Modelos Lineares Generalizados - Verificação do Ajuste do Modelo

Erica Castilho Rodrigues

3 de Junho de 2016

Estatística de Pearson Generalizada

Testes de Hipóteses usando a função Deviance

Estatística de Pearson Generalizada

Estatística de Pearson Generalizada

- Uma outra medida usada para verificar o ajuste do modelo.
- Essa estatística é dada por

$$X_p^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{Var(Y_i)}$$

onde $Var(Y_i)$ é a função de variância estimada sob o modelo que está sendo ajustado aos dados.

- Para o Poisson e Binomial a estatística fica

$$X_p^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

que é a Estatística Qui-Quadrado usual.

- Essa estatistica tem a seguinte distribuição assintótica

$$X_p^2 \sim \chi^2_{n-p}$$

onde

  - n é o tamanho da amostra;
  - p é o número de parâmetros do modelo.

- A Deviance é mais usada do que a Estatística de Pearson Generalizada.
- Isso acontece porque para a Deviance temos que:
  - seu valor sempre dimui quando acrescentamos variáveis no modelo;
  - o mesmo não é verdade para a Estatística de Pearson.

Testes de Hipóteses usando a função Deviance

- Vimos que podemos fazer testes sobre o vetor β utilizando a distribuição assintótica

$$\mathbf{b} \sim N(\beta, I(\beta^{-1}) \, .$$

- Uma alternativa:
  - comparar o ajuste de dois modelos;
  - o modelo com a variável e o modelo sem a variável.
- Um modelo deve estar contido no outro.
- A diferença deve ser apenas a variável incluída/retirada.
- A distribuição de probabilidade deve ser a mesma.
- A função de ligação deve ser a mesma.

- Vamos chamar o modelo mais simples (menos variáveis) de M_0.
- O modelo mais complexo (mais variáveis) será M_1.
- Para o modelo M_0 temos a hipótese nula de que

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0 = \begin{bmatrix} \beta_1 \\ \vdots \beta_q \end{bmatrix} .$$

- Para o modelo M_1 temos a hipótese alternativa

$$H_1 : \boldsymbol{\beta} = \boldsymbol{\beta}_0 = \begin{bmatrix} \beta_1 \\ \vdots \beta_p \end{bmatrix} .$$

- Observe que $q < p < n$.

- Podemos testar

$$H_0 \quad vs \quad H_1$$

usando a diferença das Deviances dos dois modelos

$$\Delta D = D_0 - D_1 = 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_0, \mathbf{y})\right] - 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_1, \mathbf{y})\right].$$

- Se os modelos estão bem ajustados temos que

$$D_0 \sim$$

- Podemos testar

$$H_0 \quad vs \quad H_1$$

usando a diferença das Deviances dos dois modelos

$$\Delta D = D_0 - D_1 = 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_0, \mathbf{y})\right] - 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_1, \mathbf{y})\right].$$

- Se os modelos estão bem ajustados temos que

$$D_0 \sim \chi^2_{(n-p)} \quad D_0 \sim$$

- Podemos testar

$$H_0 \quad vs \quad H_1$$

usando a diferença das Deviances dos dois modelos

$$\Delta D = D_0 - D_1 = 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_0, \mathbf{y})\right] - 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_1, \mathbf{y})\right].$$

- Se os modelos estão bem ajustados temos que

$$D_0 \sim \chi^2_{(n-p)} \quad D_0 \sim \chi^2_{(n-q)}.$$

- Portanto

$$\Delta D = D_0 - D_1 \sim$$

- Podemos testar
$$H_0 \quad vs \quad H_1$$
usando a diferença das Deviances dos dois modelos
$$\Delta D = D_0 - D_1 = 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_0, \mathbf{y})\right] - 2\left[l(\mathbf{b}_{max}, \mathbf{y}) - l(\mathbf{b}_1, \mathbf{y})\right].$$
- Se os modelos estão bem ajustados temos que
$$D_0 \sim \chi^2_{(n-p)} \quad D_0 \sim \chi^2_{(n-q)}\,.$$
- Portanto
$$\Delta D = D_0 - D_1 \sim \chi^2_{(n-q)-(n-p)} \text{ ou seja } \chi^2_{p-q}\,.$$

- Hipóteses a serem testadas:
  - H_0: a diferença entre M_0 e M_1 não é significativa;
  - H_1: a diferença enrte os modelos é significativa.
- Se ΔD não é um valor atípico na distribuição χ^2_{p-q}:
  - podemos aceitar H_0 permanecer com o modelo mais simples;
  - a diferença de ajuste entre os modelos não é significativa.
- H_0 é rejeitada para valores grandes ou pequenos de ΔD?

- Hipóteses a serem testadas:
  - H_0: a diferença entre M_0 e M_1 não é significativa;
  - H_1: a diferença enrte os modelos é significativa.
- Se ΔD não é um valor atípico na distribuição χ^2_{p-q}:
  - podemos aceitar H_0 permanecer com o modelo mais simples;
  - a diferença de ajuste entre os modelos não é significativa.
- H_0 é rejeitada para valores grandes ou pequenos de ΔD? Grandes.
- Como fica a região crítica?
  - se $\Delta D < \chi^2_c$ não rejeitamos H_0 permanecemos com o modelo M_0;
  - se $\Delta D > \chi^2_c$ rejeitamos H_0 e ficamos com o modelo M_1.

- A aproximação assintótica da distribuição de ΔD é melhor do que de D.
- Se temos um parâmetro de ruído para estimar,
  - nem sempre a Deviance poderá ser obtida diretamente dos dados;
  - precisa ainda do parâmetro de ruído.
- Vimos no caso Normal, por exemplo, que

$$D = \frac{\sum_i (y_i - \hat{y}_i)^2}{\sigma^2}$$

  precisamos ainda estimar σ^2.
- Vejamos como isso é feito no exemplo a seguir.

Exemplo

- Considere o modelo linear normal

$$E(Y_i) = \mu_i = \mathbf{x}_i^T \boldsymbol{\beta} \,.$$

- Já vimos que a Deviance desse modelo é dada por

$$D = \frac{\sum_i (y_i - \hat{y}_i)^2}{\sigma^2}$$

- Vamos usar a seguinte notação:
  - $\hat{y}_i(0)$ é o valor ajustado pelo modelo M_0;
  - $\hat{y}_i(1)$ é o valor ajustado pelo modelo M_1.

Exemplo

- A Deviance do modelo M_0 (tem q parâmetros) fica

$$D_0 =$$

Exemplo

- A Deviance do modelo M_0 (tem q parâmetros) fica

$$D_0 = \frac{\sum_i (y_i - \hat{y}_i(0))^2}{\sigma^2}$$

e do modelo M_1 (que tem p parâmetros)

$$D_1 =$$

Exemplo

- A Deviance do modelo M_0 (tem q parâmetros) fica

$$D_0 = \frac{\sum_i (y_i - \hat{y}_i(0))^2}{\sigma^2}$$

e do modelo M_1 (que tem p parâmetros)

$$D_1 = \frac{\sum_i (y_i - \hat{y}_i(1))^2}{\sigma^2}.$$

- Temos ainda que

$$D_0 \sim$$

Exemplo

- A Deviance do modelo M_0 (tem q parâmetros) fica

$$D_0 = \frac{\sum_i (y_i - \hat{y}_i(0))^2}{\sigma^2}$$

e do modelo M_1 (que tem p parâmetros)

$$D_1 = \frac{\sum_i (y_i - \hat{y}_i(1))^2}{\sigma^2} \, .$$

- Temos ainda que

$$D_0 \sim \chi^2_{n-q} \quad D_1 \sim$$

Exemplo

- A Deviance do modelo M_0 (tem q parâmetros) fica

$$D_0 = \frac{\sum_i (y_i - \hat{y}_i(0))^2}{\sigma^2}$$

e do modelo M_1 (que tem p parâmetros)

$$D_1 = \frac{\sum_i (y_i - \hat{y}_i(1))^2}{\sigma^2} \,.$$

- Temos ainda que

$$D_0 \sim \chi^2_{n-q} \quad D_1 \sim \chi^2_{n-p} \quad \Delta D \sim \chi^2_{p-q} \,.$$

- Para não termos que encontrar σ^2 vamos usar a razão

$$F = \frac{\Delta D/(p-q)}{D_1/(n-p)} \sim$$

- Para não termos que encontrar σ^2 vamos usar a razão

$$F = \frac{\Delta D/(p-q)}{D_1/(n-p)} \sim F_{p-q,n-p} \, .$$

- Dessa maneira, F fica

$$F = \frac{\Delta(\sum_i (y_i - \hat{y}_i(0))^2 - \sum_i (y_i - \hat{y}_i(1))^2)/(p-q)}{(\sum_i (y_i - \hat{y}_i(1))^2)/(n-p)} \sim$$

- Para não termos que encontrar σ^2 vamos usar a razão

$$F = \frac{\Delta D/(p-q)}{D_1/(n-p)} \sim F_{p-q,n-p} \,.$$

- Dessa maneira, F fica

$$F = \frac{\Delta(\sum_i (y_i - \hat{y}_i(0))^2 - \sum_i (y_i - \hat{y}_i(1))^2)/(p-q)}{(\sum_i (y_i - \hat{y}_i(1))^2)/(n-p)} \sim F_{p-q,n-p}.$$

- Como o σ^2 é cancelado nessa razão, torna-se desncessário estimá-lo.
- Rejeitamos H_0 quando F é grande.

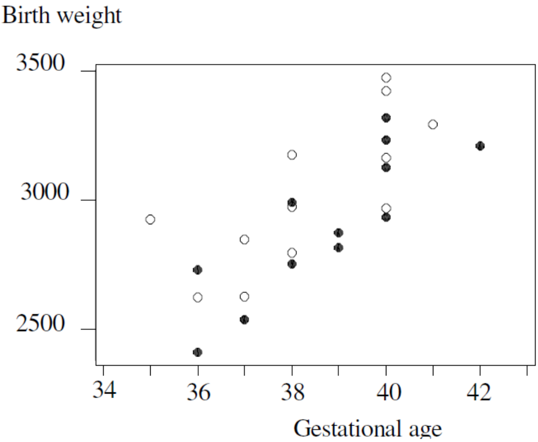
Exemplo

- A tabela a seguir mostra os dados do peso e a idade de gestação de bebês em um hospital.

	Boys		Girls	
	Age	Birthweight	Age	Birthweight
	40	2968	40	3317
	38	2795	36	2729
	40	3163	40	2935
	35	2925	38	2754
	36	2625	42	3210
	37	2847	39	2817
	41	3292	40	3126
	40	3473	37	2539
	37	2628	36	2412
	38	3176	38	2991
	40	3421	39	2875
	38	2975	40	3231
Means	38.33	3024.00	38.75	2911.33

Exemplo (continuação)

- A figura a seguir mostra o gráfico de dispersão entre as duas variáveis.



Exemplo (continuação)

- Os bebês estão divididos em dois grupos:
  - masculino e feminino.
- Como podemos escrever o modelo com essas duas variáveis?

Exemplo (continuação)

- Os bebês estão divididos em dois grupos:
  - masculino e feminino.
- Como podemos escrever o modelo com essas duas variáveis?
- A variável sexo entra como Dummy.
- O modelo sem interação fica

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i \quad \epsilon_i \sim^{iid} N(0, \sigma^2)$$

onde

Exemplo (continuação)

- Os bebês estão divididos em dois grupos:
  - masculino e feminino.
- Como podemos escrever o modelo com essas duas variáveis?
- A variável sexo entra como Dummy.
- O modelo sem interação fica

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i \quad \epsilon_i \sim^{iid} N(0, \sigma^2)$$

  onde
  - Y_i é o peso do bebê;
  - X_i é idade de gestação
  - Z_i é uma indicadora que representa sexo (1 - masculino, 0 - feminino).

Exemplo (continuação)

- Queremos verificar a necessidade de incluir o termo de interação.
- O modelo com interação é dado por

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i + \beta_3 X_i Z_i \quad \epsilon_i \sim^{iid} N(0, \sigma^2)$$

- Vamos denotar por
  - M_0: modelo sem interação;
  - M_1: o modelo com interação.
- Queremos verificar se o ganho de ajuste de M_1 em relação a M_0 é significativo.

Exemplo (continuação)

- A Soma dos Quadrados dos Resíduos está relacionada com a Deviance da segunte maneira

$$SQE = \sum_i (y_i - \hat{y}_i)^2 =$$

Exemplo (continuação)

- A Soma dos Quadrados dos Resíduos está relacionada com a Deviance da segunte maneira

$$SQE = \sum_i (y_i - \hat{y}_i)^2 = \sigma^2 D.$$

- Para os modelos temos que

$$SQE_0 = 658770.8 \quad SQE_1 = 652424.5$$

ou seja

$$D_0 =$$

Exemplo (continuação)

- A Soma dos Quadrados dos Resíduos está relacionada com a Deviance da segunte maneira

$$SQE = \sum_i (y_i - \hat{y}_i)^2 = \sigma^2 D.$$

- Para os modelos temos que

$$SQE_0 = 658770.8 \quad SQE_1 = 652424.5$$

ou seja

$$D_0 = \frac{658770.8}{\sigma^2} \quad D_1 =$$

Exemplo (continuação)

- A Soma dos Quadrados dos Resíduos está relacionada com a Deviance da segunte maneira

$$SQE = \sum_i (y_i - \hat{y}_i)^2 = \sigma^2 D.$$

- Para os modelos temos que

$$SQE_0 = 658770.8 \quad SQE_1 = 652424.5$$

ou seja

$$D_0 = \frac{658770.8}{\sigma^2} \quad D_1 = \frac{652424.5}{\sigma^2}.$$

Exemplo (continuação)

- Temos que $n = 24$ logo

$$F = \frac{\Delta(\sum_i (y_i - \hat{y}_i(0))^2 - \sum_i (y_i - \hat{y}_i(1))^2)/(p-q)}{(\sum_i (y_i - \hat{y}_i(1))^2)/(n-p)}$$

$$= \frac{(SQE_0 - SQE_1)/(p-q))}{SQE_1/(n-p)} =$$

Exemplo (continuação)

- Temos que $n = 24$ logo

$$F = \frac{\Delta(\sum_i (y_i - \hat{y}_i(0))^2 - \sum_i (y_i - \hat{y}_i(1))^2)/(p-q)}{(\sum_i (y_i - \hat{y}_i(1))^2)/(n-p)}$$

$$= \frac{(SQE_0 - SQE_1)/(p-q))}{SQE_1/(n-p)} =$$

$$\frac{(658770.8 - 652424.5)/(4-3)}{652424.5/(24-4)} = 0,19 \, .$$

Exemplo (continuação)

- Devemos comparar esse valor com a

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por

$$F_c = 4.35 \quad \text{pois}$$

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por

$$F_c = 4.35 \quad \text{pois} P(F_{1,20} > 4,35) = 0,05 \,.$$

- A região crítica é dada por

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por

$$F_c = 4.35 \quad \text{pois} P(F_{1,20} > 4,35) = 0,05 \,.$$

- A região crítica é dada por
  - se $F < F_c$, não rejeitamos H_0
  - se $F > F_c$, rejeitamos H_0 .
- Conclusão do teste:

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por
$$F_c = 4.35 \quad \text{pois} P(F_{1,20} > 4,35) = 0,05 \,.$$
- A região crítica é dada por
  - se $F < F_c$, não rejeitamos H_0
  - se $F > F_c$, rejeitamos H_0 .
- Conclusão do teste:
  - $F_{obs} = 0,19 < 4,35$ não rejeitamos H_0;

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por
$$F_c = 4.35 \quad \text{pois} P(F_{1,20} > 4,35) = 0,05 \, .$$
- A região crítica é dada por
  - se $F < F_c$, não rejeitamos H_0
  - se $F > F_c$, rejeitamos H_0 .
- Conclusão do teste:
  - $F_{obs} = 0,19 < 4,35$ não rejeitamos H_0;
  - não é necessário incluir termo de interação no modelo;

Exemplo (continuação)

- Devemos comparar esse valor com a $F_{1,20}$.
- Rejeitamos H_0, quando F é grande.
- Fixando $\alpha = 0,05$, o valor crítico dessa distribuição é dado por

$$F_c = 4.35 \quad \text{pois} P(F_{1,20} > 4,35) = 0,05 \,.$$

- A região crítica é dada por
  - se $F < F_c$, não rejeitamos H_0
  - se $F > F_c$, rejeitamos H_0 .
- Conclusão do teste:
  - $F_{obs} = 0,19 < 4,35$ não rejeitamos H_0;
  - não é necessário incluir termo de interação no modelo;
  - conclusão: o efeito da idade no peso é o mesmo para meninos e meninas.

Exemplo

- Um pesquisador quer verificar qual a dose ideal de inseticida para matar insetos.
- Diferentes doses são usadas para grupos de uma mesma espécie.
- Vamos usar a seguinte notação:
  - d_i: dose do inseticida;
  - m_i: número de insetos que receberam a dose;
  - y_i: número de insetos mortos dentre os m_i que receberam o inseticida;
  - p_i: proporção de insetos mortos.

Exemplo (continuação)

- A tabela a seguir mostra os dados coletados

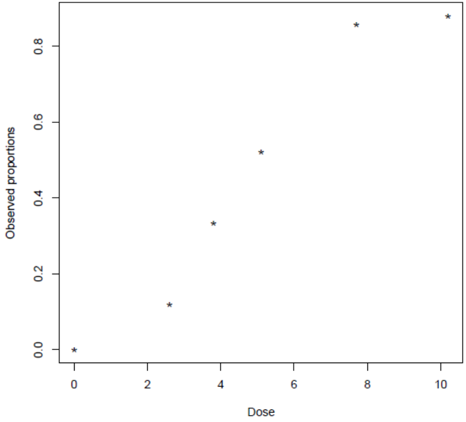
Dose (d_i)	m_i	y_i	p_i
0,0	49	0	0,00
2,6	50	6	0,12
3,8	48	16	0,33
5,1	46	24	0,52
7,7	49	42	0,86
10,2	50	44	0,88

Exemplo (continuação)

- O pesquisador deseja determinar quais as doses tais que
  - 50% dos insetos são mortos (LD_{50});
  - 90% dos insetos são mortos (LD_{90}).
- Podem usar esse dado para aplicação em campo.

Exemplo (continuação)

- A figura a seguir mostra o gráfico dispersão entre:
  - doses de inseticida (d_i) e proporção de insetos mortos (p_i).



Exemplo (continuação)

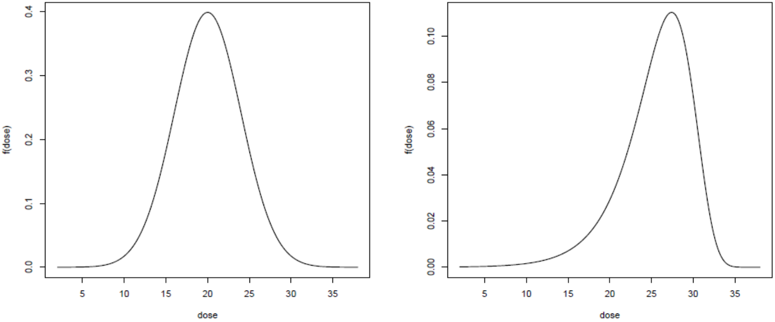
- O gráfico tem um aspecto sigmoidal.
- Esse formato pode nos guiar na escolha da função de ligação.
- Esse tipo de ensaio é chamado de dose-resposta.
- Dois aspectos devem ser considerados:
  - a dose da droga (inseticida, fungicida, herbicida, medicamento);
  - o indivíduo que recebe a droga (inseto, planta, fungo, paciente).
- A reposta do indivíduo é binária:
  - responde (1) ou não responde (0) ao tratamento.

Exemplo (continuação)

- A resposta dependerá do nível da dosagem aplicada.
- Cada indivíduo tem um nível a partir do qual responde ao tratamento.
- Esse valor é chamdo de tolerância do indivíduo.
- Essa tolerância varia de um indivíduo para o outro dentro da população.
- Portanto é uma variável aleatória e vamos denotá-la por U.

Exemplo (continuação)

- A figura seguir mostra exemplos de distribuição da tolerância.



Exemplo (continuação)

- Vamos denotar por $f(u)$ a função de densidade da tolerância.
- Seja d a dose ministrada à toda população.
- Quais indivíduos responderão à droga?
- Aqueles tais que
$$U < d\,.$$
- A probabilidade de um indivíduo escolhido ao acaso responda ao tratamento é

$$\pi(d) =$$

Exemplo (continuação)

- Vamos denotar por $f(u)$ a função de densidade da tolerância.
- Seja d a dose ministrada à toda população.
- Quais indivíduos responderão à droga?
- Aqueles tais que

$$U < d \, .$$

- A probabilidade de um indivíduo escolhido ao acaso responda ao tratamento é

$$\pi(d) = P(U < d) =$$

Exemplo (continuação)

- Vamos denotar por $f(u)$ a função de densidade da tolerância.
- Seja d a dose ministrada à toda população.
- Quais indivíduos responderão à droga?
- Aqueles tais que
$$U < d\,.$$
- A probabilidade de um indivíduo escolhido ao acaso responda ao tratamento é
$$\pi(d) = P(U < d) = \int_{-\infty}^{d} f(u)du\,.$$

Exemplo (continuação)

- Para valores pequenos de d quanto deve valer $\pi(d)$?

Exemplo (continuação)

- Para valores pequenos de d quanto deve valer $\pi(d)$?

$$\pi(d) \approx 0 \, .$$

- Para valores grandes de d quanto deve valer $\pi(d)$?

Exemplo (continuação)

- Para valores pequenos de d quanto deve valer $\pi(d)$?

$$\pi(d) \approx 0 \, .$$

- Para valores grandes de d quanto deve valer $\pi(d)$?

$$\pi(d) \approx 1 \, .$$

- π é uma função crescente ou decrescente de d?

Exemplo (continuação)

- Para valores pequenos de d quanto deve valer $\pi(d)$?

$$\pi(d) \approx 0 .$$

- Para valores grandes de d quanto deve valer $\pi(d)$?

$$\pi(d) \approx 1 .$$

- π é uma função crescente ou decrescente de d?
- Crescente, quanto maior a dose maior a probabilidade de resposta.

Exemplo (continuação)

- No exemplo dos insetos queremos encontrar um modelo razoável de como $\pi(d)$ varia com d.
- E a partir disso encontrar os valores de doses tais que
  - 50% dos indivíduos respondem à droga (LD_{50});
  - 90% dos indivíduos respondem à droga (LD_{90}).
- Seja Y_i a variável aleatória que denota o número de insetos mortos.
- Seja π_i a probailidade de um inseto do i-ésimo grupo morrer.
- Qual a distribuição de Y_i?

$$Y_i \sim$$

Exemplo (continuação)

- No exemplo dos insetos queremos encontrar um modelo razoável de como $\pi(d)$ varia com d.
- E a partir disso encontrar os valores de doses tais que
  - 50% dos indivíduos respondem à droga (LD_{50});
  - 90% dos indivíduos respondem à droga (LD_{90}).
- Seja Y_i a variável aleatória que denota o número de insetos mortos.
- Seja π_i a probailidade de um inseto do i-ésimo grupo morrer.
- Qual a distribuição de Y_i?

$$Y_i \sim Bin(\pi_i, m_i) \ .$$

Exemplo (continuação)

- Vamos usar a função de ligação canônica.
- Qual ligação é essa?

Exemplo (continuação)

- Vamos usar a função de ligação canônica.
- Qual ligação é essa? Logística.
- Isso significa que:

$$\pi =$$

Exemplo (continuação)

- Vamos usar a função de ligação canônica.
- Qual ligação é essa? Logística.
- Isso significa que:

$$\pi = \frac{1}{1 + e^{\eta_i}} \text{ ou}$$

Exemplo (continuação)

- Vamos usar a função de ligação canônica.
- Qual ligação é essa? Logística.
- Isso significa que:

$$\pi = \frac{1}{1+e^{\eta_i}} \text{ ou } \log\left(\frac{\pi_i}{1-\pi_i}\right) = \eta_i \, .$$

- Vamos ajustar o seguinte modelo

$$Y_i \sim$$

Exemplo (continuação)

- Vamos usar a função de ligação canônica.
- Qual ligação é essa? Logística.
- Isso significa que:

$$\pi = \frac{1}{1 + e^{\eta_i}} \text{ ou } \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \eta_i \, .$$

- Vamos ajustar o seguinte modelo

$$Y_i \sim Bin(\pi_i, m_i) \quad \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 d_i \, .$$

Exemplo (continuação)

- O script usado para ajustar o modelo foi o seguinte:

```
x=c(0,2.6,3.8,5.1,7.7,10.2)
m=c(49,50,48,46,49,50)
y=c(0,6,16,24,42,44)

dados=data.frame(x=x,y=y,m=m)

modelo=glm(cbind(y,m-y)~x, family="binomial",
data=dados)
```

- Precisamos criar dois vetores:
  - um com o número de sucesos e outro com número de fracassos.

```
cbind(y,m-y)
```

Exemplo (continuação)

- O resumo do ajuste encontra-se a seguir.

```
> summary(modelo)


Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.22566    0.36992  -8.720   <2e-16 **
x            0.60513    0.06781   8.923   <2e-16 **

    Null deviance: 163.745  on 5  degrees of freedo
Residual deviance:  10.258  on 4  degrees of freedo
AIC: 33.479
```

Exemplo (continuação)

- Qual modelo estimado?

Exemplo (continuação)

- Qual modelo estimado?

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3.22 + 0.60(d_i)\,.$$

- Qual interpretação do β_1?

Exemplo (continuação)

- Qual modelo estimado?

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3.22 + 0.60(d_i)\ .$$

- Qual interpretação do β_1?
- Vamos tirar a exponencial dos dois lados

$$\left(\frac{\pi_i}{1-\pi_i}\right) =$$

Exemplo (continuação)

- Qual modelo estimado?

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3.22 + 0.60(d_i)\ .$$

- Qual interpretação do β_1?
- Vamos tirar a exponencial dos dois lados

$$\left(\frac{\pi_i}{1-\pi_i}\right) = e^{-3,22+0,60(d_i)} = e^{-3,22} * e^{0,60(d_i)}$$

- O que acontece se aumentarmos d_i em uma unidade

$$\left(\frac{\pi_i}{1-\pi_i}\right) =$$

Exemplo (continuação)

- Qual modelo estimado?

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3.22 + 0.60(d_i)\ .$$

- Qual interpretação do β_1?
- Vamos tirar a exponencial dos dois lados

$$\left(\frac{\pi_i}{1-\pi_i}\right) = e^{-3{,}22+0{,}60(d_i)} = e^{-3{,}22} * e^{0{,}60(d_i)}$$

- O que acontece se aumentarmos d_i em uma unidade

$$\left(\frac{\pi_i}{1-\pi_i}\right) = e^{-3{,}22} * e^{0{,}60(d_i+1)} = e^{-3{,}22} * e^{0{,}60(d_i)} e^{0{,}60}$$

a chance fica multiplicada por $e^{0{,}6} = 1{,}82$.

Exemplo (continuação)

- Isso equivale a aumentar

Exemplo (continuação)

- Isso equivale a aumentar 82%.
- O termo
$$\left(\frac{\pi_i}{1-\pi_i}\right)$$
é denominado **chances** (odds) e mede o quanto o sucesso é mais provável que o fracasso.
- Exemplo se a razão é 5, significa que a probabilidade de sucesso é 5 vezes maior que a probabilidade de fracasso.
- Conclusão: para cada aumento em uma unidade da dose, a chance é multiplicada por e^{β_1} que nesse caso equivale a aumentar 82%.
- Esse interpretação só é possível na ligação canônica.

Exemplo (continuação)

- Vamos encontrar agora os valores estimados das doses letais.
- Lembre que:
  - LD_{50} dose tal que 50% dos insetos são mortos;
  - LD_{90} dose tal que 90% dos insetos são mortos.
- Vimos que o modelo estimado é dado por

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3,22 + 0,60(d_i)\ .$$

- Vamos isolar d_i

$$d_i = \left(\log\left(\frac{\pi_i}{1-\pi_i}\right) + 3.22\right) / 0,60$$

Exemplo (continuação)

- LD_{50} corresponde ao valor de d_i tal que $p =$

Exemplo (continuação)

- LD_{50} corresponde ao valor de d_i tal que $p = 50\%$;
- LD_{90} corresponde ao valor de d_i tal que $p =$

Exemplo (continuação)

- LD_{50} corresponde ao valor de d_i tal que $p = 50\%$;
- LD_{90} corresponde ao valor de d_i tal que $p = 90\%$.
- Portanto

$$LD_{50} =$$

Exemplo (continuação)

- LD_{50} corresponde ao valor de d_i tal que $p = 50\%$;
- LD_{90} corresponde ao valor de d_i tal que $p = 90\%$.
- Portanto

$$LD_{50} = \left(\log\left(\frac{0,5}{1-0,50}\right) + 3.22\right)/0,60 = 5,37$$

$$LD_{90} =$$

Exemplo (continuação)

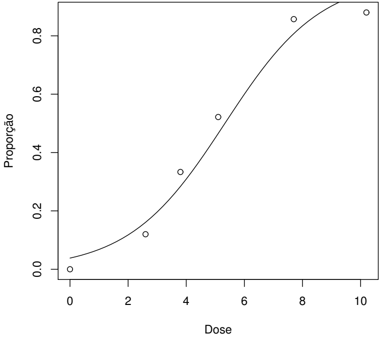
- LD_{50} corresponde ao valor de d_i tal que $p = 50\%$;
- LD_{90} corresponde ao valor de d_i tal que $p = 90\%$.
- Portanto

$$LD_{50} = \left(\log\left(\frac{0,5}{1-0,50}\right) + 3.22\right)/0,60 = 5,37$$

$$LD_{90} = \left(\log\left(\frac{0,9}{1-0,90}\right) + 3.22\right)/0,60 = 9,03$$

Exemplo (continuação)

- A figura a seguir mostra o gráfico de dispersão dos dados com a curva ajustada sobreposta.



Exemplo (continuação)

- Como podemos verificar se o modelo está bem ajustado?

Exemplo (continuação)

- Como podemos verificar se o modelo está bem ajustado? Deviance.
- O seguinte comando retorna a Deviance do modelo

```
> modelo$deviance
[1] 10.25832
```

- Com qual distribuição de referência devemos com para esse valor?

Exemplo (continuação)

- Como podemos verificar se o modelo está bem ajustado? Deviance.
- O seguinte comando retorna a Deviance do modelo

```
> modelo$deviance
[1] 10.25832
```

- Com qual distribuição de referência devemos com para esse valor?
- Com uma distribuição χ^2_{n-p}, no nosso caso χ^2_4.

Exemplo (continuação)

- Rejeitamos H_0 para valores

Exemplo (continuação)

- Rejeitamos H_0 para valores altos da deviance.
- Portatno o p-valor é dado por

Exemplo (continuação)

- Rejeitamos H_0 para valores altos da deviance.
- Portatno o p-valor é dado por

$$P(\chi_4^2 \geq 10.25832)$$

```
1-pchisq(10.25,4)
[1] 0.03642058
```

- Conclusão:

Exemplo (continuação)

- Rejeitamos H_0 para valores altos da deviance.
- Portatno o p-valor é dado por

$$P(\chi_4^2 \geq 10.25832)$$

```
1-pchisq(10.25,4)
[1] 0.03642058
```

- Conclusão: rejeitamos H_0 e concluímos que o modelo não está bem ajustado.

Exemplo (continuação)

- Queremos agora verificar se, de fato, a dose é significativa para explicar a resposta.
- Isso equivale a comparar os modelos:
  - M_0:

Exemplo (continuação)

- Queremos agora verificar se, de fato, a dose é significativa para explicar a resposta.
- Isso equivale a comparar os modelos:
  - M_0: $\eta_i = \beta_0$ (modelo nulo, só com intercepto);
  - M_1:

Exemplo (continuação)

- Queremos agora verificar se, de fato, a dose é significativa para explicar a resposta.
- Isso equivale a comparar os modelos:
  - M_0: $\eta_i = \beta_0$ (modelo nulo, só com intercepto);
  - M_1: $\eta_i = \beta_0 + \beta_1 d_i$.
- Como $n = 6$ os graus de liberdade dos modelos são:
  - M_0:

Exemplo (continuação)

- Queremos agora verificar se, de fato, a dose é significativa para explicar a resposta.
- Isso equivale a comparar os modelos:
    - M_0: $\eta_i = \beta_0$ (modelo nulo, só com intercepto);
    - M_1: $\eta_i = \beta_0 + \beta_1 d_i$.
- Como $n = 6$ os graus de liberdade dos modelos são:
    - M_0: $n - p = 6 - 1 = 5$;
    - M_1:

Exemplo (continuação)

- Queremos agora verificar se, de fato, a dose é significativa para explicar a resposta.
- Isso equivale a comparar os modelos:
  - M_0: $\eta_i = \beta_0$ (modelo nulo, só com intercepto);
  - M_1: $\eta_i = \beta_0 + \beta_1 d_i$.
- Como $n = 6$ os graus de liberdade dos modelos são:
  - M_0: $n - p = 6 - 1 = 5$;
  - M_1: $n - p = 6 - 2 = 4$.

Exemplo (continuação)

- A tabela a seguir mostra a Deviance e os graus de liberdade para cada um dos modelos:

Modelo	Graus de Liberada	Deviance
$\eta_i = \beta_0$	5	163,74
$\eta_i = \beta_0 + \beta_1 d_i$	4	10,26

- A diferença entre as Deviances é dada por

$$\Delta D = 163,74 - 10,26 = 153,48 \,.$$

- Sabemos que

$$\delta D \sim$$

Exemplo (continuação)

- A tabela a seguir mostra a Deviance e os graus de liberdade para cada um dos modelos:

Modelo	Graus de Liberada	Deviance
$\eta_i = \beta_0$	5	163,74
$\eta_i = \beta_0 + \beta_1 d_i$	4	10,26

- A diferença entre as Deviances é dada por

$$\Delta D = 163,74 - 10,26 = 153,48\,.$$

- Sabemos que

$$\delta D \sim \chi^2_1\,.$$

- Rejeitamos H_0 para valores grande ou pequenos de ΔD?

Exemplo (continuação)

- A tabela a seguir mostra a Deviance e os graus de liberdade para cada um dos modelos:

Modelo	Graus de Liberada	Deviance
$\eta_i = \beta_0$	5	163,74
$\eta_i = \beta_0 + \beta_1 d_i$	4	10,26

- A diferença entre as Deviances é dada por

$$\Delta D = 163,74 - 10,26 = 153,48\,.$$

- Sabemos que

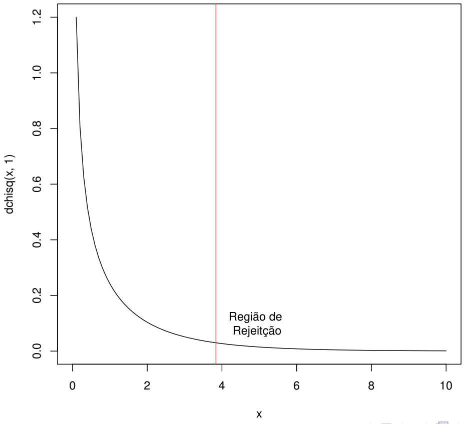
$$\delta D \sim \chi^2_1\,.$$

- Rejeitamos H_0 para valores grande ou pequenos de ΔD? Grandes.
- A região crítica é do tipo
  - $\Delta D < \chi^2_c \Rightarrow$ não rejeitamos H_0 e ficamos com o modelo M_0;
  - $\Delta D > \chi^2_c \Rightarrow$ rejeitamos H_0 e ficamos com o modelo M_1.

Exemplo (continuação)

- O valor crítico da χ^2_1 é 3,84, pois

$$P(\chi^2_1 > 3,84) = 0,05\,.$$



Exemplo (continuação)

- Conclusão:
  - $\Delta D = 153,48 > 3.84 \Rightarrow$ rejeitamos H_0;
  - isso singifica que a variável explicativa deve entrar no modelo;
  - o ganho ao acrescentar essa variável é expressivo.
- Poderíamos testar a inclusão de mais variáveis no modelo.