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 and you should do five including Q1 and Q2. The expectation is that you will submit written work, though you are welcome to type up your answers as well. Q5 includes a number of graphs; ideally you should print these out and include them with your submission. The submission date is Friday 24 October; the end of week four.

Q1 - marginal and conditional distributions

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

Y \ X	a	b
1	$\frac{1}{16}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{8}$

Q2 - working out entropy

For the above distribution work out $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, $H(X, Y)$, $H(Y) - H(Y|X)$ and $I(X; Y)$.

Q3 - working out entropy

The World Series is a competition held each year in North America between two baseball teams. The series consists of between four and seven games, terminating if either team wins four games. Thus, the set of outcomes includes sequences like AAAA, ABAAA and ABABABA. Let X be the random variable representing the outcome and Y the number of games played. Assuming the teams are equally matched and the games are independent, what are $H(X)$, $H(Y)$, $H(X|Y)$ and $H(Y|X)$?

Q4 - the average entropy

Work out the average entropy for the distribution with two events $\{x_1, x_2\}$ and $p(x_1) = p$ and $p(x_2) = 1 - p$ under the assumption that each value of p is equally likely.

Q5 - Bias in estimating information

Estimating entropy is hard; worse, the obvious estimator

$$H(X) = -\sum_x \tilde{p}(x) \log_2 \tilde{p}(x) \tag{1}$$

where

$$\tilde{p}(x) = \frac{\#(\text{trials giving } x)}{\text{total trials}} \tag{2}$$

is biased. Write a short programme to graph this for eight equally likely outcomes. In this case the entropy should be three but you should simulate estimating the entropy from n trials; in other words, pick from the eight items n times, calculate $\tilde{p}$ and estimate $H(X)$; do this multiple times for each n and plot the estimated entropy against n. Does Laplace smoothing help; Laplace smoothing is an alternative estimator for the probability

$$\tilde{p}_\alpha(x) = \frac{\#(\text{trials giving } x) + \alpha}{\text{total trials} + d\alpha} \tag{3}$$

where d is the number of outcomes, eight in our case, and α is a parameter usually taken as lying between zero and one. Try this a a few value of α, for example 0.25, 0.5 and 1.0.

Q6 - 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 lines, 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.