



Analiza rezultatelor cercetării - statistică Inferențială

**Universitatea
Politehnica
Timișoara**

Drd. Timea CISMA

Dată	Temă atelier
20.03.2024	Chestionarul – scop,obiective și proiectare
03.04.2024	Implementarea și aplicarea chestionarului
18.04.2024	Analiza rezultatelor cercetării – statistică descriptivă
24.04.2024	Analiza rezultatelor cercetării – statistică inferențială

Statistică

- Ramură a matematicii care implică colectarea, analiza, interpretarea, prezentarea și organizarea datelor
- Folosește calculul probabilităților, studiază fenomenele și procesele de tip colectiv din punct de vedere cantitativ

Statistică descriptivă = informații despre datele disponibile

Statistică inferențială = relația dintre datele disponibile

Tipul datelor

Datele diferă în funcție de natura lor și tipul de valori pe care le pot avea:

Date categoriale (calitative)

- Nominale – fără ordine (gen, culoare...)
- Ordinale – cu ordine dar cu intervale care nu sunt constate (nivelul de educație, statutul socio-economic..)
- Binare – pot lua doar 2 valori (da/nu, prezent/absent..)

Date numerice (cantitative)

- Discrete – valori distincte (număr de mașini, copii..)
- Continue – orice număr (înălțime, greutate)

Statistică inferențială

- În timp ce statisticile descriptive rezumă caracteristicile unui eșantion, statisticile inferențiale folosesc eșantionul pentru a face presupuneri despre o populație mai mare
- Scopul principal este de a trage concluzii despre o populație mai largă
- **Variabila dependentă** este observată sau măsurată în studiu
- **Variabila independentă** este manipulată

Testarea ipotezelor

- Scopul este de a compara populațiile sau de a evalua relațiile între variabile
- Ipotezele sunt testate folosind teste statistice.
- Ipoteza nulă și ipoteza alternativă – afirmații concurente
 - **Ipoteza nulă (H0):** afirmă că nu există nicio diferență, efect sau relație între variabilele testate
 - **Ipoteza alternativă (H1):** contestă ipoteza nulă și afirmă că există o diferență, efect sau relație variabilele testate

Valoarea p

- Indică cât este de probabil să se obțină rezultate cel puțin egale sau mai mari decât observația reală dacă ipoteza nulă este corectă (dacă ipoteza nulă este susținută sau respinsă)
- Cu cât este valoarea p mai mică, cu atât este mai puternică evidența că ipoteza nulă trebuie să fie respinsă
- O valoare mică a p-ului (≤ 0.05) înseamnă că ipoteza nulă este respinsă. Aceasta este o dovadă puternică că ipoteza nulă este invalidă.
- O valoare mare a p-ului (> 0.05) înseamnă că ipoteza alternativă este slabă, deci ipoteza nulă nu poate fi respinsă

Exemplu:

Valoarea p este 0.012, înseamnă că există p șansă de 1.2% ca rezultatele să fie aleatorii

Valoarea p este 0.9, înseamnă că există p șansă de 90% ca rezultatele să fie aleatorii

Teste statistice

1. Teste de comparație

Evaluează dacă există diferențe în mediile aritmetice sau medianele a două sau mai multe grupuri

Test	Comparație de	Numărul de eșantioane
T-test	Medie aritmetică	2 grupuri
ANOVA	Medie aritmetică	3 sau mai multe grupuri
Wilcoxon	Mediană	2 grupuri
Kruskal-Wallis H	Mediană	3 sau mai multe grupuri

Teste statistice

2. Teste de corelație

Determină în ce măsură două variabile sunt asociate

Test	Utilizat la variabile
Pearson's	Numerice
Spearman's	Ordinale, numerice
Chi square test	Nominale, ordinale

Teste statistice

3. Teste de regresie

Demonstrează dacă modificările în variabila independentă (predictor) determină modificări în variabila dependentă (rezultat)

Test	Utilizat la variabile:
Regresia lineară simplă	1 variabilă dependentă și 1 variabilă independentă
Regresia linieară multiplă	1 variabilă dependentă și mai multe variabile independente
Regresia binară	Variabila dependentă binară

Importare de date

Exemplu: Cum importăm în Excel chestionarul online (Exemplu Google Forms)

Explorarea practicilor economiei circulare în universități: o perspectivă a € ☆

Questions Responses **121** Settings

121 responses [View in Sheets](#)

Not accepting responses

Message for respondents
This form is no longer accepting responses

Summary Question Individual

- Get email notifications for new responses
- Select destination for responses
- Unlink form
- Download responses (.csv) **3**
- Print all responses
- Delete all responses

1 Accesăm Responses, din formularul nostru Google Forms

2 Click pe cele trei puncte, unde se deschide un meniu

3 Click Download Responses

Excel

Cum transformăm datele în valori numerice

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Automate' with the 'Tell me' search bar. The 'Home' ribbon is visible, showing options for 'Wrap Text', 'Merge & Centre', 'General', 'Conditional Formatting', 'Format as Table', 'Cell Styles', 'Insert', 'Delete', and 'Format'. The 'Find & Replace' group is also visible, with the 'Replace...' option highlighted by a red circle and the number 1. The main area displays a table with the following data:

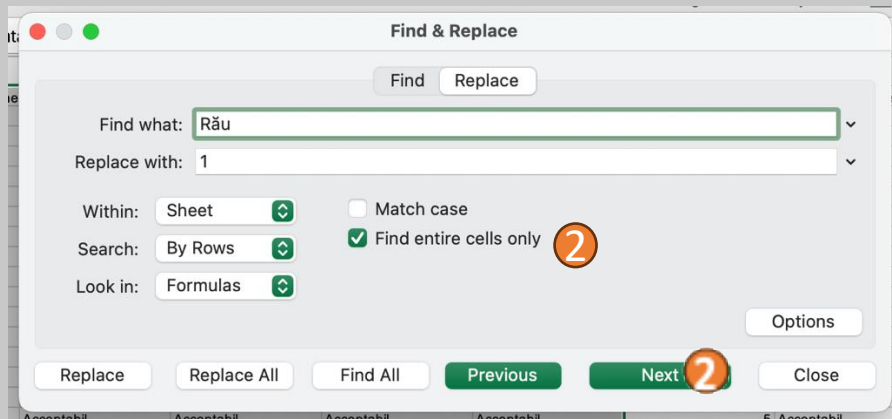
	J	K	L	M	N	O
ntă este	În opinia ta, cât de bine	În opinia ta, cât de bine	În opinia ta, cât de bine	În opinia ta, cât de bine	În opinia ta, cât de bine	În opinia ta, cât de bine
5	Acceptabil	Acceptabil	Acceptabil	Bine	Acceptabil	Destul e rău
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil
3	Bine	Bine	Foarte rău	Foarte rău	Destul e rău	Acceptabil
5	Acceptabil	Acceptabil	Destul e rău	Destul e rău	Destul e rău	Destul e rău
5	Bine	Bine	Bine	Acceptabil	Acceptabil	Acceptabil
5	Acceptabil	Acceptabil	Destul e rău	Acceptabil	Destul e rău	Acceptabil
5	Acceptabil	Destul e rău	Acceptabil	Nu știu	Destul e rău	Destul e rău
5	Acceptabil	Acceptabil	Acceptabil	Nu știu	Nu știu	Nu știu
5	Destul e rău	Destul e rău	Destul e rău	Destul e rău	Destul e rău	Destul e rău
5	Acceptabil	Acceptabil	Nu știu	Nu știu	Nu știu	Nu știu
5	Bine	Bine	Bine	Bine	Nu știu	Bine
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil
5	Acceptabil	Acceptabil	Acceptabil	Destul e rău	Foarte rău	Destul e rău
4	Foarte bine	Foarte bine	Foarte bine	Acceptabil	Foarte bine	Foarte bine
5	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil
5	Bine	Bine	Bine	Acceptabil	Acceptabil	Acceptabil
5	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu
5	Acceptabil	Acceptabil	Bine	Bine	Bine	Bine
5	Destul e rău	Destul e rău	Foarte rău	Acceptabil	Foarte rău	Foarte rău
5	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu	Nu știu
5	Destul e rău	Destul e rău	Destul e rău	Destul e rău	Destul e rău	Acceptabil
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Nu știu	Acceptabil
5	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil	Acceptabil
5	Acceptabil	Bine	Bine	Bine	Foarte bine	Acceptabil

1

Click pe lupă iar apoi Replace...

Excel

Cum transformăm datele în valori numerice



1

Înlocuim Rău cu 1
Destul de rău cu 2
Acceptabil cu 3
Bine cu 4
Foarte bine cu 5

2

Selectăm **Find entire cells only** – pentru a modifica doar celula cu exact acel text

3

Selectăm **Next** după fiecare modificare

Excel

T-test (test de comparație a mediei aritmetice)

Folosit pentru a evalua dacă există o diferență semnificativă între două grupuri

Exemplu:

Se evaluează dacă există o diferență semnificativă în performanța a două echipe de vânzări care lucrează cu diferite strategii de vânzare.

Ipoteza Nulă (H0): Nu există o diferență semnificativă în performanța de vânzare între echipa care utilizează strategia tradițională și echipa care utilizează strategia nouă.

Ipoteza Alternativă (H1): Există o diferență semnificativă în performanța de vânzare între echipa care utilizează strategia tradițională și echipa care utilizează strategia nouă

Excel

T-test (test de comparație a mediei aritmetice)

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Data'. The 'Data Analysis' button in the 'Data' ribbon is highlighted with a red circle '2'. The 'Data Analysis' dialog box is open, and 't-Test: Paired Two Sample for Means' is selected with a red circle '3'. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	Grup 1	Grup 2																				
2		10	8																			
3		8	5																			
4		3	3																			
5		4	6																			
6		5	7																			
7		8	5																			
8		6	3																			
9		7	2																			

1

Accesăm secțiunea Data

2

Selectăm Data Analysis

3

Selectăm t-Test: Paired Two Sample for Means

Excel

T-test (test de comparație a mediei aritmetice)

AutoSave Home Insert Draw Page Layout Formulas **Data** Review View Automate

Get Data (Power Query) Refresh All Queries & Connections Properties Workbook Links Stocks Currencies Geography Sort Filter Clear Reapply Advanced

	A	B	C
1	Grup 1	Grup 2	
2	10	8	
3	8	5	
4	3	3	
5	4	6	
6	5	7	
7	8	5	
8	6	3	
9	7	2	

t-Test: Paired Two Sample for Means

Input

Variable 1 Range: 1 OK Cancel

Variable 2 Range: 2

Hypothesized Mean Difference:

Labels

Alpha:

Output options

Output Range: 3

New Worksheet Ply:

New Workbook

1 Așezăm prima coloană în Variable 1 Range

2 Așezăm a doua coloană în Variable 2 Range

3 Selectăm locația pentru rezultat

Excel

T-test (test de comparație a mediei aritmetice)

A	B	C	D	E	F
Grup 1	Grup 2				
10	8		t-Test: Paired Two Sample for Means		
8	5				
3	3			Variable 1	Variable 2
4	6		Mean	6,375	4,875
5	7		Variance	5,410714286	4,410714286
8	5		Observations	8	8
6	3		Pearson Correlation	0,332637801	
7	2		Hypothesized Mean Difference	0	
			df	7	
			t Stat	1,655031853	
			P(T<=t) one-tail	0,070945801	
			t Critical one-tail	1,894578605	
			P(T<=t) two-tail	0,141891603	
			t Critical two-tail	2,364624252	

P(T<=t) one tail = 0.07
Mare mare ca 0.05 →
nu există suficiente
dovezi pentru a
respinge ipoteza nulă,
așadar aceasta este
susținută

Ipoteza Nulă (H0): Nu există o diferență semnificativă în performanța de vânzare între echipa care utilizează strategia tradițională și echipa care utilizează strategia nouă.

Ipoteza Alternativă (H1): Există o diferență semnificativă în performanța de vânzare între echipa care utilizează strategia tradițională și echipa care utilizează strategia nouă

Excel

ANOVA (test de comparație a mediei aritmetice)

Folosit pentru a evalua dacă există o diferență semnificativă între trei sau mai multe grupuri în ceea ce privește o variabilă de interes

Exemplu:

Se evaluează dacă există o diferență semnificativă în gradul de satisfacție ale angajaților din departamentul de vânzări, marketing și financiar.

Ipoteza Nulă (H0): Nu există diferențe semnificative în gradul de satisfacție ale angajaților între departamentele de vânzări, marketing și financiar.

Ipoteza Alternativă (H1): Există diferențe semnificative în gradul de satisfacție ale angajaților între departamentele de vânzări, marketing și financiar.

Excel

ANOVA (test de comparație a mediei aritmetice)

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Data'. The 'Data Analysis' button in the 'Data Tools' group is highlighted with a red circle '2'. A 'Data Analysis' dialog box is open, showing a list of analysis tools. 'Anova: Single Factor' is selected and highlighted with a red circle '3'. The dialog box also has 'OK' and 'Cancel' buttons.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Grup 1	Grup 2																			
2		10	8																		
3		8	5																		
4		3	3																		
5		4	6																		
6		5	7																		
7		8	5																		
8		6	3																		
9		7	2																		
10																					
11																					

1

Accesăm secțiunea Data

2

Selectăm Data Analysis

3

Selectăm Anova: Single Factor

Excel

ANOVA (test de comparație a mediei aritmetice)

A	B	C
Grupa 1	Grupa 2	Grupa 3
98	99	92
76	67	43
75	89	55
45	84	69
66	84	98
89	79	97
82	69	65
98	23	70
100	59	60
56	99	88

Anova: Single Factor

Input

Input Range: 1

Grouped By:

Columns
 Rows

Labels in first row

Alpha:

Output options

Output Range: 2
 New Worksheet Ply:
 New Workbook

OK
Cancel

1

Așezăm coloanele în secțiunea Input Range

2

Selectăm locația pentru afișarea rezultatului

Excel

ANOVA (test de comparație a mediei aritmetice)

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
98	9	687	76,33333333	340,75		
99	9	653	72,55555556	494,527778		
92	9	645	71,66666667	361,5		
ANOVA						
Source of Variati	SS	df	MS	F	P-value	F crit
Between Gro	110,518519	2	55,2592593	0,1385201	0,87133679	3,40282611
Within Groups	9574,22222	24	398,925926			
Total	9684,74074	26				

$P(T \leq t)$ one tail = 0.87
Mare mare ca 0.05 →
nu există suficiente
dovezi pentru a
respinge ipoteza nulă,
așadar aceasta este
susținută

Ipoteza Nulă (H0): Nu există diferențe semnificative în gradul de satisfacție ale angajaților între departamentele de vânzări, marketing și financiar.

Ipoteza Alternativă (H1): Există diferențe semnificative în gradul de satisfacție ale angajaților între departamentele de vânzări, marketing și financiar.

Teste de corelație

Folosită pentru a determina în ce măsură două variabile sunt asociate

Exemplu:

Se evaluează dacă există o relație între timpul petrecut pe rețelele de socializare și nivelul de stres al studenților

Ipoteza Nulă (H0): Nu există nicio corelație între orele petrecute pe rețelele de socializare și nivelul de stres al studenților

Ipoteza Alternativă (H1): Există o corelație între orele petrecute pe rețelele de socializare și nivelul de stres al studenților

Excel

Teste de corelație

The screenshot shows the Microsoft Excel interface. The ribbon is set to the 'Data' tab. A red circle with the number '1' highlights the 'Data' tab. In the 'Data Analysis' group on the right side of the ribbon, a red circle with the number '2' highlights the 'Data Analysis' button. A 'Data Analysis' dialog box is open, showing a list of analysis tools. A red circle with the number '3' highlights the 'Correlation' option in the list. The background spreadsheet shows data in columns A and B, with a green selection box around cells E3 and E4.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	Grup 1	Grup 2																				
2		10	8																			
3		8	5																			
4		3	3																			
5		4	6																			
6		5	7																			
7		8	5																			
8		6	3																			
9		7	2																			
10																						
11																						

1

Accesăm secțiunea Data

2

Selectăm Data Analysis

3

Selectăm Correlation

Excel

Teste de corelație

Orele petrec	Nivelul de stres
4	2
5	5
6	3
3	4
4	3
6	3
7	3
8	4
10	5

Correlation

Input

Input Range: 1

Grouped By: Columns Rows

Labels in first row

Output options

Output Range: 2

New Worksheet Ply:

New Workbook

OK Cancel

1

Așezăm coloanele
în secțiunea Input
Range

2

Selectăm locația
pentru afișarea
rezultatului

Excel

Teste de corelație

	<i>Column 1</i>	<i>Column 2</i>
<i>Column 1</i>	1	
<i>Column 2</i>	0,42253144	1

0.43 → corelație moderată și pozitivă

Ipoteza Nulă (H0): Nu există nicio corelație între orele petrecute pe rețelele de socializare și nivelul de stres al studenților

Ipoteza Alternativă (H1): Există o corelație între orele petrecute pe rețelele de socializare și nivelul de stres al studenților

Excel

Teste de corelație

Valoarea	
0.8 la 1	Corelație foarte puternică și pozitivă
0.6 la 0.8	Corelație puternică și pozitivă
0.4 la 0.6	Corelație moderată și pozitivă
0.2 la 0.4	Corelație slabă și pozitivă
0.0 la 0.2	Corelație foarte slabă și pozitivă
0.0	Nu există corelație
0.0 la -0.2	Corelație foarte slabă și negativă
- 0.2 la -0.4	Corelație slabă și negativă
-0.4 la -0.6	Corelație moderată și negativă
-0.6 la -0.8	Corelație puternică și negativă
-0.8 la -1	Corelație foarte puternică și negativă

Analiza de regresie

Folosit pentru a evalua dacă modificările în variabila independentă (predictor) determină modificări în variabila dependentă (rezultat)

Exemplu:

Se evaluează dacă există o relație negativă între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Ipoteza Nulă (H0): Nu există nicio corelație între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Ipoteza Alternativă (H1): Există o corelație între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Excel

Analiza de regresie

The screenshot shows the Microsoft Excel interface. The 'Data' tab is selected in the ribbon, indicated by a red circle '1'. The 'Data Analysis' button in the 'Analysis Tools' group is highlighted with a red circle '2'. The 'Data Analysis' dialog box is open, showing a list of analysis tools. 'Regression' is selected and highlighted with a red circle '3'. The background spreadsheet contains data for two groups.

	A	B	C	D	E	F	G	H
1	Grup 1	Grup 2						
2		10	8					
3		8	5					
4		3	3					
5		4	6					
6		5	7					
7		8	5					
8		6	3					
9		7	2					
10								
11								

1

Accesăm secțiunea Data

2

Selectăm Data Analysis

3

Selectăm Regression

Excel

Analiza de regresie

	A	B	C	D	E	F	G	H	I	J	K
1											
2	1	1									
3	3	3									
4	4	4									
5	5	5									
6	5	5									
7	3	3									
8	2	1									
9	1	1									
10	4	5									
11	5	4									
12	3	3									
13	2	3									
14	5	5									
15	4	5									
16	3	3									
17	4	4									
18	5	3									
19											
20											
21											
22											
23											
24											
25											
26											
27											

Regression

Input

Input Y Range: 1

Input X Range: 2

Labels Constant is Zero

Confidence Level: 95 %

Output options

Output Range: 3

New Worksheet Ply:

New Workbook

Residuals

Residuals Residual Plots

Standardized Residuals Line Fit Plots

Normal Probability

Normal Probability Plots

OK Cancel

- 1 Așezăm variabila dependentă în Input Y Range
- 2 Așezăm variabila independentă în Input X Range
- 3 Selectăm locația pentru afișarea rezultatului

Excel

Analiza de regresie

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0,856994376							
R Square	0,734439361							
Adjusted R Square	0,716735318							
Standard Error	0,731633061							
Observations	17							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	22,20599009	22,20599009	41,48427437	1,11191E-05			
Residual	15	8,029304029	0,535286935					
Total	16	30,23529412						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	0,633699634	0,474854894	1,334512167	0,201941282	-0,378429615	1,645828882	-0,378429615	1,645828882
X Variable 1	0,831501832	0,129098579	6,440828702	1,11191E-05	0,556334724	1,106668939	0,556334724	1,106668939

Excel

Analiza de regresie

SUMMARY OUTPUT	
<i>Regression Statistics</i>	
Multiple R	0,856994376
R Square	0,734439361
Adjusted R Square	0,716735318
Standard Error	0,731633061
Observations	17

Multiple R: coeficientul de corelație, indică cât de puternică este relația liniară (1 înseamnă o relație pozitivă perfectă, iar o valoare de zero înseamnă nicio relație)

R Square: Coeficientul de determinare, indică câte puncte cad pe linia de regresie (0.73 înseamnă că 73% din valorile se potrivesc cu modelul)

Standard Error: Distanța medie pe care valorile observate o au față de linia de regresie.

Excel

Analiza de regresie

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	22,20599009	22,20599009	41,48427437	1,11191E-05
Residual	15	8,029304029	0,535286935		
Total	16	30,23529412			

Analiza de varianță - furnizează informații despre nivelele de variabilitate în cadrul unui model de regresie și formează baza pentru teste de semnificație

Significance F: oferă o idee despre cât de statistic semnificative sunt rezultatele

Valoare mai mică de 0.05 (5%) = modelul în regulă,

Valoare mai mare de 0.05 = recomandarea alegerii unei alte variabile independente

Excel

Analiza de regresie

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	0,633699634	0,474854894	1,334512167	0,201941282	-0,378429615	1,645828882	-0,378429615	1,645828882
X Variable 1	0,831501832	0,129098579	6,440828702	1,11191E-05	0,556334724	1,106668939	0,556334724	1,106668939

Coefficients: $Y = m * X + b$

$$Y = 0.83 * X + 0.63$$

P-value: Indică dacă ipoteza nulă este susținută sau respinsă

p-values $\leq 0.05 \rightarrow$ Ipoteza nulă este respinsă

Există o corelație între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Ipoteza Nulă (H0): Nu există nicio corelație între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Ipoteza Alternativă (H1): Există o corelație între preocuparea față de sustenabilitate și cumpărarea de produse verzi

Vă mulțumesc!

Drd. Timea Cisma

Timea.cisma@student.upt.ro