1. Let X be binomially distributed with parameter p so that

$$f(x;p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 < p < 1,$$

where p is unknown.

   (a) Show this is a canonical exponential family with canonical parameter $\theta = \log\{p/(1-p)\}$;

   (b) Find the log-normalizing function $c(\theta)$;

   (c) Find the mean and variance of the canonical statistic by differentiation;

   (d) Find the mean value parameter;

   (e) Write down the maximum likelihood equation and its solution for p.

2. Let $X = (X_1, \ldots, X_n)$ be independent and Poisson distributed with

$$\mathbf{E}(X_j) = \lambda N_j$$

where $\lambda > 0$ is unknown and N_j are known constants. Models of this type arise in risk studies where N_j are the number of individuals at risk in group j, X_j the number of events, e.g. accidents or casualties, and λ is the risk rate.

   (a) Show that the above model defines a canonical exponential family with canonical parameter $\log \lambda$;

   (b) Identify the canonical sufficient statistic and find its mean and variance;

   (c) Find the mean value parameter;

   (d) Find the maximum likelihood estimate $\hat{\lambda}$ of λ;

   (e) Find the Fisher information for λ and the variance of $\hat{\lambda}$.

3. Consider a random sample of size n from a population of individuals, classified according to their genotypes in a biallelic system $X = (X_{AA}, X_{Aa}, X_{aa})$ where X_{AA} denotes the number of individuals with genotype AA etc. The sample then follows a multinomial distribution

$$f(x;p) = \binom{n}{x_{AA}, x_{Aa}, x_{aa}} p_{AA}^{x_{AA}} p_{Aa}^{x_{Aa}} p_{aa}^{x_{aa}}.$$

Under the assumption of so-called *Hardy–Weinberg equilibrium*, the individual probabilities are given as

$$p_{AA} = \mu^2, \quad p_{Aa} = 2\mu(1-\mu), \quad p_{aa} = (1-\mu)^2,$$

where $0 < \mu < 1$.

(a) Show that the assumption of Hardy–Weinberg equilbrium makes this a linear exponential family with canonical sufficient statistic equal to $t(X) = 2X_{AA} + X_{Aa}$;

(b) Find the canonical parameter;

(c) Find the log-normalizing function;

(d) Find the mean value parameter;

(e) Write down the likelihood equation and find the MLE for μ.

4. Let $f(x)$ be any non-negative function on the real line and define

$$f(x;\theta) = f(x)e^{\theta x - c(\theta)}, \theta \in \Theta,$$

where

$$\Theta = \left\{\theta \mid \int f(x)e^{\theta x}\,dx < \infty\right\}.$$

If $\Theta \neq \emptyset$ this is known as the *natural exponential family generated by* f and most distributions are therefore 'members of an exponential family'.

(a) Find Θ in the case where

$$f(x) = \begin{cases} \frac{1}{1+x^4} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(b) Show that the likelihood equation has no solution if $x > 1/\sqrt{2}$.
Hint: Find $\mathbf{E}_0(X)$ by direct integration and use that the mean value mapping is everywhere strictly increasing.

(c) What is the MLE of θ for $x > 1/\sqrt{2}$?

(d) Consider next the family generated by

$$g(x) = \begin{cases} \frac{1}{1+x^2} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

and find also here the canonical parameter space.

(e) Show that the likelihood equation always has a solution in the family generated by $g(x)$.
Hint: Use that the function $x/(1+x^2)$ is not integrable over the entire half-line $(0, \infty)$.