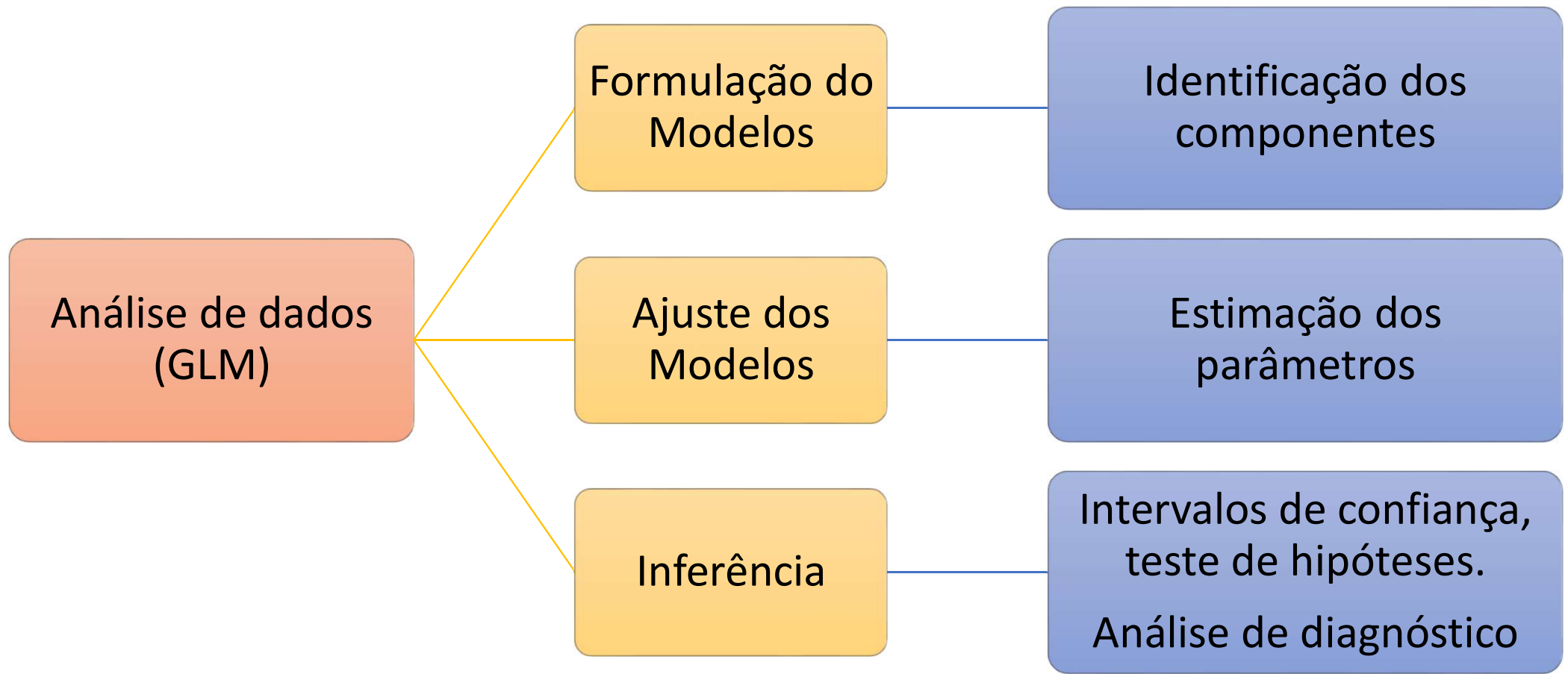


Medidas de Discrepância

Função desvio (Deviance)

Ajuste dos GLM's



Medida de discrepância nos GLM's

Para compreender análise de resíduos e os diagnósticos

Medida de discrepância no contexto dos GLM's

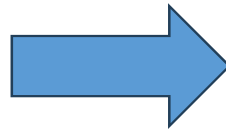
Função desvio
(Deviance)

Estatística de
Pearson

Medidas de Discrepância

O ajuste de um GLM a um conjunto de observações y pode ser considerado como uma substituição das observações (y) por valores estimados ($\hat{\mu}$).

O quanto $\hat{\mu}$ e y se diferem?



Medidas de discrepância

Admitindo-se uma **combinação satisfatória da distribuição de Y e da função de ligação**, o objetivo é determinar quantos termos são necessários na **estrutura linear** para uma descrição dos dados

Medidas de Discrepância

Quanto maior o **número de variáveis explanatórias** pode conduzir que melhor seria a explicação dos dados daquele modelo.

Porém, há um **aumento de complexidade na interpretação**.

Por outro lado, um modelo com poucas variáveis é facilmente interpretável. Mas pode não explicar bem a variabilidade dos dados.

Modelo intermediário, entre um modelo muito complexo (muito) explicativo e um modelo simples(pouco explicativo), mas muito interpretável.

Quantas variáveis na Estrutura Linear?

Pela teoria estatística um modelo pode ter até n parâmetros, se ele tem n observações, logo podemos definir:

- Modelo Nulo : é o modelo mais simples, contém apenas intercepto, isto é $g(\mu_i) = \beta_0$ 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 $\widetilde{\mu}_1 = y_i$, sendo o modelo mais geral em que os dados são perfeitamente ajustados.
- **Modelo Proposto: Modelo intermediário entre o nulo e o saturado.**

Função Desvio (ou Deviance)

O problema de seleção de variáveis explanatórias é determinar a **utilidade de um parâmetro extra no modelo proposto** ou então, verificar a falta de ajuste induzida pela sua omissão.

O objetivo de discriminar entre modelos alternativos, medidas de discrepância para medir o ajuste de um modelo.

Nelder e Wedderburn (1972) propuseram uma medida de discrepância para medir o ajuste de um modelo, a *deviance* escalonada (ou desvio escalonado):

$$S_p = 2 \cdot (\widehat{\ell}_n - \widehat{\ell}_p)$$

Em que $\widehat{\ell}_n$ e $\widehat{\ell}_p$ são os máximos da função log-verossimilhança para os modelos saturado e proposto, respectivamente.

Função Desvio (ou Deviance)

Assim, para os modelos saturado e proposto temos que:

$$\widehat{\ell}_n = \phi^{-1} \sum_{i=1}^n [y_i \tilde{\theta}_i - b(\tilde{\theta}_i)] + \sum_{i=1}^n c(y_i, \phi)$$

e

$$\widehat{\ell}_p = \phi^{-1} \sum_{i=1}^n [y_i \hat{\theta}_i - b(\hat{\theta}_i)] + \sum_{i=1}^n c(y_i, \phi)$$

Em que $\hat{\theta}_i = q(\hat{\mu}_i)$ e $\tilde{\theta}_i = q(y_i)$ são as EMV's do parâmetro canônico sob os modelos saturado e proposto

Função Desvio (ou Deviance)

Então, tem-se que

$$S_p = \phi^{-1} \left\{ 2 \sum_{i=1}^n [y_i(\tilde{\theta}_i - \hat{\theta}_i)] + b(\hat{\theta}_i) - b(\tilde{\theta}_i) \right\} = \phi^{-1} D_p$$

Em que

$$D_p = 2 \sum_{i=1}^n [y_i(\tilde{\theta}_i - \hat{\theta}_i)] + b(\hat{\theta}_i) - b(\tilde{\theta}_i)$$

É denominada deviance residual, ou simplesmente deviance

Observações

1 – A *deviance* pode ser computada para qualquer GLM a partir da EMV de μ , isto é, $\hat{\mu} = g^{-1}(X\hat{\beta})$.

2- Uma vez que $\widehat{\ell}_p \leq \widehat{\ell}_n$, isto é, $D_p \geq 0$, de forma que, quando pior o ajuste do modelo proposto maior é a *deviance*. Logo, medida que as variáveis explanatórias entram no componente sistemático a *deviance* decresce até se tornar zero para o modelo saturado.

Observações

3- Para o teste, define-se o número de graus de liberdade da *deviance* do modelo por $\nu = n - p$, isto é, como o número de observações menos o posto da matriz do modelo proposto

4 – Em alguns casos especiais como nos modelos normal e log-linear, a *deviance* iguala-se a estatísticas comumente usadas nos testes de ajuste.



PUC Minas
Virtual