

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Estatística

Modelos Lineares Generalizados

Modelo Gama

Douglas de Paula Nestlehner

Fevereiro, 2022

Capítulo 1

Problema Apresentado

Para os dados presentes na tabela abaixo envolvendo o tempo de coagulação do sangue considerando dois agentes de coagulação em diferentes concentrações do plasma.

- Ajuste um MLG gama para estes dados. Explique cada passo e discuta seus resultados. Considere diferentes funções de ligação
- Compare os modelos ajustados. Usando a sua intuição escolha o melhor modelo. Explique esta escolha.

Concentração	Agent 1	Agent 2
5	118	69
10	58	35
15	42	26
20	35	21
30	27	18
40	25	16
60	21	13
80	19	12
100	18	12

Tabela 1.1: Tempo de Coagulação

Informações pertinentes:

- Variável resposta: Tempo de coagulação (Quantitativa: 18 observações)
- Covariáveis:
 - Concentração: variável quantitativa;
 - Agentes: variável qualitativa; dois atributos: Agent 1 e Agent 2 (9 observações cada)

1.1 Ajustando os modelos

O problema pede que seja ajustado um modelo linear generalizado Gama, considerando diferentes funções de ligação. Assim sendo, considere as funções de ligação: Reciproca, Identidade e Log, ajustando três modelos o qual fiz uma comparação mais a diante.

1.1.1 Função de Ligação Reciproca

Considerando a função de ligação Reciproca, um possível modelo linear generalizado Gama a ser ajustado é:

$$\frac{1}{\mu} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

X_1 : Covariável Agents, que assume valor igual a 1 se $X_1 = \text{Agent 1}$, e assume valor igual a 0 se $X_1 = \text{Agent 2}$

X_2 : Covariável Concentração, em que optei por assumir o logaritmo da concentração, pois assim a relação entre tempo de coagulação e concentração é melhor aproximada por uma relação linear. Então $X_2 = \log(\text{Concentração})$

μ : Média da variável resposta, $E[Y]$

Na forma matricial:

$$\eta = X\beta$$

Em que:

$$\eta = \frac{1}{\mu}; X = \begin{pmatrix} 1 & 1 & \log(\text{Concentração} \cdot 1) \\ 1 & 1(9\text{primeiras linhas}) & \log(\text{Concentração} \cdot 2) \\ \vdots & \vdots & \vdots \\ 1 & 0(9\text{ultimas linhas}) & \log(\text{Concentração} \cdot 18) \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Ajustando o modelo com auxílio do R, obtive os seguintes valores para β 's

$$\beta_0 = -0.01058$$

$$\beta_1 = -0.01087$$

$$\beta_2 = 0.01776$$

Obs: O ideal seria considerar outros modelos com outros parâmetros como : interação, termo quadrático, etc., no intuito de encontrar o melhor modelo, porem para não deixar a atividade muito extensa, optei por analisar apenas o modelo mais simples

Para ajuste considerando interação entre as covariáveis, adicionaria uma outra coluna na matriz X obedecendo o critério $X_1 * X_2$, como X_1 é categórica que assume apenas valores 0 e 1, teria como resultado apenas $\log(\text{Concentração}_i)$ ou 0.

1.1.2 Função de Ligação Identidade

Considerando a função de ligação identidade, um possível modelo linear generalizado Gama a ser ajustado é:

$$\mu = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Na forma matricial:

$$\eta = X\beta$$

Em que:

$$\eta = \mu; X = \begin{pmatrix} 1 & 1 & \log(\text{Concentração } 1) \\ 1 & 1(9\text{primeiras linhas}) & \log(\text{Concentração } 2) \\ \vdots & \vdots & \vdots \\ 1 & 0(9\text{ultimas linhas}) & \log(\text{Concentração } 18) \end{pmatrix} \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Ajustando o modelo com auxilio do R, obtive os seguintes resultados para β 's:

$$\beta_0 = 71.010$$

$$\beta_1 = 8.724$$

$$\beta_2 = -13.530$$

1.1.3 Função de Ligação: log

Considerando a função de ligação log, um possível modelo linear generalizado Gama a ser ajustado é:

$$\log(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Na forma matricial:

$$\eta = X\beta$$

Em que:

$$\eta = \log(\mu); X = \begin{pmatrix} 1 & 1 & \log(\text{Concentração } 1) \\ 1 & 1(9\text{primeiras linhas}) & \log(\text{Concentração } 2) \\ \vdots & \vdots & \vdots \\ 1 & 0(9\text{ultimas linhas}) & \log(\text{Concentração } 18) \end{pmatrix} \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Ajustando o modelo com auxilio do R, obtive os seguintes resultados para β 's:

$$\beta_0 = 4.9763$$

$$\beta_1 = 0.4703$$

$$\beta_2 = -0.5848$$

1.2 Comparação de Resultados

Utilizando as medidas já calculadas no ajuste do modelo: AIC e deviance que são medidas de qualidade do modelo ajustado, obtive os seguintes resultados:

Modelo	AIC	Deviance
Reciproca	103.0749	0.3004207
Identidade	131.6139	1.4510353
Log	104.2763	0.3210963

Tabela 1.2: Comparação medidas de qualidade

Em ambas as medidas, o modelo que utilizou a função de ligação recíproca $\eta = \frac{1}{\mu}$ obteve medidas menores do que os demais modelos que utilizou-se outras funções de ligação, indicando que foi o melhor modelo ajustado (em relação aos outros dois modelos). Confirmando o que já vimos em aula, que quando se trata de modelos lineares generalizados Gama, a função de ligação mais indicada é a recíproca.