

Seminário em Engenharia Matemática

Sandra Ramos

Título:

Construção paramétrica da região de *screening* óptima em classificação supervisionada

Resumo:

Num problema de *screening*, para além da classificação de um novo indivíduo segundo os possíveis desfechos de uma variável categórica Y , é possível calcular um conjunto de probabilidades preditivas, chamadas características operacionais. Estas probabilidades permitem avaliar a qualidade do procedimento de *screening*, o que faz com que este seja, em muitas situações práticas, mais apelativo do que as técnicas discriminantes clássicas. O procedimento assenta na obtenção de uma região de especificação, $C_{\mathbf{X}}$, baseada num vector de características \mathbf{X} , tal que o indivíduo é seleccionado se o seu vector de especificação pertencer a essa região. A abordagem óptima da metodologia de *screening* selecciona $C_{\mathbf{X}}$ que maximiza a probabilidade de um indivíduo ser sucesso sempre que o seu vector de especificação pertencer a essa região.

Neste trabalho propõe-se uma metodologia bayesiana para a construção da região de especificação óptima em classificação supervisionada. O método proposto é baseado na observação de uma família de m pares de covariáveis, admitindo-se que a distribuição do par \mathbf{X}_i , $i=1,2,\dots,m$, condicional aos vários desfechos da variável Y é normal bivariada.

O desempenho do método proposto é avaliado numa aplicação a dados de níveis de expressão genética de pares de genes, utilizados para classificar doentes oncológicos em duas categorias da doença. Os resultados são comparados com os obtidos a partir de técnicas discriminantes clássicas.