UW Math Circle December 6, 2006

Here is the "teaser" problem from last week:

Exercise 1. Each time you buy a box of Cheeri-Omegas cereal, you win a prize! There are n types of prizes, and the cereal company puts a random type of prize into each box. How many boxes of cereal will you need to buy before you expect to have at least one prize of each type?

What is a reasonable guess here — $2n$? n^2? something else? To solve this problem, in addition to some new probability theory, we will need to look at some special kinds of series (sums).

Also — here are some nice geometric probability problems that can be solved by using only last week's techniques.

Exercise 2. Arnold and Bethula both agree to meet at the mall between 2 PM and 3 PM. Each one shows up at a random time in that interval, but they pick their arrival times independently. Let's assume that each one is willing to wait 20 minutes for their friend, after which they will leave. What's the probability that they will meet? (Hint: how is this geometric?)

Exercise 3. Pick 3 points at random on the circumference of a circle. What is the probability that they form an acute triangle?

1 Random Variables

Last week, we talked about random events, but another way to view probability theory is through the use of random variables. A random variable X represents a numerical value associated with the outcome of some experiment.

For example, if X is the random variable corresponding to the value rolled on a fair die, then X can take on any of the values $\{1,2,\ldots,6\}$, and

$$Pr[X=1]=1/6, Pr[X=2]=1/6,\ldots,Pr[X=6]=1/6.$$

Just like events, a pair X,Y of random variables is either *independent* or *dependent*. We say that the variables are independent if and only if

$$\text{for all } x \text{ and } y,\ Pr[X=x \text{ and } Y=y]=Pr[X=x]\cdot Pr[Y=y]. \tag{1}$$

To say that two variables are dependent just means they are not dependent. Informally, variables are independent if the value of one does not affect the value of the other.

An example of two independent variables is to throw a red die and a blue die, let X_{red} be the number on top of the red die, and let X_{blue} be the number on top of the blue die. Then X_{red} and X_{blue} are independent. An example of two dependent variables is to throw one die, let X_{top} be the number on the top of the die, and let X_{bottom} be the number on the bottom of the die. Then X_{top} and X_{bottom} are independent.

Exercise 4. Prove, using the definition of dependence given in (1), that X_{top} and X_{bottom} are dependent.

Random variables can be added, multiplied, thrown into functions and generally treated exactly like usual variables, but the value of the resulting expression becomes probabilistic, i.e., its value will differ in each trial of the experiment.

For example, using the dice "probability table" from last week, we see

$$Pr[X_{blue}+X_{red}=2]=1/36, \qquad Pr[X_{blue}+X_{red}=7]=6/36.$$

On the other hand (since the opposite faces of a die always add up to 7), we have

$$Pr[X_{top}+X_{bottom}=7]=1.$$

1.1 Expected Values

The *expectation (or expected value)* $E[X]$ of a random variable X is defined to be its average value:

$$E[X]=\sum_x x\cdot Pr[X=x],$$

where the sum is taken over all possible values x that X can take on.

For example, the expected value of rolling a die is

$$1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=\frac{7}{2}.$$

Exercise 5. Consider an unfair coin with probability p of coming up heads. Let C be 1 if the coin comes up heads, and 0 if it comes up tails. What is $E[C]$?

Exercise 6. What is the expected value of $(X_{red} + X_{blue})$? What about $(X_{top} + X_{bottom})$?

The *central limit theorem* says that (under some technical conditions), if you repeat an experiment many times and measure a variable X, then the average value of X will approach its expectation in the limit. (The proof of this is quite hard.) More precisely, suppose that on the ith trial of the experiment, the value of X is x_i; then the central limit theorem says that

$$\lim_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} = E[X].$$

In gambling and games, expected values are very useful since they tell you, roughly speaking, who will make money in the long run. For example, suppose I charge you \$4 to roll a die, and then I give you \$ X where X is the number on top. Because the *expected* number on top of the die is 3.5, you are expected to make a net profit of −\$4+\$3.5 = −\$0.5. In other words, it is bad for you to play it.

Exercise 7 ("The St. Petersburg Paradox"). Consider the following game. I give you a fair coin, and you flip it until it comes up heads. (So for example the sequence of flips could be T,T,T,T,T,T,H.) Let n be the number of times the coin came up tails; I will pay you 2^n dollars. What is the expected amount of money that you will win? In other words, what is a fair price to charge for this game?

1.2 Linearity of Expectation

The following useful fact will help us to crack the Cheeri-Omegas problem.

Property 1. *If X and Y are any random variables, then*

$$E[X+Y] = E[X] + E[Y].$$

Proof. The proof basically consists of two parts: first, we need a formula for $Pr[X+Y=i]$ in terms of the quantities $Pr[X = j \text{ and } Y = k]$; second we will do some algebraic manipulation.

How can $X+Y=i$ be true? Well, it is the same as saying that $X = j$ for some integer j, and $Y = i-j$. Hence

$$Pr[X+Y=i] = \sum_j Pr[X = j \text{ and } Y = i-j].$$

Using the definition of expectation, we get

$$E[X+Y] = \sum_i i \cdot \sum_j Pr[X = j \text{ and } Y = i-j].$$

Rearranging, we get

$$\begin{aligned} E[X+Y] &= \sum_i (j + (i-j)) \cdot \sum_j Pr[X = j \text{ and } Y = i-j] \\ &= \sum_i \sum_j jPr[X = j \text{ and } Y = i-j] + \sum_i \sum_j (i-j)Pr[X = j \text{ and } Y = i-j]. \end{aligned} \quad (2)$$

The first term in the right-hand side of Equation (2) can be rewritten as follows:

$$\sum_i \sum_j jPr[X = j \text{ and } Y = i-j] = \sum_j j \sum_i Pr[X = j \text{ and } Y = i-j] = \sum_j jPr[X = j] = E[X].$$

We similarly want to show that the second term on the right-hand side of Equation (2) is $E[Y]$. Let $k = i - j$, then we get

$$\sum_i \sum_j (i-j)Pr[X = j \text{ and } Y = i-j] = \sum_k \sum_j kPr[X = j \text{ and } Y = k].$$

Then we proceed as before:

$$\sum_k \sum_j kPr[X = j \text{ and } Y = k] = \sum_k k \sum_j Pr[X = j \text{ and } Y = k] = \sum_k kPr[Y = k] = E[Y].$$

So the right-hand side of Equation (2) is equal to $E[X] + E[Y]$. □

2 Beyond Arithmetic and Geometric Series

You are probably familiar with the formulas for arithmetic and geometric series, which are:

$$a+(a+d)+(a+2d)+\cdots+(a+kd)=(k+1)\cdot\frac{a+(a+kd)}{2} \qquad a+ar+ar^2+ar^3+\cdots+ar^k=a\cdot\frac{r^{k+1}-1}{r-1}.$$

Here a is the initial term, there are $(k+1)$ terms in total, and d and r are respectively the common difference and ratio. In sigma notation the above formulas are

$$\sum_{i=0}^{k}(a+id)=(k+1)\cdot\frac{a+(a+kd)}{2} \qquad \sum_{i=0}^{k}a\cdot r^i=a\cdot\frac{r^{k+1}-1}{r-1}.$$

By taking the limit as $k\to\infty$, we get a formula for the sum of an *infinite* geometric series, provided that $-1<r<1$:

$$a+ar+ar^2+\cdots=a/(1-r), \qquad \text{or} \qquad \sum_{i=0}^{\infty}ar^i=a/(1-r).$$

Today we will need to work with a series that is both arithmetic *and* geometric. What do we mean by this? Well, in a arithmetic series the ith term is $a+id$, in a geometric series the ith term is $a\cdot r^i$, but let's define an arithmetic-geometric sequence to be one whose ith term is $r^i\cdot(a+id)$. In a sense, it's like we have *both* a common ratio and a common difference. (The convention I am using is that in all of these sequences, the initial term a corresponds to setting $i=0$.)

So, what is the formula for $\sum_{i=0}^{\infty}r^i\cdot(a+id)$?

One proof goes along the following lines — but let's do it for geometric series first because it is simpler. Let S denote our finite geometric series,

$$S:=a+ar+ar^2+\cdots+ar^{k-1}+ar^k.$$

Something special happens when we multiply S by r : since $r\cdot ar^i=ar^{i+1}$, it appears that all of the terms get "shifted." This is what I mean:

$$\begin{array}{rcccccccccccccc} S &=& a &+& ar &+& ar^2 &+& ar^3 &+& \cdots &+& ar^{k-1} &+& ar^k \\ rS &=& & & ar &+& ar^2 &+& ar^3 &+& \cdots &+& ar^{k-1} &+& ar^k &+& ar^{k+1} \end{array}$$

Now if we subtract the two series, we get

$$\begin{array}{rrcccccccccccccccc} & S &=& a &+& ar &+& ar^2 &+& ar^3 &+& \cdots &+& ar^{k-1} &+& ar^k \\ + & -rS &=& & & -ar &+& -ar^2 &+& -ar^3 &+& \cdots &+& -ar^{k-1} &+& -ar^k &+& -ar^{k+1} \\ \hline & S-rS &=& a & & & & & & & & & & & & &+& -ar^{k+1} \end{array}$$

So we get $S-rS=a-ar^{k+1}$, or if you rearrange and solve for S, then $S=a\cdot(r^{k+1}-1)/(r-1)$ which was what we expected.

Exercise 8. Determine a closed form expression for

$$\sum_{i=0}^{k}r^i\cdot(a+id)$$

using this technique. If $-1<r<1$ then what is the limit as $k\to\infty$?

Exercise 9. It is known that $\sum_{i=0}^{k}i^2=i(i+1)(2i+1)/6$. Hence determine a closed form expression for

$$\sum_{i=0}^{k}r^i\cdot i^2$$

using this technique. If $-1<r<1$ then what is the limit as $k\to\infty$?

2.1 The Harmonic Series

The series $1+1/2+1/3+\cdots$ comes up so frequently in mathematics that it gets a special name, *the harmonic series.* We define the nth *harmonic number*, denoted H_n, to be the sum of the first n terms of the harmonic series:

$$H_n:=\sum_{i=1}^{n}\frac{1}{i}=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{n}.$$

The *Euler-Mascheroni constant*, which is denoted by γ, is a real number whose value is approximately 0.577. Although we would need some calculus and other fancy stuff to prove it, it turns out that $\ln(n)+\gamma$ is a pretty good approximation to H_n. (Here ln represents logarithm to the base $e \approx 2.718$.) In fact as $n \to \infty$ the difference between these values goes to zero. Here are some small examples:

n	H_n	$\ln(n)+0.577$
3	1.833	1.676
10	2.929	2.880
30	3.995	3.978
100	5.187	5.182

Using elementary methods, we can instead show a simpler version of the bounds which argues that H_n is "roughly" $\log(n)$.

Exercise 10. Show that $H_n \le \log_2(n)+1$. Also show that $H_n \ge (1+\log_2(n))/2$.

3 Cheeri-Omegas Solution

Consider a particular trial of the Cheeri-Omegas experiment. Let T_1 be the number of boxes before you get your first type of prize (i.e., T_1 is always 1). After you get your first type of prize, let T_2 be the number of additional boxes before you get your second distinct type of prize. In general, for each $i < n$, after you get your ith distinct type of prize, let T_{i+1} be the number of additional boxes you buy before you get your $(i+1)$th distinct type of prize.

Note that the total number T of boxes that you need to buy before you get all prizes is exactly $(T_1+T_2+\cdots+T_n)$. So by linearity of expectation,

$$E[T] = E[\sum_{i=1}^{n} T_i] = \sum_{i=1}^{n} E[T_i].$$

Next, we need to compute $E[T_i]$ for each i. Note that T_i can be equivalently defined in the following way:

> "Take a fair coin with probability $(i-1)/n$ of coming up tails. Let T_i be the number of flips you need to perform before you get your first head."

(Here each coin flip corresponds to buying a box of cereal, tails represents a type of prize we already have, and heads represents a new type of prize.) For convenience let p_i denote $(i-1)/n$. We can explicitly compute $Pr[T_i = j]$: this is the probability that the first $(j-1)$ flips are tails, and then the jth is heads, which is

$$Pr[T_i = j] = p_i^{j-1} \cdot (1-p_i).$$

So by the definition of expectation,

$$E[T_i] = \sum_{j=1}^{\infty} j \cdot p_i^{j-1} \cdot (1-p_i).$$

This is just an arithmetic-geometric series, and we can simplify it to

$$E[T_i] = 1/(1-p_i).$$

So in total we expect to buy

$$E[T] = \sum_{i=1}^{n} E[T_i] = \sum_{i=1}^{n} 1/(1-p_i) = \sum_{i=1}^{n} n/(n-i+1) = n\sum_{i=1}^{n} 1/(n-i+1) = nH_n$$

boxes.

Here are two last (unrelated and very hard) problems for you:

Exercise 11 ("Gambler's Ruin"). Consider a line of squares numbered $(0,1,\ldots,N)$. You start on square k. On each turn you flip a coin; if it is heads you move left (decrease your position by 1), and if it is tails you move right (increase your position by 1). Stop once you hit 0 or N. What is the probability that you end at N? (Answer: k/N.) Show that the expected number of steps until you finish is $k\cdot(N-k)$.

Exercise 12. Every time you go to the store, you buy a 1 gram sack of "Dr. Demento's Magic Mud." This mud is very special, because it is a mixture of water, dirt and gold! However, the mud is very inconsistent; in every sack, the amount (in kilograms) of gold is a random variable between 0 and 1, with all values equally likely. The question is, in expectation, how many trips to the store will you make before you have at least 1 gram of gold? (Answer: e. Hint: use the formula $e^x = \sum_{i=0}^{\infty} x^i/(i!)$ from last week.)