1. Questão 1

   a) Temos que:

   $$f_Y(y) = e^{-x_i\beta}e^{-y_ie^{-x_i\beta}} = \exp\left\{-y_ie^{-x_i\beta} - x_i\beta\right\} = \exp\left\{\phi\left(y_i\theta - b(\theta)\right) + c(y_i,\theta)\right\}\mathbb{1}_A(y_i),$$

   em que $\theta = -e^{-x_i\boldsymbol{\beta}}$, $\phi = 1$, $b(\theta) = -x_i\boldsymbol{\beta}$, $c(y_i, \theta) = 0$, $A = (0, \infty)$.

   b) Temos que:

   $$D(\boldsymbol{y}, \boldsymbol{\mu}) = 2\sum_{i=1}^{n}\left(l(\mu_i^{(0)}, y_i) - l(\mu_i, y_i)\right)$$

   Do item anterior temos que $\beta = -\dfrac{1}{x_i}\ln(-\theta) \to b(\theta) = -\ln(-\theta)$. Logo $\mathcal{E}(Y_i) = b'(\theta) = -\dfrac{1}{\theta} = e^{x_i\beta}$. Por outro lado, sob o modelo saturado, temos que:

   $$l(\mu_i^{(0)}, y_i) = -\ln\mu_i - \frac{y_i}{\mu_i} \to S(\mu_i) = -\frac{1}{\mu_i} + \frac{y_i}{\mu_i^2} \to \mu_i^{(0)} = y_i$$

   Portanto,

   $$D(\boldsymbol{y}, \boldsymbol{\mu}) = 2\sum_{i=1}^{n}\left(-\ln y_i - 1 + x_i\beta + y_ie^{-x_i\beta}\right) = 2\sum_{i=1}^{n}\left(y_ie^{-x_i\beta} + x_i\beta - \ln y_i - 1\right)$$

   c) Temos que:

   $$\begin{aligned} l(\boldsymbol{\beta}) &= -\beta\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}y_ie^{-x_i\beta}, S(\beta) = -n\overline{x} + \sum_{i=1}^{n}y_ix_ie^{-x_i\beta} \\ H(\beta) &= -\sum_{i=1}^{n}y_ix_i^2e^{-x_i\beta}, I(\beta) = \sum_{i=1}^{n}x_i^2 \end{aligned}$$

   Sob as condições de regularidade, temos que $\widehat{\boldsymbol{\beta}} \approx N\left(\beta, \left[\sum_{i=1}^{n}x_i^2\right]^{-1}\right)$, para n suficientemente grande.

d) Temos que:

$$\lambda = -2\left(-n\overline{Y} + n\widehat{\beta}\overline{x} + \sum_{i=1}^{n} Y_i e^{-x_i\widehat{\beta}}\right) = 2\left[n\left(\overline{Y} - \widehat{\beta}\overline{x}\right) - \sum_{i=1}^{n} Y_i e^{-x_i\widehat{\beta}}\right]$$

Sob H_0 e as condições de regularidade, $\lambda \approx \chi^2_{(1)}$ para n suficientemente grande. Assim, rejeita-se H_0 se $P(X > \lambda_0|H_0) < \alpha$, em que $X \sim \chi^2_{(1)}$ e $\lambda_0 = 2\left[n\left(\overline{y} - \widetilde{\beta}\overline{x}\right) - \sum_{i=1}^{n} y_i e^{-x_i\widetilde{\beta}}\right]$

2. Questão 2

a) Temos que:

$$f_Y(y) = \exp\{y \ln p + \ln(1-p)\} \mathbb{1}_A(y) = \exp\{\phi(y\theta - b(\theta)) + c(y,\theta)\} \mathbb{1}_A(y),$$

em que $\theta = \ln p$, $b(\theta) = -\ln(1-p)$, $\phi = 1$, $c(y,\theta) = 0$, $A = \{0, 1, 2, ...\}$. Ademais ($p = e^\theta$)

$$\begin{aligned} b(\theta) &= -\ln\left(1 - e^\theta\right) \to b'(\theta) = \frac{e^\theta}{1 - e^\theta} \to \mathcal{E}(Y) = \mu = \frac{p}{1-p} \to \frac{1}{\mu} = \frac{1}{p} - 1 \\ &\to p = \frac{\mu}{\mu + 1} \end{aligned}$$

b) Temos que:

$$f_Y(y) = \exp\left\{y \ln\left(\frac{\mu}{\mu+1}\right) - \ln(1+\mu)\right\} \mathbb{1}_A(y) = \exp\{\phi(y\theta - b(\theta)) + c(y,\theta)\} \mathbb{1}_A(y),$$

em que $\theta = \ln\left(\frac{\mu}{\mu+1}\right) \to \mu = \frac{1}{e^{-\theta} - 1}$, $b(\theta) = -\ln(1+\mu) = -\theta - \ln\left(1 - e^\theta\right)$, $\phi = 1$, $c(y,\theta) = 0$, $A = \{0, 1, 2, ...\}$.

c) Temos que:

$$\begin{aligned} Y_i &\overset{ind.}{\sim} \text{geométrica}(\mu_i) \\ \ln\left(\frac{\mu_i}{\mu_i + 1}\right) &= \sum_{j=1}^{p} x_{ji}\beta_j \end{aligned}$$

d) Temos que:

$$
\begin{aligned}
l(\beta) &= -\sum_{i=1}^{n} \ln\left(\mu_i + 1\right) + \sum_{i=1}^{n} y_i \left(\ln \mu_i - \ln(\mu_i + 1)\right) \\
&= n \ln(1 + e^{\beta}) + \left(\beta - \ln(1 + e^{\beta})\right) \sum_{i=1}^{n} y_i \\
S(\beta) &= n \frac{e^{\beta}}{1 + e^{\beta}} + n\overline{y} \left(1 - \frac{e^{\beta}}{1 + e^{\beta}}\right) = \frac{n}{1 + e^{\beta}} \left(\overline{y} + e^{\beta}\right) \\
H(\beta) &= \frac{n}{\left(1 + e^{\beta}\right)^2} \left(\left(1 + e^{\beta}\right) e^{\beta} - \left(\overline{y} + e^{\beta}\right) e^{\beta}\right) \\
&= \frac{e^{\beta} n}{\left(1 + e^{\beta}\right)^2} \left(\left(1 + e^{\beta}\right) - \left(\overline{y} + e^{\beta}\right)\right) = \frac{n e^{\beta}}{\left(1 + e^{\beta}\right)^2} \left(1 - \overline{y}\right) \\
I(\beta) &= \frac{\left(1 - e^{\beta}\right) e^{\beta} n}{\left(1 + e^{\beta}\right)^2}.
\end{aligned}
$$

Sob as condições de regularidade, temos que $\widehat{\boldsymbol{\beta}} \approx N\left(\beta, \left[\frac{\left(1 + e^{\beta}\right)^2}{\left(1 - e^{\beta}\right) e^{\beta} n}\right)\right]$, para n suficientemente grande.

3. Questão 3

    a) A estrutura linear para cada covariável para ser apropriada, com exceção das covariáveis "estradas" (para a qual uma estrutura não paramétrica pode ser melhor) e "renda' (para a qual uma curva do segundo grau pode ser mais apropriada). Por outro lado, dado que a resposta é positiva, o modelo gama parece ser mais apropriada. Em suma, ambos parecem apropriados, com uma vantagem para o modelo gama, e considerando a possibilidade de mudar o preditor.

    b) Para o modelo normal, o RS parece apresentar heteroscedasticidade e não normalidade, apesar de um número aceitável de candidatos a outliers e ausência de correlação. Isso indica um mal ajuste desse modelo. Por outro lado, o RCD do modelo gama apresenta normalidade, uma homocedasticidade aproximada, bom comportamento do preditor + função de ligação e ausência de correlação. OS RQ's concordam com os comentários anteriores. Assim, o modelo gama está bem ajustado, enquanto que o modelo normal, não.

    c) Pelo exposto no item b) e pelos resultados da Tabela 1) não restam dúvidas do que o M2 é preferível ao M1. Com efeito, com exceção do p-valor, todas as outras medidas indicam, da Tabela) que o M2 é melhor que o M1.

    d) Pelas Tabelas 2 e 3 podemos ver que somente a covariável "estradas" não se mostrou significativa. Ademais, em ordem decrescente, os maiores impactos no consumo são provocados pela "licença", pela "renda" e, ao final, pela taxa. O impactos da "licença" é positivo, enquanto que os das outras duas são negativos.

4. Questão 4

a) Como se trata de contagem, a resposta é positiva e percebe-se um comportamento semelhante a uma equação do 2o grau entre a resposta e Nitrofen, o modelo parece ter sido uma boa escolha. Entretanto, como a variância e a média não são, em geral, próximas entre si, um modelo com outra relação entre a variância e a média possa ser considerado.

b) O RCD apresenta heterocedasticidades, uma aparente não normalidade, apesar de ausência de correlação e poucos candidatos a outliers. O RQA também indica mal ajuste, pela comportamento sistemático. No entanto, o comportamento preditivo foi bastante satisfatório, ao menos no que tange as predições pontuais. No geral, o modelo não se ajustou bem aos dados.

c) Os parâmetros são, ambos significativos, como pode ser ver pelos respectivos p-valores dos testes de nulidade. Não ajustaria modelo reduzido.

d) Temos que $\mathcal{E}(Y_{ij}|x_{ij} = 0) = \beta_0$ é o valor esperado do total de ovos eclodidos para uma quantidade nula de Nitrofen. Por outro lado, $-\dfrac{\beta_1}{2\beta_2}$ é a quantidade de nitrofen que leva ao número máximo de ovos eclodidos.