

CE225 - Modelos Lineares Generalizados

Cesar Augusto Taconeli

11 de setembro, 2018

Aula 8 - Análise de deviances: comparação e avaliação de modelos

Análise de deviance

- **Modelo nulo:** é o modelo mais simples possível, contendo apenas intercepto ($g(\mu_i) = \beta_0$), tal que $\hat{\mu}_i = \hat{\mu} = g^{-1}(\hat{\beta}_0)$, ou seja, atribui igual média a todas as observações;
- **Modelo saturado:** é o modelo em que se assume um parâmetro por observação, tal que $\hat{\mu}_i = y_i$, sendo o modelo mais geral em que os dados são perfeitamente ajustados;
- **Modelo proposto:** qualquer modelo intermediário entre o nulo e o saturado.
- Embora o modelo saturado e o modelo nulo não sejam de interesse prático, ambos servem como base para avaliação e comparação de modelos propostos.

- A **deviance escalonada** de um modelo proposto é definida por:

$$S(\mathbf{y}, \hat{\boldsymbol{\mu}}) = -2[l(\hat{\boldsymbol{\mu}}; \mathbf{y}) - l(\mathbf{y}; \mathbf{y})]. \quad (1)$$

- Resgatando a log-verossimilhança para a família exponencial:

$$l(\boldsymbol{\theta}; \mathbf{y}) = \sum_{i=1}^n \left[\frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i; \phi) \right], \quad (2)$$

temos:

$$S(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2 \sum_{i=1}^n \frac{[y_i \tilde{\theta}_i - b(\tilde{\theta}_i)]}{a(\phi)} - 2 \sum_{i=1}^n \frac{[y_i \hat{\theta}_i - b(\hat{\theta}_i)]}{a(\phi)}, \quad (3)$$

em que $\tilde{\theta}_i$ e $\hat{\theta}_i$ são as estimativas de θ_i sob os modelos saturado e proposto, respectivamente.

Análise de deviance

- Caso mais geral, quando $a(\phi) = \phi/\omega_i$, temos:

$$S(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2 \sum_{i=1}^n \frac{\omega_i \left[y_i (\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right]}{\phi} = \frac{D(\mathbf{y}; \hat{\boldsymbol{\mu}})}{\phi}, \quad (4)$$

em que $D(\mathbf{y}; \hat{\boldsymbol{\mu}})$ é denominada **deviance residual**, ou simplesmente **deviance**.

- Uma vez que $l(\mathbf{y}, \hat{\boldsymbol{\mu}}) \leq l(\mathbf{y}, \mathbf{y})$, $D(\mathbf{y}; \hat{\boldsymbol{\mu}}) \geq 0$, de forma que, quanto pior o ajuste do modelo proposto, maior a deviance.

Tabela 1: Deviances para alguns MLGs

Distribuição	Deviance
Normal	$\sum_{i=1}^n (y_i - \hat{\mu}_i)^2$
Poisson	$2 \sum_{i=1}^n \left[y_i \ln \left(\frac{y_i}{\hat{\mu}_i} \right) + (\hat{\mu}_i - y_i) \right]$
Binomial	$2 \sum_{i=1}^n \left[y_i \ln \left(\frac{y_i}{m_i \hat{\mu}_i} \right) + (m_i - y_i) \ln \left\{ \frac{\left(1 - \frac{y_i}{m_i}\right)}{(1 - \hat{\mu}_i)} \right\} \right]$
Gama	$2 \sum_{i=1}^n \left[\ln \left(\frac{\hat{\mu}_i}{y_i} \right) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} \right]$
Normal inversa	$\sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{y_i \hat{\mu}_i^2}$

Análise de deviance e teste da razão de verossimilhanças

- Considere dois modelos propostos M_0 e M_1 , em que M_0 é um caso particular de M_1 (obtido por alguma restrição nos parâmetros de M_1 , usualmente fixando em zero alguns dos parâmetros de M_1).
- Considere $\phi = 1$. O teste da razão de verossimilhança aplicado à hipótese nula (de que a restrição aplicada em M_0 é válida) fica definido pela estatística:

$$\begin{aligned} & -2[l(\hat{\mu}_0; \mathbf{y}) - l(\hat{\mu}_1; \mathbf{y})] = \\ & -2[l(\hat{\mu}_0; \mathbf{y}) - l(\mathbf{y}; \mathbf{y})] - \{-2[l(\hat{\mu}_1; \mathbf{y}) - l(\mathbf{y}; \mathbf{y})]\} = \\ & D(\mathbf{y}; \hat{\mu}_0) - D(\mathbf{y}; \hat{\mu}_1), \end{aligned}$$

em que $l(\hat{\mu}_0; \mathbf{y})$ e $l(\hat{\mu}_1; \mathbf{y})$ são as log-verossimilhanças maximizadas sob os modelos restrito (M_0) e irrestrito (M_1).

Análise de deviance e teste da razão de verossimilhanças

- Sob a hipótese nula (referente a M_0), a diferença das deviances (estatística do TRV) tem distribuição (assintótica) χ^2 com $p_1 - p_0$ graus de liberdade, em que p_1 e p_0 ($p_0 < p_1$) referem-se aos números de parâmetros estimados em M_1 e M_0 .

Análise de deviance - parâmetro de dispersão desconhecido

- Se ϕ é desconhecido, então deve-se obter uma estimativa consistente ($\hat{\phi}$) que pode ser baseada no modelo irrestrito M_1 .
- Neste caso, a comparação de M_0 e M_1 baseia-se na estatística:

$$F = \frac{(D(\mathbf{y}; \hat{\mu}_0) - D(\mathbf{y}; \hat{\mu}_1)) / (p_1 - p_0)}{\hat{\phi}}, \quad (5)$$

que, sob a hipótese nula (referente a M_0), tem distribuição F-Snedecor com $p_1 - p_0$ e $n - p_1$ graus de liberdade.

Análise de deviance - parâmetro de dispersão desconhecido

- Caso se esteja testando uma sequência de modelos encaixados, então recomenda-se aplicar, em todos os testes, a estimativa de ϕ fornecida pelo modelo com maior número de termos dentre os considerados.
- Neste caso, sendo p_{max} o número de parâmetros estimados no modelo com mais termos, a distribuição de referência para testar a hipótese nula de equivalência de M_1 e M_0 é a F-Snedecor com $p_1 - p_0$ e $n - p_{max}$ graus de liberdade.

Análise de deviance

- A análise de deviance é uma generalização da análise de variância aplicada a uma sequência de MLGs encaixados (obtidos sequencialmente impondo sucessivas restrições aos parâmetros do modelo original).
- A cada passo, são acrescentados (ou excluídos) do modelo efeitos de variáveis explicativas, fatores, de interações. . .
- Numa tabela, apresenta-se a sequência de modelos ajustados, as correspondentes deviances, as diferenças entre deviances e os resultados dos testes associados.
- Para uma sequência de modelos encaixados $M_{p_1}, M_{p_2}, \dots, M_{p_r}$, com mesmas distribuição e função de ligação, e com dimensões $p_1 < p_2 < \dots < p_r$, então temos as deviances satisfazendo $D_{p_1} > D_{p_2} > \dots > D_{p_r}$.

Análise de deviance

- Se o parâmetro de dispersão é conhecido, o teste da qualidade de ajuste de um MLG com p parâmetros estimados pode ser feito com base na deviance.
- Sob a hipótese nula, de que o modelo se ajusta bem aos dados, $D(\mathbf{y}; \hat{\boldsymbol{\mu}})$ tem distribuição χ^2 com $n - p$ graus de liberdade.
- Assim, o modelo proposto é rejeitado, para um nível de significância α , se $D(\mathbf{y}; \hat{\boldsymbol{\mu}}) > \chi^2_{n-p}(1 - \alpha)$.

- Uma alternativa ao teste de qualidade de ajuste baseado na deviance é o teste de Pearson, baseado na seguinte estatística:

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\widehat{Var}(y_i)}, \quad (6)$$

que também tem distribuição assintótica χ^2 com $n - p$ graus de liberdade se o modelo proposto é correto e o parâmetro de dispersão conhecido.

Nota: A aproximação χ^2 para os testes de qualidade de ajuste nem sempre funciona bem, ainda que para grandes amostras. Vamos avaliar esse fato através de simulação.