

Biostatistică – Medicină Generală

Lucrarea de laborator Nr.4

Scop: la sfârșitul laboratorului veți ști:

- Să folosiți foaia de calcul Excel pentru a executa tabele de incidență și pentru a calcula indicatorii esențiali asociați acestora
- Să efectuați și să interpretați graficul Scatter cu ajutorul pachetului EXCEL
- Să calculați coeficienți de corelație Pearson

1. Tabele de incidență executate manual folosind programul Microsoft Excel

Avantajul programului EXCEL în prelucrarea datelor este acela că multe din calculele inerente care intervin se pot face automat. Dacă, de exemplu, conținutul a două celule trebuie adunat într-o a treia celulă, se specifică programului doar că trebuie adunate numerele din cele două celule și că trebuie depus rezultatul în a treia. Programul scrie automat suma. Dacă se modifică unul din numerele adunate sau amândouă, suma se actualizează automat și acest lucru este valabil pentru toate celulele al căror conținut se calculează din altele.

A. Tabele de dependență a factorilor. Pentru exemplificare, să construim un tabel de incidență pentru care nu se introduc decât numele celor doi factori, categoriile Pozitiv și Negativ, și cele patru celule de bază ale tabelului. Reamintim că într-un tabel de incidență, sau cu dublă intrare, avem doi factori de clasificare a pacienților, un factor pe orizontală și unul pe verticală.

Anume, 120 de pacienți au fost testați în ce privește prezența autoanticorpilor antinucleari (AAN) și prezența celulelor lupice (LE, prezente la foarte mulți pacienți cu Lupus Eritematos Sistemis). Introduceți datele de mai jos astfel ca "AAN" să apară în celula C1, iar "LE" în celula A3, numărul 63 în celula C3, 19 în celula D3, 8 în celula C4 și 30 în celula D4.

Apoi scrieți formulele din celelalte celule, în ordinea următoare:

- În celula C5, formula =c3+c4
- În celula D5, formula =d3+d4
- În celula E3, formula =c3+d3
- În celula E4, formula =c4+d4
- În celula E5, formula =e3+e4
- În celula H2, formula =(c3+d4)/(c4+d3)
- În celula H3, formula =(c3+d4)/e5
- În celula H4, formula =(c3*d4)/(c4*d3)
- În celula H5, formula =(c3*e4)/(c4*e3)

	A	B	C	D	E	F	G	H
1			AAN					
2			Pozitiv	Negativ	Total		CD=	=(c3+d4)/(c4+d3)
3	LE	Pozitiv	63	19	=c3+d3		CP=	=(c3+d4)/e5
4		Negativ	8	30	=c4+d4		OR=	=(c3*d4)/(c4*d3)
5	Total		=c3+c4	=d3+d4	=e3+e4		RR=	=(c3*e4)/(c4*e3)

Apoi în celulele G2, G3, G4 și G5, scrieți așa cum vedeți în tabelul de mai sus, pentru a fi mai clar ce reprezintă fiecare din rezultatele pe care programul le-a afișat pe coloana H în celulele corepunzătoare, adică în H2, H3, H4 și H5.

După introducerea formulelor (adică a acelor texte care încep cu semnul =), se observă, așa cum se știe, că programul, în loc să pună text în celule, pune rezultate. Acum, se pot calcula criteriile de dependență CD, CP, OR și RR și la alte tabele, cum sunt cele din exercițiile de la sfârșitul protocolului, doar prin introducerea celor patru date esențiale pe care le-ați introdus și în tabelul de mai sus, restul celulelor fiind automat calculate de program.

În cazul tabelului de mai sus, există o puternică tendință de dependență între prezența AAN și LE la cei 120 de pacienți. Acum, ori de câte ori vom avea de efectuat un astfel de tabel, vom schimba doar celulele necalculate, adică numele factorilor de clasificare și cele patru valori pe baza cărora s-au făcut calculele.

Exercițiu. Schimbați numele factorilor și datele cu cele din tabelul de mai jos, astfel:

- În celula C1, în locul textului AAN, scrieți Vanilla (dacă pacientul a consumat sau nu înghețată de vanilie)
- În celula A3, în locul textului LE, scrieți Bolnav (dacă pacientul s-a îmbolnăvit sau nu).
- În celulele C3, C4 și D3, D4, scrieți în ordine numerele 43, 11, 3 și 18.

		Vanilla	
		Pozitiv	Negativ
Bolnavi	Pozitiv	43	3
	Negativ	11	18

Veți constata că la fiecare introducere a unui nou număr, valorile tuturor celulelor calculate se recalculază, inclusiv cele ale indicatorilor CD, CP, OR și RR. Din nou veți constata că există o legătură puternică între cei doi factori de clasificare, deoarece CD, OR și RR au valori mult mai mari ca 1, iar CP are valori mult peste 50%.

B. Tabele pentru evaluarea testelor clinice. Alte cazuri în care este util să folosim tabele ca cele de mai sus sunt cele în care dorim să evaluăm calitatea unui test la care supunem pacienții și care trebuie să fie capabil să discearnă între pacienții care prezintă sau nu prezintă o afecțiune, un simptom, un semn sau altă caracteristică necesară în procesul diagnosticării.

Ca exemplu, cu secvența **Insert** ⇒ **Worksheet**, inserați o nouă foaie de calcul și realizați tabelul de incidență următor pentru verificarea performanțelor unui test clinic, comparativ cu un test sigur: 92 de bolnavi și 45 de indivizi sănătoși, au fost testați cu un test nou, propus pentru folosirea în clinică, rezultatele obținute fiind cele de mai jos. Atenție: puteți scrie o formulă (text care începe cu semnul =), numai după ce celulele care apar în formulă au fost completate. Deci, veți introduce întâi valorile de la C3, C4, D3 și D4, apoi formulele de la C5 și D5, apoi formulele de la E3, E4 și E5, și abia apoi, restul formulelor.

Concluzia generală referitoare la testul de mai sus este că testul propus este un test valoros care poate fi folosit în clinică cu succes ca un test ajutător util în diagnosticarea bolii de care suferă cei 92 de pacienți. Aceasta, deoarece sensibilitatea și specificitatea testului sunt ambele foarte bune, iar rata fals pozitivă și rata fals negativă, ambele sunt mici. Astfel de concluzii va trebui să știți să trageți dacă veți primi un alt tabel asemănător.

	A	B	C	D	E	F	G	H
1			Test Sigur					
2			Bolnavi (B)	Sănătoși (S)	Total			
3	Test Propus	Pozitivi(P)	84	7	=c3+d3			
4		Negativi(N)	8	38	=c4+d4			
5	Total		=c3+c4	=d3+d4	=e3+e4			

6							
7	Sn	=c3/c5		FPR	=d3/d5	VPP	=c3/e3
8	Sp	=d4/d5		FNR	=c4/c5	VPN	=d4/e4

Dacă va trebui să introduceți date referitoare la o altă situație, tot ceea ce va trebui să faceți este să înlocuiți doar cele patru numere de la C3, C4, D3 și D4, restul, vor fi calculate automat.

2. Graficul punctual (Scatter)

Acest grafic se realizează numai în cazul în care avem două coloane de valori, care sunt **vecine** în tabel și sunt valori **numerice**. Este unul din cele mai informative grafice (adică conține o cantitate de informație importantă), dar și dificultatea interpretării lui este mai mare.

Pentru realizarea unui grafic punctual SCATTER, coloanele cu valori numerice, trebuie selectate. În tabelul CIROZA, pe coloanele VSH1H și VSH2H, sunt înregistrate valorile pentru viteza de sedimentare a hematiilor la o oră și respectiv, două ore.

The image shows a screenshot of an Excel spreadsheet and the 'Chart Wizard - Step 1 of 4 - Chart Type' dialog box. The spreadsheet has columns AB, AC, AD, and AE. The Chart Wizard dialog box is open, showing 'XY (Scatter)' as the selected chart type. The dialog box also shows 'Standard Types' and 'Custom Types' tabs, and a 'Chart sub-type' section with four options: 'Scatter with markers only', 'Scatter with smooth lines and markers', 'Scatter with smooth lines and no markers', and 'Scatter with straight lines and markers'. The 'Scatter with markers only' option is selected. The dialog box also has a 'Press and Hold to View Sample' button and 'Cancel', '< Back', 'Next >', and 'Finish' buttons.

Pentru realizarea graficului, deschideți tabelul CIROZA, dacă nu este deschis. Apoi căutați coloanele pe care sunt înregistrate valorile VSH, coloane care sunt denumite VSH1H și VSH2H. Apoi selectați valorile de pe ambele coloane. Începeți selecția din celula AC1 și trageți și peste coloana AD, în jos, până ajungeți cu selecția la linia 235, adică ultima cu date înregistrate, așa cum se vede în figurile din stânga, unde sunt reprezentate începutul (stânga) și sfârșitul zonei selectate(dreapta).

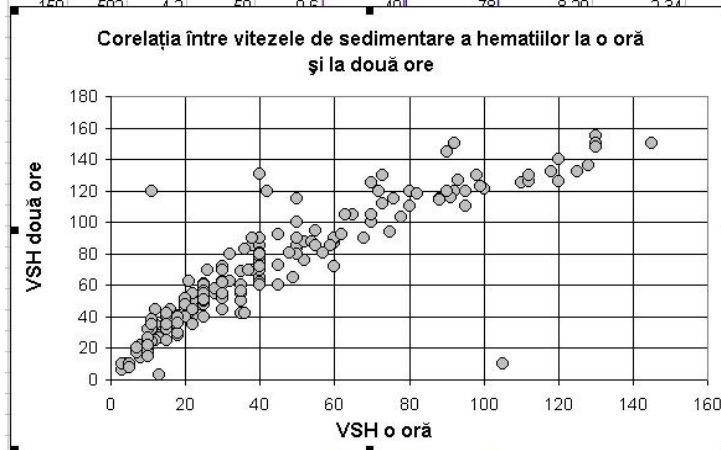
Apoi, executați clic pe iconița asistentului de grafice și în fereastra pentru pasul 1 alegeți ca în figura de mai sus executând câte un clic pe XY (Scatter) și pe modelul cu puncte dar fără linii, care vedeți că este selectat în figură. Apoi apăsați butonul **Next**. La pasul doi nu faceți nici o opțiune, doar apăsați butonul **Next**. La pasul 3, scrieți în cele trei casete de dialog, titlul graficului și numele celor două coloane. Titlul este **“Corelația între vitezele de sedimentare la o oră și la două ore.”**. În celelalte două casete scrieți **VSH1H** și **VSH2H**.

Apoi apăsați butonul **Next**. La pasul 4 apăsați doar butonul **Finish**.

În graficul care apare, ajustați dimensiunile, micșorați textele și colorați după dorință, așa cum ați procedat și la graficul histogramă. Graficul este gata.

Se observă că valorile din cele două coloane sunt corelate, deoarece punctele graficului sunt așezate într-un nor alungit și subțire (cu puține excepții). Dacă punctele se așază absolut aleator, valorile din cele două coloane nu se corelează.

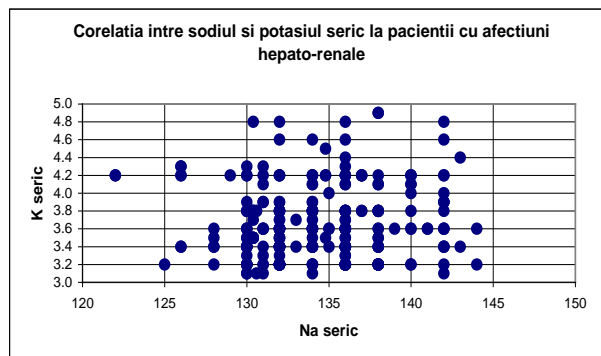
X	Y	Z	AA	AB	AC	AD	AE	AF	AG
TGO	LDH	ALBUMIN	PRTR	GLICEMIE	VSH1H	VSH2H	PROTEINE	GAMAGLOE	IGA
59	339	3.5	80	0.8	18	28	7.80	2.26	42
237	738	4.0	70	1.2	120	140	8.30	2.38	20
109	522	3.6	100		38	85	6.30	1.45	94
150	600	4.0	50	0.6	40	70	8.00	2.24	6



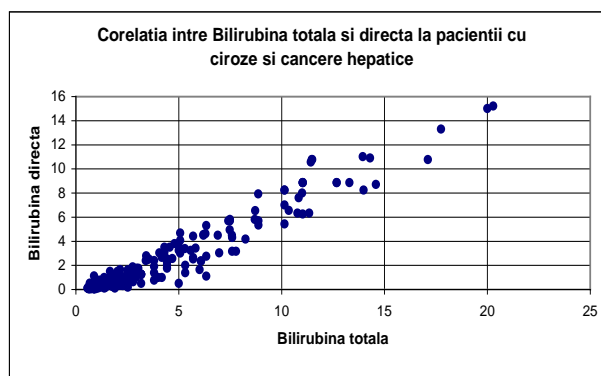
În ce privește graficul punctual (SCATTER), veți urmări totdeauna trei tipuri diferite de informație pe care acesta ni le poate oferi:

- dacă cele două serii de valori sunt corelate (**Corelația**).
- dacă eșantionul pe care s-au făcut măsurătorile este omogen sau este eterogen (**Omogenitatea**).
- anumite informații despre simetria distribuției celor două serii de valori (**Simetria**).

Exemple:

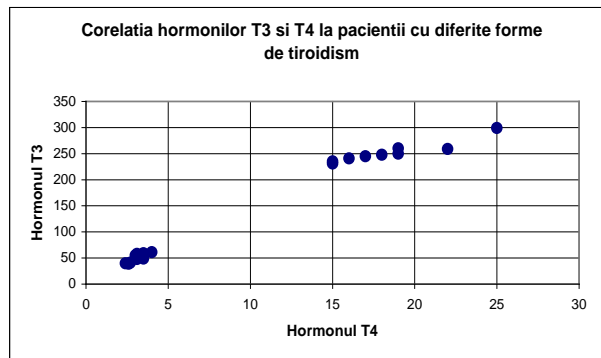


1. În graficul de mai sus se observă că între valorile sodiului seric și potasiului seric ale pacienților cu afecțiuni hepato-renale **nu există nici o corelație**, deoarece punctele sunt distribuite aproximativ întâmplător. Nu se poate observa o tendință clară ca punctele graficului să fie așezate într-un fel anume, în afară de tendința naturală ca densitatea să fie mai mare în centru, tendință care se poate observa în general, la majoritatea datelor în medicină și biologie și nu numai. Pe de altă parte, tendința punctelor de a se așeza într-un singur nor, arată omogenitatea seriei de valori.



2. În graficul de mai sus, se observă că între valorile bilirubinei totale și bilirubinei directe ale pacienților cu ciroze și cancer hepatic există o **corelație puternică**, deoarece datele sunt distribuite **aproximativ de-a lungul unei linii drepte**. De asemenea, eșantionul este **omogen**, ca și în cazul graficului 1, lucru care se observă din faptul că norul de puncte nu are tendința de a se divide în doi nori diferiți. Pe de altă parte se vede că norul este mult mai dens în stânga jos. Deci, valorile din seria bilirubinei totale (orizontală), sunt distribuite asimetric (valori mai multe în stânga =

asimetrie dreapta). La fel, valorile din seria bilirubinei directe (verticală), sunt mai dense jos, la valori mici, deci sunt distribuite tot asimetric (**asimetrie dreapta**).



3. În acest caz, se observă că între valorile hormonilor tiroidieni T4 și T3 ale pacienților cu diferite forme de tiroidism există o corelație, datele fiind distribuite tot de-a lungul unei linii, chiar dacă este discontinuă. De asemeni, se observă că distribuția lor este grupată: o grupă cu valori mici pentru cei doi hormoni indicând hipotiroidienii și o grupă cu valori mari pentru cei doi hormoni indicând hipertiroidienii, adică o tendință clară de separare în doi nori diferiți. Spunem în acest caz că eșantionul este **eterogen** sau **neomogen**.

Observație. La fiecare grafic de acest tip este bine să fie urmărite cele trei tipuri de informație pe care poate să ni le ofere (**Corelație, Omogenitate și Simetrie**).

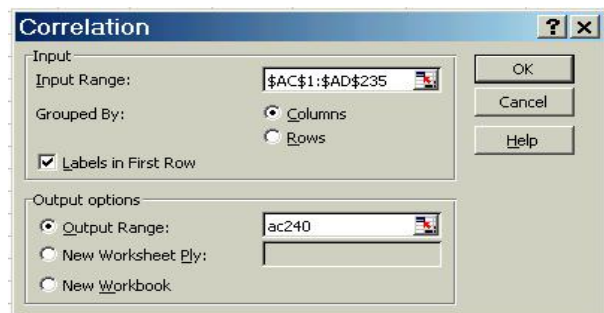
Unele ne dau o informație mai utilă privind corelația, altele ne informează mai bine în ce privește omogenitatea sau simetria, deci nu totdeauna se pot obține dintr-un singur grafic toate cele trei tipuri de informație la fel de ușor. Aprecierile sunt totdeauna subiective și depind de deprinderile pe care utilizatorul și le-a format lucrând cu cât mai multe și mai variate exemple.

3. Corelații

Pentru a verifica existența unei corelații între datele din două coloane numerice, se folosește funcția **Correlation**. Pentru aceasta, este necesar ca cele două coloane să se afle în tabelul de date una lângă alta. Dacă nu este așa, ele trebuie mutate pe o altă foaie de lucru, astfel încât să fie vecine.

Deocamdată, să ne ocupăm de cazul simplu când avem coloanele vecine. De exemplu, în tabelul CIROZA, coloanele VSH1 și VSH2, adică viteza de sedimentare a hematiilor la o oră și, respectiv, la două ore, sunt vecine, anume pe coloanele AC și AD. Valorile sunt pe prima coloană de la AC2 la AC235, iar pe coloana AD, de la AD2 la AD235.

	V	W	X	Y	Z	AA	AB	AC	AD
1	COLEST	TGP	TGO	LDH	ALBUM	IPRTR	GLICEM	VSH1H	VSH2H
2			37	59	339	3.5	80	0.8	18
3	150	129	237	736	4.0	70	1.2	120	140
4	340	56	109	522	3.6	100		38	85
5	187	62	150	502	4.2	50	0.6	40	78
6	230	31	55		3.5	70	0.6	21	63
7	174	44	91	645	3.7	70	1.4	16	45
8		65	187	468	3.5	55	1.1	40	71
9	186	47	105	588	3.3	70	0.7	15	40
10	186	44	84	547	4.0	60	0.8	30	61
11	240	51	98	394	3.4	100	1.2	30	60
12	227	59	150	567	3.2	70	1.1	40	70
13	140	56	100	418	2.6	90	1.0		
14			153	286	458	2.1	60	0.8	8
15		45	80	722	3.2	50	1.1	5	10



Pentru a calcula coeficientul de corelație al lui Pearson între două serii de date numerice se procedează astfel:

Deschideți tabelul CIROZA cu secvența **File->Open**. Alegeți opțiunea **Tools** din meniul principal al programului EXCEL. Din lista care se deschide, alegeți **Data Analysis**. Din fereastra care se deschide, alegeți **Correlation**

În caseta de dialog **Input Range**, scrieți **AC1:AD235**. Bifați caseta de validare **Labels in First Row**. Dacă nu este bifată, bifați și butonul de opțiune numit **Output Range**, și scrieți în caseta de dialog corespunzătoare **AC240**. În final, apăsați **OK**.

Rezultatul, este apariția unui tabel pe trei coloane și trei linii, coeficientul de corelație căutat fiind în celula AD242. Se observă că valorile din cele două coloane sunt strâns corelate, deoarece coeficientul de corelație este 0,89

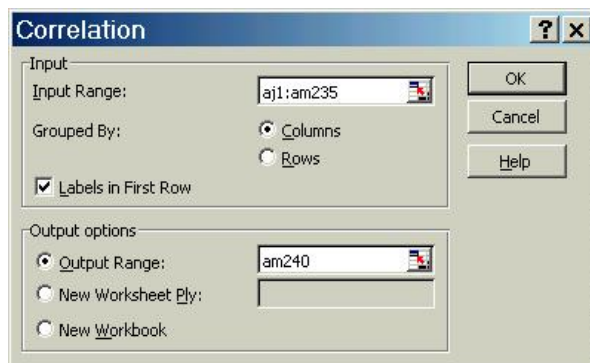
În celula AD241, precum și în celula AE242, programul a înscris numărul 1,00. La AD241 este coeficientul de corelație al coloanei VSH1H cu ea însăși. Orice serie de valori are coeficientul de corelație cu ea însăși egal cu 1,00. La fel la AE242, coeficientul de corelație al VSH2H cu ea însăși. **Această afișare este superfluă, necesară și nu neinteresează.**

	AF	AG	AH	AI	AJ	AK	AL	AM	AN
1	GAMAGLOB	IGA	IGM	IGG	C3	HB	L	TR	UREE
2	2.26	420	300	352	40	12.00	7800	160000	32
3	2.38	200	320	410	52	11.00	7600		
4	1.45	940	200	230		12.69	7600	170000	28
5	2.34	60	200	384		10.52	7000	180000	20
6	1.53	380	230	260	36	12.09	9800	50000	
7	1.94	400	480	510	38	11.10	7800	70000	61
8	1.83	800	230	260		10.71	9500	100000	65
9	2.70	900	220	280	72	13.09	8500	150000	
10	1.90	640	650	384		12.22	7800		20
11	1.70	440	290	350		13.80	6500		
12	2.80	240	250	400		14.28	6200	100000	37
13	2.00	380	160	384					24
14		210	320	280		13.08	8500	180000	17
15	1.53	200	280	320		11.24	7000	85000	19
16	1.75	160	570	176					27
17	2.65	328	900	560					19
18	2.30	778	650	128					20
19	1.36	200	340	380	64	11.50	7600	150000	
20	2.08	672	190	480					28
21	0.91	180	190	160		10.71	7800	150000	
22	2.40	400	700	180		10.71	5800	150000	
23	2.24	312	248	254	56	12.29	8000	180000	
24	2.24	400	380	436					24

Se pot calcula coeficienții de corelație între mai mult de două coloane, cu condiția ca acestea să fie vecine.

De exemplu, coloanele AJ, AK, AL, AM, conțin date despre complementul C3, hemoglobină, leucocite și trombocite. Pentru a calcula dintr-o dată toți coeficienții de corelație între trei sau mai multe coloane care sunt vecine, procedați astfel:

Verificați care sunt numele coloanelor și care sunt liniile unde încep valorile și unde se termină. În cazul nostru, valorile încep la AJ2 și se termină la AM25. Neglijați faptul că unele valori lipsesc programul va lua în considerare numai pacienții la care avem valori.



Alegeți opțiunea **Tools** din meniul principal al programului EXCEL. Din lista care se deschide, alegeți **Data Analysis**. Din fereastra care se deschide, alegeți **Correlation**, apoi apăsați **OK**.

Completați datele ca în figura de mai sus. După ce ați completat identic fereastra **Correlation**, apăsați butonul **OK**.

Veți constata că la celula AM240, spre dreapta și în jos, a fost listat un tabel cu coeficienți de corelație care se citește simplu, urmărind intersecția unei linii cu o coloană a tabelului.

De exemplu, între trombocite și hemoglobină, găsim pe ultima linie, pe coloana a doua numerică, coeficientul de corelație Pearson este $r=0,06$, adică foarte slab. Nu există corelație între valorile din cele două coloane.

	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR
226			14.28	7600	180000					
227	200		11.63	8000	150000					
228	420		9.66	10600	185000					
229						28		147.80	2.50	
230	368		12.50	4700	200000	24		130.00	4.10	143.00
231	308		13.63	7000						
232	220	72	10.84	9300		60				
233	270		11.64	9200	230000					
234	236	38	12.81	6000	165000	33	0.88	143.40	5.10	60.00
235	900	88	10.52	6600	210000	27		139.00	3.80	134.00
236										
237										
238										
239										
240										
241						C3	HB	L	TR	
242						C3	1.00			
243						HB	0.02	1.00		
244						L	0.05	-0.05	1.00	
245						TR	0.61	0.06	0.07	1.00
246										

Pe diagonală sunt valori egale cu 1,00, deoarece reprezintă corelațiile fiecărei coloane cu ea însăși.

Acest tabel listat începând de la celula AM240, se numește **Matricea coeficienților de corelație** pentru cele 4 coloane. Dintr-o astfel de matrice, ne interesează acei coeficienți care au valori mai mari, și care indică existența unei corelații, precum și cei care au valori apropiate de 0, ceea ce indică absența unei corelații. De exemplu, în matricea de mai sus, la AN244, coeficientul de 0,66, sau 66%, indică o corelație între C3 și TR. Coeficientul de 0,02, sau 2% de la AN242, indică lipsa unei corelații între C3 și Hb.

4. Exerciții și chestiuni de examen

1. Studiați dependența dintre prezența autoanticorpilor antinucleari (AAN) și a celulelor lupice (LE) la bolnavii de lupus (tabelul LUPUS).

Se vor urmări indicatorii CD, CP, OR și RR pentru aprecierea dependenței între cei doi factori (prezența celulelor lupice și prezența anticorpilor antinucleari). Pe foaia de examen, vor fi trecute atât tabelul rezultat, cât și valorile indicatorilor CD, CP, OR, RR. Atenție, programul Epi calculează și interpretează numai OR și RR. Indicatorii CD și CP, trebuie calculați.

2. Verificați dependența între îmbolnăviri (ILL) și consumul de cafea (Coffee) și de înghețată de vanilie (Vanilla), la indivizii din tabelul OSWEGO. Folosiți indicatorii CD, CP, OR.

Se vor urmări indicatorii CD, CP, OR și RR pentru aprecierea dependenței între cei doi factori (consumul cafelei și îmbolnăviri). Apoi se repetă analog punând în caseta de dialog Exposure Variable numele coloanei înghețata de vanilie (VANILLA), iar în caseta de dialog Outcome Variable se pune numele coloanei ILL. Pe foaia de examen, vor fi trecute atât tabelul rezultat la fiecare comandă, cât și valorile indicatorilor CD, CP, OR, RR. Atenție, programul Epi calculează și interpretează numai OR și RR. Indicatorii CD și CP, trebuie calculați și interpretați fără ajutorul programului EPI.

3. Un test de detectare a prezenței cancerului pulmonar este aplicat la 100 de pacienți care au cancer pulmonar, precum și la 100 de pacienți care au boli pulmonare dar nu au cancer. Dintre pacienții cu cancer, testul nou a diagnosticat corect 87, restul de treisprezece au ieșit negativi (adică nu ar avea cancer). Dintre pacienții cu alte afecțiuni decât cancerul, 85 au ieșit negativi, dar 15 au ieșit pozitivi (adică după testul în discuție, ar avea cancer). Calculați sensibilitatea și specificitatea acestui test.

Veți efectua un tabel de incidență ca la punctul 3 al acestui protocol de laborator. Sensibilitatea unui test este raportul dintre numărul celor diagnosticați corect ca pozitivi (real pozitivi) și numărul total al celor pozitivi (bolnavi). În cazul nostru, diagnosticat corect ca pozitiv înseamnă a avea cancer după testul propus (87 pacienți) iar bolnavi sunt cei care au cancer (100). Deci sensibilitatea este

$$S_n = \frac{B^+}{B} = \frac{87}{100} = 0,87$$

Specificitatea este raportul dintre numărul celor diagnosticați de test în mod corect ca sănătoși (sau fără cancer) și numărul total al celor sănătoși (sau fără cancer). Deci:

$$Sp = \frac{S^-}{S} = \frac{85}{100} = 0,85$$

Testul care are indicatorii senzitivitate și specificitate 0,87 și respectiv 0,85 este un test foarte util și deci valoros din punct de vedere clinic.

4. Din 200 de pacienți care au fost expuși la o intoxicație cu factorul activ al unui medicament, 118 au acuzat o stare de disconfort, amețeli și vertij. Alți 150 de pacienți având aceeași afecțiune ca și primii, nu au fost tratați cu medicamentul în cauză, totuși, 58 au manifestat aproximativ aceleași simptome. Deoarece simptomele manifestate de pacienți se întâlnesc destul de des în afecțiunea de care suferă ei, nu se știe dacă medicamentul a cauzat o amplificare a simptomelor. Bazat pe datele de mai sus să se aprecieze cu ajutorul indicatorului OR, dacă medicamentul a cauzat amplificarea simptomelor.

5. Un test de detectare a prezenței cancerului de sân este aplicat la 200 de paciente care au cancer de sân, precum și la 150 de paciente care au alte tumori dar nu au cancer. Dintre pacientele cu cancer, testul nou a diagnosticat greșit 22. Dintre pacientele cu alte tumori decât cancerul, 135 au ieșit negative. Calculați senzitivitatea și specificitatea acestui test.

6. Deschideți tabelul **CARDIO** și calculați matricea coeficienților de corelație între coloanele **TAMAXINT**, **TAMININT**, **TAMAX2S**, **TAMIN2S**, **TAMAX3L**, **TAMIN3L**, care reprezintă tensiunile sistolică și diastolică ale pacienților la internare, la două săptămâni de tratament și la 3 luni la controlul postspitalizare. Scrieți pe o foaie de hârtie, acei coeficienți care arată o corelație mai puternică, precum și pe aceia care arată că nu avem corelație sau avem o corelație foarte slabă.

ATENȚIE! În acest caz, unii coeficienți, deși sunt calculați de program, nu sunt utili din punct de vedere practic. De exemplu, nu interesează coeficientul de corelație între **TAMAXINT** și **TAMIN2S**. Într-adevăr, tensiunea sistolică la internare are sens să o corelăm cu tensiunea diastolică la internare sau cu tensiunea sistolică la 2 săptămâni sau la 3 luni, în rest nu are sens să o corelăm cu alte coloane.

7. Verificați cu ajutorul unui grafic **Scatter**, dacă la pacienții din tabelul **OBEZ**, valorile greutateii și înălțimii sunt corelate. Este omogen lotul de pacienți din punctul de vedere al înălțimii și al greutateii?

8. Verificați cu ajutorul unor coeficienți de corelație dacă la pacienții din tabelul **OBEZ**, se corelează vârsta, greutatea și înălțimea.