1. Questão 1

   a) Temos que:

$$
\begin{aligned}
f(y_i^*) &= P(Y_i^* = y_i^*) = P(Y_i = m_i y_i^*) = \binom{m_i}{m_i y_i^*} \mu_i^{m_i y_i^*} (1-\mu_i)^{m_i(1-y_i^*)} \mathbb{1}_{\left\{0, \frac{1}{m_i}, \frac{2}{m_i}, ,\ldots, 1\right\}}(y_i^*) \\
&= \binom{m_i}{m_i y_i^*} \mu_i^{m_i y_i^*} (1-\mu_i)^{m_i(1-y_i^*)} \mathbb{1}_A(y_i^*).
\end{aligned}
$$

   Além disso,

$$
\mathcal{E}(Y_i^*) = \frac{1}{m_i} m_i \mu_i = \mu_i, \mathcal{V}(Y_i^*) = \frac{1}{m_i^2} m_i \mu_i (1-\mu_i) = \frac{\mu_i(1-\mu_i)}{m_i}.
$$

   b) Temos que:

$$
\begin{aligned}
f(y_i^*) &= \exp\left\{ m_i y_i^* \ln\left(\frac{\mu_i}{1-\mu_i}\right) + m_i \ln(1-\mu_i) + \ln\left[\binom{m_i}{m_i y_i^*}\right] \right\} \mathbb{1}_A(y_i^*) \\
&= \exp\left\{ m_i \left(y_i^* \ln\left(\frac{\mu_i}{1-\mu_i}\right) + \ln(1-\mu_i) \right) + \ln\left[\binom{m_i}{m_i y_i^*}\right] \right\} \mathbb{1}_A(y_i^*) \\
&= \exp\left\{ \phi_i \left[y\theta - b(\theta) \right] + c(y, \phi_i) \right\} \mathbb{1}_A(y),
\end{aligned}
$$

   em que $\phi_i = m_i$, $\theta_i = \ln\left(\frac{\mu_i}{1-\mu_i}\right) \rightarrow \mu_i = \frac{1}{1+e^{-\theta_i}}$,
   $b(\theta_i) = -\ln(1-\mu_i) = \theta_i + \ln(1+e^{-\theta_i})$ e $c(y, \phi_i) = \ln\left[\binom{m_i}{m_i y_i^*}\right]$.

   Por outro lado, note que

$$
\mathcal{E}(Y_i) = m_i \mu_i > m_i \mu_i - m_i \mu_i^2 = \mathcal{V}(Y_i) \rightarrow \mathcal{E}(Y_i) > \mathcal{V}(Y_i).
$$

   c) Temos que:

$$
\ln(1-\mu_{ij}) = -e^{\eta_{ij}} \rightarrow \mu_{ij} = 1 - e^{-e^{\eta_{ij}}} = 1 - e^{-e^{x_{ij}\beta}}
$$

   d) Temos que:

$$
\psi = \frac{\frac{1-e^{-e^{\beta}}}{e^{-e^{\beta}}}}{\frac{1-e^{-1}}{e^{-1}}} = \frac{1}{e-1}\left(\frac{1}{e^{e^{\beta}}-1}\right)^{-1}.
$$

2. Questão 2

   a) Temos que $Q_i = Q(\mu_i, y_i)$.:

$$\frac{\partial Q_i}{\partial \mu_i} = \frac{1}{\sigma^2}\frac{y_i - \mu_i}{V(\mu_i)}.$$

   b) Temos que:

$$\mathcal{E}\left\{\frac{\partial Q(\mu_i; Y_i)}{\partial \mu_i}\right\} = 0 \rightarrow \mathcal{E}(Y_i - \mu_i) = 0 \rightarrow \mathcal{E}(Y_i) = \mu_i.$$

   Por outro lado,

$$\begin{aligned}
\mathcal{V}\left(\frac{\partial Q(\mu_i; Y_i)}{\partial \mu_i}\right) &= \mathcal{E}\left[\left(\frac{\partial Q(\mu_i; Y_i)}{\partial \mu_i}\right)^2\right] - \left[\underbrace{\mathcal{E}\left(\frac{\partial Q(\mu_i; Y_i)}{\partial \mu_i}\right)}_{0}\right]^2 \\
&= -\mathcal{E}\left(\frac{\partial^2 Q(\mu; Y)}{\partial \mu^2}\right) \rightarrow \\
\frac{1}{\sigma^4}\frac{\mathcal{V}(Y_i)}{V^2(\mu_i)} &= -\frac{1}{\sigma^2}\mathcal{E}\left(\frac{-V(\mu_i) - (Y_i - \mu_i)V'(\mu_i)}{V^2(\mu_i)}\right) \\
&\rightarrow \mathcal{V}(Y) = \sigma^2 V(\mu_i).
\end{aligned}$$

c) Temos que:

Se y=0, então

$$Q = \frac{1}{\sigma^2}\int_0^{\mu}\frac{-t}{t(1-t)}dt = \frac{1}{\sigma^2}\left(\ln(1-t)|_0^{\mu}\right) = \frac{1}{\sigma^2}\left(\ln(1-\mu)\right).$$

Se y=1, então

$$Q = \frac{1}{\sigma^2}\int_1^{\mu}\frac{1-t}{t(1-t)}dt = \frac{1}{\sigma^2}\int_1^{\mu}\frac{1}{t}dt = \frac{1}{\sigma^2}\ln(t)|_1^{\mu} = \frac{1}{\sigma^2}\ln(\mu).$$

Logo

$$Q = \frac{1}{\sigma^2}\left(y\ln\mu + (1-y)\ln(1-\mu)\right)$$

d) Temos que $(\mu_i = \Phi(\beta) \rightarrow \dfrac{\partial\mu_i}{\partial\beta} = f(\beta), f \sim N(0,1))$

$$\begin{aligned}
Q &= \sum_{i=1}^{n} Q_i = \sum_{i=1}^{n}\left(y_i\ln\mu_i + (1-y_i)\ln(1-\mu_i)\right) \\
\frac{\partial Q}{\partial\beta} &= \sum_{i=1}^{n}\left(\frac{y_i}{\mu_i}f(\beta) - \frac{1-y_i}{1-\mu_i}f(\beta)\right) \\
&= \sum_{i=1}^{n}\left(\frac{y_i-\mu_i}{\mu_i(1-\mu_i)}\right)f(\beta) = \frac{1}{\sqrt{2\pi}}e^{-\beta^2/2}\sum_{i=1}^{n}\left(\frac{y_i-\mu_i}{\mu_i(1-\mu_i)}\right).
\end{aligned}$$

3. Questão 3

   a) Sim. Primeiramente, a resposta jaz no intervalo (0,1). Adicionalmente, a relação entre a proporção de meninas que apresentam menarca em função da idade não somente está, naturalmente, no intervalo (0,1) mas apresenta uma relação crescente na forma "sigmodal", o que sugere a utilização de uma função de ligação como sendo a fda de alguma v.a. real.

   b) Primeiramente, percebemos que o RCD (dois gráficos da primeira linha da Figura 2) apresenta uma tendência, o que indica um mal ajuste do modelo. Adicionalmente, no QQplot com envelope muitos pontos estão fora das bandas de confiança, além de apresentar uma tendência. Ademais, o RQA, nos wormplots, apresenta uma tendência (maioria dos pontos abaixo da linha horizontal) o que reforça o mal ajuste. No entanto, em relação ao comportamento preditivo, o modelo apresenta um ajuste um pouco aceitável, embora haja proporções que não foram bem preditas. Finalmente, o p-valor indica que o modelo tem um bom ajuste. Em suma, considerando todos os aspectos mencionados, o modelo não está bem ajustado. Como alternativa, podemos considerar um modelo similar, mas usando função de ligação assimétrica, no lugar da ligação logito.

   c) Sob a validade da suposição em questão, o impacto é significativo e positivo, ou seja, quanto maior a idade, maior tende a ser probabilidade de ocorrência de menarca.

   d) Temos que (definindo $p_{i+1} = \dfrac{1}{1 + e^{-(\beta_0 + \beta_1(x_i + 1 - \overline{x}))}}$):

$$\begin{aligned} \ln\left(p_{i+1}/(1-p_{i+1})\right) - \ln\left(p_i/(1-p_i)\right) &= \beta_0 + \beta_1(x_i + 1 - \overline{x}) - \beta_0 - \beta_1(x_i - \overline{x}) \\ &= \beta_1 \rightarrow \frac{p_{i+1}/(1-p_{i+1})}{p_i/(1-p_i)} = e^{\beta_1} = \psi. \end{aligned}$$

   Assim,

$$\widetilde{\psi} = e^{\widetilde{\beta}_1} = 5,10.$$

   Por outro lado,

$$\begin{aligned} \mathcal{V}(\widehat{\psi}) &= \mathcal{V}(\widehat{\beta}_1)e^{2\beta_1} \rightarrow EP(\widehat{\psi}) = EP(\widehat{\beta}_1)e^{\beta_1} \\ \widetilde{EP}(\widehat{\psi}) &= 0,31. \end{aligned}$$

   Log, $IC(\psi) = [5,10 - 1,96 * 0,31; 5,10 + 1,96 * 0,31] = [4,49; 5,71]$.

4. Questão 4

a) A resposta é uma contagem, as relações entre a resposta e cada covariável quantitativa parece ser bem representadas por curvas exponenciais. Como há indicações de existência de interações de primeira e/ou segunda ordens, mas uma quantidade, de pequena a moderada, de observações (no total) e pelos grupos (formados pelas combinações dos níveis dos fatores, covariáveis qualitativas), foram consideradas apenas algumas daquelas interações. Adicionalmente, o modelo contempla a possibilidade existência de correlação da resposta intra unidades experimentais, que é sugerida pelo gráfico de dispersão múltipla. Assim, o modelo é apropriado.

b) Os dois gráficos do resíduo de Pearson indicam, aparentemente, indicam ausência de sistematicidade, bem como de possíveis outliers ou existência de correlações. Assim, o modelo parece ter se ajustado bem aos dados.

c) Pelos p-valores associados aos parâmetros vemos que a covariável RBV não é significativa, em relação à nenhum tipo de droga. Por outro lado a covariável WBC tem efeito, no número de células, para as drogas do tipo 2 e 3, aparentemente diferentes entre si e negativos. Adicionalmente, o período tem efeitos negativos e iguais (Drogas 1 e 2) e positivo para a Droga 3. Finalmente, para medidas feitas no período 1, para valores de WBC iguais a respectiva média observada, o número esperado de células é o mesmo para os três tipos de droga.

d) Sob a validade da suposição em questão, vemos que a estimativa do parâmetro de correlação da estrutura uniforme (que representa a correlação intra-unidades amostrais) é muito próximo de 1. Mesmo não dispondo do erro-padrão associado, tal estimativa sugere uma significância de tal correlação o que, por sua vez, implica que o modelo capturou de forma adequada a correlação presente nos dados.