

## CAPÍTULO 4

# Gráfico Media Desviación

### 4.1 Introducción

Entre los métodos gráficos para el análisis de los resultados en el diseño de productos robustos cabe destacar el desarrollado por Grima (1993), consistente en un diagrama bivariante que representa las estimaciones del valor esperado de la respuesta frente a su desviación estándar. Adicionalmente a la simplicidad que proporciona un gráfico de este tipo en la búsqueda de las mejores condiciones de proceso, cuenta con la ventaja de que la representación bivariante es independiente del número de factores controlable en estudio, cosa que no sucede con otros métodos gráficos, como aquellos que hacen uso de curvas de nivel y que están restringidos a dos variables controlables.

Si algo caracteriza a este método es su versatilidad y sencillez, lo cual lo hace idóneo para su difusión como método intuitivo y fácil de usar en el análisis de los resultados de un plan de experimentación con factores ruido. Por esta razón se ha profundizado en el estudio del método, buscando aportaciones en algunos de sus aspectos formales y prácticos. En este capítulo se expone el procedimiento, basado en el gráfico Media Desviación (MD) y, para destacar la consistencia y simplicidad de su uso, se contrasta con el empleo de otros métodos que se encuentran en la literatura. Se generaliza a modelos más complejos con interacciones de hasta tres factores, se estudia la posibilidad de que sea incorporado en un paquete estándar de software estadístico y se completa con el problema de los costes y el análisis de respuesta múltiple.

## 4.2 El gráfico Media Desviación

El **gráfico media desviación** MD parte de la idea de mostrar en un diagrama bivalente los valores de la media y la desviación estándar de la respuesta, para un conjunto de valores de los factores controlables en la región de experimentación. Este gráfico hace muy simple la tarea de encontrar la mejor combinación de valores de los factores controlables, a partir de valores deseables de la media y la desviación tipo. Adicionalmente, y no menos importante, se consigue un conocimiento del comportamiento conjunto que tienen los valores del promedio y de la desviación estándar de la respuesta en la región de análisis de los factores controlables, ante la presencia de factores ruido.

La elaboración del gráfico consiste en los siguientes pasos:

- Plantear el modelo de la respuesta y estimar sus coeficientes.
- En los modelos de la media y la varianza sustituir los valores de los parámetros por los correspondientes coeficientes estimados en el paso anterior.
- “Cubrir” la región de interés con un conjunto de puntos del tipo  $(x_{i1}, x_{i2}, \dots, x_{ip})$
- Sustituir cada punto  $(x_{i1}, x_{i2}, \dots, x_{ip})$  en el modelo estimado de la media y en el de la varianza para formar los puntos  $(m_i, d_i)$  del gráfico MD
- Representar los puntos  $(m_i, d_i)$  en un diagrama bivalente

Para ilustrar el funcionamiento del gráfico nos basaremos en el ejemplo tratado por Myers y Montgomery (2002) para analizar la calidad, burdo contra detallado, en la imagen transmitida a un televisor de color; los factores controlables son el número de lengüetas en un filtro,  $x_1$ , y la frecuencia,  $x_2$ , mientras que los factores ruido representan el número de bits en la imagen,

$z_1$ , y el voltaje aplicado,  $z_2$ . El diseño empleado es un arreglo cruzado  $3^2 \times 2^2$ , Tabla 4.1.

**Tabla 4.1** Datos de la calidad de imagen

		$z_1$	-1	-1	1	1
$x_1$	$x_2$	$z_2$	-1	1	-1	1
-1	-1		33,5	41,2	25,3	32,0
-1	0		35,8	38,1	32,8	34,0
-1	1		33,1	31,8	36,2	34,0
0	-1		30,5	41,3	15,1	24,0
0	0		34,9	40,2	27,8	31,1
0	1		35,2	37,1	33,3	35,2
1	-1		21,1	34,1	0,8	15,7
1	0		27,7	38,1	15,5	26,0
1	1		32,1	38,1	26,2	32,1

El primer paso para el uso del gráfico MD consiste en plantear el modelo de la respuesta y estimar sus coeficientes. El arreglo interno se escogió de tal manera que permita estimar un modelo de segundo orden en los factores controlables del tipo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \gamma_1 z_1 + \gamma_2 z_2 + \delta_{11} x_1 z_1 + \delta_{12} x_1 z_2 + \delta_{21} x_2 z_1 + \delta_{22} x_2 z_2. \quad (4.1)$$

El ajuste por mínimos cuadrados está dado por:

$$\hat{y} = 33,389 - 4,175x_1 + 3,748x_2 + 3,3485x_1x_2 - 2,328x_1^2 - 1,867x_2^2 - 4,076z_1 + 2,985z_2 - 2,324x_1z_1 + 1,932x_1z_2 + 3,268x_2z_1 - 2,073x_2z_2. \quad (4.2)$$

El valor estimado de la varianza del error (cuadrado medio del error) es de 0,56. Todos los coeficientes del modelo estimado resultan significativos. Los modelos de la media y de la varianza de la respuesta se derivan al aplicar los operadores  $E(\cdot)$  y  $V(\cdot)$  y son:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2, \quad (4.3)$$

$$V(y) = (\gamma_1 + \delta_{11} x_1 + \delta_{12} x_2)^2 + (\gamma_2 + \delta_{21} x_1 + \delta_{22} x_2)^2 + \sigma_\varepsilon^2. \quad (4.4)$$

Los modelos estimados de la media y la varianza de la respuesta se obtienen al sustituir los coeficientes correspondientes del modelo ajustado de la respuesta, 4.2, en las dos expresiones anteriores. Con la consideración de que los factores ruido están estandarizados, i.e.  $E(Z_j) = 0$  y  $V(Z_j) = 1$ , y son independientes, o al menos incorrelacionados, se tiene:

$$\hat{E}(y) = 33,389 - 4,175x_1 + 3,748x_2 + 3,3485x_1x_2 - 2.328x_1^2 - 1,867x_2^2, \tag{4.5}$$

$$\hat{V}(y) = (-4,076 - 2,324x_1 + 3,268x_2)^2 + (2,985 + 1,932x_1 - 2,073x_2)^2 + 0,56. \tag{4.6}$$

En la región de interés  $\{(x_1, x_2) | 0 \leq x_1, x_2 \leq 1\}$  se hace un entramado de puntos  $(x_{i1}, x_{i2})$ ,  $i = 1, 2, \dots, n$ , tan denso como se desee, por ejemplo  $(-1, -1)$ ,  $(-0.857, -1)$ ,  $(-0.714, -1)$ ,  $\dots$ ,  $(1, 1)$ , y para cada pareja se calcula su media y su desviación estándar estimada a partir de las ecuaciones 4.5 y 4.6. El último paso consiste en representar en un diagrama de dispersión estos valores. La Figura 4.1 tiene el gráfico elaborado en Excel.

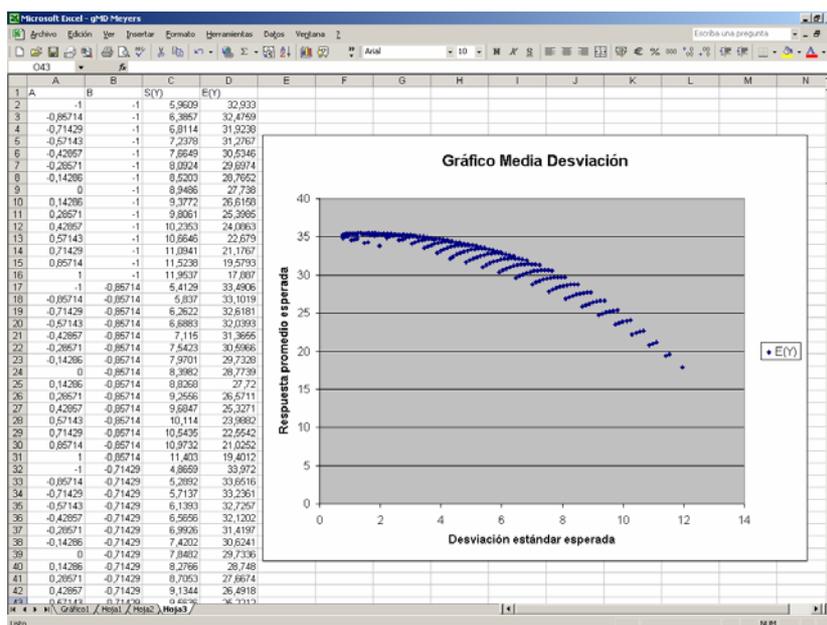


Figura 4.1 Gráfico Media Desviación elaborado en Excel.

Para sistematizar la elaboración del gráfico MD, y poder estudiar la posibilidad de que éste se incorpore en un paquete estándar de software estadístico, con algunas variantes de interés, se construyó un algoritmo (*gMD*) codificado en forma de macros para el paquete estadístico Minitab (ver Anexo 3). Se escogió esta opción de software por la gran difusión que tiene en el ámbito académico y por su creciente uso en la industria, al formar parte de las herramientas que se emplean frecuentemente en la implementación de la metodología Seis Sigma.

Adicionalmente el paquete Minitab cuenta con opciones sencillas para la identificación de puntos en sus gráficos bivariantes, lo cual es de gran utilidad en la determinación de las condiciones robustas y en el análisis de respuestas múltiples. El algoritmo permite trabajar con modelos que contienen efectos lineales y cuadráticos en los factores controlables y de ruido, interacciones dobles entre los distintos factores e interacciones triples entre dos factores controlables y uno de ruido o dos de ruido y uno controlable. El despliegue de los puntos se puede hacer únicamente para los valores de los factores en la matriz de diseño (“vértices” de la región) o sobre toda la región de experimentación (vértices e interior). Se puede también incluir costes de operación de cada nivel de los factores controlables.

Cuando se “llama” al programa se le debe indicar cuáles de las opciones mencionadas anteriormente se emplearán. Es necesario que previamente se introduzcan los coeficientes del modelo estimado de la respuesta, junto con los valores extremos que definen la región de interés (los niveles de los factores controlables en una primera etapa). Los valores de los coeficientes se introducen en la primera columna de la hoja de trabajo, en la segunda se ubican las letras que identifican a los términos correspondientes del modelo; se emplea de la letra A a la G para los factores controlables y de la S a la W para los factores ruido y se señalan los valores extremos de cada factor controlable en la primera y segunda casilla de las columnas subsecuentes. La Figura 4.2 ilustra como se introducen los datos de nuestro ejemplo.

	C1	C2-T	C3	C4	C5	C6	C7	C8	C9
	coef	ter	A	B					
1	-4,175	A	-1	-1					
2	3,748	B	1	1					
3	3,348	AB							
4	-2,328	AA							
5	-1,867	BB							
6	-4,076	S							
7	2,985	T							
8	-2,324	AS							
9	1,932	AT							
10	3,268	BS							
11	-2,073	BT							
12									

Figura 4.2 Datos de entrada, coeficientes y términos del modelo.

Tal como se mencionó, el programa *gMD* puede elaborar el gráfico únicamente para los puntos que están definidos por la matriz de diseño de los factores controlables (diseño factorial  $2^p$ ), los cuales son los “vértices” de la región de experimentación, o bien para un conjunto mayor de puntos, en la frontera y en el interior de la región de experimentación; esta segunda opción se muestra en la hoja de trabajo de la Figura 4.3. Para todos estos puntos se calculó y se representó en el diagrama bivalente su esperanza y su desviación tipo.

Minitab brinda una opción para identificar información relevante asociada a un punto o grupo de puntos en un diagrama bivalente. Al enmarcar los puntos en el gráfico éstos cambian (puntos sólidos) y aparece un recuadro con la información (en la parte superior derecha del gráfico MD). Este recuadro tiene, en el ejemplo, los valores de los factores de control (A y B) que generaron los puntos seleccionados y los valores correspondientes de la media y la desviación estándar. Se identifica fácilmente las cantidades que corresponden a las máximas respuesta promedio (entre 34,5 y 35,4) y los valores mínimos que toma la desviación tipo (entre 0,78 y 1,04). Así mismo se sabe que estas cantidades se consiguieron cuando los valores de los factores controlables se encontraban entre -1 y -0,29 para  $x_1$  (A) y entre 0,43 y 1.0 para  $x_2$  (B).

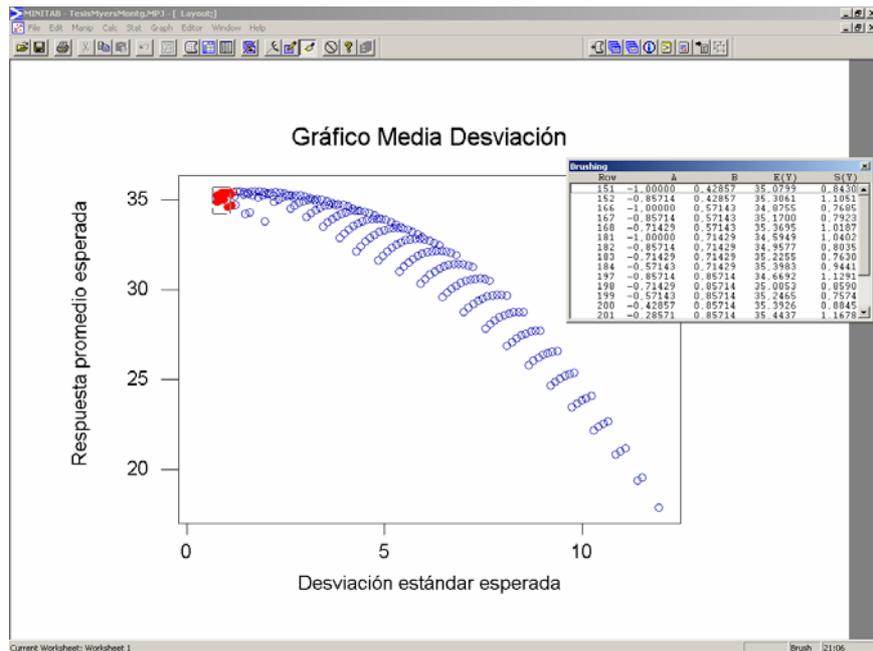


Figura 4.3 Resultados del algoritmo que construye el gráfico MD.

Las figuras 4.4 y 4.5 ilustran lo sencillo que es llevar a cabo la búsqueda de las mejores condiciones de proceso en forma secuencial. La primera de ellas se enfoca en la región que da mínima variabilidad y máxima respuesta promedio en la Figura 4.3 (puntos sólidos). Por lo tanto el entramado de puntos para la nueva figura se hace sobre la zona delimitada por los valores que generaron estos puntos sólidos ( $-1 \leq x_1 (A) \leq -0,29$  y  $0,43 \leq x_2 (B) \leq 1,0$ ).

En la Figura 4.4 los valores de la respuesta promedio ya no fluctúan entre 0 y 35, sino entre 33,8 y 35,5, y la desviación tipo solamente varía entre 0,75 y 2,6 en lugar de entre 0,75 y 12, aproximadamente.

La Figura 4.5 repite el proceso y expande la nueva región enmarcada en el gráfico de la Figura 4.4. Esta parte enmarcada es la región en la que puede existir un interés por explorar el comportamiento conjunto de la media y la desviación estándar de la respuesta, por tener la mínima variabilidad.

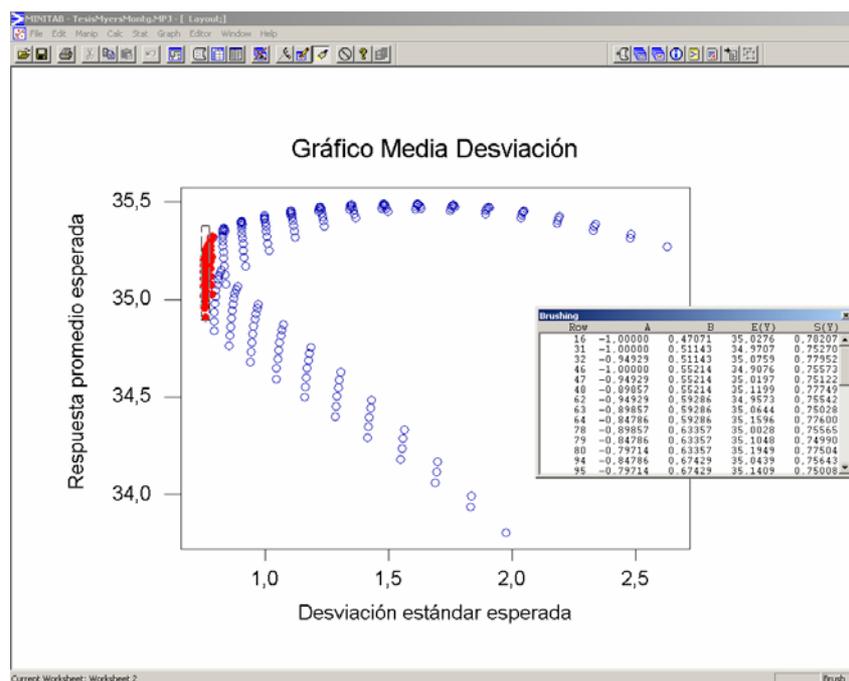


Figura 4.4 Resultados para la región  $-1 \leq x_1 \leq -0,29$  y  $0,43 \leq x_2 \leq 1,0$

Obsérvese que todos los puntos en el recuadro de la Figura 4.5 tienen prácticamente los mismos valores de la desviación y de la respuesta promedio, por lo que se decide detener el proceso de búsqueda secuencial; estos valores se encuentran en torno a 35 para la media y a 0,75 para la desviación estándar.

Meyers y Montgomery (2002) señalan que para este ejemplo la expresión de la varianza permite encontrar algebraicamente su valor mínimo en  $x_1 = -0,874$  y  $x_2 = 0,625$ , lo que proporciona aproximadamente una media estimada del proceso de 35dB y una estimación de la desviación tipo del proceso de 0,748. Esta condición robusta de operación se encuentra en la fila 113 en el recuadro de la Figura 4.5. En realidad todas las condiciones de operación que se muestran en él son condiciones robustas, esto da una gran flexibilidad al investigador. Al contar con más de una combinación posible de valores de los parámetros se pueden usar criterios adicionales (costes,

facilidad de operación, capacidad de almacenamiento de materias primas,...) en el momento de determinar las condiciones de proceso más apropiadas.

Row	A	B	E(Y)	S(Y)
81	-0,90714	0,589286	35,0312	0,74959
82	-0,88857	0,589286	35,0679	0,75461
95	-0,92571	0,607143	34,9645	0,75326
96	-0,90714	0,607143	35,0040	0,74897
97	-0,88857	0,607143	35,0418	0,74887
98	-0,87000	0,607143	35,0781	0,75298
112	-0,88857	0,625000	35,0146	0,74949
113	-0,87000	0,625000	35,0519	0,74847
114	-0,85143	0,625000	35,0877	0,75165
128	-0,87000	0,642857	35,0246	0,75033
129	-0,85143	0,642857	35,0614	0,74838
130	-0,83286	0,642857	35,0967	0,75063
144	-0,85143	0,660714	35,0340	0,75147

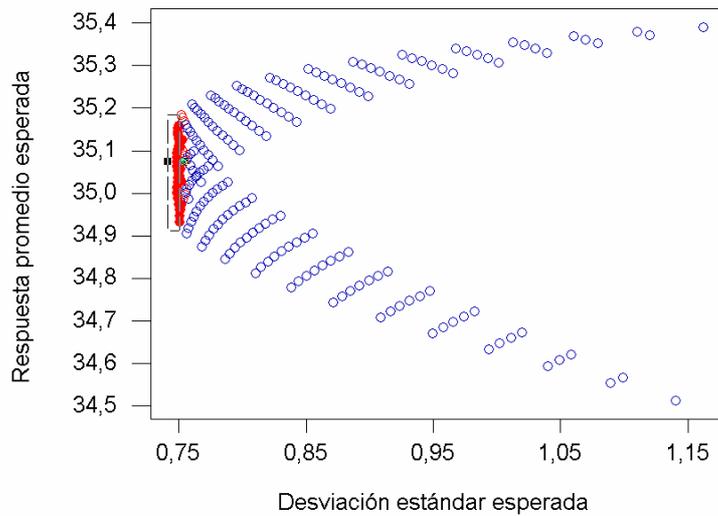


Figura 4.5 Resultados para la región  $-1 \leq x_1 \leq -0,74$  y  $0,5 \leq x_2 \leq 0,75$ .

Aunque el objetivo primario del diseño de productos robustos sea la búsqueda de las condiciones de proceso que sean insensibles a los factores ruido, la práctica de la experimentación debe conducir a un incremento en el conocimiento del fenómeno que se estudia. En ese sentido, el gráfico MD da una gran información acerca del comportamiento conjunto de la respuesta promedio y su variabilidad, en la región de experimentación, ante la posible influencia de los factores ruido.

### 4.3 Modelo para la media y la varianza de la respuesta

Para tener las expresiones analíticas de la esperanza y la varianza de la respuesta se considera un modelo para la media y otro para la varianza, deducidos ambos de un modelo de la respuesta. Asumimos que la característica de calidad que nos interesa,  $Y$ , tiene una relación funcional desconocida con los factores de diseño  $X_1, X_2, \dots, X_p$  y los factores ruido  $Z_1, Z_2, \dots, Z_q$ , del tipo  $Y = f(X_1, X_2, \dots, X_p, Z_1, Z_2, \dots, Z_q)$ . Se aproxima a esta relación a través del modelo lineal

$$\begin{aligned}
 Y = & \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j Z_j + \sum_{i=1}^q \sum_{k=i}^q \gamma_{ik} Z_i Z_k \\
 & + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k + \sum_{i=1}^p \sum_{k=i}^p \sum_{j=1}^q \lambda_{ikj} X_i X_k Z_j + \sum_{j=1}^q \sum_{i=1}^p \sum_{k=i}^p \eta_{ikj} X_j Z_i Z_k + \varepsilon.
 \end{aligned} \tag{4.7}$$

Se considera que los factores ruido son variables aleatorias independientes con  $E(Z_i) = 0$  y varianza  $V(Z_i)$  y que el término del error  $\varepsilon$  es independiente de los factores ruido con media cero y varianza constante  $\sigma_\varepsilon^2$ . El modelo contiene efectos lineales y cuadráticos en  $X$  y  $Z$ , interacciones dobles entre los distintos factores (control  $\times$  control, control  $\times$  ruido y ruido  $\times$  ruido), interacciones triples entre dos factores controlables y uno de ruido e interacciones triples entre dos factores ruido y uno controlable. Los modelos de la media y la varianza de la respuesta que se usan para construir el gráfico MD se deducen a partir del modelo de la respuesta con el uso de las propiedades siguientes:

Si  $X$  e  $Y$  son variables aleatorias y  $a$  y  $b$  constantes, entonces:

- p1.  $E(a + bX + cY) = a + bE(X) + cE(Y)$   
 p2.  $V(a + bX + cY) = b^2V(X) + c^2V(Y) + 2bc \text{cov}(X, Y)$   
 p3.  $E(X^2) = V(X) + E(X)^2$

Si  $X$  e  $Y$  son variables aleatorias independientes

- p4.  $E(XY) = E(X)E(Y)$   
 p5.  $V(XY) = V(X)V(Y) + V(X)E(Y)^2 + V(Y)E(X)^2$

$$\begin{aligned}
 \text{p6.} \quad V\left(\sum_{i=1}^n a_i X_i + \sum_{i=1}^n \sum_{j=i}^n a_{ij} X_i X_j\right) &= \sum_{i=1}^n a_i^2 V(X_i) + V\left(\sum_{i=1}^n \sum_{j=i}^n a_{ij} X_i X_j\right) \\
 &+ 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n [a_i a_{ij} V(X_i) E(X_j) + a_j a_{ij} V(X_j) E(X_i)] \\
 &+ 2 \sum_{i=1}^n a_i a_{ii} (E(X_i^3) - E(X_i^2) E(X_i))
 \end{aligned}$$

$$\begin{aligned}
 \text{p7.} \quad V\left(\sum_{i=1}^n \sum_{j=i}^n a_{ij} X_i X_j\right) &= \sum_{i=1}^n a_{ii}^2 V(X_i^2) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij}^2 V(X_i X_j) \\
 &+ 2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n [a_{ij} a_{ik} V(X_i) E(X_j) E(X_k) + \\
 &a_{ik} a_{jk} V(X_k) E(X_i) E(X_j) + a_{ij} a_{jk} V(X_j) E(X_i) E(X_k)] \\
 &+ 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n [a_{ii} a_{ij} E(X_j) (E(X_i^3) - E(X_i^2) E(X_i)) \\
 &+ a_{ij} a_{jj} E(X_i) (E(X_j^3) - E(X_j^2) E(X_j))].
 \end{aligned}$$

Si  $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m$  son variables aleatorias independientes y  $a_{11}, a_{12}, \dots, a_{nm}$  constantes, entonces:

$$\begin{aligned}
 \text{p8.} \quad V\left(\sum_{i=1}^n \sum_{j=1}^m a_{ij} X_i Y_j\right) &= \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2 V(X_i Y_j) + \\
 &2 \sum_{i=1}^n \sum_{j=1}^m \sum_{k=j+1}^m a_{ij} a_{ik} V(X_i) E(Y_j) E(Y_k) + \\
 &2 \sum_{j=1}^m \sum_{i=1}^n \sum_{k=i+1}^n a_{ij} a_{kj} E(X_i) E(X_k) V(Y_j).
 \end{aligned}$$

### Los modelos de la media y la varianza de la respuesta

El modelo 4.7 se puede describir como

$$\begin{aligned}
 Y = & \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j^* Z_j + \sum_{i=1}^p \sum_{k=i}^p \gamma_{ik} Z_i Z_k \\
 & + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k + \sum_{i=1}^p \sum_{k=i}^p \sum_{j=1}^q \lambda_{ikj} X_i X_k Z_j + \varepsilon,
 \end{aligned} \tag{4.8}$$

donde  $\gamma_{ik}^* = (\gamma_{ik} + \sum_{j=1}^p \eta_{ikj} X_j)$ . A partir de la propiedad *p1* se tiene:

$$\begin{aligned}
 E(Y) = & E(\beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j Z_j + \sum_{i=1}^p \sum_{k=i}^p \gamma_{ik}^* Z_i Z_k + \\
 & \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k + \sum_{i=1}^p \sum_{k=i}^p \sum_{j=1}^q \lambda_{ikj} X_i X_k Z_j + \varepsilon) \\
 = & \beta_0 + \sum_{i=1}^p \beta_i E(X_i) + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} E(X_i X_k) + \sum_{j=1}^q \gamma_j E(Z_j) + \sum_{i=1}^p \gamma_{ii}^* E(Z_i^2) \\
 & + \sum_{i=1}^{q-1} \sum_{k=i+1}^q \gamma_{ik}^* E(Z_i Z_k) + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i E(Z_k) \\
 & + \sum_{i=1}^p \sum_{k=1}^p \sum_{j=1}^q \lambda_{ikj} X_i X_k E(Z_j) + E(\varepsilon).
 \end{aligned}$$

como  $E(z_j) = 0$ , para cualquier  $j$ , por las propiedades *p3* y *p4* se tiene el **modelo de la media** de la respuesta:

$$E(Y) = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{i=1}^p \gamma_{ii}^* V(Z_i) \cdot \tag{4.9}$$

La expresión del modelo:

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k + \sum_{j=1}^q \gamma_j Z_j + \sum_{i=1}^q \sum_{k=i}^q \gamma_{ik}^* Z_i Z_k \\ + \sum_{i=1}^p \sum_{k=1}^q \delta_{ik} X_i Z_k + \sum_{i=1}^p \sum_{k=1}^p \sum_{j=1}^q \lambda_{ikj} X_i X_k Z_j + \varepsilon,$$

se puede reescribir en términos de las variables de ruido como:

$$Y = \theta + \sum_{j=1}^q \alpha_j Z_j + \sum_{i=1}^q \sum_{k=i}^q \gamma_{ik}^* Z_i Z_k,$$

con  $\theta$  y  $\alpha_j$  constantes dadas por:

$$\theta = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k,$$

$$\alpha_j = \gamma_j + \sum_{i=1}^p \delta_{ij} X_i + \sum_{i=1}^p \sum_{k=1}^p \lambda_{ikj} X_i X_k,$$

por lo tanto, por p2:

$$V(Y) = V\left(\sum_{j=1}^q \alpha_j Z_j + \sum_{i=1}^q \sum_{k=i}^q \gamma_{ik}^* Z_i Z_k\right) + V(\varepsilon),$$

aplicando la propiedad p6 y el supuesto de media nula para las variables  $Z_j$  se tiene:

$$V(Y) = \sum_{j=1}^q \alpha_j^2 V(Z_j) + V\left(\sum_{i=1}^q \sum_{k=i}^q \gamma_{ik}^* Z_i Z_k\right) + 2 \sum_{j=1}^q \alpha_j \gamma_{jj} E(Z_j^3) + \sigma_\varepsilon^2,$$

al usar las propiedades  $p7$  y  $p5$  se tiene el **modelo de la varianza** de la respuesta:

$$V(Y) = \sum_{j=1}^q \alpha_j^2 V(Z_j) + \sum_{j=1}^q \gamma_{jj}^{*2} V(Z_j^2) + \sum_{i=1}^{q-1} \sum_{k=i+1}^q \gamma_{ik}^{*2} V(Z_i) V(Z_k) + 2 \sum_{j=1}^q \alpha_j \gamma_{jj}^* E(Z_j^3) + \sigma_\varepsilon^2 \quad (4.10)$$

El uso de este modelo, aunque general, es muy exigente en cuanto al conocimiento que se debe tener de los distintos momentos de las variables de ruido. Si asumimos que el modelo de la respuesta no tiene términos cuadráticos en las variables  $Z_j$  entonces  $\gamma_{jj}^* = 0$  y por tanto las ecuaciones 4.2 y 4.3 se simplifican sustancialmente, quedando:

$$E(Y) = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{k=i}^p \beta_{ik} X_i X_k, \quad (4.11)$$

$$V(Y) = \sum_{j=1}^q \alpha_j^2 V(Z_j) + \sum_{i=1}^{q-1} \sum_{k=i+1}^q \gamma_{ik}^{*2} V(Z_i) V(Z_k) + \sigma_\varepsilon^2. \quad (4.12)$$

### Notación matricial

Para facilitar la obtención del gráfico MD se empleará notación matricial, con lo cual el modelo de la respuesta queda expresado de la siguiente manera:

$$Y = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{z}'\boldsymbol{\Gamma}\mathbf{z} + \mathbf{x}'\boldsymbol{\Lambda}\mathbf{z} + \mathbf{z}'\boldsymbol{\Lambda} + \mathbf{x}'\mathbf{N} + \varepsilon \quad (4.13)$$

donde

$$\mathbf{x}' = [x_1, x_2, \dots, x_p], \quad \mathbf{z}' = [z_1, z_2, \dots, z_q], \quad \boldsymbol{\beta}' = [\beta_1, \beta_2, \dots, \beta_p], \quad \boldsymbol{\gamma}' = [\gamma_1, \gamma_2, \dots, \gamma_q],$$

$$\mathbf{B} = \frac{1}{2} \begin{pmatrix} 2\beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \beta_{12} & 2\beta_{22} & \dots & \beta_{2p} \\ \dots & \dots & \dots & \dots \\ \beta_{1p} & \beta_{2p} & \dots & 2\beta_{pp} \end{pmatrix}, \quad \boldsymbol{\Gamma} = \frac{1}{2} \begin{pmatrix} 2\gamma_{11} & \gamma_{12} & \dots & \gamma_{1p} \\ \gamma_{12} & 2\gamma_{22} & \dots & \gamma_{2p} \\ \dots & \dots & \dots & \dots \\ \gamma_{1p} & \gamma_{2p} & \dots & 2\gamma_{pp} \end{pmatrix},$$

$$\Lambda = \begin{pmatrix} \delta_{11} & \delta_{12} & \dots & \delta_{1q} \\ \delta_{21} & \delta_{22} & \dots & \delta_{2q} \\ \dots & \dots & \dots & \dots \\ \delta_{p1} & \delta_{p2} & \dots & \delta_{pq} \end{pmatrix}, \quad \Lambda_j = \frac{1}{2} \begin{pmatrix} 2\lambda_{11j} & \lambda_{12j} & \dots & \lambda_{1qj} \\ \lambda_{21j} & 2\lambda_{22j} & \dots & \lambda_{2qj} \\ \dots & \dots & \dots & \dots \\ \lambda_{p1j} & \lambda_{p2j} & \dots & 2\lambda_{pqj} \end{pmatrix},$$

$$\Lambda = \begin{pmatrix} \mathbf{x}'\Lambda_1\mathbf{x} \\ \mathbf{x}'\Lambda_2\mathbf{x} \\ \dots \\ \mathbf{x}'\Lambda_q\mathbf{x} \end{pmatrix}, \quad \mathbf{N}_j = \frac{1}{2} \begin{pmatrix} 2\eta_{11j} & \eta_{12j} & \dots & \eta_{1qj} \\ \eta_{21j} & 2\eta_{22j} & \dots & \eta_{2qj} \\ \dots & \dots & \dots & \dots \\ \eta_{p1j} & \eta_{p2j} & \dots & 2\eta_{pqj} \end{pmatrix} \quad y \quad \mathbf{N} = \begin{pmatrix} \mathbf{z}'\mathbf{N}_1\mathbf{z} \\ \mathbf{z}'\mathbf{N}_2\mathbf{z} \\ \dots \\ \mathbf{z}'\mathbf{N}_p\mathbf{z} \end{pmatrix}.$$

Los modelos de la esperanza y la varianza de la respuesta, a partir de las ecuaciones 4.4 y 4.5, son:

$$E(Y) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x}, \quad (4.14)$$

$$V(Y) = \boldsymbol{\alpha}'\boldsymbol{\Psi}\boldsymbol{\alpha} + \mathbf{V}'_z\mathbf{G}\mathbf{V}_z + \sigma_\varepsilon^2, \quad (4.15)$$

donde  $\boldsymbol{\alpha} = \boldsymbol{\gamma} + \Lambda'\mathbf{x} + \Lambda$ ,  $\boldsymbol{\Psi}$  es la matriz diagonal de las varianzas de los factores ruido,  $\mathbf{V}'_z = [V(Z_1), V(Z_2), \dots, V(Z_q)]$  y

$$\mathbf{G} = \frac{1}{2} \begin{pmatrix} 0 & \gamma_{12}^{*2} & \dots & \gamma_{1q}^{*2} \\ \gamma_{12}^{*2} & 0 & \dots & \gamma_{2q}^{*2} \\ \dots & \dots & \dots & \dots \\ \gamma_{1q}^{*2} & \gamma_{2q}^{*2} & \dots & 0 \end{pmatrix}.$$

Sin pérdida de generalidad, se considera que los valores para los factores ruido  $Z_j$  están centrados en su media y que los valores codificados  $\pm 1$  corresponden a los valores  $\mu_{Z_j} \pm \sigma_{Z_j}$ , de esta forma se tiene que  $E(Z_j) = 0$  y  $V(Z_j) = 1$  en la región de experimentación. El vector de varianza  $\mathbf{V}_z$  se

convierte entonces en un vector de unos de dimensión  $q$ ,  $\mathbf{1}_q$ , y la matriz  $\Psi$  se convierte en la matriz identidad. La varianza de la respuesta queda por tanto modelada por:

$$V(y) = (\gamma + \Delta'x + \Lambda)'(\gamma + \Delta'x + \Lambda) + \mathbf{1}_q' \mathbf{G} \mathbf{1}_q + \sigma_\varepsilon^2. \quad (4.16)$$

#### 4.4 Tres casos de la literatura analizados a través del gráfico MD

Como se ha visto en el capítulo anterior, existe una gama variada de métodos alternativos para el análisis de los resultados en el diseño de productos robustos; algunos de estos métodos muestran aspectos no siempre fáciles de tratar y exigen conocimientos elevados de estadística o matemática para su comprensión y aprovechamiento cabal, por esta razón se plantea que el gráfico MD, por su sencillez y versatilidad, puede ser de gran utilidad en la industria. En esta sección se compara el método de análisis propuesto con tres casos sacados de la literatura, en los cuales se usan distintos métodos para encontrar las condiciones robustas del producto. Para cada caso al inicio se explica el análisis tal como fue desarrollado por los autores, se analiza posteriormente los datos con el método que se propone y finalmente se obtienen conclusiones.

##### 4.4.1 Moldeado por inyección

###### Planteamiento original

Steinberg y Bursztyn (1994) analizan el experimento reportado por Engels (1992) para mejorar un proceso de moldeado por inyección. La característica de interés,  $Y$ , es el porcentaje de disminución de volumen del producto obtenido, interesando que esta disminución sea estable y próxima a un valor nominal determinado. Se experimenta con 7 factores controlables (A-G) y 3 factores ruido (M-O), Tabla 4.2. El diseño en los factores controlables corresponde a un factorial  $2^{7-4}$ , el cual está cruzado con un factorial  $2^{3-1}$  en los factores ruido. El total de observaciones junto con los valores, para cada combinación de los factores controlables, de la media, la

desviación tipo y el cociente señal ruido “nominal es lo mejor” se reproducen en la Tabla 4.3; la notación *mo* en los factores ruido significa nivel alto de M y O y nivel bajo de N; la anotación (1) corresponde al tratamiento con los tres factores ruido en su nivel bajo.

**Tabla 4.2** Factores controlables y de ruido.

Factores controlables	Factores ruido
A Tiempo del ciclo	M Porcentaje de residuos
B Temperatura del molde	N Contenido de humedad
C Espesor de la cavidad	O Temperatura ambiente
D Presión	
E Velocidad de inyección	
F Tiempo de espera	
G Tamaño del acceso	

El propósito de Steinberg y Bursztyn (1994) es mostrar que se pueden dar conclusiones distintas acerca de los factores que afectan la variabilidad, al analizar el cociente señal ruido de Taguchi y la modelación explícita de los efectos de los factores ruido al usar la matriz combinada.

**Tabla 4.3** Porcentaje observado de encogimiento, junto con la media, la desviación estándar y el cociente señal ruido para cada combinación de los factores controlables.

Factores controlables							Factores ruido				$\bar{Y}$	S	SR
A	B	C	D	E	F	G	(1)	no	mo	mn			
-1	-1	-1	-1	-1	-1	-1	2,2	2,1	2,3	2,3	2,225	0,095	27,32
-1	-1	-1	1	1	1	1	2,5	0,3	2,7	0,3	1,450	1,330	0,74
-1	1	1	-1	-1	1	1	0,5	3,1	0,4	2,8	1,700	1,449	1,38
-1	1	1	1	1	-1	-1	2,0	1,9	1,8	2,0	1,925	0,095	26,06
1	-1	1	-1	1	-1	1	3,0	3,1	3,0	3,0	3,025	0,050	35,63
1	-1	1	1	-1	1	-1	2,1	4,2	1,0	3,1	2,600	1,368	5,57
1	1	-1	-1	1	1	-1	4,0	1,9	4,6	2,2	3,175	1,327	7,57
1	1	-1	1	-1	-1	1	2,0	1,9	1,9	1,8	1,900	0,081	27,33

El análisis de Engel consideró la posible existencia de una relación potencia entre la media y la desviación estándar, por lo que empleó un modelo lineal general para la desviación estándar. Las conclusiones a las que llegó consideran que el factor D tiene un gran efecto en la respuesta, así como el factor A y el E, y que sobre la variación tiene un substancial efecto el factor

F. La Figura 4.6, que representa los efectos principales del cociente SR, muestra la influencia del factor F. Un análisis de la tabla anterior nos permite llegar a la misma conclusión; los valores de la desviación estándar S, y también los del cociente SR, están claramente diferenciados en dos grupos de cuatro, cada uno de ellos asociado a uno de los niveles del factor F; en concreto, se puede concluir que cuando F está en su nivel bajo se tiene la menor variación.

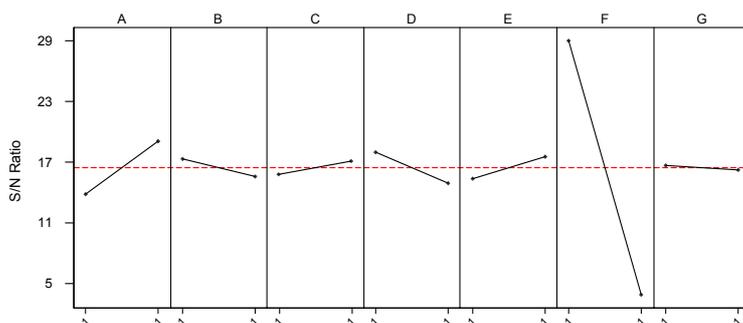


Figura 4.6 Efectos principales para el cociente SR

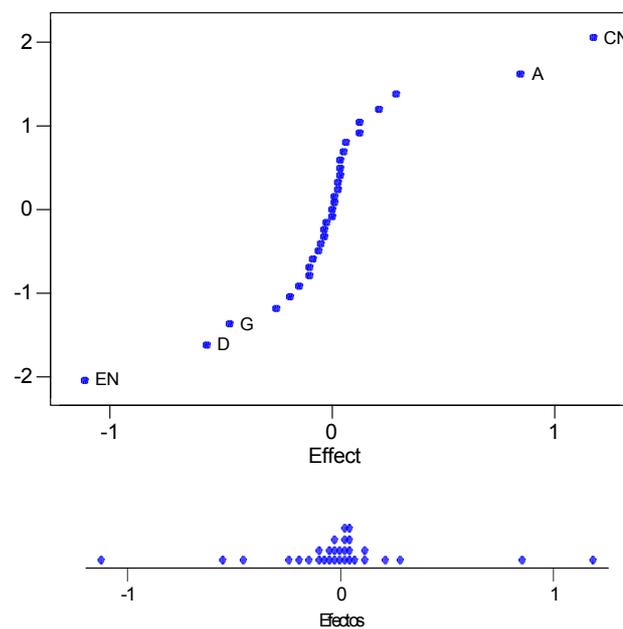
Se considera ahora el análisis de los efectos a partir de las respuestas ubicadas en la matriz combinada, consistente en un diseño factorial  $2^{10-5}$ . Tanto el arreglo interno como el externo son saturados, así que los 10 efectos principales (7 de control y 3 de ruido) y las 21 interacciones entre factores controlables y de ruido agotan todos los contrastes ortogonales en el diseño.

Tabla 4.4 Efectos principales e interacciones mayores del análisis completo del experimento del moldeado por inyección.

Efectos principales		Interacciones	
A	0.85	F	-0.04
B	-0.15	G	-0.46
C	0.13	M	-0.10
D	-0.56	N	0.00
E	0.29	O	0.03
		NC	1.18
		NE	-1.11

Se observa en la Figura 4.7 un gráfico en papel probabilístico normal y un diagrama de puntos de los contrastes para el experimento del moldeado por inyección. La Tabla 4.4 que emplean Steinberg y Bursztyn (1994) lista los efectos principales y las interacciones mayores<sup>1</sup>, tres factores de diseño tienen gran efecto principal: A, D y G y hay dos grandes interacciones: CN y EN. El factor E y N tienen un efecto principal moderado y C un efecto principal pequeño.

Un beneficio importante del análisis con arreglo combinado es que permite valorar que factores ruido afectan la variación. Ninguno de los efectos de ruido en este experimento tienen efectos principales grandes, pero N está involucrado en interacciones grandes con C y con E.

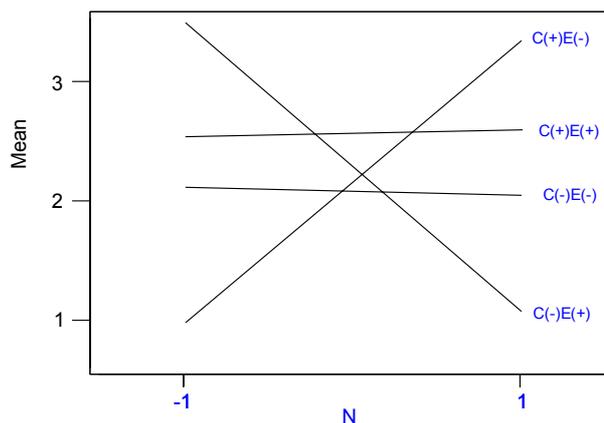


**Figura 4.7** Diagrama en papel probabilístico normal y diagrama de puntos de los 31 efectos en el control de la disminución de volumen

Las estadísticas de la Tabla 4.3 parecen identificar un claro efecto de dispersión para el factor F. ¿Por qué desaparece este efecto en el análisis del

<sup>1</sup> La tabla original tiene los siguientes valores erróneos: N=0.28, O=0.30, NC=0.90 y NE=-0.84

arreglo combinado? La Figura 4.8 muestra un diagrama de interacción conjunta, en él se observa el efecto que tiene el factor de ruido N en cada una de las combinaciones conjuntas de los niveles de los factores controlables C y E. Se aprecia las dos condiciones del proceso que son aparentemente insensibles al efecto del factor de ruido N, cuando el factor C y el E están respectivamente en su nivel bajo y cuando se encuentran ambos en el nivel alto.



**Figura 4.8** Efecto de N en cada una de las cuatro combinaciones de niveles de C y E.

La aparente contradicción entre el análisis efectuado con la matriz cruzada, en donde F(-) es la característica recomendada para tener la menor variación, y el análisis con la matriz combinada, que considera a C(+)/E(+) y a C(-)/E(-) como las condiciones robustas, se debe a que en el diseño  $2^{7-4}$  están confundidos el factor F con la interacción CE; de aquí la ilusión de un efecto de dispersión para F. El ejemplo muestra que el análisis resumen por hileras puede llevar a conclusiones erróneas. Sí, como en este ejemplo, uno de los factores ruido tiene un efecto pequeño pero interactúa con dos factores de diseño, una matriz de diseño de resolución III puede confundir totalmente los verdaderos efectos de dispersión.

### **Análisis con el gráfico MD**

El modelo usado contiene los términos que emplearon Steinberg y Bursztyn (1994) para detectar las condiciones robustas, independientemente de los criterios que siguieron con la estructura de alias para elegir estos términos. Para ellos tres factores de diseño, A, D y G, y dos interacciones, CN y EN, tienen efecto grande, con lo cual el modelo resultante es:

$$\hat{y} = 2,25 + 0,425A - 0,281D - 0,231G + 0,587CN - 0,556EN . \quad (4.17)$$

Con él se obtuvieron los 32 puntos del gráfico MD que se muestra en la Figura 4.9, usando la opción que tiene el algoritmo para mostrar únicamente los “vértices” (condiciones experimentales en la matriz de diseño) de la región de experimentación en la que se está analizando la respuesta promedio y la desviación estándar. Lo hacemos de esta manera para que sea comparable el resultado del análisis previo con el que proporciona el gráfico MD.

Las 32 condiciones experimentales conforman dos grupos bien diferenciados, uno con una desviación estándar aproximadamente de 0,32 y el otro con un valor de dispersión cercano a 1,19. En la Tabla 4.5 se aprecia la información del grupo con la variación mínima, donde el factor C y el factor E tienen al mismo tiempo o bien su nivel bajo o bien su nivel alto. Esta es la característica que Steinberg y Bursztyn (1994) mencionan como robusta en su análisis, pero no menciona cuál es el valor nominal deseado del porcentaje de disminución de volumen del producto obtenido, no obstante, por los valores en la tabla sabemos que la disminución máxima está en torno a 3,1 y la mínima cercana a 1,3. Obsérvese la ventaja que tiene el gráfico MD al permitir encontrar de forma simple las condiciones robustas y proveer al mismo tiempo diferentes opciones para el valor esperado de la respuesta, a través de los factores de ajuste.

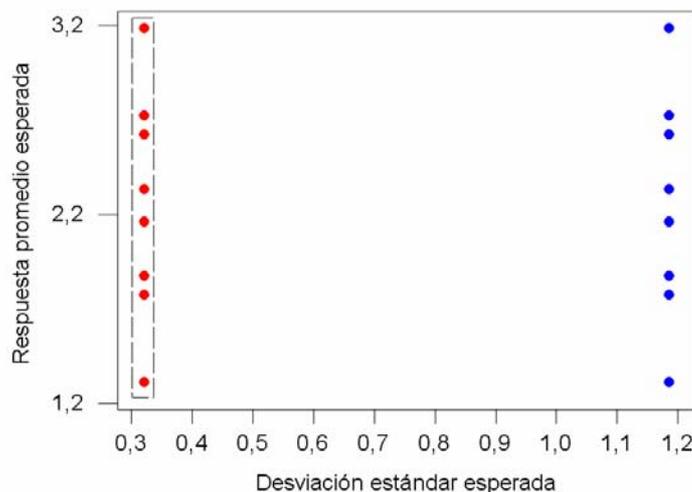


Figura 4.9 Gráfico MD para el control de la disminución de volumen, con el modelo 4.18.

Tabla 4.5 Condiciones experimentales con mínima varianza para el modelo 4.18.

Row	A	C	D	E	G	E(Y)	S(Y)
1	-1	-1	-1	-1	-1	2,337	0,31963
2	1	-1	-1	-1	-1	3,187	0,31963
5	-1	-1	1	-1	-1	1,775	0,31963
6	1	-1	1	-1	-1	2,625	0,31963
11	-1	1	-1	1	-1	2,337	0,31963
12	1	1	-1	1	-1	3,187	0,31963
15	-1	1	1	1	-1	1,775	0,31963
16	1	1	1	1	-1	2,625	0,31963
17	-1	-1	1	-1	1	1,875	0,31963
18	1	-1	-1	-1	1	2,725	0,31963
21	-1	-1	1	-1	1	1,313	0,31963
22	1	-1	1	-1	1	2,163	0,31963
27	-1	1	-1	1	1	1,875	0,31963
28	1	1	-1	1	1	2,725	0,31963
31	-1	1	1	1	1	1,313	0,31963
32	1	1	1	1	1	2,163	0,31963

### 4.4.2 Tasa de filtración

#### Planteamiento original

Myers, Kim y Griffiths (1997) retoman el caso tratado por Montgomery el mismo año, acerca de un experimento factorial en una planta a nivel piloto para estudiar los factores que influyen la tasa de filtración de una sustancia química adherente. Se variaron cuatro factores: presión,  $X_1$ , concentración

de formaldehído,  $X_2$ , tasa de agitado,  $X_3$ , y temperatura,  $Z$ . Los ingenieros de proceso están interesados en maximizar la tasa de filtración. Sin embargo, cuando el proceso esté en fase de producción la temperatura será difícil de controlar, por lo que se considera factor de ruido. Todos los factores variaron en niveles  $\pm 1$  en un arreglo factorial  $2^4$ . Los niveles  $\pm 1$  asociados a la temperatura se asume que son  $\pm\sigma_Z$  representando la variabilidad de la temperatura en el proceso. El diseño del experimento se basa en el arreglo combinado de la Tabla 4.6, donde se muestra también los valores de la tasa de filtración.

**Tabla 4.6** Experimento para la tasa de filtración.

$X_1$	Factor			Tasa de filtración
	$X_2$	$X_3$	$Z$	
-1	-1	-1	-1	45
-1	-1	-1	1	71
1	-1	-1	-1	48
1	-1	-1	1	65
-1	1	-1	-1	68
-1	1	-1	1	60
1	1	-1	-1	80
1	1	-1	1	65
-1	-1	1	-1	43
-1	-1	1	1	100
1	-1	1	-1	45
1	-1	1	1	104
-1	1	1	-1	75
-1	1	1	1	86
1	1	1	-1	70
1	1	1	1	96

Basados en el procedimiento que sugiere Vining y Myers (1990) de modelar simultáneamente las superficies de respuesta de la media y la varianza, Myers et al. (1997) proponen restringir la búsqueda de las condiciones robustas a una región de confianza de la zona que tenga la mínima varianza del proceso. Para encontrar esta zona se basan en una expresión simplificada del modelo de la varianza de la respuesta de la ecuación 4.16, donde  $\mathbf{\Lambda} = \mathbf{0}$  y  $\mathbf{l}(\mathbf{x}) = (\boldsymbol{\gamma} + \mathbf{\Lambda}\mathbf{x})$ , de esta forma la varianza estimada del proceso generada por las variables de ruido es:

$$\hat{V}(y) = \hat{\mathbf{l}}'(\mathbf{x})\hat{\mathbf{l}}(\mathbf{x}) + \hat{\sigma}_\varepsilon^2. \quad (4.18)$$

Su valor mínimo se obtiene cuando  $\hat{\mathbf{I}}(\mathbf{x}) = \mathbf{0}$ , al resolver este conjunto de ecuaciones se genera la zona con varianza mínima. Para esta zona en  $\mathbf{x}_0$  es factible construir una región de confianza del tipo:

$$\Pr \left\{ \frac{\hat{\mathbf{I}}'(\mathbf{x}_0) [\mathbf{C}(\mathbf{x}_0)]^{-1} \hat{\mathbf{I}}(\mathbf{x}_0)}{n_z \hat{\sigma}_\varepsilon^2} \leq F_{\alpha, n_z, g.l.E} \right\} = 1 - \alpha, \quad (4.19)$$

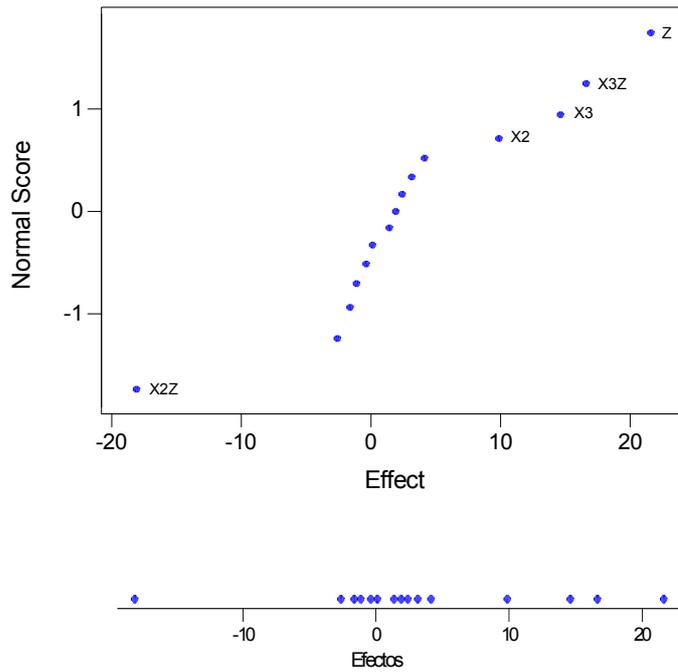
donde  $\hat{V}(\hat{\mathbf{I}}(\mathbf{x})) = \mathbf{C}(\mathbf{x}) \hat{\sigma}_\varepsilon^2$  y  $F_{\alpha, n_z, g.l.E}$  es el percentil  $(1-\alpha) \times 100$  de la distribución F.

Con los datos de la tasa de filtración se ha construido el gráfico en papel probabilístico normal y el diagrama de puntos de los efectos, Figura 4.10, en él se percibe que la presión no impacta la tasa de filtración y que el modelo de la respuesta puede ser:

$$\hat{y} = 70,06 + 4,94x_2 + 7,31x_3 + 10,81z - 9,06x_2z + 8,31x_3z. \quad (4.20)$$

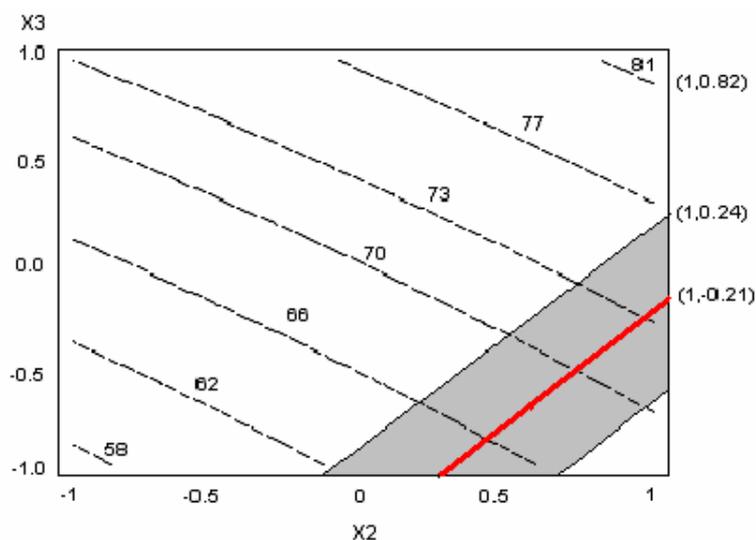
A partir de esta ecuación se deduce que la zona de mínima varianza corresponde a la línea:

$$\hat{\mathbf{I}}(\mathbf{x}) = \frac{\delta \hat{y}}{\delta z} = 10,81 - 9,06x_2 + 8,31x_3 = 0. \quad (4.21)$$



**Figura 4.10** Representación de los efectos en papel probabilístico normal y en diagrama de puntos.

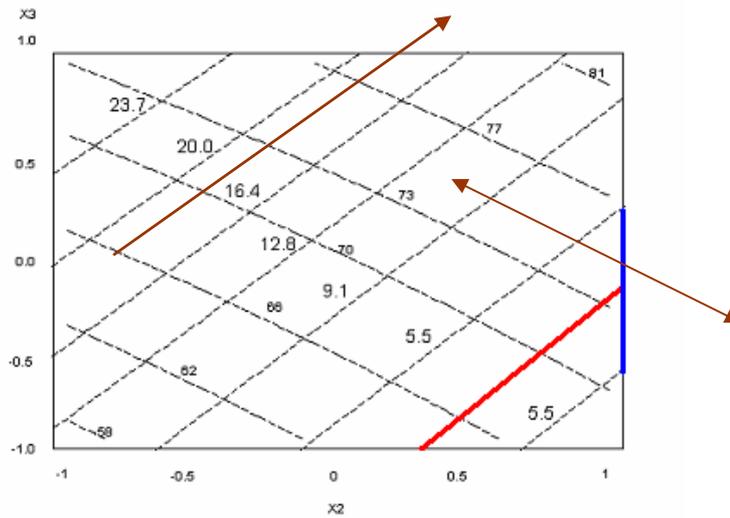
La Figura 4.11 es un diagrama de contorno para la tasa media de filtración en la región de experimentación, en ella se representa la línea de mínima varianza del proceso (línea continua) con su región de confianza al 95%, en gris. Myers et al. (1997) consideran que, para contar con la mayor tasa de filtración y la menor variabilidad, se debe operar el proceso con el nivel alto de  $x_2$  y con alguna flexibilidad alrededor de  $x_3 = 0$ . La localización más adecuada en la recta de mínima varianza se da cuando  $x_2 = 1$  y  $x_3 = -0,21$ , en cuyo caso la media estimada de la tasa de filtración es de 73,58 gal/hr.



**Figura 4.11** Diagrama de contorno para la tasa media de filtración en el cual se incluye la recta de mínima varianza del proceso con su región de confianza al 95%.

El uso simultaneo de las superficies de respuesta para la media y la varianza, a través de superponer los diagramas de contorno correspondientes a la media y la desviación estándar en la Figura 4.12, sugiere fijar los valores de operación en la región  $x_2 = 1$  y  $-0.4 \leq x_3 \leq 0,24$ , con lo cual se tiene una confianza del 95% de que los valores de la tasa de filtración del proceso excederán los 60 gal/hr.

Se aprecia en la figura que el valor de la respuesta media crece, en la región de experimentación, del vértice  $(-1; -1; 60)$  hacia el vértice  $(1; 1; 82,3)$ , en cambio la desviación estándar decrece del vértice  $(-1; 1; 28,5)$  hacia el vértice  $(1; -1; 7,9)$ , sólo que en este último caso se alcanza el mínimo dentro de la región de interés y antes de alcanzar dicho punto.



**Figura 4.12** Diagrama de contornos superpuestos de la tasa de filtración media y de la desviación estándar del proceso. La flecha ascendente señala la tendencia de la respuesta media y la otra la tendencia de la variabilidad.

### Análisis con el gráfico MD

En el análisis comparativo se usará el mismo modelo que emplearon Meyers et al. (1997), ecuación 4.20. Como el estudio que realizaron emplea en la búsqueda superficies de respuesta a través de diagramas de contorno, y no únicamente los valores de la media y la varianza en los puntos de la matriz de diseño, se elaborarán los gráficos MD necesarios para un entramado de puntos en la región de experimentación, que abarque tanto los vértices como la región interior. Note en el gráfico siguiente que la mínima dispersión y la máxima tasa de filtración no pueden tenerse al mismo tiempo.

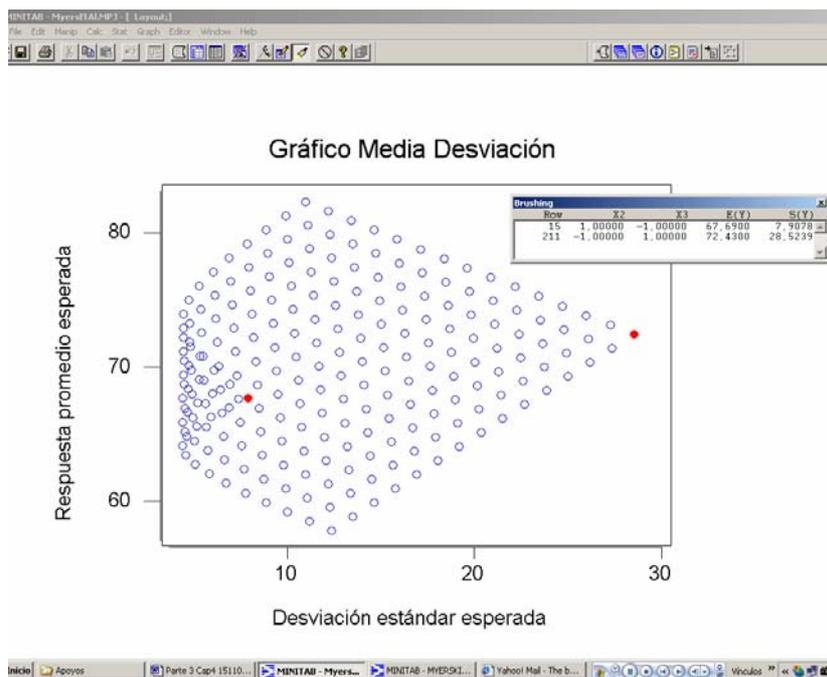
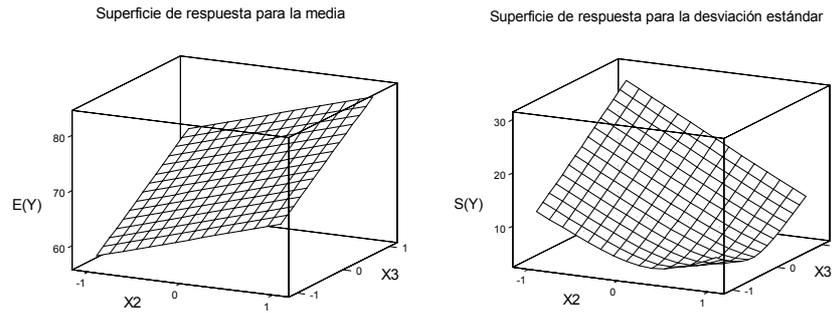


Figura 4. 13 Gráfico MD de la tasa de filtración.

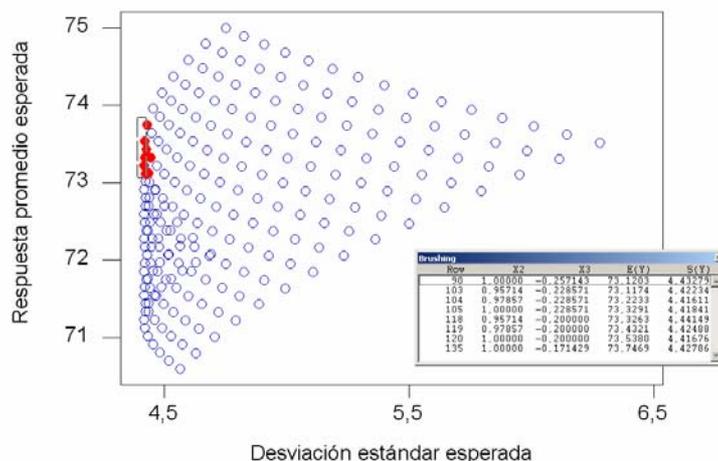
El gráfico MD, al representar el comportamiento conjunto de la media y la desviación estándar en la región de interés, permite apreciar como la desviación va disminuyendo a partir del punto ( $X_2 = -1, X_3 = 1$ ), punto sólido a la derecha del diagrama, fila 211 en la matriz de datos (ver recuadro en el gráfico), con un valor de  $S=28,52$ , hasta alcanzar su mínimo y después replegarse hacia valores mayores, punto sólido a la izquierda con coordenadas ( $X_2 = 1, X_3 = -1$ ), fila 15 en la matriz de datos, con un valor de  $S=7,9$ . Se observa también que la media tiene una distribución aproximadamente simétrica alrededor de la media de los resultados del experimento,  $\bar{y} = 70,06$ .



**Figura 4.14** Superficie de respuesta para la media y para la desviación estándar.

Al tener únicamente dos factores controlables activos podemos comparar el comportamiento que nos muestra el gráfico anterior con los gráficos de superficie de respuesta para la media y para la desviación estándar de la Figura 4.14. La curvatura que se aprecia en la superficie de respuesta de la desviación estándar se percibe claramente en el gráfico MD, e indudablemente, es más simple identificar los valores de los factores controlables que dan al mismo tiempo una respuesta promedio y una desviación estándar específica en el gráfico MD que en las superficies de respuesta asociadas.

La búsqueda de las condiciones robustas se llevó a cabo usando dos gráficos MD, el primero que corresponde a la región de experimentación  $-1 \leq x_2, x_3 \leq 1$  es el que se muestra en la Figura 4.13, en el se enmarcó los puntos que tienen la mayor tasa de filtración restringidos a que cuenten con la mínima desviación estándar, ver Figura 4.15, de esta forma sabemos por su recuadro que la zona enmarcada concierne a valores  $0,71 \leq x_2 \leq 1$  y  $-0,43 \leq x_3 \leq -0,14$ . Esta zona se analiza a mayor detalle en la Figura 4.16. Este tipo de análisis secuencial a regiones cada vez más específicas es una de las grandes ventajas del gráfico.



**Figura 4.15** Condiciones enmarcadas que tienen la mayor tasa de filtración restringidos a que cuenten con la mínima desviación estándar.

Meyers et al. (1997) encuentran que la localización más adecuada en la recta de mínima varianza se da cuando  $x_2 = 1$  y  $x_3 = -0,21$ , en cuyo caso la media estimada de la tasa de filtración es de 73,58 gal/hr. Esos valores aparecen aproximadamente en la fila 120 del recuadro. El valor de la desviación estándar del error es de 4,415 por esta razón las condiciones señaladas del recuadro son ideales por anular prácticamente toda influencia del factor ruido.

El análisis gráfico presentado por estos autores a través del uso de diagrama de contornos superpuestos, de la tasa de filtración media y de la desviación estándar del proceso, solamente es posible en la medida que se tenga dos factores controlables activos, en cambio el gráfico MD siempre se puede elaborar, independientemente del número de factores activos. Estos gráficos tienen la gran ventaja de explorar con suma facilidad las combinaciones de los factores controlables que nos proporcionan una composición aceptable en los valores de la media y la dispersión.

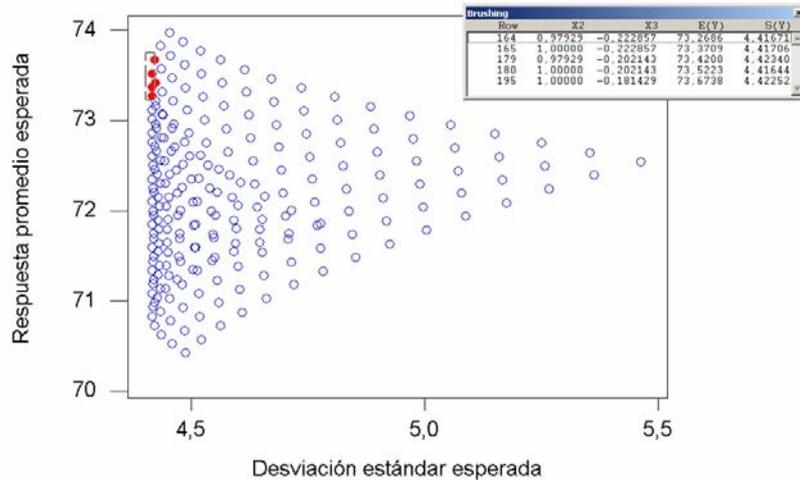


Figura 4.16 Condiciones robustas con mayor tasa de filtración.

#### 4.4.3 Resistencia al desgaste

##### Planteamiento original

Pozueta (2001) trata el caso de un diseño factorial en un proceso de fabricación de neumáticos para coche. La característica de interés  $Y$  es la resistencia al desgaste, que obviamente se desea que sea lo más grande posible. Se tienen tres factores controlables asociados al proceso (A, B, C) y dos factores (R, T) de ruido que aparecen en el momento que el usuario utiliza el coche. El diseño original es una matriz cruzada  $2^3 \times 2^2$ , los datos obtenidos junto con algunos estadísticos calculados en cada combinación de niveles en los factores controlables se presenta en la Tabla 4.7. La autora analiza este problema para identificar las causas por la que se llega a conclusiones diferentes, acerca de las condiciones robustas, dependiendo del método de análisis empleado. Al análisis que se basa en los estadísticos asociados con medidas de la variabilidad (S/R, S y  $\ln(S)$ ) le llama análisis con métricas resumen, para diferenciarlo con el análisis que emplea

modelos de superficie de respuesta, a los cuales denota como análisis con métricas no resumen.

**Tabla 4.7** Resultados de la resistencia al desgaste y estadísticos calculados en las condiciones de los factores ruido.

A	B	C	R				$\bar{y}$	S/R	S	ln(S)	
			T	-	+	-					+
-1	-1	-1		196	116	233	175	180	44,22	48,94	3,89
1	-1	-1		128	102	197	193	155	42,80	47,42	3,86
-1	1	-1		243	215	293	229	245	47,61	33,98	3,53
1	1	-1		191	148	183	138	165	44,11	25,94	3,26
-1	-1	1		165	72	214	89	135	40,25	66,40	4,20
1	-1	1		101	59	185	175	130	39,52	60,37	4,10
-1	1	1		267	167	369	197	250	46,83	89,72	4,50
1	1	1		144	104	219	233	175	43,47	61,38	4,12

En la exploración que lleva a cabo con métricas resumen utiliza el arreglo interno de la matriz cruzada y como respuesta a S/R, S y ln(S) respectivamente. Los gráficos en papel probabilístico normal y en diagrama de puntos de la Figura 4.17 sirven de base, junto con diagramas de efectos y de interacciones, para concluir que la condición que maximiza el cociente S/R es B(+) y la que minimiza S y ln(S) es B(+)C(-).

El análisis con métricas no resumen lo lleva a cabo con la matriz combinada correspondiente  $2^5$ . El modelo de la respuesta se determina a partir de los efectos que se consideran significativos al representarlos en papel probabilístico normal, ver Figura 4.18.

$$\hat{y} = 179,38 - 23,13A + 29,38B - 15,63AB - 28,62R + 28,25T + 16,38AR. \quad (4.22)$$

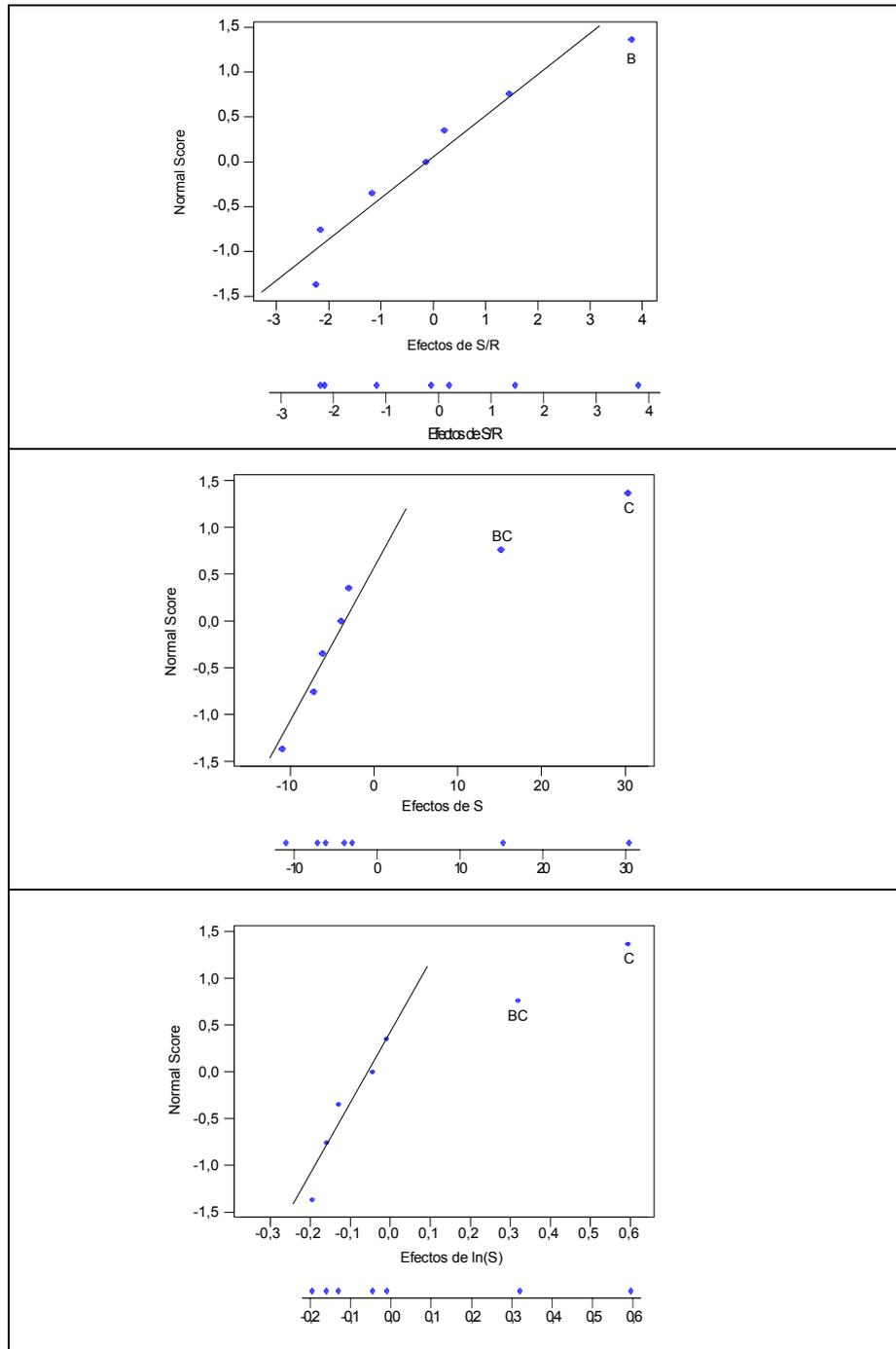
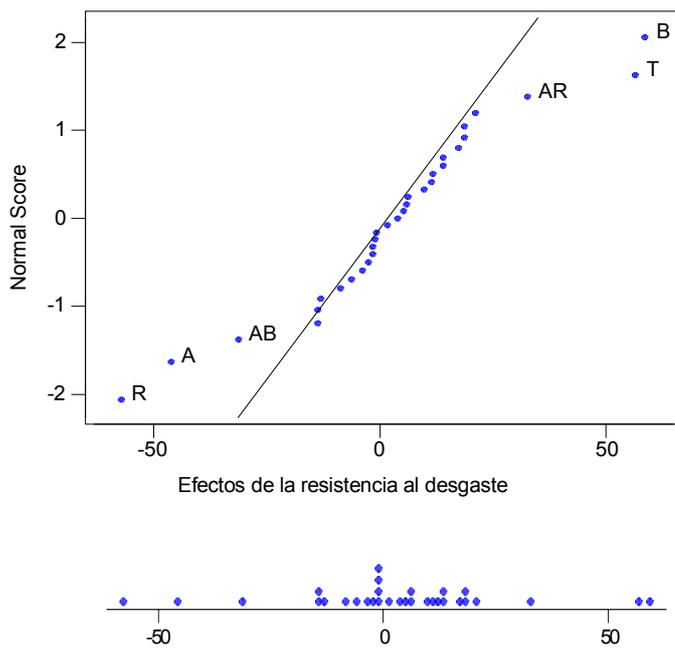


Figura 4.17 Gráficos de los efectos de S/R, S y ln(S)

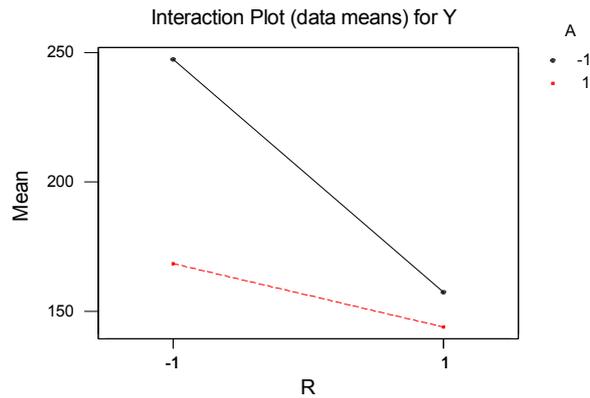


**Figura 4.18** Gráficos en papel probabilístico normal y en diagrama de puntos de los efectos de la resistencia al desgaste

El modelo únicamente contiene una interacción, AR, entre los factores ruido y los de control, esto nos indica que podemos esperar una condición robusta únicamente ante la presencia del factor de ruido R, la cual se da al ubicar el factor controlable A en su nivel alto, como se aprecia en la figura siguiente.

Una alternativa de análisis, que también explora Pozueta, consiste en tener un modelo estimado de la varianza y encontrar los valores que la minimizan en la región de experimentación. El modelo que deduce de la ecuación (4.24) es:

$$\hat{\sigma}^2(y) = 1885,47 - 937,59A + \hat{\sigma}_\epsilon^2. \tag{4.23}$$



**Figura 4.19** Efecto del factor de ruido R en presencia del factor controlable A.

Es claro que esta expresión se minimiza cuando A toma su nivel alto, lo cual es consistente con el análisis de la interacción AR. La discrepancia que existe entre este resultado y el realizado con métricas resumen (el cociente S/R da B(+) y la desviación estándar y su logaritmo natural dan B(+)C(-)) se debe, según concluye Pozueta, a que no es una buena estrategia incluir en el modelo para Y solo los efectos significativos a modo individual. Considera necesario hacer el análisis en un modelo ampliado, que se determina a partir de los efectos que surgen significativos en la regresión paso a paso correspondiente. Con esto se consigue incluir términos relevantes en la búsqueda de las condiciones robustas, en los que aparecen interacciones de los factores ruido involucrando a los factores B y C que surgen en las métricas resumen. El modelo ampliado estimado es:

$$\hat{y} = 179,38 - 23,13A + 29,38B - 15,63AB + 10,36BC - 28,62R + 28,25T + 16,38AR + 9,38CT + 9,37ACR + 8,75BCT. \quad (4.24)$$

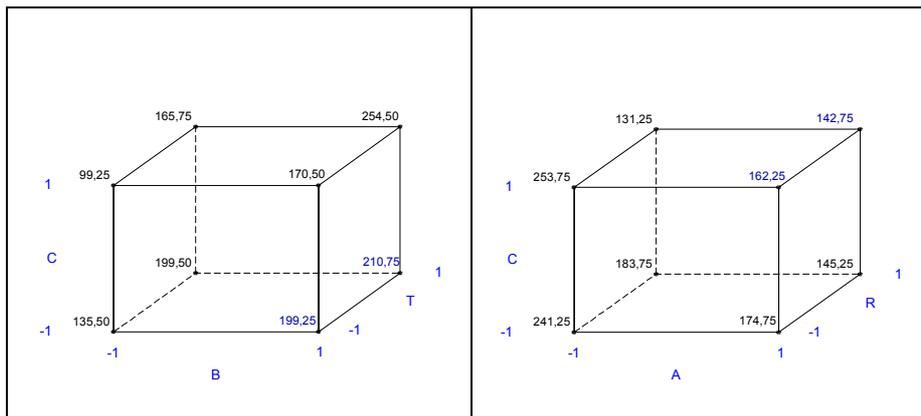
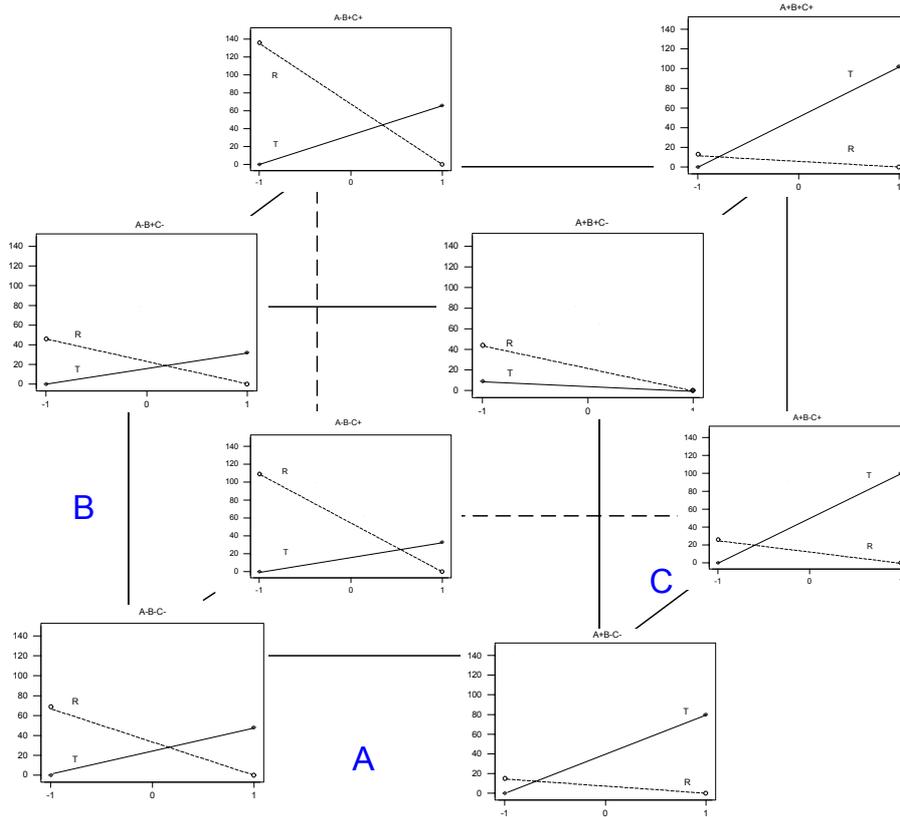


Figura 4.20 Diagramas de cubo de las interacciones ACR y BCT.

El análisis gráfico de las interacciones ACR y BCT permite concluir que la condición robusta al efecto T se da con B(+)*C*(-) en tanto que la condición A(+)*C*(+) es robusta al factor R, ver Figura 4.20, ambas condiciones de robustez son incompatibles. Como las interacciones ACR y BCT comparten el factor C, es conveniente explorar el efecto que tienen los factores ruido en cada una de las condiciones experimentales del arreglo interno. La Figura 4.21 muestra que la condición más insensible a la presencia de los efectos de los factores ruido es A(+)*B*(+)*C*(-). Esta condición une los resultados que se deducen de usar las métricas resumen *S* o ln(*S*) y el modelo de la varianza.

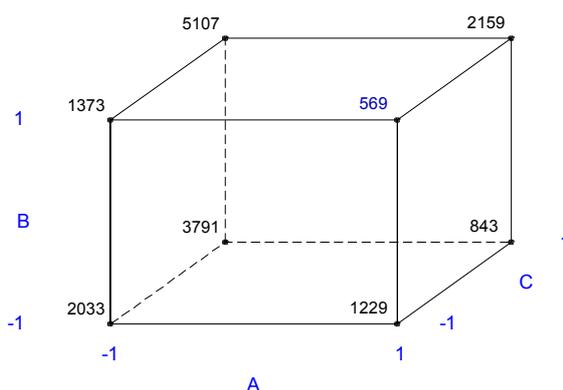


**Figura 4.21** Efecto de los factores ruido en cada una de las condiciones experimentales del arreglo interno.

El modelo estimado de la varianza, a partir del modelo ampliado 4.21, es:

$$\hat{\sigma}^2(y) = 2137,82 - 937,59A - 536,34AC + 494,37BC + 164,15B + \hat{\sigma}_\varepsilon^2. \quad (4.25)$$

Con el se puede evaluar la varianza transmitida por los factores ruido en cada una de las condiciones experimentales de los factores controlables. La Figura 4.22 presenta que la mínima variabilidad se consigue con A(+ )B(+ )C(- ) precisamente.



**Figura 4.22** Varianza estimada, a través del modelo ampliado, que es transmitida a Y por los factores ruido T y R.

### Análisis con el gráfico MD

El gráfico MD de la Figura 4.23 se construyó con el modelo ampliado de la ecuación 4.26. En el recuadro se aprecia que la menor variabilidad inducida por los factores ruido se tiene justamente con las condiciones que Pozueta encuentra como robustas: A(+) $B(+)$ C(-).

El objetivo de encontrar al mismo tiempo la máxima resistencia al desgaste de los neumáticos y la mínima variabilidad es incompatible, en realidad cuando se alcanza la menor desviación estándar se tiene una resistencia promedio muy baja.

La capacidad que ofrece de visualizar el comportamiento conjunto de la respuesta promedio y la variabilidad, inducida por los factores ruido, permite apreciar aspectos que escapan con la mayoría de los distintos métodos de solución al problema del diseño de parámetros. Como la mínima variabilidad no se encontró en el interior de la región de experimentación sino en uno de sus vértices, A(+) $B(+)$ C(-); esto indica que quizá se pueda reducir todavía más la variabilidad fuera de la región estudiada.

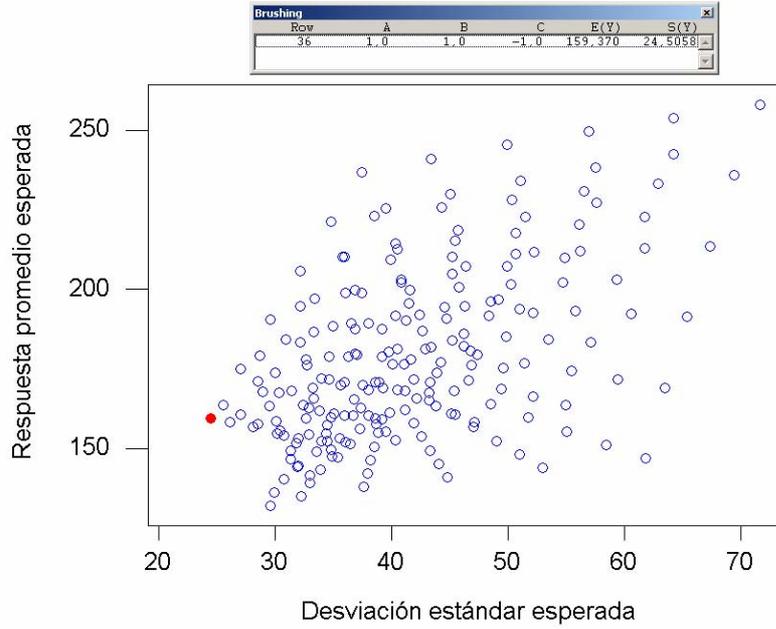


Figura 4.23 Gráfico MD para la resistencia al desgaste.

Este ejemplo ilustra claramente la diversidad de resultados a que se llega dependiendo del método utilizado, Tabla 4.8.

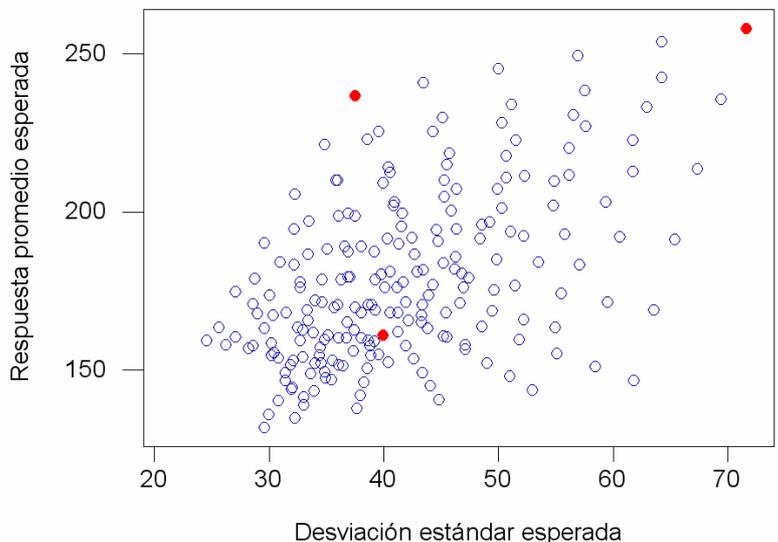
Tabla 4.8 Resultados de los diversos análisis a la resistencia al desgaste.

Diseño	Método	Resultado
Matriz cruzada	max S/R	B(+)
	min S	B(+)/C(-)
Matriz combinada	Modelo reducido	
	Media	A(+)
	Varianza	A(+)
	Modelo ampliado	
	Media	A(+)/B(+)/C(-)
	Varianza	A(+)/B(+)/C(-)
	gMD	A(+)/B(+)/C(-)

Los análisis con la matriz cruzada no son satisfactorios; si se considera el resultado que se tiene al maximizar el cociente S/R se debe usar el nivel alto

de B, no obstante con esta condición se puede tener la máxima variabilidad si seleccionamos valores inadecuados de los otros factores, por ejemplo para la combinación A(-)B(+)C(+) se da esta situación, con una desviación tipo estimada de 71.68 (punto sólido de la parte superior derecha de la Figura 4.25, ver recuadro para identificar las condiciones y las coordenadas).

Row	A	B	C	E(Y)	S(Y)
31	-1,0	1,0	-1,0	236,870	37,4572
204	1,0	0,2	1,0	161,126	39,8767
211	-1,0	1,0	1,0	258,130	71,6822



**Figura 4.24** Los puntos rojos tienen una variabilidad inadecuada y sin embargo satisfacen alguno de los resultados de la tabla 4.10.

Aunque al minimizar la desviación estándar se tiene una condición más específica, B(+)C(-), aun así se puede tener una variabilidad inadecuada (punto sólido en la parte superior izquierda del gráfico). Lo mismo puede decirse del análisis con la matriz combinada si no se selecciona acertadamente el modelo. Con el modelo reducido la condición A(+)

aparentemente es robusta, la fila 204 tiene este valor alto de A y no lo es; su punto asociado se encuentra también en la Figura 4.24.

Como sería de esperar, cuando se usa el modelo ampliado, tanto el análisis con el modelo de la respuesta como el realizado con el modelo de la varianza o el que emplea el gráfico MD llegan al mismo resultado. Empero, es indudable que los procedimientos de análisis que usó Pozueta para llegar a estos resultados son más laboriosos y exigen un nivel mayor de conocimientos si se compara con el método del gráfico MD.

#### **4.5 Análisis con más de una respuesta**

En los últimos años ha tenido un gran auge las investigaciones en torno al problema del diseño de parámetros, no obstante es poco lo que se puede encontrar orientado al análisis con más de una respuesta (ver Ames et al. (1997), Vinning (1998), Chiao y Hamada (2001) y Miró y del Castillo (2002)). Tomando en cuenta que muchos de los productos poseen más de una característica de calidad esta situación resulta extraña. Si bien es cierto que los casos con más de una respuesta de interés suelen analizarse en forma independiente para cada respuesta individual, el proceder de esta manera es la más de las veces insatisfactorio, porque las condiciones apropiadas de operación para las distintas respuestas rara vez coinciden entre sí, en particular cuando están correlacionadas. Ya Daniel (1959, página 340) se lamentaba de esta manera de actuar “Me encuentro usando métodos uno-a-la-vez en las respuestas, siendo que he eliminado el enfoque de uno-a-la-vez para los factores que influyen en la respuesta”

La posibilidad que nos da el gráfico MD de apreciar el comportamiento conjunto del promedio y la desviación estándar de la respuesta, y la facilidad para encontrar las condiciones adecuadas de operación, se pueden extender al análisis simultáneo de más de una respuesta. Con ello es posible encontrar valores que al mismo tiempo mejoren la calidad en todas las respuestas y que satisfagan distintos niveles de exigencia que se presentan en ellas. Por ejemplo, en la producción de una sustancia química puede ser más importante la densidad que la tasa de filtración.

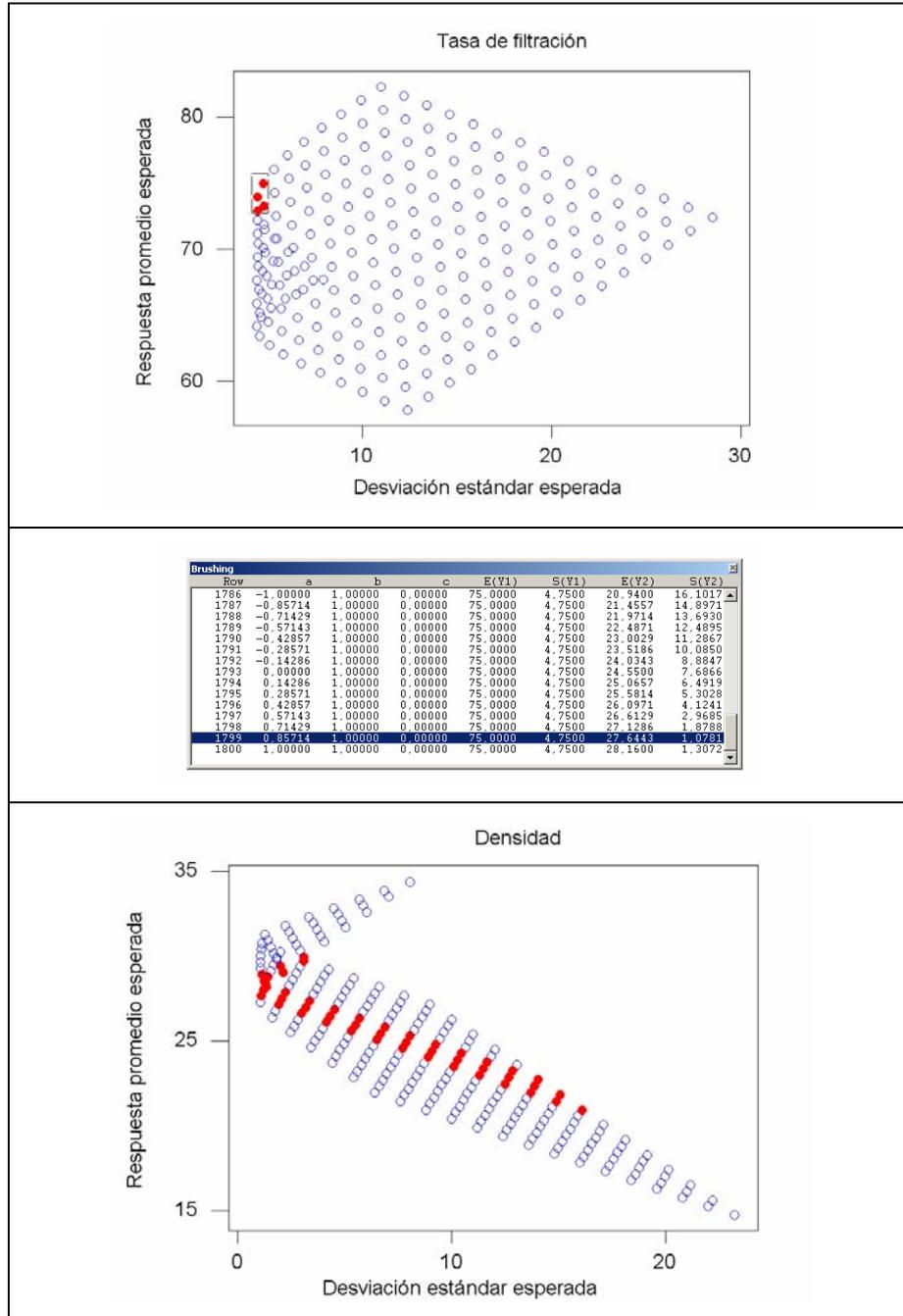
**Tabla 4. 9** Tasa de filtración y densidad de una sustancia química adherente

Z	Factor			Tasa de filtración	Densidad
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>		
-1	-1	-1	-1	45	35,9
1	-1	-1	-1	71	17,0
-1	1	-1	-1	48	26,1
1	1	-1	-1	65	41,4
-1	-1	1	-1	68	36,1
1	-1	1	-1	60	18,9
-1	1	1	-1	80	26,9
1	1	1	-1	65	43,9
-1	-1	-1	1	43	37,7
1	-1	-1	1	100	-7,8
-1	1	-1	1	45	27,8
1	1	-1	1	104	14,4
-1	-1	1	1	75	38,4
1	-1	1	1	86	-8,6
-1	1	1	1	70	28,6
1	1	1	1	96	16,2

La tabla anterior contiene los datos para el problema tratado por Myers et al. (1997) de la tasa de filtración de una sustancia adherente, se le ha añadido como segunda respuesta la densidad, que se pretende que tome un valor cercano a 27 gr/cm<sup>3</sup>, y cuyo modelo ajustado es:

$$\hat{y}_2 = 24,55 + 3,61x_1 - 6,21x_3 - 7,62z + 8,45x_1z - 7,15x_3z. \quad (4.26)$$

Los modelos estimados de las dos respuestas (ecuaciones 4.20 y 4.26) ya evidencian la imposibilidad de tener las mismas condiciones óptimas si se analizan por separado las respuestas. El modelo de la tasa de filtración no contiene al factor controlable presión, X<sub>1</sub>, en cambio el modelo de la densidad excluye la concentración de formaldehído, X<sub>2</sub>.



**Figura 4.25** Efectos en la densidad al buscar las mejores condiciones en la tasa de filtración. Se busca la máxima tasa de filtración con la mínima variabilidad.

Las Figuras 4.25 y 4.26 ilustran la posibilidad de explorar simultáneamente el comportamiento de la media y la desviación estándar de dos respuestas, en este caso la tasa de filtración y la densidad. Cuando se crean, a partir de la misma hoja de trabajo de Minitab, dos diagramas de dispersión y se marca un punto en uno de ellos, entonces automáticamente se marca los puntos en el segundo diagrama que usan las filas de la hoja de trabajo que generaron el punto marcado en el primer diagrama.

Esto se observa en la Figura 4.25, cuando en el gráfico de la Tasa de filtración se enmarcó los puntos que tiene la variabilidad mínima de la tasa de filtración y valores adecuados de la respuesta, en sólido, automáticamente se marcaron los puntos en el gráfico de la Densidad. En el recuadro se tiene la información relevante a estos puntos (valores de los factores y respuesta promedio y desviación estándar de la tasa de filtración y de la densidad), note que se marcaron cuatro puntos y el recuadro contiene más, la razón de esto es que para cada punto en el gráfico MD (pareja de valores de la respuesta promedio y su desviación estándar) puede existir más de una combinación de los factores controlables que produzcan esa respuesta promedio y esa desviación estándar. Las combinaciones de valores de los factores controlables en el recuadro generan los puntos correspondientes en el gráfico de la densidad. La fila 1799 enseña como es posible tener las mejores condiciones de robustez simultáneamente,  $S(y_1) = 4.75$  y  $S(y_2) = 1.08$ , con una tasa de filtración promedio de 75 y una densidad promedio de 27,64.

Si la densidad es más importante que la tasa de filtración bastará con explorar en su gráfico, tratando de conseguir un compromiso con el objetivo de la otra respuesta. La Figura 4.26 muestra el efecto que ocurre en la tasa de filtración al buscar las mejores condiciones en la densidad. Este modo de proceder ciertamente se puede extender a más de dos respuestas, aportando una clara visión del comportamiento de las distintas respuestas y del efecto que se ocasionan entre si.

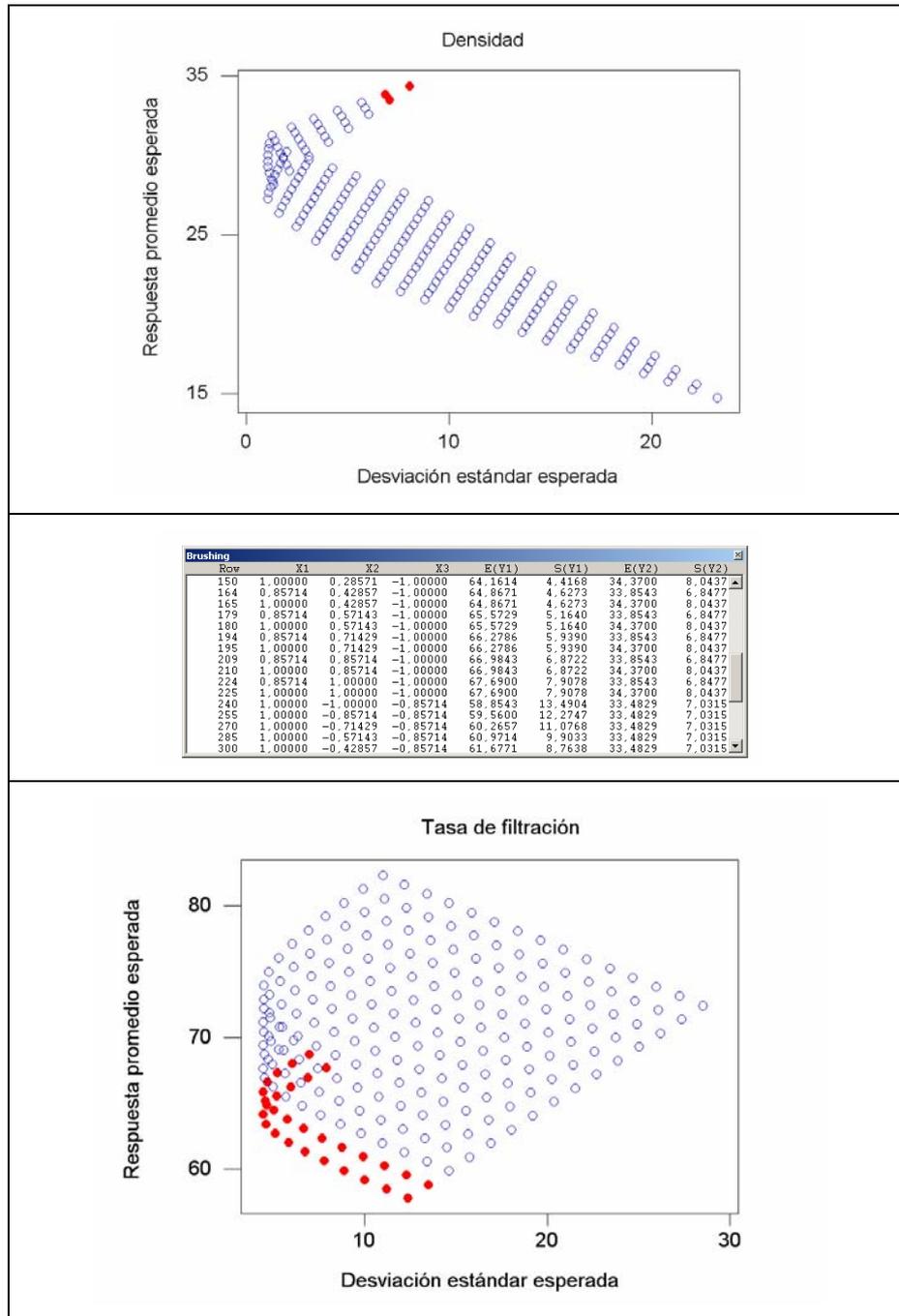


Figura 4.26 Efectos en la tasa de filtración al buscar las mejores condiciones en la densidad. Se pretende la mayor densidad.

#### 4.6 Análisis con criterios económicos

Se puede encontrar en la literatura relacionada con los diseños robustos una preocupación por encontrar métodos de experimentación económicos. Precisamente una de las críticas más recurrentes al método de Taguchi es el excesivo número de pruebas que genera la matriz de diseño y por contraste la ventaja que frecuentemente tienen en este sentido los diseños combinados o los diseños óptimos. No obstante, la búsqueda que hace Taguchi de las condiciones de proceso que sean robustas a los factores ruido se sustentan en un criterio económico, ya que minimizan una función de pérdida esperada causada por los factores ruido. Esta inquietud por considerar el impacto que tienen las variables de ruido va más allá de la fase de experimentación y tiene una sensible connotación social. Aunque frecuentemente en la práctica el uso del cociente señal ruido hace olvidar su relación con la función de pérdida.

Es de sobra conocido que la metodología de Taguchi ha contribuido a desplazar la atención a la disminución de la variación de los productos y procesos, y ya sea con el uso de ella o con el empleo de un método alternativo, lo que prima es reducir al máximo la variación. Sin embargo, consideramos que pueden existir razones prácticas o económicas que hacen inviable las condiciones de operación que proporcionan la variación mínima, y que puede ser de gran utilidad explorar condiciones cercanas a la óptima que sean factibles. En este sentido se implementó una variante al gráfico MD que diferencia los puntos con un criterio de rangos de costes. Para cada variable controlable se debe indicar el coste de operación en el nivel bajo y el coste en el nivel alto, se asume que el coste tiene un comportamiento lineal pero su modificación a otro tipo de tendencia o costes en puntos intermedios es relativamente simple.

Retomando el caso de la resistencia al desgaste, la Figura 4.27 ilustra la manera en que se introducen los costes asociados a los niveles bajo y alto de los factores controlables; estos costes son ficticios y sirven de ejemplo.

	C1	C2-T	C3	C4	C5	C6	C7	C8	C9
	coef	ter	A	B	C	cA	cB	cC	
1	-23,13	A	-1	-1	-1	5	3	6	
2	29,38	B	1	1	1	8	9	6	
3	-15,63	AB							
4	10,63	BC							
5	-28,62	S							
6	28,25	T							
7	16,38	AS							
8	9,38	CT							
9									

Figura 4.27 Datos del modelo, en las columnas C7, C8 y C9 se indican los costes asociados al nivel bajo y al nivel alto.

En el gráfico MD de la Figura 4.28 se puede observar que las seis condiciones señaladas del proceso, que nos da la mínima desviación estándar, tienen todas distintos costes de operación. Se aprecia en general que, para una variabilidad específica, se tiene un mayor coste de operación a medida que se incrementa la resistencia al desgaste.

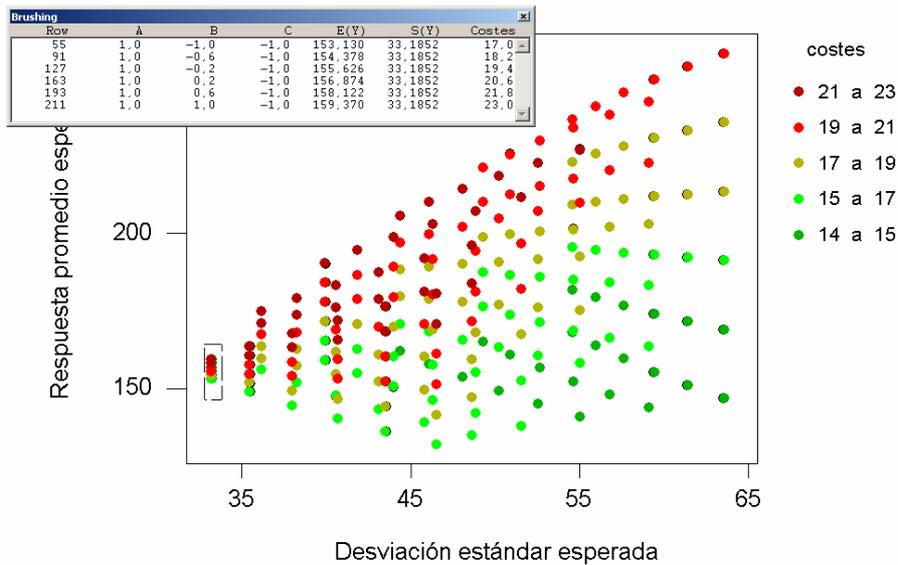


Figura 4.28 Gráfico DM con costes de operación asociados.

Uno puede pensar que dos grupos de condiciones de operación que den valores semejantes en la media y en la desviación (puntos muy cercanos en el gráfico MD) tienen costes de operación casi iguales, y eso no es siempre así. Las figuras 4.27 y 4.28 se elaboraron para el mismo modelo de la respuesta y en la misma región de experimentación pero con diferentes grupos de costes y muestran dos situaciones completamente diferentes. En la primera de ellas se tiene un patrón más o menos general de resistencia alta con costes elevados, para una variabilidad específica, y en donde dos puntos contiguos, la más de las veces, tienen costes parecidos. En la segunda todo esto cambia, no se tiene un comportamiento general en cuanto a costes y dos puntos muy próximos pueden tener valores de coste totalmente opuestos. Los tres puntos enmarcados cuentan con valores muy próximos en cuanto al valor promedio y la desviación tipo, no así en costes, que van de 16,6 a 20,6.

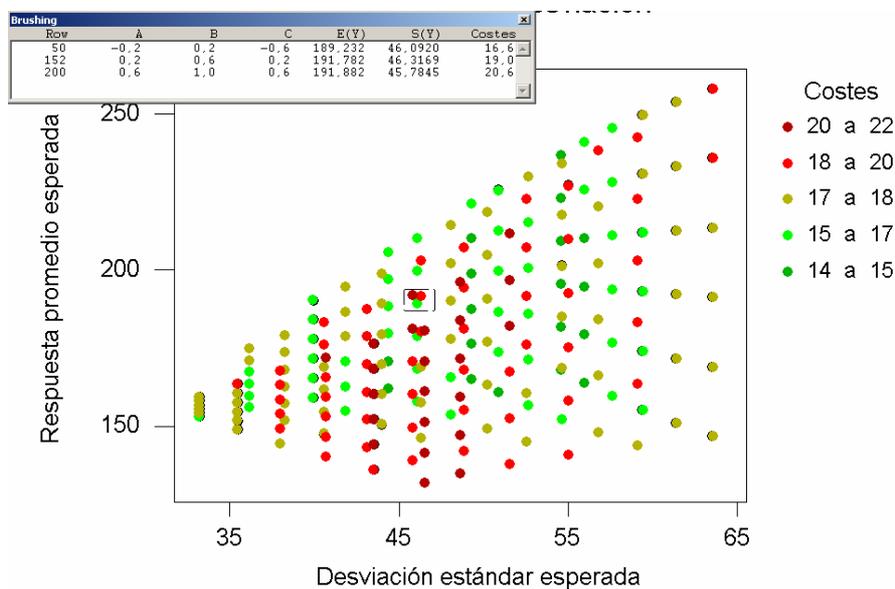


Figura 4.29 Gráfico DM con costes de operación distintos a los de la Figura 4.28.

#### 4.7 Algunas consideraciones sobre la elección del modelo de la respuesta

Como se presentó en la sección 4.3, el modelo de la respuesta:

$$Y = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\boldsymbol{\gamma} + \mathbf{z}'\boldsymbol{\Gamma}\mathbf{z} + \mathbf{x}'\boldsymbol{\Lambda}\mathbf{z} + \mathbf{z}'\boldsymbol{\Lambda} + \mathbf{x}'\mathbf{N} + \varepsilon \quad (4.27)$$

contiene efectos lineales y cuadráticos en  $X$ , efectos lineales en  $Z$ , interacciones dobles entre los distintos factores e interacciones triples entre dos factores controlables y uno de ruido o entre dos factores ruido y uno controlable. Esta expresión permite modelar una amplia gama de situaciones que se presentan en la práctica. Los modelos que se deducen para el valor esperado y la varianza (ecuaciones 4.14 y 4.16):

$$E(Y) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} \quad y$$

$$V(y) = (\boldsymbol{\gamma} + \boldsymbol{\Lambda}'\mathbf{x} + \boldsymbol{\Lambda})(\boldsymbol{\gamma} + \boldsymbol{\Lambda}'\mathbf{x} + \boldsymbol{\Lambda}) + \mathbf{1}'\mathbf{G}\mathbf{1} + \sigma_\varepsilon^2,$$

tienen una estructura muy peculiar y hacen notar que el modelo de la respuesta se encuentra compuesto por una parte que determina al valor esperado, por otra que determina a la variabilidad y por el error. La primera parte contienen únicamente factores controlables y la segunda parte contiene a todos aquellos términos que involucran factores ruido. Tanto el modelo de la media como el de la varianza no contienen a ninguna de las variables de ruido, aspecto esencial para la elaboración del gráfico MD.

En la expresión de la varianza solo el primer término es función de los factores controlables, así que el segundo y tercer término no influyen, por ser constantes, en la búsqueda de las condiciones que minimizan la variabilidad. Este hecho puede ayudar a simplificar la expresión del modelo de la respuesta eliminando el término que contiene las interacciones entre las variables de ruido ( $\mathbf{z}'\boldsymbol{\Gamma}\mathbf{z}$ ), lo cual no afecta en nada al modelo del valor esperado y abrevia la expresión de la varianza a:

$$V(y) = (\boldsymbol{\gamma} + \boldsymbol{\Lambda}'\mathbf{x} + \boldsymbol{\Lambda})(\boldsymbol{\gamma} + \boldsymbol{\Lambda}'\mathbf{x} + \boldsymbol{\Lambda}) + \sigma_\varepsilon^2. \quad (4.30)$$

Naturalmente, de existir las interacciones entre los factores ruido la expresión 4.29 dará mejores estimaciones de la variabilidad. Su valor mínimo se tiene en cualquier caso cuando  $\gamma + \Delta'x + \Lambda = 0$ . Si no existen interacciones entre los factores controlables y los de ruido ( $\Lambda = 0$ ), como es sabido, la varianza permanece constante y no es posible encontrar condiciones robustas. Una propiedad importante asociada a los términos de la expresión de la varianza, cuando el modelo no contempla interacciones entre los factores ruido ( $G = 0$ ), ni interacciones triples ( $\Lambda = 0$  y  $N = 0$ ) está dada por el vector de las pendientes de la respuesta en la dirección de las variables ruido, en este caso:

$$\gamma + \Delta'x = \frac{\delta y}{\delta z} \quad (4.31)$$

ecuación que permite una interpretación geométrica del valor de la varianza.

Como es de esperar, y como sucede en todos los casos de modelación, la determinación de los términos que constituirán el modelo es esencial para el logro de los objetivos del estudio. Hay que tener presente en el momento de la selección que el objetivo no es tener un “buen” modelo de la respuesta sino tener “buenos” modelos de la esperanza y la varianza. De las expresiones 4.27, 4.28 y 4.29 se puede concluir que un modelo adecuado de la respuesta lo será de la esperanza pero no necesariamente de la varianza. La razón es que el modelo de la varianza contiene parámetros que no se encuentran en el modelo de la respuesta; estos se producen al multiplicarse  $(\gamma + \Delta'x + \Lambda)$  por su transpuesta. En el momento de determinar los efectos significativos que conformarán los modelos, es posible que para el modelo de la respuesta no resulte significativo, un efecto por tener un valor insuficiente, pero que multiplicado por otro efecto lo sería si estuviéramos modelando la varianza, por ejemplo  $\hat{\gamma}_3 = 6,12$  no se incluiría en un modelo de la respuesta pero  $\hat{\gamma}_3 \hat{\delta}_{11} = 57,34$  si se tomaría en cuenta para modelar la varianza. Como la varianza se modela indirectamente a través del modelo de la respuesta al eliminar  $\hat{\gamma}_3$  se elimina la posibilidad de que  $\hat{\gamma}_3 \hat{\delta}_{11}$  participe en la varianza estimada.

Es conveniente hacer la siguiente reflexión: al eliminar  $\hat{\gamma}_3$  puede suceder que en realidad el valor del parámetro sea cero, más también es posible que no lo sea pero que se esté actuando correctamente ya que su valor es insignificante para efectos prácticos de modelación de la respuesta. En el primer caso no afecta que no aparezca  $\hat{\gamma}_3\hat{\delta}_{11}$  en el modelo de la varianza, porque el valor del producto de los parámetros es cero. Sin embargo, en el segundo caso estaremos modelando inadecuadamente la varianza. El ejemplo de la resistencia al desgaste trata precisamente de esta situación, el modelo que se construye únicamente con los efectos que se consideran rutinariamente significativos al usar papel probabilístico normal es inadecuado, fue necesario ampliar el modelo para poder encontrar las condiciones robustas.

Entre otras, Pozueta (2001) considera las siguientes posibilidades para elegir el modelo de la varianza:

- Usar en el modelo de la respuesta los efectos que surgen significativos en la regresión paso a paso y deducir a partir de este el modelo de la varianza.
- Encontrar la expresión del modelo de la varianza a partir de la ecuación 4.12. Cada coeficiente de este modelo está relacionado con coeficientes del modelo de la respuesta, por esto se pueden estimar sustituyendo los coeficientes del modelo estimado de Y. Seleccionar los coeficientes significativos del modelo de la varianza usando un diagrama en papel probabilístico normal.

La primera opción es la que usó en el ejemplo de la resistencia al desgaste. La segunda alternativa tiene la desventaja de ser muy engorrosa si el número de factores es relativamente grande y asume de antemano que los factores que no son significativos tampoco son inertes.

Miller et al. (1993, página 33), sin recurrir a estos argumentos, usan en su modelo, de la distorsión en el engranaje de la transmisión de un coche, los términos de  $X_4$  y  $Z_1$  que no resultaron significativos al analizarse con papel

probabilístico normal. La explicación que esgrimen es que “Es más difícil interpretar la interacción significativa  $X_4Z_1$  sin la presencia de estos factores. Sin embargo, si el propósito del modelo es estrictamente de predicción  $X_4$  y  $Z_1$  deben desecharse”.

Es habitual encontrar en la literatura que el modelo estimado de la varianza se obtiene al sustituir en su expresión los coeficientes correspondientes del modelo estimado de la respuesta; el caso de la tasa de filtración es un ejemplo de ello. Obviamente sería preferible contar con un diseño expreso para estimar el modelo de respuesta y otro para estimar el modelo de la varianza, pero esto generalmente resulta impráctico y costoso. Sin embargo no se puede soslayar la posibilidad de tener un mal modelo de la varianza si se derivó del modelo de la respuesta. Se considera que una vez encontradas las condiciones de operación más adecuadas se necesita hacer experimentos confirmatorios, práctica por demás ampliamente recomendada. Si fallan las estimaciones quizá se deba a un modelo inadecuado de la varianza y por tanto habría que elegir un nuevo modelo. Sugerimos que si se incluyen efectos que originalmente se consideraron insignificantes, es conveniente tener en cuenta el error tipo II asociado a ellos (ver capítulo 5).

#### **4.8 Conclusiones**

Como se ha visto en el capítulo 3 existe una gran diversidad de enfoques para tratar el problema del diseño de parámetros, lamentablemente, a excepción de la metodología de Taguchi, pocos de ellos se encuentran en uso, quizá debido a que no han logrado incorporarse, más que esporádicamente, a los libros de texto y a las reglas de actuación, por requerir para su práctica de conocimientos estadísticos y matemáticos avanzados.

Se considera que es necesario contar con métodos que sean fáciles de comprender y de aplicar, y en ese sentido el gráfico MD presenta grandes ventajas:

- El empleo de modelos para la respuesta está muy difundido a través de los modelos estadísticos lineales.
- La deducción de las expresiones de la esperanza matemática y la varianza de la respuesta en fórmulas cerradas, cuyos valores se representan en el diagrama bivalente, son factibles de deducir.
- Es viable implementarlo de manera computacional, con variantes tales como: representación en los vértices o en toda la región de interés, manejo o no de costes y representación de más de una respuesta.
- Da a conocer con suma facilidad el comportamiento conjunto de la media y la dispersión en la región o regiones de interés, con lo cual no solo se encuentra la condición robusta sino que se adquiere un conocimiento más profundo del proceso.
- Muchas veces evidencia la imposibilidad de obtener al mismo tiempo el valor de la respuesta deseado y la mínima variabilidad, por lo cual este diagrama tiene la ventaja de que ofrece toda la información disponible para decidir cual es la combinación de valores de los factores de diseño que más conviene.
- Siempre es un diagrama bivalente, independientemente del número de factores controlables, de factores ruido, o del modelo que se considere.
- Cuando se compara con otros métodos, como los que se estudiaron en la sección 4.6, se aprecia la gran sencillez con la que se llega a los mismos resultados.
- La posibilidad de analizar de forma gráfica más de una respuesta y comprender el efecto que tienen entre si es de indudable valor.
- El uso de costes amplía considerablemente las opciones para encontrar condiciones adecuadas y viables de proceso.

- Es frecuente tener más de una combinación de valores de los factores controlables que de valores adecuados para la media y la desviación tipo, con lo cual se puede seleccionar de ellos la combinación que más favorezca otros criterios.

#### 4.9 Aportaciones

- Se formalizaron y dedujeron las expresiones matemáticas de la esperanza y la varianza en modelos que presenten cierta complejidad. Esto facilita su implementación computacional.
- Se elaboró un algoritmo que obtiene las expresiones de la esperanza matemática y la varianza de la respuesta, y que, implementado en el lenguaje de programación que incorpora Minitab, construye el gráfico automáticamente.
- Se contrastó el gráfico MD con distintos métodos de la literatura para valorar sus ventajas y desventajas.
- Se evidenció el potencial del método gráfico para analizar más de una respuesta.
- Se adecuó al gráfico la posibilidad de considerar costes, introduciendo un código de colores en los puntos del diagrama, con lo que se extiende su capacidad de encontrar condiciones razonables.