GAUSSIAN PROCESSES
EXERCISE SHEET 5: ENTROPY AND CLT

Exercise 1 (Two identities)**.** *The Fisher information is translation invariant: $J(X+a) = J(X)$, and scales as $J(aX) = a^{-2}J(X)$.*

Exercise 2 (De Bruijn's identity)**.** *Let X be a random variable with finite variance and differentiable density. Let Y be a standard Gaussian.*

(1) Denote $Z_t = X + \sqrt{t}Y$. Prove that $\frac{d}{dt}H(Z_t) = \frac{1}{2}J(Z_t)$. (Hint: Prove that the density p_t of Z_t satisfies $\frac{\partial}{\partial t}p_t(x) = \frac{1}{2}p_t''(x)$.)

(2) Denote $Z_t = e^{-t}X + \sqrt{1-e^{-2t}}Y$. Prove that $\frac{d}{dt}H(Z_t) = J(Z_t) - 1$. (Hint: Prove that the density p_t of Z_t satisfies $\frac{\partial}{\partial t}p_t(x) = p_t''(x) + (xp_t(x))'$.)

Here $H(X) := \mathbb{E}\log p(X)$.

Exercise 3 (Smoothed variables are nice)**.** *Let X be any $\mathbb{R}$-valued random variable of finite variance. Let Y be some independent r.v.*

- *Suppose that Y has a smooth density, prove that $X+Y$ has a smooth density.*
- *Suppose that Y has an everywhere positive density, prove that $X+Y$ has a positive density.*

In particular, deduce that if Y is a Gaussian then $X+Y$ has continuously differentiable density that is uniformly bounded away from zero on any compact interval, has an uniformly bounded derivative, finite Fisher information and moreover its score-function has a uniformly bounded 4-th moment.

Exercise 4 (Cramér-Rao lower bound)**.** *Let $p(x,\theta)$ be a family of probability distributions, indexed by a real parameter θ. Define the score function $\rho_\theta(x)$ by*

$$\rho_\theta(x) := \frac{\partial p(x,\theta)}{\partial\theta}\frac{1}{p(x,\theta)},$$

and suppose that it is well-defined. The Fisher information is then defined by

$$J(\theta) = \mathbb{E}_\theta\rho_\theta^2(X),$$

where $\mathbb{E}_\theta$ means that we take the expectation with respect to the density $p(\cdot,\theta)$.

Let now T be an unbiased estimator of θ, based on X, i.e. T is a measurable function of X such that $\mathbb{E}_\theta T(X) = \theta$. Prove that

$$\mathbb{E}_\theta\left[(T(X)-\theta)^2\right] \geq J(\theta)^{-1}.$$

How can you now interpret the fact that Gaussian distribution has minimal Fisher information for fixed-variance distributions?