# Notes for Recitation 10 (4/17/98)

## 1 A Dice Problem

There is a game involving three dice and two players. The dice are not normal; rather, they are numbered as shown in Figure 1. Each hidden face has the same number as the opposite, exposed face. As a result, each die has only three distinct numbers, and each number comes up $\frac{1}{3}$ of the time.

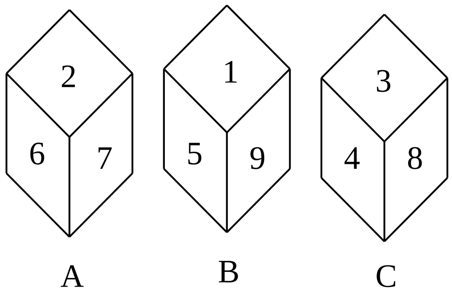


**Figure 1**: *This figure shows the strange numbering of the three dice in the Dice Problem. The number on each concealed face is the same as the number on the exposed, opposite face.*

In the game, the first player can choose any one of the three dice. Then the second player chooses one of the two remaining dice. They both roll and the player with the higher number wins. Which of the three dice should player one choose? That is, which of the three dice is best?

For example, die B is attractive, because it has a 9, the highest number overall; on the other hand, it also has a 1, the lowest number. Intuition gives no clear answer! We can solve the problem with the same four-step method used in the Monty Hall problem.

**Claim 1.1** Die A beats die B more than half of the time.

*Proof.* The claim concerns the experiment of throwing dice A and B.

*Step 1: Find the Sample Space.* The sample space for this experiment is indicated by the tree diagram in Figure 2.

*Step 2: Define Events of Interest.* We are interested in the event that die A comes up greater than die B. The outcomes in this event are marked "A" in the figure.

*Step 3: Compute Outcome Probabilities.* To find outcome probabilities, we first assign probabilities to edges in the tree diagram. Each number comes up with probability $\frac{1}{3}$, regardless of the value of the other die. Therefore, we assign all edges probability $\frac{1}{3}$. The probability of an outcome is the product of probabilities on the corresponding root-to-leaf path; this means that every outcome has probability $\frac{1}{9}$.

*Step 4: Compute Event Probabilities.* The probability of an event is the sum of the probabilities of the outcomes in the event. Therefore, the probability that die A comes up greater than die B is $\frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{5}{9}$.

As claimed, the probability that die A beats die B is greater than half.

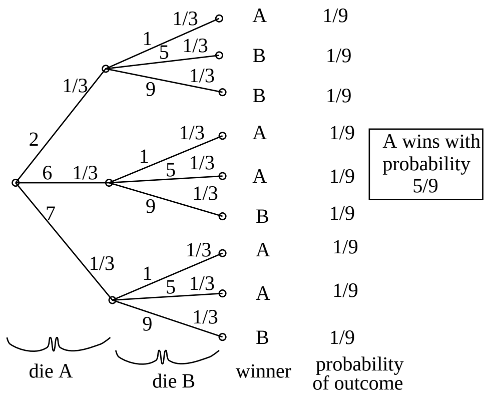


**Figure 2**: *This is the tree diagram arising when die A is played against die B. Die A beats die B with probability $\frac{5}{9}$.*

■

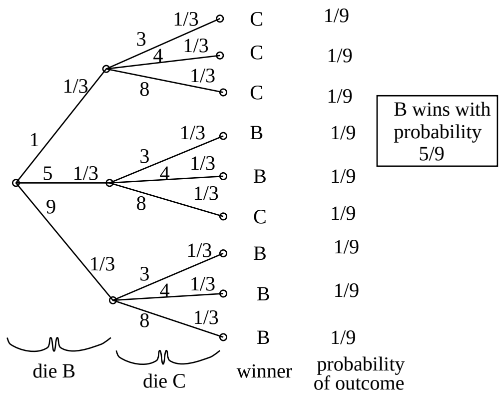


**Figure 3**: *Die B beats die C with probability $\frac{5}{9}$.*

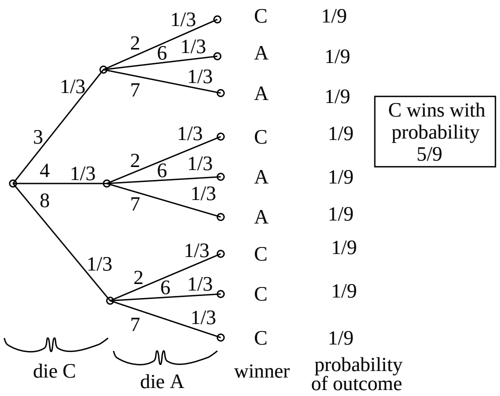


**Figure 4**: *Die C beats die A with probability $\frac{5}{9}$. Amazing!*

**Claim 1.2** Die B beats die C more than half of the time.

*Proof.* See the tree diagram in Figure 3; the proof is the same as for the preceding claim. ∎

We have shown that A beats B and that B beats C. From these results, we might conclude that A is the best die, B is second best, and C is worst. But this is totally wrong!

**Claim 1.3** Die C beats die A more than half of the time!

*Proof.* See the tree diagram in Figure 4. ∎

Die A beats B, B beats C, and C beats A! Clearly, there is no "transitive law" here! This means that no matter what die the first player chooses, the second player can choose a die that beats it with probability $\frac{5}{9}$. The player who picks first is always at a disadvantage!

## 2 Set Theoretic Identities

**Review**: in lecture we showed

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) \text{ for disjoint } A, B.$$

This first fact generalizes to any countable union:

$$\Pr(\bigcup_n A_n) = \sum_n \Pr(A_n) \text{ for pairwise disjoint } A_n.$$

We also showed

$$\begin{aligned}
\Pr(A-B) &= \Pr(A)-\Pr(A\cap B), \text{ so} && (1)\\
\Pr(A\cap B) &\le \Pr(A), && (2)\\
\Pr(\overline{B}) &= 1-\Pr(B), && (3)\\
\Pr(A\cup B) &= \Pr(A)+\Pr(B)-\Pr(A\cap B), \text{ so} && (4)\\
\Pr(A\cup B) &\le \Pr(A)+\Pr(B). && (5)
\end{aligned}$$

The last fact is known as *Boole's Inequality*, where equality holds if and only if $Pr(A \cap B) = 0$. Boole's Inequality can also be extended to any countable number of events:

**Theorem 2.1 (Boole's Inequality)** $\Pr(\bigcup_n A_n) \le \sum_n \Pr(A_n)$.

*Proof.* Note that

$$\sum_{p\in\bigcup_n A_n} \Pr(p) \le \sum_n \sum_{p\in A_n} \Pr(p)$$

because every term in the sum on the lefthand side occurs on the righthand side. But the lefthand side is $\Pr(\bigcup_n A_n)$ by definition, and the righthand side is $\sum_n \Pr(A_n)$. ∎

**Corollary 2.2** $\Pr(\bigcap_n A_n) \ge 1 - \sum_n \Pr(\overline{A_n})$.

*Proof.*

$$\begin{aligned}
\Pr(\textstyle\bigcap_n A_n) &= 1-\Pr(\overline{\textstyle\bigcap_n A_n}) && \text{by equation (3)}\\
&= 1-\Pr(\textstyle\bigcup_n \overline{A_n}) && \text{by DeMorgan's Law}\\
&\ge 1-\textstyle\sum_n \Pr(\overline{A_n}) && \text{by Boole's Inequality.}
\end{aligned}$$

∎

## 2.1 Wiring circuits

As an example, let us suppose you are wiring up a circuit. There are a total of $n$ connections to be made. From past experience we assume that any particular connection is made *incorrectly* with probability $p$, for some $0 \le p \le 1$. That is, for $1 \le i \le n$,

$$\Pr(i^{th} \text{ connection is wrong}) = p.$$

What is the probability that the circuit is wired correctly, i.e., there are no incorrect connections?

Let $A_i$ denote the event that connection $i$ is made *correctly*. Then $\overline{A_i}$ is the event that connection $i$ is made incorrectly, so $\Pr(\overline{A_i}) = p$. Now

$$\Pr(\text{all connections are OK}) = \Pr(\bigcap_{i=1}^{n} A_i) \geq 1 - \sum_{i=1}^{n} \Pr(\overline{A_i}) = 1 - np.$$

What about an upper bound? By equation (2),

$$\Pr(\text{all connections are OK}) = \Pr(\bigcap_{i=1}^{n} A_i) \leq \Pr(A_1) = 1 - p.$$

So for example, if $n = 10$ and $p = 0.01$, we get the following bounds:

$$0.9 = 1 - 10 \cdot 0.01 \leq \Pr(\text{all connections are OK}) =\leq 1 - 0.01 = 0.99.$$

So we have concluded that the chance that all connections are okay is somwhere between 90% and 99%. Although we don't have enough information to calculate the probability of individual sample points, we were able to conclude something useful by using simple facts from set theory.

Could it actually be as high as 99% ? Yes, if the errors occur in such a way that whenever you make one connection wrong, all the connections are wrong. Formally, the sample space could consist of two atomic events, one corresponding to "all OK" with probability 99/100 and one with corresponding to "none OK" with probability 1/100. Then $A_1 = A_2 = \ldots = A_{10} = \bigcap_{i=1}^{10} A_i$, so $\Pr(\text{all OK}) = \Pr(A_i) = 0.99$.

**Class Exercise**: A more natural sample space might be

$$S ::= \{0,1\}^{10}$$

where each point $s ::= (b_1, b_2, \ldots, b_{10})$ describes the bad set of connections in a circuit: $b_i = 0$ if the $i$th connection is bad and $b_i = 1$ if it is OK. Now what is $A_1$? Is it equal to $A_2$? How might we assign probabilities to the points in $S$ so $\Pr(\text{all OK}) = \Pr(A_1) = 0.99$?

**Solution**: $A_i = \{(b_1, \ldots, b_{10}) \in S : b_i = 1\}$, so $A_1 \neq A_2$. $\Pr(1^{10}) = 0.99$, $\Pr(0^{10}) = 0.01$, and probability of all other points is zero.

Could it be 90% ? Yes, suppose the errors are such that we never make two wrong connections. So the events $\overline{A_i}$ are all disjoint and the probability of getting it right is

$$\Pr(\bigcap A_i) = 1 - \Pr(\bigcup \overline{A_i}) = 1 - \sum_{i=1}^{10} \Pr(\overline{A_i}) = 1 - 10 \cdot 0.01 = 0.9.$$

**Class Exercise**: How might we assign probabilities to the points in $S$ above so $\Pr(\text{all OK}) = 0.9$ but $\Pr(A_i) = 0.99$?

**Solution**: $\Pr(s) = 0.01$ if $s$ contains 9 zeroes and 1 one, $\Pr(1^{10}) = 0.9$, and probability of all other points is zero.

## 2.2 The Nontransitive Dice Again

The three dice A, B, C had the property that on any roll of dice A and B, the probability that A beats B is 5/9, on a roll of B and C the probability that B beats C is 5/9, and likewise on a roll of C and A, the probability that C beats A is 5/9. How high can we make this probability? That is, can we construct dice so that the probability of A beating B, B beating C, C beating A are all arbitrarily high? — say 3/4 each?

To answer this, let's roll all three dice. Define $A$ to be the event that A beats B, $B$ the event that B beats C and $C$ the event that C beats A.

**Theorem 2.3**

$$\min\{\Pr(A), \Pr(B), \Pr(C)\} \le \frac{2}{3}.$$

*Proof.*

$$\begin{array}{rcll}
\Pr(A \cap B \cap C) & = & 1 - \Pr(\overline{A} \cup \overline{B} \cup \overline{C}) & \text{by equation 3 and DeMorgan} \\
& \ge & 1 - (\Pr(\overline{A}) + \Pr(\overline{B}) + \Pr(\overline{C}) & \text{by Boole's Inequality} \\
& = & \Pr(A) + \Pr(B) + \Pr(C)) - 2 & \text{by equation 3.}
\end{array}$$

But $\Pr(A \cap B \cap C) = 0$, since it is impossible that on any single roll, die A is higher than die B, *and* B is higher than C, *and* C is higher than A. Hence

$$2 \ge \Pr(A) + \Pr(B) + \Pr(C) \ge 3 \min\{\Pr(A), \Pr(B), \Pr(C)\}$$

and therefore

$$2/3 \ge \min\{\Pr(A), \Pr(B), \Pr(C)\}.$$

■

**Extra Credit Problem**: Show that the 2/3 bound can be improved to $p = (\sqrt{5} - 1)/2 = 0.618\ldots$, and that this is the minimum bound.

# 3 Conditional probability and Bayes rule

Recall the formula for conditional probability:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

As an example, consider rolling two dice. What is the probability of having rolled at least one 3, given that total of the two dice is 8?

Assume each die gives each outcome (i.e., 1,2,3,4,5, or 6) with probability $1/6$, no matter what the outcome of the other die is. Let $A$ be the event that total seen is 8, and let $B$ be the event that at least one 3 has been rolled. The sample space is $S = \{(i,j) : i,j \in \{1,2,3,4,5,6\}\}$, so that $|S| = 36$.

$$A = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}, \quad B \cap A = \{(3,5),(5,3)\},$$

Each outcome has probability $1/36$, so $\Pr(A) = 5/36$ and $\Pr(B \cap A) = 2/36$. Therefore $\Pr(B|A) = \frac{2/36}{5/36} = 2/5$.

Are $\Pr(A|B)$ and $\Pr(B|A)$ related? Yes, both relate to $\Pr(A \cap B)$:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \text{ and } \Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)},$$

so

$$\Pr(A|B)\Pr(B) = \Pr(A \cap B) = \Pr(B|A)\Pr(A).$$

This yields the famous *Bayes Rule:*

$$\Pr(A|B) = \frac{\Pr(B|A)}{\Pr(B)}\Pr(A).$$

*Example:* At a school 40% of students are male, and 60% of males and 80% of all students play sports. What is the probability that a randomly chosen student is a male, given that the student plays sports?

Let $A$ be the event that a student is male. Let $B$ be the event that a student plays sports. We are given $P(A) = 40\%$, $P(B) = 80\%$ and $\Pr(B|A) = 60\%$. We want to compute $\Pr(A|B)$. By Bayes Rule

$$\Pr(A|B) = \frac{\Pr(B|A)}{P(B)}\Pr(A) = \frac{0.6 \times 0.4}{0.8} = 30\%.$$

# 4 Team Problem

**Problem 1.** Verify that the set-theoretic probability equations above "relativize" to conditional probabilities, e.g., equation (1) relativizes to

$$\Pr(A - B \mid C) = \Pr(A \mid C) - \Pr(A \cap B \mid C).$$

**Problem 2.** In the OJ Simpson trial it was claimed by the defense that the fact that Simpson battered his wife should be excluded from the evidence as the fraction of wife batterers that murder their wives is 1/2000 (Dershowitz Lemma). In this problem we will condition this probability on the event that Nicole is dead and see that we obtain a somewhat different result.

We have the following statistical information (all for 1994):

- number of women = 125M
- number of women murdered in 1994 = 6000
- number of women murdered by spouse in 1994 = 1500
- number of women with spouse in 1994 = 50M
- number of women battered by spouse in 1994 = 3M

Being murdered by spouse counts as being battered by spouse. Also, we assume that the event of being murdered by somebody else is not influenced by whether or not the woman was being battered by her spouse.

Define the events

$S$ = Woman is murdered by spouse
$M$ = Woman is murdered
$B$ = Woman is battered by spouse

Use the given information to evaluate the following

**(a)** The Dershowitz Lemma can be reformulated as

$$\Pr(S \mid B) = 1/2000.$$

Prove it.

**(b)** Calculate $\Pr(S \mid M)$, i.e., the probability that a murdered woman was done in by her spouse.

**(c)** Calculate $\Pr(S \mid B \cap M)$, i.e., the probability that a battered, murdered woman is done in by her spouse.

**(d)** What can you conclude about Dershowitz' argument that the battery of a spouse should not be admitted as evidence because of his lemma ?

**Solution:**

$$\Pr(S \mid B) = \frac{\Pr(S \cap B)}{\Pr(B)} = \frac{\Pr(S)}{\Pr(B)} = \frac{1500/125M}{3M/125M} = \frac{1500}{3M} = 1/2000.$$

$$\Pr(S \mid M) = \frac{\Pr(S \cap M)}{\Pr(M)} = \frac{\Pr(S)}{\Pr(M)} = \frac{1500/125M}{6000/125M} = \frac{1500}{6000} = 1/4.$$

Next, note that

$$\Pr(S \mid B \cap M) = \frac{\Pr(S \cap B \cap M)}{\Pr(B \cap M)} = \frac{\Pr(S)}{\Pr(M \mid B)\Pr(B)}.$$

Then to evaluate $\Pr(M \mid B)$, we use the following fact from the previous problem:

$$\begin{aligned}\Pr(M \mid B) &= \Pr(M \cap S \mid B) + \Pr(M \cap \bar{S} \mid B) \\ &= \Pr(S \mid B) + \Pr(M \cap \bar{S}).\end{aligned}$$

In simplifying the second term we used the fact that being murdered by someone other than the spouse is not influenced by whether the victim was battered by her spouse or not, namely,

$$\Pr(M \cap \bar{S} \mid B) = \Pr(M \cap \bar{S}).$$

Now we have

$$\begin{aligned}\Pr(S \mid B \cap M) &= \frac{1500/125M}{\left(\frac{1}{2000} + \frac{6000-1500}{125M}\right)\frac{3M}{125M}} \\ &= \frac{1500 \times 125}{67000 \times 3} \\ &= \frac{125}{134} = 0.93.\end{aligned}$$