1. (a) Let $Y \sim \text{Pois}(\lambda)$. We look to find $\lambda = \lambda_U$ such that

$$\mathbb{P}(Y \leq y) = \mathbb{P}(\chi^2_{2y+2} > 2\lambda_U) = \alpha$$

By evaluating this integral, we can find a compact expression that λ_U satisfies.

$$\begin{aligned}
\alpha &\overset{set}{=} \mathbb{P}(\chi^2_{2y+2} > 2\lambda_U)\\
&= \int_{2\lambda_U}^{\infty} \frac{1}{2^{y+1}\Gamma(y+1)} x^y e^{-x/2}dx\\
&= \frac{1}{2\Gamma(y+1)} \int_{2\lambda_U}^{\infty} \left(\frac{x}{2}\right)^y e^{-x/2}dx\\
&= \frac{1}{\Gamma(y+1)} \int_{\lambda_U}^{\infty} w^{(y+1)-1}e^{-w}dw\\
&= \frac{\Gamma(y+1,\lambda_U)}{\Gamma(y+1)}
\end{aligned}$$

Here we used the substitution $w = x/2$ and the *incomplete upper gamma function* denoted by $\Gamma(\cdot,\cdot)$. Therefore, we see that the right bound for the $(1-\alpha)\%$ confidence interval, λ_U satisfies

$$\Gamma(y+1,\lambda_U) = \alpha\Gamma(y+1)$$

(b) Assuming that we observe $Y = 0$, we can calculate λ_U as follows

$$\begin{aligned}
\Gamma(0+1,\lambda_U) &= \alpha\Gamma(0+1)\\
\int_{\lambda_U}^{\infty} w^{(0+1)-1}e^{-w}dw &= \alpha\\
\int_{\lambda_U}^{\infty} e^{-w}dw &= \alpha\\
-e^{-w}\Big|_{\lambda_U}^{\infty} &= \alpha\\
e^{-\lambda_U} &= \alpha\\
\lambda_U &= -\log(\alpha)
\end{aligned}$$

Hence for a $\alpha = 0.05$ confidence region, we have $\lambda_U = -\log(0.05) \approx 2.995732$ and the interval is given by $(0, 2.995732)$.

(c) Assuming that $\lambda = te^{\beta_0}$ for some known $t \in \mathbb{R}$ and parameter β_0, we can build a confidence interval for β_0 when we observe $Y = 0$. By the same logic as above we see that

$$\begin{aligned}
\Gamma(0+1, te^{\beta_U}) &= \alpha\Gamma(0+1)\\
&\vdots\\
\exp\{-te^{\beta_U}\} &= \alpha\\
\beta_U &= \log\left(-\frac{\log(\alpha)}{t}\right)
\end{aligned}$$

For $\alpha = 0.05$ we have $\beta_U = \log\left(\frac{\lambda_U}{t}\right) = \log\left(\frac{2.995732}{t}\right)$ and our interval is given by $\left(-\infty, \log\left(\frac{0.05129}{t}\right)\right)$

2. (a) First let $\lambda_i = \exp\{\beta_0 - cy_{i-1}\}$. Then using a conditional argument along with the Markovian properties of this process, we can write the joint likelihood function as follows

$$\begin{aligned}
\mathcal{L}(\beta_0; Y_1, Y_2, \dots, Y_n) &= \mathbb{P}(Y_1 = y_1 | Y_0 = 0)\prod_{i=2}^{n}\mathbb{P}(Y_i = y_i | Y_{i-1} = y_{i-1}, \dots, Y_1 = y_1)\\
&= \frac{\lambda_1^{y_1}e^{-\lambda_1}}{y_1!}\prod_{i=2}^{n}\mathbb{P}(Y_i = y_i | Y_{i-1} = y_{i-1})\\
&= \prod_{i=1}^{n}\frac{\lambda_i^{y_i}e^{-\lambda_i}}{y_i!}
\end{aligned}$$

Using this expression, we can write the log-likelihood function as follows

$$\begin{aligned}
\ell(\beta_0; Y_1, \dots, Y_n) &= \log(\mathcal{L}(\beta_0; Y_1, \dots, Y_n))\\
&= \log\left(\prod_{i=1}^{n}\frac{\lambda_i^{y_i}e^{-\lambda_i}}{y_i!}\right)\\
&= \sum_{i=1}^{n} y_i\log(\lambda_i) - \sum_{i=1}^{n}\lambda_i - \sum_{i=1}^{n}\log(y_i!)\\
&= \sum_{i=1}^{n} y_i(\beta_0 - cy_{i-1}) - \sum_{i=1}^{n}\lambda_i - \sum_{i=1}^{n}\log(y_i!)\\
&= \beta_0\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\lambda_i - c\sum_{i=1}^{n} y_i y_{i-1} - \sum_{i=1}^{n}\log(y_i!)
\end{aligned}$$

Using this expression, we can find the maximum likelihood estimate for β_0. First notice that $\frac{d}{d\beta_0}\lambda_i = \lambda_i$ then we can write the following

$$\begin{aligned}
0 \overset{set}{=} \frac{d}{d\beta_0}\ell(\beta_0; Y_1, \dots, y_n) &= n\overline{y} - \sum_{i=1}^{n}\lambda_i\\
\sum_{i=1}^{n}\exp\{\hat{\beta}_0 - cy_i\} &= n\overline{y}\\
\exp\{\hat{\beta}_0\}\sum_{i=1}^{n}\exp\{-cy_i\} &= n\overline{y}\\
\hat{\beta}_0 = \log\left(\frac{n\overline{y}}{\sum_{i=1}^{n}\exp\{-cy_i\}}\right)
\end{aligned}$$

(b) At $c = 0$, (i.e. observations are independent) we have that

$$\hat{\beta}_0 = \log\left(\frac{n\overline{y}}{\sum_{i=1}^n \exp\{-cy_i\}}\right)\Big|_{c=0} = \log(\overline{y})$$

This matches our intuition as the maximum likelihood estimator of the Poisson rate parameter for an iid sample is given by the sample mean. In our case, with $c = 0$ this rate is $\lambda = e^{\beta_0}$. Due to invariance of MLE's, we expect our estimator of β_0 to be $\hat{\beta}_{0,ML} = \log(\hat{\lambda}_{ML}) = \log(\overline{y})$.
As $c \to \infty$, we see that $\exp\{-cy_i\} \to 0$ and $\frac{n\overline{y}}{\sum_{i=1}^n \exp\{-cy_i\}} \to \infty$ so $\hat{\beta}_0 \to \infty$. As we require $Y_i|Y_{i=1} = y$ to be a RV, as $c \to \infty$, the rate here goes to zero but recall that $\lambda > 0$ for this to be a valid density. So by necessity $\beta_0 \to \infty$ to ensure this rate does not go to 0.

(c) Assume there exists $N \in \mathbb{N}$ such that for all $i \geq N$ we have $Y_i \overset{D}{=} Y_{i-1}$. Using this, we can find the marginal distribution as follows

$$\begin{aligned}
\mathbb{P}(Y_i = k) &= \sum_{n=0}^{\infty} \mathbb{P}(Y_i = k|Y_{i-1} = n)\mathbb{P}(Y_{i-1} = n)\\
&= \mathbb{P}(Y_i = k|Y_{i-1} = 0)\mathbb{P}(Y_i = 0) + \sum_{n=1}^{\infty} \mathbb{P}(Y_i = k|Y_{i-1} = n)\mathbb{P}(Y_i = n)
\end{aligned}$$

Now notice as $c \to \infty$ $e^{\beta_0 - cy} \to 0$. Therefore, the Poisson rate is *zero* and $\mathbb{P}(Y_i = k|Y_{i-1} = n) = 0$ for $n \geq 1$ and $k \geq 1$. Moreover $\mathbb{P}(Y_i = 0|Y_{i-1} = n) = 1$ for $n \geq 1$. Hence, for $\lambda = e^{\beta_0}$ and $k \geq 1$ we have

$$\mathbb{P}(Y_i = k) = \frac{e^{-\lambda}\lambda^k}{k!}\mathbb{P}(Y_i = 0)$$

Therefore, we see that the marginal distribution of Y_i can be characterized relative to a base probability $\mathbb{P}(Y_i = 0)$. Now, first note that $\mathbb{P}(Y_i = 0|Y_{i-1} = n) = 1$ for $n \geq 1$. Using this fact we can write

$$\begin{aligned}
\mathbb{P}(Y_i = 0) &= \mathbb{P}(Y_i = 0|Y_{i-1} = 0)\mathbb{P}(Y_i = 0) + \sum_{n=1}^{\infty} \mathbb{P}(Y_i = 0|Y_{i-1} = n)\mathbb{P}(Y_{i-1} = n)\\
&= e^{-\lambda}\mathbb{P}(Y_i = 0) + \sum_{n=1}^{\infty} \mathbb{P}(Y_{i-1} = n)\\
&= e^{-\lambda}\mathbb{P}(Y_i = 0) + \sum_{n=1}^{\infty} \mathbb{P}(Y_i = n)\\
&= e^{-\lambda}\mathbb{P}(Y_i = 0) + 1 - \mathbb{P}(Y_i = 0)
\end{aligned}$$

Rearranging we see that

$$\mathbb{P}(Y_i = 0) = \frac{1}{2 - e^{-\lambda}}$$

All together we define the marginal distribution as

$$\mathbb{P}(Y_i = k) = \begin{cases} \frac{1}{2-e^{-\lambda}} & k = 0 \\ \frac{e^{-\lambda}\lambda^k}{k!\left(2-e^{-\lambda}\right)} & k \geq 1 \end{cases}$$

To evaluate the mean and variance of this RV first let $X \sim \text{Pois}(\lambda) = \text{Pois}(\exp(\beta_0))$. From here we see that

$$\mathbb{E}[Y_i] = \sum_{k=0}^{\infty} k\mathbb{P}(Y_i = k) = \sum_{k=1}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!(2-e^{-\lambda})} = \frac{1}{2-e^{-\lambda}}\mathbb{E}(X) = \frac{\lambda}{2-e^{-\lambda}}$$

$$\mathbb{E}[Y_i^2] = \sum_{k=0}^{\infty} k^2\mathbb{P}(Y_i = k) = \sum_{k=1}^{\infty} k^2\frac{e^{-\lambda}\lambda^k}{k!(2-e^{-\lambda})} = \frac{1}{2-e^{-\lambda}}\mathbb{E}(X^2) = \frac{\lambda+\lambda^2}{2-e^{-\lambda}}$$

$$\mathbb{V}(Y_i) = \mathbb{E}(Y_i^2) - [\mathbb{E}(Y_i)]^2 = \frac{(\lambda+\lambda^2)(2-e^{-\lambda}) - \lambda^2}{(2-e^{-\lambda})^2} = \frac{1}{2-e^{-\lambda}}\lambda + \frac{1-e^{-\lambda}}{2-e^{-\lambda}}\lambda^2$$

Having calculated the variance of this model, we note that the original Poisson would model have variance λ. We note that for over dispersion we require that $\mathbb{V}(Y_i) > \lambda$. That is

$$\begin{aligned} \frac{1}{2-e^{-\lambda}}\lambda + \frac{1-e^{-\lambda}}{2-e^{-\lambda}}\lambda^2 &> \lambda \\ \frac{1}{2-e^{-\lambda}} + \frac{1-e^{-\lambda}}{2-e^{-\lambda}}\lambda &> 1 \\ 1 + \lambda - e^{-\lambda} &> 2 - e^{-\lambda} \\ \lambda &> 1 \end{aligned}$$

Hence we see that this model is under dispersed for $\lambda < 1$ and overdispersed for $\lambda > 1$.