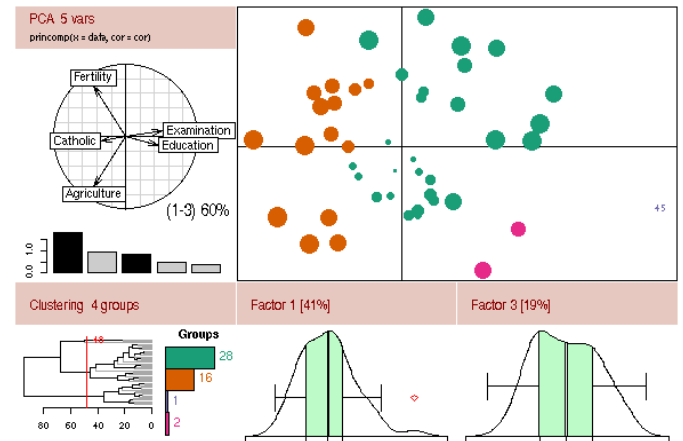


Blocs aleatoritzats i altres dissenys



$$y_{ij} = \mu + \alpha_i + e_{ij}$$



Blocs per controlar el biaix i l'error

- Hi ha factors sense interès en un estudi però que poden influir en el resultat: factors “soroll”.
- Possibles estratègies per eliminar-ne l'efecte:
 - Si desconegut i totalment incontrolable: aleatorització.
 - Si incontrolable però mesurable: restar-ne l'efecte en l'anàlisi dels resultats (cas de l'anàlisi de la covariància).
 - Si controlable experimentalment: fer blocs, comparar els tractament d'interès dins blocs fets segons nivells del(s) factor(s) de “soroll”.

Blocs aleatoritzats complets

- Sovint un factor de soroll és la pròpia unitat experimental
 - ✓ Exemples: els propis pacients en un estudi clínic, amb les seves característiques especials (salut, hàbits, ...), les parcel·les en un estudi de producció agrícola, la pròpia peça de material en estudis de duresa, etc.
- Possible estratègia:
 - *tots* els tractaments (factor(s) d'interès en l'estudi) a cada unitat experimental.
 - Ordre o lloc de tractament dins cada unitat a l'atzar. Temps de recuperació si és necessari.

CR vs RB

El disseny Completament Aleatoritzat (*Completely Randomized (CR) design*) divideix les unitats experimentals (u.e.) en a grups de tamany n_i i assigna a l'atzar els tractaments a cada grup.

El disseny en Blocs Aleatoritzats (*Randomized Block (RB) design*) divideix el grup d'unitats experimentals (N) en b grups homogenis de tamany a.

- Aquests grups s'anomenen **blocs**.
- S'assignen els tractaments a l'atzar a les u.e. Dintre de cada bloc - un tractament a una unitat de cada bloc.

CR

Tractaments

1

2

3

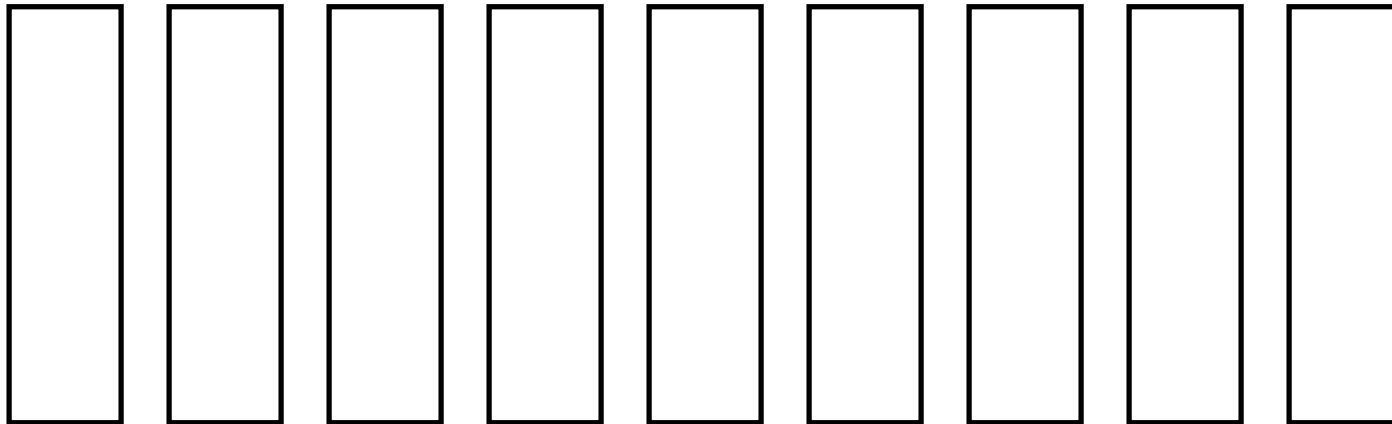
...

a

Unitats experimentals assignades a l'atzar als tractaments

RB

Blocs



Tots els tractaments apareixen una vegada a cada bloc

Exemple

- Es vol estudiar l'efecte de 4 productes químics (A, B, C i D) en la resistència a l'aigua (y) de determinades fibres tèxtils.
- Es selecciona una mostra aleatòria de material de cada bobina de teixit, i es talla en 4 peces, cadascuna de les quals rebrà a l'atzar un dels tractaments químics.
- L'experiment es repeteix 3 vegades (3 bobines).

Diagrama: Blocs (mostra de bobines)

9.9	C
10.1	A
11.4	B
12.1	D

13.4	D
12.9	B
12.2	A
12.3	C

12.7	B
12.9	D
11.4	C
11.9	A

Taula de dades

P. Quimic	Blocs (Bobines)		
	1	2	3
A	10.1	12.2	11.9
B	11.4	12.9	12.7
C	9.9	12.3	11.4
D	12.1	13.4	12.9

Model pel disseny en Blocs aleatoritzats

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad i = 1, 2, \dots, a \quad j = 1, 2, \dots, b$$

$$\sum_{i=1}^a \alpha_i = 0 \quad \sum_{j=1}^b \beta_j = 0$$

y_{ij} = observació del bloc j que ha rebut el tractament i

μ = mitjana global

α_i = efecte del tractament i

β_j = efecte del bloc j

ε_{ij} = error

Sumes de Quadrats

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_a$$

$$H_1 : \alpha_i \neq \alpha_{i'} \quad \text{per algún } i \neq i'$$

$$SS_T = SS_A + SS_B + SS_E$$

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b (Y_{ij} - \bar{Y}_{..})^2$$

$$SS_A = b \sum_{i=1}^a (\bar{Y}_{i.} - \bar{Y}_{..})^2 \quad SS_B = a \sum_{j=1}^b (\bar{Y}_{.j} - \bar{Y}_{..})^2$$

$$SS_E = \sum_{i=1}^a \sum_{j=1}^b (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2$$

Taula Anova

Font	S.S.	d.f.	M.S.	F	p-valor
Tract	SS_T	$a-1$	MS_T	MS_T/MS_E	
Bloc	SS_B	$b-1$	MS_B	MS_B/MS_E	
Error	SS_E	$(a-1)(b-1)$	MS_E		

Taula Anova

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
Blocks	7.17167	2	3.5858	40.21	0.0003
Chem	5.20000	3	1.7333	19.44	0.0017
ERROR	0.53500	6	0.0892		

Comparació de les taules Anova

Source	<i>df</i>	<i>Sum of Squares</i>
<i>Treatments</i>	$a - 1$	SS_{Tr}
<i>Error</i>	$a(n - 1)$	SS_{Error}
<i>Total</i>	$an - 1$	SS_{Total}

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}^{(CR)}$$

Source	<i>df</i>	<i>Sum of Squares</i>
<i>Blocks</i>	$b - 1$	SS_{Blocks}
<i>Treatments</i>	$a - 1$	SS_{Tr}
<i>Error</i>	$(a - 1)(b - 1)$	SS_{Error}
<i>Total</i>	$ab - 1$	SS_{Total}

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}^{(RB)}$$

Disseny en blocs aleatoritzats complet.

Comentaris finals(1)

- Suposició de no interacció, perillosa i no demostrable (en general) estadísticament.
- Cas de tractaments i/o blocs factor aleatori (*com hauria de ser a l'exemple*): són vàlides exactament les mateixes anàlisis (amb interpretació adequada).
- En general sempre més eficient que disseny totalment aleatoritzat.
- Comparacions múltiples entre tractaments: són vàlids tots els mètodes explicats pel cas d'un factor.

Disseny en blocs aleatoritzats complet.

Comentaris finals(2)

- Es poden replicar tant com es vulgui.
- Les dades faltants (missing) poden causar problemes
- L'eficiència disminueix quan augmenta el nombre de tractaments (i, per tant, la grandària del bloc)

Dissenys en Quadrats Llatins

Disseny en quadrats llatins

- Volem analitzar un factor amb t nivells (p.e. *Tractament*, amb $t = 4$ possibilitats: A, B, C, D).
- Hi ha dos factors (sense interès per ells mateixos) que sospitem que poden influir en la resposta, tots dos també amb t nivells (p.e. *Operari* i *Partida* de matèria prima): factors de bloc.
- Arrangem els factors de bloc en quadrat $t \times t$. Assignem un tractament a l'atzar a cada casella (*Operari* \times *Partida*) amb la restricció que no hi hagi cap tractament repetit a cap fila ni columna.

En un quadrat llatí hi ha tres factors:
Tractaments (t) (lletres A, B, C, ...)
Files (t)
Columnes (t)

El nombre de tractaments = el nombre de files = nombre de columnes = t .

Exemple

Una empresa de missatgeria està interessada a decidir entre cinc marques (D, P, F, C i R) de cotxe per a la pròxima compra de vehicles de la seva flota

- Les marques són comparables en preu de compra.
- La companyia vol portar a terme un estudi que els permeti comparar les marques pel que fa als costos d'operació.
- Per a això es seleccionarà a cinc conductors (files).
- A més, l'estudi es durà a terme durant un període de cinc setmanes (columnes = setmanes).

Cada setmana, a un conductor se li assigna un cotxe a l'atzar i utilitzant un dissenyen quadrat llatí.
Es registra el cost mitjà per quilòmetre al final de cada setmana i es tabula a continuació:

		Week				
		1	2	3	4	5
Drivers	1	5.83	6.22	7.67	9.43	6.57
		D	P	F	C	R
	2	4.80	7.56	10.34	5.82	9.86
		P	D	C	R	F
	3	7.43	11.29	7.01	10.48	9.27
		F	C	R	D	P
	4	6.60	9.54	11.11	10.84	15.05
		R	F	D	P	C
	5	11.24	6.34	11.30	12.58	16.04
		C	R	P	F	D

El Model per a un Quadrat Llatí

$$y_{ij(k)} = \mu + \tau_k + \rho_i + \gamma_j + \varepsilon_{ij(k)}$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, t \quad k = 1, 2, \dots, t$$

$y_{ij(k)}$ = observació de la fila i , columna j que ha rebut el tractament k

μ = mitjana global

τ_k = efecte del tractament k

ρ_i = efecte de la fila i

γ_j = efecte de la columna j

$\varepsilon_{ij(k)}$ = error

No interacció
entre files,
columnes y
tractaments

Taula Anova

Source	S.S.	d.f.	M.S.	F	p-value
Treat	SS_{Tr}	$t-1$	MS_{Tr}	MS_{Tr}/MS_E	
Rows	SS_{Row}	$t-1$	MS_{Row}	MS_{Row}/MS_E	
Cols	SS_{Col}	$t-1$	MS_{Col}	MS_{Col}/MS_E	
Error	SS_E	$(t-1)(t-2)$	MS_E		
Total	SS_T	$t^2 - 1$			

Taula Anova de l'exemple

Source	S.S.	d.f.	M.S.	F	p-value
Week	51.17887	4	12.79472	16.06	0.0001
Driver	69.44663	4	17.36166	21.79	0.0000
Car	70.90402	4	17.72601	22.24	0.0000
Error	9.56315	12	0.79693		
Total	201.09267	24			