GAUSSIAN PROCESSES
EXERCISE SHEET 4: SOME BITS OF ENTROPY

Exercise 1 (Two entropic inequalities). *Let X, Y be i.i.d. discrete random variables with entropy $H(X)$.*

- *Prove that $\mathbb{P}(X = Y) \geq 2^{-H(X)}$.*
- *Prove that $H(X+Y) \geq H(X)$. When does equality hold? Does the inequality necessarily hold when X, Y are not independent?*

Exercise 2 (Differential entropy can be infinite). *Find an example of an $\mathbb{R}$-valued random variable X with density such that $H(X) = -\infty$ and such that $H(X) = \infty$.*

Exercise 3 (A maximum entropy distribution). *Prove that the entropy of any probability distribution on $[0, \infty)$ with density and mean $m > 0$ is less than $\log_2 e + \log_2 m$. When does the equality hold, i.e. what is the maximum entropy distribution on $[0, \infty)$ of mean m?*

Exercise 4 (Entropy lower bound for optimal coding). *Prove that if $W_1, \dots, W_n$ are prefix free words on 0-s and 1-s, encoding the outcomes of a random variable X on an n-element space, then $H(X) \leq \mathbb{E}|W_X|$. You may need the following so-called Kraft inequality: for any set of permitted codewords $W_1, \dots, W_n$, we have that $\sum_{i=1}^{n} 2^{-|W_i|} \leq 1$.*

Exercise 5 (Shannon's characterization of entropy (Not Examinable)). *For each $n \in \mathbb{N}$, and probabilities $p_1, \dots, p_n$ on an n-state probability space, we consider functions $\tilde{H}_n(p_1, p_2, \dots, p_n)$ with the following properties:*

(1) For each fixed $n \in \mathbb{N}$, $\tilde{H}_n$ is continuous with respect to $(p_1, \dots, p_n)$.

(2) The entropy of the uniform distribution grows with n: for $n \geq m$,

$$\tilde{H}_n\left(\tfrac{1}{n}, \dots, \tfrac{1}{n}\right) \geq \tilde{H}_m\left(\tfrac{1}{m}, \dots, \tfrac{1}{m}\right).$$

(3) If some outcome is decomposed into two possibilities, the entropy is a weighted sum; that is, for any $n \in \mathbb{N}$ and for all probabilities $(p_1, \dots, p_{n+1})$, we have

$$\tilde{H}_{n+1}(p_1, \dots, p_n, p_{n+1}) = \tilde{H}_n(p_1, \dots, p_n + p_{n+1}) + (p_n + p_{n+1})\, \tilde{H}_2\left(\frac{p_n}{p_n + p_{n+1}}, \frac{p_{n+1}}{p_n + p_{n+1}}\right).$$

Prove that, up to a multiplicative constant, $\tilde{H}_n(p_1, \dots, p_n)$ agrees with the entropy $H(X)$ of a random variable X on $\{1, \dots, n\}$ that takes value i with probability p_i.
[Spoiler: Appendix 2 in Shannon's paper.]