

Lr 10 Rezumat

Partea a IV-a. STATISTICĂ BIVARIATĂ

Studiază simultan două variabile de același tip pentru: (1) demonstrarea independenței acestora ori (2a) determinarea formei, respectiv a modului de dependență argumentată biologic, precum și măsurarea intensității acesteia, ceea ce permite (2b) *predicția* comportării unei variabile în raport cu cealaltă variabilă.

Dependența între *variabile cantitative* se numește **corelație**, (în cazul variabilelor tip măsurătoare tratate parametric, corelația având o anumită formă), iar dependența între *variabile calitative* se numește **asociere** (într-un anumit mod).

Sinteza grafică se realizează în: (a) *tabele statistice cu dublă intrare*, numite **tabele de corelație** pentru *variabile cantitative*, respectiv **tabele de contingență** pentru *variabile calitative*, apoi în (b) *diagrame de împrăștiere*, *stereograme*, respectiv *areale în dreptunghiuri*.

Sinteza numerică se face în parametri de *corelație*, respectiv *asociere*, ceea ce va permite predicția comportării unei variabile în funcție de valorile, respectiv variantele celeilalte variabile.

Capitolul 7. TRATAREA SIMULTANĂ A DOUĂ DIMENSIUNI

7.1. Sinteza grafică bidimensională

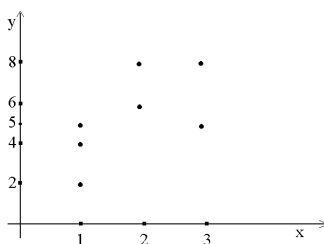
Serie bivariată:

x: 1 1 1 2 2 2 2 3 3 3

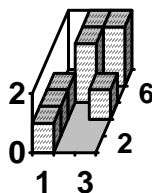
y: 2 4 5 6 6 8 8 5 8 8

Tabel cu dublă intrare:

x = y:	1	2	3
8	0	2	2
6	0	2	0
5	1	0	1
4	1	0	0
2	1	0	0



Diagramă de împrăștiere



Stereohistogramă

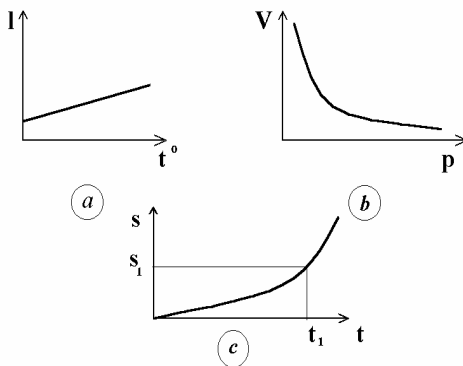
7.2. Dependenta funcțională și dependenta statistică

O **dependență funcțională** este o relație **univocă** de la x către y (unei valori fixate a lui x îi corespunde o singură valoare a lui y). Se notează $y = f(x)$. Apare în fizică, astronomie, chimie. O **dependență statistică** este o relație neunivocă "în spatele" căreia apare ca tendință o relație funcțională. Apare în biologie și în multe alte domenii experimentale.

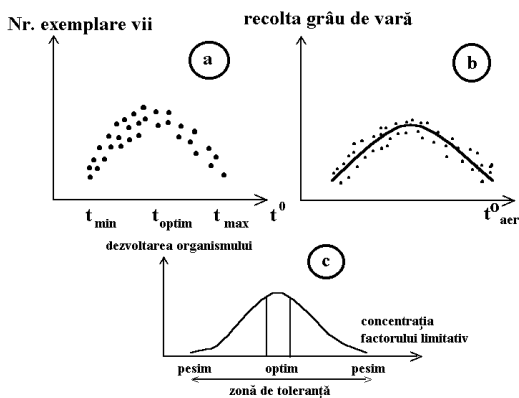
7.2.1. Exemple de dependente funcționale

Legile:

- dilatării $l = k \cdot t + k_0$;
- Boyle-Mariotte $p \cdot V = ct$;
- spațiului în căderea liberă $s = -g \cdot t^2$.



7.2.2. Exemple de dependente statistice



Legea toleranței:

În general, **linie de dependență** poate fi orice linie diferită de drepte paralele la axe.

7.2.3. Independență funcțională și independență statistică totală, respectiv reală

Independență funcțională (totală) \Leftrightarrow perechile (x, y) sunt plasate pe o dreaptă paralelă la una din axele de coordonate.

Independență statistică totală \Leftrightarrow perechile (x, y) sunt grupate în jurul unei drepte paralele la una din axele de coordonate.

Independență statistică reală \Leftrightarrow perechile (x, y) formează un nor eliptic (în particular circular) cu axele paralele la axele de coordonate.

7.2.4. Postulat asupra legăturii între planul fenomenologic și cel al datelor observate

Prin demers statistic complet putem să dovedim cu certitudine doar lipsa unei dependențe fenomenologice (manifestată prin datele respective). "O corelație (respectiv o asociere) între variabile implică existența unei dependențe în plan fenomenologic" funcționează doar ca o posibilitate, nu în mod necesar. De aceea, în cazul unei dependențe în plan fenomenologic, biologul va trebui să vină cu argumente de specialitate pentru susținerea acesteia, iar *statistica va ajuta doar la exprimarea formei corelației** și la măsurarea *intensității acesteia*. Fără argumente biologice o corelație descoperită în date poate fi o **corelație aparentă (falsă)**.

7. 2. 5. Aplicarea postulatului

1° Dovedirea independenței

se face prin diagrame de împrăștiere în care norul de puncte este plasat ca în 7. 2. 3.

2° Susținerea formei unei dependențe (propusă de biolog) și măsurarea intensității acesteia

(1) Susținerea unei anumite forme - prin diagrama de împrăștiere. (2) Determinarea ecuației formei respective (denumită linie de regresie) - printr-o metodă de regresie. (3) Măsurarea intensității gradului de corelație după forma respectivă - printr-un indice de corelație specific formei. (1) sn **identificare** sau **modelare**, (2) sn **ajustare** și (3) sn **validare**.

* (respectiv a modului de asociere în cazul variabilelor calitative)

7.3. Sinteza numerică bidimensională

7.3.1. Forme de corelație

1) $y = a + b \cdot x$	$b \neq 0$	<i>Ecuatia graficului unei: linii drepte neparalele cu axa Ox sau Oy</i>
2) $y = a + b/x$	$a > 0, b \neq 0$	<i>hiperbole</i>
3) $y = a \cdot b^x$	$a, b > 0$	<i>funcții exponențiale</i>
4) $y = a \cdot \log_b x$	$a, b > 0$	<i>funcții logaritmice</i>
5) $y = a \cdot x^b$	$a, b > 0$	<i>funcții putere</i>
6) $y = a + b \cdot x + c \cdot x^2$	$c \neq 0$	<i>parabole de gradul II</i>
7) $y = a \cdot (1 - e^{-b \cdot x})$	$a, b > 0$	<i>curbe de saturație</i>
8) $y = 1 / (a + b \cdot e^{-x})$	$a, b > 0$	<i>curbe logistice*</i>

1) *Linia dreaptă oblică care trece prin origine* ($y = b \cdot x, b \neq 0$) exprimă ideea de **dependență direct proporțională**: « la o creștere a variabilei x , de un anumit număr de ori corespunde o creștere a variabilei y , (**direct**) **proporțională** (de același număr de ori) ».

2) *Hiperbola* exprimă ideea de **dependență invers proporțională**: « la o creștere a variabilei x , de un anumit număr de ori corespunde o descreștere a variabilei y , de același număr de ori ».

7.3.2. Ajustarea unei anumite forme la datele experimentale

Sn linie (dreaptă sau curbă) de regresie linia de forma aleasă care se "potrivește" cel mai bine datelor în raport cu un anumit model și un anumit criteriu dinainte stabilite. Exemple: dreapta de regresie, hiperbola de regresie. Modele și criteriile:

1° Modelul 1 - regresia lui y în x

Modelul: x este la îndemâna experimentatorului, iar variabila y depinde de x după o funcție de o anumită formă $f(x)$ la care se adaugă o variație aleatoare ε de medie zero și dispersie finită și relativ mică ($y = f(x) + \varepsilon$). Un **criteriu** este metoda celei mai mici sume de pătrate (de distanțe între puncte și corespondentele lor pe linia de regresie care au aceleași abscise x) și sn **metoda celor mai mici pătrate**.

2° Alte modele de regresie

Modelul 2 - **regresia lui x în y** , modelul 3 - **regresia ortogonală**.

7.3.3. Ajustarea liniară - dreapta de regresie

Formulele teoretice și **formulele de calcul rapid și exact:**

Dreapta de regresie a lui y în x are ecuația (graficului) $Y = a + b \cdot x$ în care

$$b = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{N \cdot \sum x \cdot y - \sum x \cdot \sum y}{N \sum x^2 - (\sum x)^2}$$

$$a = \bar{y} - b \cdot \bar{x} = \frac{\sum y - b \cdot \sum x}{N}$$

b sn **coeficient de regresie (a lui y în x)**.

1° Proveniența denumirii de regresie

Francis Galton (1886) a observat că fiii (la maturitate) se abat de la înălțimea medie mai puțin decât tații, deci că fiii **regresează** către medie. Sintagma *linie (dreaptă sau curbă) de regresie* este improprie. Se păstrează din motive de tradiție. Cititorul trebuie să se gândească, de fapt la ideea de **linie de dependență, corelație, predicție, estimatie sau de tendință**.

2° Ajustări liniarizabile

Curbe liniarizabile: hiperbolele, curbele exponențiale, logaritmice, putere, logistice.

3° Rostul unei ajustări:

în cazul regresiei lui y în x , posibilitatea de a prognoza valori ale lui y în funcție de valori ale lui x atât prin **interpolare**, cât și prin **extrapolare**.

7.3.4. Măsurarea gradului de corelație după o anumită formă prin indice de corelație specific

Sn **indice de corelație de forma $Y = f(x)$** (a lui y în x) și se notează η_{yx} ($Y = f(x)$) expresia:

* (curbă **sigmoidă**, adică de forma unui "s" alungit)

$\sqrt{\frac{S_{Y\bar{y}}^2}{S_{y\bar{y}}^2}}$, în care $S_{Y\bar{y}}^2 = \sum (Y - \bar{y})^2$ este **variația explicată** de linia de regresie

$$Y = f(x),$$

iar $S_{y\bar{y}}^2 = \sum (y - \bar{y})^2$ este **variația totală**.

Proprietate de aditivitate: "Variația totală = variația explicată + variația reziduală."

7.3.5. Coeficientul de determinație și coeficientul de corelație liniară

Sn **covariația** seriei bidimensionale (x, y) de volum N , expresia $\sum (x - \bar{x}) \cdot (y - \bar{y})$

și **covarianța** seriei (notată **cov** (x,y)) covariația divizată prin volumul N .

1° Coeficientul de determinație

Sn **coeficient de determinație** (al unei serii bidimensionale) raportul

$$R^2 = \frac{cov^2(x, y)}{var(x) \cdot var(y)}.$$

Pătratul indicelui de corelație în raport cu dreapta de regresie este egal cu coeficientul de determinație: $\eta_{yx}^2 (Y = a + bx) = R^2 \Rightarrow R^2$ variază între 0 și 1.

2° Coeficientul de corelație liniară (Bravais-Pearson)

Formula teoretică și

formula de calcul rapid și exact:

Sn **coeficientul de corelație liniară (Bravais-Pearson)** (al unei serii bidimensionale) raportul:

$$R = \frac{cov(x, y)}{S_x \cdot S_y} = \frac{N \cdot \sum x \cdot y - \sum x \cdot \sum y}{\sqrt{(N \cdot \sum x^2 - (\sum x)^2) \cdot (N \cdot \sum y^2 - (\sum y)^2)}}, \text{ în care } S_x,$$

S_y sunt abaterile standard ale seriei x , respectiv y .

1) Pătratul coeficientului de corelație liniară R este coeficientul de determinație.

2) $-1 \leq R \leq 1$.

3) $R > 0 \Leftrightarrow$ **corelație liniară directă** și $(3') R < 0 \Leftrightarrow$ **corelație liniară inversă**.

4) $R = 1 \Leftrightarrow$ **corelație liniară directă și perfectă** \Leftrightarrow toate punctele sunt așezate pe dreapta de regresie, aceasta având panta $b > 0$ și

4') $R = -1 \Leftrightarrow$ **corelație liniară inversă și perfectă** \Leftrightarrow toate punctele sunt așezate pe dreapta de regresie, aceasta având panta $b < 0$.

7.3.6. Calcul simultan, rapid și exact al dreptei de regresie și al coeficientului de corelație liniară

Se utilizează tabelul:

x	y	x^2	y^2	xy
.	.			
.	.			
.	.			
Σx	Σy	Σx^2	Σy^2	Σxy

și formulele de calcul pentru b , a și R .

Proгноza prin ecuația de regresie

7.4. Test de semnificație pentru coeficientul de corelație liniară r

$H_0: \rho = 0$ (nu există corelație liniară)

$H_A: \rho \neq 0$ (există corelație liniară)

Statistica testului: r .

Decizia statistică: Se respinge H_0 dacă r calculat pe baza eșantionului este mai mare în modul decât valoarea tabelată în Anexa 5, în coloana corespunzătoare nivelului de semnificație ales și linia corespunzătoare numărului de grade de libertate $\nu = n$ (numărul de perechi (x,y) din eșantion) - 2*.

Observație: Acceptarea ipotezei nule nu înseamnă că nu există (nici o formă de) corelație, ci doar că nu există corelație liniară, putând exista, eventual, o altă formă neliniară de corelație.

* Se scad două grade de libertate deoarece se consideră fixate mediile valorilor x , respectiv y .

L₁₀ Teste, exerciții și probleme

TG10. Durata 180" pe calculator.

Alegeti afirmatia corecta:

1. O legatura intre variabile cantitative se numeste asociere, iar intre variabile calitative, corelatie.
2. O legatura intre variabile cantitative se numeste corelatie, iar intre variabile calitative, asociere.
3. O legatura intre variabile cantitative se numeste corelatie, iar intre variabile calitative, contingenta.

Predictia comportarii unei variabile in raport cu valorile ori variantele altei variabile este posibila daca si numai daca:

1. intre cele doua variabile exista o legatura functionala sau cel putin statistica
2. intre cele doua variabile exista o legatura functionala
3. cele doua variabile sunt independente una fata de cealalta

In majoritatea fenomenelor biologice, ecologice, intalnim legaturi:

1. functionale
2. statistice

Fie afirmatiile:

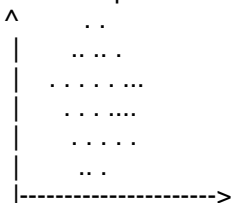
A='exista o legatura, o conexiune in fenomen'

B='exista o corelatie (grafica, numerica) in datele numerice care exprima fenomenul'

Alegeti propozitia adevarata:

1. $B \implies A$
2. $A \iff B$
3. $A \implies B$

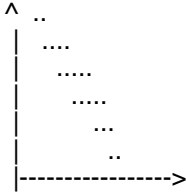
O diagrama de imprastiere de forma:



dovedeste:

1. dependenta statistica liniara
2. independenta statistica
3. dependenta directa

O diagrama de imprastiere de forma:



exprima o eventuala:

1. dependenta statistica inversa si liniara
2. independenta statistica
3. dependenta directa

Coeficientul de corelatie liniara, R , este nul in caz de:

1. independenta
2. existenta a unei alte forme de corelatie decat cea liniara
3. independenta (statistica sau functionala) sau daca exista o alta forma de corelatie decat cea liniara

TC10. Durata 410''.

1. Sinteza grafică în cazul statisticii descriptive bivariate se realizează prin tabele de _____ pentru variabile cantitative și prin tabele de _____ pentru variabile calitative.
2. În cazul dependenței _____, pe diagrama de împrăștiere punctele vor forma un nor în jurul unei anumite linii de _____.
3. Recunoaștem independența _____ când norul de puncte este plasat în jurul unei drepte paralele la una din axe, iar independența _____ se manifestă printr-un nor de puncte circular sau eliptic.
4. Prin demers complet putem dovedi cu certitudine doar lipsa unei dependențe _____.
5. Susținerea unei forme de dependență și măsurarea intensității acesteia se execută în următoarele 3 etape: 1. _____; 2. _____; 3. _____.
6. Variația _____ este suma dintre variația explicată (de linia de regresie) și variația _____.

7. Cu cât _____ este mai mic, cu atât r va trebui să fie mai _____ pentru a susține statistic existența unei corelații liniare afirmate din considerente biologice.

Exerciții sau probleme rezolvate

EGA 1.

Se da urmatorul sir bidimensional reprezentand un esantion aleator de 4 perechi de valori x, y :

x	y
1	1
3	3
5	5
5	6

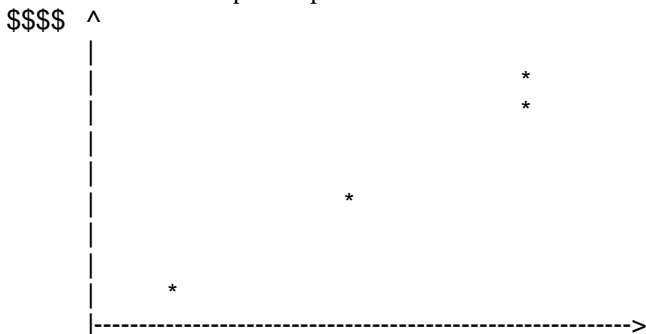
Știind ca în fenomen există o legătură între cele două variabile:

1. să se reprezinte sirul într-o digramă de împrăștiere;
 2. să se calculeze r , coeficientul de corelație liniară și dreapta de regresie a lui y în x ;
 3. consultând tabela corespunzătoare să se testeze semnificația lui r ;
 4. să se prognozeze, dacă este cazul, valoarea lui y pentru $x_0 = 7$.
- (Rezultatele vor fi rotunjite la 2 zecimale.)

Rezolvare:

LEGENDA: \$=0.(3) puncte. \$\$\$ din oficiu. Numarul de \$ se imparte la 30.

1. Desenăm norul de puncte pentru a vedea dacă forma corelației este liniară:



Putem considera că forma corelației este liniară. Este o corelație liniară directă ($r > 0$), dreapta de regresie fiind ascendentă ($b > 0$), dar punctele nu se așează toate pe dreapta de regresie ($r < 1$), corelația nefiind perfectă.

2. Calcul manual corelație și regresie liniară. (Rezultatele finale rotunjite la 2 zecimale.) $n = 4$

x	y	x ²	y ²	xy
1	1	1	1	1
3	3	9	9	9
5	5	25	25	25
5	6	25	36	30
Sx	Sy	Sx ²	Sy ²	Sxy
\$ 14	\$ 15	\$ 60	\$ 71	\$ 65

$$n \cdot Sxy - Sx \cdot Sy = 4 \cdot 65 - 14 \cdot 15 = 50$$

$$n \cdot Sx^2 - Sx \cdot Sx = 4 \cdot 60 - 14 \cdot 14 = 44$$

$$n \cdot Sy^2 - Sy \cdot Sy = 4 \cdot 71 - 15 \cdot 15 = 59$$

$$r = (n \cdot Sxy - Sx \cdot Sy) / \sqrt{(n \cdot Sx^2 - Sx \cdot Sx) \cdot (n \cdot Sy^2 - Sy \cdot Sy)} = 50 / \sqrt{44 \cdot 59} =$$

$$50 / 50.95096 = .9813358 \approx .98$$

$$b = (n \cdot Sxy - Sx \cdot Sy) / (n \cdot Sx^2 - Sx \cdot Sx) = 50 / 44 = 1.136364 \approx 1.14$$

$$a = (Sy - b \cdot Sx) / n = (15 - 1.136364 \cdot 14) / 4 = -.2272727 \approx -.23$$

3. Semnificatia lui r:

$$\text{\$ Numar de grade de libertate, niu} = n - 2 = 4 - 2 = 2$$

$$\text{\$ Deoarece } |r| = .98 > r \text{ tabelat (pt. alfa} = 0,05 \text{ si niu} = 2) = .95$$

exista o corelatie LINIARA semnificativ diferita de 0.

(altfel scris, $r = .98 \Leftrightarrow 0^{\dagger}$ ori $R \Leftrightarrow 0^*$.)

4. Predictie. Pt. $x_0 = 7$

$$\text{\$ } y_0 = a + b \cdot x_0 = -.23 + 1.14 \cdot 7 = 7.75 \approx 7.75$$

EGA 2.

Se da urmatorul sir bidimensional reprezentand un esantion aleator de 6 perechi de valori x,y:

x	y
1	1
1	2
2	1
2	2
1.5	1.5
1.3	1.4

Stiind ca in fenomen exista o legatura intre cele doua variabile:

1. sa se reprezinte sirul intr-o digrama de imprastiere;

* Semnul "=" înseamn` "≈".

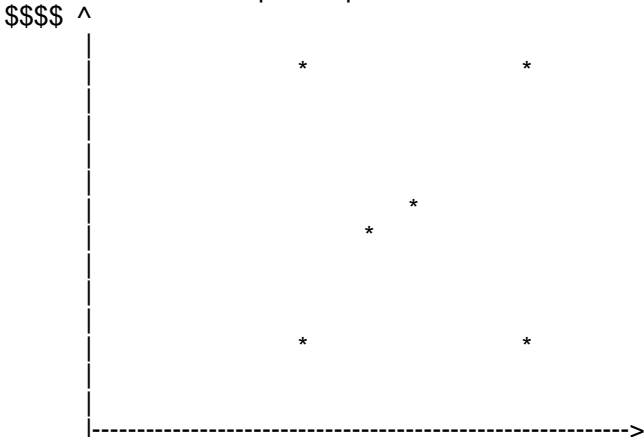
† Semnul "<" înseamn` "≠".

2. sa se calculeze r , coeficientul de corelatie liniara si dreapta de regresie a lui y in x ;
 3. consultand tabela corespunzatoare sa se testeze semnificatia lui r ;
 4. sa se prognozeze, daca este cazul, valoarea lui y pentru $x_0 = 0.5$.
- (Rezultatele vor fi rotunjite la 2 zecimale.)

Rezolvare:

LEGENDA: \$=0.(3) puncte. \$\$\$ din oficiu. Numarul de \$ se imparte la 30.

1. Desenam norul de puncte pentru a vedea daca forma corelatiei este liniara:



Norul aproximativ circular pentru acest eșantion indică lipsa oricărei corelații. Acest eșantion nu evidențiază corelația afirmată în fenomen sau aceasta nu există. Pentru alegerea între cele două variante este necesară repetarea observației sau a experimentului cu date mult mai multe, precum și reconsiderarea argumentelor biologice care susțin legătura respectivă.

2. Calcul manual corelatie si regresie liniare. (Rezultatele finale rotunjite la 2 zecimale.) $n = 6$

x	y	x ²	y ²	xy
1	1	1	1	1
1	2	1	4	2
2	1	4	1	2
2	2	4	4	4
1.5	1.5	2.25	2.25	2.25
1.3	1.4	1.69	1.96	1.82
Sx	Sy	Sx ²	Sy ²	Sxy
\$ 8.8	\$ 8.9	\$ 13.94	\$ 14.21	\$ 13.07

$$\sum n \cdot S_{xy} - S_x \cdot S_y = 6 \cdot 13.07 - 8.8 \cdot 8.9 = 9.999983E-02^*$$

$$\sum n \cdot S_x^2 - S_x \cdot S_x = 6 \cdot 13.94 - 8.8 \cdot 8.8 = 6.199994$$

$$\sum n \cdot S_y^2 - S_y \cdot S_y = 6 \cdot 14.21 - 8.9 \cdot 8.9 = 6.050007$$

$$r = (\sum n \cdot S_{xy} - S_x \cdot S_y) / \sqrt{(\sum n \cdot S_x^2 - S_x \cdot S_x) \cdot (\sum n \cdot S_y^2 - S_y \cdot S_y)} =$$

$$9.999983E-02 / \sqrt{6.199994 \cdot 6.050007} =$$

$$9.999983E-02 / 6.124541 = 1.632773E-02 \approx 0.02$$

$$b = (\sum n \cdot S_{xy} - S_x \cdot S_y) / (\sum n \cdot S_x^2 - S_x \cdot S_x) = 9.999983E-02 / 6.199994 =$$

$$1.612902E-02 \approx 0.02$$

$$a = (S_y - b \cdot S_x) / n = (8.9 - 1.612902E-02 \cdot 8.8) / 6 = 1.459677 \approx 1.46$$

3. Semnificatia lui r:

$$\text{\$ Numar de grade de libertate, niu} = n - 2 = 6 - 2 = 4$$

$$\text{\$ Deoarece } |r| = 0.02 < r \text{ tabelat (pt. alfa} = 0,05 \text{ si niu} = 4) = 0.811$$

nu exista o corelatie LINIARA semnificativ diferita de 0.

(altfel scris, $r = 0.02 \approx 0$ ori $R = 0$.)

Corelatia nefiind LINIARA in mod cel putin semnificativ ($R = 0$),
nu are sens prognoza prin DREAPTA de regresie.

3.

Un biolog a găsit urma labei din spate a unui urs. Aceasta avea 26 cm. Dorind să estimeze lungimea corpului ursului respectiv, a tranchilizat 9 urși și le-a măsurat lungimea labei din spate, precum și lungimea corpului. A obținut următoarea serie bidimensională:

Lungimea labei din spate, în cm	16	18	18	19	22	22	23	25	27
Lungimea corpului, în cm	181	202	192	215	234	225	237	245	256

Cum a procedat în continuare și ce valoare a obținut pentru lungimea corpului ursului ?

Rezolvare:

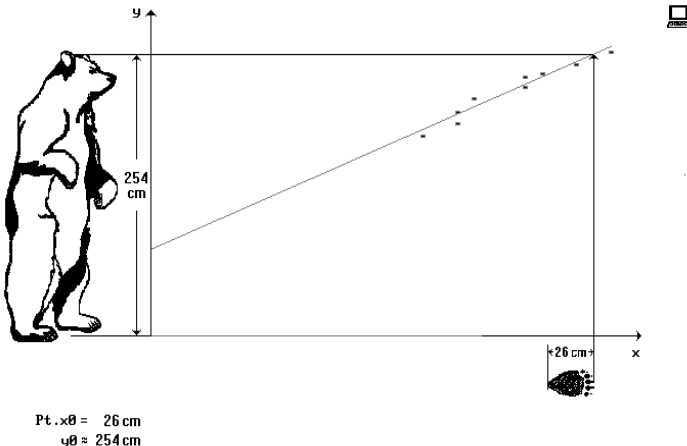
Între lungimile a diferite segmente ale unui organism viu există, de regulă, o legătură directă. Desenând norul celor 9 puncte se observă că putem aproxima forma legăturii printr-o dreaptă (vezi figura următoare). De aceea, vom calcula dreapta de regresie a lui y în x , y fiind lungimea corpului, iar x lungimea labei din spate. Pentru a măsura gradul de grupare a punctelor în jurul drepte de regresie calculăm și r , coeficientul de corelație liniară. Deoarece cele nouă perechi de puncte formează un eșantion, testăm semnificația lui r . Dacă r este cel puțin semnificativ (diferit de zero) putem considera corelația liniară, cel

* $9.999983E-02$ înseamnă $9.999983 \cdot 10^{-2} = 9.999983 / 100 = 0.0999983 \approx 0,1$.

puțin semnificativă și va avea sens să facem o prognoză (estimare) prin dreapta de regresie determinată pe baza celor 9 puncte.
 Calculul se derulează ca în exercițiile anterioare.

$b \approx 6.78; a \approx 77.63; r \approx 0,97.$

Semnificatia lui r:
 Număr de grade de libertate, $v = n - 2 = 9 - 2 = 7$. Deoarece $r = 0,97 > r$ tabelat (pentru $\alpha = 0,001$ și $v = 7$) = 0,898, există o corelație liniară înalt semnificativ diferită de 0 (altfel scris:
 $r = 0,97 > 0$ ***.)



Predicție: Pentru $x_0 = 26 \Rightarrow y_0 = a + b \cdot x_0 = 77,63 + 6,78 \cdot 26 = 253,91 \approx 254$.
 Deci ursul cu lungimea labei din spate de 26 de cm are, cel mai probabil, lungimea de 254 cm.

Exerciții sau probleme propuse

EGA 4

Se da urmatorul sir bidimensional reprezentand un esantion aleator de 8 perechi de valori x,y:

x	y
3	1
3	2
8	6
10	12
11	18
14	24
17	26
40	28

Stiind ca in fenomen exista o legatura intre cele doua variabile

1. sa se reprezinte sirul intr-o diagrama de imprastiere.
 2. Sa se calculeze r , coeficientul de corelatie liniara, si dreapta de regresie a lui y in x .
 3. Consultand tabela corespunzatoare sa se testeze semnificatia lui r .
 4. Sa se prognozeze, daca este cazul, valoarea lui y pentru $x_0 = 27$.
- (Rezultatele finale vor fi rotunjite la 2 zecimale.)

EGA

Pentru fiecare şir bivariat din următoarele patru:

5.		6.		7.		8.	
x	y	x	y	x	y	x	y
0	0	2	4	1.1	1	0	0
0	1	2	3	2	2.3	1	2
1	0	4	2	3	3.6	2	4
1	1	4	3	4	4	3	2
		6	1	5	5	4	0
		1	6	5.5	6		

stiind ca in fenomen exista o legatura intre cele doua variabile:

0. sa se reprezinte sirul intr-o diagrama de imprastiere;
 1. sa se calculeze dreapta de regresie a lui y in x si r , coeficientul de corelatie liniara;
 2. consultand tabela corespunzatoare sa se testeze semnificatia lui r ;
 3. sa se prognozeze, daca este cazul, valoarea lui y pentru $x_0 = 2$.
- (Rezultatele finale vor fi rotunjite la 2 zecimale.)

9.

În silvicultură, pentru estimarea masei lemnoase (adică a volumului de lemn) dintr-un brad, se măsoară diametrul acestuia la înălțimea de 1,30m de la sol. Volumul se calculează pe baza diametrului măsurat și a înălțimii bradului. Aceasta din urmă, fiind dificil de măsurat la brazii nedoborâți, poate fi prognozată pe baza legăturii statistice liniare existente între diametru și înălțime. Să se prognozeze înălțimea unui brad nedoborât cu diametrul de 64 cm (la 1,30m de sol), dacă pentru 10 brazi doborâți s-au obținut următoarele diametre, respectiv înălțimi:

Diametrul (la 1,30m de la sol), în cm	52	56	60	64	68	72	56	60	72	68
Înălțimea, în m	24	27	26	32	31	35	24	31	36	34

10.

O problemă interesantă, în special pentru biochimiști, se găsește în [2] pag. 103: "Într-o experiență la lucrările practice de «Metabolism», anul IV Biochimie, s-au obținut următoarele rezultate reprezentând perechile (x,y) de concentrații, densități optice:

x	1	2	3	4	5	6	7	8	9	10
y	0,09	0,14	0,2	0,275	0,345	0,41	0,475	0,57	0,64	0,72

S-a cerut determinarea concentrației într-un ser care a indicat concentrația optică $y_1 = 0,52$."

Înainte de dozării (stabilirii dozei, concentrației) asigurați-vă că dreapta de etalonare exprimă o corelație semnificativă.