Question 1

Suppose that the lifetime of a particular lightbulb is known to be distributed as $\text{Exp}(\theta)$, i.e. an exponential distribution with parameter θ. Suppose that a group of n friends bought a multipack containing n of these lightbulbs and each person keeps one lightbulb to use at home. In a few years' time, they get together and share how long their lightbulbs lasted, and these measurements (lifetimes) are written down as $x_1, x_2, \ldots, x_n$.

(a) Write down the probability density function of the $\text{Exp}(\theta)$ distribution.

(b) Given the sample of measurements $x_1, x_2, \ldots, x_n$, write down the likelihood function for θ based on these measurements.

(c) Write down the log-likelihood of θ given the measurements $x_1, x_2, \ldots, x_n$.

(d) Find the maximum likelihood estimate $\widehat{\theta}$.

(e) Are you sure that $\widehat{\theta}$ is a maximum, or could it be a minimum? If you have not already done so in (d), provide proof that $\widehat{\theta}$ is a maximum/minimum.

Question 2 (Knowledge of partial derivatives required)

Suppose the random variables $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ are assumed to be independent and identically distributed as $\text{N}\left(\mu, \sigma^2\right)$, where μ and σ^2 are unknown. Suppose further that $\mathbf{X}$ is observed as $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

(a) Compute the likelihood $L(\mu, \sigma^2|\mathbf{x})$.

(b) Compute the log-likelihood $\log L(\mu, \sigma^2|\mathbf{x})$.

(c) Compute the partial derivative $\frac{\partial}{\partial\mu} \log L(\mu, \sigma^2|\mathbf{x})$.

(d) Set the partial derivative in (c) equal to 0 and solve for μ. Show that this value of μ maximises $L(\mu, \sigma^2|\mathbf{x})$ globally for fixed σ^2.

(e) Compute the partial derivative $\frac{\partial}{\partial\sigma^2} \log L(\mu, \sigma^2|\mathbf{x})$. **Hint**: it may help to set $z = \sigma^2$ and then compute the partial derivative with respect to z.

(f) Set the partial derivative in (e) equal to 0 and solve for σ^2. Show that this value of σ^2 is a (local) maximum for $L(\mu, \sigma^2|\mathbf{x})$, for fixed μ.

(g) Show that the value of σ^2 found in (f) maximises the likelihood $L(\mu, \sigma^2|\mathbf{x})$ globally for fixed μ.

(h) Write down the maximum likelihood estimators for μ and σ^2 given the random variables $\mathbf{X}$.

Question 3 (R question)

The data set `deliverycosts.txt` consists of 20 values in two columns named `weight` and `cost`. The `weight` column represents the weight (in kg) of a parcel and `cost` represents the cost (in British pounds) for delivering a package of that given weight from London to Edinburgh via an unnamed courier company.

Create an R Markdown document to answer the following questions:

(a) Read in the data contained in `deliverycosts.txt`, and plot `cost` vs `weight` using a scatterplot. Does there appear to be a linear relationship between the two variables?

(b) Use the `lm` function to fit a linear model on `weight` and `cost`, where `weight` is the predictor variable and `cost` is the regressor variable. Plot the line showing the estimated linear relationship. **Note:** you may wish to transform or rescale `weight` and/or `cost` before fitting the linear model.

(c) Using any output from the `lm` function as well as any additional plots, briefly discuss how well the data fits the proposed model.

(d) Given the results in Steps 2 and 3, if you think a better linear model can be found, perhaps by modifying the predictor/regressor variables and the data set, do Steps 2 and 3 again for your modified linear model. If you feel there are any unusual data points, provide justification.

(e) Describe the results of your investigation into the relationship between `weight` and `cost`, and describe the relationship (if any) that exists between the two variables. Provide an interpretation of any estimated parameter values for the linear model.