COMP2610/COMP6261 - Information Theory

Tutorial 4: Entropy and Information

Young Lee and Bob Williamson
Tutors: Debashish Chakraborty and Zakaria Mhammedi

Week 5 (21st - 25th August), Semester 2, 2017

1. Suppose Y is a geometric random variable, Y ~ Geom(y) . i.e., Y has probability function

$$P(Y = y) = p(1-p)^{y-1},\ y = 1, 2, \ldots$$

Determine the mean and variance of the geometric random variable.

2. A standard deck of cards contains 4 *suits* — ♡, ♢, ♣, ♠ ("hearts", "diamonds", "clubs", "spades") — each with 13 *values* — A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K (The A, J, Q, K are called "Ace", "Jack", "Queen", "King"). Each card has a *colour*: hearts and diamonds are coloured red; clubs and spades are black. Cards with values J, Q, K are called *face cards*.

Each of the 52 cards in a deck is identified by its value v and suit s and denoted vs. For example, 2♡, J♣, and 7♠ are the "two of hearts", "Jack of clubs", and "7 of spades", respectively. The variable c will be used to denote a card's colour. Let $f = 1$ if a card is a face card and $f = 0$ otherwise.

A card is drawn at random from a thoroughly shuffled deck. Calculate:

(a) The information $h(c = \text{red}, v = \text{K})$ in observing a red King

(b) The conditional information $h(v = \text{K}|f = 1)$ in observing a King given a face card was drawn.

(c) The entropies $H(S)$ and $H(V|S)$.

(d) The mutual information $I(V; S)$ between V and S.

(e) The mutual information $I(V; C)$ between the value and colour of a card using the last result and the *data processing inequality*.

3. Recall that for a random variable X, its variance is

$$\text{Var}[X] = \text{E}[X^2] - (\text{E}[X])^2.$$

Using Jensen's inequality, show that the variance must always be nonnegative.

4. Let X and Y be independent random variables with possible outcomes $\{0, 1\}$, each having a Bernoulli distribution with parameter $\frac{1}{2}$, i.e.

$$p(X = 0) = p(X = 1) = \frac{1}{2}$$
$$p(Y = 0) = p(Y = 1) = \frac{1}{2}.$$

(a) Compute $I(X; Y)$.

(b) Let $Z = X + Y$. Compute $I(X; Y|Z)$.

(c) Do the above quantities contradict the data-processing inequality? Explain your answer.