DISEÑO FACTORIAL FRACCIONARIO

Profa. Dra. Kelly Johana Dussán Medina
kelly.medina@unesp.br

1

Cuando las variables son muchas

El número de ensayos necesarios para realizar un diseño factorial completo de 2^k aumenta rápidamente con k.

Con siete factores, por ejemplo, un diseño completo requeriría no menos de $2^7 = 128$ ensayos.

En primer lugar, el número de interacciones de alto orden aumenta drásticamente con el número de factores.

2

2

			Orden				
k	1ª	2ª	3ª	4ª	5ª	6ª	7ª
3	3	3	1	-	-	-	-
4	4	6	4	1	-	-	-
5	5	10	10	5	1	-	-
6	6	15	20	15	6	1	-
7	7	21	35	35	21	7	1

- En la mayoría de los casos, estas interacciones tienen valores pequeños y carecen de importancia práctica.
- Como en la expansión en serie de una función, los efectos principales tienden a ser mayores que las intersecciones de dos factores, que a su vez son más importantes que las interacciones de tres factores, y así sucesivamente.

3

3

En segundo lugar, a medida que aumenta el número de factores, también aumentan las posibilidades de que uno o varios de ellos no afecten significativamente a la respuesta. Para no correr el riesgo de excluir factores que pueden resultar importantes, en esta fase debemos estudiar el mayor número posible de variables.

Podemos hacerlo sin aumentar el número de ensayos utilizando **diseños fraccionados** en lugar de factoriales completos.

4

4

$2^4 = 16$

Ensayo	1	2	3	4	Respuesta
1	-	-	-	-	52
2	+	-	-	-	61
3	-	+	-	-	124
4	+	+	-	-	113
5	-	-	+	-	85
6	+	-	+	-	66
7	-	+	+	-	185
8	+	+	+	-	192
9	-	-	-	+	98
10	+	-	-	+	86
11	-	+	-	+	201
12	+	+	-	+	194
13	-	-	+	+	122
14	+	-	+	+	139
15	-	+	+	+	289
16	+	+	+	+	286

$2^{4-1} = 8$

Ensayo	1	2	3	4 (1 2 3)	Respuesta
1	-	-	-	-	52
10	+	-	-	+	86
11	-	+	-	+	201
4	+	+	-	-	113
13	-	-	+	+	122
6	+	-	+	-	66
7	-	+	+	-	185
16	+	+	+	+	286

5

5

Cómo construir una media fracción

La nota **123** indicará la columna de signos obtenida multiplicando las columnas correspondientes a los tres primeros factores. Esta columna es idéntica a la del factor **4**. Por tanto, podemos escribir,

$$4 = 123$$

- Para obtener las relaciones entre los distintos contrastes, utilizaremos dos propiedades de la multiplicación de columnas de signos.
- La primera es que multiplicando una columna por sí misma, es decir, elevando al cuadrado todos sus elementos, se obtiene siempre una columna que sólo contiene signos positivos.
- Esta nueva columna, a su vez, cuando se aplica a cualquier otra columna, la deja inalterada.

6

6

- En otras palabras, es el elemento de identidad de nuestra álgebra, por lo que utilizaremos la letra I para representarlo

$$1\ 1 = 2\ 2 = 3\ 3 = 4\ 4 = \mathrm{I}$$

- La segunda propiedad simplemente reconoce que la multiplicación de columnas es conmutativa y asociativa

$$1\ 2\ 3 = (1)\ (23) = (23)\ (1) = (12)\ (3) = 321 = (2)\ (31) = \dots$$

- Para obtener las relaciones entre los distintos contrastes, multiplicamos la expresión que define la fracción por algún producto de columnas y aplicamos las propiedades.

7

7

- Cuando queremos saber a qué equivale un determinado contraste, sólo tenemos que encontrar la manera de que aparezca solo en un lado de la ecuación:

$$4 = 123$$

- Por ejemplo, si queremos saber qué contraste tiene los mismos signos que L_2, podemos aislar el factor **2** del lado derecho multiplicando **123** por el producto **13**, porque así convertiremos en identidades el **1** y el **3** que ya están en la ecuación. Por supuesto, tenemos que multiplicar también el otro lado, para que se mantenga la relación de igualdad:

$$(13)\ (4) = (13)\ (123) = (11)\ (33)\ (2) = (\mathrm{I})\ (\mathrm{I})\ (2) = 2$$

En el lado izquierdo de la ecuación tenemos ahora el producto **134** y concluimos que $L_2 = L_{134}$

8

8

- En terminología estadística, decimos que el uso de la media fracción confunde el efecto principal **2** con la interacción **134**.
- El valor del contraste calculado, l_2 (ou l_{134}), es en realidad una estimación de la suma de los dos efectos.
- Para mostrar que el contraste calculado confunde los dos efectos y estima su suma, se utiliza la siguiente notación

$$l_2 = 2 + 134$$

A continuación, se presentan todas las relaciones entre los contrastes calculados en la fracción 2^{4-1} y los efectos obtenidos con el diseño completo 2^4 (los denominados **estándar de confusión**):

Relación entre las columnas	Contrastes de la media fracción 2^{4-1} en términos de efectos del completo 2^4
1 = 234	$l_1 = l_{234} \rightarrow 1 + 234$
2 = 134	$l_2 = l_{134} \rightarrow 2 + 134$
3 = 124	$l_3 = l_{124} \rightarrow 3 + 124$
4 = 123	$l_4 = l_{123} \rightarrow 4 + 123$
12 = 34	$l_{12} = l_{34} \rightarrow 12 + 34$
13 = 24	$l_{13} = l_{24} \rightarrow 13 + 24$
14 = 23	$l_{14} = l_{23} \rightarrow 14 + 23$
I = 1234	$l_I \rightarrow M + ½ (1234)$

Ejercício 1

Escribe las expresiones algebraicas para calcular los efectos **2** y **134** en el factorial completo 2^4 y demuestra que el contraste $\boldsymbol{L_2}$ calculado en la media fracción (2^{4-1}) corresponde realmente a la suma de estos dos efectos.

11

11

Concepto de resolución

Factoriales de resolución IV

- El diseño 2^{4-1} es un ejemplo de factorial fraccional de resolución cuatro. En un factorial de resolución cuatro, los efectos principales no se mezclan con las interacciones de los dos factores, sino que ellas se mezclan entre sí.
- La notación utilizada para representar la resolución de un diseño es un índice en números romanos. Por ejemplo, escribimos 2_{IV}^{4-1}.
- La resolución de un factorial viene determinada por sus relaciones generadoras. El número de factores que constituyen el término más corto de estas relaciones es, por definición, la resolución del diseño.

12

12

- Para definir una media fracción, sólo se necesita una relación generatriz. En nuestro ejemplo, esta relación I = 1 2 3 4 contiene cuatro factores, por lo que la resolución del factorial 2^{4-1} es cuatro.

Factoriales fraccionarios con resolución V

Una media fracción de una planificación 2^{5-1} se construye a partir de la relación:

$$5 = 1\,2\,3\,4 \;\rightarrow\; I = 1\,2\,3\,4\,5$$

Se trata, por tanto, de una media fracción de resolución 5, para la que podemos utilizar la notación 2_V^{5-1}.

Los efectos principales sólo se mezclan con las interacciones de cuatro factores, mientras que las interacciones de dos factores se mezclan con las de tres.

13

13

Relación entre las columnas	Contrastes
1 = 2345	$l_1 = l_{2345} \rightarrow$ 1 + 2345
2 = 1345	$l_2 = l_{1345} \rightarrow$ 2 + 1345
3 = 1245	$l_3 = l_{1245} \rightarrow$ 3 + 1245
4 = 1235	$l_4 = l_{1235} \rightarrow$ 4 + 1235
5 = 1234	$l_5 = l_{1234} \rightarrow$ 5 + 1234
12 = 345	$l_{12} = l_{345} \rightarrow$ 12 + 345
13 = 245	$l_{13} = l_{245} \rightarrow$ 13 + 245
14 = 235	$l_{14} = l_{235} \rightarrow$ 14 + 235
15 = 234	$l_{15} = l_{234} \rightarrow$ 15 + 234
23 = 145	$l_{23} = l_{145} \rightarrow$ 23 + 145
24 = 135	$l_{24} = l_{135} \rightarrow$ 24 + 135
25 = 134	$l_{25} = l_{134} \rightarrow$ 25 + 134
34 = 125	$l_{34} = l_{125} \rightarrow$ 34 + 125
35 = 124	$l_{35} = l_{125} \rightarrow$ 35 + 125
45 = 123	$l_{45} = l_{123} \rightarrow$ 45 + 123
I = 12345	$l_I \rightarrow$ M + ½ (12345)

14

14

Fracciones medias con resolución máxima

- Para construir las medias fracciones presentadas hasta ahora, utilizamos el efecto de interacción de orden superior para determinar los signos de columna de uno de los factores.
- En el primer ejemplo, partimos de un factorial de 2^3 y utilizamos la interacción **1 2 3** para definir los niveles de la cuarta variable, utilizando la relación I = 1 2 3 4. Esto nos llevó a una fracción de resolución 4.
- En el segundo ejemplo, partimos de un factorial 2^4 y, utilizando la relación I = 1 2 3 4 5, llegamos a una media fracción de resolución 5.
- Podríamos definir los signos de la variable 5 en la fracción 2^{5-1} mediante la relación 5 = 1 2 3. En este caso, la relación generadora pasaría a ser I = 1235 y, en consecuencia, la resolución bajaría a 4.

15

15

- Como las medias fracciones estudiadas hasta ahora se basan en la mayor interacción posible, tienen la mayor resolución para el número de factores considerados, por eso suelen ser las mejores.
- En general, para construir una fracción 2^{k-1} de máxima resolución, debemos hacer lo siguiente:

1. Escribir el diseño completo para k-1 variables;
2. Asignar a la variable restante los signos de interacción 1 2 3 .... (k).

16

16

Estudio de variables

- Hasta ahora sólo hemos tratado con medias fracciones, en las que hacemos la mitad de los ensayos del diseño completo. Dependiendo del número de factores, esta fracción puede ser demasiado grande.
- Con 5 variables, el diseño sólo tendría 8 ensayos y correspondería a una fracción 2^{5-2}. Para construir la matriz, empezaríamos con un factorial 2^3 basado en tres de las 5 variables. Luego necesitaríamos dos relaciones generadoras para definir los niveles de las dos variables restantes.
- Partimos de las relaciones **4 = 1 2 3** e **5 = 1 2**, lo que equivale a hacer I = 1234 e I = 125. Como el término más pequeño de estas relaciones contiene tres factores, el diseño tiene resolución tres, y su notación completa es 2_{III}^{5-2}.

17

17

- Obviamente, este tipo de diseños es más económica. Por otra parte, produce contrastes que mezclan efectos principales con la interacción de dos factores. Esto complica el análisis de los resultados, pero es posible que algunos de estos contrastes sean lo suficientemente pequeños como para permitirnos descartar las variables correspondientes.

Diseños saturados

- Estudiar más factores con cada vez menos ensayos.
- Dado un cierto número de ensayos, debe haber un número máximo de factores que podemos estudiar con ellas. Cuando se alcanza este límite, decimos que el diseño está **saturado**.

18

18

$$2_{III}^{7-4}$$

Ensayo	1	2	3	4	5	6	7	%aciertos
1	-	-	-	+	+	+	-	56
2	+	-	-	-	-	+	+	66
3	-	+	-	-	+	-	+	51
4	+	+	-	+	-	-	-	52
5	-	-	+	+	-	-	+	54
6	+	-	+	-	+	-	-	70
7	-	+	+	-	-	+	-	42
8	+	+	+	+	+	+	+	64
Equivalencias:	4 = 12	5 = 13	6 = 23	7 = 123				

19

19

Ejercicio 2

Calcule los valores de contraste correspondientes a los efectos principales y analícelos mediante gráficos normales.

Factor	Nivel +	Nivel -
Técnica	Globo	Volea
Frecuencia	Baja	Alta
Hora	Dia	Noche
Revestimiento	Polvo de Ladrillo	Concreto
Lado	Directo	Izquierdo
Camisa	Con	Sin
Raqueta	Media	Grande

Ensayo	1	2	3	4	5	6	7	%aciertos
1	-	-	-	+	+	+	-	56
2	+	-	-	-	-	+	+	66
3	-	+	-	-	+	-	+	51
4	+	+	-	+	-	-	-	52
5	-	-	+	+	-	-	+	54
6	+	-	+	-	+	-	-	70
7	-	+	+	-	-	+	-	42
8	+	+	+	+	+	+	+	64
Equivalencias:	4 = 12	5 = 13	6 = 23	7 = 123				

20

20

DISEÑOS SATURADOS PLACKETT E BURMAN

Fundamentos Teóricos

- Si disponemos de las condiciones para realizar 8, 16, 32, ..., 2^m ensayos, podemos utilizar los diseños saturados para estudiar la influencia de hasta 7, 15, 31, ..., 2^{m-1} factores.
- Otra clase de diseños fraccionados emplea un total de 12, 20, 24, 28, ..., ensayos para investigar simultáneamente hasta 11, 19, 23, 27, ..., factores.
- Estos diseños, propuestos por R.L. Plackett e J.P. Burman, permiten estimar todos los k = n-1 efectos principales (donde "n" representa el número de ensayos) con variancia mínima.

- Primera columna de la matriz de diseño

N (ensayos)	k (factores)	Señales
8	7	+ + + - + - -
12	11	+ + - + + + - - - + -
16	15	+ + + + - + - + + - - + - - -
20	19	+ + - - + + + + - + - + - - - - + + -
24	23	+ + + + + - + - + + - - + + - - + - + - - - -
28	27	+ + - - - - + + + + - + + + - + + - - - + - + - - + -

- La segunda fila (columna) se obtiene a partir de la primera desplazando los elementos de la fila una posición hacia abajo y colocando el último elemento en la primera posición. Y así sucesivamente.
- Se añade una fila de signos -1, completando el diseño.

23

23

Existen 4 formas equivalentes de construir matrices:

Permutación para abajo		
+	-	+
+	+	-
-	+	+

Permutación para arriba		
+	+	-
+	-	+
-	+	+

Permutación a la izquierda		
+	+	-
+	-	+
-	+	+

Permutación a la derecha		
+	+	-
-	+	+
+	-	+

24

24

Diseño fraccional saturado de Plackett y Burman para el estudio de 11 variables con 12 experimentos

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
1	+1	-1	+1	-1	-1	-1	+1	+1	+1	-1	+1
2	+1	+1	-1	+1	-1	-1	+1	+1	+1	+1	-1
3	-1	+1	+1	-1	+1	-1	-1	-1	+1	+1	+1
4	+1	-1	+1	+1	-1	+1	-1	-1	-1	+1	+1
5	+1	+1	-1	+1	+1	-1	+1	-1	-1	-1	+1
6	+1	+1	+1	-1	+1	+1	-1	+1	-1	-1	-1
7	-1	+1	+1	+1	-1	+1	+1	-1	+1	-1	-1
8	-1	-1	+1	+1	+1	-1	+1	+1	-1	+1	-1
9	-1	-1	-1	+1	+1	+1	-1	+1	+1	-1	+1
10	+1	-1	-1	-1	+1	+1	+1	-1	+1	+1	-1
11	-1	+1	-1	-1	-1	+1	+1	+1	-1	+1	+1
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

25

Las columnas son todas ortogonales, y esta simetría permite determinar individualmente los efectos principales de cada factor, suponiendo que los efectos de interacción son despreciables.

Aunque en un diseño saturado con "n" ensayos es posible estudiar hasta "n-1" factores, es aconsejable elegir un número menor para que las columnas no utilizadas desempeñen el papel de **variables inertes** y puedan utilizarse **para estimar el error asociado a los contrastes**.

En el caso del diseño Plackett-Burman, se recomienda que el número de factores reales no supere "n-4".

Una desventaja es que las relaciones entre los contrastes calculados y los efectos de un factorial completo son bastante complejas. Esto hace que sea muy difícil elegir los ensayos adicionales necesarios para no se confundan los efectos.

26

Número de matrices adecuadas en función del número de factores

Factores	# de ensayos matriz PB	Factores	# de ensayos matriz PB
4	8 12 ≥ 16	11	16 20 ≥ 24
5	12 16 ≥ 20	12	16 20 ≥ 24
6	12 16 ≥ 20	13	20 24 ≥ 28
7	12 16 ≥ 20	14	20 24 ≥ 28
8	12 16 ≥ 20	15	20 24 ≥ 28
9	16 20 ≥ 24	16	20 24 ≥ 28
10	16 20 ≥ 24	≥ 17	Siempre al menos 4 ensayos más que el número de factores

27

27

4
- $2^4 = 16$
- $2^{4-1} = 8$
- PB 8
- PB 12
- PB 16

5
- $2^5 = 32$
- $2^{5-1} = 16$
- PB 12
- PB 16
- PB 20
- PB 24
- PB 28
- PB 32

6
- $2^6 = 64$
- $2^{6-1} = 32$
- $2^{6-2} = 16$
- PB 12
- PB 16
- PB 20
- PB 24
- PB 28
- PB 32
- PB 36
- PB 40

7
- $2^7 = 128$
- $2^{7-1} = 64$
- $2^{7-2} = 32$
- $2^{7-3} = 16$
- PB 12
- PB 16
- PB 20
- PB 24
- PB 28
- PB 32
- PB 36
- PB 40

28

28

Table 1. Plackett-Burman experimental design matrix for screening composition of growth medium, BG-11 for *Synechocystis aquatilis*

$k = 8$

Trial	Level and concentration of variable (g/L)							
	X_1 $NaNO_3$	X_2 K_2HPO_4	X_3 $MgSO_4.7H_2O$	X_4 $CaCl_2.2H_2O$	X_5 Citric acid	X_6 Ferric ammonium citrate	X_7 EDTA	X_8 Na_2CO_3
T_1	+1 (2.25)	-1 (0.02)	+1 (0.1125)	-1 (0.018)	-1 (0.003)	-1 (0.003)	+1 (0.0015)	+1 (0.03)
T_2	+1 (2.25)	+1 (0.06)	-1 (0.0375)	+1 (0.054)	-1 (0.003)	-1 (0.003)	-1 (0.0005)	+1 (0.03)
T_3	-1 (0.75)	+1 (0.06)	+1 (0.1125)	-1 (0.018)	+1 (0.009)	-1 (0.003)	-1 (0.0005)	-1 (0.01)
T_4	+1 (2.25)	-1 (0.02)	+1 (0.1125)	+1 (0.054)	-1 (0.003)	+1 (0.009)	-1 (0.0005)	-1 (0.01)
T_5	+1 (2.25)	+1 (0.06)	-1 (0.0375)	+1 (0.054)	+1 (0.009)	-1 (0.003)	+1 (0.0015)	-1 (0.01)
T_6	+1 (2.25)	+1 (0.06)	+1 (0.1125)	-1 (0.018)	+1 (0.009)	+1 (0.009)	-1 (0.0005)	+1 (0.03)
T_7	-1 (0.75)	+1 (0.06)	+1 (0.1125)	+1 (0.054)	-1 (0.003)	+1 (0.009)	+1 (0.0015)	-1 (0.01)
T_8	-1 (0.75)	-1 (0.02)	+1 (0.1125)	+1 (0.054)	+1 (0.009)	-1 (0.003)	+1 (0.0015)	+1 (0.03)
T_9	-1 (0.75)	-1 (0.02)	-1 (0.0375)	+1 (0.054)	+1 (0.009)	+1 (0.009)	-1 (0.0005)	+1 (0.03)
T_{10}	+1 (2.25)	-1 (0.02)	-1 (0.0375)	-1 (0.018)	+1 (0.009)	+1 (0.009)	+1 (0.0015)	-1 (0.01)
T_{11}	-1 (0.75)	+1 (0.06)	-1 (0.0375)	-1 (0.018)	-1 (0.003)	+1 (0.009)	+1 (0.0015)	+1 (0.03)
T_{12}	-1 (0.75)	-1 (0.02)	-1 (0.0375)	-1 (0.018)	-1 (0.003)	-1 (0.003)	-1 (0.0005)	-1 (0.01)

Values in parentheses are concentrations in g/L of each variable in BG-11; Level of micronutrients in all experiments was kept constant.

$N = 12$

todos os contrastes da última linha são iguais a -1

29

	2^7 (128)	2^{7-1} (64)	2^{7-2} (32)	2^{7-3} (16)	2^{7-4} (8)	PB 8	PB 12	PB16	PB 20	PB 24
E	**-9,7**	**-9,6**	**-9,6**	**-9,6**	-7,2	-12,0	**-8,87**	**-9,63**	**-11,21**	**-9,57**
Fr	**8,9**	**8,9**	**8,9**	**9,0**	7,0	11,35	**10,57**	**8,28**	**10,09**	**8,25**
Co	**-8,7**	**-8,6**	**-8,6**	**-8,6**	-8,7	-8,65	**-9,03**	**-8,60**	**-9,79**	**-8,73**
F1	**-10,6**	**-10,6**	**-10,6**	**-10,6**	-8,7	-12,80	**-10,20**	**-10,60**	**-12,29**	**-10,70**
V1	0,3	0,3	0,2	0,8	-1,6	-3,85	3,60	-1,20	1,05	-0,82
F2	1,4	1,4	1,1	2,2	3,7	5,30	0,07	-1,37	1,41	0,60
V2	0,0	0,0	0,0	-0,7	-2,2	3,90	0,17	3,40	1,19	-0,53
Error Estándar	0,1	0,2	1,4	2,4	-	-	3,03	1,56	1,59	1,35
Media	81,7	81,7	81,7	81,7	81,7	81,73	81,7	81,75	81,5	81,62

30

Ejercicio 3

Plackett & Burman de 16 ensayos para el rendimiento de respuesta Y (%) de un proceso de purificación de enzimas con 7 factores. Calcule los efectos, el error estándar y concluya.

Ensayos	Rend (%)
1	84,0
2	90,3
3	84,2
4	68,8
5	84,1
6	56,4
7	90,1
8	75,2

Ensayos	Rend (%)
9	83,4
10	92,0
11	72,5
12	73,2
13	94,2
14	86,4
15	83,1
16	90,1

31