

Modelos Lineares Generalizados - 2024.2 - Profa. Terezinha Ribeiro

Lista Prática 1

OBSERVAÇÕES:

- (a) A solução da Lista Prática 1 deve ser feita em formato de relatório.
- (b) Além das respostas dos problemas, a organização e escrita formal das respostas serão critérios de pontuação.
- (c) Não devem ser incluídos prints de saídas do R. Todos os resultados do relatório devem ser apresentados em tabelas, figuras e texto. Tabelas e figuras incluídas no relatório necessariamente devem ser comentadas/interpretadas ao longo do texto do relatório. Tabelas e figuras sem comentários não serão pontuados.
- (d) A entrega do relatório da Lista Prática 1 deve ser feito até às 23h59m do dia 31/01/25.
- (e) O formato de entrega pode ser presencial com texto impresso ou enviado para o email terezinha.ribeiro@unb.br em formato PDF.
- (f) Não será obrigatório o envio do script R. A correção da lista será feita exclusivamente com os resultados apresentados no relatório.

Problema 1.

Nambe Mills, localizada em Santa Fé, Novo México, é uma fabricante de utensílios de mesa. Após a fundição, os itens produzidos pela Nambe Mills são moldados, lixados e polidos. Em 1989, como uma forma de racionalizar a produção de seus produtos, a empresa registrou o tempo total de lixamento e polimento, bem como o diâmetro de cada item. Os dados *nambeware* da biblioteca `GLMsData` referem-se ao

- **Type:** tipo do item (Bowl, CassDish, Dish, Plate, Tray);
- **Diam:** diâmetro do produto em polegadas;
- **Time:** tempo total de lixamento e polimento em minutos;
- **Price:** preço em dólares do item.

A variável resposta é o preço em dólares do produto. Faça uma análise descritiva dos dados através de gráficos e medidas-resumo. Inicialmente, proponha um ajuste via MLG normal para estes dados. Comece com ligação identidade. Faça seleção das covariáveis relevantes. Faça análise de diagnóstico, verificando se existem afastamentos sérios das suposições do modelo. Em seguida, proponha ajustes MLGs Gama e Normal Inversa para diferentes funções de ligação. Faça as respectivas análises de diagnóstico e se existem pontos atípicos que influenciam no ajuste. Ao final, indique qual o modelo ajustado que conduz ao melhor ajuste aos dados e faça as devidas interpretações.

Para ler os dados, faça:

```
require(GLMsData); data(nambeware); attach(nambeware)
```

Problema 2.

No arquivo **recrutas.txt** estão descritos dados referentes ao número de infecções de ouvido (**ninfec** - variável resposta) observadas em uma amostra aleatória de $n = 287$ recrutas norte-americanos em 1990, e as seguintes variáveis explicativas:

- **habito**: hábito de nadar (ocasional ou frequente);
- **local**: local de nadar (praia ou piscina),
- **fetaria**: faixa etária do recruta (15-19, 20-24, 25-29),
- **genero**: gênero do recruta (feminino ou masculino).

Faça uma análise descritiva dos dados através de gráficos e medidas-resumo. Ajuste um modelo de regressão Poisson para estes dados. Faça seleção das covariáveis relevantes. Interprete o modelo ajustado e faça uma análise de diagnóstico. Verifique se existem afastamentos sérios das suposições do modelo. Em seguida, ajuste um modelo de regressão binomial negativo. Faça seleção das covariáveis relevantes. Faça a análise de diagnóstico. Verifique se existem afastamentos sérios das suposições do modelo e se existem pontos atípicos que influenciam no ajuste. Escolha o modelo final para os dados e apresente todas as interpretações relevantes.

Para ler os dados use

```
recrutas <- read.table("recrutas.txt", header=T); attach(recrutas)
```

Informe previamente os fatores:

```
habito <- as.factor(habito); local <- as.factor(local);  
fetaria <- as.factor(fetaria); genero <- as.factor(genero)
```

Problema 3.

No arquivo **pulso.txt** são descritas as variáveis pulsação em repouso (1: normal, 2: alta), hábito de fumar (1: sim, 2: não) e peso (em kg) de 92 adultos do sexo masculino. Faça uma análise descritiva dos dados. Ajuste MLGs Bernoulli considerando diversas funções de ligação para explicar a probabilidade de pulsação alta dadas as demais variáveis. Escolha o melhor modelo ajustado e verifique se existem desvios da suposição do modelo. Interprete os resultados de acordo com a função de ligação escolhida. Para ler os dados use

```
pulso <- scan("pulso.txt", list(pulsacao=0, habitof=0, peso=0))  
attach(pulso)  
pulsacao <- as.factor(pulsacao); habitof <- as.factor(habitof)
```

Problema 4.

Os conjuntos de dados apresentados nos arquivos **dose1.txt**, **dose2.txt** e **dose3.txt** são provenientes de um experimento de dose-resposta conduzido para avaliar a influência dos extratos vegetais “aquoso frio de folhas” (dose do tipo 1), “aquoso frio de frutos” (dose do tipo 2) e de um extrato químico (dose do tipo 3) na morte de um determinado tipo de caramujo. Para cada conjunto de dados separadamente, ajuste MLGs binomiais (variando a ligação) para modelar a proporção de caramujos mortos em função de cada tipo de dose. Para o melhor ajuste encontrado (use envelopes como critério), apresente estimativas pontuais para a dose letal DL50. Para ler os dados use

```
dose1 <- scan("dose1.txt", list(dose1=0, care1=0, carm1=0))  
dose2 <- scan("dose2.txt", list(dose2=0, care2=0, carm2=0))  
dose3 <- scan("dose3.txt", list(dose3=0, care3=0, carm3=0))  
attach(dose1); attach(dose2); attach(dose3)
```

As quantidades $dose_i$, $care_i$ e $carm_i$, $i = 1, 2, 3$, denotam a dose, o número de caramujos expostos, o número de caramujos mortos, respectivamente, para o experimento com o i -ésimo tipo de dose.