



# **UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO**

INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA  
ÁREA ACADÉMICA DE INGENIERÍA

## **DISEÑOS EXPERIMENTALES DE BLOQUES INCOMPLETOS Y APLICACIONES EN LA INDUSTRIA**

MONOGRAFÍA

QUE PARA OBTENER EL TÍTULO DE INGENIERO  
INDUSTRIAL

PRESENTA:

P.L.I.I. ISIDRO JESÚS GONZÁLEZ HERNÁNDEZ

DIRECTOR: DRA. MIRIAM M. ÁLVAREZ SUÁREZ

## **DEDICATORIAS**

### ***A UAEH:***

*Estoy orgulloso de ser parte de esta gran institución la que me brindo los elementos suficientes para mi formación académica, la cual forma egresados competitivos y eficientes a nivel Nacional.*

### ***A MIS PROFESORES:***

*Por su amistad, conocimientos y experiencias que me han brindaron a lo largo de mi formación profesional, los cuales han contribuido en mi desarrollo personal y profesional.*

### ***A MI ASESORA:***

*Por el apoyo, paciencia y tiempo que me dedico para realizar este trabajo y agradecerle su  
confianza.*

# ÍNDICE

		Página
<b>INTRODUCCIÓN</b>		1
<b>JUSTIFICACIÓN</b>		4
<b>OBJETIVO GENERAL Y ESPECÍFICOS</b>		7
<b>DEFINICIONES</b>	<b>DEFINICIONES DE PRINCIPALES CONCEPTOS</b>	8
<b>CAPÍTULO 1</b>	<b>DISEÑO DE EXPERIMENTOS</b>	13
1.1	Introducción	13
1.2	Tipos de variabilidad	16
1.3	Planificación de un experimento	18
1.4	Resumen de los principales conceptos	25
1.5	Principios básicos en el diseño de experimentos	30
1.6	Algunos diseños experimentales clásicos	32
1.6.1	Diseño completamente aleatorizado	33
1.6.2	Diseño en bloques o con un factor bloque	33
1.6.3	Diseños con dos o más factores bloque	34
1.6.4	Diseños con dos o más factores	35
1.6.5	Diseños factoriales a dos niveles	36
<b>CAPÍTULO 2</b>	<b>DISEÑO DE BLOQUES INCOMPLETOS</b>	38
2.1	Introducción	38

2.2	Bloques incompletos de tratamientos para reducir el tamaño de los bloques	39
2.3	Diseños de bloques incompletos balanceados (BIB)	42
2.4	Cómo aleatorizar los diseños de bloques incompletos	43
2.5	Análisis de diseños BIB	45
2.6	Diseño renglón-columna para dos criterios de bloque	55
2.7	Reducción del tamaño de experimento con diseños parciales balanceados (BIPB)	58
<b>CAPÍTULO 3</b>	<b>DISEÑO DE BLOQUES INCOMPLETOS: DISEÑOS RESOLUBLES Y CÍCLICOS</b>	64
3.1	Introducción	64
3.2	Diseños resolubles para ayudar a manejar el experimento	65
3.3	Diseños resolubles renglón-columna para dos criterios de bloque	70
3.4	Los diseños cíclicos simplifican la construcción del diseño	75
3.5	Diseños $\alpha$ para diseños resolubles versátiles a partir de una construcción cíclica	78
<b>CAPÍTULO 4</b>	<b>APLICACIÓN A LA INDUSTRIA DE DISEÑOS DE BLOQUES INCOMPLETOS Y SU</b>	85

	<b>EFICIENCIA FRENTE A OTROS DISEÑOS.</b>	
4.1	Introducción	85
4.2	Eficiencia de los diseños de bloques incompletos	86
4.3	Aplicaciones de Diseños Experimentales de Bloques Incompletos en el sector industrial	87
<b>Apéndice 1A.1</b>	Algunos diseños de bloques incompletos balanceados.	100
<b>Apéndice 1A.2</b>	Algunos diseños de cuadrados latinos incompletos.	102
<b>Apéndice 1A.3</b>	Estimaciones de mínimos cuadrados para diseño BIB.	105
<b>Apéndice 2A.1</b>	Planes para diseños cíclicos.	108
<b>Apéndice 2A.2</b>	Arreglos para diseños $\alpha$	109
	<b>CONCLUSIONES</b>	110
	<b>BIBLIOGRAFÍA</b>	116

## ÍNDICE DE TABLAS

Página

<b>Tabla 1</b>	Diferencia entre ambos tipos de bloqueo anidados y jerarquerizados.	35
<b>Tabla 2.1</b>	Análisis de varianza para un diseño de bloque incompleto balanceado.	48
<b>Tabla 2.2</b>	Conversión porcentual del metilglucósidos con acetileno a acetileno a alta presión, en un diseño de bloque incompleto balanceado.	50
<b>Tabla 2.3</b>	Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloque incompleto balanceado.	51
<b>Tabla 2.4</b>	Estimaciones de mínimos cuadrados de las medias de presión para la conversión porcentual de multiglucósidos, en un diseño de bloques incompletos balanceado.	52
<b>Tabla 2.5</b>	Análisis intrabloques para un diseño de bloques renglón-columna incompleto balanceado con tratamientos ortogonales por renglones.	58
<b>Tabla 3.1</b>	Análisis intrabloques para un diseño de bloques incompleto balanceado resoluble.	69

<b>Tabla 3.2</b>	Análisis intrabloques para un diseño de bloques incompleto renglón-columna anidado.	73
<b>Tabla 4.1</b>	Pureza observada de un producto químico en un factorial $2^3$ parcialmente confundido.	90
<b>Tabla 4.2</b>	Coeficientes para contrastes en un diseño de tratamiento factorial $2^3$ .	91
<b>Tabla 4.3</b>	Cálculos para los efectos factoriales no confundidos con bloques.	91
<b>Tabla 4.4</b>	Los contrastes confundidos AB, AC, y BC calculados con los totales de todas las observaciones de tratamientos.	92
<b>Tabla 4.5</b>	Diferencia calculada para cada par de bloques.	92
<b>Tabla 4.6</b>	Análisis de varianza para la pureza de un producto químico en un factorial $2^3$ parcialmente confundido.	93
<b>Tabla 4.7</b>	Conversión porcentual del metilglucósidos con acetileno a acetileno a alta presión, en un diseño de bloque completo aleatorizado.	98
<b>Tabla 4.8</b>	Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloque completo aleatorizado.	98

## ÍNDICE DE CUADROS

		Página
<b>Cuadro 1</b>	Diseño de bloques incompletos con cuatro tratamientos en bloques de tres unidades.	40
<b>Cuadro 2</b>	Diseño de bloques renglón-columna incompleto balanceado de 4 X 7.	56
<b>Cuadro 3</b>	Diseño de bloques incompleto parcialmente balanceado con seis tratamientos en bloques de cuatro unidades.	60
<b>Cuadro 4</b>	Primeros y segundos asociados para los seis tratamientos en el diseño de bloques parcialmente balanceado del cuadrado 1.3	61
<b>Cuadro 5</b>	Diseño de bloques incompleto balanceado resoluble para evaluar productos alimenticios.	67
<b>Cuadro 6</b>	Diseño de bloques incompleto renglón columna anidado de 3 x 3 para las muestras de la población de insectos.	72
<b>Cuadro 7</b>	Diseño cíclico para seis tratamientos en bloques de tres unidades.	76
<b>Cuadro 8</b>	Diseño cíclico generado a partir de los bloques iniciales.	78
<b>Cuadro 9</b>	Diseño $\alpha$ para 12 tratamientos en tres grupos de réplicas.	81
<b>Cuadro 10</b>	Ilustración de bloques contiguos en replicas.	83
<b>Cuadro 11</b>	Diseño renglón-columna latinizado.	84

## INTRODUCCIÓN

Hace casi ochenta años el trabajo pionero de Sir Ronald Fisher mostró cómo los métodos estadísticos y en particular el diseño de experimentos podían ayudar a solventar el problema de cómo descubrir y entender las complejas relaciones que pueden existir entre varias variables y más aún, alcanzar este objetivo a pesar de que los datos están contaminados con error experimental. Desde que se comenzaron a utilizar en agricultura y biología, estas técnicas se han seguido desarrollando y su uso se ha generalizado a las ciencias físicas, sociales, la ingeniería y la industria.

Más recientemente, su efecto catalítico en los procesos de investigación y aprendizaje, ha demostrado el importante papel que ha jugado en la revolución industrial, en calidad y productividad, liderada por Japón.

Los experimentos en la industria moderna son más complicados, porque son muchos los factores que son susceptibles de controlarse y que afectan a la calidad de los productos, de aquí que son muchas las combinaciones de dichos factores que se deben probar para obtener resultados válidos y consistentes.

El diseño estadístico de experimentos permite optimizar la información generada acerca del sistema, en relación a los objetivos planteados. En otras palabras, el diseño de experimentos es la aplicación del método científico para generar conocimiento acerca de un sistema o proceso.

El diseño de experimentos ha sido consolidado poco a poco en la industria actual como un conjunto de herramientas estadísticas y de ingeniería, que permiten lograr la máxima eficiencia de los procesos con el mínimo costo.

Con el diseño de experimentos se pueden resolver problemas como los siguientes:

- Desarrollar productos que sean resistentes a diferentes condiciones de uso.
- Encontrar las condiciones de proceso que dan por resultado valores óptimos de las características de calidad del producto.
- Elegir entre diferentes materiales, métodos de producción o proveedores, con el fin de mejorar la calidad y reducir costos.
- Estudiar la relación que existe entre los diferentes factores que intervienen en un proceso y las características de calidad del producto. Este es el problema general que resuelve el diseño de experimentos.

Generalizando, los modelos de diseños de experimentos son modelos estadísticos que podemos aplicar para:

- Determinar qué variables tienen mayor influencia en los valores de respuesta,
- Determinar el mejor valor de las variables para tener un valor cercano al valor de respuesta deseado y,
- Determinar el mejor valor de las variables para que el valor de la respuesta tenga la menor variabilidad.

Los pormenores de cálculo para el análisis estadístico de los distintos diseños experimentales varían de un diseño a otro, aunque muchos procedimientos estadísticos usados en el análisis son comunes a la mayoría de los diseños

existentes. Esto se debe a que los procedimientos mismos generalmente se relacionan con los diseños del tratamiento específico, cada uno de los cuales puede aparecer en varias configuraciones del diseño experimental.

La intención de esta Monografía es la de introducir algunos procedimientos útiles para diversos estudios en los que no es posible considerar una réplica completa de todos los tratamientos y se hace necesario bloquizar unidades experimentales. Los ***Diseños de Bloques Incompletos*** ofrecen una solución que nos permite disminuir la varianza del error y proporciona comparaciones más precisas entre tratamientos de lo que es posible con un diseño de bloques completos.

En el Capítulo 1 se exponen las ideas generales del diseño de experimentos, el Capítulo 2 contiene los conceptos básicos de los diseños de bloques incompletos, el Capítulo 3 se ha dedicado a los diseños de bloques incompletos resolubles y cíclicos y en el Capítulo 4 se presentan algunos ejemplos de aplicaciones de los diseños de bloques incompletos a la industria y su eficiencia frente a los diseños de bloques completos. En cada uno de los Capítulos se incluyen ejemplos de aplicaciones a problemas industriales. Por último se ofrecen algunas conclusiones y recomendaciones para el uso de estos diseños tan útiles en los problemas de la industria.

## JUSTIFICACIÓN

En el campo de la industria es práctica común realizar pruebas o experimentos con la intención de que al hacer cambios en los materiales, métodos o condiciones de operación de un proceso se puedan detectar, resolver o minimizar los problemas de calidad.

El *diseño experimental* es el arreglo de unidades experimentales utilizado para controlar el error experimental, a la vez que acomoda los tratamientos. Existe en la literatura una amplia variedad de arreglos diseñados para controlar el error experimental y se observa una tendencia natural a diseñar los experimentos de acuerdo con diseños ya existentes. No obstante, desarrollar un diseño de experimentos que satisfaga la demanda del experimento que se está realizando no es tarea fácil, aunque debe ser el objetivo principal de cualquier estudio.

El logro de la máxima información, precisión y exactitud en los resultados, junto con el uso más eficiente de los recursos existentes, es un principio a seguir en la elección del diseño adecuado del experimento.

La asociación entre la asignación de tratamientos y el diseño experimental se puede realizar en dos contextos diferentes: sin bloques y con bloques. Por todos es conocido que el incorporar bloques o tratamientos replicados en un experimento, provoca una mejor estimación del error experimental además de reducirlo; sin embargo, a veces esto no es posible.

El número de bloques en un estudio de investigación afecta la precisión de las estimaciones de las medias de los tratamientos y la potencia de las pruebas estadísticas para detectar las diferencias entre las medias de los grupos en tratamiento. Sin embargo, el costo de conducir estudios de investigación restringe las réplicas a un número razonable. Luego, el número de bloques o de

réplicas está determinado por las restricciones prácticas que se pueden asignar al problema.

Es importante destacar, que a pesar de lo anteriormente expuesto, el investigador debe estar consciente que al reducir el número de tratamientos por bloque o el número de réplicas por tratamiento, estará influyendo directamente el valor de la varianza del error experimental ( $\sigma^2$ ), en el tamaño de la diferencia entre dos medias de tratamiento, en el nivel de significación de la prueba ( $\alpha$ ) y en la potencia de la prueba ( $1 - \beta$ ).

Existen fórmulas para calcular el número de réplicas necesarias para valores de  $\sigma^2$ ,  $\alpha$ ,  $\beta$  y la diferencia de medias a detectar, sin embargo, son estimaciones y aproximaciones. Con frecuencia se determinan con base en las estimaciones de la varianza obtenidas en estudios previos del mismo tipo y no a las del estudio real que se usarán para calcular los intervalos de confianza y las pruebas de hipótesis.

La réplica de un experimento proporciona los datos para estimar la varianza del error experimental. La bloquización proporciona un medio para reducir el error experimental. Sin embargo, las réplicas y la bloquización por sí solos no garantizan estimaciones válidas de la varianza del error experimental o de las comparaciones entre tratamientos.

Fisher, 1926, señaló que la sola aleatorización proporciona estimaciones válidas de la varianza del error para los métodos de inferencia estadística justificados para la estimación y pruebas de hipótesis en el experimento. La aleatorización es la asignación aleatoria de los tratamientos a las unidades experimentales.

La eficiencia relativa mide la efectividad de la bloquización para reducir la varianza del error experimental en el diseño de experimentos. En la práctica la

eficiencia relativa se mide para determinar la eficiencia del diseño usado en realidad respecto a otro más sencillo que pudo usarse pero no se hizo.

La varianza de la media de un tratamiento es una medida de la precisión de las medias de tratamiento estimadas. Esta precisión se controla mediante la magnitud de  $\sigma^2$  y el número de réplicas, que en cierto grado están bajo el control del investigador, ya que éste puede aumentar la precisión en la estimación de la media, aumentando el número de réplicas. Sin embargo, esto aumentaría el costo de la investigación. Hay otra forma de incrementar la precisión del experimento reduciendo  $\sigma^2$  a través de varias actividades de control local, como la bloquización.

La importancia que tiene la bloquización en el diseño de experimentos sobre todo en la industria, donde las condiciones de replicación no son fáciles y en la mayoría de los casos, imposible, es lo que nos impulsó a escoger este tema para la realización de esta monografía sobre los **“diseños de bloques incompletos”** para estudiar con sencillez, claridad y profundidad el diseño de experimentos y específicamente, los diseños de bloques incompletos, incluyendo las técnicas y las estrategias necesarias para planear y analizar adecuadamente las pruebas experimentales, y conocer mejor la realidad del objeto de estudio y se mejore en forma eficaz el desempeño de productos y procesos.

En ella se desarrollan temas tan importantes como: elementos de inferencia estadística, experimentos con un solo factor, diseños en bloques, diseños factoriales, planeación de un experimento, así como la eficiencia de los mismos para el tratamiento de casos particulares.

La explicación de cada técnica se apoya en casos reales y se resuelven ejemplos que ayudan a una mejor comprensión de cuándo y cómo aplicarla.

Por todo lo anterior, esta monografía es un valioso instrumento que permitirá al lector adquirir algunos conocimientos que le conducirán a un desempeño profesional exitoso.

## **OBJETIVO GENERAL Y ESPECÍFICOS**

### **❖ Objetivo General de la presente monografía es el siguiente:**

“Realizar una presentación de las bases teóricas y prácticas de los diseños experimentales de bloques incompletos tanto balanceados como parcialmente balanceados y sus posibles aplicaciones en la industria”

### **❖ Objetivos Específicos son los siguientes:**

- Conceptualizar las bases fundamentales del Diseño de Experimentos y su importancia en la investigación.
- Definir los conceptos generales y específicos de utilización de los diseños de bloques incompletos balanceados (DBIB) y diseños de bloques incompletos parcialmente balanceados (DBIPB).
- Aplicación a la industria de Diseños de Bloques Incompletos y su eficiencia frente a otros diseños.
- Manejar el software SPSS para su posible utilización a problemas industriales.

## DEFINICIONES DE LOS PRINCIPALES CONCEPTOS

Nota: Los conceptos se presentan en forma como aparecen en el trabajo

**Estadística:** Rama de las matemáticas que se ocupa de reunir, organizar y analizar datos numéricos y que ayuda a resolver problemas como el diseño de experimentos y la toma de decisiones.

**Experimento:** Es un cambio en las condiciones de operación de un sistema o proceso, que se hace con el objetivo de medir el efecto de tal cambio sobre una o varias propiedades del producto. Dicho experimento permite aumentar el conocimiento acerca del mismo.

**Diseño de experimentos:** Consiste en planear un conjunto de pruebas experimentales, de tal manera que los datos generados puedan analizarse estadísticamente, para obtener conclusiones válidas y objetivas acerca del sistema o proceso.

**Diseño experimental:** Es el arreglo de las unidades experimentales utilizado para controlar el error experimental, a la vez que acomoda los tratamientos.

**Unidad experimental:** Es la entidad física o sujeto expuesto al tratamiento independientemente de otras unidades.

- 2) son los objetos, individuos, intervalos de espacio o tiempo sobre los que se experimenta.

**Tamaño del Experimento:** es el número total de observaciones recogidas en el diseño.

**Tratamiento:** Son el conjunto de circunstancias creadas para el experimento, en respuesta a la hipótesis de investigación y son el centro de la misma.

2) Es una combinación específica de los niveles de los factores en estudio. Son, por tanto, las condiciones experimentales que se desean comparar en el experimento. En un diseño con un único factor son los distintos niveles del factor y en un diseño con varios factores son las distintas combinaciones de niveles de los factores.

**Variable:** Que puede variar, que está sujeta a variación.

2) Mat. Que se puede hacer variar a voluntad.

**Variable de interés o respuesta:** es la variable que se desea estudiar y controlar su variabilidad.

**Observación experimental:** es cada medición de la variable respuesta.

**Diseño equilibrado o balanceado:** es el diseño en el que todos los tratamientos son asignados a un número igual de unidades experimentales.

**Error experimental:** Describe la variación entre las unidades experimentales tratadas de manera idéntica e independiente.

**Varianza:** es el promedio de los cuadrados de las desviaciones,  $(x_i - \bar{x})^2$  de cada elemento,  $x_i$ , respecto a la media,  $\bar{x}$ :

**Bloqueo:** Bloquear es controlar en forma adecuada a todos los factores que puedan afectar la respuesta observada.

**Factor:** son las variables independientes que pueden influir en la variabilidad de la variable de interés.

**Factor bloque:** es un factor en el que no se está interesado en conocer su influencia en la respuesta pero se supone que ésta existe y se quiere controlar para disminuir la variabilidad residual.

**Factor tratamiento:** es un factor del que interesa conocer su influencia en la respuesta.

**Aleatorio, a:** Dependiente de un proceso fortuito.

- 2) Mat. Que depende al azar, sometido a leyes de probabilidad. Variable, función aleatoria, cuyo valor es aleatorio.

**Aleatorización:** Consiste en hacer corridas experimentales en orden aleatorio y con material seleccionado también aleatoriamente.

**Permutación:** Las permutaciones son las distintas formas en que se pueden ordenar los  $n$  elementos de un conjunto.

**Interacción de factores:** existe interacción entre dos factores  $F_i$  y  $F_j$  si el efecto de algún nivel de  $F_i$  cambia al cambiar de nivel en  $F_j$ . Esta definición puede hacerse de forma simétrica y se puede generalizar a interacciones de orden tres o superior.

**Niveles:** cada uno de los resultados de un factor. Según sean elegidos por el experimentador o elegidos al azar de una amplia población se denominan factores de **efectos fijos** o factores de **efectos aleatorios**.

**Frecuencia:** Repetición de un acto o suceso.

- 2) Tecn. Número de observaciones estadísticas correspondientes a un suceso dado.
- 3) Mat. en estadística, el número de veces que ocurre un cierto suceso. También se denomina frecuencia absoluta, en contraposición con la

frecuencia relativa, que consiste en la proporción de veces que ocurre dicho suceso con relación al número de veces que podría haber ocurrido.

**Réplica:** Estd. Repetición independiente del experimento básico. Dicho de otra manera más específica, cada tratamiento se aplica de manera independiente a dos o más unidades experimentales.

**Correlación:** f. relación recíproca o mutua entre dos o más cosas.

- 2) ling. Conjunto de dos series de fonemas opuestas por un mismo rasgo distintivo.
- 3) Relación que se establece entre ellas.
- 4) Existencia de mayor o menor dependencia mutua entre dos variables aleatorias.
- 5) En estadística, relación entre las dos variables de una distribución bidimensional. Se mide mediante el coeficiente de correlación,  $p$ .

**Estocástico:** adj. Que se debe al azar, que remite al azar.

- 2) Mat. Que remite al campo de cálculo de probabilidades.

**Proceso estocástico:** proceso en el que un sistema cambia de forma aleatoria entre diferentes estados, a intervalos regulares o irregulares.

**Ortogonalidad de factores:** dos factores  $F_i$  y  $F_j$  con  $I$  y  $J$  niveles, respectivamente, son ortogonales si en cada nivel  $i$  de  $F_i$  el número de observaciones de los  $J$  niveles de  $F_j$  están en las mismas proporciones. Esta propiedad permite separar los efectos simples de los factores en estudio.

**Parámetro:** Mat. Letra que designa en una ecuación una magnitud dada, pero a la que se pueden atribuir valores diferentes.

- 2) Estd. Número que se obtiene a partir de los datos de una distribución estadística y que sirve para sintetizar alguna característica relevante de la misma.

**Sesgo:** de sesgo II) adj. torcido, cortado o situado oblicuamente.

- 2) fig. grave o torcido en el semblante.

**Sesgado:** de sesgo II) adj. p. us. Término estadístico que significa falta de simetría con respecto a la distribución normal.

# CAPÍTULO 1

## DISEÑO DE EXPERIMENTOS

### 1.1 Introducción.

Los modelos de “Diseño de experimentos” (DOE en inglés) son modelos estadísticos clásicos cuyo objetivo es averiguar si determinados factores influyen en la variable de interés y, si existe influencia de algún factor, cuantificarla. Ejemplos donde habría que utilizar estos modelos son los siguientes:

- En el rendimiento de un determinado tipo de máquinas (unidades producidas por día) se desea estudiar la influencia del trabajador que la maneja y la marca de la máquina.
- Se quiere estudiar la influencia del tipo de pila eléctrica y de la marca en la duración de las pilas.
- Una compañía telefónica está interesada en conocer la influencia de varios factores en la variable de interés “la duración de una llamada telefónica”. Los factores que se consideran son los siguientes: hora a la que se produce la llamada; día de la semana en que se realiza la llamada; zona de la ciudad desde la que se hace la llamada; sexo del que realiza la llamada; tipo de teléfono (público o privado) desde el que se realiza la llamada.

- Una compañía de software está interesada en estudiar la variable “porcentaje que se comprime un fichero al utilizar un programa que comprime ficheros” teniendo en cuenta el tipo de programa utilizado y el tipo de fichero que se comprime.
- Se quiere estudiar el rendimiento de los alumnos en una asignatura y, para ello, se desean controlar diferentes factores: profesor que imparte la asignatura; método de enseñanza; sexo del alumno.

## 1.2 Metodología al Diseño de Experimentos.

La metodología del diseño de experimentos se basa en la experimentación. Es conocido que si se repite un experimento, en condiciones indistinguibles, los resultados presentan variabilidad que puede ser grande o pequeña. Si la experimentación se realiza en un laboratorio donde la mayoría de las causas de variabilidad están muy controladas, el error experimental será pequeño y habrá poca variación en los resultados del experimento. Pero si se experimenta en procesos industriales o administrativos, la variabilidad es grande en la mayoría de los casos.

El objetivo del diseño de experimentos es estudiar si utilizar un determinado tratamiento produce una mejora en el proceso o no. Para ello se debe experimentar utilizando el tratamiento y no utilizándolo. Si la variabilidad experimental es grande, sólo se detectará la influencia del uso del tratamiento cuando éste produzca grandes cambios en relación con el error de observación.

La metodología del Diseño de Experimentos estudia cómo variar las condiciones habituales de realización de un proceso empírico para aumentar la probabilidad de detectar cambios significativos en la respuesta, de esta forma se obtiene un mayor conocimiento del comportamiento del proceso de interés.

Para que la metodología de diseño de experimentos sea eficaz es fundamental que el experimento esté bien diseñado.

Un experimento se realiza por alguno de los siguientes motivos:

- Determinar las principales causas de variación en la respuesta.
- Encontrar las condiciones experimentales con las que se consigue un valor extremo en la variable de interés o respuesta.
- Comparar las respuestas en diferentes niveles de observación de variables controladas.
- Obtener un modelo estadístico-matemático que permita hacer predicciones de respuestas futuras.

La utilización de los modelos de diseño de experimentos se basa en la experimentación y en el análisis de los resultados que se obtienen en un experimento bien planificado. En muy pocas ocasiones es posible utilizar estos métodos a partir de datos disponibles o datos históricos, aunque también se puede aprender de los estudios realizados a partir de datos recogidos por observación, de forma aleatoria y no planificada. En el análisis estadístico de datos históricos se pueden cometer diferentes errores, los más comunes son los siguientes:

- Inconsistencia de los datos. Los procesos cambian con el tiempo, se producen cambios en el personal (cambios de personas, mejoras del personal por procesos de aprendizaje, motivación), cambios en las máquinas (reposiciones, reparaciones, envejecimiento). Estos cambios tienen influencia en los datos recogidos, lo que hace que los datos históricos sean poco fiables, sobre todo si se han recogido en un amplio espacio de tiempo.
- Variables con fuerte correlación. Puede ocurrir que en el proceso existan dos o más variables altamente correlacionadas que pueden llevar a situaciones confusas. Por ejemplo, en el proceso hay dos variables  $X_1$  y  $X_2$  fuertemente correlacionadas que influyen en la respuesta, pero si en los datos que se tienen aumentan al mismo tiempo el valor de las dos

variables no es posible distinguir si la influencia es debida a una u otra o a ambas variables (confusión de los efectos). Otra situación problemática se presenta si solo se dispone de datos de una variable (por ejemplo de  $X_1$  y no de  $X_2$ ), lo que puede llevar a pensar que la variable influyente es la  $X_1$  cuando, en realidad, la variable influyente es la  $X_2$  (variable oculta).

- El rango de las variables controladas es limitado. Si el rango de una de las variables importantes e influyentes en el proceso es pequeño, no se puede saber su influencia fuera de ese rango y puede quedar oculta su relación con la variable de interés o los cambios que se producen en la relación fuera del rango observado. Esto suele ocurrir cuando se utilizan los datos recogidos al trabajar el proceso en condiciones normales y no se experimenta (cambiando las condiciones de funcionamiento) para observar el comportamiento del proceso en situaciones nuevas.

### **1.3 Tipos de variabilidad.**

Uno de los principales objetivos de los modelos estadísticos y, en particular, de los modelos de diseño de experimentos, es controlar la variabilidad de un proceso estocástico que puede tener diferente origen. De hecho, los resultados de cualquier experimento están sometidos a tres tipos de variabilidad cuyas características son las siguientes:

- **Variabilidad sistemática y planificada.**

Esta variabilidad viene originada por la posible dispersión de los resultados debida a diferencias sistemáticas entre las distintas condiciones experimentales impuestas en el diseño por expreso deseo del experimentador. Es el tipo de variabilidad que se intenta identificar con el diseño estadístico.

Cuando este tipo de variabilidad está presente y tiene un tamaño importante, se espera que las respuestas tiendan a agruparse formando grupos (clusters).

Es deseable que exista esta variabilidad y que sea identificada y cuantificada por el modelo.

- **Variabilidad típica de la naturaleza del problema y del experimento.**

Es la variabilidad debida al ruido aleatorio. Este término incluye, entre otros, a la componente de variabilidad no planificada denominada error de medida. Es una variabilidad impredecible e inevitable.

Esta variabilidad es la causante de que si en un laboratorio se toman medidas repetidas de un mismo objeto ocurra que, en muchos casos, la segunda medida no sea igual a la primera y, más aún, no se puede predecir sin error el valor de la tercera. Sin embargo, bajo el aparente caos, existe un patrón regular de comportamiento en esas medidas: todas ellas tenderán a fluctuar en torno a un valor central y siguiendo un modelo de probabilidad que será importante estimar.

Esta variabilidad es inevitable pero, si el experimento ha sido bien planificado, es posible estimar (medir) su valor, lo que es de gran importancia para obtener conclusiones y poder hacer predicciones.

Es una variabilidad que va a estar siempre presente pero que es tolerable.

- **Variabilidad sistemática y no planificada.**

Esta variabilidad produce una variación sistemática en los resultados y es debida a causas desconocidas y no planificadas. En otras palabras, los resultados están siendo sesgados sistemáticamente por causas

desconocidas. La presencia de esta variabilidad supone la principal causa de conclusiones erróneas y estudios incorrectos al ajustar un modelo estadístico.

Como se estudiará posteriormente, existen dos estrategias básicas para tratar de evitar la presencia de este tipo de variabilidad: la aleatorización y la técnica de bloques.

Este tipo de variabilidad debe de intentar evitarse y su presencia lleva a conclusiones erróneas.

#### **1.4 Planificación de un experimento.**

La experimentación forma parte natural de la mayoría de las investigaciones científicas e industriales, en muchas de las cuales, los resultados del proceso de interés se ven afectados por la presencia de distintos factores, cuya influencia puede estar oculta por la variabilidad de los resultados muestrales. Es fundamental conocer los factores que influyen realmente y estimar esta influencia. Para conseguir ésto es necesario experimentar, variar las condiciones que afectan a las unidades experimentales y observar la variable respuesta. Del análisis y estudio de la información recogida se obtienen las conclusiones.

La forma tradicional que se utilizaba en la experimentación, para el estudio de estos problemas, se basaba en estudiar los factores uno a uno, ésto es, variar los niveles de un factor permaneciendo fijos los demás. Esta metodología presenta grandes inconvenientes:

- Es necesario un gran número de pruebas.
- Las conclusiones obtenidas en el estudio de cada factor tiene un campo de validez muy restringido.
- No es posible estudiar la existencia de interacción entre los factores.
- Es inviable, en muchos casos, por problemas de tiempo o costo.

Las técnicas de **diseño de experimentos** se basan en estudiar simultáneamente los efectos de todos los factores de interés, son más eficaces y proporcionan mejores resultados con un menor coste.

A continuación se enumeran las etapas que deben seguirse para una correcta planificación de un diseño experimental, etapas que deben ser ejecutadas de forma secuencial. También se introducen algunos conceptos básicos en el estudio de los modelos de diseño de experimentos.

Las etapas a seguir en el desarrollo de un problema de diseño de experimentos son las siguientes:

1. Definir los objetivos del experimento.
2. Identificar todas las posibles fuentes de variación, incluyendo:
  - factores tratamiento y sus niveles,
  - unidades experimentales,
  - factores nuisance (molestos): factores bloque, factores ruido y covariables.
3. Elegir una regla de asignación de las unidades experimentales a las condiciones de estudio (tratamientos).
4. Especificar las medidas con que se trabajará (la respuesta), el procedimiento experimental y anticiparse a las posibles dificultades.
5. Ejecutar un experimento piloto.
6. Especificar el modelo.
7. Esquematizar los pasos del análisis.
8. Determinar el tamaño muestral.
9. Revisar las decisiones anteriores. Modificarlas si se considera necesario.

Las etapas del listado anterior no son independientes y en un determinado momento puede ser necesario volver atrás y modificar decisiones tomadas en algún paso previo.

A continuación se hace una breve descripción de las decisiones que hay que tomar en cada una de las etapas enumeradas. Sólo después de haber tomado estas decisiones se procederá a realizar el experimento.

### 1. Definir los objetivos del experimento.

Se debe hacer una lista completa de las preguntas concretas a las que debe dar respuesta el experimento. Es importante indicar solamente cuestiones fundamentales ya que tratar de abordar problemas colaterales puede complicar innecesariamente el experimento.

Una vez elaborada la lista de objetivos, puede ser útil esquematizar el tipo de conclusiones que se espera obtener en el posterior análisis de datos.

Normalmente la lista de objetivos es refinada a medida que se van ejecutando las etapas del diseño de experimentos.

### 2. Identificar todas las posibles fuentes de variación.

Una **fente de variación** es cualquier “cosa” que pueda generar variabilidad en la respuesta. Es recomendable hacer una lista de todas las posibles fuentes de variación del problema, distinguiendo aquellas que, a priori, generarán una mayor variabilidad. Se distinguen dos tipos:

- **Factores tratamiento:** son aquellas fuentes cuyo efecto sobre la respuesta es de particular interés para el experimentador.
- **Factores “nuisance”:** son aquellas fuentes que no son de interés directo pero que se contemplan en el diseño para reducir la variabilidad no planificada.

A continuación se precisan más estos importantes conceptos.

## I Factores y sus niveles.

Se denomina **factor tratamiento** a cualquier variable de interés para el experimentador cuyo posible efecto sobre la respuesta se quiere estudiar.

Los **niveles** de un factor tratamiento son los tipos o grados específicos del factor que se tendrán en cuenta en la realización del experimento.

Los factores tratamiento pueden ser **cualitativos o cuantitativos**.

Ejemplos de factores **cualitativos** y sus niveles respectivos son los siguientes:

- proveedor (diferentes proveedores de una materia prima),
- tipo de máquina (diferentes tipos o marcas de máquinas),
- trabajador (los trabajadores encargados de hacer una tarea),
- tipo de procesador (los procesadores de los que se quiere comparar su velocidad de ejecución),
- un aditivo químico (diferentes tipos de aditivos químicos),
- el sexo (hombre y mujer),
- un método de enseñanza (un número determinado de métodos de enseñanza cuyos resultados se quieren comparar).
- 

Ejemplos de factores **cuantitativos** son los siguientes:

- tamaño de memoria (diferentes tamaños de memoria de ordenadores),
- droga (distintas cantidades de la droga),
- la temperatura (conjuntos de temperaturas seleccionadas en unos rangos de interés).

Debe tenerse en cuenta que en el tratamiento matemático de los modelos de diseño de experimento, los factores cuantitativos son tratados como cualitativos y

sus niveles son elegidos equiespaciados o se codifican. Por lo general, un factor no suele tener más de cuatro niveles.

Cuando en un experimento se trabaja con más de un factor, se denomina:

**Tratamiento** a cada una de las combinaciones de niveles de los distintos factores.

**Observación** es una medida en las condiciones determinadas por uno de los tratamientos.

**Experimento factorial** es el diseño de experimentos en que existen observaciones de todos los posibles tratamientos.

## II Unidades experimentales.

Son el material donde evaluar la variable respuesta y al que se le aplican los distintos niveles de los factores tratamiento.

Ejemplos de unidades experimentales son:

- en informática, ordenadores, páginas web, buscadores de internet,
- en agricultura, parcelas de tierra,
- en medicina, individuos humanos o animales,
- en industria, lotes de material, trabajadores, máquinas.

Cuando un experimento se ejecuta sobre un período de tiempo de modo que las observaciones se recogen secuencialmente en instantes de tiempo determinados, entonces los propios instantes de tiempo pueden considerarse unidades experimentales.

Es muy importante que las unidades experimentales sean representativas de la población sobre la que se han fijado los objetivos del estudio. Por ejemplo, si se utilizan los estudiantes universitarios de un país como unidades experimentales, las

conclusiones del experimento no son extrapolables a toda la población adulta del país.

### III Factores “nuisance”: bloques, factores ruido y covariables.

En cualquier experimento, además de los factores tratamiento cuyo efecto sobre la respuesta se quiere evaluar, también influyen otros factores, de escaso interés en el estudio, pero cuya influencia sobre la respuesta puede aumentar significativamente la variabilidad no planificada. Con el fin de controlar esta influencia pueden incluirse en el diseño nuevos factores que, atendiendo a su naturaleza, pueden ser de diversos tipos.

**Factor bloque.** En algunos casos el factor nuisance puede ser fijado en distintos niveles, de modo que es posible controlar su efecto a esos niveles. Entonces la forma de actuar es mantener constante el nivel del factor para un grupo de unidades experimentales, se cambia a otro nivel para otro grupo y así sucesivamente. Estos factores se denominan factores de bloqueo (factores-bloque) y las unidades experimentales evaluadas en un mismo nivel del bloqueo se dice que pertenecen al mismo **bloque**. Incluso cuando el factor nuisance no es medible, a veces es posible agrupar las unidades experimentales en bloques de unidades similares: parcelas de tierra contiguas o períodos de tiempo próximos probablemente conduzcan a unidades experimentales más parecidas que parcelas o períodos distantes.

Desde un punto de vista matemático el tratamiento que se hace de los factores-bloque es el mismo que el de los factores-tratamiento en los que no hay interacción, pero su concepto dentro del modelo de diseño de experimentos es diferente. Un factor-tratamiento es un factor en el que se está interesado en conocer su influencia en la variable respuesta y un factor-bloque es un factor en el que no se está interesado en conocer su influencia pero se incorpora al diseño del experimento para disminuir la variabilidad residual del modelo.

**Covariable.** Si el factor nuisance es una propiedad cuantitativa de las unidades experimentales que puede ser medida antes de realizar el experimento (el tamaño de un fichero informático, la presión sanguínea de un paciente en un experimento médico o la acidez de una parcela de tierra en un experimento agrícola). El factor se denomina **covariable** y juega un papel importante en el análisis estadístico.

**Ruido.** Si el experimentador está interesado en la variabilidad de la respuesta cuando se modifican las condiciones experimentales, entonces los factores nuisance son incluidos deliberadamente en el experimento y no se aísla su efecto por medio de bloques. Se habla entonces de factores ruido.

En resumen, las posibles fuentes de variación de un experimento son:

<b>Fuente</b>	<b>Tipo</b>
Debida a las condiciones de interés (Factores tratamiento)	Planificada y sistemática
Debida al resto de condiciones controladas (Factores “nuisance”)	Planificada y sistemática
Debida a condiciones no controladas (error de medida, material experimental, ... )	No planificada, pero ¿sistemática?

### 3. Elegir una regla de asignación de las unidades experimentales a las condiciones de estudio (“tratamientos”).

La **regla de asignación o diseño experimental** especifica qué unidades experimentales se observarán bajo cada tratamiento. Hay diferentes posibilidades:

- diseño factorial o no,
- anidamiento,
- asignación al azar en determinados niveles de observación,
- el orden de asignación, etc.

En la práctica, existen una serie de diseños estándar que se utilizan en la mayoría de los casos.

**4. Especificar las medidas que se realizarán (la “respuesta”), el procedimiento experimental y anticiparse a las posibles dificultades.**

**Variable respuesta o variable de interés.** Los datos que se recogen en un experimento son medidas de una variable denominada variable respuesta o variable de interés.

Es importante precisar de antemano cuál es la variable respuesta y en qué unidades se mide. Naturalmente, la respuesta está condicionada por los objetivos del experimento. Por ejemplo, si se desea detectar una diferencia de 5 g. en la respuesta de dos tratamientos no es apropiado tomar medidas con una precisión próxima al gramo.

A menudo aparecen dificultades imprevistas en la toma de datos. Es conveniente anticiparse a estos imprevistos pensando detenidamente en los problemas que se pueden presentar o ejecutando un pequeño experimento piloto. Enumerar estos problemas permite en ocasiones descubrir nuevas fuentes de variación o simplificar el procedimiento experimental antes de comenzar.

También se debe especificar con claridad la forma en que se realizarán las mediciones: instrumentos de medida, tiempo en el que se harán las mediciones, etc.

## 5. Ejecutar un experimento piloto.

Un experimento piloto es un experimento que utiliza un número pequeño de observaciones. El objetivo de su ejecución es ayudar a completar y chequear la lista de acciones a realizar. Las ventajas que proporciona la realización de un pequeño experimento piloto son las siguientes:

- permite practicar la técnica experimental elegida e identificar problemas no esperados en el proceso de recogida de datos,
- si el experimento piloto tiene un tamaño suficientemente grande puede ayudar a seleccionar un modelo adecuado al experimento principal,
- los errores experimentales observados en el experimento piloto pueden ayudar a calcular el número de observaciones que se precisan en el experimento principal.

## 6. Especificar el modelo.

El modelo matemático especificado debe indicar la relación que se supone que existe entre la variable respuesta y las principales fuentes de variación identificadas en el paso 2. Es fundamental que el modelo elegido se ajuste a la realidad con la mayor precisión posible.

El modelo más habitual es el modelo lineal:

$$Y = \sum_{i=1}^k \alpha_i + \varepsilon.$$

En este modelo la respuesta viene dada por una combinación lineal de términos que representan las principales fuentes de variación planificada más un término residual debido a las fuentes de variación no planificada. Los modelos que se estudian en este texto se ajustan a esta forma general. El experimento piloto puede ayudar a comprobar si el modelo se ajusta razonablemente bien a la realidad.

Los modelos de diseño de experimentos, según sean los factores incluidos en el mismo, se pueden clasificar en: modelo de efectos fijos, modelo de efectos aleatorios y modelos mixtos. A continuación se precisan estas definiciones.

**Factor de efectos fijos** es un factor en el que los niveles han sido seleccionados por el experimentador. Es apropiado cuando el interés se centra en comparar el efecto sobre la respuesta de esos niveles específicos.

Ejemplo: un empresario está interesado en comparar el rendimiento de tres máquinas del mismo tipo que tiene en su empresa.

**Factor de efectos aleatorios** es un factor del que sólo se incluyen en el experimento una muestra aleatoria simple de todos los posibles niveles del mismo. Evidentemente se utilizan estos factores cuando tienen un número muy grande de niveles y no es razonable o posible trabajar con todos ellos. En este caso se está interesado en examinar la variabilidad de la respuesta debida a la población entera de niveles del factor.

Ejemplo: una cadena de hipermercados que tiene en plantilla 300 trabajadores de caja está interesada en estudiar la influencia del factor trabajador en la variable “tiempo en el cobro a un cliente”.

**Modelo de efectos fijos** es un modelo en el que todos los factores son factores de efectos fijos.

**Modelo de efectos aleatorios** es un modelo en el que todos los factores son factores de efectos aleatorios.

**Modelo mixto** es un modelo en el que hay factores de efectos fijos y factores de efectos aleatorios.

## **7. Esquematizar los pasos del análisis estadístico.**

El análisis estadístico a realizar depende de:

- los objetivos indicados en la etapa 1,
- el diseño seleccionado en la etapa 3,
- el modelo asociado que se especificó en la etapa 5.

Se deben esquematizar los pasos del análisis a realizar que deben incluir:

- estimaciones que hay que calcular,
- contrastes a realizar,
- intervalos de confianza que se calcularán
- diagnosis y crítica del grado de ajuste del modelo a la realidad.

## **8. Determinar el tamaño muestral.**

Calcular el número de observaciones que se deben tomar para alcanzar los objetivos del experimento.

Existen, dependiendo del modelo, algunas fórmulas para determinar este tamaño. Todas ellas sin embargo requieren el conocimiento del tamaño de la variabilidad no planificada (no sistemática y sistemática, si es el caso) y estimarlo a priori no es fácil, siendo aconsejable sobreestimarla. Normalmente se estima a partir del experimento piloto y en base a experiencias previas en trabajos con diseños experimentales semejantes.

## **9. Revisar las decisiones anteriores. Modificar si es necesario.**

De todas las etapas enumeradas, el proceso de recolección de datos suele ser la tarea que mayor tiempo consume, pero es importante realizar una planificación

previa, detallando los pasos anteriores, lo que garantizará que los datos sean utilizados de la forma más eficiente posible.

Es fundamental tener en cuenta que:

“Ningún método de análisis estadístico, por sofisticado que sea, permite extraer conclusiones correctas en un diseño de experimentos mal planificado”.

Recíprocamente, debe quedar claro que el análisis estadístico, es una etapa más que está completamente integrada en el proceso de planificación.

“El análisis estadístico no es un segundo paso independiente de la tarea de planificación. Es necesario comprender la totalidad de objetivos propuestos antes de comenzar con el análisis. Si no se hace así, tratar que el experimento responda a otras cuestiones a posteriori puede ser (lo será casi siempre) imposible”.

Pero no sólo los objetivos están presentes al inicio del análisis sino también la técnica experimental empleada. Una regla de oro en la experimentación y que debe utilizarse es la siguiente:

“No invertir nunca todo el presupuesto en un primer conjunto de experimentos y utilizar en su diseño toda la información previa disponible”.

Finalmente indicar que todas las personas que trabajan en el experimento se deben implicar en el mismo, esto es:

“Toda persona implicada en la ejecución del experimento y en la recolección de los datos debe ser informada con precisión de la estrategia experimental diseñada”.

## 1.5 Principios básicos en el diseño de experimentos.

Al planificar un experimento hay tres principios básicos que se deben tener siempre en cuenta:

- El principio de aleatorización.
- El bloqueo.
- La factorización del diseño.

Los dos primeros (aleatorizar y bloquear) son estrategias eficientes para asignar los tratamientos a las unidades experimentales sin preocuparse de qué tratamientos considerar. Por el contrario, la factorización del diseño define una estrategia eficiente para elegir los tratamientos sin considerar en absoluto como asignarlos después a las unidades experimentales.

### **Aleatorizar**

“Aleatorizar todos los factores no controlados por el experimentador en el diseño experimental y que pueden influir en los resultados serán asignados al azar a las unidades experimentales”.

Ventajas de aleatorizar los factores no controlados:

- Transforma la variabilidad sistemática no planificada en variabilidad no planificada o ruido aleatorio. Dicho de otra forma, aleatorizar previene contra la introducción de sesgos en el experimento.
- Evita la dependencia entre observaciones al aleatorizar los instantes de recogida muestral.
- Valida muchos de los procedimientos estadísticos más comunes.

## **Bloquear**

“Se deben dividir o particionar las unidades experimentales en grupos llamados bloques de modo que las observaciones realizadas en cada bloque se realicen bajo condiciones experimentales lo más parecidas posibles.

A diferencia de lo que ocurre con los factores tratamiento, el experimentador no está interesado en investigar las posibles diferencias de la respuesta entre los niveles de los factores bloque”.

Bloquear es una buena estrategia siempre y cuando sea posible dividir las unidades experimentales en grupos de unidades similares.

La ventaja de bloquear un factor que se supone que tienen una clara influencia en la respuesta pero en el que no se está interesado, es la siguiente:

- Convierte la variabilidad sistemática no planificada en variabilidad sistemática planificada.

Con el siguiente ejemplo se trata de indicar la diferencia entre las estrategias de aleatorizar y de bloquear en un experimento.

### **Ejemplo 1.1**

Se desea investigar las posibles diferencias en la producción de dos máquinas, cada una de las cuales debe ser manejada por un operario.

En el planteamiento de este problema la variable respuesta es “la producción de una máquina (en un día)”, el factor-tratamiento en el que se está interesado es el “tipo de máquina” que tiene dos niveles y un factor nuisance es el “operario que maneja la máquina”. En el diseño del experimento para realizar el estudio se

pueden utilizar dos estrategias para controlar el factor “operario que maneja la máquina”.

**Aleatorizar:** se seleccionan al azar dos grupos de operarios y se asigna al azar cada grupo de operarios a cada una de las dos máquinas. Finalmente se evalúa la producción de las mismas.

**Bloquear:** se introduce el factor-bloque “operario”. Se elige un único grupo de operarios y todos ellos utilizan las dos máquinas.

¿Qué consideraciones se deben tener en cuenta al utilizar estas dos estrategias? ¿Qué estrategia es mejor?

### **La factorización del diseño.**

“Un diseño factorial es una estrategia experimental que consiste en cruzar los niveles de todos los factores tratamiento en todas las combinaciones posibles”.

Ventajas de utilizar los diseños factoriales:

- Permiten detectar la existencia de efectos interacción entre los diferentes factores tratamiento.
- Es una estrategia más eficiente que la estrategia clásica de examinar la influencia de un factor manteniendo constantes el resto de los factores.

### **1.6 Algunos diseños experimentales clásicos.**

Un diseño experimental es una regla que determina la asignación de las unidades experimentales a los tratamientos. Aunque los experimentos difieren unos de otros en muchos aspectos, existen diseños estándar que se utilizan con mucha frecuencia. Algunos de los más utilizados son los siguientes:

### 1.6.1 Diseño completamente aleatorizado.

El experimentador asigna las unidades experimentales a los tratamientos al azar. La única restricción es el número de observaciones que se toman en cada tratamiento. De hecho si  $n_i$  es el número de observaciones en el  $i$ -ésimo tratamiento,  $i = 1, \dots, l$ , entonces, los valores  $n_1, n_2, \dots, n_l$  determinan por completo las propiedades estadísticas del diseño. Naturalmente, este tipo de diseño se utiliza en experimentos que no incluyen factores bloque.

El modelo matemático de este diseño tiene la forma:

$$\text{Respuesta} = \text{Constante} + \text{Efecto tratamiento} + \text{Error}$$

### 1.6.2 Diseño en bloques o con un factor bloque.

En este diseño el experimentador agrupa las unidades experimentales en bloques, a continuación determina la distribución de los tratamientos en cada bloque y, por último, asigna al azar las unidades experimentales a los tratamientos dentro de cada bloque.

En el análisis estadístico de un diseño en bloques, éstos se tratan como los niveles de un único factor de bloqueo, aunque en realidad puedan venir definidos por la combinación de niveles de más de un factor nuisance.

El modelo matemático de este diseño es:

$$\text{Respuesta} = \text{Constante} + \text{Efecto bloque} + \text{Efecto tratamiento} + \text{Error}$$

El diseño en bloques más simple es el denominado **diseño en bloques completos**, en el que cada tratamiento se observa el mismo número de veces en cada bloque.

El diseño en bloques completos con una única observación por cada tratamiento se denomina **diseño en bloques completamente aleatorizado** o, simplemente, **diseño en bloques aleatorizado**.

Cuando el tamaño del bloque es inferior al número de tratamientos no es posible observar la totalidad de tratamientos en cada bloque y se habla entonces de **diseño en bloques incompletos**.

### 1.6.3 Diseños con dos o más factores bloque.

En ocasiones hay dos (o más) fuentes de variación lo suficientemente importantes como para ser designadas factores de bloqueo. En tal caso, ambos factores bloque pueden ser **cruzados o anidados**.

Los factores bloque están **cruzados** cuando existen unidades experimentales en todas las combinaciones posibles de los niveles de los factores bloques.

**Diseño con factores bloque cruzados.** También denominado **diseño fila-columna**, se caracteriza porque existen unidades experimentales en todas las celdas (intersecciones de fila y columna).

El modelo matemático de este diseño es:

$$\text{Respuesta} = \text{Constante} + \text{Efecto bloque fila} + \text{Efecto bloque columna} + \text{Efecto tratamiento} + \text{Error}$$

Los factores bloque están **anidados** si cada nivel particular de uno de los factores bloque ocurre en un único nivel del otro factor bloque.

**Diseño con factores bloque anidados o jerarquizados.** Dos factores bloque se dicen anidados cuando observaciones pertenecientes a dos niveles distintos de

un factor bloque están automáticamente en dos niveles distintos del segundo factor bloque.

**Tabla 1. Diferencia entre ambos tipos de bloqueo anidados y jerarquerizados.**

Bloques Cruzados					Bloques Anidados				
			Bloque 1				Bloque 1		
B		1	2	3	B		1	2	3
L	1	+	+	+	L	1	+		
O	2	+	+	+	O	2	+		
Q	3	+	+	+	Q	3	+		
U					U	4		+	
E					E	5		+	
						6		+	
						7			+
						8			+
						9			+

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

#### 1.6.4 Diseños con dos o más factores.

En algunas ocasiones se está interesado en estudiar la influencia de dos (o más) factores tratamiento, para ello se hace un diseño de filas por columnas. En este modelo es importante estudiar la posible interacción entre los dos factores. Si en cada casilla se tiene una única observación no es posible estudiar la interacción entre los dos factores, para hacerlo hay que replicar el modelo, esto es, obtener k observaciones en cada casilla, donde k es el número de réplicas.

El modelo matemático de este diseño es:

$$\text{Respuesta} = \text{Constante} + \text{Efecto Factor Fila} + \text{Efecto Factor Columna} + \text{Efecto Interacción} + \text{Error}$$

Generalizar los diseños completos a más de dos factores es relativamente sencillo desde un punto de vista matemático, pero en su aspecto práctico tiene el inconveniente de que al aumentar el número de factores aumenta muy rápidamente el número de observaciones necesario para estimar el modelo. En la práctica es muy raro utilizar diseños completos con más de factores.

Un camino alternativo es utilizar **fracciones factoriales** que son diseños en los que se supone que muchas de las interacciones son nulas, esto permite estudiar el efecto de un número elevado de factores con un número relativamente pequeño de pruebas. Por ejemplo, el diseño en **cuadrado latino**, en el que se supone que todas las interacciones son nulas, permite estudiar tres factores de  $k$  niveles con solo  $k^2$  observaciones. Si se utilizase el diseño equilibrado completo se necesitan  $k^3$  observaciones.

#### **1.6.5 Diseños factoriales a dos niveles.**

En el estudio sobre la mejora de procesos industriales (control de calidad) es usual trabajar en problemas en los que hay muchos factores que pueden influir en la variable de interés. La utilización de experimentos completos en estos problemas tiene el gran inconveniente de necesitar un número elevado de observaciones, además puede ser una estrategia ineficaz porque, por lo general, muchos de los factores en estudio no son influyentes y mucha información recogida no es relevante. En este caso una estrategia mejor es utilizar una técnica secuencial donde se comienza por trabajar con unos pocos factores y según los resultados que se obtienen se eligen los factores a estudiar en la segunda etapa.

Los **diseños factoriales**  $2^k$  son diseños en los que se trabaja con  $k$  factores, todos ellos con dos niveles (se suelen denotar + y -). Estos diseños son adecuados para tratar el tipo de problemas descritos porque permiten trabajar con un número elevado de factores y son válidos para estrategias secuenciales.

Si  $k$  es grande, el número de observaciones que necesita un diseño factorial  $2^k$  es muy grande ( $n = 2^k$ ). Por este motivo, las **fracciones factoriales**  $2^{k-p}$  son muy utilizadas, éstas son diseños con  $k$  factores a dos niveles, que mantienen la propiedad de ortogonalidad de los factores y donde se suponen nulas las interacciones de orden alto (se confunden con los efectos simples) por lo que para su estudio solo se necesitan  $2^{k-p}$  observaciones (cuanto mayor sea  $p$  menor número de observaciones se necesita pero mayor confusión de efectos se supone).

En los últimos años Taguchi ha propuesto la utilización de fracciones factoriales con factores a tres niveles en problemas de control de calidad industriales.

## **CAPÍTULO 2**

### **DISEÑO DE BLOQUES INCOMPLETOS**

#### **2.1 Introducción**

En ocasiones es necesario bloquizar unidades experimentales en grupos más pequeños que una réplica completa de todos los tratamientos que se usarían con bloques completos aleatorizado o un diseño de cuadrado latino. El diseño de bloques incompletos se usa para disminuir la varianza de error experimental y proporcionar comparaciones más precisas entre tratamientos de lo que es posible con el diseño de bloques completos. En este trabajo se presenta una descripción general de algunos grupos importantes de diseño de bloques incompletos, se muestran el método de aleatorización y los métodos básicos de análisis para diseño de bloques incompletos balanceados (BIB) y bloques incompletos parcialmente balanceados (BIPB), tomando en cuenta también la eficiencia de los diseños.

## 2.2 Bloques incompletos de tratamientos para reducir el tamaño de los bloques

Por alguna razón los experimentos pueden requerir una reducción del tamaño de bloque. Los diseños de bloques completos pueden reducir las varianzas del error experimental estimadas, pero esta reducción puede ser insuficiente, pues el número de tratamientos pueden ser tan grandes que resultan imprácticos para reducirla. Además, con el agrupamiento natural de las unidades experimentales en bloques quizá no se obtengan las unidades por bloques necesarias para el número de tratamientos de un diseño de bloques completos. En el siguiente ejemplo, el número limitado de cámaras de control ambiental evita la ejecución de una réplica completa de todos los tratamientos en una corrida de cámaras disponibles.

### **Ejemplo 2.1                    Generación de semillas de tomate a alta temperatura constante**

Es usual que los tomates se produzcan durante los meses de invierno en las regiones áridas tropicales, la producción invernal se siembra a fines del verano cuando las temperaturas del suelo pueden exceder 40°C, lo que sobrepasa los 35°C, temperatura máxima sugerida de germinación.

**Objetivo de investigación:** Un científico de plantas deseaba determinar a que intervalos de temperatura podía esperar la inhibición de la germinación de las semillas de tomate para un grupo de cultivos.

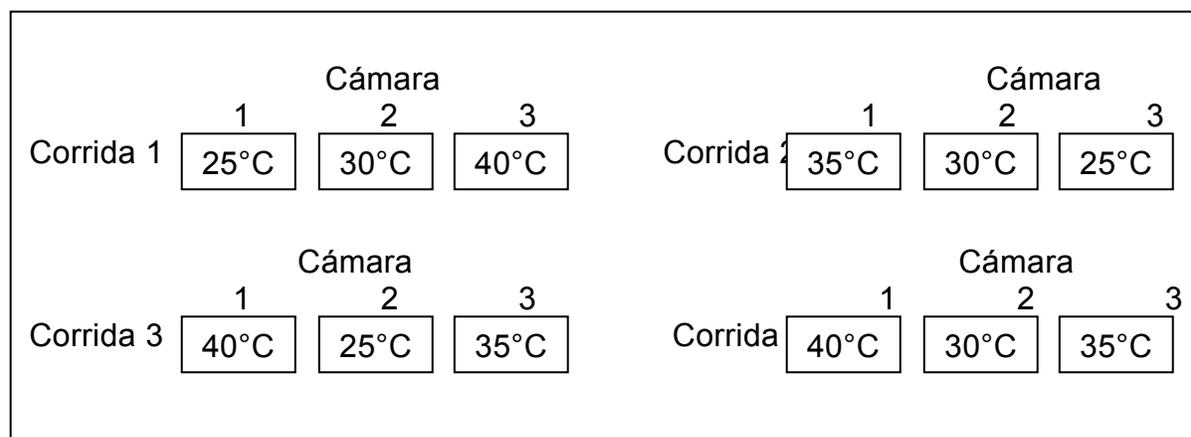
**Diseño del tratamiento:** se eligieron cuatro temperaturas para representar un intervalo común para el área de cultivo en consideración: 25°C, 30°C, 35°C y 40°C. La semilla de tomate se sometió a una temperatura constante cámaras de ambiente controlado.

**Diseño experimental:** una cámara sería una unidad experimental, pues la réplica verdadera de cualquier tratamiento de temperatura requería una corrida independiente del tratamiento en una cámara. Cualquier número de factores puede contribuir a la variación de la respuesta entre corridas, ya que las condiciones de todo el experimento debían repetirse para una corrida réplica, por lo que se consideró esencial bloquizar las corridas.

El bloque completo y la réplica del experimento requerían cuatro cámaras; sin embargo el científico sólo disponía de tres. Como el bloque natural de una corrida tenía menos cámaras (unidades experimentales) que tratamientos, construyó un diseño de bloques incompletos.

El cuadro 1 muestra un diagrama de diseño, en el cual se probaron tres temperaturas diferentes en cada una de las cuatro corridas. Las corridas representan bloques incompletos de los tres tratamientos de temperatura y los tratamientos se asignaron al azar a las cámaras para la corrida. Algunas características especiales de este diseño se analizan en la siguiente sección.

**Cuadro 1** Diseño de bloques incompletos con cuatro tratamientos en bloques de tres unidades.



Fuente: Dr. J. Coons, Departamento de Botany, Eastern Illinois University.

Otros ejemplos: Los lotes de material para la investigación industrial sirven de bloques, pero puede no haber material suficiente en uno solo para todos los tratamientos experimentales; los criterios para agrupar sujetos pueden provocar que el número de sujetos parecidos sea insuficiente en cada grupo para asignar los tratamientos planeados para el estudio, la variedad agrícola de las pruebas con frecuencia contiene un gran número de especies y los bloques completos no son prácticos para reducir la varianza del error; los diseños de los bloques incompletos son la elección adecuada para los experimentos del tipo de estos ejemplos.

La guía principal para el tamaño del bloque es tener un conjunto homogéneo de unidades experimentales para obtener comparaciones precisas entre los tratamientos. Los diseños de bloques incompletos fueron introducidos por Yates (1936a, 1936b) para experimentos en los que el número de unidades experimentales por bloque es menor que el número de tratamientos. Los diseños se desarrollaron ante la necesidad de experimentos que incluyeran el conjunto relevante de tratamientos para estudiar la hipótesis de investigación a un cuando estaban restringidos a tamaños de bloque lógicos.

Es posible hacer una clasificación del diseño de bloques incompleto en dos grandes grupos: aquellos arreglados como bloques incompletos aleatorizados con un criterio de bloque y aquellos con arreglos basados en el cuadro latino con dos criterios de bloque. Los diseños también puede ser balanceados, donde cada tratamiento se aparee un número igual de veces con los demás tratamientos en un mismo bloque, considerando todos los bloques en el experimento. Un diseño parcialmente balanceado ocurre cuando se tiene diferente número de pares de tratamientos en el mismo bloque o cuando algún par de tratamientos nunca ocurre en el mismo bloque. En este tema se presenta una descripción general de los diseños de bloques incompletos y una introducción al análisis de los datos correspondientes a estos diseños.

### 2.3 Diseños de bloques incompletos balanceados (BIB)

#### El diseño BIB compara todos los tratamientos con igual precisión

El **diseño de bloques incompletos balanceado** es un arreglo tal que todos los tratamientos tienen igual número de réplicas y cada par de tratamientos se presenta en el mismo bloque un número igual de veces en algún lugar del diseño, el balance obtenido con el mismo número de ocurrencias de todos los pares de tratamientos en el mismo bloque tiene como resultado una precisión igual en todas las comparaciones entre los pares de medias de tratamiento.

El diseño de bloque incompleto tiene  $r$  réplicas de  $t$  tratamientos en  $b$  bloques de  $k$  unidades experimentales con  $K < t$  y el número total de unidades experimentales en  $N = rt = bk$ , en el cuadro 1 por ejemplo, el experimento con tomates descrito tiene  $b = 4$  bloques de  $k = 3$  unidades experimentales, cada uno de los  $t = 4$  tratamientos tiene  $r = 3$  réplicas y existe un total de  $N = bk = 4 \cdot 3 = 12$ , o sea,  $N = rt = 3 \cdot 4 = 12$  unidades experimentales. Por inspección se observa que cada par de tratamientos ocurre dos veces en los bloques, el par  $(25^\circ, 30^\circ)$  está en los bloques 1 y 2 y el par  $(30^\circ, 35^\circ)$  en los bloques 2 y 4.

El número de bloques donde ocurre cada par de tratamientos es  $\lambda = r(k-1)(t-1)$ , donde  $\lambda < r < b$ . el valor entero  $\lambda$  se deriva del hecho de que cada tratamiento está apareado con los otros  $t-1$  tratamientos en algún lugar del diseño  $\lambda$  veces. Existen  $\lambda(t-1)$  pares para un tratamiento específico en el experimento, el mismo tratamiento aparece en  $r$  bloques con  $k-1$  de los otros tratamientos, y cada tratamiento aparece en  $r(k-1)$  pares. Por lo tanto:

$$\lambda(t-1) = r(k-1) \quad \text{o} \quad \lambda = r(k-1)/(t-1)$$

Es decir en el diseño del experimento con tomates del cuadro 1  $\lambda = 3(3-1)/(4-1) = 2$ .

Se puede construir un diseño de bloques incompletos balanceados mediante la asignación de las combinaciones adecuadas de tratamientos a cada uno de los  $b = \binom{t}{k}$  bloques, con frecuencia es posible obtener el balance con menos de  $\binom{t}{k}$  bloques. No existe un método único para construir todas las clases de diseños de bloques incompletos pero existen métodos para construir algunos. La construcción de estos diseños ha sido tema de diversas investigaciones matemáticas que han arrojado un vasto arreglo de diseños de bloques balanceados y parcialmente balanceados.

El apéndice 1A.1 proporciona algunos planes para pocos tratamientos y se pueden encontrar otras tablas con diseños útiles para muchas situaciones prácticas en Cochran y Cox (1957) y en Fisher y Yates (1963). En el capítulo 3 se ilustrarán con más detalle varias categorías de diseño de bloques incompletos balanceados tradicionales, en tanto este capítulo presenta una introducción a su estructura básica con ejemplos de su aplicación y análisis.

## 2.4 Cómo aleatorizar los diseños de bloques incompletos

Una vez construido el diseño básico con los números de código de los tratamientos, los pasos para la aleatorización son las siguientes:

**Paso 1.** Aleatorizar el arreglo de los bloques conformados por grupos de los números de código asignados a los tratamientos.

**Paso 2.** Aleatorizar el arreglo de los números de tratamiento dentro de cada bloque.

**Paso 3.** Aleatorizar la asignación de los tratamientos a los números de código en el plan.

A continuación se ilustra la aleatorización con el plan de diseño básico para  $t = 4$  tratamientos en  $b = 4$  bloques de  $k = 3$  unidades experimentales cada uno. Antes de la aleatorización, el plan es:

Bloque			
1	<u>1</u>	<u>2</u>	3
2	<u>1</u>	<u>2</u>	4
3	<u>1</u>	<u>3</u>	4
4	<u>2</u>	<u>3</u>	4

**Paso 1.** Suponiendo que los bloques para el experimento son las corridas de las tres cámaras de cultivo con tres tratamientos de temperatura usadas en el ejemplo 2.1 los grupos de tratamiento (1, 2, 3), (1, 2, 4), (1, 3, 4) y (2, 3, 4) deben asignarse de manera aleatoria a las corridas, se elige una permutación aleatoria de los números 1 a 4 y se asignan los cuatro bloques a las cuatro corridas; con la permutación 2, 4, 1, 3 la asignación es:

Bloque				Bloque original
1	<u>1</u>	<u>2</u>	4	2
2	<u>2</u>	<u>3</u>	4	4
3	<u>1</u>	<u>2</u>	3	1
4	<u>1</u>	<u>3</u>	4	3

**Paso 2.** Se asignan al azar los números de código de tratamiento a las cámaras de cultivo en cada corrida. Se elige una permutación de los números 1 a 4 para cada cámara y se omite el número de tratamiento ausente en la corrida. Las siguientes son las cuatro permutaciones junto con la asignación a cada cámara (A, B, C) en cada corrida:

Bloque	Cámara			Permutación
	A	B	C	
1	2	4	1	2 4 / 3 1
2	3	4	2	3 4 / 1 2
3	1	2	3	/4 1 2 3
4	1	4	3	1 4 3 / 2

**Paso 3.** Suponiendo también que los tratamientos son las temperaturas 25°, 30°, 35° y 40°C, una permutación aleatoria 2, 4, 3, 1 produce una asignación de las temperaturas para los números de código de tratamiento con la sustitución 2→25°C, 4→30°C, 3→35°C y 1→40°C en la tabla anterior.

Corrida	Cámara		
	A	B	C
1	25°C	30°C	40°C
2	35°C	30°C	25°C
3	40°C	25°C	35°C
4	40°C	30°C	35°C

## 2.5 Análisis de diseños BIB

### Modelo estadístico para los diseños BIB

El modelo lineal estadístico para un diseño de bloques incompleto balanceado es:

$$\gamma_{ij} = \mu + \tau_i + \rho_j + e_{ij}$$

$$i= 1,2,\dots,t \quad j= 1,2,\dots,b \quad (1)$$

donde  $\mu$  es la media general,  $\tau_i$  es el efecto fijo del  $i$ -ésimo tratamiento,  $\rho_j$  es el efecto fijo del  $j$ -ésimo bloque, y las  $e_{ij}$  son errores experimentales aleatorios independientes con media 0 y varianza  $\sigma^2$ .

Recordando que se tiene  $r$  réplicas de los  $t$  tratamientos en  $b$  bloques incompletos de  $k$  unidades experimentales, el número total de observaciones es  $N = rt = bk$ , donde cada par de tratamientos aparece junto en  $\lambda = r(k-1)/(t-1)$  bloques del experimento.

Los efectos de tratamiento y bloques no son ortogonales en el diseño de bloques incompletos porque no aparecen todos los tratamientos en cada bloque; por lo tanto, no sería correcto para el diseño incompleto calcular la participación de suma de cuadrados para los tratamientos igual que para los diseños de bloque incompletos; tampoco las medias de tratamiento observadas proporcionarían estimaciones no sesgadas de  $\mu_i = \mu + \tau_i$ . Los parámetros estimados y las sumas de los cuadrados de los tratamientos para los diseños de bloque incompletos balanceados se calculan con formulas relativamente directas, el desarrollo de las estimaciones de mínimos cuadrados se encuentra en el apéndice 1A.3.

### **Particiones de sumas de cuadrados para diseños BIB**

Las particiones de sumas de cuadrados se pueden derivar al considerar modelos alternativos completos y reducidos para el diseño. Las soluciones de las ecuaciones normales se obtienen para el modelo completo,  $\gamma_{ij} = \mu + \tau_i + \rho_j + e_{ij}$  con las estimaciones,  $\hat{\gamma}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\rho}_j$ , para calcular la suma de cuadrados del error experimental para el modelo completo:

$$SCE_f = \sum_i \sum_j (\gamma_{ij} - \hat{\gamma}_{ij})^2 = \sum_i \sum_j (\gamma_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\rho}_j)^2 \quad (2)$$

donde  $SCE_f$ : Suma de Cuadrados para el Error (modelo completo).

Las soluciones a las ecuaciones normales para el modelo reducido,  $\gamma_{ij} = \mu + \tau_i + \rho_j + e_{ij}$ , con estimaciones,  $\hat{\gamma}_{ij} = \hat{\mu} + \hat{\rho}_j$ , se usan para calcular las sumas de cuadrados del error experimental para el modelo reducido;

$$SCE_r = \sum_i \sum_j (\gamma_{ij} - \hat{\gamma}_{ij})^2 = \sum_i \sum_j (\gamma_{ij} - \hat{\mu} - \hat{\rho}_j)^2 \quad (3)$$

donde  $SCE_r$ : Suma de Cuadrados para el Error (modelo reducido).

La diferencia  $SCE_r - SCE_f$  es la reducción en la suma de cuadrados que resulta al incluir  $\tau_i$  en el modelo completo, se trata de la suma de cuadrados debida a los tratamientos después de considerar los efectos de bloque en el modelo, esto se conoce como SC tratamiento (ajustada) e implica que los efectos de bloque también se toman en cuenta al estimar los efectos de tratamiento en el modelo completo. Para los diseños de bloque incompletos balanceados, la suma de cuadrados de los tratamientos ajustada,  $SCE_r - SCE_f$ , se puede calcular directamente como:

$$SCT(ajustada) = \frac{k \sum_{i=1}^t Q_i^2}{\lambda t} \quad (4)$$

donde  $SCT$ : Suma de Cuadrados para Tratamientos.

Con  $t - 1$  grados de libertad. La cantidad  $Q_i$  es un total de tratamiento ajustado que se calcula con:

$$Q_i = y_{i.} - \frac{1}{k} B_i \quad (5)$$

donde  $B_i = \sum_j^b n_{ij} y_j$  es la suma de todos los totales de bloques que incluyen el  $i$ -ésimo tratamiento y  $n_{ij} = 1$  si el tratamiento  $i$  aparece en el bloque  $j$  y  $n_{ij} = 0$  de otra manera. Esta corrección al total del tratamiento tiene el efecto neto de eliminar los efectos de bloque de ese total.

La suma de cuadrados para los bloques se deriva del modelo reducido al ignorar los tratamientos en el modelo con  $SCB(\text{ajustada}) = SC \text{ total} - SCE_r$ . Los efectos de tratamiento no se consideran cuando se estiman los efectos de bloque, y la suma de cuadrados recibe el nombre de suma de cuadrados *no ajustada*.

La partición aditiva de  $SC$  total es:

$$SC \text{ total} = SCB(\text{no ajustada}) + SCT(\text{ajustada}) + SCE_r \quad (6)$$

donde  $SC$ : Suma de Cuadrados

$SCB$ : Suma de Cuadrados para Bloques

$SCT$ : Suma de Cuadrados para Tratamientos

$SCE_r$ : Suma de Cuadrados para el Error (modelo completo).

La tabla 2.1 muestra una descripción del análisis de varianza para las particiones de suma de cuadrados. Muchos programas de computadora (SPSS, SAS, S+ y otros) tienen la capacidad de calcular las particiones correctas de suma de cuadrados, las estimaciones de mínimos cuadrados de las medias de tratamiento, los contrastes entre las medias de mínimos cuadrados y sus errores estándar, para el diseño de bloques incompletos.

**Tabla 2.1 Análisis de varianza para un diseño de bloque incompleto balanceado**

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios
Total	$N - 1$	$\sum (\gamma_{ij} - \bar{\gamma}_{..})^2$	
Bloques	$b - 1$	$\kappa \sum (\bar{\gamma}_{.j} - \bar{\gamma}_{..})^2$	<b>CMB</b> (no adj.)
Tratamientos	$t - 1$	$\frac{\kappa \sum Q_i^2}{\lambda t}$	<b>CMT</b> (aj.)
Error	$N - t - b - 1$	Restando	<b>CME</b>

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.316

Donde *CMB*: Cuadrados Medios para Bloques

*CMT*: Cuadrados Medios para Tratamientos

*CME*: Cuadrados Medios para el Error

### **Ejemplo 2.2 Vinilación de metilglucósidos**

Se ha encontrado que el resultado de la adición de acetileno a un metilglucósido en presencia de una base, a alta presión, es la producción de varios éteres monovinílicos, proceso que se conoce como vinilación. Los éteres monovinílicos son convenientes para polimerización en muchas aplicaciones industriales.

**Objetivo de investigación:** Los químicos deseaban obtener más información específica sobre el efecto de la presión en el porcentaje de conversión del metilglucósido en isómeros de monovinil.

**Diseño de tratamiento:** Con base en trabajos anteriores, se seleccionaron presiones dentro del intervalo que producía la conversión máxima. Se eligieron 5 presiones para estimar una ecuación de respuesta: 250, 325, 400, 475 y 550 psi.

**Diseño del experimento:** como sólo se disponía de tres cámaras de alta presión para una corrida de las condiciones experimentales, fue necesario bloquizar las corridas por que no podía haber una variación sustancial de una corrida a otra producida por nuevas preparaciones de las cámaras para el experimento. Los químicos establecieron un diseño de bloques incompleto balanceado (corridas), cada uno con tres unidades experimentales (cámaras presurizadas) y se usaron tres presiones diferentes en cada corrida; el diseño obtenido tenía seis replicas de cada tratamiento de presión.

Las presiones usadas en cada corrida y las conversiones porcentuales a isómeros de monovinil se muestran en la tabla 2.2 y la suma aditiva de la partición de suma de cuadrados del metilglucósido se muestra en la tabla 2.3.

**Tabla 2.2** Conversión porcentual del metilglucósidos con acetileno a acetileno a alta presión, en un diseño de bloque incompleto balanceado.

Corrida	Presion (psi)					$y_j$
	250	325	400	475	550	
1	16	18	–	32	–	66
2	19	–	–	46	45	110
3	–	26	39	–	61	126
4	–	–	21	35	55	111
5	–	19	–	47	48	114
6	20	–	33	31	–	84
7	13	13	34	–	–	60
8	21	–	30	–	52	103
9	24	10	–	–	50	84
10	–	24	31	37	–	92
$y_i.$	113	110	188	228	311	950
$*B_i.$	507	542	576	577	648	
$^{\dagger}Q_i$	-56.0	-70.7	-4.0	35.7	95.0	

\*Ejemplo:  $B_1 = y_{.1} + y_{.2} + y_{.6} + y_{.7} + y_{.8} + y_{.9}$   
 $^{\dagger}Q_i = y_{i.} - \frac{1}{3} B_i = 113 - \frac{1}{3} (507) = -56.0$

Fuente: Drs J. Berry y A. Deutschman, University of Arizona

### Inferencias para las medias de tratamiento

Las estimaciones de los mínimos cuadrados de las medias de tratamiento y sus errores estándar estimados se muestran en la tabla 2.4 y la estimación de mínimos cuadrados para una media de tratamiento  $\mu_i$  es  $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i$ , donde:

**Tabla 2.3** Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloque incompleto balanceado.

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr<F
Total	29	5576.67			
Bloques (no. Aj.)	9	1394.67	154.96		
Presión (aj.)	4	3688.58	922.14	29.90	.000
Error	16	493.42	30.84		

Fuente: Drs J. Berry y A. Deutschman, University of Arizona

$$\hat{\mu} = \gamma_{..} \quad \text{y} \quad \hat{\tau}_i = \frac{\kappa Q_i}{\lambda t} \quad (7)$$

Por ejemplo, de la tabla 2.2,  $Q_1 = y \quad \hat{\mu} = \gamma_{..} = 950/30 = 31.67$ , de manera que:

$$\hat{\tau}_i = \frac{\kappa Q_i}{\lambda t} = \frac{3}{(3)(5)} (-56.00) = -11.20 \quad \text{y} \quad \hat{\mu} = 31.67 - 11.20 = 20.47$$

### Errores estándar para las medidas de tratamiento

El error estándar para una estimación de la media de tratamiento es:

$$S_{\hat{\mu}_i} = \sqrt{\frac{CME}{rt} \left( 1 + \frac{kr(t-1)}{\lambda \tau} \right)} = \sqrt{\frac{30.84}{(6)(5)} \left( 1 + \frac{(3)(6)(4)}{(3)(5)} \right)} = 2.44 \quad (8)$$

Una estimación de un intervalo de confianza del 95% para una media de tratamiento en la tabla 2.4 es  $\hat{\mu}_i \pm t_{0.025,16}(S_{\hat{\mu}_i})$ , donde  $t_{0.025,16} = 2.120$ . El error estándar de la diferencia estimada entre dos medias de tratamiento,  $\hat{\mu}_i = \hat{\mu}_j$ , es:

$$S_{(\hat{\mu}_i - \hat{\mu}_j)} = \sqrt{\frac{2kCME}{\lambda t}} = \sqrt{\frac{(2)(3)30.84}{(3)(5)}} = 3.51 \quad (9)$$

Tabla 2.4 Estimaciones de mínimos cuadrados de las medias de presión para la conversión porcentual de multiglucósidos, en un diseño de bloques incompletos balanceado.

Presión (psi)	$\hat{\mu}_i$	$s_{\hat{\mu}_i}$
250	20.47	2.44
325	17.53	2.44
400	30.87	2.44
475	38.80	2.44
250	50.67	2.44

Fuente: Drs J. Berry y A. Deutschman, University of Arizona

### Pruebas de hipótesis sobre las medias de tratamientos

El estadístico  $F_0$  para probar la hipótesis nula que no hay diferencias entre las medias de tratamiento es:

$$F_0 = \frac{CMT(adj.)}{CME} = \frac{922.14}{30.84} = 29.90 \quad (10)$$

donde *CMT*: Cuadrados Medios para Tratamientos

*CME*: Cuadrados Medios para el Error

con  $Pr > F = .000$  (tabla 2.3). El valor crítico para un nivel de significancia de .05 es  $F_{.05,4,16} = 3.01$ ; por lo tanto, existen diferencias significativas entre las presiones con respecto a la conversión de metilglucósidos en productos de Vinilación.

### Contrastes entre las medias de tratamientos

Los contrastes se calculan con las estimaciones de mínimos cuadrados de las medias de tratamiento como sigue:

$$c = \sum_{i=1}^t d_i \hat{\mu}_i \quad (11)$$

Con error estándar estimado de:

$$S_c = \sqrt{\frac{kCME}{\lambda t} \left( \sum_{i=1}^t d_i^2 \right)} \quad (12)$$

El estadístico  $t_0 = c/s_c$  se puede usar para probar la hipótesis nula  $H_0: C = 0$  con valor crítico basado en el estadístico  $t$  Student con  $N - t - b + 1$  grados de libertad.

La suma de cuadrados de 1 grado de libertad para el *contraste* se puede calcular con las medias de mínimos cuadrados  $\hat{\mu}_i$  como:

$$SCC = \frac{\lambda t \left( \sum d_i \hat{\mu}_i \right)^2}{k \sum d_i^2} \quad (13)$$

donde SCC: Suma de Cuadrados para Contrastes.

El estadístico  $F_0 = SCC/CME$  se usa para probar la hipótesis nula  $H_0: C = 0$  con valor crítico  $F_{\alpha,1,(N-t-b+1)}$ .

Donde *CME*: Cuadrados Medios para el Error

El tratamiento de presión para este experimento es un factor cuantitativo con cinco niveles y la regresión de la conversión porcentual de metilglucósido a presión usando contrastes polinomiales ortogonales para la presión describe el efecto de la presión sobre la tasa de conversión.

## Recuperación de la información de tratamientos a partir de la comparación de bloques

El análisis del diseño de bloques incompletos ilustrado hasta aquí estima los efectos de tratamientos con base en la información del tratamiento contenida en los bloques; esto se conoce como análisis *intra*bloques. Los diseños de bloques incompletos son no ortogonales debido a que no todos los tratamientos aparecen en todos los bloques, y las comparaciones entre bloques contienen cierta información sobre las comparaciones de tratamientos. Yates (1940a) demostró que esta información *inter*bloques se puede recuperar con un análisis interbloques y combinarse con la información del análisis intrabloque.

### Los contrastes de bloques contienen contraste de tratamientos

La información sobre las comparaciones de tratamientos contenida en una comparación entre bloques se puede ilustrar con las dos primeras corridas del experimento descrito en el ejemplo 2.2. Los datos de la conversión porcentual de metilglucósidos en las dos primeras corridas son.

Corrida	Presion (psi)					Media
	250	325	400	475	550	
1	16	18	–	32	–	22
2	19	–	–	46	45	37

Un contraste entre las medias de las dos corridas, 37 y 22, también es un contraste entre dos de los tratamientos de presión, 550 y 325 psi. Todas las comparaciones de los bloques contienen comparaciones similares de los tratamientos. El objetivo es recuperar esta información entre bloques, o

interbloques, sobre tratamientos y combinarla con la información dentro de los bloques, o intrabloques.

El tema de recuperación de la información interbloques se menciona aquí solo para indicar la disponibilidad del método, puede encontrarse un estudio exhaustivo en Kempthorne (1952), Cochran y Cox (1957), John (1971), John (1987) y John y Williams (1995).

La información del análisis interbloques se incorpora al análisis intrabloques con un estimador de los efectos de tratamiento que combina los estimadores intra e interbloques de los efectos de tratamiento; si el bloque ha sido eficaz para reducir el error experimental, la estimación interbloques contribuye sólo con una cantidad pequeña de información a las estimaciones combinadas, pero si los efectos del bloque son pequeños, la información recuperada con la estimación interbloques se reduce casi al análisis con la recuperación de la información interbloques se reduce casi al análisis normal sin el ajuste con bloques.

## **2.6 Diseño renglón-columna para dos criterios de bloque**

Cuando existe la necesidad de controlar la variación con más de un criterio de bloque, el diseño de cuadrado latino es un diseño de bloques completos usado para controlar la variación entre las unidades experimentales con dos factores de bloque; pero puede ser impracticable en algunas situaciones, pues el número de unidades experimentales que requiere,  $N = t^2$ , puede exceder las restricciones del material experimental o el número de tratamientos puede exceder el tamaño de los bloques disponibles.

Cuando se requieren dos criterios de bloque para el experimento es posible usar diseños renglón-columna con los renglones, las columnas o ambos como bloques incompletos; estos diseños se arreglan en  $p$  renglones y  $q$  columnas de

unidades experimentales. Consideremos el experimento clásico para probar cuatro llantas de automóvil en las cuatro posiciones de cuatro automóviles con un diseño de cuadrado latino y suponiendo que el equipo de investigación quiere evaluar  $t = 7$  tratamientos de llantas, todavía se tiene la necesidad de controlar la variación debida a la posición de la llanta y al automóvil, pero los autos solo tiene cuatro posiciones para probar siete llantas. Para este experimento se puede usar el diseño renglón-columna con un conjunto incompleto de tratamientos en cada columna, mostrado en el cuadro 2.

**Cuadro 2** Diseño de bloques renglón-columna incompleto balanceado de 4 X 7.

Posición	<i>Automóvil</i>						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	3	4	5	6	7	1	2
(2)	5	6	7	1	2	3	4
(3)	6	7	1	2	3	4	5
(4)	7	1	2	3	4	5	6

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.321

Los autos se usan como bloques incompletos con cuatro tratamientos evaluados en sus cuatro posiciones y las posiciones son bloques completos ya que cada tratamiento se evalúa en cada posición. Después de una inspección cuidadosa, se observa que cada par de tratamientos ocurre en un automóvil dos veces en algún lugar del experimento y, por lo tanto, el diseño de bloques incompleto está balanceado. El diseño tiene un balance natural respecto a las posiciones por que constituyen bloques completos. Este ejemplo es un caso en el que claramente era necesario un diseño de bloque incompleto porque había un número insuficiente de posiciones disponibles para probar todos los tratamientos a la vez.

## Los diseños ortogonales por renglón tiene una replica completa en cada uno

Dado que el diseño renglón-columna del cuadro 2 es un diseño de bloques completos para los renglones e incompleto y balanceado para las columnas, el diseño se conoce como un diseño ortogonal por renglón (John, 1987). Como cada tratamiento se presenta en cada renglón, los tratamientos son ortogonales a los renglones.

Youden (1937, 1940) desarrolló renglones de cuadrados latinos incompletos, conocidos ahora como *cuadrados de Youden*, omitiendo dos o más renglones del diseño de cuadrado latino. Los parámetros del diseño son  $t = b$ ,  $r = k$  y  $\lambda = k(k - 1)/(t - 1)$ . En el apéndice 1A.2 se muestran algunos planes de algunos cuadrados de Youden para experimentos pequeños; otros se pueden encontrar en Cochran Cox (1957) y Peterson (1985).

El diseño ortogonal por renglones del cuadro 2 es un cuadrado de Youden que tiene  $r = k = 4$  renglones como replicas,  $b = 7$  columnas y  $t = 7$  tratamientos, con  $\lambda = 2$  para los bloques de columna incompletos, las columnas son bloques incompletos de  $k = r = 4$  unidades y los renglones son bloques completos que contienen cada uno de los  $t = 7$  tratamientos.

### Descripción del análisis de diseños de renglón-columna

El modelo lineal para el diseño renglón-columna es:

$$y_{ijm} = \mu + \tau_i + p_j + \gamma_m + e_{ijm}$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, k \quad m = 1, 2, \dots, b \quad (14)$$

donde  $\mu$  es la media general,  $\tau_i$  es el efecto del tratamiento,  $p_j$  es el efecto del renglón,  $\gamma_m$  es el efecto de la columna y  $e_{ijm}$  es el error experimental aleatorio.

Cada renglón contiene una réplica completa de todos los tratamientos, y los tratamientos son ortogonales a los renglones; las columnas también son ortogonales a los renglones y los totales de tratamiento se ajustan sólo para los bloques de columna incompletos para proporcionar estimaciones no sesgadas de medias de tratamiento y una prueba F válida para los efectos de tratamiento. El análisis de varianza intrabloques para los diseños ortogonales por renglón solo difiere del diseño de bloques incompletos balanceados en cuanto a la suma de las particiones de sumas de cuadrados para los renglones, el resto de los aspectos del análisis permanece igual.

**Tabla 2.5 Análisis intrabloques para un diseño de bloques renglón-columna incompleto balanceado con tratamientos ortogonales por renglones.**

Fuente de variación	Grados de libertad	Suma de cuadrados
Total	$N - 1$	$\sum (\gamma_{ijm} - \bar{\gamma} \dots)^2$
Renglones (replicas)	$k - 1$	$t \sum (\bar{\gamma}_{.j} - \bar{\gamma} \dots)^2$
Columnas (no ajustadas)	$b - 1$	$k \sum (\gamma_{.m} - \bar{\gamma} \dots)^2$
Tratamientos (ajustados)*	$t - 1$	$k \sum Q_i^2 / \lambda t$
Error	$(t - 1)(k - 1)(b - 1)$	Restando

\* $Q_i = y_{i..} - (B_i/k)$  donde  $B_i$  es la suma de totales de bloque de aquellas columnas que incluyen el tratamiento  $i$ .

## 2.7 Reducción del tamaño de experimento con diseños parcialmente balanceados (BIPB)

No es posible construir diseños balanceados para todas las situaciones experimentales que requieren bloques incompletos, en algunos casos el número de réplicas necesario puede ser prohibitivo; por tanto, con frecuencia se construyen diseños parcialmente balanceados que requieren menos réplicas.

El número mínimo de replicas requerido para el diseño balanceado es  $r = \lambda (t - 1) / (k - 1)$ : suponiendo un experimento con tratamiento  $t = 6$ , se requiere bloques de tamaño  $k = 4$ . El diseño balanceado mostrado en el apéndice 1A.1 requiere 10 replicas o  $rt = 60$  unidades experimentales y es muy posible que no disponga de 60 unidades experimentales o que el costo del experimento con 60 unidades sea demasiado elevado.

### **Ocurrencias desiguales de partes de tratamientos**

Bose y Nair (1939) propusieron el diseño de bloques incompleto parcialmente balanceado, este diseño tiene algunas pares de tratamientos que aparecen en más bloques que otros pares, por lo que algunas comparaciones de tratamientos tendrán mayor precisión que otras. Es más sencillo usar un diseño de bloques balanceados que proporcione la misma precisión para todas las comparaciones entre tratamientos, pero si los recursos están limitados y no se pueden obtener las réplicas suficientes, el diseño parcialmente balanceado es una alternativa atractiva cuando el diseño balanceado requiere un número excesivo de unidades experimentales.

Examinando el diseño de bloques incompleto parcialmente balanceado para seis tratamientos en bloques de cuatro unidades mostrado en el cuadro 3, es posible observar que si bien el diseño balanceado del apéndice 1A.1 requiere 10 replicas para obtener un balance con  $\lambda = 6$ , que el diseño parcialmente balanceado tiene dos replicas y requiere doce unidades experimentales entre bloques de cuatro unidades.

**Cuadro 3** Diseño de bloques incompleto parcialmente balanceado con seis tratamientos en bloques de cuatro unidades.

Bloque 1	1	4	2	5
Bloque 2	2	5	3	6
Bloque 3	3	6	1	4

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.323

Algunos pares de tratamientos están en dos bloques, mientras que otros pares sólo en uno: los pares de tratamiento (1,4), (2,5) y (3,6) están en dos de los bloques mientras que el resto aparecen en uno solo y por tanto los pares de tratamientos que se presentan juntos en dos bloques se comparan con una precisión un poco mayor que los que están en un solo bloque. Las diferentes precisiones para las comparaciones de tratamiento es el sacrificio pagado por un experimento más pequeño, pero la diferencia en precisión no es tan grande como para evitar el uso de diseño de bloques parcialmente balanceados.

Al aumentar el número de réplicas puede aumentar la precisión de las comparaciones de tratamientos; si las dos réplicas que proporciona el experimento inicial no son suficientes, otra repetición del mismo experimento proporcionará cuatro réplicas que, de resultar eficientes, todavía representan ganancia en términos de reducción de costos en comparación con un diseño balanceado completo.

### **Clases asociadas para las ocurrencias de pares de tratamientos**

Un diseño parcialmente balanceado cada tratamiento es miembro de dos o más clases asociadas; una clase asociada es un grupo de tratamientos donde

cada par de tratamientos ocurre en  $\lambda_i$  bloques, los pares de tratamientos que ocurren en  $\lambda_i$  bloques se conocen como los  $i$ -ésimos asociados.

El diseño del cuadro 3 tiene dos clases asociadas, los pares de tratamientos (1,4), (2,5) y (3,6) son asociados primero con  $\lambda_1 = 2$ , cada par ocurre en dos bloques y cada tratamiento tiene  $n_1 = 1$  primeros asociados. El resto de los pares de tratamientos son segundos asociados con  $\lambda_2 = 1$ , existen  $n_2 = 4$  segundos asociados para cada tratamiento; por ejemplo, los segundos asociados del tratamiento 1 son los tratamientos 2, 3, 5 y 6, porque están con el tratamiento 1 en algún bloque del diseño. En el cuadrado 4 se muestran todos los conjuntos de asociados.

**Cuadrado 4** Primeros y segundos asociados para los seis tratamientos en el diseño de bloques parcialmente balanceado del cuadrado 2.3

Tratamientos	Asociados	
	Primeros	Segundos
1	4	2, 3, 5, 6
2	5	1, 3, 4, 6
3	6	1, 2, 4, 5
4	1	2, 3, 5, 6
5	2	1, 3, 4, 6
6	3	1, 2, 4, 5

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.324

### Notas sobre el análisis de diseños BIPB

Una ventaja de los diseños balanceados sobre los principalmente balanceados es que los cálculos manuales del análisis de varianza son un poco más sencillos. Antes del advenimiento de los programas estadísticos modernos, era imperativo disponer de fórmulas de cálculo manual para el análisis de datos, pero ahora es posible calcular las particiones de sumas de cuadrados adecuadas

para el análisis de varianza, al igual que estimaciones no sesgadas de las medias de tratamiento y sus errores estándar para la mayoría de los diseños con los programas disponibles. Por lo anterior, el diseño parcialmente balanceado es una alternativa real siempre que existan comparaciones razonablemente precisas de todos los pares de tratamientos.

### **Particiones de sumas de cuadrados para diseños BIPB**

El modelo lineal para el diseño de bloques incompletos parcialmente balanceado es el mismo que para el diseño de bloques incompletos parcialmente balanceado mostrado en la ecuación (1). Como los tratamientos no son ortogonales a los bloques, las particiones ortogonales de sumas de cuadrados de nuevo se derivan ajustando el modelo completo,  $\gamma_{ij} = \mu + \tau_i + \rho_j + e_{ij}$ , para obtener  $SCF_f$  y el modelo reducido alternativo sin efectos de tratamiento,  $\gamma_{ij} = \mu + \rho_j + e_{ij}$ , para obtener  $SCE_r$ .

Las sumas de cuadrados para tratamientos ajustados por los bloques se derivan como SC tratamiento (ajustada) =  $SCE_r - SCF_f$  y la suma de cuadrados de bloque no ajustada para los tratamientos es la misma que la mostrada en la tabla 2.1 para diseño balanceado.

Las formulas para calcular las estimaciones de mínimos cuadrados de los efectos de tratamiento, las medidas de tratamiento y la suma de cuadrados de tratamiento ajustada no son directas como las de los diseño balanceado, ya que existe más de una clase asociada para cada tratamiento y deben hacerse ajustes más complejos a los totales de tratamiento por fortuna, muchos programas estadísticos disponibles pueden realizar los cálculos y no es necesario usar las fórmulas para los tediosos cálculos manuales de un análisis exhaustivo de los datos. Los detalles de los cálculos manuales se pueden encontrar en Cochran y

Cox (1957) y los detalles del desarrollo se pueden encontrar en John (1971) y John (1987).

# **CAPÍTULO 3**

## **DISEÑO DE BLOQUES INCOMPLETOS: DISEÑOS RESOLUBLES Y CÍCLICOS**

### **3.1 Introducción**

En el capítulo 2 se presenta la descripción general y el análisis de los diseños de bloques incompletos. En este capítulo se explican varias clases de diseños de bloques incompletos, incluyendo los diseños resolubles con bloques agrupados en réplicas completas de tratamientos. También se presentan los diseños de cíclicos que se pueden construir sin usar extensos planes sistemáticos y los diseños  $\alpha$ , que amplían el número de los diseños resolubles aprovechables para los experimentos.

### 3.2 Diseños resolubles para ayudar a manejar el experimento

Los **diseños resolubles** tiene bloques agrupados de manera que cada grupo constituye una réplica completa de los tratamientos, el agrupamiento en réplicas completas es eficaz para administrar un experimento.

Una de las primeras aplicaciones de los diseños resolubles fue con pruebas de cultivo con platas colocadas en las parcelas de una granja experimental, los labradores querían probar un gran número de líneas genéticas y hacer todas las comparaciones por pares entre ellas con la misma precisión. Fue necesario ordenar las parcelas en bloques más pequeños que la réplica completa para reducir la varianza del error experimental aún más de lo que es posible con los diseños de bloques completos. El diseño de bloques incompletos resoluble era atractivo por que además de reducir el tamaño de los bloques para obtener mayor precisión, también permitía al investigador controlar estos grandes estudios en campo basándose en las réplicas.

Los diseños resolubles son útiles en la práctica cuando no es posible realizar el experimento completo al mismo tiempo, pues los experimentos con diseño resoluble pueden llevarse a cabo en etapas, con una o más réplicas terminadas en cada etapa. Además, si por alguna razón el experimento termina antes de tiempo, habrá un número igual de réplicas en todos los tratamientos.

Los diseños resolubles son arreglos en  $r$  grupos de réplicas de  $s$  bloques con  $k$  unidades por bloque. En los diseños de bloques incompletos balanceados resolubles, el número de tratamientos es un múltiplo del número de unidades por bloque,  $t = sk$ , y el número total de bloques satisface la relación  $b = rs \geq t + r - 1$ .

### **Ejemplo 3.1 Un diseño resoluble para pruebas con productos alimenticios**

Examinemos un experimento para probar nueve variables de un producto alimenticio en el mismo día con sujetos humanos como jueces. Varios factores asociados con el experimento hacen que un diseño de bloques incompleto sea atractivo, se puede esperar que un juez discrimine de manera adecuada hasta cuatro o cinco muestras de productos en cualquier momento, pero por problemas de programación, no es posible tener un número suficiente de jueces disponibles el mismo día para las réplicas de tratamiento; por lo tanto, deben prepararse de nuevo los alimentos cada día de pruebas y es necesario garantizar que la variación en la preparación de un día a otro no interfiera con las comparaciones de tratamientos en el experimento.

El cuadro 5 muestra una prueba diseñada para tres jueces en la que cada uno evaluará tres variables del producto en un día dado. Se programan tres jueces durante cuatro días de pruebas, ellos constituyen un bloque incompleto de tres tratamientos y cada uno se asigna al azar a uno de los bloques de tres tratamientos y prueba tres productos presentados al azar para su evaluación.

El diseño es resoluble con una réplica completa del experimento realizado un día dado y está balanceado, pues cada tratamiento ocurre una vez en el mismo bloque con cada uno de los tratamientos en el experimento. Los jueces solo evalúan tres productos a la vez, lo que reduce considerablemente la variación dentro del bloque en comparación con un juez que evalúa nueve productos a la vez.

**Cuadro 5** Diseño de bloques incompleto balanceado resoluble para evaluar productos alimenticios.

<i>Día 1</i>		<i>Día 2</i>		<i>Día 3</i>		<i>Día 4</i>	
<i>Juez</i>		<i>Juez</i>		<i>Juez</i>		<i>Juez</i>	
1	(1, 2, 3)	4	(1, 4, 7)	7	(1, 5, 9)	10	(1, 6, 8)
2	(4, 5, 6)	5	(2, 5, 8)	8	(2, 6, 7)	11	(2, 4, 9)
3	(7, 8, 9)	6	(3, 6, 9)	9	(3, 4, 8)	12	(3, 5, 7)

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.340

### Diseños de retícula balanceada

Los diseños de retícula (lattice) balanceados son un grupo muy conocido de diseños resolubles propuestos por Yates (1937). El número disponible de estos cuadrados de retícula está limitado debido a que el número de tratamientos debe ser cuadrado exacto  $t = k^2$ , los diseños requieren  $r = (k + 1)$  réplicas y  $b = k(k + 1)$  bloques para un balance completo con  $\lambda = 1$ , cada grupo de réplicas contiene  $s = k$  bloques de  $k$  unidades experimentales cada uno y el número  $k$  de unidades por bloque debe ser el número primo o una potencia de un número primo. Se pueden encontrar planes para diseños de retícula balanceados con  $t = 9, 16, 25, 49, 64$  y  $81$  en Cochran y Cox (1957) o Peterson (1985).

El diseño del cuadro 5 es una retícula balanceada con nueve tratamientos en bloques de tres unidades que tiene cuatro grupos de réplicas con un total de 12 bloques. Observe que cada par de tratamientos ocurre en un bloque  $\lambda = 1$  vez en alguna parte del diseño balanceado.

El balance completo requiere  $(k + 1)$  réplicas con diseños de retícula y, como son diseños resolubles, uno o más grupos de réplicas se pueden eliminar para reducir un diseño de retícula parcialmente balanceado. Los diseños con  $r = 2, 3$  o

4 grupos de réplicas respectivos se conocen como retículas *simples*, *triples* o *cuádruples*.

El factor de eficiencia promedio para un diseño de retícula balanceado es:

$$E = \frac{(k+1)(r-1)}{r(k+2) + (k+1)}$$

y por las retículas simple y triple, los factores de eficiencia promedio respectivos son  $(k+1)/(k+3)$  y  $(2k+2)/(2k+5)$ .

### **Diseños de retícula rectangulares.**

La restricción en el número de tratamientos o variables de las retículas cuadradas es importante. Los diseños de retícula rectangulares desarrollados por Harshbarger (1949) proporcionaron diseños resolubles con números de tratamientos intermedios proporcionados por los diseños de retícula cuadrados.

Los diseños de retícula rectangulares son arreglos de  $t = s(s - 1)$  tratamientos en bloques con  $k = (s - 1)$  unidades y los números de tratamientos están cerca del punto medio entre los que proporcionan las retícula cuadradas. Cochran y Cox (1957) construyeron planes para retículas rectangulares con  $t = 12, 20, 30, 42, 56, 72$  y  $90$  tiramientos. Los diseños con  $r = 2$  y  $r = 3$  replicas se conocen como retículas rectangulares simple y triple, respectivamente.

Existen otros diseños de bloques incompletos balanceados resolubles sin las restricciones de las retículas cuadradas o rectangulares; sin embargo, son para un conjunto limitado de tamaños de bloque y números de tratamientos.

### Descripción del análisis de diseños resolubles

El modelo lineal para el diseño resoluble es:

$$y_{ijm} = \mu + \tau_i + \gamma_j + \rho_{m(j)} + e_{ijm} \quad (15)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, r \quad m = 1, 2, \dots, s$$

donde  $\mu$  es la media general,  $\tau_i$  es el efecto del tratamiento,  $\gamma_j$  es el efecto del grupo de replicas,  $\rho_{m(j)}$  es el efecto del bloque anidado dentro de la replica y  $e_{ijm}$  es el error experimental aleatorio.

El ajuste secuencial de los modelos alternativos descritos en el capítulo 2 proporciona una partición de suma de cuadrados ortogonal, pero en este caso se requiere una partición de suma de cuadrados adicional para los grupos de réplicas. Los bloques están anidados dentro de los grupos de réplicas se unen como SC(bloques/réplicas). La tabla 3.1 contiene una descripción del análisis de varianza intrabloques.

**Tabla 3.1 Análisis intrabloques para un diseño de bloques incompleto balanceado resoluble**

Fuente de variación	Grados de libertad	Suma de cuadrados
Total	$N - 1$	$\sum (\gamma_{ijm} - \bar{\gamma} \dots)^2$
Replicas	$r - 1$	$sk \sum (\bar{\gamma}_{.j} - \bar{\gamma} \dots)^2$
Bloques (no ajustados) dentro de réplicas	$r(s - 1)$	$k \sum (\gamma_{..m} - \bar{\gamma} \dots)^2$
Tratamientos (ajustados)	$t - 1$	$\frac{k \sum Q_i^2}{\lambda t}$
Error	$N - t - rs + 1$	Restando

$Q_i = y_{i\dots} - (B_i/k)$ , y  $B_i$  es la suma de totales de bloque que incluyen  $i$ -ésimo tratamiento.

Los tratamientos son ortogonales a los grupos de réplicas ya que cada tratamiento aparece una vez en cada grupo de réplica y los totales de tratamientos ajustados sólo por los efectos de bloque proporcionan estimaciones no sesgadas de las medidas de tratamiento y una prueba  $F$  válida para los efectos de tratamiento.

### **3.3 Diseños resolubles renglón-columna para dos criterios de bloque**

Algunas situaciones de investigación requieren de criterios de bloque y no permiten completos de tratamientos para los renglones o las columnas necesarios para los cuadrados ortogonales de Youden. Los diseños renglón-columna tienen  $t$  tratamientos colocados en bloques de  $k = pq$  unidades y los tratamientos de cada bloque se arreglan en un diseño renglón-columna de  $p \times q$ .

Existe un número limitado de diseños de bloques incompleto renglón-columna balanceados para los experimentos que requieren que los tamaños de bloque para los renglones y columnas sean menores que el número de tratamientos, estos diseños necesitan un gran número de bloques para cumplir con este requerimiento de balance.

#### **Ejemplo 3.2 Diseño renglón-columna anidado para muestra de población de insectos**

Examinemos un estudio epidemiológico sobre la transmisión de un parásito microbico a los humanos a través de un insecto en un área agrícola tropical. El parásito se transmite a los humanos mediante la picadura del insecto, pero si el insecto no está infectado por el parásito, puede adquirirlo cuando pica a un humano infectado. Un proyecto de salud pública es la plantación de pruebas de efectividad de una droga que se sabe que varía el ciclo de vida del parásito en los

humanos, un entomólogo que trabaja en el proyecto piensa obtener muestras de las poblaciones de insectos para supervisar el efecto que el programa tiene en la reducción de la tasa de infección.

Se programaron nueve tipos de hábitats dentro y cerca de las cuatro plantaciones agrícolas para supervisar el proyecto, los entornos son lugares como corrientes de ríos, pueblos cercanos a las plantaciones, campos agrícolas y aguas estancadas, y el entomólogo sólo recolectará muestras de insectos en tres de ellos en un día. Se estableció un diseño de muestreo renglón-columna para controlar la variación potencial causada por muestreo en distintos días y horas del día en las cuatro plantaciones, el diseño se muestra en el cuadro 6 tiene nueve tratamientos de entorno con muestreo en un diseño renglón-columna de 3 X 3 en cuatro plantaciones. Los tres renglones del arreglo de 3 X 3 corresponden a la hora del día en que se toman las muestras, y las tres columnas se refieren a los tres días de recolección en cada plantación.

Por inspección de cada par de tratamientos o entornos se puede verificar que el diseño se toma una muestra a la vez de cada uno en el mismo día, cada par de tratamientos debe aparecer una vez en el mismo renglón y una vez en la misma columna en alguna parte tener un diseño renglón-columna balanceado. El diseño se puede resolver con las plantaciones como réplicas y una réplica completa de los nueve tipos de hábitat en cada plantación; es un diseño anidado en donde los renglones y las columnas se anidan dentro de los grupos de réplicas.

### **Diseños de retículas cuadradas balanceadas**

Estos diseños, también propuestos por Yates (1940), son diseños renglón-columna anidados resolubles, que están entre los primeros desarrollos de diseños de bloques incompletos; el ejemplo del cuadro 6 es uno de ellos. La **retícula cuadrada balanceada** tiene  $t = k^2$  tratamientos dispuestos en un diseño renglón-

columna de  $k \times k$ , un arreglo de  $k \times k$  consiste en una de las réplicas completas de tratamientos con  $p = q = k$ . el diseño balanceado requiere  $r = (k + 1)$  réplicas para que el balance sea completo, de manera que cada par de tratamientos se encuentre una vez en el mismo renglón y una vez en la misma columna.

En algún momento del experimento. Cuando  $k$  es impar, un valor de repetición de  $r = \frac{1}{2}(k + 1)$  proporcionará un semibalace tal que cualquier par de tratamientos esté una vez en el mismo renglón o una vez en la misma columna.

En Cochran y Cox (1957) y en Peterson (1985) se pueden encontrar planos para diseños cuadrados de retícula  $t = 9, 16, 25, 49, 64, 81, 121$  y  $169$ . Una presentación más amplia de estos diseños se encuentra en Kempthorne (1952) y Federer (1955).

**Cuadro 6** Diseño de bloques incompleto renglón columna anidado de  $3 \times 3$  para las muestras de la población de insectos.

<table style="margin: auto;"> <thead> <tr> <th></th> <th colspan="3"><i>Día</i></th> </tr> <tr> <th><i>Hora</i></th> <th>(1)</th> <th>(2)</th> <th>(3)</th> </tr> </thead> <tbody> <tr> <td>(1)</td> <td>1</td> <td>2</td> <td>3</td> </tr> <tr> <td>(2)</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>(3)</td> <td>7</td> <td>8</td> <td>9</td> </tr> </tbody> </table> <p align="center">Plantación 1</p>		<i>Día</i>			<i>Hora</i>	(1)	(2)	(3)	(1)	1	2	3	(2)	4	5	6	(3)	7	8	9	<table style="margin: auto;"> <thead> <tr> <th></th> <th colspan="3"><i>Día</i></th> </tr> <tr> <th><i>Hora</i></th> <th>(4)</th> <th>(5)</th> <th>(6)</th> </tr> </thead> <tbody> <tr> <td>(4)</td> <td>1</td> <td>4</td> <td>7</td> </tr> <tr> <td>(5)</td> <td>2</td> <td>5</td> <td>8</td> </tr> <tr> <td>(6)</td> <td>3</td> <td>6</td> <td>9</td> </tr> </tbody> </table> <p align="center">Plantación 2</p>		<i>Día</i>			<i>Hora</i>	(4)	(5)	(6)	(4)	1	4	7	(5)	2	5	8	(6)	3	6	9
	<i>Día</i>																																								
<i>Hora</i>	(1)	(2)	(3)																																						
(1)	1	2	3																																						
(2)	4	5	6																																						
(3)	7	8	9																																						
	<i>Día</i>																																								
<i>Hora</i>	(4)	(5)	(6)																																						
(4)	1	4	7																																						
(5)	2	5	8																																						
(6)	3	6	9																																						
<table style="margin: auto;"> <thead> <tr> <th></th> <th colspan="3"><i>Día</i></th> </tr> <tr> <th><i>Hora</i></th> <th>(7)</th> <th>(8)</th> <th>(9)</th> </tr> </thead> <tbody> <tr> <td>(7)</td> <td>1</td> <td>6</td> <td>8</td> </tr> <tr> <td>(8)</td> <td>9</td> <td>2</td> <td>4</td> </tr> <tr> <td>(9)</td> <td>5</td> <td>7</td> <td>3</td> </tr> </tbody> </table> <p align="center">Plantación 3</p>		<i>Día</i>			<i>Hora</i>	(7)	(8)	(9)	(7)	1	6	8	(8)	9	2	4	(9)	5	7	3	<table style="margin: auto;"> <thead> <tr> <th></th> <th colspan="3"><i>Día</i></th> </tr> <tr> <th><i>Hora</i></th> <th>(10)</th> <th>(11)</th> <th>(12)</th> </tr> </thead> <tbody> <tr> <td>(10)</td> <td>1</td> <td>9</td> <td>5</td> </tr> <tr> <td>(11)</td> <td>6</td> <td>2</td> <td>7</td> </tr> <tr> <td>(12)</td> <td>8</td> <td>4</td> <td>3</td> </tr> </tbody> </table> <p align="center">Plantación 4</p>		<i>Día</i>			<i>Hora</i>	(10)	(11)	(12)	(10)	1	9	5	(11)	6	2	7	(12)	8	4	3
	<i>Día</i>																																								
<i>Hora</i>	(7)	(8)	(9)																																						
(7)	1	6	8																																						
(8)	9	2	4																																						
(9)	5	7	3																																						
	<i>Día</i>																																								
<i>Hora</i>	(10)	(11)	(12)																																						
(10)	1	9	5																																						
(11)	6	2	7																																						
(12)	8	4	3																																						

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.344

### Descripción del análisis para diseños renglón-columna anidados

El modo lineal para cualquier diseño renglón-columna anidado resoluble es:

$$y_{ijlm} = \mu + \beta_m + \rho_{j(m)} + \gamma_{l(m)} + \tau_i + e_{ijlm} \quad (16)$$

$$i = 1, 2, \dots, t \quad j = 1, 2, \dots, p \quad l = 1, 2, \dots, q \quad m = 1, 2, \dots, r$$

donde  $\mu$  es la media general,  $\tau_i$  es el efecto del tratamiento,  $\beta_m$  es el efecto de grupo de réplica,  $\rho_{j(m)}$  y  $\gamma_{l(m)}$  son el renglón y la columna respectivos anidados en los efectos de grupos de réplicas y  $e_{ijlm}$  es el error experimental aleatorio.

Cada tratamiento está en cada grupo de réplicas y son ortogonales a ellos, los renglones y columnas también son ortogonales a los grupos de réplicas. Las particiones de sumas de cuadrados para las réplicas, los renglones dentro de las réplicas y las columnas dentro de las réplicas se pueden calcular de manera usual, como se muestra en el análisis de varianza descrito en la tabla 3.2. Los tratamientos no son ortogonales a los renglones o columnas de los grupos de réplicas y deben ajustarse por los efectos de renglón y de columna.

**Tabla 3.2** Análisis intrabloques para un diseño de bloques incompleto renglón-columna anidado

Fuente de variación	Grados de libertad	Suma de cuadrados
Total	$N - 1$	$\sum (y_{ijlm} - \bar{y} \dots)^2$
Grupos (replicas)	$r - 1$	$t \sum (\bar{y} \dots_m - \bar{y} \dots)^2$
Columnas (no ajustadas) dentro de grupos	$r(q - 1)$	$p \sum (\bar{y} \dots_{lm} - \bar{y} \dots_m)^2$
Renglones (no ajustadas) dentro de grupos	$r(p - 1)$	$q \sum (\bar{y} \dots_{j,m} - \bar{y} \dots_m)^2$
Tratamientos (ajustados)	$t - 1$	$\frac{pq(t-1)}{rt(p-1)(q-1)} k \sum Q_i^2$
Error	$r(p-1)(q-1)-(t-1)$	Restando

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p. 345

Los totales de tratamiento ajustado necesarios para las estimaciones de mínimos cuadrados de los efectos de tratamientos y las sumas de los cuadrados de tratamiento ajustados son:

$$Q_i = \gamma_{i\dots} + r\bar{\gamma}_{\dots} - \frac{1}{p}R_i - \frac{1}{q}C_i \quad (17)$$

Donde  $\gamma_{i\dots}$  es el total para el  $i$ -ésimo tratamiento  $\bar{\gamma}_{\dots}$  es la gran media para el experimento,  $R_i$  es la suma de los totales de renglones que incluyen al  $i$ -ésimo tratamiento y  $C_i$  es la suma de los totales de columnas que incluyen al  $i$ -ésimo tratamiento.

La estimación de los mínimos cuadrados del efecto de tratamiento es:

$$\hat{\tau}_i = \frac{pq(t-1)Q_i}{rt(p-1)(q-1)} \quad (18)$$

y la estimación de los mínimos cuadrados de la media de tratamiento es  $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i$ , donde  $\bar{\mu} = \bar{\gamma}_{\dots}$  es la estimación de la media general.

El factor de eficiencia para el diseño renglón-columna anidado es:

$$E = \frac{t(p-1)(q-1)}{pq(t-1)} \quad (19)$$

donde el factor de eficiencia para la retícula cuadrada balanceada es  $E = (k-1)/(k+1)$ , con  $p = q = k$  y  $t = k^2$ ; por ejemplo, con  $k = 3$  del cuadro 6, el factor de eficiencia es  $E = 2/4 = 0.50$ . la retícula cuadrada balanceada tendría que reducir la varianza del error experimental en un 50% para estimar las comparaciones entre las medias de tratamiento de manera tan precisa como un diseño de bloques completos del mismo tamaño.

### 3.4 Los diseños cíclicos simplifican la construcción del diseño

Los diseños de bloques incompletos complejos pueden añadir más complicaciones al proceso si hay necesidad de consultar con frecuencia la distribución del diseño. Por lo que debe contarse con tablas amplias de planes de diseño para usar los diseños de bloques incompletos que se han presentado hasta ahora, establecer un diseño en la instalación física necesaria para el experimento exige una vigilancia constante, debe tenerse cuidado de evitar errores cuando se asignan los tratamientos a las unidades experimentales y al registrar los datos.

Los diseños cíclicos ofrecen cierto grado de simplificación en la construcción del diseño y en el establecimiento del proceso experimental y como se generan a partir de un bloque inicial de tratamientos, es necesario un esquema completo de los planes de distribución del experimento. Una vez que se conoce el bloque inicial, se genera el plan del experimento y se asignan al azar los tratamientos a las etiquetas numéricas del diseño.

#### Como construir un diseño cíclico

La designación cíclica se refiere al método usado para construir los diseños, al igual que al esquema de asociación entre tratamientos. Los tratamientos se asignan a los bloques a través de una sustitución cíclica de etiquetas de tratamientos de un bloque generador inicial.

El método requiere etiquetar los tratamientos como  $(0, 1, 2, \dots, t - 1)$ , el ciclo comienza con un bloque de tamaño  $k$  requerido por el experimento y los bloques subsecuentes se obtienen sumando 1 a la etiqueta de tratamiento de bloque anterior, si una etiqueta de tratamiento excede  $(t - 1)$ , se reduce mediante el módulo  $t$ ; el ciclo continua hasta que regresa a la configuración del bloque inicial.

Considerando el diseño para 6 tratamientos en bloques de tres unidades generadas por sustitución cíclica del cuadro 7, se etiquetan los seis tratamientos (0, 1, 2, 3, 4, 5) para el diseño, que se construye con el bloque inicial (0, 1, 3); el segundo bloque se construye sumando 1 a cada etiqueta de tratamiento del primer bloque, resultando en (1, 2, 4) como configuración del segundo bloque; se suma 1 a cada etiqueta del segundo bloque y se obtiene (2, 3, 5) como los tratamientos para el tercer bloque.

**Cuadro 7 Diseño cíclico para seis tratamientos en bloques de tres unidades**

Bloque	
1	(0, 1, 3)      Bloque inicial
2	(1, 2, 4)
3	(2, 3, 5)
4	(3, 4, 0)
5	(4, 5, 1)
6	(5, 0, 2)

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.346

La transición del bloque 3 al bloque 4 en el ciclo requiere reducir una etiqueta en el bloque 4 mediante el modulo de  $t = 6$ , por que si se suma 1 a cada etiqueta en el bloque 3, la configuración del bloque 4 será (3, 4, 6) donde  $5 + 1 = 6$  excede  $(t - 1) = 5$  por lo tanto, 6 se debe reducir según el modulo 6, que es  $6 = 0(\text{mod } 6)$ , y 0 sustituye a 6 en la configuración del bloque 4, es decir, (3, 4, 0); en la transición del bloque 5 al 6 se realiza una sustitución similar.

Dentro del ciclo completo mostrado en el cuadro, se construyo un conjunto completo de  $b = 6$  bloques, los números forman un ciclo completo de etiquetas de tratamiento hacia abajo de las columnas, las etiquetas de tratamiento en la primera posición del bloque tienen un ciclo de 0 a 5 mientras que en la segunda posición el ciclo comienza con la etiqueta 1 y la tercera posición con la etiqueta 3, entonces el

diseño está completo con los bloques de 1 a 6 pues el séptimo miembro del ciclo sería la configuración del bloque inicial (0, 1, 3).

### **Construcción de diseños renglón-columna con el método cíclico**

Los diseños cíclicos se pueden usar para los diseños renglón-columna si el experimento requiere de dos criterios de bloque, el único requisito es que el número de réplicas sea un múltiplo de tamaño de bloque, o sea,  $r = ik$ , donde  $i$  es un valor entero. El diseño del cuadro 7, con  $t = 6$  y  $r = k = 3$  generado a partir del bloque inicial (0, 1, 3), se puede usar como diseño renglón-columna con seis renglones y tres columnas.

Se trata de un diseño ortogonal por columna, ya que cada tratamiento aparece en cada columna; si se mantiene la integridad del renglón y la columnas al aleatorizar, entonces la suma de cuadrados para columnas y renglones es una partición de la suma de cuadrados total y en el análisis de los tratamientos se ajustan por los bloques de los renglones.

### **Tablas para construir diseños cíclicos**

En la tabla 2A.1 del apéndice 2A.1 se puede encontrar la tabla compacta de los bloques iniciales necesaria para los diseños cíclicos de relativa eficiencia con  $4 \leq t \leq 15$  tratamientos, tomada de John (1987), quien también elaboró otra tabla para  $16 \leq t \leq 30$  tratamientos; pueden encontrarse más tablas en John (1996).

Para un tamaño de bloque  $k - 1$  tratamientos del bloque inicial son iguales para cualquier número  $t$  de tratamientos. El último tratamiento del bloque se toma con el valor correspondiente en la columna para  $t$  en la tabla 2A.1 del apéndice; por

ejemplo, con  $k = 3$ ,  $r = 3$ , el bloque inicial para  $t = 5$  es  $(0, 1, 2)$  y es  $(0, 1, 3)$  para  $t = 6$ .

Si se requiere un diseño con más replicas ( $r > k$ ), la tabla proporciona un bloque generador de segundo o tercer nivel, en ella se construye un conjunto de bloques a partir de cada uno mediante la sustitución cíclica. El diseño con  $t = 6$  tratamientos en bloques de tamaño  $k = 3$  generados por el bloque inicial  $(0, 1, 3)$  tenían  $r = 3$  réplicas, el diseño con  $r = 6$  réplicas se construye agregando el bloque inicial  $(0, 2, 1)$  que se encuentra en la tabla 2A.1 en la línea con  $k = 3$  y  $r = 6$ ; el diseño obtenido tiene  $b = 12$  bloques, la mitad de ellos generados por cada uno de los bloques iniciales como se muestra en el cuadrado 8. Este diseño está parcialmente balanceado con  $\lambda_1 = 3$  y  $\lambda_2 = 2$  para las dos clases asociadas.

**Cuadro 8      Diseño cíclico generado a partir de los bloques iniciales**

Bloque		Bloque	
1	$(0, 1, 3)$ Bloque inicial	1	$(0, 2, 1)$ Bloque inicial
2	$(1, 2, 4)$	2	$(1, 3, 2)$
3	$(2, 3, 5)$	3	$(2, 4, 3)$
4	$(3, 4, 0)$	4	$(3, 5, 4)$
5	$(4, 5, 1)$	5	$(4, 0, 5)$
6	$(5, 0, 2)$	6	$(5, 1, 0)$

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.348

### **3.5 Diseños $\alpha$ para diseños resolubles versátiles a partir de una construcción cíclica**

Los diseños  $\alpha$  son diseños resolubles desarrollados por Paterson y Williams (1957) en respuesta a la necesidad de las respuestas de campo reglamentadas para variedades de cosechas agrícolas en el Reino Unido, en ellas el número de variedades y réplicas los fijaba un reglamento y no estaba bajo el control del

experimentador. El gran número de variedades necesito diseños de bloques incompletos y diseños resolubles para un control adecuado en el campo, pues los diseños resolubles existentes no siempre admiten el número de variedades o los tamaños de bloque para las pruebas.

### **Características de los diseños $\alpha$**

Estos diseños no tienen limitantes en cuanto al tamaño del bloque excepto por la restricción de que el número de tratamientos,  $t$ , debe ser el múltiplo de tamaño de bloque  $k$ , de manera  $t = sk$ , para tener un diseño resoluble con tamaño de bloque iguales. Las tablas de planes de diseños completos son innecesarias, ya que se puede generar por sustitución cíclica a partir de un arreglo inicial de números en forma muy parecida a los diseños cíclicos.

El desarrollo de los diseños  $\alpha$  ha aumentado el número y la flexibilidad de los diseños resolubles disponibles para pruebas experimentales. Los diseños desarrollados por Peterson y Williams (1976) para las pruebas reglamentadas incluían diseños con  $r = 2, 3$  o  $4$  replicas con  $t \leq 100$  tratamientos y tamaños de bloque en  $4 \leq k \leq 16$ .

Los diseños están parcialmente balanceados con dos o tres clases asociadas, los diseños con dos clases asociadas con  $\lambda_1 = 0$  y  $\lambda_2 = 1$  se conocen como diseños  $\alpha(0, 1)$ , y los que tiene tres clases asociadas con  $\lambda_1 = 0$ ,  $\lambda_2 = 1$  y  $\lambda_3 = 2$  se denotan como  $\alpha(0, 1, 2)$ . La mayor parte de los diseños considerados para las pruebas de variedad fueron diseñados  $\alpha(0, 1)$ . Debido a la manera que se construyen, no existen los diseños balanceados ni los diseños  $\alpha(1, 2)$ .

El limite superior para la eficiencia de los diseños  $\alpha$  que proporcionaron Paterson y Williams (1976) es:

$$E = \frac{(t-1)(r-1)}{(t-1)(r-1) + r(s-1)} \quad (20)$$

**Como construir los diseños  $\alpha$**

La construcción de un diseño para  $t = sk$  tratamientos con  $r$  réplicas comienza con un arreglo de  $k \times r$  designado como el arreglo *generador*  $\alpha$ . Cada columna del arreglo generador  $\alpha$  se usa para construir  $s - 1$  columnas adicionales por sustitución cíclica y el nuevo arreglo de  $k \times rs$  recibe el nombre de arreglo intermedio  $\alpha^*$ .

Consideremos un arreglo generador  $\alpha$  de  $4 \times 3$  para  $t = 12$  tratamientos,  $k = 4$  unidades por bloque,  $r = 3$  replicas y  $s = 3$  bloques por replica:

**Arreglos generador  $\alpha$**

Columna		
1	2	3
0	0	0
0	0	2
0	2	1
0	1	1

Cada columna se usa para generar  $s - 1 = 2$  columnas adicionales por sustitución cíclica con  $\text{mod}(s) = \text{mod}(3)$  para derivar el arreglo intermedio  $\alpha^*$ ; el resultado será  $r = 3$  arreglos de tamaño  $k \times s$ . en este ejemplo, habrá tres arreglos de  $4 \times 3$ , los arreglos intermedios  $\alpha^*$  generados de cada una de las tres columnas del arreglo  $\alpha$  son:

<b>Arreglos intermedios <math>\alpha^*</math></b>								
Arreglo generado a partir de								
Columna1			Columna2			Columna3		
0	1	2	0	1	2	0	1	2
0	1	2	0	1	2	2	0	1
0	1	2	2	0	1	1	2	0
0	1	2	1	2	0	1	2	0

Por último, las etiquetas se obtienen sumando  $S = 3$  a cada elemento del segundo renglón del arreglo intermedio  $\alpha^*$ ,  $2s = 6$  para cada elemento del tercer renglón, y  $3s = 9$  al cuarto renglón. Las columnas de los arreglos son los bloques del diseño y cada conjunto de  $s$  columnas generado con una columna arreglo generador  $\alpha$  constituye un réplica completa; en el cuadro 9 se muestra el diseño  $\alpha(0, 1, 2)$  obtenido.

**Cuadro 9**      **Diseño  $\alpha$  para 12 tratamientos en tres grupos de réplicas**

Réplica	<u>I</u>			<u>II</u>			<u>III</u>		
Bloque	1	2	3	1	2	3	1	2	3
	0	1	2	0	1	2	0	1	2
	3	4	5	3	4	5	5	3	4
	6	7	8	8	6	7	7	8	6
	9	10	11	10	11	12	10	11	9

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.350

Las etiquetas de tratamiento para el diseño son  $(0, 1, 2, \dots, t - 1)$  y los tratamientos reales se asignan al azar a las etiquetas. Además, la aleatorización procede con una asignación aleatoria de los bloques del diseño a los bloques reales dentro de los grupos de réplicas y una asignación aleatoria de los tratamientos a las unidades experimentales reales dentro de los bloques físicos.

Una tabla de 11 arreglos  $\alpha$  básicos reproducida de Paterson, Williams y Hunter (1978) se muestra en la tabla 2A.2 del apéndice. Se puede generar un total de 47 diseños  $\alpha$  a partir de 11 arreglos. Los arreglos se pueden usar para generar diseños con  $20 \leq t \leq 100$ , números de tratamientos, con  $r = 2, 3$  o  $4$  réplicas y la restricción usual de  $t = sk$ . Los arreglos producirán diseños con tamaños de bloques  $k \geq 4$  con la condición  $k \leq s$ .

Observando el primer arreglo que aparece en la tabla 2A.2 del apéndice para  $s = k = 5$ , vemos que el arreglo tiene  $k = 5$  renglones y  $r = 4$  columnas.

0	0	0	0
0	1	4	2
0	2	3	4
0	3	2	1
0	4	1	3

Este arreglo en particular se puede usar para generar un diseño  $\alpha$  con  $r = 2$  réplicas si se usan sólo sus dos primeras columnas; de la misma manera, se pueden generar diseño de tres o cuatro replicas con las primeras tres o cuatro columnas, respectivamente. Un diseño con tamaño de bloque  $k = 5$  se obtiene con los cinco renglones del arreglo, un diseño con tamaño de bloque  $k = 4$  se construye con los primeros cuatro primeros renglones, con  $s = 5$  se puede construir un diseño para  $t = 25$  tratamientos en bloques de  $k = 5$  unidades con  $r = 2, 3$  o  $4$  réplicas. Además, con  $s = 5$  es posible construir un diseño para  $t = 20$  tratamientos en bloques de  $k = 4$  unidades con  $r = 2, 3$  o  $4$  réplicas.

### **Diseños de bloque incompletos resolubles latinizados**

Los diseños resolubles presentados hasta hora tienen estructuras de bloque anidadas. La aleatorización de los bloques incompletos es independiente dentro de cada réplica, ya sea que el diseño tenga un solo criterio de bloque o dos en los renglones y columnas.

En algunos casos, los bloques de una réplica pueden tener alguna relación con los bloques de otra, estas relaciones existen en experimentos donde los bloques en diferentes réplicas son contiguos en su espacio, como se ilustra en el cuadro 10.

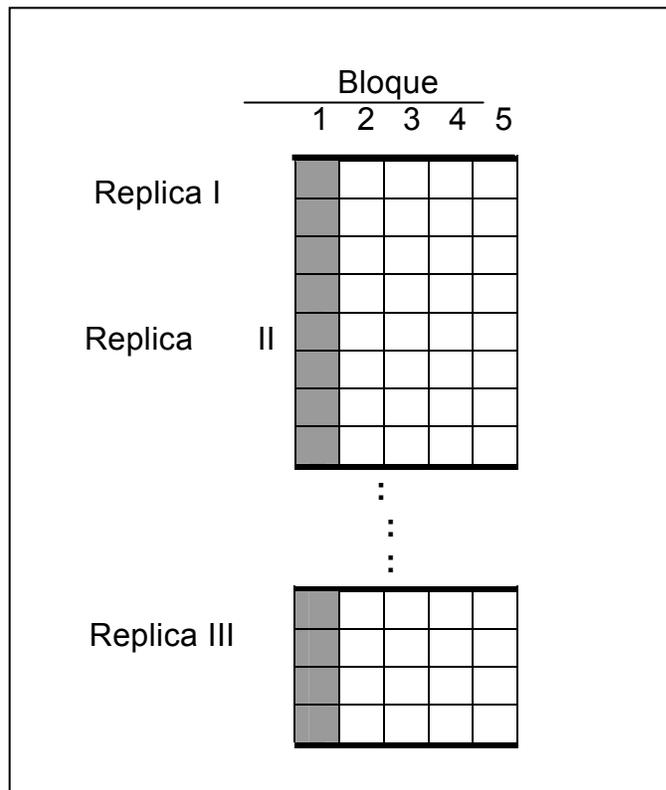
Williams (1986) presento los resultados en diseños más generales donde los tratamientos no aparecen más de una vez en los bloques largos y diseños latinizados para incluir el bloque renglón- columna. Los diseños  $\alpha$  se pueden usar para generar los diseños latinizados más generales como lo ilustra John y Williams (1995); el cuadro 11 se muestra un diseño renglón-columna latinizado, observe que los tratamientos no más de una vez en cada bloque largo (columnas). John y Williams (1955) hacen un análisis exhaustivo de los diseños resolubles y cíclicos.

**Cuadro 11    Diseño renglón-columna latinizado**

	Bloque			
	1	2	3	4
Replica I	0	1	2	3
	8	5	10	7
	4	9	6	11
Replica II	6	0	3	5
	1	7	4	2
	11	10	9	8
Replica III	9	8	0	1
	7	6	5	4
	2	3	11	10

Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.352

**Cuadro 10 Ilustración de bloques contiguos en replicas**



Fuente: R. O. Kuehl (2001), "Diseño de experimentos." México. D.F.: Ed. Matemáticas THOMSON. (2ª ed.). p.351

Si la distribución del espacio es la mostrada en el cuadrado, entonces cualquier bloque, como el bloque 1, en cada replica se encuentra en la misma línea de espacio como bloque 1 en todas las demás réplicas. El resultado es la ocurrencia de un bloque "largo" en dirección vertical de la distribución del cuadro 10.

Si los bloques se aleatorizan por separado en cada réplica, existe la posibilidad de que todas las réplicas de uno o más tratamientos ocurran en el mismo bloque "largo", lo que producirá una circunstancia indeseable en el caso de perturbaciones no previstas que negaran el uso del bloque "largo". Para evitar esta posibilidad, Harshbarger y Davis (1952) propusieron los diseños con retícula latinizados, en los que la retícula rectangular se estructuró de manera que cada tratamiento ocurra sólo una vez en cada bloque largo.

# CAPÍTULO 4

## APLICACIÓN A LA INDUSTRIA DE DISEÑOS DE BLOQUES INCOMPLETOS Y SU EFICIENCIA FRENTE A OTROS DISEÑOS.

### 4.1 Introducción

La eficiencia de un diseño con respecto a otro se mide mediante la comparación de varianzas para las estimaciones de las diferencias entre medias de tratamiento en los diseños. Por ejemplo, la varianza de la diferencia entre dos medias de tratamiento para el diseño de bloques completo aleatorizado (BCA) es  $2\sigma_{rcb}^2/r$ , ¿Es esa varianza menor que su contraparte del diseño totalmente aleatorizado  $2\sigma_{cr}^2/r$ ? La eficiencia relativa del diseño BCA con respecto al diseño totalmente aleatorizado se puede determinar porque se puede obtener una estimación de  $\sigma^2$  para este último a partir de los datos en el diseño BCA. Entonces es posible evaluar la efectividad del bloque para reducir la varianza del error experimental.

## 4.2 Eficiencia de los diseños de bloques incompletos

No existe la misma facilidad si se requiere determinar la eficiencia de un diseño de bloques incompleto balanceado (BIB) con respecto al diseño BCA porque no es posible calcular una estimación de  $\sigma^2$  para el diseño BCA; sería bueno determinar si un tamaño de bloque más pequeño del diseño BIB daría como resultado una menor varianza del error experimental. Todavía se usa la razón de las varianzas para una deficiencia entre dos medias de tratamiento para comparar el diseño BIB con el diseño BCA, la diferencia es que mide sólo la eficiencia potencial del diseño BIB porque no es posible estimar  $\sigma^2$  para el diseño BCA.

### Factor de eficiencia para los diseños de bloques incompletos

La varianza de la diferencia entre dos medias de tratamiento en el diseño BIB es  $2k\sigma_{bib}^2 / \lambda t$ . Si este diseño y el BCA tienen el *mismo número de tratamientos y réplicas*, la eficiencia del diseño BIB respecto al diseño BCA es la razón de las varianzas:

$$Eficiencia = \frac{(2\sigma_{rcb}^2 / r)}{(2k\sigma_{bib}^2 / \lambda t)} = \frac{\sigma_{rcb}^2}{\sigma_{bib}^2} * \frac{\lambda t}{rk} \quad (21)$$

La cantidad  $E = \lambda t / rk$  se llama factor de eficiencia para el diseño de bloques completo balanceado y proporciona una indicación de la pérdida causada al usar un diseño de bloque incompleto sin reducir  $\sigma^2$ . El diseño BIB es más preciso en la comparación de dos medias de tratamiento que el diseño BCA si

$$\frac{\sigma_{bib}^2}{\sigma_{rcb}^2} < \frac{\lambda t}{rk} \quad (22)$$

El factor de eficiencia es un límite inferior de la eficiencia del diseño BIB con respecto al diseño BCA para un experimento BCA para un experimento con el

mismo número de réplicas y la misma varianza del error  $\sigma^2$ . Dados un diseño BIB con  $t = 4$ ,  $r = 3$ ,  $k = 3$  y  $\lambda = 2$  y un diseño BCA con  $t = 4$  y  $r = 3$ , ambos requieren el mismo número de unidades experimentales, pero tienen diferentes arreglos de bloque. El diseño BIB es más preciso que el diseño BCA si:

$$\frac{\sigma_{bib}^2}{s_{rcb}^2} < \frac{2(4)}{3(3)} = 0.89$$

En otras palabras,  $\sigma_{bib}^2$  tendría que ser cerca de un 11% más pequeño que  $\sigma_{rcb}^2$  para que el diseño BIB tuviera la misma precisión que el diseño BCA con el mismo número de réplicas.

La intención de un diseño de bloques incompleto es reducir la varianza del error e incrementar la precisión de las comparaciones entre las medidas de tratamiento, la meta sería reducir  $\sigma^2$  para los diseños BIB de manera que se logre la desigualdad de la ecuación (22). Un diseño de bloques incompleto exitoso reducirá la varianza del error  $\sigma_{bib}^2$  será menor que  $\sigma_{rcb}^2$ .

### 4.3 Aplicaciones de Diseños Experimentales de Bloques Incompletos en el sector industrial

Los diseños de bloques incompletos se utilizan para reducir la varianza del error experimental, ya que a veces puede no ser posible realizar una réplica completa para factoriales  $2^k$  con muchos factores en un bloque completo. *Si la materia prima no es suficiente en un lote de producción para manejar todos los tratamientos, se puede usar cada lote de materia prima como un bloque incompleto.* Se pueden elaborar bloques de tamaño reducido para factoriales  $2^k$  aprovechando la construcción de los contrastes de sus efectos y *sacrificando información de los tratamientos para aumentar la precisión.*

Los diseños de bloque incompletos para factoriales  $2^k$  se construyen de manera que uno o más contrastes de tratamiento sean idénticos a los contrastes de bloque. Se dice que el efecto del tratamiento se **confunde** por completo con los bloques y no es posible distinguirlo del efecto de esos bloques.

- El diseño factorial  $2^k$  con muchos factores, cada uno a nivel “alto” o “bajo”, se puede usar para detectar los factores importantes en el proceso con un mínimo de unidades experimentales; es posible detectar las tendencias principales con factores de dos niveles para identificar los factores potencialmente importantes.
- En consecuencia, los factoriales  $2^k$  se usan con frecuencia en las primeras etapas de experimentación para detectar los factores que son candidatos potenciales para una investigación detallada.
- El nivel bajo de un factor *cuantitativo* se denota con “0” y el nivel alto con “1”. De manera equivalente, las dos categorías del factor cualitativo se pueden codificar como “0” y “1”.
- El “efecto principal” de un factor es básicamente el cambio promedio en la respuesta cuando el factor va de '-' a '+’.
- Los “efectos de las interacciones” son cambios en la salida del proceso motivados por dos ó más factores que están interactuando entre sí.
- El número de las combinaciones de tratamientos necesario para tener resultados completos sería igual a  $2^k$ .
- Así, un experimento factorial  $2^k$  que tiene 3 factores ( $k = 3$ ) necesitará 8 combinaciones de tratamientos, mientras que si tenemos 4 factores ( $k=4$ ), necesitará 16 combinaciones. Podemos apreciar que el número de corridas

necesario para completar un experimento factorial, aún cuando son sólo dos niveles por cada factor, se hace enormemente grande.

Como la mitad de los tratamientos tiene coeficiente (+) y la mitad coeficiente (-) para todo efecto en los factoriales  $2^k$ , los tratamientos se pueden dividir en 2 grupos con la **regla pares-impares**.

La suma de cuadrados se calcula de manera usual para el **análisis de varianza** excepto que se excluye la partición de la suma de cuadrados para el efecto de interacción confundido con los bloques.

En general, se elige el efecto de interacción de mayor orden factorial  $2^k$  para confundirlo con los bloques. En el caso de experimentos con muchos factores, los efectos principales, las interacciones de dos factores y otras interacciones de menor orden son los efectos de mayor interés.

#### **Ejemplo 4.1 Pureza de un producto químico**

Se pensó que la pureza de un producto químico tenía influencia de 3 factores, tasa de agitación (A), concentración del compuesto base (B) y concentración del reactivo (C). El químico estableció un experimento con un diseño factorial con factores a 2 niveles para un arreglo factorial  $2^3$ .

**Diseño del experimento:** el químico deseaba 3 réplicas del experimento, pero sólo podía realizar 4 corridas del proceso químico en un día. Por lo tanto, cada réplica debía correrse en dos bloques incompletos (días).

Un contraste de efecto para un factorial  $2^3$  consiste en cuatro combinaciones de tratamientos con coeficiente + y cuatro con -. Para evitar la total confusión de un efecto, el químico confundió una interacción de 2 factores distinta en cada réplica. Es decir:  $b = 2$  y  $r = 3$  realizando la disposición de modo que en cada

réplica se encontraran todos los tratamientos pero distribuidos en los dos bloques. En la primera réplica aparece la interacción BC confundida, en la segunda réplica aparece la interacción AC confundida y en la última réplica aparece la interacción AB confundida.

El análisis de varianza se realizó mediante el paquete estadístico SPSS, versión 11.0.

A continuación aparecen las tablas de los tratamientos con los datos recolectados y finalmente el Análisis de Varianza del diseño utilizado.

**Tabla 4.1 Pureza observada de un producto químico en un factorial 2<sup>3</sup> parcialmente confundido.**

BC Confundido		AC Confundido		AB Confundido					
+1		-1		+1		-1			
(1)	25	ab	43	abc	39	(1)	26	a	43
bc	34	c	30	b	29	c	32	b	43
abc	42	ac	40	(1)	27	ab	52	ac	40
a	25	b	33	ac	40	c	34	bc	36
Bloque 1		Bloque 2		Bloque 3		Bloque 5		Bloque 6	
Réplica 1		Réplica 2		Réplica 3					

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

Las sumas de cuadrados para los efectos no confundidos con bloques, A, B, C y ABC, se pueden calcular de manera normal mediante la ecuación:

$$SC(AB...) = \frac{r}{2^n} (l_{AB...})^2$$

utilizando los coeficientes de los contrastes de la tabla 4.2.

Tabla 4.2 Coeficientes para contrastes en un diseño de tratamiento factorial 2<sup>3</sup>.

Efectos Factoriales									
TRATAMIENTO	I	A	B	C	AB	AC	BC	ABC	Y
(1)	+	-	-	-	+	+	+	-	32
A	+	+	-	-	-	-	+	+	35
B	+	-	+	-	-	+	-	+	28
Ab	+	+	+	-	+	-	-	-	31
C	+	-	-	+	+	-	-	+	48
Ac	+	+	-	+	-	+	-	-	39
Bc	+	-	+	+	-	-	+	-	28
Abc	+	+	+	+	+	+	+	+	29
Divisor	8	4	4	4	4	4	4	4	
Efecto	33.8	-0.5	-9.5	4.5	2.5	-3.5	-5.5	2.5	
SC		0.5	180.5	40.5	12.5	24.5	60.5	12.5	

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

Tabla 4.3 Cálculos para los efectos factoriales no confundidos con bloques.

	(1)	a	b	ab	c	ac	bc	abc			
Medias	26	35	32	47	32	40	36	44			
	SC										
A	-	+	-	+	-	+	-	+	40	8	600.0
B	-	-	+	+	-	-	+	+	26	8	253.5
C	-	-	-	-	+	+	+	+	12	8	54.0
ABC	-	+	+	-	+	-	-	+	-6	8	13.5

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

**Tabla 4.4** Los contrastes confundidos AB, AC, y BC calculados con los totales de todas las observaciones de tratamientos.

	(1)	a	b	ab	c	ac	bc	abc	
<b>Medias</b>	78	105	95	141	96	120	108	132	
	$\Sigma K_i$								
<b>AB</b>	+	-	-	+	+	-	-	+	18
<b>AC</b>	+	-	+	-	-	+	-	+	-24
<b>BC</b>	+	+	-	-	-	+	+	+	-30

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

**Tabla 4.5** Diferencia calculada para cada par de bloques

	Réplica I		Réplica II		Réplica III	
	BC confundido		AC confundido		AB confundido	
	Bloque 1	Bloque 2	Bloque 1	Bloque 2	Bloque 1	Bloque 2
<b>Totales</b>	126	146	135	155	161	153
<b>Diferencia</b>	-20		-20		8	

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

Tabla 4.6 Análisis de varianza para la pureza de un producto químico en un factorial  $2^3$  parcialmente confundido

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F_0
Réplicas	2	111.00	55.50	
Bloques dentro de las réplicas	3	106.00	36.00	
A	1	600.00	600.00	40.6
B	1	253.50	253.50	17.2
C	1	54.00	54.00	3.7
ABC	1	13.50	13.50	<1
AB (Réplicas I, II)	1	6.25	6.25	<1
AC (Réplicas I, III)	1	1.00	1.00	<1
BC (Réplicas II, III)	1	6.25	6.25	<1
Error	11	162.50	14.77	
Total	23	1316.00		

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

Como se puede apreciar ninguna de las interacciones tuvieron diferencias significativas y sólo los efectos principales A (tasa de agitación) y B (concentración del compuesto base) resultaron significativos. La varianza del error fue muy baja (14.77) y permitió detectar las verdaderas diferencias entre los factores A y B por separado, con el consiguiente ahorro de material al no tener que utilizar las réplicas completas, lo cual demuestra la efectividad del uso de un diseño de bloques incompletos para detectar los efectos en un experimento en este caso, de un factorial  $2^3$ .

**Ejemplo 4.2 Eficiencia de un diseño de bloques incompletos con respecto a otro diseño de bloques completos aleatorizado para Vinilación de metilglucósidos.**

Retomaremos el ejemplo 2.2 del capítulo 2 donde se presenta un diseño de experimentos de bloques incompletos para Vinilación de metilglucósidos, solo se toma la tabla 2.2 y 2.3 del capítulo 2 para: compararlas con un diseño de bloques completos para Vinilación de metilglucósidos, determinar la eficiencia y precisión de un diseño BIB.

**Ejemplo:** Se ha encontrado que el resultado de la adición de acetileno a un metilglucósido en presencia de una base, a alta presión, es la producción de varios éteres monovinílicos, proceso que se conoce como vinilación. Los éteres monovinílicos son convenientes para polimerización en muchas aplicaciones industriales.

**Objetivo de investigación:** Los químicos deseaban obtener más información específica sobre el efecto de la presión en el porcentaje de conversión del metilglucósido en isómeros de monovinil.

**Diseño de tratamiento:** Con base en trabajos anteriores, se seleccionaron presiones dentro del intervalo que producía la conversión máxima. Se eligieron 5 presiones para estimar una ecuación de respuesta: 250, 325, 400, 475 y 550 psi.

## Diseño de Bloques Incompletos para Vinilación de metilglucósidos

**Diseño del experimento:** como sólo se disponía de tres cámaras de alta presión para una corrida de las condiciones experimentales, fue necesario bloquizar las corridas por que no podía haber una variación sustancial de una corrida a otra producida por nuevas preparaciones de las cámaras para el experimento. Los químicos establecieron un diseño de bloques incompleto balanceado (corridas), cada uno con tres unidades experimentales (cámaras presurizadas) y se usaron tres presiones diferentes en cada corrida; el diseño obtenido tenía  $r = 6$  replicas,  $k = 3$  unidades experimentales,  $t = 5$  tratamientos y  $\lambda = 3$ .

Las presiones usadas en cada corrida y las conversiones porcentuales a isómeros de monovinil se muestran en la tabla 2.2 y la suma aditiva de la partición de suma de cuadrados del metilglucósido se muestra en la tabla 2.3.

### Diseño del Bloque Incompleto

Presión (psi)

	250	325	400	475	550
Bloque 1	16	18	–	32	–
Bloque 2	19	–	–	46	45
Bloque 3	–	26	39	–	61
Bloque 4	–	–	21	35	55
Bloque 5	–	19	–	47	48
Bloque 6	20	–	33	31	–
Bloque 7	13	13	34	–	–
Bloque 8	21	–	30	–	52
Bloque 9	24	10	–	–	50
Bloque 10	–	24	31	37	–

**Tabla 2.2 Conversión porcentual del metilglucósidos con acetileno a acetileno a alta presión, en un diseño de bloque incompleto balanceado.**

Corrida	Presión (psi)					$y_j$
	250	325	400	475	550	
1	16	18	–	32	–	66
2	19	–	–	46	45	110
3	–	26	39	–	61	126
4	–	–	21	35	55	111
5	–	19	–	47	48	114
6	20	–	33	31	–	84
7	13	13	34	–	–	60
8	21	–	30	–	52	103
9	24	10	–	–	50	84
10	–	24	31	37	–	92
$y_{i.}$	113	110	188	228	311	950
$*B_i.$	507	542	576	577	648	
$^{\dagger}Q_i$	-56.0	-70.7	-4.0	35.7	95.0	

\*Ejemplo:  $B_1 = y_{.1} + y_{.2} + y_{.6} + y_{.7} + y_{.8} + y_{.9}$   
 $^{\dagger}Q_i = y_{i.} - \frac{1}{3} B_i = 113 - \frac{1}{3} (507) = -56.0$

Fuente: Drs J. Berry y A. Deutschman, University of Arizona

**Tabla 2.3 Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloque incompleto balanceado.**

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	Pr<F
Total	29	5576.67			
Bloques (no. Aj.)	9	1394.67	154.96		
Presión (aj.)	4	3688.58	922.14	29.90	.000
Error	16	493.42	<b>30.84</b>		

Fuente: Drs J. Berry y A. Deutschman, University of Arizona

### Diseño de Bloques Completos para Vinilación de metilglucósidos

**Diseño del experimento:** fue necesario bloquizar las corridas por que no podía haber una variación sustancial de una corrida a otra producida por nuevas preparaciones de las cámaras para el experimento. El diseño del experimento resultante fue un diseño de bloques completo aleatorizado (corridas), con  $b= 6$  bloques,  $r= 6$  replicas y  $t= 5$  tratamientos.

Las presiones usadas en cada corrida y las conversiones porcentuales a isómeros de monovinil se muestran en la tabla 4.7 y la suma aditiva de la partición de suma de cuadrados del metilglucósido se muestra en la tabla 4.8.

#### Diseño del Bloque Completo

	Presión (psi)				
	250	325	400	475	550
<b>Bloque 1</b>	16	18	39	32	45
<b>Bloque 2</b>	19	26	21	46	61
<b>Bloque 3</b>	20	19	33	35	55
<b>Bloque 4</b>	13	13	34	47	48
<b>Bloque 5</b>	21	10	30	31	52
<b>Bloque 6</b>	24	24	31	37	50

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

**Tabla 4.7 Conversión porcentual del metilglucósidos con acetileno a acetileno a alta presión, en un diseño de bloques completos aleatorizado.**

Corrida	Presión (psi)					$y_j$
	250	325	400	475	550	
1	16	18	39	32	45	150
2	19	26	21	46	61	173
3	20	19	33	35	55	162
4	13	13	34	47	48	155
5	21	10	30	31	52	144
6	24	24	31	37	50	166
$y_i$	113	110	188	228	311	950

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

**Tabla 4.8 Análisis de varianza para la conversión porcentual de metilglucósidos, en un diseño de bloques completos balanceado.**

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F	$Pr < F$
Total	29	5576.66			
Replicas	5	114.66	22.93		
Presión	4	4736.33	1184.08	32.63	.000
Error	20	725.66	<b>36.28</b>		

Fuente: I. G. Hernández (2006), CIAII. Universidad Autónoma del Estado de Hidalgo

### Eficiencia del ejemplo

Para determinar la eficiencia del diseño de bloques incompletos del ejemplo aplicaremos la formulas 21 del punto 4.2 donde hace referencia al factor eficiencia.

$$Eficiencia = \frac{(2\sigma_{rcb}^2 / r)}{(2k\sigma_{bib}^2 / \lambda t)} = \frac{\sigma_{rcb}^2}{\sigma_{bib}^2} * \frac{\lambda t}{rk} \quad (21)$$

$$Eficiencia = \frac{36.28}{30.84} * \frac{(3)(5)}{(6)(3)} = 0.98$$

La eficiencia del diseño de bloques incompletos con respecto al diseño de bloques completos es del 98%.

En este ejemplo, el diseño de bloques incompletos es exitoso porque reduce la varianza del error donde  $\sigma_{bib}^2$  es menor que  $\sigma_{rcb}^2$ ; los resultados de los errores de varianza del ejemplo son:

$$\sigma_{bib}^2 = 30.84$$

$$\sigma_{rcb}^2 = 36.28$$

La intención del ejemplo es demostrar que un diseño de bloques incompletos reduce la varianza del error e incrementa la precisión de las comparaciones entre las medias de tratamiento ya que el objetivo de utilizar diseños de bloques incompletos es reducir  $\sigma^2$ .

**Apéndice 1A.1    Algunos diseños de bloques incompletos balanceados**

**Plan 1A.1**     $t = 1, k = 3, r = 6, b = 10, \lambda = 3, E = 0.83$

(1,2,3), (1,2,5), (1,4,5), (2,3,4), (3,4,5), reps 1,2,3

(1,2,4), (1,3,4), (1,3,5), (2,3,5), (2,4,5), reps 4,5,6

**Plan 1A.2**     $t = 6, k = 3, r = 5, b = 10, \lambda = 2, E = 0.80$

(1,2,5), (1,2,6), (1,3,4), (1,3,6), (1,4,5),

(2,3,4), (2,3,5), (2,4,6), (3,5,6), (4,5,6)

**Plan 1A.3**     $t = 6, k = 4, r = 10, b = 15, \lambda = 6, E = 0.90$

(1,2,3,4), (1,4,5,6), (2,3,5,6) reps 1,2

(1,2,3,5), (1,2,4,6), (3,4,5,6) reps 3,4

(1,2,3,6), (1,3,4,5), (2,4,5,6) reps 5,6

(1,2,4,5), (1,3,5,6), (2,3,4,6) reps 7,8

(1,2,5,6), (1,3,4,6), (2,3,4,5) reps 9,10

**Plan 1A.4**     $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78$

(1,2,4), (2,3,5), (3,4,6), (4,5,7), (1,5,6),

(2,6,7), (1,3,7)

**Plan 1A.5**     $t = 8, k = 4, r = 7, b = 14, \lambda = 3, E = 0.86$

(1,2,3,4), (5,6,7,8) rep 1                      (1,2,7,8), (3,4,5,6) rep 2

(1,3,6,8), (2,4,5,7) rep 3                      (1,4,6,7), (2,3,5,8) rep 4

(1,2,5,6), (3,4,7,8) rep 5                      (1,3,5,7), (2,4,6,8) rep 6

(1,4,5,8), (2,3,6,7) rep 7

**Plan 1A.6**     $t = 9, k = 3, r = 4, b = 12, \lambda = 1, E = 0.75$

(1,2,3), (1,4,7), (1,5,9), (1,6,8), (2,4,9), (2,5,8)

(2,6,7), (3,4,8), (3,5,7), (3,6,9), (4,5,6), (7,8,9)

**Plan 1A.7**  $t = 9, k = 4, r = 8, b = 18, \lambda = 1, E = 0.84$

(1,2,3,4), (1,3,8,9), (1,4,6,7), (1,5,7,8), (2,3,6,7),  
 (2,4,5,8), (2,6,8,9), (3,5,7,9), (4,5,6,9) reps 1,2,3,4  
 (1,2,4,9), (1,2,5,7), (1,3,6,8), (1,5,6,9), (2,3,5,6),  
 (2,7,8,9), (3,4,5,8), (3,4,7,9), (4,6,7,8) reps 5,6,7,8

**Plan 1A.8**  $t = 10, k = 3, r = 9, b = 30, \lambda = 2, E = 0.74$

(1,2,3), (1,4,6), (1,7,9), (2,5,8), (2,8,10)  
 (3,4,7), (3,9,10), (4,6,9), (5,6,10), (5,7,8), reps 1,2,3  
 (1,2,4), (1,5,7), (1,8,10), (2,3,6), (2,5,9)  
 (3,4,8), (3,7,10), (4,5,9), (5,7,10), (6,7,8), reps 4,5,6  
 (1,3,5), (1,6,8), (1,9,10), (2,4,10), (2,6,7)  
 (2,7,9), (3,5,6), (3,8,9), (4,5,10), (4,7,8), reps 7,8,9

**Plan 1A.9**  $t = 10, k = 4, r = 6, b = 15, \lambda = 2, E = 0.83$

(1,2,3,4), (1,2,5,6), (1,3,7,8), (1,4,9,10), (1,5,7,9),  
 (1,6,8,10), (2,3,6,9), (2,4,7,10), (2,5,8,10), (2,7,8,9),  
 (3,5,9,10), (3,6,7,10), (3,4,5,8), (4,5,6,7), (4,6,8,9)

**Plan 1A.10**  $t = 10, k = 5, r = 9, b = 18, \lambda = 4, E = 0.89$

(1,2,3,4,5), (1,2,3,6,7), (1,2,4,6,9), (1,2,5,7,8), (1,3,6,8,9), (1,3,7,8,10),  
 (1,4,5,6,10), (1,4,8,9,10), (1,5,7,9,10), (2,3,4,8,10), (2,3,5,9,10), (2,4,7,8,9),  
 (2,5,6,8,9), (2,6,7,9,10), (3,4,6,7,10), (3,4,5,7,9), (3,5,6,8,9), (4,5,6,7,8)

**Plan 1A.11**  $t = 11, k = 5, r = 5, b = 11, \lambda = 2, E = 0.88$

(1,2,3,5,8), (1,2,4,7,11), (1,3,6,10,11), (1,4,8,9,10),  
 (1,5,6,7,9), (2,3,4,6,9), (2,5,9,10,11), (2,6,7,8,10),  
 (3,4,5,7,10), (3,7,8,9,11), (4,5,6,8,11)

## Apéndice 1A.2 Algunos diseños de cuadrados latinos incompletos

**Tabla 1A** Planes para diseños de cuadrados latinos incompletos derivados de cuadrados latinos completos.

t	k	r	b	$\lambda$	E	Plan
4	3	3	4	2	0.89	*
	3	6	8	4	0.89	**
	3	9	12	6	0.89	**
5	4	4	5	3	0.94	*
	4	8	10	6	0.94	**
6	5	5	6	4	0.96	*
	5	10	12	8	0.96	**
7	6	6	7	5	0.97	*
8	7	7	8	6	0.98	*
9	8	8	9	7	0.98	*
10	9	9	10	8	0.99	*
11	10	10	11	9	0.99	*

\*Construido a partir de un cuadrado latino de  $t \times t$  por omisión de la última columna

\*\*Por repetición del plan para  $r = t - 1$ , que se construyó con un cuadrado latino de  $t \times t$  sin la última columna

### Planes para otros diseños de cuadrados latinos incompletos

**Plan 1B.1**  $t = 5, k = 2, r = 4, b = 10, \lambda = 1, E = 0.63$

Reps I y II

1	2	3	4	5
2	5	4	1	3

Reps III y IV

1	2	3	4	5
3	4	2	5	1

**Plan 1B.2**  $t = 5, k = 3, r = 6, b = 10, \lambda = 3, E = 0.83$

Reps I y II

1	2	3	4	5
2	1	4	5	3
3	5	2	1	4

Reps III y IV

1	2	3	4	5
2	3	4	5	1
4	5	1	2	3

**Plan 1B.3**  $t = 7, k = 2, r = 6, b = 21, \lambda = 1, E = 0.58$

Reps I y II

1	2	3	4	5	6	7
2	6	4	7	1	5	3

Reps III y IV

1	2	3	4	5	6	7
3	4	5	6	7	1	2

Reps V y VI

1	2	3	4	5	6	7
4	3	6	5	2	7	1

**Plan 1B.4**  $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78$

7	1	2	3	4	5	6
1	2	3	4	5	6	7
3	4	5	6	7	1	2

**Plan 1B.5**  $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78$

Reps I y II

1	2	3	4	5	6	7	8	9
2	8	4	7	6	1	3	9	5

Reps V y VI

1	2	3	4	5	6	7	8	9
4	6	2	5	7	8	9	1	3

Reps III y IV

1	2	3	4	5	6	7	8	9
3	5	6	9	8	7	1	4	2

Reps VII y VIII

1	2	3	4	5	6	7	8	9
5	4	8	6	3	9	2	7	1

**Plan 1B.6**  $t = 9, k = 4, r = 8, b = 18, \lambda = 3, E = 0.84$

Reps I, II, III y IV

1 2 3 4 5 6 7 8 9  
 4 6 8 1 7 9 3 2 5  
 6 8 9 3 1 4 2 5 7  
 7 9 1 2 8 5 6 4 3

Reps V, VI, VII y VIII

1 2 3 4 5 6 7 8 9  
 2 3 4 9 1 8 6 5 7  
 5 6 7 2 9 1 4 3 8  
 7 5 1 9 6 3 8 4 2

**Plan 1B.7**  $t = 9, k = 4, r = 8, b = 18, \lambda = 3, E = 0.84$

Reps I, II, III y IV

1 2 3 4 5 6 7 8 9  
 2 6 8 3 1 4 9 5 7  
 3 8 5 9 7 2 1 4 6  
 7 4 9 2 3 5 6 1 8  
 8 1 2 6 4 7 3 9 5

Reps V, VI, VII y VIII

1 2 3 4 5 6 7 8 9  
 2 6 5 3 7 8 4 9 1  
 3 5 1 2 9 7 8 4 6  
 5 1 4 8 2 3 9 6 7  
 9 8 6 7 4 5 1 3 2

**Plan 1B.8**  $t = 10, k = 3, r = 9, b = 30, \lambda = 2, E = 0.74$

Reps I, II y III

1 2 3 4 5 6 7 8 9 10  
 2 5 7 1 8 4 9 10 3 6  
 3 8 4 6 7 9 1 2 10 5

Reps IV, V y VI

1 2 3 4 5 6 7 8 9 10  
 2 3 4 9 7 8 10 1 5 6  
 4 6 8 5 1 9 3 10 2 7

Reps VII, VIII y

1 2 3 4 5 6 7 8 9 10  
 IX  
 3 7 8 2 6 1 9 4 10 5  
 5 6 9 10 3 8 2 7 1 4

### Apéndice 1A.3 Estimaciones de mínimos cuadrados para diseño BIB

El modelo lineal para el diseño de bloque incompleto balanceado es

$$\begin{aligned} \gamma_{ij} &= \mu + \tau_i + \rho_j + e_{ij} \\ i &= 1, 2, \dots, t \quad j = 1, 2, \dots, b \end{aligned} \quad (1A.1)$$

donde  $\mu$  es media general,  $\tau_i$  es el efecto de tratamiento,  $\rho_j$  es el efecto de bloque y las  $e_{ij}$  son errores experimentales aleatorios independientes con media 0 y varianza  $\sigma^2$ .

Sea  $n_{ij}$  una variable indicadora con  $n_{ij} = 1$  si el  $i$ -ésimo tratamiento está en el  $j$ -ésimo bloque y  $n_{ij} = 0$  de otra manera.

Las estimaciones de mínimos cuadrados para los parámetros en el modelo completo minimizan la suma de cuadrados para el error experimental:

$$Q = \sum_{i=1}^t \sum_{j=1}^b n_{ij} e_{ij}^2 = \sum_{i=1}^t \sum_{j=1}^b n_{ij} (\gamma_{ij} - \mu - \tau_i - \rho_j)^2 \quad (1A.2)$$

Las ecuaciones normales para el diseño de bloques incompleto balanceado son:

$$\mu: \quad N\hat{\mu} + r \sum_{i=1}^t \hat{\tau}_i + k \sum_{j=1}^b \hat{\rho}_j = y_{..} \quad (1A.3)$$

$$\tau_i: \quad r\hat{\mu} + r\hat{\tau}_i + \sum_{j=1}^b n_{ij} \hat{\rho}_j = y_{i.} \quad i = 1, 2, \dots, t \quad (1A.4)$$

$$\rho_j: \quad k\hat{\mu} + \sum_{i=1}^t n_{ij} \hat{\tau}_i + k\hat{\rho}_j = y_{.j} \quad j = 1, 2, \dots, b \quad (1A.5)$$

Cada una de las sumas del conjunto de  $t$  ecuaciones para las  $\tau_i$  y el conjunto de  $b$  ecuaciones para las  $p_j$  es igual a la ecuación para  $\mu$ , esto tiene como resultado dependencias lineales entre las ecuaciones. Las restricciones  $\sum_i \hat{\tau}_i = 0$  y  $\sum_j \hat{p}_j = 0$  se pueden usar para proporcionar una solución única.

Al imponer las restricciones, la estimación de  $\mu$  es  $\hat{\mu} = y./N$  a partir de la ecuación (1A.3). Las ecuaciones para  $p_j$  ecuaciones (1A.5), se usan para eliminar las  $\hat{p}_j$  de las ecuaciones  $\tau_i$  en la ecuación (1A.4). Después de eliminar las ecuaciones con  $\hat{\tau}_i$  son:

$$rk\hat{\tau}_i - r\hat{\tau}_i - \sum_{j=1}^b \sum_{p=1}^t n_{ij}n_{pj}\hat{\tau}_p = ky_i - \sum_{j=1}^b n_{ij}y_{.j} \quad (1A.6)$$

El lado derecho de la ecuación (1A.6) es  $kQ_i$  es el total de tratamiento ajustado que se usa para calcular la suma de cuadrados de tratamiento ajustada. Como:

$$\sum_{j=1}^b n_{ij}n_{pj} = \lambda, \quad \text{para } p \neq i \quad (1A.7)$$

y  $n_{pj}^2 = n_{pj}$  ( $n_{pj} = 0$  o  $1$ ), la ecuación (1A.6) se puede escribir como:

$$r(k-1)\hat{\tau}_i - \lambda \sum_{\substack{p=1 \\ p \neq i}}^t \hat{\tau}_p = kQ_i \quad (1A.8)$$

Como se impuso la condición  $\sum_i \hat{\tau}_i$  sobre la solución, la situación:

$$-\hat{\tau}_i = \sum_{\substack{p=1 \\ p \neq i}}^t \hat{\tau}_p$$

Se puede convertir en la ecuación (1A.8). Con la igualdad  $\lambda(t - 1) = r(k - 1)$ , la ecuación para  $\hat{t}_i$  es :

$$\lambda t \hat{t}_i = kQ_i \quad (1A.9)$$

$$\hat{t}_i = \frac{kQ_i}{\lambda t} \quad i = 1, 2, \dots, t$$

## Apéndice 2A.1 Planes para diseños cíclicos

Tabla 2A.1 Bloques iniciales para generar diseños de bloques incompletos parcialmente balanceados para  $4 \leq t \leq 15$ ,  $r \leq 10$

k	r	Primeros k – 1 tratamientos	k-ésimo tratamiento, t =											
			4	5	6	7	8	9	10	11	12	13	14	15
2	2	0	1	1	1	1	1	1	1	1	1	1	1	1
	4	0	2	2	2	3	3	3	3	3	3	5	4	4
	6	0	1	3	3	2	2	2	2	5	5	2	6	2
	8	0	1	4	5	4	4	4	4	2	2	4	3	7
	10	0	2	1	4	5	5	5	5	4	4	3	5	5
3	3	0 1	2	2	3	3	3	3	4	4	4	4	4	4
	6	0 2	1	3	1	3	7	6	7	7	5	7	7	8
	9	0 1	3	2	3	3	4	5	3	3	6	4	6	5
4	4	0 1 3	–	2	2	6	7	7	6	7	7	9	7	7
	8	0 1 4	–	2	2	6	7	8	2	6	6	6	6	6
	5	0 1 2 4	–	–	5	5	7	7	7	7	7	7	9	10
5	10	0 2 3 4	–	–	5	5	7	8	9	8				
		0 2 3 6									7	11	12	10
6	6	0 1 2 3 6	–	–	–	5	5	5	5	10	10	10	10	10
7	7	0 1 2 3 4 7	–	–	–	–	5	5	9	9	9	9	9	10
8	8	0 1 2 3 4 6 8	–	–	–	–	–	5	9	9	9	9	11	11
9	9	0 1 2 3 4 5 7 9	–	–	–	–	–	–	8	8	8	10	10	10
10	10	0 1 2 3 4 5 6 9 10	–	–	–	–	–	–	–	7	7	7	12	12

Reproducido con el permiso de J. A. John (1981), "Efficient Cyclic Desgns", Journal of Royal Statical Society, B, 43, 76-80.

## Apéndice 2A.2 Arreglos para diseños $\alpha$

Tabla 2A.2 Arreglos generados de  $(k \times r)$  para diseños  $\alpha$

$s = k = 5$				$s = k = 6$				$s = k = 7$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	4	2	0	1	5	4	0	1	3	2
0	2	3	4	0	3	2	5	0	2	6	4
0	3	2	1	0	2	3	1	0	4	5	1
0	4	1	3	0	4	1	2	0	3	2	6
				0	5	1	3	0	5	1	3
								0	6	4	5
$s = k = 8$				$s = k = 9$				$s = k = 10$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	6	0	1	8	7	0	1	9	5
0	3	7	1	0	3	6	4	0	3	6	9
0	5	3	4	0	7	2	3	0	5	7	2
0	2	5	3	0	2	3	5	0	4	5	6
0	4	1	6	0	4	1	6	0	6	3	1
0	6	0	2	0	5	7	2	0	7	2	4
0	7	6	5	0	6	5	1	0	8	4	7
				0	8	4	7	0	9	8	2
								0	2	6	3
$s = 11 \ k = 9$				$s = 12 \ k = 8$				$s = 13 \ k = 7$			
0	0	0	0	0	0	0	0	0	0	0	0
0	1	6	7	0	1	2	3	0	1	4	10
0	4	8	1	0	7	5	1	0	3	8	11
0	9	7	5	0	9	6	4	0	9	2	1
0	2	3	6	0	4	11	8	0	12	10	6
0	5	1	3	0	11	3	10	0	8	5	12
0	6	5	10	0	10	4	7	0	6	7	8
0	3	9	4	0	5	1	6				
0	7	4	1								
$s = 14 \ k = 7$				$s = 15 \ k = 6$							
0	0	0	0	0	0	0	0				
0	1	8	10	0	1	8	7				
0	9	10	7	0	3	12	14				
0	11	13	2	0	7	2	5				
0	2	6	1	0	10	3	11				
0	5	1	12	0	14	3	8				
0	3	1	11								

Reproducido de H. D. Patterson, E. R. Williams y E. A. Hunter (1978), "Block designs for variety trials", Journal of Agricultural Science 90, tabla 2, p. 399, con permiso de Cambridge University Press.

## CONCLUSIONES

### Elección de un diseño de experimento

- Un experimento exitoso incluye el diseño del experimento correcto para contestar las preguntas de investigación específicas y proporcionar estimaciones precisas de las medidas de tratamiento y los contrastes de interés para el investigador.
- El conocimiento del investigador sobre las condiciones del experimento y el material, es indispensable para construir los bloques de tamaño y composición adecuados. Cada experimento representa sus propios retos y condiciones; ningún conjunto de reglas proporciona el diseño correcto para todo experimento exitoso con la aplicación inteligente y con los principios de diseño para el material experimental.
- Las pautas que guían la selección o construcción de un diseño para un problema de investigación específico, es hacer todo lo posible para elegir un diseño adecuado para el problema de investigación.
- Cuando los diseños estándar no se ajustan al problema de investigación, es necesario encontrar una solución de diseño innovadora, con un enfoque más informal para el diseño de experimentos (Mead, 1988). La razón más importante para pensar en un diseño informal es evitar la trampa de cambiar el diseño del tratamiento a fin de ajustarse a un diseño existente en particular, en ninguna circunstancia sería inteligente cambiar el problema de investigación para ajustarse a un diseño experimental.

### **Elección de diseños de bloques incompletos**

- Con frecuencia se necesita un diseño de bloques incompletos para obtener las mejores condiciones posibles para las comparaciones entre tratamientos con grupos tan homogéneos de unidades experimentales como sea posible.
- Los diseños de bloques incompletos tuvieron su origen en pruebas experimentales que incluyeron un gran número de tratamientos, como en las pruebas de variedades.
- El objetivo principal de los diseños de bloques incompletos es reducir el tamaño de bloque para reducir la varianza del error experimental y tener experimentos que se puedan manejar con base en grupos de réplicas; en las pruebas agrícolas, los tamaños de bloque son flexibles siempre que su reducción logre disminuir la varianza del error.
- En la mayoría de los casos, se dispone de diseños apropiados para las necesidades de los experimentos, y se cuenta con un gran número de planes publicados para los diseños de bloques incompletos balanceados, producto de investigaciones sobre construcción de diseños.
- Los métodos sistemáticos para construir diseños de bloques incompletos balanceados requiere desarrollos matemáticos que están fuera del alcance de este trabajo, pero se pueden encontrar exposiciones de las investigaciones sobre la existencia y construcción de diseños en John (1971), John (1987) y en John y Williams (1995).

### **Dónde encontrar planes de diseños**

- Los programas de computadora proporcionan el método más conveniente para desarrollar un plan de diseño de bloque incompleto; CycDesigN (Whitaker, Williams y John, 1998) y ALPHA + (Williams y Talbot, 1993) son ejemplos de programas que generan los diseños presentados en los capítulos 2 y 3.

### **El factor de eficiencia aumenta con el tamaño de bloque**

- Se espera que la varianza del error experimental disminuya conforme lo hace el tamaño del bloque, la disminución en  $\sigma^2$  con tamaños de bloques más pequeños depende del éxito logrado al colocar las unidades experimentales en el arreglo correcto de bloques, pero sin embargo, una inspección cuidadosa del factor de eficiencia sugiere adoptar un tamaño de bloque tan grande como sea posible. La medida de eficiencia para un diseño de bloque incompleto balanceado se puede expresar como:

$$E = \frac{\lambda t}{rk} = \frac{t}{t-1} \left( 1 - \frac{1}{k} \right)$$

(C.1)

El valor de  $E$  aumenta con el tamaño de bloque  $k$ , una relación similar se cumple para los diseños de bloque incompletos parcialmente balanceados.

### **Los recursos determinan el tamaño de bloque**

- Los recursos disponibles determinan la elección del tamaño del bloque porque en algunos experimentos el tamaño de bloque está restringido por

los instrumentos disponibles como ocurrió en el experimento de vinilación de metilglucósidos del ejemplo 2.2. El investigador disponía sólo de tres cámaras para probar cinco tratamientos de presión, con  $t = 5$ ,  $r = 6$ ,  $k = 3$  y  $\lambda = 3$ , el factor de eficiencia para el diseño de bloques incompleto balanceado era  $E = 3(5)/6(3) = 0.83$ . Una precisión equivalente para un diseño de bloques incompletos con cinco cámaras de presión requería de una reducción del 17% en  $\sigma^2$  para el diseño, pero la única elección disponible para el investigador era el diseño de bloques incompletos con tres tratamientos de bloque.

- Algunas condiciones experimentales permiten una mayor flexibilidad al seleccionar el tamaño de bloque, como las pruebas de variedades que se colocan en parcelas de una granja experimental. En este caso, el tamaño de bloque debe ser lo más grande que se pueda mientras mantenga la precisión necesaria para las comparaciones de variedades.

#### **Los diseños parcialmente balanceados aumentan la eficiencia en el uso de recursos.**

- Los diseños parcialmente balanceados se pueden usar en lugar de los diseños balanceados para optimizar el uso de recursos. Examinando el diseño parcialmente balanceado del cuadro 3, los tamaños de bloque se restringieron a cuatro unidades experimentales, en tanto que un diseño balanceado habría requerido diez réplicas de los seis tratamientos y 15 bloques.
- 
- El diseño parcialmente balanceado proporciona más precisión en algunas comparaciones que en otras debido a que no todos los pares de tratamientos ocurren en el mismo número de bloques; sin embargo, con una

elección prudente del diseño es posible tener un diseño parcialmente balanceado que proporcione casi un equilibrio en todas las comparaciones.

- El diseño parcialmente balanceado del cuadro 3, que tenía algunos pares de tratamientos en  $\lambda_1 = 2$  bloques y otros en  $\lambda_2 = 1$  bloque, se puede demostrar que la eficiencia de las comparaciones respectivas entre primeros y segundos asociados es  $E_1 = 1.00$  y  $E_2 = 0.86$ , con una eficiencia promedio de  $E = 0.88$  (Clatworthy, 1973). Existe una precisión casi igual para ambos conjuntos de asociados y sólo se sacrifica una pequeña parte de esa precisión con el diseño parcialmente balanceado.

### **Diseños resolubles y cíclicos**

- Los diseños resolubles tienen bloques agrupados de manera que cada grupo constituye una réplica completa de los tratamientos, cuestión muy eficaz para administrar un experimento ya que reduce el tamaño de los bloques para obtener mayor precisión permitiendo controlar experimentos muy grandes.
- Otra de las ventajas de los diseños resolubles es que pueden llevarse a cabo en etapas, con una o más replicas terminadas en cada etapa.
- Los diseños de retícula (lattice) balanceados son un grupo particular muy conocidos de diseños resolubles propuestos por Yates (1936b), aunque es muy limitado debido a que el número de tratamientos debe ser un cuadrado exacto.
- Los diseños de retícula rectangulares desarrollados por Harshbarger (1949) proporcionan diseños resolubles con números de tratamientos intermedios a los proporcionados por los diseños de retícula cuadrados.

- Los diseños balanceados o los resolubles pueden, en algunos casos, agotar los recursos disponibles debido al número de réplicas requerido, luego es posible usar diseños parcialmente balanceados más flexibles en lugar de los diseños balanceados con poca pérdida de precisión, pero tal vez no se disponga de tablas de diseños parcialmente balanceados porque muchas aparecen en publicaciones antiguas.
- Los diseños cíclicos ofrecen cierto grado de simplificación en la construcción del diseño y en el establecimiento del proceso experimental y como se generan a partir de un bloque inicial de tratamientos, es necesario un esquema completo de los planes de distribución del experimento.
- Quizá los diseños cíclicos sean la mejor alternativa para producir un diseño eficiente, pues proporcionan diseños de muchos tamaños y su construcción es sencilla a partir de uno o dos bloques iniciales de tratamientos. Las tablas de bloques iniciales para los diseños cíclicos se encuentran en publicaciones un poco más recientes.

### **Recomendación**

- Todos los problemas prácticos tienen particularidades que deben estudiarse antes de que se adopten los métodos estadísticos más efectivos para resolverlos. Por consecuencia, cada problema nuevo debe ser tratado por si mismo y con un cierto respeto ya que podríamos obtener la solución correcta de un problema equivocado. Las técnicas estadísticas son más efectivas cuando se combinan con el apropiado conocimiento del tema al que se aplican. Los métodos son una ayuda importante no un sustituto del conocimiento del problema por parte del investigador.

**BIBLIOGRAFÍA**

Bose, R. C., Clatworthy, W. H., and Shrikhand, S. S. (1954). "Tables of partially balanced designs with two associate cases." North Carolina Agricultural Experiment Station. Technical Bulletin, no. 107.

Bose, R. C., and Nair, K. R. (1939). "Partially balanced incomplete block designs." *Sankhya* 4, 337-372.

Box, G. E. P., Hunter, W. G, and Hunter, J. S. (1978). *Statistics for experimenters. An introduction to design, data analysis, and model building*. New York: Wiley.

Box, G. E. P., Hunter, W. G. & Hunter, S. J. (2001). *Estadística para investigadores*. Barcelona, España: Editorial Reverte, S. A.

Clatworthy, W. H. (1973). "Tables de two-associate-class partially balanced designs." National Abureau of Standards, Applied Mathematics Series, no. 63. Washington, D.C.

Cochran, W. G., and Cox, G. M. (1957). *Experimental Designs*, 2d ed. New York: Wiley.

Cox, D. R. (1958). *Planning of Experiments*. New York: Wiley.

Federer, W. T. (1955). *Experimental design: Theory and application*. New York: Macmillan.

Fisher, R. A., Yates, F. (1963). *Statistical tables for biological, agricultural and medical research*, 6<sup>th</sup> ed. Edinburgh: Oliver and Boyd.

Freíd, J. E. & Simon G. A. (2002). *Estadística elemental* (8<sup>va</sup> ed.). Edo. de México: Prentice Hall.

Gil, S. I. & Zárate de Lara G. P. (2000). *Métodos estadísticos* (6<sup>ta</sup> ed.). México, D. F.: Trillas.

Harshbarger, B. (1949). "Triple rectangular lattices." *Biometrics* 5, 1-13.

Harshbarger, B., and Davis, L. L. (1952). "Latinized rectangular lattices." *Biometrics* 8, 73-84.

Hildebrand, D. K. & Lyman, R. O. (1998). *Estadística aplicada a la administración y a la economía*. Edo. de México: Prentice Hall.

John, J. A. (1966). "Cyclic incomplete block designs." *Journal of the Royal Statistical Society, B* 28, 345-360.

John, J. A. (1987). *Cyclic designs*. London: Champan and Hall.

John, J. A., Williams, E. R. (1995). *Cyclic and computer-generated designs*, 2d ed., London: Champan and Hall.

John, P. W. M. (1971). "Statistical design and analysis of experiments." New York: Macmillan.

Kempthorne, O. (1952). *The design and analysis of experiments*. New York: Wiley.

Kuehl, R. O. (2001). *Diseño de experimentos, Principios estadísticos para el diseño y análisis de investigación* (2<sup>a</sup> ed.). México. D.F.: Ed Matemáticas THOMSON. (pp.310-390)

Manson, R. L., Gunst, R. F., and Hess, J. L. (1989). *Statistical design and analysis of experiments with applications to engineering and science*. New York: Wiley.

Mead, R. (1988). *The design of experiments. Statistical principles for practical application*. Cambridge, UK: Cambridge University Press.

Patterson, H. D., and Williams, E. R. (1976). *A new class of resolvable incomplete block designs*. *Biometrika* 63, 83-92.

Patterson, H. D., Williams, E. R., and Hunter, E. A. (1978). "Block designs for variety trials." *Journal of Agricultural Science* 90, 395-400.

Petersen, R. G. (1985). *Design and analysis of experiments*. New York: Marcel Dekker.

Piegnatiello, J. J., and Ramberg, J. S. (1985). *Journal of Quality Technology* 17, 198-206. Discussion of article by R. N. Kacker (1985). "Off-line quality control, parameter design, and the Taguchi method." *Journal of quality technology* 17, 176-188.

Whitaker, E. J., Williams, E. R., and John, J. A. (1998). *CycDesign*, CSIRO forestry and Forest Products, Canberra, Australia, and The University of Waikato, Hamilton, New Zealand.

Williams, E. R. (1986). Row and column designs with contiguous replicates. *Australian Journal of Statistical* 28, 154-163.

Williams, E. R., and Talbot, M. (1993). *ALPHA+: Experimental design for variety trials. Design user manual*. CSIRO, Canberra and SASS, Edinburgh.

Yates, F. (1936a). "Incomplete randomized blocks." *Annals of Eugenics* 7, 121-140.

Yates, F. (1936b). "A new method of arranging variety trials involving a large number of varieties." *Journal of Agricultural Science* 26, 424-455.

Yates, F. (1940a). "The recovery of interblock information in balanced incomplete block designs." *Annls of Eugenics* 10, 317-325.

Yates, f. (1940b). "Lattice squares." *Journal of agricultural Science* 30, 672-687.

Youden, W. J. (1937). "Use of incomplete block replications in estimating tobaccomosaic virus." *Contribuciones from Boyce Thomson Institute* 9, 41-48.

Youden, W. J. (1940). "Experimental designs to increase accuracy of greenhouse studies." *Contributions from Boyce Thomson Institute* 11, 219-228.