****

**MINISTERUL EDUCAȚIEI, CULTURII ȘI CERCETĂRII**

**AL REPUBLICII MOLDOVA**

**Universitatea Tehnică a Moldovei**

**Facultatea Calculatoare, Informatică şi Microelectronică**

**Departamentul Informatică şi Ingineria Sistemelor**

**Elaborat: Calancea Catalin**

**Grupa: MI-222**

**Raport**

**pentru lucru individual**

**Tema:Teste statistice si simulari.**

***la cursul de “Probabilitate și Statistică Analitică”***

Verificat:

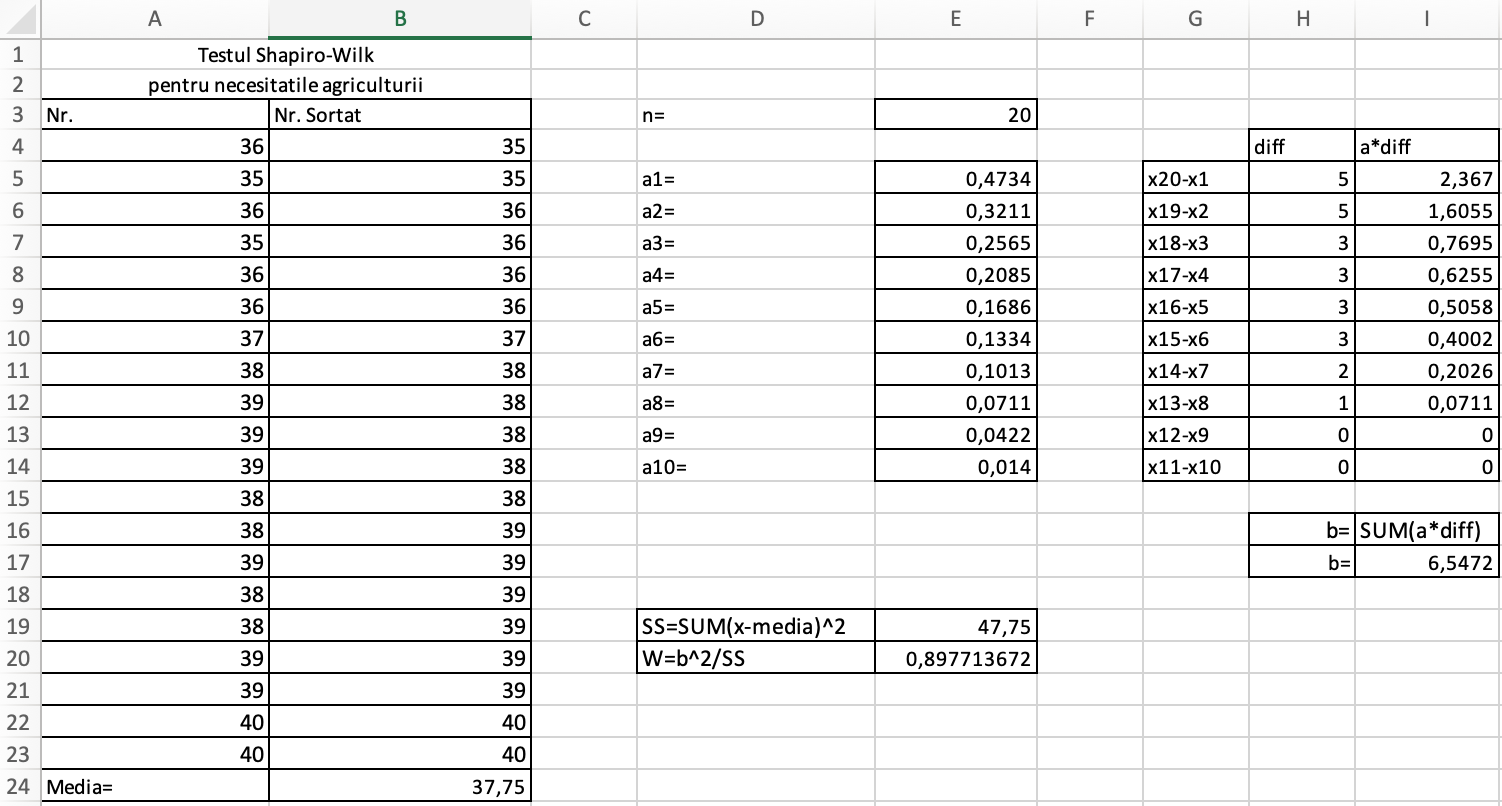
**M. Mantaluță,** *asistentă univ.*

Facultatea FCIM, UTM

**Chișinău – 2022**

| **Sarcina 1:** Analiza statistică a datelor despre indicatorii principali privind utilizarea apei (pentru necesitatile agriculturii , pentru irigare) pentru perioada anilor 2002-2021, în Republica Moldova, folosind testul Shapiro-Wilk și testul χ2 (chi-pătrat) Pearson.  **Indicatorii principali privind utilizarea apei pentru necesitatile agriculturii 2002-2021** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** |
| ..pentru necesitatile agriculturii | 36 | 35 | 36 | 35 | 36 | 36 | 37 | 38 | 39 | 39 | 39 | 38 | 38 | 39 | 38 | 38 | 39 | 39 | 40 | 40 |

* **Testul Shapiro-Wilk (pentru necesitatile agriculturii):**



* Observăm că valoarea W (0,8977) este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mică decât 0.05 deci concluzionăm că ipoteza nu este acceptată.

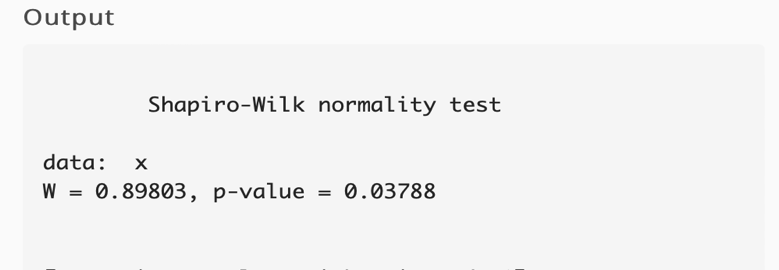
***Codul in R***

x <- c(36,35,36,35,36,36,37,38,39,39,39,38,38,39,38,38,39,39,40,40)

shapiro.test(x)

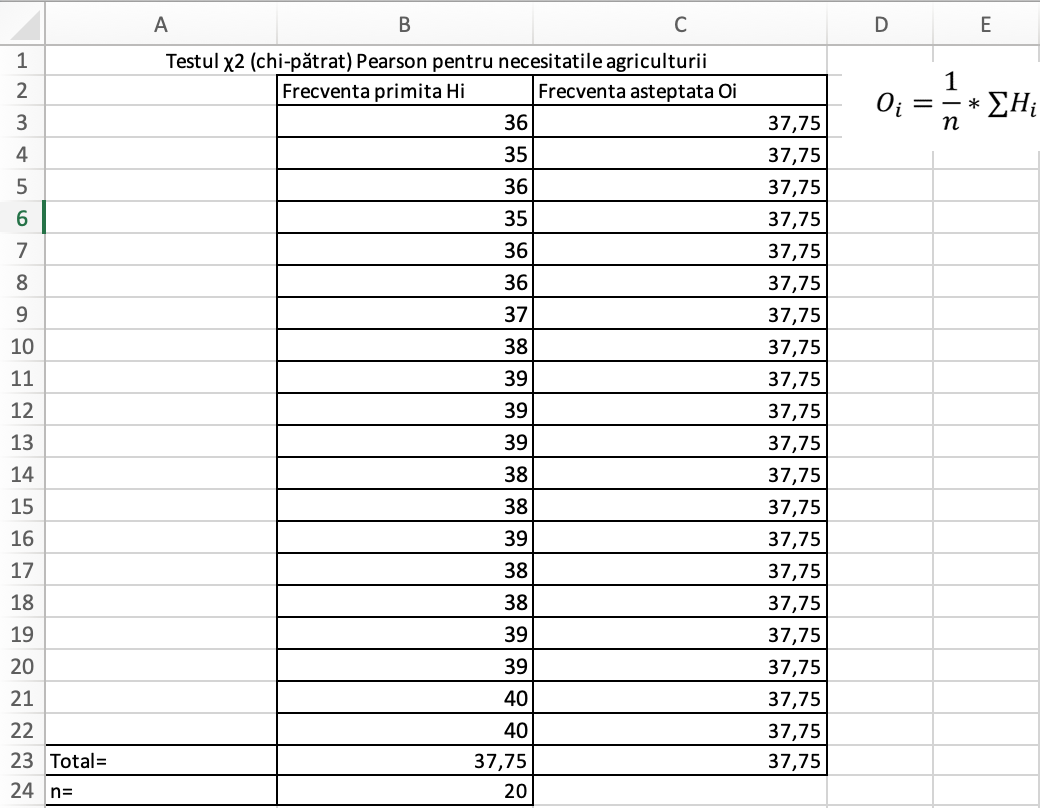
qqnorm(x)

qqline(x)





* **Testul χ2 (chi-pătrat) Pearson** **(pentru necesitatile agriculturii):**



k=n-1 = 20-1=19 (df, *Gradul de libertate)*

+ + + + ++ + + + + ++ 1,2649

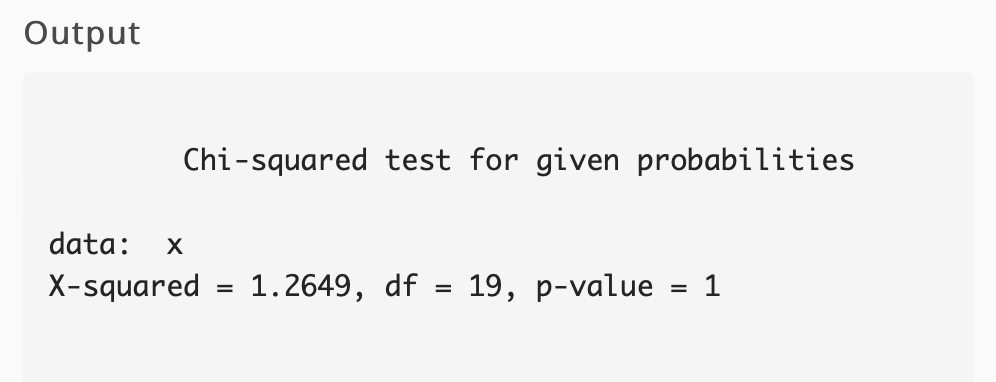
Valoarea calculata (1,2649) < Valoarea critica 0,05 ; 19 (30,1435)

* Ipoteza nulă se acceptă fiindcă valoarea calculată este mai mica decât valoarea critică identificată în cadrul tabelului chi-pătrat**.**

***Codul in R***

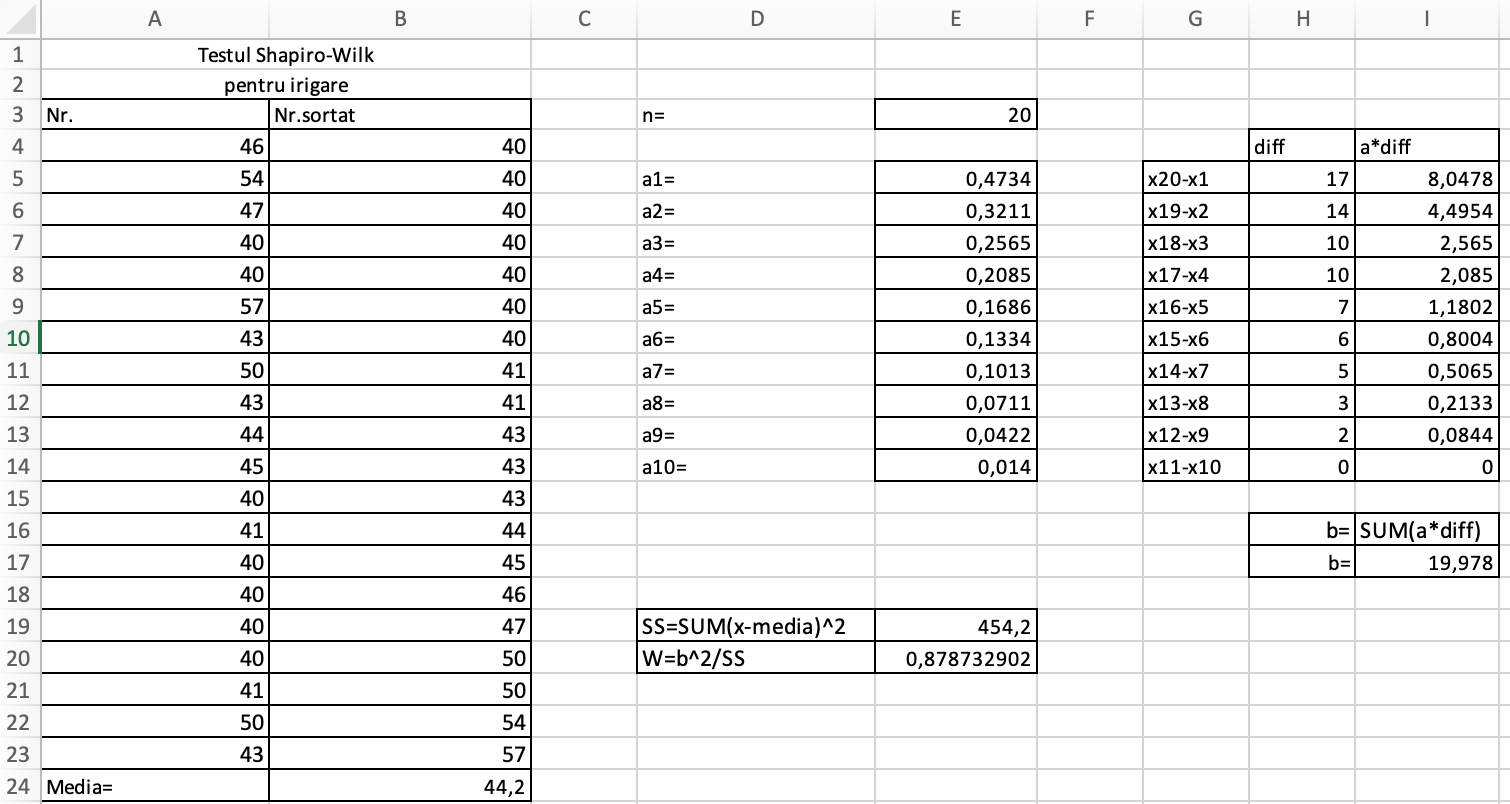
x <- c(36,35,36,35,36,36,37,38,39,39,39,38,38,39,38,38,39,39,40,40)

chisq.test(x)



| **Indicatorii principali privind utilizarea apei pentru irigare 2002-2021** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** |
| ..pentru irigare | 46 | 54 | 47 | 40 | 40 | 57 | 43 | 50 | 43 | 44 | 45 | 40 | 41 | 40 | 40 | 40 | 40 | 41 | 50 | 43 |

* **Testul Shapiro-Wilk (pentru irigare):**



* Observăm că valoarea W (0,8787) este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mică decât 0.05 deci concluzionăm că ipoteza nu este acceptată.

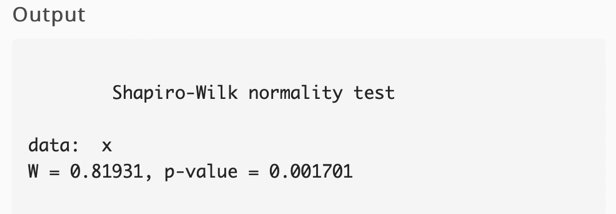
***Codul in R***

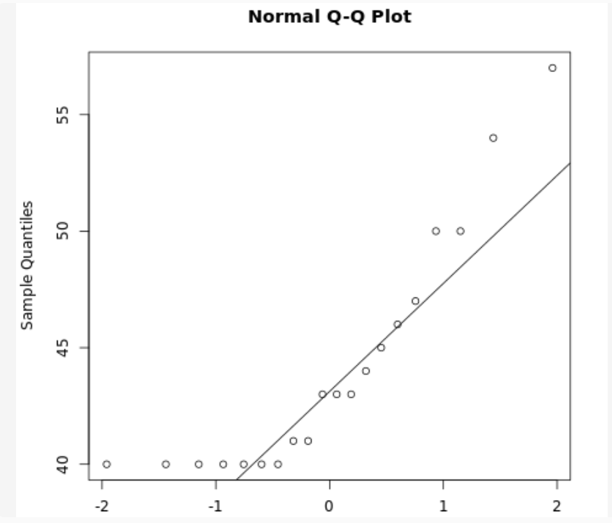
x <- c(46,54,47,40,40,57,43,50,43,44,45,40,41,40,40,40,40,41,50,43)

shapiro.test(x)

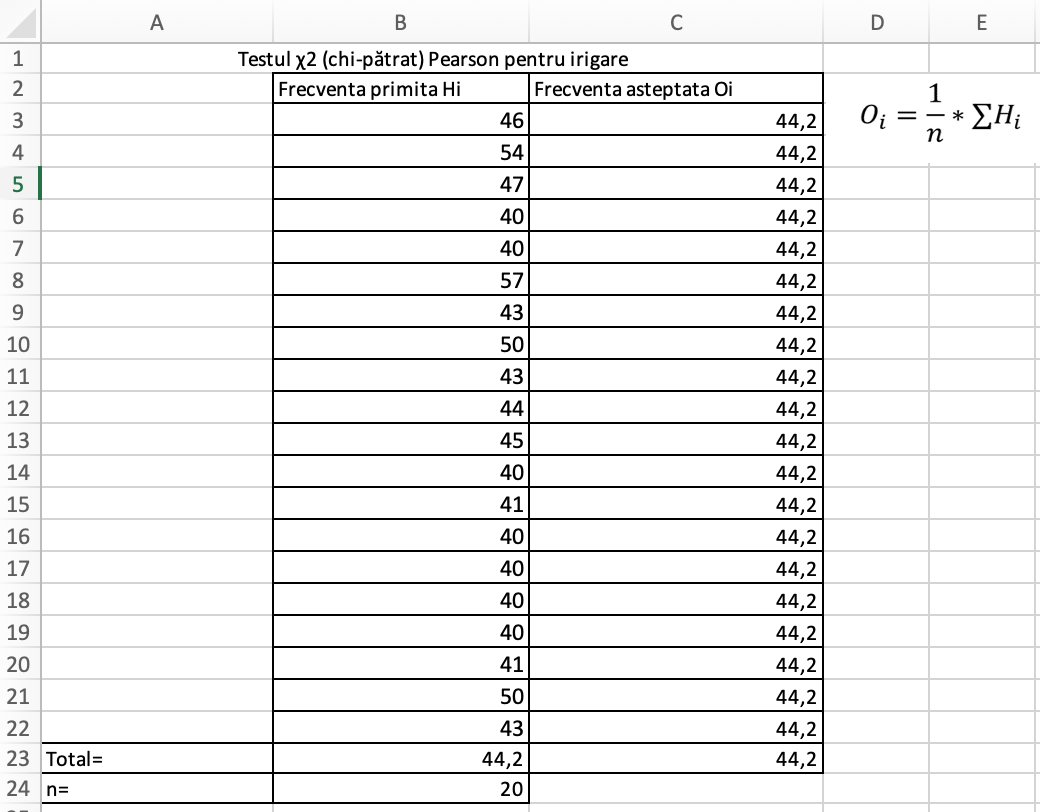
qqnorm(x)

qqline(x)





* **Testul χ2 (chi-pătrat) Pearson** **(pentru irigare):**



+ + + + ++ + + + + ++ 11,023

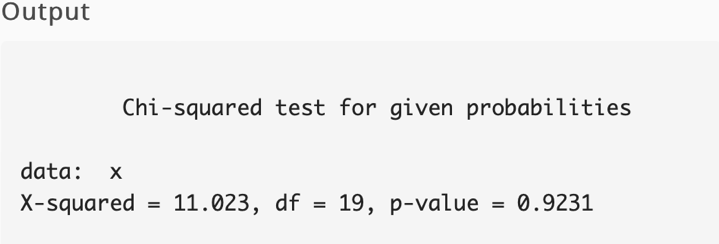
Valoarea calculata (11,023) < Valoarea critica 0,05 ; 19 (30,1435)

* Ipoteza nulă se acceptă fiindcă valoarea calculată este mai mica decât valoarea critică identificată în cadrul tabelului chi-pătrat**.**

***Codul in R***

x <- c(46,54,47,40,40,57,43,50,43,44,45,40,41,40,40,40,40,41,50,43)

chisq.test(x)



**Sarcina 2:** Analiza statistică a datelor privind viteza medie lunara (iunie, septembrie) a vântului în Republica Moldova, pentru perioada anilor 2002-2021, folosind testul Shapiro-Wilk și testul tStudent.

| **Viteza medie a vintului pe luniile (iunie si septembrie), Statia meteorologica (Briceni) si anii 2002-2021** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Briceni** | | | | | | | | | | | | | | | | | | | |
| **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** |
| Iunie | 2,3 | 2,2 | 2,2 | 2,0 | 1,9 | 1,7 | 1,7 | 2,1 | 2,1 | 2,3 | 1,8 | 1,7 | 1,9 | 1,8 | 1,8 | 1,9 | 1,9 | 1,7 | 1,7 | 1,9 |
| Septembrie | 1,6 | 2,0 | 1,6 | 1,5 | 2,3 | 1,8 | 1,7 | 1,5 | 2,0 | 1,7 | 2,0 | 2,4 | 1,9 | 2,0 | 1,6 | 2,4 | 1,8 | 1,8 | 1,9 | 1,9 |

* **Testul Shapiro-Wilk (viteza media a vintului in luna iunie):**

## 

* Observăm că valoarea W (0,9705) nu este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mare decât 0.05 deci concluzionăm că ipoteza este acceptată.

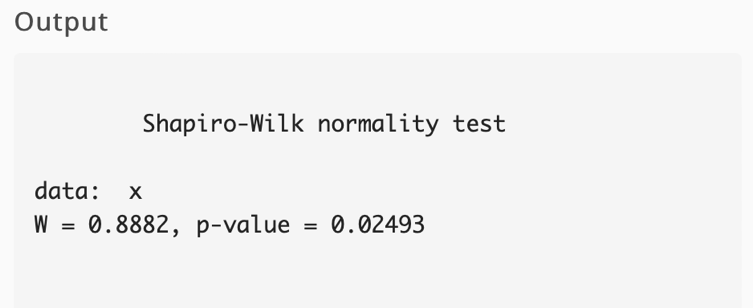
***Codul in R***

x <- c(2.3,2.2,2.2,2,1.9,1.7,1.7,2.1,2.1,2.3,1.8,1.7,1.9,1.8,1.8,1.9,1.9,1.7,1.7,1.9)

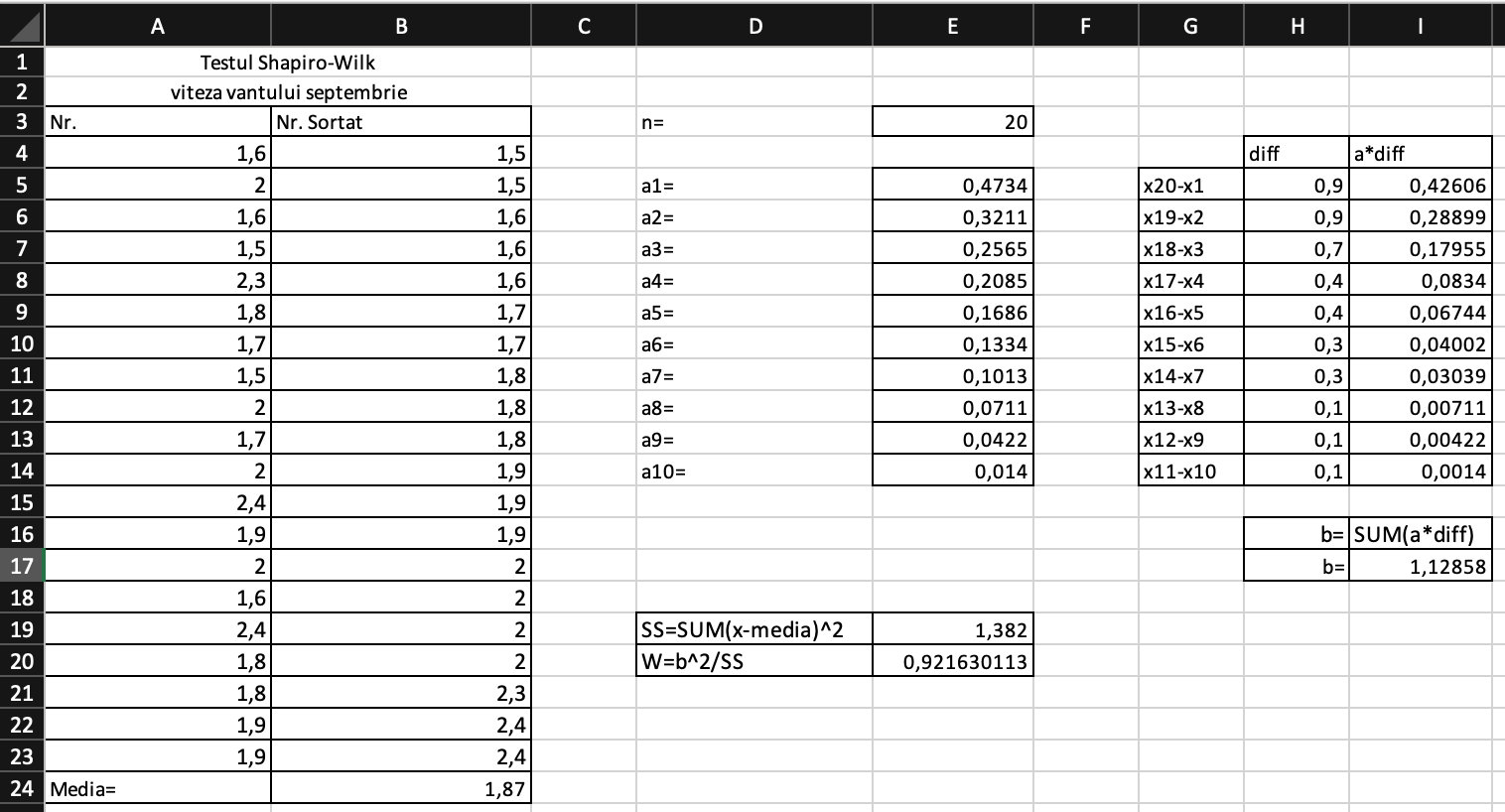
shapiro.test(x)

qqnorm(x)

qqline(x)

******

* **Testul Shapiro-Wilk (viteza media a vintului in luna septembrie):**



* Observăm că valoarea W (0,9216) nu este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mare decât 0.05 deci concluzionăm că ipoteza este acceptată.

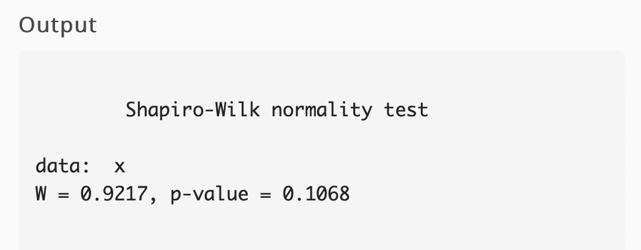
***Codul in R***

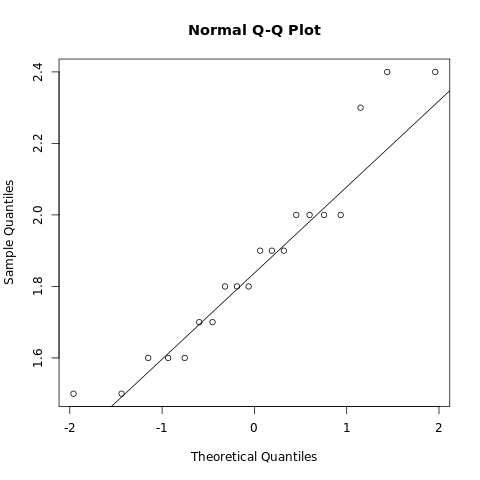
x <- c(1.6,2,1.6,1.5,2.3,1.8,1.7,1.5,2,1.7,2,2.4,1.9,2,1.6,2.4,1.8,1.8,1.9,1.9)

shapiro.test(x)

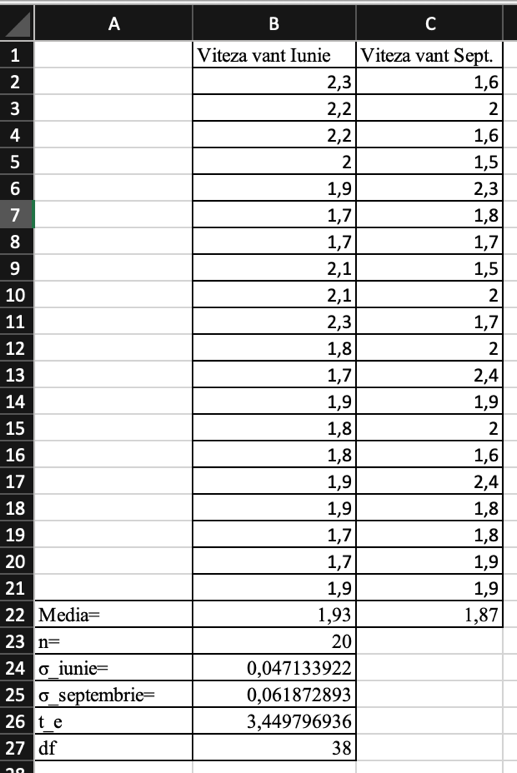
qqnorm(x)

qqline(x)





* **Testul t-Student (viteza media a vintului in luna iunie,septembrie ):**



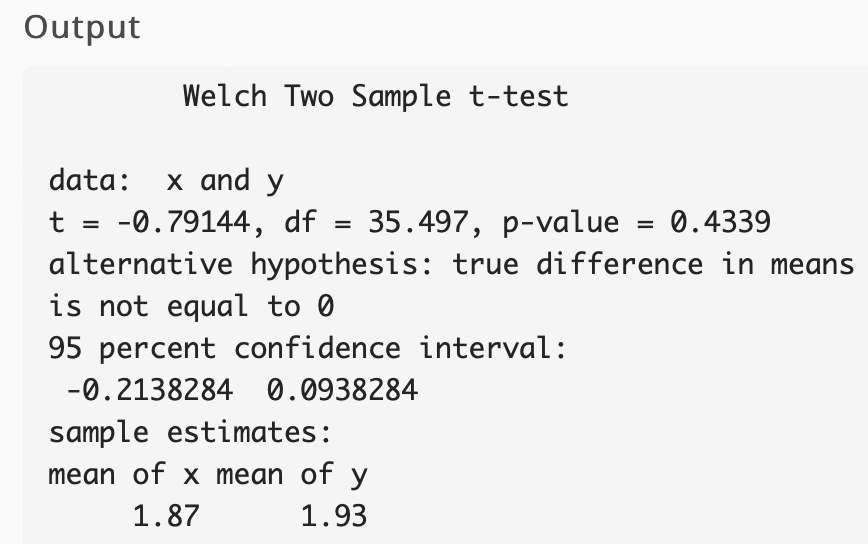
Calcularea gradului de liberate:

***Codul in R***

x <- c(1.6,2,1.6,1.5,2.3,1.8,1.7,1.5,2,1.7,2,2.4,1.9,2,1.6,2.4,1.8,1.8,1.9,1.9)

y <- c(2.3,2.2,2.2,2,1.9,1.7,1.7,2.1,2.1,2.3,1.8,1.7,1.9,1.8,1.8,1.9,1.9,1.7,1.7,1.9)

t.test(x,y)



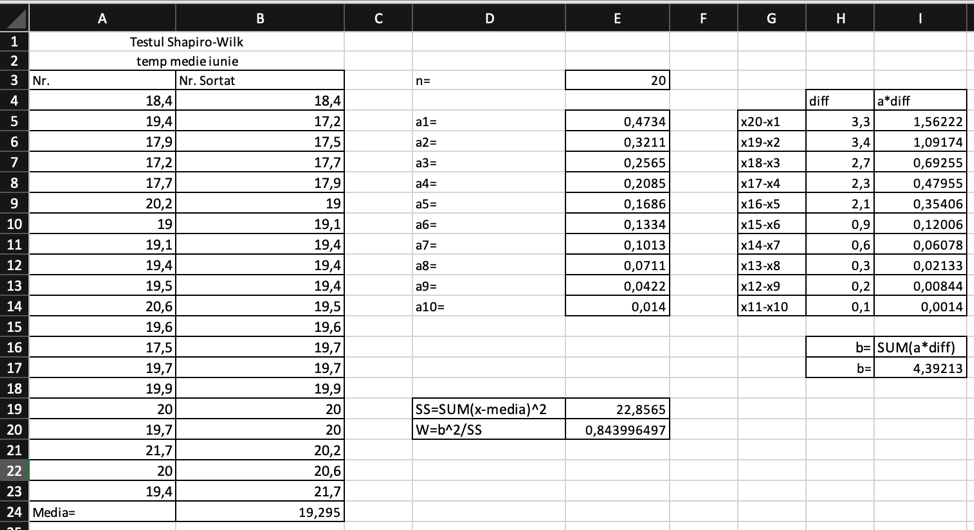
* p-value > 0.05 => Pentru testul t-Student ipoteza nulă se acceptă fiindcă p-value = 0.4339 și este mai mare decât 0.05.

| **Temperatura medie a aerului pe luniile (iunie si septembrie), Statia meteorologica (Briceni) si anii 2002-2021** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Briceni** | | | | | | | | | | | | | | | | | | | |
| **2002** | **2003** | **2004** | **2005** | **2006** | **2007** | **2008** | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **2021** |
| Iunie | 18,4 | 19,4 | 17,9 | 17,2 | 17,7 | 20,2 | 19,0 | 19,1 | 19,4 | 19,5 | 20,6 | 19,6 | 17,5 | 19,7 | 19,9 | 20,0 | 19,7 | 21,7 | 20,0 | 19,4 |
| Septembrie | 14,4 | 14,4 | 14,2 | 15,9 | 15,5 | 14,5 | 13,6 | 16,7 | 13,9 | 16,7 | 17,1 | 12,9 | 15,8 | 17,8 | 17,0 | 16,0 | 16,3 | 16,1 | 17,8 | 13,5 |

* Valoarea critică t(1.684) < astfel putem concluziona că ipoteza nulă nu este acceptată.

**Sarcina 3:**Prelucrarea statistică a datelor privind temperatura medie lunara(iunie, septembrie) a aerului în Republica Moldova Chișinău, pentru perioada anilor 2002-2021, cu ajutorul testului Shapiro-Wilk și testului Fisher F-test.

* **Testul Shapiro-Wilk (pentru temperatura media a aerului in luna iunie):**



* Observăm că valoarea W (0,8439) este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mica decât 0.05 deci concluzionăm că ipoteza nu este acceptată.

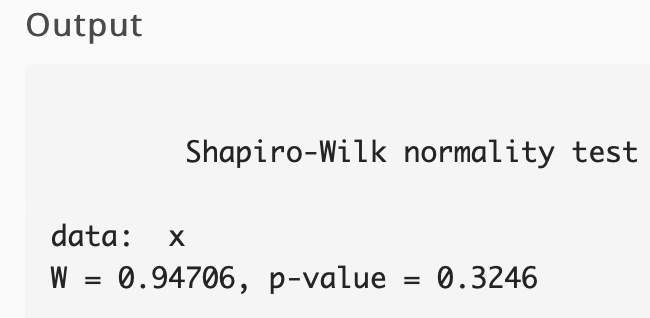
***Codul in R***

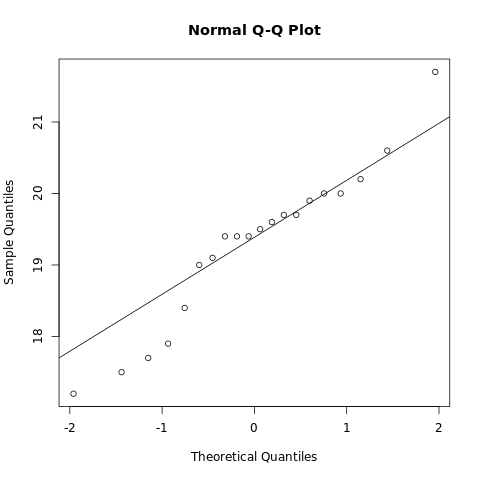
x <-c(18.4,19.4,17.9,17.2,17.7,20.2,19,19.1,19.4,19.5,20.6,19.6,17.5,19.7,19.9,20,19.7,21.7,20,19.4)

shapiro.test(x)

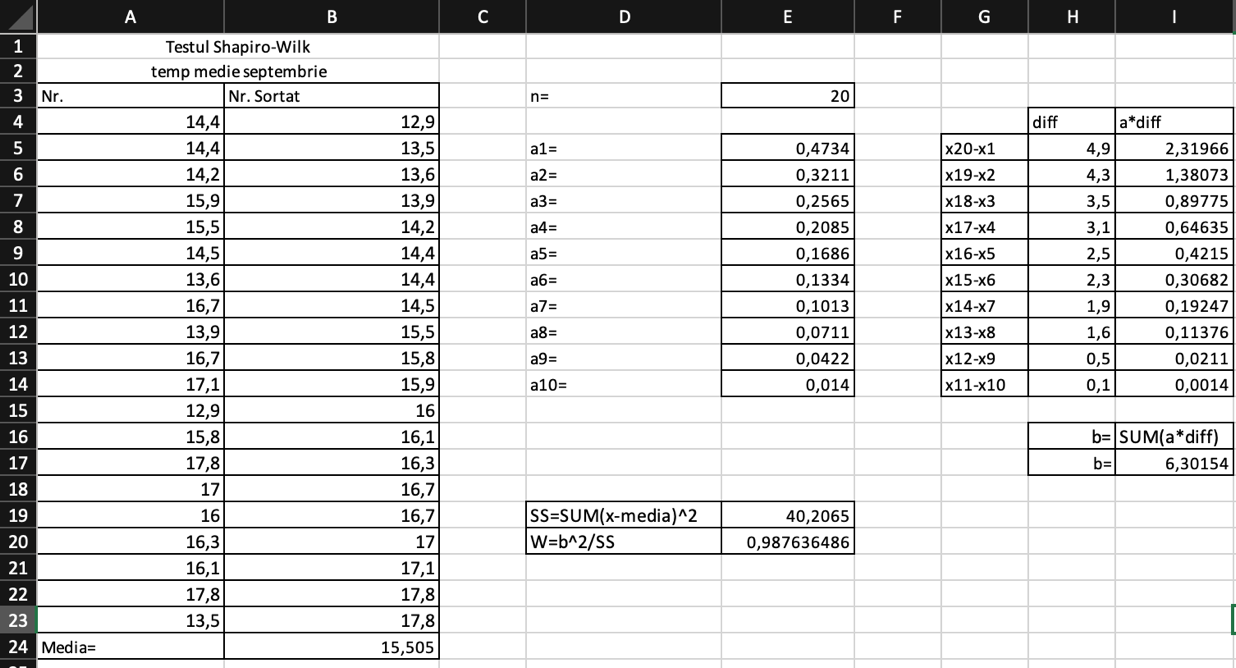
qqnorm(x)

qqline(x)





* **Testul Shapiro-Wilk (pentru temperatura media a aerului in luna septembrie):**



* Observăm că valoarea W (0,9876) nu este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mare decât 0.05 deci concluzionăm că ipoteza este acceptată.

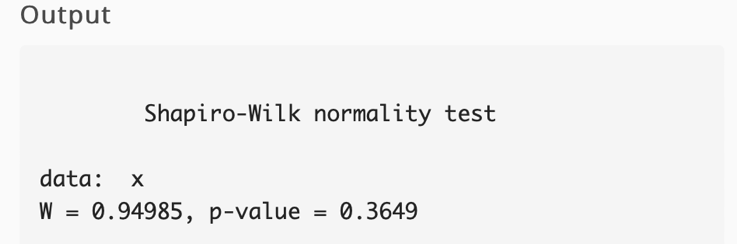
***Codul in R***

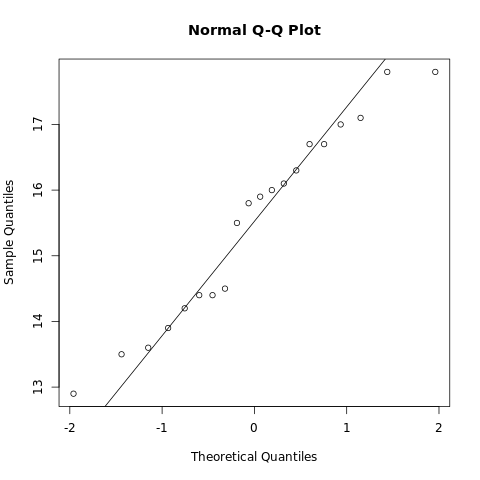
x <- c(14.4,14.4,14.2,15.9,15.5,14.5,13.6,16.7,13.9,16.7,17.1,12.9,15.8,17.8,17,16,16.3,16.1,17.8,13.5)

shapiro.test(x)

qqnorm(x)

qqline(x)





* **Testul Fisher f-Test (pentru temperatura media a aerului in lunile iunie , septembrie):**



=1,1676

=2,0788

=1,2205

=2,2005

Calculăm gradul de libertate:

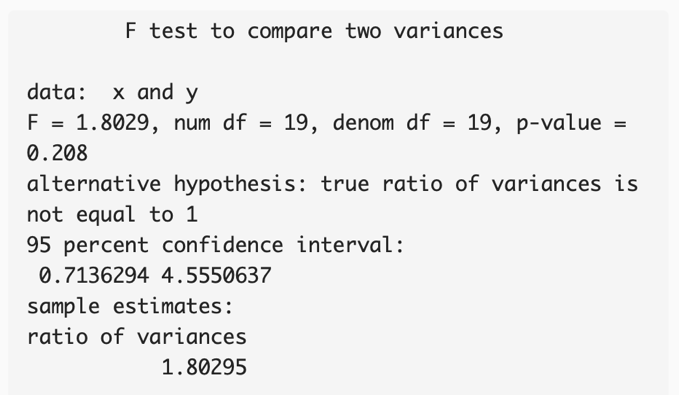
* Valoarea critică F( > F(calculată -) astfel putem concluziona că ipoteza nulă este acceptată.

***Codul in R***

x <- c(14.4,14.4,14.2,15.9,15.5,14.5,13.6,16.7,13.9,16.7,17.1,12.9,15.8,17.8,17,16,16.3,16.1,17.8,13.5)

y <- c(18.4,19.4,17.9,17.2,17.7,20.2,19,19.1,19.4,19.5,20.6,19.6,17.5,19.7,19.9,20,19.7,21.7,20,19.4)

var.test(x,y)



| **Institutii de invatamant superior ,Fara frecventa, Masculin si 2000/01-2019/20** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2000/01** | **2001/02** | **2002/03** | **2003/04** | **2004/05** | **2005/06** | **2006/07** | **2007/08** | **2008/09** | **2009/10** | **2010/11** | **2011/12** | **2012/13** | **2013/14** | **2014/15** | **2015/16** | **2016/17** | **2017/18** | **2018/19** | **2019/20** |
| Masculin | 11 469 | 12 274 | 14 613 | 15 918 | 16 643 | 20 787 | 20 676 | 18 595 | 17 128 | 14 373 | 13 518 | 13 724 | 14 685 | 14 718 | 14 086 | 12 445 | 11 765 | 10 227 | 9 768 | 8 772 |

**Sarcina** 4**:** Aplicarea testului Shapiro-Wilk și testului Kolmogorov-Smirnov privind numărul de studenți în institutiile de învățământ superior pe forme de învățământ(Fara Frecvență) și sexe (Masculin, Feminin), pentru anii de studii 2000/01-2019/20, în Republica Moldova.

* **Testul Shapiro-Wilk (pentru studii fara frecventa masculin 2000/01-2019/20):**

## 

* Observăm că valoarea W (1,0023) nu este în afara limitei intervalului din cadrul tabelului 2 (0,905) , deci putem constata că valoare p este mai mare decât 0.05 deci concluzionăm că ipoteza este acceptată.

***Codul in R***

x <- c(11469,12274,14613,15918,16643,20787,20676,18595,17128,14373,13518,13724,14685,14718,14086,12445,11765,10227,9768,8772)

shapiro.test(x)

qqnorm(x)

qqline(x)

## 

## 

| **Institutii de invatamant superior, Fara frecventa, Feminin, 2000/01-2019/20** | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **2000/01** | **2001/02** | **2002/03** | **2003/04** | **2004/05** | **2005/06** | **2006/07** | **2007/08** | **2008/09** | **2009/10** | **2010/11** | **2011/12** | **2012/13** | **2013/14** | **2014/15** | **2015/16** | **2016/17** | **2017/18** | **2018/19** | **2019/20** |
| Feminin | 13 014 | 15 131 | 16 833 | 19 531 | 23 044 | 25 474 | 26 120 | 24 854 | 23 488 | 17 288 | 16 553 | 16 392 | 17 520 | 18 215 | 17 503 | 15 688 | 14 293 | 12 739 | 12 359 | 11 905 |

* **Testul Shapiro-Wilk (pentru studii fara frecventa feminin 2000/01-2019/20):**

## 

x <- c(13014,15131,16833,19531,23044,25474,26120,24854,23488,17288,16553,16392,17520,18215,17503,15688,14293,12739,12359,11905)

shapiro.test(x)

qqnorm(x)

qqline(x)

## 

## 

* **Testul Kolmogorov – Smirnov (pentru studii fara frecventa 2000/01-2019/20):**

## 

Dstat = 0.15710802

Dcrit = 0,000125371

* Ipoteza nulă este respinsă deoarece Dcrit< Dstat
* Codul în limbajul R:

x <- c(11469,12274,14613,15918,16643,20787,20676,18595,17128,14373,13518,13724,14685,14718,14086,12445,11765,10227,9768,8772)

y <- c(13014,15131,16833,19531,23044,25474,26120,24854,23488,17288,16553,16392,17520,18215,17503,15688,14293,12739,12359,11905)

plot(ecdf(x))

xlim = range(c(x,y))

col = "blue"

plot(ecdf(y))

add = TRUE

lty = "dashed"

col = "red"

ks.test(x,y)

## 

## 

* Reieșind din informația obținută după execuția funcției în limbajul R putem afirma ferm că ipoteza nulă este respinsă deoarece a=0.05 iar p < a.

## Concluzie:In urma realizarii lucrului individual , am aplicat testele statistice , la care am aplicat simularile in codul R , pentru unele teste a fost nevoie de a ordona valorile crescator , dupa care am cautat coeficientii pentru unele teste si le-am anexat in anexa pentru a compara valorile primite cu cele din tabel , si sa deducem daca ipoteza poate fi sau nu acceptata. In cazul unor teste am aplicat studiu pentru doua variabile , cum ar fi viteza medie a vantului pentru iunie si septembrie , sau studiu in institutile superioare fara frecventa a genului masculine si feminin.Am folosit testul Kolmogorov-Smirnov pentru a compara proba de distributie , dupa care am comparat Distributia calculate si distributia critica.

## Anexa

## 

## Figura 1. Coeficinetii pentru testul Shapiro-Wilk

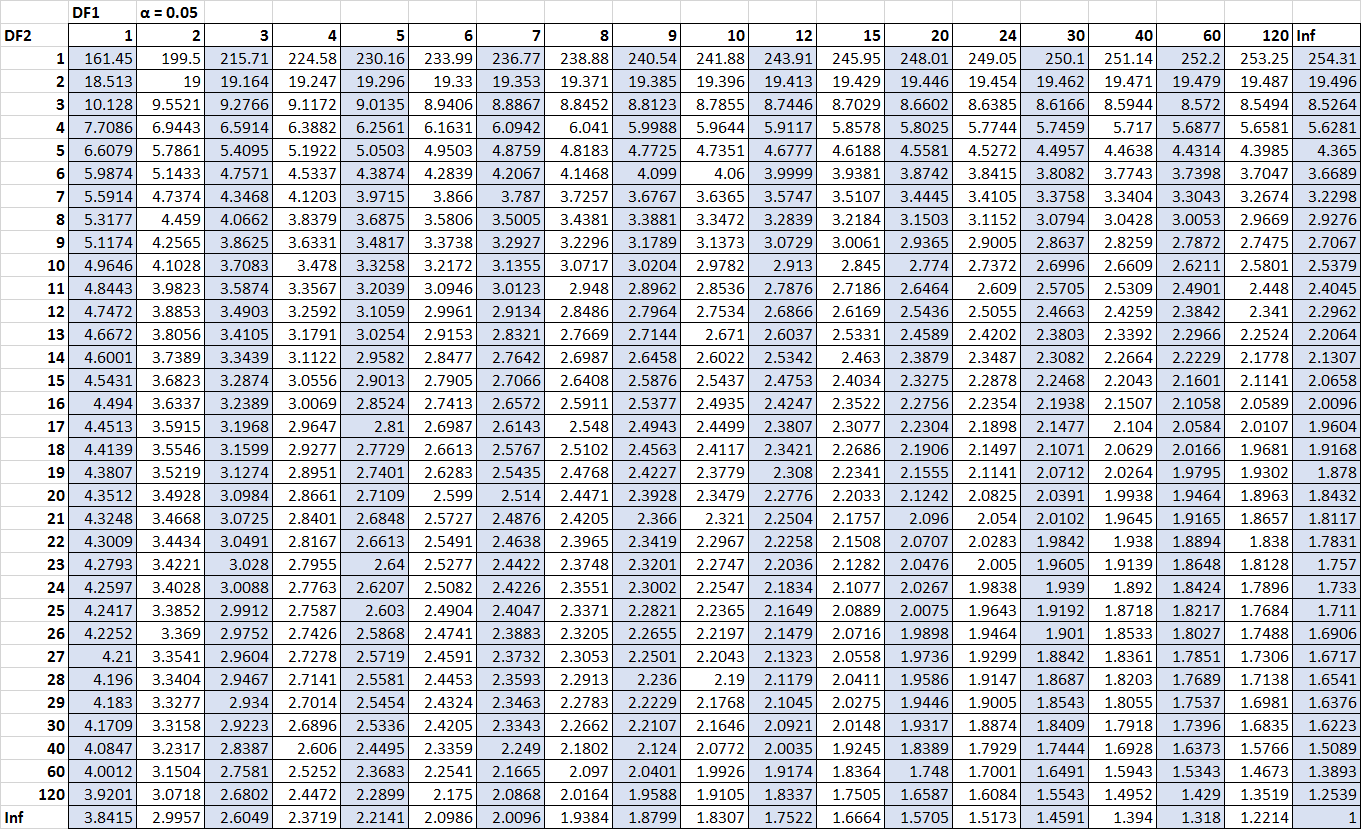
## 

## Figura 2. p-value coefficient pentru testul Shapiro-Wilk

## Figura 3. Valoarea critica pentru testul Testul χ2 (chi-pătrat) Pearson

## T Table - T Table

## Figura 4. Tabel de distribuție pentru Testul t-Student



## Figura 5. Tabel de distribuție pentru Testul Fisher