

TEMA 8: Diseño Estadístico de Experimentos Aplicado a la Teoría de la Calidad

- 1 Diseño de experimentos para la mejora del proceso de producción
- 2 Diseños factoriales fraccionarios 2^{k-p}
- 3 Diseños Box-Behnken 3^{k-p}
- 4 Diseños de superficie de respuesta no factoriales
 - Diseños estándar disponibles
- 5 Otros diseños alternativos
 - Cuadrados latinos
 - Diseños mixtos y superficies triangulares
 - Diseños para superficies y mixturas con restricciones
 - Construcción de diseños D- y A-óptimos

1. Diseño de experimentos para la mejora del proceso de producción

Según se vio en el Tema 1, existen diferentes enfoques en la introducción del concepto de calidad. En todos ellos hay rasgos comunes presentes tales como:

- La calidad es una medida del nivel en que se satisfacen requerimientos y expectativas del cliente.
- La calidad no es un concepto estático, puesto que las expectativas de los clientes pueden cambiar.
- La calidad incluye el cumplimiento de especificaciones que deben satisfacer productos o servicios para cubrir las necesidades del cliente.

El concepto de calidad según la filosofía de Taguchi se basa en la pérdida que este producto causa a la sociedad. La pérdida es definida en términos de dos aspectos:

- La pérdida causada por la variabilidad del proceso.
- La pérdida causada por los perjuicios que ocasiona el consumo de productos con poca calidad.

La calidad del diseño es independiente del concepto de calidad que se considere. Más concretamente, la calidad de diseño aumenta cuando se consigue determinar las condiciones internas y externas que producen una disminución de la variabilidad del proceso y una eliminación del sesgo, o error de aproximación al valor objetivo (valor nominal). Por ejemplo, en la metodología de Taguchi se determinan los niveles de los factores controlables (parámetros del diseño) que hacen el diseño robusto frente a factores externos de ruido, no controlables, determinados por las condiciones medio-ambientales. Taguchi identifica tres fuentes de variabilidad:

- (i) Factores de ruido externo o variación en las condiciones ambientales.
- (ii) Factores de ruido interno o deterioro.
- (iii) Factores de ruido asociados a la elaboración de las unidades (por ejemplo, los factores que intervienen en la variabilidad en materias primas, equipos de manufactura, etc.) y que ocasionan diferencias en productos elaborados con las mismas especificaciones.

Existen diferentes enfoques para establecer el diseño de factores controlables, es decir, factores que intervienen en la elaboración de las unidades y en el funcionamiento del proceso (fuentes de variabilidad (ii) y (iii)) y factores no controlables, asociados a las condiciones medio-ambientales (fuentes de variabilidad (i)). Los enfoques principales son:

- (I) No considerar el control de factores de ruido durante el experimento. En lugar de esto, considerar réplicas de la variable respuesta bajo las posibles combinaciones de niveles de factores controlables y medir la variabilidad del proceso en términos de la desviación estándar muestral en cada punto del experimento.
- (II) Identificar los factores de ruido antes de iniciar la experimentación con el proceso. Incluir entonces los factores de ruido con los factores controlables en la matriz del diseño. Por ejemplo, se puede considerar un diseño factorial fraccionario con k factores, donde k incluye los factores controlables y de ruido, construido a partir de un diseño factorial completo con $k - p$ factores y 2^{k-p} tríadas. Se evaluarán las interacciones entre factores controlables (o del diseño) y factores de ruido, y se identificarán los niveles de los factores controlables que minimizan los efectos de los factores de ruido.
- (III) Considerar una matriz del diseño para factores controlables (matriz de factores internos) y una matriz del diseño para factores no controlables (matriz para factores externos). Observar la variable respuesta (característica de interés) en cada experimento bajo las posibles combinaciones de condiciones medio-ambientales (factores externos). El análisis se efectúa en términos de m mediciones obtenidas en cada experimento, donde m denota el número de filas de la matriz del diseño externo (Metodología de Taguchi).

El enfoque seguido en la aplicación de la teoría de diseño de experimentos en la industria, para el control de la calidad de la producción, se halla motivado por el objetivo que se persigue, que es bastante diferente del interés o de los objetivos planteados en el contexto estadístico para el desarrollo de esta teoría. Más concretamente, en la industria el objetivo primordial es usualmente extraer la máxima cantidad de información no sesgada en relación con los factores que afectan al proceso de producción a partir de la menor cantidad posible de observaciones (para minimizar costos). Sin embargo, en el desarrollo estadístico de las técnicas de diseño de experimentos el objetivo primordial es probar si un factor concreto produce un efecto significativo, desde el punto de vista estadístico, sobre la variable de interés dependiente o variable respuesta. Este último objetivo ha motivado el desarrollo de técnicas de análisis de la varianza (ANOVA/MANOVA) para analizar los efectos de interacciones de orden superior entre factores. En el ámbito de la industria, los efectos de interacciones entre factores se consideran usualmente despreciables, puesto que el interés fundamental reside en el análisis del mayor número posible de factores principales con el mínimo gasto. En este sentido, la consideración de factores de interacción de orden muy elevado sólo complica el análisis y el proceso de identificación de efectos de factores principales.

La diferencia entre objetivos ha marcado el enfoque seguido en el desarrollo y la aplicación de las técnicas de diseño de experimentos en los dos ámbitos. Así se tiene que en el ámbito estadístico se consideran diseños con pocos factores; por ejemplo, en los textos clásicos sobre ANOVA se considera el análisis de diseños con 5 o 6 factores como máximo (con más de seis factores resulta impracticable). El desarrollo se centra en el estudio de contrastes de significación robustos y válidos. Este ha sido el objetivo de la asignatura Diseño de Experimentos de primer curso de la Licenciatura en Ciencias y Técnicas Estadísticas. Sin embargo, en los textos dirigidos a la aplicación del diseño de experimentos a la industria se consideran diseños con muchos factores (e.g. 16 o 32) y los efectos de factores de interacción no pueden ser analizados. El objetivo primordial en estos textos es la derivación de estimaciones insesgadas de los efectos de factores principales, y quizás, de los factores de interacción doble, a partir del mínimo número de observaciones posible.

En general, en el desarrollo y la aplicación de las técnicas de diseño de experimentos se intenta abordar dos problemas:

- (i) Cómo diseñar un experimento.
- (ii) Cómo analizar los resultados de un experimento.

Este último punto se aborda mediante la aplicación de técnicas de análisis de la varianza (ANCOVA/MANCOVA, el Modelo Lineal General, modelos de regresión generales, ANCOVA mixta, componentes de la varianza, etc.), que han sido estudiadas en la asignatura Diseño de Experimentos de primer curso de la Licenciatura en Ciencias y Técnicas

Estadísticas, así como en otras asignaturas de esta licenciatura).

En este tema nos centraremos en el desarrollo del primer punto, para proporcionar los elementos que entran en juego en diferentes tipos de diseños, que contemplan los aspectos que son deseables en relación con el objetivo de identificar la combinación óptima de niveles de factores que afectan a la calidad del producto y del proceso. Notéase que en este tema se estudian los diseños más aplicados en la industria, salvo los diseños robustos, formulados en la metodología de Taguchi, que serán introducidos en el Tema 9.

2. Diseños factoriales fraccionarios 2^{k-p}

Según se ha comentado en la sección anterior, en el estudio de las condiciones que afectan al proceso de producción y al producto, un objetivo primordial es estudiar el mayor número de factores posibles. La consideración de dos niveles en cada factor suele proporcionar información suficiente en muchos casos. Para minimizar costos, teniendo en cuenta que el estudio de factores de interacción no es el objetivo fundamental, se sacrifica, en parte, el análisis de dichos factores y se consideran los diseños factoriales fraccionarios con dos niveles.

En un diseño 2^{k-p} intervienen k factores. Sin embargo, la matriz del diseño está constituida por 2^{k-p} filas, donde p denota el número de factores generados a partir de las columnas que definen los factores de interacción en un diseño factorial (completo) con $k-p$ factores y dos niveles. Es decir, se considera como matriz del diseño la matriz que define el diseño factorial completo con 2^{k-p} filas y cuyas $k-p$ primeras columnas corresponden a factores principales. Se asocian entonces los p restantes factores principales a las columnas correspondientes a factores de interacción. Puede ocurrir que con los p factores principales restantes no se cubran todas las columnas de la matriz del diseño original.

Ejemplo: Diseño 2^{11-4} . En este tipo de diseño fraccionario ortogonal se parte de la matriz del diseño factorial 2^4 con cuatro factores y dos niveles. La matriz del diseño

factorial fraccionario con 11 factores y 16 tríadas viene dada por

<i>Experimento</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>K</i>
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
13	-1	-1	1	1	1	-1	-1	1	1	1	-1
14	-1	-1	1	-1	1	1	1	-1	-1	1	-1
15	-1	-1	-1	1	-1	1	1	1	-1	1	1
16	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1

Este diseño permite analizar un gran número de factores con un número relativamente pequeño de datos. Sin embargo, el análisis de interacciones de orden superior a dos no es posible, puesto que dichos factores de interacción son ‘confundidos’ con factores principales del diseño fraccionario. Más concretamente, los factores E, F, G, H e I se asocian a las columnas de los factores de interacción tres ABC, ABD, ACD y BCD , y al factor principal I se le hace corresponder la columna del factor de interacción cuatro $ABCD$ en la matriz del diseño original. Los factores J y K se identifican con factores de interacción doble. Consecuentemente, en el análisis únicamente aparecen factores de interacción doble.

Un diseño con resolución R es aquél en el que los factores de interacción de orden l no se identifican o confunden con factores de interacción de orden inferior a $R-l$. En el ejemplo anterior se tiene $R=III$ y $l=1$, por tanto, los efectos de factores principales no se confunden entre sí; sin embargo, factores principales son confundidos con factores de interacción doble. Los efectos de factores de interacción doble también se confunden entre sí. El diseño factorial fraccionario de cinco factores en ocho tríadas, es decir, el diseño factorial fraccionario 2^{5-2} , posee también resolución III . En diseños con resolución IV ($R=IV$, $l=1$) los factores principales no se identifican con factores de interacción doble; sin embargo, factores de interacción doble se pueden identificar entre sí. Un ejemplo es el diseño factorial fraccionario 2^{4-1} . En diseños con resolución V , los factores principales

o los factores de interacción doble no se identifican entre sí; sin embargo, factores de interacción doble son identificados o confundidos con factores de interacción triple.

Los diseños fraccionarios con resolución IV son frecuentemente usados, puesto que el balance entre información útil y número de tríadas es positivo. Los diseños fraccionarios con resolución III proporcionan gran cantidad de información en relación con el número de tríadas, aunque en ocasiones pueden ser poco significativos debido al elevado número de identificaciones y factores que se confunden. Generalmente, cuando se deben incluir un número elevado de factores, se suele utilizar en una primera fase un diseño con resolución III y, después de identificar los factores cuyos efectos no son significativos, pasar a un diseño con resolución IV para los factores restantes. En general, los diseños con resolución V resultan demasiado caros en la mayoría de las situaciones.

Tras realizar las identificaciones correspondientes al nivel de ‘confusión’ considerado (resolución del diseño), se ha de determinar cuáles son los factores que producen un efecto significativo en la respuesta y, de dichos factores, cuáles son los factores que producen un efecto real en la respuesta (entre factores que son confundidos o identificados). Se suelen considerar los siguientes criterios básicos:

- Determinar como efecto real para dos factores de interacción identificados el de aquel factor que implique a factores principales con efectos más significativos sobre la variable respuesta.
- Determinar como real el efecto de factores principales frente a factores de interacción de orden superior o igual a tres.

Cuando los datos no nos permiten identificar con claridad los efectos reales de factores principales frente a factores de interacción doble se puede considerar un *folding* sobre el diseño, que consiste en los siguientes pasos:

- Se realizan dos tablas de análisis de la respuesta, correspondientes a la matriz del diseño original y a la matriz del diseño obtenida al sustituir en la matriz original niveles bajos por niveles altos y viceversa.
- Se obtienen de las tablas de análisis de la respuesta asociadas a cada diseño, estimaciones de los efectos de cada factor (a partir de la observación de la respuesta para cada combinación de niveles de los factores en cada matriz).
- Se estiman los efectos de los factores principales mediante la semisuma de las correspondientes estimaciones obtenidas en las dos tablas de análisis de la respuesta consideradas. Análogamente, se estiman los efectos de los factores de interacción doble mediante la semidiferencia de las estimaciones previas obtenidas.

Ejemplo. Se considera la siguiente matriz del diseño factorial 2^3 , es decir, del diseño con 3 factores a dos niveles:

<i>Experimento</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
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

Se definen entonces los siguientes diseños factoriales fraccionarios:

- Diseño 2^{4-1} , para cuatro factores, A, B, C, D .

Se asigna el factor D a la columna siete, es decir, se considera la identificación

$$D = ABC.$$

Por tanto,

$$AB = CD \quad (\text{columna 4})$$

$$AC = BD \quad (\text{columna 5})$$

$$AD = BC \quad (\text{columna 6}).$$

- Diseño 2^{5-2} , para cinco factores, A, B, C, D, E .

Se asignan los factores D y E a las columnas cuatro y cinco, respectivamente. Se tienen entonces las identificaciones

$$A = BD = CE \quad (\text{columna 1})$$

$$B = AD \quad (\text{columna 2})$$

$$C = AE \quad (\text{columna 3})$$

$$D = AB \quad \text{columna 4}$$

$$E = AC \quad (\text{columna 5})$$

$$BC = DE \quad (\text{columna 6})$$

$$BE = CD \quad (\text{columna 7}) \tag{1}$$

- Diseño 2^{6-3} , para seis factores, A, B, C, D, E, F .

Se asignan los factores D, E y F a las columnas cuatro, cinco y seis, respectivamente.

Se tienen entonces las identificaciones

$$\begin{aligned}
A &= BD = CE & (\text{columna 1}) \\
B &= AD = CF & (\text{columna 2}) \\
C &= AE = BF & (\text{columna 3}) \\
D &= AB = EF & \text{columna 4} \\
E &= AC = DF & (\text{columna 5}) \\
F &= BC = DE & (\text{columna 6}) \\
AF &= BE = CD & (\text{columna 7})
\end{aligned} \tag{2}$$

- Diseño 2^{7-4} , para siete factores, A, B, C, D, E, F, G .

Se asignan los factores D, E, F y G a las columnas cuatro, cinco, seis y siete, respectivamente. Se tienen entonces las identificaciones

$$\begin{aligned}
A &= BD = CE = FG & (\text{columna 1}) \\
B &= AD = CF = EG & (\text{columna 2}) \\
C &= AE = BF = DG & (\text{columna 3}) \\
D &= AB = EF = CG & \text{columna 4} \\
E &= AC = BG = DF & (\text{columna 5}) \\
F &= AG = BC = DE & (\text{columna 6}) \\
G &= AF = BE = CD & (\text{columna 7})
\end{aligned} \tag{3}$$

Aplicación. A continuación se considera un diseño 2^{7-4} para determinar la calidad de la soldadura de una junta realizada con un proceso que no parece seguro. Se consideran los siguientes factores:

- A: Contenido de aleación del material puro.
- B: Diámetro del electrodo.
- C: Corriente del circuito.
- D: Dióxido de carbono.
- E: Voltaje del soldador.
- F: Velocidad de la corriente en el tubo de alimentación.
- G: Precalentamiento.

Se consideran dos niveles para cada uno de los factores anteriores y se observa la resistencia de la soldadura bajo las diferentes combinaciones de los factores considerados, obteniéndose las siguientes medidas:

$$147,2; 91,3; 72,7; 87,4; 84,1; 78,2; 94,6; 138,8.$$

Los efectos estimados para los factores considerados a partir de la tabla de análisis de la respuesta asociada al diseño 2^{7-4} , experimento E_1 , son

- Efecto estimado para factor el A: -0.72.
- Efecto estimado para el factor B: -1.82.
- Efecto estimado para el factor C: -0.72.
- Efecto estimado para el factor D: 37.38.
- Efecto estimado para el factor E: 19.88.
- Efecto estimado para el factor F: 30.18.
- Efecto estimado para el factor G: -5.12.

Según lo anterior, en un primer análisis se pensaría en los factores D , E y F como factores influyentes o significativos. Sin embargo, hay que tener en cuenta que se han considerado las siguientes identificaciones:

$$\begin{aligned} D &= AB = EF = CG \\ E &= AC = BG = DF \\ F &= AG = BC = DE, \end{aligned}$$

de acuerdo con un diseño fraccionario. Si suponemos que los factores A, B, C y G no producen efectos significativos en la respuesta, se tiene entonces que no definirán factores de interacción con efectos significativos y, por tanto, las identificaciones anteriores se reducen a

$$\begin{aligned} D &= EF \\ E &= DF \\ F &= DE. \end{aligned}$$

Para determinar cuáles son los efectos reales, si se deben a factores de interacción doble o si se deben a factores principales, se necesita información adicional. La consideración de réplicas en cada experimento no ayudaría demasiado. Se considera entonces un *fold-ing* sobre el diseño experimental fraccionario 2^{7-4} . Se observa de nuevo la firmeza de la

soldadura bajo las condiciones establecidas por un segundo experimento E_2 definido por la matriz del diseño que resulta de intercambiar en la matriz del diseño 2^{7-4} niveles altos por niveles bajos. Las mediciones obtenidas son

89,8; 137,4; 82,7; 72,2; 71,9; 87,3; 144,6; 93,2,

y las correspondientes estimaciones de efectos vienen dadas por

- Efecto estimado para el factor A: -3.72.
- Efecto estimado para el factor B: -1.58.
- Efecto estimado para el factor C: -0.28.
- Efecto estimado para el factor D: -37.72.
- Efecto estimado para el factor E: 18.28.
- Efecto estimado para el factor F: 31.22.
- Efecto estimado para el factor G: 2.18.

A partir de las estimaciones de los efectos de cada factor, obtenidas en los experimentos E_1 y E_2 , se calcula una aproximación de los efectos de factores principales y de los efectos de factores de interacción doble como sigue:

	Experimento 1 (E_1)	Experimento 2 (E_2)	Ef. principales $\frac{E_1+E_2}{2}$	Ef. interacción $\frac{E_1-E_2}{2}$
$\bar{Y}$	99,29	97,39	98,34	0,95
1	-0,72	-3,72	-2,22	1,50
2	-1,82	-1,58	-1,70	0,12
3	-0,72	-0,28	-0,50	-0,22
4	37,38	-37,72	-0,17	37,55
5	19,88	18,28	19,08	0,80
6	30,18	31,22	30,70	-0,52
7	-5,12	2,18	-1,47	-3,65

Por tanto, se concluye, en un primer análisis, que los factores que deben ser considerados son los factores E y F y su interacción, es decir, el factor EF . Puesto que el objetivo es obtener la máxima firmeza o dureza, se considerarían los niveles altos de E y F .

Existen dos criterios básicos para la selección de un diseño fraccionario, que consisten en considerar:

- (I) *Criterio de diseño basado en la máxima no confusión*: Diseños de máxima resolución que maximizan el número de factores de interacción de orden inferior o igual al valor crítico que no se confunden con factores de interacción de orden crítico.

- (II) *Criterio de diseño basado en la mínima aberración:* Diseños de máxima resolución que minimizan el número de pares de factores de interacción de orden crítico que se confunden.

3. Diseños Box-Behnken 3^{k-p}

Algunos estudios requieren analizar factores que presentan tres niveles. Por ejemplo, si se sospecha que el efecto de factores sobre la variable de respuesta (variable dependiente) no es lineal, entonces se necesita considerar, al menos, tres niveles para contrastar efectos lineales y cuadráticos (e interacciones) para dichos factores. En otras ocasiones, los factores tienen una naturaleza categórica, presentando más de dos categorías.

El procedimiento a seguir en la generación de diseños 3^{k-p} es muy similar al descrito para diseños 2^{k-p} . Más concretamente, se parte de la matriz de un diseño a tres niveles completo. Entonces se utilizan las columnas que definen los factores de interacción para definir nuevos factores principales en el diseño fraccionario, es decir, se confunden los nuevos factores principales con factores de interacción del diseño original.

Ejemplo. Diseños 3^{3-1} . En este diseño se parte de un diseño con dos factores A y B con tres niveles. El nuevo factor considerado C se identifica o confunde con el factor de interacción AB .

Experimento	A	B	C
1	0	0	0
2	0	1	2
3	0	2	1
4	1	0	2
5	1	1	1
6	1	2	0
7	2	0	1
8	2	1	0
9	2	2	2

Los valores del factor C son calculados mediante la ecuación

$$C = 3 - \text{mod}_3(A + B), \quad (4)$$

donde $\text{mod}_3(L)$ representa el resto de dividir el número L entre 3, es decir, representa un valor congruente con L módulo tres. En realidad, la identidad (4), que define el nuevo factor C , se expresa de forma equivalente en términos de la igualdad

$$\text{mod}_3(A + B + C) = 0,$$

que puede ser interpretada mediante la relación

$$ABC = 0,$$

que indica el nivel de ‘confusión’ de este diseño.

En algunos casos la identidad que genera el diseño se define en términos del valor 2 que aparece como un factor de multiplicación. Por ejemplo, en el diseño factorial fraccionario con 4 factores a 3 niveles, en 9 bloques, definido en términos de 27 experimentos, el generador del diseño es el factor $ABCD$ y los generadores de bloques son AB y AC^2 . La identidad fundamental, en términos de la cual se define el diseño, es

$$0 = \text{mod}_3(B + 2C + D + 2E + F).$$

Según se ha comentado, el objetivo de los diseños fraccionarios es analizar el máximo número de factores en el mínimo número de tríadas. Este objetivo se halla reflejado en la construcción de los diseños Box-Behnken. Los diseños completos que generan estos diseños 3^{k-p} no son sencillos, puesto que provienen de la combinación de diseños factoriales a dos niveles con diseños por bloques incompletos y presentan una estructura bastante compleja en el proceso de identificación de factores. Sin embargo, estos diseños son bastante económicos y, por tanto, particularmente útiles cuando la ejecución de experimentos resulta muy costosa.

El análisis de este tipo de diseños es muy similar al análisis desarrollado para diseños fraccionarios a dos niveles. Sin embargo, para cada efecto hay que contrastar el efecto lineal y cuadrático (no lineal).

4. Diseños de superficie de respuesta no factoriales

En los diseños estudiados anteriormente se requiere que todos los factores sean considerados a dos o tres niveles. En muchos casos, estos diseños no son lo suficientemente flexibles; en particular, cuando algunas de las combinaciones de los niveles de los factores no están permitidas. Por ejemplo, dos o más factores no pueden estar simultáneamente al nivel más alto. También, por razones de eficiencia interesa explorar la región experimental de interés en puntos que no pueden ser representados mediante un diseño factorial.

En esta sección se ofrece una alternativa que permite resolver estos problemas. Más concretamente, se consideran diseños que permiten la estimación (ajuste) de superficies de respuesta que satisfacen la ecuación general

$$Y = b_0 + b_1X_1 + \dots + b_kX_k + b_{12}X_1X_2 + b_{13}x_1X_3 + \dots + b_{k-1,k}X_{k-1}X_k + b_{11}X_1^2 + \dots + b_{kk}X_k^2. \quad (5)$$

El modelo anterior se ajustaría a los valores observados de una variable respuesta Y , que incluye:

- (1) Los efectos principales de los factores, $X_1, \dots, X_k$.
- (2) Sus interacciones, $(X_1X_2, X_1X_3, \dots, X_{k-1}X_k)$.

(3) Sus componentes cuadráticas, $(X_1^2, \dots, X_k^2)$.

No se realizan suposiciones en relación con los niveles de los factores, pudiéndose analizar cualquier conjunto de valores continuo de los factores. Hay algunas consideraciones en relación con la eficiencia y el sesgo del diseño que han llevado a la definición de algunos diseños estándar (que se comentarán brevemente en la subsección siguiente), que son frecuentemente utilizados cuando se pretende ajustar superficies de respuesta. Sin embargo, en el ámbito del diseño de superficies con ciertas características y de diseños D- y A-óptimos, estos diseños estándar no son aplicables en la práctica. En el planteamiento anterior se realiza un análisis donde no se hacen suposiciones sobre la estructura de los datos, es decir, sobre el rango de valores de los factores, o de su combinación; por tanto, con este marco se puede analizar cualquier diseño para ajustar a los datos el modelo general anterior.

En la construcción de diseños basados en el modelo general (5), es importante tener en cuenta ciertas propiedades que son aconsejables en un diseño, tales como:

- La ortogonalidad: Una característica deseable en cualquier diseño es que las estimaciones de los factores principales y de interacción de interés sean ortogonales.
- Un objetivo primordial en la industria es la extracción de la mayor cantidad posible de información insesgada; la isotropía o rotabilidad del diseño es una característica aconsejable en relación con este aspecto. Es decir, es deseable que la cantidad de información no varíe cuando se produce una rotación de los puntos del diseño original.

En relación con la obtención de estas características, se puede comprobar que dependen del número de puntos centrales del diseño y de la llamada distancial axial (α), que es la distancia de los puntos estrella al centro del diseño. Se puede probar que un diseño es rotable o isotrópico si

$$\alpha = n_c^{1/4},$$

siendo n_c el número de puntos en la porción factorial del diseño. Un diseño construido a partir de (5) es ortogonal si

$$\alpha = \left\{ \left[(n_0 + n_c + n_s)^{1/2} - n_c^{1/2} \right] n_c / 4 \right\}^{1/4},$$

siendo n_0 y n_s el número de puntos centrales y el número de puntos estrella del diseño, respectivamente.

Para conseguir las dos características ortogonalidad y rotatividad, se debe ajustar la distancia axial primero para conseguir la rotabilidad, y entonces, añadir los puntos centrales de forma que

$$n_0 \approx 4n_c^{1/2} + 4 - 2k,$$

para conseguir la ortogonalidad, siendo k el número de factores en el diseño. Finalmente, las fórmulas anteriores deben ser modificadas si se consideran bloques en el diseño, en términos de los puntos centrales y no centrales que se encuentran en la porción factorial del diseño, así como en términos de los puntos centrales y no centrales estrella.

4.1. Diseños estándar disponibles

La porción factorial de estos diseños se construye normalmente a partir de diseños factoriales fraccionarios con dos niveles. El diseño es completado con puntos centrales y estrella. La porción factorial fraccionaria del diseño suele ser de resolución V o superior. Esta resolución no es estrictamente necesaria. Así, se tiene que cuando la ejecución de los experimentos resulta muy cara, o bien cuando no se necesita una gran potencia estadística en el contraste de adecuación del modelo, se puede utilizar una resolution III en la porción factorial fraccionaria.

El análisis de estos modelos es similar al análisis de los diseños 3^{k-p} a partir del ajuste del modelo (5) a los datos.

5. Otros diseños alternativos

En esta sección se refieren brevemente algunas otras modalidades de diseños que son frecuentemente utilizados en procesos de producción en la industria.

5.1. Cuadrados latinos.

Se utilizan cuando los factores presentan más de dos niveles y se sabe que no existen interacciones entre los factores, o bien que son despreciables. Se consideran entonces cuadrados latinos que definen las posibles combinaciones de los niveles de los factores principales para estimar sus efectos.

Ejemplo. Para tres factores con cuatro niveles se tiene

	1	2	3	4
1	A	B	D	C
2	D	C	A	B
3	B	D	C	A
4	C	A	B	D

que define una de las posibles combinaciones para estimar los efectos de los factores principales. Se conoce con el nombre de cuadrado latino 4×4 . Nótese que para un diseño factorial completo habría 64 experimentos si se consideran todos los factores de interacción.

Otra variante es el cuadrado greco-latino, que se obtiene por superposición de dos cuadrados latinos, y el cuadrado hiper-grecolatino, que se obtiene por superposición de tres o más cuadrados latinos. Un ejemplo de este tipo de diseños se construye cuando se superponen los tres posibles cuadrados latinos asociados al diseño de factores con cuatro niveles.

5.2. Diseños mixtos y superficies triangulares

Surgen cuando se han de analizar mixturas de componentes que deben sumar una constante. La fórmula más frecuente para resumir la información de proporciones de mixturas es mediante gráficos triangulares.

Por ejemplo, si se tiene una mixtura de tres componentes A , B y C , cualquier mixtura de estas tres componentes puede ser representada mediante un punto en un sistema de coordenadas triangular definido por las tres variables:

A	B	C
1	0	0
0	1	0
0	0	1
0,5	0,5	0
0,5	0	0,5
0	0,5	0,5

La suma de cada mixtura es uno. Por tanto, los valores para las componentes en cada mixtura pueden ser interpretados como proporciones. Si se representan estos datos en un gráfico tridimensional, se comprueba que los puntos forman un triángulo. Sólo los puntos dentro del triángulo donde la suma de los valores de las componentes es uno son mixturas válidas. Por tanto, se puede representar el triángulo para visualizar de forma resumida los valores de las componentes (proporciones) para cada mixtura.

Se puede añadir una cuarta dimensión al triángulo, perpendicular a las tres primeras. Los valores de la variable dependiente o respuesta pueden representarse entonces utilizando esta cuarta dimensión. Nótese que la superficie de respuesta puede ser representada en un gráfico tridimensional, o bien en un gráfico de contorno bidimensional.

5.3. Diseños para superficies y mixturas con restricciones

Según se ha mencionado en el contexto de los diseños mixtos, en estudios aplicados la región experimental con frecuencia está restringida, es decir, no todos los niveles o variantes de los factores pueden ser combinados entre sí. Se dispone de un algoritmo, debido a Piepel (1988) y Snee (1985), para encontrar o determinar los vértices y centroides para

tales regiones con ligaduras o restricciones. Este algoritmo se implementa para ligaduras o restricciones lineales del tipo

$$A_1X_1 + A_2X_2 + \dots + A_qX_q + A_0 \geq 0.$$

En la ecuación anterior $A_0, \dots, A_q$ representan los parámetros para las restricciones lineales de los q factores considerados y $X_1, \dots, X_q$ representan los valores (niveles) de los q factores. Este tipo de ligaduras son consideradas en una gran variedad de situaciones.

5.4. Construcción de diseños D- y A-óptimos

Los métodos para la construcción de diseños D- y A-óptimos proporcionan diferentes opciones para seleccionar de una lista de puntos válidos (candidatos dados por combinaciones de niveles o valores de los factores) aquellos puntos que permitirán extraer la máxima cantidad de información de la región experimental, dado el modelo que se pretende ajustar a los datos. Se necesita suministrar la lista de puntos candidatos y especificar el tipo de modelo que se espera ajustar a los datos y el número de tríadas del experimento. Se construye entonces un diseño con el número deseado de casos, que proporcionará el mayor grado posible de ortogonalidad entre las columnas de la matriz del diseño. Se han desarrollado asimismo algoritmos para la búsqueda de este tipo de diseños óptimos. Sin embargo, este aspecto no se halla contemplado en el temario de esta asignatura.

La nomenclatura ‘D- y A-óptimos’ proviene del objetivo contemplado en estos diseños en relación con maximizar el determinante y la diagonal (minimizar los elementos fuera de la diagonal), respectivamente, de la matriz del diseño para obtener máxima ortogonalidad o independencia entre los factores.