Diogo Albuquerque nº2015248499
Renato Carvalho nº 2015235034
Renato Santos nº2015237457

Para esta meta foi-nos pedido a realização de hipóteses tendo por base os dados recolhidos na primeira meta. Para tal, depois de analisados os dados obtidos, foram formadas duas perguntas e, para cada uma, diferentes hipóteses.

Questões:

1. A diferença de tamanho máximo do array ordenado para N=1000, entre o QuickSort, MergeSort e InsertionSort para EPS = 1/N, 5/N, 10/N é significativa?

Para responder à questão é necessário especificar quais as variáveis dependentes e independentes em estudo. Sendo assim é de referir que as variáveis independentes são o logaritmo aplicado e o valor de EPS (valor da probabilidade ocorrência de erro na amostragem) na obtenção dos dados e que, a variável dependente é o tamanho máximo de uma subsequência ordenada. Tendo isto em conta formaram-se três conjuntos de hipóteses cada um com uma hipótese nula e uma alternativa sendo mutualmente exclusivas.

O primeiro conjunto:

H^A_0: Os três algoritmos produzem médias iguais ($\mu Q = \mu M = \mu I$);

H^A_1: Pelo menos 2 algoritmos apresentam uma média diferente;

Segundo conjunto:

H^E_0: Para os diferentes valores de EPS as médias obtidas são iguais ($\mu 0.001 = \mu 0.005 = \mu 0.01$);

H^E_1: A média é diferente para pelo menos dois valores de EPS diferentes;

Terceiro conjunto:

H^{AE}_0: As variáveis independentes não têm interação;

H^{AE}_1: As variáveis independentes têm interação.

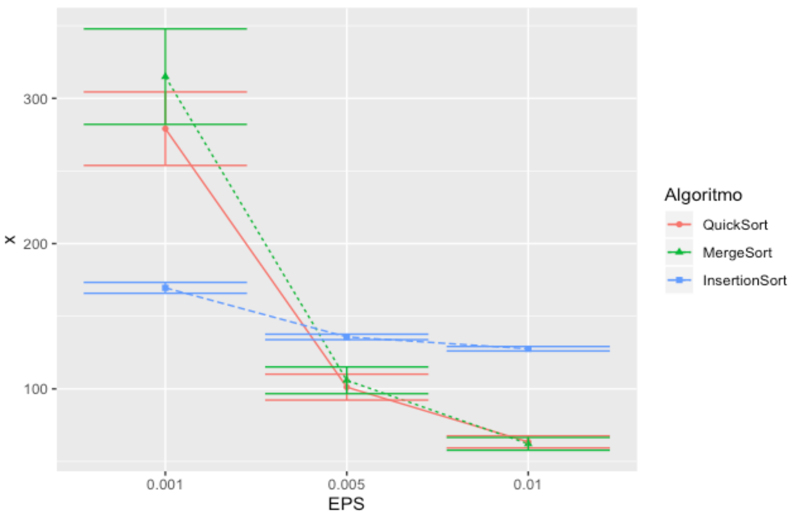

Figura 1 - Interação entre os algoritmos e a variação de EPS com intervalos de confiança das amostras relativas ao tamanho máximo de uma subsequência ordenada (x).

Analisando a ‘Figura 1’ podemos observar que o algoritmo ‘InsertionSort’ apresenta resultados médios inferiores aos restantes algoritmos para EPS=0.001 e, para os restantes valores de EPS, apresenta resultados médios do tamanho máximo de uma subsequência superiores. Podemos então esperar que existe interação entre as duas variáveis independentes em questão.

De forma a avaliarmos as hipóteses geradas, vamos recorrer ao Two-Way ANOVA de forma a concluirmos qual das hipóteses excluímos. Como temos hipóteses em que não interessa a interação (primeiros dois conjuntos de hipóteses) e para quando interessa (terceiro conjunto), iremos realizar o ANOVA das duas formas, primeiro sem interação e depois com interação.

```
> #Sem interação
> aov.out = aov(TamanhoMaximo ~ Algoritmo+EPS, data = Tabela_A_Variar_EPS_Global)
> summary(aov.out)
             Df  Sum Sq Mean Sq F value  Pr(>F)
Algoritmo     2   46532   23266   6.171 0.00218 **
EPS           2 5007679 2503839 664.066 < 2e-16 ***
Residuals   901 3397194    3770
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figura 2 - Tabela Two-Way ANOVA sem interação – questão 1

Com base nos dados obtidos (figura 2) concluímos que, como os valores de P para os algoritmos (0.00218) e EPS (2e-16) são inferiores ao nível de significância de 5% rejeitamos as hipóteses nulas HA0 e HE0.

Como as hipóteses de cada conjunto são mutualmente exclusivas afirmamos que as hipóteses alternativas HA1 e HE1 são satisfeitas.

Após realizarmos o Two-Way ANOVA sem interação, realizaremos o Two-Way ANOVA com interação.

```
> #Com interação
> aov.out = aov(TamanhoMaximo ~ Algoritmo*EPS, data = Tabela_A_Variar_EPS_Global)
> summary(aov.out)
               Df  Sum Sq Mean Sq F value   Pr(>F)
Algoritmo       2   46532   23266   10.75 2.43e-05 ***
EPS             2 5007679 2503839 1157.17  < 2e-16 ***
Algoritmo:EPS   4 1456304  364076  168.26  < 2e-16 ***
Residuals     897 1940890    2164
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figura 3 - Tabela Two-Way ANOVA com interação – questão 1

Na figura 3 podemos observar que o valor de P da interação das variáveis independentes é inferior ao nível de significância de 5% o que nos permite rejeitar a hipótese nula HAE0 e, uma vez que as hipóteses de cada conjunto são mutualmente exclusivas afirmamos que a hipótese alternativa HAE1 é satisfeita.

Como só podemos concluir alguma coisa do Two-Way ANOVA se os pressupostos forem satisfeitos, iremos realizar o teste de normalidade de resíduos e o teste de homogeneidade da variância.

```
> shapiro.test(aov.out$res)

        Shapiro-Wilk normality test

data:  aov.out$res
W = 0.7645, p-value < 2.2e-16
```

Figura 4 - Teste de normalidade dos resíduos – questão 1

```
> bartlett.test(TamanhoMaximo~interaction(Algoritmo,EPS), data=Tabela_A_Variar_EPS_Global)

        Bartlett test of homogeneity of variances

data:  TamanhoMaximo by interaction(Algoritmo, EPS)
Bartlett's K-squared = 1539.1, df = 8, p-value < 2.2e-16
```

Figura 5 - Teste de homogeneidade da variância – questão 1

Dado que, para ambos os testes efetuados (Figuras 4 e 5), os valores de p são muito pequenos (inferiores a 0.05) conclui-se que os pressupostos da análise ANOVA não se verificam. Pelo que não podemos concluir nada através desta análise, sendo necessário recorrer a métodos não paramétricos alternativos. O método usado será o teste de randomização sem restrições de permutações.

```
> print(pvalueA/5000)
[1] 0
> print(pvalueE/5000)
[1] 0
> print(pvalueAE/5000)
[1] 0
```

Figura 6 - Teste de randomização sem restrições de permutações – questão 1

Com a análise do teste presente na figura 6 podemos concluir que rejeitamos a hipótese nula do algoritmo o que significa que pelo menos dois algoritmos têm médias diferentes e rejeitamos a hipótese nula para os valores de EPS, ou seja, a média é diferente para pelo menos dois valores de EPS diferentes.

Concluímos também que existe interação entre os algoritmos e a variação de EPS no que respeita ao tamanho máximo da subsequência ordenada o que nos permite-nos responder à questão efetuada inicialmente de forma afirmativa.

Por outras palavras, para os resultados obtidos neste estudo, a escolha de um par algoritmo-valor de EPS tem uma influência significativa no tamanho máximo da subsequência ordenada obtida.

2. A diferença do tamanho máximo do array ordenado entre o QuickSort, MergeSort e InsertionSort para N = 1000, N=3000 e N = 6000 é significativa?

Para responder à questão é necessário especificar quais as variáveis dependentes e independentes. Neste caso as variáveis independentes são o logaritmo aplicado e o valor de N (tamanho máximo do array a ser ordenado) na obtenção dos dados e que, a variável dependente é o tamanho máximo de uma subsequência ordenada. Tendo isto em conta formaram-se três conjuntos de hipóteses cada um com uma hipótese nula e uma alternativa sendo mutuamente exclusivas.

Primeiro conjunto:
$H^A{}_0$: Os três algoritmos produzem médias iguais ($\mu Q = \mu M = \mu I$);
$H^A{}_1$: Pelo menos 2 algoritmos apresentam uma média diferente;

Segundo conjunto:
$H^N{}_0$: Para os diferentes valores de N as médias produzidas são iguais ($\mu 1000 = \mu 3000 = \mu 6000$);
$H^N{}_1$: A média é diferente para pelo menos dois valores de EPS diferentes;

Terceiro conjunto:
$H^{AN}{}_0$: As variáveis independentes não têm interação;
$H^{AN}{}_1$: As variáveis independentes têm interação.

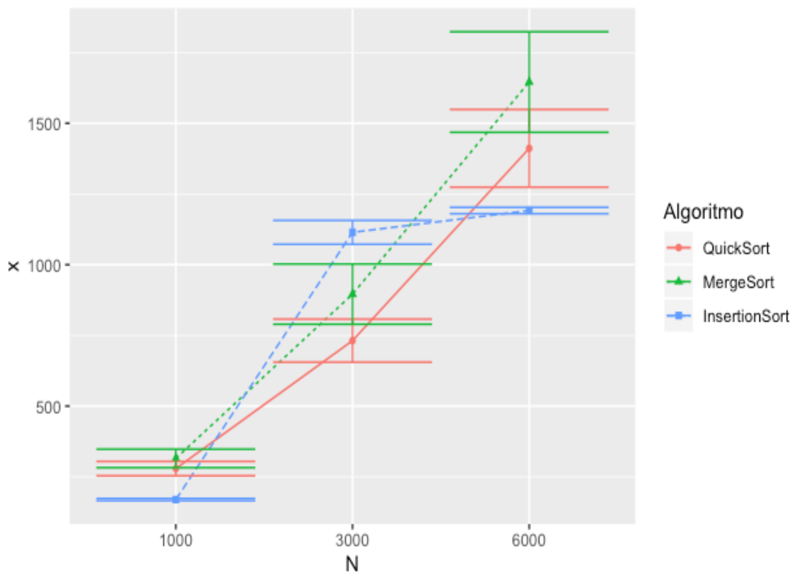

Figura 7 - Interação entre os algoritmos e a variação de N com intervalos de confiança das amostras relativas ao tamanho máximo de uma subsequência ordenada (x).

Podemos constatar na figura 7 que os algoritmos QuickSort e MergeSort apresentam melhores resultados, ou seja, médias do tamanho máximo de uma subsequência ordenada superiores para os valores de N=1000 e N=6000 quando comparados com os resultados do InsertSort. No então para N=3000 este último algoritmo apresenta valores superiores. Utilizando a figura como previsão das conclusões a que se poderão chegar mais à frente podemos especular que haverá interação entre as variáveis.

De forma a avaliarmos as hipóteses geradas, vamos recorrer ao Two-Way ANOVA de forma a concluirmos qual das hipóteses excluímos. Como temos hipóteses em que não interessa a interação (primeiros dois conjuntos de hipóteses) e para quando interessa (terceiro conjunto), iremos realizar o ANOVA das duas formas, primeiro sem interação e depois com interação.

```
> #Sem interação
> aov1.out = aov(TamanhoMaximo ~ N+Algoritmo, data = Tabela_A_Variar_N_Global)
> summary(aov1.out)
             Df    Sum Sq   Mean Sq F value  Pr(>F)    
N             2 204261747 102130874 1070.04 < 2e-16 ***
Algoritmo     2   3757363   1878681   19.68 4.3e-09 ***
Residuals   898  85710326     95446                    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figura 8 - Tabela Two-Way ANOVA sem interação - questão 2

Segundo os resultados obtidos no teste Two-Way ANOVA sem interação presentes na figura 8 os valores de P que dizem respeito aos algoritmos bem como os que dizem respeito aos valores de N são inferiores ao nível de significância escolhido (5%).

Tendo isto em conta as hipóteses nulas H^{A}_{0} e H^{N}_{0} são rejeitadas. Pelo facto de as hipóteses de cada conjunto serem mutualmente exclusivas as hipóteses H^{A}_{1} e H^{N}_{1} são satisfeitas.

Após realizarmos o Two-Way ANOVA sem interação, realizaremos o Two-Way ANOVA com interação.

```
> #Com interação
> aov1.out = aov(TamanhoMaximo ~ N*Algoritmo, data = Tabela_A_Variar_N_Global)
> summary(aov1.out)
             Df    Sum Sq   Mean Sq F value   Pr(>F)    
N             2 204261747 102130874 1293.79  < 2e-16 ***
Algoritmo     2   3757363   1878681   23.80 8.51e-11 ***
N:Algoritmo   4  15138689   3784672   47.94  < 2e-16 ***
Residuals   894  70571637     78939                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figura 9 - Tabela Two-Way ANOVA com interação – questão 2

Com base nos dados obtidos no teste Two-Way ANOVA com interação presentes na figura 9 concluímos que o valor de P para a interação entre os algoritmos e os valores de N (2e-16) é inferior ao nível de significância de 5% e, portanto, a hipótese nula H^{AN}_{0} é rejeitada.

Tal como foi explicado em cima, isto significa que a hipótese alternativa H^{AN}_{1} é satisfeita.

Como só podemos concluir alguma coisa do Two-Way ANOVA se os pressupostos forem satisfeitos, iremos realizar o teste de normalidade de resíduos e o teste de homogeneidade da variância.

```
> shapiro.test(aov1.out$res)

        Shapiro-Wilk normality test

data:  aov1.out$res
W = 0.80008, p-value < 2.2e-16
```

Figura 10 - Teste de normalidade dos resíduos – questão 2

```
> bartlett.test(TamanhoMaximo~interaction(Algoritmo,N), data=Tabela_A_Variar_N_Global)

        Bartlett test of homogeneity of variances

data:  TamanhoMaximo by interaction(Algoritmo, N)
Bartlett's K-squared = 1398, df = 8, p-value < 2.2e-16
```

Figura 11 - Teste de homogeneidade da variância – questão 2

Depois de efetuados o teste de normalidade de resíduos (figura 10) e o teste de homogeneidade da variância (figura 11) os seus resultados mostram que, pelo facto de em ambos os casos os valores serem muitos pequenos (inferiores a 0.05), o que significa que os pressupostos da análise ANOVA não se verificam.

Concluindo-se então que irá ser necessário recorrer a métodos não paramétricos alternativos. O método usado será o teste de randomização sem restrições de permutações.

```
> print(pvalueA/5000)
[1] 0
> print(pvalueN/5000)
[1] 0
> print(pvalueAN/5000)
[1] 0
```

Figura 12 - Teste de randomização sem restrições de permutações – questão 2

Com a análise dos resultados do teste de randomização sem restrições de permutações visível na figura acima concluímos que a primeira e segunda hipóteses nulas elaboradas para responder à questão (H^A_0 e H^N_0 respetivamente) são refutadas. O que resulta na satisfação das suas hipóteses alternativas H^A_1 e H^N_1 respetivamente que no dizem que pelo menos dois algoritmos têm médias diferentes e a média dos tamanhos máximos de uma subsequência ordena é diferente para pelo menos dois valores de N diferentes.

A figura 12 permite-nos afirmar que existe interação entre os algoritmos e a variação dos valores de N no que respeita ao tamanho máximo da subsequência ordenada uma vez que a hipótese nula H^{AN}_0 é refutada.

Portanto podemos concluir e assim responder à pergunta 2 que, para os resultados obtidos neste estudo, a escolha de um par algoritmo-valor de N tem uma influência significativa no tamanho máximo da subsequência ordenada obtida.