

COMP2610/COMP6261 - Information Theory

Tutorial 4: Entropy and Information

Week 4, Semester 2, 2021

1. Suppose Y is a geometric random variable, $Y \sim \text{Geom}(p)$. i.e., Y has probability function,

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

Determine the mean and variance of the geometric random variable.

2. The World Series is a seven-game series that terminates as soon as either team wins four games. Let X be the random variable that represents the outcome of a World Series between teams A and B; possible values of X are AAAA, BABABAB, and BBBAAAA. Let Y be the number of games played, which ranges from 4 to 7. Assuming that A and B are equally matched and that the games are independent, calculate

- a) $H(X)$
- b) $H(Y)$
- c) $H(Y | X)$
- d) $H(X | Y)$

3. Recall that for a random variable X , its variance is $\text{Var}[X] = E[X^2] - (E[X])^2$. 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.