Problems for Recitation 19

1 Bayes' Rule

Bayes' Rule says that if A and B are events with nonzero probabilities, then:

$$\Pr\{A \mid B\} \cdot \Pr\{B\} = \Pr\{B \mid A\} \cdot \Pr\{A\}$$

a. Prove Bayes' Rule.

$P(A \cap B) =$

b. A weatherman walks to work each day. Some days it rains:

$$\Pr\{\text{rains}\} = 0.30$$

Sometimes the weatherman brings his umbrella. Usually this is because he predicts rain, but he also sometimes carries it to ward off bright sunshine.

$$\Pr\{\text{carries umbrella}\} = 0.40$$

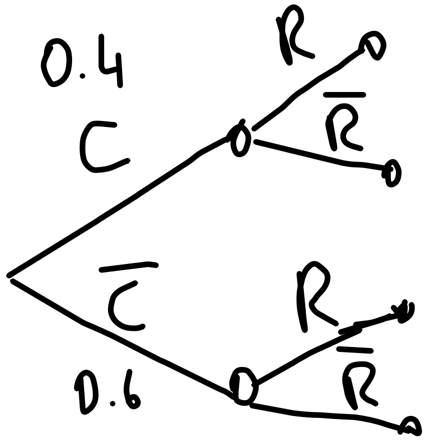
As a weatherman, he usually doesn't get caught out in a storm without protection:

$$\Pr\{\text{carries umbrella} \mid \text{rains}\} = 0.80$$

Suppose you see the weatherman walking to work, carrying an umbrella. What is the probability of rain? Use Bayes' Rule.

$$Pr(R|C) = \frac{Pr(C|R)\cdot Pr(R)}{Pr(C)}$$

$$\boxed{Pr(R\cap C) = Pr(C|R)\,Pr(R)} = 0.3\cdot 0.8 = 0.24$$

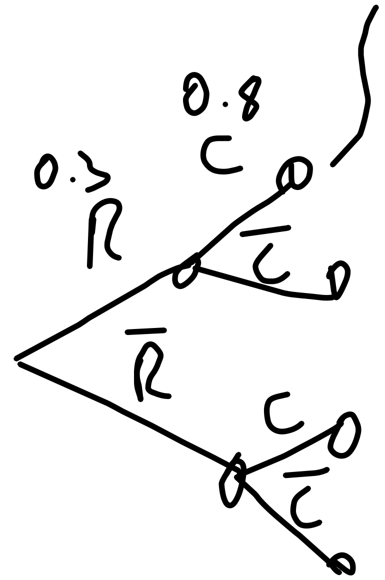


$$\boxed{Pr(C\cap R) = Pr(R|C)\,Pr(C)}$$

? !

$$= Pr(C|R)\cdot Pr(R)$$

$$\rightarrow Pr(R|C) = \frac{Pr(C|R)\cdot Pr(R)}{Pr(C)}$$



2 DNA Profiles

Suppose that we create a a national database of DNA profiles. Let's make some (overly) simplistic assumptions:

- Each person can be classified into one of 20 billion different "DNA types". (For example, you might be type #13,646,572,661 and the person next to you might be type #2,785,466,098.) Let $T(x)$ denote the type of person x.
- Each DNA type is equally probable.
- The DNA types of Americans are mutually independent.

a. A congressman argues that there are only about 300 million Americans, so even if a profile for every American were stored in the database, the probability of even one coincidental match would be very small.

Recall from lecture that if there are N days in a year and m people in a room, then the probability that no two people in the room have the same birthday is about $e^{-m^2/(2N)}$. Using this fact, what is the probability that two people's DNA profiles would match if every person's profile were stored in the database?

Pr(

b. After this database is implemented, some DNA is found at a crime scene. The DNA is sequenced and a person with matching DNA is found through the database and accused of the crime. At the trial the defense attorney argues that, by the birthday principle, the probability that there are multiple people whose DNA is identical is a virtual certainty, and so the jury cannot conclude beyond a reasonable doubt that the defendant is the criminal.

What is the flaw in this argument? Under what circumstances could you conclude based on DNA evidence alone that there is no doubt that the defendant committed the crime? (assuming no errors in the DNA tests, a comprehensive database, etc. etc.)

3 The Immortals

There were n Immortal Warriors born into our world, but in the end *there can be only one.* The Immortals' original plan was to stalk the world for centuries, dueling one another with ancient swords in dramatic landscapes until only one survivor remained. However, after a thought-provoking discussion of probabilistic independence, they opt to give the following protocol a try:

1. The Immortals forge a coin that comes up heads with probability p.
2. Each Immortal flips the coin once.
3. If *exactly one* Immortal flips heads, then he or she is declared The One. Otherwise, the protocol is declared a failure, and they all go back to hacking each other up with swords.

a. One of the Immortals (the Kurgan from the Russian steppe) argues that as n grows large, the probability that this protocol succeeds must tend to zero. Another (McLeod from the Scottish highlands) argues that this need not be the case, provided p is chosen *very carefully.* What does your intuition tell you?

$$p \overset{?}{=} \frac{1}{N}$$

b. What is the probability that the experiment succeeds as a function of p and n?

c. How should p, the bias of the coin, be chosen in order to maximize the probability that the experiment succeeds? (You're going to have to compute a derivative!)

Example: $\Omega=\{1,2,3,4\}$, $(\mathcal{A}=2^{\Omega})$, $P(k)=1/4 \ \forall k\in\Omega$.

$A=\{1,2\}$, $B=\{1,3\}$, $C=\{2,3\}$.

(1) Are A, B indep?

(2) Are A, B, C mut. indep? pairwise indep?

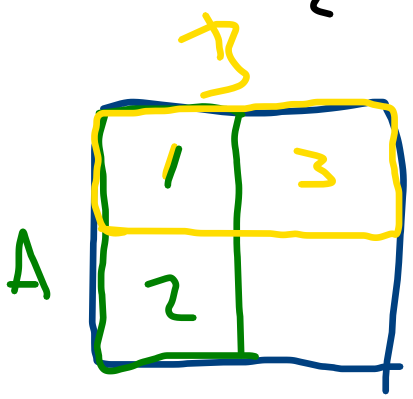
(3) Are A, B cond. indep. given C?

(1) $P(A\cap B) \overset{?}{=} P(A)P(B)$

$P(1) \overset{?}{=} P(1\cup 2)\cdot P(1\cup 3)$

$\frac{1}{4} \overset{?}{=} \frac{1}{2}\cdot\frac{1}{2}$

$\frac{1}{4} \overset{\checkmark}{=} \frac{1}{4}$



(2) $P(A\cap B\cap C) \overset{?}{=} P(A)P(B)P(C)$

$\emptyset \neq \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}$

$\left.\begin{matrix} A\cap B \\ A\cap C \\ B\cap C \end{matrix}\right\}$ disjoint

(3) $P(A\cap B\,|\,C) \overset{?}{=} P(A|C)\cdot P(B|C)$

$\frac{P(A\cap B\cap C)}{P(C)} \overset{?}{=} \frac{P(A\cap C)}{P(C)}\cdot\frac{P(B\cap C)}{P(C)}$

$\emptyset \overset{?}{=} \frac{\frac{1}{4}}{\frac{1}{2}}\cdot\frac{\frac{1}{4}}{\frac{1}{2}}$

$\emptyset \overset{?}{=} \frac{1}{2}\cdot\frac{1}{2}$

$\emptyset \neq \frac{1}{4}$