

EXERCICIS D'INTRODUCCIÓ A LA QUIMIOMETRIA AMB FULL DE CÀLCUL

Josep Lluís Beltran
Elisabeth Bosch
Ramón Companyó
Anna Izquierdo-Ridorsa
Martí Rosés
Romà Tauler

Departament de Química Analítica



EXERCICIS D'INTRODUCCIÓ A LA QUIMIOMETRIA AMB FULL DE CÀLCUL

Josep Lluís Beltran
Elisabeth Bosch
Ramón Companyó
Anna Izquierdo-Ridorsa
Martí Rosés
Romà Tauler

Departament de Química Analítica

Publicacions i Edicions



UNIVERSITAT DE BARCELONA



INTRODUCCIÓ

El *Diccionari de Química Analítica* publicat recentment pel TERMCAT defineix la Quimiometria com la branca de la Química que s'ocupa de l'aplicació de tècniques matemàtiques als procediments del mesurament per tal d'obtenir dels resultats la màxima quantitat d'informació química i física.

Ningú no discuteix avui que la Quimiometria constitueix una part imprescindible dels programes universitaris de Química Analítica i, de fet, en tots els manuals avançats s'inclou un o alguns capítols dedicats al tema. L'assignatura Ampliació de Química Analítica, que és obligatòria per a tots els estudiants de Química de la UB, inclou en el seu programa alguns temes d'introducció a la Quimiometria. Atès que aquests temes impliquen un contingut important d'estadística i càlcul numèric, cal ampliar els conceptes introduïts a les classes de teoria amb seminaris assistits per ordinador que han de consistir en la realització d'exercicis pràctics relacionats amb els punts suara esmentats.

Les qüestions i problemes recollits en aquest text responen a la necessitat de posar a l'abast dels estudiants un recull d'exercicis que els ajudin a comprendre i aprofundir alguns dels conceptes bàsics de la Quimiometria. Els temes seleccionats responen a la part de la disciplina que tracta sobre l'aplicació a la Química Analítica de diferents mètodes estadístics i d'optimització i s'agrupen en quatre blocs que són: 1) Hipòtesis i Tests Estadístics; 2) Anàlisi de la Variància i Disseny d'Experiments; 3) Optimització d'Experiments, i 4) Calibratge Analític Univariant. La resolució dels problemes proposats es realitza mitjançant l'aplicació de fulls de càlcul, de les eines estadístiques que aquests tenen incorporades, i, si s'escau, de l'aplicació de programes d'ordinador especialment dissenyats per a la realització dels exercicis preparats. En el text es presenten alguns exercicis resolts i uns altres per resoldre. El full de càlcul escollit per a la resolució pràctica dels exercicis proposats en aquest manual ha estat l'Excel de Microsoft. És per això que en aquest text guia es donen instruccions d'ús d'algunes de les eines que proporciona l'Excel, tal com les ofereix el mateix programa.

La innovació que va representar el fet de posar en marxa aquests seminaris, les dificultats que comporta la seva organització per a un nombre gran d'estudiants en un període relativament curt de temps, així com la necessitat de disposar, per part d'alumnes i professors, d'un material escrit que possibilités l'optimització de l'aprenentatge, ens va orientar cap a la preparació d'aquest text guia.

Volem posar de manifest i agrair l'esforç realitzat per les persones que han contribuït a posar a punt els exercicis i problemes d'aquesta assignatura, els professors implicats, i especialment Jordi Iglesias Yagüe, que ha tingut cura de la diagramació, maquetació i edició tècnica d'aquest text. A més volem fer constar el nostre agraïment al Gabinet d'Avaluació i Innovació Universitària (GAIU) per la subvenció rebuda per a la redacció d'una primera versió d'aquest llibre.

ÍNDEX

CAPÍTOL 1

1. Tests estadístics	3
1.1. Ús del full de càlcul Excel per a l'aplicació dels tests estadístics	6
1.1.1. Exemple d'utilització del full de càlcul Excel per a la comparació dels resultats de dues sèries de mesures	7
1.1.2. Exemple d'utilització del full de càlcul Excel per verificar si un conjunt de dades segueix o no una distribució normal	10
1.2. Exercicis d'aplicació	12
Bibliografia	18

CAPÍTOL 2

2. Anàlisi de la variància (ANOVA)	19
2.1. ANOVA d'un factor	20
2.1.1. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA d'un factor	21
2.2. ANOVA de dos factors	22
2.2.1. ANOVA de dos factors sense replicats	23
2.2.2. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA de dos factors sense replicats	24
2.2.3. ANOVA de dos factors amb replicats	25
2.2.4. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA de dos factors amb replicats	27
2.3. Exercicis d'aplicació	29
Bibliografia	33

CAPÍTOL 3

3. Optimització d'experiments analítics	35
3.1. Algorismes Simplex	35
3.1.1. Simplex Simple o de Grandària Fixada	35
3.1.2. Simplex de Grandària Variable	36
3.2. Funcionament del programa	38
3.3. Finestres i comandaments	39
3.3.1. Finestra d'Inici	39
3.3.2. Finestra d'Experiments amb Resposta Quadràtica	39
3.3.3. Finestra del Simplex	40
3.3.4. Comandaments del Menú Fitxer	41
3.3.5. Comandaments del Menú Zoom	41
3.3.6. Comandaments del Menú Configuració	41
3.3.7. Comandaments del Menú Ajuda	42
3.3.8. Finestra d'Entrar Vèrtexs	42

3.3.9. Finestra d'Escollir Vèrtexs	42
3.3.10. Finestra d'Avis	43
3.4. Experiments	43
3.4.1. Experiments Externs	44
3.4.2. Experiments amb Resposta Quadràtica	44
3.4.3. Separació d'Isòmers de l'Octà per Cromatografia de Gasos	44
3.5. Exercicis d'aplicació	45
Bibliografia	47
Annex	48
Elaboració de gràfics de superfície i de contorn	48
Utilització de l'Eina Solver	50
 CAPÍTOL 4	
4. Calibratge lineal	51
4.1. Regressió lineal per mínims quadrats	51
4.2. Límits de detecció i de quantificació	53
4.3. Correlació entre variables	54
4.4. Utilització del full de càlcul Excel	55
4.4.1. Exemple d'utilització	55
4.5. Exercicis d'aplicació	60
Bibliografia	62
 ANNEX	 63

CAPÍTOL 2

2. ANÀLISI DE LA VARIÀNCIA (ANOVA)

En el capítol anterior s'han estudiat i aplicat diferents tests estadístics per comparar dos resultats (o parelles de resultats) obtinguts en condicions diferents. Però en moltes situacions cal poder comparar més de dos resultats. La tècnica estadística més emprada en aquests casos és **l'anàlisi de la variància**. Alguns exemples als quals s'aplicaria l'anàlisi de la variància serien:

- a) estudi de l'efecte del temps d'emmagatzematge d'una mostra de sèrum sobre el contingut proteic d'aquesta mostra;
- b) comparació dels resultats obtinguts per diferents laboratoris en l'anàlisi d'una mateixa mostra fent servir el mateix mètode analític;
- c) estudi del contingut en traces de Pb d'una aigua potable emmagatzemada en un dipòsit, a partir de mostres preses aleatòriament.

En aquests exemples hi ha dues possibles fonts de variació dels resultats obtinguts. En tots els casos una d'elles és l'error aleatori associat al propi procediment analític. Aquest error aleatori es pot quantificar a partir de mesures replicades d'una mateixa mostra. L'altra possible font de variació és el temps d'emmagatzematge en l'exemple (a), els diferents laboratoris en l'exemple (b) i la presa de mostra en l'exemple (c). Aquesta segona font de variació correspon al factor l'efecte del qual s'està avaluant. Molt sovint a aquest factor se l'anomena **factor controlat**, per diferenciar-lo de l'error aleatori inherent a totes les mesures experimentals. No obstant això, no és d'ús general, ja que en l'exemple (c) el factor que s'estudia també és aleatori. Tot i això, aquesta serà la nomenclatura que es farà servir en aquest capítol.

L'anàlisi de la variància és una eina estadística que permet separar i estimar les diferents causes de variació dels resultats: permet separar i quantificar la variació en les dades deguda a canvis en el factor controlat, de la variació deguda a l'error aleatori. Un cop es coneixen aquestes dues magnituds, es poden comparar estadísticament entre si. Si hi ha diferències significatives entre les dues fonts de variació, es podrà concloure que canvis en el factor controlat impliquen canvis en el valor de la propietat que es mesura. En l'exemple (a) això implicaria que les condicions d'emmagatzematge de la mostra no són les adients perquè el contingut proteic varia al llarg del temps; en l'exemple (b) això ens indicaria que un o més dels laboratoris probablement està cometent algun error sistemàtic en l'anàlisi; i en l'exemple (c) la presència de diferències significatives ens indicaria que el contingut del dipòsit no és homogeni i que, com a conseqüència, l'error de presa de mostra és significativament superior al de l'anàlisi. Si no hi ha diferències significatives entre les dues fonts de variació, això voldrà dir que el factor controlat no té cap efecte sobre la propietat que es mesura.

En els exemples anteriors s'estudiava l'efecte d'un únic factor controlat. L'ANOVA que es porta a terme en aquests casos s'anomena **ANOVA d'un factor**. En aquests tipus d'ANOVA es requereix la realització de mesures replicades de cada anàlisi per poder quantificar l'error aleatori.

En altres casos pot ser necessari l'estudi de més d'un factor controlat sobre les dades experimentals. En aquests casos es porta a terme l'anomenat **ANOVA de dos factors**. Si es volguessin estudiar simultàniament n factors controlats, s'hauria de fer un **ANOVA de n factors**.

Exemples en els quals es requereix un ANOVA de dos factors són els següents:

a) exercici simultani d'intercalibració de laboratoris i de mètodes d'anàlisi: diferents laboratoris analitzen una mateixa mostra amb mètodes diferents;

b) estudi del temps i de la temperatura d'emmagatzematge d'una mostra de sèrum de la qual es vol determinar el contingut en traces de Cu(II).

2.1. ANOVA d'un factor

Suposem que es vol portar a terme una comparació dels resultats obtinguts per 5 laboratoris diferents en l'anàlisi d'una mateixa mostra fent servir el mateix mètode analític. Cada laboratori fa 3 anàlisis replicades de la mostra. Normalment els resultats experimentals es representen en una taula, tal i com s'indica a continuació:

	Rèpl. 1	Rèpl. 2	Rèpl. 3	M_i
Lab. 1	10.0	11.5	10.3	10.6
Lab. 2	12.0	10.5	9.3	10.6
Lab. 3	10.1	13.4	12.5	12.0
Lab. 4	9.6	7.0	8.0	8.2
Lab. 5	14.1	14.2	12.5	13.6
M_{global}				11.0

A la taula s'hi ha afegit la mitjana per a cada laboratori (M_i) i la mitjana global (M_{global}). A partir de tots aquests valors, ja es pot calcular l'error aleatori i l'error associat al factor controlat. Les expressions que es fan servir, pel cas general d'analitzar r valors diferents del factor controlat (en l'exemple són laboratoris) amb n replicats per a cadascun d'ells, són les següents:

Variància associada a l'error aleatori:

$$(s_{\text{aleat}})^2 = \sum (x_{ik} - M_i)^2 / r(n-1) = SS_{\text{aleat}} / r(n-1)$$

on:

$i = 1, \dots, r$ es refereix als laboratoris.

$k = 1, \dots, n$ es refereix als replicats per laboratori.

x_{ik} és una mesura individual que correspon al laboratori i i al replicat k .

Variància associada al factor controlat:

$$(s_{lab})^2 = n \sum (M_i - M_{global})^2 / (r-1) = SS_{lab} / (r-1)$$

Variància global:

$$(s_{global})^2 = \sum \sum (x_{ik} - M_{global})^2 / (rn-1) = SS_{global} / (rn-1)$$

Un cop s'han obtingut els valors de les variàncies associades a l'error aleatori i al factor controlat, es realitza un test F per tal de veure si la variància deguda al factor controlat és significativament superior a la deguda a l'error aleatori o no. Si es compleix la hipòtesi nul·la es podrà concloure que els resultats experimentals no estan afectats pel valor del factor controlat. Si, pel contrari, es compleix la hipòtesi alternativa, això indicarà que el resultat experimental depèn del factor controlat.

Tota la informació obtinguda en l'ANOVA es resumeix en una taula, tal i com s'indica a continuació:

Font de variació	Suma de quadrats	Graus de llibertat	Variància	Valor de F
Factor controlat	SS_{lab}	$(r-1)$	$SS_{lab} / (r-1)$	$F = s_{lab}^2 / s_{aleat}^2$
Error aleatori	SS_{aleat}	$r(n-1)$	$SS_{aleat} / r(n-1)$	
Global	SS_{global}	$rn-1$	$SS_{global} / (rn-1)$	

2.1.1. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA d'un factor

S'ha d'entrar en el menú **Herramientas** i en el submenú **Análisis de datos**. Un cop en aquest submenú s'ha de triar l'opció **Análisis de varianza de un factor**. Apareix una finestra que està dividida en dues parts: una amb el títol **ENTRADA** i l'altra amb el títol **OPCIONES DE SALIDA**. En la primera part es demana tota la informació necessària per fer córrer el programa:

- **Rango de entrada:** correspon a les cel·les on hi ha les dades experimentals; també s'hi poden incloure les cel·les que contenen les etiquetes identificatives.

- **Agrupado por:** si cada valor del factor controlat correspon a una fila, aleshores s'ha de marcar on posa **Filas**; si cada valor del factor controlat correspon a una columna, s'ha de marcar on posa **Columnas**.

- **Títulos en la primera columna** (o fila, depenent de com està agrupat): si hem posat etiquetes identificatives ho hem de marcar, i així el programa sabrà que aquella columna o fila no conté resultats.

- **Alfa:** hem de donar el nivell de significació que volem per realitzar el test F.

En la segona part (**OPCIONES DE SALIDA**) es demana on es vol que el programa tregui els resultats. Els pot treure (a) en el mateix full on hem introduït les dades experimentals, (b) en un full nou i (c) en un llibre nou. El més simple és fer servir l'opció (a). S'ha de marcar aquesta opció i s'ha d'indicar en el requadre adjunt quina és la cel·la superior esquerra a partir de la qual es vol que apareguin els resultats.

A continuació es mostren els resultats obtinguts amb Excel amb les dades de l'exemple anterior:

	Rèpl. 1	Rèpl. 2	Rèpl. 3
Lab. 1	10	11.5	10.3
Lab. 2	12	10.5	9.3
Lab. 3	10.1	13.4	12.5
Lab. 4	9.6	7	8
Lab. 5	14.1	14.2	12.5

Análisis de varianza de un factor (resumen)

Grupos	Cuenta	Suma	Promedio	Varianza
Lab. 1	3	31.8	10.6	0.63
Lab. 2	3	31.8	10.6	1.83
Lab. 3	3	36.0	12.0	2.91
Lab. 4	3	24.6	8.2	1.72
Lab. 5	3	40.8	13.6	0.91

Análisis de varianza

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico de F
Entre grupos	47.76	4	11.94	7.4625	0.00472195	3.47804985
Dentro de los grupos	16	10	1.6			
Total	63.76	14				

A partir dels resultats obtinguts ($F_{\text{calculada}} > F_{\text{tabulada}}$ pel nivell de significació escollit, que en aquest cas ha estat de 0.05) es pot concloure que hi ha diferències significatives entre laboratoris. Això pot implicar que un o més dels laboratoris analitzats està cometent un error sistemàtic. Aquesta comparació també es pot portar a terme entre el nivell de significació experimental per a la F trobada (indicat per “Probabilidad” a la taula de resultats) i el nivell de significació escollit (0.05 en aquest cas). En aquest exemple, el nivell de significació experimental és inferior a l'escollit, la qual cosa indica que hi ha diferències significatives entre laboratoris.

2.2. ANOVA de dos factors

En aquest cas s'estudia l'efecte de dos factors controlats diferents sobre el resultat. Un exemple d'aplicació d'un ANOVA de dos factors seria en un exercici simultani d'intercalibració de laboratoris i mètodes d'anàlisi, on una mateixa mostra és analitzada per diferents laboratoris amb mètodes diferents.

En una ANOVA de dos factors es poden donar dues situacions diferents:

a) no es tenen replicats. En l'exemple anterior això implica que cada laboratori analitza només un cop la mostra per cada mètode;

b) es tenen replicats. Cada laboratori analitza n vegades la mostra per cada mètode.

2.2.1. ANOVA de dos factors sense replicats

En aquesta situació la variació global associada als resultats es pot separar en tres aportacions diferents, dues degudes als factors controlats i la tercera que conté l'error residual. Aquest error residual engloba l'error aleatori i l'aportació deguda a la possible interacció entre els dos factors controlats. Atès que no s'han fet replicats, l'error aleatori no es pot calcular directament, i es fa servir l'error residual com a mesura d'aquest. Òbviament aquesta suposició només serà certa en cas d'absència d'interacció entre factors.

A la taula següent es donen, a títol d'exemple, els resultats experimentals obtinguts per 4 laboratoris diferents en l'anàlisi d'una mostra amb tres mètodes diferents:

	Mètode 1	Mètode 2	Mètode 3	M_i
Laboratori 1	84	79	83	82
Laboratori 2	80	77	78	78
Laboratori 3	83	80	80	81
Laboratori 4	79	79	78	79
M_j	82	79	80	
				$M_{global} = 80$

M_i és la mitjana per a cada laboratori i, per $i = 1, \dots, r$

M_j és la mitjana per a cada mètode j, per $j = 1, \dots, c$

A la taula s'hi han afegit les mitjanes per cada laboratori, per cada mètode i la mitjana global. A partir de tots aquests valors ja es pot calcular la contribució de cada factor controlat a la variació global i l'error residual.

Variància entre laboratoris:

$$(s_{lab})^2 = c \sum (M_i - M_{global})^2 / (r-1) = SS_{lab} / (r-1)$$

Variància entre mètodes:

$$(s_{mèt})^2 = r \sum (M_j - M_{global})^2 / (c-1) = SS_{mèt} / (c-1)$$

Variància global:

$$(s_{global})^2 = \sum \sum (x_{ij} - M_{global})^2 / (rc-1) = SS_{global} / (rc-1)$$

Variància associada a l'error residual:

Es determina per diferència:

$$SS_{\text{residual}} = SS_{\text{global}} - SS_{\text{lab}} - SS_{\text{mèt}}$$

$$\text{graus de llibertat} = (rc-1) - ((c-1) + (r-1)) = rc - c - r + 1 = (r-1)(c-1)$$

Un cop obtinguts aquests valors de les variàncies, es realitzen dos tests F per determinar si hi ha diferències significatives o no entre la variància deguda a cadascun dels factors controlats i la de l'error residual.

Tota la informació obtinguda en l'ANOVA es resumeix en una taula, tal i com s'indica a continuació:

Font de variació	Suma de quadrats	Graus de llibertat	Variància	Valor de F
Entre laboratoris	SS_{lab}	$(r-1)$	$SS_{\text{lab}} / (r-1)$	$S_{\text{lab}}^2 / S_{\text{residual}}^2$
Entre mètodes	$SS_{\text{mèt}}$	$(c-1)$	$SS_{\text{mèt}} / (c-1)$	$S_{\text{mèt}}^2 / S_{\text{residual}}^2$
Error residual	SS_{residual}	$(r-1)(c-1)$	$SS_{\text{residual}} / (r-1)(c-1)$	
Global	SS_{global}	$(rc-1)$	$SS_{\text{global}} / (rc-1)$	

2.2.2. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA de dos factors sense replicats

S'ha d'entrar en el menú **Herramientas** i en el submenú **Análisis de datos**. Un cop en aquest submenú s'ha de triar l'opció **Análisis de varianza de dos factores con una sola muestra por grupo**. Apareix una finestra que està dividida en dues parts, amb els títols **ENTRADA** i **OPCIONES DE SALIDA**, respectivament. En la primera part es demana tota la informació necessària per a fer córrer el programa:

- **Rango de entrada:** correspon a les cel·les on hi ha les dades experimentals; també s'hi poden incloure les cel·les que contenen les etiquetes identificatives.

- **Títulos:** si hem posat etiquetes identificatives ho hem de marcar, i així el programa sabrà que la primera columna i la primera fila no contenen resultats.

- **Alfa:** hem de donar el nivell de significació que volem per realitzar el test F.

En la segona part (**OPCIONES DE SALIDA**) es demana on es vol que el programa tregui els resultats. Els pot treure (a) en el mateix full on hem introduït les dades experimentals, (b) en un full nou i (c) en un llibre nou. El més simple és fer servir l'opció (a). S'ha de marcar aquesta opció i s'ha d'indicar en el requadre adjunt quina és la cel·la superior esquerra a partir de la qual es vol que apareguin els resultats.

A continuació es mostren els resultats obtinguts amb Excel amb les dades de l'exemple anterior:

	Mètode 1	Mètode 2	Mètode 3
Laboratori 1	84	79	83
Laboratori 2	80	77	78
Laboratori 3	83	80	80
Laboratori 4	79	79	78

Análisis de varianza de dos factores con una sola muestra por grupo

RESUMEN	Cuenta	Suma	Promedio	Varianza
Laboratori 1	3	246	82	7
Laboratori 2	3	235	78.3333333	2.33333333
Laboratori 3	3	243	81	3
Laboratori 4	3	236	78.6666667	0.33333333
Mètode 1	4	326	81.5	5.66666667
Mètode 2	4	315	78.75	1.58333333
Mètode 3	4	319	79.75	5.58333333

Análisis de varianza

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico de F
Filas	28.6666667	3	9.55555556	5.83050847	0.03275623	4.75705519
Columnas	15.5	2	7.75	4.72881356	0.05848241	5.14324938
Error	9.83333333	6	1.63888889			
Total	54	11				

En l'exemple les files corresponen a laboratoris i les columnes a mètodes. Si es compara el valor de F tabulat per al nivell de significació escollit i el calculat a partir de l'ANOVA, es pot concloure que hi ha diferències significatives entre laboratoris i que no hi ha diferències significatives entre mètodes. Això implica que algun (o alguns) dels laboratoris està cometent un error sistemàtic probablement en alguna etapa del tractament de la mostra, comuna a tots els mètodes. Tal i com ja s'ha indicat prèviament, aquesta comparació també es pot portar a terme entre el nivell de significació experimental per a la F trobada i el nivell de significació escollit.

2.2.3. ANOVA de dos factors amb replicats

En aquest cas la variació global associada als resultats experimentals es pot separar en quatre aportacions diferents, dues degudes als factors controlats, una deguda a l'error aleatori i una deguda a la interacció entre factors. Ara l'error aleatori ja es pot calcular directament a partir dels replicats. En canvi, la variació deguda a la interacció s'ha de determinar per diferència.

A la taula següent es donen els resultats experimentals obtinguts en l'exemple següent: En un exercici d'intercalibració de laboratoris s'envien tres mostres diferents de carbó a tres laboratoris perquè en determinin el contingut d'arsènic. Cada laboratori ha de fer dos replicats de cada mostra. Si els resultats són els que figuren a la taula, decideix si hi ha diferències entre els resultats dels diferents laboratoris.

	Mostra 1		Mostra 2		Mostra 3		M _i
Laboratori 1	5.1	5.1	5.8	5.4	6.5	6.1	5.66
Laboratori 2	5.3	5.4	5.4	5.9	6.6	6.7	5.88
Laboratori 3	5.3	5.1	5.2	5.5	6.5	6.4	5.66
M _j	5.22		5.53		6.46		
					M _{global} = 5.74		

M_i és la mitjana per a cada laboratori, per $i = 1, \dots, r$

M_j és la mitjana per a cada mostra, per $j = 1, \dots, c$

Cada laboratori fa n replicats de cada mostra, $k = 1, \dots, n$

A la taula s'hi han afegit les mitjanes per a cada laboratori, per a cada mostra i la mitjana global. A partir de tots aquests valors ja es pot calcular la contribució de cada font de variació (factors controlats, error aleatori, interacció) a la variació global.

Variància entre laboratoris:

$$(s_{lab})^2 = nc \sum (M_i - M_{global})^2 / (r-1) = SS_{lab} / (r-1)$$

Variància entre mostres:

$$(s_{mostres})^2 = nr \sum (M_j - M_{global})^2 / (c-1) = SS_{mostres} / (c-1)$$

Variància associada a l'error aleatori:

$$(s_{aleat})^2 = \sum \sum \sum (x_{ijk} - M_{ij})^2 / rc(n-1) = SS_{aleat} / rc(n-1)$$

Variància global:

$$(s_{global})^2 = \sum \sum \sum (x_{ijk} - M_{global})^2 / (nrc-1) = SS_{global} / (nrc-1)$$

Variància associada a la interacció:

Es determina per diferència:

$$SS_{interacció} = SS_{global} - SS_{lab} - SS_{mostra} - SS_{aleat}$$

$$\text{graus de llibertat} = (nrc-1) - ((c-1) + (r-1) + (rc(n-1))) = (r-1)(c-1)$$

$$(s_{interacció})^2 = SS_{interacció} / ((r-1)(c-1))$$

Un cop obtinguts aquests valors, es poden realitzar tres tests F diferents:

a) si es vol veure si hi ha diferències significatives entre laboratoris:

$$F = (s_{\text{lab}})^2 / (s_{\text{aleat}})^2$$

b) si es vol veure si hi ha diferències significatives entre mostres:

$$F = (s_{\text{mostres}})^2 / (s_{\text{aleat}})^2$$

c) si es vol veure si hi ha interacció entre els dos factors controlats:

$$F = (s_{\text{interacció}})^2 / (s_{\text{aleat}})^2$$

Tota la informació obtinguda en l'ANOVA es resumeix en una taula, tal i com s'indica a continuació:

Font de variació	Suma de quadrats	Graus de llibertat	Variància	Valor de F
Entre laboratoris	SS_{lab}	$r-1$	$SS_{\text{lab}} / (r-1)$	$s_{\text{lab}}^2 / s_{\text{aleat}}^2$
Entre mostres	SS_{mostres}	$c-1$	$SS_{\text{mostres}} / (c-1)$	$s_{\text{mostres}}^2 / s_{\text{aleat}}^2$
Interacció	$SS_{\text{interacció}}$	$(r-1)(c-1)$	$SS_{\text{interacció}} / ((r-1)(c-1))$	$s_{\text{interacció}}^2 / s_{\text{aleat}}^2$
Error aleatori	SS_{aleat}	$rc(n-1)$	$SS_{\text{aleat}} / (rc(n-1))$	
Global	SS_{global}	$rcn-1$	$SS_{\text{global}} / (rcn-1)$	

2.2.4. Ús del full de càlcul Excel per a l'aplicació de l'ANOVA de dos factors amb replicats

S'ha d'entrar en el menú **Herramientas** i en el submenú **Análisis de datos**. Un cop en aquest submenú s'ha de triar l'opció **Análisis de varianza de dos factores con varias muestras por grupo**. Apareix una finestra que està dividida en dues parts, amb els títols **ENTRADA** i **OPCIONES DE SALIDA**, respectivament. En la primera part es demana tota la informació necessària per a fer córrer el programa:

- **Rango de entrada:** correspon a les cel·les on hi ha les dades experimentals; també s'han d'incloure les cel·les que contenen les etiquetes identificatives.

- **Fila por muestra:** cada replicat s'ha de donar en una fila diferent. Aquí demanen quants replicats es tenen per a cada anàlisi. El nombre de replicats ha de ser el mateix per a totes les anàlisis.

- **Alfa:** hem de donar el nivell de significació que volem per a realitzar el test F.

En la segona part (**OPCIONES DE SALIDA**) es demana on es vol que el programa tregui els resultats. Els pot treure (a) en el mateix full on hem introduït les dades experimentals, (b) en un full nou i (c) en un llibre nou. El més simple és fer servir l'opció (a). S'ha de marcar aquesta opció i s'ha d'indicar en el requadre adjunt quina és la cel·la superior esquerra a partir de la qual es vol que apareguin els resultats.