Práctica 5.5: Entropy, large deviations, and coding

Ejercicio 1. Let $\mu = \{\mu_1, \dots, \mu_n\}$ be a probability measure on $\{1, \dots, n\}$. We call

$$H(\mu) = -\sum_{i=1}^{n} \mu_i \log(\mu_i)$$

its Shannon entropy. Prove that $0 \le H(\mu) < \log(n)$.

Ejercicio 2. Relative entropy Given μ and ν two probability measures on $\{1, \dots, n\}$, we call

$$H(\mu|\nu) = \sum_{i=1}^{n} \mu_i \log(\mu_i/\nu_i)$$

the relative entropy of μ with respect to ν. Prove that $H(\mu|\nu) \ge 0$.

Ejercicio 3. Stirling Prove that

$$ee^{-n}n^n \le n! \le nee^{-n}n^n.$$

Hint: Use that $\log(\lfloor x \rfloor) \le \log(x) \le \log(\lceil x \rceil)$, $x \ge 1$.

Ejercicio 4. Variational representation of relative entropy Prove that

$$\begin{aligned} H(\mu|\nu) &= \sup_f \Big\{ \sum_{i=1}^{n} \mu_i f_i - \log \big(\sum_{i=1}^{n} \nu_i e^{f_i} \big) \Big\} \\ &= \sup_f \big\{ \mathbb{E}_\mu[f] - \log \big(\mathbb{E}_\nu[e^f] \big) \big\}, \end{aligned}$$

where the supremum is over all functions $f : \{1, \dots, n\} \to \mathbb{R}$. Use this variational representation to prove that $H(\mu|\nu) \ge 0$.

Ejercicio 5. Consider the alphabet $\mathcal{A} = \{a, b, c\}$, and two possible codings that associate to each letter in $\mathcal{A}$ a word in the alphabet $\{0, 1\}$,

Code 1	Code 2
$a \to 0$	$a \to 01$
$b \to 010$	$b \to 001$
$c \to 10$	$c \to 10$

Prove that some messages cannot be reconstructed from the coded message when the first code is used, while this is always possible if the second code is applied.

Ejercicio 6. (Laplace principle) Consider $g > 0$ and f continuous functions; prove that

$$\lim_{n \to +\infty} \frac{1}{n} \log \int_a^b g(x) e^{nf(x)} dx = \max_{x \in [a,b]} f(x) \,.$$

Ejercicio 7. (Contraction principle) Let A be a finite set, μ_n a sequence of probability measures on A defined by

$$\mu_n(a) = \frac{e^{-n\alpha(a)}}{\sum_{b \in A} e^{-n\alpha(b)}}$$

and $f : A \to B$ a function from A to another finite set B. Here $\alpha : A \to \mathbb{R}^+$ are positive weights associated to each element of A. Prove that for any $S \subseteq A$ we have

$$\lim_{n\to+\infty} \frac{1}{n} \log \mu_n(S) = -\inf_{a\in S} \bar{\alpha}(a)\,,$$

where $\bar{\alpha}(a) = \alpha(a) - \min_{b\in A} \alpha(b)$. Let ν_n be the sequence of probability measures on B defined by $\nu_n(b) = \sum_{a:f(a)=b} \mu_n(a)$. Prove that for any $W \subseteq B$

$$\lim_{n\to+\infty} \frac{1}{n} \log \nu_n(W) = -\inf_{b\in W} \beta(b)\,,$$

where $\beta(b) = \inf_{a:f(a)=b} \bar{\alpha}(a)$.

Ejercicio 8. (Cramer from Sanov) Let μ, ν probability measures on $\{1, 2, \ldots, n\}$. Prove that

$$\inf_{\nu:\sum_i i\nu_i = x} H(\nu|\mu) = \sup_{\lambda} \left\{ \lambda x - \log \sum_i \mu_i e^{\lambda i} \right\}.$$

Ejercicio 9. Show that for the probability measure on $\mathbb{N}$

$$\mu(n) = \frac{1}{Zn(\log n)^{1+\varepsilon}}$$

where Z is a normalization constant, we have $H(\mu) = +\infty$.

Ejercicio 10. (Fekete's lemma) Let a_n be a real sequence that is subadditive, i.e. such that $a_{n+m} \leq a_n + a_m$; show then that the limit $\lim_{n\to+\infty} \frac{a_n}{n}$ exists and coincides with $\inf_n \frac{a_n}{n}$.

Ejercicio 11. Prove a large deviations principle for the pair empirical measure of samples from an i.i.d. sequence.