

Capítulo 2:

Estado del arte

2 *Estado del arte*

Las ideas de Taguchi obtuvieron una gran aceptación en Occidente, sin embargo no quedaba claro que el método propuesto daba lugar a productos o procesos robustos y con buen comportamiento funcional. Así, si se aceptaba que para medir la calidad de un producto era necesaria una función de pérdidas que tuviera en cuenta las variaciones funcionales que pudiera experimentar el comportamiento de este producto una vez en manos del cliente, no quedaba claro que siguiendo las pautas recogidas en el capítulo anterior referentes al “Método de Taguchi” se lograra en todos los casos el objetivo de optimizar la función de pérdidas seleccionada. Este y otros aspectos dieron lugar a numerosos estudios en un intento de mejorar diferentes aspectos del método.

Sin pretender ser exhaustivos a la hora de citar todos los temas que rodean el diseño de parámetros para la consecución de productos y procesos robustos, presentaremos aquellos que aparecen con más frecuencia cuando se aplican estas técnicas en la industria. Estos temas los clasificamos en los siguientes grupos siguiendo las etapas de la metodología de Taguchi expuesta en el capítulo 1:

- Selección de la métrica a estudiar;
- Selección de la matriz de diseño o condiciones experimentales;
- Análisis de datos y detección de efectos significativos;
- Selección de condiciones óptimas

El problema a tratar puede ser planteado de la siguiente forma: Dada una característica de calidad Y que depende a través de una función f desconocida de dos tipos de factores, de control y ruido, se desea elaborar una estrategia o metodología que contribuya a aumentar el conocimiento sobre la relación existente entre la característica de calidad y los factores estudiados, y que ayude a seleccionar las condiciones de los factores de control que optimizan esta característica, con el doble objetivo de mantenerla tan cerca del valor nominal y hacerla lo más robusta posible al efecto de factores ruido.

Para llevar a cabo esta estrategia se experimenta seleccionando determinados factores de control (X 's) y ruido (Z 's), siendo conscientes de que quedan fuera otros factores que pueden afectar a esta respuesta y que denominamos W 's.

$$Y = f(X, Z, W)$$

X = Factores de Control seleccionados ; Z = Factores Ruido seleccionados

W = Factores no seleccionados

$$Y = g(X, Z) + h$$

(2.1)

La proyección de f en el espacio generado por los factores X 's y Z 's da lugar a una función g quedando la componente h en el espacio complementario. La expresión h recoge la relación entre f y los factores W 's no considerados (pudiendo estar implicados además los factores X 's y Z 's si están combinados con estos W 's).

En la práctica se desea dar respuesta a las siguientes preguntas:

- ¿Qué condiciones experimentales seleccionar para X 's y Z 's?;
- ¿Qué métrica utilizar para estudiar Y ?;
- ¿Cómo se identifican los efectos significativos?;
- ¿Cómo se seleccionan las condiciones óptimas?.

Para ello, se acostumbra a aproximar la función g en (2.1) por medio de polinomios asumiendo que esta aproximación lineal es buena y la componente aleatoria restante no contiene información relevante. Además, se está interesado en estimar los valores de:

- $E(Y/Z \text{ variable}) = E(g(X,Z)/Z \text{ variable})$
- $V(Y/Z \text{ variable}) = V(g(X,Z)/Z \text{ variable})$

2.1 Selección de la métrica o “performance measure”

Las métricas propuestas en la literatura pueden ser clasificadas en dos grupos:

- *Métricas resumen*;
- *Métricas no-resumen*.

Las métricas resumen integran los valores obtenidos de la característica Y al dejar fijas las condiciones de los factores de control y variar las condiciones de ruido. Así, son métricas resumen $E(Y/Z \text{ variable})$, $Var(Y/Z \text{ variable})$ o las S/R propuestas por Taguchi. Estas métricas se expresan sólo en función de los factores de control y al optimizarlas puede encontrar las condiciones en los factores de control donde Y es robusta a variaciones en los factores ruido.

Con estas métricas se busca las mejores condiciones de control sin plantearse conocer cómo es la relación $g(X,Z)$ de partida.

Las métricas no-resumen estudian directamente la característica Y para distintas condiciones de los factores de control y ruido e intentan conocer al máximo la relación $g(X,Z)$ en (2.1). A partir de la relación completa se estudia, en las distintas condiciones de los factores de control, el efecto de los factores ruido en la respuesta intentando minimizar este efecto.

A continuación exponemos las métricas más utilizadas siguiendo esta clasificación justificando su uso.

2.1.1 Métricas resumen

La mayor parte de las métricas resumen que se recogen en la literatura surgen como alternativas de análisis de los datos a partir de un diseño basado en alguna de las matrices producto recomendadas por Taguchi. Por lo tanto se evalúan para cada condición fija de los factores de control a partir de los valores obtenidos al variar las condiciones de ruido.

Métricas S/R y métricas $PerMIA$

Tanto las métricas S/R , como el procedimiento en dos pasos para optimizar el problema a partir de los factores de control X_d y X_a definidos en el capítulo anterior, propuestos por Taguchi fueron polémicos desde el comienzo ya que no quedaba claro que siguiendo las pautas establecidas se llegara a minimizar alguna función de pérdidas.

Leon, Shoemaker y Kackar (LSK en adelante) publicaron en 1987 un interesante artículo analizando las métricas S/R y definiendo métricas resumen alternativas, las denominadas $PerMIA$ “Performance Measure Independent of Adjustment”. Estas son métricas que no dependen de los factores X_a y que minimizan igualmente una función de pérdidas. LSK (1987) derivan una manera de obtener una función $PerMIA$ para cada tipo de problema dependiendo de:

- La forma de la función de transferencia $g(X,Z)$ definida en (2.1);

- La forma de la función de pérdidas.

Analizando el método de Taguchi, LSK (1987) demuestran que si se toma una métrica similar a la S/R_N sugerida por Taguchi (en el caso de que la respuesta Y tiene un valor nominal objetivo),

$$PerMIA = 10 \log \left(\frac{E^2(Y)}{Var(Y)} \right) \quad (2.2)$$

a partir del procedimiento en dos pasos recomendado por Taguchi, y mencionado en el capítulo anterior, se puede llegar a minimizar una función de pérdidas como la que sigue

$$L(y, \tau) = k(y - \tau)^2 \quad \tau = \text{Valor nominal para } Y \quad (2.3)$$

siempre y cuando la función de transferencia $g(X,Z)$ en (2.1) toma la siguiente forma,

$$Y = \mu(X_a, X_d) \varepsilon(Z, X_d)$$

$$E(Y) = \mu(X_a, X_d); \quad V(Y) = \mu(X_a, X_d)^2 V[\varepsilon(Z, X_d)] = \mu^2 P(X_d)$$

$$\begin{aligned} X_d &= \text{F. de control que afectan a la dispersión y tal vez a la localización de } Y \\ X_a &= \text{F. de control que sólo afectan a la localización de } Y \\ Z &= \text{Factores ruido} \\ P &= \text{Función de los factores } X_d \end{aligned} \quad (2.4)$$

Este tipo de relación ente la característica Y y los factores de control y ruido implican que la varianza transmitida a Y por los factores ruido es proporcional a su valor esperado.

Los dos pasos a seguir para optimizar el problema serían:

- Seleccionar los factores de control X_{d^*} que maximizan $PermiaI$;
- Seleccionar los factores de control X_{a^*} de tal forma que

$$\mu(X_{d^*}, X_{a^*}) = \frac{\tau}{1 + \sigma^2(X_{d^*})}$$

LSK (1987) demuestran que el estadístico S/R_N que propone Taguchi es una estimación de la métrica presentada en (2.2), pero no es el único estadístico que minimiza la función de pérdidas en (2.3); de hecho hay otras funciones *PerMIA* que lo hacen;

La selección de una *PerMIA* depende totalmente de la función de transferencia y de la función de pérdidas.

Por ejemplo, si para la misma función de transferencia definida en (2.4), y que suponen que la varianza sea proporcional al valor esperado, se trabaja en escala logarítmica $\text{Log}(Y)$ y se toma como función de pérdidas

$$L(y, \tau) = k[\log(y) - \log(\tau)]^2 \quad \tau = \text{Valor nominal para } Y \quad (2.5)$$

se logra minimizar ésta con la *PerMIA*

$$\text{PerMIA} = \text{Log}[\text{Var}(\log(Y))] \quad (2.6)$$

función propuesta también por Box (1988). LSK (1987) y Box (1988) añaden que si no hay preferencias entre una u otra función de pérdidas se aconseja utilizar como métrica la definida en (2.6) antes que la S/R_N debido a que la primera tiene mejores propiedades estadísticas. Por lo tanto si se cree dependencia entre *media* y *varianza* se aconseja utilizar la transformación logarítmica y trabajar a continuación directamente con la *varianza* o una transformación logarítmica de ésta.

Para el caso particular en que *media* y *varianza* son independientes (tal y como se asume en la mayoría de situaciones prácticas) y la función de transferencia puede ser expresada como

$$Y = \mu(X_a, X_d) + \varepsilon(N, X_d) \quad \text{con } E(Y) = \mu(X_a, X_d); \\ V(Y) = V[\varepsilon(N, X_d)] = P(X_d) \quad (2.7)$$

si la función de pérdidas es del tipo (2.3) la métrica *PerMIA* recomendada evalúa directamente la *varianza* de la respuesta:

$$\text{PerMIA} = \text{Var}(Y) \quad (2.8)$$

Además, para este caso la métrica S/R_N propuesta por Taguchi no es una métrica *PerMIA* ya que no es independiente de la *media*

$$S / R_N = 10 \log \left(\frac{E^2(Y)}{\text{Var}(Y)} \right) = 10 [\log(\hat{\mu}^2) + \log(\hat{\sigma}^2)] \quad (2.9)$$

y por consiguiente, el procedimiento en dos pasos no lleva al objetivo de minimizar la función de pérdidas.

De igual manera, el resto de métricas S/R propuestas por Taguchi pueden resultar buenas alternativas de trabajo si se cumplen condiciones particulares de la función de transferencia y de la función de pérdidas pero lo que no se puede hacer es un uso indiscriminado de ellas.

Media y Variabilidad por separado

Como alternativa al uso generalizado de métricas S/R , el uso de una *PerMIA* requiere unos conocimientos previos que, en general, no se tienen en el entorno práctico donde se desarrollan los experimentos.

Debido a estas razones, Lucas et al. (1985) y Box (1988) entre otros, proponen la modelización por separado de la *Media* y la *Variabilidad* (o *Log(s)* para estabilizar la variabilidad). En aquellos casos en que estas dos funciones no sean independientes, Box (1988) propone el uso de transformaciones de la variable original antes de aplicar estas dos métricas. La transformación a elegir está supeditada al grado de conocimiento de la dependencia entre *media* y Variabilidad; así para el caso particular tratado anteriormente en que Y adopta la forma en (2.4) y por lo tanto *media* y *varianza* son proporcionales, $\hat{\sigma}/\hat{\mu}$, estimador del coeficiente de variación de Y , es una función de variabilidad independiente de $\hat{\mu}$ que es proporcional a $s_{\log(y)}$. Por lo tanto una alternativa a trabajar con la S/R propuesta por Taguchi, $\log\left(\frac{\bar{Y}^2}{s^2}\right)$, sería la de transformar logarítmicamente la variable original y trabajar con $\log s_{\log(y)}$, métrica más sencilla y con mejores propiedades de eficiencia.

Por lo tanto, si se desea trabajar por separado con la localización y la dispersión de la respuesta tal y como se hace en la mayoría de situaciones prácticas, se han de tener en cuenta varios aspectos, frecuentemente ignorados en la práctica (Ullman (1988), con el fin de eliminar al máximo la posible dependencia entre *media* y dispersión:

- La selección inicial de la característica de interés ;
- El uso de transformaciones de esta variable Y ;
- La selección de la métrica a utilizar en el análisis.

Uso de transformaciones

En cuanto a la selección de una transformación, no queda claro que la transformación $\text{Log}(Y)$ sea la más adecuada, se puede pensar incluso en no transformar. De hecho, es difícil seleccionar una transformación si no se tiene conocimientos previos sobre la relación teórica subyacente y a la vez resulta peligroso hacerlo mal ya que este hecho puede complicar el estudio.

Así, aunque la transformación *logarítmica* sea muy utilizada de cara a una modelización por mínimos cuadrados, Box y Fung (1983) proponen una estrategia para situaciones en que no se tiene conocimiento previo de una posible relación teórica entre *media* y dispersión y se desea trabajar con métricas resumen se basa en seleccionar una transformación de Y , Y^λ , a partir de los “ λ -plots” para los modelos *Media* y *Log(s)* de la variable transformada. Así, se selecciona aquel valor “ λ^* ” que cumpla los dos principios siguientes:

- Parsimonia: los modelos en (2.10) y (2.11), para la *Media* y la *Variabilidad* tienen pocos efectos significativos;
- Separación: *Media* y *Variabilidad* son independientes.

$$\hat{\mu}_{Y^\lambda} = \bar{Y}^\lambda = F_1(X_a, X_d) \quad (2.10)$$

$$\log[\hat{V}(Y^\lambda)] = \log(s_{Y^\lambda}^2) = F_2(Z, X_d) \quad (2.11)$$

De entre todas las transformaciones equivalentes, y para no complicar la interpretación de los datos, se han de seleccionar aquellas transformaciones que tengan sentido en el contexto de trabajo.

Aunque se ha demostrado empíricamente que las transformaciones que estabilizan la *varianza* usualmente logran aditividad en los modelos, en algunas aplicaciones industriales, tales como diseño y optimización de circuitos, la superficie de respuesta es tan altamente no-lineal, que no se puede lograr ni parsimonia ni separación. Para algunos de estos casos donde no se puede lograr ambos objetivos, Box y Hill (1974) proponen elegir en primer lugar la transformación que da parsimonia y a continuación, mediante pesos empíricos, acomodar la heterogeneidad de las varianzas.

Modelo General Lineal (GLM)

Mc Cullagh y Nelder (1989), Grego (1993) y Hamada y Nelder (1997) proponen la utilización del modelo general lineal, GLM, en la modelización conjunta o por separado de la *media* y la *variabilidad* de la respuesta. Esta metodología admite un mayor número

de distribuciones para la respuesta Y (pertenecientes a la familia exponencial 1-parámetro) y para aplicarla es necesario conocer:

- La distribución del error de la respuesta (esto determina los pesos que serán asignados a cada punto en el ajuste del modelo);
- La “función link” que define la escala en la cual los efectos de las variables explicativas se combinan de forma aditiva.

Un caso particular ocurre cuando las varianzas muestrales siguen una distribución gamma y la “función link” es la función logarítmica. En tal caso hay una equivalencia asintótica entre trabajar con los métodos ordinarios de regresión mínimo cuadrática a partir de la transformación logarítmica que estabiliza la *varianza* y trabajar con GLM: Los efectos obtenidos por ambos métodos son indistinguibles para pequeños valores de σ_i^2 aunque la aproximación por GLM es más eficiente si las hipótesis sobre la distribución del modelo son ciertas, Atkinson (1982).

En general, si se tiene conocimiento de que la distribución teórica del error de la respuesta pertenece a esta familia puede plantearse el analizar los datos con el GLM, y si no, no hay razón para ello.

Métricas resumen en un entorno práctico

En la práctica, las inevitables restricciones económicas que rodean la experimentación obligan a elaborar estrategias secuenciales partiendo de diseños factoriales fraccionales que permiten aproximar localmente las funciones objetivo por medio de polinomios de a lo sumo segundo orden (en la mayoría de los casos no se logra obtener los términos cuadráticos puros) cuyos coeficientes se estiman por mínimos cuadrados ordinarios. Los términos cruzados permiten detectar la necesidad de ampliar los modelos a otros de orden superior en la siguiente etapa experimental.

$$\bar{Y}^\lambda = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad \varepsilon - iid -> N(0, \sigma_1) \quad (2.12)$$

$$\log(s_{Y^\lambda}^2) = \beta_0^* + \sum_{i=1}^k \beta_i^* X_i + \sum_{i \neq j} \beta_{ij}^* X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad \varepsilon - iid -> N(0, \sigma_2) \quad (2.13)$$

El experimentador que decida trabajar por separado con las métricas de localización y dispersión estará asumiendo que ambas funciones son independientes y deberá hacer lo posible para conseguirlo.

En aquellos casos en que se conozca la existencia de una relación teórica entre, por ejemplo la *media* y la *varianza*, será más fácil encontrar una transformación de la

variable original Y que rompa tal relación y que a su vez haga que la aproximación polinómica sea aceptable (aunque no siempre sea esto posible).

Si por el contrario no se sabe de la existencia de tal relación y se decide utilizar algunas de las sugerencias presentadas anteriormente, es muy probable que no se detecte mejora alguna con las transformaciones ya que esto suceda se necesita un número elevado de réplicas (número elevado de condiciones de ruido) a la vez de un amplio rango de variación en los datos (en las distintas condiciones de control) (Gunter(1987)). Además, si los diseños utilizados son diseños factoriales altamente fraccionados el método de Box y Fung (1983) es poco efectivo ya que las interacciones están confundidas con efectos principales y resulta difícil conseguir el objetivo de “separación”.

Otros método como el de Nair y Pregibon (1988) basado en una herramienta gráfica, “mean-variance plot”, que ayuda en la detección de relaciones entre *media* y *varianza*, funciona bien sólo si hay pocos efectos que afectan a la dispersión.

Por lo tanto la situación más frecuente con la que nos encontraremos en la práctica será aquella que partiendo de una matriz producto generada por diseños fraccionados se desee estimar por medio de polinomios de segundo orden incompletos dos modelos que representan a la *media* y $\log(s)$,

$$\bar{Y} = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad \varepsilon - iid - > N(0, \sigma_1) \quad (2.14)$$

$$\log(s) = \beta_0^* + \sum_{i=1}^k \beta_i^* X_i + \sum_{i \neq j} \beta_{ij}^* X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad \varepsilon - iid - > N(0, \sigma_2) \quad (2.15)$$

Es decir, se asume independencia entre *media* y variabilidad modelando ambas funciones separadamente por métodos sencillos (se utiliza la transformación logarítmica más por criterios teóricos que por criterios prácticos).

2.1.2 Métricas no-resumen

Una desventaja importante del análisis de datos a partir de métricas resumen es que al resumirse el comportamiento de la respuesta ante los factores ruido se pierde información sobre el efecto de los factores ruido en la respuesta y la posible dependencia de este efecto de determinados factores de control. Es decir, con las métricas resumen se logra tener una visión global pero se pierde el detalle que ayuda a entender el problema.

La alternativa de análisis consiste en modelar directamente la superficie de respuesta en función de los factores de control y de los factores ruido a la vez (función $g(X;Z)$ en 2.1)). Welch, Yu, Kang y Sacks (1990) fueron los primeros autores que

propusieron esta idea que fue respaldada posteriormente por autores del prestigio de Shoemaker, Tsui y Wu (1991), Box y Jones (1992) y Lorenzen (1992) entre otros.

La estrategia en este caso se plantea de la manera siguiente:

- a) Estimar localmente el modelo que liga la respuesta Y y los factores de control (X 's) y ruido (Z 's) seleccionados

$$\hat{Y} = \hat{g}(X, Z) + e \tag{2.16}$$

$X = \text{F. de Control seleccionados}; Z = \text{F. Ruido seleccionados}$

- b) Formular una medida de robustez para Y ;
- c) Seleccionar los niveles de X 's que hacen más robusta Y .

De una manera similar a lo que se hacía con métricas resumen, la aproximación a la función $g(X, Z)$ se realiza mediante polinomios. Esta aproximación si se parte de “k” factores de control X_i , y “r” factores ruido Z_j , puede ser expresada como

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{ij} X_i Z_l + \varepsilon$$

$i, j = 1, \dots, k \quad l, m = 1, \dots, r \quad \varepsilon - iid -> N(0, \sigma)$ (2.17)

$$E(X_i) = E(Z_l) = 0$$

Una vez estimado el modelo en (2.17) por mínimos cuadrados ordinarios, los factores de control que surgen significativos se clasifican en dos grupos:

- Los factores X_a no contienen efectos significativos interaccionando con los factores ruido;
- Los factores X_d tienen efectos interaccionando con los factores ruido.

De esta forma el modelo en (2.17) se puede expresar como:

$$\hat{Y} = P_1(X_a, X_d) + P_2(X_d, Z) + e \tag{2.18}$$

donde P_1 es un polinomio que contiene los efectos significativos que comprenden a los factores de control X_a y P_2 es un polinomio que recoge los efectos significativos de los factores ruido y de las interacciones entre los factores de control X_d y los factores ruido.

De esta forma, el análisis del comportamiento de la superficie de respuesta se lleva a cabo desde dos puntos de vista: La localización y la variabilidad de Y .

Estudio de la localización de la respuesta en las condiciones de control:

$$E(\hat{Y}) = P_1(X_a, X_d) \quad (2.19)$$

Estudio de la variabilidad transmitida por los factores ruido en cada condición de los factores de control. Existen dos métodos para llevar a cabo este estudio:

- A partir del estudio de las interacciones entre factores de control y factores ruido;
- A partir de una métrica que minimice una función de pérdidas. Por ejemplo, si se toma la *varianza* directamente se obtiene:

$$Var(\hat{Y}) = Var[P_2(X_d, Z)] + Var(e) \quad (2.20)$$

Por lo tanto, los factores X_d se utilizan para minimizar la variabilidad generada por los factores ruido y los factores X_a para ajustar el valor esperado de la respuesta.

Sean cualesquiera el método utilizado para obtener robustez, los dos métodos pasan por una estimación del modelo en (2.17) y por lo tanto, la capacidad de detectar condiciones robustas dependen de la bondad de la estimación de la respuesta media, hecho que ya ha sido citado por otros autores (Shoemaker et al. (1991)). Sin embargo, cuando se modela la *varianza* directamente utilizando métricas resumen, la estimación de la función de varianzas no depende de una mala estimación de la función de la *media*.

A partir de la expresión en (2.20) podemos deducir que siendo P_2 un polinomio con términos cruzados entre los factores de control y ruido, la función $Var(Y)$ será de segundo orden en los factores de control. Por lo tanto, podemos sospechar que las aproximaciones por métricas resumen a $Var(Y)$ pueden no funcionar bien si se utilizan modelos lineales sencillos. Por otra parte, si *media* y *variabilidad* no son independientes entre si, la aproximación a $g(X;Z)$ por polinomios es posible que no recoja bien el comportamiento de la variabilidad de Y ante variaciones de los factores Z 's.

2.1.3 Conclusiones sobre el uso de métricas resumen y no-resumen

Hemos clasificado el conjunto de métricas a utilizar en el análisis de datos provenientes de experimentos ejecutados para la obtención de condiciones robustas en dos tipos: *métricas resumen* y *métricas no-resumen*.

Para poder utilizar las *métricas resumen* se necesita que las condiciones de ruido a las que se someten las condiciones experimentales en los factores de control sean las mismas, ello se consigue partiendo de una matriz producto. Este tipo de diseño admite los dos tipos de análisis.

Si el diseño de partida está definido a partir de una matriz producto, existen 3 métodos para analizar la variabilidad transmitida por los factores ruido:

- 1) Estimar directamente $V(Y)$ con métricas resumen;
- 2) Estimar el modelo para Y con *métricas no-resumen* y analizar las interacciones entre factores de control y ruido;
- 3) Estimar el modelo para Y con *métricas no-resumen* y deducir la expresión de $V(Y)$ a partir de este modelo.

Aspectos más importantes del uso de *métricas resumen* o *métricas no-resumen*:

- Las métricas resumen, S/R ; utilizadas por Taguchi para analizar los problemas de robustez fueron el origen de numerosas críticas: Resulta difícil justificar su uso e interpretar los resultados;
- LSK (1987) demuestran que algunas de las S/R de Taguchi están asociadas a la minimización de alguna función de pérdidas pero no se pueden usar de forma indiscriminada ya que no siempre resultan ser métricas independientes del ajuste o *media*;
- LSK (1987) recomiendan el uso de una *PerMIA* para reducir la variabilidad. La selección de una *PerMIA* depende de la función de transferencia y de la función de pérdidas. No siempre se asegura la existencia de una *PerMIA* y a veces el método para estimar la *PerMIA* es complicado. Algunas de las *PerMIA* recomendadas son:

Relación entre media y variabilidad	Función de Pérdidas	PerMIA*	Alternativas
Proporcionales	$L(y)=k(y-\tau)^2$	$10 \log (E^2(Y)/\text{Var}(Y))$	S/R
Proporcionales	$L(y)=k(\log(y)-\log(\tau))^2$	$\text{Log}(\text{Var}(\log(y)))$	S/R
Independientes	$L(y)=k(y-\tau)^2$	$\text{Var}(Y)$	(**)

(*)Las PerMIAS recomendadas tienen mejor propiedades estadísticas que las S/R alternativas.

(**) En este caso no se debe utilizar S/R

- Para trabajar directamente con $\text{Var}(Y)$, de cara a conseguir robustez, se requiere que las funciones *media* y *variabilidad* sean independientes, existiendo diferentes métodos para detectar la dependencia y tratar de corregirla;
- Si la distribución teórica de la *varianza* pertenece a la familia exponencial 1-parámetro, se puede optar por la modelización vía GLM;
- En general, sin no hay información previa que pueda llevar a seleccionar una métrica más sofisticada, la información de los datos se revela mejor con métodos simples de análisis. Es mejor trabajar en término de la *Media* y la *Desv. Típica* o $\log(s)$ que utilizar métodos más complicados ya que con experimentaciones reducidas no se detectarán mejoras (transformando los datos o modelando vía GLM);
- Con la métricas no-resumen se logra tener mayor conocimiento de la relación existente entre Y y todos los factores a estudio que con las métricas resumen;
- Si se analiza la robustez a partir de la estimación del modelo para Y , la detección de condiciones robustas a la variabilidad provocada por las condiciones de ruido está supeditada a la bondad de ajuste del modelo para la localización.

2.2 Selección de la matriz de diseño

La matriz de diseño contiene las condiciones en que se han de fijar los factores de control y ruido para realizar la experimentación. En la práctica, las restricciones económicas obligan a economizar el número de experimentos lo cual otorga una gran importancia a este tema.

Primeramente presentaremos la selección de la matriz en función del modelo teórico que se pretende estimar, y a continuación hablaremos del orden de la experimentación teniendo en cuenta este tema.

2.2.1 Matriz producto y matriz ampliada

Aunque el objetivo final del diseño de parámetros es encontrar aquellas condiciones de los factores de control X 's que hacen a la característica de calidad Y lo más robusta posible a los factores ruido Z 's, manteniendo el valor de Y dentro de unos márgenes establecidos, cada diseño permite explorar la relación de Y con los factores X 's y Z 's de una manera diferente. Por ello, antes de seleccionar la matriz de diseño es primordial plantear la hipótesis sobre la estructura del modelo teórico que se pretende investigar.

Como hemos visto en los apartados anteriores, existen principalmente dos tipos de planteamientos para el modelo teórico:

- Modelo resumen: Estudia, en cada condición de los factores de control, el valor esperado de la respuesta y la variabilidad debida a los factores ruido. Para realizar este estudio se resume la información a lo largo de las condiciones de ruido. Por ejemplo si se trabaja con *media y variabilidad* se acostumbra a utilizar los modelos:

$$E(Y) = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad (2.21)$$

$$\log(\sigma) = \beta_0^* + \sum_{i=1}^k \beta_i^* X_i + \sum_{i \neq j} \beta_{ij}^* X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad (2.22)$$

- Modelo no-resumen: Se estudia el comportamiento de la respuesta Y para cada condición experimental determinada por los factores de control y ruido. Más adelante se explota la relación entre los factores X 's y Z 's para llegar a condiciones robustas.

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l + \varepsilon \quad (2.23)$$

$i, j = 1, \dots, k \quad l, m = 1, \dots, r$

De acuerdo a la anterior división, y tomando como referencia para ilustrar los ejemplos los diseños factoriales a dos niveles, existen, a grandes rasgos, dos métodos de plantear la matriz de diseño:

- Matriz producto: Es la matriz planteada por Taguchi y puede ser interpretada como un diseño $2^{k-p} \times 2^{r-q} = 2^{(k+r)-(p+q)}$, resultante de realizar un

producto cartesiano de un diseño 2^{k-p} para los factores de control y un diseño 2^{r-q} para los factores ruido;

- **Matriz ampliada:** Es una matriz de tantas columnas como número total de factores incluyendo tanto a los factores de control como a los ruido. Puede ser interpretada como un diseño $2^{(k+r)-x}$.

Veámoslo con un ejemplo.

- **Ejemplo 1:** En un tratamiento térmico se trabaja con una característica de calidad asociada a la deformación que sufren las piezas y se desea encontrar condiciones de proceso donde los factores ruido afecten lo menos posible a la deformación. Se ha decidido experimentar con 3 factores de control y 2 factores ruido. Los factores de control son: Tiempo inicial de precalentamiento en el horno (A), tiempo total en el horno (B) y la temperatura máxima del horno (C). Los factores ruido son: La posición de la pieza en el horno (R) y la temperatura ambiente (S) a la que se enfría la pieza una vez sacada del horno. El presupuesto permite realizar 16 experimentos.

Con la restricción de un máximo de 16 experimentos, se ha optado por seleccionar 4 “prototipos de proceso” y someterlos cada uno a 4 condiciones de ruido. Se puede ver que esta estrategia natural de someter a los “prototipos” a las mismas condiciones de ruido para seleccionar el que se comporte mejor corresponde con un diseño del tipo $2^{3-1} \times 2^2$, donde los prototipos se seleccionan a partir de un diseño 2^{3-1} con $C = AB$ y las condiciones de ruido a partir de un diseño completo 2^2 .

Tabla 2.1 Matriz producto. Diseño 2^{3-1} con $C=AB$ para los factores de control y 2^2 para los factores de ruido

	Tiempo inicial (A)	Tiempo total (B)	Temperatura (C)	Posición Horno (R)	-1	1	-1	1
				Tiempo espera (S)	-1	-1	1	1
P1	-1	-1	1		x	x	x	x
P2	1	-1	-1		x	x	x	x
P3	-1	1	-1		x	x	x	x
P4	1	1	1		x	x	x	x

Estas pruebas admiten una reordenación en la forma estándar tal y como aparecen en la Tabla 2.2. Por lo tanto este diseño se puede ver como un diseño 2^{5-1} de Resolución III con generados $C=AB$.

Tabla 2.2 Matriz ampliada 2^{5-1} para 3 factores de control y 2 de ruido obtenida a partir de la matriz producto en la Tabla 2.1

	Tiempo Inicial (A)	Tiempo Total (B)	Temperatura (C)	Posición Horno (R)	Tiempo espera (S)
P1	-1	-1	1	-1	-1
P2	1	-1	-1	-1	-1
P3	-1	1	-1	-1	-1
P4	1	1	1	-1	-1
P1	-1	-1	1	1	-1
P2	1	-1	-1	1	-1
P3	-1	1	-1	1	-1
P4	1	1	1	1	-1
P1	-1	-1	1	-1	1
P2	1	-1	-1	-1	1
P3	-1	1	-1	-1	1
P4	1	1	1	-1	1
P1	-1	-1	1	1	1
P2	1	-1	-1	1	1
P3	-1	1	-1	1	1
P4	1	1	1	1	1

Si en el momento de seleccionar la matriz de diseño se hubieran seguido las pautas habituales de hacerlo a partir de una única matriz 2^{5-1} pero de máxima resolución se hubiera llegado a un diseño similar al que se presenta en la Tabla 2.3 de Resolución V.

Tabla 2.3 Matriz ampliada 2^{5-1} para 3 factores de control y 2 de ruido. S=ABCR

	Tiempo inicial (A)	Tiempo total (B)	Temperatura (C)	Posición Horno (R)	Tiempo espera (S)
P1	-1	-1	-1	-1	1
P2	1	-1	-1	-1	-1
P3	-1	1	-1	-1	-1
P4	1	1	-1	-1	1
P5	-1	-1	1	-1	-1
P6	1	-1	1	-1	1
P7	-1	1	1	-1	1
P8	1	1	1	-1	-1
P1	-1	-1	-1	1	-1
P2	1	-1	-1	1	1
P3	-1	1	-1	1	1
P4	1	1	-1	1	-1
P5	-1	-1	1	1	1
P6	1	-1	1	1	-1
P7	-1	1	1	1	-1
P8	1	1	1	1	1

Si en esta situación reordenamos las condiciones experimentales con el fin de que aparezcan según la estructura de una matriz producto, obtenemos la presentación que aparece en la Tabla 2.4. Así se puede ver que en este diseño se están seleccionando un número mayor de condiciones de proceso (P1-P8) y que se está trabajando con 4 condiciones diferentes de ruido; ahora bien, 4 de estos procesos se someten a 2 de estas condiciones ruido y el resto a las 2 restantes. Por lo tanto, los 8 procesos no son

comparables en cuanto a robustez ya que no se someten a las mismas condiciones, en todo caso se podría hacer un estudio por bloques.

Tabla 2.4 Matriz ampliada 2^{5-1} S=ABCR representada como una matriz producto

				Posición Horno (R)	-1	1	-1	1
	Tiempo Inicial (A)	Tiempo Total (B)	Temperatura (C)	Tiempo espera (S)	-1	-1	1	1
P1	-1	-1	-1			x	x	
P2	1	-1	-1		x			x
P3	-1	1	-1		x			x
P4	1	1	-1			x	x	
P5	-1	-1	1		x			x
P6	1	-1	1			x	x	
P7	-1	1	1			x	x	
P8	1	1	1		x			x

Comparando los dos diseños, el representado en las Tabla 2.1 y Tabla 2.2 por una parte y el representado en la Tabla 2.3 y Tabla 2.4 por otra, lo primero que llama la atención es la manera tan diferente en que se gastan los grados de libertad. En la matriz producto se gastan muchos recursos en probar diferentes condiciones de ruido para cada condición de proceso, mientras que en la matriz ampliada (seleccionada por criterios de máxima resolución) se reparten de forma más equilibrada los recursos; como consecuencia en el primer caso probamos 4 condiciones de proceso mientras que en el segundo caso 8. Por otra parte, si se deseara experimentar en 8 condiciones de proceso a partir de una matriz producto que incluyera 4 condiciones de ruido se necesitarían 32 condiciones experimentales en lugar de 16.

Esta diferencia en asignación de los grados de libertad tiene implicaciones estadísticas en cuanto al tipo de efectos que pueden ser estimados por uno u otro tipo de diseños cuando se analizan los datos con *métricas no-resumen*.

Si el diseño se ha seleccionado a partir de una matriz producto en general:

- Se llega a menos condiciones en los factores de control que se si selecciona a partir de una matriz ampliada. Como consecuencia de este hecho el número de efectos asociados a los factores de control que pueden ser estimados cuando se trabaja con *métricas no-resumen* es menor;
- Siempre permiten estimar libre de confusiones las interacciones entre factores de control y factores ruido (reordenando previamente las observaciones en una matriz ampliada).

En cambio, si los diseños se seleccionan a partir de una matriz ampliada teniendo en cuenta a todos los factores involucrados y siguiendo criterios de máxima resolución, se pueden obtener diseños con más efectos estimables en los factores de control pero en muchas situaciones las interacciones entre factores de control y factores ruido no estarán

libres de confusiones (en el ejemplo presentado los dos diseños permiten estimaciones libres de confusiones pero esto no es siempre así).

Por otra parte, un diseño a partir de una matriz ampliada seleccionada por criterios de máxima resolución, no permite analizar los “prototipos” a partir de *métricas resumen* ya que las condiciones a las que se someten los prototipos son diferentes y por lo tanto los resultados obtenidos a partir de las *métricas resumen* no son comparables.

Taguchi propone la selección del diseño a partir de una matriz producto y no por criterios de resolución y este hecho ha provocado numerosas críticas, las más comunes relacionadas con el hecho de que estos diseños necesitan una gran cantidad de condiciones experimentales para conseguir una resolución determinada en los factores de control (Shoemaker, Tsui y Wu (1991), Sacks y Welch (1992) y Box (1992) entre otros), y si se trabaja con *métricas no-resumen* es posible seleccionar diseños con un número total de condiciones experimentales menor.

A pesar de las críticas recibidas con motivo del coste asociado a su ejecución, estos diseños gozan de una gran aceptación en el entorno industrial. Podemos justificar parte de esta aceptación en la disposición intuitiva de las pruebas. En cuanto al coste, este se puede reducir drásticamente si el diseño no se ejecuta de forma totalmente aleatoria.

Cuando el problema del número de condiciones experimentales es grave, Taguchi (1987) propone la utilización de un único “factor ruido compuesto” obtenido a partir de los factores ruido a estudio, que recogería entre todas las condiciones de ruido posibles las dos que provocan las condiciones extremas en la respuesta. Esta estrategia, que conceptualmente tiene un buen planteamiento, sin embargo presenta serias dificultades prácticas ya que resulta difícil conocer a priori cuáles son estas condiciones.

2.2.2 Alternativas a la Matriz Producto

A raíz de las primeras publicaciones del trabajo de Taguchi, diversos autores han trabajado en temas relacionados con métodos de asignación de factores a las columnas de una matriz ampliada con tal de que el diseño resultante permitiera estimar aquellos efectos que son importantes a la hora de detectar condiciones robustas.

Si se asume que la relación entre la respuesta Y y los factores a estudio puede ser aproximada por un modelo lineal de primer orden con términos cruzados:

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l + \varepsilon$$

$i, j = 1, \dots, k \quad l, m = 1, \dots, r$

- Si interesa estudiar la robustez o lo que es lo mismo, el efecto de los factores ruido en las distintas condiciones de los factores de control, es necesario estimar los efectos que contengan algún factor ruido: θ_l , θ_{lm} , δ_{il} (θ_{lm} no afecta a la selección de las condiciones robustas sólo al valor esperado en estas condiciones);
- Si interesa estudiar la respuesta *media* en diferentes condiciones de los factores de control es necesario estimar los efectos que contengan únicamente factores de control (β_0 , β_i , β_{ij}).

Por lo tanto si se trabaja con diseños factoriales a dos niveles y se da prioridad a la detección de condiciones robustas, estos diseños han de permitir estimar los efectos principales y las interacciones entre los factores de control y los factores ruido. Queda claro que no es suficiente seleccionar diseños basados en los criterios de máxima resolución y mínima aberración ya que estos diseños tratan por igual a los factores de control y a los factores ruido (Shoemaker y Tsui (1992)).

Addelman (1962 y 1964) posee trabajos, previos a la época en que estamos hablando, donde propone diseños factoriales para un conjunto de factores clasificados en dos categorías. Estos diseños (para dos, tres cuatro y cinco niveles de los factores) se clasifican en tres clases:

- Diseños Addelman de clase 1: De un conjunto de n factores, permiten estimar libre de confusiones todos los efectos principales y todas las interacciones dobles entre un subconjunto de k factores;
- Diseños Addelman de clase 2: De un conjunto de n factores, permiten estimar libre de confusiones todos los efectos principales, todas las interacciones dobles entre un subconjunto de k factores y todas las interacciones dobles entre el resto de $n-k$ factores;
- Diseños Addelman de clase 3: De un conjunto de n factores, permiten estimar libre de confusiones todos los efectos principales y todas las interacciones dobles que contienen algún factor de un subconjunto de k factores seleccionados previamente.

Inspirándose en los diseños propuestos por Addelman, Jones (1990) propone un conjunto de diseños tales que, fijando el número de factores de control a estudio, se obtiene el número máximo de factores ruido que pueden estudiarse para un número determinado de condiciones experimentales, de tal forma que el diseño propuesto permite estimar, como mínimo, todos los efectos principales y todas las interacciones dobles entre factores de control y factores ruido; en cada caso particular, las interacciones de otro tipo pueden estar confundidas. Los diseños obtenidos a partir de la matriz producto son un caso particular de estos diseños pero se puede lograr el mismo objetivo con menos condiciones experimentales.

Jones (1990) además propone diferentes diseños dependiendo de si la función a optimizar es $M(x)$, relacionada con la cercanía de la respuesta al objetivo, $V(x)$, relacionada con la variabilidad de la respuesta debido a los factores ruido, o $R(x)$, una media ponderada de las dos expresiones anteriores evaluadas en las condiciones de los factores de control. Por ejemplo, si se desea optimizar $R(x)$ se necesitan diseños que permitan estimar los efectos principales y todas las interacciones entre factores de control y factores ruido.

En el caso en que se trabaje con modelos lineales de segundo orden, Jones (1990) presenta los diseños “*Addelman de clase 3 aumentados*” construidos a partir de los diseños de clase 3 de Addelman (1962) añadiendo puntos centrales y estrella; los diseños “*Box-Behnken modificados*”, (construidos como su nombre indica a partir de los diseños Box-Behnken); los diseños “*Koshal modificados*” y los diseños “*de resolución III* de Hartley modificados*”.

Wu y Zhu (1999) más recientemente aportan tablas de diseños donde, una vez fijado el número de condiciones experimentales y el número total de factores, dependiendo de cómo se alojen los factores entre el grupo de control o el grupo de factores ruido, se presentan varias alternativas de diseños en los que es posible estimar, si no todas, si algún grupo de interacciones entre factores de control y factores ruido. En estas tablas hay diseños que compiten entre sí en cuanto a sus características, y su selección dependerá de la importancia que se le de a priori a los efectos que permiten estimar. Los diseños de Jones (1990) son casos particulares de los aquí presentados.

2.2.3 Diseños no aleatorizados

Hasta ahora no hemos hecho ningún comentario sobre el orden en que se deberían de realizar las pruebas y suponiendo que un diseño con n condiciones experimentales requiere n experimentos u operaciones independientes a la hora de la ejecución. Sin embargo, en la práctica no siempre esto es así; de hecho en muchas más ocasiones de lo que cabría pensar, los experimentos no se aleatorizan sino que se realizan en un orden que supone un ahorro económico importante de experimentación y todo ello a costa de mover al mínimo los niveles de ciertos factores. Como consecuencia, a la hora de evaluar el coste del diseño, el número de condiciones experimentales pasa a un segundo plano y lo que verdaderamente cobra importancia es el número de movimientos de los factores o lo que podríamos llamar el número de operaciones diferentes que se realizan.

Algunos diseños que recogen esta idea son los diseños denominados “split-plot” estudiados de forma muy amplia por Cox (1958) y a los que Jones (1991), Box y Jones (1992) y Bisgaard (2000) entre otros, les dan un tratamiento especial al aplicarlos al entorno de diseño de parámetros.

Los diseños *split-plot* aplicados al diseño de productos y procesos robustos, tienen una apariencia similar a los diseños obtenidos a partir de una matriz producto pero sin embargo se ejecutan de manera diferente a lo que sería un diseño totalmente aleatorizado, con el consiguiente ahorro económico de ejecución. Taguchi (1987)

menciona explícitamente esta alternativa de experimentación en algunos de sus trabajos originales. Sin embargo, en la mayoría de trabajos publicados donde se utilizan matrices producto no se acostumbra a mencionar el orden de experimentación, por lo que se asume que son diseños totalmente aleatorizados y por lo tanto con una mala relación entre el número de efectos que pueden ser estimados y el coste de la experimentación.

El orden de ejecución no afecta a la estimación de los efectos sino que provoca diversas fuentes de error experimental, por lo tanto afecta a la significación de los efectos. Por lo tanto, si se reordenan las condiciones experimentales de una matriz producto en una única matriz extendida con el fin de estimar un modelo completo en los factores de control y ruido, no todos los efectos serán estimados con la misma precisión y no se podrán ordenar los efectos por su magnitud.

El escenario que rodea la selección de este tipo de diseños alternativos sería el siguiente: se desean realizar una serie de pruebas experimentales dentro de la metodología de diseño de parámetros pero se tienen restricciones económicas serias en cuanto al número de pruebas a realizar. Cabría la posibilidad de obtener el valor del comportamiento de la respuesta en un mayor número de condiciones experimentales si se agrupan estas ya que supondría una gran reducción de costes.

Aunque es posible realizar las pruebas de una gran variedad de formas, si no se sigue un “mínimo de orden” en el orden de realización, el análisis posterior de los datos comienza a ser dificultoso. Por ello se estudia principalmente tres métodos de ejecución de los experimentos: *totalmente aleatoria*, *split-plot* y *strip-plot*.

Los tres métodos pueden ser aplicados a la matriz producto del ejemplo 1 presentada en la Tabla 2.1. Esta matriz dará lugar a 16 valores de la respuesta para 16 condiciones diferentes de experimentación; sin embargo, el coste total necesario para obtener estos valores puede llegar a ser extraordinariamente variable dependiendo del método utilizado a la hora de ejecutar las condiciones, veámoslo:

- **Diseño totalmente aleatorizado:** Se seleccionan de forma aleatoria las 16 condiciones de la matriz producto en la Tabla 2.1, o equivalentemente una de las 16 filas de la matriz ampliada en la Tabla 2.2, y se ejecutan una a una. Esto supone como máximo 16 movimientos de las condiciones de proceso y 16 de ruido. La Tabla 2.5 presenta una posible selección.

Tabla 2.5 Diseño totalmente aleatorizado a partir de la matriz producto en Tabla 2.1

Orden de realización	Tiempo Inicial (A)	Tiempo Total (B)	Temperatura (C)	Posición Horno (R)	Tiempo espera (S)
5	-1	-1	1	-1	-1
1	1	-1	-1	-1	-1
3	-1	1	-1	-1	-1
10	1	1	1	-1	-1
2	-1	-1	1	1	-1
12	1	-1	-1	1	-1
8	-1	1	-1	1	-1
16	1	1	1	1	-1
4	-1	-1	1	-1	1
14	1	-1	-1	-1	1
9	-1	1	-1	-1	1
11	1	1	1	-1	1
6	-1	-1	1	1	1
15	1	-1	-1	1	1
13	-1	1	-1	1	1
7	1	1	1	1	1

- Diseño split-plot: se puede asociar al “plot” tanto los factores de control como los factores ruido:
 - Plot: factores de control. Se seleccionan de forma aleatoria las 4 filas de la matriz producto en la Tabla 2.1 y en cada una de las 4 condiciones del tratamiento térmico se fabrican todas las piezas que van a ser utilizadas posteriormente en las 4 condiciones de ruido. Una vez fabricadas estos grupos de 4 piezas, las piezas dentro de cada grupo se asignan de forma aleatoria a las 4 condiciones de ruido y se realizan las 16 pruebas de forma aleatoria. Resultado: 4 operaciones de proceso y 16 operaciones de ruido.

Tabla 2.6 Diseño *split-plot* a partir de la matriz en Tabla 2.1. Plot: F. de control

Orden de realización	T. Inicial (A)	T. Total (B)	Temperatura (C)		Orden de realización	Posición Horno (R)	Tiempo espera (S)
2	-1	-1	1	→	5	-1	-1
					13	1	-1
					3	-1	1
					1	1	1
1	1	-1	-1	→	7	-1	-1
					14	1	-1
					8	-1	1
					11	1	1
3	-1	1	-1	→	4	-1	-1
					16	1	-1
					9	-1	1
					15	1	1
4	1	1	1	→	6	-1	-1
					10	1	-1
					12	-1	1
					2	1	1

- Plot: factores ruido. Se seleccionan de forma aleatoria las 16 condiciones de los factores de control y se seleccionan de forma aleatoria las 4 columnas de la matriz producto en la Tabla 2.1. Se trabaja en una condición de ruido con las piezas provenientes de 4 condiciones distintas de proceso. Resultado: 16 operaciones de proceso y 4 operaciones de ruido.

Tabla 2.7 Diseño *split-plot* a partir de la matriz en Tabla 2.1. Plot: f. ruido

Orden de realización	T. Inicial (A)	T. Total (B)	Temperatura (C)		Orden de realización	Posición Horno (R)	Tiempo espera (S)
5	-1	-1	1	}	4	-1	-1
1	1	-1	-1				
3	-1	1	-1				
10	1	1	1				
2	-1	-1	1	}	1	1	-1
12	1	-1	-1				
8	-1	1	-1				
16	1	1	1				
4	-1	-1	1	}	2	-1	1
14	1	-1	-1				
9	-1	1	-1				
11	1	1	1				
6	-1	-1	1	}	3	1	1
15	1	-1	-1				
13	-1	1	-1				
7	1	1	1				

La asignación de factores al “plot” no tiene porqué venir dada en función de la categoría factores de control o factores ruido, se podría pensar en tomar una selección de factores de ambos grupos. De hecho, puesto que lo que origina esta peculiar manera de ejecución son criterios de coste se podría realizar una clasificación entre factores “difíciles de mover” y el resto y a partir de esta clasificación seleccionar el diseño (Bisgaard (2000) y Wu y Hamada (2000) tiene algunas sugerencias bajo esta filosofía).

- Diseño *strip-plot*: Tanto los factores de control como los factores ruido se asocian a un plot. En este caso, las piezas fabricadas como en el caso del *split-plot* con los factores de control como plot, se agrupan en cuatro grupos para ser sometidas a 4 condiciones de ruido seleccionadas de forma aleatoria según las columnas de la misma matriz. Resultado: 4 operaciones de proceso y 4 operaciones de ruido. La Tabla 2.8 presenta una posible selección.

Tabla 2.8 Diseño *strip-plot* a partir de la matriz en Tabla 2.1

Orden de realización	T. Inicial (A)	T. Total (B)	Temperatura (C)		Orden de realización	Posición Horno (R)	Tiempo espera (S)
2	-1	-1	1		4	-1	-1
1	1	-1	-1		1	1	-1
3	-1	1	-1		2	-1	1
4	1	1	1		3	1	1

Por lo tanto, partiendo de las condiciones experimentales que determina el diseño factorial 2^3 seleccionado, hemos llegado a diferentes diseños que, aun teniendo las mismas condiciones de partida:

- Se ejecutan de manera muy diferente y por lo tanto tienen presupuestos de ejecución muy variable;
- Se analizan de manera diferente (como veremos más adelante).

Jones (1991) compara la eficiencia de los diseños totalmente aleatorizados, diseños bloqueados totalmente aleatorizados, diseños “split-plot” y diseños “strip-plot”. La Tabla 2.9 muestra un resumen de esta comparación. Notar que aunque por criterios estadísticos de eficiencia es más recomendable en un diseño “split-plot” asignar el “plot” a los factores ruido que a los factores de control, en la práctica el coste asociado a uno u otro diseño puede ser tan diferente como para determinar el diseño a seleccionar.

Tabla 2.9 Comparación de movimientos en las condiciones de control, M_c , y ruido, M_r , para los distintos diseños así como la eficiencia en la estimación

	Totalmente Aleatorizado	Split-Plot Plot = C	Split-Plot Plot = R
Totalmente Aleatorizado $M_c = 2^{k+r}, M_r = 2^{k+r}$			
Split-Plot Plot = Factor Control: (C) $M_c = 2^k, M_r = 2^{k+r}$	+ : R, R*R C*R - : C, C*C		= : C*R + : R, R*R - : C, C*C
Split-Plot Plot = Factor Ruido: (R) $M_c = 2^{k+r}, M_r = 2^r$	+ : C, C*C, C*R - : R, R*R	= : C*R + : C, C*C - : R, R*R	
Strip-Plot Plot: C y R $M_c = 2^k, M_r = 2^r$	+ : C*R - : C, C*C, R, R*R	+ : C*R - : R, R*R = : C, C*C	+ : C*R - : C, C*C = : R, R*R

+ : Efectos que son estimados mejor por el diseño especificado en la fila que por el especificado en la columna

- : Efectos que son estimados peor por el diseño especificado en la fila que por el especificado en la columna

= : Efectos que son estimados con igual eficiencia por el diseño especificado en la fila que por el especificado en la columna

2.2.4 Conclusiones sobre la selección de la Matriz de diseño

Resumimos este apartado en los siguientes puntos:

- Existen principalmente 2 métodos de seleccionar una matriz de diseño: A partir de la matriz producto (propuesta por Taguchi) o a partir de una matriz extendida conteniendo a todos los factores de control y ruido;
- Los diseños ejecutados a partir de una matriz producto permiten el análisis de los datos utilizando *métricas resumen* y *métricas no-resumen*. En este caso utilizando *métricas no-resumen* se pueden estimar libre de confusiones las interacciones entre factores de control y factores ruido. Si se desea que estos diseños tengan resolución elevada se necesita un número elevado de condiciones experimentales;
- Si el diseño se ejecuta a partir de una matriz extendida sólo se puede analizar con *métricas no-resumen*. Si este diseño se selecciona con criterios de máxima resolución probablemente no permitirá la estimación libre de confusiones de las interacciones entre factores de control y ruido;
- Jones (1990) y Wu y Zhu (1999) entre otros plantean diseños a partir de una matriz extendida tratando de una manera muy particular las interacciones entre factores de control y factores ruido. La matriz producto es un caso particular de estos diseños pero en general se llega a diseños con menos condiciones experimentales;

- Una restricción importante en la selección del diseño es el presupuesto económico que se tiene para la realización de las pruebas; este presupuesto determina, aunque no directamente, el número de factores a seleccionar y el total de experimentos que pueden realizarse.
- Existen varias maneras de ejecutar las condiciones que definen una matriz producto dando lugar a tres categorías de diseños de coste de ejecución muy diferente: totalmente aleatorizado, split-plot y strip-plot. Si se analizan estos diseños con métricas no-resumen hay que tener en cuenta las diferentes fuentes del error experimental;
- En la práctica, a la hora de seleccionar un diseño se ha de tener en cuenta diversas alternativas teniendo en cuenta: El número de factores de control y de ruido con los que se desea experimentar, el número de condiciones u operaciones diferentes de ejecución, el coste de cada condición experimental en los factores de control y ruido, los efectos que pueden ser estimados y el poder de detección de efectos significativos.
- Aunque se ha avanzado enormemente en este campo creemos que es necesario continuar la investigación en esta línea de cara a facilitar la selección de diseños en la práctica. Para ello se debería tener en cuenta las propuestas de Wu (1999), Jones (1990), Bisgaard (2000) y Wu y Hamada (2000).

2.3 Análisis de datos y detección de efectos significativos

Cuando la experimentación se ha llevado a cabo partiendo de diseños factoriales, la estimación de los efectos se realiza por mínimos cuadrados ordinarios y los métodos más utilizados para la detección de los efectos significativos están basados en *t-test* individuales y en gráficos en papel probabilístico normal o semi normal (Daniel, 1959). En el primer caso se requiere tener una estimación del error estándar del efecto y en el segundo no es necesario conocer este error, pero sólo se pueden llevar al gráfico efectos con el mismo error estándar.

Si la experimentación se ha realizado partiendo de una matriz producto, las estimaciones mínimo cuadradas de los efectos sobre la respuesta *Y* que se obtienen a partir de *métricas resumen* o *métricas no-resumen* son las mismas e independientes del orden de ejecución de las pruebas; sin embargo las estimaciones del error experimental, y por consiguiente del error estándar de los efectos, sólo son iguales si el diseño se ha ejecutado aleatoriamente.

Por consiguiente se puede estimar los efectos por el método que se desee pero sin tener en cuenta el grado de significación si el diseño no es aleatorizado; para obtener el grado de significación se puede estimar cada fuente de error a partir de una análisis de la

varianza, ANOVA, y aplicarlo a cada grupo de efectos. La Tabla 2.10 resume estas situaciones.

Tabla 2.10 Estimación del grado de significación de los efectos

	M. resumen (M. Cuadrados Ordinarios)	Métrica no-resumen (M. Cuadrados Ordinarios)	ANOVA
Totalmente aleatorizado	Correcto	Correcto	-----
Split-plot (Plot = F. Control)	Correcto	Incorrecto	2 grupos: <ul style="list-style-type: none"> • C, C*C • R, R*R, C*R
Split-plot (Plot = F. Ruido)	Incorrecto	Incorrecto	2 grupos: <ul style="list-style-type: none"> • R, R*R • C, C*C, C*R
Strip-plot	Incorrecto	Incorrecto	3 grupos: <ul style="list-style-type: none"> • C, C*C • R, R*R • C*R

Si las pruebas se han ejecutado según un diseño en *split-plot* o *strip-plot*, existen dos fuentes de error diferentes para el primer caso, la que viene del “plot” y la que viene del “subplot”, y tres para el último, la del “plot”, el “subplot” y la de la interacción entre “plot” y “subplot”. Por consiguiente, si se utilizan los métodos habituales de estimación mínimo cuadrática asumiendo que el diseño es totalmente aleatorizado:

- Si se utilizan *métricas resumen* la estimación que se obtiene del error estándar de los efectos (todos ellos asociados a factores de control) será correcta si el diseño es aleatorizado o es un *split-plot* con el “plot” los factores de control;
- Si se utilizan *métricas no-resumen*, la estimación del error estándar de los efectos sólo es correcta si se ha aleatorizado el diseño.

Para la detección de efectos significativos se pueden utilizar *t-test* o gráficos en papel probabilístico normal teniendo cuidado de realizar un gráfico para cada grupo de efectos con el mismo error estándar (Jones (1991) y Bisgaard(2000)).

Si, como ocurre habitualmente, se parte de prototipos fabricados seguidos sin mover las condiciones de proceso, es decir como un *split-plot* con el “plot” los factores de control, es mucho más sencillo analizar los datos con una *métrica resumen* que hacerlo con una *no-resumen* que incluya a todos los factores en el modelo, ya que este último análisis implica la estimación de varias fuentes de variación y tratar los efectos de una manera más sofisticada.

En la práctica se desconoce la magnitud del error experimental y los experimentos son llevados a cabo sin realizar réplicas genuinas por lo que se ha de recurrir a métodos de estimación del error que dependen del diseño seleccionado.

Si se parte de un diseño factorial ejecutado de forma aleatoria, la estrategia de selección propuesta por Taguchi de agrupar los efectos más pequeños para estimar el error experimental, tiende a incrementar el error de tipo I, es decir, tiende a considerar significativos efectos que no lo son. Stephens y Mouw (1999) aplicando el método de selección propuesto por Taguchi demuestran que en un diseño con 15 efectos aleatorios es muy probable seleccionar 3 de ellos como significativos con $\alpha=0.01$.

Para el mismo diseño con 15 efectos aleatorios, Box (1988) varía ligeramente el método propuesto por Taguchi y, partiendo de los 7 efectos más pequeños para estimar el error, muestra lo probable que es detectar como significativos entre 3 y 6 de los 8 factores restantes. Por esta razón este autor recomienda utilizar el principio de “sparsity” y tomar aquellos efectos que representan a interacciones de orden superior a la hora de estimar el error experimental. Este mismo autor recomienda utilizar información a priori sobre el número de efectos significativos en experiencias similares para seleccionar los efectos a partir de los gráficos de probabilidad Bayes (Box y Meyer, 1986).

Finalmente, como resultado del análisis, se seleccionarán aquellos efectos significativos para la métrica a estudio. Cuando se trabaja con *métricas resumen* del tipo *media* y *varianza*, los factores significativos se clasifican en:

- Factores de control “ X_d ”: Factores de control que afectan a la medida de variabilidad;
- Factores “ X_a ”: Factores de control que afectan sólo al ajuste y no afectan a la variabilidad;

En cambio, cuando se trabaja con *métricas no-resumen* se estudian tanto los factores de control como los ruidos por ello la clasificación de los factores significativos es como sigue:

- Factores de control que interactúan con factores ruidos (equivalente a los “ X_d ” anteriores);
- Factores ruidos “ Z_d ” que afectan a la respuesta e interactúan con los factores de control “ X_d ”;
- Factores de control que no interactúan con factores ruidos pero afectan a la respuesta (equivalente a los “ X_a ” anteriores);

- Factores ruido “ Z_a ” que afectan a la respuesta pero no interaccionan con factores de control.

Independientemente de la métrica utilizada, las condiciones óptimas a la variabilidad transmitida por los factores ruido implicará a factores de control del tipo “ X_d ”.

2.3.1 Selección de condiciones óptimas del modelo

En el capítulo 1 se hizo una descripción de cómo se seleccionaban las condiciones óptimas a partir de la metodología propuesta por Taguchi por un procedimiento en dos pasos. Si se trabaja con S/R , la selección de los niveles de los factores significativos que maximizan esta métrica se hace a partir de los gráficos de los efectos. Ya hemos visto que la optimización de S/R no tiene porqué estar ligada siempre a una optimización de una función de pérdidas y que en la mayoría de las situaciones se tiende a trabajar de forma separada con métricas para localización y para la dispersión, es decir asumiendo mientras no se tenga evidencia de lo contrario, que ambas medidas son independientes..

Cuando se utilizan *métricas resumen* alternativas a las propuestas por Taguchi, *media* y $\log(s)$ por ejemplo, se llega a un conjunto de factores de control que aparecen como significativos en alguna de las anteriores métricas o en ambas. Si la optimización se hace por separado, las condiciones óptimas se seleccionan a partir de los gráficos de los efectos sobre la *media* y sobre $\log(s)$.

Cuando el análisis de los datos se realiza a partir de *métricas no-resumen*, modelando directamente la respuesta Y en función de los factores de control y ruido mediante aproximaciones polinómicas, nos encontramos que el modelo general estimado

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l + \varepsilon$$

$i, j = 1, \dots, k \quad l, m = 1, \dots, r$

tiene una parte que permanecerá fija una vez seleccionados los niveles de los factores de control

$$\beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j \tag{2.24}$$

$i, j = 1, \dots, k$

y tiene una parte “variable”, donde el valor esperado de la respuesta en las condiciones de control depende del nivel en el que se encuentren ciertos factores ruido:

$$\sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l + \varepsilon \quad (2.25)$$

$i, j = 1, \dots, k \quad l, m = 1, \dots, r$

Suponiendo que queremos optimizar por separado ambas expresiones:

- La selección de los niveles óptimos de los factores significativos que afectan a la *media* se realiza en general a partir de los gráficos de los efectos;
- La selección de los niveles óptimos de los factores significativos que afectan a la *dispersión* se puede realizar por dos métodos principalmente:
 - A partir de los gráficos de interacciones de factores de control y factores ruido que surgen significativos en (2.25), seleccionando los niveles de los factores de control en los que el efecto de los factores ruido se minimiza;
 - Minimizando la expresión de la Varianza(Y) que se obtiene a partir de la parte correspondiente del modelo para Y;

$$V(Y) = V\left(\sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{ij} X_i Z_l\right) + \sigma^2 \quad (2.26)$$

$i, j = 1, \dots, k; \quad l, m = 1, \dots, r$

Como veremos más adelante, la existencia de interacciones entre factores de control y factores ruido es un requisito necesario pero no suficiente para la existencia de condiciones robustas, y las soluciones obtenidas por cada uno de los dos métodos, en contra de lo que podría esperar, no siempre coinciden.

Independientemente del tipo de análisis que se realice (a partir de *métricas resumen* o modelando directamente la respuesta), cuando en la selección de condiciones óptimas se desea tener en cuenta los dos criterios de media y variabilidad a la vez puede haber conflicto entre ambos y por ellos diversos autores recomiendan distintas estrategias. Jones (1991) construye una métrica nueva a partir de las dos anteriores ponderando el peso de cada una de las partes a partir de un coeficiente “ λ ” difícil de seleccionar en un entorno real de experimentación. Grima (1993) propone analizar las distintas soluciones alternativas en un gráfico bivalente “*media - varianza*” que permite tener una idea global de cada situación.

La dificultad que entraña dar un peso a una parte sobre la otra hace que la mayoría de autores sigan manteniéndose en las dos métricas originales a la hora de seleccionar las condiciones y se estudien varias soluciones alternativas utilizando criterios externos

al propio estudio para decantarse por una solución u otra (criterios de coste por ejemplo).

Selección de condiciones robustas al error experimental

En todos los análisis presentados hasta el momento se ha asumido que el modelo de partida era cierto y que los errores o discrepancias entre los valores observados y los que predice el modelo siguen una distribución normal con varianza constante e independiente de cada factor.

Cuando se dispone de réplicas, es posible detectar problemas de heterocedasticidad, y por consiguiente oportunidades de seleccionar condiciones robustas a este ruido, estudiando la variabilidad en cada condición experimental por métodos similares a los expuestos anteriormente.

Si no hay réplicas, algunos autores han estudiado métodos para minimizar $V(\epsilon)$ a partir de un tratamiento especial de los residuos (Ferrer y Romero, 1995 y Fuller and Bisgaard, 1995). Se trataría de trabajar con una transformación de los residuos

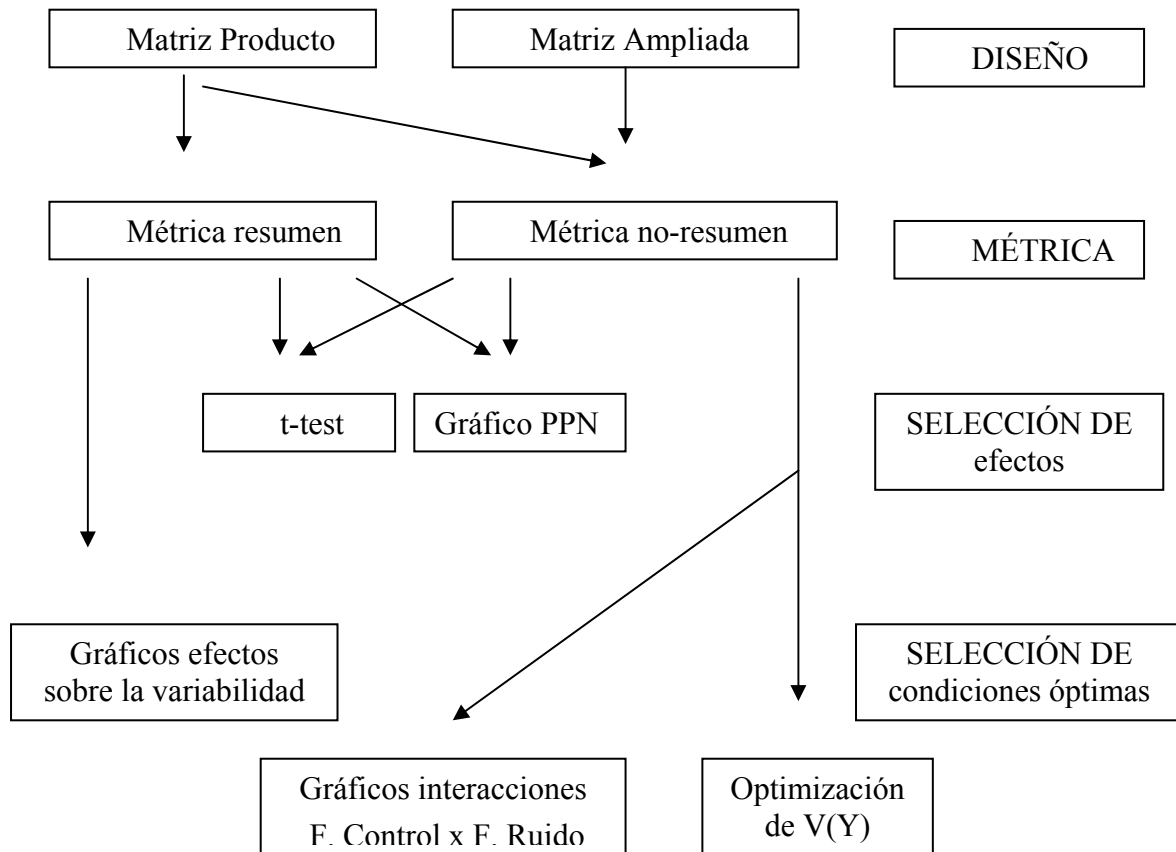
$$Y_i = \hat{Y}_i + e_i$$

$$w_i \begin{cases} \ln e_i^2 + 1.27 & \text{si } e_i \neq 0 \\ \ln(0.5 \min \|e_i\|)^2 + 1.27 & \text{si } e_i = 0 \end{cases} \quad (2.27)$$

y modelar esta respuesta en función de los factores de control con el objetivo de encontrar condiciones robustas al ruido externo.

$$w_i = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \epsilon \quad (2.28)$$

2.4 Controversias



La presentación que hemos llevado a cabo en este capítulo pretendía mostrar que la metodología más habitual utilizada para la consecución de productos o procesos robustos mediante la experimentación abarca principalmente:

- 2 tipos de métricas: Métricas resumen y Métricas no-resumen;
- 2 tipos de matriz de diseño: Matriz producto y Matriz extendida;
- Varias maneras de ejecutar los experimentos: Totalmente aleatorizada, Split-plot y Strip-plot;
- Básicamente 2 maneras de seleccionar los efectos significativos: T-test individuales y gráficos en papel probabilístico normal;
- 2 maneras de seleccionar las condiciones robustas: Gráficos de interacciones entre factores de control y ruido cuando se estudia el modelo de localización

para Y o gráficos de los efectos significativos cuando se analiza directamente la variabilidad.

Antes de tomar partida por una u otra manera de seleccionar los diseños, analizarlos y seleccionar las condiciones óptimas, sería conveniente comparar los métodos bajo las mismas condiciones de trabajo:

- Si el diseño de partida es un diseño en forma de matriz producto se pueden comparar los efectos significativos encontrados al analizarlo a partir de métricas resumen o métricas no-resumen;
- Si el diseño de partida es un diseño en forma de matriz producto se pueden comparar las condiciones robustas seleccionadas utilizando métricas resumen con las obtenidas a partir métricas no-resumen bien con gráficos de interacciones o bien a partir de la estimación de $V(Y)$.

Aunque teóricamente los resultados obtenidos en las dos comparaciones han de ser similares, en la práctica esto no siempre es así y prueba de ello son los dos ejemplos que presentaremos en el siguiente capítulo.

Estas diferencias que aparecen en estos ejemplos han sido la semilla de la investigación que hemos llevado a cabo y que va encaminada a encontrar las causas que originan que las conclusiones sobre un estudio de robustez son sensibles al método analítico seleccionado.

Una vez encontradas las causas de estas diferencias tomaremos partida a la hora de elaborar unas pautas de actuación de cara a experimentar con el fin de obtener productos o procesos robustos.

