Assignment 1. Consider the family of Bernoulli distributions for a binary random variable \( x = 0, 1 \) given by

\[
p(x; \beta) = \beta^x (1 - \beta)^{1-x},
\]

where \( \beta \in (0, 1) \) is the parameter. Prove that it is an exponential family. Give the sufficient statistics (Hint: it is one dimensional) and express the natural parameter as a function of \( \beta \). Give the cumulant function as a function of the natural parameter.

Assignment 2. 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

\[
p_\eta(x) = \exp\left[ \langle \phi(x), \eta \rangle - A(\eta) \right]
\]

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 3. 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 4. The probability density function of a Laplace distribution (aka double exponential distribution) with location parameter \( \mu \) and scale \( b \) is given by

\[
p(x \mid \mu, b) = \frac{1}{b} \exp\left( -\frac{|x - \mu|}{b} \right).
\]

Find the maximum likelihood estimates of the location parameter and the scale parameter given an i.i.d. sample \( T^m = \{ x_i \in \mathbb{R} \mid i = 1, \ldots, m \} \).

Assignment 5. Prove that under the regularity assumptions given in the lecture, the Fisher information of a parametric distribution family \( p_\theta(x) \) can be equivalently written as

\[
I(\theta) = \mathbb{E}_{\theta}\left[ \frac{d}{d\theta} \log p_\theta(x) \right] \quad \text{and} \quad I(\theta) = -\mathbb{E}_{\theta}\left[ \frac{d^2}{d\theta^2} \log p_\theta(x) \right].
\]

a) Prove that \( \mathbb{E}_{\theta}\left[ \frac{d}{d\theta} \log p_\theta(x) \right] = 0 \) and conclude the first equivalent definition above.
b) Use integration by parts and the logarithmic “trick” \( \frac{d}{d\theta}p_\theta(x) = p_\theta(x) \frac{d}{d\theta} \log p_\theta(x) \) to prove the second equivalent definition above.

**Assignment 6.** Consider the cumulant function \( A(\eta) \) of an exponential family

\[
p_\eta(x) = \exp\left[\phi(x)\eta - A(\eta)\right]
\]

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

\[
\frac{d}{d\eta} A(\eta) = \mathbb{E}_\eta[\phi(x)].
\]

b) Prove that its second derivative is non-negative and conclude that \( A(\eta) \) is a convex function.