

**CURS – IV –**

**ANOVA**

## Metode statistice de analiza factorilor de variabilitate în experimentul biologic (ANOVA)

Să cercetăm, în continuare, problema comparării mai multor selecții provenite din populații pe care le știm ca fiind normal repartizate, de exemplu concentrațiile plasmatică realizate de tablete care conțin diferiți excipienți, dar care au aceeași substanță activă, în aceeași doză.

Vrem să verificăm ipoteza compusă că acestea provin de fapt din aceeași populație, având media  $\mu$  și dispersia  $\sigma$ , deci că excipienții folosiți nu influențează semnificativ cedarea și absorbția substanței active:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

față de ipoteza alternativă că cel puțin două medii nu sunt egale.

O variantă de rezolvare a problemei ar fi compararea mediilor de selecție două câte două prin metodele prezentate anterior.

Fie, de exemplu, relația între mediile de selecție  $\bar{x}_1 < \bar{x}_2 < \bar{x}_3 < \bar{x}_4$ . Este evident greșit a aplica o relație de tranzitivitate și a spune că

$$\mu_1 = \mu_2 \text{ și } \mu_2 = \mu_3 \text{ și } \mu_3 = \mu_4 \Rightarrow \mu_1 = \mu_2 = \mu_3 = \mu_4.$$

Motivul acestei erori este legat în primul rând de violarea unui principiu de bază al teoriei selecției: alegerea la întâmplare a selecțiilor. Ori compararea loturilor după criteriul “a posteriori”, al mărimii mediilor de selecție este într-adevăr o abatere de la acest principiu.

Mai mult, nici măcar  $\mu_1 = \mu_4$  nu implică în acest caz  $\mu_2 = \mu_3$  din cauza dependenței rezultatului testelor de relațiile între dispersiile populațiilor din care provin selecțiile.

Ca urmare, problema comparării mai multor selecții (loturi) trebuie abordată prin alte metode care să facă compararea tuturor selecțiilor în același timp. O astfel de abordare se bazează pe compararea dispersiilor de selecție și se numește analiză dispersională.

Analiza dispersională este o altă metodă fundamentală a statisticii care, în plus față de mijloacele de calcul a “tendinței centrale” a rezultatelor experimentelor repetate, caracterizează mai ales variabilitatea acestora și factorii ce o determină.

Variabilitatea se poate datora existenței unor factori cu influențe sistematice, a unor factori aleatori de fluctuație mai pronunțată și, în final, factori locali, inevitabili, determinând o fluctuație mai mică, definită “ca fluctuație experimentală”. Analiza dispersională își propune separarea “variabilității totale” în: variabilitatea datorată factorilor sistematici, variabilitatea factorilor cu efecte aleatoare, plus o variabilitate “reziduală” (diferența până la variabilitatea totală), care reprezintă de fapt variabilitatea experimentală. Din aceste variabilități se evaluează dispersiile parțiale corespunzătoare diferiților factori, calculându-se semnificația rapoartelor lor prin aplicarea testului F.

Principial datele experimentale se grupează în funcție de diferite criterii și se urmăresc efectele asupra variabilității în funcție de aceste criterii, efecte care se cuantifică în raport cu variabilitatea reziduală.

Analiza dispersională este cunoscută în aplicațiile de biofarmacie și farmacocinetică sub denumirea de ANOVA (de la Analysis of Variance).

### **Analiza funcțională unifactorială**

Cea mai simplă analiză dispersională, numită analiză dispersională unidimensională sau unifactorială (numită în literatura engleză și “one-way ANOVA”) sau “experiment complet aleator”, “experiment cu grupuri paralele”, corespunde testului t de analiză a două eșantioane independente și compară două sau mai multe grupuri.

De exemplu, n pacienți sunt grupați în k scheme de tratament. Putem să comparăm efectele a două medicamente administrate la mai multe grupuri de voluntari, la care se poate adauga și un grup “placebo”. Voluntarii se distribuie aleator în toate grupurile. După măsurarea unui parametru dat, se testează ipoteza nulă că toate valorile parametrului testat sunt egale în populația corespunzătoare diferitelor tratamente testate, deci tratamentele sunt echivalente între ele.

În ipoteza că toate grupurile aparțin aceleiași populații, ideea testului este aceea că variabilitatea în interiorul grupurilor trebuie să fie de același ordin cu variabilitatea între mediile grupurilor.

În consecință, dispersia totală, evaluată ca suma a pătratelor diferențelor între valorile individuale și media întregii populații selectate  $SS_T$ , este separată într-o parte datorită variației între grupuri (within), sau variabilității “interioare” și o parte datorită variabilității “dintre” (between) grupuri:  $SS_T = SS_W + SS_B$ .

Dacă numărul de grupuri este  $k$  și numărul de subiecți în grupul  $i$  este  $n_i$  această egalitate poate fi explicată după cum urmează:

$$SS_T = \sum_i^n \sum_{j=1}^{n_i} (x_{ij} - \bar{X})^2 \quad (1)$$

unde  $\bar{X} = \text{media mare} = \frac{\sum_i \sum_j x_{ij}}{\sum_i n_i} = \frac{\sum_1^k n_i \bar{X}_i}{\sum_1^k n_i} = \frac{\sum x}{N}$  și  $\bar{X}_i$  este media grupului  $i$ .

Fixând grupul  $i$  putem scrie

$$\sum_j^{n_i} (x_{ij} - \bar{X})^2 = \sum_j^{n_i} [(x_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X})]^2 = \sum_j^{n_i} (x_{ij} - \bar{X}_i)^2 + \sum_j^{n_i} (\bar{X}_i - \bar{X})^2 + 2 \sum_j^{n_i} (x_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X})$$

Ultimul termen este egal cu

$$2(\bar{X}_i - \bar{X}) \left( \sum_j^{n_i} x_{ij} - n_i \bar{X}_i \right) = 2(\bar{X}_i - \bar{X})(n_i \bar{X}_i - n_i \bar{X}_i) = 0$$

și (1) devine

$$SS_T = \sum_i^k n_i (\bar{X}_i - \bar{X})^2 + \sum_i^k \sum_j^{n_i} (x_{ij} - \bar{X}_i)^2 = SS_B + SS_W \quad (2)$$

relație cunoscută ca *identitatea analizei dispersionale*.

Considerăm variabilele aleatoare:

$$s_T^2 = \frac{SS_T}{\sum_i n_i - 1} = \frac{SS_T}{N - 1}, \quad s_B^2 = \frac{SS_B}{k - 1} \quad \text{și} \quad s_W^2 = \frac{SS_W}{\sum_i n_i - k} = \frac{SS_W}{N - k}.$$

Ținând cont de regula generală demonstrată mai înainte că,  $(n-1)\delta_x^2 \approx \sigma^2 \chi^2(n-1)$  avem,

$$SS_W = \sum_{i,j} (x_{ij} - \bar{x}_i)^2 = \sum_i (n_i - 1) \frac{\sum_j (x_{ij} - \bar{x}_i)^2}{n_i - 1} = \sum_i (n_i - 1) s_i^2$$

$$\text{Deci, } s_W^2 = \frac{SS_W}{\sum_i (n_i - 1)} \approx \sigma^2 \chi^2(\sum_i (n_i - 1)) = \sigma^2 \chi^2(N - k)$$

Analog,  $s_B^2 \approx \sigma^2 \chi^2(k-1)$  și deci raportul  $F = \frac{\frac{SS_E}{k-1}}{\frac{SS_W}{N-k}} = \frac{S_B^2}{S_W^2}$  este distribuit

$F(k-1, N-k)$ .

Se observă că de fapt  $\frac{SS_B}{k-1} = \frac{\sum_i n_i (\bar{X}_i - \bar{X})^2}{k-1} = s_x^2$  reprezintă dispersia de selecție ponderată a mediilor de grup față de marea medie.

Abaterile mediilor grupurilor față de media generală depind atât de hazardul măsurătorilor cât și de factori ce țin de însăși natura grupurilor. Abaterile în interiorul grupurilor sunt independente de acești factori, deoarece fiecare valoare măsurată este raportată la însăși media grupului respectiv. Ele reprezintă fluctuații aleatoare.

Variabilitatea în interiorul grupurilor reprezintă diferența între variabilitatea totală și variabilitatea între grupuri.

Pentru simplificarea calculelor în aplicațiile practice s-au introdus notațiile  $\sum_i \sum_j^{n_i} x_{ij} = \sum x$  și  $\sum_i \sum_j^{n_i} x_{ij}^2 = \sum x^2$  și formulele precedente se aduc la forme echivalente

ce presupun un volum mai mic de calcule, după cum urmează:

$$\begin{aligned} SS_T &= \sum (x_{ij} - \bar{X})^2 = \sum x_{ij}^2 - 2\bar{X} \sum x_{ij} + N\bar{X}^2 = \sum x_{ij}^2 - 2 \frac{(\sum x_{ij})^2}{N} + N \frac{(\sum x_{ij})^2}{N^2} = \\ &= \sum x_{ij}^2 - \frac{(\sum x_{ij})^2}{N} = \sum x^2 - \frac{(\sum x)^2}{N} \end{aligned}$$

$$SS_B = \sum_1^k n_i (\bar{X}_i - \bar{X})^2 = \sum_1^k \frac{\left( \sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i} - \frac{\left( \sum_i \sum_j x_{ij} \right)^2}{\sum_1^k n_i} = \sum_1^k \frac{\left( \sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i} - \frac{(\sum x)^2}{\sum_1^k n_i}$$

### **Exemplu: Compararea parametrilor farmacocinetici ai unui medicament după administrarea pe mai multe căi**

Să considerăm constanta de eliminare (Tabelul 1) pentru un medicament administrat într-o singură doză oral și i.v., iar i.m. timp de trei zile la diferite grupuri. Este de așteptat ca eliminarea să fie independentă de calea de administrare.

Vom compara constantele de eliminare după administrarea orală și i.m.

Tabelul 1 Constanta de eliminare  $k_e * 10^4 \text{ min}^{-1}$

	Oral prima doza	i.m. o doza	i.m. a-5-doza	
	106	55	84	
	109	40	105	
	160	109	174	
$\sum x_{ij}$	375	204	363	$\sum x = 942$
$\sum x_{ij}^2$	48717	16506	48357	$\sum x^2 = 113580$
$n_i$	3	3	3	N=9
$n_i - 1$	2	2	2	$\sum (n_i - 1) = 6$
$\bar{x}_i$	125	68	121	

$$SS_T = \sum (x_{ij} - \bar{X})^2 = \sum x^2 - \frac{(\sum x)^2}{N} = 113580 - \frac{942^2}{9} = 14984$$

$$SS_B = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 = \sum_{i=1}^k \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i} - \frac{\left(\sum_i \sum_j x_{ij}\right)^2}{\sum_i n_i} = \frac{375^2}{3} + \frac{204^2}{3} + \frac{363^2}{3} - \frac{942^2}{9} =$$

$$46875 + 13872 + 43923 - 98596 = 6074$$

$$SS_W = SS_T - SS_B = 14984 - 6074 = 8910$$

$$s_W^2 = \frac{SS_W}{N - k} = \frac{8910}{6} = 1485, \quad s_B^2 = \frac{SS_B}{k - 1} = \frac{6074}{2} = 3037$$

$$F(k - 1, \sum n_i - k) = F(2, 6) = \frac{s_B^2}{s_W^2} = \frac{3037}{1485} = 2,07 \quad \text{valoare aflată în zona de acceptare}$$

$$(f_{2,6;97,5} = 7,26 \text{ și } f_{2,6;99} = 10,92)$$

Aplicăm în continuare același procedeu, introducând și administrarea i.v.

Tabelul 2. Constanta de eliminare  $k_e * 10^4 \text{ min}^{-1}$

		i.m. o doza	i.m. a-5-doza	i.v.	
	106	55	84	63	
	109	40	105	70	
	160	109	174		
$\sum x_{ij}$	375	204	363	133	$\sum x = 1075$
$\sum x_{ij}^2$	48717	16506	48357	8869	$\sum x^2 = 122449$
$n_i$	3	3	3	2	N=11
$n_i - 1$	2	2	2	1	$\sum (n_i - 1) = 7$
$\bar{x}_i$	125	68	121	66	

$$SS_T = \sum (x_{ij} - \bar{X})^2 = \sum x^2 - \frac{(\sum x)^2}{N} = 122449 - \frac{1075^2}{11} = 17392$$

$$SS_B = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 = \sum_{i=1}^k \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i} - \frac{\left(\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}\right)^2}{\sum_{i=1}^k n_i} = \frac{375^2}{3} + \frac{204^2}{3} + \frac{363^2}{3} + \frac{133^2}{2} - \frac{1075^2}{11} =$$

$$46875 + 13872 + 43923 + 8844 - 105056 = 8428$$

$$SS_W = SS_T - SS_B = 17392 - 8428 = 8964$$

$$s_W^2 = \frac{SS_W}{N - k} = \frac{8964}{7} = 1280, \quad s_B^2 = \frac{SS_B}{k - 1} = \frac{8428}{3} = 2809$$

$$F(k - 1, \sum n_i - k) = F(2, 6) = \frac{s_B^2}{s_W^2} = \frac{2809}{1280} = 2,19 \quad \text{valoare aflată în zona de acceptare}$$

$$(f_{2,6;97,5} = 7,26 \text{ și } f_{2,6;99} = 10,92)$$

### Condiții necesare pentru aplicarea analizei dispersionale.

#### a) Modelul variabilei aleatoare supusă măsurătorilor

Subliniem ceea ce am spus sau am presupus ca și condiții prealabile pentru a putea aplica testul prezentat:

1) Pentru ca sumele calculate să fie repartizate  $\chi^2$  este obligatoriu ca selecțiile să provină din populații repartizate normal. Dacă acest lucru nu se întâmplă, rămâne să fie aplicate teste neparametrice, după cum s-a prezentat anterior.

2) Analiza dispersională compară dispersia între grupuri cu dispersia totală din interiorul grupurilor pentru punerea în evidență a efectelor sistemice. Dispersiile în populațiile din care provin grupurile s-au presupus egale. Pentru verificarea acestei ipoteze se poate aplica testul F dispersiilor de selecție luate două câte două, sau se poate aplica testul Bartlett.

3) Aditivitatea efectelor de intra și intervariabilitate, fenomen care depinde de natura intimă a fenomenului măsurat.

Ca urmare rezultatul masuratorilor este o variabila aleatoare care se poate scrie:

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij} \text{ cu } E(\varepsilon_{ij}) = 0 \text{ și } D(\varepsilon_{ij}) = \sigma_e^2 \text{ unde } \mu \text{ este media generală a populației,}$$

$D(\alpha) = \sigma_\alpha^2$ ,  $\alpha_i$  este un factor fix ce variază de la un grup la altul și este egal cu diferența între  $\mu$  și media grupului  $\mu_i = \mu + \alpha_i$ , iar  $\varepsilon_{ij}$  este eroarea de măsurare.

Media a  $n$  determinări în interiorul unui grup este  $\bar{X}_i = \mu + \alpha_i + \frac{\sum \varepsilon_{ij}}{n} = \mu + \alpha_i + \bar{\varepsilon}_i$  deoarece  $\alpha_i$  și  $\mu$  sunt constante.

Media generală va fi:  $M = \mu + \frac{\sum \alpha_i}{k} + \frac{\sum \sum \varepsilon_{ij}}{kn} = \mu + \bar{\alpha} + \bar{\varepsilon}$  deoarece  $\alpha_i$  variază de la un grup la altul iar  $\mu$  este același pentru întreaga populație.

Mai departe  $x_{ij} - \bar{X}_i = (\mu + \alpha_i + \varepsilon_{ij}) - (\mu + \alpha_i + \bar{\varepsilon}_i) = \varepsilon_{ij} - \bar{\varepsilon}_i$  și dispersia în interiorul

grupului va fi:  $s_i^2 = \frac{\sum (x_{ij} - \bar{X}_i)^2}{n_i - 1} = \frac{\sum (\varepsilon_{ij} - \bar{\varepsilon}_i)^2}{n_i - 1}$  iar media ei este

$$E(s_i^2) = E\left(\frac{\sum (x_{ij} - \bar{X}_i)^2}{n_i - 1}\right) = E\left(\frac{\sum (\varepsilon_{ij} - \bar{\varepsilon}_i)^2}{n_i - 1}\right) = \sigma_e^2 \text{ (din faptul că așa cum s-a arătat mai$$

înainte  $E(s_x^2) = \sigma^2$ ,  $s_x^2$  fiind un estimator nedeplasat al dispersiei).

Sumând pentru toate grupurile se obține

$$E(s_w^2) = E\left(\frac{\sum (n_i - 1) S_i^2}{N - k}\right) = \frac{\sum (n_i - 1) E(S_i^2)}{N - k} = \frac{N - k}{N - k} \sigma^2 = \sigma_e^2$$

Deci  $s_w^2$  este un estimator nedeplasat al lui  $\sigma^2$  indiferent dacă ipoteza  $H_0$  este adevărată sau nu.

Dacă  $n_1 = n_2 = \dots = n_k = n$

$$s_B^2 = \frac{SS_B}{k - 1} = \frac{1}{k - 1} \sum n_i (\bar{X}_i - \bar{X})^2 = \frac{n}{k - 1} \left[ \sum (\alpha_i - \bar{\alpha})^2 + \sum (\bar{\varepsilon}_i - \bar{\varepsilon})^2 \right]$$

$$n \frac{\sum (\bar{\varepsilon}_i - \bar{\varepsilon})^2}{k - 1} = ns_{\bar{\varepsilon}}^2 \text{ și } E(ns_{\bar{\varepsilon}}^2) = n \frac{\sigma_e^2}{n} = \sigma_e^2$$

În ceea ce privește  $\alpha$  avem două cazuri în funcție de cum a fost ales:

a)  $\alpha$  este o variabilă aleatoare  $N(0, \sigma_\alpha^2)$ , atunci  $s_B^2 = ns_\alpha^2 + s_e^2$  și  $E(s_B^2) = n\sigma_\alpha^2 + \sigma_e^2$



b)  $\alpha_i$  factori ficși ,  $\alpha_i = \mu - \mu_i \Rightarrow \sum \alpha_i = 0$  și  $\bar{\alpha} = 0$ , atunci

$$s_B^2 = n \frac{\sum (\alpha_i - 0)^2}{k-1} + s_e^2 \text{ și } E(s_B^2) = n \frac{\sum \alpha_i^2}{k-1} + \sigma_e^2$$

Deci ne referim la grupe oarecare, alese întâmplător din toate selecțiile posibile din întreaga populație,  $\alpha$  este o variabilă aleatoare care aparține  $N(0, \sigma_\alpha^2)$ , dar nu mai avem  $\sum \alpha_i = 0$  pentru o grupă de selecție oarecare. Este cazul când concluziile ce ne interesează se referă, în principal, la întreaga producție și nu la șarjele cercetate.

Calculare similare arată că, atunci când  $n_1 = n_2 = \dots = n_k = n$ ,  $s_B^2 = n s_\alpha^2 + s_e^2$  și  $E(s_B^2) = \sigma_e^2 + n \sigma_\alpha^2$ .

Ca urmare,  $s_B^2 = \frac{SS_B}{k-1}$  este o estimatie nedeplasată a lui  $\sigma_e^2$  numai atunci când ipoteza  $H_0$  este adevărată.

### **b) Clasificarea ierarhică în trepte. Scindarea dispersiei totale în dispersii parțiale ale treptelor.**

Dacă vom considera eroarea unei metode de analiză, este de așteptat ca aceasta să fie egală cu suma erorilor parțiale ale fazelor analizei.

Scindarea erorii totale în erorile parțiale ale fazelor oferă posibilitatea punerii în evidență a fazelor ce determină diminuarea reproductibilității și a locului unde trebuie intervenit pentru îmbunătățirea metodei.

#### **Analiza dispersională multifactorială**

În analiza anterioară criteriul de împărțire în grupuri a fost unul singur – calea de administrare sau perioada de administrare. În continuare ne vom ocupa de cazul când vrem să determinăm ponderea a doi factori în același timp, folosind datele obținute într-un singur experiment.

Pentru simplificarea scrierii prezentăm pentru început cazul când testăm k medicamente continuând aceeași substanță activă, administrate în n perioade diferite.

Se consideră că efectele determinate de cei doi factori analizați sunt variabile aleatoare independente. Având doi factori, se testează două ipoteze de nul. Dacă vom așeza datele într-o matrice, liniile și coloanele corespunzând respectiv pentru valori date ale primului și ale celui de al doilea factor, o primă ipoteză se referă la egalitatea mediilor

liniilor, iar a doua la egalitatea mediilor coloanelor. Ipoteza alternativă presupune existența unor diferențe între linii sau respectiv între coloane.

		Medicament				
		1	2	...	k	
Perioada Tratament	1	$X_{11}$	$X_{12}$		$X_{1k}$	$\bar{x}_{1.}$
	2	$X_{21}$	$X_{22}$		$X_{2k}$	$\bar{x}_{2.}$
	.					
	.					
	n	$X_{n1}$	$X_{n2}$		$X_{nk}$	$\bar{x}_{n.}$
		$\bar{x}_{.1}$	$\bar{x}_{.2}$		$\bar{x}_{.k}$	

unde  $\bar{x}_{.j}$  reprezintă media valorilor din coloana i, iar  $\bar{x}_{i.}$  reprezintă media valorilor din linia j.

### Modelul variabilei aleatoare.

Valorile experimentale le considerăm ca rezultanta unor efecte aditive corespunzător liniilor, coloanelor și erorilor întâmplătoare:

$$x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

unde  $\alpha_i$  este partea lui  $x_{ij}$  datorată liniei (schemei de administrare),  $\beta_j$  reprezintă contribuția coloanei (forma medicamentoasă), iar  $\varepsilon_{ij}$  este eroarea experimentală.

### Scindarea sumei pătratelor abaterilor.

Suma pătratelor abaterilor valorilor individuale față de media generala M este:

$$SS_T = \sum \sum (x_{ij} - M)^2, \text{ expresie ce se poate scrie și în forma:}$$

$$SS_T = \sum \sum (x_{ij} - M)^2 = \sum \sum [(\bar{x}_{L.} - M) + (\bar{x}_{.C} - M) + (x - \bar{x}_{L.} - \bar{x}_{.C} + M)]^2 = \sum \sum (\bar{x}_{L.} - M)^2 + \sum \sum (\bar{x}_{.C} - M)^2 + \sum \sum (x - \bar{x}_{L.} - \bar{x}_{.C} + M)^2 = SS_L + SS_C + SS_R$$

Produsele mixte nu s-au mai trecut deoarece sunt nule.

Am scindat așadar variabilitatea totală într-o componentă dată de linii, o componentă dată de coloane și o componentă reziduală.

Corespunzător rezultatelor prezentate la analiza unifactorială, aici se poate arăta că:

$$E(SS_T) = (n_C n_L - 1)\sigma^2 + n_L \sum \alpha_i^2 + n_C \sum \beta_j^2$$

$$E(SS_L) = (n_C - 1)\sigma^2 + n_L \sum \alpha_i^2$$

$$E(SS_C) = (n_L - 1)\sigma^2 + n_c \sum \beta_j^2$$

$$E(SS_R) = (n_C - 1)(n_L - 1)\sigma^2$$

În cazul analizei bifactoriale apar două ipoteze de nul pe care vrem să le testăm:

$H_0^{(1)}$  : toate formele de tratament (liniile) sunt echivalente, adică  $\alpha_i = 0$  ;

$H_0^{(2)}$  : toate medicamentele (coloanele) sunt echivalente, adică  $\beta_j = 0$  .

Indiferent dacă ipotezele sunt adevărate sau nu,  $E(s_R^2) = \sigma^2$  .

Dacă cele două ipoteze sunt adevărate:  $s_e^2 = s_R^2$ ,  $s_C^2$  și  $s_L^2$  sunt estimațiile nedepășate ale lui  $\sigma^2$  .

Pentru a testa ipoteza  $H^{(1)}$  este natural să considerăm raportul  $\frac{s_L^2}{s_e^2}$ , iar pentru ipoteza  $H^{(2)}$  raportul  $\frac{s_C^2}{s_e^2}$  .

În cazul în care pentru fiecare combinație de cauze se fac mai multe determinări folosind mai mulți voluntari, variabilitatea se poate scindea în patru componente:

$$SS_T = SS_L + SS_C + SS_W + SS_R$$

**Exemplul 2: Variația unui parametru farmacocinetic atât în funcție de calea de administrare cât și în funcție de numărul de administrare**

Constanta de eliminare  $k_e * 10^4 \text{ min}^{-1}$

	Oral prima doza	i.m. o doza	i.m. a-5-doza	i.v.	
	106	55	84	63	
	109	40	105	70	
	160	109	174		
$\sum x_i$	375	204	363	133	$\sum x = 1075$
$\sum x_i^2$	48717	16506	48357	8869	$\sum x^2 = 122449$
$n_i$	3	3	3	2	N=11
$n_i - 1$	2	2	2	1	$\sum (n_i - 1) = 7$
$\bar{x}_i$	125	68	121	66	

$$SS_T = \sum x^2 - \frac{(\sum x)^2}{N} = 17392$$

Mai departe considerăm două grupuri: voluntarii cărora medicamentul li s-a administrat o dată și voluntarii care au primit 5 doze.

$$SS_B(nr.doze) = \frac{363^2}{3} + \frac{(375 + 204 + 133)^2}{8} - 105056 = 2235$$

$$SS_W = SS_T - SS(nr.doze) = 17392 - 2235 = 15157$$

$$\frac{SS_W}{N-k} = \frac{15157}{9} = 1684, \quad \frac{SS_B}{k-1} = \frac{15157}{1} = 15157$$

$$F(k-1, \sum n_i - k) = F(1,9) = \frac{15157}{1684} = 9 \text{ care aparține zonei de acceptare pentru riscul}$$

$\alpha < 0,01$  ( $f_{1,9,99} = 10,56$ ) dar nu aparține zonei de acceptare pentru  $\alpha < 0,05$  ( $f_{1,9,95} = 5,12$ ).

$$SS_T = \sum (x_i - \bar{x})^2 = \sum x^2 - \frac{(\sum x)^2}{N} = 122449 - \frac{1075^2}{11} = 17392$$

Un alt factor care poate influența concentrația în sânge este calea de administrare.

$$SS_B(cale) = \sum_{i=1}^k n_i (\bar{x}_i - \bar{X})^2 = \sum_{i=1}^k \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i} - \frac{\left(\sum_i \sum_j x_{ij}\right)^2}{\sum_i n_i} = \frac{375^2}{3} + \frac{204^2}{3} + \frac{133^2}{2} - \frac{1075^2}{11} =$$

$$46875 + 13872 + 43923 + 8844 - 105056 = 8428$$

$$SS_B(nr.doze) = \frac{363^2}{3} + \frac{(375 + 204 + 133)^2}{8} - 105056 = 2235$$

$$SS_W = SS_T - SS_B(cale) - SS_B(nr.doze) = 17392 - 8428 - 2235 = 6792$$

În acest fel am obținut următoarele rezultate:

Sursa de variație	GL	SS	Raportul F
Între căile de administrare	3	8428	f(3,6)=1,24
Între numărul de doze	1	2235	f(1,6)=0,32
Variația intragrupuri	6	6792	
Variația totală	10	17392	

și comparând cu valorile din tabelele de repartiție Fisher se constată că poate fi acceptată ipoteza că toate grupurile cercetate aparțin aceleiași populații statistice. Deci, nici calea de administrare și nici numărul de doze nu influențează rata eliminării.

### Exemplul 3: Compararea biodisponibilității unei substanțe active administrate în preparate diferite și în perioade diferite de tratament.

Trei medicamente au fost administrate la nouă subiecți într-un studiu de biodisponibilitate obținându-se ariile de sub curbă prezentate în tabelul de mai jos:

Perioada	1	2	3	Suma
Pacient 1	B=107	C=102	A=99	308
Pacient 2	A=100	C=106	B=89	295
Pacient 3	B=98	A=90	C=128	316
Pacient 4	C=71	B=54	A=63	188
Pacient 5	A=92	B=111	C=107	310
Pacient 6	C=113	A=115	B=91	319
Pacient 7	B=169	A=107	C=195	551
Pacient 8	C=88	B=95	A=77	260
Pacient 9	A=122	C=168	B=155	445
Suma pe perioadă	I: 160	II: 1028	III: 1004	$\sum x = 2992$
Suma pe medicament	A: 945	B: 969	C: 1078	$\sum x^2 = 364720$
Media pe medicament	105	107,7	119,8	

Separăm acum varianta în patru componente: pacienți, perioadă, medicament și eroarea în interiorul tuturor grupurilor.

$$SS_T = \sum x^2 - \frac{(\sum x)^2}{N} = 33162,1$$

$$SS_B(\text{pacienti}) = \sum_{i=1}^9 \frac{(\sum \text{linie})^2}{3} - \frac{(\sum x)^2}{N} = 29834,1$$

$$SS_B(\text{medicament}) = \sum_{i=1}^3 \frac{(\sum \text{medicament})^2}{9} - \frac{(\sum x)^2}{N} = 1116,5$$

$$SS_B(\text{perioada}) = \frac{\sum I^2 + \sum II^2 + \sum III^2}{9} - \frac{(\sum x)^2}{N} = 264,3$$

$$SS_W = SS_T - SS_B(\text{pacienti}) - SS_B(\text{medicament}) - SS_B(\text{perioada}) = 1947,2$$

Pe aceasta cale obținem:

Sursa de variație	DF	Suma pătratelor	Media	Raportul F
Între pacienți	8	29834,1	3729,3	
Între medicamente	2	116,1	58,05	3,15
Între perioade	2	264,3	132,15	0,75
Variația intragrupuri	14	1947,2	139,09	
Total	26	33162,1		

Deoarece  $f_{2,14;0,99} = 3,74$  nu există diferențe semnificative între cele două medicamente după administrare unică sau administrări multiple.