Q7

(a) Derive Newton's Method

$$\begin{aligned}
L(\theta|Y) &= \prod_{i=1}^{n} P(y_i|\theta) \\
&= \prod_{i=1}^{n} \frac{(\theta+1)^{y_i} e^{-(\theta+1)}}{y_i!} \\
l(\theta|Y) &= log(\theta+1)\sum_{i=1}^{n} y_i - n(\theta+1) - \sum_{i=1}^{n} log(y_i!)
\end{aligned}$$

First and Second derivative is

$$\begin{aligned}
\frac{\partial l(\theta|Y)}{\partial \theta} &= \frac{\sum_{i=1}^{n} y_i}{\theta+1} - n \\
\frac{\partial^2 l(\theta|Y)}{\partial \theta^2} &= -\frac{\sum_{i=1}^{n} y_i}{(\theta+1)^2}
\end{aligned}$$

So we can optimize θ by iteratation

$$\begin{aligned}
\theta^{(t+1)} &= \theta^{(t)} - \left[\frac{\partial^2 l(\theta^{(t)}|Y)}{\partial {\theta^{(t)}}^2}\right]^{-1} \frac{\partial l(\theta^{(t)}|Y)}{\partial \theta^{(t)}} \\
&= 1 + 2\theta^{(t)} - \frac{n(\theta+1)^2}{\sum_{i=1}^{n} y_i}
\end{aligned}$$

(b) Scoring method

Since $y_i \sim Poisson(\theta+1)$, $E[y_i] = \theta + 1$. Then,

$$\begin{aligned}
I(\theta) = E\left[-\frac{\partial^2 l(\theta|Y)}{\partial \theta^2}\right] &= \frac{\sum_{i=1}^{n} E[y_i]}{(\theta+1)^2} \\
&= \frac{n}{\theta+1}
\end{aligned}$$

So we can optimize θ by iteratation

$$\begin{aligned}
\theta^{(t+1)} &= \theta^{(t)} - \left[I(\theta^{(t)})\right]^{-1} \frac{\partial l(\theta^{(t)}|Y)}{\partial \theta^{(t)}} \\
&= \frac{\sum_{i=1}^{n} y_i}{n} - 1
\end{aligned}$$

(c) Derive EM

Treat s_i's as missing data

$$
\begin{aligned}
L(\theta|S,Y) &= \prod_{i=1}^{n} P(S,Y|\theta) \\
&= \prod_{i=1}^{n} P(Y|S,\theta)P(S|\theta) \\
&= \prod_{i=1}^{n} \frac{e^{-1}}{(y_i - s_i)!}\frac{\theta^{s_i}e^{-\theta}}{s_i!} \\
&= e^{-n(1+\theta)} \cdot \theta^{\sum_{i=1}^{n} s_i} \cdot \prod_{i=1}^{n} \frac{1}{(y_i - s_i)! \cdot s_i!}
\end{aligned}
$$

Then Log likelihood is

$$
l(\theta|S,Y) = -n(1+\theta) + \sum_{i=1}^{n} s_i log(\theta)
$$

0.1 E-step

$$
\begin{aligned}
Q(\theta|\theta^{(t)}) &= E[l(\theta|Y_{com})|Y_{obs},\theta^{(t)}] \\
&= -n(1+\theta) + \sum_{i=1}^{n} E[s_i|Y_{obs},\theta^{(t)}]log(\theta)
\end{aligned}
$$

PMF of $s_i|Y_{obs},\theta^{(t)}$ is

$$
\begin{aligned}
P(s_i|Y_{obs}=y_i,\theta^{(t)}) &= \frac{P(s_i,y_i|\theta^{(t)})}{P(y_i|\theta^{(t)})} \\
&= \frac{{\theta^{(t)}}^{s_i}e^{-\theta^{(t)}}}{s_i!} \cdot \frac{e^{-1}}{(y_i - s_i)!} \cdot \left[\frac{(\theta^{(t)}+1)^{y_i}e^{-(\theta^{(t)}+1)}}{y_i!}\right]^{-1} \\
&= \frac{y_i!}{(y_i - s_i)!s_i!}\left(\frac{\theta^{(t)}}{\theta^{(t)}+1}\right)^{s_i}\left(\frac{1}{\theta^{(t)}+1}\right)^{y_i - s_i}
\end{aligned}
$$

Thus $s_i|Y_{obs}=y_i,\theta^{(t)} \sim Binomial\left(y_i, \frac{\theta^{(t)}}{\theta^{(t)}+1}\right)$, Then $E[s_i|Y_{obs},\theta^{(t)}] = y_i \cdot \frac{\theta^{(t)}}{\theta^{(t)}+1}$

0.2 M-step

$$
Q(\theta|\theta^{(t)}) = -n(1+\theta) + log\theta \cdot \sum_{i=1}^{n} y_i \cdot \frac{\theta^{(t)}}{\theta^{(t)}+1}
$$

Maximize the Qfunction by letting the first derivative 0

$$\begin{aligned}
\frac{\partial Q(\theta|\theta^{(t)})}{\partial \theta} &=^{let} 0 \\
&= -n + \frac{\sum_{i=1}^{n} y_i \cdot \frac{\theta^{(t)}}{\theta^{(t)}+1}}{\theta} \\
\theta^{(t+1)} &= \frac{\sum_{i=1}^{n} y_i}{n} \cdot \frac{\theta^{(t)}}{\theta^{(t)}+1}
\end{aligned}$$

We can get the $\hat{\theta}$ by iterate the E-step and M-step

(d) EM algorithm has unique solution

EM algorithm converges when $\theta^{(t+1)} = \theta^{(t)}$

$$\begin{aligned}
\hat{\theta} &= \frac{\sum_{i=1}^{n} y_i}{n} \cdot \frac{\hat{\theta}}{\hat{\theta}+1} \\
\hat{\theta} &= \frac{\sum_{i=1}^{n} y_i}{n} - 1
\end{aligned}$$

which does not depend on $\theta^{(t)}$. it means that regardless inital value EM-algorithm converge to unique solution

(e) find the convergence rate of the EM

as $M(\theta) = \frac{\sum_{i=1}^{n} y_i}{n} \cdot \frac{\theta}{\theta+1}$

$$\begin{aligned}
DM(\theta) &= \frac{\partial M(\theta)}{\partial \theta} \\
&= \frac{\sum_{i=1}^{n} y_i}{n \cdot (\theta+1)^2}
\end{aligned}$$

(f) Compute $V_{obs}(\hat{\theta})$ using Louis' method

from (c) $s_i|Y_{obs} = y_i, \theta^{(t)} \sim Binomial\left(y_i, \frac{\theta^{(t)}}{\theta^{(t)}+1}\right)$. Then,

$$\begin{aligned}
E[s_i|Y_{obs}, \theta] &= y_i \cdot \frac{\theta}{\theta+1} \\
V[s_i|Y_{obs}, \theta] &= y_i \cdot \frac{\theta}{(\theta+1)^2}
\end{aligned}$$

First compute the $I_{com} = E[-l''(\theta|Y_{com})|Y_{obs}, \theta]$

$$
\begin{aligned}
E[-l''(\theta|Y_{com})|Y_{obs}, \theta] &= \frac{\partial^2 (n(1+\theta) - log\theta \sum_{i=1}^{n} E[s_i|Y_{obs}, \theta]))}{\partial \theta^2} \\
&= \frac{1}{\theta^2} \sum_{i=1}^{n} y_i \cdot \frac{\theta}{\theta+1} \\
&= \frac{1}{\theta(\theta+1)} \sum_{i=1}^{n} y_i
\end{aligned}
$$

Second compute $Var(l'(\theta|Y_{com})|Y_{obs}, \theta)$

$$
\begin{aligned}
Var(l'(\theta|Y_{com})|Y_{obs}, \theta) &= Var\left(-n + \frac{1}{\theta} \sum_{i=1}^{n} s_i|Y_{obs}, \theta\right) \\
&= \frac{1}{\theta^2} \sum_{i=1}^{n} Var(s_i|Y_{obs}, \theta) \\
&= \frac{1}{\theta^2} \sum_{i=1}^{n} y_i \cdot \frac{\theta}{(\theta+1)^2} \\
&= \frac{1}{\theta(\theta+1)} \sum_{i=1}^{n} y_i
\end{aligned}
$$

as $I_{obs} = I_{com} - Var(l'(\theta|Y_{com})|Y_{obs}, \theta)$

$$
\begin{aligned}
I_{obs} = I_{com} - Var(l'(\theta|Y_{com})|Y_{obs}, \theta) &= \frac{1}{\theta(\theta+1)} \sum_{i=1}^{n} y_i - \frac{1}{\theta(\theta+1)} \sum_{i=1}^{n} y_i \\
&= \frac{1}{(\theta+1)^2} \sum_{i=1}^{n} y_i
\end{aligned}
$$

Thus,

$$
V_{obs} = I_{obs}^{-1} = \frac{(\theta+1)^2}{\sum_{i=1}^{n} y_i}
$$

Then,

$$
V_{obs}(\hat{\theta}) = \frac{(\hat{\theta}+1)^2}{\sum_{i=1}^{n} y_i}
$$

where $\hat{\theta} = \frac{\sum_{i=1}^{n} y_i}{n} - 1$

$$
\begin{aligned}
V_{obs}(\hat{\theta}) &= \frac{(\hat{\theta}+1)^2}{\sum_{i=1}^{n} y_i} \\
&= \frac{1}{n^2} \sum_{i=1}^{n} y_i
\end{aligned}
$$

(f) Compute $V_{obs}(\hat{\theta})$ using SEM

$V_{obs}(\hat{\theta}) = \Big[(1 - DM(\hat{\theta}))I_{com}\Big]^{-1}$ from (e) and (f) we already computed $DM(\hat{\theta})$ and I_{com}

$$\begin{aligned}
V_{obs}(\hat{\theta}) &= \Big[(1 - DM(\hat{\theta}))I_{com}\Big]^{-1} \\
&= \left[\left(1 - \frac{\sum_{i=1}^{n} y_i}{n(1+\hat{\theta})^2}\right) \frac{1}{\hat{\theta}(\hat{\theta}+1)} \sum_{i=1}^{n} y_i\right]^{-1}
\end{aligned}$$

where $\hat{\theta} = \frac{\sum_{i=1}^{n} y_i}{n} - 1$. Then,

$$\begin{aligned}
V_{obs}(\hat{\theta}) &= \left(1 - \frac{n}{\sum_{i=1}^{n} y_i}\right) \frac{n}{\frac{\sum_{i=1}^{n} y_i}{n} - 1} \\
&= \left(\frac{\sum_{i=1}^{n} y_i - n}{\sum_{i=1}^{n} y_i}\right) \frac{n^2}{\sum_{i=1}^{n} y_i - n} \\
&= \frac{n^2}{\sum_{i=1}^{n} y_i}
\end{aligned}$$

PMF of $s_i|Y_{obs}, \Theta^{(t)}$ is

$$\begin{aligned}
P(s_i|Y_{obs} = y_i, \Theta^{(t)}) &= \frac{P(s_i, y_i|\Theta^{(t)})}{P(y_i|\Theta^{(t)})} \\
&= \frac{{\theta^{*(t)}}^{s_i} e^{-\theta^{*(t)}}}{s_i!} \cdot \frac{(1-\alpha)^{y_i - s_i} e^{-(1-\alpha)}}{(y_i - s_i)!} \cdot \left[\frac{(\theta^{*(t)} + 1 - \alpha)^{y_i} e^{-(\theta^{*(t)} + 1 - \alpha)}}{y_i!}\right]^{-1} \\
&= \frac{y_i!}{(y_i - s_i)! s_i!} \left(\frac{\theta^{*(t)}}{\theta^{*(t)} + 1 - \alpha}\right)^{s_i} \left(\frac{1}{\theta^{*(t)} + 1 - \alpha}\right)^{y_i - s_i}
\end{aligned}$$