Probabilistic Foundations of Artificial Intelligence

Solutions to Problem Set 1

Oct 13, 2017

1. Conditional Probabilities

For each statement below, either prove it is true, or give a counterexample showing it is false. In the following, we assume that all events have non-zero probability.

(a) If $P(a|b,c) = P(b|a,c)$, then $P(a|c) = P(b|c)$

(b) If $P(a|b,c) = P(a)$, then $P(b|c) = P(b)$

(c) If $P(a|b) = P(a)$, then $P(a|b,c) = P(a|c)$

Solution

(a). True.

From Bayes' rule, we get

$$P(a,b,c) = P(a|b,c)P(b|c)P(c) \tag{1}$$

and

$$P(a,b,c) = P(b|a,c)P(a|c)P(c) \tag{2}$$

From the question we have $P(a|b,c) = P(b|a,c)$, and therefore we can rewrite (1) as $P(a,b,c) = P(b|a,c)P(b|c)P(c)$. Combining with (2) we get $P(a|c) = P(b|c)$.

(b). False.

The statement is equivalent to: $a \perp (b,c) \Rightarrow b \perp c$, which is false. See Figure 1 for a counterexample (description below).

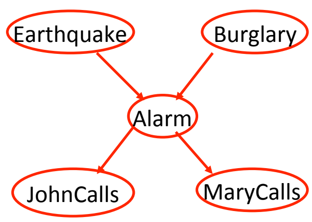

Figure 1: Example from lecture slides: casual parametrization

Counterexample. If $a = JohnCalls$, $b = Burglary$, $c = Earthquake$, then

$$a \perp (b, c) \mid Alarm.$$

However, the event $Burglary$ is dependent with $Earthquake$ if $Alarm$ is observed:

$$b \not\perp c \mid Alarm$$

Therefore, we have identified an example where the statement is false.

(c). False.

Counterexample. Suppose $a \perp b$, each taking 0 and 1 with probability 0.5, and $c = ab$. $P(a = 0|b = 0) = 1/2$. But, when $c = 0$, $P(a = 0) = 2/3$ and $P(a = 0) > P(a = 0|b = 0)$. Therefore the statement $a \perp b \mid c$ is false.

2. Finding the fake coin

Suppose you are given a bag containing n unbiased coins. You are also told that $n - 1$ of these coins are normal, that is, they have a head on one side and a tail in the other. The remaining one is fake and has heads on both sides.

(a) Suppose you pick a coin from the bag uniformly at random, you flip it, and get a head. Given this result, what is the probability that the coin you picked is the fake one? (Note that we ask for a conditional probability.)

(b) Suppose you continue flipping the same coin for a total of k times and you get k heads. Given this result, what is the probability that you picked the fake coin?

(c) Now, suppose you devise the following method to determine if the coin is fake or not. You flip it k times, after which you conclude that it is the fake one if all k flips have resulted in heads, else you conclude that it is normal. What is the probability that using this method you arrive at a wrong conclusion? (Note that this time we ask for an unconditional probability.)

Solution

(a) Define a random variable C corresponding to the coin that takes values N (normal) or F (fake) and a random variable E corresponding to the outcome of the flip that takes values H (heads) or T (tails). From the problem description, we can write down the following probabilities:

$$\begin{aligned} P(C = N) &= \frac{n-1}{n} \\ P(C = F) &= \frac{1}{n} \\ P(E = H \mid C = N) &= 0.5 \\ P(E = H \mid C = F) &= 1 \end{aligned}$$

To compute the probability $P(C = F \mid E = H)$, we can use Bayes' rule:

$$\begin{aligned}
P(C = F \mid E = H) &= \frac{P(E = H \mid C = F)P(C = F)}{P(E = H)} \\
&= \frac{P(E = H \mid C = F)P(C = F)}{P(E = H \mid C = N)P(C = N) + P(E = H \mid C = F)P(C = F)} \\
&= \frac{1 * \frac{1}{n}}{0.5 * \frac{n-1}{n} + 1 * \frac{1}{n}} \\
&= \frac{2}{n+1}
\end{aligned}$$

(b) Since the outcomes of coin flips are conditionally independent given the coin, if we denote by E_k the random variable that corresponds to the outcome of k flips with value H^k when k heads occur, then we have that

$$\begin{aligned}
P(E_k = H^k \mid C = N) &= 2^{-k} \\
P(E_k = H^k \mid C = F) &= 1.
\end{aligned}$$

Using Bayes' rule similarly to the previous question we have

$$\begin{aligned}
P(C = F \mid E_k = H^k) &= \frac{P(E_k = H^k \mid C = F)P(C = F)}{P(E_k = H^k)} \\
&= \frac{P(E_k = H^k \mid C = F)P(C = F)}{P(E_k = H^k \mid C = N)P(C = N) + P(E_k = H^k \mid C = F)P(C = F)} \\
&= \frac{1 * \frac{1}{n}}{2^{-k} * \frac{n-1}{n} + 1 * \frac{1}{n}} \\
&= \frac{1}{1 + (n-1)2^{-k}}
\end{aligned}$$

Figure 2 shows how the probability of having picked the fake coin among $n = 50$ total coins increases as the number of observed heads k gets larger.

(c) A wrong conclusion is reached when we pick a normal coin and claim it is fake (i.e., we observe k heads) or we pick the fake coin and claim it is normal (i.e., observe tails at least once). Note that the latter event is not possible (happens with probability 0) by definition of the fake coin. Therefore, the probability of a wrong conclusion is

$$\begin{aligned}
P(C = N, E_k = H^k) &= P(E_k = H^k \mid C = N)P(C = N) \\
&= \frac{n-1}{n} 2^{-k},
\end{aligned}$$

which is depicted in Figure 3 for $n = 50$.

Suppose you throw a dice repeatedly until you get a 6.

(a) What is the probability of finding a sequence of length n?

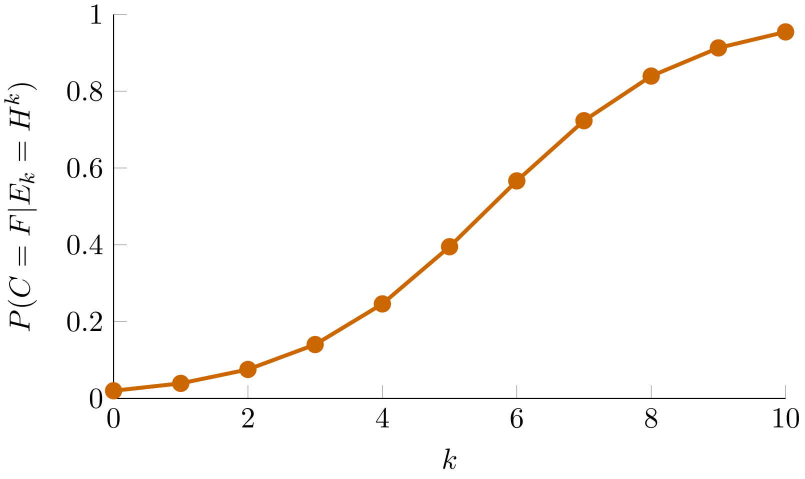

Figure 2: Probability of having picked the fake coin given k observed heads ($n = 50$).

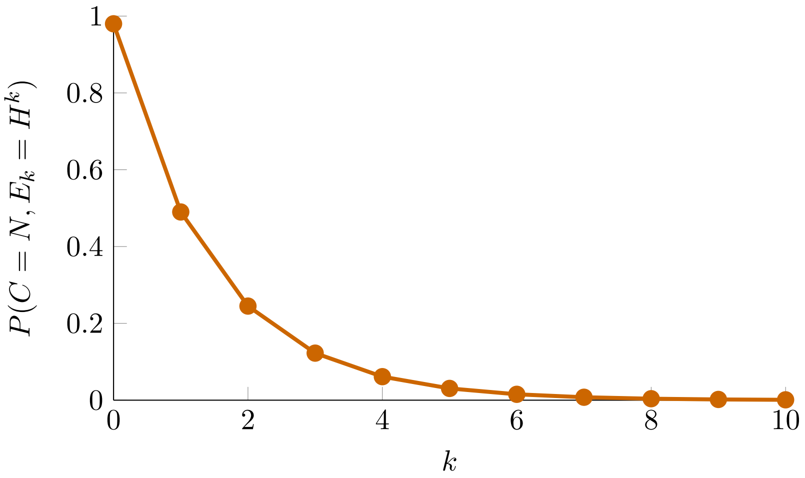

Figure 3: Probability of reaching a wrong conclusion using k flips ($n = 50$).

(b) What is the expected value of the sequence length?

(c) What is the expected number of 3s we observe?

Solution

(a) A sequence of length n is obtained only if the first $n-1$ throws are not a 6 and the n-th throw is exactly a 6.

$$P(L_n) = \left(\frac{5}{6}\right)^{(n-1)} \frac{1}{6}$$

(b) The expected length is:

$$
\begin{aligned}
E[L_n] &= \sum_{i=1}^{\infty} i \left(\frac{5}{6}\right)^{(i-1)} \frac{1}{6} \\
&= \frac{1}{6} \sum_{i=1}^{\infty} i \left(\frac{5}{6}\right)^{(i-1)} \\
&= \frac{1}{6} \frac{1}{(1-(5/6))^2} = 6
\end{aligned}
$$

(c) By symmetry, the expected number of 3s or any other number (except 6) should be the same. We denote with X_i the total number of i in the sequence. Also denote X as the total numbers from 1 to 5 in the sequence. Therefore, $X = \sum_{i=1}^{5} X_i$.

For any sequence of length n, the random variable X takes the value $X = n - 1$. Hence,

$$
\begin{aligned}
E[X] &= E[L_n] - 1 \\
&= \sum_{i=1}^{5} E[X_i] \\
&= 5E[X_3] \\
E[X_3] &= \frac{1}{5} \left(E[L_n] - 1\right) = 1
\end{aligned}
$$