

5194:  
2

UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO



FACULTAD DE ESTUDIOS SUPERIORES ZARAGOZA



ESPECIALIDAD EN FARMACIA INDUSTRIAL

**DISEÑO DE MEZCLAS.  
UNA HERRAMIENTA ESTADÍSTICA EN LA INVESTIGACIÓN  
FARMACÉUTICA.**

TESINA QUE PARA OBTENER EL GRADO DE ESPECIALISTA EN PROCESOS FARMACÉUTICOS  
PRESENTA:

**Q.F.B. RICARDO ZAMORA RAMÍREZ**

ASESORES

**M. EN C. ARMANDO CERVANTES SANDOVAL  
DRA. RAQUEL LÓPEZ ARELLANO**

MÉXICO, D F MAYO DE 2000



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## INDICE POR TEMA

I. RESUMEN	1
II. INTRODUCCIÓN	2
III. REGRESIÓN LINEAL	3
A. FUNDAMENTOS	3
B. REGRESIÓN LINEAL MÚLTIPLE	4
1. PRUEBAS DE HIPÓTESIS DE LA REGRESIÓN LINEAL MÚLTIPLE	
a) CRITERIO PARA CONSIDERAR LA SIGNIFICANCIA DE LA REGRESIÓN	4
b) CRITERIO PARA LA SELECCIÓN DE LAS VARIABLES	5
c) COEFICIENTE DE DETERMINACIÓN $R^2$	6
2. CRITERIOS DE SELECCIÓN DE MODELOS	6
a) ERROR PURO Y FALTA DE AJUSTE	
b) GRÁFICA DE RESIDUALES	8
IV. SUPERFICIE DE RESPUESTA	
A. MODELACIÓN EMPÍRICA	10
B. DISEÑOS COMPUESTOS CENTRALES	11
C. DISEÑOS DE BOX-BEHNKEN	12
D. REPRESENTACIÓN GRÁFICA	13
V. BÚSQUEDA DE PUNTOS ÓPTIMOS	
A. LOCALIZACIÓN DEL PUNTO ESTACIONARIO	15
B. PROCEDIMIENTO DE PENDIENTE ASCENDENTE	16
C. DETERMINACIÓN DE LAS COORDENADAS DEL PUNTO ESTACIONARIO DE UN SISTEMA DE RESPUESTA	16
1. DESARROLLO DE LA FORMA CANÓNICA DE LA ECUACIÓN	18
2. ANÁLISIS RIDGE	19

<b>VI. DISEÑOS EXPERIMENTALES DE MEZCLAS</b>	
<i>A. DEFINICIÓN</i>	21
<i>B. CONSIDERACIONES</i>	21
<i>C. DISEÑO SIMPLEX Y CENTROIDE SIMPLEX</i>	24
<b>VII. MODELOS PARA LOS DISEÑOS DE MEZCLAS</b>	
<i>A. MODELACIÓN LINEAL</i>	27
<i>B. MODELO CUADRÁTICO</i>	27
<i>C. MODELO CÚBICO ESPECIAL</i>	27
<i>D. MODELO CÚBICO COMPLETO</i>	28
<i>E. INTERPRETACIÓN DE LOS COEFICIENTES ESTIMADOS</i>	29
<b>VIII. ENFRENTÁNDOSE CON LOS PROBLEMAS DE MEZCLAS</b>	
<i>A. PROBLEMAS SIN RESTRICCIONES</i>	32
1. PLANES DE GAMMON	32
2. PLANES DE LAMBRAKIS	33
<i>B. PROBLEMAS CON RESTRICCIONES</i>	35
1. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE INFERIOR	35
INTRODUCCIÓN A LOS L-PSEUDOCOMPONENTES	36
2. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE SUPERIOR.	39
INTRODUCCIÓN DE LOS U-SEUDOCOMPONENTES	39
3. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE SUPERIOR Y UN LÍMITE SUPERIOR	42
INTRODUCCIÓN A LOS DISEÑOS DE VÉRTICES	42
<b>IX. INTERPRETACIÓN DE LOS MODELOS AJUSTADOS QUE HACER UNA VEZ GENERADO EL MODELO</b>	44
PROCEDIMIENTO PARA ESCOGER EL MODELO CORRECTO	47
<b>X. VARIABLES DE PROCESO INCLUIDOS EN EXPERIMENTOS CON MEZCLAS</b>	48

## **XI. EJEMPLOS DE APLICACIONES DE LOS DISEÑOS DE MEZCLAS EN FARMACIA**

<b>A. DISEÑO DE MEZCLAS PARA UNA FORMULACIÓN DE LIBERACIÓN PROLONGADA DE TEOFILINA</b>	50
1. PLANTEAMIENTO DEL PROBLEMA	50
2. SALIDAS EN SAS PARA WINDOWS	52
3. CÁLCULO DE LAS PROPORCIONES DE LOS COMPONENTES EN SUS VALORES ORIGINALES	53
4. INTERPRETACIÓN DE LA SALIDA DE SAS PARA WINDOWS	54
<b>B. OPTIMIZACIÓN DE UNA FORMULACIÓN DE TABLETAS USANDO UN DISEÑO MEZCLAS</b>	55
1. PLANTEAMIENTO DE PROBLEMA	55
2. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA T90% EN BUFFER	55
INTERPRETACIÓN DE LOS RESULTADOS DE T 90% EN BUFFER	57
3. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA VARIACIÓN DE PESO	57
INTERPRETACIÓN DE RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA VARIACIÓN DE PESO	59
4. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA RESISTENCIA A LA RUPTURA	59
INTERPRETACIÓN DE LOS RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA RESISTENCIA A LA RUPTURA	60
5. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA FRIABILIDAD	61
INTERPRETACIÓN DE LA SALIDAS EN SAS PARA WINDOWS PARA PARA FRIABILIDAD	62
6. RESULTADOS GENERALES	63
INTERPRETACIÓN DE LOS RESULTADOS GENERALES	64

## **XII. DISCUSIÓN DE RESULTADOS** 65

## **XIII. CONCLUSIONES** 69

**APENDICE A**  
*DETERMINACIÓN DE INCONSISTENCIAS*

**APENDICE B**  
*TRANSFORMACIÓN DE RESTRICCIONES*

**APENDICE C**  
*DETERMINACIÓN DE VERTICES EXTREMOS*

**APENDICE D**  
*MANEJO DE DATOS EN EL PAQUETE SAS PARA WINDOWS 6.0*

## I. RESUMEN

Si bien es cierto que la técnica estadística diseño de mezclas no es totalmente desconocida, dado que se describe en varios libros y existen paquetes estadísticos que la presentan, no está siendo aprovechada en todo su potencial en el área farmacéutica, debido a que se desconoce en toda su extensión los fundamentos teóricos que la sustentan y su potencial de aplicación, es por ello que se realizó una investigación, recopilación, selección y depuración de información en diferentes bibliotecas y hemerotecas para que dicha información, una vez esquematizada y ordenada, se plasmara en un trabajo que sirva de guía para aplicar este diseño experimental a un problema específico

Por consiguiente, esta tesina presenta la información teórica necesaria para entender los casos en que se deben aplicar cada uno de los diseños de mezclas, la manera de trabajarlos en el paquete estadístico SAS y como interpretar los resultados obtenidos, así como ejemplos de su utilización.

Los resultados obtenidos muestran que los diseños de mezclas son una herramienta de bastante utilidad en la industria farmacéutica donde pueden ser aplicados en varias áreas como son: desarrollo, para preformulación y en las fases finales de optimización de formulaciones, en el área analítica, en cromatografía de líquidos o en producción por su capacidad de emplearse incluyendo variables de procesos.

Las conclusiones más importantes de este trabajo es que no existen técnicas estadísticas absolutas y dependerá del tipo de problema que se esté enfrentando para la elección de una u otra. Sin embargo, es posible complementarlas y facilitar de esta manera el conseguir los resultados buscados.

Si el problema es determinar una proporción de constituyentes para obtener un producto con ciertas características de calidad, el diseño de mezclas es una opción muy interesante a considerarse en las fases exploratorias iniciales e incluso en experimentos finales.

## I. RESUMEN

Si bien es cierto que la técnica estadística diseño de mezclas no es totalmente desconocida, dado que se describe en varios libros y existen paquetes estadísticos que la presentan, no está siendo aprovechada en todo su potencial en el área farmacéutica, debido a que se desconoce en toda su extensión los fundamentos teóricos que la sustentan y su potencial de aplicación, es por ello que se realizó una investigación, recopilación, selección y depuración de información en diferentes bibliotecas y hemerotecas para que dicha información, una vez esquematizada y ordenada, se plasmara en un trabajo que sirva de guía para aplicar este diseño experimental a un problema específico.

Por consiguiente, esta tesina presenta la información teórica necesaria para entender los casos en que se deben aplicar cada uno de los diseños de mezclas, la manera de trabajarlos en el paquete estadístico SAS y como interpretar los resultados obtenidos, así como ejemplos de su utilización.

Los resultados obtenidos muestran que los diseños de mezclas son una herramienta de bastante utilidad en la industria farmacéutica donde pueden ser aplicados en varias áreas como son: desarrollo, para preformulación y en las fases finales de optimización de formulaciones; en el área analítica, en cromatografía de líquidos o en producción por su capacidad de emplearse incluyendo variables de procesos.

Las conclusiones más importantes de este trabajo es que no existen técnicas estadísticas absolutas y dependerá del tipo de problema que se esté enfrentando para la elección de una u otra. Sin embargo, es posible complementarlas y facilitar de esta manera el conseguir los resultados buscados.

Si el problema es determinar una proporción de constituyentes para obtener un producto con ciertas características de calidad, el diseño de mezclas es una opción muy interesante a considerarse en las fases exploratorias iniciales e incluso en experimentos finales.



## II. INTRODUCCIÓN

La estadística es una herramienta que se ha convertido en un apoyo primordial en la industria farmacéutica, a tal grado que la FDA (Food and Drug Administration) mediante los GMP's (Good Manufacturing Practices) y GLP's (Good Laboratory Practices) recomienda su uso de forma rutinaria para el control de procesos, el establecimiento de proyectos de desarrollo, para el tratamiento e interpretación de nuevas formas de dosificación de activos.

Esbozando el uso de la estadística, en el área farmacéutica, podemos mencionar que se emplea en control de calidad, en pruebas de estabilidad, en validación y desarrollo de métodos analíticos, en protocolos de diseños preclínicos, en validación de procesos y en desarrollo farmacéutico: siendo en estos dos últimos puntos donde se puede enfatizar su utilidad mediante el uso de diseño de experimentos y procesos de optimización.

Dada la importancia actual por optimizar los recursos con que se cuenta en la industria y en las universidades es importante conocer la existencia de herramientas que nos apoyen en la localización de la mejor respuesta a las variables estudiadas, de una forma segura, rápida y con el menor número de experimentos.

Dentro de la metodología de superficie de respuesta se encuentra el diseño de mezclas; esta técnica se utiliza con escasa frecuencia en nuestro país, debido probablemente a que es poco conocida, y también porque su análisis era muy complejo (facilitándose actualmente con el uso de la computadora) pero no por ello deja de ser una herramienta potencialmente útil en farmacia.

Los *objetivos particulares* de este trabajo son:

- \* Analizar los fundamentos, ventajas y restricciones de los diseños estadísticos de mezclas.
- \* Mostrar la aplicabilidad de los diseños estadísticos de mezclas en la investigación farmacéutica.
- \* Aplicar los diseños de mezclas mediante el software estadístico SAS para windows.

El objetivo general de esta tesina es: conjuntar los elementos teóricos y prácticos sobre el diseño de mezclas para mostrar el potencial de esta herramienta estadística en el área farmacéutica, mostrando su aplicación mediante ejemplos utilizando el software estadístico SAS para Windows.

Consecuentemente:

*Este material proporcionará los elementos necesarios para comprender y decidir sobre la adecuada aplicación del diseño de mezclas en problemas específicos en la investigación farmacéutica, facilitando el análisis e interpretación de resultados mediante el uso del software estadístico SAS para Windows*

### III. REGRESIÓN LINEAL

#### A. FUNDAMENTOS.

En este capítulo se presentan los fundamentos de la regresión lineal múltiple, herramienta indispensable para generar un modelo de comportamiento de una o más variables independientes sobre una variable dependiente, también se establecen los criterios de selección de las variables; así como los criterios de selección del modelo mismo. En el capítulo IV se presenta la metodología de superficie de respuesta; la modelación empírica, los diseños compuestos centrales y de Box-Behnken así como la representación gráfica para determinar la respuesta como una superficie sólida; posteriormente en el capítulo V se presentan las metodologías de búsqueda de puntos óptimos de manera numérica, como localizar el punto estacionario y las técnicas para localizar la mejor respuesta. Resumiendo en los capítulos III, IV y V se presentan las bases teóricas que facilitan la comprensión de los diseños de mezclas y el lector está en posibilidad de discernir si le parece conveniente recordar estos fundamentos teóricos o si prefiere directamente adentrarse en el mundo de los diseños de mezclas que se presentan a partir del capítulo VI.

Antes de iniciar con la empresa de leer este capítulo conviene recordar ciertos conceptos que facilitarían su entendimiento. En estadística, tradicionalmente se utilizan letras griegas para representar parámetros de la población y letras en español para representar muestras estadísticas; por consiguiente  $\mu$  es la media de la población,  $\sigma$  es la desviación estándar de la población,  $\sigma^2$  es la variancia de la población,  $\bar{Y}$  es la media de la muestra,  $s$  es la desviación estándar de la muestra y  $s^2$  es la variancia de la muestra.

Las variables con subíndices son notaciones empleadas para identificar inequívocamente, un miembro específico de un conjunto de mediciones, si nos referimos al conjunto completo de mediciones como  $X$ ,  $Y$  y  $Z$  o algunas otras letras del alfabeto, colocamos un subíndice a la derecha y abajo de la letra de referencia para identificar una medida particular del conjunto, por ejemplo  $X_5$  se refiere al quinto miembro del conjunto  $X$ . A menudo se emplean la letra  $i$  y  $k$  como subíndices para identificar el  $i$ -ésimo elemento y el  $k$ -ésimo elemento de un conjunto.

Las letras  $N$  y  $n$  se refieren a los elementos de una población y a los elementos de la muestra respectivamente.

El coeficiente de correlación se denota con la letra  $r$  y representa el grado de asociación entre dos variables.

El coeficiente de determinación,  $r^2$ , es un número que varía entre 0 y 1, y representa la proporción de la variación total presente en los valores de  $Y$  que es explicada por la ecuación de regresión; así mismo las letras  $\hat{\beta}_0$  y  $\hat{\beta}_1, \dots, \hat{\beta}_k$  son estimadores de los coeficientes de regresión  $\beta_0, \beta_1, \dots, \beta_k$  presentes en la ecuación de regresión y  $b_0, b_1, \dots, b_k$  son los valores de ajuste, y se obtienen a partir de la muestra (predichos).

## B. REGRESIÓN LINEAL MÚLTIPLE<sup>(1)(2)</sup>

Con esta técnica se desea determinar la relación entre una o más variables de regresión  $X$  y la respuesta  $Y$ . El problema general consiste en ajustar un modelo

$$Y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad \text{Desde } i = 1, 2, \dots, n \quad \dots \dots \dots (1)$$

Donde:

$\beta_0, \beta_1, \dots, \beta_k$  son parámetros (llamados coeficientes de regresión)

$X_{i1}, X_{i2}, \dots, X_{ik}$  son valores fijos.

$\varepsilon_i$  son errores independientes con distribución normal  $(0, \sigma^2)$

Usualmente los parámetros desconocidos ( $\beta_k$ ) se denominan coeficientes de regresión y pueden determinarse mediante **mínimos cuadrados**. Donde  $\varepsilon_i$ , denominado error aleatorio, presenta una media igual a cero y varianza  $\sigma^2$ .

### 1. PRUEBAS DE HIPÓTESIS DE LA REGRESIÓN LINEAL MÚLTIPLE

#### a) CRITERIO PARA CONSIDERAR LA SIGNIFICANCIA DE LA REGRESIÓN

A menudo se desea probar la hipótesis de que tan significativos son los parámetros del modelo de regresión lo cual se logra probando que los coeficientes son diferentes a cero, es decir:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_a: \beta_i \neq 0, \text{ donde } i = 1, 2, \dots, k$$

Rechazar  $H_0$  implica que al menos una de las variables del modelo contribuye significativamente al ajuste. La prueba de esta hipótesis es una generalización del modelo utilizado en regresión lineal simple, donde la suma total de cuadrados ( $SC_T$ ) se descompone en la suma de cuadrados de regresión ( $SC_R$ ) y en la suma de cuadrados del error ( $SC_e$ ).

$$SC_T = SC_R + SC_e \quad \dots \dots \dots (2)$$

Consecuentemente el valor de  $F$  estimado se obtiene de la ecuación:

$$F = \frac{SC_R / k}{SC_e / (n - k - 1)} = \frac{MC_R}{MC_e} \quad \dots \dots (3)$$

Valor que se compara con tablas de  $F$  con  $k$  y  $n-k-1$  grados de libertad. Donde  $k$  se refiere al número de coeficientes del modelo

**b) CRITERIO PARA LA SELECCIÓN DE LAS VARIABLES**

Es importante probar las hipótesis respecto a los coeficientes de regresión individuales; tales pruebas son útiles para evaluar cada variable de regresión en el modelo. En ocasiones el modelo puede ser más efectivo si se le introducen variables adicionales o, quizá si se eliminan una o más variables que se encuentran en el mismo.

Introducir variables al modelo de regresión provoca que la suma de cuadrados de la regresión aumente y que la del error disminuya. Por lo que se debe decidir si el incremento de la suma de cuadrados de la regresión es suficiente para garantizar el uso de la variable adicional en el modelo. Además si agregamos una variable poco importante al modelo podemos aumentar el cuadrado medio del error, disminuyendo así la utilidad del mismo

Las hipótesis para probar la significancia de cualquier coeficiente individual, por ejemplo  $\beta_i$  son:

$$H_0: \beta_i = 0$$

$$H_a: \beta_i \neq 0$$

Y la estadística apropiada para probar la hipótesis es:

$$t_\alpha = \frac{\beta_i}{\sqrt{CM_e C_{ii}}} \dots \dots \dots (4)$$

Donde:

$$H_0: \beta_i = 0 \text{ se rechaza si } |t_\alpha| > t_{\alpha/2, n-k-1}$$

Así como  $C_{ii}$  se refiere al  $i$ -ésimo elemento de la matriz  $(X'X)^{-1}$  que a su vez es la inversa de la diagonal  $(X'X)$ .

El estimador de mínimos cuadrados es  $\hat{\beta}$  y éste a su vez es una variable aleatoria, además la distribución  $\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$  se debe a que es una combinación lineal de observaciones  $Y_j$ . Por lo tanto, la varianza del coeficiente de regresión  $\hat{\beta}_i$  es igual a  $\sigma^2$  veces el  $(i+1)$ -ésimo elemento de la diagonal  $(X'X)^{-1}$ , es decir  $C_{ii}$ . Así cada coeficiente de regresión tiene la propiedad distribucional

$$\hat{\beta}_i \sim N(\beta_i, \sigma^2 C_{ii})$$

c) **COEFICIENTE DE DETERMINACIÓN  $r^2$**

Después de haber encontrado la recta de regresión, se debe de investigar que tan bien se ajusta el modelo a los datos mediante el cálculo de  $r^2$ . Este coeficiente se construye con base en dos cantidades. La primera es la suma de cuadrados del error ( $SC_e$ ), la cual representa la suma de las distancias de los datos a la recta que mejor se ajusta. La segunda cantidad es la suma de cuadrados alrededor de la media  $\bar{Y}$  y se conoce como la suma de cuadrados totales ( $SC_T$ ).

El valor de  $r^2$  se define como:

$$r^2 = \frac{SC_T - SC_e}{SC_T} \dots\dots\dots ( 5 )$$

Este valor representa el porcentaje de la suma de cuadrados total que es explicada por la relación lineal. Conviene aclarar que  $r^2$  es un buen indicador de la calidad del ajuste de la recta de regresión, pero no se debe usar como un criterio único de selección del modelo.

**2. CRITERIOS DE SELECCIÓN DE MODELOS**

Con anterioridad se mencionó que los valores de  $t$ ,  $r^2$  y  $F$  que se obtuvieron con la regresión lineal múltiple no son suficientes para evaluar la precisión de un modelo, por lo cual se requieren de otros elementos para verificar el ajuste de dichos modelos.

a) **ERROR PURO Y FALTA DE AJUSTE**

Con el fin de investigar si el modelo refleja el comportamiento de los datos correctamente, la suma de cuadrados del error se divide en dos partes: el error puro de la repetición y la falta de ajuste de la suma de cuadrados del error puro

La suma de cuadrados del error puro se calcula con los valores de las repeticiones en el punto central del diseño de superficie de respuesta. La suma de cuadrados del error puro se calcula de la forma

$$SC_{EP} = \sum_{i=1}^n \sum_{j=1}^{r_i} \left( Y_{ij} - \bar{Y}_i \right) \dots\dots\dots ( 6 )$$

Donde:

$Y_{ij}$  es la  $i$ -ésima observación del diseño  $i$ -ésimo

$\bar{Y}_i$  es la media de las  $r_i$  observaciones del diseño  $i$ -ésimo

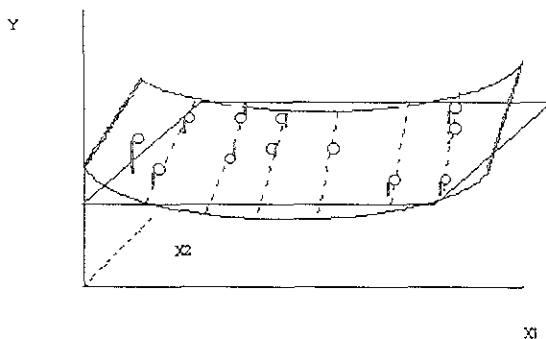
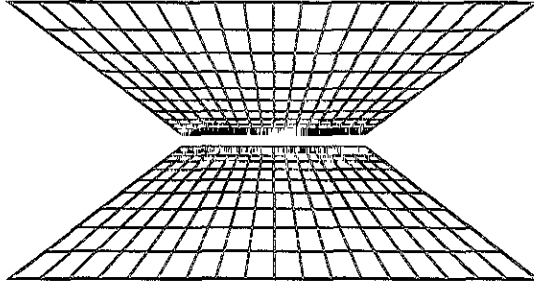


FIG 2 RESIDUALES CON RESPECTO A UNA MEDIA EN TRES DIMENSIONES

En este capítulo se recordaron los fundamentos de la regresión lineal múltiple, lo cual es de utilidad para generar y ajustar un modelo polinomial con dos o mas variables independientes y una variable de respuesta, los criterios que se deben usar para considerar la significancia e importancia de los parámetros obtenidos de dicho modelo y así observar la correlación que existe entre el modelo y los datos mediante el cálculo de  $r^2$ ; así también la evaluación de la idoneidad del ajuste del modelo generado mediante regresión lineal con el cálculo de la falta de ajuste y gráficas de residuales que compara las desviaciones de los datos reales con los estimados por el modelo. Con estas herramientas se facilitará la comprensión del siguiente capítulo que es la metodología de superficie de respuesta

## IV. SUPERFICIE DE RESPUESTA<sup>(1)(2)(3)</sup>



La metodología de superficie de respuesta (MSR) es la alternativa más adecuada en experimentos cuyo objetivo es establecer la relación entre los factores y la variable dependiente, dentro de una región experimental. La MSR está conformada por un conjunto de herramientas estadísticas que se aplican de acuerdo con las siguientes etapas:

1. Diseño y recopilación de datos experimentales, los cuales se usan en la determinación de una ecuación que se emplea para hacer predicciones
2. Aplicación de las técnicas de regresión lineal múltiple para obtener la mejor ecuación que represente el comportamiento de los datos.
3. Análisis de la superficie ajustada mediante gráficas de contorno y otras técnicas matemáticas y numéricas.

Dentro de estas herramientas, los diseños experimentales tienen la ventaja de explorar la relación entre los factores y la variable dependiente en toda la región experimental, y no solamente en las fronteras.

### A. MODELACIÓN EMPÍRICA.

En la mayoría de los problemas de MSR se desconoce la forma de la relación entre la respuesta y las variables independientes; por ello, el primer paso en la MSR consiste en determinar una aproximación apropiada en la relación funcional real entre  $Y$  y las variables independientes. Por lo general se emplea un polinomio de orden bajo, donde si la respuesta es descrita adecuadamente por una función lineal de las variables independientes, la función de aproximación en el modelo de primer orden es:

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad \dots \dots \dots (9)$$

Cuando existen curvaturas en el sistema debe utilizarse un polinomio de mayor grado; por ejemplo, el modelo de segundo orden:

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \beta_{ii} X_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k \beta_{ij} X_i X_j + \varepsilon \dots \dots \dots (10)$$

El modelo de la ecuación 10, a pesar de parecer sencillo, es un modelo bastante flexible y con coeficientes adecuados tiene la capacidad de caracterizar una gran variedad de superficies

Algunas de la propiedades que debería tener un diseño experimental de segundo orden es:

- Poder estimar los coeficientes del modelo cuadrático
- Tener un número pequeño de experimentos
- Facilidad para detectar falta de ajuste
- Uso de bloques

Para satisfacer el primero de estos criterios cada variable debe tener al menos tres niveles. El diseño más sencillo que cumple con estos requisitos es el factorial  $3^k$  (-1, 0, +1), la desventaja de este diseño es que para valores grandes de  $k$  se requieren un número demasiado elevado de experimentos, (tabla 2).

No. de factores $k$	No. de Exp. en un diseño $3^k, N$	No. de coeficientes en la ecuación cuadrática
2	9	6
3	27	10
4	81	15
5	243	21
6	729	28
7	2187	36

TABLA 2 INCREMENTO DEL NÚMERO DE EXPERIMENTOS A MEDIDA QUE AUMENTA EL NÚMERO DE FACTORES

### B. DISEÑOS COMPUESTOS CENTRALES

Otra clase de diseños que tienen la ventaja de requerir menos experimentos que los  $3^k$  son los diseños compuestos centrales. Estos se construyen con base en factoriales con dos niveles (lo cual permite la estimación de efectos principales e interacciones) Además incluyen un conjunto de puntos en los ejes (llamados puntos estrella) los cuales junto con el punto central (por lo general repetido), permiten estimar términos cuadráticos puros, la combinación de los puntos centrales y estrella requiere de tres niveles de cada variable independiente, denotados por  $-\alpha$ , 0 y  $+\alpha$ .



La magnitud alfa (distancia del origen al punto de intersección) y las repeticiones de los puntos centrales están en relación con la rotabilidad y la variación uniforme, respectivamente.

El primero de ellos indica que la precisión de las predicciones de la ecuación cuadrática solamente depende de la distancia de dicho punto al origen y no de su dirección, este criterio define el valor de  $\alpha$  (el valor de  $\alpha$  para tener rotabilidad depende del número de puntos en la porción factorial del diseño). Generalmente  $\alpha = \sqrt[3]{n}$ .

El segundo expresa que la precisión en las predicciones es la misma en la parte media que en las orillas y fija el número de puntos centrales.

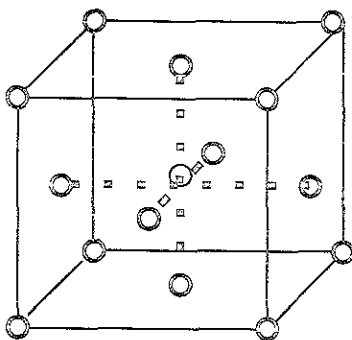


FIG. 3 ESQUEMA DE UN DISEÑO COMPUESTO CENTRAL.

### C. DISEÑOS DE BOX-BEHNKEN

Otra alternativa para la estimación de superficies de respuestas es el uso de diseños de *Box-Behnken*. Estos tienen dos ventajas sobre los diseños compuestos centrales; la primera es que utilizan menos experimentos (este ahorro es mínimo cuando se tiene de 5 a 7 factores) y la segunda es que en estos diseños existen solamente tres niveles (+1, 0, -1). Se considera que estos diseños son rotables (o casi rotables).

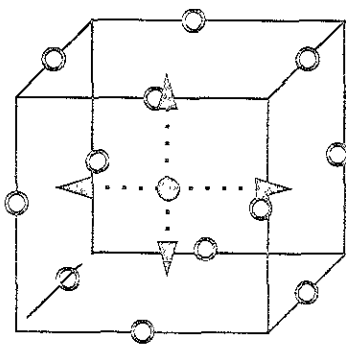


FIG. 4 DISEÑO BOX-BEHNKEN

## D. REPRESENTACIÓN GRÁFICA

Normalmente se evalúa a  $Y$  como una función de las variables independientes, es decir,  $Y = f(X_1, X_2, \dots, X_k) + \varepsilon$ , donde  $\varepsilon$  representa el error observado en la respuesta. Si la respuesta esperada se denota por  $E(Y) = f(X_1, X_2, \dots, X_k)$  entonces la superficie representada por  $\eta = f(X_1, X_2, \dots, X_k)$  se denomina superficie de respuesta. Es posible representar gráficamente la respuesta, por ejemplo, para un problema en el cual se desea maximizar el rendimiento  $Y$  de un proceso que está determinado por dos factores  $X_1$  (*temperatura*) y  $X_2$  (*presión*) obtenemos la *figura 5*, donde la respuesta se representa como una superficie sólida en un espacio tridimensional.

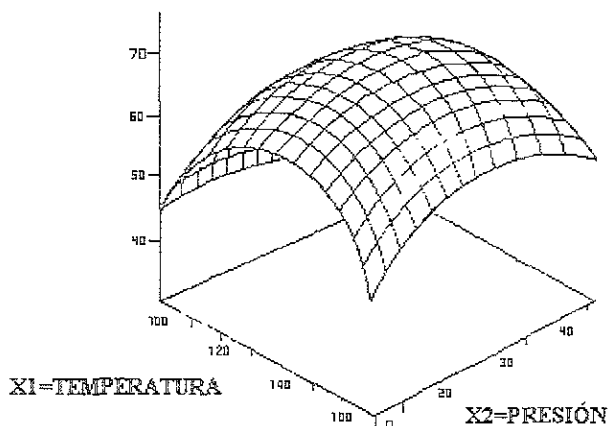


FIGURA 5 SUPERFICIE DE RESPUESTA PARA LOS FACTORES PRESIÓN Y TEMPERATURA

Con la finalidad de visualizar mejor la superficie de respuesta, a menudo se gráficán los contornos de dicha superficie en forma bidimensional; en la cual se trazan líneas de respuesta en el plano  $X_1, X_2$ . Cada contorno corresponde a una altura específica de la superficie de respuesta, tal gráfica es útil para estudiar los niveles de  $X_1$  y  $X_2$  que dan por resultado cambios en la forma o altura de la superficie, y por consiguiente con este tipo de gráficas se facilita visualizar las relaciones cuando existen más de dos factores, (*figura 6*).

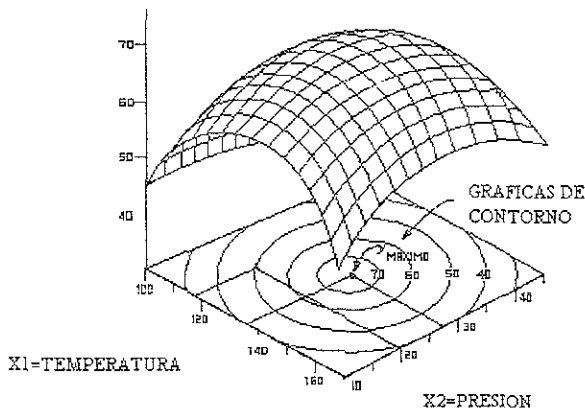


FIGURA 6 GRÁFICA DE CONTORNOS PARA LOS NIVELES DE  $X_1$  Y  $X_2$

La metodología de superficie de respuesta permite establecer la relación entre variables dependientes e independientes mediante la generación de un modelo polinomial de primer o segundo grado; dicho modelo se puede establecer a través de diversos diseños de experimentos como los factoriales  $3^k$ , los diseños compuestos centrales y los diseños de *Box-Behnken*; y posteriormente se establece la posibilidad de representar la respuesta a través de gráficas de contornos lo que facilitará observar las tendencias de los resultados al modificar los niveles de los factores. Con estas herramientas se está en posibilidad de comprender los diseños de mezclas, pero antes de pasar a ellos se presenta en el siguiente capítulo las técnicas para localizar los puntos óptimos (máximos o mínimos) de manera numérica con el fin de conocerlos y consecuentemente localizar de manera más precisa los puntos de respuesta buscados.

## V. BÚSQUEDA DE PUNTOS ÓPTIMOS<sup>(3)</sup>

### A. LOCALIZACIÓN DEL PUNTO ESTACIONARIO

Existen varios procedimientos para determinar las condiciones bajo las cuales obtendremos el mejor valor de respuesta; esta respuesta podría ser un rendimiento máximo o un nivel de calidad alto, también es posible desear obtener un costo mínimo o el menor número de defectos en un proceso. En cualquiera de estos casos se estarán buscando las condiciones óptimas.

El primer paso en el proceso de encontrar las condiciones óptimas es identificar las variables que presentan la mayor influencia en la respuesta, las cuales generalmente son pocas y se identifican con facilidad, el segundo paso consiste en expresar la respuesta de interés como una función de las variables, para lo cual usualmente se utiliza un modelo polinomial de primer orden, este modelo nos dará las bases para realizar las pruebas iniciales, y también hará posible modelar los valores del área de respuesta estimada sobre la región experimental, dicha representación tomará la forma de una gráfica de contorno del área estimada, dentro de esta gráfica las líneas de contorno son establecidas mediante la conexión de dos ejes ( $X_1$  y  $X_2$ ) en la región experimental que produce los mismos valores de  $Y$ , lo anterior se representa en la *figura 7* donde dos factores (tiempo de reacción y temperatura de reacción) modifican el rendimiento de una reacción química, es posible observar en esta figura bajo que condiciones de los factores se obtiene el mejor rendimiento.

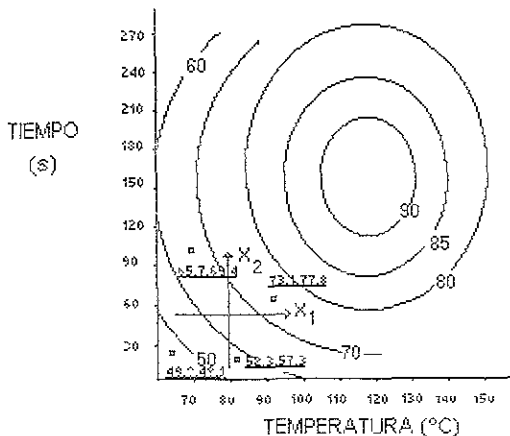


FIG 7 BÚSQUEDA DEL PUNTO ÓPTIMO

## ***B. PROCEDIMIENTO DE PENDIENTE ASCENDENTE***

Este procedimiento consiste en desarrollar una serie de experimentos en la ruta del máximo incremento sobre la respuesta. El procedimiento comienza precisamente con aproximar una área de respuesta usando una ecuación de 1<sup>er</sup> grado, y posteriormente al moverse sobre esta área se obtendrán incrementos de los valores de respuesta (se deberá tener en mente que existe una intensa dependencia de la dirección de la pendiente sobre los valores utilizados de las variables independientes), hasta aproximarse a una región donde se presentará una curvatura de dicha área, el incremento de la respuesta eventualmente se nivelará dentro de un punto máximo en una dirección particular. Si se continua sobre esta dirección y se localiza un decremento de la respuesta, se deberá de realizar una nueva serie de experimentos y a partir de ella generar un modelo de primer orden; esta secuencia de ensayos continuará hasta que se haga evidente que no existen incrementos adicionales en la respuesta. Una vez detectada una curvatura, se comienzan a pintar modelos cuadráticos (con interacciones) por medio de diseños compuestos centrales o de Box-Behnken.

## ***C. DETERMINACIÓN DE LAS COORDENADAS DEL PUNTO ESTACIONARIO DE UN SISTEMA DE RESPUESTA.***

Una región estacionaria es definida como una región donde la inclinación del área, o los gradientes a lo largo de los ejes de las variables, es menor comparada con el estimado del error experimental. El punto estacionario de una región estacionaria es el punto en el cual la inclinación del área de respuesta es cero cuando es tomada en todas direcciones.

Suponiendo que se desea determinar los niveles de  $X_1, X_2, \dots, X_k$  que optimizan la respuesta predicha. Este óptimo, si existe será el conjunto de  $X_1, X_2, \dots, X_k$ , tal que las derivadas parciales  $\frac{\partial Y}{\partial X_1} = \frac{\partial Y}{\partial X_2} = \dots = \frac{\partial Y}{\partial X_k} = 0$ . Dicho punto se denomina punto estacionario.

El punto estacionario podría representar:

*Una respuesta máxima*

*Una respuesta mínima*

*Un punto silla*

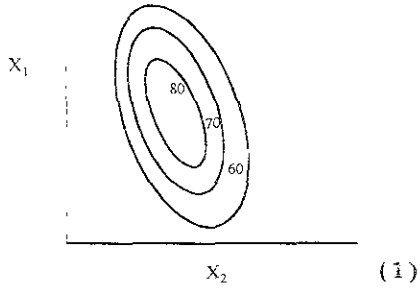
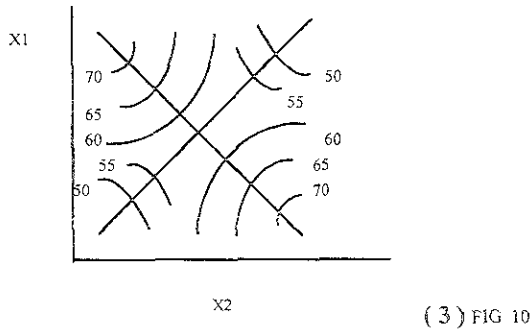
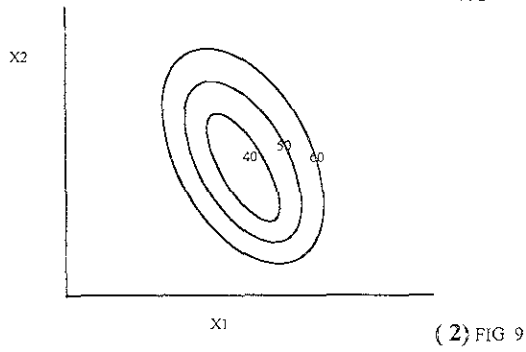


FIG 8



FIGURAS 8, 9, 10 PUNTOS MAXIMO, MÍNIMO Y PUNTO SILLA PARA 2 FACTORES

Se puede obtener una solución general para el punto estacionario Usando la notación matricial para el modelo de segundo orden se obtiene:

$$Y = \beta_0 + x'b + x'Bx \quad \dots \dots \dots (11)$$

Donde:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} \quad B = \begin{bmatrix} & & & \\ & b_{11} & \frac{b_{12}}{2} & \dots & \frac{b_{1k}}{2} \\ & & b_{22} & \dots & \frac{b_{2k}}{2} \\ & & & \ddots & \vdots \\ & & & & b_{kk} \end{bmatrix}$$

*Simétrica*

Los elementos del vector  $b$  son los coeficientes estimados de los términos de 1<sup>er</sup> orden de la ecuación 11, y los elementos de la matriz simétrica  $B_{K \times K}$  son los coeficientes estimados de los términos de segundo orden. Donde los términos de la diagonal principal son los coeficientes cuadráticos puros  $\beta_{ii}$  y los que están fuera de ella son términos mixtos  $\beta_{ij}/2$ ;  $i \neq j$ . Las derivadas parciales de  $\hat{Y}(X)$  con respecto a  $X_1, X_2, \dots, X_k$  son:

$$\begin{array}{l} \frac{\partial \hat{Y}(X)}{\partial X_1} = b_1 + 2b_{11}X_1 + \sum_{j=2}^k b_{1j}X_j \\ \frac{\partial \hat{Y}(X)}{\partial X_2} = b_2 + 2b_{22}X_2 + \sum_{j=2}^k b_{2j}X_j \\ \frac{\partial \hat{Y}(X)}{\partial X_k} = b_k + 2b_{kk}X_k + \sum_{j=1}^{k-1} b_{kj}X_j \end{array} \quad \left| \quad b + 2BX \right.$$

Al igualar cada una de las  $K$  derivadas a cero y resolviendo para cada uno de los valores de las  $X_i$  se encuentran las coordenadas de los puntos estacionarios que a su vez serán los valores de los elementos del vector  $X_0$  dados por:

$$X_0 = -\frac{B^{-1}b}{2} \quad \text{Donde } B^{-1} \text{ es la inversa de la matriz } B \text{ de la ecuación 11.}$$

## 1. DESARROLLO DE LA FORMA CANÓNICA DE LA ECUACIÓN

Una vez obtenido el punto estacionario, es necesario caracterizar la superficie de respuesta en la vecindad inmediata de este punto. Por caracterizar se entiende determinar a cual de las situaciones antes descritas (máximo, mínimo o punto silla) corresponde este punto, así como la sensibilidad relativa de la respuesta a las variables  $X_1, X_2, \dots, X_k$ .

La forma más directa de hacer esto consiste en examinar la gráfica de contornos del modelo ajustado. Es conveniente primero transformar el modelo en un nuevo sistema de coordenadas con el origen en el punto estacionario denominado  $x_0$  y entonces rotar (girar)

los ejes de este sistema hasta que sean paralelos a los ejes principales de la superficie de respuesta ajustada, como se muestra en la *figura 11*.

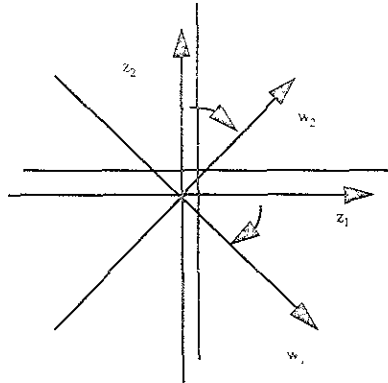


FIG 11 FORMA EN QUE SE ROTAN (GIRAN) LOS EJES DE ESTE SISTEMA HASTA QUE SEAN PARALELOS A LOS EJES PRINCIPALES DE LA SUPERFICIE DE RESPUESTA AJUSTADA

De esta manera obtenemos un modelo ajustado de la forma:

$$\bar{y} = \bar{y}_0 + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \dots + \lambda_k w_k^k \quad \dots \dots \dots (12)$$

Donde las ( $w_i$ ) son las variables independientes transformadas y las ( $\lambda_i$ ) son los valores propios. La naturaleza de la superficie de respuesta puede determinarse a partir del punto estacionario, el signo y la magnitud de ( $\lambda_i$ ). Si todas las  $\lambda$  son positivas, entonces  $x_0$  es un punto de respuesta mínima, si todas las  $\lambda$  son negativas, entonces  $x_0$  es un punto de respuesta máxima y si las  $\lambda$  tienen distintos signos  $x_0$  corresponde a un punto silla.

#### D. ANÁLISIS RIDGE

Durante el análisis del área de respuesta ajustada se podría descubrir que el punto estacionario no está dentro de la región experimental, pero si se sospecha que el valor óptimo de la respuesta se encuentra en los bordes de la región experimental, es posible utilizar el análisis *Ridge* para localizar este punto.

En general, el análisis *Ridge* se utiliza para encontrar un valor máximo absoluto (o un mínimo) de la respuesta estimada de  $\hat{Y}$  en esferas concéntricas de radio variante,  $R_L$  ( $L=1, 2, \dots, k$ ) los cuales son centrados en  $(X_1, X_2, \dots, X_k) = (0, 0, \dots, 0)$  y son contenidos dentro de la región experimental. Lo anterior se muestra en la *figura 12* donde los tres círculos se trazan con diferente radio.



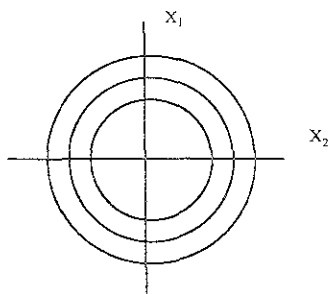


FIGURA 12 CIRCULOS CONCENRICOS DEL ANÁLISIS RIDGE EN LOS CUALES SE LOCALIZAN VALORES MÁXIMOS DE LA RESPUESTA

Asumiendo que el modelo ajustado sobre la región de las  $k$  variables codificadas  $X_1, X_2, \dots, X_k$  es de segundo orden y se expresa por la ecuación 11; suponiendo además que el punto estacionario está en los bordes de una esfera de radio  $R$ , se restringirá el examen a encontrar las coordenadas de las variables que maximicen  $Y$  sujetas a la condición:

$$\sum_{i=1}^k X_i^2 = R^2$$

Una vez encontradas esas coordenadas para un valor particular de  $R$ , se podría cambiar el valor de  $R$  y repetir el procedimiento para varios valores elegidos de  $R$  y graficándolos con sus apropiadas coordenadas en  $X_1, X_2, \dots, X_k$  contra  $\hat{Y}$ ; esto generaría gráficas de valores de  $\hat{Y}$  para varias distancias del punto central.

En este capítulo se muestran las técnicas de determinación de puntos óptimos es decir las condiciones bajo las cuales se obtendrá el mejor valor de respuesta un máximo o un mínimo según sea el caso; esto consistirá primero en determinar las variables que mayor influencia tienen sobre el proceso, para posteriormente expresar dicha respuesta como una función de las variables mediante un modelo polinomial de primer orden y desarrollar una serie de experimentos sobre la ruta de máximo incremento que se puede observar mediante una gráfica de contornos y con un incremento en la pendiente de la ecuación hasta llegar al punto donde no existirán más incrementos, es decir que existe una inclinación del área llamado punto estacionario, una vez determinado este punto se procede a caracterizar la superficie de respuesta en su vecindad inmediata mediante un cambio de variables, a esto se le denomina análisis canónico y es de utilidad para conocer si el punto estacionario era un máximo, un mínimo o un punto silla. Otra técnica para localizar el punto estacionario es mediante el análisis *Ridge*, que consiste en generar esferas concéntricas de diferentes radios. Todas estas herramientas le son de utilidad para comprender como es que se consigue localizar los valores de las variables independientes que generarán el valor más apropiado de la variable dependiente. Es importante señalar que la técnica que utiliza el paquete SAS para Windows, por todo el soporte teórico con que cuenta, es el análisis *Ridge*

## VI. DISEÑOS EXPERIMENTALES DE MEZCLAS

Hasta ahora se han presentado diseños de superficie de respuesta donde los factores involucrados son independientes; sin embargo existen otros tipos de diseños de superficie de respuesta en los cuales los factores involucrados no son independientes, este es el caso de los diseños de mezclas, en donde al modificar algún componente de la mezcla se tendrá que modificar otro u otros de los componentes de la mezcla, en estos diseños se considera que la variable dependiente sólo es resultado de la cantidad de los elementos de la mezcla.

Retomando lo anterior, al evaluar la capacidad de una ecuación matemática para representar una superficie de respuesta en un diseño de mezclas, se supone que existe alguna correlación de una función  $Y = f(X_1, X_2, \dots, X_q)$ , la cual define la dependencia de  $Y$  en las proporciones  $X_1, X_2, \dots, X_q$  de los componentes. Esta función es continua en  $X_i$ , y es representada por un polinomio de primer o segundo grado.

### A. DEFINICIÓN<sup>(4)</sup>

Los diseños de superficie de respuesta para mezclas son una técnica estadística donde se busca conseguir una respuesta óptima (máximos o mínimos) en función de los elementos constitutivos de la mezcla.

En un programa experimental de este tipo se desea modelar el área experimental de combinaciones de la mezcla para:

- \* Predecir una respuesta debida a algún o algunos de los ingredientes de la mezcla
- \* Obtener alguna medida de la influencia en la respuesta de cada componente o de sus combinaciones.

Un programa experimental de mezclas consiste de  $n$  ensayos, en los cuales se observa la respuesta del  $i$ -ésimo ensayo denotado como  $Y_i$ ; en donde se asume que la varianza de la media  $\bar{Y}$  es la misma para todas las respuestas, observando que la  $i$ -ésima observación está compuesta por la media más el error  $\varepsilon_i$ , esto es  $Y_i = \mu + \varepsilon_i$ , con  $i = 1, 2, \dots, n$ ,  $1 \leq i \leq n$ , donde los valores del error  $\varepsilon_i$  no están correlacionados y tienen una media  $\mu$  y varianza  $\sigma^2$ .

### B. CONSIDERACIONES<sup>(2)(5)(9)</sup>

Una característica de los experimentos con mezclas es que las proporciones de los factores pueden estar medidas en fracción volumen, fracción peso, fracción mol o porcentaje. Si la porción del  $i$ -ésimo componente es  $x_i$  y existen  $q$  componentes en la mezcla, las proporciones deben satisfacer las restricciones:

$$A. \quad 0.0 \leq x_i \leq 1.0; \quad 1 \leq i \leq q$$

$$B. \quad \sum_{i=1}^q x_i = 1 \quad i = 1, 2, \dots, q$$

Debido a estas restricciones la región experimental o área de interés es un simplex regular con  $\{q-1\}$  dimensiones (simplex es una región donde se puede trazar una línea recta y todos los puntos de la recta quedan dentro de una misma región).

Los diseños de experimentos para mezclas tienen dos características que los hacen diferentes al resto de los diseños experimentales:

- 1) La respuesta obtenida de una mezcla depende de la proporción y no de la cantidad total de la mezcla.
- 2) La composición de los componentes no puede ser variada de forma independiente de los otros componentes de la mezcla.

Es conveniente aclarar que en la experimentación con mezclas continúan vigentes los conceptos básicos del diseño tradicional:

- Los datos se generan a partir de un diseño de experimentos de la recolección de respuestas cuidadosamente medidas.
- Un modelo generalmente un polinomio puede ser empleado para describir los datos experimentales.
- Las gráficas de superficie de respuesta pueden ser empleadas para encontrar la región de respuesta óptima.
- Pueden seleccionarse una serie de puntos experimentales adicionales con el fin de verificar el ajuste del modelo.

Si se desea analizar una mezcla con dos componentes tenemos las siguientes restricciones  $0.0 \leq x_1 \leq 1.0$  y  $0.0 \leq x_2 \leq 1.0$  y  $x_1 + x_2 = 1.0$ , lo cual limita el empleo de los diseños experimentales factoriales. Comparando el área experimental para un diseño  $2^2$  ésta incluye el área y las esquinas de un cuadrado, en el análisis de mezclas nuestra área experimental para una mezcla binaria corresponderá a los puntos de la línea  $x_1 = 1 - x_2$ . Cuando la mezcla es de tres componentes, la región experimental restringida corresponde a un triángulo equilátero, y para interpretarlo podemos utilizar coordenadas trilineales, donde cada uno de los lados de la gráfica representa una mezcla que carece en absoluto de uno de los tres componentes (el componente indicado en el vértice opuesto). Las nueve líneas de malla en cada dirección indican incrementos de 10% en los componentes respectivos (figura 14).

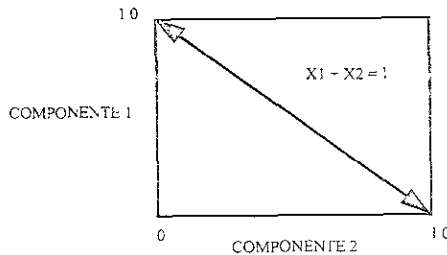


FIG. 13. ÁREA EXPERIMENTAL PARA UNA MEZCLA BINARIA QUE CORRESPONDE A LOS PUNTOS DE LA LÍNEA  $x_1 = 1 - x_2$ .

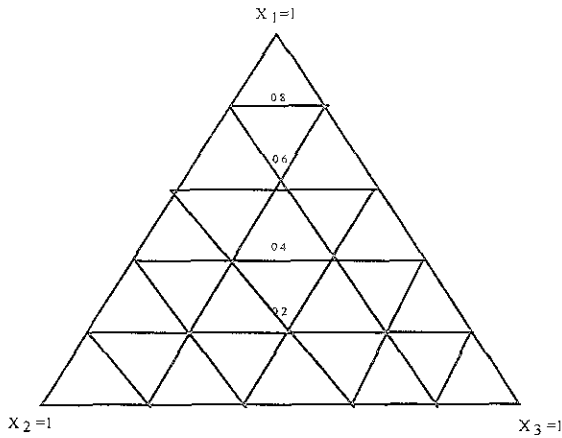
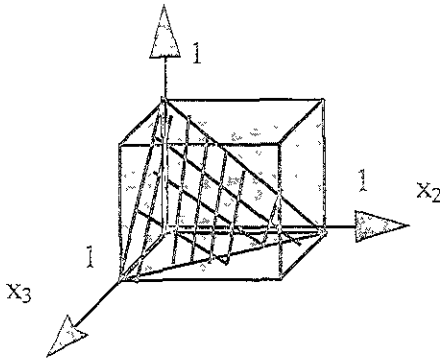


FIG. 14. MEZCLA DE TRES COMPONENTES. LA REGIÓN EXPERIMENTAL RESTRINGIDA CORRESPONDE A UN TRIÁNGULO EQUILÁTERO

Para una mezcla con cuatro elementos la región de exploración corresponderá a un plano que pasa equidistante de los tres componentes en un espacio tridimensional, como se observa en la figura 15.



El cual se escribe con la notación

$$\{(x_1, x_2, x_3) \mid x_i = 0, \frac{1}{2}, 1, i = 1, 2, 3 \quad x_1 + x_2 + x_3 = 1\}$$

FIGURA 15. REGIÓN EXPLORATIVA PARA UNA MEZCLA DE 4 COMPONENTES

### C. DISEÑO SIMPLEX Y CENTROIDE SIMPLEX<sup>(1)(4)</sup>

Los diseños simplex se emplean para estudiar los efectos de los componentes de mezclas en la variable de respuesta, porque la respuesta en los diseños de mezclas es obtenida en una región simplex completa; esto significa que deberá seleccionarse un modelo que represente de manera uniforme todo este espacio simplex; el diseño que tiene esta propiedad de uniformidad es el simplex en red de la forma  $\{q, m\}$ , en donde  $q$  se refiere al número de componentes y  $m$  es el número de combinaciones binarias sobre los vértices de la figura geométrica de respuesta (las proporciones asumidas por cada componente toman valores equiespaciados  $m+1$  desde 0 hasta 1).  $X_i = 0, 1/m, 2/m, \dots, 1$  donde  $i = 1, 2, \dots, q$ , y se emplean todas las posibles combinaciones (mezclas) de las proporciones de la ecuación.

Por ejemplo si suponemos que un sistema consta de 3 componentes ( $q = 3$  y  $m = 2$ ) cada componente llevará las proporciones:

$$X_i = 0, 1/2, 1 \text{ Donde } i = 1, 2 \text{ y } 3$$

y la red simplex consiste de las seis proporciones:

$$(X_1, X_2, X_3) = (1, 0, 0), (0, 1, 0), (0, 0, 1),$$

$$(1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2)$$

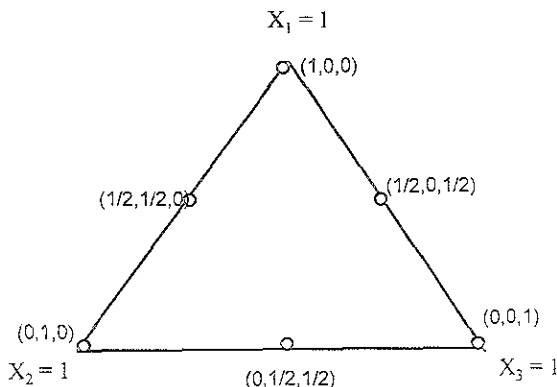


FIG. 16. DISEÑO EN RED CON 6 PUNTOS DE EXPERIMENTACIÓN EN UNA MEZCLA DE COMPONENTES 3.

Donde a los tres primeros se les denominará como los componentes puros y los tres siguientes corresponden a mezclas binarias, localizadas en los puntos centrales de los tres lados del triángulo

Una alternativa al diseño de red simplex es el diseño **centroide simplex**, este diseño tienen los puntos de la red del diseño de red simplex mas un punto central en el triángulo en donde los componentes tienen un valor equivalente a la unidad dividida por el número de componentes en estudio. En este tipo de diseños de  $q$  componentes hay  $2^q - 1$  puntos, que corresponden a las  $q$  permutaciones de  $(1,0,0,\dots,0)$ ,  $(q/2)$  permutaciones de  $(1/2,1/2,0,\dots,0)$ ,  $(q/3)$  permutaciones de  $(1/3,1/3,0,\dots,0)$  y el centroide global  $(1/q,1/q,\dots,1/q)$ .

El punto central se localiza mediante la determinación del nivel medio de todos los niveles de la mezcla; es decir en este punto los componentes tienen un valor equivalente a la unidad dividida por el número de componentes en el estudio.

El análisis del centroide simplex es similar al de red simplex asumiendo la misma restricción que la suma de todos los niveles de los factores debe de ser igual a 1.

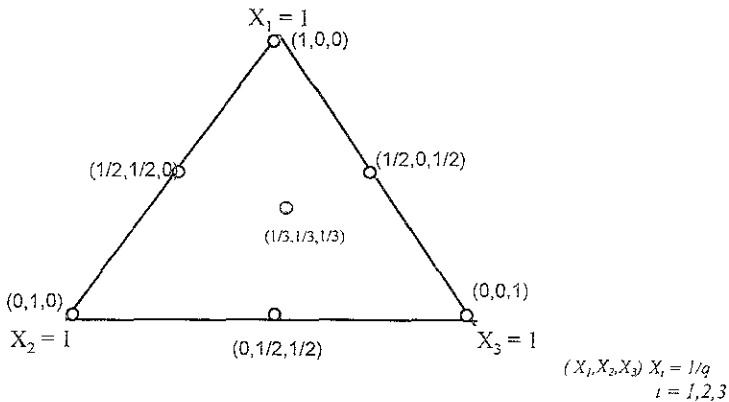


FIG 17 DISEÑO DEL CENTROIDE SIMPLEX

El diseño centroide simplex puede presentar variantes que incluyen combinaciones lineales de los componentes en estudio que caen interiormente dentro del área que forma el triángulo o región de estudio.

El ejemplo de la *figura 18* corresponde a un centroide agregado con tres factores que presenta 3 puntos en el interior del triángulo; correspondientes a los puntos  $(1/2, 1/4, 1/4)$ ,  $(1/4, 1/2, 1/4)$  y  $(1/4, 1/4, 1/2)$

Con estos diseños se puede dilucidar con mayor precisión la forma del área de respuesta dentro del área del triángulo.

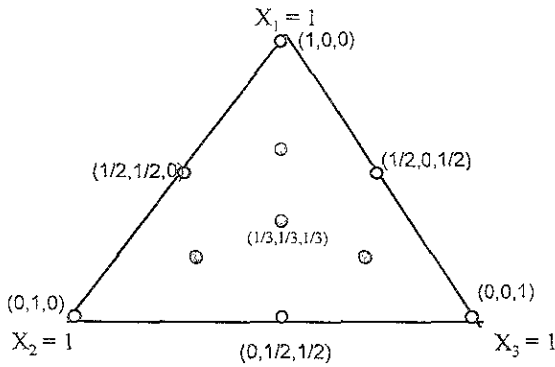


FIG. 18 COMBINACIONES DENTRO DEL ÁREA RESPUESTA, VARIANTES DE LOS DISEÑOS CENTRALES

Este capítulo introduce a los diseños de mezclas, definida como una técnica estadística que tiene por objetivo obtener una respuesta óptima (máxima o mínima) en función de los componentes de una mezcla, consecuentemente los niveles de dichos elementos o ingredientes no son independientes, generándose con ello dos restricciones, una que indica que la cantidad máxima de un componente no es posible que sea mayor de la unidad, y que la suma de los componentes de la mezcla no debe exceder la unidad (o el 100 %). Por consiguiente, el espacio donde se realizarán los experimentos corresponderá a un simplex regular, el cual en el caso de tres componentes corresponde a un triángulo equilátero. Al explorar una superficie de respuesta que corresponde a un simplex se deben utilizar los diseños simplex en red, los centroides simplex o alguna de sus variantes, generándose con ello una matriz de experiencias, con ayuda de los resultados de estos experimentos se busca un polinomio de respuesta que permita predecir respuestas para ciertas proporciones de los componentes; tema que se presenta en el próximo capítulo.

## VII. MODELOS PARA LOS DISEÑOS DE MEZCLAS<sup>(4)(6)(7)</sup>

La forma canónica del polinomio de mezclas es derivada por la aplicación de la restricción  $X_1 + X_2 + \dots + X_k = 1$  para simplificar los términos del polinomio estándar; es decir, se elimina el término constante de las ecuaciones polinomiales. Por ejemplo, para dos componentes  $X_1$  y  $X_2$ ; el polinomio estándar de primer grado es:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \quad \dots \dots (13)$$

Sin embargo, como  $X_1 + X_2 = 1$ , se puede reemplazar  $\beta_0$  por  $\beta_0(X_1 + X_2)$  en  $Y$  para obtener:

$$Y = (\beta_0 + \beta_1) X_1 + (\beta_0 + \beta_2) X_2 \quad \dots \dots (14)$$

$$Y = \beta_1' X_1 + \beta_2' X_2 \quad \dots \dots (15)$$

De esta manera se remueve el término  $\beta_0$ . En el caso del polinomio de segundo grado, los términos cuadráticos  $\beta_{11} X_1^2$  y  $\beta_{22} X_2^2$  son removidos del modelo, así como el término constante  $\beta_0$ . Por consiguiente los modelos de mezcla tienen menos términos que los polinomios estándar; a este tipo de formas canónicas de polinomios se les denomina modelos tipo Scheffé.

En general los modelos más importantes de mezclas son:

### A. MODELO LINEAL

$$Y = \sum_{i=1}^k \beta_i X_i \quad \dots \dots (16)$$

El modelo lineal se usará en los casos en que la mezcla de los componentes sea aditivo y la calidad del producto se defina como una combinación lineal de sus proporciones.

### B. MODELO CUADRÁTICO

$$Y = \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^j \sum_{i < j} \beta_{ij} X_i X_j + \varepsilon \quad \dots \dots (17)$$

El modelo cuadrático se emplea si existe interacción (antagonismo o sinergismo) entre los constituyentes de la mezcla, y por consiguiente la calidad fuese superior o inferior a la que se hubiera obtenido con la combinación lineal de sus proporciones

### C. MODELO CÚBICO ESPECIAL

$$Y = \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^j \sum_{i < j} \beta_{ij} X_i X_j + \sum_{i=1}^{k-1} \sum_{i < j} \sum_{i < j < k} \beta_{ijk} X_i X_j X_k + \varepsilon \quad \dots \dots (18)$$



### D. MODELO CÚBICO COMPLETO

$$Y = \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^j \sum_{i < j}^k \beta_{ij} X_i X_j + \sum_{i=1}^j \sum_{i < j}^k \delta_{ij} X_i X_j (X_i - X_j) + \sum_{i=1}^j \sum_{i < j}^k \sum_{i < j < k}^k \beta_{ijk} X_i X_j X_k + \varepsilon \dots (19)$$

El término cúbico  $\beta_{123}$  identifica la diferencia en los valores de la variable dependiente con la mezcla  $X_1=1/3, X_2=1/3, X_3=1/3$  y el valor pronosticado en dicho punto con el modelo cuadrático.

Los modelos lineal, cuadrático y cúbico completo son generalmente asociados con  $\{q, 1\}$ ,  $\{q, 2\}$  y  $\{q, 3\}$  del diseño de red simplex, porque establecido el número de componentes  $\{q\}$  sólo se requerirá una combinación binaria para generar el modelo lineal, dos para el modelo cuadrático y 3 para el modelo cúbico. Con respecto a la ecuación cúbica especial se observa que se reduce a la forma de un polinomio de tercer grado que posee las combinaciones ternarias de los componentes  $i, j$  y  $k$ , por consiguiente representa un polinomio al menos de grado de 2, donde es posible calcular el siguiente número de términos:

$$N = \frac{q(q^2 + 5)}{6} \dots\dots\dots(20)$$

Mientras que la fórmula para calcular el número de términos para el modelo cúbico completo es:

$$N = \frac{q(q+1)(q+2)}{6} \dots\dots\dots(21)$$

En la *tabla 2* se indica el número de puntos de respuesta, el número de factores y el número de espacios por factor:

	<i>Cuadrático</i>	<i>Cúbico especial</i>	<i>Cúbico</i>	<i>Cuártico</i>
Número de espacios ( <i>m</i> )	2	2	3	4
Número de factores ( <i>q</i> )	Número de puntos de respuesta ( <i>k</i> )			
3	6	7	10	15
4	10	14	20	35
5	15	25	35	70
6	21	41	56	126
8	36	92	120	330
10	55	175	220	715

TABLA 2. NÚMERO DE PUNTOS DE RESPUESTA, NÚMERO DE FACTORES Y EL NÚMERO DE ESPACIOS POR FACTOR

En general el número de puntos en el diseño de red simple es:

$$N = \frac{(q+m-1)!}{m!(q-1)!} \dots \dots (22)$$

La ecuación anterior no es aplicable para el simplex cúbico especial, para este modelo se utiliza la siguiente ecuación:

$$N = \frac{q(q+1)}{2} + \frac{q(q+1)(q-2)}{6} \dots \dots (23)$$

La forma canónica del polinomio en  $k$  componentes que tiene los puntos calculados del diseño de centroide simple es:

$$Y = \sum_{i=1}^k \beta_i X_i + \sum_{i < j} \sum \beta_{ij} X_i X_j + \sum_{i < j < k} \sum \beta_{ijk} X_i X_j X_k + \beta_{1,2, \dots, k} X_1 X_2 \dots X_k \dots (24)$$

Las gráficas facilitan la visualización del diseño de red simple que corresponde, así como las áreas de experimentación.

### E. INTERPRETACIÓN DE LOS COEFICIENTES ESTIMADOS<sup>(4)(2)</sup>

Los coeficientes  $\beta_i$  del modelo lineal representan la respuesta esperada del componente  $X_i=1, X_j=0; j \neq i$ . Y define la altura del área de la mezcla en el vértice simple denotado por  $X_i=1$ . La porción  $\sum_{i=1}^k \beta_i X_i$  de cada modelo es llamada porción de la combinación lineal y representa la combinación de los componentes estrictamente aditiva.

Los coeficientes binarios ( $\beta_{ij}$ ) son estimadores de la curvatura del área de respuesta (llamado frecuentemente sinergismo o antagonismo) en la combinación de 2 componentes.

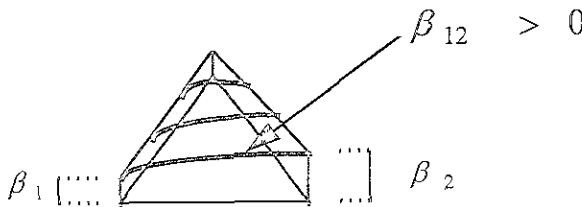


FIG. 19. SIN. ROLISMO O ANTAGONISMO EN LA COMBINACION DE 2 COMPONENTES

Una prueba para encontrar sinergismo en una combinación de 2 componentes consiste en probar la hipótesis  $H_0: \beta_{ij} = 0$  contra  $H_a: \beta_{ij} \neq 0$ ; la evaluación se efectúa mediante la ecuación:

$$t = \frac{\beta_{ij}}{e.e.(\beta_{ij})} \dots \dots \dots (25)$$

Donde e.e. es el error estándar del coeficiente  $\beta_{ij}$ .

Comparando el valor calculado de  $t$  contra el de tablas; si  $t_{cal} > t_{tab}$  inferimos con el nivel de significancia establecido que el valor de la respuesta de los componentes combinados es mayor que si utilizamos los componentes de manera individual.

Los términos estimados de mayor grado como  $\beta_{ijk}$  o  $\beta_{ijk}X_iX_j(X_i-X_j)$  describen desviaciones adicionales en la forma del plano dentro del área de respuesta, no sólo sobre los vértices

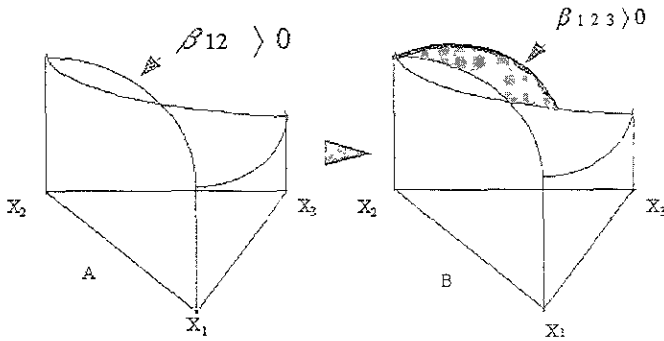


FIG 20 A) CURVATURA DEL AREA DE RESPUESTA EN LA COMBINACION DE 3 COMPONENTES B) CURVAS DENTRO DEL AREA

Cuando los datos son colectados sólo de los puntos del  $\{q, m\}$  red simplex (así como en los puntos del diseño simplex centroide) los coeficientes estimados en el polinomio canónico son una simple función de los valores observados de respuesta. Esto es porque el número de términos en los modelos es igual al número de puntos en el correspondiente diseño en red. Para demostrar esto suponemos que tenemos un  $\{3,2\}$  red simplex, y definimos como  $\bar{Y}$  la media con  $r$  observaciones replicadas, colectadas en  $X_i = 1, X_j = 0; i \neq j, i = 1, 2, 3, \dots, k$ . adicionalmente,  $\bar{Y}_{ij}$  es la media de las  $r_{ij}$  observaciones colectadas en la mezcla binaria 50% y 50% ( $X_i = 1/2, X_j = 1/2, \dots, X_k = 0$  de todos  $i < j < k$ ) de los componentes  $i$  y  $j$

Los parámetros estimados de los coeficientes del modelo cuadrático se obtienen usando:  
 $\beta_{11} = 4 Y_{11} - 2(Y_i + Y_j)$  (nótese que los estimadores  $\beta_{ij}$  sólo colectan los datos a lo largo del eje conectando los vértices  $X_i=1, X_j=0; i \neq j$  y  $X_j=1, X_i=0; i \neq j$ ).

Las cantidades escalares 4 y 2 en la fórmula no dependen de los valores de  $r_i$  y  $r_{ij}$ ; pero sí provienen de los valores de  $X_i$  y  $X_j$ . Es importante mencionar que la ecuación sólo puede utilizarse para calcular los coeficientes en los modelos cuadráticos y cúbicos especiales.

En este capítulo se establece que los modelos generados a través de los diseños de mezclas tienen un menor número de términos que los polinomios estándar y son denominados tipo Scheffé, dentro de los modelos más importantes están el lineal, cuadrático, cúbico especial y cúbico completo; donde el primero de ellos se utiliza en los casos en que la mezcla de los componentes genere un comportamiento aditivo en la respuesta (se asocia al diseño simplex  $\{q, 1\}$ ) el segundo si existe alguna interacción (sinergismo o antagonismo) en la combinación binaria de los componentes (se asocia al diseño simplex  $\{q, 2\}$ ) y el tercero y cuarto comparten términos cúbicos y evalúan comportamientos de la mezcla en combinaciones de tres ingredientes (se asocia a los diseños simplex  $\{q, 3\}$ ). En este momento se cuenta con los elementos teóricos suficientes para enfrentarse a los diseños de mezclas, tema del siguiente capítulo.



Para transformar la serie anterior de experimentos a un plan de Gammon se transformarán únicamente los primeros tres de la siguiente forma:

$$0 \text{ por } \frac{1}{3-1} = \frac{1}{2} \text{ y } 1 \text{ por } 0.$$

Por lo que se obtienen para los primeros 3 experimentos la *tabla 4*:

Num. de Exp.	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
1	0	1/2	1/2
2	1/2	0	1/2
3	1/2	1/2	0

TABLA 4 TRANSFORMACION DE LOS PRIMEROS 3 EXPERIMENTOS DE LA TABLA 3 A UN PLAN DE GAMMON

Es posible representar los cambios anteriores como puntos de un triángulo, generando los puntos de la *figura 21*:

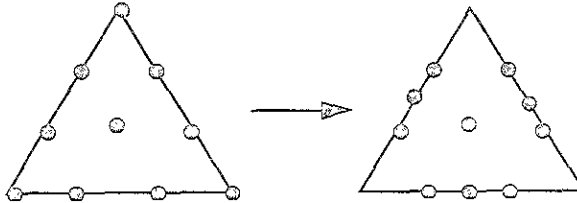


FIG. 21 REPRESENTACIÓN GRÁFICA DE LA TRANSFORMACIÓN DE LOS EXPERIMENTOS DE SCHEFFÉ A UN PLAN DE GAMMON PARA UNA MEZCLA TERNARIA

## 2. PLANES DE LAMBRAKIS<sup>(6)</sup>

Como se mencionó anteriormente la principal crítica a los planes de Scheffé es que incluyen la evaluación de los componentes puros. En muchas ocasiones se tendrá más interés en la evaluación de combinaciones donde participen todos los componentes. Ejemplificando el caso de una mezcla de 3 constituyentes se requiere evaluar la superficie dentro del triángulo y no considerar la respuesta sobre los componentes puros o sobre las mezclas de 2 constituyentes. En este tipo de situaciones los diseños de *Lambrakis* son de utilidad.

Mediante este arreglo se propone un comportamiento lineal de  $k$  valores 0, en donde  $N-k$  valores no nulos  $V_i$ , son remplazados de la siguiente forma:

$$0 \text{ por } \frac{1}{k(M+1)} \dots \dots \dots (26)$$

$$\text{y } V_i \text{ por } V_i \frac{M}{M-1} \dots \dots \dots (27)$$

Donde  $M$  se refiera al grado del polinomio

*Ejemplo.* Si requiere evaluar una mezcla cuaternaria que presenta el siguiente plan de Scheffé con 10 puntos, en sus variables codificadas:

NUM. DE EXP.	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	NUM. DE VALORES NULOS
1	1	0	0	0	3
2	0	1	0	0	3
3	0	0	1	0	3
4	0	0	0	1	3
5	1/2	1/2	0	0	2
6	1/2	0	1/2	0	2
7	1/2	0	0	1/2	2
8	0	1/2	1/2	0	2
9	0	1/2	0	1/2	2
10	0	0	1/2	1/2	2

TABLA 5 PUNTOS PARA UNA MEZCLA CUATERNARIA CON UN MODELO DE SCHEFFÉ CON 10 PUNTOS EN SUS VARIABLES CODIFICADAS

La modificación de *Lambrakis* a la serie anterior de experimentos se efectúa de la siguiente forma:

A) Para las líneas con tres valores nulos:

$$0 \longrightarrow \frac{1}{k(M+1)} = \frac{1}{3(2+1)} = \frac{1}{9}$$

$$1 \longrightarrow 1 \frac{M}{M+1} = \frac{2}{2+1} = \frac{2}{3}$$

B) Para las líneas con dos valores nulos

$$0 \longrightarrow \frac{1}{2(2+1)} = \frac{1}{6}$$

$$1 \longrightarrow \left(\frac{1}{2}\right) \frac{2}{2+1} = \frac{1}{3}$$

Por consiguiente el *plan de Lambrakis* para la mezcla cuaternaria se establece en la siguiente tabla.

NUM. DE EXPERIMENTO	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
1	2/3	1/9	1/9	1/9
2	1/9	2/3	1/9	1/9
3	1/9	1/9	2/3	1/9
4	1/9	1/9	1/9	2/3
5	1/3	1/3	1/6	1/6
6	1/3	1/6	1/3	1/6
7	1/3	1/6	1/6	1/3
8	1/6	1/3	1/3	1/6
9	1/6	1/3	1/6	1/3
10	1/6	1/6	1/3	1/3

TABLA 6 TRANSFORMACIÓN DE UN MODELO DE SCHEFFÉ A UN PLAN DE LAMBRAKIS PARA UNA MEZCLA CUATERNARIA EN SUS VARIABLES CODIFICADAS

## *B. PROBLEMAS CON RESTRICCIONES<sup>(4)(5)</sup>*

Como se mencionó con anterioridad, debido a las restricciones de los diseños de mezclas la región experimental es un simplex, pero frecuentemente debido a ciertas restricciones adicionales que presentan las proporciones de los componentes no se estará en completa libertad de explorar todo este espacio simplex completo (el total del rango  $0 \leq X_i \leq 1.0$  para todos los componentes). Es claro que siempre que existan restricciones el área de trabajo se verá reducida, lo que incrementará la precisión de los modelos estimados.

Existen diferentes situaciones en las cuales existen restricciones, pero todas podemos incluirlas en los tres casos que a continuación se describen:

- a. *restricciones cuando todos o algunos de los componentes presentan un límite inferior*
- b. *restricciones cuando todos o algunos de los componentes presentan un límite superior.*
- c. *restricciones cuando todos o algunos de los componentes presentan límites tanto inferiores como superiores.*

### **1. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE INFERIOR.**

En este caso se requiere que al menos una cierta proporción de un componente esté presente en todas las combinaciones, esto elimina el caso cuando  $X_i = 0$ .

Para ejemplificar se considera el caso de 3 componentes con las siguientes restricciones:

$$X_1 \geq 0.35, X_2 \geq 0.20 \text{ y } X_3 \geq 0.15$$

En la región resultante requerirá que al menos una porción de  $X_1 = 0.35$  del componente 1 esté presente en cada combinación, con al menos una porción de  $X_2 = 0.20$  del componente 2, combinado con al menos  $X_3 = 0.15$  del componente 3; por supuesto no todos los componentes pueden asumir el límite inferior simultáneamente porque en ese caso la suma de  $X_1 + X_2 + X_3 = 0.70$  es menor que la unidad y no es válido para los diseños de mezclas

El espacio posible de las mezclas correspondientes a estas restricciones se reduce a un *subespacio de la región simplex original* y el subespacio resultante es también un *simplex*. Lo anterior obliga a redefinir a los componentes originales hacia *L-pseudocomponentes* mediante:

$$X'_i = \frac{X_i - L_i}{1 - \sum_{i=1}^q L_i} \quad \dots \quad (28)$$



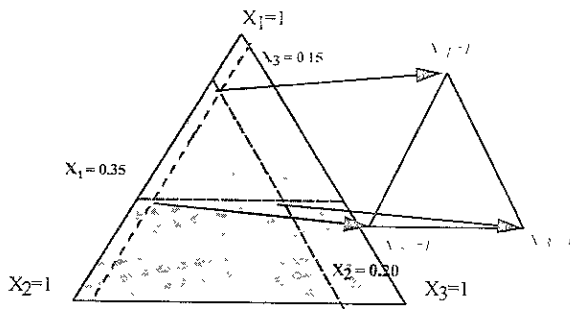


FIG. 22 EL ESPACIO POSIBLE DE LAS MEZCLAS CORRESPONDIENTES A LAS RESTRICCIONES ESTA REPRESENTADO POR EL TRIANGULO INTERIOR EL CUAL TAMBIÉN ES UN SIMPLEX

a) INTRODUCCIÓN A LOS L-PSEUDOCOMPONENTES.

Debido a que los límites inferiores de los componentes conducen a una subregión simplex más pequeña del simplex original, será evidente la necesidad de redefinir las coordenadas de la subregión en términos de pseudocomponentes con la finalidad de facilitar la construcción del diseño; dichos pseudocomponentes son definidos como combinaciones de los componentes originales. Sin embargo, se deberá recordar que estamos trabajando con pseudocomponentes y si se desea hacer inferencias respecto a los componentes se deberán transformar los pseudocomponentes a los componentes originales para generar el modelo deseado.

Para demostrar como se definen los L-pseudocomponentes en términos de los componentes originales y sus límites inferiores, diremos que en general el sistema consiste de q componentes y que  $L_i \geq 0$  es denotado por límite inferior de los componentes i,  $i = 1, 2, \dots, q$ , la restricción de límite inferior se expresa de manera general como:

$$X_i \geq L_i \geq 0 \quad \text{y} \quad \sum_{i=1}^q L_i < 1 \quad \text{para } i = 1, 2, \dots, q$$

donde algunas de las  $L_i$  podría ser igual a cero. Si se resta el limite inferior  $L_i$  de  $X_i$  y se divide la diferencia por  $1 - (\sum L_i)$ , el L-pseudocomponente  $X'_i$  se definirá usando la transformación como:

$$X'_i = \frac{X_i - L_i}{1 - L} \quad \dots \dots \dots (29)$$

$$\text{Donde: } L = \sum_{i=1}^q L_i < 1$$

A forma de ilustración, estableciendo los pseudocomponentes de las restricciones establecidas anteriormente se obtendrá:

$$x'_1 = \frac{x_1 - 0.35}{0.30}, \quad x'_2 = \frac{x_2 - 0.20}{0.30}, \quad x'_3 = \frac{x_3 - 0.15}{0.30}$$

El factor espacio de la *figura 22* es un simplex regular en los pseudocomponentes  $X'_i$ , donde  $X'_1 + X'_2 + X'_3 = 1$ . El rango de cada  $X'_i$  en los términos del simplex original es  $1-L$ .

La orientación del L-pseudocomponente simplex es la misma orientación de los componentes del simplex original

Las coordenadas  $X'_i=1, X'_j=0, j \neq i$  en los componentes originales son:

$$\begin{aligned} (X'_1, X'_2, X'_3) = (1,0,0) &= (X_1, X_2, X_3) = (1 - L_2 - L_3, L_2, L_3) \\ &= (0,1,0) = (L_1, 1 - L_1 - L_3, L_3) \\ &= (0,1,0) = (L_1, L_2, 1 - L_1 - L_2) \end{aligned}$$

La construcción del diseño en L-pseudocomponentes se ilustra a continuación, en donde se elige un modelo polinomial de segundo grado para modelar el área sobre la región  $X'_i$ , así como un simplex  $\{3,2\}$  en red, el diseño  $X'_1=1, 1/2, 0$  se muestra en la *tabla 7*.

L-Pseudocomponentes			Componentes Originales		
$X'_1$	$X'_2$	$X'_3$	$X_1$	$X_2$	$X_3$
1	0	0	0.65	0.20	0.15
0.5	0.5	0	0.50	0.35	0.15
0	1	0	0.35	0.50	0.15
0	0.5	0.5	0.35	0.35	0.30
0	0	1	0.35	0.20	0.45
0.5	0	0.5	0.50	0.20	0.30

TABLA 7. L-PSEUDOCOMPONENTES Y COMPONENTES ORIGINALES PARA UNA MEZCLA DE 3 COMPONENTES CON LÍMITES INFERIORES

Los  $X_i$  correspondientes a los componentes originales de la red se obtienen mediante la ecuación:

$$X_i = L_i + (1-L)X'_i \quad \dots \quad (30)$$

Así que, para  $i = 1, L_1 = 0.35$  y  $L = 0.70$ , se tiene que evaluar  $X_1$  correspondiente a  $X'_1 = 1.0, X_1 = 0.35 + (0.30)(0.50) = 0.65$ ; calculando de la misma manera  $X'_1 = 0.5$ , se obtiene que  $X_1 = 0.5$ . Se debe notar que el rango de valores para cada  $X_i$  es de  $L-1$ , esto es;  $X_i$  va de  $L_i$  a  $L_i + (1-L)$  como  $X'_i$  va de cero a uno.

Una vez que las combinaciones de la mezcla son definidas en el sistema original a partir de los L-pseudocomponentes, el siguiente paso es coleccionar los valores de respuesta observados en el diseño obtenido para generar un modelo en términos de los L-pseudocomponentes, y a partir de éste es posible obtener un modelo en términos de componentes originales.

El modelo polinomial de segundo grado en los L-pseudocomponentes es:

$$\eta = \gamma_1 x'_1 + \gamma_2 x'_2 + \gamma_3 x'_3 + \gamma_{12} x'_1 x'_2 + \gamma_{13} x'_1 x'_3 + \gamma_{23} x'_2 x'_3 \dots (31)$$

Donde el modelo correspondiente en los términos originales deberá ser de la forma:

$$\eta = \gamma_1 \frac{(x_1 - L_1)}{1 - L} + \gamma_2 \frac{(x_2 - L_2)}{1 - L} + \gamma_3 \frac{(x_3 - L_3)}{1 - L} + \gamma_{12} \frac{(x_1 - L_1)(x_2 - L_2)}{(1 - L)^2} + \gamma_{13} \frac{(x_1 - L_1)(x_3 - L_3)}{(1 - L)^2} + \gamma_{23} \frac{(x_2 - L_2)(x_3 - L_3)}{(1 - L)^2}$$

o

$$\eta = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 \dots (32)$$

Donde las  $\beta$ 's pueden ser expresadas en términos de las  $\gamma$ 's, como:

$$\beta_1 = \frac{\gamma_{12} L_2 (L_1 - 1) + \gamma_{13} L_3 (L_1 - 1) + \gamma_{23} L_2 L_3}{(1 - L)^2} + \frac{\gamma_1 - \sum_{i=1}^3 \gamma_i L_i}{1 - L} \dots (33)$$

$$\beta_2 = \frac{\gamma_{12} L_1 (L_2 - 1) + \gamma_{13} L_3 L_1 + \gamma_{23} L_3 (L_2 - 1)}{(1 - L)^2} + \frac{\gamma_2 - \sum_{i=1}^3 \gamma_i L_i}{1 - L} \dots (34)$$

$$\beta_3 = \frac{\gamma_{12} L_1 L_2 + \gamma_{13} L_1 (L_3 - 1) + \gamma_{23} L_2 (L_3 - L_1)}{(1 - L)^2} + \frac{\gamma_3 - \sum_{i=1}^3 \gamma_i L_i}{1 - L} \dots (35)$$

$$\beta_{ij} = \frac{\gamma_{ij}}{(1 - L)^2}; \quad i, j = 1, 2, 3. \quad i < j$$

Para los modelos de los L-pseudocomponentes las interpretaciones de los parámetros  $\gamma_i$  y  $\gamma_{ij}$ ;  $i, j = 1, 2$  y  $3$ ,  $i < j$ , se realiza de la misma manera que se realiza en los componentes originales. El  $\gamma_{ij}$ ;  $i = 1, 2$  y  $3$  representan la altura de los vértices  $X'_i$  en la subregión triangular; y los  $\gamma_{ij}$ ,  $i < j$ , representa desviaciones de la superficie plana en las combinaciones binarias  $X'_i X'_j$ .

## 2. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE SUPERIOR<sup>(9)</sup>

### a) INTRODUCCIÓN DE LOS U-PSEUDOCOMPONENTES

Cuando dos o más componentes están restringidos por límites superiores,  $X_i \leq U_i$ , se sugiere trabajar con U-pseudocomponentes, definidos como:

$$U_i = \frac{U_i - X_i}{U - 1}, i = 1, 2, \dots, q \quad \dots\dots\dots(36)$$

Donde  $U = \sum_{i=1}^q U_i > 1$ . La región de los U-pseudocomponentes,  $U_i, i=1, 2, \dots, q$ , es un *simplex invertido*, el cual se llama *U-simplex*. Por ejemplo, si se tiene una mezcla de tres componentes con las siguientes restricciones:

$$X_1 \leq 0.70 = U_1, \quad X_2 \leq 0.60 = U_2, \quad X_3 \leq 0.80 = U_3$$

El U-pseudocomponente corresponderá al triángulo invertido de la *figura 23*; en donde se observa que los vértices,  $U_i = 1, i = 1, 2, \dots, q$  de los U-pseudocomponentes podrían extenderse más allá de los bordes del simplex original. Cuando esto pasa, la región experimental de mezclas posibles es común (definido como la intersección) entre el original y el simplex invertido. Esta región no es un simplex. El U-simplex es la región experimental sólo cuando se encuentra enteramente dentro de la región simplex, lo que sucede si y sólo si:

$$\sum_{i=1}^q U_i - U_{mn} \leq 1$$

Donde  $U_{mn}$  es el mínimo de los  $q$  límites superiores.

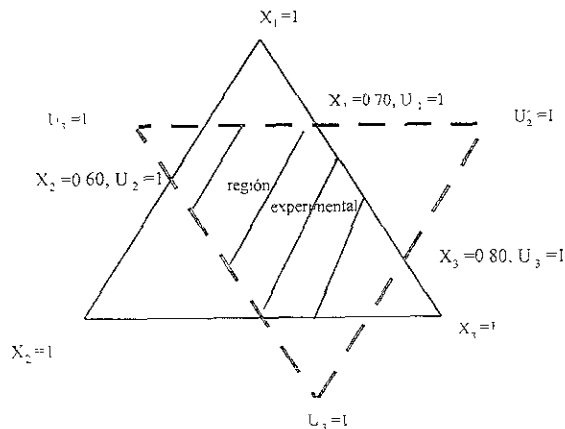


FIGURA 23. LOS VÉRTICES U-PSEUDOCOMPONENTAL PODRÁN EXTENDERSE MÁS ALLÁ DE LOS BORDES DEL SIMPLEX ORIGINAL.

Cuando el U-simplex está situado enteramente en el simplex original, el denominador.  $U-1$ , en la *Ecuación 37* para  $U_i$ , es el rango experimental de los valores de  $U_i$ . Por esa razón de conveniencia se define el tamaño de la línea del U-simplex como:

$$R_u = \sum_{i=1}^q U_i - 1 \quad \dots\dots (37)$$

Donde  $0 < R_u \leq q-1$ . Por supuesto, cuando el U-simplex se encuentra en el simplex original  $R_u \leq 1/2$ .

Los límites superiores de algunos o de todas las  $X_i$  genera límites inferiores implícitos,  $L_i^*$ , para todas las  $X_i$ , estos serán:

$$L_i^* = U_i - R_u \quad \dots\dots(38)$$

Además, cuando un vértice del U-simplex se encuentra fuera de los límites del simplex original, el resultado son límites inferiores implícitos para  $X_i$ , donde el vértice  $X_i = 1$  es opuesto al vértice distante del U-simplex, el cual es negativo porque  $L_i^* < 0$ .

Por ejemplo, los límites inferiores para  $X_1$ ,  $X_2$  y  $X_3$  de la *ecuación 38* son:

$$L_1^* = 0.70 - 1.10 = -0.40 \quad L_2^* = 0.60 - 1.10 = -0.50$$

$$L_3^* = 0.80 - 1.10 = -0.30$$

Donde  $R_u = (0.7+0.6+0.8)-1 = 2.1-1=1.1$ . Encontrándose que todas las  $L_i^*$  son negativas, lo que significa que un vértice del U-simplex (específicamente el  $U_i=1$  vértice opuesto al  $X_i = 1$  vértice) se encuentra fuera de los límites del simplex original. Como se observa en la *figura 23*.

Los límites superiores y los límites inferiores implícitos en las  $X_i$ , se definen como:

$$U_i^* = L_i + R_L \quad \dots\dots(39)$$

Donde  $R_L = 1 - \sum_{i=1}^q L_i$ , son usados para determinar si los límites son consistentes, y con ello se establece la forma (número de vértices, ejes y caras) y el tamaño de la región experimental de las mezclas posibles.

Cuando se presentan límites superiores y el espacio U-simplex se encuentre dentro de la región simplex original, se utilizan algunos de los diseños estándar en red para investigar el área de experimentación, los datos colectados de los puntos del diseño pueden ser ajustados tanto con un modelo en términos de  $X_i$  o con un modelo en términos de  $U_i$ . Sin embargo, como la orientación del U-simplex es opuesto al simplex original se deberá considerar al momento de realizar interpretaciones con los coeficientes en el modelo ajustado con  $U_i$ .

Es posible ilustrar el uso de la transformación de los *U-pseudocomponentes* mediante la *ecuación*  $X_i = U_i - R_u U_i$ . Se analiza el siguiente ejemplo con una mezcla de tres componentes, donde

$$X'_1 \leq 0.4, X'_2 \leq 0.6 \text{ y } X'_3 \leq 0.3$$

Dado que  $(U_1+U_2+U_3-U_{\min}) = (0.4+0.6+0.3-0.3) = 1.0 \leq 1.0$ , el U-simplex se encuentra en los límites del triángulo  $X_i$  original, como se observa en la *figura 24*.

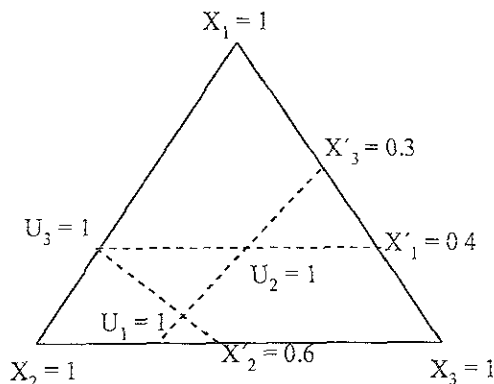


FIGURA 24 EL U-SIMPLEX DEFINIDO POR  $X_1 \leq 0.4, X_2 \leq 0.6$  Y  $X_3 \leq 0.3$

Los U-pseudocomponentes, usando el denominador  $R_u = U-1 = 0.3$  son:

$$U_1 = \frac{0.4 - X_1}{0.3}, U_2 = \frac{0.6 - X_2}{0.3}, U_3 = \frac{0.3 - X_3}{0.3}$$

Suponiendo un diseño simplex centroide con siete puntos; Los puntos en  $X_1, X_2$  y  $X_3$ , correspondientes a los siete puntos en  $U_1, U_2$  y  $U_3$  se listan en la *tabla 8*. Los valores de los  $X_i$  se obtuvieron de:

$$X_1 = 0.4 - 0.3 U_1, X_2 = 0.6 - 0.3 U_2, X_3 = 0.3 - 0.3 U_3$$

U-PSEUDOCOMPONENTES			COMPONENTES ORIGINALES		
$U_1$	$U_2$	$U_3$	$X_1$	$X_2$	$X_3$
1	0	0	0.10	0.60	0.30
0	1	0	0.40	0.30	0.30
0	0	1	0.40	0.60	0
1/2	1/2	0	0.25	0.45	0.30
1/2	0	1/2	0.25	0.60	0.15
0	1/2	1/2	0.40	0.45	0.15
1/3	1/3	1/3	0.30	0.50	0.20

TABLA 8 VALORES DE  $X_1, X_2, X_3$  CORRESPONDIENTES A LOS PUNTOS  $U_1, U_2$  Y  $U_3$  PARA UNA MEZCLA DE TRES COMPONENTES DONDE  $X_1 \leq 0.4, X_2 \leq 0.6$  Y  $X_3 \leq 0.3$  EN UN DISEÑO SIMPLEX CON 7 PUNTOS

### 3. RESTRICCIONES CUANDO TODOS O ALGUNOS DE LOS COMPONENTES PRESENTAN UN LÍMITE SUPERIOR Y UN LÍMITE INFERIOR<sup>(9)</sup>

#### a) INTRODUCCIÓN A LOS DISEÑOS DE VÉRTICES EXTREMOS.

Existen problemas de mezclas en que sus componentes están acotados superior e inferiormente, para cada  $X_i$ , lo que origina la siguiente restricción  $0 \leq L_i \leq X_i \leq U_i \leq 1$ . Por ejemplo cuando es necesario en una formulación de tabletas adicionar un lubricante, pero un exceso del mismo ocasionaría problemas. En esta situación, la región experimental es irregular y los pseudocomponentes no se pueden utilizar para definir el diseño. Una alternativa es utilizar los diseños de vértices de *McLean* y *Anderson*; los cuales consisten de todos los vértices de la región experimental irregular.

Cuando sólo una o dos de las proporciones de los componentes son restringidas en su valor; la forma del área resultante no es difícil de imaginar, sin embargo, si todos los componentes presentan restricciones la forma del espacio resultante tomará la forma de un hiperpoliedro convexo y frecuentemente será más complicado que un simplex.

En estas situaciones cada uno de los bordes de la región restringida será usada para establecer los puntos del diseño, los bordes particulares serán elegidos dependiendo de la forma y el grado de la ecuación que será usada para modelar el área de respuesta.

Para localizar los puntos del diseño se requiere al menos algunos de los vértices extremos de la región así como los puntos medios y centroides; estos últimos también se utilizan para calcular la falta de ajuste del modelo de segundo grado.

Existen fórmulas que son usadas para calcular el número de vértices en estos diseños, pero un requisito para poder emplearlas es que los límites superiores e inferiores sean consistentes, es decir, cuando es posible utilizar el límite inferior o superior de cada uno de los componentes. Por ejemplo en las restricciones:

$$0.0 \leq X_1 \leq 0.1$$

$$0.1 \leq X_2 \leq 0.2$$

$$0.6 \leq X_3 \leq 0.8$$

En ellas se observan que no son consistentes por que  $X_3$  no puede tomar el valor inferior de 0.6; el mínimo valor que podrá tomar es,  $X_3 = (1 - U_1 - U_2) = 0.7$

Para establecer la consistencia de las restricciones, detectar inconsistencias, ajustar dichas inconsistencias y hacer las restricciones consistentes se puede revisar el apéndice A y B respectivamente

Una vez definidas las restricciones como consistentes se utiliza la siguiente ecuación para establecer el número de bordes:

$$Nd = C \dots\dots\dots(40)$$

Y para calcular las coordenadas de los vértices extremos de la región restringida se pueden utilizar los algoritmos de *Mclean y Anderson*, o de *Snee y Marquardts*. La primera de estas técnicas aparece en el apéndice C.

Conociéndose el número de vértices, el número de bordes, así como las coordenadas correspondientes a cada vértice, se conseguirá saber el número de caras que tendrá el hiperpoliedro del área donde se estará trabajando; y dicha información será de utilidad para generar una lista de puntos candidatos a partir de los cuales se ajustará alguno de los diseños tipo Scheffé; siempre auxiliados por el criterio ya mencionado, donde se considera que los vértices extremos serán utilizados para estimar los coeficientes lineales, mientras que los puntos a mitad de los ejes son de ayuda para estimar las combinaciones binarias y finalmente si se desea ajustar un modelo cúbico especial será necesario establecer los puntos centroides de las caras.

Este capítulo presenta una serie de técnicas para trabajar con los problemas de mezclas, estos se aplican considerando primero si existe algún tipo de restricción y en segundo el grado del modelo que se desea establecer (lineal, cuadrático, etc.) en el caso de problemas sin restricciones se utilizan los diseños simplex y centroide simplex para modelar el área de respuesta, sin embargo es posible no estar interesado en conocer toda esta área de respuesta completa donde obviamente están los componentes en su forma pura y en tal situación existen dos diseños posibles; los planes de *Gammon* que evalúan la respuesta de los constituyentes al menos en combinaciones binarias, y los planes de *Lambrakis* que evalúan la respuesta de los mismos al menos en combinaciones triples. En el caso de problemas con restricciones existen tres tipos de diseños: cuando los componentes presentan límites inferiores, superiores y tanto límites superiores como inferiores, dependiendo del caso en que se esté trabajando se emplean ciertas transformaciones especiales para generar el modelo que describa el comportamiento de la respuesta, siempre considerando que para establecer las conclusiones finales es imprescindible obtener los datos en sus valores originales



## IX. INTERPRETACIÓN DE LOS MODELOS AJUSTADOS<sup>(6)</sup>

### A. ¿QUÉ HACER UNA VEZ GENERADO EL MODELO?

Si se desea analizar los datos de un experimento con mezclas; inicialmente se deberá considerar si el objetivo es obtener algún modelo matemático con el propósito de describir la forma del área de respuesta sobre el espacio simplex o si se está más interesado en determinar los roles que juegan (esto es, la medida de los efectos) de los componentes individuales; sin embargo es posible en muchas ocasiones realizar ambos objetivos con el mismo análisis.

Considerando que se ha obtenido un polinomio para modelar el área de respuesta de un diseño de mezclas, el siguiente paso será describir a través de una prueba de hipótesis, si alguno de los términos en el modelo de Scheffé son estadísticamente iguales a cero. Lo cual se realiza frecuentemente incluyendo a todos los términos de un determinado grado.

*Por ejemplo:* si se ha obtenido un modelo cúbico para una mezcla de tres componentes de la forma:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \beta_{123} X_1 X_2 X_3$$

El primer contraste de hipótesis, será:

*Hipótesis nula  $H_0$ :* la respuesta no depende de los componentes de la mezcla;  $\beta_1 = \beta_2 = \beta_3$

*Hipótesis alterna  $H_a$ :* La respuesta depende de los componentes de la mezcla

Cuando la hipótesis nula resulta ser verdadera, los tres coeficientes lineales  $\beta_1 = \beta_2 = \beta_3$ , son iguales a un valor constante (llamado  $\beta_0$ ). Lo que implica que en una mezcla de tres componentes todos los coeficientes tienen la misma altura sobre el triángulo. Como se observa en la *figura 25*.

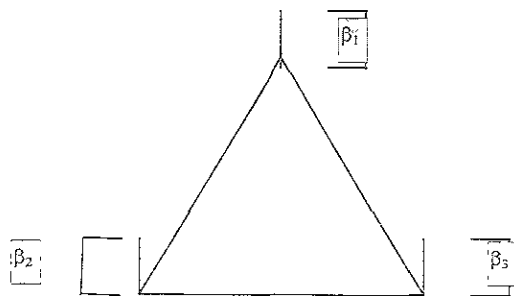


FIGURA 25. LOS TRES COEFICIENTES LINEALES  $\beta_1 = \beta_2 = \beta_3$  SON IGUALES CON UN VALOR CONSTANTE LLAMADO  $\beta_0$

La prueba para la hipótesis nula es generada mediante la relación:

$$F = \frac{SCR / (p - 1)}{SCE / (N - p)} \dots\dots (41)$$

$p$  = número de combinaciones diferentes

$N$  = número total de observaciones

Donde:

$$SCR = \sum_{U=1}^N (\hat{Y}_U - \bar{Y})^2 \dots\dots(42)$$

$$SCE = \sum_{U=1}^N (Y_U - \hat{Y}_U)^2 \dots\dots(43)$$

El valor  $F$  relacionado es comparado con el valor  $F_{(p-1, N-p, \alpha)}$  y la hipótesis nula es rechazada con un  $\alpha$  nivel de significancias si el valor de  $F$  calculado es mayor que el de tablas.

En caso de rechazar  $H_0$  y concluir que la respuesta depende de los componentes de la mezcla se realiza la determinación del coeficiente de ajuste para deducir si el modelo explica la suficiente cantidad de variación en los valores de la respuesta. La cantidad de variación es calculada mediante:

$$R_A^2 = 1 - \frac{SC_E / (N - p)}{SC_T / (N - 1)} \dots\dots(44)$$

Cuando los valores de  $R_A^2$  excede 0.9; significa que el error de la varianza estimada obtenida del análisis del modelo ajustado es menor del 10 %, y es considerado razonable.

En caso de determinar un valor razonable de  $R_A^2$ , el siguiente paso corresponderá en cuestionarse si es indispensable trabajar con un modelo cúbico o si es posible obtener datos igualmente confiables de un modelo de menor grado; se deberá preferir trabajar con modelos de menor grado porque es más fácil trabajar durante la predicción de la respuesta, y es menos complicado el comportamiento de los componentes de la mezcla.

La reducción del modelo se puede realizar de varias maneras; puede ser efectuada por una simple suma de términos o mediante un contraste entre los términos, y posteriormente usar ese contraste como los nuevos términos en el modelo, incluso es posible eliminar ciertos términos del modelo. Este último procedimiento es el más obvio a efectuarse y consiste en

remover los términos de mayor grado que no son significativos, lo cual se consigue mediante:

$$t = \frac{\hat{\beta}_y}{e.e.(\hat{\beta}_y)} \dots \dots (45)$$

Donde:

$\hat{\beta}_y$  es el valor estimado de  $\beta_y$  y e.e. es el error estándar del valor estimado de  $\hat{\beta}_y$ .

Si  $t_{cal}$  resulta ser menor que  $t$  de tablas se dice que este término no es diferente de cero. Esto puede ser corroborado con la comparación de las  $R_A^2$  que no presentará un cambio significativo al retirar el término; de igual manera es posible realizar un contraste obteniendo una  $F$  de cambio mediante la ecuación:

$$F = \frac{(SC_{R^{FC1}} - SC_{R^{FC2}}) / q}{SCE_{FC1} / (N - T)} \dots \dots (46)$$

La  $F$  de cambio se compara con tablas de  $F$  y si  $F$  de cambio es menor que el valor localizado en las tablas de  $F$  no se rechaza la hipótesis de que el coeficiente es igual a cero.

Otra manera de dilucidar posibles cambios entre los modelos es mediante la generación de gráficas de residuales donde no se deberán apreciar cambios significativos en la distribución y la distancia de los residuales con respecto a la media.

Anteriormente se mencionó que otro camino para reducir modelos es mediante la combinación de algunos de los términos en el modelo. Dicha combinación es realizada mediante la suma de las proporciones de los términos cuando sus coeficientes reflejan un tipo de combinación similar. Por ejemplo, si se tiene el modelo lineal:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Y los estimados  $\beta_1$  y  $\beta_2$  son aproximadamente iguales en magnitud y sus errores estándar son aproximadamente iguales entre sí, esto conducirá a reducir el modelo de la siguiente forma:

$$Y = \beta_1 (X_1 + X_2) + \beta_3 X_3$$

La interpretación de los coeficientes mínimos cuadrados de Scheffé es diferente a la de los métodos cuadrados normales (sin mezclas). En problemas de mezclas, si el coeficiente  $\beta_i$  no es significativo, esto no implica que el coeficiente  $X_i$  no tenga un efecto significativo en la

variable dependiente. Así mismo, si el valor de  $\beta_i$  es cero, esto implica que el valor estimado de la variable dependiente es cero en  $X_i=1.0$ .

## 1. PROCEDIMIENTO PARA ESCOGER EL MODELO CORRECTO<sup>(2)</sup>

Como se mencionó con anterioridad, si existe sinergismo o antagonismo al mezclar componentes, es común encontrar superficies de respuesta no lineales, por ello cuando se tiene un estimador de la varianza (calculada a partir de las repeticiones de los puntos del diseño) se recomienda el siguiente procedimiento para encontrar el mejor diseño.

Primero tratar de ajustar al modelo lineal:

$$Y = \sum_{i=1}^k \beta_i X_i$$

Y determinar los estimadores de  $\beta_i$ , los valores de  $R^2_A$ , la prueba de bondad de ajuste y las gráficas de residuales. Si el modelo resulta ser correcto se procederá a determinar si es posible la reducción de términos.

En caso que la prueba de falta de ajuste sea significativa y los residuales presenten alguna tendencia, tratar de ajustar el modelo a uno cuadrático y determinar las mismas pruebas de ajuste del modelo. En caso de no lograr el ajuste, intentarlo para los modelos cúbico o cúbico especial, siempre considerándose que a medida que se establezcan modelos con mayor número de términos se requerirá aumentar la cantidad de experimentos.

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i < j} \beta_{ij} X_i X_j + \varepsilon$$

$$Y = \sum_{i=1}^k \beta_i X_i + \sum_{i < j} \beta_{ij} X_i X_j + \sum_{i < j < k} \beta_{ijk} X_i X_j X_k$$

Una vez que el usuario estableció el tipo de diseño de acuerdo al problema que esté trabajando (con o sin restricciones), y se generó una matriz de experimentación correspondiente al diseño elegido generándose un modelo de descripción de la respuesta, es importante saber interpretar dicho modelo; es por ello que en este capítulo se presenta como hacerlo, como primer paso se deberá efectuar una prueba de hipótesis para saber si la respuesta depende o no de los componentes de la mezcla, concluyéndose la dependencia de la respuesta por los componentes se procederá a dilucidar si el modelo explica la suficiente cantidad de variación en los valores de la respuesta mediante el cálculo de  $R^2_A$  y si éste excede el 0.9 significa que el modelo es adecuado. El siguiente paso será entonces pensar en la posibilidad de disminuir el grado del modelo o el número de términos del mismo, porque es más fácil trabajar y entender el comportamiento de los componentes durante la predicción de la mejor respuesta con modelos más sencillos.

## X. VARIABLES DE PROCESO INCLUIDAS EN EXPERIMENTOS CON MEZCLAS<sup>(2)(4)</sup>

En estos experimentos el tipo estándar del diseño es el arreglo en red (un  $\{q, m\}$  diseño en red o centroide simplex) determinando los puntos del arreglo factorial en cada una de las variables del proceso. Por ejemplo, para tres componentes se obtienen las proporciones  $X_1$ ,  $X_2$  y  $X_3$ ; suponiendo que también existen 2 variables del proceso denotadas por  $Z_1$  y  $Z_2$ , y cada variable del proceso a dos niveles

Si se ajusta un modelo lineal de 3 términos con los datos colectados de los puntos en  $\{3, 2\}$  red simplex, podemos realizar la combinación  $\{3, 2\} \times 2^2$  arreglo factorial. Los experimentos se especifican en la *tabla 9*, donde el número total de experimentos en el diseño de las variables del proceso (esto es, 4) y el número correspondiente en el diseño con mezclas (esto es, 4).

Conviene señalar que cualquier diseño con mezclas se puede combinar con cualquier diseño de variables del proceso, repitiendo todos los experimentos del primer tipo de diseño (esto es, con mezclas) en cada una de las combinaciones del diseño de variables del proceso.

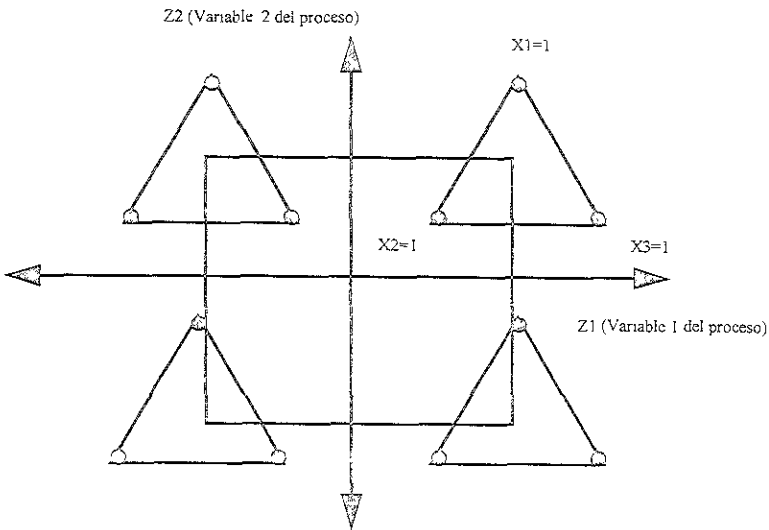


FIGURA 26. DISEÑO DE MEZCLAS CON TRES COMPONENTES CON DOS VARIABLES DEL PROCESO

Los experimentos a efectuarse serían los siguientes:

COMPONENTES DE LA MEZCLA				VARIABLES DE PROCESO	
Número del experimento	$X_1$	$X_2$	$X_3$	$Z_1$	$Z_2$
1	1	0	0	-	-
2	0	1	0	-	-
3	0	0	1	-	-
4	1/3	1/3	1/3	-	-
5	1	0	0	+	-
6	0	1	0	+	-
7	0	0	1	+	-
8	1/3	1/3	1/3	+	-
9	1	0	0	-	+
10	0	1	0	-	+
11	0	0	1	-	+
12	1/3	1/3	1/3	-	+
13	1	0	0	+	+
14	0	1	0	+	+
15	0	0	1	+	+
16	1/3	1/3	1/3	+	+

TABLA 9. CANTIDADES CODIFICADAS DE LOS 16 EXPERIMENTOS PARA UN DISEÑO DE MEZCLAS CON TRES COMPONENTES COMBINADO CON UN DISEÑO DE DOS VARIABLES DEL PROCESO

Al multiplicar ambos modelos se obtiene el siguiente polinomio de respuesta:

$$Y = \beta_1\gamma_0X_1 + \beta_1\gamma_1Z_1X_1 + \beta_1\gamma_2Z_2X_1 + \beta_1\gamma_{12}Z_1Z_2X_1 + \beta_2\gamma_0X_2 + \beta_2\gamma_1Z_1X_2 + \beta_2\gamma_2Z_2X_2 + \beta_2\gamma_{12}Z_1Z_2X_2 + \beta_3\gamma_0X_3 + \beta_3\gamma_1Z_1X_3 + \beta_3\gamma_2Z_2X_3 + \beta_3\gamma_{12}Z_1Z_2X_3$$

En este capítulo se ofrece una alternativa de uso de los diseños de mezclas y consiste en combinarlo con variables de proceso; esto se realiza incluyendo a cada proporción de los componentes establecida cada uno de los niveles de la variable de estudio, y así mismo incluyendo cada una de estas combinaciones en el modelo generado. Con ello se conseguirá establecer la mejor respuesta de alguna proporción de los componentes de la mezcla a una variable de proceso.

# XI. EJEMPLOS DE APLICACIONES DE LOS DISEÑOS DE MEZCLAS EN FARMACIA

## A. DISEÑO DE MEZCLAS PARA UNA FORMULACIÓN DE LIBERACIÓN PROLONGADA DE TEOFILINA.

### 1. PLANTEAMIENTO DEL PROBLEMA

EN ESTE EJEMPLO SE UTILIZA UN DISEÑO DE VÉRTICES (FIGURA 27) CON TRES DIFERENTES DERIVADOS DE CELULOSA (HIDROXIPROPIL CELULOSA (HPC), HIDROXIPROPILMETIL CELULOSA (HPMC) Y CELULOSA MICROCRISTALINA (MC) PARA PREPARAR TABLETAS QUE CONTENÍAN 10 % DE TEOFILINA MEDIANTE UNA GRANULACIÓN HÚMEDA, DETERMINÁNDOSE EL EFECTO DE CADA DERIVADO DE CELULOSA SOBRE LA PROLONGACIÓN DE LA LIBERACIÓN DE TEOFILINA, PARA ELLO SE UTILIZÓ COMO VARIABLES DE RESPUESTA EL CÁLCULO DE LA CONSTANTE DE LA VELOCIDAD DE LIBERACIÓN (MÍNIMA  $k$ ) ASÍ COMO EL TIEMPO MEDIO DE LIBERACIÓN (MÁXIMO MTD) DE LAS FORMULACIONES PRELIMINARES OBSERVADAS EN LA TABLA 10, DONDE SE DILUCIDA QUE LOS MEJORES RESULTADOS SE LOCALIZAN DENTRO DEL ÁREA EXPERIMENTAL INICIAL (FIGURA 27) QUE CORRESPONDIERON A LAS FORMULACIONES 3, 4 Y 5 POR CONSIGUIENTE SE ESTABLECIÓ UN NUEVO DOMINIO EXPERIMENTAL (FIGURA 28) CON UN CENTROIDE SIMPLEX DE 10 FORMULACIONES QUE SE INCLUYEN EN LA TABLA 12; CON ESTOS RESULTADOS SE CONSTRUYÓ UN MODELO DE SEGUNDO GRADO MEDIANTE EL PAQUETE SAS FOR WINDOWS (LA FORMA DE INTRODUCIR LOS DATOS EN ESTE PAQUETE SE MUESTRA EN EL APENDICE D) Y SE OBSERVÓ QUE LOS MEJORES RESULTADOS DE LAS VARIABLES DE RESPUESTA FUERON CONSEGUIDAS POR LAS FORMULACIONES B Y D.

Los experimentos preliminares se efectuaron con los siguientes límites:

$$0.1 \leq X_1 \leq 0.2$$

$$0.1 \leq X_2 \leq 0.2$$

$$0.7 \leq X_3 \leq 0.9$$

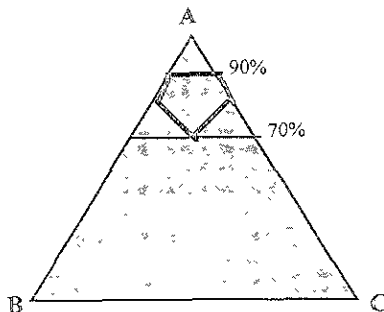


FIG. 27 DISEÑO DE VÉRTICES CON TRES DIFERENTES DERIVADOS DE CELULOSA (HPC) (HPMC) Y (MC)

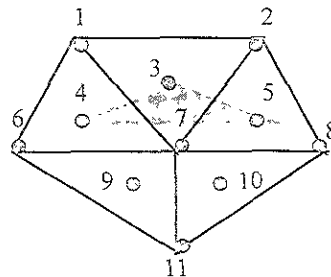


FIGURA 28 NUEVO DOMINIO EXPERIMENTAL (CON UN CENTROIDE SIMPLEX)

NÚMERO DE FORMULACIÓN	MC (X <sub>3</sub> )	HPMC (X <sub>2</sub> )	HPC (X <sub>1</sub> )	MDT MINUTOS	K (MIN <sup>-1</sup> )
1	90	10	0	28.2	0.053
2	90	0	10	16.3	0.591
3	87	6.5	6.5	26.7	0.558
4	83	13.5	3.5	54.2	0.019
5	83	3.5	13.5	43.0	0.024
6	80	20	0	8.4	0.014
7	80	10	10	17.1	0.078
8	80	0	20	10.5	0.138
9	77	15	8	9.7	0.049
10	77	8	15	18.5	0.063
11	70	15	15	18.2	0.048

TABLA 10 DISEÑO DE VERTICES CON TRES DIFERENTES DERIVADOS DE CELULOSA (HIDROXIPROPIL CELULOSA (HAPC), HIDROXIPROPILMETIL CELULOSA (HPMC) Y CELULOSA MICROCRISTALINA (MC) PROPORCIONES PRELIMINARES DE LAS FORMULACIONES Y SUS CORRESPONDIENTES RESULTADOS

Las 11 formulaciones iniciales se presentan en la *tabla 10* donde se observa que las formulaciones 4, 5 y 6 presentan los mejores resultados de tiempo medio de liberación, así como de la constante de liberación, por consiguiente se plantea un nuevo dominio experimental utilizando el triángulo que forman las formulaciones antes mencionadas (*figura 28*) y aplicando un diseño centroide simplex  $\{3, 2\}$  es decir con tres componentes se buscará establecer un polinomio de segundo orden por lo que será necesario tener 10 formulaciones (*tabla 11 y 12*). Para establecer las proporciones de los excipientes en sus porcentajes originales se emplea la *ecuación 30*, por ejemplo para proporción  $X_1 = 1/3$ ,  $X_2 = 1/3$ ,  $X_3 = 1/3$ ,

$$\begin{aligned} \text{los nuevos límites corresponden a :} & \quad 3.5 \leq X_1 \leq 13.5 \\ & \quad 3.5 \leq X_2 \leq 13.5 \\ & \quad 73.0 \leq X_3 \leq 93.0 \end{aligned}$$

los porcentajes originales serán:

$$X_1 = L_1 + (1-L) X'_1 \quad X_3 = (0.83 + (1-0.9)1/3) * 100$$

$$X_3 = \underline{86.33}, \quad X_2 = \underline{6.835} \quad \text{y} \quad X_1 = \underline{6.835}$$

FORMULACIÓN	X <sub>3</sub>	X <sub>2</sub>	X <sub>1</sub>	X <sub>3</sub> %	X <sub>2</sub> %	X <sub>1</sub> %
3	1	0	0	93	3.5	3.5
4	0	1	0	83	13.5	3.5
5	0	0	1	83	3.5	13.5
G	1/2	1/2	0	88	8.5	3.5
C	1/2	0	1/2	88	3.5	8.5
A	0	1/2	1/2	83	8.5	8.5
B	1/3	1/3	1/3	86.33	6.835	6.835
E	2/3	1/6	1/6	89.66	5.166	5.166
F	1/6	2/3	1/6	84.66	10.166	5.166
D	1/6	1/6	2/3	84.66	5.166	10.166

TABLA 11 CENTROIDE SIMPLEX DE 10 FORMULACIONES DONDE SE INCLUYEN LAS PROPORCIONES DE LOS TRES DIFERENTES COMPONENTES Y SUS CORRESPONDIENTES VALORES ORIGINALES



FORMULACIÓN	X <sub>3</sub> %	X <sub>2</sub> %	X <sub>1</sub> %	MTD EN MINUTOS
3	100	0	0	22.40
4	0	100	0	54.20
5	0	0	100	43.00
G	50	50	0	51.50
C	50	0	50	127.50
A	0	50	50	77.08
B	33.33	33.33	33.33	144.60
E	66.67	16.67	16.67	80.00
F	16.67	66.67	16.67	86.05
D	16.67	16.67	66.67	102.10

TABLA 12 CENTROIDE SIMPLEX DE 10 FORMULACIONES DONDE SE INCLUYEN LAS PROPORCIONES Y RESULTADOS DE LA SEGUNDA AREA RESTRINGIDA

## 2. SALIDAS DEL SAS PARA WINDOWS

SOURCE	DF	SUM SQUARES	MEANS SQUARES	F VALUE	PROB>F
MODEL	6	73903.02	12317.17	35.80	0.002
ERROR	4	1376.26	344.06		
TOTAL	10	75279.29			

TABLA 13 ANALISIS DE VARIANZA

ROOT MSE	18.55	R-SQUARE	0.98
DEP MEAN	78.84	ADJ R-SQ	0.95
C.V.	23.53		

TABLA 14 CALCULO DE LA CANTIDAD DE VARIACION CALCULADA MEDIANTE ( $R^2$ )

### Parameter Estimates

VARIABLE	DF	ESTIMATE	ERROR	PARAMETER	PROB> T
X <sub>1</sub>	1	17.22	17.88	0.963	0.390
X <sub>2</sub>	1	52.38	17.88	2.929	0.043
X <sub>3</sub>	1	39.67	17.88	2.218	0.091
X <sub>1</sub> X <sub>2</sub>	1	102.45	82.48	1.242	0.282
X <sub>1</sub> X <sub>3</sub>	1	425.96	82.50	5.163	0.006
X <sub>2</sub> X <sub>3</sub>	1	167.28	82.48	2.028	0.113

TABLA 15 CALCULO DE LOS PARAMETROS DEL MODELO DE SEGUNDO GRADO

### Ridge of Optimum

Type of ridge=MAXIMUM

Coding Coefficients for the Independent Variables

X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	PRED	STDERR
0.32	0.28	0.39	117.99	9.30
0.33	0.33	0.33	113.57	9.25
0.33	0.23	0.43	121.43	9.52
0.35	0.18	0.45	120.00	9.79
0.36	0.19	0.46	124.44	10.00
0.38	0.14	0.48	127.24	10.83
0.40	0.09	0.50	129.88	12.05
0.43	0.05	0.52	132.37	13.63
0.45	0.004	0.54	134.76	15.57
0.48	0.047	0.511	72.16	13.64

TABLA 16. CALCULO DE MAXIMO VALOR DE MTD EN SUS VARIABLES CODIFICADAS

### 3. CÁLCULO DE LAS PROPORCIONES DE LOS COMPONENTES EN SUS VALORES ORIGINALES.

Para transformar las proporciones de los L-pseudocomponentes correspondientes al máximo valor de MTD ( $X_1=0.45$ ,  $X_2=0.004$ ,  $X_3=0.54$  y  $MTD=134.76$ ) de la *tabla 15* a los valores originales se utiliza la *ecuación 30*; sustituyendo los datos correspondientes de la proporción de cada componente en dicha ecuación se obtiene:

$$X_3 = 0.83 + (1-0.9) 0.45$$

$$X_3 = 0.875 \text{ o } 87.5 \%$$

$$X_2 = 0.035 + (1-0.9) 0.004$$

$$X_2 = 0.0354 \text{ o } 3.54 \%$$

$$X_1 = 0.035 + (1-0.9) 0.54$$

$$X_1 = 0.089 \text{ o } 8.9 \%$$

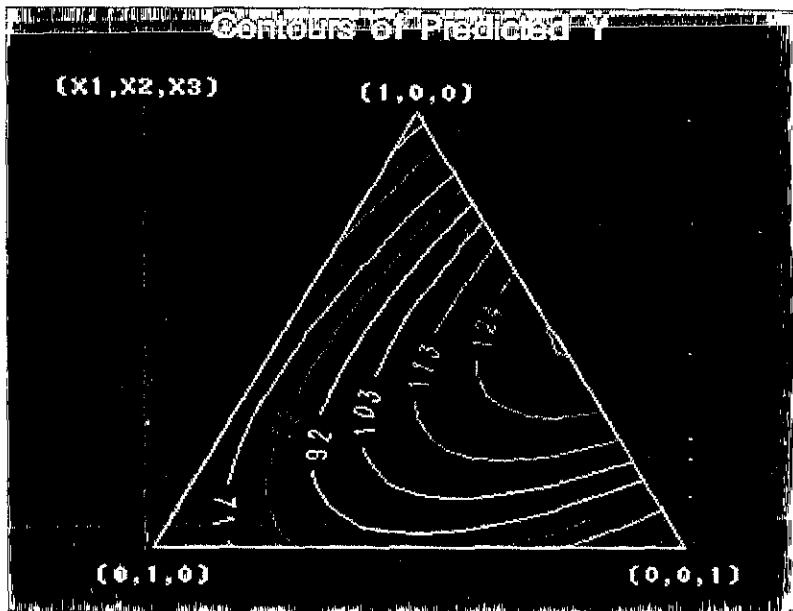


FIGURA 29 GRÁFICA DE CONTORNOS DE LA SALIDA DE SAS PARA WINDOWS

#### d) INTERPRETACIÓN DE LA SALIDAS EN SAS PARA WINDOWS

En la *tabla 12* se presenta el análisis de varianza de los resultados correspondientes a la segunda área restringida, donde se hace patente la gran relación que existe entre los constituyentes de la mezcla y la variable de respuesta debido a que el valor de significancia de los datos es de 0.002; así también la cantidad de variación explicada en la *tabla 13* es de 95 %, por lo que se considera que el modelo cuadrático que se genera a partir de los valores estimados de la *tabla 15* describe de manera apropiada el comportamiento de los datos. Por consiguiente es posible determinar la formulación que genera el valor óptimo de MTD; mediante una grafica de contorno o mediante métodos numéricos como el análisis RIDGE, es importante recordar que las proporciones óptimas de la *tabla 16* corresponden a los L-pseudocomponentes y es necesario transformarlas a las proporciones de los componentes originales, la cual resultó ser 87.5 % HPC, 3.54% de HPMC y 8.9% de MC, proporción que es concordante con el valor que se observa en la gráfica de contorno de la *figura 29*

Como se observa a través de este ejemplo los diseños de mezclas pueden utilizarse primero como una fase exploratoria para tener una aproximación a una zona de respuesta con los mejores resultados y a partir de ella establecer otro diseño que dilucide de forma más precisa la formulación con mejores características, obteniendo con ello un menor número de ensayos y consiguiéndose de una manera sencilla y rápida el resultado deseado

## B. OPTIMIZACIÓN DE UNA FORMULACIÓN DE TABLETAS USANDO UN DISEÑO DE MEZCLAS

### 1. PLANTEAMIENTO DEL PROBLEMA.

EN ESTE EJEMPLO SE REALIZA LA OPTIMIZACIÓN DEL PERFIL DE DISOLUCIÓN DE UNA TABLETA DE LIBERACIÓN CONTROLADA DE NAFTIDROFURIL UTILIZANDO UN DISEÑO CENTROIDE SIMPLEX. LOS EXCIPIENTES INVESTIGADOS FUERON AVICEL PH 102 (CELULOSA MICROCRISTALINA), TABLETOSE EP (LACTOSA) Y FOSFATO DE CALCIO DIHIDRATADO. REALIZÁNDOSE COMO PRIMER PASO UNA GRANULACIÓN CON EL PRINCIPIO ACTIVO CON GOMA XANTANA, GOMA GUAR, LACTOSA D80 Y UNA SOLUCIÓN DE PVP AL 5.0%. LOS PARÁMETROS DE RESPUESTA FUERÓN: VELOCIDAD DE LIBERACIÓN, RESISTENCIA A LA RUPTURA, FRIABILIDAD Y VARIACIÓN DE PESO OBTENIÉNDOSE COMO RESULTADO UN MODELO MATEMÁTICO DE PREDICCIÓN DE RESPUESTA CON EL MODELO DE SCHÉFFE Y UNA GRÁFICA DE CONTORNO LO QUE AYUDÓ A ENCONTRAR LA FORMULACIÓN CON LAS MEJORES CARACTERÍSTICAS.

FORMULA	AVICEL PH 102 $X_1$	TABLETOSE EP $X_2$	FOSFATO DE CALCIO DIHIDRATADO $X_3$	T90% EN BUFFER (MIN)	RESISTENCIA A LA RUPTURA (N)	FRIABILIDAD %	VARIACION DE PESO (%)
1	1	0	0	197	127	0.65	0.67
2	0	1	0	110	63	1.18	1.04
3	0	0	1	324	83	0.77	1.82
4	1/2	1/2	0	67	83	0.75	0.75
5	1/2	0	1/2	362	88	0.70	0.47
6	0	1/2	1/2	312	64	0.91	1.67
7	1/3	1/3	1/3	214	69	0.74	0.77
8	2/3	1/6	1/6	206	101	0.53	0.35
9	1/6	2/3	1/6	171	59	0.89	0.87
10	1/6	1/6	2/3	344	74	0.77	0.88

TABLA 17 PROPORCIONES DE LAS MEZCLAS Y RESULTADOS DE LAS VARIABLES DE RESPUESTA

### 2. RESULTADOS OBTENIDOS EN SAS FOR WINDOWS PARA T90% EN BUFFER

*General Linear Models Procedure*  
 Number of observations in data set = 10  
 The SAS System

General Linear Models Procedure

Dependent Variable: Y  
Sum of Mean

Source	DF	Sum Squares	Means Squares	F Value	Pr>F
Model	6	623175.961	103862.660	297.81	0.0001
Error	4	1395.039	348.760		
Total	10	624571.000			

TABLA 18 ANALISIS DE VARIANZA PARA 190% EN BUFFER

Root MSE	18.67511	R-square	0.998
Dep. Mean	230.7000	Adj. R-sq	0.944
C.V.	8.0949		

TABLA 19 CALCULO DE LA CANTIDAD DE VARIACION CALCULADA MEDIANTE  $(R_s^2)$

Parameter	Estimate	T for H0: Parameter=0	Pr >  T	Std Error of Estimate
X <sub>1</sub>	198.139	0.0004	18.010	
X <sub>2</sub>	114.049	6.33	0.003	18.006
X <sub>3</sub>	328.991	18.27	0.001	18.011
X <sub>1</sub> *X <sub>2</sub>	-402.395	-4.85	0.008	83.043
X <sub>1</sub> *X <sub>3</sub>	351.491	4.23	0.013	83.070
X <sub>2</sub> *X <sub>3</sub>	331.313	3.99	0.016	83.046

TABLA 20 CALCULO DE LOS PARAMETROS DEL MODELO DE SEGUNDO GRADO PARA 190% EN BUFFER

Ridge of Optimum Response  
Type of ridge=MAXIMUM

PROPORCIÓN	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	_PRED_	_STDERR_
1	0.333	0.333	0.334	244.515	9.401
2	0.303	0.294	0.404	272.272	9.451
3	0.276	0.252	0.473	295.758	9.507
4	0.255	0.205	0.540	315.123	9.550
5	0.248	0.152	0.600	330.771	9.837
6	0.265	0.093	0.643	343.724	10.965
7	0.298	0.035	0.667	355.513	13.100

TABLA 21 CALCULO DEL MAXIMO VALOR DE 190% EN BUFFER

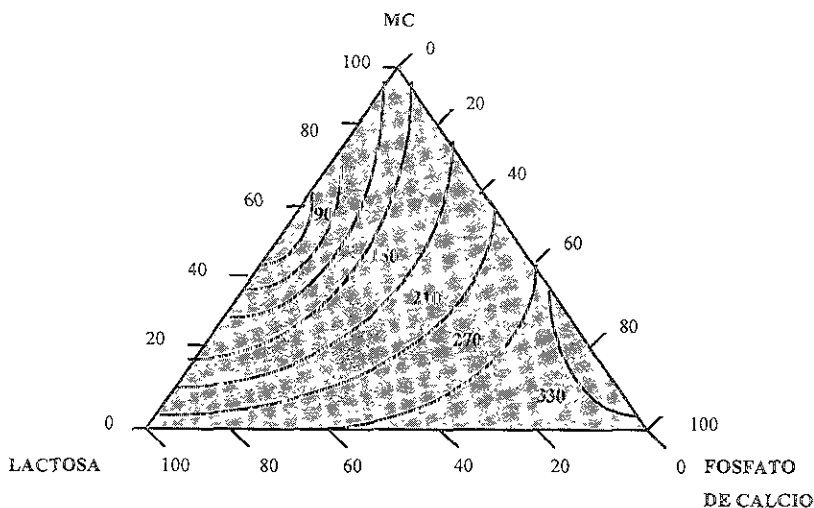


FIG 30 SUPERFICIE DE CONTORNO DEL TIEMPO CUANDO 90% DE NAFTIDROFURIL ES LIBERADO EN BUFFER DE FOSFATOS

### a) INTERPRETACION DE RESULTADOS DE T 90 % EN BUFFER

En la tabla 18 se presenta el análisis de varianza de los resultados correspondientes a los valores de  $T_{90\%}$  en buffer, donde se manifiesta la gran relación que existe entre los constituyentes de la mezcla y la disolución debido a que el valor de significancia de los datos es de 0.001, así también la cantidad de variación explicada en la tabla 13 es de 95 %, por lo que se considera que el modelo cuadrático que se genera a partir de los valores estimados de la tabla 15 describe de manera apropiada el comportamiento de los datos de disolución; el máximo valor calculado para este parámetro corresponderá a la proporción 7 de la tabla 21 y es de 355 minutos. Se observa también en la gráfica de contornos figura 30 que a medida que se incremente la cantidad de fosfato de calcio la disolución se verá retardada.

### 3. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA VARIACIÓN DE PESO.

#### Analysis of Variance

Source	DF	Squares	Square	F Value	Prob>F
Model	6	10.542	1.757	58.568	0.001
Error	4	0.120	0.030		
U Total	10	10.662			

TABLA 22 ANALISIS DE VARIANZA PARA VARIACION DE PESO

Root MSE	0.1732	R-square	0.989
Dep Mean	0.929	Adj R-sq	0.972
C.V	18.648		

TABLA 23 CALCULO DE LA CANTIDAD DE VARIACION CALCULADA MEDIANTE  $(R^2)$

Parameter	Estimate	T for H <sub>0</sub> : Parameter=0	Pr >  T	Std Error of Estimate
X <sub>1</sub>	0.680	4.07	0.015	0.167
X <sub>2</sub>	1.049	6.28	0.003	0.167
X <sub>3</sub>	1.786	10.69	0.004	0.167
X <sub>1</sub> *X <sub>2</sub>	-0.834	-1.08	0.340	0.770
X <sub>1</sub> *X <sub>3</sub>	-3.599	-4.67	0.010	0.770
X <sub>2</sub> *X <sub>3</sub>	0.458	0.59	0.584	0.770

TABLA 24 CALCULO DE LOS PARAMETROS DEL MODELO DE SEGUNDO GRADO PARA VARIACION DE PESO

Ridge of Optimum Response  
Type of ridge=MINIMUM

PROPORCIÓN	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	PRED	STDERR
1	0.333	0.333	0.336	0.730	0.088
2	0.381	0.321	0.297	0.660	0.087
3	0.425	0.293	0.283	0.602	0.088
4	0.458	0.250	0.291	0.551	0.088
5	0.484	0.202	0.314	0.502	0.091
6	0.506	0.152	0.342	0.451	0.098
7	0.525	0.102	0.373	0.398	0.109
8	0.542	0.052	0.406	0.343	0.126
9	0.558	0.002	0.440	0.284	0.147

TABLA 25 CALCULO DEL MINIMO VALOR PARA VARIACION DE PESO

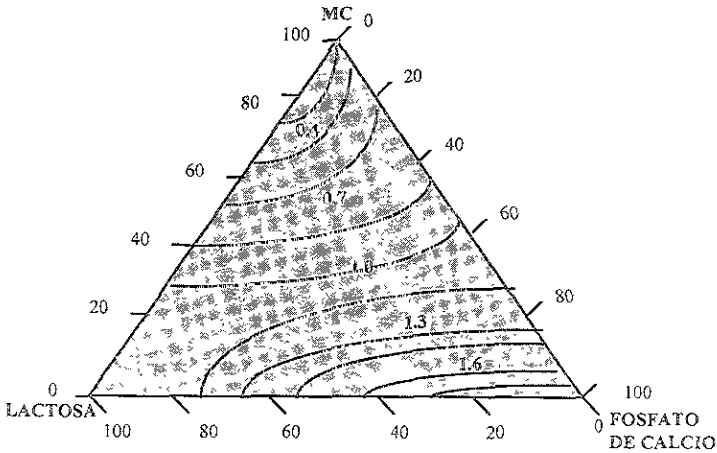


FIG 31 SUPERFICIE DE CONTORNO PARA LOS RESULTADOS DE VARIACION DE PESO

a) INTERPRETACIÓN DE LOS RESULTADOS CORRESPONDIENTES A VARIACIÓN

## DE PESO.

En la *tabla 22* se hace patente la relación que existe entre las proporciones de los constituyentes de la mezcla y la variación de peso de las tabletas mediante un análisis de varianza porque el valor de significancia de los datos es de 0.001, y se considera que el modelo cuadrático que se genera a partir de los valores estimados de la *tabla 25* describe de manera apropiada el comportamiento de los datos de variación de peso porque la variación explicada que se observa en la *tabla 24* es de 97 %, así mismo para este parámetro el valor mínimo es de 0.28 % y corresponde a la proporción 9 de la *tabla 26*. En la grafica de contornos de la *figura 31* se observa que a medida que se incrementen las proporciones de celulosa microcristalina y fosfato de calcio se disminuirá la variación de peso de las tabletas.

### 4. RESULTADOS OBTENIDOS EN SAS PARA WINDOWS PARA RESISTENCIA A LA RUPTURA.

*The SAS System*  
*General Linear Models Procedure*  
 Dependent Variable: Y  
 Sum of Mean

Source	DF	Squares	Square	F Value	Pr > F
Model	6	70289.456	11714.909	423.90	0.0001
Error	4	110.544	27.636		
Uncorrected Total	10	70400.000			

TABLA 26 ANALISIS DE VARIANZA PARA RESISTENCIA A LA RUPTURA

Root MSE	5.257	R-square	0.998
Dep Mean	81.600	Adj. R-sq	0.996
C.V.	6.442		

TABLA 27 CALCULO DE LA CANTIDAD DE VARIACION CALCULADA MEDIANTE (R<sup>2</sup>)

Parameter	Estimate	T for H0: Parameter=0	Pr >  T	Std Error of Estimate
X <sub>1</sub>	128.731	25.39	0.001	5.070
X <sub>2</sub>	61.283	12.09	0.003	5.069
X <sub>3</sub>	84.477	16.66	0.001	5.100
X <sub>1</sub> *X <sub>2</sub>	-53.098	-2.27	0.086	23.377
X <sub>1</sub> *X <sub>3</sub>	-66.832	-2.86	0.046	23.383
X <sub>2</sub> *X <sub>3</sub>	-41.606	-1.78	0.150	23.377

TABLA 28 CALCULO DE LOS PARAMETROS DEL MODELO DE SEGUNDO GRADO PARA VARIACION DE PESO



*Ridge of Optimum Response*

PROPORCIÓN	$X_1$	$X_2$	$X_3$	PRED	STDERR
1	0.333	0.333	0.334	73.499	2.629
2	0.380	0.315	0.305	75.834	2.638
3	0.429	0.304	0.267	78.404	2.658
4	0.479	0.300	0.221	81.243	2.703
5	0.531	0.302	0.170	84.381	2.828
6	0.582	0.310	0.108	87.849	3.120
7	<b>0.633</b>	<b>0.323</b>	<b>0.044</b>	<b>91.669</b>	<b>3.661</b>

TABLA 29 CALCULO DEL MÁXIMO VALOR PARA RESISTENCIA A LA RUPTURA

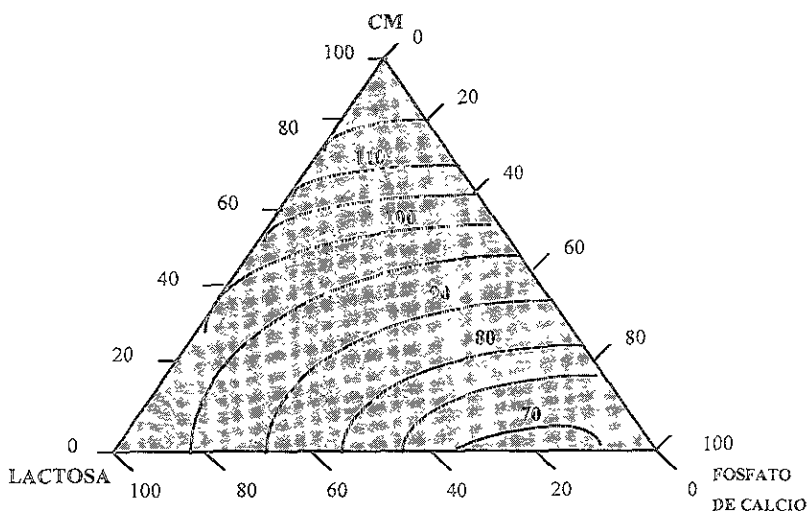


FIG 32 SUPERFICIE DE CONTORNO PARA LOS RESULTADOS DE RESISTENCIA A LA RUPTURA

a) INTERPRETACIÓN DE LOS RESULTADOS CORRESPONDIENTES RESISTENCIA A LA RUPTURA.

La relación que existe entre las proporciones de los constituyentes de la mezcla y la resistencia a la ruptura de la tabletas es estadísticamente significativa porque la *tabla 26* de análisis de varianza posee un valor de significancia de 0.0001, y se considera que el modelo cuadrático que se genera a partir de los valores estimados de la *tabla 28* describe de manera apropiada el comportamiento de los datos porque la variación explicada en la *tabla 27* es de 99 %. así mismo para la resistencia a la ruptura el valor máximo es de 91 N y corresponde a la *proporción 7* de la *tabla 28*. En la grafica de contornos de la *figura 32* se observa que las propiedades de dureza de la tableta se verán mejoradas a medida que se incremente la proporción de celulosa microcristalina.

## 5. RESULTADOS OBTENIDOS EN SAS PARA FRIABILIDAD

*The SAS System*  
*General Linear Models Procedure*  
*Analysis of Variance*

Source	DF	Sum Squares	Mean Square	F Value	Prob>F
Model	6	6.487	1.081	294.58	0.001
Error	4	0.015	0.037		
U Total	10	6.520			

TABLA 30 ANALISIS DE VARIANZA PARA FRIABILIDAD

Root MSE	0.061	R-square	0.998
Dep Mean	0.789	Adj R-sq	0.994
C.V.	7.679		

TABLA 31 CALCULO DE LA CANTIDAD DE VARIACION CALCULADA MEDIANTE ( $R^2$ )

Parameter	Estimate	T for H0: Parameter=0	Pr >  T	Std Error of Estimate
$X_1$	0.625	10.693	0.004	0.058
$X_2$	1.186	20.300	0.001	0.058
$X_3$	0.785	13.433	0.002	0.058
$X_1 \times X_2$	-0.787	-2.922	0.043	0.269
$X_1 \times X_3$	-0.148	-0.550	0.612	0.269
$X_2 \times X_3$	-0.307	-1.139	0.318	0.269

TABLA 32 CALCULO DE LOS PARAMETROS DEL MODELO DE SEGUNDO GRADO PARA FRIABILIDAD

*Ridge of Optimum Response*  
*Type of ridge=MINIMUM*

PROPORCIÓN	$X_1$	$X_2$	$X_3$	_PRED_	_STDERR_
1	0.333	0.333	0.334	0.730	0.030
2	0.377	0.310	0.313	0.709	0.030
3	0.424	0.296	0.286	0.693	0.031
4	0.472	0.276	0.252	0.678	0.031
5	0.522	0.267	0.211	0.665	0.031
6	0.573	0.263	0.164	0.652	0.032
7	0.625	0.264	0.110	0.641	0.034
8	0.677	0.270	0.052	0.631	0.039

TABLA 33 CALCULO DEL MINIMO VALOR PARA FRIABILIDAD

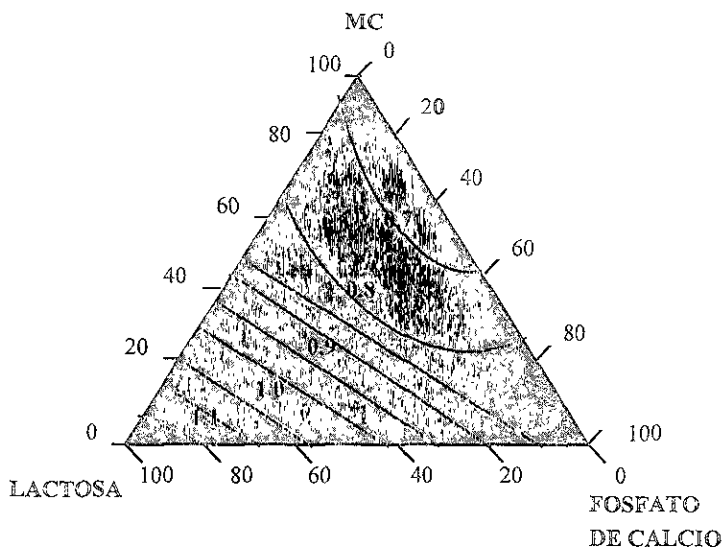


FIGURA 32 SUPERFICIE DE CONTORNO PARA LOS RESULTADOS DE FRIABILIDAD

#### a) INTERPRETACIÓN DE RESULTADOS DE FRIABILIDAD.

En la *tabla 30* se presenta el análisis de varianza de los resultados correspondientes a los valores de friabilidad, donde se manifiesta la gran relación que existe entre los constituyentes de la mezcla y la disolución debido a que el valor de significancia de los datos es de 0.001, así también la cantidad de variación explicada en la *tabla 31* es de 99 %, por lo que se considera que el modelo cuadrático que se genera a partir de los valores estimados de la *tabla 32* describe de manera apropiada el comportamiento de los datos de friabilidad; el mínimo valor calculado para este parámetro correspondera a la proporción 8 de la *tabla 32* y es de 0.63. En la gráfica de superficies de contorno (*figura 32*) se observa que a medida que se incrementen la proporción de celulosa microcristalina y lactosa las tabletas tendrán mejores características de friabilidad.

Según el análisis, al disminuir la proporción de fosfato de calcio disminuye la friabilidad, probablemente porque éste componente ( $x_3$ ) está en dos interacciones antagónicas  $x_2 x_3$  y  $x_1 x_3$ ; a pesar de que sus coeficientes no son realmente significativos.

## 6. RESULTADOS GENERALES

RESPUESTA	ADJ R-SQ	POLINOMIO OBTENIDO
FRIABILIDAD %	0.994	$0.625 X_1 + 1.186 X_2 + 0.785 X_3 - 0.787 X_1 X_2 - 0.148 X_1 X_3 + 0.307 X_2 X_3$
VARIACION DE PESO %	0.972	$0.68 X_1 + 1.049 X_2 + 1.786 X_3 - 0.834 X_1 X_2 - 3.598 X_1 X_3 + 0.458 X_2 X_3$
RESISTENCIA A LA RUPTURA (N)	0.997	$129.09 X_1 + 61.645 X_2 + 83.474 X_3 - 53.508 X_1 X_2 - 69.943 X_1 X_3 - 44.745 X_2 X_3$
VELOCIDAD DE DISOLUCION (MIN)	0.994	$198.139 X_1 + 114.049 X_2 + 328.99 X_3 - 402.395 X_1 X_2 + 351 X_1 X_3 + X_2 X_3 331.313$

TABLA 34 RESUMEN DE LOS POLINOMIOS Y COEFICIENTES DE AJUSTE OBTENIDOS MEDIANTE SAS PARA WINDOWS

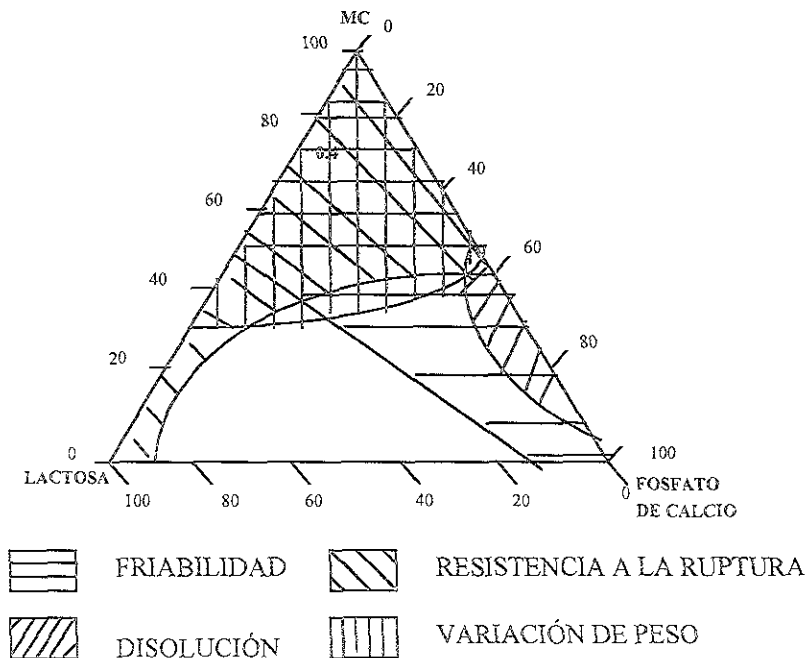


FIGURA 34 GRÁFICAS SUPERPUESTAS DE LAS ÁREAS CON LAS MEJORES RESPUESTAS

## a) INTERPRETACION DE RESULTADOS GENERALES

En la *tabla 17* se observa cada una de las formulaciones propuestas en sus variables codificadas y sus resultados sobre las variables de respuesta, con lo cual se generaron los polinomios de respuesta y sus respectivas coeficientes de variación explicada, cada uno de los cuales es superior al 90 % por lo que se considera que dichos polinomios explican de manera suficiente el comportamiento de los resultados, por consiguiente se generan las respectivas gráficas de contorno. En la gráfica correspondiente a la friabilidad (*figura 33*) se observa que a medida que se incrementa las proporciones de celulosa microcristalina y fosfato de calcio se esperarán mejores características de friabilidad de las tabletas, así también se observa algo parecido para la grafica de variación de peso (*figura 31*) y para la grafica de resistencia a la ruptura (*figura 32*); contrariamente en la gráfica de tiempo cuando 90% de naftidrofuril es liberado se observa que a medida que se incrementa la cantidad de fosfato de calcio se obtendrán mejores características de liberación (*figura 30*); tomado los contornos de mejor respuesta de cada una de las gráficas y superponiéndolas se puede generar la *gráfica 34* a partir de la cual se establece el área donde se deberá buscar la proporción de cada excipiente que nos brindara las tabletas con las mejores características de cada uno de los parámetros evaluados; y la cual resulta ser de 36.0 % de Avicel pH 102, 60.0 % de Fosfato de calcio y 4.0 % de Lactosa (este porcentaje es posible observarse en la gráfica)..

## XII. DISCUSIÓN

Los diseños de mezclas son una técnica estadística que se encuentra dentro de la metodología de superficie de respuesta y tiene por objetivo obtener un producto con ciertas características de calidad en función de los componentes de una mezcla; se diferencia de los demás diseños en que los niveles de los factores no son independientes, porque al disminuir o aumentar la cantidad de alguno de los componentes será necesario modificar la cantidad de algún otro u otros, consecuentemente los polinomios que se generan a través de esta técnica son más sencillos (carecen de un término independiente  $\beta_0$ ) y se denominan tipo Scheffé; dichos polinomios serán utilizados para modelar el comportamiento de una respuesta o serie de respuestas sobre un área o espacio experimental correspondiente a un simplex regular; para poder explorar este espacio se utilizan los diseños simplex en red, los centroides simplex o alguna de sus variantes; estos diseños serán los que determinen el número y el tipo de combinaciones o experimentos a efectuarse (matriz de experiencias).

Es importante tomar en consideración que de acuerdo al grado del polinomio a generar se asocia un tipo de diseño específico, es decir el polinomio de primer grado se asocia al diseño simplex  $\{q, 1\}$ , donde  $q$  se refiere al número de componentes y el segundo término se refiere al número de combinaciones binarias, por consiguiente este diseño únicamente evaluará comportamientos aditivos de los componentes sobre la respuesta o más específicamente el efecto de los componentes puros. El modelo de segundo grado se asocia al diseño simplex  $\{p, 2\}$  y por ello además de los componentes puros evaluará combinaciones binarias (1/2) de los mismos consiguiendo con ello establecer si existen sinergismos o antagonismos de los constituyentes sobre la respuesta. El modelo cúbico completo se asocia a diseño  $\{p, 3\}$  y establece la existencia de posibles curvaturas dentro del área del espacio simplex por su capacidad de evaluar combinaciones de tres ingredientes (1/3). Será evidente que entre mayor sea el grado del polinomio que se desea establecer, mayor será el número de experimentos que se deberán realizar.

Para aplicar los diferentes tipos de diseños de mezclas a un problema específico se deberán hacer una serie de consideraciones, una de las más importantes es la existencia de restricciones; si no existen, es decir si es posible evaluar cada constituyente en cualquier concentración el trabajo se verá facilitado y sólo será necesario establecer el grado del polinomio y aplicar el correspondiente diseño simplex en red o centroide simplex. Sin embargo, podría no ser de interés evaluar concentraciones de los constituyentes en su forma pura, en tal situación existen dos opciones de diseños a elegir: los planes de *Gammon* que tienen la capacidad de evaluar a los constituyentes al menos en combinaciones binarias, y los planes de *Lambrakis* que tiene la capacidad de evaluar combinaciones ternarias; estos diseños pueden resultar de especial interés. Por ejemplo si se deseara establecer un sistema de tres cosolventes para obtener la máxima solubilidad de un fármaco o una mezcla ternaria de solventes como fase móvil en un sistema cromatográfico de líquidos.

Cuando se presentan problemas con restricciones, es decir cuando es posible utilizar algún ingrediente o todos únicamente en ciertas proporciones o concentraciones se deberá trabajar

de diferente manera, como primer paso se deberá aclarar el tipo de restricción que se está enfrentando las cuales pueden ser de tres tipos: cuando los componentes presentan límites superiores, cuando presentan límites inferiores y cuando presentan límites inferiores y superiores. El primer caso se aplica cuando se trabaje con antioxidantes y/o conservadores en cuyos casos es común utilizarlos por debajo de ciertas proporciones. El segundo caso corresponderá cuando se esté formulando un comprimido con dos o tres ingredientes activos (p.e. vitamínicos, antibióticos, etc.) que deberán estar presentes al menos en ciertas proporciones. El tercer caso será bastante frecuente porque la gran mayoría de los excipientes que acompañan a las diferentes formas farmacéuticas se utilizan en ciertos intervalos de concentración. Por ejemplo, en comprimidos están los lubricantes, los deslizantes y desintegrantes; en las suspensiones están los viscosantes y los humectantes.

En general se recomienda utilizar modelos de segundo grado para tener un número de experimentos razonable y por que es el mayor grado que genera el paquete SAS para Windows, sin embargo existen diversos paquetes específicos de mezclas que pueden generar modelos de mayor grado y complejidad (estos se encuentran en la bibliografía de este trabajo).

Dependiendo de cada caso se emplean ciertas transformaciones especiales para establecer si tales restricciones son consistentes (es decir si son posibles) como para transformarlas en consistentes, determinando posteriormente el número de puntos (proporciones de cada componente) a experimentar para generar un modelo; a su vez este modelo deberá ser modificado para obtenerse en sus valores originales. El polinomio generado de descripción de respuesta deberá ser evaluado mediante diversas pruebas de hipótesis primeramente para determinar la dependencia de la respuesta por los componentes de la mezcla, posteriormente se determinará si dicho modelo explica la suficiente cantidad de variación en los valores de la respuesta ( $R^2_A \geq 0.9$ ) para, finalmente, determinar si existe la posibilidad de simplificarlo disminuyendo su grado o el número de términos.

Como se dilucida en los resultados, los diseños de mezclas presentan bastantes ventajas de utilización en el área farmacéutica, sobre todo en desarrollo, donde se pueden utilizar para realizar experimentos preliminares de preformulación, tal es el caso de ejemplo 1, donde se establece un área de experimentación inicial (que corresponde a un simplex por las restricciones de los componentes) para posteriormente transformarse en un espacio más pequeño, obteniéndose un polinomio a partir del cual se establecen las proporciones de los excipientes que generarán la mejor formulación; es importante puntualizar que normalmente en desarrollo, se trabaja con formulaciones de 6 o más componentes por consiguiente, podría parecer engorroso el manejar tantos experimentos, pero existe la posibilidad de mantener algunos de ellos constantes, detalle que se observa en el mismo *ejemplo 1* donde se efectúa un granulado inicial con las concentraciones tanto del principio activo como los excipientes iniciales constantes, y las proporciones variadas son las correspondientes a los excipientes del mezclado posterior.

Otro caso es un problema en el cual se deseaba diseñar una suspensión de *carbamazepina* como principio activo, y donde la primera fase consistió en establecer un diseño de mezclas de tres viscosantes (goma Xantana, Carboximetil Celulosa (mediana viscosidad) y carbopol) para establecer una viscosidad específica y posteriormente trabajar sobre los otros 4 componentes de la formulación manteniendo constante las de los tres iniciales, consiguiendo una disminución en el número de experimentos necesarios para una formulación de 7 constituyentes.

En el *ejemplo 2* se establece claramente como es posible eliminar esa aparente “desventaja” de los diseños de mezclas de sólo poder evaluar una variable de respuesta (en farmacia usualmente se trabaja con más de 2 variables de respuesta) puesto que se obtienen modelos para cada una de las 4 variables de respuesta (friabilidad, resistencia a la ruptura, variación de peso y disolución) y mediante superposición de la superficie de contorno de cada una de las mejores áreas se genera una área con proporciones de los componentes que cumplirán con las características de calidad deseadas en el producto.

Se debe enfatizar que los diseños de mezclas no sólo se utilizan en desarrollo farmacéutico, sino también es posible emplearlos en el establecimiento de mezclas de detergentes y sanitizantes para mejorar la limpieza de material y equipo en planta y microbiología, o como se muestra en el *capítulo 9* se pueden incluir en el estudio variables de proceso y con ello utilizarlos en el área analítica en cromatografía de líquidos en donde se establece una combinación de solventes para la fase móvil y varias velocidades de flujo para mejorar la resolución de la sustancia analizada; así como para optimizar alguna formulación recién desarrollada o ya establecida localizándose las mejores condiciones de proceso.

Como se ha observado, existen diferentes tipos diseños de mezclas, por consiguiente se sugieren la siguiente serie de pasos para resolver un problema que involucre estos diseños:

1. Definir los objetivos del experimento; qué se desea observar, podría pretenderse establecer la influencia que tienen los componentes sobre la (s) respuesta (s), u obtener una combinación de los componentes para conseguir una respuesta óptima.
2. Seleccionar los componentes de la mezcla o algún factor (variable de proceso). En este caso de acuerdo al experimentador se podrían fijar algunos de ellos para reducir el número de experiencias a realizar.
3. Identificar alguna restricción en los componentes de la mezcla para especificar la región dentro de la cual se realizarán los experimentos.
4. Especificar la o las respuestas a ser medidas.
5. Proponer un modelo apropiado para modelar los datos de la respuesta como función de los componentes de la mezcla o algunos de los otros factores seleccionados para el estudio. En este punto se sugiere utilizar modelos de grado 2 porque son los modelos



que genera el SAS para Windows; y porque son modelos sencillos que requieren menor número de experimentos.

6. Seleccionar alguno de los diseños de mezclas que serán apropiados para ajustar el modelo propuesto y además permitan determinar la falta de ajuste.

7. Realizar la matriz de experimentos que se obtuvo a partir del diseño de mezclas.

8. Generar el modelo a partir de las respuestas.

9. Realizar las pruebas de adecuación del modelo como  $R_A^2$  y las diferentes pruebas de hipótesis para disminuir el grado del polinomio o el número de coeficientes presentes en él.

10. Realizar la determinación de la proporción o región de proporciones que permitirán conseguir la mejor respuesta o serie de respuestas mediante análisis RIDGE y/o gráficas de contorno.

11. Realizar ensayos de verificación de puntos óptimos para comprobar la eficacia del modelo de predicción.

Para conseguir que sea más didáctica la manera en que se resolverá un problema de diseños de mezclas se presenta el *cuadro 1*, donde la secuencia de actividades se realiza a través de cuadros de decisión, y es de utilidad para resolver prácticamente cualquier problema en el cual estén inmiscuidas combinaciones de componentes para obtener un producto con ciertas características.

El uso de los diseños de mezclas dependerá de la experiencia y del ingenio del usuario; pero si el lector utiliza las recomendaciones que se proporcionan en este trabajo seguramente se le facilitará la resolución de los problemas en donde se involucren diseños de mezclas.

### XIII. CONCLUSIONES

La diferencia de los diseños de mezclas con respecto a los demás diseños de superficie de respuesta es que los factores son los componentes o ingredientes y en consecuencia no son independientes entre si.

Los diseños de mezclas tienen la ventaja de requerir menor cantidad de experimentos para conseguir localizar la respuesta buscada (mínimo o máximo).

El espacio de experimentación de los diseños de mezclas corresponde a un simplex regular y para explorarlo se deben utilizar los diseños simplex en red, los centroides simplex o alguna de sus variantes; el primero de estos diseños se emplea cuando se desea analizar un problema en donde es posible tener cada uno de los componentes en una cantidad del 100 %, sin embargo se recomienda añadir puntos centrales (centroide simplex) con la finalidad de corroborar la adecuación del modelo.

Cuando se considere que no tiene sentido evaluar a los constituyentes en su forma pura y se desea evaluarlos al menos en combinaciones binarias se deberá de utilizar los planes de Gannom, así mismo cuando carezca de interés el evaluar los componentes puros y en combinaciones binarias se puede recurrir a los planes de Lambrakis que evalúa los componentes al menos en combinaciones terciarias.

Es posible encontrar problemas de mezclas con restricciones, es decir que alguno de los componentes presentan límites inferiores o límites superiores e incluso ambos; en el primer caso se utilizan los L-pseudocomponentes, en el segundo los U-pseudocomponentes y en el tercero los diseños de vértices, siempre considerando que es indispensable determinar que las restricciones sean consistentes (es decir, que se pueden utilizar cada uno de los límites superiores e inferiores de las restricciones) y si no es así transformarlas en consistentes.

Los modelos que se generan con los diseños de mezclas tienen un menor número de términos que los polinomios estándar y de ellos los más usuales son: el lineal, el cuadrático, el cúbico especial y el cúbico completo; el criterio para seleccionar alguno de estos modelos se fundamenta en la existencia de interacciones entre los constituyentes de la mezcla. Por consiguiente, el primero de los modelos mencionados se empleará cuando la mezcla de los componentes presente un comportamiento aditivo en la respuesta; el segundo, si existe alguna interacción en la combinación binaria de los componentes de la mezcla; el tercero y cuarto evalúan el comportamiento de una combinación de tres componentes. El segundo paso después de generar cualesquiera de los modelos anteriores es, a través de una prueba de hipótesis, determinar si la respuesta depende de los componentes de la mezcla y cuando ello resulta ser afirmativo se realiza la determinación del coeficiente de ajuste para decidir si el modelo explica la suficiente cantidad de variación en los valores de la respuesta, cuando dicho valor excede 0.9 se considera razonable y como siguiente paso se procederá a tratar de simplificar el modelo generado

Los diseños de mezclas son una herramienta estadística que se puede aplicar en varias áreas de la industria farmacéutica. Por ejemplo, en desarrollo (para desarrollar nuevas formulaciones o mejorar las existentes), en el área de producción (optimización de procesos) y también es posible emplearlos en el área analítica (para optimizar mezclas de fases móviles y/o temperaturas de columnas en HPLC). Como los diseños de mezclas se pueden utilizar cuando se elabore un producto a partir de algunos constituyentes el área de aplicación es bastante amplia, desde la industria de la panificación hasta la petroquímica.

Una desventaja de estos diseños es que únicamente son capaces de evaluar aparentemente una respuesta, sin embargo como se observó en el segundo ejemplo es posible evaluar más de una respuesta, mediante la superposición de las zonas de mejor respuesta y consiguiendo con ello obtener las proporciones en las cuales el producto cumplirá con las características de calidad diseñadas, siempre considerando que se deberá de corroborar dicha información mediante la experimentación.

Como se ha podido notar, utilizar los diseños de mezclas no es complicado, sobre todo si cuenta con una herramienta que le facilite la elaboración de los cálculos como son los paquetes computacionales, existen varios que son de bastante utilidad, el empleado en esta tesis fue el SAS para Windows versión 6.11 porque es un paquete fácil de utilizar ya que la mayor parte de las personas está familiarizado con el ambiente Windows; sin embargo este paquete no es capaz de calcular polinomios mayores de segundo grado.

El generar polinomios hasta de segundo grado no es una total desventaja puesto que son fáciles de interpretar y ofrecen una idea, en la mayoría de los casos, bastante aproximada del comportamiento de la respuesta.

## . APÉNDICE A

### *DETERMINACIÓN DE INCONSISTENCIAS*

Las *inconsistencias* pueden ser determinadas mediante las fórmulas:

$$0 \leq L_i \leq x_i \leq U_i \leq 1, i=1,2,\dots,q \text{ y } \sum_{i=1}^q x_i = \dots\dots\dots(1)$$

Las restricciones de la ecuación 1 son consistentes cuando cada una de las combinaciones posibles de  $x_i=1,2,\dots,p$  y cada proporción del componente (no necesariamente todos simultáneamente) contiene el límite inferior ( $x_i=L_i$ ) y todas las proporciones de los componentes contienen el límite superior. Por ejemplo para las restricciones:

$$\begin{aligned} 0.0 &\leq x_1 \leq 0.1 \\ 0.1 &\leq x_2 \leq 0.2 \\ 0.6 &\leq x_3 \leq 0.8 \end{aligned}$$

No son consistentes o son inconsistentes, por que  $x_3$  no puede tomar el valor de 0.6; el valor mínimo que puede tomar es  $x_3 = (1-U_1-U_2) = 0.7$ .

Para determinar inconsistencias en las restricciones e identificarlas, primero se debe calcular el rango de cada restricción mediante:

$$R_i = U_i - L_i, \quad i = 1,2,\dots,q$$

Y después

$$R'_i = \frac{R_i}{1 - \sum_{i=1}^q L_i}, \quad i = 1,2,\dots,q \dots\dots\dots(2)$$

Representa el rango del L-pseudocomponente,  $x'_i$ . Ahora, si para alguna  $i$ ,  $R'_i > 1$ , esto es,

$$R_i > 1 - \sum_{i=1}^q L_i \quad \text{o} \quad U_i + \sum_{j=1}^q L_j > \dots\dots\dots(3)$$

Donde  $U_i$ , es inaceptable; tal como el caso  $U_1$  en el ejemplo antes descrito, donde  $R'_1 = 0.5/0.4 = 1.2 > 1$ . También, el límite menor  $L_i$ , es inaceptable si para alguna  $i$ ,

$$\sum_{i=1}^q R_i > 1 \quad \text{o} \quad L_i + \sum_{i=1}^q U_i < \dots\dots\dots(4)$$

Otra forma de detectar restricciones inconsistentes causadas por una  $L_i$  inaceptable es usando los U-pseudocomponentes,  $u_i$ ,  $i=1,2,..,g$ , donde U-simplex se define como.

$$R_u = \sum_{i=1}^g U_i, \dots\dots\dots(5)$$

Si alguna  $R_i > R_u$ , el  $L_i$  es inaceptable.

Cuando las restricciones son inconsistentes y por consiguiente se requiere ajustarlas para hacerlas consistentes, el ajuste se deben hacer sobre los límites que no son aceptables.

## APÉNDICE B

### TRANSFORMACIÓN DE RESTRICCIONES

Cuando se cambian las restricciones inconsistentes a consistentes frecuentemente se modifica la forma y el tamaño de la región experimental. Para ilustrar esto se consideran las restricciones:

$$0.2 \leq X_1 \leq 0.4 \quad 0.2 \leq X_2 \leq 0.6 \quad 0.18 \leq X_3 \leq 0.7$$

Donde los rangos de los componentes son  $R_1 = 0.4 - 0.2 = 0.2$ ,  $R_2 = 0.4$  y  $R_3 = 0.52$ , también:

$$R_L = 1 - 0.58 = 0.42, \quad R_U = 1.7 - 1 = 0.7$$

Y como  $R = 0.52 > R_L$ , esto significa que el límite superior,  $U_3 = 0.7$ , no es alcanzable y por consiguiente es inconsistente. Si  $U_3$  es reemplazado por el límite superior implícito mediante,  $U^*_3 = L_3 + R_L = 0.18 + 0.42 = 0.6$ , las restricciones se transformarían en consistentes y  $R_U = 0.6$ .

También es posible forzar las restricciones haciéndolas consistentes mediante la alteración del límite inferior. Por ejemplo si se deseara mantener a  $U_3 = 0.7$  como el límite superior de  $X_3$ , para conseguir que las restricciones sean consistentes se requerirá alterar los límites inferiores  $L_1 = 0.2$  a  $L_1 = 0.1$  o disminuyendo  $L_2 = 0.2$  a  $L_2 = 0.1$  o disminuyendo ambos  $L_1$  y  $L_2$  en una cantidad igual de 0.05, es decir  $L_1 = 0.15$  y  $L_2 = 0.15$ . Por consiguiente se deberá revisar cada uno de las posibles opciones de restricción para determinar como afectarán las restricciones la forma y tamaño de la región experimental.

Una técnica que es de utilidad para determinar las restricciones más convenientes es mediante la estimación de la varianzas y covarianzas  $\text{var}(\hat{\beta}) = (X'X)^{-1} \hat{\sigma}^2$  de los parámetros estimados de los diferentes modelos obtenidos de cada área, y se trabajará con el modelo que presente la menor varianza.

*Fórmulas para determinar el número de vértices extremos, ejes y caras de la región restringida*

Una fórmula que puede ser usada para calcular el número de bordes d-dimensionales ( $d = 0, 1, \dots, q-2$ ) de una región, definida por las restricciones consistentes,

$$X_1 + X_2 + \dots + X_q = 1, \quad 0 \leq L_i \leq X_i \leq U_i \leq 1$$

Es determinada mediante

$$N_d = C(q, q-d-1) + \sum_{r=1}^{q-d-1} L(r)C(q-r, q-r-d-1) - \sum_{r=d+1}^q [L(r) + E(r)]C(r, r-d-1) \dots (5)$$

Para entender los términos de la fórmula se proponen los siguientes pasos; primero determinar los rangos.  $R_i = U_i - L_i$ , y  $R_L = 1 - \sum_{i=1}^q L_i$  y  $R_U = \sum_{i=1}^q U_i - 1$ .

Donde  $R_p$  es el mínimo de  $R_L$  y  $R_U$ . El segundo paso será dividir las  $C(q, r)$  combinaciones de los rangos dentro de tres mutuamente exclusivas y exhaustivas subpruebas donde:  $r = 1, 2, \dots, q$

$L(r)$  = El número de combinaciones de los rangos de los componentes que suman un número que es menor a  $R_p$ .

$E(r)$  = El número de los combinaciones de los rangos de los componentes igual a  $R_p$

$G(r)$  = El número de combinaciones de los rangos de los componentes que suman un número mayor de  $R_p$ .

$$C(q, r) = L(r) + E(r) + G(r)$$

## APÉNDICE C

### *DETERMINACIÓN DE LOS VÉRTICES EXTREMOS MEDIANTE MCLEAN Y ANDERSON*

Las coordenadas de los vértices extremos del poliedro pueden ser definidas mediante el procedimiento propuesto de *McLean y Anderson* a través de:

1. Listar todas las posibles combinaciones de los valores de  $L_i$  y  $U_i$  para  $q-1$  componentes y obtener el valor del componente en blanco. Este procedimiento produce  $n^{2-1}$  combinaciones. Con tres componentes, por ejemplo, donde  $L_1, L_2$  y  $L_3$  son los límites inferiores y  $U_1, U_2$  y  $U_3$  son los límites superiores, las combinaciones de los niveles del componente son  $L_1L_2$ ,  $L_1U_2$ ,  $U_1L_2$ ,  $U_1U_2$ , donde la proporción para el componente 3 está en blanco; el procedimiento se repite  $q$  veces, permitiendo que cada componente tenga un nivel en cero y sea capaz de ser calculado, obteniéndose una lista que consistirá de  $q \times (2^{q-1})$  posibles combinaciones. Este se puede realizar ayudándose de un diseño factorial de dos niveles.

2. A través de todas las posibles combinaciones generadas encontrar todos los posibles espacios en blanco posibles, esto es, llenar dentro de los niveles (necesariamente encontrados dentro de las restricciones del factor buscado) que hará el total de los niveles para que la suma de las combinaciones sea igual a la unidad. Cada combinación admisible de los niveles de todos los  $q$  componentes se definirán como vértice extremo.

Después de determinar los vértices extremos del poliedro convexo, el siguiente paso es definir los puntos centrales de las caras y de los ejes. Esto es que un punto centroide se localizará en cada una de las caras en dos dimensiones, tres dimensiones, ...,  $d$ -dimensiones, donde  $d \leq q-2$ , así también se deberá determinar el centroide total del poliedro.

El centroide de las caras en dos dimensiones es localizado mediante aislamiento de todos los vértices que tienen  $q-3$  niveles de factor idénticos y mediante el promedio de cada uno de los tres factores remanentes. Todos los centroides faltantes son encontrados de forma similar usando todos los vértices que tienen  $q-r-1$  niveles de factor idénticos para una de las caras  $r$ -dimensionales donde  $r \leq d \leq q-2$ .

Para ilustrar lo anterior se ejemplifican las siguientes restricciones:

$$0.40 \leq x_1 \leq 0.80$$

$$0.10 \leq x_2 \leq 0.50$$

$$0.05 \leq x_3 \leq 0.30$$

$$0.05 \leq x_4 \leq 0.30$$



Para estas restricciones se generaran las  $4 \times (2^{4-1}) = 32$  posibles combinaciones de los componentes que podrían o no ser admisibles, comenzando la lista con 8 vértices admisibles que se presentan y enumeran como, (1), (2), ..., (8) en la tabla 1.

	$x_1$	$x_2$	$x_3$	$x_4$		$x_1$	$x_2$	$x_3$	$x_4$
(1)	0.4	0.1	0.05		(3)	0.4	0.05	0.05	0.05
(3)	0.4	0.1	0.3	0.2	(4)	0.4	0.25	0.05	0.3
	0.4	0.5	0.05	0.05	(5)	0.4	0.25	0.3	0.05
(2)	0.4	0.5	0.3			0.4		0.3	0.3
	0.8	0.1	0.05	0.05	(2)	0.8	0.1	0.05	0.05
	0.8	0.1	0.3			0.8		0.05	0.3
	0.8	0.5	0.05			0.8		0.3	0.05
	0.8	0.5	0.3			0.8		0.3	0.3
	0.4	0.1		0.05	(2)	0.8	0.1	0.05	0.05
(6)	0.4	0.1	0.2	0.03	(7)	0.055	0.1	0.05	0.3
(3)	0.4	0.5	0.05	0.05	(8)	0.055	0.1	0.3	0.05
	0.4	0.5		0.3			0.1	0.3	0.3
(2)	0.8	0.1	0.05	0.05	(3)	0.04	0.5	0.05	0.05
	0.8	0.1		0.3			0.5	0.05	0.3
	0.8	0.5		0.05			0.5	0.3	0.05
	0.8	0.5		0.3			0.5	0.3	0.3

TABLA 1. PROCEDIMIENTO PROPUESTO POR MCLEAN Y ANDERSON PARA DEFINIR LAS COORDENADAS DE LOS VERTICES EXTREMOS DEL POLIEDRO

Las dos caras son encontradas mediante el agrupamiento de los vértices del polihedro dentro del grupo de tres o más vértices, donde cada vértice tienen el mismo valor de  $x_1$  para uno de los componentes. Para el vértice (2), (3), (4) y (7) presentan  $x_3=0.05$  y estos cuatro vértices definen una cara.

Las 6 caras bidimensionales restrictivas del polihedro y las coordenadas del centroide de cada cara son listadas como los puntos 9 al 14 en la tabla 2. El centroide global, cuyas coordenadas son definidas como el promedio de las coordenadas de los ocho vértices corresponde al punto 15.

PUNTOS	TIPO DE BORDES	PROPORCIONES DE LOS COMPONENTES DE LA MEZCLA				COMBINACION DE VÉRTICES
		X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	
1	Vértice	0.40	0.10	0.30	0.20	
2	Vértice	0.80	0.10	0.05	0.05	
3	Vértice	0.40	0.50	0.05	0.05	
4	Vértice	0.40	0.25	0.05	0.30	
5	Vértice	0.40	0.25	0.30	0.05	
6	Vértice	0.40	0.10	0.20	0.30	
7	Vértice	0.55	0.10	0.05	0.30	
8	Vértice	0.55	0.10	0.30	0.05	
9	cara centroide	0.58	0.24	0.05	0.13	2,3,4,5
10	cara centroide	0.54	0.24	0.17	0.05	2,3,5,8
11	cara centroide	0.40	0.24	0.18	0.18	1,3,4,5,6
12	cara centroide	0.54	0.10	0.18	0.18	1,2,6,7,8
13	cara centroide	0.45	0.15	0.10	0.30	4,6,7
14	cara centroide	0.45	0.15	0.30	0.10	1,5, 8
15	centroide global	0.49	0.19	0.16	0.16	1,2,..... , 8

TABLA 2 COORDENADAS DE LOS 15 PUNTOS DEL DISEÑO PARA LAS RESTRICCIONES DEL EJEMPLO INICIAL

## APÉNDICE D

### MANEJO DE DATOS EN EL PAQUETE SAS

1.0 En la pantalla de *program editor* (Figura 1) teclear los datos de la siguiente manera:

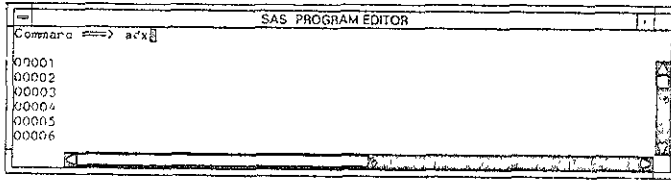


FIGURA A

Data sauser.problemal;

Input x1	x2	x3	y;
Cards;			
100	0	0	22.40
0	100	0	54.20
0	0	100	43.00
50	50	0	51.50
50	0	50	127.50
0	50	50	77.08
33.33	33.33	33.33	144.60
66.67	16.67	16.67	80.00
16.67	66.67	16.67	86.05
16.67	16.67	66.67	102.10;

run;

2.0 Dar un clic derecho al mouse y aparecerá una ventana con las siguientes opciones:

File	Edit	View	Locals	Globals	Help
------	------	------	--------	---------	------

3.0 Posesionarse en *Globals* y aparecerán una serie de opciones; elegir *Invoke application* y posteriormente en *Data analysis* y en *Experimental design*; esta última opción despliega la pantalla del *ADX (Design Directory)* en el cual se presentan 4 tipos de diseños de experimentos: *Two level*, *Response Surface*, *Orthogonal Array* y *Mixture*; ubicarse en esta última (*Mixture*) y aparecera un cuadro de dialogo, en el cual se deberá especificar el diseño, tipo de diseño, la descripción del diseño, número de corridas, número de bloques, factores, respuestas y tipo de resolución. (ver figuras A, B y C)

4.0 Establecidos los parámetros anteriores posesionarse en *File* y se desplegara una serie de opciones, elegir guardar y proceder a guardarlo con el mismo nombre con que se cargo en *program editor* (ejemplo1).

5.0 La pantalla de *ADX* se cerrara y es posible observar en la ventana de *out put* los experimentos que se deben realizar aleatorizados.

6.0 Regresar a la pantalla de *ADX* y abrir el archivo guardado como ejemplo1, seleccionar *Examine* y en la ventana de *out put* del SAS se observaran los datos procesados por el paquete.

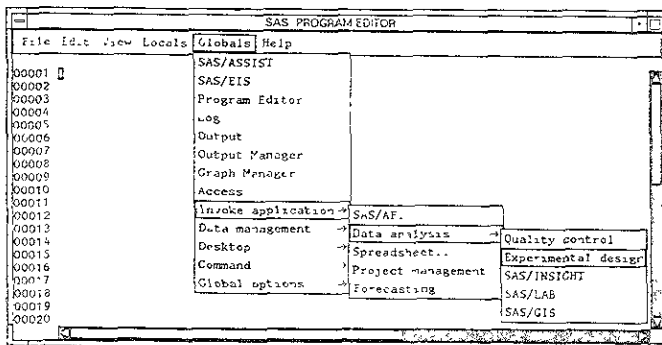


FIGURA B

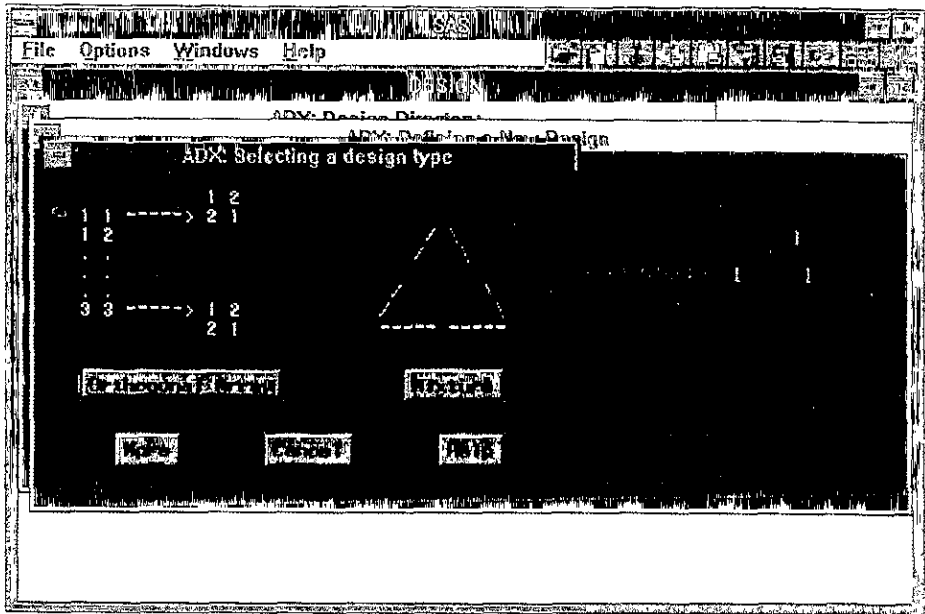


FIGURA C

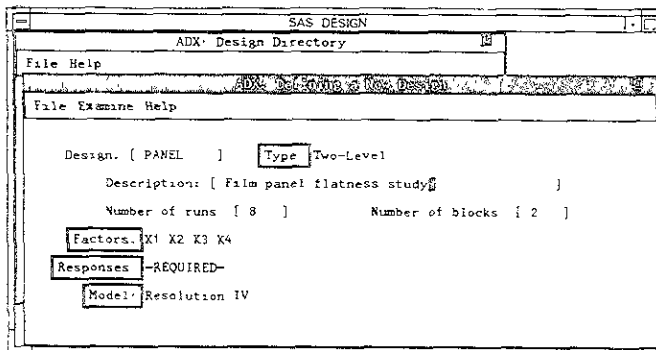


FIGURA E

## BIBLIOGRAFIA

1. MONTGOMERY DOUGLAS C., "DISEÑO Y ANALISIS DE EXPERIMENTOS", EDITORIAL IBEROAMERICANA, MÉXICO, 1991
2. LAWSON JOHN, MADRIGAL JOSE L., ERJAVEC JOHN, "ESTRATEGIAS EXPERIMENTALES PARA EL MEJORAMIENTO DE LA CALIDAD INDUSTRIAL", EDITORIAL IBEROAMERICANA, MÉXICO, 1992.
3. KHURI ANDRE I, CORNELL JHON A., "RESPONSE SURFACES DESIGN AND ANALYSIS", EDITION BY JOHN WILEY AND SONS, NEY YORK, 1990.
4. CORNELL JHON A., "EXPERIMENTS WITH MIXTURES", SECOND EDITION, PUBLICATION BY WILEY-INTERSCIENCE, NEW YORK, 1990.
5. MASON ROBERTA, GUNST RICHARD F., HESS JAMES L., "STATISTICAL DESIGN AND ANALYSIS OF EXPERIMENTS", PUBLICATION BY JOHN WILEY AND SONS INC., NEW YORK, 1989.
6. LORENZEN THOMAS J, ANDERSON VIRGIL L., "DESIGN OF EXPERIMENTS", PUBLICATION BY MARCEL DEKKER INC., NEW YORK, 1993.
7. ATKINSON A. C. AND DONEY G., "OPTIMUM EXPERIMENTAL DESIGN", PUBLICATION BY OXFORD UNIVERSITY, NEW YORK, 1992.
8. MEAD R., "THE DESIGNS OF STATISTICAL EXPERIMENTS", PUBLICATION BY CAMBRIDGE UNIVERSITY, NEW YORK, 1990
9. KRISTL A., VOJNOVIE D., KARBA D., MRHAR A., "MIXTURE DESIGN OF THEOPHYLLINE RETARD FORMULATION", INTERNATIONAL JOURNAL PHARMACEUTICS 100 (1-3). 33-40, (1993)
10. P.J. WAALER, C. GRAFFNER AND W. MULLER, "OPTIMIZATION OF A MATRIX TABLET FORMULATION USING A MIXTURE DESIGN", ACTA PHARM NORD. 4 (1): 9-16, (1992)
11. SAS PROCEDURES GUIDE, VERSION 6, THIRD EDITION, SAS INSTITUTE INC. NC USA.
12. LEON LACHMAN, HERBERT A. LIEBERMAN, JOSEPH L. KANING, "THE

*THEORY AND PRACTICE OF INDUSTRIAL PHARMACY*, THIRD EDITION, PUBLICATION BY LEA AND FEBIGER, PHILADELPHIA, USA 1986.

- 13 *GEORGE E. P. BOX, NORMAN R. DRAPER, "EMPIRICAL MODEL-BUILDING AND RESPONSE SURFACES"*, EDITION BY JOHN WILEY AND SONS INC, NEW YORK, USA, 1987.
14. *PIEPEL GREGORY F., CORNELL J. A. "MIXTURE APPROACHES: EXAMPLES, DISCUSSION AND RECOMMENDATIONS"*, JOURNAL OF QUALITY TECHNOLOGY 26 (3): 177-196, (1994).
15. *KHURI ANDRE, "RESPONSE SURFACE MODELS WITH MIXTURES EFFECTS"*, JOURNAL OF QUALITY TECHNOLOGY 28 (2). 177-186, (1996).
16. *WATTS G. DONALD, "UNDERSTANDING CANONICAL ANALYSIS"*, JOURNAL OF QUALITY TECHNOLOGY 27 (1) 40-44, (1995)
- 17 *MONTGOMERY DOUGLAS C, VOTH SHEILLE R. "MULTICOLLINEARITY AND LEVERAGE IN MIXTURE EXPERIMENTS"*, JOURNAL OF QUALITY TECHNOLOGY 26 (2) 96-108, (1994).