

CAPITULO VII

DISEÑO DE EXPERIMENTOS: DISEÑOS CON DOS O MÁS FACTORES

En la sección anterior se investigó la diferencia entre k tratamientos (o niveles) de un solo factor, (ej. resistencia de una viga). Ahora por ejemplo, un ingeniero está interesado en los efectos de la temperatura y el tiempo del ciclo en la brillantez del teñido de telas producidas en una fábrica de tejidos. Podría estudiar el proceso, a dos niveles de temperatura, (350 y 400 °F); y dos niveles de tiempo de ciclo (40 y 50). La experimentación en este caso debe incluir 4 combinaciones de factores de nivel. Para este propósito, y en los que intervienen dos o más factores, los diseños factoriales son los más eficientes.

Al inicio de un trabajo de optimización, cuando todavía no se tiene un buen conocimiento del comportamiento del proceso a optimizar, generalmente la lista de factores o variables que pueden influir en el proceso son muchas. El objetivo principal en una primera etapa de optimización es la de identificar aquellas variables que tengan gran influencia en el proceso. Para esta fase, los diseños factoriales son los más recomendados.

7.1. Diseño factorial 2^k

Un diseño factorial es aquel en el que se investigan todas las posibles combinaciones de los niveles de los factores en cada ensayo completo o réplica del experimento. Los niveles representan los valores que pueden tomar las variables o factores. Si consideramos por ejemplo 2 niveles, el diseño se denomina diseño factorial a dos niveles o diseño factorial 2^k . El número total de experimentos a llevarse a cabo viene definido por la relación:

$$N = 2^k$$

donde:

k = número de variables;

N = número total de experimentos.

Con el siguiente ejemplo se ilustra adecuadamente el desarrollo de un diseño factorial.

7.2. Ejemplo a evaluar mediante Diseño Factorial

Una empresa confecciona productos metálicos niquelados. Sus compradores no están completamente satisfechos con las características del niquelado en las piezas metálicas. En la empresa se constituye un equipo de evaluación que identifica el problema como incapacidad de lograr la capa de níquel con el espesor adecuado, siendo esta dificultad el resultado de las condiciones del proceso productivo.

Se desarrollan en la empresa estudios de experimentación para incrementar la calidad del producto niquelado. En las evaluaciones preliminares, quedó plenamente comprobado que el espesor de la capa de níquel es la causa de los reclamos de los compradores. El espesor de la capa de níquel se evalúa utilizando un perfilómetro, instrumento que proporciona mediciones exactas y estables del espesor de la capa de níquel.

7.2.1 Selección de factores

A pesar de que son muchos los factores que intervienen en el proceso de niquelado, tales como temperatura de la solución, tiempo de recubrimiento, pH, contenido de fósforo, contenido de níquel en solución, voltaje, densidad de corriente, etc.; para los propósitos de este ejemplo, se selecciona las variables de tiempo y temperatura del baño para la evaluación experimental. Esto hecho no invalida la posibilidad de utilizar la cantidad de variables que se estime necesario. Por lo tanto, las otras variables del proceso se mantendrán constantes durante la experimentación.

Tanto la temperatura como el tiempo de inmersión en el baño son variables cuantificables, cuyos valores se fijan a **dos** niveles según lo siguiente:

Factor	Nivel (-)	Nivel (+)
Temperatura, °C	16	32
Tiempo, seg.	4	12

Adicionalmente, por experiencias y conocimiento del proceso, se espera que ocurra una fuerte interacción entre estas dos variables; esto es, que las variaciones ocurridas en una de las variables influyan en la otra y estos efectos interactuados se reflejen en el espesor del níquel.

7.2.2 Tratamiento de datos.

Con dos factores fijados a dos niveles, (+ y -), se decide utilizar un diseño factorial completo, en donde $N = 2^k = 2^2 = 4$ experimentos. Los valores de las variables a experimentar se *codifican* con valores +1 y -1, como se indica en la Tabla 7.1.

Tabla 7.1: Matriz de diseño con valores actuales y codificados para pruebas de niquelado.

Prueba #	Valores actuales		Valores codificados	
	A Temp. °C	B Tiempo, seg.	A	B
1	16	4	-1	-1
2	16	12	-1	+1
3	32	4	+1	-1
4	32	12	+1	+1

La sección de la Tabla con valores actuales sirve de guía para conducir los experimentos, y la sección con los valores codificados para los propósitos de cálculos y evaluación de datos.

En general, los experimentos deben ejecutarse lo más aleatoriamente posible, sin seguir necesariamente el orden en el que la Tabla de diseño dicta. En este caso particular, la temperatura es una variable que no facilita la aleatorización, pues es difícil modificarla rápidamente a los valores en que se necesita experimentar. En este caso, debe tomarse medidas para evitar inducir variabilidad diferente a cada prueba.

Con los recursos que se dispone, se decide efectuar 5 pruebas con cada una de las 4 combinaciones, midiendo el espesor de la capa de níquel como la variable respuesta. En lo posible, estas repeticiones se efectúan aleatoriamente; es decir, no repitiendo la misma combinación 5 veces continuadas. Los resultados obtenidos por la experimentación se resumen en la Tabla 7.2.

Tabla 7.2: Datos de evaluación del proceso de niquelado.

Prueb a #	A Temp. °C	B Tiempo Seg.	Espesor, micropulg.					Promedio
1	16 (-1)	4 (-1)	116.1	116.9	112.6	118.7	114.9	115.8
2	16 (-1)	12 (+1)	116.5	115.5	119.2	114.7	118.3	116.8
3	32 (+1)	4 (-1)	106.7	107.5	105.2	107.1	106.5	106.7
4	32 (+1)	12 (+1)	123.2	125.1	124.5	124.0	124.7	124.3

De las observaciones de los resultados resulta evidente que la mejor combinación de variables corresponde a 32 °C y 12 seg. la que proporciona el mejor espesor de níquel. Sin embargo, una "elección a ganador" por simple inspección de los resultados es una actitud altamente negativa. Ello no permite evaluar cual de las variables tiene mayor efecto y trunca, tanto posibles sub-optimizaciones, como predecir la respuesta de valores no experimentados con las variables.

El procedimiento de evaluación se continua determinando el promedio de los valores altos y el promedio de los valores bajos para cada variable. Luego se calcula el efecto Δ , que representa la diferencia entre los valores altos y bajos:

Promedio	A Temp. °C	B Tiempo Seg.	
(-)	116.3	111.3	Promedio de espesor para valores bajos
(+)	115.5	120.6	Promedio de espesor para valores altos
	-0.8	+9.3	Efecto Δ

Promedio general = 115.93

De esos cálculos se observa el efecto del factor temperatura equivalente a -0.8 micropulgadas. Esto significa que el aumento la temperatura de 16 a 32 °C ocasiona una disminución en el espesor en 0.8 micropulgadas. Similarmente, el factor efecto tiempo es de +9.3; es decir, el aumento de tiempo de residencia dentro del baño de 4 a 12 seg. ocasiona un aumento en el espesor en 9.3 micropulgadas. Comparando efectos, el factor tiempo resulta ser aproximadamente 12 veces más importante que el factor temperatura en el espesor de la capa de níquel.

A continuación se elabora gráficos con los valores promedios obtenidos, como se observa en las Figuras 7.1. En estas se debe destacar el hecho de la interacción que existe entre los dos factores en estudio. Esto se evidencia por el hecho que en ambos gráficos no existe paralelismo entre las rectas trazadas.

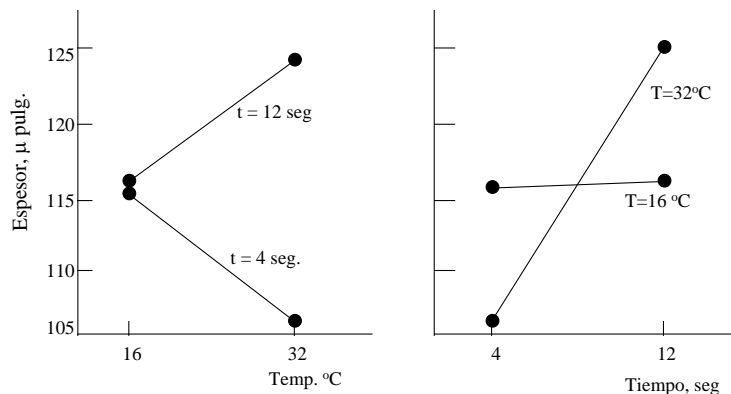


Figura 7.1: Gráfico de interacción para temperatura y tiempo en el proceso de níquelado.

A fin de cuantificar los efectos de interacción, ya develados con los gráficos de la Figura 7.1, se elabora una Tabla con las variables codificadas y adicionando una columna de interacción *AB*. Esta columna se obtiene multiplicando los valores codificados correspondientes a los dos factores. Esto se observa en la Tabla 7.3.

Tabla 7.3: Matriz codificada con columna de interacción

Prueba #	A	B	AB	Promedio
1	-1	-1	+1	115.8
2	-1	+1	-1	116.8
3	+1	-1	-1	106.7
4	+1	+1	+1	124.3

Se puede determinar matemáticamente el efecto de interacción entre los dos factores, utilizando la columna *AB* y la respuesta obtenida para cada combinación:

Valores	Interacción AB	
(-)	$\frac{116.8+106.7}{2}=111.75$	Promedio de interacción para valores bajos
(+)	$\frac{115.8+124.3}{2}=120.05$	Promedio de interacción para valores altos
	$120.05 - 111.75 = 8.3$	Efecto Δ

El alto valor del efecto Δ para la interacción AB , indica la fuerte interacción que existe entre esas dos factores. El siguiente paso es calcular los "medios efectos" para cada columna. Estos "medios efectos" serán usados en la ecuación de predicción a definirse posteriormente.

	A (Temp.)	B (tiempo)	AB
$\Delta/2$	$-0.8/2 = -0.4$	$9.3/2 = 4.65$	$8.3/2 = 4.15$

En las Figuras, 7.2 y 7.3, se resumen los datos obtenidos hasta el momento.

La variable B (tiempo) y la interacción AB muestran tener gran influencia en el proceso de niquelado, en cambio, la variable A (temperatura) no tiene mucha significación en el proceso.

A continuación se debe evaluar cuales de los factores e interacciones son importantes. Para ello se dispone de tres procedimientos:

- Análisis gráfico;
- Análisis de la Varianza, ANAVA;
- Análisis de regresión.

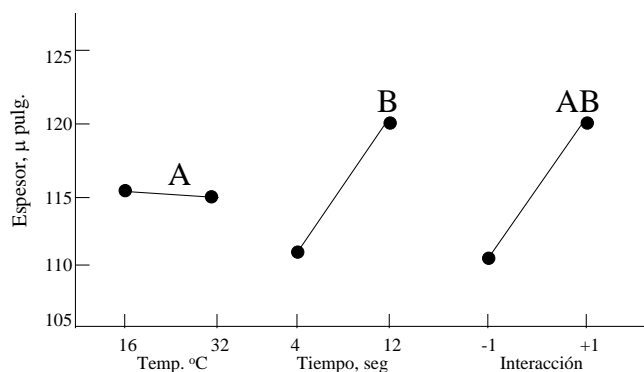


Figura 7.2: Promedios para cada efecto.

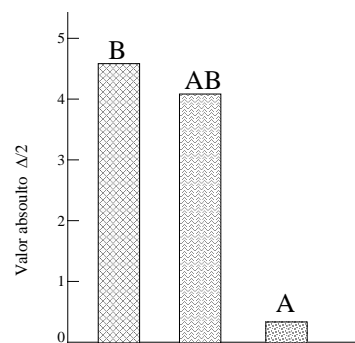


Figura 7.3: Diagrama PARETO para valores $\Delta/2$ o medios efectos.

7.2.3 Análisis gráfico

Este análisis se basa en la simple inspección de los gráficos desarrollados luego del análisis de datos efectuados. Mediante este procedimiento se hace el análisis y selección de los factores más importantes y se usaran para definir la ecuación de predicción. El diagrama de Pareto muestra que factores tienen mayor importancia; sin embargo, un profundo conocimiento técnico del proceso en estudio puede ser gravitante para predecir la importancia de un factor.

La forma general de la ecuación de predicción es:

$$\hat{Y} = \bar{Y} + \left(\frac{\Delta_A}{2}\right)A + \left(\frac{\Delta_B}{2}\right)B + \left(\frac{\Delta_{AB}}{2}\right)AB + \dots$$

donde:

\hat{Y} = la respuesta de predicción o Y estimado

\bar{Y} = promedio de los valores de respuesta.

$\Delta A/2$ = medio efecto para el factor A

$\Delta B/2$ = medio efecto para el factor B

$\Delta AB/2$ = medio efecto para el factor AB

Consecuentemente, de la observación de los resultados gráficos se concluye que los factores B (tiempo) e interacción $A-B$ (tiempo - temperatura) tienen importancia significativa en el proceso de niquelado. Sin embargo, el modelo matemático incluirá el factor temperatura a pesar que no muestra incidencia significativa en el proceso. Esto en razón a que técnicamente se reconoce que la temperatura es importante en el proceso de niquelado. Con estas consideraciones, el modelo matemático queda definido de la siguiente manera:

$$\hat{Y} = 115.93 - 0.4 A + 4.65 B + 4.15 AB$$

7.2.4 Verificación del modelo

A continuación se prueba el modelo con valores ya experimentados. Así por ejemplo, si consideramos $A = 32^\circ\text{C}$ y $B = 4$ seg., entonces el modelo responde:

$$\hat{Y} = 115.93 - 0.4 (+1) + 4.65 (-1) + 4.15 (+1)(-1) = 106.73$$

Ahora, si $A = 32^\circ\text{C}$ y $B = 12$ seg., entonces:

$$\hat{Y} = 115.93 - 0.4 (+1) + 4.65 (+1) + 4.15 (+1)(+1) = 124.33$$

Nótese que los valores actuales de las variables son reemplazados en la ecuación de predicción con los valores codificados.

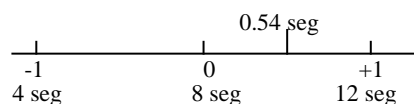
Se observa que la ecuación de predicción responde correctamente a combinaciones de valores ya experimentados. Esto da seguridad para que el modelo pueda ser utilizado para predecir combinaciones no probadas y/o predecir un valor objetivo.

Supóngase ahora que se fija como objetivo el tener una capa de níquel de 120 micropulgadas. Ya que se dispone de 2 variables y se decide técnicamente que la temperatura debe fijarse a 32°C , la ecuación puede usarse según:

$$120 = 115.93 - 0.4 (+1) + 4.65 (B) + 4.15 (+1)(B)$$

$$B = 0.508$$

Ya que este valor de B es un valor codificado del tiempo, para obtener el valor real se procede según lo siguiente:

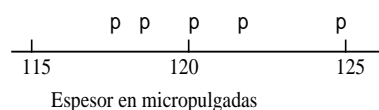


El tiempo real se obtendrá de un simple cálculo de interpolación. Para este valor codificado de $B = 0.508$ se obtiene un tiempo de 10.04 seg. Por lo tanto, con 10.04 seg. y 32°C , el niquelado deberá tener una capa de 120 micropulgadas de espesor.

Se debe proceder ahora con las pruebas confirmativas para determinar la validez total del modelo obtenido. Las pruebas de confirmación se llevarán a cabo con 32°C de temperatura del baño y 10 seg. de tiempo de inmersión de la pieza en el baño de níquel.

Con esos valores fijados de los factores, se realizan 5 pruebas de verificación y se obtienen los siguientes resultados de micropulgadas de espesor de la capa de níquel:

120.5; 122.7; 117.4; 125.3; 117.0.



Ya que hay una distribución razonable alrededor de 120, se concluye que el modelo obtenido es válido.

7.3 Análisis de la varianza

El análisis gráfico, no-estadístico, se complementa con el estudio de análisis de la varianza, lo cual aporta el criterio objetivo para completar la evaluación. Este análisis se desarrolla con un ejemplo.

Se realiza una investigación sobre el efecto de dos factores en la velocidad de reacción de un proceso, sean estas la concentración de un reactivo (Z_1) y la cantidad de un catalizador, (Z_2). Los niveles de los factores son los siguientes:

Factores	Nivel (-)	Nivel (+)
Z_1 : Reactivo (%)	15	20
Z_2 : Catalizador (sacos)	1	2

Con estos valores se ejecuta un diseño experimental con $2^2 = 4$ experimentos y tres replicas en cada prueba. La Tabla 7.4 señal el diseño a utilizar.

Tabla 7.4: Matriz de diseño experimental para 2 factores

N	Factores actuales		Factores codificados		
	Z_1	Z_2	X_1	X_2	X_1X_2
1	15	1	-1	-1	+1
2	20	1	+1	-1	-1
3	15	2	-1	+1	-1
4	20	2	+1	+1	+1

Se realizan los experimentos aleatoriamente con réplicas también aleatorizadas, y los resultados se muestran en la siguiente Tabla.

Tabla 7.5: Matriz de diseño experimental para 2 factores con 3 réplicas.

N	Factores			Respuestas			Total respuestas	\bar{Y}_r
	X_1	X_2	X_1X_2	I	II	III		
1	-1	-1	+1	28	25	27	80	26.67
2	+1	-1	-1	36	32	32	100	33.33
3	-1	+1	-1	18	19	23	60	20.00
4	+1	+1	+1	31	30	29	90	30.00

Promedio general $\bar{Y} = 27.50$

Los promedios de bajo y alto nivel para los factores e interacción y el efecto Δ se muestran a continuación:

Promedio	X_1	X_2	X_1X_2
(-)	23.335	30	26.665
(+)	31.665	25	28.335
Efecto Δ	8.33	-5	1.67

El análisis de la varianza se inicia planteando la siguiente ecuación que representa a la variabilidad total de los datos en sus partes componentes:

$$SS_{Total} = SS_{Efectos} + SS_{Error}$$

donde

SS_{Total} = Suma total de cuadrados.

$SS_{Efectos}$ = Suma de cuadrados debida a los tratamientos.

SS_{Error} = Suma de cuadrados debida al error.

Para el análisis de la varianza de un **diseño factorial donde con varias réplicas** en los puntos del diseño, la suma total de cuadrados viene definido por:

$$SS_{Total} = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^r Y_{ijk}^2 - \frac{Y^2}{Nr} \quad \text{donde:}$$

N = número de pruebas experimentales;

r = número de réplicas en el diseño.

O sea, SS_{Total} es igual a la suma de cada una de las respuestas elevadas al cuadrado, menos la suma de todas las respuestas elevada al cuadrado dividida el número total de experimentos por el número de réplicas efectuadas

En el ejemplo:

$$\begin{aligned} SS_{Total} &= (28^2 + 25^2 + 27^2 + \dots + 30^2 + 29^2) - \frac{(28 + 25 + 27 + \dots + 30 + 29)^2}{4 * 3} \\ &= 9398 - \frac{(330)^2}{12} \\ &= 323 \end{aligned}$$

La suma de cuadrados de los efectos e interacciones se da por la siguiente relación:

$$SS_{Efecto} = \frac{\left(\sum_{i=1}^N X_{ij} Y_j \right)^2}{Nr}$$

En el ejemplo:

$$SS_{X_1} = \frac{[(100 + 90) - (80 + 60)]^2}{4 * 3} = 208.33$$

$$SS_{X_2} = \frac{[(90 + 60) - (80 + 100)]^2}{4 * 3} = 75.00$$

$$SS_{X_1 X_2} = \frac{[(80 + 90) - (100 + 60)]^2}{4 * 3} = 8.33$$

$$SS_{Efectos} = SS_{X_1} + SS_{X_2} + SS_{X_1 X_2} = 208.33 + 75.00 + 8.33 = 291.66$$

De la ecuación $SS_{Total} = SS_{Efectos} + SS_{Error}$, se obtiene SS_{Error} : de la siguiente forma:

$$SS_{Error} = SS_{Total} - SS_{Efectos} = 323 - 291.66 = 31.34$$

A través del cálculo del *efecto A* se observa que la variable X_1 , (concentración del reactivo) tiene incidencia importante en el proceso. La manera precisa y estadística de medir la importancia de esa variable es por el Teorema de Cochran, que se resume en la siguiente expresión:

$$F_o = \frac{MS_{Efectos}}{MS_{Error}}$$

donde:

$$MS_{Efectos} = SS_{Efectos} / gl_T$$

$$MS_{Error} = SS_{Error} / gl_E$$

de aquí:

gl_T = grados de libertad de los efectos e interacciones. Igual a 1 en los diseños factoriales a dos niveles ($2-1=1$)

gl_E = grados de libertad de la suma de cuadrados del error. Igual a:

$$\sum_{i=1}^{n_r} n_{r_i} - 1$$

Esto es, la sumatoria del número de réplicas menos 1

Una variable o una interacción es significativa si se cumple la siguiente condición:

$$F_O > F(\alpha; gl_T, gl_E)$$

En el ejemplo:

$$gl_T = 2 - 1 = 1$$

$$gl_E = (3-1) + (3-1) + (3-1) + (3-1) = 8$$

$$MS_{Efectos(X1)} = 208.33/1 = 208.33$$

$$MS_{Efectos(X2)} = 75/1 = 75$$

$$MS_{Efectos(X1X2)} = 8.33/1 = 8.33$$

$$MS_{Error} = 31.34/8 = 3.92$$

Con todos estos datos se construye la Tabla de Análisis de la Varianza. Tabla 7.6

Tabla 7.6: Análisis de la Varianza para el ejemplo.

Fuente	SS	g.l.	MS	F_O
X_1	208.33	1	208.33	53.15
X_2	75.00	1	75.00	19.13
X_1X_2	8.33	1	8.33	2.13
Error	31.34	8	3.92	
Total	323.00			

Para un nivel de significación $\alpha = 0.01$; $gl_T = 1$; $gl_E = 8$ entonces $F(0.01; 1, 8) = 11.26$

La condición $F_O > F(\alpha; gl_T, gl_E)$ se cumple para X_1 y X_2 . Por lo tanto, solo las variables X_1 y X_2 tienen incidencia significativa en el proceso. Con esto se formula el modelo matemático y es como sigue:

$$\hat{Y} = 27.50 + \frac{8.33}{2} X_1 - \frac{5}{2} X_2$$

$$\hat{Y} = 27.5 + 4.165 X_1 - 2.5 X_2$$

Nótese que los valores a sustituir para las variables corresponde a los valores codificados +1 y -1

Obtenido este modelo, debe ser verificado experimentalmente para comprobar su validez.

7.3.1 Análisis de residuos

Para saber cuan distanciados están los valores que se predicen con el modelo de los valores experimentales, se efectúa el análisis de residuos. Este análisis considera los residuales igual a:

$$Y - \hat{Y}$$

y que se resumen en la Tabla 7.7.

Tabla 7.7: Análisis de residuos para el ejemplo.

			Y_r					$R = (Y_r - \hat{Y})$		
N	X_1	X_2	Y_1	Y_2	Y_3	\bar{Y}	\hat{Y}	R_{Y1}	R_{Y2}	R_{Y3}
1	-1	-1	28	25	27	26.67	25.835	+2.165	-0.835	+1.165
2	+1	-1	36	32	32	33.33	34.165	+1.835	-2.165	-2.165
3	-1	+1	18	19	23	20.00	20.835	-2.835	-1.835	+2.165
4	+1	+1	31	30	29	30.00	29.165	+1.835	+0.835	-0.165

Con estos datos se calcula la suma de cuadrados residual del modelo, SSM_R

$$SSM_R = \sum_{i=1}^{N_r} \frac{(Y - \hat{Y})^2}{Nr - I}$$

donde:

\hat{Y} = Y estimado o respuesta según el modelo.

Y = Y observado o respuesta experimental

Nr = número total de experimentos.

I = número de parámetros del modelo matemático.

Nr-I = grados de libertad del residuo = g/R

En el ejemplo:

$$SSM_R = \frac{(2.165)^2 + (-0.835)^2 + \dots + (-0.165)^2}{(4)(3) - (3)} = \frac{39.67}{9} = 4.41$$

Para determinar estadísticamente si el modelo matemático hallado representa adecuadamente a los datos experimentales, se realiza la prueba F:

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{4.41}{3.92} = 1.124$$

El modelo es adecuado si:

$$F_o < F(\alpha; g/R, g/E)$$

Si en el ejemplo: $\alpha = 0.05$; $g/R = 9$; $g/E = 8$

$$F(0.05; 9, 8) = 3.39$$

Entonces el modelo matemático es adecuado.

7.3.2 Decodificación del modelo a escala natural

En el ejemplo, el modelo codificado obtenido fue:

$$\hat{Y} = 27.5 + 4.165 X_1 - 2.5 X_2$$

es decir, en este modelo los valores de las variables X_1 y X_2 se representan por valores (+1) y (-1). Para que en el modelo se pueda reemplazar los valores reales de las variables, se calcula lo siguiente:

Centro del diseño

$$Z_{x1}^o = \frac{(20+15)}{2} = 17.5 \quad Z_{x2}^o = \frac{(2+1)}{2} = 1.5$$

Radio del Diseño

$$\Delta Z_{X_1} = \frac{(20-15)}{2} = 2.5 \quad \Delta Z_{X_2} = \frac{(2-1)}{2} = 0.5$$

Relación E

$$E_{X_1} = \frac{Z_{X_1}^o}{\Delta Z_{X_1}} = \frac{17.5}{2.5} = 7 \quad E_{X_2} = \frac{Z_{X_2}^o}{\Delta Z_{X_2}} = \frac{1.5}{0.5} = 3$$

lo cual se resume en el siguiente cuadro:

	X_1	X_2
Nivel Inferior (-)	15	1
Nivel Superior (+)	20	2
Centro del diseño, (Z^o)	17.5	1.5
Radio del Diseño (ΔZ)	2.5	0.5
Relación E	7	3

El modelo decodificado se calcula según:

$$\hat{Y}_D = a_o + a_1 X_1 + a_2 X_2$$

y se obtiene:

$$a_o = \bar{Y} - \left(\frac{\Delta_{X_1}}{2} \right) E_{X_1} - \left(\frac{\Delta_{X_2}}{2} \right) E_{X_2}$$

$$= 27.5 - (4.165)(7) - (-2.5)(3) = 5.845$$

$$a_1 = \frac{\left(\frac{\Delta_{X_1}}{2} \right)}{\Delta Z_{X_1}} = \frac{4.165}{2.5} = 1.666$$

$$a_2 = \frac{\left(\frac{\Delta_{X_2}}{2} \right)}{\Delta Z_{X_2}} = \frac{-2.5}{0.5} = -5.0$$

Con lo que el modelo decodificado se expresa:

$$\hat{Y}_D = 5.845 + 1.666 Z_1 - 5.0 Z_2$$

En términos generales:

$$a_o = \bar{Y} - \sum_{j=1}^k \left(\frac{\Delta_j}{2} \right) E_j - \sum_{ij=1}^k \left(\frac{\Delta_{ij}}{2} \right) E_{ij}; \quad i \neq j$$

$$a_j = \frac{\left(\frac{\Delta_j}{2} \right)}{\Delta Z_j} - \sum_{ij=1}^{k-1} \frac{\left(\frac{\Delta_{ij}}{2} \right)}{\Delta Z_j} \quad ; \quad i \neq j$$

$$a_{ij} = \frac{\left(\frac{\Delta_{ij}}{2} \right)}{(\Delta Z_i \Delta Z_j)}$$

7.4. Diseños factoriales fraccionados a dos niveles

Los diseños factoriales completos son buenos en el sentido que proveen estimados para los efectos de todos los factores, así como los efectos de posibles interacciones. Una desventaja es que estos diseños requieren mayores cantidades de experimentos cuando más de cuatro factores se necesitan evaluar. Por ejemplo, no es inusual necesitar desarrollar un experimento con 10 factores a dos niveles. Un diseño factorial completo requeriría $2^{10} = 1,024$ experimentos, lo cual son demasiados para ejecutarse.

Como una alternativa a la posibilidad de tener gran cantidad de variables, se presentan los diseños factoriales fraccionados, en los que se ejecuta solo una fracción del total de experimentos que se

requieren. Lo negativo del método es que parte de la información sobre interacciones se llega a perder.

Supóngase que se tiene tres factores, A, B, y C y se les quiere evaluar mediante un diseño factorial. Para esto se tiene dos alternativas:

- Un diseño factorial completo con 8 experimentos;
- Un diseño factorial fraccionado con 4 experimentos.

Estas dos alternativas se presentan en las Tablas 7.8 y 7.9. En esas Tablas, AB, AC, BC y ABC representan la interacción entre los tres factores.

Obsérvese que en la Tabla 7.9 la columna A es igual a la columna BC; B = AC; C = AB. Ya que los valores de tales columnas son iguales, no es posible determinar si el efecto calculado de la columna A es debido al factor A o debido a la interacción BC. Usualmente las interacciones son relativamente pequeñas comparadas con los efectos de los factores, lo que en parte justifica el uso de los diseños fraccionados.

Tabla 7.8: Diseño factorial completo con tres factores a dos niveles, $N = 2^3 = 8$

N	A	B	C	AB	AC	BC	ABC
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

Tabla 7.9: Diseño factorial fraccionado con tres factores a dos niveles, en 4 pruebas.

N	A	B	C=A*B	AB	AC	BC	ABC
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

Consideremos ahora por ejemplo que se tiene cuatro factores A, B, C, y D; a evaluar un diseño factorial fraccionado con 8 pruebas. Un diseño completo incluiría $N = 2^4 = 16$ pruebas. Para elaborar la matriz codificada de diseño, obsérvese la Tabla 7.8 con $2^3 = 8$. Es obvio en esta Tabla que la columna de interacción ABC es poco probable de producirse con un efecto significativo; de modo que para 4 factores la Tabla 7.8 puede reescribirse como se observa en la Tabla 7.10, considerando las respectivas interacciones.

Tabla 7.10: Diseño factorial fraccionado con cuatro factores a dos niveles, 8 pruebas.

N	A	B	C	D=ABC	AB=CD	AC=BD	BC=AD
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

En general, un diseño factorial fraccionado está definido por: 2^{k-p} donde:

k = número de factores a estudiar ;

p = número de factores a confundir.

7.5. Diseño factorial 2^k con réplica en el punto central del diseño

El diseño factorial 2^k asume una relación lineal entre las variables y la respuesta. De ser cierta la linealidad, el modelo debe predecir adecuadamente todos los puntos del diseño, inclusive el punto central. Aunque no se verifique linealidad perfecta, el modelo funcionará bastante bien si la linealidad se cumple de manera aproximada. De no predecir adecuadamente el modelo, especialmente en el punto central, se afirma que el modelo no es suficiente para explicar las respuestas en tal región y es posible asumir la existencia de curvatura. Se debe realizar pruebas en el punto central del diseño para dar un estimado del efecto de curvatura.

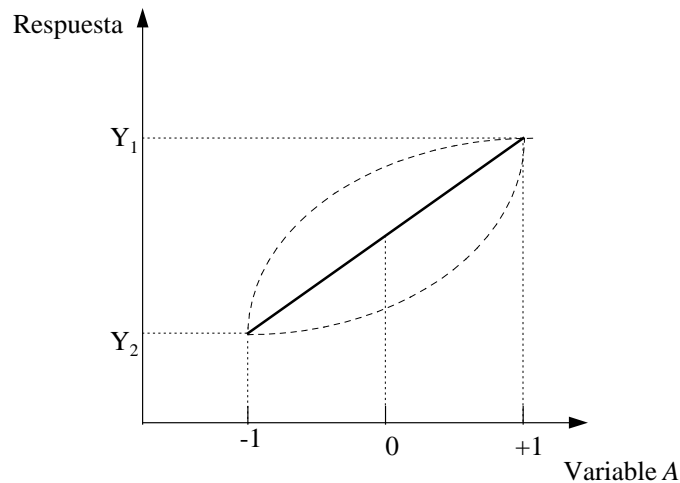


Figura 7.4: Modelo lineal fijado a nivel bajo y alto, con probable curvatura en el centro.

7.5.1 Ejemplo de aplicación

Un ingeniero químico se encuentra estudiando el rendimiento de un proceso. Existen dos variables de interés: tiempo y temperatura de reacción. Debido a que el ingeniero duda de la suposición de linealidad, decide utilizar un diseño 2^2 , con una sola prueba en cada punto externo del diseño y 5 réplicas en el punto central.

El diseño y los resultados se muestran en la Figura 7.5.

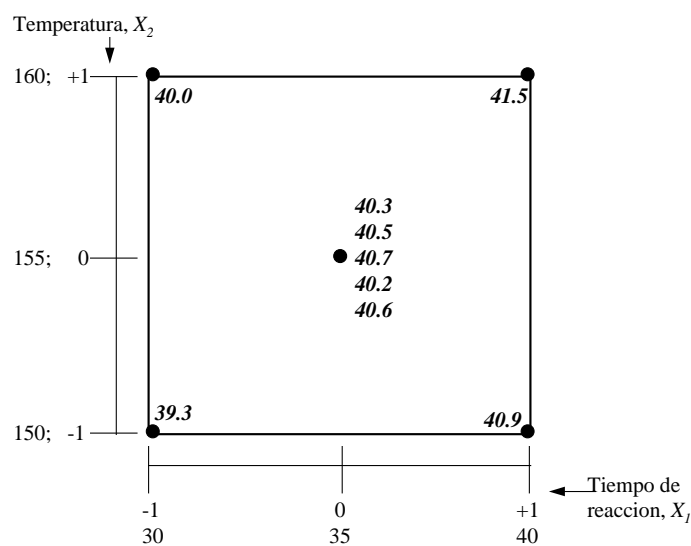


Figura 7.5: Diseño factorial con réplica en el centro del diseño.

La Tabla 7.11 muestra los valores actuales y codificados del diseño.

Tabla 7.11: Matriz del diseño factorial con réplicas en el centro y respuestas experimentales.

Prueba #	Z_1 Tiempo	Z_2 Temper.	X_1	X_2	X_1X_2	Respuesta Y
1	30	150	-1	-1	+1	39.3
2	40	150	+1	-1	-1	40.9
3	30	160	-1	+1	-1	40.0
4	40	160	+1	+1	+1	41.5
5	35	155	0	0		40.3
5	35	155	0	0		40.5
5	35	155	0	0		40.7
5	35	155	0	0		40.2
5	35	155	0	0		40.6

Promedio general (exceptúa los puntos en el centro) = 40.425

Los promedios de bajo y alto nivel para los factores e interacción y el efecto Δ se muestran a continuación:

Promedio	X_1	X_2	X_1X_2
(-)	39.65	40.1	40.45
(+)	41.2	40.75	40.40
Efecto Δ	1.55	0.65	-0.05

Obsérvese que los puntos replicados en el centro no son considerados en el cálculo de los efectos; mas bien serán utilizados para hallar la suma de cuadrados del error.

7.5.1 Análisis de la varianza

El análisis de la varianza difiere algo del efectuado cuando se efectúan réplicas en cada punto del diseño. Implica el cálculo de:

$SS_{Efectos}$ = Suma de cuadrados debida a los tratamientos.

SS_{Error} = Suma de cuadrados debida al error.

$SS_{Curvatura}$ = Suma de cuadrados debida a la curvatura.

La suma de cuadrados de los efectos e interacciones se da por la siguiente relación:

$$SS_{Efecto} = \frac{\left(\sum_{i=1}^N X_{ij} Y_j \right)^2}{Nr}$$

La suma de cuadrados debida al error se calcula con la siguiente expresión:

$$SS_{Error} = \sum_{i=1}^{n_o} \left(Y_i^o - \bar{Y}^o \right)^2$$

donde:

Y_i^o = réplicas en el punto central del diseño.

n_o = número de réplicas en el centro.

\bar{Y}^o = promedio de todas las réplicas.

La suma de cuadrados debida a la curvatura se calcula con la siguiente expresión:

$$SS_{Curvatura} = \frac{N n_o (\bar{Y} - \bar{Y}^o)^2}{N + n_o}$$

donde:

\bar{Y} = promedio de puntos exteriores del diseño.

En el ejemplo, la suma de cuadrados debida a los tratamientos es:

$$SS_{X_1} = \frac{[(41.5 + 40.9) - (40.0 + 39.3)]^2}{(4)(1)} = 2.4025$$

$$SS_{X_2} = \frac{[(40.0 + 41.5) - (40.9 + 39.3)]^2}{(4)(1)} = 0.4225$$

$$SS_{X_1 X_2} = \frac{[(39.3 + 41.5) - (40.9 + 40.0)]^2}{(4)(1)} = 0.0025$$

La suma de cuadrados del error:

$$\bar{Y}^o = 40.46$$

$$SS_{Error} = (40.3 - 40.46)^2 + (40.5 - 40.46)^2 + (40.7 - 40.46)^2 + (40.2 - 40.46)^2 + (40.6 - 40.46)^2 = 0.172$$

$$gl_E = (n_o - 1) = 5 - 1 = 4; \quad (\text{grados de libertad del error})$$

La suma de cuadrados debida a la curvatura:

$$SS_{Curvatura} = \frac{(4)(5)(40.425 - 40.46)^2}{4 + 5} = 0.00272$$

Tabla 7.12: Análisis de la Varianza para el ejemplo.

Fuente	SS	g.l.	MS	F_o
X_1	2.4025	1	2.4025	55.87
X_2	0.4225	1	0.4225	9.83
$X_1 X_2$	0.0025	1	0.0025	0.058
Curvatura	0.00272	1	0.00272	0.0063
Error	0.172	4	0.043	
Total	3.00222	8		

Evaluando el estadístico F versus F_o , se tiene:

$$F(\alpha; gl_T, gl_E) = F(0.05; 1, 4) = 7.71$$

Una variable, una interacción ó la curvatura es significativa si se cumple la siguiente condición:

$$F_o > F(\alpha; gl_T, gl_E)$$

Como esta condición se cumple para X_1 y X_2 , el modelo matemático queda definido:

$$\hat{Y} = 40.425 + \frac{1.55}{2} X_1 + \frac{0.65}{2} X_2$$

$$\hat{Y} = 40.425 + 0.775 X_1 + 0.325 X_2$$

7.5.2 Análisis de residuos

N	X ₁	X ₂	Y _r	\hat{Y}	R = (Y _r - \hat{Y})
1	-1	-1	39.3	39.325	-0.025
2	+1	-1	40.9	40.875	0.025
3	-1	+1	40.0	39.975	0.025
4	+1	+1	41.5	41.525	0.025

$$SSM_R = \sum_{i=1}^{N_r} \frac{(Y - \hat{Y})^2}{Nr - I}$$

$$SSM_R = \frac{(-0.025)^2 + \dots + (0.025)^2}{(4)(1) - (3)} = 0.0025$$

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{0.0025}{0.043} = 0.06$$

El modelo es adecuado si: $F_o < F(\alpha; g/R, g/E)$. En el ejemplo: $\alpha = 0.05$; $g/R = 1$; $g/E = 4$

$F(0.05; 1, 4) = 7.71$. Entonces el modelo matemático ajusta adecuadamente los datos experimentales.

7.5.3 Ejemplo de aplicación: efecto de curvatura importante

Al inicio de la etapa de optimización de un proceso de flotación de un mineral sulfurado de cobre, un investigador elige tres variables que, según su criterio puede mejorar la recuperación metálica de cobre. Estas variables son la dosificación del colector, pH y porcentaje de sólidos según:

Factor	Nivel (-)	Nivel (+)
Z ₁ : Dosificación de colector	0.02	0.06
Z ₂ : pH	10	11
Z ₃ : % de sólidos	27.5	33.5

El investigador decide realizar un diseño $2^3 = 8$ experimentos con tres réplicas en el centro. La Tabla 7.13 muestra el diseño y sus resultados.

Tabla 7.13: Matriz del diseño factorial con réplicas en el centro y respuestas.

N	Z ₁	Z ₂	Z ₃	X ₁	X ₂	X ₃	X ₁ X ₂	X ₁ X ₃	X ₂ X ₃	X ₁ X ₂ X ₃	Respuesta Y
1	0.02	10	27.5	-1	-1	-1	+1	+1	+1	-1	94.0
2	0.06	10	27.5	+1	-1	-1	-1	-1	+1	+1	94.0
3	0.02	11	27.5	-1	+1	-1	-1	+1	-1	+1	94.6
4	0.06	11	27.5	+1	+1	-1	+1	-1	-1	-1	92.2
5	0.02	10	33.5	-1	-1	+1	+1	-1	-1	+1	92.5
6	0.06	10	33.5	+1	-1	+1	-1	+1	-1	-1	92.5
7	0.02	11	33.5	-1	+1	+1	-1	-1	+1	-1	93.2
8	0.06	11	33.5	+1	+1	+1	+1	+1	+1	+1	92.1
9	0.04	10.5	30.5	0	0	0					92.5
9	0.04	10.5	30.5	0	0	0					92.4
9	0.04	10.5	30.5	0	0	0					92.4

Promedio general (exceptúa puntos en el centro) $\bar{Y} = 93.1375$

Con los datos obtenidos se calcula el efecto Δ para cada una de las variables e interacciones, como sigue:

	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$
Δ	-0.87	-0.22	-1.12	-0.87	0.32	0.37	.0325

Tabla 7.14: Análisis de la Varianza para el ejemplo.

Fuente	SS	g.l.	MS	F_o	
X_1	1.5313	1	1.5313	459.375	Colector
X_2	0.1013	1	0.1013	30.375	pH
X_3	2.5313	1	2.5313	759.375	% sólidos
X_1X_2	1.5313	1	1.5313	459.375	Colector-pH
X_1X_3	0.2112	1	0.2112	63.375	Colector-%sól.
X_2X_3	0.2813	1	0.2813	84.375	pH-%Sólid.
$X_1X_2X_3$	0.2112	1	0.2112	63.375	Colector-pH-% sól.
Curvatura	1.0819	1	1.0819	324.556	
Error	0.0067	2	0.0033		
Total	7.4873	10			

$$F(\alpha; g|_T, g|_E) = F(0.01; 1, 2) = 98.5$$

$$\hat{Y} = 93.1375 - \frac{0.87}{2}X_1 - \frac{1.12}{2}X_3 - \frac{0.87}{2}X_1X_2$$

$$\hat{Y} = 93.1375 - 0.435X_1 - 0.56X_3 - 0.435X_1X_2$$

Siendo el efecto de la curvatura muy significativo, se concluye que la región óptima se encuentra en el centro del diseño. Con estos resultados, el investigador decide aplicar un diseño de escalamiento. Además, en razón de la significación de la interacción X_1X_2 es de esperar que el pH también tenga significación en el proceso, por lo que se le deberá incluir en la etapa de escalamiento.

Análisis de residuos

N	X_1	X_3	X_1X_2	Y_r	\hat{Y}	$R = (Y_r - \hat{Y})$
1	-1	-1	+1	94.0	93.7	0.3
2	+1	-1	-1	94.0	93.7	0.3
3	-1	-1	-1	94.6	94.575	0.025
4	+1	-1	+1	92.2	92.825	-0.625
5	-1	+1	+1	92.5	92.575	-0.075
6	+1	+1	-1	92.5	92.575	-0.075
7	-1	+1	-1	93.2	93.45	-0.25
8	+1	+1	+1	92.1	91.70	-0.4

$$SSM_R = \sum_{i=1}^{N_r} \frac{(Y - \hat{Y})^2}{Nr - I}$$

$$SSM_R = \frac{0.20125}{8-4} = 0.20125$$

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{0.20125}{0.0033} = 60.985$$

El modelo es adecuado si: $F_o < F(\alpha; g|R, g|E)$. En el ejemplo: $\alpha = 0.01$; $g|R = 4$; $g|E = 2$

$F(0.05; 4, 2) = 99.25$. Entonces el modelo matemático ajusta adecuadamente los datos experimentales.

7.6 Ejemplo de diseño factorial fraccionado

En la investigación para optimizar un proceso de flotación de un mineral sulfurado de cobre, se investiga la influencia de cuatro factores: tiempo de molienda, pH, dosificación del colector y dosificación del espumante. En el siguiente cuadro se muestran los valores a considerar durante el experimento:

Factor	Nivel (-)	Nivel (0)	Nivel (+)
Z_1 : Tiempo de molienda, min	5.0	6.5	8.0
Z_2 : pH	8.5	9.5	10.5
Z_3 : Colector, lb/Ton (Z-200)	0.1	0.15	0.20
Z_4 : Espumante, lb/ton (Dowfroth 200)	0.2	0.30	0.40

Para las 4 variables ($k=4$), se puede utilizar un diseño $2^{4-1} = 8$ experimentos. El diseño se lleva a cabo con tres réplicas en el centro. La variable codificada X_4 se reemplaza por los valores codificados correspondientes a la interacción $X_1X_2X_3$ ya que esta interacción tiene muy pocas probabilidades de ser importante.

Tabla 7.14: Matriz del diseño factorial fraccionado con réplicas en el centro y respuestas.

N	X_1	X_2	X_3	X_4	X_1X_2	X_1X_3	X_1X_4	Z_1 t, min.	Z_2 pH	Z_3 Colect, lb/ton	Z_4 Espum. lb/ton	Y
1	-1	-1	-1	-1	+1	+1	+1	5	8.5	0.1	0.2	87.7
2	+1	-1	-1	+1	-1	-1	+1	8	8.5	0.1	0.4	90.4
3	-1	+1	-1	+1	-1	+1	-1	5	10.5	0.1	0.4	87.5
4	+1	+1	-1	-1	+1	-1	-1	8	10.5	0.1	0.2	92.0
5	-1	-1	+1	+1	+1	-1	-1	5	8.5	0.2	0.4	84.0
6	+1	-1	+1	-1	-1	+1	-1	8	8.5	0.2	0.2	86.4
7	-1	+1	+1	-1	-1	-1	+1	5	10.5	0.2	0.2	85.0
8	+1	+1	+1	+1	+1	+1	+1	8	10.5	0.2	0.4	88.2
9	0	0	0	0				6.5	9.5	0.15	0.3	88.9
9	0	0	0	0				6.5	9.5	0.15	0.3	88.7
9	0	0	0	0				6.5	9.5	0.15	0.3	88.5

Promedio general (exceptúa puntos en el centro) $\bar{Y} = 87.65$

Promedio para pruebas en el centro $\bar{Y}^o = 88.70$

De la Tabla 7.14 se deduce que $X_4 = X_1X_2X_3$, $X_1X_2 = X_3X_4$; $X_1X_3 = X_2X_4$; $X_1X_4 = X_2X_3$

Con los datos obtenidos se calcula el efecto Δ para cada una de las variables e interacciones, como sigue:

	X_1	X_2	X_3	X_4	X_1X_2	X_1X_3	X_1X_4
Δ	3.2	1.05	-3.5	-0.25	0.65	-0.1	0.35

Siguiendo el proceso de cálculo arriba señalado, es decir calculando:

$$SS_{Efecto} = \frac{\left(\sum_{i=1}^N X_{ij} Y_j \right)^2}{Nr}$$

$$SS_{Error} = \sum_{i=1}^{n_o} \left(Y_i^o - \bar{Y}^o \right)^2$$

$$SS_{Curvatura} = \frac{N n_o \left(\bar{Y} - \bar{Y}^o \right)^2}{N + n_o}$$

se construye la Tabla del análisis de la varianza.

Tabla 7.15: Análisis de la Varianza para el ejemplo.

Fuente	SS	g.l.	MS	F_o	
X_1	20.48	1	20.48	512.0	Tiempo
X_2	2.205	1	2.205	55.125	pH
X_3	24.50	1	24.50	612.5	Colector
X_4	0.125	1	0.125	3.125	Espumante
X_1X_2	0.845	1	0.845	21.125	
X_1X_3	0.32	1	0.32	8.0	
X_1X_4	0.245	1	0.245	6.125	
Curvatura	2.405	1	2.405	61.125	
Error	0.08	2	0.04		
Total	51.205	10			

$$F(\alpha; gl_T, gl_E) = F(0.01; 1, 2) = 98.5$$

El modelo matemático es:

$$\hat{Y} = 87.65 + \frac{3.2}{2} X_1 - \frac{3.5}{2} X_3$$

$$\hat{Y} = 87.65 + 1.65 X_1 - 1.75 X_3$$

Análisis de residuos

N	X_1	X_3	Y_r	\hat{Y}	$R = (Y_r - \hat{Y})$
1	-1	-1	87.7	87.8	-0.1
2	+1	-1	90.4	91.0	-0.6
3	-1	-1	87.5	87.8	-0.3
4	+1	-1	92.0	91.0	1.0
5	-1	+1	84.0	84.3	-0.3
6	+1	+1	86.4	87.5	-1.1
7	-1	+1	85.0	84.3	0.7
8	+1	+1	88.2	87.5	0.7

$$SSM_R = \sum_{i=1}^{N_r} \frac{\left(Y - \hat{Y} \right)^2}{Nr - I}$$

$$SSM_R = \frac{3.74}{8-3} = 0.748$$

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{0.0748}{0.04} = 18.70$$

El modelo es adecuado si: $F_0 < F(\alpha; g|R, g|E)$; $F(0.05; 5, 2) = 19.30$. Entonces el modelo matemático ajusta adecuadamente los datos experimentales.

7.7. Diseños de Plackett - Burman

Este diseño desarrollado por Plackett y Burman en 1946 permite estudiar un gran número de variables, reduciéndolas luego a un grupo pequeño para posterior experimentación. Es un tipo especial de diseño factorial fraccionado, donde se puede estudiar k variables en N pruebas experimentales, donde $N > k$ y otorga la posibilidad de experimentar con $N-1$ variables. El modelo contempla el uso de las denominadas *variables ficticias* (no representan a ninguna variable experimental) y se *utilizan para determinar la varianza del error experimental*. Por esta última razón, es conveniente introducir en el diseño por lo menos **2 variables ficticias**.

Para construir un diseño de $P-B$, se toma un vector de tamaño n determinado y con el se generan los restantes vectores. Por ejemplo, si $n = 8$ combinaciones, el vector generador consiste de $n-1 = 7$ valores que son : (+ + + - + - -). La matriz de diseño se completa como sigue:

- Utilizar el vector generador como la primera fila, n_1
- La segunda fila, n_2 , se llena tomando el primer valor de n_1 y colocándolo en el último lugar de la fila n_2 y deslizando luego los demás valores de n_1 delante de ese primer valor.
- La tercera fila, n_3 , se construye utilizando el último valor de la fila n_2 y colocándolo en el último lugar de la fila n_3 y deslizando a partir de ese los restantes valores de n_2 .
- Continuar hasta que todas las $(n-1)$ filas se completen.
- Completar la fila n con valores negativos (-1)

Los vectores generadores para los diversos diseños P-B son como sigue:

$n = 8$	+	+	+	-	+	-	-																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
---------	---	---	---	---	---	---	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Obsérvese que los valores de n son múltiplos de 4 y el total de variables a experimentar (inclusive las *variables ficticias*) es igual a $n-1$.

Supóngase que se quiere experimentar con 4 variables ($k=4$), el número de pruebas es $n = 8$ y el total de variables a considerar es $n-1 = 7$, según la Tabla 7.16. Obsérvese que si son 4 variables reales y 7 el total de variables a considerar, entonces 3 variables serán ficticias.

Tabla 7.16: Diseño P-B con 4 variables reales, 3 variables ficticias y 8 pruebas

N	X_1	X_2	X_3	X_4	F_1	F_2	F_3
1	+	+	+	-	+	-	-
2	+	+	-	+	-	-	+
3	+	-	+	-	-	+	+
4	-	+	-	-	+	+	+
5	+	-	-	+	+	+	-
6	-	-	+	+	+	-	+
7	-	+	+	+	-	+	-
8	-	-	-	-	-	-	-

Ahora supóngase que se desea experimentar con 8 variables. Se escoge un diseño P-B con 12 pruebas y 11 variables. La matriz codificada se da en la Tabla 7.17.

Tabla 7.17: Diseño P-B con 8 variables reales, 3 variables ficticias y 12 pruebas

N	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	F_1	F_2	F_3
1	+	+	-	+	+	+	-	-	-	+	-
2	+	-	+	+	+	-	-	-	+	-	+
3	-	+	+	+	-	-	-	+	-	+	+
4	+	+	+	-	-	-	+	-	+	+	-
5	+	+	-	-	-	+	-	+	+	-	+
6	+	-	-	-	+	-	+	+	-	+	+
7	-	-	-	+	-	+	+	-	+	+	+
8	-	-	+	-	+	+	-	+	+	+	-
9	-	+	-	+	+	-	+	+	+	-	-
10	+	-	+	+	-	+	+	+	-	-	-
11	-	+	+	-	+	+	+	-	-	-	+
12	-	-	-	-	-	-	-	-	-	-	-

7.8.1. Ejemplo de diseño P-B

En el estudio de separación de Molibdeno del concentrado bulk Cu-Mo se parte de las siguientes premisas:

- El NaHS actúa como depresor de la chalcopirita (CuFeS_2). Se investiga el efecto del NaHS en términos de la estabilidad del HS^- , para lo cual se controla el potencial redox de la solución.
- El nitrógeno actúa en la flotación como agente protector que evita la oxidación del NaHS. Se evalúa si el N_2 ejerce algún efecto en la eficiencia de separación.
- Se evalúa también el tiempo de acondicionamiento y el tiempo de flotación.

Se considera los siguientes tres criterios de optimización:

- Eficiencia de separación del Molibdeno, Y_1 .
- Eficiencia de concentración, Y_2 .
- Consumo de NaHS, Y_3 .

En el siguiente cuadro se muestran los valores a considerar durante el experimento:

Factor	Nivel (-)	Nivel (+)
Z ₁ : Potencial redox, mV.	- 400	-200
Z ₂ : Carbón activado, lb/ton	0	2
Z ₃ : Gas	Nitrógeno	Oxígeno
Z ₄ : Tiempo de acondicionamiento, min.	2	5
Z ₅ : Tiempo de flotación, min.	5	10

Tabla 7.18: Diseño P-B con para flotación de Molibdeno.

N	X ₁	X ₂	F ₁	X ₃	X ₄	X ₅	F ₂	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅	Y ₁ .	Y ₂ .	Y ₃ .
1	+	+	+	-	+	-	-	-200	2	N ₂	5	5	19.00	1.48	2.90
2	+	+	-	+	-	-	+	-200	2	O ₂	2	5	2.30	0.15	6.50
3	+	-	+	-	-	+	+	-200	0	N ₂	2	10	10.00	0.65	3.30
4	-	+	-	-	+	+	+	-400	2	N ₂	5	10	84.00	22.97	5.80
5	+	-	-	+	+	+	-	-200	0	O ₂	5	10	15.10	1.77	7.80
6	-	-	+	+	+	-	+	-400	0	O ₂	5	5	39.80	4.98	15.40
7	-	+	+	+	-	+	-	-400	2	O ₂	2	10	74.60	16.53	14.90
8	-	-	-	-	-	-	-	-400	0	N ₂	2	5	45.90	5.13	7.00

7.8.1.1. Evaluación para eficiencia de separación, Y₁.

Con los datos obtenidos se calcula el efecto Δ para cada una de las variables para eficiencia de separación, Y_1 .

	X ₁	X ₂	F ₁	X ₃	X ₄	X ₅	F ₂
Δ	-49.48	17.48	-0.96	-6.70	6.20	19.10	-4.62

Según este cuadro, el factor más importante es el potencial redox, (X_1). El paso de -200 mV. a -400 mV mejora 49.48 % la recuperación.

Siguiendo el proceso de cálculo se determina:

$$SS_{\text{Efecto}} = \frac{\left(\sum_{i=1}^N X_{ij} Y_j \right)^2}{Nr}$$

$$SS_{\text{Error}} = \sum_{i=1}^{n_o} (EF)^2$$

donde:

EF = efecto de las variables ficticias.

$$SS_{\text{Error}} = (-0.96)^2 + (-4.62)^2 = 22.27$$

$$MS_{Error} = \frac{SS_{Error}}{nF}$$

donde:

nF = número de variables ficticias.

$$MS_{Error} = \frac{22.27}{2} = 11.14$$

Con todos esos cálculos que se construye la Tabla del análisis de la varianza.

Tabla 7.19: Análisis de la Varianza para el ejemplo de flotación de Molibdeno.

Fuente	SS	g.l.	MS	F_o	
X_1	4895.55	1	4895.55	439.45	Potencial redox,
X_2	596.85	1	596.85	53.58	Carbón activado
F_1	1.90	1	1.90	0.17	
X_3	91.80	1	91.80	8.26	Gas
X_4	78.75	1	78.75	7.07	Tiempo acondic
X_5	735.36	1	735.36	66.01	Tiempo flotac.
F_2	42.78	1	42.78	3.84	
Error	22.27	2	11.14		
Total	6435.34	9			

$$F(\alpha; gl_T, gl_E) = F(0.05; 1, 2) = 18.51$$

El modelo matemático es:

$$\hat{Y} = 36.34 - 24.74 X_1 + 8.64 X_2 + 9.59 X_5$$

Análisis de residuos

N	X_1	X_2	X_5	Y_r	\hat{Y}	$R = (Y_r - \hat{Y})$
1	+	+	-	19.00	10.65	8.35
2	+	+	-	2.30	10.65	-8.35
3	+	-	+	10.00	12.55	-2.55
4	-	+	+	84.00	79.31	4.70
5	+	-	+	15.10	12.55	2.55
6	-	-	-	39.80	42.85	-3.05
7	-	+	+	74.60	79.31	-4.70
8	-	-	-	45.90	42.05	3.06

$$SSM_R = \sum_{i=1}^{N_r} \frac{(Y - \hat{Y})^2}{Nr - I}$$

$$SSM_R = \frac{215.24}{8 - 4} = 53.81$$

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{53.81}{11.14} = 4.82$$

El modelo es adecuado si: $F_o < F(\alpha; g|R, g|E); F(0.05; 4, 2) = 19.25$. Entonces el modelo matemático ajusta adecuadamente los datos experimentales.

7.8.1.2. Evaluación para eficiencia de concentración del molibdeno, Y_2

Tabla 7.20: Diseño P-B con para flotación de Molibdeno.

N	X ₁	X ₂	F ₁	X ₃	X ₄	X ₅	F ₂	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅	Y ₂
1	+	+	+	-	+	-	-	-200	2	N ₂	5	5	1.48
2	+	+	-	+	-	-	+	-200	2	O ₂	2	5	0.15
3	+	-	+	-	-	+	+	-200	0	N ₂	2	10	0.65
4	-	+	-	-	+	+	+	-400	2	N ₂	5	10	22.97
5	+	-	-	+	+	+	-	-200	0	O ₂	5	10	1.77
6	-	-	+	+	+	-	+	-400	0	O ₂	5	5	4.98
7	-	+	+	+	-	+	-	-400	2	O ₂	2	10	16.53
8	-	-	-	-	-	-	-	-400	0	N ₂	2	5	5.13

Con los datos obtenidos se calcula el efecto Δ para cada una de las variables para eficiencia de concentración del molibdeno, Y_2

	X ₁	X ₂	F ₁	X ₃	X ₄	X ₅	F ₂
Δ	-11.54	7.30	-1.44	-1.85	2.04	7.40	1.11

$$SS_{Error} = (-1.44)^2 + (-1.11)^2 = 3.32$$

$$MS_{Error} = \frac{3.32}{2} = 1.66$$

Con todos esos cálculos que se construye la Tabla del análisis de la varianza.

Tabla 7.21: Análisis de la Varianza para el ejemplo de flotación de Molibdeno.

Fuente	SS	g.l.	MS	F _o	
X ₁	266.34	1	266.34	160.44	Potencial redox,
X ₂	106.58	1	106.58	106.58	Carbón activado
F ₁	4.18	1	4.18	2.52	
X ₃	6.84	1	6.84	4.12	Gas
X ₄	8.28	1	8.28	4.99	Tiempo acondic
X ₅	109.37	1	109.37	65.88	Tiempo flotac.
F ₂	2.46	1	2.46	1.48	
Error	3.32	2	1.66		
Total	507.38	9			

$$F(\alpha; gl_T, gl_E) = F(0.05; 1, 2) = 18.51$$

El modelo matemático es:

$$\hat{Y} = 6.63 - 5.77 X_1 + 3.65 X_2 + 3.70 X_3$$

Análisis de residuos

N	X_1	X_2	X_3	Y_r	\hat{Y}	$R = (Y_r - \hat{Y})$
1	+	+	-	1.48	0.81	0.67
2	+	+	-	0.15	0.81	-0.66
3	+	-	+	0.65	0.91	-0.26
4	-	+	+	22.97	19.75	3.32
5	+	-	+	1.77	0.91	0.26
6	-	-	-	4.98	5.05	-0.07
7	-	+	+	16.53	19.75	-3.32
8	-	-	-	5.13	5.05	0.08

$$SSM_R = \frac{21.77}{8-4} = 5.44$$

$$F_o = \frac{SSM_R}{MS_{Error}} = \frac{5.44}{1.66} = 3.28$$

El modelo es adecuado si: $F_o < F(\alpha; gl_R, gl_E)$; $F(0.05; 4, 2) = 19.25$. Entonces el modelo matemático ajusta adecuadamente los datos experimentales.

7.8.1.3. Evaluación para consumo de NaHS, Y_3

Tabla 7.22: Diseño P-B con para flotación de Molibdeno.

N	X_1	X_2	F_1	X_3	X_4	X_5	F_2	Z_1	Z_2	Z_3	Z_4	Z_5	Y_3
1	+	+	+	-	+	-	-	-200	2	N ₂	5	5	2.90
2	+	+	-	+	-	-	+	-200	2	O ₂	2	5	6.50
3	+	-	+	-	-	+	+	-200	0	N ₂	2	10	3.30
4	-	+	-	-	+	+	+	-400	2	N ₂	5	10	5.80
5	+	-	-	+	+	+	-	-200	0	O ₂	5	10	7.80
6	-	-	+	+	+	-	+	-400	0	O ₂	5	5	15.40
7	-	+	+	+	-	+	-	-400	2	O ₂	2	10	14.90
8	-	-	-	-	-	-	-	-400	0	N ₂	2	5	7.00

Con los datos obtenidos se calcula el efecto Δ para cada una de las variables para consumo de NaHS, Y_3

	X_1	X_2	F_1	X_3	X_4	X_5	F_2
Δ	5.65	-0.85	2.35	6.40	0.05	0.00	-0.40

$$SS_{Error} = (2.35)^2 + (-0.40)^2 = 5.68$$

$$MS_{Error} = \frac{5.68}{2} = 2.84$$

Con todos esos cálculos que se construye la Tabla del análisis de la varianza.

Tabla 7.23: Análisis de la Varianza para el ejemplo de flotación de Molibdeno.

Fuente	SS	g.l.	MS	F_o	
X_1	63.85	1	63.85	22.47	Potencial redox,
X_2	1.45	1	1.45	0.51	Carbón activado
F_1	11.05	1	11.05	3.89	
X_3	81.92	1	81.92	28.83	Gas
X_4	0.00	1	0.00	0.00	Tiempo acondic
X_5	0.00	1	0.00	0.00	Tiempo flotac.
F_2	0.32	1	0.32	0.11	
Error	5.68	2	2.84		
Total	164.26				

$$F(\alpha; gI_T, gI_E) = F(0.025; 1, 2) = 38.51$$

Con lo cual se concluye que no hay variable significativa a ese nivel de significación. De todos modos, el gas que se utiliza, influye en el consumo del NaHS, siendo importante en la reducción de costos del proceso.

7.8.1.4. Conclusiones

De las tres evaluaciones se concluye que el potencial redox tiene importante influencia en el proceso, y que a valores más negativos (-400 mV) es beneficiosa para la eficiencia de separación, así como para la eficiencia de concentración del molibdeno, influyendo también en el consumo del NaHS.

También se concluye que el uso del nitrógeno es beneficioso porque reduce el consumo de NaHS.

7.9. Operaciones Evolutivas (EVOP) con diseños Factoriales.

Casi nunca es cierto que los procesos industriales sean operados en las mejores condiciones. A menudo el trabajo a escala reducida en el que se basa el diseño de muchos procesos, sólo proporciona una burda aproximación a las condiciones óptimas a utilizar en la producción industrial. La operación evolutiva (EVOP, *Evolutionary Operation*), es una forma continua de operar un proceso productivo de forma que además de producir, genere información sobre como mejorarlo a base de ejecutar diseños factoriales durante la producción. Para evitar cambios apreciables en las características del producto se efectúan pequeños cambios en los niveles de las variables de proceso. Para determinar los efectos de estos cambios, será pues necesario repetir varias veces los experimentos elementales y promediar los resultados.

El método EVOP esta ideado para ser llevado a cabo por los *operadores del proceso* en la planta, mientras se continúa manufacturando *producto satisfactorio*. Por lo tanto, las circunstancias son muy diferentes a las laboratorio o planta piloto, donde:

- Hay que dedicar dinero y tiempo para investigar;
- Se puede variar muchos factores simultáneamente en experimentos cuidadosamente llevados a cabo por técnicos especializados.
- La producción de material invendible no es problema.

En esta situación es aconsejable en general estudiar el máximo número de factores en cada diseño manteniendo la relación (factores estudiados)/(experimentos elementales realizados) tan alta como sea posible. Si en cierta fase de la investigación hay que realizar 16 experimentos elementales, en general es mejor utilizarlos para estudiar cuatro variables en un diseño factorial 2^4 ó mejor aún, estudiar 5 factores en media fracción de un factorial 2^5 , que estudiar 3 factores en un diseño 2^3 duplicado.

Como contraste, considerese la situación en las que se utilizan la operación evolutiva. En ellas:

- a. Como la relación señal/ruido se debe mantener baja, se necesita generalmente un gran número de experimentos elementales para detectar los efectos de los cambios.
- b. Esos experimentos elementales no son mas que operaciones necesarias para la fabricación, que se hubiesen hecho de cualquier modo y no representan un aumento de coste apreciable.

- c. En la línea de producción, las cosas se deben mantener muy simples y normalmente sólo es factible variar dos o tres factores en fase de investigación.

En estas circunstancias, es razonable utilizar diseños factoriales 2^2 ó 2^3 replicados, a menudo añadiendo un punto central.

A medida que se va disponiendo de los resultados, se van actualizando continuamente las medidas y los efectos estimados y se presentan en un tablón informativo, que puede ser una pizarra corriente.

En un estudio reportado en la literatura, se plantea como objetivo disminuir el coste por tonelada de cierto producto de una planta petroquímica. En una fase de la investigación dos variables consideradas importantes son:

1. La tasa de reflujo de una columna de destilación;
2. La relación de flujo de reciclado.

Se esperaba que los cambios de la magnitud planeada introducirían efectos transitorios que desaparecerían en 6 horas. Se dejaron 18 horas más de operación en condiciones estacionarias para hacer las mediciones necesarias.

La respuesta que se evalúa es el costo medio por tonelada. El diseño empleado fue un factorial de 2^2 con un punto central. Los resultados al final de la Fase I se muestran en la Tabla 7.24 y son medidas obtenidas luego de 5 repeticiones

Tabla 7.24: Experimentación *EVOP* de una planta petroquímica. Fase I

Relación Reflujo	Flujo Reciclado.	Costo por tonelada
6.7	7.5	92
6.7	8.0	86
7.1	7.5	95
7.1	8.0	91
6.9	7.75	92

	Relación Reflujo	Flujo Reciclado.	<i>Interacción</i>
Efectos y errores tipo	4.0 ± 1.5	5.0 ± 1.5	1.0 ± 1.5

En este momento se consideró que había suficientes pruebas para justificar un cambio a menor relación de flujo y mayor reciclado. La fase II se abordó con cinco ciclos más, que confirma que se consiguen costos más bajos, como se observa en la Tabla 7.25.

Tabla 7.25: Experimentación *EVOP* de una planta petroquímica. Fase II

Relación Reflujo	Flujo Reciclado.	Costo por tonelada
5.9	8.0	83
5.9	8.5	80
6.7	8.0	84
6.7	8.5	81
6.3	8.25	82

	Relación Reflujo	Flujo Reciclado.	Interacción
Efectos y errores tipo	1.0 ± 1.0	3.0 ± 1.0	0.0 ± 1.0

Con los cambios realizados se ensayaron con valores aún más altos para reciclado, lo cual motivó la fase III, Tabla 7.26; esta fase concluyó después de 4 ciclos y condujo a la conclusión de que el costo más bajo se obtiene con un valor de reflujo cercano a 6.3 y un reciclado de alrededor de 8.5, con un costo por tonelada de 80 libras.

Tabla 7.26: Experimentación *EVOP* de una planta petroquímica. Fase I

Relación Reflujo	Flujo Reciclado.	Costo por tonelada
5.9	8.25	84
5.9	8.75	86
6.7	8.25	83
6.7	8.75	85
6.3	8.5	80

	Relación Reflujo	Flujo Reciclado.	Interacción
Efectos y errores tipo	-1.0 ± 1.5	2.0 ± 1.5	0.0 ± 1.5