Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2009)

Problem Set 2: Solutions
Due September 23, 2009

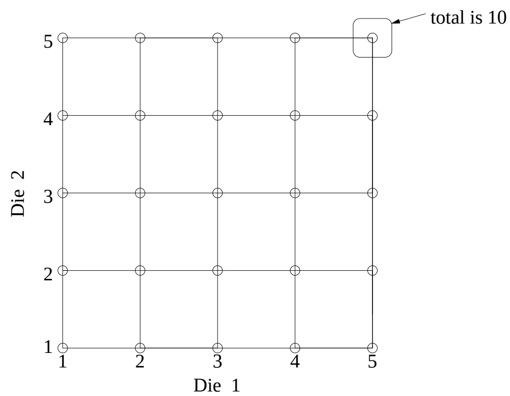
1. (a) We need to consider the sample space. The sample space is the set of ordered pairs (x_1, x_2) with $1 \leq x_1, x_2 \leq 6$. Each point in the sample space is equally likely, and there are 6 "favorable" outcomes, hence the probability of doubles is $\boxed{\frac{1}{6}}$.

   (b) Once we are told that the sum is 4 or less, our sample space becomes

$$\{(1,1); (1,2); (1,3); (2,1); (2,2); (3,1)\}$$

   and hence the probability of doubles is $\boxed{\frac{1}{3}}$.

2. (a) i. $\boxed{\text{No}}$



Overall, there are 25 different outcomes in the sample space. For a total of 10, we should get a 5 on both rolls. Therefore $A \subset B$, and

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1$$

We observe that to get at least one 5 showing, we can have 5 on the first roll, 5 on the second roll, or 5 on both rolls, which corresponds to 9 distinct outcomes in the sample space. Therefore

$$\mathbf{P}(B) = \frac{9}{25} \neq \mathbf{P}(B|A)$$

ii. $\boxed{\text{No}}$ Given event A, we know that both roll outcomes must be 5. Therefore, we could not have event C occur, which would require at least one 1 showing. Formally, there are 9 outcomes in C, and

$$\mathbf{P}(C) = \frac{9}{25}$$

But

$$\mathbf{P}(C|A) = 0 \neq \mathbf{P}(C)$$

(b) i. $\boxed{\text{No}}$ Out of the total 25 outcomes, 5 outcomes correspond to equal numbers in the two rolls. In half of the remaining 20 outcomes, the second number is higher than the first one. In the other half, the first number is higher than the second. Therefore,

$$\mathbf{P}(F) = \frac{10}{25}$$

There are eight outcomes that belong to event E:

$$E = \{(1,2),(2,3),(3,4),(4,5),(2,1),(3,2),(4,3),(5,4)\}.$$

To find $\mathbf{P}(F|E)$, we need to compute the proportion of outcomes in E for which the second number is higher than the first one:

$$\mathbf{P}(F|E) = \frac{1}{2} \neq \mathbf{P}(F)$$

ii. $\boxed{\text{Yes}}$ Conditioning on event D reduces the sample space to just four outcomes

$$\{(2,5),(3,4),(4,3),(5,2)\}$$

which are all equally likely. It is easy to see that

$$\mathbf{P}(E|D) = \frac{2}{4} = \frac{1}{2}, \qquad \mathbf{P}(F|D) = \frac{2}{4} = \frac{1}{2}, \qquad \mathbf{P}(E \cap F|D) = \frac{1}{4} = \mathbf{P}(E|D)\mathbf{P}(F|D)$$

3. (a)

$$\mathbf{P}(\text{find in A and in A}) = \mathbf{P}(\text{in A}) \cdot \mathbf{P}(\text{find in A}|\text{in A}) = 0.4 \cdot 0.25 = 0.1$$

$$\mathbf{P}(\text{find in B and in B}) = \mathbf{P}(\text{in B}) \cdot \mathbf{P}(\text{find in B}|\text{in B}) = 0.6 \cdot 0.15 = 0.09$$

Oscar should search in Forest A first.

(b) Using Bayes' Rule,

$$\begin{aligned}\mathbf{P}(\text{in A}|\text{not find in A}) &= \frac{\mathbf{P}(\text{not find in A}|\text{in A}) \cdot \mathbf{P}(\text{in A})}{\mathbf{P}(\text{not find in A}|\text{in A}) \cdot \mathbf{P}(\text{in A}) + \mathbf{P}(\text{not find in A}|\text{in B}) \cdot \mathbf{P}(\text{in B})} \\ &= \frac{(0.75)\cdot(0.4)}{(0.4)\cdot(0.75)+(1)\cdot(0.6)} = \frac{1}{3}\end{aligned}$$

(c) Again, using Bayes' Rule,

$$\begin{aligned}\mathbf{P}(\text{looked in A}|\text{find dog}) &= \frac{\mathbf{P}(\text{find dog}|\text{looked in A}) \cdot \mathbf{P}(\text{looked in A})}{\mathbf{P}(\text{find dog})} \\ &= \frac{(0.25)\cdot(0.4)\cdot(0.5)}{(0.25)\cdot(0.4)\cdot(0.5)+(0.15)\cdot(0.6)\cdot(0.5)} = \frac{10}{19}\end{aligned}$$

(d) In order for Oscar to find the dog, it must be in Forest A, not found on the first day, alive, and found on the second day. Note that this calculation requires conditional independence of not finding the dog on different days and the dog staying alive.

$$\begin{aligned}\mathbf{P}(\text{find live dog in A day 2}) &= \mathbf{P}(\text{in A}) \cdot \mathbf{P}(\text{not find in A day 1}|\text{in A}) \\ &\quad \cdot \mathbf{P}(\text{alive day 2}) \cdot \mathbf{P}(\text{find day 2}|\text{in A}) \\ &= 0.4 \cdot 0.75 \cdot (1 - \frac{1}{3}) \cdot 0.25 = 0.05\end{aligned}$$

4. (a) Suppose we choose old widgets. Before we choose any widgets, there are $1000 \cdot 0.15 = 150$ defective old widgets. The probability that we choose two defective widgets is

$$\begin{aligned}\mathbf{P}(\text{two defective}|\text{old}) &= \mathbf{P}(\text{first is defective}|\text{old}) \cdot \mathbf{P}(\text{second is defective}|\text{first is defective, old}) \\ &= \frac{150}{1000}\frac{149}{999} = 0.02247\end{aligned}$$

Now let's consider the new widgets. Before we choose any widgets, there are $1500 \cdot 0.05 = 75$ defective old widgets. Similar to the calculations above,

$$\begin{aligned}\mathbf{P}(\text{two defective}|\text{new}) &= \mathbf{P}(\text{first is defective}|\text{new}) \cdot \mathbf{P}(\text{second is defective}|\text{first is defective, new}) \\ &= \frac{75}{1500}\frac{74}{1499} = 0.002568\end{aligned}$$

By the total probability law,

$$\begin{aligned}\mathbf{P}(\text{two defective}) &= \mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) \\ &\quad + \mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new}) \\ &= \frac{1}{2} \cdot 0.02247 + \frac{1}{2} \cdot 0.002568 = 0.01243.\end{aligned}$$

Note that this number is very close to what we would get if we ignored the effects of removing one defective widget before choosing the second widget:

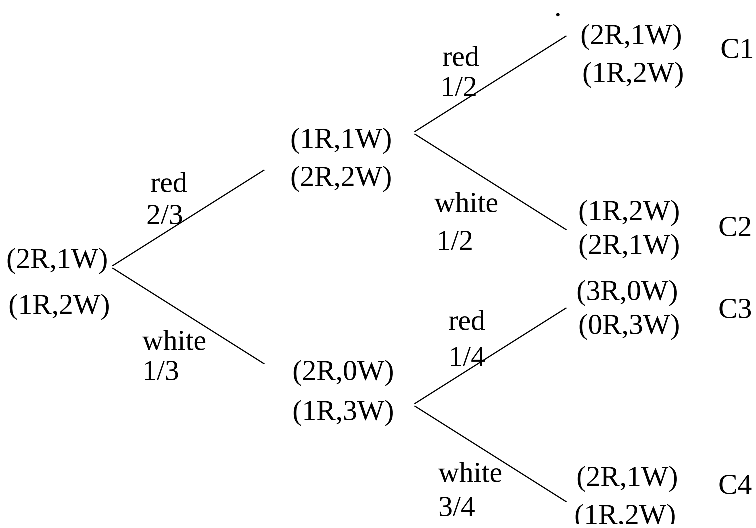
$$\begin{aligned}\mathbf{P}(\text{two defective}) &= \mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) \\ &+ \mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new}) \\ &\approx \frac{1}{2} \cdot 0.15^2 + \frac{1}{2} \cdot 0.05^2 = 0.0125.\end{aligned}$$

(b) Using Bayes' rule,

$$\begin{aligned}\mathbf{P}(\text{old}|\text{two defective}) &= \frac{\mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old})}{\mathbf{P}(\text{old}) \cdot \mathbf{P}(\text{two defective}|\text{old}) + \mathbf{P}(\text{new}) \cdot \mathbf{P}(\text{two defective}|\text{new})} \\ &= \frac{\frac{1}{2} \cdot 0.02247}{\frac{1}{2} \cdot 0.02247 + \frac{1}{2} \cdot 0.002568} = 0.8974\end{aligned}$$

5. A tree diagram will be helpful for this problem. Initially, there are two red and one white balls in bin 1, which we denote by (2R,1W). Similarly, bin 2 has (1R,2W). When we reach into bin 1, we pull out a red ball with probability 2/3. From the initial configuration of red and while balls, the system may have evolved in four possible ways. We call these cases C_1, C_2, C_3, and C_4, as shown in the tree diagram. Case C_1 refers to both balls transferred being red, C_2 refers to red followed by a white, etc., corresponding to the four branches of the tree above. From the tree diagram, we deduce

$$\mathbf{P}(C_1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}, \qquad \mathbf{P}(C_2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}, \qquad \mathbf{P}(C_3) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}, \qquad \mathbf{P}(C_4) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}.$$



Using the Total Probability Theorem in the universe created by selecting bin 1 or bin 2,

$$
\begin{aligned}
\mathbf{P}(\text{red}|\text{bin 1 chosen}) &= \sum_{i=1}^{4} \mathbf{P}(\text{red}|\text{bin 1 chosen}, C_i)\mathbf{P}(C_i|\text{bin 1 chosen}) \\
&= \sum_{i=1}^{4} \mathbf{P}(\text{red}|\text{bin 1 chosen}, C_i)\mathbf{P}(C_i) = \frac{2}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{1}{3} + 1\cdot\frac{1}{12} + \frac{2}{3}\cdot\frac{1}{4} = \frac{7}{12} \\
\mathbf{P}(\text{red}|\text{bin 2 chosen}) &= \sum_{i=1}^{4} \mathbf{P}(\text{red}|\text{bin 2 chosen}, C_i)\mathbf{P}(C_i) = \frac{1}{3}\cdot\frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3} + 0\cdot\frac{1}{12} + \frac{1}{3}\cdot\frac{1}{4} = \frac{5}{12}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{P}(\text{red}) &= \mathbf{P}(\text{red}|\text{bin 1 chosen})\mathbf{P}(\text{bin 1 chosen}) + \mathbf{P}(\text{red}|\text{bin 2 chosen})\mathbf{P}(\text{bin 2 chosen}) \\
&= \frac{7}{12}\cdot\frac{1}{2} + \frac{5}{12}\cdot\frac{1}{2} = \frac{1}{2}
\end{aligned}
$$

Finally,

$$
\mathbf{P}(\text{bin 1 chosen}|\text{red}) = \frac{\mathbf{P}(\text{red}|\text{bin 1 chosen})\cdot\mathbf{P}(\text{bin 1 chosen})}{\mathbf{P}(\text{red})} = \frac{7}{12}
$$

Note that this is slightly larger than $\frac{1}{2}$.

6. (a) We proceed as follows:

$$
\begin{aligned}
\mathbf{P}(A \cap (B \cup C)) &= \mathbf{P}((A \cap B) \cup (A \cap C)) \\
&= \mathbf{P}(A \cap B) + \mathbf{P}(A \cap C) - \mathbf{P}(A \cap B \cap C) \\
&\overset{*}{=} \mathbf{P}(A)\mathbf{P}(B) + \mathbf{P}(A)\mathbf{P}(C) - \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) \\
&= \mathbf{P}(A)\left[\mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(B)\mathbf{P}(C)\right] \\
&= \mathbf{P}(A)\mathbf{P}(B \cup C),
\end{aligned}
$$

where the equality marked with $*$ follows from the independence of A, B, and C.

(b) Proof 1: If A and B are independent, then A^c and B^c are also independent (see Problem 1.43, page 63 for the proof).

For any two independent events U and V, DeMorgan's Law implies

$$\begin{aligned}\mathbf{P}(U \cup V) &= \mathbf{P}((U^c \cap V^c)^c) = 1 - \mathbf{P}(U^c \cap V^c) = 1 - \mathbf{P}(U^c) \cdot \mathbf{P}(V^c) \\ &= 1 - (1 - \mathbf{P}(U))(1 - \mathbf{P}(V)).\end{aligned}$$

We proceed to prove the statement by induction. Letting $U = A_1$ and $V = A_2$, the base case is proven above. Now we assume that the result holds for any n and show that it holds for $n+1$. For independent $\{A_1, \ldots, A_n, A_{n+1}\}$, let $B = \cup_{i=1}^{n} A_i$. It is easy to show that B and A_{n+1} are independent. Therefore,

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_{n+1}) &= 1 - (1 - \mathbf{P}(B)) \cdot (1 - \mathbf{P}(A_{n+1})) \\ &= 1 - \prod_{i=1}^{n+1} (1 - \mathbf{P}(A_i)),\end{aligned}$$

which completes the proof.

Proof 2: Alternatively, we can use the version of the DeMorgan's Law for n events:

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_n) &= \mathbf{P}((A_1^c \cap A_2^c \cap \ldots \cap A_n^c)^c) \\ &= 1 - \mathbf{P}(A_1^c \cap A_2^c \cap \ldots \cap A_n^c).\end{aligned}$$

But we know that $A_1^c, A_2^c, \ldots, A_n^c$ are independent. Therefore

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2 \cup \ldots \cup A_n) &= 1 - \mathbf{P}(A_1^c)\mathbf{P}(A_2^c) \ldots \mathbf{P}(A_n^c) \\ &= 1 - \prod_{i=1}^{n} (1 - \mathbf{P}(A_i)).\end{aligned}$$