CE225 - Modelos Lineares Generalizados

Cesar Augusto Taconeli

11 de julho, 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 seja inviável e o modelo nulo não seja de interesse, ambos servem como base para avaliação e comparação de modelos propostos.

Análise de deviance

- A **deviance escalonada** de um modelo proposto é definida como a estatística da razão de verossimilhança relativa ao modelo saturado:

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

- Resgatando a log-verossimilhança para a família exponencial de dispersão:

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

temos:

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

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:

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

em que $D(\boldsymbol{y}; \hat{\boldsymbol{\mu}})$ é a deviance do modelo proposto.

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

Análise de deviance

Tabela 1: Deviances para alguns modelos mais usuais

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).
- Considerando $\phi = 1$, o teste da razão de verosimilhança aplicado à hipótese nula (de que a restrição aplicada em M_0 é válida) fica definido pela estatística:

$$-2[l(\hat{\boldsymbol{\mu}}_{\mathbf{0}}; \boldsymbol{y}) - l(\hat{\boldsymbol{\mu}}_{\mathbf{1}}; \boldsymbol{y})] =$$

$$-2[l(\hat{\boldsymbol{\mu}}_{\mathbf{0}}; \boldsymbol{y}) - l(\boldsymbol{y}; \boldsymbol{y})] - \{-2[l(\hat{\boldsymbol{\mu}}_{\mathbf{1}}; \boldsymbol{y}) - l(\boldsymbol{y}; \boldsymbol{y})]\} =$$

$$D(\boldsymbol{y}; \hat{\boldsymbol{\mu}}_{\mathbf{0}}) - D(\boldsymbol{y}; \hat{\boldsymbol{\mu}}_{\mathbf{1}}),$$

em que $l(\hat{\boldsymbol{\mu}}_{\mathbf{0}}; \boldsymbol{y})$ e $l(\hat{\boldsymbol{\mu}}_{\mathbf{1}}; \boldsymbol{y})$ são as log-verossimilhanças maximizadas sob os modelos restrito (M_0) e irrestrito M_1.

Análise de deviance e teste da razão de verossimilhanças

- Essa estatística assume maiores valores a medida que o modelo restrito (M_0) proporciona pior ajuste do que M_1.
- Sob a hipótese nula, a diferença das deviances (estatística do TRV) tem distribuição (assintótica) qui-quadrado com $p_1 - p_0$ graus de liberdade, em que p_1 e p_0 ($p_0 < p_1$) são os 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 deve ser baseada ns estatística:

$$F = \frac{(D(\boldsymbol{y}; \hat{\boldsymbol{\mu}}_{\mathbf{0}}) - D(\boldsymbol{y}; \hat{\boldsymbol{\mu}}_{\mathbf{1}}))/(p_1 - p_0)}{\hat{\phi}}, \tag{5}$$

que, sob a hipótese nula (que M_0 equivale a M_1), tem distribuição F 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 deve-se usar, em todos os testes, a estimativa de ϕ fornecida pelo modelo com mais 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 hipótese nula de equivalência de M_1 e M_0 usamos a distribuição F 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 modelos encaixados (obtidos sequencialmente impondo sucessivas restrições aos parâmetros do modelo original).
- A cada passo, são acrescentados efeitos de variáveis explicativas, fatores e suas interações.
- Numa tabela, apresenta-se a sequência de modelos ajustados, as correspondentes deviances, as diferenças entre deviances e os testes associados.
- A análise de deviance garante que, para uma sequência de modelos encaixados $M_{p_1}, M_{p_2}, ..., M_{p_r}$, com mesmas distribuição e função de ligação, e com dimensões $p_1 < p_2 < ... < p_r$, então temos as deviances satisfazendo $D_{p_1} > D_{p_2} > ... > D_{p_r}$.

Análise de deviance

- Se o parâmetro de dispersão é conhecido, o teste da qualidade de ajuste de um modelo proposto com p parâmetros pode ser feito com base na deviance.

- Sob a hipótese nula, de que o modelo proposto é correto, $D(\boldsymbol{y}; \hat{\boldsymbol{\mu}})$ tem distribuição qui-quadrado com $n - p$ graus de liberdade.

- Assim, o modelo proposto é rejeitado, para um nível de significância α, se $D(\boldsymbol{y}; \hat{\boldsymbol{\mu}}) > \chi^2_{n-p}(1-\alpha)$.

Análise de deviance

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

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

que tem distribuição assintótica qui-quadrado com $n - p$ graus de liberdade se o modelo proposto é correto e o parâmetro de dispersão conhecido.

Nota: A aproximação qui-quadrado para os testes de qualidade de ajuste é restritiva não apenas por ser um resultado assintótico, mas também por não funcionar bem quando a distribuição de y apresenta elevada dispersão.