad \approx \quad 0.5973$$

1.2 Group Problem

There are two people. Simultaneously, they both begin rolling fair, independent dice. Each person stops as soon as he or she rolls a five for the first time. What is the expected number of rolls made by the person who keeps going longer?

Let the random variables X_1 and X_2 be the number of rolls made by the two players, and let $X = \max(X_1, X_2)$. Starting with Fact 1, we have:

$$\begin{aligned}
\mathrm{Ex}[X] &= \sum_{k=0}^{\infty} \Pr\{X > k\} \\
&= \sum_{k=0}^{\infty} \Pr\{X_1 > k \cup X_2 > k\} \\
&= \sum_{k=0}^{\infty} 1 - \Pr\{X_1 \le k \cap X_2 \le k\} \\
&= \sum_{k=0}^{\infty} 1 - \Pr\{X_1 \le k\} \cdot \Pr\{X_2 \le k\} \\
&= \sum_{k=0}^{\infty} 1 - \Pr\{X_1 \le k\}^2 \\
&= \sum_{k=0}^{\infty} 1 - (1 - \Pr\{X_1 > k\})^2 \\
&= \sum_{k=0}^{\infty} 1 - \left(1 - \left(\frac{5}{6}\right)^k\right)^2 \\
&= \sum_{k=0}^{\infty} 2\left(\frac{5}{6}\right)^k - \left(\frac{5}{6}\right)^{2k} \\
&= 2\sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^k - \sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^{2k} \\
&= \frac{2}{1-\frac{5}{6}} - \frac{1}{1-\left(\frac{5}{6}\right)^2} \\
&= 12 - \frac{36}{11} \\
&= \frac{96}{11} \\
&= 8.727272\ldots
\end{aligned}$$

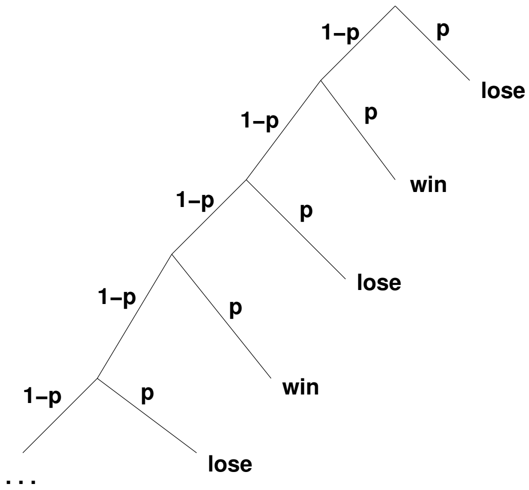
2 Analyzing Random Processes

Sometimes you're confronted with a random process involving a sequence of events. For example, in lecture we saw the "Russian roulette duel". Then you need to find the probability of some event, such as the event that the first guy gets shot. Here's a good approach to such problems, using techniques shown briefly in lecture.

1. Start drawing a tree diagram. Of course, the tree may be infinite, so stop as soon as the overall structure of the tree is apparent.

2. There are then two ways to approach computing the probability of the event.

   (a) Sometimes, you can express the probability of the event you're considering as an infinite sum. For example, in lecture we had a tree that looked like this:

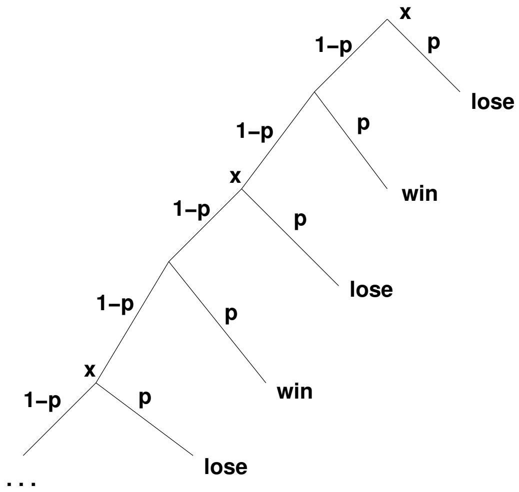
   

   We wanted the probability that the first player loses, which is:

   $$p + p(1-p)^2 + p(1-p)^4 + p(1-p)^6 + \ldots \quad = \quad \frac{1}{2-p}$$

   (b) Sometimes a direct summation is very messy. In this case, look for an *infinite subtree that contains itself.* Create a variable to represent the probability that your event occurs, starting from the root of the subtree. Mark the root of each instance of the subtree with that variable. Repeat this process until your tree contains no infinite path of unmarked vertices. Then you should be able to express your variables in terms of each other, solve the resulting equations, and determine the probability of your event.

   For the tree in lecture, the entire tree contains itself. We mark each vertex that is a root of a copy of this tree with the variable x.



There are now no infinite unmarked paths. We get the equation:

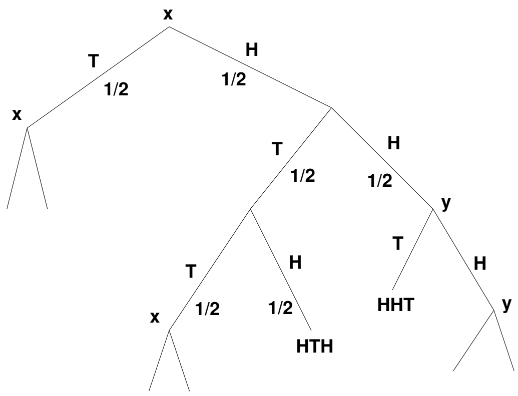
$$x = p + (1-p)^2 x$$

Solving, we find $x = 1/(2-p)$. This is the probability that the first player loses starting at the root of the tree, which is the event we're interested in.

2.1 Group Problem

Suppose that you start flipping a fair coin. What is the probability that the sequence HHT comes up before the sequence HTH?

We draw the tree diagram. (Note that once HH is observed, HHT is guaranteed to appear before HTH.) The entire tree is a subtree of itself. We mark the root of each copy with an x. There is still an infinite unmarked path, so we identify another subtree that contains itself and mark the root of each copy with a y. This gives:



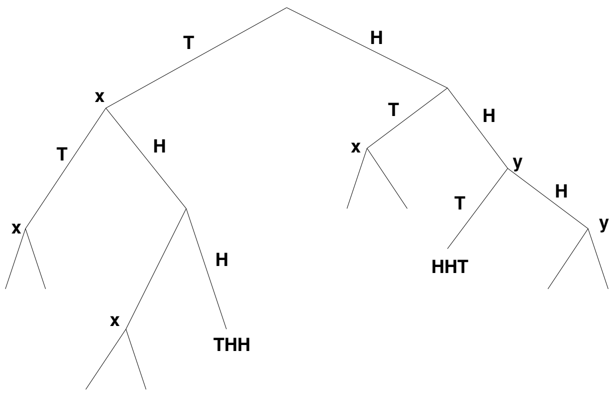
There are now no infinite subtrees. We get the equations:

$$\begin{aligned} x &= \frac{1}{2} \cdot x + \frac{1}{8} \cdot x + \frac{1}{4} \cdot y \\ y &= \frac{1}{2} + \frac{1}{2} y \end{aligned}$$

Solving, we find that $y = 1/4$ and so $x = 2/3$. The probability that HHT precedes HTH is $x = 2/3$.

What is the probability that THH comes up before HHT?

We draw a tree diagram and mark vertices as follows:



This gives the equations:

$$\begin{aligned} x &= \frac{1}{2} \cdot x + \frac{1}{4} \cdot x + \frac{1}{4} \\ y &= \frac{1}{2} \cdot y \end{aligned}$$

Solving, we find that $x = 1$ and $y = 0$. (We could have had an insight and deduced that this must be so, but it's good to have a mechanical method to fall back on!) The probability that THH precedes HHT is not:

$$\begin{aligned} \Pr\{\text{THH precedes HHT}\} &= \frac{1}{2} \cdot x + \frac{1}{4} \cdot x + \frac{1}{4} \cdot y \\ &= \frac{3}{4} \end{aligned}$$

In fact, there is no best length-three sequence; each is preceded by another with probability greater than 1/2.