Worksheet

Many of these problems are taken from the excellent text book Cover and Thomas.

Useful formulas

- **Shannon's entropy**: $H(X) = -\sum_i p(x_i) \log_2 p(x_i)$

- **joint entropy**: $H(X, Y) = -\sum_{i,j} p(x_i, y_j) \log_2 p(x_i, y_j)$

- **conditional entropy**: $H(X|Y) = -\sum_{i,j} p(x_i, y_j) \log_2 p(x_i|y_j)$

- **the chain rule for entropy**: $H(X, Y) = H(X|Y) + H(Y)$

- **mutual information**:

$$I(X, Y) = \sum_{i,j} p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

- $I(X, Y) = H(X) + H(Y) - H(X, Y)$

- $I(X, Y) = H(X) - H(X|Y)$

- $I(X, Y) = H(Y) - H(Y|X)$

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 $I(X; Y)$.

comsm0075.github.io

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.