ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 2 Information Theory and Coding

Solutions to homework 1 Oct 08, 2007

PROBLEM 1. Note that $E_0 = E_1 \cup E_2 \cup E_3$.

(a) (1) For disjoint events, $P(E_0) = P(E_1) + P(E_2) + P(E_3)$, so $P(E_0) = 3/4$.

(2) For independent events, $1 - P(E_0)$ is the probability that none of the events occur, which is the product of the probabilities that each one doesn't occur. Thus $1 - P(E_0) = (3/4)^3$ and $P(E_0) = 37/64$.

(3) If $E_1 = E_2 = E_3$, then $E_0 = E_1$ and $P(E_0) = 1/4$.

(b) (1) From the Venn diagram in Fig. 1, $P(E_0)$ is clearly maximized when the events are disjoint, so $\max P(E_0) = 3/4$.

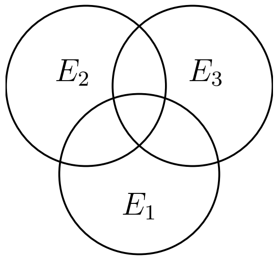


Figure 1: Venn Diagram for problem 1 (b)(1)

(2) The intersection of each pair of sets has probability 1/16. As seen in Fig. 2, $P(E_0)$ is maximized if all these pairwise intersections are identical, in which case $P(E_0) = 3(1/4 - 1/16) + 1/16 = 5/8$. One can also use the formula $P(E_0) = P(E_1)+P(E_2)+P(E_3)-P(E_1\cap E_2)-P(E_1\cap E_3)-P(E_2\cap E_3)+P(E_1\cap E_2\cap E_3)$, and notice that all the terms except the last is fixed by the problem, and the last term cannot be made more than $\min_{i,j} P(E_i \cap E_j) = 1/16$.

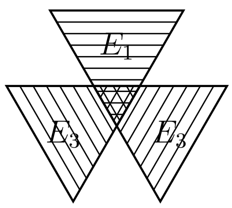


Figure 2: Venn Diagram for problem 1 (b)(2)

PROBLEM 2. Let L be the event that the loaded die is picked and H the event that the honest die is picked. Let A_i be the event that i is turned up on the first roll, and B_i be the event that i is turned up on the second roll. We are given that $P(L) = 1/3, P(H) = 2/3$; $P(A_1 \mid L) = 2/3$; $P(A_i \mid L) = 1/15 \quad 2 \leq i \leq 6$; $P(A_i \mid H) = 1/6 \quad 1 \leq i \leq 6$. Then

$$P(L \mid A_1) = \frac{P(L, A_1)}{P(A_1)} = \frac{P(A_1 \mid L)\,P(L)}{P(A_1 \mid L)\,P(L) + P(A_1 \mid H)\,P(H)} = \frac{2}{3}.$$

This is the probability that the loaded die was picked conditional on the first roll showing a 1. For two rolls we make the assumption from the physical mechanism involved in rolling a die that the outcome on the two successive rolls of a given die are independent. Thus $P(A_1B_1 \mid L) = (2/3)^2$ and $P(A_1B_1 \mid H) = (1/6)^2$. It follows as before that

$$P(L \mid A_1B_1) = \frac{8}{9}.$$

PROBLEM 3.

(a) Since every door is picked with the same probability, the probability that you pick the door with the car behind it is $\frac{1}{3}$.

(b) Let the door with the car be numbered 1 and the door with the goats be numbered $2, 3$. Let E be the number of the door picked the first time. E takes values in $\{1, 2, 3\}$ with uniform probability. Let W represent the event that you get the car if you switch to the other unopened door.

$$\begin{aligned}
\Pr(\mathrm{W}) &= \sum_{i=1}^{3} \Pr(\mathrm{W}, E = i) \\
&= \sum_{i=1}^{3} \Pr(\mathrm{W}|E = i)\Pr(E = i)
\end{aligned}$$

If you picked door 1 which has the car, you would lose the car if you switch i.e., $\Pr(\mathrm{W}|E = 1) = 0$. If you picked either door 2 or 3, the unopened door is door 1, so you win if you switch to this door, i.e., $\Pr(\mathrm{W}|E = i) = 1$ for $i = 2, 3$.
Substituting the conditional probabilities, we get $\Pr(\mathrm{W}) = \frac{2}{3}$. You increase your chances of getting the car from $\frac{1}{3}$ to $\frac{2}{3}$ by switching to the other unopened door.

PROBLEM 4.

(a)

$$\begin{aligned}
E[X+Y] &= \sum_{x,y} (x+y) P_{XY}(x,y) \\
&= \sum_{x,y} x P_{XY}(x,y) + \sum_{x,y} y P_{XY}(x,y) \\
&= \sum_{x} x P_X(x) + \sum_{y} y P_Y(y) \\
&= E[X] + E[Y].
\end{aligned}$$

Note that independence is not necessary here and that the argument extends to non-discrete variables if the expectation exists.

(b)

$$\begin{aligned}
E[XY] &= \sum_{x,y} xy P_{XY}(x,y) \\
&= \sum_{x,y} xy P_X(x) P_Y(y) \\
&= \sum_{x} x P_X(x) \sum_{y} y P_Y(y) \\
&= E[X]\, E[Y].
\end{aligned}$$

Note that the statistical independence was used on the second line. Let X and Y take on only the values ± 1 and 0. An example of uncorrelated but dependent variables is

$$P_{XY}(1,0) = P_{XY}(0,1) = P_{XY}(-1,0) = P_{XY}(0,-1) = \frac{1}{4}.$$

An example of correlated and dependent variables is

$$P_{XY}(1,1) = P_{XY}(-1,-1) = \frac{1}{2}.$$

(c) Using (a), we have

$$\begin{aligned}\sigma^2_{X+Y} &= E\big[(X - E[X] + Y - E[Y])^2\big] \\ &= E[(X - E[X])^2] + 2E[(X - E[X])(Y - E[Y])] + E[(Y - E[Y])^2].\end{aligned}$$

The middle term, from (a), is $2(E[XY] - E[X]E[Y])$. For uncorrelated variables that is zero, leaving us with $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$.

PROBLEM 5.

(a) Note that the event $N = n$ is the same as the coin falling tails $n-1$ times followed by it falling heads. Since the coin flips are independent and they are fair, we get $\Pr(N = n) = 2^{-(n-1)}2^{-1} = 2^{-n}$. Using Bayes' rule:

$$\Pr(N = n | N \in \{n, n+1\}) = \frac{\Pr(N = n)}{\Pr(N \in \{n, n+1\})} = \frac{2^{-n}}{2^{-n} + 2^{-(n+1)}} = 2/3$$

(b) The only way we find 1 franc in the chosen box is when $N = 1$ and we have chosen the box with the smaller amount of money. The other box thus contains 3 francs.

(c) If we find 3^n francs in the chosen box, we know that N is either n (and the other box contains 3^{n-1} francs) or $n+1$ (and the other box contains 3^{n+1} francs). Using part (a), $N = n$ with probabity $2/3$, and $N = n+1$ with probability $1/3$. Thus the expected money in the other box is

$$\tfrac{2}{3}3^{n-1} + \tfrac{1}{3}3^{n+1} = \tfrac{11}{9}3^n \quad \text{francs.}$$

(d) Indeed, no matter what we find in the chosen box, the expected amount in the other box is more then the amount found in the chosen box (3 vs 1 as in part (b) or 11/9 times as in part (c)). Wt thus have, with X and Y representing the amount in the two boxes,

$$E[X|Y] > Y \qquad \text{and} \qquad E[Y|X] > X.$$

This appears to be a paradox if we take expectations again to obtain

$$E[X] > E[Y] \qquad \text{and} \qquad E[Y] > E[X].$$

However, some thought reveals that $E[X]$ and $E[Y]$ do not exist, and so the last equation is without content: Since $\Pr(N = n) = 2^{-n}$, the expected amount of money in the box with the smaller amount is $\sum_{n\geq 1} 2^{-n}3^{n-1}$ which is a divergent series.

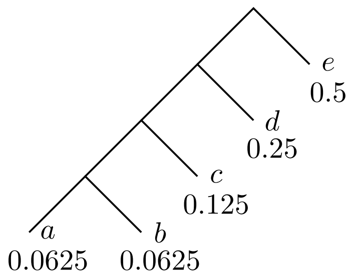
Problem 6.

(a) Code I is prefix-free, Code II is not.

(b) Both codes are uniquely decodable: Code I because it is instantaneous, Code II because the 1's at the beginning of each code word act as markers that separates the codewords and the decoding can be performed by counting the 0's between the 1's.

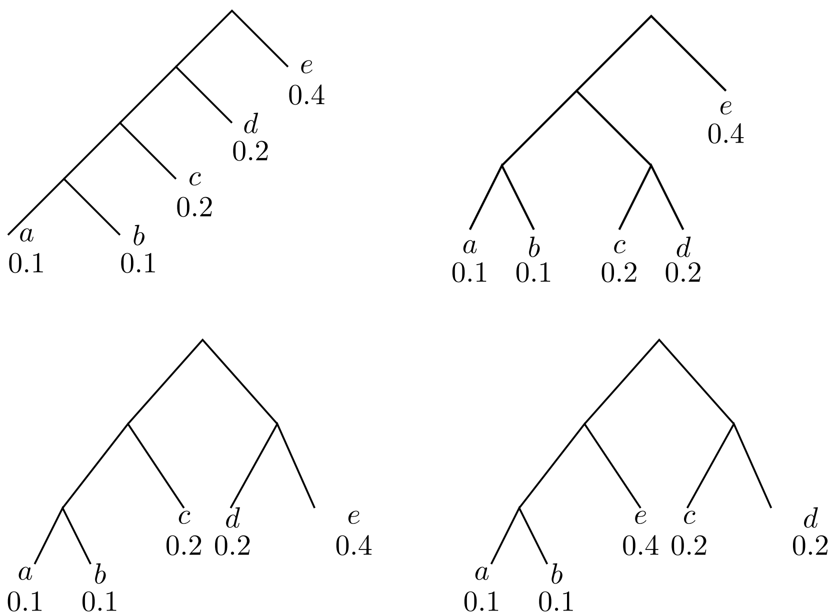
Problem 7.
The figures explicitly specify the optimal mapping from the source letter to the source codeword. $H, \bar{L}$ respectively denote entropy of the source and the expected length for the mapping. The number in brackets indicate probabilities.

(a) The optimal tree is given by



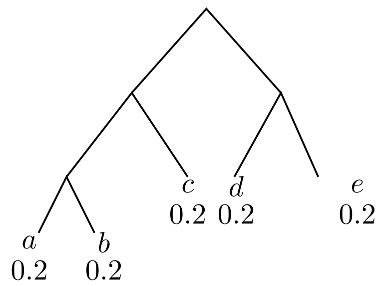
with $H = 1.875$ and $\bar{L} = 1.875$.

(b) We have four possible optimal tree structures corresponding to this distribution:



For any of these trees $\bar{L} = 2.20$, whereas $H \approx 2.122$.

(c) The optimal tree for this distribution is



with $H \approx 2.322$ and $\bar{L} = 2.40$.