Worksheet

Many of these problems are taken from the excellent text book Cover and Thomas. Although the questions do vary a bit in difficulty each is worth two marks. Q3 is probably harder than Q4.

Q1 - marginal and conditional distributions

Work out the marginal probability distributions and the $x = a$ conditional probability distribution $P(Y|X = a)$ for

$Y \backslash X$	a	b
1	$\frac{1}{3}$	$\frac{1}{6}$
2	0	$\frac{1}{4}$
3	$\frac{1}{8}$	$\frac{1}{8}$

Q2 - working out entropy

A fair coin is flipped until the first head occurs. Let X denote the number of flips required.

1. Find the entropy $H(X)$ in bits. The following expressions may be useful:

$$\begin{aligned}\sum_{n=0}^{\infty} r^n &= \frac{1}{1-r} \\ \sum_{n=0}^{\infty} nr^n &= \frac{r}{(1-r)^2}\end{aligned} \tag{1}$$

2. A random variable X is drawn according to this distribution. Find a sequence of yes-no questions of the form, 'Is X contained in the set S?'. Compare $H(X)$ to the expected number of questions required to determine X. For the most efficient sequence, that is the sequence the shortest expected number of questions, these two numbers will be the same. You will find that the most efficient sequence is very straightforward!

Q3 - A puzzle which lends itself to information type reasoning

Suppose that you have n coins, among which there may or may not be one counterfeit coin. If there is a counterfeit coin it will weigh either less or more than the other coins. The coins are weighed using a balance, any number of coins van be put on each side of the balance, though obviously you will want the same number on each side.

1. Find an upper bound on the number of coins n so that k weighings will find the counterfeit coin, if any, and correctly declare it to be heavier or lighter.

2. What is the coin-weighing strategy for $k = 3$ weighings and 12 coins,

Q4 - Working out entropy and information

Let $p(x, y)$ be given by $p(0, 0) = p(0, 1) = p(1, 1) = 1/3$ and $p(1, 0) = 0$. Find $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, $H(X, Y)$, $H(Y) - H(Y|X)$ and $I(X; Y)$.

Q5 - A question about information in the brain

Answer just one of these two questions, each is worth equal marks but the second is much harder than the first, so you'd be better off doing the first unless you are particularly interested in this topic. Both papers are available in the paper repository in the github.

1. The original idea of estimating neural information by binning spike trains was spread across several papers, but one of the main references is Strong et al. (1998). One aspect of this paper we didn't discuss is the use of extrapolation to estimate the information as the number of samples becomes large based on the behaviour for smaller numbers of samples. Can you give a short, up to five line, summary of what this involves.

2. In Nemenman et al. (2004) there is a deep commentary on how information in neural data is computed. This is a very difficult paper and the mathematics towards the end is hard. The aim of this question is to read the paper and offer a three or four line overall summary of what the paper is trying to do.

References

Nemenman, I., Bialek, W., and van Steveninck, R. d. R. (2004). Entropy and information in neural spike trains: Progress on the sampling problem. *Physical Review E*, 69(5):056111.

Strong, S. P., Koberle, R., van Steveninck, R. R. d. R., and Bialek, W. (1998). Entropy and information in neural spike trains. *Physical review letters*, 80(1):197.