Massachusetts Institute of Technology
6.042J/18.062J, Fall '02: Mathematics for Computer Science
Professor Albert Meyer and Dr. Radhika Nagpal

Solutions to In-Class Problems — Week 10, Fri

Problem 1. You are organizing a neighborhood census and instruct your census takers to knock on doors and note the sex of any child that answers the knock. Assume that there are two children in a household and that girls and boys are equally likely to be children and to open the door.

A sample space for this experiment has outcomes that are triples whose first element is either B or G for the sex of the elder child, likewise for the second element and the sex of the younger child, and whose third coordinate is E or Y indicating whether the *e*lder child or *y*ounger child opened the door. For example, (B, G, Y) is the outcome that the elder child is a boy, the younger child is a girl, and the girl opened the door.

(a) Let T be the event that the household has two girls, and O be the event that a girl opened the door. List the outcomes in T and O.

Solution. $T = \{GGE, GGY\}, O = \{GGE, GGY, GBE, BGY\}$ ■

(b) What is the probability $\Pr\{T \mid O\}$, that both children are girls, given that a girl opened the door?

Solution. $1/2$ ■

(c) Where is the mistake in the following argument?

> If a girl opens the door, then we know that there is at least one girl in the household. The probability that there is at least one girl is
>
> $$1 - \Pr\{\text{both children are boys}\} = 1 - (1/2 \times 1/2) = 3/4. \tag{1}$$
>
> So,
>
> $$\begin{aligned} &\Pr\{T \mid \text{there is at least one girl in the household}\} && (2)\\ &= \frac{\Pr\{T \cap \text{there is at least one girl in the household}\}}{\Pr\{\text{there is at least one girl in the household}\}} && (3)\\ &= \frac{\Pr\{T\}}{\Pr\{\text{there is at least one girl in the household}\}} && (4)\\ &= (1/4)/(3/4) = 1/3. && (5)\end{aligned}$$

> Therefore, given that a girl opened the door, the probability that there are two girls in the household is $1/3$.

Solution. The argument is a correct proof that

$$\Pr\{T \mid \text{there is at least one girl in the household}\} = 1/3.$$

The problem is that the event, H, that the household has at least one girl, namely,

$$H ::= \{\texttt{GGE}, \texttt{GGY}, \texttt{GBE}, \texttt{GBY}, \texttt{BGE}, \texttt{BGY}\},$$

is not equal to the event, O, that a girl opens the door. These two events differ:

$$H - O = \{\texttt{BGE}, \texttt{GBY}\},$$

and their probabilities are different. So the fallacy is in the final conclusion where the value of $\Pr\{T \mid H\}$ is taken to be the same as the value $\Pr\{T \mid O\}$. Actually, $\Pr\{T \mid O\} = 1/2$. ■

(WE DIDN'T GET TO THE NEXT TWO PROBLEMS IN CLASS ON FRIDAY.)

Problem 2. Suppose there are 100 people in a room. Assume that their birthdays are independent and uniformly distributed. As stated in lecture notes, with probability $> 99\%$ there will be two that have the same birthday.

Now suppose you find out the birthdays of all the people in the room except one—call her "Jane"—and find all 99 dates to be different.

(a) What's wrong with the following argument:

> With probability greater than 99%, some pair of people in the room have the same birthday. Since the 99 people we asked all had different birthdays, it follows that with probability greater than 99% Jane has the same birthday as some other person in the room.

Solution. Here's the problem with the argument: Let A be the event that some two people in the room have the same birthday. Let B be the event that the 99 people we asked all have different birthdays.

It is true that $\Pr\{A\} > 0.99$ (that is indeed the probability spoken about in the lecture notes). However, that is the *a priori* probability, i.e., assuming all the birthdays are uniform and independent, with no other constraints. The argument above makes the erroneous assumption that event A has probability of at least 99% even once we know that event B holds. But once we know that event B holds, the birthdays are no longer independent. Thus $\Pr\{A \mid B\}$ is not necessarily equal to $\Pr\{A\}$ (in our case, they are actually quite different, as will be computed in part (b)). ■

(b) What is the actual probability that Jane has the same birthday as some other person in the room?

Solution. Let S be the set of birthdays of the 99 people in the room other than Jane. By assumption, $|S| = 99$. Let b be the date of Jane's birthday. Since b is uniformly distributed over a set of size 365, and b is independent of all the birthdays in S, we have

$$\Pr\{A \mid B\} = \Pr\{b \in S\} = \frac{|S|}{365} = \frac{99}{365} \approx 27.1\%,$$

where B is the event that the 99 people we asked all have different birthdays. ■

Problem 3. Suppose you repeatedly flip a fair coin until you see the sequence HHT or the sequence HTT. What is the probability you see the sequence HTT before you see the sequence HHT? *Hint:* Try to find the probability that HHT comes before HTT conditioning on whether you first toss an H or a T. The answer is not $1/2$.

Solution. Let A be the event that HTT appears before HHT, and let $p ::= \Pr\{A\}$.

Suppose our first toss is T. Since neither of our patterns starts with T, the probability that A will occur from this point on is still p. That is, $\Pr\{A \mid \texttt{T}\} = p$.

Suppose our first toss is H. To find the probability that A will now occur, that is, to find $r ::= \Pr\{A \mid \texttt{H}\}$, we consider different cases based on the subsequent throws.

Suppose the next toss is H, that is, the first two tosses are HH. Then neither pattern appears if we continue flipping H, and when we eventually toss a T, the pattern HHT will then have appeared first. So in this case, event A will never occur. That is $\Pr\{A \mid \texttt{HH}\} = 0$.

Suppose the first two tosses are HT. If we toss a T again, then we have tossed HTT, so event A has occurred. If we next toss an H, then we have tossed HTH. But this puts us in the same situation we were in after rolling an H on the first toss. That is, $\Pr\{A \mid \texttt{HTH}\} = r$.

Summarizing this we have:

$$\Pr\{A\} = \Pr\{A \mid \texttt{T}\}\Pr\{\texttt{T}\} + \Pr\{A \mid \texttt{H}\}\Pr\{\texttt{H}\} \qquad \text{(Law of Total Probability)} \tag{6}$$

$$p = p\frac{1}{2} + r\frac{1}{2} \qquad \text{so} \tag{7}$$

$$p = r. \tag{8}$$

Continuing, we have

$$\Pr\{A \mid \texttt{H}\} = \Pr\{A \mid \texttt{HT}\}\Pr\{\texttt{T}\} + \Pr\{A \mid \texttt{HH}\}\Pr\{\texttt{H}\} \qquad \text{(Law of Total Probability)} \tag{9}$$

$$r = \Pr\{A \mid \texttt{HT}\}\frac{1}{2} + 0 \cdot \frac{1}{2} \tag{10}$$

$$\Pr\{A \mid \texttt{HT}\} = \Pr\{A \mid \texttt{HTT}\}\Pr\{\texttt{T}\} + \Pr\{A \mid \texttt{HTH}\}\Pr\{\texttt{H}\} \qquad \text{Law of Total Probability} \tag{11}$$

$$\Pr\{A \mid \texttt{HT}\} = 1 \cdot \frac{1}{2} + r\frac{1}{2} \tag{12}$$

$$r = \left(\frac{1}{2} + \frac{r}{2}\right)\frac{1}{2} \qquad \text{by (10) \& (12)} \tag{13}$$

$$r = \frac{1}{3}. \tag{14}$$

So HTT comes before HHT with probability $p = r = \frac{1}{3}$.

It is an amazing fact that these kind of events are another example of *intransitivity* of the kind we saw with some special dice in Notes 10. Namely, if you may pick *any* pattern of three tosses such as HTT, then I can pick a pattern of three tosses such as HHT. If we then bet on which pattern will appear first in a series of tosses, the odds will be in my favor. In particular, even if you instead picked the "better" pattern HHT, there is another pattern I can pick that has a more than even chance of appearing before HHT. So this intransitivity phenomenon comes up for real—you could make money on it betting in a bar—and not just for some contrived dice. ■