

# ABORDAGEM BAYESIANA NÃO PARAMÉTRICA DO PROBLEMA DE TRIAGEM

*Sandra Ramos, Antónia Amaral Turkman, Marília Antunes*

CEUAL/IPP, CEUAL/FCUL, CEUAL/FCUL

e-mail: `sfr@isep.ipp.pt`

`antonia.turkman@fc.ul.pt`

`marilia.antunes@fc.ul.pt`

**Resumo:** O procedimento de triagem envolve a construção de uma região de especificação  $C_{\mathbf{X}}$ , no espaço  $d$ -dimensional, de modo a que um indivíduo futuro com um vector de características em  $C_{\mathbf{X}}$  tenha maior probabilidade de ser identificado como um *sucesso* (a resposta  $Y$  pertence a uma região conhecida  $C_Y$ ). Na abordagem preditiva bayesiana a obtenção da região  $C_{\mathbf{X}}$  é baseada num critério óptimo assente na maximização de  $P(Y \in C_Y | \mathbf{X} \in C_{\mathbf{X}}; D)$ , restringida à classe das regiões  $C_{\mathbf{X}}$  com probabilidade preditiva de triagem  $\alpha = P(\mathbf{X} \in C_{\mathbf{X}} | D)$  fixa. Em literatura já publicada, a construção da região  $C_{\mathbf{X}}$  baseia-se habitualmente em modelos paramétricos para  $(Y, \mathbf{X})$  e ou modelos semiparamétricos mais flexíveis.

Neste trabalho, propõe-se uma abordagem totalmente não paramétrica que considera um modelo bayesiano, baseado em mistura por um processo de Dirichlet dependente para  $Y | \mathbf{X}$  e em mistura por um processo de Dirichlet para  $\mathbf{X}$ . A metodologia proposta é ilustrada através de simulações e de conjuntos de dados reais descritos na literatura. São apresentados resultados de comparações de desempenho entre a abordagem proposta e outras abordagens já existentes.

**palavras-chave:** triagem; modelação não paramétrica; abordagem bayesiana.

## Agradecimentos

Este trabalho foi parcialmente financiado pela Fundação para a Ciência e Tecnologia - FCT: Projecto *PTDC/MAT/118335/2010* e *Pest – OE/MAT/UI0006/2014*.