

INSTITUTO SUPERIOR DE AGRONOMIA
Modelos Matemáticos e Aplicações (2020-21)

Teste – Modelos Lineares Generalizados e Modelos Lineares Mistos

31 de Maio, 2021

Duração: 2h30

I [9 valores]

Um estudo visa estimar o número de bagos por cacho de uvas (variável de contagem BE, do inglês *berries*) a partir de três outras variáveis: o peso do cacho (variável Bw, de *bunch weight*, em *g*) e duas variáveis observáveis em imagens a duas dimensões obtidas por robôs que percorrem a vinha, a saber, o número de bagos por cacho visíveis nas imagens (variável BEv, de contagem) e a área do cacho nas imagens (variável Ba, de *bunch area*, em *cm²*). O conjunto de dados usado para ajustar o modelo incluía observações sobre 75 cachos de cada uma de 5 castas, num total de 375 observações, mas como se pretendia um modelo a aplicar a qualquer casta, as observações foram tratadas em conjunto.

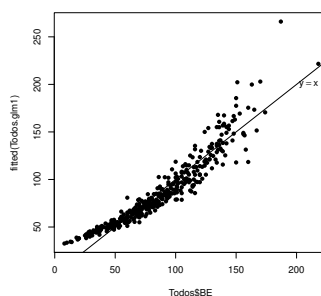
Eis alguns indicadores de síntese:

```
> summary(Todos[,c("BE", "BEv", "Ba", "Bw")])
      BE      BEv      Ba      Bw
Min.   : 8.0   Min.   : 8.0   Min.   : 10.60  Min.   : 10.6
1st Qu.: 61.0  1st Qu.: 34.0  1st Qu.: 54.52  1st Qu.: 86.0
Median : 85.0  Median : 44.0  Median : 74.12  Median : 133.6
Mean   : 87.7  Mean   : 44.7  Mean   : 74.44  Mean   : 137.2
3rd Qu.: 113.5 3rd Qu.: 55.0  3rd Qu.: 90.67  3rd Qu.: 174.8
Max.   : 218.0 Max.   : 83.0  Max.   : 154.62  Max.   : 351.0
```

1. Tendo em conta a natureza da componente aleatória BE diga, justificando, qual a distribuição de probabilidades mais adequada para a descrever (de entre as estudadas nas aulas).
2. Independentemente da resposta na alínea anterior, foram ajustados dois Modelos Lineares Generalizados com variável resposta Poisson, mas diferindo na sua função de ligação. Eis os resultados:

<pre>> summary(Todos.glm1) Call: glm(formula = BE ~ BEv + Bw + Ba, family = poisson(link = log), data = Todos) Coefficients: Estimate Std. Error z value Pr(> z) (Intercept) 3.3338620 0.0201937 165.094 <2e-16 BEv 0.0166706 0.0007675 21.721 <2e-16 Bw 0.0029365 0.0001973 14.881 <2e-16 Ba -0.0011815 0.0005132 -2.302 0.0213 --- Null deviance: 5970.14 on 374 degrees of freedom Residual deviance: 676.76 on 371 degrees of freedom AIC: 3012.5</pre>	<pre>> summary(Todos.glm2) Call: glm(formula = BE ~ BEv + Bw + Ba, family = poisson(link = identity), data = Todos) Coefficients: Estimate Std. Error z value Pr(> z) (Intercept) -4.07425 1.14234 -3.567 0.000362 BEv 1.38669 0.07167 19.348 < 2e-16 Bw 0.34403 0.02071 16.613 < 2e-16 Ba -0.23393 0.05079 -4.606 4.1e-06 --- Null deviance: 5970.14 on 374 degrees of freedom Residual deviance: 267.95 on 371 degrees of freedom AIC: 2603.7</pre>
--	---

- (a) Descreva em pormenor o modelo indicado à esquerda (modelo `Todos.glm1`).
- (b) Eis a nuvem de pontos correspondentes aos bagos por cacho (eixo horizontal) e correspondentes valores estimados pelo modelo `Todos.glm1` (à esquerda), com a recta $y = x$. Comente-o.



- (c) Indique o número médio de bagos que o modelo ajustado à direita (modelo `Todos.glm2`) faria corresponder a um cacho de peso 20 g e com, na respectiva imagem, 15 cm^2 de área e 10 bagos visíveis. Comente, sabendo também que o valor previsto pelo outro modelo é 34.521.
 - (d) Qual dos dois modelos escolheria, com base na informação disponível? Justifique a sua resposta.
 - (e) Considere uma modificação ao modelo `Todos.glm2` (à direita): considerar que a distribuição da sua componente aleatória é Normal. Comente esse modelo. Como seria possível comparar este novo modelo com o modelo `Todos.glm2`?
3. Os modelos acima ajustados incluem um preditor que exige a pesagem manual dos bagos (`Bw`). Procurou-se um modelo cuja componente sistemática dependa apenas de medições sobre as imagens recolhidas de forma automática. Foi ajustado um modelo Poisson, com função ligação identidade, e apenas dois preditores: `BEv` e `Ba`. O desvio residual obtido foi 547.3. Efectue um Teste de Razão de Verossimilhanças para saber se este novo modelo tem uma qualidade de ajustamento significativamente pior do que o modelo correspondente com 3 preditores. Comente.