

Veți lucra cu fișierul de date "Date studenti.xls" de pe drive, din folderul **Date de lucru**, care este descris în documentul "Date Studenti descriere.pdf".

Setul de date propus reprezintă o colecție organizată de date culese voluntar de la o serie de studenți din cadrul Facultății de Medicină din anul I de studiu.

Pentru început, facem o mică introducere în tema dată (o sinteză a cursului).

Teste pentru variabile numerice (normal distribuite):

1. TESTUL t NEPERECHE

condiție:

- seriile de valori provin din determinări pe **indivizi diferiți**, adică **seriile sunt independente**
- se compară două valori medii (μ_1 și μ_2) obținute pe **serii independente** cu valori numerice **distribuite normal**

ipotezele statistice:

- ipoteza de nul, H_0 : Între cele două valori medii nu sunt diferențe (diferențele sunt nesemnificative)
 - formula: $\mu_1 = \mu_2$
- ipoteza alternativă, H_a : Între cele două valori medii există diferențe
 - formula: $\mu_1 \neq \mu_2$ ipoteză alternativă bilaterală
 $\mu_1 > \mu_2$ sau $\mu_1 < \mu_2$ ipoteză alternativă unilaterală

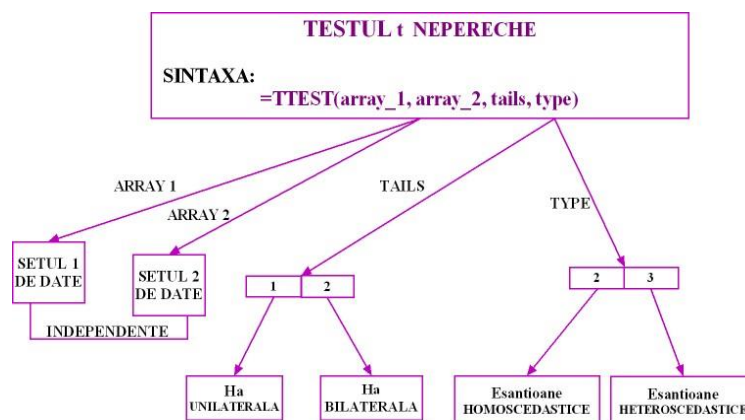


figura 1. Schema explicativă a comenzii din excel, aferentă testului t nepereche

2. TESTUL t PERECHE

condiție:

- seriile de valori provin din determinări pe **aceiași indivizi** în condiții diferite, adică **seriile sunt dependente**
- se compară două valori medii obținute pe **serii perechi și variabile distribuite normal**

ipotezele statistice:

- ipoteza de nul, H_0 : Între cele două valori medii nu sunt diferențe (media diferentelor notată μ_d este nula; diferențele sunt ne semnificative)
 - formula: $\mu_d = 0$
- ipoteza alternativă, H_a : Între cele două valori medii există diferențe (media diferentelor este nenula)
 - formula: $\mu_d \neq 0$ ipoteză alternativă bilaterală
 $\mu_d > 0$ sau $\mu_d < 0$ ipoteză alternativă unilaterală

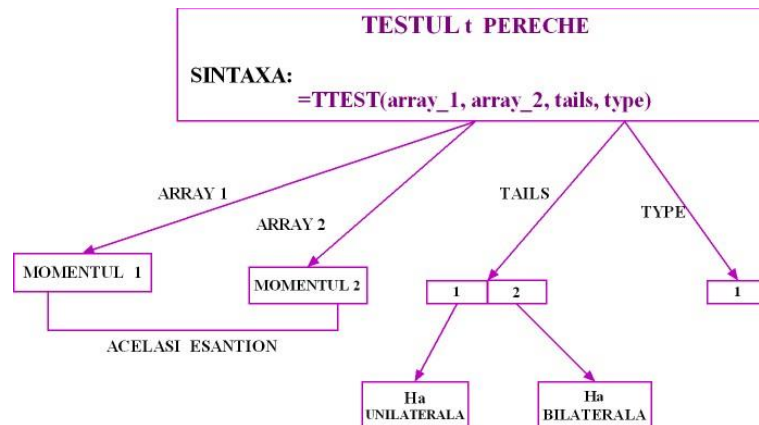


figura 2. Schema explicativă a comenzii din excel, aferentă testului t pereche

În urma aplicării oricărui test statistic vom obține valoarea probabilității ca diferențele să fie întâmplătoare (p), după care luăm decizia statistică astfel:

Teste statistice uzuale (I): teste t pereche și nepereche
Aplicații în MS Excel și/sau EpiInfo

sl.dr. Lavinia Moleriu, sl. dr. Anca Tudor

$p \geq 0,05 \rightarrow H_0$ se accepta, diferențele sunt ne semnificative

$p < 0,005 \rightarrow H_0$ se respinge, diferențele sunt semnificative

$p < 0,01 \rightarrow H_0$ se respinge, diferențele sunt foarte semnificative

$p < 0,001 \rightarrow H_0$ se respinge, diferențele sunt extrem de semnificative

Exerciții propuse:

1. Diferă înălțimea fetelor de cea a băieților?

Aveți de rezolvat următoarele sarcini pentru fiecare exercițiu:

- Formularea ipotezelor de cercetare
- Formularea ipotezelor statistice
- Calculul indicatorilor de tendință centrală și de dispersie
- Alegerea și aplicarea testului statistic
- Decizia statistică
- Interpretarea rezultatelor

2 sugestii de rezolvare în Excel:

- pag.114-117 din “BTimar-MetodeBiostatisticaAplicate-2016.pdf”
 - sortați datele după variabila sex
- alegeți opțiunea: Data → Data Analysis → alegeți testul t nepereche (figura 3)

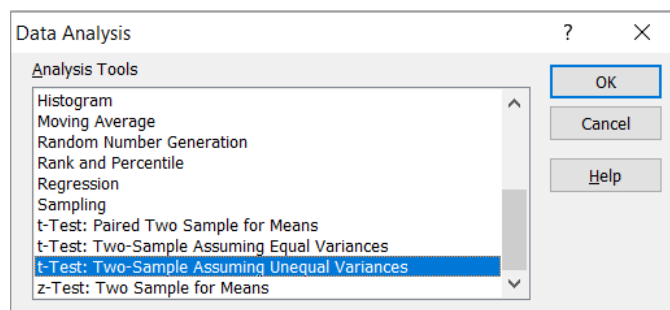


figura 3. Testul t nepereche cu varianțe inegale

Teste statistice uzuale (I): teste t pereche și nepereche
Aplicații în MS Excel și/sau EpiInfo

sl.dr. Lavinia Moleriu, sl. dr. Anca Tudor

- selectați valorile înălțimii (figura 4)

t-Test: Two-Sample Assuming Unequal Variances

Input

Variable 1 Range:

Variable 2 Range:

Hypothesized Mean Difference:

☐ Labels

Alpha:

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

OK Cancel Help

figura 4. Introducerea datelor în fereastra de calcul

- rezultatele vor arăta ca în figura 5

	masculin	feminin
t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	Variable 2
Mean	178,2380952	165,9148936
Variance	48,09047619	46,20999075
Observations	21	47
Hypothesized Mean Difference	0	
df	38	
t Stat	6,811399547	
P(T<=t) one-tail	2,21534E-08	
t Critical one-tail	1,68595446	
P(T<=t) two-tail	4,43068E-08	
t Critical two-tail	2,024394164	

figura 5. Rezultatele aplicării testului

Din rezultate extragem valoarea probabilității $p=4,43068E-08$, ceea ce înseamnă că $p<0,001$, adică diferențele între înălțimile studenților după sexe sunt extrem de semnificative; media pentru studenții de sex masculin e semnificativ crescută față de media înălțimii studentelor.

sugestie de rezolvare în EpiInfo:

pag.104-108 din “BTimar-MetodeBiostatisticaAplicate-2016.pdf”

2. Diferă notele de la admitere după opinia despre biostatistică (dificilă/nedificilă)?

- rezolvați cu ce program doriți și scrieți răspunsurile la sarcini.

3. Sunt mulțumiți studenții din anul I de masa pe care o au?

Sugestie: comparați masa reală cu masa0 (cea ideală), cu programul Excel (în EpiInfo este mai laborios...).

Răspunsurile le editați într-un document word, cu explicații și/sau capturi de ecran, fișier denumit după șablonul ”Nume.prenume.MG_grupa”, pe care îl veți trimite la adresa de e-mail a cadrului didactic de predare de lucrări practice.

Vă puteți documenta din cursurile “noteCurs infoSupl” și 02 în format pdf, din folderul ”Diapozitive”, din cartea de curs ”Mihalas-Lungeanu-Biostatistica-EVB2011.pdf” și din “BTimar-MetodeBiostatisticaAplicate-2016.pdf” de pe drive.

Solutii:

2. Diferă notele de la admitere după opinia despre biostatistică (dificilă/nedificilă)?

Aveți de rezolvat următoarele sarcini pentru fiecare exercițiu:

a. Formularea ipotezelor de cercetare

dorim sa vedem daca difera semnificativ notele studentilor de la admitere dupa opinia lor despre biostatistica (considerate dificila / nedificila)

b. Formularea ipotezelor statistice

ipoteza de nul H_0 : **diferențele** între media de la admitere pentru cei care considera biostatistica dificila si media de la admitere pentru cei care considera biostatistica nedificila **sunt nesemnificative**

$$\mu_D = \mu_N$$

ipoteza alternative H_a : diferențele sunt semnificative

bilaterală: $\mu_D \neq \mu_N$

unilaterală: $\mu_D > \mu_N$ sau $\mu_D < \mu_N$

c. Calculul indicatorilor de tendință centrală și de dispersie

Data Analysis → Descriptive statistics (excel)

MEANS ADMITERE Biostat_dificil (EpiInfo)

d. Alegerea și aplicarea testului statistic

testul t nepereche: cu variante egale, deoarece după aplicarea testului Fisher am obținut $p > 0.05$

adica ipoteza de nul, H_0 : $s_D = s_N$

MEANS ADMITERE Biostat_dificil (EpiInfo)

e. Decizia statistică

$p = 0.852 > 0.05 \rightarrow H_0$ se accepta, diferențele sunt nesemnificative

f. Interpretarea rezultatelor

Media notelor de la admitere pentru cei care considera biostatistica dificila este nesemnificativ mai mare decat media notelor de la admitere pentru cei care nu considera biostatistica dificila.

3. *Sunt mulțumiți studenții din anul I de masa pe care o au?*

a. Formularea ipotezelor de cercetare

dorim sa vedem daca difera semnificativ masa studentilor de cea pe care si-ar dori sa o aiba (ideala)

b. Formularea ipotezelor statistice

ipoteza de nul H_0 : **diferentele sunt nesemnificative** (media diferentelor dintre masa si masa0 e nula)
 $\mu_D=0$

ipoteza alternative H_a : diferentele sunt semnificative

bilaterală: $\mu_D \neq 0$

unilaterală: $\mu_D > 0$ sau $\mu_D < 0$

c. Calculul indicatorilor de tendință centrală și de dispersie

Data Analysis → Descriptive statistics (excel)

d. Alegerea și aplicarea testului statistic

testul t pereche deoarece avem de comparat valori provenite de la aceeasi studenti, in doua conditii diferite

Data Analysis → t-Test: Paired Two Sample for Means

e. Decizia statistică

$p=0.0044 < 0.01 \rightarrow H_0$ se respinge, H_a se accepta, diferentele sunt foarte semnificative

f. Interpretarea rezultatelor

Masa studentilor este semnificativ crescuta fata de cea ideala.