

P A R T E 3

Estudio crítico de los distintos enfoques al problema del diseño de parámetros. El gráfico media-desviación (MD)

Introducción

En esta parte 3 de la tesis se estudia el uso del gráfico media-desviación propuesto por Grima (1993) para el análisis de los resultados en el diseño de productos robustos. Con el objeto de ubicar en su contexto esta alternativa de análisis hemos dividido en dos capítulos la parte 3, en el primero se expone brevemente algunas de las propuestas al problema del diseño robusto de parámetros y en el segundo se estudia el gráfico como una opción de análisis versátil y sencilla.

El primer capítulo de la parte 3 pretende aportar una visión global de los diferentes enfoques encaminados a la solución del problema del diseño de parámetros. Las citas que se dan en los diferentes apartados pueden servir para profundizar en el tema. El método de Taguchi se desarrolla de forma más extensa por ser el método pionero y la razón de ser de muchos de los métodos alternativos. Los diferentes enfoques se han agrupado, en la medida de lo posible, con base en la selección del diseño, la selección de la métrica y el análisis formal.

El capítulo dedicado al gráfico media-desviación expone su uso y la manera de elaborarlo, se destaca su consistencia y su simplicidad comparándolo con otros métodos de análisis y se muestran las opciones que tiene para considerar costes y respuestas múltiples.

CAPÍTULO 3

Distintos enfoques al problema del diseño de parámetros

3.1 Introducción

El diseño de parámetros, o **diseño robusto de parámetros**, es una técnica de mejora de la calidad propuesta por Genichi Taguchi que tiene por objeto reducir la variación presente en productos y procesos. La idea esencial reside en que es posible reducir la variabilidad que generan variables fuera de control, al seleccionar valores adecuados en las variables susceptibles de controlarse; de esta forma se tiene un producto o proceso **robusto** (insensible) a los cambios inevitables en dichas variables incontrolables. Esta robustez a las variables incontrolables solamente se puede incorporar en la etapa primaria de diseño del producto. El uso de la palabra diseño en la denominación de diseño de parámetros no tiene el sentido de diseño de experimentos sino de selección de los niveles de las variables controlables para conseguir robustez en los productos o procesos.

El entusiasmo despertado en los Estados Unidos por las prácticas de calidad japonesas a principios de los años 80 y las dos conferencias Mohonk (1984, 1985), organizadas por la Quality Assurance Center de los AT&T Bell Labs, fueron la base para que las ideas de Taguchi se dieran a conocer en la comunidad estadística occidental. A la par de su creciente uso en la industria surgieron las críticas, algunas de ellas tan intrascendentes como la originalidad de la idea de robustez o si los diseños ortogonales empleados por Taguchi son los diseños clásicos de Plackett y Burman de 1946. A partir de entonces las críticas han venido acompañadas de propuestas alternativas, desde estrategias experimentales hasta formas de analizar los datos, entre ellas cabe destacar el uso de transformaciones, Box (1988), el empleo de

métricas más sencillas (s^2 , s o $\log(s)$), las medidas del desempeño que son independientes del ajuste (PerMIA) Leon et al (1987), la modelación explícita de los efectos de los factores ruido a través de la matriz combinada con un enfoque de superficie de respuesta, Vining y Myers (1990), Lucas (1989), Shoemaker et al (1989) y Welch et al (1990), los métodos gráficos, Grima (1993), Myers et al. (1997), los modelos lineales generalizados, Grego (1993), Hamada y Nelder (1997), los planes experimentales, Mitchell y Box (1974), Lucas (1976), Atkinson (1995) y Romero (2002), y los enfoques de programación no lineal, Fathi (1991), Del Castillo y Montgomery (1993).

3.2 Método de Taguchi

Taguchi hace una separación de los factores del proceso o producto en **factores controlables** y **factores ruido**. Los factores controlables son aquellas variables susceptibles de fijarse a niveles deseados durante el proceso. Los factores ruido son aquellos que influyen en el proceso pero que al no poderse controlar, por ser muy costoso o difícil, producen variabilidad. Un aspecto clave en el diseño de parámetros es el introducir los factores ruido en el diseño de experimentos para estudiar la relación que tienen con los factores controlables. Por lo tanto se asume que, aunque incontrolables en el sistema, los factores ruido pueden controlarse para propósitos experimentales.

La idea central de encontrar los niveles de los factores controlables para reducir la variabilidad transmitida por variables de ruido (conseguir la robustez ante la presencia de los factores ruido) demanda la existencia de ciertas interacciones entre los factores ruido y los controlables.

La Figura 3.1 ilustra esta noción, si se usa el nivel alto del factor controlable X se consigue aproximadamente la misma respuesta promedio, independientemente del valor que tenga (en la región analizada) el factor ruido Z . Se suele utilizar en la literatura la notación X_1, X_2, \dots para los factores controlables y Z_1, Z_2, \dots para los factores ruido.

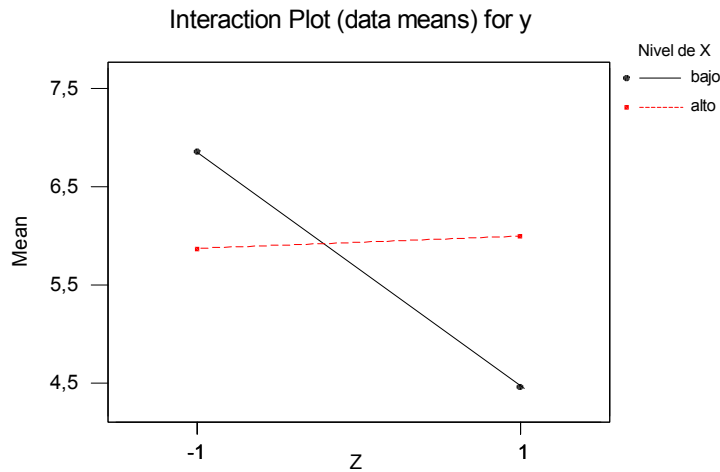


Figura 3.1 Interacción entre el factor controlable X y el factor ruido Z

En la metodología de Taguchi la parte experimental se realiza empleando un diseño ortogonal para los factores controlables, denominado **arreglo interno**, que es cruzado con otro diseño ortogonal para los factores ruido, llamado **arreglo externo**. Los niveles de los factores ruido en el arreglo externo se colocan por fila y no por columna. Cada fila en el arreglo interno junto con cada columna en el arreglo externo conforman las condiciones de cada uno de los experimentos que se realizan, los resultados de los cuales se presentan en una matriz llamada **arreglo cruzado**.

Para ilustrar la estructura en que se suelen presentar estos arreglos, se muestran en la Tabla 3.1 los datos correspondientes a un experimento que Miller, Sitter, Wu y Long (1993) analizan para conocer la distorsión que tienen el engranaje y los piñones de la transmisión de un coche, durante un tratamiento de endurecimiento con calor. El estudio incluye cinco factores controlables y tres factores ruido.

El diseño factorial 2^{5-1} que tiene de encabezado los factores controlables es el arreglo interno, la matriz con los datos es el arreglo cruzado, y el diseño 2^3 , en forma transpuesta, que se encuentra arriba del arreglo cruzado es el arreglo externo. La prueba que se realizó con las condiciones indicadas en la

primera fila del arreglo interno ($X_1 = X_2 = X_3 = X_4 = -1$ y $X_5 = 1$) y las condiciones de la primera columna del arreglo externo ($Z_1 = Z_2 = Z_3 = -1$) dio un resultado de 8,0

Tabla 3.1 Datos de distorsión por calentamiento

Z_1	-1	1	-1	1	-1	1	-1	1
Z_2	-1	-1	1	1	-1	-1	1	1
Z_3	-1	-1	-1	-1	1	1	1	1
X_1	-1	-1	-1	-1	1	1	1	1
X_2	-1	-1	-1	-1	1	1	1	1
X_3	-1	-1	-1	-1	1	1	1	1
X_4	-1	-1	-1	-1	1	1	1	1
X_5	-1	-1	-1	-1	1	1	1	1
	8,0	-0,5	-1,0	13,0	12,0	1,5	14,0	14,0
	0,0	0,0	4,0	1,0	8,0	-4,0	-2,0	-0,5
	-0,5	5,0	3,5	1,0	8,0	12,5	-1,0	3,0
	4,5	-3,0	-2,0	-6,5	0,5	19,0	0,5	4,0
	5,5	2,0	6,0	10,5	16,5	2,0	6,0	4,5
	1,0	0,0	3,0	8,5	4,5	9,5	-1,0	4,5
	3,0	13,5	6,0	-2,0	4,5	5,5	18,5	-1,5
	4,5	-1,5	1,5	13,0	0,0	6,0	-2,0	19,5
	5,0	-1,5	3,5	2,0	1,0	6,5	6,5	9,5
	8,0	0,0	2,0	-1,0	2,0	4,0	8,5	-1,0
	3,5	1,5	10,5	4,0	13,0	0,5	6,5	-0,5
	1,0	-3,0	8,0	8,0	7,5	-1,0	7,5	-0,5
	12,0	6,5	8,0	9,5	7,5	7,0	12,0	12,5
	8,0	-3,0	13,5	-5,0	6,5	4,0	21,0	11,0
	17,0	11,5	3,5	-4,0	7,0	5,0	15,5	6,0
	14,0	9,5	-3,0	6,0	-1,5	10,5	8,0	1,5

Como es necesario efectuar 8 pruebas para cada fila del arreglo interno se necesitó de $16 \times 8 = 128$ experimentos. Este ejemplo ilustra uno de los aspectos duramente criticados de la metodología de Taguchi, el número en ocasiones excesivamente grande de pruebas que se tienen que realizar.

Muchos de los diseños que sugiere Taguchi, tanto para el arreglo interno como para el arreglo externo, son diseños factoriales altamente fraccionados, lo cual no permite obtener estimaciones de las interacciones entre los factores controlables, esta desventaja será comentada posteriormente.

Para el análisis de los resultados Taguchi clasifica los distintos problemas de diseño de parámetros dependiendo del objetivo que se quiere obtener con la respuesta (minimizarla, maximizarla, u obtener un valor nominal) y dependiendo de la categoría calcula un índice, llamado **cociente señal**

ruido, S/R, para medir la variabilidad inducida por los factores ruido. En la Tabla 3.2 se indica la expresión algebraica de estos índices.

Tabla 3.2 Cociente señal ruido, S/R.

Objetivo para el valor de la respuesta	S/R
“cuanto menor mejor”	$-10 \log \left[\frac{1}{n} \sum y_i^2 \right]$
“cuanto más grande mejor”	$-10 \log \left[\frac{1}{n} \sum 1/y_i^2 \right]$
“nominal es lo mejor”	$10 \log (\bar{y}^2 / s^2)$

Se calcula un cociente S/R para cada grupo de condiciones en el arreglo interno (para cada fila), así que la sumatoria se realiza en los puntos del arreglo externo. La Tabla 3.3 muestra el arreglo interno junto con los valores asociados de la media, la desviación estándar y el cociente S/R para el caso comentado y con el objetivo cuanto menor mejor.

Tabla 3.3 Arreglo interno con los estadísticos obtenidos para el ejemplo de la distorsión que tienen el engranaje y los piñones de la transmisión de un coche.

X ₁	X ₂	X ₃	X ₄	X ₅	\bar{Y}	S	S/R
-1	-1	-1	-1	1	7,6	6,6	-19,8
1	-1	-1	-1	-1	0,8	3,7	-11,0
-1	1	-1	-1	-1	3,9	4,5	-15,3
1	1	-1	-1	1	2,1	7,7	-17,5
-1	-1	1	-1	-1	6,6	4,8	-18,1
1	-1	1	-1	1	3,8	3,8	-14,3
-1	1	1	-1	1	5,9	7,0	-18,9
1	1	1	-1	-1	5,1	7,6	-18,9
-1	-1	-1	1	-1	4,1	3,5	-14,4
1	-1	-1	1	1	2,8	3,8	-13,1
-1	1	-1	1	1	4,9	4,8	-16,4
1	1	-1	1	-1	3,4	4,7	-15,0
-1	-1	1	1	1	9,4	2,5	-19,7
1	-1	1	1	-1	7,0	8,5	-20,5
-1	1	1	1	-1	7,7	6,8	-20,0
1	1	1	1	1	5,6	6,1	-18,0

Taguchi define la pérdida de calidad a través de una **función de pérdida** de no calidad y centra el análisis en minimizar esta función, que represente la

pérdida financiera en que se incurre por no cumplir con el objetivo que tiene la característica de calidad. Para el caso “nominal es lo mejor” suele usar la función de pérdida $L(y, \tau) = K(y - \tau)^2$, donde y es la característica de calidad medida, τ es el valor nominal que se pretende alcanzar, y K es una constante que se puede determinar si se conoce el coste de la función en un valor y_0 . La expresión de $L(y, \tau)$ indica que no se incurre en pérdida cuando se tiene el valor nominal y que a medida que nos distanciamos de este valor la pérdida se incrementa. Para la situación “cuanto menor mejor” el objetivo consiste en tener un valor de nominal τ de cero y por esto la función es $L(y) = Ky^2$. Con el objetivo “cuanto más grande mejor” la función de pérdida que se considera es $L(y) = K/y^2$.

El minimizar la función de pérdida es equivalente a maximizar el cociente S/R respectivo. Los cocientes S/R estimados se analizan usando el método de análisis de la varianza, con ello se identifica el grupo de condiciones de los parámetros controlables que dan a la respuesta un comportamiento robusto frente a los factores ruido.

Una vez conseguidas las circunstancias robustas se separan de los factores controlables los que no afectan al cociente S/R, conocidos como **factores de ajuste**. Los cambios en los valores de estos factores influyen en el valor de la respuesta sin afectar la variabilidad, por tal razón se recurre a ellos para obtener las mejores condiciones que permitan alcanzar el comportamiento esperado en la respuesta, conservando la condición robusta. Taguchi recomienda hacer pruebas confirmatorias para validar los resultados obtenidos (mínima variabilidad y valor de la respuesta señalado).

3.3 Comentarios al método de Taguchi

La metodología de Taguchi, desde su aparición en Occidente ha sido controvertida, no obstante existe unanimidad en reconocer sus aportaciones. Para Kacker (1985) la contribución de Taguchi al diseño de parámetros se puede dividir en cuatro categorías, ordenadas de acuerdo a su mérito: filosofía de calidad, metodología orientada a la ingeniería, diseño de experimentos y análisis de datos. Como mencionaremos más adelante las

principales críticas se centran precisamente en las dos últimas categorías: el diseño de experimentos y el análisis de los datos. Entre sus aportaciones más significativas se destaca su énfasis en la variabilidad, el modelar los factores ruido, el popularizar la idea de robustez y las innovaciones que introdujo en la modelación.

3.3.1 Aportaciones

Énfasis en la variabilidad

- Desde el punto de vista estadístico el problema deja de centrarse primordialmente en la media de la característica de calidad. Para Wu (1992) la reducción de la variación es su contribución más significativa. Aunque es cierto que la variabilidad y su control siempre ha sido una preocupación, el énfasis que le da Taguchi a esta tarea y la orientación hacia su disminución en la fase del diseño del producto ha sido innovadora.
- Este énfasis ha propiciado investigaciones en el área de los efectos de dispersión, Box y Meyer (1986), Zunica y Romero (1988), Ferrer y Romero (1994), Bergman y Hynén (1997) y Ferrer (2002), y en el modelado de la varianza, Lucas (1989), Shoemaker et al. (1989) y Welch et al. (1990), Vining y Myers (1990), Grego (1993), Lucas (1994), y Hamada y Nelder (1997).

Modelar los factores ruido

- Para Vining y Meyers (1992) el introducir sistemáticamente los factores ruido en los diseños experimentales, para conocer las relaciones que tienen con los factores controlables, es una contribución vital, que permite en muchas ocasiones crear productos y procesos robustos con una reducción sensible en los costes.
- Las medidas contra la variación son diversas (tecnología más sofisticada, mayor exigencia hacia los proveedores, mejores niveles de

capacitación del personal, etcétera) y el éxito del diseño de parámetros depende de la habilidad en identificar los factores involucrados y la existencia de ciertas interacciones entre los factores controlables y los de ruido, lo cual no es siempre posible.

Popularizar la idea de robustez

- Aunque es cierto que el efecto de robustez se buscó en aplicaciones pioneras del sector agrícola y de la industria alimenticia, no hay duda de que Taguchi ha difundido el importante papel que tienen los estudios de robustez en el diseño de los productos y procesos industriales, fomentando y expandiendo, adicionalmente, el empleo de los diseños experimentales y reduciendo la separación existente entre ingeniería y estadística.

Innovaciones

- Un aspecto innovador de la metodología de Taguchi es el ampliar la idea de variable explicativa, al introducir los conceptos de variable: (factor) controlable, ruido, señal, ajuste e indicativa.
- El diseñar para modelar de manera simultánea la media y la variabilidad es, por otra parte, una idea novedosa que ha producido aportaciones en esta forma de enfocar los problemas a través de la metodología de superficie de respuesta, Lucas (1989), Welch et al. (1990), Vining y Myers (1990), Tsui y Wu (1991), y el uso de programación no lineal, Fathi, Y. (1991), Del Castillo y Montgomery (1993), Kim y Lin (1998), K oksoy y Doganaksoy (2003).
- En la indagaci n de arreglos ortogonales que sean econ micos en el n mero de observaciones, Taguchi favorece el uso de los arreglos¹ L_{18} ($2 \cdot 3^7$), L_{18} ($6 \cdot 3^6$) y L_{36} ($2^{11} \cdot 3^{12}$), los cuales necesitan un n mero menor

¹ En la notaci n $L_a(b^c)$, a es el n mero de filas en el arreglo, b el n mero de niveles y c el n mero de factores. Es posible tener arreglos para factores que tienen diferentes n meros de nivel, por ejemplo $L_{18}(2^1 \cdot 3^7)$.

de observaciones que las combinaciones de factoriales fraccionales 2^{k-p} y 3^{k-p} . No obstante, dado lo complejo de los patrones de alias que tienen, su uso se ha centrado en el filtrado de factores (screening design).

3.3.2 Críticas

Así como hay un reconocimiento prácticamente general a las principales contribuciones de Taguchi, también hay críticas bastante extendidas, siendo las más comunes las atribuidas al número excesivo de condiciones experimentales, al no considerar las interacciones entre factores controlables, al uso de cocientes señal ruido ineficientes, al celo excesivo en el proceso de optimización y a los análisis en ocasiones erróneos.

Número excesivo de condiciones experimentales

- Esta crítica se produce esencialmente por el método de cruzar el arreglo interno con el externo, el problema señalado se complica si se emplean factores a más de dos niveles, práctica común en la metodología de Taguchi. Entre las alternativas que se presentan al gran esfuerzo experimental que en ocasiones produce el arreglo cruzado se tiene el uso de los diseños en parcelas divididas, Box y Jones (1992), Bisgaard (2000), Bisgaard y Kulahci (2001) y el empleo de la matriz combinada, Shoemaker, Tsui y Wu (1991).
- En descarga hay que decir que la economía de la fase de experimentación depende en gran medida del proceso de aleatorización y, aunque obviamente no cambia el número $n \times m$ de observaciones, se puede reducir grandemente el esfuerzo experimental requerido con el uso de parcelas divididas. La matriz cruzada no niega la noción de aleatorización, sin embargo a menudo se conduce el experimento como parcela dividida sin tener en cuenta su estructura y la estructura del error, lo cual lleva a conclusiones engañosas. Aunque el diseño habitual en parcelas divididas tiene a los factores controlables como factores de clasificación (tratamientos de la parcela completa), puede darse otro tipo de arreglo en conveniencia con los costes de experimentación y las

condiciones de operación, Shoemaker, Tsui, y Wu (1991), Box y Jones (1992), Bisgaard y Kulachi (2001).

- A favor del uso del arreglo cruzado se tiene lo sencillo que resulta el diseño de los dos arreglos por separado y lo simple que es el análisis por fila. Pozueta (2001) considera una ventaja el repetir las filas del arreglo interno en las distintas condiciones de ruido, porque de esta forma se captura la influencia de variables no consideradas.

No considerar las interacciones entre factores controlables

- Taguchi, aparentemente, para minimizar el problema del número excesivo de condiciones experimentales, recurre a diseños ortogonales altamente fraccionados en los arreglos interno y externo, en los que las interacciones están confundidas con los efectos principales. Esto demanda el supuesto de que son despreciables los efectos de las interacciones entre los factores controlables, $C \times C$, y los efectos de las interacciones entre los factores ruido², $N \times N$.
- A favor de este supuesto Phadke (1992) arguye que de existir interacciones $C \times C$ se necesita incrementar grandemente el número de experimentos para poder estudiarlas, y si estos efectos son grandes seguramente interactúan con las condiciones experimentales haciendo muy difícil que el óptimo encontrado se conserve durante la puesta de operación en la planta. Por tal razón se recomienda hacer esfuerzos para suprimir o minimizar las interacciones $C \times C$, esto aparentemente se logra si se actúa cuidadosamente en la selección de la característica de calidad que será la respuesta a medir, en la elección de los factores controlables y sus niveles y en la opción de cociente S/R que se escoja.

² Las interacciones $N \times N$ en realidad tienen un papel irrelevante en la búsqueda de las condiciones robustas, como se verá en el capítulo 4.

Para Taguchi la tarea más importante que tienen que realizar los ingenieros en el diseño de parámetros consiste en seleccionar la característica de calidad que será la respuesta a medir, de tal manera que elimine las interacciones $C \times C$. Los experimentos confirmatorios permitirán verificar la aditividad (ausencia de interacciones) en el modelo. En este sentido Lucas (1994) argumenta que existe evidencia empírica del “principio de los efectos principales”, en donde los efectos principales son más importantes que los efectos de orden superior, esto se ve confirmado con el uso exitoso que tienen este tipo de diseños.

- Con respecto a los experimentos confirmatorios hay ejemplos en los que estos fallan y es muy difícil garantizar que en todas las circunstancias se puede encontrar siempre una característica del proceso relevante y que sea aditiva a los efectos de los factores controlables, como sugiere Taguchi.
- Box (1992) tajantemente dice que no se puede defender lógicamente la necesidad de experimentar para conocer que factores tienen efecto principal significativo, asumiendo al mismo tiempo que se sabe cuáles factores son los que tienen interacciones.

Cocientes señal ruido ineficiente

- Si se ha hecho un gran esfuerzo experimental al usar el arreglo cruzado, obteniendo en ocasiones un número excesivo de observaciones, se considera que es un desperdicio el colapsar los datos en un cociente S/R, en ocasiones con grandes pérdidas de información para el análisis. Por ejemplo, ignorar el efecto que tienen los factores en la respuesta y privar de un conocimiento más amplio del desempeño del proceso.
- Box (1988) critica severamente los cocientes S/R, considerando los criterios de relevancia, eficiencia, robustez y transformación de datos. Muestra por ejemplo que el cociente S/R “cuanto mayor es mejor” puede confundir los efectos de localización con los de dispersión; que bajo el supuesto de datos normales, independientes e idénticamente distribuidos este cociente es ineficiente como medida de localización si

se le compara con la media muestral y que es excepcionalmente no robusto en presencia de datos atípicos.

- Un aspecto delicado en el uso del cociente S/R se basa en el supuesto de que el factor de ajuste tiene un efecto multiplicativo en la respuesta; como el supuesto no se puede garantizar Box (1988) propone como alternativa evaluar un rango de transformaciones y hacer el análisis en términos de la transformación, lo cual da una gran simplificación.
- Hamada y Nelder (1997) van más allá y recomiendan el uso de los Modelos Lineales Generalizados, los cuales son susceptibles de emplearse en situaciones donde las transformaciones no garantizan la divisibilidad y la parsimonia en el modelo; por ejemplo, cuando los datos son conteos o frecuencias, con la ventaja adicional que al no transformar los datos se preservan las dimensiones originales.

Un defecto, conceptualmente más general, que hace Nelder (1992) al uso del cociente S/R estriba en que se invierte el proceso normal de seleccionar primero el modelo y a partir de sus parámetros estimados elaborar para la predicción cantidades que compendian (**métricas resumen** en la notación de Pozueta (2001)) con estimaciones de su incertidumbre. La razón empírica para no crear primero las métricas resumen es que el modelado para ellas suele ser más complicado que el de la respuesta básica. Si se consigue un modelo más preciso se logra una optimización más confiable y finalmente un mejor diseño del producto o proceso. Es más fácil que el ingeniero tenga conocimientos previos para modelar directamente la respuesta de interés que para modelar el cociente S/R.

- En defensa de los cocientes S/R, Phadke (1992) subraya que en su selección no se persigue una transformación de los datos que estabilice la varianza, sino el identificar la relación ideal entre el factor señal y la característica de calidad y poder evaluar así la sensibilidad a los factores ruido.

Celo excesivo en el proceso de optimización

- Si bien es cierto que el uso de los experimentos confirmatorios puede dar lugar a un proceso iterativo en la búsqueda de las condiciones óptimas, la metodología de Taguchi ha desatendido las virtudes de planear de manera secuencial los experimentos, cosa por demás común en la experimentación y en particular en el enfoque de superficie de respuestas (filtrado de variables, cambio de región, incremento del diseño, ajuste a modelos más elaborados y exploración en la región experimental), en donde no interesa únicamente encontrar la combinación óptima de factores sino, sobre todo, tener una comprensión del sistema, Easterling (1985), Box (1988).

Análisis erróneos

- Aunque el hecho de hacer análisis erróneos no es muchas veces privativo del método sino más bien de quien lo emplea, existen en la literatura críticas a los métodos de análisis que emplea Taguchi. Steinberg y Burtsztyl (1994), Box y Fung (1986), Hurly (1994) y otros, revelan con ejemplos como el análisis propuesto por Taguchi lleva tanto a conclusiones erróneas acerca de que factores de diseño tienen influencia en la dispersión, como a recomendaciones que no encuentran la asignación óptima de los parámetros.

A pesar de la gran cantidad de críticas y las numerosas aportaciones que aparecen frecuentemente como métodos alternativos en las revistas especializadas, el método de Taguchi se emplea y demanda continuamente, quizá por su relativa simplicidad³, lo poco difundido que están los métodos alternativos en la industria y porque ninguno de estos forma un corpus orientado expresamente al diseño de parámetros.

Un aspecto que se debe tener en cuenta en el diseño de parámetros es lo que Lucas (1994) llama “Hacer algo mejor que un proceso robusto”, si la influencia de los factores ruido en la respuesta es muy grande se pueden

³ Para Lucas (1992) la elegancia de las contribuciones de Taguchi radica en su sencillez esencial.

tener mejores resultados diferenciando el producto de acuerdo a condiciones más limitadas en los factores ruido, esto se observa por ejemplo en el sector automovilístico, donde es mejor tener diferentes tipos de aceite para el motor en función del tipo de carburante que contar con un aceite robusto para todo tipo de motores.

3.4 Planes experimentales

La búsqueda de opciones alternativas a los diseños cruzados que palien los defectos que se les achacan (requerir generalmente de una gran cantidad de trabajo experimental y en general no permitir estimar las interacciones entre los factores controlables), han generado diversas aportaciones que adecuan métodos estadísticos ya existentes, como los diseños en parcelas divididas (split-plot), la metodología de superficie de respuesta y los planes experimentales óptimos.

3.4.1 Diseños en parcelas divididas

El número generalmente excesivo de pruebas que demanda la estructura de la matriz cruzada no se ve necesariamente reflejado en una gran cantidad de esfuerzo experimental. Esto sucede porque en ocasiones no se ejecuta el diseño en una forma completamente al azar, sino como un diseño en parcelas divididas⁴ con restricción en la aleatorización. Esto se observa cuando para cada fila del arreglo interno se construye un prototipo, y este se prueba en las condiciones de los factores de ruido señaladas en el arreglo externo.

Box, G. y Jones

Como se sabe, una manera de reducir la cantidad de esfuerzo experimental consiste en emplear los diseños en parcelas divididas. Se acostumbra usar estos diseños cuando resulta complejo o costoso aleatorizar completamente

⁴ Box y Jones (1992) comentan que aparentemente Taguchi algunas veces ignoraba al momento de hacer el análisis, que el diseño se había conducido en forma de parcela dividida.

el orden de ejecución de los experimentos. Box y Jones (1992) analizan esta posibilidad cuando se cuenta con una matriz cruzada, y para ello dan dos opciones: que los factores controlables conformen los factores de clasificación (tratamientos de la parcela completa) y los tratamientos de las subparcelas contengan los factores ruido (como se hace usualmente en el método de Taguchi), o que los factores ruido estén asociados a los factores de clasificación, con los factores controlables asignados a las subparcelas.

No es intrascendente donde se ubiquen los factores controlables y por ende los de ruido, ya que generalmente en los diseños en parcelas divididas los efectos principales de los factores de clasificación se estiman de manera menos precisa que los efectos principales de las subparcelas o de las interacciones. Esto hace razonable el considerar a los factores ruido como factores de clasificación, dado que no suele interesar los efectos principales de estos factores, y no importaría que se estimen con menos precisión. Esta forma de proceder no es la que acostumbra la metodología de Taguchi, Michaels (1964), Box y Jones (1992), Bisgaard (2000), Bisgaard y Kulahci (2001).

Box y Jones (1992) muestran que los diseños en parcelas divididas son mejores que los diseños completamente aleatorizados, tanto en lo referente al número de operaciones experimentales⁵ exigidas como a la eficiencia del arreglo. Esto último hace que se estime con mayor precisión las importantes interacciones entre los factores controlables y los factores ruido. Una posible desventaja, de los diseños en parcelas divididas, es que tienen dos estructuras de error diferentes, y por tanto los grados de libertad de la estimación de la varianza del error para el diseño completamente aleatorio se dividen en dos, lo que puede dar lugar a que se tengan pocos grados de libertad para la prueba de los factores de clasificación.

También proponen otros diseños que pueden reducir el trabajo experimental y elevar la eficiencia del arreglo, en particular analizan el diseño de bloque

⁵ El número de operaciones experimentales tiene que ver con el número de veces que hay que reiniciar las condiciones experimentales. En un diseño completamente al azar se debe reiniciar las condiciones experimentales antes de cada ensayo.

en franja (strip-block), el cual tiene más ventajas que el diseño en parcelas divididas y por consiguiente que el diseño cruzado.

Bisgaard

Para el arreglo interno y para el arreglo externo se acostumbra usar diseños factoriales fraccionados. Entre los aspectos que se debe tener en cuenta durante la elección de estos diseños se tiene la resolución que adquiere el arreglo cruzado y los patrones de confusión; Bisgaard (2000) estudia estos aspectos para los diseños factoriales fraccionados a dos niveles. Una manera que sugiere para disminuir el número de operaciones experimentales reside en confundir ciertas interacciones, parcial o completamente, lo que se logra con la técnica de parcelas divididas con diseños confundidos (*split plot confounding*), los diseños así obtenidos tienen una resolución mayor entre los factores de las subparcelas que la que se logra con el diseño cruzado.

En función de los patrones de confusión se puede escoger entre un diseño completamente al azar y un diseño en parcelas divididas, cuidando que las interacciones entre los factores controlables y los factores ruido no queden confundidos con interacciones de dos factores, en este sentido un diseño de resolución III puede llegar a ser una buena alternativa.

Una propiedad interesante que menciona Bisgaard se refiere a la resolución del diseño, si los arreglos interno y externo son de resolución mayor o igual a 3, entonces, las relevantes interacciones entre los factores de clasificación y los factores de las subparcelas ($C \times N$) estarán confundidas con interacciones de más de dos factores.

Para el análisis de los diseños sin réplicas, se da un método para separar los efectos estimados que están asociados a las parcelas completas de los que corresponden a las subparcelas, con ellos se puede analizar la significación de los efectos usando para cada grupo una representación en papel probabilístico normal. Si no se hace esta separación y se analizan los efectos estimados como un diseño completamente al azar, se pueden dejar de percibir efectos significativos al graficarlos en el papel probabilístico normal.

3.4.2 Arreglo combinado

Shoemaker, Tsui y Wu

Box, Bisgaard y Fung (1988) en su artículo “An Explanation and Critique of Taguchi’s Contributions to Quality Engineering” alaban del enfoque japonés el uso extensivo e innovador de los diseños de experimentos estadísticos, pero critican muchas de las técnicas de diseño y análisis que usan, por ser ineficientes e innecesariamente complicadas. Señalan que se llevan a cabo investigaciones en la Universidad de Wisconsin para encontrar formas más eficientes de diseñar los experimentos. Shoemaker, Tsui y Wu (1991) proponen como alternativa al uso del arreglo cruzado el considerar juntos los factores controlables y los factores ruido en un diseño convencional, a este le llaman **diseño combinado**. Ya Lucas (1989) y Welch et al. (1990) habían usado un arreglo de estas características, al proponer la metodología de superficie de respuestas como un enfoque alternativo al método de Taguchi, pero no habían abundado en sus características y ventajas.

Las mejoras que con este tipo de diseño se tienen son: un número frecuentemente inferior de observaciones requeridas, una gran flexibilidad para seleccionar los efectos estimables, con lo que se pueden modelar las interacciones entre los factores controlables, y lo que es más importante, el contar en numerosas ocasiones con diseños de resoluciones superiores a los que daría el arreglo cruzado.

Shoemaker et al. (1991) para ejemplificar las ventajas del arreglo combinado analizan diversas posibilidades para un experimento en el que se quiere controlar la variabilidad en la capa de silicón que se pone en ciertos circuitos integrados, se tiene 8 factores controlables a dos niveles cada uno y dos factores ruido, uno a dos niveles y el otro a cuatro. El diseño para el arreglo interno es un factorial 2^{8-4} de resolución IV y el diseño del arreglo externo es un 2×4 . Con los 128 valores de la matriz cruzada se pueden estimar los doce efectos principales de los factores, las 32 interacciones dobles entre los factores de control y los factores ruido, $C \times N$, ninguna interacción doble entre los factores de control, $C \times C$, ni entre los factores

ruido y todas las interacciones triples entre estos últimos factores. Como se comentó anteriormente, Taguchi prima las interacciones $C \times N$ sobre las interacciones $C \times C$ y $N \times N$, lo que representa en casos como este un desperdicio de la labor experimental.

Es un hecho que en la mayoría de los experimentos está presente el principio de escasez de efectos, **effect sparsity**, (solo un número pequeño de los factores que se analiza tienen efecto significativo sobre la variable respuesta) con base en ello Shoemaker et al. lograron reducir considerablemente el número de experimentos, a través del arreglo combinado, al sacrificar algunas interacciones que los ingenieros juzgaron a priori insignificantes o intrascendentes. De esta manera se redujo a la mitad el número de experimentos a partir de una selección cuidadosa de los generadores del diseño. Los efectos estimables que obtienen son los 12 efectos principales, 12 efectos de las interacciones $C \times C$, 19 de las $C \times N$, 2 del tipo $N \times N$ y algunas interacciones triples. Los resultados obtenidos con este diseño combinado fueron cualitativamente los mismos que arrojó el análisis con el arreglo cruzado, pero con un ahorro significativo en costes y esfuerzos experimentales y con un conocimiento más amplio del comportamiento y relación de los factores con la respuesta.

Myers, Khuri y Vinning

El adaptar el arreglo cruzado a un diseño combinado permite el uso de cualquiera de los diversos tipos de diseño que uno encuentra en la metodología de superficie de respuesta: diseños factoriales, diseños factoriales fraccionales, diseños compuestos, diseños de Box-Behnken, diseños D-óptimos, etc. Consideremos el ejemplo de Myers et al. (1992), en el cual un ingeniero está interesado en conducir un experimento con tres factores controlables y dos factores ruido, el diseño cruzado seleccionado tiene para los factores controlables X_1 , X_2 y X_3 un arreglo interno $L_9(3^3)$ (cada uno de los factores está a tres niveles) y para las variables ruido Z_1 y Z_2 el arreglo externo es un factorial 2^2 , ver Tabla 3.4. Para la ejecución de este diseño cruzado se necesitan 36 experimentos si se realiza completamente al azar.

Tabla 3.4 Arreglo interno y externo de la matriz cruzada del ejemplo de Myers et al. (1992)

Arreglo interno			Arreglo externo	
X ₁	X ₂	X ₃	Z ₁	Z ₂
-1	-1	-1	-1	-1
-1	0	0	1	-1
-1	1	1	-1	1
0	-1	0	1	1
0	0	1		
0	1	-1		
1	-1	1		
1	0	-1		
1	1	0		

El diseño combinado que se plantea como alternativa al diseño cruzado es un diseño centrado compuesto, Box (1954), con tres experimentos centrales, en la Tabla 3.5 se muestra este diseño. Con el fin de ahorrar espacio la primera fila de la tabla equivale a un diseño 2^{5-1} de resolución V (16 experimentos).

Tabla 3.5 Diseño combinado: Diseño centrado compuesto

X ₁	X ₂	X ₃	Z ₁	Z ₂	
±1	±1	±1	±1	±1	Diseño 2^{5-1} de resolución V (16 experimentos)
-1	0	0	0	0	
1	0	0	0	0	
0	-1	0	0	0	
0	1	0	0	0	
0	0	-1	0	0	
0	0	1	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	

En este diseño combinado se tienen 25 experimentos, lo que da un ahorro de 11 experimentos, y tal vez lo más importante es que se pueden estimar los efectos lineales y cuadráticos de los factores controlables junto con sus interacciones, además se pueden estimar los efectos lineales y las interacciones de los factores ruido, y también las importantes interacciones entre los factores controlables y los factores ruido.

Lorenzen y Villalobos

Para que sea eficiente el uso del arreglo cruzado cuando se aplican diseños altamente fraccionados en los arreglos interno y externo, se requiere que no existan interacciones significativas $C \times C$ con el fin de evitar confusiones con los efectos principales. Para Taguchi esto se puede conseguir con una elección correcta de la característica de calidad que se usará como respuesta. Lorenzen y Villalobos (1990) al comparar diferentes enfoques para el diseño robusto de parámetros concluyen que aun asumiendo que las interacciones entre los factores controlables son despreciables, el arreglo combinado tiene una detectabilidad superior para los efectos principales, con las ventajas adicionales de que los supuestos se pueden verificar y se tiene mejores propiedades de robustez.

Kacker

En contra del uso del arreglo combinado Kacker (1992) menciona tres defectos que tiene este enfoque:

- Al basarse en la regresión clásica, el supuesto de varianza constante es irreal en muchos casos, debido a que es improbable que todas las fuentes de variación (factores ruido) se incluyan como variables regresoras.
- Falla también el supuesto subyacente de que los ingenieros pueden proporcionar una lista priorizada de los factores y las interacciones importantes, no siendo posible usar la propuesta de Shoemaker et al. (1991) para reducir el número de pruebas a través del arreglo combinado eliminando interacciones irrelevantes.

- El método es muy sensible a los datos perdidos.

Pozueta

Pozueta (2001) al comparar las ventajas e inconvenientes del arreglo cruzado menciona que este es más sencillo e intuitivo para el usuario, al existir diseños por separado para los factores de control y los factores ruido, además la propia disposición de la matriz predispone para un análisis sencillo por filas, y siempre se puede pasar de un diseño cruzado a un arreglo combinado para detectar posibles discrepancias entre los análisis correspondientes.

Lucas

Lucas (1994) no ve desventajas en el hecho de ignorar los efectos de las interacciones de los factores controlables, ya que muchos de los procesos de filtrado de variables (screening designs) se basan en diseños de resolución III de Plackett-Burman, donde los efectos principales se encuentran confundidos con las interacciones de dos factores. Como es sabido, en estos diseños se estudian un gran número de factores con un número reducido de experimentos, esto último se logra gracias a la estructura de estos diseños, los cuales han demostrado ampliamente su eficiencia. Asimismo, empíricamente se detecta en los experimentos de superficie de respuestas que los efectos principales lineales suelen ser más importantes que los efectos de orden superior, aun en situaciones donde los experimentadores consideran que las interacciones y la curvatura pueden ser significativas.

Como alternativa al arreglo cruzado, Lucas (1994) y Borkowski y Lucas (1997), recomiendan usar una estructura especial para el diseño combinado que denominan **diseño compuesto de resolución mixta**, constituido por un diseño 2^{k-p} en el que la parte correspondiente a los factores controlables es de resolución mayor o igual a V, la parte asociada a los factores ruido tiene resolución mayor o igual a III, y las interacciones $C \times N$ no están confundidas con efectos principales o interacciones $C \times C$, a este diseño se le añaden $2p$ puntos estrella y N_0 puntos centrales, siendo p el número de

factores controlables. Este tipo de diseño permite estimas de todos los parámetros del siguiente modelo:

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \beta_{ii} X_i^2 + \sum_{i=1}^{p-1} \sum_{k=i+1}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j Z_j + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k + \varepsilon \quad (3.1)$$

donde X_1, X_2, \dots, X_p son los factores controlables y Z_1, Z_2, \dots, Z_q los factores ruido. Con ello se tienen los efectos que interesan con un arreglo cruzado cuando los factores controlables tienen tres niveles y dos los factores ruido, más los efectos de las interacciones C×C. Aunque da la impresión que los diseños compuestos de resolución mixta requieren de un elevado número de puntos, Lucas muestra, para diferentes números de factores en los arreglos interno y externo, que es comparable el número de experimentos que requiere el mínimo arreglo cruzado de Taguchi con el que necesita un diseño compuesto de resolución mixta equivalente (con tres puntos centrales). En contra de algunas propuestas, recomienda no usar los diseños centrados compuestos más pequeños porque tienen una eficiencia (D-eficiencia y/o G-eficiencia) muy baja.

3.4.3 Diseños óptimos

Un concepto esencial desarrollado por Fisher en relación a los diseños experimentales, y que ha marcado su desarrollo, es que la estructura del diseño determina si la respuesta a muchas de nuestras preguntas es accesible o no. Es conocido que un diseño factorial 2^k no permite ajustar un modelo que tenga términos cuadráticos en algunos de los factores, lo cual si es factible con un diseño factorial 3^k o un diseño centrado compuesto.

Atkinson y Donev

Si consideramos el modelo de regresión lineal usual⁶ $E(\mathbf{Y}) = \mathbf{F}\boldsymbol{\beta}$, donde \mathbf{Y} es el vector $n \times 1$ de respuestas, $\boldsymbol{\beta}$ es el vector $p \times 1$ de parámetros

⁶ Se considera que los errores del modelo son independientes, aditivos, y distribuidos normalmente con varianza σ^2

desconocidos y \mathbf{F} es la matriz de diseño extendida, donde las filas de \mathbf{F} son funciones de las m variables explicativas x_i , que dan origen a la matriz $n \times m$ de diseño⁷ \mathbf{X} , entonces el estimador mínimo cuadrático de $\boldsymbol{\beta}$ es:

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}'\mathbf{F})^{-1} \mathbf{F}'\mathbf{y} \quad (3.2)$$

Y la matriz de covarianzas del estimador es:

$$\text{var } \hat{\boldsymbol{\beta}} = \sigma^2 (\mathbf{F}'\mathbf{F})^{-1} \quad (3.3)$$

Esta última expresión nos indica que la variabilidad que presenten las estimas de los parámetros, y por tanto la precisión de las mismas, depende de la matriz de información $\mathbf{F}'\mathbf{F}$ y en particular del diseño empleado.

A fines de la década de los cincuenta Kieffer y Wolfowitz sientan las bases de la teoría de los **diseños experimentales óptimos**. En estos diseños se buscan n puntos de la región de experimentación que permitan la mejor estimación de los parámetros del modelo lineal considerado, bajo cierto criterio de optimización. Un aspecto interesante es que estos n puntos, o condiciones experimentales, no tienen que ser necesariamente distintos. Con el diseño óptimo se pretende encontrar la matriz de diseño \mathbf{X} que cumpla con un criterio específico de optimización, la mayoría de las veces el criterio está en función de la matriz de información, $\mathbf{F}'\mathbf{F}$, o de la matriz de covarianzas del estimador mínimo cuadrático de $\boldsymbol{\beta}$.

Para simplificar el problema de encontrar el diseño óptimo se ignora la restricción de que el número de pruebas en cada punto del diseño debe ser entero, esto frecuentemente se puede hacer al remplazar el diseño \mathbf{X} de n

⁷ Por ejemplo, si se tiene solo una variable explicativa x , la matriz de diseño $\mathbf{X}' = [-1, 1/3, 1/3, 1]$, y el modelo que se desea ajustar es $E(Y_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$, entonces la matriz de diseño extendida tiene la siguiente forma:

$$\mathbf{F} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1/3 & 1/9 \\ 1 & 1/3 & 1/9 \\ 1 & 1 & 1 \end{bmatrix}$$

pruebas por un **diseño continuo** ξ ; una vez que se encuentra el diseño óptimo continuo ξ^* se puede encontrar un buen diseño para n pruebas a partir de una aproximación entera. Por ejemplo, si la región de interés para el diseño es el intervalo $[-1, 1]$ y se desea ajustar el modelo de primer orden:

$$E(Y) = \beta_0 + \beta_1 x \quad (3.4)$$

entonces el diseño (D-óptimo) que minimiza la varianza de $\hat{\beta}_1$ es aquel que ubica la mitad de los pruebas en el punto $x = -1$ y la otra mitad en $x = 1$, esto se suele denotar como:

$$\xi = \left\{ \begin{array}{cc} -1 & 1 \\ 1/2 & 1/2 \end{array} \right\} \quad (3.5)$$

En general, un diseño continuo en el que se hacen pruebas en k puntos de la región de diseño se escribe

$$\xi = \left\{ \begin{array}{c} x_1 \ x_2 \ \dots \ x_k \\ w_1 \ w_2 \ \dots \ w_k \end{array} \right\} \quad (3.6)$$

siendo la primera fila los valores de los factores en los k puntos de diseño y la segunda fila los pesos asociados (proporción de puntos). En el ejemplo no tenemos problemas para identificar cuantas pruebas se asignan a cada punto del diseño óptimo cuando el número de pruebas es par. No es lo mismo cuando el número es non, para $n=3$ se tiene que colocar una prueba en uno de los extremos del intervalo, y en el otro los restantes, con lo que no se consigue un diseño D-óptimo. Los diseños para un número específico de pruebas se conocen como **diseños exactos**, en la práctica todos los diseños son exactos, y frecuentemente se pueden encontrar por aproximaciones enteras al diseño óptimo continuo.

Federov

En la búsqueda del diseño óptimo se debe encontrar la mejor ubicación de los puntos en la región de diseño y una medida de distribución que representa la proporción de pruebas que se asignan a los puntos, los pesos

w_1, w_2, \dots, w_k de la expresión 3.6 componen la medida de distribución. Federov (1972) encuentra un diseño óptimo bajo los criterios D y G para ajustar un polinomio de segundo grado con dos factores, en la región cuadrada con vértices $(\pm 1, \pm 1)$, la medida de distribución asigna a cada uno de los vértices un 14,6% de las pruebas, un 8% a cada punto medio de los bordes y un 9,6% en el centro. Con 13 pruebas es posible tener un diseño exacto que sea una buena aproximación a este diseño óptimo, colocando dos pruebas en cada vértice ($13 \times 14,6\% = 1,9$) y una en el centro y en cada punto medio de los bordes, ver Figura 3.1.

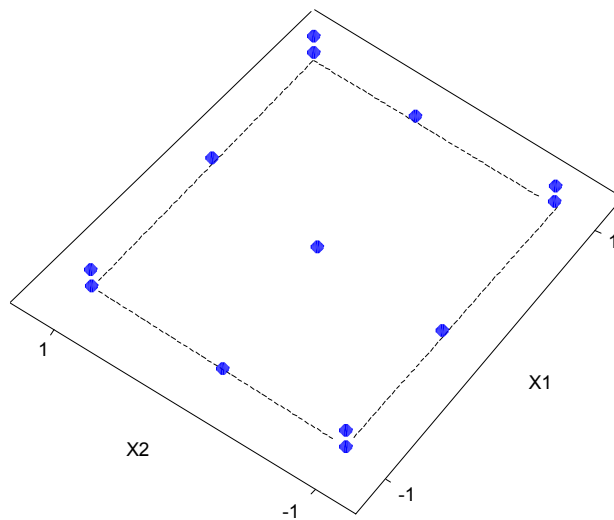


Figura 3.1 Diseño con 13 pruebas a partir del diseño D/G óptimo de Federov

Un aspecto interesante que se observa en el diseño de la figura anterior es que los n puntos, o condiciones experimentales, no tienen necesariamente que ser distintos. Los diseños graficados en la Figura 3.2 están tomados de Atkinson y Donev (1992) y corresponden a un diseño exacto D-óptimo para ajustar un modelo de superficie de respuesta de segundo orden en dos factores, en la misma región de diseño que el ejemplo de Federov, y para seis pruebas. Lo atractivo de estos diseños es que los puntos encontrados no se encuentran todos en los vértices, como nos tienen acostumbrado los

diseños factoriales 3^k , y en el ejemplo b) los puntos que no están en los vértices tampoco se encuentran en los puntos medios de los bordes o en el centro.

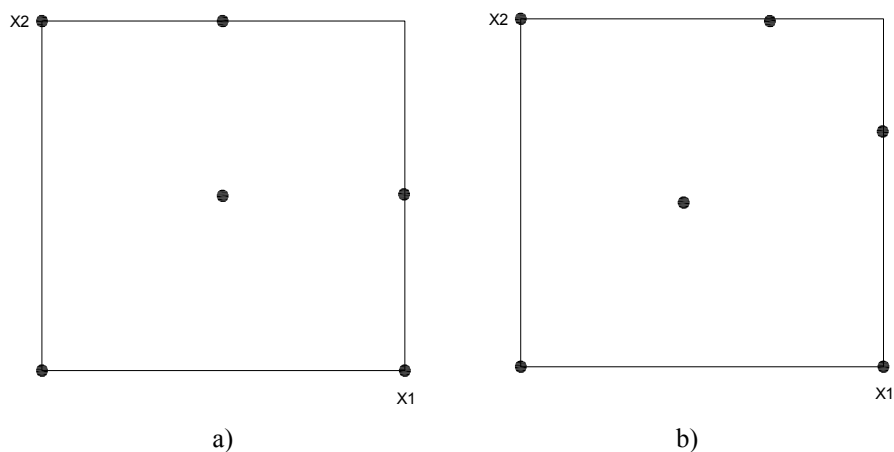


Figura 3.2 Dos ejemplos de un diseño exacto D-óptimo con seis pruebas para ajustar una superficie de respuesta de segundo orden

Los criterios de diseño óptimo se denotan con nombres alfabéticos, y por esta razón el tema es conocido como **óptimalidad alfabética**. El más común es el diseño D-óptimo, que maximiza el determinante de la matriz de información. La interpretación estadística que tiene este hecho es la de minimizar el volumen de la región de confianza para β , con ello se obtienen las estimas más precisas de los parámetros del modelo. Se suele escribir la matriz de información de un diseño continuo como $M(\xi)$.

Un criterio asociado al D-óptimo es el G-óptimo, el cual minimiza el máximo valor de la varianza estandarizada⁸ de la respuesta ajustada, en la región de diseño X , esto es:

$$\min_{\xi} \bar{d}(\xi) = \min[\max_x d(x, \xi)] \quad (3.7)$$

⁸ La varianza estandarizada se define como:

$$d(x, \xi) = n \text{Var}(\hat{y}(x)) / \sigma^2$$

Existe un teorema de equivalencia que permite determinar si un diseño propuesto es D-óptimo; basta con confirmar si $\bar{d}(\xi) = p$, con p igual al número de parámetros en el modelo.

Para comparar la eficiencia que tiene un diseño cualquiera con respecto al óptimo se construyen índices (reciben el nombre del criterio), por ejemplo, el índice de eficiencia usando el criterio D, D-eficiencia, para un diseño ξ está dada por

$$D_{ef} = \left[\frac{|M(\xi)|}{|M(\xi^*)|} \right]^{1/p} \quad (3.8)$$

Este valor representa el porcentaje de esfuerzo experimental que se aprovecha con dicho experimento.

Si interesa que un subconjunto s de parámetros del modelo se estime lo más preciso que sea posible, se reemplaza el criterio D-óptimo por el D_s-óptimo. El criterio de optimización consiste en maximizar $|\mathbf{F}'\mathbf{F}|/|\mathbf{F}_2'\mathbf{F}_2|$, para lo cual se necesita reescribir el modelo de la siguiente forma

$$E(\mathbf{Y}) = \mathbf{F}\boldsymbol{\beta} = \mathbf{F}_1\boldsymbol{\beta}_1 + \mathbf{F}_2\boldsymbol{\beta}_2 \quad (3.9)$$

siendo $\boldsymbol{\beta}_1$ un vector $s \times 1$.

Box, M. y Draper

Hoel (1965) muestra que los diseños factoriales 3^k son D-óptimos, con las k variables codificadas tomando valores de -1 a 1 y niveles -1, 0, 1. Lo mismo se puede decir de los diseños factoriales 2^k , con niveles -1 y 1, Box, M. y Draper, N. (1971). Los modelos que permiten ajustar los diseños 2^k y 3^k son excesivos porque generalmente las interacciones de orden mayor a 2 no son significativas, y en este sentido parece lógico preguntarse por los mejores diseños cuando se especifica el número de puntos disponibles y tal vez alguna restricción en el tipo de diseño. Box y Draper encuentran para el diseño factorial fraccional 2^{3-1} los diseños D-óptimos cuando el número de pruebas es de 4, 5 o 6 y lo mismo para el diseño 3^2 cuando el número de pruebas es un número entre 6 y 18. Consideran preferible usar un diseño D-

óptimo que emplear indiscriminadamente un diseño factorial clásico, entre las ventajas que señalan se encuentra que:

- El experimentador se ve obligado a pensar en el modelo más adecuado antes de que se lleve a cabo la experimentación, este modelo puede ser lineal o no lineal.
- El número de pruebas no está restringido a una potencia de 2 o de 3, lo cual ofrece una gran flexibilidad y ahorro.
- Los diseños factoriales están limitados a regiones cuboidales, lo que puede ser una restricción importante en algunas aplicaciones, en cambio los D-óptimos pueden encontrarse en regiones de diseño de cualquier tipo.
- La búsqueda de diseños D-óptimos permite manejar modelos en los que la matriz de covarianzas no es de la forma $\sigma^2 I$, lo que amplía las posibilidades de la modelación.

Mitchell y Box

Mitchell y Box (1974) desarrollaron un algoritmo simple y flexible, DETMAX, para la construcción de un diseño experimental D-óptimo, que hasta la fecha se sigue usando, con algunas variantes. La idea básica consiste en iniciar con n pruebas seleccionadas al azar, e ir mejorando el diseño con base en aumentar una prueba que incremente al máximo $|F'F|$ y eliminar una prueba que ocasione la mínima disminución en el valor del determinante.

Lucas

Lucas (1976) compara tres clases de diseños, que comúnmente se usan en el ajuste de modelos de superficie de respuesta cuadráticos, los diseños compuestos de Box y Wilson, DC, los diseños de Box y Behnken, DBB, y los diseños *Uniform Shell*, US, de Doehlert. Los criterios de comparación son la D-eficiencia y la G-eficiencia. En general, los diseños US no son tan

eficientes como los DC o los DBB. Dependiendo del número de factores, en ocasiones los diseños compuestos son mejores que los de Box y Behnken, y viceversa.

Welch

Welch (1982) propone un algoritmo, basado en el método de búsqueda *Branch and Bound*, para construir diseños experimentales D-óptimos, que permita trabajar con valores de n suficientemente pequeños. Como el diseño D-óptimo puede ser inapropiado cuando el número de observaciones no es grande con respecto al número de parámetros, Welch (1984) generaliza el algoritmo DETAMAX de Mitchell para criterios G-óptimo y V-óptimo⁹.

Atkinson

Atkinson (1995) considera cuatro problemas que no se acostumbra a resolver con la metodología de superficie de respuesta y que son susceptibles de tratarse con diseños óptimos:

- Región de diseño restringida. Una cualidad, con repercusiones prácticas, que tienen los algoritmos para diseños óptimos es el poder construir diseños para un número arbitrario n de pruebas y, adicionalmente, en regiones de diseño que no son típicas. Atkinson y Donev (1992) encuentran un diseño exacto D-óptimo con 20 pruebas para un modelo de segundo orden, donde la región de diseño está restringida porque existen combinaciones de los dos factores que harían que el mecanismo de encendido que se desea estudiar no funcionara, o por el contrario se dañara o destruyera, ver Figura 3.3.

⁹ Minimiza la varianza promedio, i.e. $\min_{\xi} \sum_{j=1}^m d(x_j, \xi) / n = \bar{d}$

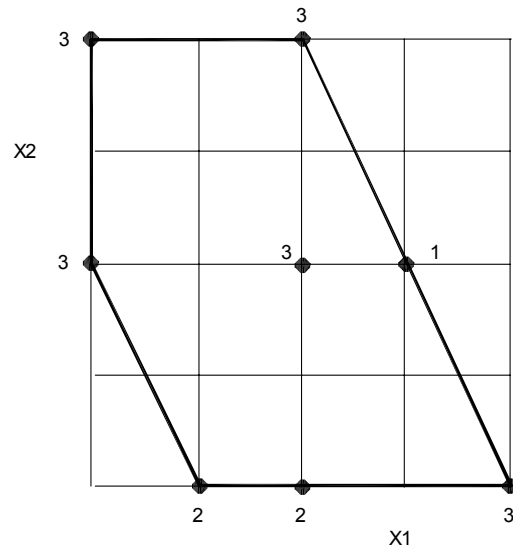


Figura 3.3 Ejemplo de un diseño D-óptimo con 20 pruebas para el análisis del comportamiento de encendido de un motor de coche

- Diseños con factores continuos y con factores discretos. En lugar de repetir el diseño correspondiente a los factores cuantitativos, en cada uno de los niveles de los factores cualitativos, como se suele hacer, lo cual demanda un gran esfuerzo experimental, se propone tratar el problema directamente con los métodos de diseño óptimo.
- Diseños de superficie de respuesta en bloques. Se pueden conseguir diseños adecuados, modificando el algoritmo para diseños exactos, con la finalidad de contar con un número determinado de pruebas en cada nivel del factor.

Atkinson comenta también la aplicabilidad de los diseños óptimos en los siguientes problemas:

- Diseños de mezclas, en donde la restricción¹⁰ que se impone a los factores define un simplex, resultando factible buscar en esta región un diseño óptimo.
- Diseños para la estimación de la varianza. Modelando simultáneamente la media y la varianza, a través del modelo

$$y = f'(\mathbf{x})\boldsymbol{\beta} + \sigma \{Var(g'(z)\boldsymbol{\alpha})\} \varepsilon \quad (3.10)$$

donde ε es una variable aleatoria normal estándar.

- Diseños localmente D-óptimos para modelos no lineales, Diseños para modelos lineales generales, Diseños óptimos Bayesianos, y Diseños compuestos.

Mays y Easter

Mays y Easter (1997) hacen un análisis para comparar diseños factoriales replicados estándar con diseños óptimos (D-óptimo e I-óptimo¹¹), para superficies de respuesta con dos variables de diseño, a dos o tres niveles, en presencia de una estructura de varianza heterogénea causada por la presencia de efectos de dispersión, y con modelos lineales de primero y segundo orden. Los resultados de su análisis, con cuatro estructuras de varianza a considerar, indican que los diseños típicos en general no son óptimos y su eficiencia decrece en la medida en que aumenta la severidad de los efectos de dispersión. Aunque la eficiencia de los diseños estándar para el modelo de segundo orden es generalmente mayor que para el modelo de primer orden, es recomendable usar diseños alternativos.

En 1999 Mays estudia los diseños centrados compuestos óptimos en presencia de efectos de dispersión. Los diseños centrados compuestos, CCD, propuestos por Box y Wilson (1951) para ajustar modelos de segundo

¹⁰ La restricción es del tipo $\sum_{j=1}^m x_j = 1 \quad (x_j \geq 0)$

¹¹ El criterio I-óptimo, conocido como varianza integrada, minimiza la varianza de predicción promediada en la región de diseño.

orden, como es sabido, representan una alternativa a los diseños factoriales 3^k y reducen considerablemente el número de pruebas cuando se tiene un gran número de factores. Mays encuentra la asignación óptima de N ensayos experimentales a las ubicaciones del CCD, con los criterios D-óptimo e I-óptimo, con tres estructuras de varianza (lineal, cuadrática y cuadrática inversa), cinco grados de dispersión y para 2, 3, y 4 variables. Analiza la eficiencia de estos diseños óptimos CCD en referencia a los diseños CCD convencionales.

Romero Z.

Romero Z. (2002) extiende el problema de encontrar diseños experimentales óptimos al ámbito del diseño robusto de parámetros. Desarrolla un procedimiento para obtener diseños Ds-óptimos continuos, asumiendo irrelevantes las interacciones entre los factores de ruido¹², para el modelo que contiene efectos lineales y cuadráticos en los factores controlables X_i , efectos lineales en los factores ruido Z_k , e interacciones dobles entre los distintos factores (control \times control, control \times ruido y ruido \times ruido), esto es:

$$E(Y) = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j Z_j + \sum_{i=1}^q \sum_{k=i+1}^q \gamma_{ik} Z_i Z_k + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k \quad (3.11)$$

Encuentra diseños de tamaño mínimo que permiten estimar todos los parámetros relevantes del modelo dejando al menos 5 grados de libertad residuales, además indica en cuales de los puntos encontrados se pueden hacer réplicas sin perder la optimalidad (hasta cinco réplicas), y proporciona también para cada diseño los valores de las varianzas de las estimaciones de los parámetros, donde el único valor desconocido es el de la varianza residual (la cual se puede estimar a partir de las réplicas)

El procedimiento se basa en reducir el problema inicial, de gran complejidad, a la búsqueda de dos de tres pesos óptimos, que se asignan a

¹² En el siguiente capítulo se verá que estas interacciones no influyen en la búsqueda de las soluciones robustas

los tres subconjuntos de puntos del diseño¹³ 3^p , que soportan el diseño óptimo. A partir de una adaptación del algoritmo DETMAX de Mitchell construye planes experimentales Ds-óptimos exactos, para dos o tres factores de control y hasta tres factores de ruido. El número de pruebas de diseño garantiza una Ds-eficiencia de al menos 90 % y una G-eficiencia superior al 50%.

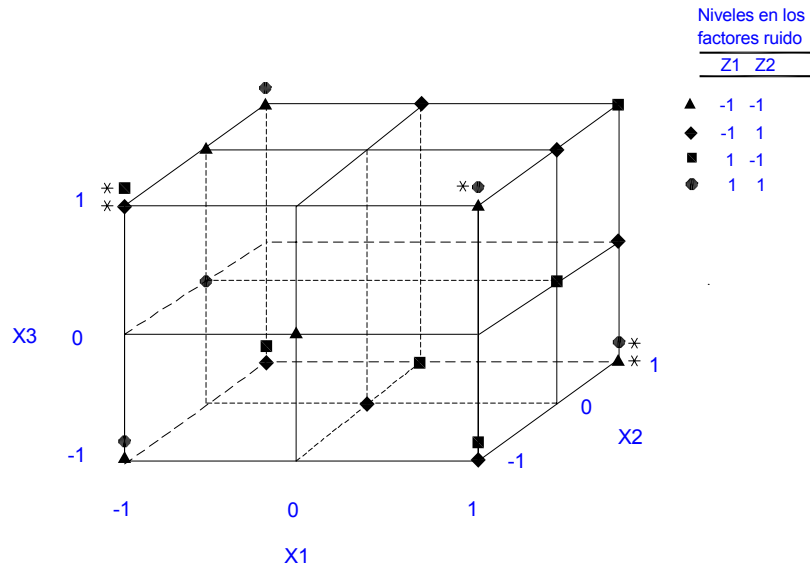


Figura 3.4 Diseño Ds-óptimo con 24 pruebas, para estimar los efectos relevantes del modelo (3.11) con tres factores controlables y dos factores ruido, las cinco condiciones experimentales con asterisco son las posibles réplicas.

El modelo (3.11) con 3 factores controlables y 2 factores ruido tiene 19 parámetros relevantes, si este modelo se quisiera estimar con una matriz cruzada de Taguchi se requeriría de 108 pruebas, con una Ds-eficiencia del 88%, en cambio el diseño Ds-óptimo exacto que se muestra en la Figura 3.4 solo necesita de 24 pruebas (19+5) si no se hacen réplicas. El diseño óptimo con las réplicas (29 pruebas) tiene una Ds-eficiencia del 93,7%

¹³ p es el número de factores de control en el modelo

Box, G.

Box, G. (1982) duda de la utilidad de los enfoques de optimización para definir los diseños experimentales, entre sus argumentos menciona la relevancia de la decisión, la pertinencia del modelo y el criterio de selección:

- Relevancia de la decisión. Las decisiones que más influyen en los resultados de una investigación no dependen realmente de criterios estadísticos. Estas decisiones conciernen con la elección de la región de interés, la selección de las variables que participarán, sus rangos y localizaciones y las transformaciones. Una vez elegida la región, y bajo el supuesto de que el modelo ajusta perfectamente y que solo interesan las propiedades de la varianza, cualquier conjunto de puntos que cubra esta región, de manera razonablemente uniforme, será igual de adecuado.
- Pertinencia del modelo. El aforismo “Todos los modelos son incorrectos, algunos modelos son útiles” es particularmente cierto para las funciones de tipo polinomial que no pretenden más que graduar localmente la verdadera función. Empíricamente se ha observado que las aproximaciones no necesitan ser muy buenas para que el método de superficie de respuesta trabaje apropiadamente.
- Criterio de selección. La función de información en sí misma es la medida más directa de la conveniencia del diseño. La mejor acción, en este caso, es seleccionar el diseño directamente de aquellos que den la función de información pertinente, y no con un criterio que encuentra un valor extremo. El problema del diseño no es tanto una cuestión de escoger el diseño que incrementa la información total, sino de extender la información total en la forma deseada.

3.5 Análisis

El uso del cociente señal ruido para el análisis de los resultados ha sido duramente criticado, entre las críticas más recurrentes se tiene: el colapsar un número excesivo de observaciones en un solo cociente, en ocasiones con grandes pérdidas de información en el análisis, el desconocer el tipo de influencia que tienen los factores en la respuesta, el no poder tratar datos que son conteos o frecuencias, y el asumir que el factor de ajuste tiene un efecto multiplicativo en la respuesta. Las opciones alternativas que se dan se pueden agrupar en función de uno de los dos enfoques que se dan a la solución del problema: usar un estadístico diferente o modelar directamente la respuesta. En el primer grupo quedarían las medidas del desempeño que son independientes del ajuste y las transformaciones de los datos, el segundo congregaría los métodos de superficie de respuesta, los métodos lineales generalizados y los métodos de programación no lineal.

3.5.1 Medidas del desempeño que son independientes del ajuste

Leon, Shoemaker y Kackar

Leon, Shoemaker y Kackar (1987) identificaron que existen muchos problemas reales en donde el cociente S/R no es independiente de los factores de ajuste, y el uso de ellos puede llevar a una selección de los niveles de los parámetros de control distante del óptimo. Proponen otro tipo de estadísticos diferentes a los cocientes S/R, que toman en cuenta la existencia de factores de ajuste, y que llaman “medidas del desempeño independientes del ajuste” (performance measures independent of adjustment, PerMIA). El cociente señal ruido de Taguchi es una medida del desempeño que funciona de manera deficiente cuando no es independiente de los parámetros de ajuste.

Cuando existen factores de ajuste, se puede derivar la PerMIA de la función de pérdida conocida y de la forma general del modelo de la función de transferencia $f(\mathbf{X}, \mathbf{Z})$. Se considera que la respuesta Y es función (función de transferencia) de los factores de control \mathbf{X} y de los factores de ruido \mathbf{Z} ,

estos últimos considerados aleatorios; de aquí que la respuesta sea aleatoria. A semejanza con las ideas de Taguchi, se descomponen los factores de control en dos grupos, $\mathbf{X} = (\mathbf{d}, \mathbf{a})$, y se emplea un procedimiento de optimización de dos pasos, en el primero se encuentran los factores de control que afectan a la dispersión, \mathbf{d} , y posteriormente se seleccionan los factores de ajuste, \mathbf{a} .

Se demuestra que el cociente S/R para el caso “nominal es lo mejor” no depende de \mathbf{a} , esto es que \mathbf{a} afecta la localización pero no la dispersión cuando la función de transferencia es del tipo multiplicativa,

$$Y = \mu(\mathbf{d}, \mathbf{a})\varepsilon(\mathbf{Z}, \mathbf{d}) \quad (3.12)$$

donde $E(Y) = \mu(\mathbf{d}, \mathbf{a})$ es una función estrictamente monótona, en este caso el procedimiento de Taguchi minimiza la pérdida cuadrática. Este tipo de función de transferencia exige que los componentes del error no contengan factores de ajuste, y su expresión implica que la varianza de la respuesta Y transmitida por los factores de ruido es proporcional a su valor esperado.

Si se desea trabajar con la escala logarítmica y se asume la misma función de transferencia, y por función de pérdida,

$$L(y, \tau) = K[\log(y) - \log(\tau)]^2, \quad (3.13)$$

siendo τ el valor nominal de la respuesta, entonces el estadístico PerMIA, que minimiza la función de pérdida con el procedimiento de dos pasos está dado por:

$$\text{PerMIA} = \log[\text{Var}(\log(Y))]. \quad (3.14)$$

Esta medida fue sugerida por Box (1986) como una alternativa al cociente S/R de Taguchi, la cual tiene mejores propiedades estadísticas, pero puede tener problemas cuando la función de pérdida tiene un valor cercano a cero para τ .

Si se tiene una expresión aditiva, en lugar de multiplicativa, para la función de transferencia, del tipo,

$$Y = \mu(\mathbf{d}, \mathbf{a}) + \varepsilon(\mathbf{Z}, \mathbf{d}), \quad (3.15)$$

con $E[\varepsilon(\mathbf{Z}, \mathbf{d})] = 0$ y $L(y, \tau) = K(y - \tau)^2$, entonces en lugar del cociente S/R se debe usar $\text{PerMIA} = \text{Var}(Y)$ en el proceso de optimización.

3.5.2 Transformación de los datos

Box, G.

Un aspecto delicado en el uso del cociente S/R se basa en el supuesto de que el factor de ajuste tiene un efecto multiplicativo en la respuesta. Box (1988) analiza el caso “nominal es lo mejor” y muestra que si la transformación logaritmo estabiliza la varianza entonces

$$S/R \approx \text{const} - \text{const} \ln S_{\log y} \quad (3.16)$$

Como el supuesto no se puede garantizar, Box propone como alternativa evaluar transformaciones que garanticen parsimonia en el modelo e independencia entre la media y la desviación estándar, Box y Cox (1964) y Box y Fung (1983). Al hacer el análisis en términos de la transformación se logra una gran simplificación. Una vez escogido el criterio de desempeño adecuado, se debe ser cuidadoso en la manera de estimarlo. Por ejemplo, si la medida de desempeño seleccionada es $1/\mu$ entonces el estimador máximo verosímil es

$$1/\bar{y} \quad \text{y no} \quad \left(\sum_i 1/y_i \right) / n$$

En relación a los casos cuanto más grande mejor y cuanto menor mejor Box los considera extremadamente ineficientes. Ambos criterios son ideas de localización, por lo cual ante supuestos de datos independientes distribuidos en forma normal, la media muestral es la medida de localización con mejores propiedades estadísticas.

3.5.3 Metodología de Superficie de Respuesta

Vining y Myers

Un gran número de las aportaciones alternativas al método de Taguchi, para el diseño de productos robustos, se ha basado en la metodología de Superficie de Respuesta (MSR). Vining y Myers (1990) proponen considerar a la media y la varianza como respuestas de interés, y aplicar el método de respuesta dual desarrollado por Myers y Carter (1973) para alcanzar un valor deseado para la media y otro para la varianza.

El modelo sugerido para cada una de las respuestas es polinomial de segundo orden. El método dual optimiza una respuesta primaria sujeta a una restricción adecuada en el valor de la respuesta secundaria. Los tres casos que trata Taguchi quedarían conformados de la siguiente manera:

- “nominal es lo mejor”, significa mantener μ en el valor nominal, τ , mientras se minimiza σ^2 .
- “cuanto más grande mejor”, significa hacer μ tan grande como sea posible, mientras se controla σ^2 .
- “cuanto menor mejor”, significa hacer μ tan pequeño como sea posible, mientras se controla σ^2 .

El procedimiento que proponen consta de cuatro pasos

1. Hacer experimentos de búsqueda en la respuesta primaria, como el método de la máxima pendiente, para determinar la región de interés apropiada.
2. Usar cualquiera de los diseños clásicos de segundo orden de superficie de respuesta, para ajustar el modelo de la respuesta primaria.
3. Analizar los resultados del experimento con la metodología de respuesta dual.

4. Hacer al menos un experimento confirmatorio en la localización encontrada.

En la situación “nominal es lo mejor”, consideran su método más general y más natural, superando al procedimiento PerMIA, ya que este último no es aplicable cuando los factores de control no se pueden dividir en factores de ajuste y factores que afectan la dispersión. Bajo el supuesto de error aditivo en el modelo, se puede usar el método que ellos sugieren para proponer posibles factores de ajuste, dado que se desarrollan modelos separados para la media y la varianza.

Myers, Vinning y Khuri (1992), señalan que conjuntamente con la estimación de las condiciones óptimas la metodología de superficie de respuestas da una alta prioridad a la búsqueda de una mejor comprensión del proceso. Esto se consigue explorando simultáneamente las superficies de respuesta generadas por la media y la varianza de la respuesta. Los modelos de la media y la varianza los derivan del modelo para la respuesta, usando las propiedades de los operadores $E(\cdot)$ y $Var(\cdot)$. Resaltan que es más importante en la industria el modelar la varianza del proceso que modelar la media, algo que no pasa comúnmente.

En la búsqueda de las condiciones óptimas, además de la optimización restringida del método dual que sugieren en su anterior artículo, proponen el empleo de gráficos de contorno superpuestos, un ejemplo de ello se trata en el capítulo 4.

Las ventajas que dan a la MSR sobre el método de Taguchi son:

- Los arreglos combinados que se usan en esta metodología para incluir los factores ruido proporcionan la estimabilidad deseada, y generalmente se tienen ahorros en el plan experimental al usar estos arreglos.
- El enfoque de superficie de respuesta dual da una comprensión superior del proceso, permiten el uso de diseños en forma secuencial y el interesado puede emplear cualquier métrica en la respuesta.

- Se puede determinar qué factores de control tienen efecto sobre la dispersión y que factores de ruido incrementan la variabilidad y no están sujetos a control por no interactuar con los factores de control.

Lucas

Lucas (1994) para lograr un proceso robusto recomienda analizar un diseño de Taguchi como un diseño de superficie de respuesta, de esta manera se puede aprovechar todos los procedimientos de diagnóstico que se han desarrollado para esta metodología. No ve inconvenientes en usar los diseños altamente fraccionados, y frecuentemente saturados, que emplea la metodología de Taguchi, empero recomienda no saturar completamente el diseño, guardando cuatro o cinco grados de libertad para diagnósticos y estimación del error.

Un concepto interesante que menciona el artículo de Lucas es el poder conseguir en ocasiones algo mejor que un proceso robusto. La elaboración de diferentes productos, que respondan cada uno a usuarios o grupos de condiciones específicas, puede dar mejores resultados que únicamente un producto robusto. El análisis a través de superficies de respuesta permite percibir esta situación al incluir tanto los factores de control como los de ruido. El análisis de Taguchi no permite conocer este tipo de información, al ignorar los efectos principales en las variables del arreglo externo se limita la identificación de efectos significativos en las variables ruido, y son precisamente estas las que pueden indicarnos que la diferenciación del producto puede ser exitosa. También menciona la idea expresada por Nelder (1988) de que al analizar el experimento, suele convenir transformar los datos de tal manera que los efectos de las interacciones sean desdeñables, facilitando la interpretación.

Khattree

Khattree (1996) extiende el trabajo de Myers, Khuri y Vinning (1992) al caso en que se requiere usar bloques porque algunos de los factores de ruido no se pueden combinar en un mismo esquema experimental; Taguchi denomina estas variables factores indicativos. Para el caso en que se tienen

dos bloques presenta un modelo para cada grupo de respuestas, en notación matricial estos son:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}_1\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1 \\ \mathbf{y}_2 &= \mathbf{X}_2\boldsymbol{\beta} + \mathbf{Z}_2\boldsymbol{\gamma}_2 + \boldsymbol{\varepsilon}_2 \end{aligned} \quad (3.17)$$

donde \mathbf{X}_1 y \mathbf{X}_2 son las matrices de diseño para las variables de control, \mathbf{Z}_1 y \mathbf{Z}_2 son matrices para las variables de ruido y los vectores $\boldsymbol{\beta}$, $\boldsymbol{\gamma}_1$ y $\boldsymbol{\gamma}_2$ representan los coeficientes de regresión correspondientes. Sólo existe un $\boldsymbol{\beta}$ para ambos modelos para poder dar la misma predicción cuando se tiene el mismo grupo de variables de control. El proceso de estimar los parámetros de los modelos se simplifica grandemente si los diseños experimentales en los dos bloques son idénticos y ortogonales a \mathbf{Z}_1 y \mathbf{Z}_2 , en este caso, y bajo el supuesto de homocedasticidad, la estimación se hace por mínimos cuadrados, combinando los datos de los dos diseños en uno solo, de la manera siguiente:

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{Z}_1 & \mathbf{0} \\ \mathbf{X} & \mathbf{0} & \mathbf{Z}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} \quad (3.18)$$

Myers, Kim y Griffiths

Myers, Kim y Griffiths (1997) basados en el procedimiento que sugiere Vining y Myers (1990) de modelar simultáneamente las superficies de respuesta de la media y la varianza, proponen la búsqueda de condiciones robustas dentro de la región de confianza que se puede construir para la zona que tiene la mínima varianza del proceso. Si se asume para la respuesta el modelo

$$\mathbf{y} = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{x}\Delta\mathbf{z} + \boldsymbol{\varepsilon}$$

Esta zona de mínima varianza se estima a partir de $\hat{\mathbf{I}}(\mathbf{x}) = \hat{\boldsymbol{\gamma}} + \hat{\boldsymbol{\Delta}}'\mathbf{x} = \mathbf{0}$. Para diseños en los cuales los efectos principales de ruido son mutuamente

ortogonales con las interacciones control ruido, se puede encontrar una región de confianza en \mathbf{x}_0 , del tipo

$$\Pr \left\{ \frac{\hat{\mathbf{I}}(\mathbf{x}_0) [\mathbf{C}(\mathbf{x}_0)]^{-1} \hat{\mathbf{I}}(\mathbf{x}_0)}{n_z \hat{\sigma}_\varepsilon^2} \leq F_{\alpha, n_z, g.l.E} \right\} = 1 - \alpha \quad (3.19)$$

con $\hat{V}(\hat{\mathbf{I}}(\mathbf{x})) = \mathbf{C}(\mathbf{x}) \hat{\sigma}_\varepsilon^2$.

Hay que mencionar que, en la mayoría de las propuestas basadas en la MSR, la estimación del modelo de la varianza se basa en el modelo estimado de la respuesta, por lo que es especialmente importante que este modelo esté bien ajustado; esto origina que las decisiones acerca de los valores óptimos de los factores de control sean muy sensibles a la manera como el modelo de la respuesta fue identificado. Se tiene ejemplos en los que la minimización directa de la varianza obtenida a partir del modelo ajustado no detecta efectos importantes de variabilidad que si detecta el análisis convencional de Taguchi; el análisis gráfico de las interacciones entre los factores de control y los de ruido en ocasiones no indican los niveles de los factores de control que minimizan la variable respuesta. Pozueta (2001) recomienda una estrategia de tres pasos, tendente a evitar errores en la selección de la parte asociada a los factores de ruido en el modelo.

Las recomendaciones que se hacen en los procedimientos de Taguchi, para la selección cuidadosa de la característica de calidad que se considerará como respuesta, es aplicable también en la MSR, debido a que una elección inadecuada puede inducir no linealidad, haciendo muy difícil encontrar un modelo ajustado parsimonioso. De la misma manera, la selección de los factores debe ser tal que simplifique la relación con la respuesta, sin olvidar las técnicas de transformación de datos para simplificar el modelo.

3.5.4 Modelo Lineal Generalizado

Grego

Para el problema de optimización simultánea de la media y la variación de un proceso, Grego (1993) considera el modelo lineal de la respuesta

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (3.20)$$

donde \mathbf{e} es un vector de variables aleatorias e_i que se distribuyen normalmente con media cero y con una estructura del error

$$\text{Var}(e_i) = \exp(\boldsymbol{\gamma}'\mathbf{z}_i) \quad (3.21)$$

Los factores x_i influyen a la media y los factores z_k a la varianza del proceso, algunos x_i coinciden con ciertos z_k si existen factores que afectan a la media y a la varianza. Para analizar los modelos (3.20) y (3.21) uno puede ajustar los modelos

$$\bar{\mathbf{Y}} = \mathbf{X}\boldsymbol{\beta} \quad (3.22)$$

y

$$\ln(\mathbf{s}^2) = \mathbf{Z}\boldsymbol{\gamma} \quad (3.23)$$

No se modela directamente \mathbf{s}^2 porque cada s_i^2 tiene una distribución Gama¹⁴. Una posibilidad de modelar respuestas que no tienen distribución Normal es usando un modelo lineal generalizado, GLM, McCullagh y Nelder (1989). Grego propone el empleo de ellos para modelar directamente la varianza como una alternativa al empleo de la transformación logarítmica, en concreto con el uso del modelo multiplicativo Gama para la distribución del error y el logaritmo como función de enlace (*link*).

Como método de selección de factores activos, el empleo de modelos lineales generalizados para modelar \mathbf{s}^2 , en contraposición con el modelado de $\ln(\mathbf{s}^2)$ que hace uso de la teoría normal, tiene la ventaja de que todos los resultados aparecen con la métrica original en lugar de la escala

¹⁴ La distribución de s_i^2 es una ji-cuadrada, que es un caso particular de la distribución Gama.

transformada, con lo que se facilita la interpretación del modelo resultante, una cualidad importante en todo buen modelo. Suplementariamente, en la práctica el modelo multiplicativo Gama para el error funciona de modo satisfactorio con muestras pequeñas.

Hamada y Nelder

Hamada y Nelder (1997) destacan que los datos provenientes de la característica de calidad frecuentemente representan conteos o cocientes de conteos, en cuyo caso los supuestos del modelo lineal clásico no se cumplen. Suele suceder que se transforman los datos buscando varianza constante, aunque no siempre se consiga la linealidad y la normalidad; por ejemplo, si los errores siguen una distribución de Poisson entonces no existe una transformación que satisfaga los supuestos simultáneamente. Estos problemas se pueden evitar si se emplean modelos lineales generales al especificar la componente sistemática del modelo y, al mismo tiempo, permite seleccionar separadamente una distribución adecuada para el componente del error en el modelo. Esto se logra transformando los valores hipotéticos de la media en lugar de los datos con la función *link* adecuada.

Para modelar conjuntamente la media y la dispersión usan un par de modelos lineales generales enlazados, uno para la media y el otro para la dispersión. Cada uno tiene su propia variable respuesta, su función de enlace que determina la escala en la cual se asume que los efectos de las variables explicativas son aditivos y su conjunto de variables que contribuyen al predictor lineal. Factores que son importantes para la dispersión pueden o no concurrir en el modelo para la media; esto no sucede en los procedimientos que emplean MSR cuando se deriva el modelo de la varianza a partir del modelo de la media. El algoritmo para ajustar los modelos es una extensión del algoritmo GLM estándar, el cual consiste en un procedimiento iterativo, que inicia con el ajuste para el modelo de la media asumiendo que los valores ajustados para la dispersión son conocidos, de aquí se estima el modelo para la dispersión usando los valores ajustados de la media para formar la variable respuesta; los ajustes se van alternando entre el modelo de la media y el de la dispersión hasta que la convergencia se logra.

La clase de modelos que se pueden abordar es lo suficientemente amplia para tratar las situaciones donde la variable respuesta es continua, o es el producto de conteos, todo esto de forma unificada.

El empleo de GLM en ocasiones elimina interacciones removibles, con lo que se consigue aditividad en los factores. Esto facilita la selección de los niveles que optimizan la respuesta esperada, y también, si se está usando diseños factoriales fraccionales, se simplifica la identificación de los efectos significativos al desaparecer interacciones que son alias.

3.5.5 Programación no lineal

Fathi

Fathi (1991) propone un enfoque de programación no lineal para resolver el problema del diseño de parámetros. Expresa el problema como un modelo para el diseño de parámetros (PDM), que consiste en encontrar un conjunto de valores para $\mu_1, \mu_2, \dots, \mu_n$ que minimicen σ_y^2 sujeto a cierta restricción, esto es:

$$\begin{array}{ll} \text{minimizar} & z = \sigma_y^2 \\ \mu_1, \mu_2, \dots, \mu_n & \\ \text{sujeto a} & f(\mu_1, \mu_2, \dots, \mu_n) = \tau, \\ & (\mu_1, \mu_2, \dots, \mu_n) \in M. \end{array} \quad (3.24)$$

Si se conoce la forma analítica de la función $\sigma_y^2(\mu_1, \mu_2, \dots, \mu_n)$ entonces el problema se reduce a un modelo de optimización no lineal restringido. Como la expresión funcional de σ_y^2 suele ser desconocida se tiene que recurrir a técnicas de estimación, de las cuales menciona básicamente tres: Simulación de Montecarlo, Experimentos diseñados estadísticamente y Aproximación funcional. Fathi desestima los dos primeros debido a que el procedimiento de Montecarlo tiene, para el problema planteado, requisitos

computacionales excesivos, el segundo método demanda menos esfuerzo computacional pero proporciona poca asistencia para encontrar la solución óptima; por esta razón considera que el método de Taguchi potencialmente lleva a aproximaciones crudas y a conclusiones erróneas, lo que justifica el empeño que hace esta metodología en hacer experimentos confirmatorios.

Asumiendo que la función f es continua y dos veces derivable en cualquier punto, se estima dicha función usando series de Taylor, alrededor de un punto dado ξ ; la estimación de σ_y^2 en cualquier valor μ se logra con la expresión analítica de la varianza de la función aproximada f , evaluada en μ . Esta estimación de la varianza de la respuesta es razonablemente exacta en $\mu = \xi$ y pierde exactitud a medida que μ se aleja de ξ .

Para resolver el PDM se propone un procedimiento iterativo que denomina método de aproximación sucesivo de varianza cuadrática, SQVAM. Para los puntos iniciales del algoritmo se recomienda encontrarlos de manera rudimentaria, como puede ser con un diseño grandemente fraccionado en el espacio de $\mu_1, \mu_2, \dots, \mu_n$. El primer paso del procedimiento consiste en construir el PDM reemplazando a $\sigma_y^2(\mu_1, \mu_2, \dots, \mu_n)$ con la expresión $\text{Var}[q(x)]$, donde $q(x)$ es la aproximación de $f(x)$. El segundo paso usa un algoritmo de programación de aproximación cuadrática sucesiva (GRG o SUMT).

Fathi recomienda verificar la solución óptima μ^{op} , encontrando en forma independiente el valor de $\sigma_y^2(\mu^{op})$; por ejemplo, a través de un modelo de simulación Montecarlo; en caso de no coincidir los valores de la varianza en el óptimo se debe regresar al segundo paso del procedimiento SQVQM. En general el procedimiento no converge al punto mínimo por usar una aproximación cuadrática a f , no obstante se sugiere el uso repetido del método empleando varios puntos de inicio, que pueden ser determinados con diseños altamente fraccionados en el espacio de $\mu_1, \mu_2, \dots, \mu_n$.

Entre las características más significativas del SQVAM se tiene:

- No requerir de medidas del desempeño, como PerMIA o cociente señal ruido, como medidas intermedias para encontrar el óptimo, sino hacerlo directamente optimizando la función objetivo, la cual puede representar la pérdida esperada.
- Permitir que cualquier relación funcional entre las variables independientes se incorpore al modelo, como una restricción adicional. De esta manera sólo se incrementan los requisitos computacionales, sin imponer restricciones a su aplicabilidad.
- Manejar varias características de calidad a la vez, construyendo como función objetivo una combinación lineal adecuada de las respectivas varianzas.
- No restringirse a dos o tres niveles para las variables, como es el caso de la mayoría de los enfoques basados en diseños experimentales.

Del Castillo y Montgomery

Del Castillo y Montgomery (1993) retoman el método de respuesta dual propuesto por Vining y Myers (1990), el cual en esencia es un método de programación no lineal, dado que ambas respuestas son usualmente funciones cuadráticas. Usan el algoritmo del gradiente reducido generalizado, GRG, porque es un método robusto del tipo primal, con lo cual la búsqueda de la solución óptima se realiza a través de la región factible, y la facilidad para obtener una rutina GRG lo hace ideal para implementarlo. Este procedimiento no requiere que ambas funciones sean cuadráticas, y la región delimitada por las restricciones puede ser cuboidal o esférica.

Las mejoras que ha tenido el método de respuesta dual en la década pasada se han hecho en la forma de plantear el problema, y en los avances en los algoritmos de optimización; Kim y Lin (1998) dan una información pormenorizada de ellas. Uno de los mayores inconvenientes que tienen los enfoques al problema (optimizar la media sujeto a mantener la desviación estándar por debajo de un valor especificado) es la restricción de la segunda

respuesta, dado que normalmente se desconoce un valor aceptable para ella. Köksoy y Doganaksoy (2003) proponen una forma más flexible de plantear el problema considerándolo un caso de programación no lineal multiobjetivo. De esta manera se tiene:

“cuanto más grande mejor”:

$$\begin{aligned} & \{\max \hat{\mu}, \min \hat{\sigma}\} \\ & \text{sujeto a } x \in R \end{aligned}$$

“cuanto menor mejor”:

$$\begin{aligned} & \{\min \hat{\mu}, \min \hat{\sigma}\} \\ & \text{sujeto a } x \in R \end{aligned}$$

donde R puede ser una región cuboidal o esférica. El caso “nominal lo mejor” se puede tratar iterando una solución basada en los casos anteriores. Köksoy y Doganaksoy adaptan el software NIMBUS (*Nondifferentiable Interactive Multiobjective BUndle-based Optimization System*) para ilustrar su método con tres casos de la literatura.

3.6 Estudio de efectos sobre la dispersión

El énfasis que el método de Taguchi hace en la reducción de la variabilidad como parte integral de los procesos de mejora de la calidad, sin duda, ha propiciado un creciente interés en el modelo estructural de la varianza y en la estimación de los parámetros asociados a dicho modelo, los que se denominan **efectos de dispersión**. La forma en que Taguchi estudia el efecto que sobre la media y la dispersión tienen los factores se basa en replicar las condiciones experimentales, comúnmente a través de introducir en el diseño los factores de ruido.

Una forma de identificar los efectos de dispersión, a partir de experimentos replicados, consiste en evaluar la varianza muestral para cada condición experimental; a partir de estas varianzas, o el logaritmo de ellas, como

respuesta se identifican los efectos significativos. Una desventaja evidente es el elevado número de observaciones que se generan cuando se desea probar un número razonable de factores, razón por la cual algunos autores han propuesto métodos para identificar los factores de dispersión basados en experimentos sin réplicas. Algunas referencias de estos métodos son Box y Meyer (1986), Zunica y Romero V. (1988), Ferrer y Romero V. (1994), Bergman y Hynén (1997) y Ferrer (2002).

Box, G. y Meyer

Box y Meyer (1986), proponen un método económico para identificar en la etapa primaria de filtrado los factores de localización y los factores de dispersión activos, asumiendo el supuesto común en esta etapa de la existencia de escasez de efectos, *effect sparsity*, (solo un número pequeño de los factores que se analiza tienen efecto significativo sobre la variable respuesta). Este método, para experimentos factoriales 2^{k-p} no replicados, se desarrolla en dos etapas:

- Se estima el modelo lineal para la media de la respuesta y se obtienen los residuos, con ello se logra eliminar los efectos de posición y, por tanto, los residuos contienen la información acerca de la variabilidad.
- El efecto que en la dispersión tiene cada factor se determina a partir del logaritmo del cociente entre las varianzas de los residuos asociadas a los dos niveles del factor, $\ln(S^2_{i+} / S^2_{i-})$. Si las varianzas de los residuos en los niveles alto y bajo del factor son iguales, el cociente se distribuye aproximadamente normal con media cero. Al tener factoriales no replicados, se emplean gráficos para el diagnóstico, por ejemplo papel probabilístico normal y pruebas de significación no formales.

Nair y Pregibon

Nair y Pregibon (1988) hacen una extensión del método de Box y Meyers al caso con réplicas y lo comparan con un método propuesto por Bartlett y Kendall (1946). Encuentran que este último tiene un desempeño razonable para identificar el modelo pero es ineficiente cuando hay pocos efectos no

nulos; en cambio la extensión del método de Box y Meyer es altamente eficiente bajo el supuesto de escasez de efectos, pero no es recomendable usarlo cuando el supuesto no es razonable, como sucede cuando la selección de los factores se basa considerablemente en apreciaciones subjetivas de expertos.

Bergman y Hynén

Con el objeto de eliminar la dependencia que tiene el método de Box y Meyer con el número de factores de posición (el poder de identificar los efectos de dispersión depende de qué efectos de posición se tiene), Bergman y Hynén (1997) proponen un método que usa las observaciones originales y donde el estadístico propuesto sigue una distribución F, lo cual reduce la dependencia con los métodos de discriminación subjetivos, como el papel probabilístico normal.

El método de Box y Meyers tiene la virtud de trabajar con datos no replicados, con todas las ventajas que esto acarrea; sin embargo, se encuentra limitado a diseños factoriales a dos niveles. Esta restricción impide en la práctica estudiar los efectos de dispersión en una gran cantidad de datos que no provienen de experiencias diseñadas, especialmente cuando algunas de las variables fluctúan de forma continua en la muestra.

Zunica y Romero V.

Zunica y Romero V. (1988) abordan el problema de estudiar los efectos sobre la dispersión, con datos que pueden tener características que son normalmente muy restrictivas: de muestras pequeñas, en ausencia de repeticiones, sin provenir de situaciones experimentales, que fluctúan de forma continua en la muestra. Proponen un modelo que permite identificar y estimar no solo los efectos de dispersión sino también los de posición, adaptan el método de dos etapas "*Two-Step Procedure*", TSP, propuesto por Harvey (1976) para estimar un modelo que asume efectos multiplicativos sobre la varianza de la variable respuesta, del tipo:

$$\begin{aligned} y_j &= \boldsymbol{\beta}' \mathbf{x}_j + u_j & j = 1, 2, \dots, N \\ \text{var}(u_j) &= \sigma_j^2 = \exp(\boldsymbol{\alpha}' \mathbf{z}_j) \end{aligned} \quad (3.25)$$

donde \mathbf{z}_j es el vector de las variables para las cuales se desea conocer el efecto que tienen en la dispersión, dicho modelo incluye como caso particular el modelo acostumbrado de regresión lineal. El método bietápico de estimación, TSP, consiste en estimar α mediante una regresión ordinaria usando como respuesta el logaritmo natural de cuadrado de los residuos estimados de la regresión de y sobre las x_j , y como variables regresoras las variables z_j

Comparan, en un contexto asintótico, la estimación que da el método TSP con la que se obtiene al usar el procedimiento alternativo de estimación que sugieren, consistente en aplicar iteradamente el procedimiento común de mínimos cuadrados ponderados, MCPI, con la finalidad de tomar en cuenta el carácter heterocedástico del modelo; también comparan estos métodos de estimación con el procedimiento Máximo Verosímil, MV, encontrando resultados muy próximos entre los estimadores MCPI y MV, y en algunos casos difiriendo ambos con los estimadores TSP.

Para mostrar las posibilidades de identificación de los efectos de dispersión y de posición, aplican los tres métodos de estimación a dos problemas, uno proveniente de la industria automovilística, en donde se obtienen estimaciones muy parecidas con los tres métodos. El segundo ejemplo lo hacen a partir de los datos de un trabajo de Taguchi y Wu (1985), el cual fue estudiado por Box y Meyer (1986), además del efecto que encuentran estos autores Zunica y Romero V. identifican otro efecto altamente significativo sobre la dispersión, ilustrando con ello el potencial que tiene la metodología propuesta, en particular el procedimiento MCPI.

Ferrer y Romero V.

Ferrer y Romero V. (1993, 1994, 1995) estudian las propiedades de los tres estimadores para el caso de pocas observaciones, lo que sin duda es más útil desde un punto de vista de su empleo en la práctica industrial. Encuentran que el mejor estimador para este caso es el TSP, aunque resulta sesgado,

razón por la cual proponen un nuevo estimador insesgado, al que denominan AVMT, producto de la media aritmética de los estimadores TSP y MV. Estudian el efecto que tiene la presencia de datos anómalos en el sesgo y la varianza de los tres estimadores, concluyendo que el estimador TSP es el más recomendable, cuando se cuenta con pocas observaciones, al unir tres propiedades esenciales:

- Mayor eficiencia, al tener menor variabilidad.
- Menor sensibilidad a valores anómalos.
- Extrema sencillez de cálculo, al requerir únicamente del uso de mínimos cuadrados.

Ferrer

El artículo de Ferrer (2002) hace una revisión amplia del desarrollo de los estudios de los efectos de dispersión, indicando ventajas e inconvenientes de ellos. Comenta los modelos de funciones de varianza del tipo

$$\begin{aligned} E(Y_j) &= f(X_i, \boldsymbol{\beta}) \\ \text{var}(Y_j) &= \sigma_j^2 g^2(E(Y_j), Z_k, \boldsymbol{\alpha}) \end{aligned} \quad (3.26)$$

donde las Y_j son independientes, $\boldsymbol{\beta}$ es el vector de efectos de posición, $\boldsymbol{\alpha}$ es el vector de efectos de dispersión, X_i y Z_k las variables que influyen sobre la media y la dispersión, f es la función de respuesta media, g es la función de varianza y σ es un parámetro de escala desconocido.

De los modelos de función de varianza destacan los modelos logarítmicos, como el estudiado por Zunica y Romero V. (1988). Ferrer describe para este modelo las propiedades de los estimadores TSP, MCPI y MV cuando se tienen muestras pequeñas, también menciona el estudio hecho cuando se tiene la presencia de datos anómalos, en donde se concluye que los sesgos que se presentan en las estimaciones depende del efecto de dispersión real. Señala que el análisis de efectos de posición y dispersión es un problema intrínsecamente difícil y que para el estudio de las propiedades de los

estimadores TSP, MCPI y MV, en el caso de pocas observaciones, es necesario investigar el comportamiento de los estimadores en modelos más complejos, y estudiar el problema del sesgo de estimación en los diseños factoriales fraccionados a dos niveles cuando se emplean los métodos de estimación iterativos.

3.7 Conclusiones

Son innegables las aportaciones que ha hecho la metodología de Taguchi a la mejora de la calidad: con su innovador concepto de diseño robusto de parámetros, su énfasis en la variabilidad presente en productos y procesos, la orientación que da a la disminución de esta en la fase de diseño, con su idea de diseñar para modelar de manera simultánea la media y la variabilidad, el introducir sistemáticamente los factores ruido en los diseños experimentales para conocer las relaciones que tienen con los factores controlables, la difusión del importante papel que tienen los estudios de robustez en el diseño de los productos y procesos industriales, fomentando y expandiendo adicionalmente el empleo de los diseños experimentales y reduciendo la separación existente entre ingeniería y estadística.

Más polémicos son los defectos que se le encuentran a su metodología, como son: el número excesivo de condiciones experimentales, el no considerar las interacciones entre factores controlables, el uso de cocientes señal ruido ineficientes, el celo excesivo en el proceso de optimización y los análisis en ocasiones erróneos.

Estas críticas han dado lugar a interesantes y valiosas aportaciones, desde nuevos planes experimentales hasta métodos de identificación de factores, algunas de estas aportaciones adecuan métodos estadísticos ya existentes, como los diseños en parcelas divididas (split-plot), la metodología de superficie de respuestas y los planes experimentales óptimos, otras son aportaciones que han ampliado la metodología estadística, como las medidas del desempeño que son independientes del ajuste y los estudios de efectos sobre la dispersión.

