

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
PROVA DE RECUPERAÇÃO DE ESTATÍSTICA COMPUTACIONAL - 13/11/2024

Questão 1. Um álbum tem espaço para N figurinhas. Cada pacote que você compra, para preencher o álbum, vem com quatro figurinhas. Todas as figurinhas são igualmente prováveis de serem obtidas. Você está interessado em saber o número de pacotes que deve comprar para completar o álbum. Assim, como exemplo, considere que $N = 5$ e que as figurinhas sejam 1, 2, 3, 4, 5. Suponha que as figurinhas que você obteve, na ordem das compras dos pacotes, foram: $\{4, 2, 5, 3\}$, $\{3, 4, 2, 5\}$ e $\{1, 1, 5, 3\}$. Portanto, neste caso, precisou-se comprar 3 pacotes.

- (a) Para $N = 20$, considere o seguinte experimento: pacotes serão comprados até que o álbum esteja completo; registra-se, então, o número total de pacotes comprados. Simule este experimento 100 mil vezes e registre o resultado de cada simulação. O que a média desses resultados representa? Para esta mesma simulação, calcule a média dos resultados que são maiores do que 5 e interprete este resultado.
- (b) Crie uma função cuja entrada seja N e cuja saída seja a estimativa a média dos resultados obtidos. Analise a função para $N = 5, 10, 15, 20$. Comente os resultados obtidos.

Questão 2. Papagaio-do-mar é o nome comum dado às aves charadriiformes da família dos alcídeos, pertencentes ao gênero *Fratercula*. Existem três espécies de papagaios-do-mar conhecidas: *arctica*, *corniculata* e *cirrhata*. O conjunto `papagaio.txt` apresenta informações sobre o peso (em gramas), o tamanho (em centímetros), a envergadura da asa (em centímetros), o tamanho do bico (em centímetros) e a espécie de 500 papagaios-do-mar.

- (a) A partir do conjunto `papagaio.txt`, faça uma análise gráfica a fim de conhecer cada espécie de papagaio-do-mar. Escreva um pequeno texto comentando sobre suas descobertas.
- (b) Para a espécie *corniculata*, há uma ave com um dado faltante. Crie um modelo de regressão para descobrir este dado faltante.
- (c) Crie um modelo KNN para classificar as espécies de papagaio-do-mar. Utilize $k = 1, 3, 4$ vizinhos. Qual desses modelos se saiu melhor? Justifique sua resposta.

Questão 1: (a) 20 pontos; (b) 20 pontos.

Questão 2: (a) 20 pontos; (b) 20 pontos; (c) 20 pontos.