Assignment 1. Prove that the family of univariate normal distributions $\mathcal{N}(\mu, \sigma)$ with density
$$p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
is an exponential family
$$p_\eta(x) = \exp[\langle \phi(x), \eta \rangle - A(\eta)]$$
with sufficient statistics $\phi(x) = \begin{bmatrix} x \\ x^2 \end{bmatrix}$. Deduce a formula expressing the natural parameter vector $\eta$ in terms of $\mu$ and $\sigma$.

Assignment 2. The Kullback-Leibler divergence for probability densities $p(x)$ and $q(x)$ is defined by
$$D_{KL}(p \parallel q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} \, dx.$$Compute the KL-divergence for two univariate normal distributions $\mathcal{N}(\mu, \sigma)$ and $\mathcal{N}(\tilde{\mu}, \tilde{\sigma})$.

Assignment 3. The probability density function of a Laplace distribution (aka double exponential distribution) with location parameter $\mu$ and scale $b$ is given by
$$p(x \mid \mu, b) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right).$$Find the maximum likelihood estimates of the location parameter and the scale parameter given an i.i.d. sample $T^n = \{x_i \in \mathbb{R} \mid i = 1, \ldots, m\}$.

Assignment 4. Consider the cumulant function $A(\eta)$ of an exponential family
$$p_\eta(x) = \exp[\phi(x)\eta - A(\eta)]$$for a discrete random variable $x \in \mathcal{X}$. It is defined by
$$A(\eta) = \log \sum_{x \in \mathcal{X}} \exp[\phi(x)\eta].$$Notice that we consider for simplicity that $\phi(x)$ and $\eta$ are scalars.

a) Prove that its first derivative is given by
$$\frac{d}{d\eta} A(\eta) = \mathbb{E}_{x \sim p_\eta}[\phi(x)].$$

b) Prove that its second derivative is non-negative and conclude that $A(\eta)$ is a convex function.
Assignment 5. Consider the family of univariate normal distributions $\mathcal{N}(\mu, 1)$ with density

$$p(x; \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}.$$ 

Compute its Fisher information $I(\mu)$. Suppose you want to estimate the mean $\mu$ from a sample $T^m$. What sample size $m$ is needed (asymptotically) to ensure that the estimated mean will be in the $\epsilon$ interval around the true mean with probability 99%?

Assignment 6. Prove the equality $\mathbb{E}_\theta \left[ \frac{d}{d\theta} \log p_\theta(x) \right] = 0$ and conclude that the Fisher information is the variance $I(\theta) = \nabla_\theta \left[ \frac{d}{d\theta} \log p_\theta(x) \right]$ of the random variable $\frac{d}{d\theta} \log p_\theta(x)$. 