Massachusetts Institute of Technology
6.042J/18.062J: Mathematics for Computer Science
Professor David Karger

Tutorial 11
May 5, 1997

Solutions for Tutorial 11

Problem 1. Warmup

Let $A_1, \ldots, A_n$ be a collection of disjoint sets such that their union is equal to the sample space. The following is true:

$P(B) = P(B \cap A_1) + \ldots + P(B \cap A_n)$

Use this to show that if $A_1, \ldots, A_n$ is a disjoint collection that covers the sample space, then for any events B, C such that $P(C) \neq 0$

$P(B \mid C) = P(B \cap A_1 \mid C) + ... + P(B \cap A_n \mid C)$

Observe the useful corollary:

$P(B \mid C) = P(B \cap A \mid C) + P(B \cap \bar{A} \mid C)$

Recall that

$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)}$$

$$\begin{aligned}
P(B \mid C) &= \frac{P(B \cap C)}{P(C)} \\
&= \frac{P(B \cap C \cap A_1) + \ldots + P(B \cap C \cap A_n)}{P(C)} \\
&= \frac{P(B \cap A_1 \cap C) + \ldots + P(B \cap A_n \cap C)}{P(C)} \\
&= P(B \cap A_1 \mid C) + \ldots + P(B \cap A_n \mid C)
\end{aligned}$$

Problem 2. *In* Polya's Urn Model *we begin with an urn having b black and r red balls. Let $c > 0$ be an arbitrary natural number. At each time t ($t = 1, 2, \cdots$), we draw one ball from the urn, and put $c + 1$ balls of the same color into the urn. (So there are $b + r + c$ balls in the urn afterwards).*

Show that the probability of drawing a sequence of t balls, so that the last ball is black is $b/(b + r)$.

(Hint: argue that the probability of a sequence is independent of the ordering. $P(BRB) = P(RBB)$, etc. Then the probability that a sequence starts with B is the same as the probability that the sequence ends with B for all sequences of length t.)

First, show that if we have a sequence of n picks from the urn, and m of them turn out to be black balls, that the probability that we choose m black and n red balls does not depend on the order in which we choose them:

The probability of choosing the ith black ball at the jth pick, is $\frac{b+(i-1)c}{b+r+(j-1)c}$. Similarly, the probability of choosing the ith red ball at the jth pick is $\frac{r+(i-1)c}{b+r+(j-1)c}$. In both cases, the total number of balls to select from on the jth pick is $b+r+(j-1)c$ which is not dependent on what's been picked in the past. Also, the number of black balls to select from only depends on how many black balls have been selected in the past, not on when they were seleted.

The probability, then, of choosing m black balls and $n-m$ red balls is

$$\frac{(b)(b+c)(b+2c)\dots(b+(m-1)c)\cdot(r)(r+c)(r+2c)\dots(r+(n-m-1)c)}{(b+r)(b+r+1)(b+r+2)\dots(b+r+(n-1)c)}$$

The numerator of the fraction can be rearranged in any way, so the probability of choosing a sequence of m black balls and $n-m$ red balls is irrelevant of the order in which they're chosen.

Now, enumerate every possible set of ways of choosing n black and red balls, and the probability of choosing that sequence. We want to select all of the sequences which end with a black ball. Note however, that since the probability of choosing a sequence only depends on how many black balls are in the sequence, not the order in which they're chosen, that the probabilty of choosing a sequence is equal to the probability of choosing the reverse of the sequence (ie, $RBBRB$ has the same probability as $BRBBR$). Upon reversing, the set of all sequences ending with a black ball becomes the set of all sequences starting with a black ball. We know this probability is $\frac{b}{b+r}$.

Problem 3. Mathematicians

(a) *Suppose that A and C are to fight in a pistol duel. We know that A's chance of hitting his target is 0.3 and C's is 0.5. Assuming that C shoots first, what's the probability that A survives?*

P_{CA}(A survives) $=$ P(C misses A)(P(A hits C) $+$ P(A misses C)P(C misses A)$\dots$))

$P_{CA}(A) = 0.5[0.3 + (0.7)(0.5)[0.3 + (0.7)(0.5)[\dots]]]$

We can solve this as a sum $P_{CA}(A) = (0.5)\sum_{i=0}^{\infty}(0.3)^{i+1}(0.5 \times 0.7)^i$

or $P_{CA}(A) = (0.5)(0.3 + 0.7P_{CA}(A))$ so $P_{CA}(A) = \frac{0.15}{0.65} = 0.23$.

Now suppose that B enters the fray and A, B, and C participate in a three-cornered duel. A's and C's chances of hitting are same as before while B never *misses. They are to fire at their choice of target (assume B and C shoot at the more "dangerous" adversary if there is a choice) in succession in the order A, B, C cyclically until one man is left standing.*

(b) What's the probability that A survives if he targets B first?

If A hits B, then we have the shooting match we saw above with C and A. If A misses B, then B kills C, and A has one shot at B before B kills him.

$P(\text{A survives}) = P(\text{A misses B})(P(\text{B hits C})P(\text{A hits B})) + P(\text{A hits B})P_{CA}(\text{A survives}) = 0.7((1.0)(0.3)) + (0.3)(0.23) = 0.27.$

(c) What's the probability that A survives if he targets C first?

A can survive only if he misses C, B kills C, and then A kills B. If A hits C the first time, then B will kill A.

$P(\text{A survives}) = P(\text{A misses C})P(\text{B hits C})P(\text{A hits B}) = (0.7)(1.0)(0.3) = 0.21$

(d) If A (and A alone) has the option of firing into the air (targeting no one) on the first round, what's the probability that he survives?

A kills no one, B kills C and then A gets one shot at B.

$P(\text{A survives}) = P(\text{B hits C})P(\text{A hits B}) = 0.3$

(e) What should A's strategy be in order to maximize his chances of surviving?

Shoot into the air.

Problem 4. EXTRA!

The IRS has decided to audit k tax forms among those submitted. It doesn't know how exactly how many forms have been submitted (although it has many more than k). It will will make its selection in a one-time pass in chronological order over the received tax forms. Since it does not want to provide any incentive for taxpayers to file their forms at particular times, it wants each tax form to be equally likely to be audited.

It begins by tentatively selecting the first k forms received. As it visits each subsequent form, it has the option to replace one of the k currently selected forms by the latest one, so that it always has a selection of exactly k forms.

Devise a randomized strategy for the IRS, so that at any time during the one-time pass, every tax return visited to that point in the pass is equally likely to be in the current audit list.

Hint: Consider $k = 1$, and then try to generalize.