

ANÁLISE DISCRIMINANTE

- 1) Utilizando o arquivo de dados alunos, referente à avaliação de candidatos em um teste de seleção, aplique a técnica de análise discriminante, considerando a variável result as variáveis: inglês, int_tex, raclog e redação.

total de observações: 200

3 grupos: 1 80 últimos colocados
 2 80 classificação intermediária
 3 40 primeiros colocados

- a) Teste a igualdade de média das variáveis grupos;
 - b) Qual a variável mais discrimina no primeiro modelo no modelo? Qual a justificativa?
 - c) Avalie os resultados da análise
- 2) Considere o arquivo gpa referente às variáveis GPA (graduate grade point average) e gmat (graduate management aptitude test) usadas para ajudar a decidir sobre a admissão de candidatos aos programas de pós-graduação de uma instituição de ensino. A variável dependente apresenta as seguintes categorias: 1 – admissão, 2 – não admissão, 3 – dúvida. Verifique o poder de discriminação de cada variável em uma análise de discriminante. Analise a qualidade do modelo ajustado.
 - 3) Considere o arquivo Biscobis referente a uma amostra de 100 empresas clientes de uma grande empresa que é fornecedora no setor industrial . Considere as variáveis:

Variável dependente: x14 = perfil de compra (1 = nova compra 2=primeira recompra 3=outras).

Variáveis independentes: avaliação de 0 a 10 de atributos da Biscobis:

X₁ = rapidez na entrega do produto

X₂ = nível de preço

X₃ = flexibilidade de preço

X₄ = imagem do fornecedor

X₅ = serviço como um todo

X₆ = imagem da força de vendas

X₇ = qualidade do produto

Faça o processamento da análise discriminante. Faça uma análise completa das principais informações obtidas e avalie a precisão da classificação a posteriori dos elementos do banco de dados.

- 4) Considere o arquivo `prcancer` referente a uma amostra de 53 homens. O tratamento e prognóstico de câncer depende de quanto a doença se espalhou. Uma das regiões em que o câncer pode se espalhar refere-se aos nódulos linfáticos. Se os mesmos forem atingidos, o prognóstico é geralmente mais pobre do que em caso negativo. Por isso é desejável estabelecer o quanto antes se os nódulos são cancerosos. Para certos tipos de câncer, cirurgia exploratória é feita só para determinar se os nódulos são cancerosos, uma vez que isso determinará qual o tratamento necessário. Se for possível prever se os nódulos são afetados ou não com base nos dados sem a realização de cirurgia, considerável desconforto e gasto poderão ser evitados. Os dados referem-se a 53 homens com câncer de próstata. Para cada paciente, temos: idade (**age**), serum acid phosphatase (**acid** – um valor de laboratório que é elevado se o tumor se espalhou para certas regiões), o estágio da doença (**stage** – uma indicação do avanço da doença), o grau do tumor (**grade** – uma indicação da agressividade) e os resultados do raio x (**xray**), assim como se o câncer se espalhou para os nódulos da região linfática na fase da cirurgia. O problema é prever se os nódulos são positivos para câncer com base nos valores das variáveis que podem ser medidas sem cirurgia (variável dependente: **node**). As variáveis `xray`, `stage` e `grade` são categóricas, codificadas como 0 e 1. O valor 1 sempre indica a pior situação (raio x positivo, estágio avançado e agressividade). Aplique a técnica análise de discriminante. Interprete os resultados obtidos e responda às questões: