Applied math Exercise 2

Erik Paskalev

September 10, 2024

1 Suppose we roll an unbiased die. Let X denote the number of eyes we have rolled. Furthermore Y is a random variable with value 0 or 1, depending on whether the number of eyes is even or odd. Compute $H(X), H(Y)$ and $H(X|Y)$.

Answer: The entropy of X is

$$H(X) = -\sum_{i=1}^{6} \frac{1}{6} log_2(\frac{1}{6}) = log_2(6) \approx 2.584 \text{ bits} \tag{1}$$

Since the number of even and odd numbers on a die is equal probability distribution of Y is $(\frac{1}{2}, \frac{1}{2})$, meaning H(Y) is:

$$H(Y) = -\sum_{i=1}^{2} \frac{1}{2} log_2(\frac{1}{2}) = log_2(2) = 1 \text{ bit} \tag{2}$$

The entropy of the conditional probability distribution $H(X|Y)$ is:

$$H(X|Y) = -\sum_{i=1}^{2}\sum_{j=1}^{6} P(X = x_j, Y = y_i) log_2(P(X = x_j, Y = y_i))P(Y = y_i) =$$

$$= -\sum_{i=1}^{3} P(X = x_{2*j}|Y = 0) log_2(P(X = x_{2*j}|Y = 0))P(Y = 0)-$$

$$- \sum_{i=1}^{3} P(X = x_{2*j-1}|Y = 1) log_2(P(X = x_{2*j-1}|Y = 1))P(Y = 1) =$$

$$= -6 * \frac{1}{6} log_2(\frac{1}{6})\frac{1}{2} = \frac{log_2(6)}{2} \approx 1.292 \text{ bits}$$

$$(3)$$

2   Cory and Dudley both toss three fair coins. We call the number of times 'heads' pops up respectively C and D. The total amount of times 'heads' is tossed is denoted by T, so T = C + D.

(a) Compute the entropy of C.

Answer: C is the sum of three independent coin tosses meaning it is a Binomial distribution:

$$C = \{\binom{3}{0}\frac{1}{8}, \binom{3}{1}\frac{1}{8}, \binom{3}{2}\frac{1}{8}, \binom{3}{3}\frac{1}{8}\} = \{\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}\} \qquad (4)$$

So the entropy of C is:

$$H(C) = -2\frac{1}{8}log_2(\frac{1}{8}) - 2\frac{3}{8}log_2(\frac{3}{8}) =$$

$$= \frac{1}{4}log_2(8) + \frac{3}{4}log_2(\frac{8}{3}) \approx 1.811 \text{ bits}$$

$$(5)$$

(b) Determine the probability distribution of T.

Answer: T is a sum of two independent probability distributions C and D which themselves are sums of three independent coin tosses. That means T the sum of six independent coins flips making it a Binomial distribution:

$$C = \{\binom{6}{0}\frac{1}{64}, \binom{6}{1}\frac{1}{64}, \binom{6}{2}\frac{1}{64}, \binom{6}{3}\frac{1}{64}, \binom{6}{4}\frac{1}{64}, \binom{6}{5}\frac{1}{64}, \binom{6}{6}\frac{1}{64}\} =$$
$$= \{\frac{1}{64}, \frac{6}{64}, \frac{15}{64}, \frac{20}{64}, \frac{15}{64}, \frac{6}{64}, \frac{1}{64}\}$$

(6)

(c) Calculate the entropy of T.

Answer: The entropy of T is:

$$H(T) = -2\frac{1}{64}log_2(\frac{1}{64}) - 2\frac{6}{64}log_2(\frac{6}{64}) - 2\frac{15}{64}log_2(\frac{15}{64}) - \frac{20}{64}log_2(\frac{20}{64}) =$$
$$= \frac{1}{32}log_2(32) + \frac{3}{16}log_2(\frac{16}{3}) + \frac{15}{32}log_2(\frac{64}{15}) + \frac{5}{16}log_2(\frac{16}{5}) \approx 2.333 \text{ bits}$$

(7)

(d) True or false: if you know C, the uncertainty of T halves. Hint: determine $H(T|C)$.

Answer: The conditional probability distribution $P(T|C)$ is:

T | C	0	1	2	3
0	$\frac{1}{64}$	0	0	0
1	$\frac{3}{64}$	$\frac{3}{64}$	0	0
2	$\frac{3}{64}$	$\frac{9}{64}$	$\frac{3}{64}$	0
3	$\frac{1}{64}$	$\frac{9}{64}$	$\frac{9}{64}$	$\frac{1}{64}$
4	0	$\frac{3}{64}$	$\frac{9}{64}$	$\frac{3}{64}$
5	0	0	$\frac{3}{64}$	$\frac{3}{64}$
6	0	0	0	$\frac{1}{64}$

Table 1: Visualization of P(T | C).

Thus the entropy of $H(T|C)$ is:

$$H(T|C) = -\sum_{i=1}^{3}\sum_{j=1}^{6}P(T = t_j|C = c_i)log_2(P(T = t_j|C = c_i))P(C = c_i) =$$

$$= -4 * \frac{1}{64}log_2(\frac{1}{64}) - 8 * \frac{3}{64}log_2(\frac{3}{64}) - 4 * \frac{9}{64}log_2(\frac{9}{64}) \approx$$

$$\approx 3.622 \text{ bits}$$

(8)

3   At a best-of-five tennis match the winner is whoever wins three sets first. Suppose that players A and B are as likely to win a set, so that a set is won by either A or B with probability $\frac{1}{2}$. Let X be the random variable that describes the course of each possible match, such as AAA, ABBAA or ABBB (but not ABAB, since that match hasn't finished yet). The number of sets needed to decide the winner of the match is given by the random variable Y. Hence Y only attains the values 3, 4 and 5.

(a) Determine all possible outcomes of X and give the corresponding probabilities.

Answer: Since the probability of victory of both sides is $\frac{1}{2}$ then each game is equivalent to a coin flip: All games of length 3: AAA - $\frac{1}{8}$, BBB - $\frac{1}{8}$. All games of length 4: BAABA - $\frac{1}{16}$, ABAA - $\frac{1}{16}$, AABA - $\frac{1}{16}$, ABBB - $\frac{1}{16}$, BABB - $\frac{1}{16}$, BBAB - $\frac{1}{16}$. All games of length 5 can be created combinatorially by taking each game of length 4 and adding an A or B (depending on who won the set) at one of 4 positions to create a new match. However, doing this double counts each game so in total the number of unique games are $6*4/2 = 12$ games of length 5 each with probability $\frac{1}{32}$.

(b) Determine the possible outcomes of Y and give the corresponding probabilities.

Answer: The value of Y depends on the length of each game so:

$$P(Y = 3) = 2 * \frac{1}{8} = \frac{2}{8}$$

$$P(Y = 4) = 6 * \frac{1}{16} = \frac{3}{8} \tag{9}$$

$$P(Y = 5) = 12 * \frac{1}{32} = \frac{3}{8}$$

(c) Compute H(X)

$$H(X) = -2 * \frac{1}{8} * log_2(\frac{1}{8}) - 6 * \frac{1}{16}log_2(\frac{1}{16} - 12 * \frac{1}{32}log_2(\frac{1}{32}) = \tag{10}$$
$$= 4.125 \text{ bits}$$

(d) Compute H(Y)

$$H(Y) = -\frac{2}{8}log_2(\frac{2}{8}) - 2 * \frac{3}{8}log_2(\frac{3}{8}) \approx 1.561 \text{ bits} \tag{11}$$

(e) Calculate $H(X|Y)$ in 3 decimals.

$$H(X|Y) = -\sum_{i=3}^{5} H(X|Y = y_i)P(Y = y_i) = \tag{12}$$
$$= -2 * \frac{1}{2}log_2(\frac{1}{2})\frac{2}{8} - 6 * \frac{1}{6}log_2(\frac{1}{6})\frac{3}{8} - 12 * \frac{1}{12}log_2(\frac{1}{12})\frac{3}{8} =$$
$$= \frac{log_2(2)}{4} + \frac{3log_2(6)}{8} + \frac{3log_2(12)}{8} \approx 2.563 \text{ bits}$$

(f) Calculate $H(Y|X)$.

Answer: Since Y is defined as a function of X then $H(f(X)|X) = 0$ bits.

4 Research has proven that $\frac{7}{10}$ of the men has dark hair and that $\frac{1}{4}$ of the women is blond. Furthermore it is known that $\frac{4}{5}$ of the blonde women marries a dark-haired man. Suppose that the number of men and women is equal, and that every man marries precisely one woman. Furthermore assume that every person has hair and that dark and blond are the only possible colours. Denote the hair colour of the man by X, and of the woman by Y.

(a) Determine the probability distributions of both X and Y (i.e. give the probability of each possible outcome).

Answer: X and Y are Bernouli distributed.

$$P(X) = (\frac{7}{10}, \frac{3}{10})$$
$$P(Y) = (\frac{3}{4}, \frac{1}{4}) \tag{13}$$

(b) Determine the conditional probability distributions $P(X|Y = blond)$ and $P(X|Y = dark)$.

Answer: We know $P(X = dark|Y = blond) = \frac{4}{5}$ and $P(X = blond|Y = blond) = \frac{1}{5}$ meaning $P(X|Y = blond) = (\frac{4}{5}, \frac{1}{5})$. To find $P(X|Y = dark)$ we need $P(X, Y)$:

P(X,Y)	dark	blonde
dark	$\frac{9}{20}$	$\frac{1}{5}$
blonde	$\frac{3}{10}$	$\frac{1}{20}$

Table 2: Visualization of P(X,Y).

Meaning $P(X|Y = dark) = (\frac{9}{20}\frac{4}{3}, \frac{3}{10}\frac{4}{3}) = (\frac{3}{5}, \frac{2}{5})$

(c) Calculate H(X) in 3 decimals.

$$H(X) = -\frac{7}{10}log_2(\frac{7}{10}) - \frac{3}{10}log_2(\frac{3}{10}) \approx 0.881 \tag{14}$$

(d) Show that the uncertainty of the man's hair colour becomes larger if it is known that his wife has dark hair. Round up to 3 decimals.

Answer: to determine this we need to calculate the entropy of $P(X|Y = dark)$**:**

$$H(X|Y = dark) = -\frac{3}{5}log_2(\frac{3}{5}) - \frac{2}{5}log_2(\frac{2}{5}) \approx 0.970 > H(X) \approx 0.881 \tag{15}$$

(e) How much information about the hair colour of the man is revealed by the hair colour his wife? Round up to 3 decimals.

Answer: This is a generalization of the last task requiring the computation of the entropy of $P(X|Y)$

$$H(X|Y) = \sum_{i=1}^{2} H(X|Y = y_i)P(Y = y_i) \approx$$
$$\approx 0.970 * \frac{3}{4} - \frac{4}{5}log_2(\frac{4}{5})\frac{1}{4} - \frac{1}{5}log_2(\frac{1}{5})\frac{1}{4} \approx 0.908 \tag{16}$$

5 **Suppose we describe the climate in a country by writing down a few weather types together with the probability of each type. In the table below these information is given for two exotic countries. We denote their probability distributions as PF and PR. (See table in textbook) A lost explorer is situated in one of these countries. He throws a bottle in the ocean on which he has written the weather of the past 21 days: RSSRSTSRSRTSSS-RTRTSSS.**

(a) Estimate the probabilities on every wheather type in the country the explorer dwells in. Denote this probability distribution by P_E.

Answer: The best estimate for the probability of each weather occurring is the data itself and since each day is independent of every other then: $P_E = (\frac{11}{21}, \frac{6}{21}, \frac{5}{21})$

(b) Calculate the Kullback-Leibler distances $D(P_F, P_E)$ and $D(P_R, P_E)$ in 3 decimals.

Answer:

$$D(P_F, P_E) = \frac{1}{3}log_2(\frac{1}{3}\frac{21}{11}) + \frac{1}{2}log_2(\frac{1}{2}\frac{21}{6}) + \frac{1}{6}log_2(\frac{1}{6}\frac{21}{5}) \approx 0.100 \qquad (17)$$

$$D(P_R, P_E) = \frac{1}{2}log_2(\frac{1}{2}\frac{21}{11}) + \frac{2}{5}log_2(\frac{2}{5}\frac{21}{6}) + \frac{1}{10}log_2(\frac{1}{10}\frac{21}{5}) \approx 0.035 \qquad (18)$$

(c) You have found the bottle and decide to send out a rescue brigade. To which country would you send this brigade according to the answer found in b?

Answer: KL-divergence measures the distance in entropy from one distribution to another meaning the lower the distance the closer to the true distribution we are. In this case P_E is closer to P_R than P_F so it is more likely the lost explorer is in Farawayea than Remotelystan.

6 Archeologists have dug up a clay tablet with the following text:

$$\text{III} \int \Pi \int \int \cup \int \cup \int \Pi \int \Pi \int \int \cup \cup \int \int \cup \qquad (19)$$

There is some doubt about the language in which this text has been written. Some historians think it is a primitive variant of Gallic, while others think it might be Keltic. Earlier the archeologists found fragments of both Gallic and Keltic, consisting of 200 symbols. The symbol $\int$ occured 107 times in the Gallic text, while Π appeared 57 times. In the Keltic text these respective numbers are 90 and 56. Assume that both languages consist of only 3 symbols.

(a) Estimate the letter frequencies in both languages and in the text written on the clay tablet.

Answer:

$$P_G = (\frac{107}{200}, \frac{57}{200}, \frac{36}{200})$$
$$P_K = (\frac{90}{200}, \frac{56}{200}, \frac{54}{200}) \qquad (20)$$
$$P_T = (\frac{10}{20}, \frac{5}{20}, \frac{5}{20})$$

(b) Determine the symmetric Kullback-Leibler distances between the probability distribution of the clay tablet and the distributions corresponding to the earlier found texts, in 3 decimals. In what language is the text on the tablet probably written?

Answer:

$$D_{sym}(P_G, P_T) = \frac{D(P_G, P_T) + D(P_T, P_G)}{2} =$$

$$= \frac{1}{2}\sum_{i=1}^{3} P(P_G = p_{gi})log_2(\frac{P(P_G = p_{gi})}{P(P_T = p_{ti})}) +$$

$$+ \frac{1}{2}\sum_{i=1}^{3} P(P_T = p_{ti})log_2(\frac{P(P_T = p_{ti})}{P(P_G = p_{gi})}) = \tag{21}$$

$$= \frac{1}{2}(\frac{107}{200}log_2(\frac{107}{100}) + \frac{57}{200}log_2(\frac{57}{50}) + \frac{36}{200}log_2(\frac{36}{50}) +$$

$$+ \frac{10}{20}log_2(\frac{100}{107}) + \frac{5}{20}log_2(\frac{50}{57}) + \frac{5}{20}log_2(\frac{50}{36})) \approx 0.021$$

$$D_{sym}(P_K, P_T) = \frac{D(P_K, P_T) + D(P_T, P_K)}{2} =$$

$$= \frac{1}{2}\sum_{i=1}^{3} P(P_K = p_{ki})log_2(\frac{P(P_K = p_{ki})}{P(P_T = p_{ti})}) +$$

$$+ \frac{1}{2}\sum_{i=1}^{3} P(P_T = p_{ti})log_2(\frac{P(P_T = p_{ti})}{P(P_K = p_{ki})}) = \tag{22}$$

$$= \frac{1}{2}(\frac{90}{200}log_2(\frac{90}{100}) + \frac{56}{200}log_2(\frac{56}{50}) + \frac{54}{200}log_2(\frac{54}{50}) +$$

$$+ \frac{10}{20}log_2(\frac{100}{90}) + \frac{5}{20}log_2(\frac{50}{56}) + \frac{5}{20}log_2(\frac{50}{54})) \approx 0.007$$

From the KL-divergence calculations we can conclude it is more likely the clay tablet is Keltic.

7 Kasparov and Federer play a chess match consisting of 3 games. The outcomes of each game are independent of one another. Every game satisfies the probability distribution of Y given by: $P(Y) = (\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$. We denote the outcomes of the three games as Y_1, Y_2 and Y_3. The possible outcomes will be abbreviated to K, F and D. If one of both players wins more games than his opponent, he wins the match. Otherwise the matchs ends unresolved. The outcome of the match will be called X. There are 3 possible outcomes of X, namely K (Kasparov wins), F (Federer wins) and U (Unresolved). It is given that $P(X = U) = \frac{11}{54}$.

(a) Give the Huffman code for Y.

Answer: K - 0, D - 10, F - 11

(b) Determine for this coding the expected number of bits necessary to encode the first set.

Answer:

$$E(\#bits) = \frac{1}{2} + \frac{1}{6} * 2 + \frac{1}{3} * 2 = 1.5 \text{ expected bits} \tag{23}$$

(c) Compute the probability that Kasparovs wins the match, i.e. $P(X = K)$. Hint: distinguish between three cases.

Answer: To compute the probability we can split all Kasparov match wins into three cases: one win, two wins, three wins. For the three wins case there is only one way of achieving such a match i.e. KKK - $\frac{1}{8}$. For the one win case there are three ways - KDD $\frac{1}{18}$, DKD - $\frac{1}{18}$, DDK - $\frac{1}{18}$. For the last case of two wins there are two subtypes of matches one where Federer wins and one where he doesn't with three different ways both can happen. $\frac{3}{12} + \frac{3}{24}$ in total the sum of all cases is: $P(X = K) = \frac{2}{3}$.

(d) Calculate the entropy H(X) in 3 decimals.

Answer:

$$P(X) = (\frac{2}{3}, \frac{7}{54}, \frac{11}{54})$$
$$H(X) = -\frac{2}{3}log_2(\frac{2}{3}) - \frac{7}{54}log_2(\frac{7}{54}) - \frac{11}{54}log_2(\frac{11}{54}) \approx 1.239 \text{ bits}$$

(24)

(e) Show that the uncertainty about the outcome of the match increases if Federer wins the first set, that means if $Y_1 = F$. Give your answer in 3 decimals.

Answer:

$$H(X|Y_1 = F) = -\sum_{i=1}^{3} P(X = x_i|Y_1 = F)log_2(P(X = x_i|Y_1 = F))$$
$$= -P(X = K|Y_1 = F)log_2(P(X = K|Y_1 = F))-$$
$$- P(X = F|Y_1 = F)log_2(P(X = F|Y_1 = F))-$$
$$= P(X = D|Y_1 = F)log_2(P(X = D|Y_1 = F))$$

$$P(X = K|Y_1 = F) = \frac{P(X = K, Y_1 = F)}{P(Y_1 = F)} = \frac{\frac{1}{24}}{\frac{1}{6}} = \frac{1}{4}$$

(25)

$$P(X = F|Y_1 = F) = \frac{P(X = F, Y_1 = F)}{P(Y_1 = F)} = \frac{\frac{15}{36*6}}{\frac{1}{6}} = \frac{15}{36}$$

$$P(X = D|Y_1 = F) = \frac{P(X = D, Y_1 = F)}{P(Y_1 = F)} = \frac{\frac{2}{36}}{\frac{1}{6}} = \frac{1}{3}$$

$$H(X|Y_1 = F) = -\frac{1}{4}log_2(\frac{1}{4}) - \frac{15}{36}log_2(\frac{15}{36}) - \frac{1}{3}log_2(\frac{1}{3}) \approx$$
$$\approx 1.554 \text{ bits} > H(X) \approx 1.239 \text{ bits}$$

8 Adrian and Bruno have arrived at Schiphol. Adrian is going to tally how many KLM planes land, while Bruno will tally the number of planes there land before the first KLM plane lands. After four planes however, they have lost their interest and stop counting. We call the number of planes they have counted respectively A and B. For example, if only the third plane was of KLM, we have A = 1 and B = 2. Assume that the probability of a KLM aircraft is equal to $\frac{1}{2}$.

(a) Determine the probability distribution of B.

Answer: $P(B) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$

(b) Calculate the entropy H(B).

$$H(B) = -\frac{1}{2}log_2(\frac{1}{2}) - \frac{1}{4}log_2(\frac{1}{4}) - \frac{1}{4}log_2(\frac{1}{8}) = 1.75 \text{ bits} \tag{26}$$

(c) Give the Huffman coding of B.

Answer: $B - (0, 10, 110, 111)$

(d) Determine the expected number of bits to encode an outcome with the code found in (c).

Answer:

$$E(\#bits) = \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} = 1.75 \text{ expected bits} \tag{27}$$

(e) Show that the uncertainty about B increases if Adrian tells he has counted only one plane. (You thus have to compute $H(B|A = 1)$.)

Answer:

$$H(B|A = 1) = -\sum_{i=0}^{4} P(B = b_i|A = 1)log_2(P(B = b_i|A = 1)) = \tag{28}$$

$$= -4 * \frac{1}{4}log_2(\frac{1}{4}) = 2 \text{ bits} > H(B) = 1.75$$

9 In a train there is an equal number of students and commuters. Half of the commuters travels first class. Moreover one of every twelve travelers is a student that travels first class. We define the random variable T as the type (either student or commuter) of an arbitrary traveler, and the random variable K as the class in which (s)he travels. The joint distribution of T and K can be visualized by the following table:

	T = student	T = commuter	total
K = 1	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$
K = 2	$\frac{5}{12}$	$\frac{1}{4}$	$\frac{2}{3}$
total	$\frac{1}{2}$	$\frac{1}{2}$	1

Table 3: Visualization of P(K,T).

(a) Fill in this table.

(b) Determine $H(T|K = 1)$.

Answer:

$$P(K) = (\frac{1}{3}, \frac{2}{3})$$

$$H(T|K = 1) = -\sum_{i=1}^{2} P(T|K = 1) log_2(P(T|K = 1)) = \quad (29)$$

$$= -\frac{1}{4} log_2(\frac{1}{4}) - \frac{3}{4} log_2(\frac{3}{4}) \approx 0.811 \text{ bits}$$

(c) Is the entropy H(T) larger than, smaller than, or equal to the conditional entropy $H(T|K)$? (You can answer this question without doing any computations! If you however choose to do some calculations, you may use that $H(T|K = 2) \approx 0.954$ and you may give $H(T|K)$ in 3 decimals.)

Answer:

$$H(T) = 1 \text{ bit}$$

$$H(T|K) = \sum_{i=1}^{2} H(T|K = k_i)P(K = k_i) = \tag{30}$$

$$\approx 0.811 * \frac{1}{3} + 0.954 * \frac{2}{3} \approx 0.906 < H(T) = 1$$

(d) In a compartment are as many students as commuters. Is this compartment more likely to be first or second class? Motivate your answer by means of relative entropy. (An exact answer is not required.)

Answer: We can answer this my comparing the probability distribution of people in this cart with the expected 1st and 2nd class cars. The distribution of this car is the same as T: $P(T) = (\frac{1}{2}, \frac{1}{2})$

$$H(K = 1) = (\frac{1}{4}, \frac{3}{4})$$

$$H(K = 2) = (\frac{5}{8}, \frac{3}{8})$$

$$D(T, K = 1) = \frac{1}{2}log_2(2) + \frac{1}{2}log_2(\frac{2}{3}) \approx 0.207 \tag{31}$$

$$D(T, K = 2) = \frac{1}{2}log_2(\frac{4}{5}) + \frac{1}{2}log_2(\frac{4}{3}) \approx 0.046 < D(T, K = 1) \approx 0.207$$

meaning it's more likely this is a 2nd class car.

10 Pete plays bingo. In four rounds the numbers 11, 25, 31 and 62 are announced in arbitrary order. Two of those numbers (25 and 62) are on the bingo card of Pete. With F and L we denote the first respectively last round in which the numbers of Pete are announced. Obviously F < L. For example, if '11, 62, 25, 31' is the order in which the numbers are announced, then F = 2 and L = 3.

(a) Show that the probability distribution of L is given by:

$$P(L = 2) = \frac{1}{6}, P(L = 3) = \frac{1}{3}, P(L = 4) = \frac{1}{2} \tag{32}$$

Answer: In order for L = 2 we need both chosen by Pete numbers to be the first 2 and both not chosen to be last there are 2 different orientations both groups can take meaning there are 4 total sequences. The total number of possible sequences is 4! so $P(L = 2) = \frac{4}{4!} = \frac{1}{6}$. **For L = 4 one of the numbers chosen by Pete has to be in position 4 and the others can be in any orientation which gives** $2 * 3!$ **for the total number of valid position since for each orientation of the first three we can swap the 2 numbers chosen by Pete and we would get a new sequence.** $P(L = 4) = \frac{2*3!}{4!} = \frac{1}{2}$. **For L = 3 we can subtract the other 2 options** $P(L = 3) = 1 - P(L = 2) - P(L = 4) = \frac{1}{3}$.

(b) Determine the entropy of L.

Answer:

$$H(L) = -\frac{1}{6}log_2(\frac{1}{6}) - \frac{1}{3}log_2(\frac{1}{3}) - \frac{1}{2}log_2(\frac{1}{2}) \approx 1.459 \text{ bits} \tag{33}$$

(c) Give the Huffman coding of L.

Answer: $L - (11, 10, 0)$

(d) Determine the expected number of bits for this encoding.

Answer:

$$E(\#bits) = \frac{1}{2} + 2 * \frac{1}{3} + 2 * \frac{1}{6} = 1.5 \text{ expected bits} \tag{34}$$

(e) Compute $I(L|F = 1)$, i.e. the information that the outcome F = 1 reveals about L. Interpret your answer.

Answer:

$$I(L|F = 1) = H(L) - H(L|F = 1)$$

$$H(L|F = 1) = -\sum_{i=2}^{4} P(L = l_i|F = 1)log_2(P(L = l_i|F = 1)) = \tag{35}$$

$$= -3 * \frac{4}{2 * 3!}log_2(\frac{4}{2 * 3!}) = log_2(3) \approx 1.584$$

$$I(L|F = 1) \approx 1.459 - 1.584 \approx -0.125 \text{ bits}$$

A negative bit value indicates that knowing the fact that F = 1 reduces our certainty regarding the value of L.

11 Suppose that X and Y are fair tetrahedral dice. The four sides of X and Y display the complex numbers 1, -1, i, -i. Note that complex numbers will be formally introduced in Chapter 5. For the moment it is enough to make use of the equality $i^2 = -1$. Let P be the product of the outcomes, i.e. $P = X * Y$.

(a) Calculate the entropy of X.

Answer:

$$H(X) = -4 * \frac{1}{4}log_2(\frac{1}{4}) = log_2(4) = 2 \text{ bits} \tag{36}$$

(b) Give the Huffman coding for X. Is this coding optimal?

Answer: $X - (00, 01, 10, 11)$

$$H(\#bits) = 4 * 2 * \frac{1}{4} = 2 \text{ expected bits} = H(X) = 2 \tag{37}$$

meaning this encoding is optimal.

(c) Give all possible outcomes of P together with their probabilities.

Answer:

P	1	-1	i	-i
	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Table 4: Visualization of P.

(d) How much information does X reveal about P?

Answer:

$$I(P|X) = H(P) - H(P|X)$$
$$H(P) = log_2(4) = 2 \text{ bits}$$
$$H(P|X) = -\sum_{i=1}^{4}\sum_{j=1}^{4} P(P = p_j|X = x_i)log_2(P(P = p_j|X = x_i))P(X = x_i) =$$
$$= \sum_{i=1}^{4} 2 * \frac{1}{4} = 2 \text{ bits} I(P|X) = 2 - 2 = 0 \text{ bits} \tag{38}$$

meaning knowing the value of X reveals no information about P.