Information Theory
Practice Set 2

Lakshmi Prasad Natarajan

Solutions are not to be returned

Reading Exercise

1. *(Review of probability theory).* Suppose $X_1, \ldots, X_n$ are random variables such that

$$X_i \text{ is independent of } (X_1, \ldots, X_{i-1}) \text{ for all } i = 2, \ldots, n.$$

   Argue that $X_1, \ldots, X_n$ are independent random variables.

2. Theorem 2.6.6 from Cover and Thomas.

3. *(The maximum a posteriori (MAP) rule).*
   Consider the problem of detecting X based on a jointly distributed random variable Y. Let $g : \mathcal{Y} \to \mathcal{X}$ denote the detection rule. We want to identify the rule that minimizes the probability of error P_e.

   Please ensure that you understand the correctness of the following steps.

   (a) First, we observe the following

$$\begin{aligned} P_e &= P[X \neq g(Y)] \\ &= \sum_y p(y)\, P[X \neq g(y) \ |Y = y] \\ &= \sum_y p(y)\, \left(1 - P[X = g(y) \ |Y = y] \right) \\ &= \sum_y p(y) \left(1 - p_{X|Y}(g(y)|y) \right) \end{aligned}$$

   (b) Next, observe that if, for each $y \in \mathcal{Y}$, $g(y)$ is chosen so that $\left(1 - p_{X|Y}(g(y)|y)\right)$ is minimized, then this g minimizes P_e.

   (c) Hence, the rule that minimizes P_e is the following

$$g_{\mathsf{MAP}}(y) = \arg\max_{x \in \mathcal{X}} p_{X|Y}(x|y).$$

   (d) $p_{X|Y}$ is called the *a posteriori probability*, and the rule $g_{\mathsf{MAP}}(y)$ mentioned above is called the *MAP rule.*

   (e) If for a given y there is more than one choice of x that maximizes $p_{X|Y}(x|y)$, then any of these choices of x can be set to be equal to $g_{\mathsf{MAP}}(y)$.

4. Solve Problem 2.32 from Cover & Thomas using the MAP rule.

Practice Set

1. Problems from Cover & Thomas (at the end of Chapter 2):

   2.1 (a) and (b), 2.2, 2.3, 2.5, 2.7 (a), 2.8, 2.14 (a)–(c), 2.17: explain the inequalities (a)–(e), 2.27, 2.37.