

Modelos Log-lineares

Tabelas tridimensionais

Maria Teresa Leão Costa

UnB – IE Departamento de Estatística

Análise de Dados Categorizados



- Modelos log-lineares podem representar vários padrões de independência e associação.
- Termos de associação de dois fatores descrevem a odds ratio condicional entre as variáveis.

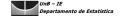


Modelos Log-lineares para Tabelas Tridimensionais

- Considere uma tabela de contingência tridimensional contendo a classificação cruzada das variáveis X, Y e Z.
 - tabelas parciais distribuição de X e Y para cada nível de Z, por exemplo.
 - tabelas marginais
- Designando as probabilidades conjuntas das células por:

$$\pi_{ijk}$$
, $i=1,...,I$, $j=1,...,J$, $k=1,...,K$ onde:

$$\sum_{i} \sum_{j} \sum_{k} \pi_{ijk} = 1$$





Tipos de Independência

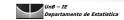
Independência Mútua:

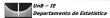
 As três variáveis são mutuamente independentes quando:

$$\pi_{ijk} = \pi_{i++}\pi_{+j+}\pi_{++k}, para\ todo\ i, j\ e\ k$$

 Na escala logarítmica independência mútua é equivalente ao seguinte modelo log-linear:

$$\log m_{iik} = \mu + \lambda_i^X + \lambda_i^Y + \lambda_k^Z$$





■ Independência Conjunta:

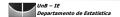
A variável Y é conjuntamente independente de X e Z quando:

$$\pi_{ijk} = \pi_{i+k}\pi_{+j+}$$
, para todo $i, j e k$

■ O modelo log-linear equivalente é dado por:

$$\log m_{iik} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ}$$

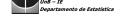
■ Definições similares aplica-se para X ser conjuntamente independente das variáveis Y e para Z ser conjuntamente independente de X e Y.



Tipos de Independências

- Se X e Y são independentes na tabela parcial para a k-ésima categoria de Z, então X e Y são condicionalmente independentes no k-ésimo nível de Z.
- Seja $\pi_{ij|k} = \pi_{ijk} / \pi_{++k}$, i = 1, ..., I, j = 1, ..., J a distribuição conjunta de X e Y no k-ésimo nível de Z. Então independência condicional no k-ésimo nível de Z é dado por:

$$\pi_{ii|k} = \pi_{i+|k}\pi_{+i|k}$$
 para todo i e j





Independência Marginal:

 As variáveis X e Y são marginalmente independentes se:

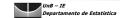
$$\pi_{ij+} = \pi_{i++} \pi_{+j+}$$
 $i = 1, 2, ..., I \quad j = i = 1, 2, ..., J$





<u>Independência</u> Condicional e Marginal

- Vimos que associação parcial pode ser bem diferente da associação marginal. Além disso, vimos que independência condicional de X e Y dado Z não implica em independência marginal de X e Y.
- Ambas as independências condicional e marginal se verificam se um dos tipos de independências estudas anteriormente se aplicam.

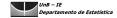




Ausência de Interação entre os Três Fatores

- Nos modelos log-lineares descritos anteriormente existem três, dois e um par de variáveis condicionalmente independentes.
- O termo λ_{ij}^{XY} indicam que as variáveis são condicionalmente dependentes.
- Um modelo que permita que todas as três variáveis sejam condicionalmente dependentes é dado por:

$$\log m_{iik} = \mu + \lambda_i^X + \lambda_i^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{ik}^{YZ} + \lambda_{ik}^{YZ}$$





Modelos Log-Lineares para Tabelas Tridimensionais - Modelos Hierárquicos

- Os preditores lineares para os modelos tem uma estrutura análoga aos modelos de fatoriais de análise de variância.
- O termo com índice simples são análogos ao efeitos principais, os termos com duplo índice são análogos a interação entre dois fatores.





Modelos Log-Lineares para Tabelas Tridimensionais - Modelos Hierárquicos

 A soma dos parâmetros para qualquer índice é igual a zero. Isto é,

$$\sum_{i} \lambda_{i}^{X} = \sum_{i} \lambda_{j}^{Y} = \sum_{k} \lambda_{k}^{Z} = \sum_{i} \lambda_{ij}^{XY} = \sum_{i} \lambda_{ij}^{XY} = \cdots = \sum_{k} \lambda_{ijk}^{XYZ} = 0$$

 O modelo log-linear geral para uma tabela tridimensional é:

$$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ} + \lambda_{ik}^{XZ} + \lambda_{ijk}^{XYZ}$$





Modelos Log-Lineares para Tabelas Tridimensionais - Modelos Hierárquicos

 Alguns modelos log-lineares para tabelas tridimensionais

Modelo Log – linear Símbolo

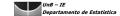
$$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z \tag{X,Y,Z}$$

$$\log m_{iik} = \mu + \lambda_i^X + \lambda_i^Y + \lambda_k^Z + \lambda_{ii}^{XY}$$
 (XY, Z)

$$\log m_{iik} = \mu + \lambda_i^X + \lambda_i^Y + \lambda_k^Z + \lambda_{ii}^{XY} + \lambda_{ik}^{YZ}$$
 (XY, YZ)

$$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ} + \lambda_{ik}^{XZ} \qquad (XY, YZ, XZ)$$

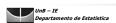
$$\log m_{ijk} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ} + \lambda_{ik}^{XZ} + \lambda_{ijk}^{XYZ} \quad (XYZ)$$





Modelos Log-Lineares para Tabelas Tridimensionais - Modelos Hierárquicos

Os modelos log-lineares apresentados na tabela anterior são chamados de modelos hierárquicos. Isto significa se o modelo contém efeitos de ordem maior ele deve também conter efeitos de ordem menor. Por exemplo, quando o modelo contém λ_{ij}^{xy} ele deve conter λ_i^x e λ_j^y.





Modelos Log-Lineares para Tabelas Tridimensionais - Interpretação dos Parâmetros

 Para interpretar modelos log-lineares, descreveremos as associações marginais e parciais usando as razões de chances. A tabela marginal {π_{ij+}} usa um conjunto de (I-1) (J-1) razões de chances tal que.

$$\theta_{ij}^{XY} = \frac{\pi_{ij+}\pi_{i+1,j+1,+}}{\pi_{i+1,j+}\pi_{i,j+1,+}}, 1 \le i \le I-1, 1 \le j \le J-1$$

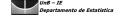




Modelos Log-Lineares para Tabelas Tridimensionais - Interpretação dos Parâmetros

 Para interpretar modelos log-lineares, descreveremos as associações marginais e parciais usando as razões de chances. A tabela marginal {π_{ij+}} usa um conjunto de (I-1) (J-1) razões de chances tal que.

$$\theta_{ij}^{XY} = \frac{\pi_{ij+}\pi_{i+1,j+1,+}}{\pi_{i+1,j+}\pi_{i,j+1,+}}, 1 \le i \le I-1, 1 \le j \le J-1$$



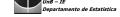


Tridimensionais - Interpretação dos Parâmetros

Para um nível fixo k de Z, a razão de chances correspondente é dada por:

$$\theta_{ij(k)}^{XY} = \frac{\pi_{ijk}\pi_{i+1,j+1,k}}{\pi_{i,i+1,k}\pi_{i+1,i,k}}, 1 \le i \le I-1, 1 \le j \le J-1$$

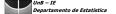
Descreve a associação condicional X e Y.
 Similarmente, as associações condicionais Xe Z é descrita por (I-1)(K-1) razões de chances {θ_{i(j)k}} para cada nível de J de Y

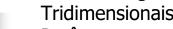


Tridimensionais - Interpretação dos Parâmetros

- E as associações condicionais Y e Z é descrita por (J-1)(K-1) razões de chances $\{\theta_{(i)jk}\}$ para cada nível de I de X.
- Os parâmetros do modelo log-linear são funções das razões de chances condicionais. As relações são mais simples para uma variável binária. Para tabelas 2x2x2, substituindo-se o modelo saturado na razão de chances condicional obteremos:

$$\lambda_{111}^{XYZ} = \frac{1}{8} \log \left(\frac{\theta_{11(1)}}{\theta_{11(2)}} \right) = \frac{1}{8} \log \left(\frac{\theta_{1(1)1}}{\theta_{1(2)1}} \right) = \frac{1}{8} \log \left(\frac{\theta_{(1)11}}{\theta_{(2)11}} \right)$$





Tridimensionais - Interpretação dos Parâmetros

• Para a restrição de soma zero sobre $\{\lambda_{ijk}^{XYZ}\}$. Cada λ_{ijk}^{XYZ} é igual a zero quando a razão de chances entre as duas variáveis é a mesma para cada nível da terceira variável. E portanto,

$$\lambda_{11}^{XY} = \frac{1}{4} \log \theta_{11(k)}$$

