Homework 3

1. For $X \sim f(x;\theta)$, show that

$$E\left\{\frac{\partial}{\partial\theta}\log f(x;\theta)\right\} = 0, \quad \text{and} \quad -E\left\{\frac{\partial^2}{\partial\theta^2}\log f(x;\theta)\right\} = E\left[\left\{\frac{\partial}{\partial\theta}\log f(x;\theta)\right\}^2\right].$$

2. Suppose we have $y_i \overset{\text{ind}}{\sim} Poisson(\mu_i), i = 1, \cdots, n$.

   (a) Show that the Poisson distribution belongs to the exponential dispersion family to identify θ_i, $b(\theta_i)$, $a(\phi)$ under the exponential dispersion family form given in Lecture note.

   (b) Derive the canonical link function $g(\mu_i)$ to be modeled via $\eta_i = \mathbf{x}_i\boldsymbol{\beta}, i = 1, \cdots, n$ for the Poisson regression.

3. Write your own `code` (both in `R` and `Python`) to estimate the parameters $\boldsymbol{\beta}$ in the Poisson regression based on Newton-Raphson method. Namely, we have the following model for $(y_i, \mathbf{x}_i), i = 1, \cdots, n$:

$$y_i \mid \mathbf{x}_i \sim Poisson(\mu_i), \qquad i = 1, \cdots, n$$

   where

$$g(\mu_i) = \eta_i = \mathbf{x}_i\boldsymbol{\beta},$$

   with g being the canonical link function obtained in Problem 2-(b).

4. Apply your code and report your parameter estimates of the Poisson regression to "Smoking and Lung Cancer data" that contains information from a Canadian study of mortality by age and smoking status, downloadable from

   `https://grodri.github.io/glms/datasets/#smoking`

   where you can also find the details about the data.