

**IFTM - INSTITUTO FEDERAL DO TRIÂNGULO MINEIRO  
ENGENHARIA DE COMPUTAÇÃO**

**Yuri David Silva Duarte**

**PROBABILIDADE E ESTATÍSTICA  
TRABALHO FINAL**

**Uberaba  
2024**

**YURI DAVID SILVA DUARTE**

**PROBABILIDADE E ESTATÍSTICA  
TRABALHO FINAL**

Projeto de pesquisa apresentado ao curso de Engenharia de Computação do Instituto Federal do Triângulo Mineiro para obtenção do diploma de graduação

Orientador: Jorge Alencar

**Uberaba  
2024**

```

# Bibliotecas

import pandas as pd
import math
import matplotlib.pyplot as plt

# Estudante

print("\tTrabalho de Probabilidade e Estatística\n")

print("\tCapítulo 9: Tópicos Especiais\n")

print("Estudante: Yuri David Silva Duarte\n")

print("Professor: Jorge Alencar\n")

print("Data: 11/12/24\n")

input("Pressione Enter para continuar...")

# Entrada de dados

variaveis = ["Var. X", "Var. Y", "Var. Z", "Var. W"]

amostraA = [
    [7.00, 3.20, 4.70, 1.40],
    [6.40, 3.20, 4.50, 1.50],
    [6.90, 3.10, 4.90, 1.50],
    [5.50, 2.30, 4.00, 1.30],
    [6.50, 2.80, 4.60, 1.50],
    [5.70, 2.80, 4.50, 1.30],
    [6.30, 3.30, 4.70, 1.60],
    [4.90, 2.40, 3.30, 1.00],
    [6.60, 2.90, 4.60, 1.30],
    [5.20, 2.70, 3.90, 1.40],
    [5.00, 2.00, 3.50, 1.00]
]

df = pd.DataFrame(amostraA, columns=variaveis)

print("\n\tAmostra A\n\n",df, end="\n\n")

amostraB = [
    [6.30, 3.30, 6.00, 2.50],
    [5.80, 2.70, 5.10, 1.90],
    [7.10, 3.00, 5.90, 2.10],
    [6.30, 2.90, 5.60, 1.80],
    [6.50, 3.00, 5.80, 2.20],

```

```

        [7.60, 3.00, 6.60, 2.10],
        [4.90, 2.50, 4.50, 1.70],
        [7.30, 2.90, 6.30, 1.80],
        [6.70, 2.50, 5.80, 1.80],
        [7.20, 3.60, 6.10, 2.50],
        [6.50, 3.20, 5.10, 2.00],
        [6.40, 2.70, 5.30, 1.90],
        [6.80, 3.00, 5.50, 2.10],
        [5.70, 2.50, 5.00, 2.00],
        [5.80, 2.80, 5.10, 2.40],
        [6.40, 3.20, 5.30, 2.30],
        [6.50, 3.00, 5.50, 1.80],
        [7.70, 3.80, 6.70, 2.20],
        [7.70, 2.60, 6.90, 2.30],
        [6.00, 2.20, 5.00, 1.50],
        [6.90, 3.20, 5.70, 2.30]
    ]

df = pd.DataFrame(amostraB, columns=variaveis)

print("\tAmostra B\n\n",df, end="\n\n")

varYA = [3.20, 3.20, 3.10, 2.30, 2.80, 2.80, 3.30, 2.40, 2.90, 2.70, 2.00]

print("\tVáriavel Y da Amostra A\n", varYA, end="\n\n")

varYB = [
    3.30, 2.70, 3.00, 2.90, 3.00, 3.00, 2.50, 2.90, 2.50, 3.60,
    3.20, 2.70, 3.00, 2.50, 2.80, 3.20, 3.00, 3.80, 2.60, 2.20,
    3.20
]

print("\tVáriavel Y da Amostra B\n", varYB, end="\n\n")

varWA = [1.40, 1.50, 1.50, 1.30, 1.50, 1.30, 1.60, 1.00, 1.30, 1.40, 1.00]

print("\tVáriavel W da Amostra A\n", varWA, end="\n\n")

varWB = [
    2.50, 1.90, 2.10, 1.80, 2.20, 2.10, 1.70, 1.80, 1.80, 2.50,
    2.00, 1.90, 2.10, 2.00, 2.40, 2.30, 1.80, 2.20, 2.30, 1.50,
    2.30
]

print("\tVáriavel W da Amostra B\n", varWB, end="\n\n")

input("Pressione Enter para continuar...")

```

```

# Questão 1

## A partir da Amostra A e da Amostra B, determine se a população A e a
população B
## possuem a mesma variância populacional da Variável 01 a uma significância
de 10%.

print("\n\tQuestão 1\n")

### Valor da significância
significancia = 10 %%

### Média de YA
mediaYA = sum(varYA) / len(varYA)

### Variância Amostral de YA
varianciaAmostralYA = 0

for i in varYA:
    varianciaAmostralYA += (i - mediaYA) ** 2

varianciaAmostralYA /= (len(varYA) - 1)

### Média de YB
mediaYB = sum(varYB) / len(varYB)

### Variância Amostral de YB
varianciaAmostralYB = 0

for i in varYB:
    varianciaAmostralYB += (i - mediaYB) ** 2

varianciaAmostralYB /= (len(varYB) - 1)

### Cálculo do estimador F

F = varianciaAmostralYA / varianciaAmostralYB

### Ponto de corte

# TODO: Pegar da tabela

F1 = 1 / 2.77 #F(a,b)

F2 = 2.77 #F(b,a)

### Saída de dados

```

```

print(f"Média da variável Y da Amostra A: {mediaYA:.4f}\n")

print(f"Variância Amostral da variável Y da Amostra A:
{varianciaAmostralYA:.4f}\n")

print(f"Média da variável Y da Amostra B: {mediaYB:.4f}\n")

print(f"Variância Amostral da variável Y da Amostra B:
{varianciaAmostralYB:.4f}\n")

print(f"F = S²YA / S²YB = {varianciaAmostralYA:.4f} /
{varianciaAmostralYB:.4f} = {F:.4f}\n")

# TODO: Entender se isso tá certo

print(f"F1 = F(a,b) = {F1:.4f}\n")

print(f"F2 = F(b,a) = {F2:.4f}\n")

### Teste de região crítica (pg 334 da referência 1)
if(F < F1 or F > F2):
    print("As variâncias são diferentes\n")

else:
    print("As variâncias são iguais\n")

input("Pressione Enter para continuar...")

# Questão 2

## A partir da Amostra A e da Amostra B, determine se a população A e a
população B
## possuem a mesma média populacional da Variável 01 a uma significância de
10%.
## (Note que o aluno pode ter que usar a informação obtida no item anterior
para tomar
## essa decisão).

print("\n\tQuestão 2\n")

significancia = 10

### Eu sei que as variancias são iguais, pelo exercicio anterior
### amostras Independentes variancias desconhecidas e iguais: ^
### usamos o estimador media amostral

```

```

### Estimador da média amostral
D = mediaYA - mediaYB

### Estimador da variância
EstimadorVar = (((len(varYA) - 1) * varianciaAmostralYA) + ((len(varYB) - 1)
* varianciaAmostralYB)) / ((len(varYA) - 1) + (len(varYB) - 1))

### H0 :  $\mu_D = 0$ 
ud = 0

### Tabela de t-student
### Usando a distribuição de t-Student
T = (D - ud) / (math.sqrt(EstimadorVar * (len(varYA) + len(varYB))))

### T possui a distribuição t-Student com ((nx - 1) + (ny - 1)) graus de
liberdade.
grauDeLiberdade = ((len(varYA) - 1) + (len(varYB) - 1))

### Peguei este valor da tabela com (grau de liberdade) x (significancia /
2)
t = 2.042

### Saída de dados

print(f"Estimador da média amostral = {D:.4f}\n")

print("Considerando a hipótese nula sendo H0 :  $\mu_D = 0$ \n")

print(f"Usando a distribuição de t-Student para T = {T:.4f}\n")
print(f"Obtemos t = {t:.4f}\n")

print("Concluimos que:",end=" ")
if(T < (t * -1) or T > (t)):
    print("As médias são diferentes\n")

else:
    print("As médias são iguais\n")

input("Pressione Enter para continuar...")

# Questão 3

## Para verificar o efeito da Variável 01 sobre a Variável 02,
## use a Amostra A para obter a reta ajustada entre as variáveis.
## Construa o diagrama de dispersão, baseando-se nos pares de valores
## fornecidos e, em seguida, desenhe a reta ajustada. Baseando-se apenas

```

```

## no gráfico, você diria que o ajuste é adequado?

print("\n\tQuestão 3\n")

### Cálculo da média da variável W da amostra B
mediaWA = sum(varWA) / len(varWA)

### Cálculo do parâmetro Beta
numerador = 0
denominador = 0

for i in range(len(varYA)):
    numerador += (varYA[i] * varWA[i]) - (len(varYA) * mediaYA * mediaWA)
    denominador += (varYA[i] ** 2) - ((len(varYA)) * (mediaYA ** 2))

betaObs = numerador / denominador

### Cálculo do parâmetro Alfa
alfaObs = mediaWA - (betaObs * mediaYA)

### Colocando os pontos das variáveis no gráfico
plt.plot(varYA, varWA, 'o')

### Gerando a reta
eixox = [min(varYA), max(varYA)]
eixoy = [(betaObs * min(varYA)) + alfaObs, (betaObs * max(varYA)) + alfaObs]

plt.plot(eixox, eixoy)

### Organizando o gráfico
plt.title('População no modelo Regressão Linear Simples')
plt.xlabel('Variável Y')
plt.ylabel('Variável W')

### Saída de dados

print(f"A média da variável Y da amostra A é: {mediaYA:.4f}\n")

print(f"A média da variável W da amostra A é: {mediaWA:.4f}\n")

print(f"O valor do parâmetro Beta é: {betaObs:.4f}\n")

print(f"O valor do parâmetro Alfa é: {alfaObs:.4f}\n")

print(f"y = Alfa + (Beta * X) = {alfaObs:.4f} + ({betaObs:.4f} * x)\n")

print("Agora vamos olhar o gráfico\n")

```



```
input("Pressione Enter para continuar...")

plt.show()

print("\nAnalisando o gráfico, a reta foi ajustada adequadamente!\n")

input("Pressione Enter para encerrar...")

#TODO: Comando pyinstaller --onefile --icon="Icons/favicon.ico"
Code/trabalhoYuri.py
```