Diseño de Experimentos - 3007340
DOE - Diseño factoriales 2^{k-p} - Parte II

Nelfi González Alvarez

Profesora Asociada Escuela de Estadística

e-mail: ngonzale@unal.edu.co

Facultad de Ciencias, Universidad Nacional de Colombia Sede Medellín

UNIVERSIDAD NACIONAL DE COLOMBIA

Escuela de Estadística

Semestre 02 de 2021

Contenido I

Contenido

El concepto de resolución

Recuerde que (Gutiérrez y de la Vara Salazar, 2012):

- En un experimento factorial fraccionado sólo pueden estimarse las sumas (o restas) de efectos aliados.
- La interpretación de los efectos aliados se hace fácilmente si puede suponerse que sólo uno de los efectos aliados es importante, de modo que el efecto total del grupo de alias se puede atribuir a este único efecto importante.
- Por tanto, siempre que sea posible se debe elegir diseños fraccionados en los cuales los efectos potencialmente importantes sean alias de efectos que de antemano son irrelevantes.

De modo que:

Bajo el principio de jerarquía en el que los efectos principales son más importantes que las interacciones dobles, y estos a su vez son más importantes que las interacciones triples, etc., se deben usar fracciones que tengan alta resolución.

A mayor resolución se puede evaluar más claramente lo que sucede con los efectos potencialmente importantes.

Definición 1.1 (Gutiérrez y de la Vara Salazar (2012))

Un diseño factorial fraccionado 2^{k-p} es de resolución R si los efectos formados por la interacción de p factores no son alias de efectos de interacción que tengan menos de $R-p$ factores.

De la anterior definición tenemos,

- **Diseños de resolución III:** Si los efectos principales no son alias entre ellos pero algunos son alias de interacciones dobles. Ej. Un 2^{3-1} con relación definidora $I = ABC$ es de resolución III, pues su estructura alias es

$$A+BC, \quad B+AC, \quad C+AB$$

- **Diseños de resolución IV:** Los efectos principales no están aliados entre ellos ni con interacciones dobles, pero algunas interacciones dobles están aliadas entre ellas. Por ejemplo, la fracción 2^{6-2} con generadores $ACEF$ y $BDEF$, por tanto con relación definidora $I = ACEF = BDEF = ABCD$, y estructura alias,

$A+CEF+ABDEF+BCD$	$B+ABCEF+DEF+ACD$	$C+AEF+BCDEF+ABD$
$D+ACDEF+BEF+ABC$	$E+ACF+BDF+ABCDE$	$F+ACE+BDE+ABCDF$
$AB+BCEF+ADEF+CD$	$AC+EF+ABCDEF+BD$	$AD+CDEF+ABEF+BC$
$AE+CF+ABDF+BCDE$	$AF+CE+ABDE+BCDF$	$BE+ABCF+DF+ACDE$
$BF+ABCE+DE+ACDF$	$ADF+CDE+ABE+BCF$	$ABF+BCE+ADE+CDF$

- **Diseños de resolución V:** Los efectos principales y las interacciones dobles sólo están aliados con interacciones triples o de orden superior. Ej. Un2^{5-1} con relación definidora $I = ABCDE$ es de resolución V, pues su estructura alias es

$A+BCDE$	$AB+CDE$	$BD+ACE$
$B+ACDE$	$AC+BDE$	$BE+ACD$
$C+ABDE$	$AD+BCE$	$CD+ABE$
$D+ABCE$	$AE+BCD$	$CE+ABD$
$E+ABCD$	$BC+ADE$	$DE+ABC$

- En general, *en los experimentos factoriales fraccionados 2^{k-p} la resolución está dada por el efecto de la relación definidora con el menor número de letras.* Así pues, como en los factoriales fraccionados 2^{k-1} la relación definidora es el mismo generador, entonces *la resolución es igual al número de letras del generador,* por eso en los diseños 2^{3-1}, 2^{4-1} y 2^{5-1} la resolución es respectivamente III, IV, y V.

Contenido

Factorial fraccionado saturado

Por lo general son experimentos factoriales fraccionados de resolución III en el que se estudian los efectos principales de los k factores usando solo $k+1$ corridas, por tanto no hay grados de libertad para el MSE del modelo correspondiente,

$$Y = \beta_0 + \sum_{j=1}^{k} \beta_j X_j + E, \text{ con } E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right). \tag{1}$$

Por ejemplo, un factorial fraccionado 2^{7-4}_{III} con generadores ABD, ACE, BCF y $ABCG$,

A	B	C	$D=AB$	$E=AC$	$F=BC$	$G=ABC$
−1	−1	−1	+1	+1	+1	−1
+1	−1	−1	−1	−1	+1	+1
−1	+1	−1	−1	+1	−1	+1
+1	+1	−1	+1	−1	−1	−1
−1	−1	+1	+1	−1	−1	+1
+1	−1	+1	−1	+1	−1	−1
−1	+1	+1	−1	−1	+1	−1
+1	+1	+1	+1	+1	+1	+1

La relación definidora es

$$I = ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG = ABEF = BEG = \\ AFG = DEF = ADEG = BDFG = CEFG = ABCDEFG$$

Observe que el efecto con menor número de letras es de 3, por tanto la resolución es III.
Con la ayuda de R podemos verificar esto con la estructura alias reducida.

```
> #Generando fracción 2^(7-4) con D=AB, E=AC, F=BC, G=ABC
> diseño3=FrF2(nruns=8,nfactors=7,generators=c("AB","AC","BC","ABC"),randomize=FALSE,alias.info=3)
> diseño3
   A  B  C  D  E  F  G
1 -1 -1 -1  1  1  1 -1
2  1 -1 -1 -1 -1  1  1
3 -1  1 -1 -1  1 -1  1
4  1  1 -1  1 -1 -1 -1
5 -1 -1  1  1 -1 -1  1
6  1 -1  1 -1  1 -1 -1
7 -1  1  1 -1 -1  1 -1
8  1  1  1  1  1  1  1
class=design, type= FrF2.generators

> design.info(diseño3)$aliased
$legend
[1] "A=A" "B=B" "C=C" "D=D" "E=E" "F=F" "G=G"

$main
[1] "A=BD=CE=FG=BCG=BEF=CDF=DEG" "B=AD=CF=EG=ACG=AEF=CDE=DFG" "C=AE=BF=DG=ABG=ADF=BDE=EFG" "D=AB=CG=EF=ACF=AEG=BCE=BFG"
[6] "E=AC=BG=DF=ABF=ADG=BCD=CFG" "F=AG=BC=DE=ABE=ACD=BDG=CEG" "G=AF=BE=CD=ABC=ADE=BDF=CEF"

$fi2
character(0)

$fi3
[1] "ABD=ACE=AFG=BCF=BEG=CDG=DEF"
```

Observe que efectos principales no están aliados entre sí pero sí con algunos efectos de interacción doble, por esto la resolución de esta fracción es III

De acuerdo a la estructura alias y por orden de importancia, solo pueden considerarse en el MRLM efectos principales, y el número de efectos cumple la regla de ser igual a $2^{k-p}-1=7$, que es en este caso el mismo valor de $k=7$. Con este diseño sólo se pueden estudiar los efectos principales, suponiendo a priori que las interacciones dobles no son importantes, lo cual conlleva un riesgo si en realidad hay interacciones dobles importantes. Si solo se observa una vez cada uno de los tratamientos, entonces el ajuste del MRLM será saturado y deberá aplicarse los métodos de Pareto y gráfico de Daniel para la selección de efectos, pero será imposible tener al menos 8 grados de libertad en el error en el modelo predictivo.

Contenido

Otras consideraciones

- *En general no se deben correr experimentos de resolución III, a menos que se esté dispuesto a aceptar que sólo importan los efectos principales.* Excepciones:
    - Cuando se tiene una cantidad muy grande de factores a estudiar;
    - Con pocos factores, si cada corrida experimental es muy costosa.
- *Buscar la máxima resolución posible con un número razonable de corridas experimentales.* Para $5 < k \leq 15$ existen fracciones de resolución IV que no requieren más de 32 corridas experimentales.
- *Si las corridas son económicas y el proceso es masivo, optar por fracciones de resolución V.* En este tipo de procesos es posible incrementar el número de corridas experimentales hasta 64 o más.
- *Diseños factoriales con igual resolución no implican que tengan la misma habilidad para estimar efectos potencialmente importantes.* Por ejemplo, para un diseño 2^{7-2} se tiene las siguientes dos posibles relaciones definidoras:

  $$I = DEFG = ABCDF = ABCEG \text{ y } I = ABCF = ADEG = BCDEFG$$

  ambas conducen a una fracción de resolución IV, pero la primera tiene tres pares de interacciones dobles que son alias, mientras que la segunda tiene seis pares. Luego la primero permite estimar más interacciones dobles limpiamente.

Contenido

Ejemplo: Sección 10.9 en Notas de Clase

En una empresa se compró un equipo de corte. Después de dos semanas notan que el exceso de vibración en el proceso es un problema muy serio. Un equipo de mejoramiento que incluye a los ingenieros del proceso decide tratar de reducir la vibración aplicando un diseño de experimentos. El equipo identifica siete factores, todos asociados con la herramienta de corte, que pueden tener algo que ver con la cantidad de vibración:

- A: tamaño de grano del material
- B: longitud del material
- C: diámetro del material
- D: revoluciones por minuto
- E: peso de la precarga
- F: estructura del material
- G: velocidad de alimentación

Se seleccionan dos niveles para cada factor. Como la puesta en marcha del equipo y el tiempo de corrida tienen un alto costo, se decide usar un experimento de ocho corridas. Un miembro del grupo de mejoramiento con experiencia previa en máquinas similares asegura que los efectos de interacción entre los factores se pueden considerar despreciables, lo cual hace factible el uso de una fracción que sólo permite el estudio de los efectos principales. Los factores y sus niveles aparecen a continuación:

Factor	Descripción (unidades)	Niveles (bajo, alto)
A	Tamaño de grano (/pulg.)	80 − 120
B	Longitud (pulg.)	1.0 − 2.0
C	Diámetro (pulg.)	1.0 − 1.5
D	rpm(x1000)	15 − 20
E	Peso de precarga (lbs.)	1.0 − 4.0
F	Estructura material (onzas)	1.0 − 4.0
G	Velocidad alim. (pulg./min.)	2.0 − 4.0

Despreciando todos los efectos de interacción, se corre un diseño altamente fraccionado y saturado, usando una fracción 2^{7-4} de resolución III. La matriz de diseño y la vibración observada se dan a continuación:

A	B	C	D	E	F	G	Y
−1	−1	−1	+1	+1	+1	−1	77.4
+1	−1	−1	−1	−1	+1	+1	68.3
−1	+1	−1	−1	+1	−1	+1	81.9
+1	+1	−1	+1	−1	−1	−1	66.2
−1	−1	+1	+1	−1	−1	+1	42.1
+1	−1	+1	−1	+1	−1	−1	78.3
−1	+1	+1	−1	−1	+1	−1	39.0
+1	+1	+1	+1	+1	+1	+1	68.4

Los tratamientos observados fueron *def*, *afg*, *beg*, *abd*, *cdg*, *ace*, *bcf* y *abcdefg*. Es fácil verificar que esta fracción es la que previamente se presentó, es decir, tiene como generadores a ABD, ACE, BCF y ABCG, y por tanto su relación definidora: $I = ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG = ABEF = BEG = AFG = DEF = ADEG = BDFG = CEFG = ABCDEFG$. Su estructura alias reducida mostrando solo alias para efectos principales con las interacciones dobles es

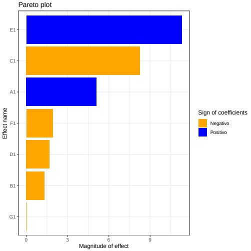
$A + BD + CE + FG$
$B + AD + CF + EG$
$C + AE + BF + DG$
$D + AB + CG + EF$
$E + AC + BG + DF$
$F + AG + BC + DE$
$G + AF + BE + CD$

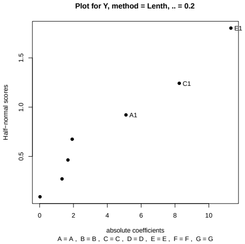
El MRLM preliminar es, $Y = \beta_0 + \sum_{j=1}^{7} \beta_j X_j + E,\ E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$.

Consultar al final del Capítulo 10 de Notas de Clase el Código R usado.

MATRIZ DE DISEÑO Y VIBRACIÓN OBSERVADA								CÁLCULOS PARA CONTRASTES						
A	B	C	D	E	F	G	Y	A	B	C	D	E	F	G
-1	-1	-1	1	1	1	-1	77.4	-77.4	-77.4	-77.4	77.4	77.4	77.4	-77.4
1	-1	-1	-1	-1	1	1	68.3	68.3	-68.3	-68.3	-68.3	-68.3	68.3	68.3
-1	1	-1	-1	1	-1	1	81.9	-81.9	81.9	-81.9	-81.9	81.9	-81.9	81.9
1	1	-1	1	-1	-1	-1	66.2	66.2	66.2	-66.2	66.2	-66.2	-66.2	-66.2
-1	-1	1	1	-1	-1	1	42.1	-42.1	-42.1	42.1	42.1	-42.1	-42.1	42.1
1	-1	1	-1	1	-1	-1	78.3	78.3	-78.3	78.3	-78.3	78.3	-78.3	-78.3
-1	1	1	-1	-1	1	-1	39.0	-39.0	39.0	39.0	-39.0	-39.0	39.0	-39.0
1	1	1	1	1	1	1	68.4	68.4	68.4	68.4	68.4	68.4	68.4	68.4
CONTRASTES								**40.8**	**-10.60**	**-66.0**	**-13.40**	**90.4**	**-15.40**	**-0.20**
EFECTOS (CONTRASTE $\div\ 2^{k-p-1}$, pues $n=1$)								**10.2**	**-2.65**	**-16.5**	**-3.35**	**22.6**	**-3.85**	**-0.05**
PARÁMETROS ESTIMADOS (EFECTOS $\div$ 2)								**5.1**	**-1.325**	**-8.25**	**-1.675**	**11.3**	**-1.925**	**-0.025**
Promedio datos $\overline{Y} = 65.2$														

```
#Regresión con variables codificadas
 > modelo=lm(Y~A+B+C+D+E+F+G)
 > summary(modelo)
 Call: lm.default(formula = Y ~ A + B + C + D + E  + F + G)
 Residuals: ALL 8 residuals are 0: no residual
 degrees of freedom!
 Coefficients:
 Estimate Std. Error t value Pr(>|t|)
 (Intercept)   65.200         NA      NA       NA
 A1             5.100         NA      NA       NA
 B1            -1.325         NA      NA       NA
 C1            -8.250         NA      NA       NA
 D1            -1.675         NA      NA       NA
 E1            11.300         NA      NA       NA
 F1            -1.925         NA      NA       NA
 G1            -0.025         NA      NA       NA
 Residual standard error: NaN on 0 degrees of
 freedom. Multiple R-squared: 1,  Adjusted  R-squared: NaN
 F-statistic: NaN on 7 and 0 DF, p-value: NA
```





El gráfico de pareto y de probabilidad media normal (con $\alpha = 0.2$) señalan como efectos importantes a retener, solo a los efectos alias asociados a E, C y A, en ese orden, de modo que se enviarán al error las SS asociadas a B, D, F y G, logrando solo un SSE aproximado con cuatro grados de libertad.

Sumas de cuadrados de los siete efectos alias interpretados como efectos principales

	A	B	C	D	E	F	G
contraste de totales	40.80	-10.60	-66.0	-13.400	90.40	-15.400	-0.200
$(\text{contraste})^2$	1664.64	112.36	4356.0	179.560	8172.16	237.160	0.040
$SS_X = (\text{contraste})^2 \div n2^{k-p}$, con $n=1$	208.08	14.045	544.5	22.445	1021.52	29.645	0.005
$SSE(B+D+F+G)$	66.14						

En R, las SS de los efectos en el modelo saturado,

```
> anova(modelo)
Analysis of Variance Table
Response: Y
          Df  Sum Sq Mean Sq F value Pr(>F)
A          1  208.08  208.08     NaN    NaN
B          1   14.04   14.04     NaN    NaN
C          1  544.50  544.50     NaN    NaN
D          1   22.44   22.44     NaN    NaN
E          1 1021.52 1021.52     NaN    NaN
F          1   29.64   29.64     NaN    NaN
G          1    0.00    0.00     NaN    NaN
Residuals  0    0.00     NaN
Warning message:
In anova.lm(modelo) :
ANOVA F-tests on an essentially perfect fit are unreliable
```

Al llevar al error los efectos B, D, F, G, el modelo predictivo es $Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_5 X_5 + E$, $E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$. Para la ec. estimada basta tomar del ajuste saturado solo las estimaciones de los parámetros que quedan en la ec. predictiva, es decir,

$$\widehat{Y} = 65.2 + 5.1X_1 - 8.25X_3 + 11.3X_5$$

Puede verificar este resultado corriendo en R el modelo predictivo:

```
> modelo2=lm(Y~A+C+E)
> summary(modelo2)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   65.200      1.438  45.351 1.41e-06 ***
A1             5.100      1.438   3.547  0.02385 *
C1            -8.250      1.438  -5.738  0.00457 **
E1            11.300      1.438   7.860  0.00142 **
---
Residual standard error: 4.066 on 4 degrees of freedom
Multiple R-squared:  0.9641, Adjusted R-squared:  0.9371
F-statistic: 35.76 on 3 and 4 DF,  p-value: 0.002393
```

El ANOVA del modelo predictivo se arma con las SS de los efectos retenidos y el SSE aproximado con las SS de los efectos despreciados.

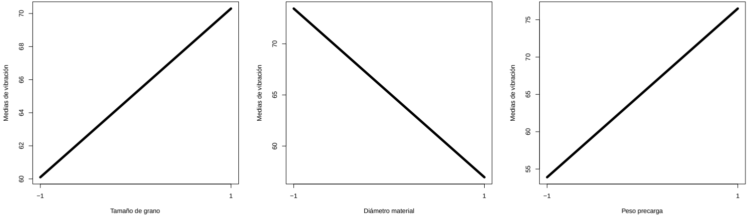
FUENTE	g.l	SS	CM	F0	Valor P
A: grano	1	208.08	208.08	12.58	0.0239
C: diámetro	1	544.50	544.50	32.93	0.0046
E: precarga	1	1021.52	1021.25	61.78	0.0140
Error	4	66.14	16.54		
Total	7	1840.24	$R^2_{adj} = 0.9371$		

Puede verificarse este resultado en R, con la función anova sobre el modelo predictivo (guardado en el objeto R `modelo2`),

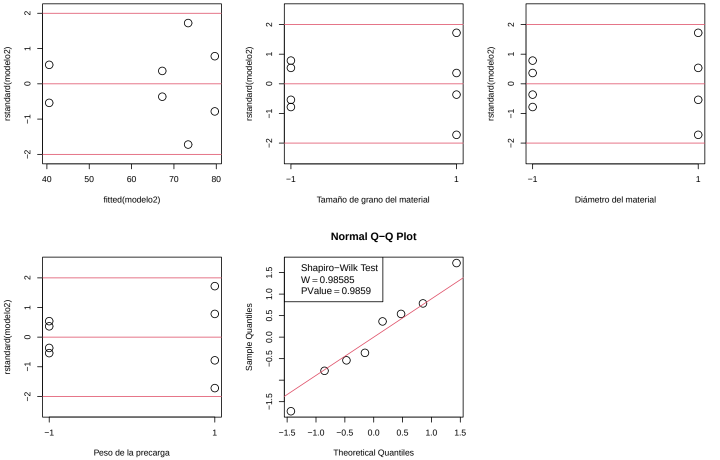
```
> anova(modelo2)
Analysis of Variance Table
Response: Y
          Df  Sum Sq Mean Sq F value   Pr(>F)
A          1  208.08  208.08  12.584 0.023854 *
C          1  544.50  544.50  32.930 0.004569 **
E          1 1021.52 1021.52  61.779 0.001416 **
Residuals  4   66.14   16.53
---
```

Aunque el SSE apenas tiene 4 grados de libertad, en este caso parece que el MSE está razonablemente estimado dado el valor del R^2_{adj}. El objetivo del estudio es identificar la combinación de los niveles de los factores donde ocurre la menor vibración de la máquina de corte. En la siguiente figura se muestran los gráficos de medias de la vibración para los efectos activos y para cada uno se localiza el valor más bajo de la respectiva línea: Tamaño de grano del material (A) en nivel bajo: 80/pulg, diámetro del material (C) en nivel alto: 1.5 pulg. y peso de precarga (E) en nivel bajo: 1 lb. Los niveles de los demás factores se eligen con criterios de economía (nivel más barato), de modo que en la ec. ajustada del modelo predictivo se obtiene como respuesta mínima predicha en los tratamientos considerados el siguiente valor (se trabaja con niveles codificados):

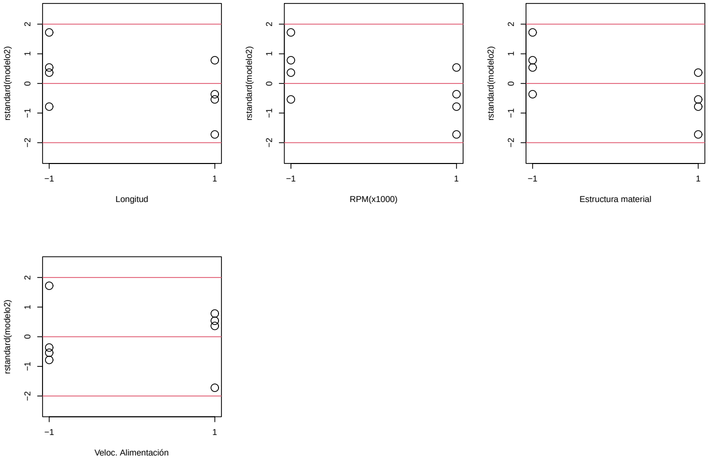
$$\widehat{Y} = 65.2 + 5.1 \times (-1) - 8.25 \times 1 + 11.3 \times (-1) = 40.55$$



Como en cualquier modelo lineal, se analizan residuos vs. respuesta ajustada y vs. factores presentes en la ecuación del modelo predictivo, para valorar si hay evidencia fuerte en contra de que los errores son de media cero y de varianza constante, adicionalmente se evalúa normalidad y se chequea que no haya patrón de carencia de ajuste. Vemos menor dispersión en los niveles bajo de los factores A: Tamaño de grano del material y E: Peso de la precarga pero mayor dispersión en el nivel alto de C: Diámetro del material. Con respecto al supuesto de normalidad, no hay evidencia fuerte en contra.



Sin embargo, también es conveniente chequear residuos vs. los otros factores eliminados. Observe que hay un patrón de tendencia lineal entre el promedio de los residuos en nivel bajo y alto de cada factor excluido, pero no podemos incluir estos factores dado que no hay grados de libertad suficientes para dejar a todos los siete efectos principales en el MRLM.



Dean, A., Voss, D., and Draguljić, D. (2017). *Design and Analysis of Experiments*, 2^{nd} Edition. Springer.

Gutiérrez Pulido, H. y de la Vara Salazar, R. (2012). *Análisis y Diseño de Experimentos*, 3^{a} Edición. McGraw-Hill.

Kuehl, R. O. (2001). *Diseño de Experimentos. Principios Estadísticos de Diseño y Análisis de Investigación*, 2^{a} Edición. Thomson Learning.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*, 5^{th} Edition. McGraw-Hill Irwin.

Montgomery, D. C. (2020). *Design and Analysis of Experiments*, 10^{th} Edition. John Wiley & Sons, Inc.