Information Theory
Practice Sets 3 and 4

Lakshmi Prasad Natarajan

Solutions are not to be returned

Warm-up exercise.

- Recall the formulae and description of
  - data processing inequality
  - definition of sufficient statistics
  - Fisher factorization theorem
  - convergence of a sequence of numbers
  - convergence in probability
  - weak law of large numbers
- Note that if $a_1, a_2, \ldots$ is a sequence of numbers with the limit a, and if $f(x)$ is any function continuous at a, then the sequence $f(a_1), f(a_2), \ldots$ converges to $f(a)$.

Practice Problems

Practice Set 3

1. From Cover & Thomas, Chapter 2:
   2.6, 2.15, 2.21, 2.23, 2.28 (only the first part of the question), 2.31, 2.39.
2. In the following scenarios, $X = (X_1, \ldots, X_n)$ where X_i are i.i.d with a distribution that is dependent on a parameter θ. Also, $T = f(X)$ is a function of X. You must show that T is a sufficient statistic of X for θ.
   (a) X_i are uniformly distributed in the interval $[-\theta, +\theta]$, and $t = (t_1, t_2) = f(x) = f(x_1, \ldots, x_n) = (\min_i x_i, \max_i x_i)$.
   (b) X_i are Gaussian with mean θ and variance 1, and $t = \sum_i x_i$. Is $t = \frac{\sum_i x_i}{n}$ a sufficient statistic?
   (c) X_i are Gaussian with mean 0 and variance θ, and $t = \sum_i x_i^2$.
3. Suppose T is a sufficient statistic of X for θ, and g is an invertible function. Is it true that $g(T)$ is also a sufficient statistic of X for θ ?

Practice Set 4

1. Cover & Thomas, Chapter 3:
   3.2, 3.6 (use the fact that log is a continuous function), 3.8, 3.9.

2. Suppose $|\mathcal{X}| < \infty$ and let $X_1, \dots, X_n$ be i.i.d on $\mathcal{X}$ with the probability mass function $p_X(x)$, $x \in \mathcal{X}$.

   (a) Let $a \in \mathcal{X}$. Use the weak law of large numbers to show that

   $$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i = a\} \xrightarrow{p} p_X(a),$$

   where $\mathbf{1}$ is the indicator function.

   (b) We now want to show that when n is large, with high probability, every element $a \in \mathcal{X}$ occurs approximately $p_X(a)$ times in the sequence $X_1, \dots, X_n$. That is, show that for any $\epsilon > 0$

   $$\lim_{n\to\infty} P\left[\bigcap_{a\in\mathcal{X}} \left\{ \left| \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i = a\} - p_X(a) \right| \le \epsilon \right\} \right] = 1$$

   `Hint`: You may want to use the union bound on probability $P[A \cup B] \le P[A] + P[B]$.

3. *Convergence in the mean square sense.* We say that random variables $Z_1, Z_2, \dots$ converge to a random variable Z in the mean square sense if

   $$\lim_{n\to\infty} \mathbb{E}\left[|Z_n - Z|^2\right] = 0.$$

   This is denoted as $Z_n \xrightarrow{m.s.} Z$.

   Show that if $X_1, X_2, \dots$ are i.i.d with finite mean μ and variance σ^2, then

   $$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{m.s.} \mu.$$
