



**CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA - CEEI
UNIDADE ACADÊMICA DE SISTEMAS E COMPUTAÇÃO - DSC**

Perguntas Referentes ao Projeto

**RELATÓRIO DE ATIVIDADES
DESENVOLVIDAS PELO(A)
MONITOR(A) Iann Carvalho Barbosa,
NA DISCIPLINA Estatística Aplicada, NO
PERÍODO LETIVO 2019.1.**

Campina Grande – 05 de Julho de 2019

1. O que é o erro em uma regressão linear? Qual a sua relação com os betas?

O erro em uma regressão linear é uma estimativa todas as influências no comportamento da variável Y que não podem ser explicadas pelo comportamento da variável X. O resultado dos betas varia o erro com a fórmula da reta de regressão linear.

2. Qual a importância de se ter independência entre variáveis?

Chamamos de colinearidade e multicolinearidade o fenômeno que indica a correlação forte entre duas variáveis, quando elas possuem uma relação muito forte, podem acarretar problemas levando a previsões e inferências erradas ou duvidosas. A independência ocorre quando a ocorrência de uma variável não afeta a probabilidade de ocorrência da outra.

3. Qual a diferença entre independência estocástica (probabilística) e independência linear (determinística)?

A independência estocástica ocorre quando os eventos são independentes e a ocorrência de um não afeta a probabilidade do outro ocorrer. Por exemplo:

$$P(A \cap B) = P(A) \times P(B)$$

a independência estocástica é quando não é correlação entre esses eventos:

$$\text{Corr}(A, B) = 0.$$

Já a independência linear ocorre quando um conjunto de vetores $\{v_1, v_2, \dots, v_k\}$

$$x_1 v_1 + x_2 v_2 + \dots + x_k v_k = 0$$

possui apenas uma solução:

$$x_1 = x_2 = \dots = x_k = 0.$$

4. Qual a importância da independência linear entre as variáveis de uma regressão linear múltipla?

Na regressão linear múltipla, assim como na regressão linear simples, é necessário uma independência dessas variáveis, porque quando elas possuem uma relação muito forte, podem acarretar problemas levando a previsões e inferências erradas ou duvidosas.

5. Quando é necessário fazer treinamento e teste? Análise de resíduos já resolve o problema?

O treinamento é necessário para problemas relacionados à estatística inferencial na criação de modelos preditivos e a análise de resíduos pode ser utilizada para verificar as diferenças entre os valores previstos e o valor esperado.

6. Normalidade, independência, homogeneidade. O que são e para que servem?

- Normalidade: Os testes de normalidade são usados para determinar se um conjunto de dados de uma dada variável aleatória é bem modelada por uma distribuição normal ou não.
- Independência (estatística): A Independência entre variáveis aleatórias ou eventos significa que a partir do resultado de um deles não é possível inferir nenhuma conclusão sobre o outro.
- O teste de homogeneidade é usado quando pretendemos testar se os dados associados às categorias de uma das variáveis se comporta semelhante nas classes ou populações definidas pelas categorias da outra variável classificatória.