Shannon's Enropy

COMSM0075 Information Processing and Brain

`comsm0075.github.io`

September 2020

Shannon's entropy

For a finite discrete distribution with random variable X, possible outcomes $\{x_1, x_2, \ldots x_n\} \in \mathcal{X}$ and a probability mass function p_X giving probabilities $p_X(x_i)$, the entropy is

$$H(X) = -\sum_{x_i \in \mathcal{X}} p_X(x_i) \log_2 p_X(x_i)$$

Shannon's entropy

The probabilities can be estimated

$$p(x_i) \approx \frac{f(x_i)}{\sum_i f(x_i)}$$

where $f(x_i)$ is the number of times x_i occurs.

Biased estimation

Note that this gives a biased estimator. If all events are equally likely then $p(x_i) = 1/n$ but the estimates of the probability won't give equal estimates and so if $H(X) = \log_2 n$, estimates of $H(X)$ from data will typically give a lower value.

Biased estimation

An artificial experiment could be done to calculate the average estimate of $H(X)$ for a particular n and a particular number of samples.