Solution Sheet 1

Last modified November 18, 2019

Problem 1

$$\mathsf{H}(X) = -\sum_{i=0}^{M} p_i \log p_i$$
$$\mathsf{H}(Y) = -\sum_{i=0}^{M} q_i \log q_i$$

Rewriting and substituting the problem definition results in

$$\mathsf{H}(X) = -\sum_{i=0}^{j-2} p_i \log(p_i) - p_{j-1}\log(p_{j-1}) - p_j \log(p_j) - \sum_{i=j+1}^{M} p_i \log(p_i),$$
$$\mathsf{H}(Y) = -\sum_{i=0}^{j-2} p_i \log(p_i) - \frac{p_{j-1}+p_j}{2}\log\left(\frac{p_{j-1}+p_j}{2}\right) - \frac{p_{j-1}+p_j}{2}\log\left(\frac{p_{j-1}+p_j}{2}\right) - \sum_{i=j+1}^{M} p_i \log(p_i).$$

Note that $\mathsf{H}(X)$ and $\mathsf{H}(Y)$ are identical except the terms outside of the sums.

Approach 1: Only the terms in $\mathsf{H}(X)$ and $\mathsf{H}(Y)$ that include p_j and p_{j-1} are of interest and we thus define

$$\tilde{\mathsf{H}}(X) = -p_j \log(p_j) - p_{j-1}\log(p_{j-1}),$$
$$\tilde{\mathsf{H}}(Y) = (-2)\,\frac{p_{j-1}+p_j}{2}\log\left(\frac{p_{j-1}+p_j}{2}\right).$$

Introduce $c = p_j + p_{j-1}$ and get

$$\tilde{\mathsf{H}}(X) = -p_j \log(p_j) - (c-p_j)\log(c-p_j),$$
$$\tilde{\mathsf{H}}(Y) = -c\log\left(\frac{c}{2}\right).$$

Find the extremum of $\tilde{\mathsf{H}}(X)$ by computing

$$\frac{\partial \tilde{\mathsf{H}}(X)}{\partial p_j} = -\log(p_j) + \log(c-p_j) = 0 \quad \Rightarrow \quad p_j = \frac{c}{2}.$$

Furthermore, $p_j = c/2$ is a maximum, which can be verified by

$$\left.\frac{\partial^2 \tilde{\mathsf{H}}(X)}{\partial p_j^2}\right|_{p_j=\frac{c}{2}} = -\frac{1}{p_j} - \left.\frac{1}{c-p_j}\right|_{p_j=\frac{c}{2}} = -\frac{4}{c} < 0.$$

Thus,

$$\max \tilde{\mathsf{H}}(X) = -\frac{c}{2}\log\left(\frac{c}{2}\right) - \left(c-\frac{c}{2}\right)\log\left(c-\frac{c}{2}\right) = -c\log\left(\frac{c}{2}\right).$$

Since $\max \tilde{\mathsf{H}}(X) = \tilde{\mathsf{H}}(Y)$, it follows that $\tilde{\mathsf{H}}(X) \le \tilde{\mathsf{H}}(Y)$, and consequently

$$\mathsf{H}(X) \le \mathsf{H}(Y).$$

Approach 2: Use the fact that

$$\ln x \le x - 1,$$

and compute

$$\begin{aligned}
\mathsf{H}(X) - \mathsf{H}(Y) &= -p_{j-1}\log(p_{j-1}) - p_j\log(p_j) + (p_{j-1}+p_j)\log\left(\frac{p_{j-1}+p_j}{2}\right) \\
&= p_{j-1}\log\left(\frac{p_{j-1}+p_j}{2p_{j-1}}\right) + p_j\log\left(\frac{p_{j-1}+p_j}{2p_j}\right) \\
&= \frac{1}{\ln 2}\left(p_{j-1}\ln\left(\frac{p_{j-1}+p_j}{2p_{j-1}}\right) + p_j\ln\left(\frac{p_{j-1}+p_j}{2p_j}\right)\right) \\
&\le \frac{1}{\ln 2}\left(p_{j-1}\left(\frac{p_{j-1}+p_j}{2p_{j-1}} - 1\right) + p_j\left(\frac{p_{j-1}+p_j}{2p_j} - 1\right)\right) \\
&= \frac{1}{\ln 2}\left(\frac{p_{j-1}+p_j}{2} - p_{j-1} + \frac{p_{j-1}+p_j}{2} - p_j\right) \\
&= \frac{1}{\ln 2}(p_{j-1}+p_j-p_{j-1}-p_j) \\
&= 0.
\end{aligned}$$

Therefore,

$$\mathsf{H}(X) \le \mathsf{H}(Y).$$

Problem 2

1. The expected value is given by

$$\begin{aligned}
\mu_N = \mathbb{E}_N[N] &= \sum_{n=1}^{\infty} np(1-p)^{n-1} \\
&= -p\sum_{n=1}^{\infty}\frac{\partial}{\partial p}(1-p)^n \\
&= -p\frac{\partial}{\partial p}\left(\sum_{n=0}^{\infty}(1-p)^n - 1\right) \\
&= -p\frac{\partial}{\partial p}\left(\frac{1}{1-(1-p)} - 1\right) = \frac{1}{p}
\end{aligned}$$

2. The entropy is given by

$$\begin{aligned}
\mathsf{H}(N) = \mathbb{E}_N[-\log(P_N(N))] &= -\sum_{n=1}^{\infty} p(1-p)^{n-1}\log(p(1-p)^{n-1}) \\
&= -\sum_{n=1}^{\infty} p(1-p)^{n-1}\left(\log(p) + (n-1)\log(1-p)\right) \\
&= -p\log(p)\sum_{n=1}^{\infty}(1-p)^{n-1} - p\log(1-p)\sum_{n=1}^{\infty}(n-1)(1-p)^{n-1} \\
&= -p\log(p)\sum_{n=0}^{\infty}(1-p)^{n} - p\log(1-p)\sum_{n=0}^{\infty}(n)(1-p)^{n} \\
&= -\frac{p\log(p)}{p} - \frac{p(1-p)\log(1-p)}{p^2} \\
&= \frac{1}{p}\left(-p\log(p) - (1-p)\log(1-p)\right) = \mu_N \cdot \mathsf{H_b}(p)
\end{aligned}$$

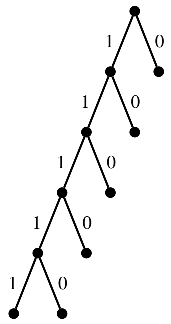
Problem 3

1. A prefix-free code is always uniquely decodable because for every finite sequence one can start decoding at the beginning of the sequence and read as many symbols as necessary to obtain a valid codeword. As

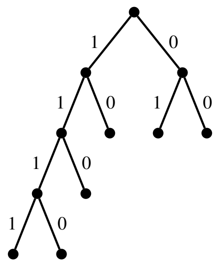
no codeword is the prefix of another codeword, this always results in one unique sequence of codewords. Prefix-free codes are also called instantaneously decodable codes.

2. a) This is a block code (all codewords have the same length) where none of the codewords are identical, and therefore the code is prefix-free.
   b) prefix-free
   c) not prefix-free, since 0 is a prefix of 01. However, the code is uniquely decodable.
   d) not prefix-free, since 1 is a prefix of 101. However, the code is uniquely decodable.
   e) This code is neither prefix-free, nor uniquely decodable. For example, 01 has two different interpretations.

3. a) $2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-5}=1$, therefore a code exists, and can be visualized with the help of the following tree



   b) $2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-4}+2^{-5}=33/32>1$, therefore a code does not exist
   c) $2^{-2}+2^{-2}+2^{-2}+2^{-3}+2^{-4}=15/16<1$, therefore a code exists, and can be visualized with the help of the following tree



Problem 4

1. Let $\mathcal{A}$ = "second ball is red" and $\mathcal{B}$ = "first ball is labeled B". We have
   $\Pr(\mathcal{A})=1/4+1/10+1/20=2/5$
   $\Pr(\mathcal{B})=1/2$
   $\Pr(\mathcal{A}\cap\mathcal{B})=1/4$
   Therefore, $\Pr(\mathcal{B}|\mathcal{A})=(1/4)/(2/5)=5/8$.

2. Since $\Pr(\mathcal{B}|\mathcal{A})\neq\Pr(\mathcal{B})$, these two events are not independent.

3. Let $\mathcal{C}$ = "first ball is labeled C". We have
   $\Pr(\mathcal{C})=1/4$
   $\Pr(\mathcal{A}\cap\mathcal{C})=1/10$
   Therefore, $\Pr(\mathcal{A}\cap\mathcal{C})=\Pr(\mathcal{A})\cdot\Pr(\mathcal{C})=1/10$ and the two events are independent.

Problem 5

1.

x	$P_X(x)$	$(P_X(x))^2$	$-\log_2(P_X(x))$
0	0.1	0.01	3.32
1	0.2	0.04	2.32
2	0.7	0.49	0.515

$\mathbb{E}[X] = \sum_x x P_X(x) = 1.6$
$\mathbb{E}[P_X(X)] = \sum_x P_X(x) P_X(x) = 0.54$
$\mathbb{E}[-\log_2(P_X(X))] = 1.157$

Only the expected value changes to $\mathbb{E}[X] = 2.6$.

2. We have

$$\begin{aligned}
\mathbb{E}_Y[\mathbb{E}_X[X|Y]] &= \sum_y \mathbb{E}_X[X|Y=y] \cdot P_Y(y) \\
&= \sum_y \sum_x x \cdot P_{X|Y}(x|y) \cdot P_Y(y) \\
&= \sum_y \sum_x x \cdot \frac{P_{X,Y}(x,y)}{P_Y(y)} \cdot P_Y(y) \\
&= \sum_x x \sum_y P_{X,Y}(x,y) \\
&= \sum_x x P_X(x) = \mathbb{E}_X[X]
\end{aligned}$$

Problem 6

1. $P_X(0) = 8/27$, $P_X(1) = 12/27$, $P_X(2) = 6/27$, $P_X(3) = 1/27$

2.

	$X=0$	$X=1$	$X=2$	$X=3$	$\sum(\cdot) = P_Y(y)$
$Y=0$	8/27	8/27	2/27	0	2/3
$Y=1$	0	4/27	4/27	1/27	1/3
$\sum(\cdot) = P_X(x)$	8/27	12/27	6/27	1/27	

3. $\mathbb{E}_X[X] = 0 \cdot 8/27 + 1 \cdot 12/27 + 2 \cdot 6/27 + 3 \cdot 1/27 = 1$
$\mathbb{E}_Y[Y] = 0 \cdot 2/3 + 1 \cdot 1/3 = 1/3$
$\mathbb{E}_{P_{X,Y}}[XY] = 1 \cdot 4/27 + 2 \cdot 4/27 + 3 \cdot 1/27 = 15/27$
$\mathbb{E}_{P_{X,Y}}[(X - \mathbb{E}_X[X])(Y - \mathbb{E}_Y[Y])] = 15/27 - 1/3 = 2/9$
Therefore, X and Y are correlated.

4. $\Pr(N \leq n) = \sum_{i=0}^{n} P_N(i)$, where $P_N(i) = 1/3 \cdot (2/3)^i$. Therefore

$$\Pr(N \leq n) = \sum_{i=0}^{n} 1/3 \cdot (2/3)^i = 1/3 \cdot \frac{(2/3)^{n+1} - 1}{2/3 - 1} = 1 - (2/3)^{n+1}$$

Problem 7

1. $Y \sim \mathcal{N}(0,1)$

2. They are not correlated since $\mathbb{E}_{P_{X,Y}}[XY] = \mathbb{E}_Z[Z] \cdot \mathbb{E}_X[|X|^2] = 0 \cdot \mathbb{E}_X[|X|^2] = 0$.

3. $\Pr(|X| \geq 1) = 2\Pr(X \geq 1) = 2Q(1)$

   $\Pr(|Y| \geq 1) = 2\Pr(Y \geq 1) = 2Q(1)$

4. $\Pr(|X| \geq 1, |Y| \geq 1) = \Pr(X \geq 1) = 2Q(1)$

5. No, since $\Pr(|X| \geq 1, |Y| \geq 1) \neq \Pr(X \geq 1) \cdot \Pr(X \geq 1)$