

## 1. a)

Para a seleção das idades de forma aleatória, utilizei um código em python usando a biblioteca random, como mostra o texto abaixo:

```
atores = [
    44, 41, 62, 52, 41, 34,
    34, 52, 41, 37, 38, 34,
    32, 40, 43, 56, 41, 39,
    49, 57, 41, 38, 42, 52,
    51, 35, 30, 39, 41, 44,
    49, 35, 47, 31, 47, 37,
    57, 42, 45, 42, 44, 62,
    43, 42, 48, 49, 56, 38,
    60, 30, 40, 42, 36, 76,
    39, 53, 45, 36, 62, 43,
    51, 32, 42, 54, 52, 37,
    38, 32, 45, 60, 46, 40,
    36, 47, 29, 43
]
entradas = []
for i in range(int(len(atores)/2)):
    entrada = random.choice(atores)
    actores.remove(entrada)
    entradas.append(entrada)
    print(entrada, end = ' ')
print()
```

- A amostra utilizada para os testes foi a seguinte:

(30, 32, 34, 34, 34, 35, 35, 36, 36, 36, 37, 37, 37, 38, 38, 39, 40, 40, 41, 41, 41, 42, 42, 43, 43, 43, 43, 44, 45, 47, 49, 51, 51, 52, 52, 53, 57, 62)

$H_0 = 33.5$   $H_1 \geq 33.5$

- Dos 38 valores observados na amostra, 36 são maiores que 33.5
- Usando a distribuição binomial com  $n = 38$ ,  $p = 0.5$ , encontramos:

Valor-p =  $P\{S \geq S_{obs}\} = P\{S \geq 33.5\} = 100.00\%$

```
import scipy.stats as stats
from scipy.stats import binom

x, n, p = 2, 38, 0.5

prob = 1 - binom.cdf(x, n, p)
print(prob)
Valor-p = 2.699380274862051e-09
```

Portanto rejeita  $H_0$ , pois para que  $H_0$  seja verdadeiro o nível de significância teria de ser menor 0.0000002%.

## b)

Ainda utilizando a amostra do exercício anterior, e utilizando a biblioteca stats para o cálculo do teste de Wilcoxon, redigi o seguinte código:

```
import scipy.stats as stats

idadeAtores = [
    30, 32, 34, 34, 34, 35, 35,
    36, 36, 36, 37, 37, 37, 38, 38, 39, 40, 40,
    41, 41, 41, 42, 42, 43, 43, 43, 43, 44, 45,
    47, 49, 51, 51, 52, 52, 53, 57, 62
]
```

```
difIdade = [i - 33.5 for i in idadeAtores]
prob = stats.wilcoxon(difIdade, alternative='greater')
print(prob[1])
```

E obtive o seguinte resultado:

Valor-p = 7.530616130679846e-10

Portanto, o teste mostra forte evidência que os atores tenham mais de 33.5 anos de idade. Assim como no teste anterior.

c)

Sejam X e Y idade dos atores e atrizes; vamos testar  $H_0: F_X = F_Y$  versus  $H_1: F_X > F_Y$

Utilizando a biblioteca stats, para efetuar os cálculos e usando duas amostras aleatórias das idades (seguindo a mesma logica da letra A), obtemos:

```
import scipy.stats as stats

Atores = [29, 32, 32, 34, 34, 35, 36, 36, 37, 37, 38, 38, 39, 40, 40, 40, 41, 41,
          41, 41, 42, 42, 43, 43, 44, 44, 45, 49, 49, 51, 52, 52, 52, 53, 57, 62, 62, 76]

Atrizes = [17, 19, 21, 24, 24, 25, 26, 27, 27, 28, 28, 29, 29, 29, 29, 30, 31, 31,
           33, 34, 34, 34, 35, 35, 35, 36, 38, 39, 41, 41, 43, 45, 54, 60, 61, 63, 74, 80]

resultado = stats.ranksums(Atores, Atrizes, 'greater')
print(resultado[1])
» Valor-p = 0.00012013720713463176
```

Portanto, vamos rejeitar  $H_0$ , visto que para ter igualdade entre as idades o nível de significância tem de ser menor que 0.01%.

2. a)

Para a seleção dos países da América do Sul de maneira aleatória, foi utilizado o seguinte código:

```
def paisesAleatorios(Paises, taxas,qtde):
    indices = sample(range(0, len(Paises)), qtde)
    rand_paises = []
    rand_taxas = []

    for i in indices:
        rand_paises.append(Paises[i])
        rand_taxas.append(taxas[i])
    print('\n',rand_paises)
    return rand_taxas

caminho = '8\Ex-08-dados.xlsx'
mediana = 2

dadosAS = pd.read_excel(caminho, sheet_name='America do Sul')
paisesAS = dadosAS.loc[:, 'País'].values.astype(str)
taxasAS = dadosAS.loc[:,2018].values.astype(float)

rand_taxasAS = paisesAleatorios(paisesAS, taxasAS, 10)
```

E obtemos a seguinte amostra:

['Guiana', 'Venezuela', 'Argentina', 'Colômbia', 'Peru', 'Paraguai', 'Chile', 'Brasil', 'Uruguai', 'Ecuador']

Assumindo as hipóteses de:

H0 = 2      H1 &gt; 2

E utilizando o teste de Wilcoxon, obtemos:

```
dif_texasAS = [taxa - 2 for taxa in rand_texasAS]
probAS = wilcoxon(dif_texasAS, alternative='greater')
```

Valor-p = 0.2783203125

Portanto, não podemos rejeitar H0, pois o nível de significância para aceitar H1 é muito alto (acima de 5%), ou seja, podemos rejeitar H0 com 82% de certeza.

b) H0 = 2      H1 &gt; 2

```
dadosEU = pd.read_excel(caminho, sheet_name='Europa')
paísesEU = dadosEU.loc[:, 'País'].values.astype(str)
texasEU = dadosEU.loc[:, 2018].values.astype(float)

rand_texasEU = paísesAleatorios(paísesEU, texasEU)

dif_texasEU = [taxa - 2 for taxa in rand_texasEU]
probEU = wilcoxon(dif_texasEU, alternative='greater')
```

**Amostra** = ['Bélgica', 'Croácia', 'Albânia', 'Áustria', 'Bósnia e Herzegovina', 'Bielorrússia', 'Armênia', 'Andorra', 'Chipre', 'Azerbaijão']

Valor-p = 0.9990234375

Portanto, também não podemos rejeitar H0, pois há 99.9% de chance de H0 ser verdadeiro.

c) H0: FX = FY versus H1: FX > FY, sendo X e Y, as taxas dos países da América e da Europa;

Utilizando o Teste da soma dos postos e as amostras obtidas nas questões anteriores, obtemos:

```
print(ranksums(rand_texasAS, rand_texasEU, alternative='greater').pvalue)
» Valor-p = 0.0015985497239672654
```

Portanto, iremos rejeitar H0 99.84% de certeza.

```
print(wilcoxon(rand_texasAS, rand_texasEU, alternative='greater').pvalue)
» Valor-p = 0.009765625
```

Veja que pelo teste de Wilcoxon também rejeitamos H0, porém com um grau de significância um pouco menor.

d) H0: FX = FY versus H1: FX > FY, sendo X e Y, as taxas dos países da América e da África;

**Amostra** = ['Burundi', 'Eritreia', 'África do Sul', 'Botsuana', 'Chade', 'Camarões', 'Benim', 'Argélia', 'Egito', 'Burkina Faso']

```
dadosAF = pd.read_excel(caminho, sheet_name='Africa')
paísesAF = dadosAF.loc[:, 'País'].values.astype(str)
texasAF = dadosAF.loc[:, 2018].values.astype(float)

rand_texasAF = paísesAleatorios(paísesAF, texasAF)

dif_texasAF = [taxa - 2 for taxa in rand_texasAF]
probAF = wilcoxon(dif_texasAF, alternative='greater')
print('\nValor-p AS > AF(Ranksums)', ranksums(rand_texasAS, rand_texasAF,
alternative='greater').pvalue)
» Valor-p AS > AF(Ranksums) = 0.9997467292576538
```

Portanto, não podemos rejeitar H0, visto que o valor-p é muito alto, de tal forma que podemos rejeitar H1 com 99.97% de certeza.

3.

```
dadosPA = pd.read_excel(caminho, sheet_name='Países')
paísesPA = dadosPA.loc[:, 'País'].values.astype(str)
texasPA18 = dadosPA.loc[:, 2018].values.astype(float)
texasPA10 = dadosPA.loc[:, 2010].values.astype(float)
```

```

print('\nPareado')
rand_texasPA18 = paisesAleatorios(paisesPA, taxasPA18, 100)
print()
rand_texasPA10 = paisesAleatorios(paisesPA, taxasPA10, 100)
print('\nNao pareado ')
taxasPA18_NPareado = paisesAleatorios(paisesPA, taxasPA18, 120, False)
print()
taxasPA10_NPareado = paisesAleatorios(paisesPA, taxasPA10, 80, False)

probPA_pareado = ranksums(rand_texasPA18, rand_texasPA10, alternative='less')
print('Teste pareado: ', probPA_pareado[1])

probPA_NPareado = ranksums(taxasPA18_NPareado, taxasPA10_NPareado, alternative='less')
print('Teste não Pareado: ', probPA_NPareado[1])

```

a) H0: FX = FY versus H1: FX < FY, sendo X e Y, as taxas dos países em 2018 e 2010;

**Amostra Pareada:** ['Japão', 'Lituânia', 'Armênia', 'Fiji', 'Níger', 'Mongólia', 'Letônia', 'Guiana', 'Moçambique', 'Irã', 'República Centro-Africana', 'El Salvador', 'Nicarágua', 'Catár', 'Gana', 'Togo', 'Botsuana', 'São Cristóvão e Nevis', 'ESwatiní', 'Palau', 'Ruanda', 'Chade', 'Guiné Equatorial', 'Brasil', 'Bélgica', 'Hungria', 'Mianmar', 'Peru', 'Sudão', 'Indonésia', 'Malásia', 'Guiné-Bissau', 'Samoa', 'Djibuti', 'Paraguai', 'Tuvalu', 'Maldivas', 'Noruega', 'Vanuatu', 'Bahrein', 'Paquistão', 'Alemanha', 'Bósnia e Herzegovina', 'Lesoto', 'Afeganistão', 'Namíbia', 'Iémen', 'Estados Unidos da América', 'Honduras', 'Bangladesh', 'Costa Rica', 'Emirados Árabes Unidos', 'Espanha', 'Cazaquistão', 'Mali', 'República da Moldávia', 'Benim', 'Haiti', 'Uzbequistão', 'Montenegro', 'Bulgária', 'Ucrânia', 'Granada', 'Cabo Verde', 'Quênia', 'São Marinho', 'Mônaco', 'Tunísia', 'Nigéria', 'Gabão', 'Grécia', 'Papua Nova Guiné', 'Tajiquistão', 'Austrália', 'Guatemala', 'Rússia', 'Croácia', 'Chipre', 'Geórgia', 'Butão', 'Suécia', 'Palestina', 'Tanzânia', 'Libéria', 'Seicheles', 'Andorra', 'Guiné', 'Somália', 'Camboja', 'Eslovênia', 'Argélia', 'México', 'Ilhas Cook', 'Itália', 'Sérvia', 'Belize', 'Vietnã', 'Venezuela', 'Índia', 'França']

Valor-p = 0.048306088643448444

Portanto, a partir do valor-p iremos rejeitar H0 com 95.27% de certeza que há uma leve redução taxa de mortalidade com base na amostra escolhida.

b) H0: FX = FY versus H1: FX < FY, sendo X e Y, as taxas dos países em 2018 e 2010;

**Amostra 2018:** ['Tailândia', 'Kiribati', 'Tonga', 'Nova Zelândia', 'Gabão', 'Noruega', 'Eslovênia', 'El Salvador', 'Suriname', 'República da Moldávia', 'Lituânia', 'Reino Unido', 'Uzbequistão', 'Malta', 'Namíbia', 'Sudão', 'Serra Leoa', 'Irã', 'Bielorrússia', 'Honduras', 'Samoa', 'Burundi', 'Mongólia', 'Turquemenistão', 'São Vicente e Granadinas', 'Iémen', 'Costa Rica', 'Estados Unidos da América', 'Andorra', 'França', 'Suécia', 'Emirados Árabes Unidos', 'Israel', 'Comores', 'Somália', 'Macedônia do Norte', 'Bahrein', 'Brasil', 'Chade', 'Sri Lanka', 'Austrália', 'Quênia', 'Palau', 'Bélgica', 'Antígua e Barbuda', 'Cabo Verde', 'Síria', 'Butão', 'Vietnã', 'Bósnia e Herzegovina', 'Islândia', 'Seicheles', 'Argentina', 'Angola', 'Croácia', 'Costa do Marfim', 'Quirguistão', 'Moçambique', 'Dominga', 'Suíça', 'Papua Nova Guiné', 'Luxemburgo', 'Rússia', 'Fiji', 'Barbados', 'Romênia', 'Paquistão', 'Holanda', 'Chile', 'Congo', 'Ucrânia', 'Peru', 'Arábia Saudita', 'ESwatiní', 'Eslováquia', 'Niue', 'Guatemala', 'Áustria', 'Índia', 'Cuba', 'Níger', 'Camboja', 'Coreia do Norte', 'Irlanda', 'Uganda', 'Tunísia', 'Gana', 'Zâmbia', 'Mali', 'Ilhas Cook', 'Brunei Darussalam', 'Japão', 'Mianmar', 'Letônia', 'Benim', 'Kuwait', 'São Marinho', 'Filipinas', 'Camarões', 'Uruguai', 'Marrocos', 'Nauru', 'São Cristóvão e Nevis', 'Bolívia', 'África do Sul', 'Nigéria', 'Djibuti', 'Líbia', 'Etiópia', 'China', 'Madagáscar', 'Panamá', 'Peru', 'Belize', 'Gâmbia', 'Ilhas Salomão', 'Eritreia', 'Armênia', 'Colômbia', 'Maurício']

**Amostra 2010:** ['Papua Nova Guiné', 'Gâmbia', 'Canadá', 'Espanha', 'Mônaco', 'Indonésia', 'Samoa', 'Geórgia', 'Suíça', 'Irã', 'República Democrática do Congo', 'Seicheles', 'São Cristóvão e Nevis', 'Camboja', 'Guiné Equatorial', 'Jordânia', 'Líbia', 'Síria', 'Benim', 'Afeganistão', 'República Dominicana', 'Croácia', 'Iémen', 'Chipre', 'Vanuatu', 'Equador', 'Senegal', 'Mali', 'Itália', 'Portugal', 'Tanzânia', 'Argélia', 'Lesoto', 'Tajiquistão', 'Lituânia', 'Austrália', 'Colômbia', 'Sérvia', 'Bangladesh', 'Fiji', 'Sudão', 'Rússia', 'Omã', 'Mongólia', 'Peru', 'Suriname', 'Argentina', 'Namíbia', 'Honduras', 'Timor-Leste', 'México', 'Áustria', 'Trindade e Tobago', 'Costa Rica', 'Bielorrússia', 'Togo', 'Alemanha', 'São Tomé e Príncipe', 'Paquistão', 'Guiné-Bissau', 'Quirguistão', 'Egito', 'Palau', 'Niue', 'Brasil', 'Zâmbia', 'Bahrein', 'Guiana', 'Polônia', 'Nepal', 'Malásia', 'Chile', 'Catár', 'Ilhas Salomão', 'Bósnia e Herzegovina', 'Paraguai', 'Eslovênia', 'Macedônia do Norte', 'Cuba', 'Noruega']

Valor-p = 0.10852362914384733

**Nome: Lucas Patrick Saturnino Nicácio**

**Matrícula: 99763**

Portanto, a partir do valor-p não podemos rejeitar  $H_0$  visto que não há evidência significativa de que a taxa de mortalidade tenha diminuído.