Problem sheet 1, Information Theory, HT 2022

Designed for the first tutorial class

Question 1 We are given a deck of n cards in order $1, 2, \cdots, n$. Then a randomly chosen card is removed and placed at a random position in the deck. What is the entropy of the resulting deck of card?

Question 2 (Polling inequalities) Let $a \geq 0, b \geq 0$ are given with $a + b > 0$. Show that

$$-(a+b)\log(a+b) \leq -a\log(a) - b\log(b) \leq -(a+b)\log(\frac{a+b}{2})$$

and that the first inequality becomes an equality iff $ab = 0$, the second inequality becomes an equality iff $a = b$.

Question 3 Let X, Y, Z be discrete random variables. Prove or provide a counterexample to the following statements:

(a) $H(X) = H(42X)$;

(b) $H(X|Y) \geq H(X|Y, Z)$;

(c) $H(X, Y) = H(X) + H(Y)$.

Question 4 Does there exist a discrete random variable X with a distribution such that $H(X) = +\infty$? If so, describe it as explicitly as possible.

Question 5 Let $\mathcal{X}$ be a finite set, f a real-valued function $f : \mathcal{X} \mapsto \mathbb{R}$ and fix $\alpha \in \mathbb{R}$. We want to maximise the entropy $H(X)$ of a random variable X taking values in $\mathcal{X}$ subject to the constraint

$$\mathbb{E}[f(X)] \leq \alpha. \tag{1}$$

Denote by U a uniformly distributed random variable over $\mathcal{X}$. Prove the following optimal solutions for the maximisation.

(a) If $\alpha \in [\mathbb{E}[f(U)],\ \max_{x\in\mathcal{X}} f(x)]$, then the entropy is maximised subject to (1) by the uniformly distributed random variable U.

(b) If f is non-constant and $\alpha \in [\min_{x\in\mathcal{X}} f(x),\ \mathbb{E}[f(U)]]$, then the entropy is maximised subject to (1) by the random variable Z given by

$$P(Z = x) = \frac{e^{\lambda f(x)}}{\sum_{y\in\mathcal{X}} e^{\lambda f(y)}} \qquad \text{for } x \in \mathcal{X}.$$

where $\lambda < 0$ is chosen such that $\mathbb{E}[f(Z)] = \alpha$.

(c) (Optional) Prove that under the assumptions of (b), the choice for λ is unique and we have $\lambda < 0$.

Question 6 *(A revision on strong law of large numbers (SLLN) in probability theory, please take this question as a reference) Let X be a real-valued random variable.*

(a) Assume additionally that X is non-negative. Show that for every $x > 0$, we have

$$\mathbb{P}(X \geq x) \leq \frac{\mathbb{E}[X]}{x}.$$

(b) Let X be a random variable of mean μ and variance σ^2. Show that

$$\mathbb{P}(|X - \mu| > \varepsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

(c) Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d random variables with mean μ and variance σ^2. Show that $\frac{1}{m}\sum_{n=1}^{m} X_n$ converges to μ in probability, i.,e., for every $\varepsilon > 0$,

$$\lim_{m\to+\infty} \mathbb{P}\left(\left|\frac{1}{m}\sum_{n=1}^{m} X_n - \mu\right| > \epsilon\right) = 0.$$

`This is a weak version of SLLN. It can be strengthen by Borel-Cantelli lemma to the often-used version:` $\mathbb{P}(\lim_{m\to+\infty} \frac{1}{m}\sum_{n=1}^{m} X_n = \mu) = 1$.

Question 7 *(Optional)* Partition the interval $[0,1]$ into n disjoint sub-intervals of length $p_1, \cdots, p_n$. Let $X_1, X_2, \cdots$ be i.i.d. random variables, uniformly distributed on $[0,1]$, and $Z_m(i)$ be the number of the $X_1, \cdots, X_m$ that lie in the i^{th} interval of the partition. Show that the random variables

$$R_m = \Pi_{i=1}^n p_i^{Z_m(i)} \text{ satisfy } \frac{1}{m}\log(R_m) \stackrel{m\to+\infty}{\longrightarrow} \sum_{i=1}^n p_i \log(p_i) \text{ with probability 1.}$$