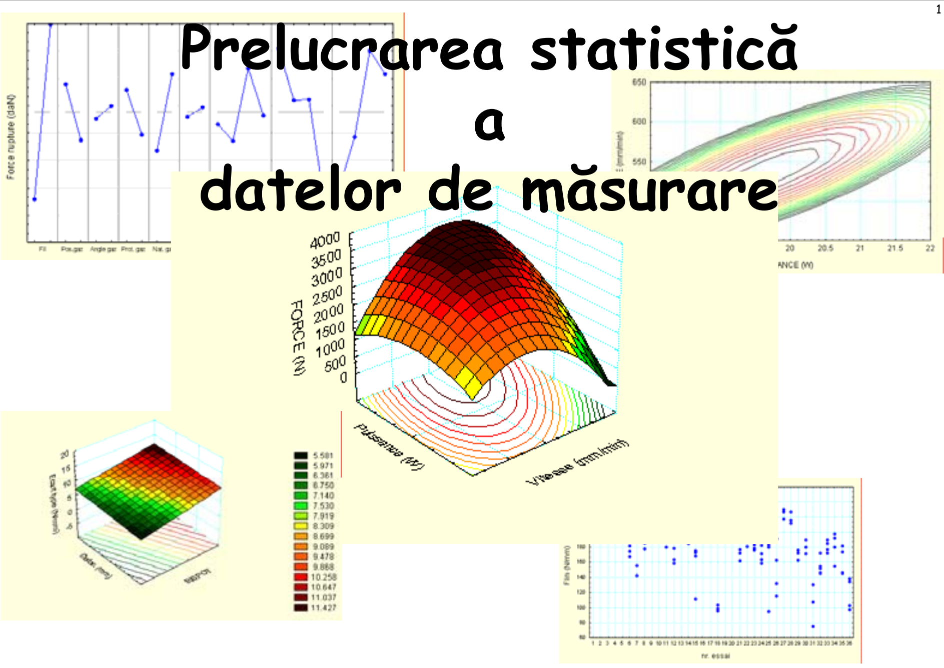
Prelucrarea statistică a datelor de măsurare



Veridicitatea rezultatelor din stiintele naturii si inginerie
$\Rightarrow$ existenta rezultatelor experimentale

Experiment $\Rightarrow$ Proces de masurare

Tema experimentala (de masurare):
ce trebuie facut?

Strategia experimentala (de masurare):
cum trebuie procedat?

Prelucrarea rezultatelor, formularea concluziilor, luarea deciziilor

Prelucrarea statistică a datelor de măsurare

Factori perturbatori ai procesului de masurare

Cauze:
- principiul sau metoda de masurare;
- mijloacele de masurare;
- mediul ambiant;
- obiectul supus masurarii;
- interactiunea obiect supus masurarii – mijloc de masurare;
- operator.

Conditii de referinta: temperatura, presiune, umiditate, vibratii etc.

Categorii de erori de masurare

- Erori grosolane $\Rightarrow$ rezultate aberante (eliminate);
- Erori sistematice $\Rightarrow$ prezinta repetabilitate; partial compensate, partial eliminate;
- Erori aleatoare $\Rightarrow$ aparitie intamplatoare; pot fi diminuate, dar nu eliminate.

Prelucrarea rezultatelor masurarilor se face in conditiile prezentei (cel putin a) erorilor aleatoare de masurare.

m $\Rightarrow$ adevarata valoare a marimii masurate; z_i $\Rightarrow$ erori aleatoare; x_i $\Rightarrow$ rezultatele masurarilor

$$x_i = m + z_i$$

Elemente de teoria erorilor

Repetand de un numar mare de ori (in conditii identice) masurarea unei marimi a carei adevatata valoare este **m**, se constata ca rezultatele (aleatoare) x_i ale masurarilor si, implicit, erorile aleatoare de masurare z_i, respecta urmatoarele axiome (postulate) :

- Principiul cauzal - z_i mici, mai frecvente ca z_i mari;
- Principiul limitativ - z_i inferioare limitei care cumuleaza toate cauzele de erori;
- Principiul distributiv - suma algebrica a erorilor tinde la zero;
- Principiul probabilistic - probabilitatea aparitiei unei erori depinde numai de marimea sa.

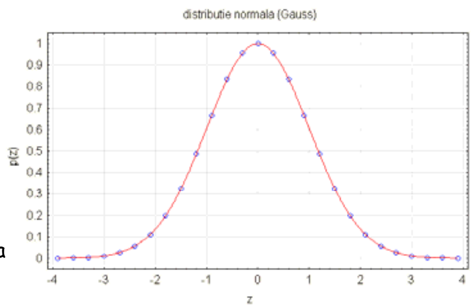
$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n}}$$ deviatie standard

$\sigma^2 \Leftrightarrow$ dispersie

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$ deviatie standard esantion

$s^2 \Leftrightarrow$ dispersie empirica

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$ medie aritmetica

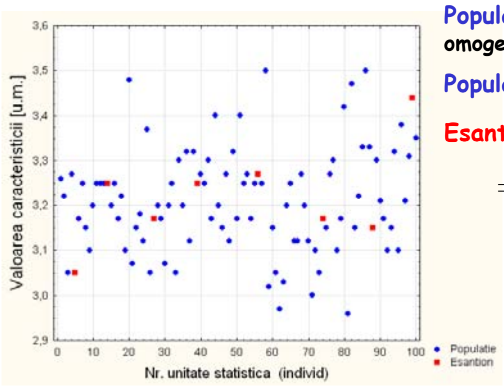


distributie normala $$p(z) = f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(z^2/2\sigma^2)}$$

Variabila aleatoare

STATISTICA ⇒ **ramura matematicii aplicate care cuprinde un grup de metode de calcul cu ajutorul carora se pot obtine informatii privind fenomenele de masa.**

NOTIUNI DE BAZA: - **populatia statistica** (colectivitatea statistica);
- **esantionul** (proba; selectie) **prelevat din populatie.**



Populatia = **totalitatea obiectelor calitativ omogene** (la care se urmareste o caracteristica)

Populatia compusa din **unitati statistice** (indivizi)

Esantionul ⇒ utilizat la estimarea caracteristicii (proprietatii) populatiei ⇒
⇒ reprezentativ (tehnica de prelevare) & & volumul esantionului

Valorile caracteristicii
sunt caracterizate prin parametri statistici:
- parametri de grupare (media aritmetica);
- parametri de imprastiere (dispersia)

Populatia ⇒ μ ; σ^2 ⇒ constante
Esantionul ⇒ ; s^2 ⇒ variabile

notiuni elementare de statistica

Valori: $x_1, x_2, \ldots, x_n$ ⇒ parametri de grupare si de imprastiere

$x_1, x_2, \ldots, x_n$ ≡ valori masurate

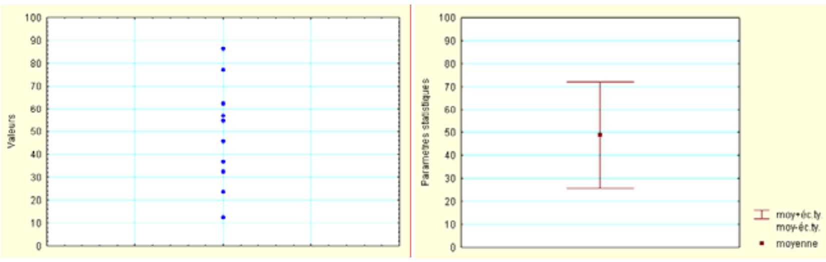
Parametru de grupare: media aritmetica

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Parametri de imprastiere: abaterea standard; dispersia

$$s = \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_n-\bar{x})^2}{n-1}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \qquad s^2$$

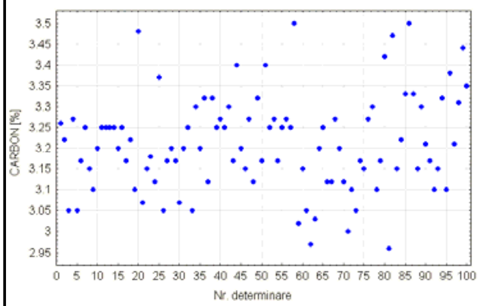
Exemplu
Valori:
23,5;
86,3;
45,8;
12,4;
32,6;
54,8;
36,7;
62,4;
77,3;
56,8



n = 10 $\bar{x}$ = 48,86 s = 23,3 s^2 = 542,89

Densitati de repartitie (1)

7

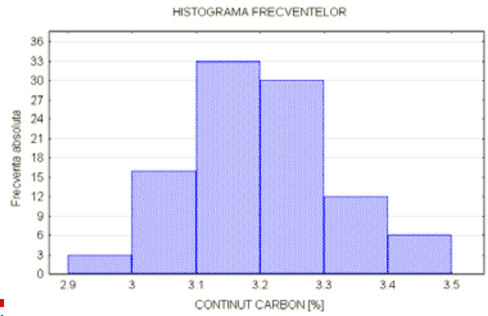


Esantion n=100 determinari

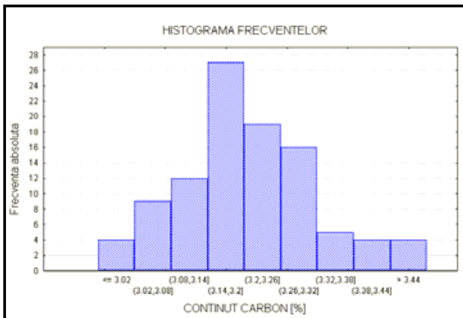
x_{min} = 2,96% x_{max} = 3,50%

HISTOGRAMA FRECVENTELOR

(2,9; 3,0] $\Rightarrow n_1=3$
(3,0; 3,1] $\Rightarrow n_2=16$
(3,1; 3,2] $\Rightarrow n_3=33$
(3,2; 3,3] $\Rightarrow n_4=30$
(3,3; 3,4] $\Rightarrow n_5=12$
(3,4; 3,5] $\Rightarrow n_5=6$

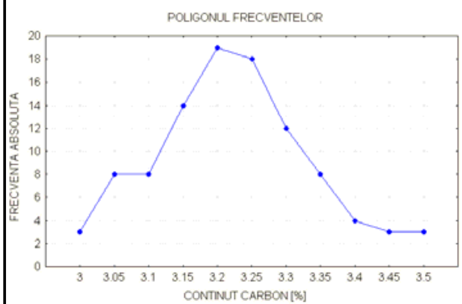


8

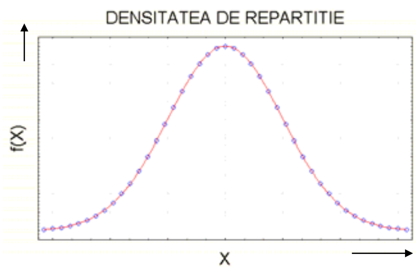


Densitati de repartitie (2)

n=100 determinari

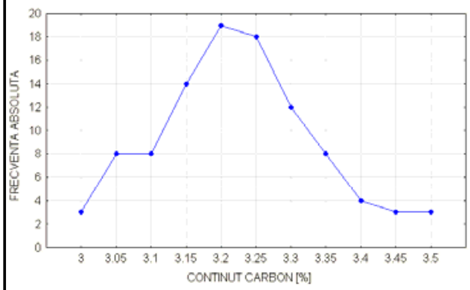


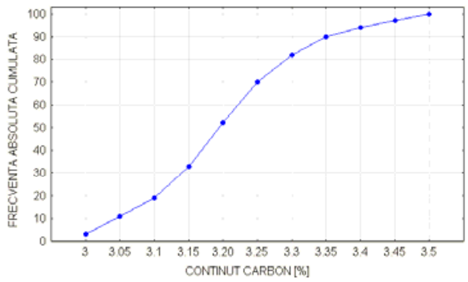
$n \to \infty$
repartitia de frecventa
densitatea de repartitie



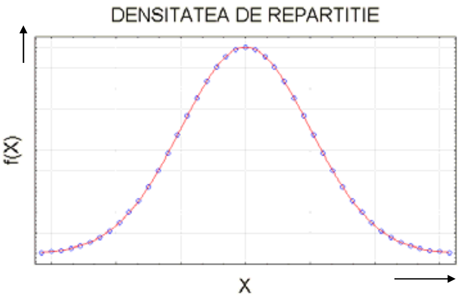
n finit		$n \to \infty$
FRECVENTA aparitiei	$\Rightarrow$	PROBABILITATEA aparitiei

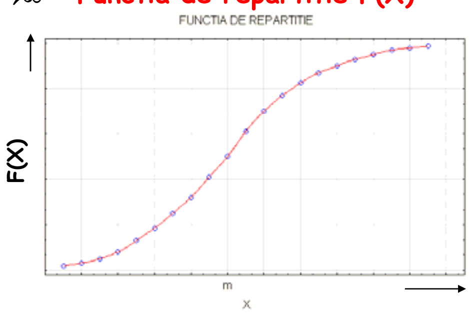
frecventa **n=100 determinari** **frecventa cumulata**





Densitatea de repartitie f(X) **n→∞** **Functia de repartitie F(X)**





Aplicatia 1 – parametri statistici (1)

1. Formularea problemei

În vederea efectuării unui studiu privind costurile implementării unui nou procedeu tehnologic, a fost realizată prelucrarea a diferite materiale prin procedeul tehnologic supus analizei. Pentru fiecare material prelucrat, a fost înregistrată productivitatea prelucrării *Qp* [mm^3/min] (volumul de material îndepărtat în unitatea de timp). Rezultatele măsurărilor sunt precizate în tab.A1.1.

Evaluarea raportului (costuri/performanţe ale procedeului de prelucrare) urmează a fi realizată pe baza unei analize statistice, fapt pentru care se cere:

A. calculul următorilor parametri statistici ai şirului valorilor obţinute în urma măsurării:
media aritmetică;
mediana;
modul;
media geometrică;
dispersia (abaterea medie pătratică experimentală);
deviaţia standard;
eroarea standard;
valorile minimă şi maximă precum şi mărimea intervalului dintre acestea;
valorile corespunzătoarele cuartilelor superioară şi inferioară precum şi lungimea intervalului intercuartilic;
valorile asimetriei şi asimetriei standard;
valorile excesului şi excesului standard;
coeficientul de variaţie (eroarea relativă);
suma tuturor valorilor.

B. reprezentarea grafică a poligonului frecvenţelor şi a histogramei frecvenţelor (atât pentru frecvenţe absolute cât şi relative) şi comentarea rezultatelor obţinute.

Aplicatia 1 – parametri statistici (2)

Tab.A1.1 Valorile măsurate ale productivității prelucrării

Nr. crt.	Q_p [mm^3/min]	Nr. crt.	Q_p [mm^3/min]	Nr. crt.	Q_p [mm^3/min]	Nr. crt.	Q_p [mm^3/min]	Nr. crt.	Q_p [mm^3/min]
1	2775	18	4080	35	3940	52	2905	69	2575
2	3365	19	2155	36	1915	53	2490	70	2525
3	3735	20	2230	37	2670	54	2635	71	2735
4	3570	21	2745	38	3900	55	2620	72	2865
5	3530	22	2855	39	3420	56	2725	73	3035
6	3155	23	3245	40	2200	57	2385	74	2125
7	2965	24	2990	41	1800	58	1875	75	2125
8	2720	25	2890	42	2670	59	2215	76	2945
9	3430	26	3265	43	2595	60	2045	77	3015
10	3210	27	3360	44	2700	61	2380	78	2585
11	3380	28	3840	45	2556	62	3415	79	2835
12	3070	29	3725	46	2120	63	3725	80	2370
13	3620	30	3955	47	2678	64	3060	81	2950
14	3410	31	3830	48	2870	65	3465	82	2790
15	3425	32	4360	49	3003	66	2605	83	2295
16	3445	33	4054	50	3381	67	2640	84	2625
17	3205	34	3605	51	2800	68	2395	85	2720

Aplicatia 1 – parametri statistici (3)

2. Rezolvare

În vederea efectuării calculelor, rezultatele măsurărilor din tab.A1.1 se notează:

$x1, x2, \ldots, xi, \ldots, xn \quad , \quad i = 1,\ldots,85$

În urma prelucrării rezultatelor experimentale, se obțin următoarele valori:

a. Media aritmetică: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{251107}{85} = 2954{,}2 \ [\mathrm{mm}^3/\mathrm{min}]$

b. Mediana: $Me = 2870$ [mm^3/min]

c. Modul: $Mo = 2670$ [mm^3/min]

d. Media geometrică: $\bar{x}_g = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} = 2898{,}77 \ [\mathrm{mm}^3/\mathrm{min}]$

e. Dispersia (abaterea medie pătratică experimentală): $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 328700 \ [\mathrm{mm}^3/\mathrm{min}]^2$

f. Deviația standard (abaterea standard): $s = \sqrt{s^2} = 573{,}323 \ [\mathrm{mm}^3/\mathrm{min}]$

g. Eroarea standard: $\frac{s}{\sqrt{n}} = \frac{579{,}682}{\sqrt{85}} = 62{,}1857 \ [\mathrm{mm}^3/\mathrm{min}]$

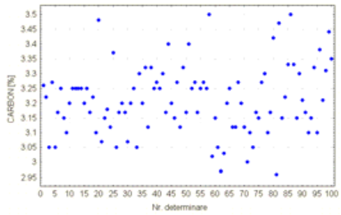
…………

j. Asimetria: $\bar{\gamma}_1 = \frac{n \cdot \sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)\cdot(n-2)\cdot s^3} = 0{,}217666$

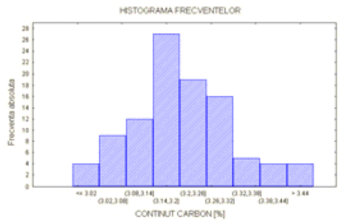
Aplicatia 1 – parametri statistici (4)

13

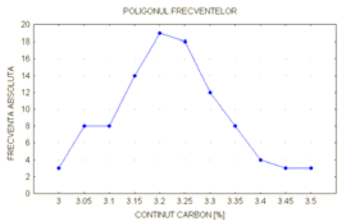
Norul de puncte (exemplu)



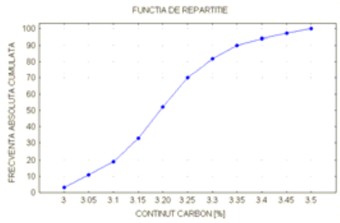
Histograma frecventelor (exemplu)



Poligonul frecventelor (exemplu)



Poligonul frecventelor cumulate(exemplu)



Elemente de teoria estimatiilor

14

$x_1, x_2, \ldots, x_n \Rightarrow$ **n** rezultate ale masurarilor, fara erori grosolane sau sistematice, distributie normala

Estimarea adevaratei valori a marimii masurate, **m**, implica:

- sa se determine o functie $f(x_1, x_2, \ldots, x_n)$ care sa furnizeze o valoare suficient de apropiata de **m** ;
- sa se determine un interval $(f-e_1;\ f+e_2)$, care cu o probabilitate impusa $P=1-\alpha$, sa contina valoarea **m**

$P \Rightarrow$ nivel de incredere (siguranta a estimatiei); $(f-e_1;\ f+e_2) \Rightarrow$ interval de incredere;

Observatie: de regula $e_1=e_2=e$, iar $f(x_1, x_2, \ldots, x_n) = \bar{x} \Rightarrow \bar{x} - e < m < \bar{x} + e$

$$e = [t(\alpha, n)\cdot s]/[n^{1/2}]$$

exemplu

X_i	35,6	35,9	36,1	36,2	36,6
f_i	1	3	3	2	1

$\bar{x}$ = 36,06; s = 0,2633; $t_{tab(0,01;\ 9)}$= 3,25 $\Rightarrow$ **e = 0,27**

Intervalul de incredere $\Rightarrow$ (36,06-0,27; 36,06+0,27) $\equiv$ **(35,79; 36,33)**

Verificarea ipotezelor statistice

Ipoteza statistica (H):

orice consideratie despre proprietatile multimii din care a fost extrasa o proba (esantion; selectie).

- Ipoteza initiala (de nul) H_0
- Ipoteza concurenta (alternativa) H_1

Emiterea ipotezei ⇒ verificarea ipotezei statistice

Etape parcurse pentru verificarea ipotezei statistice:

1. Calculul marimii θ_{calc} (pe baza datelor existente si functie de testul statistic aplicat);
2. Alegerea valorii critice θ_{tab} (din tabele adecvate);
3. Compararea $\theta_{calc} \Leftrightarrow \theta_{tab}$;
4. Acceptarea sau respingerea ipotezei de nul (pentru pragul de semnificatie α ales).

Categorii de probleme care apeleaza la verificarea ipotezelor statistice

- Eliminarea rezultatelor aberante (testul Student; testul Grubbs-Smirnov etc.)
- Verificarea normalitatii distributiei datelor (testul χ^2)
- Compararea dispersiilor (testul Fisher; testul Cochran)
- Compararea mediilor aritmetice (testul Student)

Eliminarea rezultatelor aberante

n=10 ⇒ $x_1, x_2, \ldots, x_{10}$ ⇒ 72,5; 59,4; 78,0; 68,0; 63,0; 70,1; 72,9; 68,5; 54,5; 75,6

H_0 ⇒ minim; maxim sunt rezultate aberante? α=0,05

$GS_{calc} = |x^* - \bar{x}| / s$ $\bar{x}$ = 68,25 ; s =7,36 54,5 ⇒ GS_{calc} =1,868; 78,0 ⇒ GS_{calc}=1,325

GS_{tab}=f(n,P)

P=1-α

P \ n	2	3	4	5	6	7	9	10	12	15	20
0,90	1,18	1,50	1,70	1,84	1,94	2,02	2,15	2,20	2,28	2,38	2,60
0,95	1,39	1,74	1,94	2,08	2,18	2,27	2,39	2,44	2,52	2,62	2,73
0,99	1,82	2,22	2,43	2,57	2,68	2,76	2,88	2,93	3,01	3,10	3,21

$GS_{calc} < GS_{tab}$ ⇒ 54,5 si 78,0 nu sunt rezultate aberante pentru α = 0,05

$t_{calc} = |x^* - \bar{x}| / s$ 54,5 ⇒ $\bar{x}$ = 69,78 ; s =5,89 ⇒ t_{calc} = 2,594

78,0 ⇒ $\bar{x}$ = 67,17 ; s =6,91 ⇒ t_{calc} = 1,567

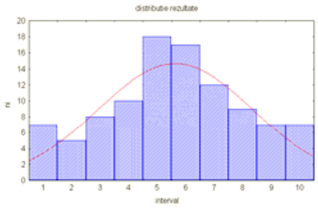
Verificarea normalitatii distributiei datelor

N=100 valori $\Rightarrow$ $x_1, x_2, \ldots, x_{100}$ $\Rightarrow$ L=10 intervale ; $\bar{x}$ = 8,63 ; s=0,127 H_0 $\Rightarrow$ distributie normala; α=0,05

$$t_i = \frac{x_i - \bar{x}}{s} = \frac{x_i - 8{,}63}{0{,}127} \Rightarrow \Phi(t_i) \Rightarrow \text{prob.integrala} \Rightarrow \underset{\Phi(+\infty)=0{,}5}{\text{din tabele}} \Rightarrow p_i = \Phi(t_i) - \Phi(t_{i-1}) \Rightarrow \chi^2_{calc} = \sum_{i=1}^{L} \frac{(n_i - N \cdot p_i)^2}{N \cdot p_i}$$

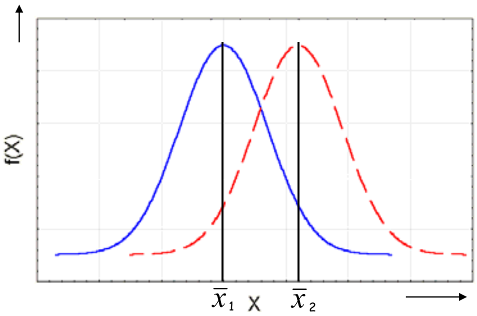
Daca: $\chi^2_{calc} > \chi^2_{tab\ (\alpha;L-3)}$ $\Rightarrow$ distributia difera de cea normala; P=1-α

Nr. interv.	interval $(x_{i-1}; x_i]$	n_i	t_i	$\Phi(t_i)$	p_i	χ^2_i
1	(-∞; 8,425]	7	-1,614	-0,4467	0,0533	0,523
2	(8,425;8,475]	5	-1,220	-0,3888	0,0579	0,108
3	(8,475;8,525]	8	-0,827	-0,2959	0,0929	0,179
4	(8,525;8,575]	10	-0,433	-0,1676	0,1283	0,624
5	(8,575;8,625]	18	-0,039	-0,0156	0,1520	0,516
6	(8,625;8,675]	17	0,354	0,1383	0,1539	0,168
7	(8,675;8,725]	12	0,748	0,2728	0,1345	0,157
8	(8,725;8,775]	9	1,142	0,3733	0,1005	0,110
9	(8,775;8,825]	7	1,536	0,4377	0,0644	0,048
10	(8,825;+∞)	7	+∞	0,5000	0,0623	0,095
		100			1,0000	2,528



$\chi^2_{tab\ (0{,}05;7)}$ =14,07

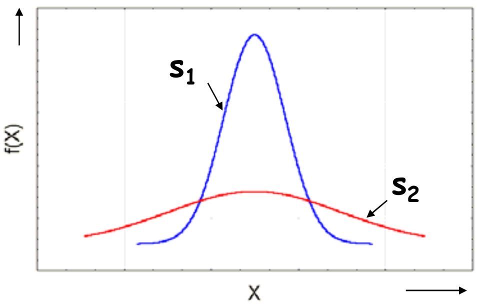
Compararea dispersiilor si a mediilor aritmetice



Densitati de repartitie ce difera prin parametrul de grupare

$$\bar{x}_1 < \bar{x}_2$$

Parametri statistici: media aritmetica, mediana, modul, media geometrica, coeficientul de variatie, amplitudinea, asimetria, deviatia standard, dispersia, excesul etc.



Densitati de repartitie ce difera prin parametrul de imprastiere

$$s_1 < s_2$$

19

Compararea dispersiilor

Compararea a doua dispersii

$x_1', x_2', \ldots, x_{n1}' \Rightarrow$ **s_1^2**; $x_1'', x_2'', \ldots, x_{n2}'' \Rightarrow$ **s_2^2**; $s_1^2 > s_2^2 \Rightarrow$ **F_{calc}** $= s_1^2/s_2^2$ daca $F_{calc} > F_{tab(\alpha;\, \nu1;\, \nu2)} \Rightarrow$ dif. semnif.

exemplu

$n_1 = 200 \Rightarrow$ **s_1^2** $= 3{,}82$

$n_2 = 15$

X_i''	44,6	45,7	46,5	48,5	49,5
f_i	2	4	4	4	1

s_2^2 $= 2{,}41$ $\alpha = 0{,}05$

$H_0 \Rightarrow$ diferenta semnificativa

F_{calc} $= s_1^2/s_2^2 = 1{,}585$ $F_{tab(\alpha;\, \nu1;\, \nu2)} =$ **$F_{tab(0,05;\, 199;\, 14)}$** $= 2{,}13$ $\Rightarrow F_{calc} < F_{tab} \Rightarrow$ **se respinge H_0** **P = 0,95**

Compararea a k dispersii

k serii de masurari ; **n** masurari/serie $\Rightarrow s_1^2; s_2^2; \ldots; s_k^2$; $(s_1^2 > s_i^2, \forall i > 1)$; **$G_{calc}$** $= s_1^2/\Sigma s_i^2$, i=1,…, n

daca $G_{calc} > G_{tab(\alpha;\, k;\, \nu)} \Rightarrow$ dif. semnif.

exemplu

$k = 6$; $n=7 \Rightarrow$ **s_1^2** $= 3{,}82$; **s_2^2** $= 1{,}70$; **s_3^2** $= 1{,}30$; **s_4^2** $= 0{,}92$; **s_5^2** $= 0{,}78$; **s_6^2** $= 0{,}81$ $\alpha = 0{,}05$ **H_0** $\Rightarrow$ dif. semnif

$G_{calc} = 3{,}82/9{,}33 = 0{,}409$ $G_{tab(0,05;6;6)} = 0{,}418$ $\Rightarrow G_{calc} < G_{tab} \Rightarrow$ **se respinge H_0** **P = 0,95**

20

Compararea mediilor aritmetice

Compararea a doua medii aritmetice

$x_1', x_2', \ldots, x_{n1}' \Rightarrow \bar{x}_1$; **s_1^2**; $x_1'', x_2'', \ldots, x_{n2}'' \Rightarrow \bar{x}_2$; **s_2^2**; Ipoteza: s_1^2 ; s_2^2 = dispersii omogene

$H_0 \Rightarrow$ diferenta semnificativa

$$t_{calc} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_{ech}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} \qquad s_{ech} = \sqrt{\frac{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1 + n_2 - 2}}$$

daca $t_{calc} > t_{tab(\alpha;\, n1+n2-2)} \Rightarrow$ diferenta semnificativa (se accepta H_0) pt. $P=1-\alpha$

exemplu

$n_1 = 25 \Rightarrow \bar{x}_1 = 23{,}56$ **s_1^2** $= 1{,}21$ $n_2 = 50 \Rightarrow \bar{x}_2 = 22{,}80$ **s_2^2** $= 1{,}56$

$H_0 \Rightarrow$ diferenta semnificativa **P = 0,99**

s_{ech} $= 1{,}20 \Rightarrow t_{calc} = 2{,}584$; $t_{tab(0,01;\, 73)} = 2{,}645$; $\Rightarrow t_{calc} < t_{tab} \Rightarrow$ **se respinge H_0** **P = 0,99**

Aplicatia 2 – verificarea ipotezelor statistice (1)

Formularea problemei

Se efectuează un studiu economic care presupune compararea din punct de vedere al performanţelor şi al costurilor a unor mărci de automobile de pe piaţa mondială. Pentru realizarea studiului, au fost selecţionate mărci de automobile fabricate în: SUA, Uniunea Europeană şi Japonia.

Cu caracteristicile autoturismelor analizate, s-a alcătuit fişierul AUTO, care cuprinde 150 de înregistrări, corespunzătoare la tot atâtea mărci de automobile.

Caracteristicile urmărite (parametrii) au fost următoarele:
- consum de benzină (mile străbătute cu un galon de benzină) – **mpg**;
- număr de cilindri (4, 6 sau 8 cilindri) – **cil**;
- deplasament – **depl**;
- putere motor – **putere**;
- timpul de accelerare de la 0 la 100 km/h – **acc**;
- anul de fabricaţie – **an**;
- masa automobilului – **masa**;
- zona de origine – **origin**;
- firma producătoare – **prod**;
- model automobil – **model**;
- preţul automobilului – **pret**;
- cod zonă de origine: SUA – 1; CE – 2; Japonia – 3 – **codorig**.

Tab. 1
Valorile puterii autovehiculelor

Nr.crt.	Putere [CP]	Nr.crt.	Putere [CP]	Nr.crt.	Putere [CP]	Nr.crt.	Putere [CP]	Nr.crt.	Putere [CP]
1	48	**31**	103	**61**	69	**91**	88	**121**	88
2	66	**32**	125	**62**	90	**92**	72	**122**	88
3	52	**33**	115	**63**	115	**93**	84	**123**	88
4	70	**34**	133	**64**	115	**94**	84	**124**	85
5	60	**35**	71	**65**	90	**95**	92	**125**	84
6	110	**36**	68	**66**	76	**96**	110	**126**	90
7	140	**37**	115	**67**	60	**97**	84	**127**	92
8	139	**38**	85	**68**	70	**98**	58	**128**	74
9	105	**39**	88	**69**	65	**99**	64	**129**	68
10	95	**40**	90	**70**	90	**100**	60	**130**	68
11	85	**41**	110	**71**	88	**101**	67	**131**	63
12	88	**42**	130	**72**	90	**102**	65	**132**	70
13	100	**43**	129	**73**	90	**103**	62	**133**	88
14	90	**44**	138	**74**	78	**104**	68	**134**	75
15	105	**45**	135	**75**	90	**105**	63	**135**	70
16	85	**46**	155	**76**	75	**106**	65	**136**	67
17	110	**47**	142	**77**	92	**107**	65	**137**	67
18	120	**48**	125	**78**	75	**108**	74	**138**	67
19	145	**49**	150	**79**	65	**109**	75	**139**	110
20	165	**50**	71	**80**	105	**110**	75	**140**	85
21	139	**51**	65	**81**	65	**111**	100	**141**	92
22	140	**52**	80	**82**	48	**112**	74	**142**	112
23	68	**53**	80	**83**	48	**113**	80	**143**	96
24	95	**54**	77	**84**	67	**114**	76	**144**	84
25	97	**55**	125	**85**	67	**115**	116	**145**	90
26	75	**56**	71	**86**	67	**116**	120	**146**	86
27	95	**57**	90	**87**	67	**117**	110	**147**	52
28	105	**58**	70	**88**	62	**118**	105	**148**	84
29	85	**59**	70	**89**	132	**119**	88	**149**	79
30	97	**60**	65	**90**	100	**120**	85	**150**	82

Aplicatia 2 – verificarea ipotezelor statistice (3)

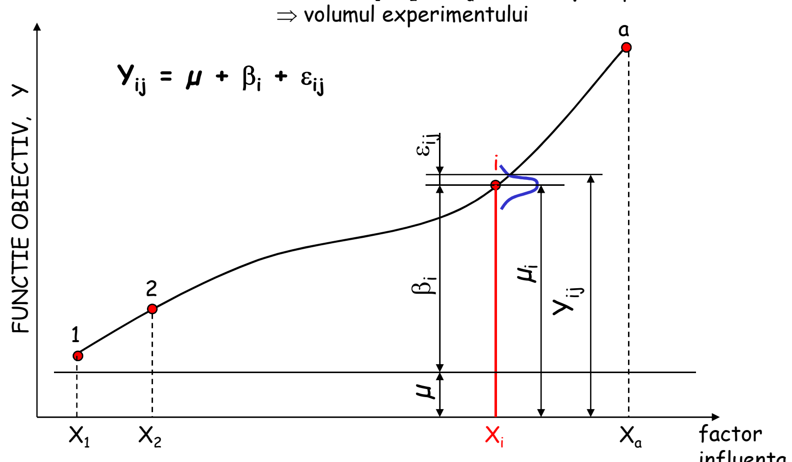
Se cere:

1. să se completeze fişierul AUTO cu cele 150 de valori ale puterii motorului (**putere**) din Tab.1;
   se apelează la instrucţiunea de editare a fişierelor şi se introduce o nouă coloană;
2. să se determine numărul de automobile apaţinând fiecărei zone de origine, precum şi valoarea medie, minimă şi maximă a consumului de benzină, pentru fiecare dintre cele trei zone de origine;
   se recurge la calculul parametrilor statistici (Summary Statistics);
3. să se stabilească dacă pentru un prag de semnificaţie α=0,05, consumurile de benzină ale automobilelor fabricate în SUA şi în CE diferă semnificativ de valoarea **mpg** = 50 şi dacă da, în ce sens?;
   se recurge la analiza unui şir de date (One Sample Analysis); se selectează din totalul valorilor cele referitoare la automobilele analizate (Select);
4. să se găsească un prag de semnificaţie pentru care consumul de benzină al automobilelor fabricate în Japonia nu diferă semnificativ de valoarea **mpg** = 31,5; pentru o siguranţă a estimaţiei de 95%; cât este intervalul de încredere al mediei aritmetice în acest caz?;
   se recurge la analiza unui şir de date (One Sample Analysis); se selectează din totalul valorilor cele referitoare la automobilele analizate (Select);
5. să se stabilească dacă din punct de vedere al masei automobilului şi al preţului acestuia, automobilele fabricate în SUA şi Japonia sunt echivalente, pentru un prag de semnificaţie α=0,05; dar din punct de vedere al timpului de accelerare?;
   se recurge la analiza a două şiruri de date (Two Sample Analysis); se selectează din totalul valorilor cele referitoare la automobilele analizate (Select);
6. să se decidă dacă în cele trei zone de origine, maşinile cu 4 cilindri sunt echivalente din punctul de vedere al caracteristicilor **mpg**, **masa**, **acc** şi **pret**; dar maşinile cu 6 cilindri, din punct de vedere al anului de fabricaţie? pentru α=0,05;
   se recurge la analiza a două şiruri de date (Two Sample Analysis); se selectează din totalul valorilor cele referitoare la automobilele analizate (Select);
7. să se stabilească dacă există vreo zonă de origine în care diferenţele de preţ între automobilele cu 4 şi cu 6 cilindri sa nu fie semnificative (pentru α=0,05);
   se recurge la analiza a două şiruri de date (Two Sample Analysis); se selectează din totalul valorilor cele referitoare la automobilele analizate (Select);
8. să se reprezinte histograma frecvenţelor valorilor consumului de benzină pentru automobilele fabricate în SUA, precizându-se în axe inclusiv unităţile de măsură şi zona de origine a automobilelor pentru care s-a făcut reprezentarea.
   se recurge la reprezentarea grafică a histogramei frecvenţelor (Frequency Histogram); se selectează din totalul valorilor cele referitoare la automobilele analizate.

Analiza dispersionala unifactoriala (1)

Este utilizata pentru verificarea semnificatiei efectelor produse de catre un factor de influenta X, asupra unei functii obiectiv Y, intr-un domeniu analizat.

Date initiale: **FO**; **FI** si **nivelele FI**: X_1; X_2;…; X_a; nr. de **replici** pe fiecare nivel, n ⇒ ⇒ volumul experimentului



$H_0 \Rightarrow \mu_1 = \mu_2 = … = \mu_i = … = \mu_a \Leftrightarrow \beta_1 = … = \beta_i = … = \beta_a$

Analiza dispersionala unifactoriala (2)

FO: y; FI: x; m ⇒ 4 nivele ale FI; n = 5 replici pe fiecare nivel

$$S_1 = \sum_{i=1}^{m}\sum_{j=1}^{n} y_{ij}^2 \qquad S_2 = n\cdot\sum_{i=1}^{n}(\bar{y}_i)^2 \qquad S_3 = m\cdot n\cdot \bar{y}_T^2 \qquad \bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$$

$s^2_{nivel} = (S_2 - S_3)/(m-1)$ $\qquad$ $s^2_e = (S_1 - S_2)/m(n-1)$

Nr. repl.	Nivele ale factorului X			
	x_1	x_2	x_3	x_4
1	56	64	45	42
2	55	61	46	39
3	62	50	45	45
4	59	55	39	43
5	60	56	43	41
suma nivel	292	286	218	210
medie nivel	58,4	57,2	43,6	42,0

Sursa dispersiei	Suma patratelor	Grade de libertate	Dispersii estimate	F_{calc}
Nivelele factorului de influenta	$S_2 - S_3 =$ 1135	$\nu = m-1=$ 3	$s^2_{nivel} =$ 378,3	$\frac{s^2_{niv}}{s^2_e}$ 29,8
Erorile aleatoare de masurare	$S_1 - S_2 =$ 203,2	$\nu_0 = m(n-1)=$ 16	$s^2_e =$ 12,7	
Dispersia totala	$S_1 - S_3 =$ 1338,2	mn-1= 19	-	

$F_{tab(\alpha;\ \nu;\ \nu 0)} = F_{tab(0,05;\ 3;\ 16)} = 5,29$ **deoarece** $F_{calc} > F_{tab}$ ⇒

factorul X are influenta semnificativa asupra lui Y

Aplicatia 3 – analiza dispersionala unifactoriala (1)

Formularea problemei

Un beneficiar producător de confecții este interesat de maximizarea rezistenței la întindere a unei noi fibre sintetice. El dorește să afle dacă procentul de bumbac din fibră afectează această rezistență și în ce mod.

Se cunoaște faptul că pentru a avea celelalte calități cerute, fibra trebuie să conțină între 10% și 40% bumbac.

Se considera ca problema propusă poate fi studiata apelandu-se la analiza dispersionala unifactoriala.

Pentru aceasta, se alege:

- ca functie obiectiv $y \equiv Rm$ rezistența la întindere a fibrei in [N/cm^2];
- ca factor de influenta x, procentul de bumbac din fibră;

în domeniul de interes pentru beneficiar, factorului de influenta i se fixeaza $a = 5$ nivele de variatie, corespunzatoare urmatoarelor continuturi de bumbac: 10%, 15%, 20%, 25% si 30%, pentru fiecare nivel efectuandu-se cate $n = 5$ determinari (replici) ale rezistenței fibrei; rezulta pentru intregul experiment un numar de $N = a{\cdot}n = 25$ masurari.

Cele 25 de determinari ale rezistenței fibrei au fost efectuate în ordine aleatoare, pentru a evita influenaa factorilor sistematici asupra rezultatelor masurarilor.

Ordinea de efectuare a incercarilor, precum si rezultatele obtinute sunt prezentate in tab.1.

Aplicatia 3 – analiza dispersionala unifactoriala (2)

Tab.1 Rezultatele masurarilor in ordinea de efectuare a acestora

Numar masurare	Continut de bumbac [%]	Rezistență [N/cm²]	Numar masurare	Continut de bumbac [%]	Rezistență [N/cm²]
1	15	17,5	14	15	20,4
2	25	57,9	15	10	16,2
3	15	16,4	16	30	99,8
4	30	107,3	17	30	105,6
5	25	67,6	18	20	25,6
6	10	14,7	19	10	16,9
7	20	21,7	20	20	24,5
8	15	15,9	21	30	101,4
9	20	22,8	22	25	58,4
10	25	64,5	23	30	109,5
11	15	22,2	24	25	68,6
12	10	12,9	25	10	13,4
13	20	16,4	-	-	-

Se cere:

a) să se decidă asupra faptului dacă procentul de bumbac din fibră influenţează semnificativ rezistenţa la întindere a acesteia, pentru o siguranţă a afirmaţiei de 95%;

b) să se stabilească pentru ce procent de bumbac din fibră se obţine valoarea maximă a rezistenţei acesteia; să se precizeze care este valoarea medie a celor cinci replici corespunzătoare acestui procent de bumbac, precum şi intervalul de încredere în care se situează adevărata valoare a rezistenţei fibrei, pentru o siguranţă a estimaţiei de 95%;

c) să se compare grafic valorile medii ale rezistenţei fibrei pentru diferitele procente de bumbac şi să se aprecieze dacă există mai multe valori ale procentului de bumbac pentru care rezistenţele la rupere ale fibrelor să nu difere semnificativ, pentru o siguranţă a estimaţiei de 95%;

d) să se stabilească dacă reziduurile prezintă o distribuţie aleatoare sau urmează o anumită tendinţă, în timpul realizării încercărilor;

e) să se stabilească toate valorile conţinutului de bumbac pentru care apar diferenţe semnificative între rezistenţele fibrelor; să se facă aceeaşi analiză utilizând testele Scheffe şi Duncan şi să se formuleze concluziile finale.

Aplicatia 3 – analiza dispersionala unifactoriala (3)

<u>Aplicarea metodologiei de calcul</u>

Pentru aplicarea metodologiei de calcul, in tab.2, este data matricea-program a experimentului, obţinută pe baza rezultatelor din tab.1.

In tab.2 apar si sumele valorilor functiei obiectiv (rezistenţa fibrei) corespunzatoare nivelelor, precum si mediile $\overline{y}_i$ ale valorilor functiei obiectiv, corespunzatoare fiecarui nivel, i al factorului de influenta *(continutul de bumbac).*

Tab.2 Matricea-program a experimentului

Numar replica	Valorile masurate ale rezistenţelor y_{ij} [N/cm²] functie de nivelul i al factorului de influenta *(procentul de bumbac din fibră)*				
j	10%	15%	20%	25%	30%
1	14,7	17,5	21,7	57,9	107,3
2	12,9	16,4	22,8	67,6	99,8
3	16,2	15,9	16,4	64,5	105,6
4	16,9	22,2	25,6	58,4	101,4
5	13,4	20,4	24,5	68,6	109,5
Suma nivel	74,1	92,4	111	317	523,6
Medie nivel $\overline{y}_i$	14,82	18,48	22,2	63,4	104,72

Valori globale: $\sum_{i=1}^{5}\sum_{j=1}^{5} y_{ij} = 699{,}22$; $\overline{y}_T = 139{,}8$

Aplicatia 3 – analiza dispersionala unifactoriala (4)

In cazul utilizarii analizei dispersionale unifactoriale, se presupune ca rezultatele masurarilor pot fi puse sub forma:

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \beta_i + \varepsilon_{ij}\ , \quad i=1,\ldots,5\ ; \quad j=1,\ldots,5\ ,$$

$\mu \Rightarrow$ centrul de grupare global al tuturor rezultatelor masurarilor y_{ij} (media aritmetica a intregii populatii y);

$\mu_i \Rightarrow$ mediile aritmetice corespunzatoare nivelelor (centrul de grupare al valorilor masurate pentru nivelul i al factorului de influenta analizat);

$y_{ij} \Rightarrow$ rezultatele experimentale, corespunzatoare celor $a \cdot n$ determinari;

$\varepsilon_{ij} \Rightarrow$ valorile erorilor aleatoare de masurare, avand repartitii normale independente in jurul valorilor μ_i , repartitii caracterizate de parametrii $(0, \sigma^2)$.

iar ipoteza de nul ce se doreste a fi verificata este de forma:

$$H_0 : \quad \beta_1 = \beta_2 = \ldots = \beta_5$$

Metodologia de calcul pentru cazul aplicarii analizei dispersionale unifactoriale, conduce la rezultatele sintetizate in tab.3.

Tab.3 Rezultatele aplicarii metodologiei de calcul pentru analiza dispersionala unifactoriala

Sursa dispersiei	Suma patratelor	Grade de libertate	Dispersii (estimate)	Criteriu statistic
Nivelele factorului de influenta	$SS_{nivel} = 30193{,}202$	a - 1= 4	$s^2_{nivel} = 7548{,}3004$	$F_{calculat} = \dfrac{s^2_{nivel}}{s^2_e}$ = 582,511
Erorile aleatoare de masurare	$SS_e = 259{,}164$	N - a = 20	$s^2_e = 12{,}9582$	
Dispersia totala	$SS_T = 30452{,}366$	N - 1= 24	-	

Aplicatia 3 – analiza dispersionala unifactoriala (5)

Valoarea criteriului F_{tab} pentru un prag de semnificatie al testului: $\alpha = 0{,}05$ se alege pentru $\nu_1 = 4$ si $\nu_2 = 20$ din tabele adecvate, rezultand:

$$F_{0,05;4;20} = 2{,}83$$

Deoarece:

$$F_{calculat} > F_{tab}\ ,$$

rezulta cu probabilitatea P = 1- α = 0,95, respectiv cu o siguranta a estimatiei de 95% ca ipoteza de nul se respinge, deci ca mediile aritmetice ale valorilor funcaiei obiectiv corespunzatoare nivelelor factorului de influenta difera semnificativ;

aceasta este echivalent cu a spune ca procentul de bumbac din fibră influenteaza semnificativ rezistența acesteia.

Valoarea R^2 calculata cu relatia este: $$R^2 = \frac{SS_{nivel}}{SS_T} = \frac{30193{,}202}{30452{,}366} = 0{,}9915$$

rezultand ca peste 99% din imprastierea rezultatelor masurarilor in jurul valorii mediei lor aritmetice poate fi explicata pe baza analizei dispersionale unifactoriale. Marimea R^2 (coeficientul de pondere) este o masura a adecvantei aplicarii metodei analizei dispersionale si respectiv a influentei continutului de bumbac asupra rezistenței fibrei.

Aplicatia 3 – analiza dispersionala unifactoriala (6)

31

Tab.4 Marimi estimate pe baza rezultatelor masurarilor

Nivel factor i	Procent bumbac [%]	Numar determinari n	Medie nivel $\bar{y}_i$ [N/cm²]	Deviatia standard [N/cm²]	Interval de incredere 95% [N/cm²]
1	10	5	14,82	0,7729166	[11,46; 18,18]
2	15	5	18,48	1,2138369	[15,12; 21,84]
3	20	5	22,20	1,5984637	[18,84; 25,56]
4	25	5	63,40	2,2487774	[60,04; 66,76]
5	30	5	104,72	1,8098066	[101,36; 108,08]

Valorile estimate ale mediilor rezistențelor fibrei corespunzatoare procentului de bumbac din aceasta, deviatiile standard pentru fiecare nivel al factorului de influenta, precum si intervalele de incredere ale mediilor corespunzatoare unei sigurante a afirmatiei de 95% (α = 0,05), sunt centralizate in tab.4.

Tab.5 Analiza tuturor perechilor de medii aritmetice prin metoda Scheffé

Numar comparatie	Nivele ale caror medii se compara (contrast)	Diferenta intre medii [N/cm²]	Existenta unor diferente semnificative (95%)
1	10% - 15%	14,82 - 18,48 = -3,66	NU
2	10% - 20%	14,82 - 22,20 = -7,38	NU
3	10% - 25%	14,82 - 63,40 = -48,58	DA
4	10% - 30%	14,82 - 104,72 = -89,90	DA
5	15% - 20%	18,48 - 22,20 = -3,72	NU
6	15% - 25%	18,48 - 63,40 = -44,92	DA
7	15% - 30%	18,48 - 104,72 = -86,24	DA
8	20% - 25%	22,20 - 63,40 = -41,20	DA
9	20% - 30%	22,20 - 104,72 = -82,52	DA
10	25% - 30%	63,40 - 104,72 = -41,32	DA

Pentru identificarea procentelor de bumbac care duc la obtinerea unor rezistențe ale fibrelor intre care nu exista diferente semnificative se poate aplica metoda Scheffé de analiza a contrastelor.

Prin aceeasta metoda se analizeaza toate comparatiile posibile intre oricare doua valori medii ale rezistențelor fibrelor, corespunzatoare utilizarii celor cinci continuturi de bumbac. Rezulta, cu o siguranta a afirmatiei de 95% concluziile sintetizate in tab.5.

Prin analiza tuturor contrastelor, rezulta ca fiind omogene (fara a produce diferente semnificative asupra functiei obiectiv) urmatoarele grupe de nivele: 10%; 15%; 20% bumbac.

Aplicatia 3 – analiza dispersionala unifactoriala (7)

32

In urma efectuarii deteminarilor experimentale si a prelucrarii statistice a rezultatelor prin metoda analizei dispersionale unifactoriale pot fi formulate urmatoarele concluzii:

1. procentul de bumbac din fibră influenteaza semnificativ rezistența acesteia la întindere *Rm*;

2. la cresterea conținutului de bumbac, in domeniul studiat, se inregistreaza o crestere a rezistenței fibrei, dupa cum urmeaza:

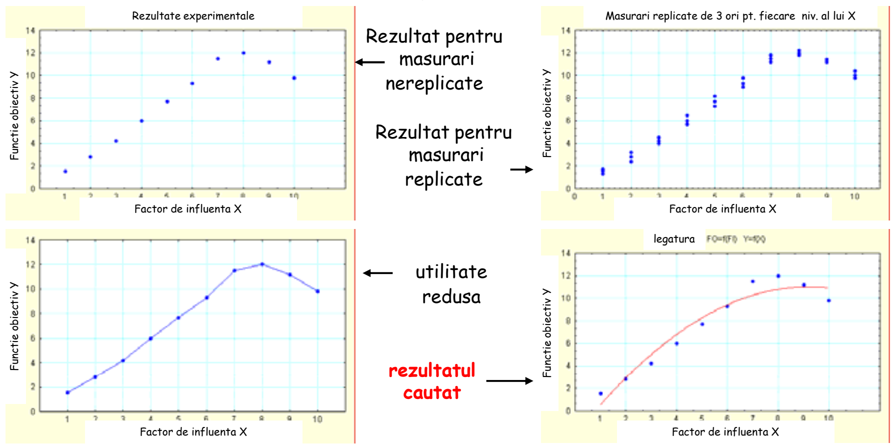
   - intre 10% si 20% bumbac crestere nesemnificativa din punct de vedere statistic;

   - la peste 20% bumbac modificarile procentului de bumbac din fibră conduc la cresteri semnificative ale rezistenței acesteia.

Metoda celor mai mici patrate. Analiza regresionala

Metoda celor mai mici patrate. Consideratii generale.

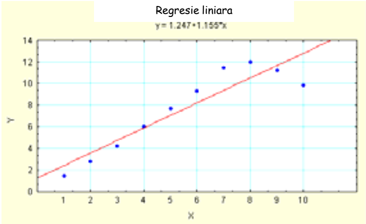
Problema de rezolvat: Sa se gaseasca cu ajutorul rezultatelor experimentale legatura: **FO=f(FI)**



Particularitatea problemei: Datorita prezentei erorilor experimentale, **nu trebuie ca punctele experimentale sa fie unite cu segmente**, ci trebuie sa se gaseasca "curba" care sa treaca **cat mai aproape** posibil **de ansamblul punctelor** experimentale, limitand pe cat posibil « zgomotul" experimental.

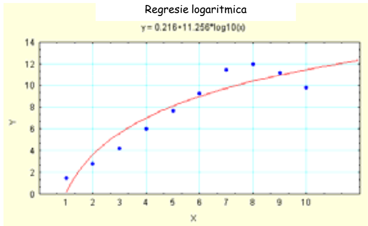
Metoda celor mai mici patrate
Cazul unei functii obiectiv de o singura variabila.

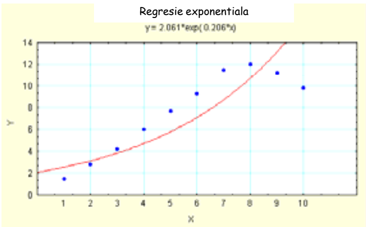
Observatie: Este necesara **definirea formei** generale **a modelului** experimental cautat, iar metoda permite particularizarea modelului (gasirea coeficientilor modelului) cu ajutorul rezultatelor experimentale .



← Model liniar

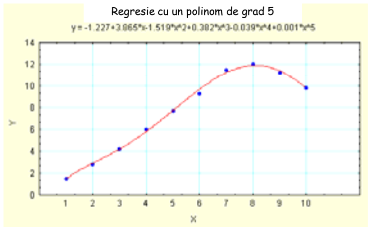
Model logaritmic →

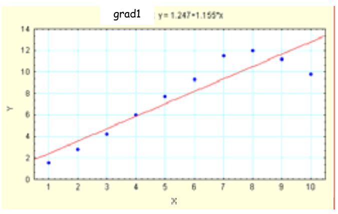




← Model exponential

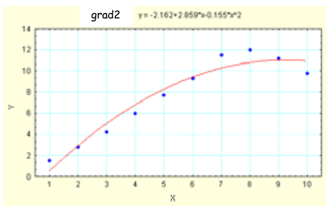
Model polinomial →

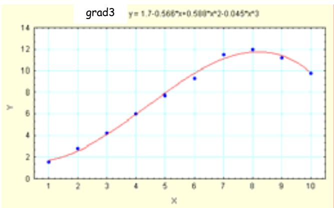




← $Y = b_0+b_1x$

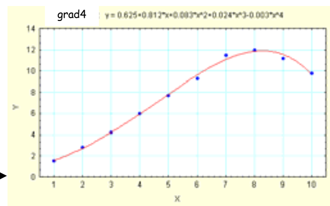
$Y = b_0+b_1x+b_2x^2$ →





← $Y = b_0+b_1x+b_2x^2+b_3x^3$

$Y = b_0+b_1x+b_2x^2+b_3x^3+b_4x^4$ →



Observatii:

Nr. coef. $b_t \leq N$

Gradul polinomului: se alege;

Poate fi identificat un **grad optimal al polinomului,** care depinde de N si de precizia de estimare dorita.

grad	b_0	b_1	b_2	b_3	b_4	b_5
1	1,247	1,155	-	-	-	-
2	-2,162	2,859	-0,155	-	-	-
3	1,700	-0,566	0,588	-0,045	-	-
4	0,625	0,812	0,083	0,024	-0,003	-
5	-1,227	3,865	-1,519	0,382	-0,039	0,001

Metodologia pentru aplicarea metodei celor mai mici patrate

1. Stabilirea FO si FI (domeniu, nivele, structura experimentului)
2. Realizarea incercarilor ⇒ rezultatele experimentale: $y_1, y_2, \ldots, y_i, \ldots, y_n$
3. Ipoteza asupra formei legaturii "f" FO↔FI
4. Explicitarea legaturii **$y=f(x_j, b_t)$** calculand coeficientii b_t ai modelului ($j=1,\ldots, k$; $t=1,\ldots, d$; y=FO; x_i=FI; b_j=coef. modelului)

$\tilde{y}_i$ ⇒Valori estimate; y_i ⇒Valori masurate;

Prin definitie: $S=s^2_{con}=\left[\sum_{i=1}^{n}(\tilde{y}_i-y_i)^2\right]/(n-1)=\left[\sum_{i=1}^{n}[f(x_j,b_t)_i-y_i]^2\right]/(n-1)$ dispersie de concordanta

Observatie: gasirea b_t minimizand dispersia de concordanta ⇒ rezolvarea sistemului de ecuatii (sistem de ecuatii normale)

$$\frac{\partial S}{\partial b_t}=0\ ,\ t=1,\ldots,d$$

$$\Rightarrow b_t$$

Calculul coeficientilor de regresie in cazul unei functii de gradul I de o singura variabila

$y=f(x_j, b_t)$ ⇒ $y=b_0 + b_1x$

Problema: gasirea lui b_0 si b_1 cu ajutorul a **n** rezultate experimentale

$$S=s^2_{con}=\left[\sum_{i=1}^{n}(\tilde{y}_i-y_i)^2\right]/(n-1)=\left[\sum_{i=1}^{n}[f(x_j,b_t)_i-y_i]^2\right]/(n-1)=\left[\sum_{i=1}^{n}[b_0+b_1x_i-y_i]^2\right]/(n-1)$$

$$\frac{\partial S}{\partial b_0}=0;\quad \frac{\partial S}{\partial b_1}=0 \quad \Rightarrow b_0 ; b_1$$

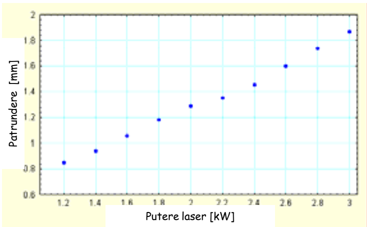
$$b_0=\frac{\sum_{i=1}^{n}y_i\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}x_iy_i}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}$$

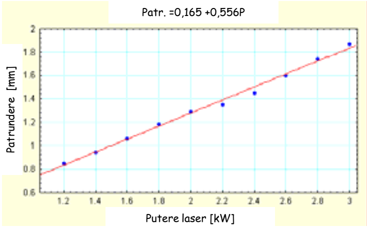
$$b_1=\frac{n\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}$$

Exemplu de aplicatie

Problema: sa se gaseasca dependenta patrunderii=f(puterea laser), pentru sudarea cap la cap a unui aliaj de aluminiu, cu ajutorul unui laser Nd:YAG, si pentru v=2m/min

$P \in [1,2,\ldots,3]$kW; n=10 incerc.





Incerc	P≡x [kW]	Patr.≡y [mm]	Incerc	P≡x [kW]	Patr. ≡y [mm]
1	1,2	0,85	6	2,2	1,35
2	1,4	0,94	7	2,4	1,45
3	1,6	1,06	8	2,6	1,60
4	1,8	1,18	9	2,8	1,74
5	2,0	1,29	10	3,0	1,87

$b_0 = 0{,}165$; $b_1 = 0{,}556$ $\Rightarrow y = b_0 + b_1 x \Rightarrow$

Patr.=0,165 + 0,556P

Incerc	patr.mas [mm]	patr.est. [mm]	Incerc	patr.mas [mm]	patr.est. [mm]
1	0,85	0,80	6	1,35	1,38
2	0,94	0,92	7	1,45	1,50
3	1,06	1,03	8	1,60	1,62
4	1,18	1,15	9	1,74	1,73
5	1,29	1,27	10	1,87	1,85

Exemplu de calcul

$y = b_0 + b_1 x$ ⇒

$$b_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

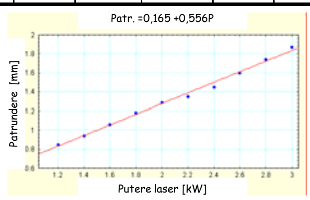
Incerc	P≡x [kW]	Patr.≡y [mm]	Incerc	P≡x [kW]	Patr. ≡y [mm]
1	1,2	0,85	6	2,2	1,35
2	1,4	0,94	7	2,4	1,45
3	1,6	1,06	8	2,6	1,60
4	1,8	1,18	9	2,8	1,74
5	2,0	1,29	10	3,0	1,87

$$n = 10; \quad \sum_{i=1}^{10} x_i = 21; \quad \sum_{i=1}^{10} y_i = 13{,}33;$$

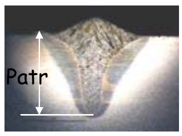
$$\sum_{i=1}^{10} x_i^2 = 47{,}4; \quad \left(\sum_{i=1}^{10} x_i\right)^2 = 441; \quad \sum_{i=1}^{10} x_i y_i = 29{,}828$$

⇒ $b_0 = 0{,}165$
$b_1 = 0{,}556$

patr = 0,165 + 0,556P



incerc	patr.est [mm]	incerc	patr.est. [mm]
1	0,80	6	1,38
2	0,92	7	1,50
3	1,03	8	1,62
4	1,15	9	1,73
5	1,27	10	1,85



Comparatie intre diferite modele liniare

incerc	P [kW]	Patr.mas.1 [mm]	Patr.mas.2 [mm]	Patr.mas.3 [mm]
1	1,2	0,85	0,85	-
2	1,4	0,94	-	-
3	1,6	1,06	-	-
4	1,8	1,18	-	-
5	2,0	1,29	-	1,29
6	2,2	1,35	-	-
7	2,4	1,45	-	1,45
8	2,6	1,60	-	-
9	2,8	1,74	-	-
10	3,0	1,87	1,87	-

incerc	Patr.mas [mm]	Patr.est1 [mm]	Patr.est2 [mm]	Patr.est3 [mm]
1	0,85	0,80	0,85	0,97
2	0,94	0,92	0,96	1,05
3	1,06	1,03	1,08	1,13
4	1,18	1,15	1,19	1,21
5	1,29	1,27	1,30	1,29
6	1,35	1,38	1,42	1,37
7	1,45	1,50	1,53	1,45
8	1,60	1,62	1,64	1,53
9	1,74	1,73	1,76	1,61
10	1,87	1,85	1,87	1,69
s^2 (dif.max)	0	102 (5)	143 (8)	869 (18)

N=10

$$patr = 0,165 + 0,556P$$

N=2

$$patr = 0,17 + 0,567P$$

N=2

$$patr = 0,49 + 0,40P$$