



2º Teste de Avaliação de Conhecimentos – 2024/1

Instruções gerais:

- Insira a data e identifique-se no cabeçalho da folha de respostas (nome e matrícula).
- A prova é composta de duas questões (A, B e C). Todas as questões devem ser desenvolvidas em folhas de papel almaço ou relatório RMarkdown. Não se esqueça de rotular corretamente as questões e seus itens.
- Lembre-se que as questões são abertas. Tenha clareza no desenvolvimento das questões e deixe claro seu raciocínio e todos os comandos que utilizou para chegar aos resultados apresentados. A avaliação não considerará apenas o resultado oferecido.

QUESTÕES

- A. Uma empresa tem adotado o seguinte procedimento para avaliar a qualidade do produto que está sendo fabricado: Sabe-se que quando o processo está sob “controle” estatístico, a probabilidade de se fabricar uma peça com defeito é, em média, igual a $p = 0,01$. Entretanto, quando o processo está “fora de controle”, esta probabilidade passa para $p = 0,03$, em média. De acordo com os resultados de cada amostra de 4 peças que é selecionada, tem que se tomar uma decisão sobre o estado de “controle” ou de “falta de controle” do processo produtivo. Suponha que você seja o Estatístico dessa empresa.
1. Construa uma regra de classificação que permita, para cada amostra observada, decidir se o processo “está ou não fora de controle”. O objetivo é minimizar os erros de classificação.
 2. Se uma determinada amostra indicar um número de peças defeituosas igual a 2, você diria que o processo está “sob controle estatístico” ou “fora de controle”? Justifique.
 3. Descreva os erros de classificação para este problema.
 4. Calcule as probabilidades dos erros de classificação definidos em (3) de acordo com a regra de classificação que você adotou em (1).
 5. Como Estatístico dessa empresa, o que você poderia dizer sobre a qualidade da regra de classificação?
- B. O conjunto de dados `biopsy{MASS}` contém observações multivariadas sobre 699 biópsias de tumores de mama, que foram classificados como "benign" (benigno) ou "malignant" (benigno ou maligno) [MANGASARIAN e WOLBERG, 1990]. As medidas das nove variáveis em cada biópsia são escores (de 1 a 10) atribuídos pelo médico coordenador, de acordo com sua avaliação clínicas. Analise esses dados. Em particular, investigue as diferenças entre os dois tipos de tumor, encontre uma regra para classificar os tumores com base apenas nas variáveis do conjunto de dados e avalie o desempenho de sua regra.
- C. Qual das técnicas estatísticas, que você aprendeu nesta disciplina despertou mais sua atenção? Explique resumidamente.