# GAUSSIAN PROCESSES
# EXERCISE SHEET 4: SOME BITS OF ENTROPY

**Exercise 1.**

Let $p(\omega) = \mathbb{P}(X = \omega)$ for all $\omega \in \Omega$. We have

$$\mathbb{P}(X = Y) = \sum_{\omega \in \Omega} \mathbb{P}(X = Y = \omega) = \sum_{\omega \in \Omega} p(\omega)^2 = 2^{\log_2 \mathbb{E} p(X)}$$

and the first claim follows as $\log_2 \mathbb{E} p(X) \geq \mathbb{E} \log_2 p(X)$ by Jensen's inequality.

For the second claim notice first that

$$\begin{aligned} H(X+Y|Y) &= \sum_y \mathbb{P}[Y = y](-\sum_z \mathbb{P}[X+Y = z|Y = y] \log_2 \mathbb{P}[X+Y = z|Y = y]) \\ &= \sum_y \mathbb{P}[Y = y](-\sum_z \mathbb{P}[X = z - y] \log_2 \mathbb{P}[X = z - y]) \\ &= \sum_y \mathbb{P}[Y = y](-\sum_x \mathbb{P}[X = x] \log_2 \mathbb{P}[X = x]) = H(X). \end{aligned}$$

On the other hand from Sheet 3 exercise 2, we see that

$$H(X+Y|Y) \leq H(X+Y)$$

with equality if and only if $Y$ and $X+Y$ are independent. This on the other hand implies that $X$ must be a constant almost surely, as proven in the lemma below. The inequality does not hold in general if $X$ and $Y$ are not independent, take $X$ to be $\pm 1$ with equal probability and let $Y = 1 - X$ so that $X \sim Y$. Then $H(X+Y) = H(0) = 0 \leq H(X) = 1$.

**Lemma 1.** *Let $X$ and $Y$ be i.i.d. random variables such that $Y \perp X+Y$. Then $X$ is a.s. a constant.*

*Proof.* Let $\varphi$ be the characteristic function of $X$. Then we have for any $a, b \in \mathbb{R}$ that

$$\varphi(a+b)\varphi(b) = \mathbb{E}e^{i(a+b)Y+ibX} = \mathbb{E}e^{iaY+ib(X+Y)} = \varphi(a)\mathbb{E}e^{ib(X+Y)} = \varphi(a)\varphi(b)^2.$$

If there exists $x$ such that $\varphi(x) = 0$, then by setting $a = b = \frac{x}{2}$ in the above inequality we get

$$\varphi(\frac{x}{2})^3 = \varphi(x)\varphi(x/2) = 0,$$

so that also $\varphi(x/2) = 0$. By induction $\varphi(x/2^n) = 0$, but this is a contradiction since $\varphi$ is continuous at 0 and $\varphi(0) = 1$. Thus $\varphi(x) > 0$ for all $x$ and we may divide by $\varphi(b)$ in the above equation to get

$$\varphi(a+b) = \varphi(a)\varphi(b)$$

for all $a, b$. Let then $b = -a$. We get $\varphi(a)\varphi(-a) = \varphi(0) = 1$. As $|\varphi(a)| \leq 1$ for all $a$, this implies that $|\varphi(a)| = 1$ for all $a$. But notice that having $|\varphi(a)| = 1$ for a given $a$ is only possible if in the triangle inequality

$$|\varphi(a)| = |\mathbb{E}e^{iaX}| \leq \mathbb{E}|e^{iaX}| = 1$$

the angle of $e^{iaX}$ is constant almost surely. This means that if $a \neq 0$, then $X$ must only take values in a set of the form $A_a := \{x_0 + k\frac{2\pi}{a} : k \in \mathbb{Z}\}$, where $x_0$ is some fixed point in the support of $X$. Now this holds for all $a$, so $X$ can only take values in $\bigcap_{a \neq 0} A_a$. Let us consider the intersection of $A_1$ and $A_a$. Assume we have $x_0 + 2\pi k = x_0 + l\frac{2\pi}{a}$ for some $k, l \neq 0$. This implies that $l/k = b$, so $b$ must be rational. Thus we see that if one chooses $b$ irrational then $x_0 + 2\pi k \notin A_b$ for any $k \neq 0$, and thus $\bigcap_{a \neq 0} A_a = \{x_0\}$. $\square$

$\square$

**Exercise 2**.

The basic idea is that the more concentrated we are, the lower the entropy. Let us therefore try to put a blow-up of the density at the origin. Take $p(x) = \frac{c}{x \log_2(x)^2}$ on $[0, 1/2]$ and 0 elsewhere, where $c > 0$ is a normalizing constant ensuring that $\int p(x)\, dx = 1$. Then

$$\begin{aligned} H(X) &= -\int_0^{1/2} \frac{c}{x \log_2(x)^2} \log_2\left(\frac{c}{x \log_2(x)^2}\right) dx \\ &= -\int_0^{1/2} c\frac{\log_2 c - 2\log_2(-\log_2(x))}{x \log_2(x)^2}\, dx + \int_0^{1/2} \frac{c}{x \log_2(x)}\, dx = -\infty. \end{aligned}$$

(Note that $\log(\log(x))' = \frac{1}{x \log x}$.)

Conversely, to make the distribution more spread, let us then do the same thing with $p(x) = \frac{c'}{x \log_2(x)^2}$ on $[2, \infty)$ which piles up mass at the infinity (actually $c' = c$ since the two integrals are related by the change of variables $x \mapsto 1/x$). We have

$$H(X) = -\int_2^{\infty} c' \frac{\log_2 c' - 2\log_2 \log_2(x)}{x \log_2(x)^2}\, dx + \int_2^{\infty} \frac{c'}{x \log_2(x)}\, dx = \infty.$$

$\square$

**Exercise 3**.

Let $X \geq 0$ be a random variable whose probability distribution has a density $p(x)$ on $[0, \infty)$ such that

$$\mathbb{E}X = \int x p(x)\, dx = m.$$

We will show that the entropy is maximised by the exponential distribution which has density

$$q(x) := \frac{1}{m} \exp\left(-\frac{x}{m}\right).$$

We have

$$H(X) = -\int_0^{\infty} p(x) \log_2(p(x))\, dx.$$

We would like to use Jensen's inequality, but this has to be done in such a way that equality will be attained by $p = q$. Let us therefore write

$$\begin{aligned} H(X) &= \int_0^{\infty} p(x) \log_2(q(x)/p(x))\, dx - \int_0^{\infty} p(x) \log_2(q(x))\, dx \\ &\leq \log_2\left(\int_0^{\infty} q(x)\, dx\right) - \int_0^{\infty} p(x)\left(-\log_2(m) - \frac{x}{m}\right) dx \\ &= \frac{1 + \log(m)}{\log(2)}. \end{aligned}$$

In order to have equality in the inequality we need to have $q(x) = p(x)$. $\square$

**Exercise 4**.

We have by Kraft's inequality that $\sum_i 2^{-|W_i|} \leq 1$, which we may also write as

$$\mathbb{E}\frac{2^{-|W_X|}}{p(X)} \leq 1.$$

Now by Jensen's inequality we have

$$1 \geq \mathbb{E}\frac{2^{-|W_X|}}{p(X)} = \mathbb{E}2^{-|W_X|-\log_2 p(X)} \geq 2^{-\mathbb{E}|W_X|+H(X)},$$

from which the claim follows. $\square$