**Estatística Multivariada 2**

**Ensino Não Presencial Emergencial - 2**

**Lista de Exercícios - 1**

**Entrega: 08/01/2021 - 29/03/2021 – 23:59 hs**

1. Um estudo deseja determinar um procedimento para discriminar duas populações multivariadas. A pesquisa possui dados suficientes para estimar as funções de densidade f1(x) e f2(x) associadas as populações π1 e π2  respectivamente. Seja ainda, C(2/1) = 50 e C(1/2) = 100. Sabe-se ainda que 20% do total de possíveis itens pertencem a π2 .
2. Apresente ECM (na sua forma geral), para atribuir um novo item a uma das duas populações.
3. Medidas observadas para um novo item tem f1(x) = 0.3 e f2(x) = 0.5 . A partir do procedimento definido em a, em que população deveria ser atiribuida esta nova observação.
4. Sejam duas populações π1 e π2 com densidades dadas por:

f1(x) = (1/2) ( 1 - | x | ) para | x | ≤ 1 ;

f2(x) = (1/2) ( 1 - | x – 0.5 | ) para –0.5 ≤ x ≤ 1.5 ;

1. Faça um esboço(em um mesmo gráfico) das densidades de cada uma das populações;
2. Estabeleça a regra de classificação que minimiza ECM se p1 = p2 e c(1/2) = c(2/1);
3. Pela regra acima, se uma nova observação x0 = -.4 é observada, em que grupo a mesma deve ser classificada; Represente no gráfico de a, esta nova observação.
4. Quais seriam os resultado de b e c no caso de p1 = 0.2. Como podemos interpretar (justificar) as diferenças de resultados obtidos nos dois casos.
5. Dois grupos de indivíduos classificados segundo a presença (G2) ou não (G1) de indicadores de senilidade (velhice). Para cada individuo uma bateria de quatro testes (X1 a X4) foi realizado cujos resultados, por grupo são apresentados abaixo:

|  |  |  |
| --- | --- | --- |
|  | Médias Grupos | |
|  | G1 – n1=37 | G2 – n2=12 |
| X1 | 12.57 | 8.75 |
| X2 | 9.57 | 5.33 |
| X3 | 11.49 | 8.50 |
| X4 | 7.97 | 4.75 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 11.2624 | 9.04060 | 7.1550 | 3.3791 |
| Sp= |  | 13.5265 | 7.3784 | 2.5014 |
|  |  |  | 11.5796 | 2.6167 |
|  |  |  |  | 5.8133 |

2.1) Assumindo que ambos grupos tem uma densidade normal, obtenha uma regra de classificação para um novo individuo (considere custos e prioris iguais).

2.2)Em que grupo seriam classificados indivíduos com as seguintes medidas:

I1 = ( 10.4, 7.3, 9.4, 6.3) I2= (13.0, 9.0, 10.6, 9.1)

1. Os dados abaixo referem-se a análise de livros de três diferentes autores. De cada autor foi considerado um determinado livro do qual foram selecionadas aleatoriamente 14 paginas. Em cada uma dessas paginas as seguintes variáveis foram observadas na ordem do arquivo de dados (Dados14.xls) são:

AUTOR - 1 – Tom Clancy – The Bear and the Dragon

2 – K. Rowling – Harry Potter e a Pedra Filosofal

3 – Leon Tolstoi – Guerra e Paz

PALAVRA – Número médio de palavras por parágrafo

CARACTERES – Número médio de caracteres por palavra

FACILIDADE – Índice de Facilidade de Leitura – Quanto maior ⇒ mais fácil

LEGIBILIDADE – Índice de Legibilidade – Quanto maior ⇒ mais legível

1. Apresente medidas descritivas apropriadas ao problema;
2. Assumindo que os dados podem ser considerados provenientes de um modelo normal, encontre uma apropriada regra discriminante para o problema.
3. Apresente pelo menos uma medida de avaliação do modelo obtidoem b?