1. The following questions all related to a random variable *y* which is distributed according to the Poisson distribution with density

$$f(y, \theta_o) = \frac{\theta_o^y e^{-\theta_o}}{y!}$$

(a) Write down the log-likelihood function $\ell(\theta, y)$ corresponding to a single observation drawn from a Poisson population.

For a single observation:

$$L(\theta) = f(y, \theta) = \frac{\theta^y e^{-\theta}}{y!}$$

$$\ell(\theta) = \ln[L(\theta)] = \ln(\theta^y) + \ln(e^{-\theta}) - \ln(y!)$$
$$= y\ln(\theta) - \theta - \ln(y!)$$

For the whole sample:

$$L_N(\theta) = \prod_{y_i=0}^{N} \frac{\theta^{y_i} e^{-\theta}}{y_i!}$$

$$\ell(\theta) = \ln \prod_{y_i=0}^{N} \frac{\theta^{y_i} e^{-\theta}}{y_i!}$$
$$= \sum_{y_i=0}^{N} [y_i \ln(\theta) - \theta - \ln(y_i!)]$$

(b) Compute the expected log-likelihood and show that it's optimized at $\theta = \theta_o$ by examining the necessary first order condition

For a single observation

$$E[\ell(\theta)] = \sum_{y=0}^{\infty} f(y, \theta_o)\, \ell(\theta)$$

$$\frac{\partial E[\ell(\theta)]}{\partial \theta} = \sum_{y=0}^{\infty} f(y) \left(\frac{y}{\theta} - 1\right) = \sum_{y=0}^{\infty} f(y) \left(\frac{y-\theta}{\theta}\right) = \frac{1}{\theta} \sum_{y=0}^{\infty} f(y)(y-\theta)$$

$$= \frac{1}{\theta} \sum_{y=0}^{\infty} f(y)\cdot y - \frac{1}{\theta} \sum_{y=0}^{\infty} f(y)\cdot \theta \quad \Rightarrow \sum_{y=0}^{\infty} f(y)\cdot y = E(y) = \theta_o$$ because it is a Poisson Distribution

$$= \frac{\theta_o}{\theta} - \frac{1}{\theta}\cdot\theta \Rightarrow \quad \frac{\theta_o}{\theta} - 1 = 0$$ to optimize

$$\Rightarrow \frac{\theta_o}{\theta} = 1 \quad \Rightarrow \quad \hat{\theta}_o = \theta$$

(c) Show that the optimum from the prior part is a maximum by showing that

$$\left.\frac{\partial^2 E[\ell(\theta, y)]}{\partial \theta^2}\right|_{\theta=\theta_o} < 0$$

since $\frac{\partial E[\ell(\theta)]}{\partial \theta} = \frac{\theta_o}{\theta} - 1$

SOC $[\theta]$ $-\frac{\theta_o}{\theta^2}$ because it is evaluated at $\theta = \theta_o$ $\quad -\frac{\theta_o}{\theta^2} = -\frac{1}{\theta_o} < 0$

In Poisson Distribution, the parameter θ_o, which is also the mean, is strictly positive.

(d) Derive an expression for the empirical expectation.

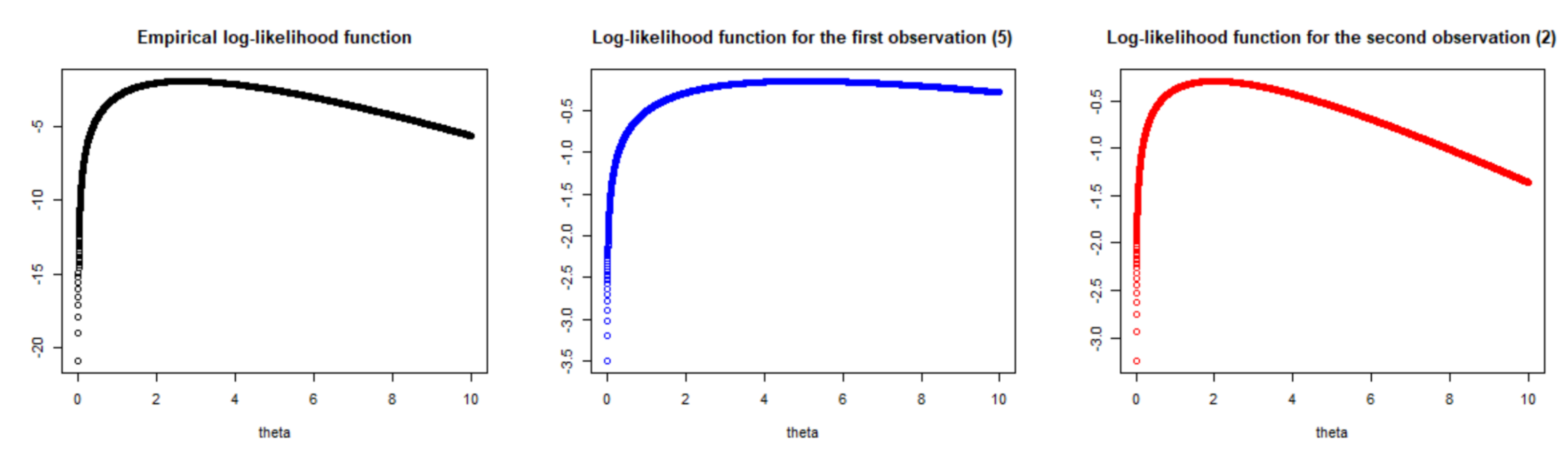
$$\hat{E}_{1000}[\ell(\theta, y)] = \sum_{y=0}^{11} \hat{F}_{1000}(y) \cdot \ell(\theta, y)$$

where $\hat{F}_{1000}(y)$ =

y	0	1	2	3	4	5	6	7	8	9	10	11	
$\hat{F}_{1000}(y)$	.077	.182	.224	.217	.146	.089	.037	.016	.008	.002	.001	.001	

and $\ell(\theta, y) = y\ln(\theta) - \theta - \ln(y!)$

(e) Using the data, plot the expected log-likehood of the sample and likelihood functions corresponding to the first two observations.



(f) Show that $\hat{\theta}_{ML}$ is the sample mean by solving the necessary first order condition.

$$\left.\frac{\partial \hat{E}_N[\ell(\theta,y)]}{\partial\theta}\right|_{\theta=\hat{\theta}_{ML}} = 0$$

Since $\hat{E}_{1000}[\ell(\theta,y)] = \sum_{y=0}^{11} \hat{F}_{1000}(y)\cdot \ell(\theta,y)$ and $\ell(\theta,y) = y\ln(\theta) - \theta - \ln(y!)$

FOC $[\theta]$ $\quad \sum_{y=0}^{11} \hat{F}_{1000}(y)\cdot\left(\frac{y}{\theta}-1\right)$

$$= \frac{1}{\theta}\sum_{y=0}^{11} \hat{F}_{1000}(y)(y-\theta)$$

$$= \frac{1}{\theta}\sum_{y=0}^{11} \hat{F}_{1000}(y)\cdot y - \frac{1}{\theta}\sum_{y=0}^{11} \hat{F}_{1000}(y)\cdot\theta$$

$$= \frac{1}{\theta}\hat{E}_{1000}(y) - 1 \quad \left(\text{because } \sum_{y=0}^{11} F_{1000}(y)\cdot y = \hat{E}_{1000}(y)\right)$$

Set the above equation to zero produces $\frac{\hat{E}_{1000}(y)}{\theta} = 1 \Rightarrow \hat{\theta}_{ML} = \hat{E}_{1000}(y)$

(g) Show that the optimization problem is particularly well-behaved (globally concave) by showing that

$$\left.\frac{\partial^2 \hat{E}_N[\ell(\theta,y)]}{\partial\theta^2}\right|_{\theta=\hat{\theta}_{ML}} < 0$$

Since the first order condition with respect to θ is $\frac{1}{\theta}\hat{E}_{1000}(y) - 1$

SOC $[\theta]$ $\Rightarrow -\frac{1}{\theta^2}\cdot\hat{E}_{1000}(y) = -\frac{1}{\hat{E}_{1000}(y)} = -\frac{1}{2.747} < 0$ ($\hat{E}_{1000}(y) = 2.747$ calculated in R)

(h) Write a function to compute the expected log-likelihood function for the sample and use an optimization routine to solve its maximizer, confirming your results in (f).