InnOvaciOnes de NegOciOs 4(2): 283 – 330, 2007
© 2007 UANL, Impreso en México (ISSN 1665-9627)

Diseños experimentales e investigación científica (Experimental designs and scientific research)

Badii, M.H, J. Castillo, M. Rodríguez, A. Wong & P. Villalpando
UANL, San Nicolás, N.L. 66450, México, mhbadii@yahoo.com.mx

Key words: Design, experiment, research, science

Abstract. The basics of the experimental designs are noted. Different features of common types of experimental designs such as the completely randomized design, the randomized block design, the Latin Square design, the split plot design and the factorial design are described. Each experimental design is illustrated by an example using real data. The application of experimental designs to the scientific research is discussed.

Palabras claves: Ciencia, diseño, experimento, investigación

Resumen. Se describen los fundamentos de los diseños experimentales. Se explican las distintas características de los diseños experimentales del uso común, tales como el diseño completamente aleatorio, e diseño de bloques al azar, el diseño de cuadro latino, el diseño de parcelas divididas y el de factorial. Para cada diseño se presenta un ejemplo con los datos reales del campo. Se discute la aplicación de estos diseños en relación con la investigación científica.

Introducción

Un diseño experimental es un esquema de cómo realizar un experimento. El objetivo fundamental de los diseños experimentales radica en el determinar si existe una diferencia significativa entre los diferentes tratamiento del experimento y en caso que la respuesta es afirmativa, cual sería la magnitud de esta diferencia. Una segunda meta de los diseños experimentales es verificar la existencia de una tendencia derivado del análisis de los datos del experimento. La diferencia principal entre los diseños experimentales radica en la forma en que se agrupan o clasifican las unidades experimentales. En todos los diseños las unidades experimentales se clasifican por tratamientos; pero en algunos, estos se clasifican preferentemente en bloques, filas, parcelas principales y otras modalidades. El análisis de varianza utiliza las medias de dichos agrupamientos, denominadas fuente de variación, para estimar varianzas o más precisamente cuadrados medios. Un cuadrado medio que estima la dispersión entre mediciones de parcelas debidas a causas aleatorias; esta se denomina error experimental. En ausencia de diferencias reales debidas a medias de los tratamientos, bloques u otras fuentes de variación, dichos cuadrados medios serán, en promedio, iguales. Sólo esporádicamente un cuadrado medio se desviará de otro de manera considerable, exclusivamente por casualidad. Cuando una prueba F indica que el cuadrado medio de una de las fuentes de variación es significativamente mayor que el cuadrado medio debido a efectos aleatorios, decimos que existen diferencias reales entre las medias de aquella fuente particular de variación; empero, recuérdese: siempre existe una probabilidad definida de que estemos equivocados en semejante conclusión. Está en manos del experimentador seleccionar las probabilidades para las cuales se encuentra dispuesto a concluir que existen efectos reales (Badii et al., 2007a-h, Rositas et al; 2007).

Es frecuente descubrir los resultados que cabría esperar con una probabilidad del 5% o menor como significativos y aquellos esperados con un 1% o menor como altamente significativos. Cuando un experimentador aplica la frase "los tratamientos son significativamente diferentes", realmente está diciendo que si la hipótesis nula es verdadera, las probabilidades de obtener tales diferencias de medias del tratamiento son sólo de un 5%. Está afirmando que no hubo tal probabilidad de ocurrencia en su experimento y que, por tanto, el resultado significativo se debió a un efecto real del tratamiento.

En este artículo se explicarán las características principales de los diseños experimentales comúnmente utilizados en la investigación de campo, se proporcionará un ejemplo de cada uno y se dará a conocer el procedimiento a seguir en el análisis de los datos. Se utiliza el mismo conjunto de datos para los primeros diseños: el diseño completamente aleatorio y el diseño de bloques completos al azar. Esto muestra las posibles ventajas de un diseño sobre el otro, manteniendo la sencillez de los cálculos, de modo que podamos concentrarnos en lo que se está haciendo y por qué.

A continuación se presentan algunos conceptos básicos relacionados con el diseño experimental. ***Experimento.*** Un procedimiento que basado en el control de las condiciones permite verificar (apoyar, rechazar o modificar) una hipótesis. Un experimento puede ser considerado como una pregunta que detectará nuevos hechos, confirmará los resultados de ensayos anteriores y dará recomendaciones de aplicación práctica (Badii et al., 2007g). ***Unidad experimental.*** La unidad material del experimento al cual se aplica el experimento. ***Control de las condiciones.*** Se trata de controlar aquellas condiciones externos a las unidades experimentales que pueden ocasionar variación o ruido en los resultados del experimento. ***Tratamiento.*** La condición específica del experimento bajo del cual está sujeto la unidad experimental. Es una de las formas que, en cantidad y calidad, el factor a estudiar toma durante el experimento. Por ejemplo, si el factor a estudiar es la cantidad de fósforo, cada una de las dosis de fósforo aplicados durante el experimento es un tratamiento. Los tratamientos a estudiar durante un experimento pueden ser una combinación de varios factores simples: si quiere estudiarse la cantidad de proteína y energía digestible requeridos para la engorda de corderos, se puede considerar tratamientos simples como 14.0% de proteína en al dieta y 2.4 Mcal de energía digestible por kg de alimento. Otro ejemplo sería que el productor de detergentes pude establecer como tratamiento el tipo de agua (dura o suave), la temperatura del agua, la duración de lavado, la marca y tipo de lavadora. En estudios sociológicos y psicológicos, los tratamientos se pueden referir a edad, sexo, grado de educación, religión y otros (Badii et al., 2007g). ***Testigo (control).*** Es un tratamiento que se compara. Si a varios grupos de animales se les administran diferentes dosis de vitaminas, pero no a un grupo testigo, el análisis estadístico dará información acerca del aumento de peso, altura y precocidad de los animales que recibieron los animales comparados con los que no la recibieron. ***Repetición.*** Cuando en un experimento se tiene un conjunto de tratamientos para poder estimar el error experimental, es necesario que dichos tratamientos aparezcan más de una vez en el experimento, para así aumentar la precisión de éste, controlar el error experimental y disminuir la desviación estándar de la media. Por tanto, repetición es el número de veces que un tratamiento aparece en el experimento (Badii et al., 2007g). ***Diseño experimental.*** Es un esquema para realizar un experimento. Los objetivos de un diseño experimental son: (1) verificar si la diferencia entre los tratamientos es una diferencia verdadera o se debe a un proceso al azar, (2) establecer tendencias entre las variables. Es el procedimiento que se sigue para asignar los tratamientos a las unidades experimentales. En un método aleatorio (asignación al azar), se asigna el tratamiento a cada unidad experimental mediante un sorteo o por medio de una tabla de números aleatorios (Badii et al., 2007g). Existen diferentes tipos de diseños experimentales basado en algunas características. En la Tabla 1 se mencionan dos tipos clásicos de diseños experimentales. ***Rasgos universales del diseño experimental***. Existen numerosos diseños experimentales cada uno adecuado para analizar cierto tipo de pregunto. Sin embargo, todos los diseños experimentales comparten los tres siguientes rasgos. 1) La selección aleatoria de las unidades experimentales. Esto evita el sesgo del muestreo. 2) El número de las repeticiones. Esto permite la cuantificación del error experimental. 3) El control local de las condiciones. Esto ayuda a la reducción del error experimental. Cabe mencionar que se puede reducir el nivel del error experimental, a parte del control local de las condiciones o variables, por medio del aumento del tamaño de la muestra y también por el apoyo del modelo de Análisis de Covarianza.

Decisión	**Hipótesis**	
	Cierto	**Falso**
Rechazar	Error tipo I (α)	No error (Poder Estadístico)
Aceptar	No error (Intervalo de Confianza)	Error tipo II (β)

Principales diseños experimentales comúnmente utilizados son: Diseños factoriales, diseño completamente aleatorio, diseños de bloques completos e incompletos y diseño de parcelas y bloques divididos, y a parte los de discriminante, cluster y serie de tiempo (Tabla 1a).

Tabla 1a. Diseños experimentales comunes con algunas de sus características relevantes.

Nombre	**Rasgos**	**Ventajas**	**Eficiencia**
Diseño Completamente al azar (Badii et al., 2007b)	Con 0 gradiente de variabilidad	1. fácil de diseñar 2. fácil de analizar 3. diferentes # de repeticiones 4. máximo g.l. para el error	100%
Diseño de Bloques al azar (Badii et al., 2007a)	Con 1 gradiente de variabilidad	1. reduce la varianza de error 2. fácil de analizar 3. más flexibilidad 4. más precisión	167%
Diseño de Cuadro Latino (Badii et al., 2007a)	Con 2 gradientes de variabilidad	1. reduce la varianza de error 2. fácil de analizar 3. más flexibilidad 4. más precisión	222%

<table>
<tr><td>(a) Diseño factorial, con asignación al azar y (b) Parcelas divididas, sin asignación al azar de la unidad experimental a la unidad de muestra. (Badii et al., 2007a)</td><td>Más de 1 factor</td><td>1. más económico
2. permite medir las interacciones</td><td>288%</td></tr>
<tr><td colspan="4"><u>Diseños multivariados:</u> Cunado se utiliza un gran número de variables</td></tr>
<tr><td colspan="2">1. Componentes principales</td><td colspan="2">* Provee ordenación y el perfil jerárquica</td></tr>
<tr><td colspan="2">2. Análisis Factor</td><td colspan="2">* Reducir el número de las variables para el análisis</td></tr>
<tr><td colspan="2">3. Análisis Discriminante (Badii et al., 2007d)</td><td colspan="2">1. Agrupar en base a la diferencia
2. Más riguroso con los supuestos de la normalidad</td></tr>
<tr><td colspan="2">4. Análisis Cluster (Badii et al., 2007c)</td><td colspan="2">1. Agrupar en base a la similitud
2. Más robusto con los supuestos de la normalidad</td></tr>
<tr><td colspan="2">5a. LISREL (Linear Structured relationship) (Rositas et al., 2007)
5b. EQ (como LISREL, pero más amigable)
5c. LPS Graph & LPS SMART</td><td colspan="2">1. Busca linealizar las interrelaciones entre las variables
2. Intercambia las variables independientes a las dependientes y vice versa</td></tr>
<tr><td colspan="2">6. Correlación Canónica</td><td colspan="2">* Interrelación entre gran número de variables</td></tr>
</table>

Error experimental. Dos unidades experimentales en el mundo natural (seres vivos, conceptos sociales, psicológicos, etc.) nunca son exactamente (100%) lo mismos. Esta diferencia se debe a dos factores: a) elementos genéticos, y b) elementos ambientales. Aún cuando estas unidades sean dos hermanos gemelos siempre surge diferencia debido al factor ambiental. A esta diferencia innata que existe entre las unidades experimentales se le denominan el error experimental o la variabilidad desconocida. ***Tipos de errores.*** Existen dos tipos de errores, a) error tipo I o α que significa el rechazar una hipótesis correcta, y b) error tipo II o β: apoyar una hipótesis falsa. Cuando se efectúa una prueba de hipótesis, puede acontecer uno de los siguientes casos (Badii et al., 2007h).

1. Diseño completamente aleatorio

Este diseño es el más sencillo, eficiente y se origina por la asignación aleatoria de los tratamientos a un conjunto de unidades experimentales previamente determinado. En este diseño usamos k tratamientos, asignándose cada uno al azar a n unidades experimentales; para cada unidad seleccionamos al azar un número de 1 a *k* para decidir que tratamiento debemos aplicar a esa unidad experimental. Si no existen restricciones, con excepción del requerimiento de igual número de unidades experimentales por tratamiento, entonces se dice que el experimento tiene un *diseño completamente aleatorio*. En este caso, todas las unidades experimentales tienen la misma probabilidad de recibir cualquiera de los tratamientos y las unidades experimentales son independientes (Anexo A).

Después que se ha efectuado el experimento, tenemos un grupo de datos consistente en las *kn* respuestas de las unidades experimentales, clasificadas en *k* grupos de acuerdo con los tratamientos que se aplicaron. Suponemos:

1) que los valores observados en cualquiera de los grupos constituyen una muestra aleatoria de todas las posibles respuestas bajo ese tratamiento para todas las unidades experimentales.
2) que la variación entre las unidades tratadas de la misma manera es igual para todos los tratamientos.
3) que las respuestas se distribuyen normalmente.

Ventajas del diseño

1. Permite flexibilidad completa (cualquier número de tratamientos y de repeticiones). Todo el material experimental disponible puede usarse.
2. El análisis estadístico es fácil (aún con diferentes números de repeticiones), o si los errores experimentales difieren de un tratamiento a otro.
3. Método de análisis aun sigue siendo sencillo, cuando existe la perdida relativa de información.
4. El diseño es capaz de estimar el error estándar por unidad experimental (error experimental) con un mayor grado de precisión.

La aleatorización completa puede ser apropiada cuando, **a)** donde el material experimental es homogéneo; b) donde es probable que una parte apreciable de las unidades se destruyan; y **c)** en experimentos pequeños en donde la mayor precisión de otros diseños no compensa la pérdida de grados de libertad del error.
Construcción de la tabla de ANOVA: consideremos el grupo de los siguientes datos:

	Grupo		**Totales**		
	1	2	3 …..	k	
	X_{11}	X_{21}	X_{31}	X_{k1}	
	X_{12}	X_{22}	X_{32}	X_{k2}	
	X_{13}	X_{23}	X_{33}	X_{k3}	
.	.	.	.		
.	.	.	.		
	X_{1n}	X_{2n}	X_{3n}	X_{kn}	
Suma	$X_{1.}$	$X_{2.}$	$X_{3.}$	$X_{k.}$	$X_{..}$
Media	$X_{1.}$	$X_{2.}$	$X_{3.}$	$X_{k.}$	$\overline{X}_{..}$

Primero debemos considerar que fuentes de variación se presentan y cómo se dividirá la variación total de acuerdo a estas fuentes. La variación total en los datos se mide con la suma total de cuadrados de las desviaciones a la media total. Una fuente de variación, las diferencias entre las medias de los grupos, se mide por la suma de cuadrados de la desviación de las medias de los grupos con respecto a la media total. La única variación restante es aquella entre las observaciones dentro de cada grupo – esto es, la variación de los elementos de cada grupo con respecto a la media de este grupo. Esto lo medimos con la suma conjunta de cuadrados de las desviaciones de las observaciones individuales a las medias de los grupos.

Unidades experimentales = Número de tratamiento x número de repetición *n* = *t*x

a) Término de corrección

$$C = \frac{\left(\sum_{i=}^{n} X_{ij}\right)^2}{rn} = \frac{(Grantotal)^2}{rn} \qquad (1)$$

b) Suma de los Cuadrados para tratamiento

$$SC_{Trat} = \frac{\sum (T_i)^2}{r_i} - C \qquad (2)$$

c) Suma de los Cuadrados para total

$$SC_{Total} = \sum_{i=1}^{rt} X_{ij} - C \qquad (3)$$

d) Suma de los Cuadrados para error

$$SC = SC_{Total} - SC_{Trat} \qquad (4)$$

Cálculo del error estándar de la diferencia entre dos tratamientos

a) de igual número de repetición $S_d = \sqrt{s^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right)}$ (5)

b) número de repeticiones diferentes $S_d = \sqrt{\frac{2S^2}{r}}$ (6)

Ejemplo 1. Se probaron 4 métodos de enseñanza en 12 estudiantes. Los estudiantes se asignan aleatoriamente a los métodos de enseñanza (tres estudiantes para cada método). La aleatorización evita la asignación sistemática de los mejores estudiantes a algún método. Supongamos que al final del período de entrenamiento los estudiantes toman una prueba estandarizada, alcanzando los resultados de la siguiente (Tabla 1b).

Tabla 1b. Datos de métodos de enseñanza.

	Grupo				
	1	**2**	**3**	**4**	
	110	111	113	118	
	109	116	108	123	
	105	109	109	125	
Suma	324	336	330	366	1356
Media	108	112	110	122	113

Prueba de la Hipótesis nula: No hay diferencia significativa entre los métodos de enseñanza.
Contra la hipótesis alterna HA: Si hay diferencia significativa entre los métodos de enseñanza.

Como primer paso al construir un análisis de varianza encontramos los totales y las medias, como se muestra en la tabla de datos. Hallamos entonces las sumas de cuadrados de la siguiente manera:

Paso 1. La suma total de cuadrados es

$$\Sigma\,\Sigma\,(X_{ij} - X_{..})^2 = (110 - 113)^2 + (109 - 113)^2 + (105 - 113)^2$$
$$+ (111 - 113)^2 + ... + (125 - 113)^2$$
$$= 428$$

Paso 2. La suma de cuadrados entre grupos es

$$n\,\Sigma\,(X_{i.} - X_{..})^2 = 3[(108 - 113)^2 + (112- 113)^2$$
$$+ (110 - 113)^2 + (122 - 113)^2]$$
$$= 3\ (116)$$
$$= 348$$

Paso 3. La suma de cuadrados dentro de los grupos se obtiene más fácilmente como la diferencia entre la suma total de cuadrados y la suma de cuadrados entre los grupos (Tabla 2)

$$\Sigma\,\Sigma\,(X_{ij} - X_{i.})^2 = 428 - 348 = 80$$

Tabla 2. Análisis de varianza para los datos.

Fuente	**gl**	**gl**	**CM**	**Fcal**	$F_{tab\ (0.05)}$
Entre grupos	3	348	116	11.6	7,591
Dentro de los grupos	8	80	10		
Totales	11	428			

Fcal > F tabulada entonces la Ho se rechaza, y por tanto, hay diferencia significativa entre los métodos de enseñanza.

Ejemplo 2. Se aplicaron tres dosis de una fertilizantes en dos épocas de primavera (P) y de otoño (O). Los resultados sobre el rendimiento de cultivo de una especie vegetal se presentan en la Tabla 3. Prueba de la hipótesis nula en la cual no existen efecto de dosis y época sobre el rendimiento de cultivo.

Tabla 3. Rendimiento de una especie vegetal basado en fertilizante y época de aplicación.

P1	O1	P2	O2	P3	O3	Testigo	
9	30	16		10	17	12	30
9	7	10	18	4	7	10	18
16	21	18	24	4	16	24	32
4	9	18	12	4	17	29	26
			19				
$\sum = 38$	**67**	**62**	**73**	**22**	**57**	**75**	**106**
$\overline{x}$ = **9.5**	**16.8**	**15.5**	**18.2**	**5.8**	**14.2**	**18.75**	**26.5**

Tabla de ANOVA				
Fuente de Var.	S.C.	gl	CM	F_c
Tratamiento	972.3	6	162	3.61*
Error	1122.9	25	44.9	
Total	209.52	31	0	

2. Cálculo de datos perdidos

La simplicidad y la flexibilidad del diseño no se ven afectadas por algunos datos perdidos. Supongamos que en un ensayo con el fertilizante de potasio nos encontramos dos grupos (A y *B*) de datos perdidos (Tabla 4).

Tabla 4. Datos perdidos en el ensayo de fertilizantes

Cálculos de datos perdidos				
K20/ha	**R1**	**R2**	**R3**	**Totales Tx**
40.4	0.00	8	7.93	**15.93 T**
60.5	8.14	8.15	7.87	24.16
80.6	7.76	B	7.74	**15.50 T**
120.8	7.17	7.57	7.8	22.54
161	7.46	7.68	7.21	22.55
	30.58 B	**31.4 B**	**38.55**	**100.48**

El procedimiento de siguientes cálculos el diseño nos permite el cálculo de datos perdidos, para su reemplazo. El procedimiento es como sigue:

$$\Pr omedio = \frac{8.00 + 7.93}{2} = 7.96$$

$$Yi = \frac{rB + tT - Gtotal}{(r-1)(t-1)}$$

$$b = \frac{3(31.40) + 5(15.50) - (100.48 + 7.96)}{(3-1)(5-1)} = 7.91$$

$$a = \frac{3(30.58) + 5(15.93) - (100.48 + 7.91)}{(3-1)(5-1)} = 7.86 \qquad (7)$$

Donde, Y_i = la fórmula para datos perdidos, *a* = primer dato perdido, *b* = segundo dato perdido.

3. Diseño de bloques al azar

En muchos problemas de investigación es necesario diseñar experimentos en los que pueda controlarse sistemáticamente la variabilidad producida por diferentes fuentes extrañas. Por ejemplo supongamos que se desea determinar si cuatro diferentes puntas producen una diferencia en las lecturas de un equipo para medir la dureza. La máquina funciona presionando la punta sobre una probeta de metal y determinando la dureza de la probeta a partir de la profundidad de la marca que se produce. Se ha decidido obtener cuatro observaciones para cada punta. Solo existe un factor–tipo de punta y el diseño de un factor completamente aleatorizado consiste en asignar aleatoriamente cada uno de los 4 X 4 = 16 ensayos a una unidad experimental, ó sea una probeta de metal y tomar las lecturas de la dureza correspondiente. Por lo tanto se requerirían de 16 probetas de metal para realizar este experimento, una para cada ensayo (Anexo A).

En principio existe un problema serio con el diseño completamente aleatorizado en esta situación. Si las probetas son ligeramente distintas en cuanto dureza, como seria el caso si provinieran de diferentes vaciados, las unidades experimentales (probetas o especimenes) contribuyen a la variabilidad observada en la lectura de dureza. Como resultado, el error experimental refleja tanto el error aleatorio como la variabilidad entre probetas.

Se desea que el error experimental sea lo más pequeño posible; en otras palabras se busca sustraer del error experimental la variabilidad producida por las probetas. Un diseño que logre esto requiere que se pruebe cada punta, una vez, en cada una de las cuatro probetas diferentes. Este diseño se conoce como diseño aleatorio por bloques completos.

La palabra completo indica que todos los tratamientos (puntas) son probadas en cada bloque (probetas). Si se usa este diseño, los bloques o probetas forman una unidad experimental más homogénea con la cual comparan las puntas. Esta estrategia de diseño mejora efectivamente la precisión en la comparación al eliminar la variabilidad entre probetas. El orden en que las cuatro puntas deben de ser probadas en cada bloque se determina aleatoriamente.

3.1 Objetivo y característica del diseño

Mantener la variabilidad entre unidades experimentales dentro de un bloque tan pequeño como sea posible y maximizar las diferencias entre bloques. Si no hay diferencia entre los bloques, este diseño no contribuirá a la precisión para detectar las diferencias de tratamientos.

Cada tratamiento es asignado el mismo número de veces a unidades experimentales dentro de un bloque, usualmente una vez (de veces más). Por regla general, es más eficiente tener una sola repetición de cada tratamiento por bloque. A fin de minimizar el error experimental, deben tomarse todas las precauciones para tratar las unidades experimentales dentro de un bloque lo más uniformemente posible. Los bloques pueden estar constituidos por áreas compactas de un campo, grupos de animales que pueden manipularse de un modo uniforme, o diferentes tiempos de aplicación de tratamientos a unidades experimentales.

3.2 Muestreo Aleatorio

Después de que las unidades experimentales han sido agrupadas en los bloques deseados, los tratamientos se asignan aleatoriamente a las unidades dentro de cada bloque, con una distribución aleatoria hecha para cada bloque; por ejemplo, los cuatro tratamientos *A*, *B*, *C* y *D* son asignados de la manera arbitrariamente (Tabla 5).

Tabla 5. Cuatro tratamientos repetidos igual número de veces en un diseño de bloques completamente al azar.

I	II	III	IV
D	A	C	C
A	D	D	B
B	C	B	D
C	B	A	A

Ejemplo 3. El experimento buscaba determinar el efecto de la implantación de una hormona, sobre la capacidad productiva de insectos (machos y hembras) en diferentes sitios (Tabla 6).

Tabla 6. Producción de insectos agrupados por tratamiento y por bloque.

		Bloque					
Tratamiento		**I**	**II**	**III**	**IV**	**Total (T_t)**	**Media X_i/Tratam**
A	**HH0**	47	52	62	51	212	53
B	**MH0**	50	54	67	57	228	57
C	**HH1**	57	53	69	57	236	59
D	**MH1**	54	65	74	59	252	63
Total/B (T_b)		208	224	272	224	928 = $\sum^{ij} X_{ij}$	
Media/B ($\overline{X_b}$)		52	56	68	56		58 = media principapl

Tratamientos = conjunto factorial, niveles de hormonas y el sexto.

Bloques = Cuatro sitios de ensayo

Repeticiones = uno/bloque

- Término de corrección:

$$C = \frac{\left(\sum X\right)^2}{rn} \quad \text{donde } C = (928)^2 / 4x4 = 53824$$

- Suma de cuadrados y cuadrados medios:

S.C. Bloques: $$SCB = \sum \frac{\sum T_j^{21}}{n} - C \qquad (8)$$

$$SCB = \frac{208^2 + ... + 224^2}{4} - 53824 = 54400 - 53824 = 576.0$$

- Suma de Cuadrados para los Tratamientos:

$$SCTrat = \frac{\sum T_1^2}{r} - C$$

$$SCTrat = \frac{212^2 + ... + 224^2}{4} - C = 54032 - 53824 = 208.0$$

- Suma Cuadrada Total:

$$SCTotal = \sum X_{IJ}^2 - C$$

$$SCTotal = (47)^2 + (52)^2 + ... + (59)^2 - C = 54678 - 53824 = 854.0$$

- Suma Cuadrada debida al Error:

$$SCE = SCTotal - SCTrat - SCB$$
$$SCE = 854 - 208 - 576 = 70 \qquad (9)$$

Puesto que cada tratamiento ocurre el mismo número de veces en cada bloque, las diferencias significativas (p<0.005) entre bloques no se deben a los tratamientos, sino a otras diferencias asociadas con los bloques (Tabla 7).

Tabla 7. ANOVA del diseño de bloques completos aleatorios.

Fuente de Variación	**gl**	**S.C**	**C.M**	F_{cal}	F_{Tab} (5%)
Bloques	3	576	192.0	24.69**	5.08
Tratamientos	3	208	69.3	3.86 N.S	
Error	9	70	7.78		
Total	15	854			

4. Diseños de parcelas divididas

El diseño de parcelas divididas es un tipo especial de diseño de bloques completos que suele utilizarse en experimentos factoriales. Se aplican niveles de uno ó más factores a las parcelas completas que se dividen en sub-parcelas, a las cuales se aplican los niveles de uno ó más factores adicionales. El diseño básico involucra la asignación de tratamientos de un factor a parcelas principales en un diseño completamente aleatorio, de bloque completos al azar o de cuadro latino. Los tratamientos de segundo factor se asignan a sub-parcelas dentro de cada parcela principal.

4.1 Aplicaciones

1. Puede usarse cuando los tratamientos relacionados con los niveles de uno o más factores necesitan mayores cantidades de material experimental en una unidad experimental que los tratamientos de otros factores.
2. El diseño puede usarse si va a incorporarse en un experimento un factor adicional para aumentar su alcance.
3. Se puede determinar las diferencias mayores entre los niveles de ciertos factores que entre los niveles de otros. En este caso, las combinaciones de los tratamientos para los factores pueden asignarse aleatoriamente a las unidades completas.
4. El diseño se usa cuando se desea mayor precisión para comparaciones entre ciertos factores, que para otras.

Modelo lineal:

$$y_{ijk} = \mu + \beta_i + \tau_j + n_{ij} + \delta_k + (\tau\delta)_{jk} + e_{ijk} \qquad (10)$$

para $i = 1,2,...,r$, $j = 1,2,...,p$, $k = 1,2,...,q$

donde: μ es un efecto general, β_i el efecto del bloque completo i, τ_j el efecto del tratamiento j sobre la parcela grande (ij), n_{ij} el elemento aleatorio de error sobre la parcela grande (ij), δ_k el efecto del subtratamiento k dentro de la parcela grande (ij), $(\tau\delta)_{jk}$ la interacción entre el tratamiento j y el subtratamiento k, e_{ijk} el error sobre la parcela chica (ijk), y y_{ijk} el valor de la característica en estudio.

4.2 Arreglo en unidades divididas

Distribución en bloques al azar

Supongamos que el factor *A* es la especie o variante *A*, donde i = 1......a (unidad grande), y el factor *B* es la densidad poblacional o variante *B*, donde j = 1... b (subunidad), y el número de repetición sea igual a k = 1... n, *entonces* el arreglo general será como indica la Tabla 8.

Tabla 8. Arreglo de unidades experimentales en bloques al azar.

		Bloques		
Especie	**Densidad**	**I**	**II**	X_{ij}
A_1	b_1 b_2	X_{111} X_{121}	X_{112} X_{122}	$X_{11.}$ $X_{12.}$
	$X_{i.k}$	$X_{1.1}$	$X_{1.2}$	$X_{1..}$
A_2	b_1 b_2	X_{211} X_{221}	X_{212} X_{222}	$X_{21.}$ $X_{22.}$
	$X_{i.k}$	$X_{2.1}$	$X_{2.2}$	$X_{2..}$
Σ/bloque	$X_{..k}$	$X_{..1}$	$X_{..2}$	$X_{...}$

La distribución de las sumas por niveles de factores en estudio será de la siguiente forma.

	Densidad		
Especie	**b_1**	**b_2**	**$X_{i..}$**
A_1	$X_{11.}$	$X_{12.}$	$X_{1..}$
A_2	X_{21}	$X_{22.}$	$X_{2..}$
Σ/densidad	**$X_{.1.}$**	**$X_{.2.}$**	**$X_{...}$**

4.3 Procedimiento de cálculos

Cálculo de las medias: Las medias tanto de los factores como su interacción se calcula de la siguiente manera:

a) Valor promedio para el factor *A* $$\overline{X}_A = \frac{\sum X_{i..}}{bn} \qquad (11)$$

b) Valor promedio para el factor *B* $$\overline{X}_j = \frac{\sum X_{j..}}{an} \qquad (12)$$

c) Valor promedio para la interacción *A* x *B* $$\overline{X}_{ij} = \frac{\sum X_{ij}}{n} \qquad (13)$$

Cálculo de grados de libertad:

Fuentes de variación	G.L.
Bloques	n – 1
Factor (A)	a – 1
Error A	(a - 1)(n - 1)
Unidad grande	an – 1
Factor (B)	b – 1
Int. A x B	(a - 1)(b - 1)
Error B	a(b - 1)(n - 1)
Total	abn – 1

Cálculo de las sumas cuadrados

a) Factor de corrección $$FC = \frac{X_{...}^2}{abn}$$

b) Suma de los cuadrados de unidades grande o de parcelas grande

$$SC_{unidad\ grande} = \frac{\sum X_{i.k}^2}{b} - FC = \frac{X_{1.1}^2 + \cdots X_{2.2}^2}{2} - FC \qquad (14)$$

c) Suma de los cuadrados de bloques (efecto del ambiente) y/o repeticiones:

$$SC_{bloques} = \frac{\sum X_{..k}^2}{ab} - FC$$

d) Suma de los cuadrados para el factor *A:*

$$SC_A = \frac{\sum X_{i..}^2}{bn} - FC \qquad (15)$$

e) Suma de los cuadrados del error *A:*

$$SC_{errorA} = SC_{und\,\mathrm{gra}\,nde} - \left(SC_{bloques} + SC_{.A}\right) \qquad (16)$$

f) Suma cuadrada total

$$SC_{total} = \sum X_{ijk}^2 - FC$$

g) Suma de los cuadrados para el factor *B*

$$SC_B = \frac{\sum X_{.j.}^2}{an} - FC \qquad (17)$$

h) Suma de los cuadrados de interacción

$$SC_{AB} = \left(\frac{\sum X_{ij.}^2}{an} - FC\right) - \left(SC_A + SC_B\right) \qquad (18)$$

i) Suma de los cuadrados del error *B*

$$SC_{error} = SC_{total} - \left(SC_{und\,.\mathrm{gra}\,nde} + SC_{.B} + SC_{.AB}\right)$$

La suma de los cuadrados de los errores *A* y *B* se consideran simplemente como la varianza del A (S_a^2) y del *B* (S_b^2) respectivamente y se demuestra como:

$$S_{errorA}^2 = S_a^2$$
$$S_{errorB}^2 = S_b^2$$

Cálculo de valores de diferencia entre las medias

a) Diferencia entre medias para niveles del factor *A*: a_2 - a_1

Valor Promedio

$$\overline{X} = \frac{X_{i..}}{bn}$$

Error estándar Estimado

$$S_{\overline{X}} = \sqrt{s_a^2 / bn} \qquad (19)$$

Error estándar de la diferencia entre dos medias

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{2s_a^2 / bn} \qquad (20)$$

b) Diferencia entre medias para niveles del factor *B*: b_2 - b_1

Valor Promedio

$$\overline{X}_j = \frac{X_{..j.}}{an}$$

Error estándar Estimado

$$S_{\overline{X}} = \sqrt{\frac{s_b^2}{an}} \qquad (21)$$

Error estándar de la diferencia entre dos medias

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{2s_b^2}{an}} \qquad (22)$$

c) Diferencia entre medias de *B* al mismo nivel de *A*: a_1b_2 - a_1b_1

Valor Promedio

$$\overline{X}_{ij} = \frac{X_{ij}}{n}$$

Error estándar Estimado

$$S_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{2s_b^2}{n}} \qquad (23)$$

Error estándar de la diferencia entre dos medias

$$S_{\overline{X}} = \sqrt{\frac{s_b^2}{n}} \qquad (24)$$

d) Diferencia entre medias de *A* al mismo nivel de *B* o a distintos niveles de *B*.

Para a_2b_1 - a_1b_1 y a_2b_2 - a_1b_1:
Valor Promedio

$$\overline{X}_{ij} = \frac{X_{ij}}{n}$$

Error estándar de la diferencia entre dos medias

$$S_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{2\left[(b-1)s_b^2 + s_a^2\right]}{bn}} \qquad (25)$$

Para que la diferencia entre 2 medias sea significativa, es necesario que se presente la siguiente desigualdad:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{\overline{X}_1-\overline{X}_2}} > t = \frac{(b-1)S_b^2 t_{0.05(G.Lb.)} + S_b^2 + (G.L.a)}{(b-1)S_b^2 + S_a^2} \qquad (26)$$

$$t = \frac{diferenciademedias}{E.E.deladiferencia} > t$$

Ejemplo 4. Producción de maíz, 2 niveles de nitrógeno y 4 abonos (Tabla 9).

Tabla 9. Arreglo de los datos de una parcela dividida.

Bloque	Nivel del nitrógeno	Abono	Producción
I	0	Barbecho (B)	13.8
	0	Cebada (C)	15.5
	0	Soya (S)	21.0
	0	Cebada-Soya (CS)	18.9
	120	Barbecho (B)	19.3
	120	Cebada (C)	22.2
	120	Soya (S)	25.3
	120	Cebada-Soya (CS)	25.9
II	0	Barbecho (B)	13.5
	0	Cebada (C)	15.0
	0	Soya (S)	22.7
	0	Cebada-Soya (CS)	18.3
	120	Barbecho (B)	18.0
	120	Cebada (C)	24.2
	120	Soya (S)	24.8
	120	Cebada-Soya (CS)	26.7
III	0	Barbecho (B)	13.2
	0	Cebada (C)	15.2
	0	Soya (S)	22.3
	0	Cebada-Soya (CS)	19.6
	120	Barbecho (B)	20.5
	120	Cebada (C)	25.4
	120	Soya (S)	28.9
	120	Cebada-Soya (CS)	27.6

Prueba de hipótesis nula en las cuales no hay efecto de fertilizantes y abonos obre el rendimiento de maíz (Tablas 10, 11 y 12).

Procedimiento de cálculo

- Suma de los cuadrados de bloques:

$$SC_{bloques} = \frac{\sum X_{..k}^2}{ab} - FC$$

$$\frac{(161.9)^2 + (163.2)^2 + (172.2)^2}{2(4)} - \frac{(497.3)^2}{24} = 7.87$$

Tabla 10. Producción de maíz ton/ acre, organizados por tratamientos y bloques.

Tratamientos		Bloques			Total
N	**Abono**	**I**	**II**	**III**	
0	**B**	13.8	13.5	13.2	40.5
	C	15.5	15.0	15.2	45.7
	S	21.0	22.7	22.3	66.0
	C-S	18.9	18.3	19.6	56.8
Total para parcela Principal N0		**69.2**	**69.5**	**70.3**	**209.0**
120	**B**	19.3	18.0	20.5	57.8
	C	22.2	24.2	25.4	71.8
	S	25.3	24.8	28.4	78.5
	C-S	25.9	26.7	27.6	80.2
Total para parcela Princ. N120		**92.7**	**93.7**	**101.9**	**288.3**
Total por bloque		**161.9**	**163.2**	**172.2**	**497.3**

Fertilizante N	Abonos				Total
Abono	**B**	**C**	**S**	**C-S**	**Factor A**
0	40.5	45.7	66.0	56.8	209.0
120	57.8	71.8	78.5	80.2	288.3
Total Factor B	98.3	117.5	144.5	137.0	497.3

- Suma de los cuadrados de nitrógeno:

$$SC_{Nitrogeno} = \frac{(209.0)^2 + (288.3)^2}{3(4)} - \frac{(497.3)^2}{24} = 262.02$$

- Suma de los cuadrados de parcelas de nitrógeno:

$$SCPG = \frac{(69.2)^2 + ...(101.9)^2}{4} - \frac{(497.3)^2}{24} = 274.92$$

- $SC_{Error}\ A = SCPG - SCB - SCN = 274.92 - 7.87 - 262.02 = 5.03$

- SC_{Abono} =

$$\frac{(98.3)^2 + (117.5)^2 + (144.5)^2 + (137.0)^2}{3(2)} - \frac{(497.3)^2}{24} = 215.26$$

- *SCN x Abono*= *(40.5)*^{2} + ….+ (80.2)^2 – (497.3)^2/24 = 18.7
- *SCPG x Sub.P Abono* = (13.8)^2 + (15.5)^2 + …. + (27.6)^2 – C = 516.12
- *SC*$_{Error}$*B* = 516.12 -274.92 -215.26 – 18.7 =7.24

Tabla 11. Resultado del análisis de varianza para el experimento del rendimiento del maíz.

Fuente Var.	**S.C.**	**gl**	**C.M.**	**F$_c$**	**F$_{tabulada}$**
Parcela grande	274.92	5			
Bloques	7.87	2	3.935	1.56 NS	19.00
Factor A (nitrógeno)	262.02	1	262.02	104.18 **	18.51
Error A	5.03	2	**2.515**		
Factor B (abono)	215.26	3	71.753	118.99 **	3.49
Interacción N x Abono	18.7	3	6.233	10.34 **	3.40
Error Factor B	7.24	12	0.603		
Total	516.12	23			

Tabla 12. Valores promedios para los niveles de factores de nitrógeno y abono.

Fertilizante N	**B**	**C**	**S**	**C-S**	$\overline{X}$ **/ N**
0	13.5	15.2	22.0	18.9	17.4
120	19.3	23.9	26.2	26.7	24.0
$\overline{X}$ /Abono	16.39	19.58	24.08	22.83	20.7

Cálculo de valores de las diferencias entre las medias

1. Diferencia entre medias para niveles de *Factor A*: $a_2 - a_1$

- *Cálculo del error estándar estimada de la diferencia entre las medias*

$$S_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{2S_a^2}{bn}} = \sqrt{\frac{2(5.03)}{12}} = 0.1956$$

- *Cálculo de la diferencia de medias*

$$a_2 - a_1 = 24.0 - 17.4 = 6.6$$

- *Prueba de significancia* (*t Student*)

$$t_c = \frac{24.0 - 17.4}{0.1956} = 7.21^{**}$$

2. Diferencia entre medias para niveles del factor B: b_2-b_1

Prueba de Tukey $\mu_1 = \mu_2 = \mu_3 = \mu_4$ (Tabla 9)

- *Cálculo del error estándar estimada de la diferencia entre las medias*

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{S_b^2}{an}} = \sqrt{\frac{(7.24)}{6}} = 1.1$$

Tabla 13. Tabla de comparaciones múltiples de medias para el factor de abono.

Comparación	**Diferencia** $\overline{X}_B - \overline{X}_A$	$S_{\overline{X}}$	q_c	**Conclusión**
S vs CS	1.25	1.1	1.138	H_0
S vs C	4.49	1.1	4.094	H_0
S vs B	7.69	1.1	7.001	H_A
CS vs C	3.24	1.1	2.956	H_0
CS vs B	6.44	1.1	5.863	H_A
C vs B	3.19	1.1	2.907	H_0

El valor q tabulada es 4.199 ($q_{0.05,12,4}$ = 4.199)

3. Diferencia entre medias del factor *B* al mismo nivel del factor A: a_1b_2 –a_1b_1

- *Cálculo del error estándar estimada de la diferencia entre las medias*

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{2S_b^2}{n}} = \sqrt{\frac{2(7.24)}{3}} = 2.196$$

- *Cálculo de la diferencia de medias*

$a_1b_2 - a_1b_1 = 22.0 - 13.5 = 8.5$
$a_2b_2 - a_2b_1 = 26.2 - 19.3 = 6.9$

- *Prueba de significancia* (*t Student*)

$$t_c = \frac{22.0 - 13.5}{2.196} = 3.87^{**}$$

$$t_c = \frac{26.2 - 19.3}{2.196} = 3.14NS$$

4. Diferencia entre medias del factor *A* al mismo nivel del factor *B* o a distintos niveles del factor *B*: $a_2b_1 - a_1b_1$; $a_2b_2 - a_1b_1$

- *Cálculo del error estándar estimada de la diferencia entre las medias*

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{2((b-1)S_b^2 + S_a^2\ )}{bn}} = \sqrt{\frac{2(3x(5.03 + 7.24))}{12}} = 2.478$$

- *Cálculo de la diferencia de medias*

$a_2b_2 - a_1b_1 = 23.9 - 13.5 = 10.4$
$a_2b_3 - a_1b_1 = 26.2 - 13.5 = 12.7$
$a_2b_3 - a_1b_1 = 26.7 - 13.5 = 13.2$

- *Prueba de significancia* (*t Student*)

$$t_c = \frac{23.9 - 13.5}{2.4789} = 4.195^{**}$$

$$t_c = \frac{26.2 - 13.5}{2.4789} = 5.123^{**}$$

$$t_c = \frac{26.7 - 13.5}{2.4789} = 5.324^{**}$$

5. Diseño de cuadro latino

En este diseño, la distribución aleatoria de los tratamientos se restringe más ampliamente mediante la agrupación de los mismos, tanto en columnas como en hileras (bloques). Así resulta posible eliminar la variabilidad del error experimental asociada con ambos efectos. Cada tratamiento ocurre el mismo número de veces (usualmente una vez) en cada hilera y columna y proporcionará una comparación más precisa de los efectos del tratamiento (Anexo A).

Un cuadro Latino requiere al menos tantas repeticiones como tratamientos existan; por tanto no resulta práctico para experimento con un gran número de tratamientos, con una sola unidad experimental por tratamiento en cada columna e hilera. Este diseño se ha usado con ventaja en muchos campos de investigación donde hay dos fuentes principales de variación en la realización de un experimento.

Los cuadros latinos más comunes van de 5x5 a 8x8; cuadros mayores de 12x12 se usan muy rara vez. En los cuadros latinos, como en los bloques al azar, a medida que aumenta el tamaño del bloque, el error experimental por unidad probablemente aumente. Los cuadros latinos pequeños proporcionan pocos grados de libertad para estimar el error experimental y así debe lograrse una disminución sustancial en el error para compensar el corto número de grados de libertad.

La aleatorización en el cuadro latino consiste en elegir un cuadro latino al azar entre todos los cuadros latinos posibles. Fisher y Yates dan el conjunto completo de cuadros latinos desde 4x4 hasta 6x6 y muestran cuadros hasta de tamaño 12x12. Cochran & Cox dan cuadros latinos de muestra desde 3x3 hasta 12x12.

- Los cuadros 2 x 2 no proporciona ningún grado de libertad del error (2-1)(2-2) = ∅.

- Los cuadros 3 x 3 sólo proporciona 2 grado de libertad del error (3-1)(3-2) = 2.

- Los cuadros 4 x 4 dan 6 grado de libertad del error (4-1)(4-2) = 6.

Este hecho elimina el uso de cuadros 2 x 2 como cuadro latino. A los cuadros 2 x 2 se los aplicará la prueba de χ^2 bajo tablas de contingencia.

- Tres cuadros 3 x 3 (9 repeticiones) proporcionará 10 grado de libertad del error, mientras que dos cuadros 4 x 4 tendrán 15 grado de libertad del error.

5.1 Procedimiento estadístico

- Término de corrección C = $\sum (X_{ij})^2 / r^2$

El procedimiento del análisis de varianza (suma de cuadrados) para un cuadro latino r x r se presenta en la Tabla 14.

Tabla 14. Cálculo de las sumas de los cuadrados y grados de libertad del diseño de Cuadro Latino.

Fuente de variación	gl	Suma de los cuadrados
Hileras	r - 1	$\sum \frac{X_{i.}^2}{r} - C$
Columnas	r - 1	$\sum \frac{X_{j..}^2}{r} - C$
Tratamientos	r – 1	$\sum (X_{ij}^2) / r - C$
Error	(r – 1) (r - 2)	S.C.T. – SCH – SCC - SC.Trat
Total	r^2 - 1	$\sum X_{ij}^2 - C$

Ejemplo 5. Un hacendado desea ensayar los efectos de cuatro fertilizantes *A*, *B*, *C*, y *D* en el rendimiento de trigo. Para eliminar fuentes de error debidas a la variabilidad en la fertilidad del suelo emplea los fertilizantes en una distribución de un cuadrado latino, donde los números indican rendimientos en dkl por unidad de área. Efectuar un análisis de varianza para determinar si hay una diferencia significativa entre los fertilizantes a niveles de significación de 0.05 y 0.01.

A 18	C 21	D 25	B 11
D 22	B 12	A 15	C 19
B 15	A 20	C 23	D 24
C 22	D 21	B 10	A 17

Primero obtenemos totales para filas y columnas.

				Totales
A 18	C 21	D 25	B 11	**75**
D 22	B 12	A 15	C 19	**68**
B 15	A 20	C 23	D 24	**82**
C 22	D 21	B 10	A 17	**70**
77	**74**	**73**	**71**	**295**

Obtenemos el total de rendimientos para cada uno de los fertilizantes.

A	**B**	**C**	**D**	**Total**
70	48	85	92	295

La variación total y las variaciones para filas, columnas y tratamientos se obtienen de forma común:

Término de corrección: C =

$$(295)^2/4 \times 4 = 5,439.06$$

Variación Total (SCT) =

$$[(18)^2+\ldots+(17)^2–(295)^2]- 5,439.06 = 5769–5439.06 = 329.94$$

Variación entre filas =

$$[(75)^2+ \ldots +(70)^2] /4- 5,439.06 = 5,468.25– 5,439.06 = 29.19$$

Variación entre columnas =

$$[(77)^2 +\ldots + (71)^2]/4 - 5,439.06=5,443.75 – 5,439.06 = 4.69$$

Variación entre tratamientos =

$$[(70)^2+\ldots+ (92)^2] /4– 5,439.06 = 5,723.25 – 5,439.06 = 284.19$$

Variación residual =

$$329.94 – (29.19 + 4.69 + 284.19) = 11.87$$

Tabla 15. Análisis de varianza sobre el rendimiento de trigo en función de aplicación de diferentes fertilizantes.

Fuente de Variación	SC	gl	CM	$F_{cal.}$
Filas	29.19	3	9.73	4.92
Columnas	4.69	3	1.563	0.79 NS
Tratamientos	284.19	3	94.73	47.9
Residuales	11.87	6	1.978	
Residuales	11.87	6	1.978	
Total	329.94	15		

Puesto que $F_{0.95,\ 3,\ 6} = 4.76$, rechazamos al nivel 0.05 la hipótesis de que hay medias de fila iguales. Se deduce que al nivel 0.05 hay diferencia en la fertilidad del suelo de una fila a otra. Puesto que el valor F para columnas es menor que 1, concluimos que no hay diferencia en la fertilidad del suelo en las columnas. El valor de F para tratamientos es 47.9 superior al valor tabulado 4.76, podemos concluir que hay diferencia altamente significativa entre fertilizantes. Con $F_{0.99,\ 3,6} = 9.78$, podemos aceptar la hipótesis que no hay diferencia en la fertilidad del suelo en las filas (o las columnas) al nivel de significación del 0.01. Sin embargo, debemos concluir que hay una diferencia entre fertilizantes al nivel 0.01.

6. Experimentos factoriales

El diseño factorial es aquel en el que el conjunto de tratamientos consiste en todas las combinaciones posibles de los niveles de varios factores. El factor, es una clase de tratamiento, y en experimentos factoriales, todo factor proporcionara varios tratamientos. Nivel, se refiere a los diferentes tratamientos dentro de un factor. El número de factores y niveles que pueden compararse en un solo experimento solo se limita por consideraciones prácticas. En general un experimento factorial permite la separación y la evaluación de los efectos de cada uno de 2 o más factores que afectan solo a una unidad experimental, además permite la detección de los efectos de interacción entre 2 o más factores.

En experimentos factorial es necesario considerar el arreglo y la distribución. Arreglos más utilizados: combinatorio, en parcelas divididas y en franjas.

Distribuciones más utilizadas. Completamente al azar, bloques al azar y cuadro latino.

Notación. Los sistemas de notación que se usan generalmente son similares, las letras mayúsculas se utilizan para designar factores, ejemplos *A*, *B*, *C*, etc. Las combinaciones de letras minúsculas y subíndices numéricos, o simplemente subíndices, se usan para denotar combinaciones de tratamientos y medias: por ejemplo a_1b_3 indica la combinación de tratamiento compuesta del primer nivel de *A* y el tercer nivel de *B,* y a la correspondiente media de tratamiento. A menudo el cero se usa para el primer nivel de un subíndice.

Frecuentemente, sucede que el experimentador esté interesado en estudiar varios factores que, actúan simultáneamente. Varias opciones podrían plantearse para realizar la investigación. Una consiste en ensayar los factores, haciendo variar uno de ellos y manteniendo fijos los demás. Otra será que cada uno de los factores varíe, ensayando varios niveles de cada uno ellos; con esta opción, el investigador puede deducir las posibles relaciones entre los factores, y estimar su efecto principal, puesto que, antes de iniciar el experimento, se desconoce cuáles de estos factores son importantes, y si ejercen su acción independientemente (Anexo A).

6.1 Utilidad de los experimentos factoriales

1. En trabajos de exploración, donde el objeto es determinar rápidamente los efectos de cada uno de cierto número de factores dentro de un intervalo especifico.
2. En investigaciones de las interacciones entre los efectos de varios factores. Por su naturaleza las interacciones no se pueden estudiar sin probar algunas de las combinaciones que se forman de los diferentes factores. Frecuentemente la información se obtiene mejor probando todas las combinaciones.
3. En experimentos diseñados para poder llegar a recomendaciones que deben aplicarse a una gran variedad de condiciones. Se pueden introducir factores auxiliares en un experimento para probar los factores principales bajo una variedad de condiciones similares a las encontradas en la población a la cual se van a aplicar dichas recomendaciones.

Por otro lado, si se acumula una cantidad considerable de información o si el objeto de la investigación es especializado, puede ser más provechoso realizar un trabajo intensivo sobre un factor único o sobre unas cuantas combinaciones de factores. Por ejemplo algunas combinaciones están dirigidas a encontrar la combinación de los niveles de los factores que producirían una máxima respuesta. Cuando los factores que se van investigar son numerosos, la principal desventaja del experimento factorial estriba en su tamaño y complejidad. La magnitud de la tarea puede reducirse teniendo únicamente una sola repetición, hasta es posible obtener casi toda la información deseada probando solamente una fracción del número total de combinaciones de tratamiento, aun cuando esto se haga con ciertos riesgos.

6.2 Experimentos factoriales 2^n

Supongamos que se ensayan dos factores, digamos humedad (*h*), y temperatura (*t*), cada uno en dos niveles, y que observamos su efecto sobre el comportamiento de insectos (Tabla 16). Sean h_o y h_1 los niveles ensayados de humedad, y t_o y t_1 los niveles ensayados de temperatura (4 combinaciones de tratamientos): h_0t_0, h_1t_0, h_0t_1 y h_1t_1 = tratamientos de un experimento factorial 2^2 en *r* bloques=un diseño de *r* bloques completos al azar (con 4 Unidades Experimentales). y_{ijk} = Rendimiento de la unidad experimental, con el nivel h_i de humedad y el nivel t_j de temperatura, en el bloque completo *k* (con *i,j* = 0 o 1, y *k* = 1,2,... *r*).

Tabla 16. Arreglo factorial en bloques completos.

Tratamientos	Bloques completos	Totales de tratamientos
	1 2 ..... r	
H_0t_0	Y_{001} Y_{002} ..... Y_{00r}	T_{00}
h_0t_0	Y_{101} Y_{102} ..... Y_{10r}	T_{10}
h_0t_1	Y_{011} Y_{012} ...... Y_{01r}	T_{01}
h_1t_1	Y_{111} Y_{112} .... Y_{11r}	T_{11}
Totales de Bloques	**B_1 B_2 ..... B_r**	**G**

El análisis estadístico, se basa en el modelo lineal siguiente:

$$y_{ijk} = \mu + \beta_k + \tau_{ij} + e_{ijk}$$
$$i = 0,1, j = 0,1; k = 1,2,....., r \qquad (27)$$

donde:

μ = Efecto general

β_k= Efecto del bloque k

τ_{ij} =Efecto del tratamiento h_it_j

e_{ijk}=Error con las propiedades usuales (media cero, varianza constante y no correlación con otros términos de error)

y_{ijk}=la característica observada

$\sum\lambda_{ij}\tau_{ij}$, =El mejor estimador de una constante entre tratamientos
λ_{ij} = Ciertas constantes reales, tales que $\sum\lambda_{ij}$=0, está dado por

$$\sum_{ij} \lambda_{ij} \tau_{ij} = \sum \lambda_{ij} y_{ij} , \qquad (28)$$

donde : $y_{ij} = T_{ij}/ r$; además:

$$Var\left\{ \sum_{ij} \lambda_{ij} \tau_{ij} \right\} \frac{\sigma^2}{r} \sum_{ij} \lambda_{ij}^2 \qquad (29)$$

Es importante observar que los tres grados de libertad para tratamientos, pueden partirse en grados de libertad individuales, al considerar los siguientes contrastes entre efectos de tratamientos (Tabla 17).

$$C_1 = 1/2(-\tau_{00} + \tau_{10} - \tau_{01} + \tau_{11}) \qquad (30)$$

$$C_2 = 1/2(-T_{00} - T_{10} + T_{01} + T_{11})$$
$$C_3 = 1/2(T_{00} - T_{10} - T_{01} + T_{11})$$

Tabla 17. Análisis de varianza para el diseño Factorial 2^2.

Fuentes de variación	Grados de libertad	Sumas de cuadrados	Cuadrados medios
Bloques	r-1	$\sum_k \frac{B_k^2}{4} - \frac{G^2}{4r}$	CMB
Tratamientos	3	$\sum_{ij} \frac{T_{ij}^2}{r} - \frac{G^2}{4r}$	CMT
Error	3(r-1)	SCE	S^2
Total	4r-1	$\sum_{ijk} y_{ijk}^2 - \frac{G^2}{4r}$	

Para un diseño factorial de 2X2 o 2^2, o sea un experimento con dos factores con dos niveles de cada factor, cualquiera de las siguientes notaciones es adecuada.

Factor	***A***						
	Forma completa			**Formas abreviadas**			
	Nivel	A_1	a_2	A_1	a_2	a_1	a_2
B	$\mathbf{b_1}$	A_1b_1	a_2b_1	(1)	a	00	10
	$\mathbf{b_2}$	A_1b_2	a_2b_2	B	ab	01	11

Los grados de libertad y las sumas de cuadrados para la varianza en un factorial 2^2 pueden particionarse en grados de libertad única e independiente y sus correspondientes sumas de cuadrados. Para un experimento factorial general, se divide los grados de libertad y las sumas de cuadrados en sub-conjuntos o componentes aditivos, no necesariamente con grados de libertad únicos.

6.3 Efectos simples, efectos principales e interacciones

Las 4 diferencias, a_2-a_1, a cada nivel de *B* y b_2-b_1 a cada nivel de *A*, se llaman efectos simples, para los datos en I de la tabla, el efecto simple de *A* al primer nivel de *B* es 2; en el efecto simple de *B* al segundo nivel de A es –6. Los resultados de promediar los efectos simples se llaman efectos principales. Se detectan con letras mayúsculas, como los factores. El efecto principal del factor A para los datos de I en la tabla es 5; el efecto principal del factor *B* para los datos en III es 6. Los efectos principales se calculan por unidades. Para un factorial 2^2, *A* y *B* están dados por las ecuaciones siguientes:

$$A = 1/2\left[(a_2b_2 - a_1b_2) + (a_2b_1 - a_1b_1)\right]$$
$$A = 1/2\left[(a_2b_2 + a_2b_1) - (a_1b_2 - a_1b_1)\right] \qquad (31)$$

$$B = 1/2\left[(a_2b_2 - a_2b_1) + (a_1b_2 - a_1b_1)\right]$$
$$B = 1/2\left[(a_2b_2 + a_1b_2) - (a_2b_1 + a_1b_1)\right] \qquad (32)$$

Los efectos principales son promedios en una variedad de condiciones, las cuales se dan dentro de los bloques, así como también entre los bloques; de este modo el factor A se replica dentro de cada bloque ya que esta presente a ambos niveles para cada nivel del factor *B*. El hecho de promediar implica que las diferencias, esto es, los efectos simples solo varían debido al azar de un nivel a otro del factor o factores. Esta es una hipótesis sometida a una prueba de significancia cuando los tratamientos están dispuestos factorialmente; la hipótesis es la de que no hay interacción entre factores.

Cuando los efectos simples para un factor difieren más de lo que pueda ser atribuible al azar, esta respuesta diferencial se llama interacción de los 2 factores. Esta relación es simétrica; esto es, la interacción de A con *B* es lo mismo que la de *B* con *A*.

La interacción de *A* y *B* se define en:

$$AB = 1/2\left[(a_2b_2 - a_1b_2) - (a_2b_1 - a_1b_1)\right]$$

$$AB = 1/2\left[(a_2b_2 + a_1b_1) - (a_1b_2 + a_2b_1)\right] \qquad (33)$$

La interacción es la mitad de la diferencia entre las sumas de las dos diagonales de la tabla 2X2 que es la mitad de la diferencia entre las sumas de los tratamientos, donde *A* y *B* están presentes a los niveles más altos y más bajos y de lo tratamientos donde uno solo esta presente al nivel más alto. Esto siempre es cierto para el factorial 2X2. La interacción mide el que no logre el efecto *A*, *o* la respuesta a *A*, de ser la misma para cada nivel de B, o al contrario, el no lograr el efecto B de ser el mismo para cada nivel de A (Tabla 18)..

En la respuesta a *A* o el aumento de a_1 a a_2 es mayor para b_2 que para b_1, esto es ha habido variación en la magnitud del incremento. En II (Tabla 18), la respuesta a *A* es un aumento en presencia de b_1 y una disminución en presencia de b_2; ha habido un cambio en la dirección del incremento. En términos de medias de tratamientos presentadas en una tabla de 2 factores, variaciones suficientemente grandes en las magnitudes de las diferencias entre medias en una columna (o fila), al pasar de columna a columna (o de fila a fila), pueden constituir una interacción. Además los cambios de rango de cualquier media de tratamiento de una columna (o fila), al pasar de columna a columna (o filas), puede constituir una interacción.

6.4. Ilustración de la interacción

La presencia o ausencia de efectos principales no nos dice nada respecto a la presencia o ausencia de interacción. La presencia o ausencia de interacción no nos dice nada respecto a la presencia o ausencia de efectos principales, pero nos dice algo acerca de la homogeneidad de los efectos simples.

Tabla 18. Determinación de efectos simples y principales.

Factor	**I *A* = Clase**				
***B*=tasa**	**Nivel**	***A_1***	**a_2**	**Media**	***A_2*-a_1**
	b_1	30	32	31	2
	b_2	36	44	40	8
	Media	33	38	35.5	5
	b_2-b_1	6	12	9	
Factor	**II A= Clase**				
***B*= tasa**	**Nivel**	***A_1***	**a_2**	**Media**	***A_2*-a_1**
	b_1	30	32	31	2
	b_2	36	26	31	-10
	Media	33	29	31	-4
	b_2-b_1	6	-6	0	
	III *A*=Clase				
***B*= tasa**	**Nivel**	***A_1***	**a_2**	**Media**	**A_2-a_1**
	b_1	30	32	31	2
	b_2	36	38	37	2
	Media	33	35	34	2
	b_2-b_1	6	6	6	

Una interacción significante es aquella que es lo suficientemente grande como para que se pueda explicar con base en el azar y la hipótesis nula de que no hay interacción. Cuando la interacción es significante, los factores no son independientes entre si; los efectos simples de un factor difieren y la magnitud de un efecto simple depende de el nivel del otro factor del termino de la interacción. Si la interacción no es significante, se puede concluir que los factores en consideración son independientes entre si; los efectos simples de un factor son los mismos para todos los niveles de los otros factores, dentro de una variación aleatoria medida por el error experimental. El efecto principal, es la mejor estimación de la diferencia común. La ilustración de efectos simples, efectos principales e interacciones se muestra en la Tabla 18.

Para I: *AB* = (8-2) en términos de los efectos simples *A*
= (12-6) en términos de los efectos simples *B*
= 3

Para II: *AB* =
(26-36-32+30) = - 6

Para III: *AB* =
(38-36-32+30) = 0

Cuando los factores son independientes, los efectos simples son iguales a los efectos principales correspondientes a un efecto principal, de tal manera que los efectos principales son las únicas cantidades necesarias para describir completamente las consecuencias de las variaciones de un factor, en un experimento factorial cada efecto principal se estima con la misma precisión que si todo el experimento se hubiese dedicado a ese solo factor.

En un experimento factorial es correcto particiones los grados de libertad de los tratamientos y la suma de cuadrados en los componentes atribuibles a los efectos principales y a interacciones aun cuando la prueba de F total de que no hay diferencia tratamientos no sea significante; los efectos principales y comparaciones de interacción son comparaciones planeadas.

Los resultados de un experimento factorial conducen a una explicación relativamente sencilla debido a la variedad y naturaleza de las comparaciones de tratamientos. Cuando los factores no son independientes, los datos necesitan un estudio detallado con la posibilidad de más experimentación.

Supongamos que ensayan dos factores, cada uno con dos niveles, con un total de cuatro combinaciones de tratamientos, los cuales pueden representarse en la Tabla 19.

Tabla 19. Las combinaciones en un arreglo factorial.

Factor	Niveles
Variedad A	a_1, a_2, $c = 1...a$
Densidad B	b_1, b_2, $j = 1...b$
Repetición n	$K = 1...b$

Si estas combinaciones de tratamiento se ensayan en un diseño de *r* bloques completos al azar, tenderemos 4*r* unidades experimentales, observando el rendimiento, $y_{ijk,}$ sobre la unidad experimental, con el nivel n_i del factor *A* y p_j del factor *B*, en el bloque completo *k* (Tabla 20).

Tabla 20. La "Tabla de doble entrada" para las combinaciones.

	b_1	b_2	b_3
a_1	a_1b_1	a_1b_2	a_1b_3
a_2	a_2b_1	a_2b_2	a_2b_3
a_3	a_3b_1	a_3b_2	a_3b_3
a_4	a_4b_1	a_4b_2	a_4b_3

Se puede tener interés en lo siguiente:
a).- ¿Cuál es el mejor nivel del factor *A*?
b).- ¿Cuál es el mejor nivel del factor *B*?
c).- Estudiar el efecto de interacción "acción conjunta", en este el estudio de ANOVA se tiene que extender.

Bajo una distribución completamente al azar se tendrán las siguientes.

Causas	G.L.
Tratamientos	$ab - 1$
Factor A	$a - 1$
Factor B	$b - 1$
Interacción A x B	$(a - 1)(b - 1)$
Error	$ab(n-1)$
Total	$abn - 1$

La Interacción es la acción conjunta de 2 o más factores, o la modificación del efecto de un factor por la acción del efecto de otro o más factores. Los efectos de factores en estudio pueden ser: *aditivos, multiplicativos e interactivos*. Cuando la *diferencia* entre los niveles de factores es igual a cero, los efectos de los factores son *aditivos* o los factores son *independientes*. En este caso las líneas de diferencias son paralelas, y se concluye que los niveles del factor *A* se comportan de manera similar en los niveles del Factor *B* (por ejemplo aumenta la densidad de especie, aumenta el rendimiento).

Efecto interactivo. Cuando los niveles de los factores en estudio no son independientes y los niveles de un factor (B por ejemplo) esta relacionada o depende del nivel del primer factor (A) como es el caso de Interacción genotipo x ambiente, en este caso se trata de efectos *interactivos* (Tabla 21). En general, cuando se estudian 2 o más factor es recomendable ver si las líneas de "tendencia" sugieren ser: a) paralelas (efectos *aditivos*), b) cruzadas (efectos *interactivos*).

Tabla 21. Tabla *a* x *b* o de doble entrada para estudiar efectos del factor A, del B e interacción AB.

Factor A	**Factor B**			**Total por Factor A ($X_{i.}$)**
	b_1	**B_2**	**B_3**	
a_1	$X_{11.}$	$X_{12.}$	$X_{13.}$	$X_{1..}$
a_2	$X_{21.}$	$X_{22.}$	$X_{23.}$	$X_{2..}$
a_3	$X_{31.}$	$X_{32.}$	$X_{33.}$	$X_{3..}$
a_4	$X_{41.}$	$X_{42.}$	$X_{43.}$	$X_{4..}$
Total/ factor B ($X_{.j.}$)	$X_{.1.}$	$X_{.2.}$	$X_{.3.}$	$X_{...}$

6.5. Cálculo de medias y el error estándar de la media

- Tratamiento: $\overline{X}_{ij.} = \frac{X_{ij.}}{n}$ $\quad S_{\overline{X}} = \sqrt{\frac{s^2}{n}}$
- Factor A: $\overline{X}_{i} = \frac{X_{i..}}{bn}$ $\quad S_{\overline{X}} = \sqrt{\frac{s^2}{bn}}$
- Factor B: $\overline{X}_{j} = \frac{X_{j..}}{an}$ $\quad S_{\overline{X}} = \sqrt{\frac{s^2}{an}}$
- General $\overline{X} = \frac{X\ldots}{abn}$

El procedimiento estadístico consiste de dos etapas de cálculos (Tabla 22):

Cálculo de SC para los factores, SC_{total}, $SC_{Tratamientos}$, y SC_{error}, :

a) Factor de corrección $FC = \frac{\left(\sum_{i=1}^{n} X_{ijl}\right)^2}{N}$

b) La suma cuadrada total $SC_{total} = \sum X_{ijk}^2 - FC$

c) Sumas de los cuadrados de tratamiento $SC_{tratamientos} = \frac{\sum X_{ij.}^2}{n} - FC$

d) La suma cuadrado del error I **------**

6.6. División de SC y el grado de libertad en las causas parciales A, B y AB

Tabla 22. Tabla de análisis de varianza.

FV	**SuSC**	**gl**	**CM**	**F**
Tratamientos	$\frac{\sum X_{ij.}^2}{n_{ij}}$ - FC	ab - 1	SCtrat/ab-1	$\frac{CM_{trat}}{CM_{Error}}$
Factor A	$SC_A = \frac{\sum X_{i..}^2}{bn} - FC$	a - 1	SCFac.A/a-1	$\frac{CM_{Fac.A}}{CM_{Error}}$
Factor B	$SC_B = \frac{\sum X_{.j.}^2}{an} - FC$	b-1	SCFac.B/b-1	$\frac{CM_{Fac.B}}{CM_{Error}}$
Interacción A X B	$SC_{AB} = \left(\frac{\sum X_{ij.}^2}{n} - FC\right) - \left(SC_A + SC_B\right)$	(a-1)(b-1)	SCInt./Gl.	$\frac{CM_{Interacion}}{CM_{Error}}$
Error I	$SC_{error} = SC_{total} - SC_{tratamientos}$	ab(n-1)	SCError/Gl	S_I^2
Error II	$SC_{error} = SC_{total} - \left(SC_{tratamientos} + SC_{bloques}\right)$	ab(n-1) - (n+1)	SCError/Gl	S_{II}^2

Bloques	$SC_{bloques} = \frac{\sum X_{..k}^2}{ab} - FC$	n-1	SCBloque/n-1	$\frac{CM_{Bloque}}{CM_{Error}}$
Total	$SC_{total} = \sum X_{ijk}^2 - FC$	abn-1	——	——

a) Suma de los Cuadrados para el factor *A:*

$$SC_A = \frac{\sum X_{i..}^2}{bn} - FC$$

b) Suma de los Cuadrados para el factor *B*:

$$SC_B = \frac{\sum X_{.j.}^2}{an} - FC$$

c) Suma de los Cuadrados de interacción

$$SC_{AB} = \left(\frac{\sum X_{ij.}^2}{n} - FC \right) - \left(SC_A + SC_B \right)$$

d) Suma de los Cuadrados de bloques $SC_{bloques} = \frac{\sum X_{..k}^2}{ab} - FC$

e) Suma de los Cuadrados del error II

$$SC_{error} = SC_{total} - \left(SC_{tratamientos} + SC_{bloques} \right)$$

Ejemplo 6. En un experimento (Tabla 23) factorial 2^3 donde se determino la longitud (mm) de una especie de insecto en 4 ecosistemas bajo la influencia de tres factores de humedad, temperatura y luminosidad (cada uno con dos niveles).

Factor A = Humedad: h_1= 40 y h_2 = 80% Nivel = 2
Factor B = Temperatura: t_1 = 20 y t_2 = 30 Nivel = 2
Factor C = Luminosidad: l_1 = 100 y l_2 = 200 Nivel = 2
No de Tratamiento = 2 X 2 X 2 = 8 Trat. = 8
No de Ecosistemas = 4
N = Tratamientos X Ecosistemas = 8 X 4 = 32

Tabla 23. Combinación de niveles de tratamientos en 4 bloques (Martínez-Garza, 1988).

Tratamiento	**Factor**			**Bloque**				**Suma**
	H	**T**	**L**	**I**	**II**	**III**	**IV**	
1	1	1	1	125.6	98.2	110.6	130.1	464.5=T_{111}
2	2	1	1	112.1	101.5	147.4	135.9	496.9=T_{211}
3	1	2	1	150.8	154.8	175.0	185.0	665.6=T_{121}
4	2	2	1	167.1	185.0	174.4	151.5	678.0=T_{221}
5	1	1	2	121.0	100.6	134.8	134.4	490.8=T_{112}
6	2	1	2	149.2	131.1	118.3	161.3	559.9=T_{212}
7	1	2	2	181.1	174.3	137.0	161.5	653.9=T_{122}
8	2	2	2	145.1	201.0	188.8	201.5	736.4=T_{222}
Suma/B				**1152.0**	**1146.5**	**1186.3**	**1261.2**	**4746 = G**

Prueba de hipótesis en la cual:
1. No hay efecto de los factores sobre el comportamiento (tamaño) de insectos.
2. No hay interacción entre los niveles de los tres factores en estudio.

Procedimiento Estadístico

1. Factor de Corrección (C)

$$C = \frac{G^2}{tr} = \frac{4746.0^2}{8X4} = \frac{22524516.00}{32} = 703{,}891.12$$

2. S.C. debida al total (S.C.Total)

$$S.C.Total = \sum y^2 - C$$
$$= 125.6^2 + 112.1^2 + ... + 201.5^2 - C$$
$$= 730{,}897.74 - 703{,}891.12 = 27{,}006.62$$

3. S.C. debida a los bloques (S.C.B.)

$$S.C.B. = \sum \frac{B^2 l}{t} - C$$
$$= \frac{1152.0^2 + ... + 1261.2^2}{8} - C$$
$$= \frac{5639499.38}{8} - C = 1{,}046.30$$

4. S.C. debida a Tratamientos (S.C.T.)

$$SCT = \sum \frac{T_{ijq}^{2}}{r} - C$$

$$= \frac{464.5^2 + 496.9^2 + ... + 736.4^2}{4} - C$$

$$= \frac{2889620.04}{4} - C = 18{,}513.89$$

Interacción Humedad x Temperatura

$$S.C(HT) = \sum \frac{T_{IJ}^{2}}{2r} - SC(H) - SC(T) - C$$

$$S.C.(HT) = \frac{955.3^2 + ... + 1414.4^2}{2X4} - 1205.41 - 1628.11 - 703891.12 = 1.35$$

Interacción Humedad x Luminosidad

Factor	Humedad		Suma/luminosidad
Luminosidad	**H_1**	**H_2**	
l_1	1130.1	1174.9	**2305.0 =L_1**
l_2	1144.7	1296.3	**2441.0=L_2**
Suma/humedad	**2274.8 =H_1**	**2471.2=H_2**	4746.0=G

$$SC(L) = \frac{L_1^2 .. + L_2^2}{4r} - C$$

$$= \frac{2305.0^2 + 2441.0^2}{4X4} - C = 578.00$$

$$SC(HL) = \sum \frac{L_{i.q}^{2}}{2r} - SC(H) - SC(L) - C$$

$$S.C.(HL) = \frac{1130.1^2 + ... + 1296.3^2}{2X4} - 1205.41 - 578.00 - 703891.12$$

$$S.C.(HL) = 356.44$$

5. S.C. de los Efectos Factoriales (S.C. Fac.)

Humedad y Temperatura

Factor	**Humedad**		**Suma/temperatura**
Temperatura	**H_1**	**H_2**	
T_1	955.3	1056.8	**2012.1=$T_{1..}$**
T_2	1319.5	1414.4	**2733.9=$T_{2..}$**
Suma/humedad	**2274.8=h_1**	**2471.2=h_2**	4746.0=G

Interacción temperatura luminosidad

Factor	**Temperatura**		**Suma/luminosidad**
Luminosidad	**T_1**	**T_2**	
l_1	961.4	1343.6	**2305.0 =** l_1
l_2	1050.7	1390.3	2441.0 = l_2
Suma/Temperatura	**2012.1 = T_1**	**2733.9=T_2**	4746.0 = G

Tabla 24. Valores promedio y error estándar correspondiente.

	Promedio	**E.estándar**
General	$X_{ijk} = \frac{X....}{abcn = N}$	
Tratamiento	$X_{ij.} = \frac{\sum X_{ij.}}{n}$	$S_X = \sqrt{\frac{S^2}{n}}$
Factor *A*	$X_{i..} = \frac{X_{i..}}{bcn}$	$S_{Xi..} = \sqrt{\frac{S^2}{bcn}}$
Factor *B*	$X_{.j.} = \frac{X_{.j.}}{acn}$	$S_X = \sqrt{\frac{S^2}{acn}}$
Factor *C*	$X_{..K} = \frac{X_{..k}}{abn}$	$S_{X..k} = \sqrt{\frac{S^2}{abn}}$

$$SC(H) = \frac{H_1^2 ... + H_2^2 ..}{4r} - C$$

$$= \frac{2274.8^2 + 2471.2^2}{4X4} - C$$

$$= 705{,}096.53 - 703{,}891.12 = 1{,}205.41$$

$$SC(T) = \frac{T_1^2 .. + T_2^2 ..}{4R} - C$$

$$= \frac{2012.1^2 + 2733.9^2}{4X4} - C = 16{,}281.11$$

Tabla 25. Análisis de varianza de un experimento factorial 2^3.						
Fuente de variación	**SC**	**gl**	**CM**	**F cal.**	**F_t 0.05**	**F_t 0.01**
Bloques	1,046.30	3	348.76	< 1		
Tratamiento	18,513.89	7	2,644.84	7.45	2.49	3.64
Humedad	1,205.41	1	1,205.41	3.40	4.32	
Temperatura	16,281.11	1	16,281.11	45.92	4.32	8.02
Hum X Temp	1.36	1	1.36	< 1		
Luminosidad	578.00	1	578.00	1.63	4.32	
Hum X Lum	356.44	1	356.44	1.01	4.32	
Tem X Lum	56.71	1	56.71	< 1		
H X T X L	34.87	1	34.87	< 1		
Error	7,446.43	21	354.59			
Total	27,006.62	31				

$$SC(TL) = \sum \frac{T_{.jq}^2}{2r} - SC(T) - SC(L) - C$$

$$SC(TL) = \frac{961.4^2 + ... + 1390.3^2}{2X4} - 16281.11 - 578.00 - 703891.12$$

$$S.C.(TL) = 56.71$$

Interacción de tres factores (Tablas 24 y 25):
SC (HTL) = SCT-SC(H)-SC(T)-SC(L)-SC(HT)-SC(HL)-SC(TL)
SC (HTP) = 18,513.89-1,205.41-16,281.11-1.35-578.00-356.44-56.71
SC (HTL)= 34.87

Cálculo de Diferencia Mínima Significativa (DMS):

$$DMS = t\alpha, n\sqrt{\frac{2s^2}{N}}$$

$$DMS = t_{0.05,21}\sqrt{\frac{2s^2}{4r}}$$

$$DMS = 2.08 * \sqrt{\frac{2(354.59)}{16}} = 13.8$$

Conclusiones

Una característica común de los experimentos en muchas disciplinas es cuando se repiten estos experimentos, los resultados de los tratamientos varían de un ensayo al otro. Obviamente, esta variación genera un grado de incertidumbre con relación a las conclusiones derivados de estos resultados. Ahora bien, existe variación innata entre las unidades experimentales debido a los factores de la herencia y del medio ambiente. Esta variación se denomina el error del experimento o el error experimental cuyo efecto debe distingue de las variaciones debido a la influencia de los tratamientos. Precisamente, es por estas razones que se usan los diseños experimentales, es decir, el uso de los diseños experimentales se debe a la necesidad del determinar la probable diferencia estadística entre diferentes tratamiento y a parte, buscar tendencias o patrones derivados de los resultados. Hay diseños experimentales estándares que se han usado durante casi un siglo en diferentes disciplinas científicas, especialmente en el área de agricultura, biología, psicología sociología, física, etc. Sin embargo, actualmente existen diseños específicos adecuados para cada rama del estudio. A parte hay paquetes especializados (SAS, SPSS, MINITAB, etc.) que permiten la conducción de diferentes tipos de diseños experimentales. El objetivo de este manuscrito es el familiarizar a los alumnos con las bases críticas de los diseños experimentales comunes por medio de un ejemplo de los datos reales. Conociendo estos fundamentos permite un acercamiento más amigable a otros tipos de diseños ejemplificados en diversos paquetes estadísticos.

Referencias

Badii, M.H., J. Castillo, K. Cortez & H. Quiroz. 2007a. Análisis de clusters. Pp. 15-36. In: M. H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, J.N. Barragán & A. E. Flores. 2007b. Análisis discriminante. Pp. 119-136. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, J. Rositas & G. Alarcón. 2007c. Uso de un método de pronóstico en investigación. Pp. 137-155. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H. J. Castillo & M. Rodríguez. 2007d. Respuesta de superficie. Pp. 187-198. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, A. Wong & J.N. Barragán. 2007e. Lattice designs. Pp. 245-290. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, F. Gorjón & R. Foroughbakhch. 2007f. Completely randomized designs. Pp. 307-334. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, R. Rositas & G. Ponce. 2007g. Experimental designs. Pp. 335-348. In: M.H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Badii, M.H., J. Castillo, J. Landeros & K. Cortez. 2007h. Papel de la estadística en la investigación científica. InnOvaciOnes de NegOciOs. 4(1): 107-145.

Brits, S. J. M. & H. H. Lemmer, 1990. An adjusted Friedman test for the nested design. Communic. Statist.- Theor. Meth. 19: 1837-1855.

Conover, W. J. & R. L. Iman, 1976. On some alternative procedures using ranks for the analysis of experimental designs. Communic. Statist-Theor. Meth. 5:1349-1368.

Cox, G. M. & W. P. Martín 1937. Iowa State Collection. Journal Science, 11: 323-331.

Daniel, W. W., 1998. Bioestadística base para el análisis de las ciencias de la salud, Tercera edición, UTEHA Noriega editores, México, 347-380.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. Annual Eugen. 7: 179-188.

Fogiel, M. 2000. Statistics. Research and Education Association, Piscataway, N. J.

Gillman, M. & R. Hails. 1997. An Introduction to Ecological Modeling Putting Practice into Theory. Blackwell Science, London.

Goodman, L.A. 1970. The multivariante analysis of qualitative data: Interactions among multiple classifications. J. Amer. Statist. Assoc. 65:226-256.

Goulden, G. H. 1952. Methods of Statistical Analysis. Second Edition. Wiley, N. Y.

Hill, R, S. 1980. A stopping rule for partitioning dendrograms. Botanical Gazette, 141: 321-324.

Huntsberger, D. V. & P. Billingsley, 1977. Elementos de Estadística Inferencial, Primera edición, Compañía Editorial Continental, México, 320-343.

Iman, R. L., S. C. Hora & W. J. Conover, 1984. Comparison of asymptotically distribution-free procedures for the análisis of complete blocks. J.Amer. Statist.Assoc. 79:674-685.

Kemp, A.W. 1989. A note on Striling's expansion for factorial n. Statist. Prob. Lett. 7: 21-22.

Kepner, J.L. & D. H. Robinson, 1988. Nonparametric methods for detecting treatment effects in repeated measures designs. J..Amer. Statist. Assoc. 83:456-461.

Keppel, G. 1991. Design and Analysis: A Researcher's Handbook 3rd ed. Prentice Hall. Englewood Cliffs New Jersey.

Kirk, R. E. 1982. Experimental design: Procedures for theBehavioral Sciences. 2nd ed. Brooks/Cole, Monterey California.

Lance, G. N. & W. T. Williams. 1967. A genral theory for classificatory sorting strategies. 1. hierarchial systems. Computer Journal, 9: 373-380.

Maxwell, S. E. 1980. Pairwise multiple comparisons in repeated measures designs. J. Educ. Statist. 5:269-287.

Maxwell, S. E. & H. D. Delaney, 1990. Desining Experiment and analyzing data. Wadsworth, Belmont, California.

O'Brien, R. G., & M. Kaiser, 1985. MANOVA meted for analyzing repeated measures designs. Psychol. Bull. 97: 316-333.

Ostle, B., 1994. Estadística Aplicada, Primera edición, Editorial Limusa, México, 447-452.

Pompa, J., L. Mucina, O. Van Tongeren & E. Van der Maarel. 1983. On the determination of optimal levels in phytosociological classifications. Vegetario, 52: 65-75.

Ratliff, R. D. & R. D. Pieper. 1982. Approaches to plant community classification for range managers. Journal of Range Management Monographs series, No. 1.

Rohlf, F. G. 1974. Methods of comparing classifications. Annual Review of Ecology and Systematics, 5: 101-113.

Rohlf, F. G. 1982. Consensus indices for comparing classifications. Mathematical Biological Sciences, 59: 131-144.

Rositas, J., M.H. Badii, J. Castillo & R. Foroughbakhch. 2007. Técnicas de investigación basadas en sistemas de modelación estructurada. Pp. 291-304. In: M. H. Badii & J. Castillo (eds.). Técnicas Cuantitativas en la Investigación. UANL, Monterrey.

Samuels, M. L., G. Casella, & G. P. McCabe 1991. Interpreting blocks and random factors. J. Amer. Statist. Assoc. 86: 798-808.

Shearer, P. R. 1973. Missing data in quantitative designs. J. Royal Statist. Soc. Ser. C. Appl. Statist. 22: 135-140.

Spiegel, M. R., 1976. Probabilidad y Estadística Serie Schaum, Primera edición, Editorial McGraw-Hill, México.

Sneath, P. H. A. & R. R. Sokal. 1973. Numerical Taxonomy. Freeman, San Francisco.

Southwood, T. R. E. & P. A. Henderson. 2000. Ecological Methods. Third Edition. Blackwell Science, London.

Anexo A

Tabla A. Tabla de ANOVA para el diseño completamente al azar*.

FV	Gl	SC	CM	$F_{observado}$
Tratamientos	t - 1	$\Sigma(X^2t/r)$ - c	SC_t/gl	CM_t/CM_E
Error	resta	resta	SC_E/gl	
Total	rt - 1	$\Sigma(X^2)$ - c	__	
*: FV = Fuente de variabilidad, Gl= Grados de libertad, SC= Suma de Cuadrados, CM = Cuadrado medio, t= # de tratamientos, r = # de repeticiones, c = Factor de corrección = $(\Sigma X)^2/N$, N = # total de repeticiones.				

Tabla B. Tabla de ANOVA para el diseño bloques al azar*.

FV	Gl	SC	CM	$F_{observado}$
Bloques	b -1	$\Sigma(X^2b/r)$ - c	SC_B/gl	CM_{Bl}/CM_E
Tratamientos	T -1	$\Sigma(X^2t/r)$ - c	SC_t/gl	CM_t/CM_E
Error	resta	resta	SC_E/gl	
Total	rt - 1	$\Sigma(X^2)$ - c	__	
*: b = # de bloques, los demás acrónimos como en la Tabla A.				

Tabla C. Tabla de ANOVA para el diseño Cuadro latino*.

FV	Gl	SC	CM	$F_{observado}$
Hilera	h - 1	$\Sigma(X^2h/r)$ - c	SC_h/gl	CM_h/CM_E
Columna	c - 1	$\Sigma(X^2c/r)$ - c	SC_c/gl	CM_c/CM_E
Tratamientos	t - 1	$\Sigma(X^2t/r)$ - c	SC_t/gl	CM_t/CM_E
Error	resta	resta	SC_E/gl	
Total	rt - 1	$\Sigma(X^2)$ - c	__	
*: h = # de bloques horizontales, c = # de bloques verticales, los demás acrónimos como en la Tablas A y B.				

Tabla C. Tabla de ANOVA para el diseño factorial*.

FV	Gl	SC	CM	$F_{observado}$
rep	r - 1	$\Sigma(X^2r/at)$ - c	SC_r/gl	CM_r/CM_E
Tratamientos	at - 1	$\Sigma(X^2at/r)$ - c	SC_t/gl	CM_t/CM_E
Factor A	a - 1	$\Sigma(X^2a/tr)$ - c	SC_a/gl	CM_a/CM_E
Factor B	b - 1	$\Sigma(X^2b/tr)$ - c	SC_b/gl	CM_b/CM_E
A X B	a-1(b-1)	resta	SC_{aXb}/gl	CM_{aXb}/CM_E
Error	resta	resta	SC_E/gl	
Total	rt - 1	$\Sigma(X^2)$ - c	__	
*: h = # de bloques horizontales o hileras, c = # de bloques verticales o columnas, los demás acrónimos como en la Tablas A, B y C.				