

CS 188 Section Handout: Probability Basics

Let $D \in \{d_0, d_1, d_2, d_3\}$ be one of four dice, where d_0 is fair. Let $R \in [1, 6]$ be the outcome of a die roll.

1 Joint probability assembly

A casino employee informs you (the inspector) about a dealer, Angelo: “He cheats a third of the time, using loaded die d_1 for which the six-dotted side shows up five times as often as each of the other sides.”

D	$P(D)$
d_0	$2/3$
d_1	$1/3$

R	D	$P(R D)$
1	d_0	$1/6$
2	d_0	$1/6$
3	d_0	$1/6$
4	d_0	$1/6$
5	d_0	$1/6$
6	d_0	$1/6$
1	d_1	$1/10$
2	d_1	$1/10$
3	d_1	$1/10$
4	d_1	$1/10$
5	d_1	$1/10$
6	d_1	$1/2$

R	D	$P(R, D)$
1	d_0	$1/9$
2	d_0	$1/9$
3	d_0	$1/9$
4	d_0	$1/9$
5	d_0	$1/9$
6	d_0	$1/9$
1	d_1	$1/30$
2	d_1	$1/30$
3	d_1	$1/30$
4	d_1	$1/30$
5	d_1	$1/30$
6	d_1	$1/6$

2 Estimation

Your informant reports back with the following statistics about another dealer, Bert: “He uses two dice, both loaded. Using d_2 , I watched him roll 15 sixes, 10 fives, 5 fours, 5 threes, 5 twos, and no ones. With d_3 , I observed 5 sixes, 10 fives, 15 fours, 15 threes, 15 twos, and 20 ones.” We can infer the following about these dice (a *maximum likelihood estimate*).

R	D	$c(R, D)$
1	d_2	0
2	d_2	5
3	d_2	5
4	d_2	5
5	d_2	10
6	d_2	15
1	d_3	20
2	d_3	15
3	d_3	15
4	d_3	15
5	d_3	10
6	d_3	5

R	D	$P(R, D)$
1	d_2	0
2	d_2	$1/24$
3	d_2	$1/24$
4	d_2	$1/24$
5	d_2	$1/12$
6	d_2	$1/8$
1	d_3	$1/6$
2	d_3	$1/8$
3	d_3	$1/8$
4	d_3	$1/8$
5	d_3	$1/12$
6	d_3	$1/24$

R	D	$P(R D)$
1	d_2	0
2	d_2	$1/8$
3	d_2	$1/8$
4	d_2	$1/8$
5	d_2	$1/4$
6	d_2	$3/8$
1	d_3	$1/4$
2	d_3	$3/16$
3	d_3	$3/16$
4	d_3	$3/16$
5	d_3	$1/8$
6	d_3	$1/16$

3 Inference by enumeration

- What is $P(R = 6 \mid \text{dealer} = A)$? Sum over d_0 and d_1 : $\frac{1}{9} + \frac{1}{6} = \frac{5}{18}$
- What is $P(R = 6 \mid \text{dealer} = B)$? Sum over d_2 and d_3 : $\frac{1}{8} + \frac{1}{24} = \frac{1}{6}$
- What is $P(D = d_0 \mid R = 6, \text{dealer} = A)$? $\frac{P(D=d_0, R=6 \mid \text{dealer}=A)}{P(R=6 \mid \text{dealer}=A)} = \frac{1/9}{5/18} = \frac{2}{5}$
- What is $P(D = d_0 \mid \text{dealer} = B)$? 0
- What is $P(D = d_2 \mid \text{dealer} = B)$? $\frac{40}{120} = \frac{1}{3}$.

Suppose you know that A works some five nights per week and B works on some other two nights.

- What is $P(D)$?

As an example, consider $P(D = d_0)$.

$$\begin{aligned}
 P(D = d_0) &= P(D = d_0, \text{dealer} = A) + P(D = d_0, \text{dealer} = B) \\
 &= P(D = d_0 \mid \text{dealer} = A)P(\text{dealer} = A) + P(D = d_0 \mid \text{dealer} = B)P(\text{dealer} = B) \\
 &= \frac{2}{3} \cdot \frac{5}{7} + 0 \cdot \frac{2}{7} \\
 &= \frac{10}{21}
 \end{aligned}$$

D	$P(D)$
d_0	$10/21$
d_1	$5/21$
d_2	$2/21$
d_3	$4/21$

- One night, you observe a dealer (A or B) roll a six. What is the probability that the die is loaded?

$$\begin{aligned}
 P(D \in \{d_1, d_2, d_3\} \mid R = 6) &= 1 - P(D = d_0 \mid R = 6) \\
 &= 1 - [P(D = d_0 \mid R = 6, A)P(A) + P(D = d_0 \mid R = 6, B)P(B)] \\
 &= 1 - [\frac{2}{5} \cdot \frac{2}{7} + 0 \cdot \frac{2}{7}] = \frac{5}{7}
 \end{aligned}$$

4 Sequence of independent events

You confront the suspected dealer and ask him to roll his current die three times in a row. He rolls the sequence $S = (6, 2, 6)$

Calculate the following likelihood probabilities:

- $P(S \mid D = d_0) = (\frac{1}{6})^3 = \frac{1}{216}$
- $P(S \mid D = d_1) = (\frac{1}{2}) \cdot (\frac{1}{2}) \cdot (\frac{1}{10}) = \frac{1}{40}$
- $P(S \mid D = d_2) = (\frac{1}{8}) \cdot (\frac{3}{8}) \cdot (\frac{3}{8}) = \frac{9}{512}$
- $P(S \mid D = d_3) = (\frac{1}{16})^2 (\frac{3}{16}) = \frac{3}{4096}$

5 Bayes' Rule

Given the evidence you acquired previously, determine which die was rolled:

$$\hat{d} = \arg \max_d P(D = d | S)$$

Is the die loaded? Which dealer are you arresting?

By Bayes rule, we have

$$P(D = d | S = s) = \frac{P(S = s | D = d)P(D = d)}{P(S = s)} \quad (1)$$

Note that for any d , the denominator of the right hand side will be the same. So, when we look for the maximal d , we need only compute the numerator for each die.

- $P(S | D = d_0)P(D = d_0) = \frac{1}{216} \cdot \frac{10}{21} = \frac{5}{2268}$
- $P(S | D = d_1)P(D = d_1) = \frac{1}{40} \cdot \frac{5}{21} = \frac{1}{168}$
- $P(S | D = d_2)P(D = d_2) = \frac{9}{512} \cdot \frac{2}{21} = \frac{3}{1792}$
- $P(S | D = d_3)P(D = d_3) = \frac{3}{4096} \cdot \frac{4}{21} = \frac{1}{7168}$

The greatest of these values arises from $D = d_1$, so the most likely scenario is that you are confronting A using the loaded d_1 . Time to call in some backup!

Note: The denominator can be expressed as the sum of numerators:

$$P(S) = \sum_d P(S|D = d)P(D = d) = \frac{5}{2268} + \frac{1}{168} + \frac{3}{1792} + \frac{1}{7168}$$

Then use Bayes' Rule to determine that there is roughly a 60% chance that you are correctly accusing dealer A . On the other hand, there is a 22% chance that dealer A was rolling the fair die, and an 18% chance that it's really dealer B . Maybe you should gather more evidence...