Information Theory I

Fall Semester 2017

Prof. Dr. A. Lapidoth

Signal and Information Processing Laboratory

Institut für Signal- und Informationsverarbeitung

Exercise 1 of September 20, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it1.html

Problem 1 — *Expectation of a Chance Variable*

In Information Theory, one is frequently required to compute the expectation of $g(X)$ where X is a chance variable taking values in $\mathcal{X}$ and $g\colon \mathcal{X} \to \mathbb{R} \cup \{-\infty, +\infty\}$ is a function whose domain is $\mathcal{X}$ and which takes values on the *extended* real line $\mathbb{R} \cup \{-\infty, +\infty\}$. In this case, to avoid terms of the form $0 \cdot \infty$, the expectation of $g(\cdot)$, rather than being defined as

$$\mathsf{E}[g(X)] = \sum_{x \in \mathcal{X}} P_X(x)\, g(x) \tag{1}$$

as for ordinary real-valued functions, must be defined as

$$\mathsf{E}[g(X)] = \sum_{x \in \operatorname{supp}(P_X)} P_X(x)\, g(x), \tag{2}$$

where $\operatorname{supp}(P_X)$ denotes the *support* of P_X and is defined as

$$\operatorname{supp}(P_X) \triangleq \{x \in \mathcal{X} \colon P_X(x) \neq 0\}.$$

The right hand-side of (2) is taken to be $+\infty$ if $\Pr[g(X) = +\infty] > 0$ and $\Pr[g(X) = -\infty] = 0$. It is understood to be $-\infty$ if $\Pr[g(X) = -\infty] > 0$ and $\Pr[g(X) = +\infty] = 0$, and it is undefined if $\Pr[g(X) = +\infty] > 0$ and $\Pr[g(X) = -\infty] > 0$.

a) Let $|\mathcal{X}| = \mathsf{L}$ and $|\operatorname{supp}(P_X)| = \mathsf{L}'$ where $\mathsf{L}' \leq \mathsf{L}$. Here $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$, i.e., the number of elements in $\mathcal{A}$.

   Find $\mathsf{E}\left[\frac{1}{P_X(X)}\right]$.

   Hint: What is $g(x)$?

b) Suppose that X and X' are chance variables taking values in $\mathcal{X}$. Write out the general expression for $\mathsf{E}[P_X(X)]$ and $\mathsf{E}[P_{X'}(X)]$.

c) Write out the general expression for $\mathsf{E}[-\log P_X(X)]$ and $\mathsf{E}[-\log P_{X'}(X)]$.

Note: We define $a + \infty = \infty$ whenever $a \neq -\infty$ and $a \cdot \infty = \operatorname{sgn}(a) \cdot \infty$ whenever $a \neq 0$.

Problem 2 — *On the Expectation of a Discrete Random Variable*

Let the random variable T take on only positive integer values. Show that

$$\mathsf{E}[T] = \sum_{v=1}^{\infty} \Pr[T \geq v].$$

Problem 3 ***Statistical Independence***

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed (IID) binary random variables. Assume $n \geq 2$, and assume

$$\Pr[X_i = 0] = \Pr[X_i = 1] = \frac{1}{2}, \quad i = 1, \ldots, n.$$

Let Z be a parity check on $X_1, \ldots, X_n$, i.e., $Z = X_1 \oplus X_2 \oplus \cdots \oplus X_n$ where $0 \oplus 0 = 1 \oplus 1 = 0$ and $0 \oplus 1 = 1 \oplus 0 = 1$.

a) Is Z statistically independent of X_1?

b) Are $Z, X_1, \ldots, X_{n-1}$ statistically independent?

c) Are $Z, X_1, \ldots, X_n$ statistically independent?

d) Is Z statistically independent of X_1 if $\Pr[X_i = 1] = p \neq \frac{1}{2}$ for all i? You may take $n = 2$ here.

Problem 4 ***Markov's Inequality and Chebyshev's Inequality***

a) *(Markov's Inequality)* For any nonnegative random variable X and any $\delta > 0$, show that

$$\Pr[X \geq \delta] \leq \frac{\mathsf{E}[X]}{\delta}.$$

Exhibit a random variable X and some δ that achieves this inequality with equality.

Hint: In the definition of expectation split the sum/integration into two parts according to whether $x \geq \delta$ or $x < \delta$.

b) *(Chebyshev's Inequality)* Let Y be a random variable with mean μ and variance σ^2. Show that for any $\epsilon > 0$,

$$\Pr[|Y - \mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}.$$

Hint: Let $X = (Y - \mu)^2$ and use a).

c) *(The Weak Law of Large Numbers)* Let $Z_1, Z_2, \ldots, Z_n$ be a sequence of IID random variables with mean μ and variance σ^2. Let $\bar{Z}_n = \frac{1}{n}\sum_{k=1}^{n} Z_k$ be the sample mean. Show that

$$\Pr[|\bar{Z}_n - \mu| \geq \epsilon] \leq \frac{\sigma^2}{n\epsilon^2}.$$

Thus, $\Pr[|\bar{Z}_n - \mu| > \epsilon] \to 0$ as $n \to \infty$. This is known as the weak law of large numbers.

Hint: Use b).