Modelagem Bayesiana e Aplicações

Márcia D'Elia Branco

Universidade de São Paulo
Instituto de Matemática e Estatística
http:www.ime.usp.br/m̃branco

Modelo binário assimétrico e TRI

Regressão binária: dados aumentados

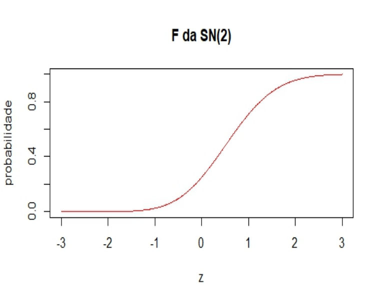
- Modelos de regressão binários, ordinais e multinomiais podem ser simplificados usando um conjunto de variáveis latentes contínuas que permitam a implementação do algoritmo de dados aumentados.
- Conforme já discutido em exemplo da aula8 desta disciplina, para a ligação probito definimos

$$z_i = x_i^t \beta + u_i \quad u_i \sim N(0,1) \;\; ind..$$

- Considere $y_i = I(z_i > 0)$, então y_i segue uma Bernoulli com probabilidade $p_i = \Phi(x_i^t \beta)$.
- Este esquema pode ser usado para outras ligações substituindo a distribuição dos erros u_i de forma adequada. Na regressão logito, usar a distribuição logística.

Ligação assimétrica

- A substituição da suposição de normalidade para os erros pela distribuição normal-assimétrica produz uma função de ligação assimétrica.
- Para curvas assimetrica a velocidade que a probabilidade se aproxima do zero não é a mesma da aproximação do 1.



Ligação assimétrica

- Essa ligação foi proposta em Bazán, Branco e Bolfarine em *Bayesian Analysis* (2006), no contexto da Teoria da Resposta ao Item.
- Uma variável aleatória U tem distribuição normal-assimétrica padrão se sua fdp é dada por

$$f(u) = 2\phi(u)\Phi(\lambda u) \qquad -\infty < \lambda < \infty.$$

- O parâmetro λ esta associado a assimetria da distribuição. $\lambda > 0 (\lambda < 0)$ representa assimetria positiva (negativa).
- Se $\lambda = 0$ resulta no modelo normal.
- A ligação foi denominda *skew-probito* e tem o modelo probito como um caso especial.

Ligação assimétrica

- Uma importante característica da distribuição normal-assimétrica é sua representação hierárquica:

$$U \mid V = v \sim N(\delta v, 1 - \delta^2)$$

$$V \sim N(0,1)I(0,1)$$

- Em que δ é uma reparametrização de λ restrita ao intervalo $(-1,1)$ e dada por

$$\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$$

- Uma propriedade útil para construção do ligação é

$$P(U \leq u \mid \lambda) = 1 - P(U \leq -u \mid -\lambda).$$

- Assim, o modelo *skew-probito* é contruído considerando $y_i = I(z_i > 0)$ com

$$z_i = x_i^t \beta + u_i \quad u_i \sim SN(0, 1, -\lambda) \ \ ind..$$

- Note que

$$P(y_i = 1) = P(z_i > 0) = P(u_i > -x_i^t \beta) = 1 - P(u_i \leq -x_i^t \beta)$$

- Pela propriedade $p_i = F(x_i^t \beta \mid \lambda)$ em que F é a f.d.a de uma $SN(0, 1, \lambda)$.

Ligação assimétrica

- Para implementação computacional é conveniente considerar a representação hierárquica da SN.
- Assim

$$
\begin{aligned}
z_i \mid v_i, \beta, \delta &\sim N(x_i^t\beta - \delta v_i, 1 - \delta^2) \\
v_i &\sim N(0,1)I(0,1)
\end{aligned}
$$

- Condicional a v_i temos um modelo similar ao modelo probito.
- Similar ao que foi feito para o probito, temos as seguintes condicionais completa

$$z_i \mid v_i, \beta, \delta, y_i = 1 \sim NT(x_i^t\beta - \delta v_i, 1 - \delta^2)I_{(z_i > 0)}$$

$$z_i \mid v_i, \beta, \delta, y_i = 0 \sim NT(x_i^t\beta - \delta v_i, 1 - \delta^2)I_{(z_i < 0)}$$

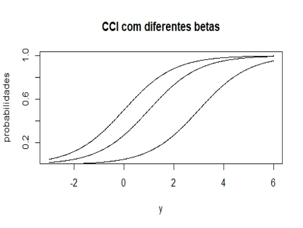
Modelos da TRI

- A TRI pode ser vista como um caso especial de Análise de Traço Latente.
- Foi introduzida nos anos 50 na área de psicometria, como alternativa a Teoria Classica dos Testes.
- Considere um Teste com J itens, respondido por I indivíduos.
- Seja $x_{ij} = 1$ se o i-ésimo indivíduo responde corretamente ao item j e $x_{ij} = 0$ caso contrário.
- Definimos $p_{ij} = P(x_{ij} = 1)$. A TRI modela essa probabilidade como função de uma quantidade latente θ_i.
- Na teoria dos testes θ_i é vista como a habilidade do i-ésimo indivíduo no conteúdo em teste.

- O modelo de Rasch ou 1PL é dado por

$$p_{ij} = \frac{\exp(\theta_i - \beta_j)}{1 + \exp(\theta_i - \beta_j)}$$

- Os parâmetros β_j são interpretados como dificuldade do item.
- Apresentamos as curvas características de 3 itens com diferentes dificuldades



Modelos da TRI

- O modelo 2PL inclui um parâmatro α_j interpretado como discriminação e é dado por

$$p_{ij} = \frac{\exp(\alpha_j(\theta_i - \beta_j))}{1 + \exp(\alpha_j(\theta_i - \beta_j))}$$

- A suposição usual sobre as variáveis latentes é $\theta_i \sim N(0,1)$ independentes.
- Para o propósito da implementação bayesiana do modelo, considera-se a reparametrização $(a_j, b_j) = (\alpha_j, -\alpha/\beta_j)$

- O modelo de 2 parâmetros considerando a ligação *skew-probito* é caracterizado de forma hierárquica por

$$z_{ij} \mid v_{ij}, a_j, b_j, \delta_j, y_i = 1 \sim N(a_j\theta_i + b_j - \delta_j v_{ij}, 1 - \delta_j^2) I(z_{ij} > 0)$$

$$z_{ij} \mid v_{ij}, a_j, b_j, \delta_j, y_i = 0 \sim N(a_j\theta_i + b_j - \delta v_{ij}, 1 - \delta_j^2) I(z_{ij} < 0)$$

$$v_{ij} \sim N(0,1) I(0,1)$$

- Para completar o modelo, especificamos distribuições a priori:

$$a_j \sim N(m_a, v_a) I(0, \infty)$$
$$b_j \sim N(0, v_b)$$
$$\delta_j \sim U(-1, 1)$$

Modelos SP na TRI

- O novo conjunto de parâmetros δ_j associados a assimetria da função de ligação, no contexto da TRI é um parâmetro de item.
- Se um item possue δ positivo (negativo) ele discrimina melhor (pior) o grupo de estudantes com menor habilidade.
- Ele recebe também o nome de parâmetro de penalização, como veremos no exemplo.
- A aplicação a seguir está no artigo Bazán et al (2014) do BJPS.

Aplicação: Teste de Matemática

- O conjunto de dados refere-se a um teste de matemática aplicada a um grupo de 974 estudantes de 4 ano de escolas rurais do Peru.
- Os alunos responderam 18 questões que foram avaliadas como certa (1) ou errada (0). O escore médio foi 8.27 e o devio padrão 4.2.
- Foram ajustados os modelos 2P-probito (PN) e 2P-skewprobito (SPN), usando a linguagem BUGS.
- As medidas de comparação de modelos consideradas são: DIC, EAIC, EBIC e SSR (soma dos quadrados dos resíduos latentes).

Aplicação: Teste de Matemática

- Todas as medidas escolhem o modelo SPN.
- Mais dois modelos foram ajustados, PSN e SNSP. Ambos substituem a suposição de normalidade para θ_i pela de normal-assimétrica.
- O item 14 foi o que apresentou o maior valor positivo para δ.
- A comparação das CCI para este item, mostra que:

  - Para estudantes com menor habilidade,
  as probabilidades de acerto são maiores no modelo simétrico.

  - Para estudantes com menor habilidade,
  as probabilidades de acerto são menores no modelo simétrico.