DISEÑO EXPERIMENTAL

UNIDAD N° 5

DISEÑO DEL EXPERIMENTO

Es necesario planear primero el experimento que se va a aplicar

- Aleatorización completa
- Bloques aleatorizados
- Cuadrados latinos
- Cuadrados grecolatinos

ALEATORIZACIÓN COMPLETA

Suponga que se tiene un experimento agrícola. Para diseñar este experimento se puede dividir la tierra en 4 x 4 = 16 parcelas y asignar cada tratamiento (indicados por las letras *A*, *B*, C, *D*) a cuatro bloques elegidos en forma totalmente aleatoria

El propósito de la aleatorización es eliminar diversas fuentes de error

D	A	C	C
B	D	B	A
D	C	B	D
A	B	C	A

ALEATORIZACIÓN COMPLETA

La aleatorización es una técnica que se utiliza para equilibrar el efecto de condiciones externas o no controlables que pueden influir en los resultados de un experimento

Por ejemplo, la temperatura ambiental, la humedad, la materia prima o los operadores pueden cambiar durante un experimento y afectar inadvertidamente los resultados de la prueba

Al aleatorizar el orden en el que se realizan las corridas experimentales, se reduce la probabilidad que las diferencias en los materiales o las condiciones del experimento sesguen considerablemente los resultados

La aleatorización también permite estimar la variación inherente de los materiales y las condiciones de manera que se puedan hacer inferencias estadísticas válidas con base en los datos del experimento

BLOQUES ALEATORIZADOS

Cuando se necesita todo un conjunto de tratamientos para cada bloque, los tratamientos *A*, *B*, C y *D* se introducen en orden aleatorio en cada uno de los bloques I, II, III y IV, y por esta razón a los bloques se les llama *bloques aleatorizados*

Este tipo de diseño se emplea para controlar *una fuente de error* o *variabilidad*: a saber, la diferencia entre los bloques

Bloques	**Tratamientos**			
I	C	B	A	D
II	A	B	D	C
III	B	C	D	A
IV	A	D	C	B

BLOQUES ALEATORIZADOS

Algunos experimentos diseñados pueden proveer información de manera efectiva cuando las mediciones son difíciles o muy costosas de hacer o pueden minimizar el efecto de variabilidad no deseada en la inferencia del tratamiento

Un diseño de bloques aleatorizados es un diseño frecuentemente utilizado para minimizar el efecto de la variabilidad cuando se asocia con unidades discretas (por ejemplo, ubicación, operador, planta, lote, tiempo)

El caso usual consiste en distribuir aleatoriamente una réplica de cada combinación de tratamientos dentro de cada bloque

Por lo general, no hay un interés intrínseco en los bloques, y se considera que éstos son factores aleatorios

La suposición habitual es que el bloque por interacción de tratamiento es cero, y esta interacción pasa a ser el término de error para probar los efectos del tratamiento

CUADRADOS LATINOS

En el diseño de cuadrado latino se controlan dos factores de bloque y se estudia un factor de tratamiento, por lo que se tienen cuatro fuentes de variabilidad que puede afectar a la respuesta observada: los tratamientos, el factor de bloque 1 (columnas), el factor de bloque 2 (renglones) y el error aleatorio

Se llama cuadrado latino por dos razones:

*es un cuadro debido a que tiene la restricción adicional de que los tres factores involucrados se prueban en la misma cantidad de niveles

*es latino porque se utilizan letras latinas para denotar a los tratamientos o niveles del factor de interés

	Factor 1			
Factor 2	D	B	C	A
	B	D	A	C
	C	A	D	B
	A	C	B	D

EJEMPLO (CUADROS LATINOS)

- PARA ESTUDIAR EL EFECTO DE LA *ILUMINACIÓN* (A=NATURAL, B=MUY FUERTE, C=ESCASA) EN LA *VELOCIDAD DE LECTURA* SE REALIZA UN EXPERIMENTO QUE CONSISTE EN CONTAR EL NÚMERO DE PALABRAS LEÍDAS EN UN MINUTO PARA DISTINTOS TIPOS DE PAPEL (B=BLANCO, C=EN COLOR, S=SATINADO) Y DIFERENTE TIPOGRAFÍA (G=LETRA GRANDE, P=LETRA PEQUEÑA, N=NORMAL). OBTENIÉNDOSE LOS RESULTADOS EXPUESTOS EN LA SIGUIENTE TABLA:

		Tipo de papel		
		satinado	blanco	color
Letra	grande	258 A	230 C	240 B
	normal	235 B	270 A	240 C
	pequeña	220 C	225 B	260 A

- ANALIZAR ESTOS DATOS Y ESTUDIAR LA POSIBLE INFLUENCIA DE LOS FACTORES *ILUMINACIÓN, TIPOS DE PAPEL* Y *DIFERENTE TIPOGRAFÍA* EN LA VARIABLE DE INTERÉS *VELOCIDAD DE LECTURA*, A UN NIVEL DE SIGNIFICANCIA DE 0,05

SOLUCIÓN

- PRIMERO SE OBTIENEN LOS TOTALES DE LOS RENGLONES (TIPO DE LETRA) Y LOS TOTALES DE LAS COLUMNAS (TIPO DE PAPEL). TAMBIÉN SE OBTIENE EL TOTAL DE PALABRAS LEÍDAS SEGÚN EL TIPO DE ILUMINACIÓN, COMO SE MUESTRA EN LA TABLA:

				TOTAL
	258 A	230 C	240 B	728
	235 B	270 A	240 C	745
	220 C	225 B	260 A	705
TOTAL	713	725	740	2178

	A	B	C	
TOTAL	788	700	690	2178

SOLUCIÓN

- LUEGO OBTENEMOS LA VARIACIÓN TOTAL Y LAS VARIACIONES DE LOS RENGLONES, LAS COLUMNAS Y LOS TRATAMIENTOS:
- VARIACIÓN TOTAL

$$V = 258^2 + 230^2 + \cdots + 260^2 - \frac{2178^2}{9} = 529414 - 527076 = 2338$$

- VARIACIÓN ENTRE RENGLONES

$$V_R = \frac{728^2}{3} + \frac{745^2}{3} + \frac{705^2}{3} - \frac{2178^2}{9} = 527344{,}667 - 527076 = 268{,}667$$

- VARIACIÓN ENTRE COLUMNAS

$$V_C = \frac{713^2}{3} + \frac{725^2}{3} + \frac{740^2}{3} - \frac{2178^2}{9} = 527198 - 527076 = 122$$

- VARIACIÓN ENTRE TRATAMIENTOS

$$V_B = \frac{788^2}{3} + \frac{700^2}{3} + \frac{690^2}{3} - \frac{2178^2}{9} = 529014{,}667 - 527076 = 1938{,}667$$

- VARIACIÓN DEBIDA AL ERROR

$$2338 - 268{,}667 - 122 - 1938{,}667 = 8{,}666$$

SOLUCIÓN

Variación	Grados de libertad	Cuadrado medio	F
Renglones, 268,667	2	$\hat{S}_{\mathrm{R}}^2 = \frac{268,667}{2} = 134,33$	$\frac{134,33}{4,333} = 31,001$
Columnas, 122	2	$\hat{S}_{\mathrm{C}}^2 = \frac{122}{2} = 61$	$\frac{61}{4,333} = 14,078$
Tratamientos, 1938,667	2	$\hat{S}_{\mathrm{I}}^2 = \frac{1938,667}{2} = 969,33$	$\frac{969,33}{4,333} = 223,708$
Residuales, 8,666	2	$\hat{S}_{\mathrm{E}}^2 = \frac{8,666}{2} = 4,333$	
Total, 2338	8		

Para los grados de libertad consideramos K (número de niveles), por lo que en los renglones, columnas y tratamientos es K-1, mientras que en la variación residual es (K-1)*(K-2)

Se tiene que $F_{0.05,2,2} = 19$. Por lo tanto, la hipótesis el tipo de letra no influye se rechaza. Con respecto al tipo de papel, se acepta que este factor no influye. En cuanto al tipo de iluminación, se rechaza la hipótesis de que no influye este factor.

CUADRADOS GRECOLATINOS

Cuando es necesario controlar *tres fuentes de error o variabilidad* se emplea un *cuadrado grecolatino*

Estos cuadrados son, en esencia, dos cuadrados latinos superpuestos uno sobre otro, usando las letras latinas *A, B,* C y *D* para uno de los cuadrados y las letras griegas $\alpha, \beta, \gamma\ y\ \delta$ para el otro

Un requerimiento adicional por satisfacer es que cada letra griega debe usarse una y sólo una vez con cada letra latina; si se satisface esta condición se dice que el cuadrado es *ortogonal*

B_γ	A_β	D_δ	C_α
A_δ	B_α	C_γ	D_β
D_α	C_δ	B_β	A_γ
C_β	D_γ	A_α	B_δ

EJEMPLO (CUADROS GRECOLATINOS)

- INTERESA SABER SI EXISTE DIFERENCIA SIGNIFICATIVA ENTRE LAS MILLAS RECORRIDAS POR GALÓN, ENTRE LAS GASOLINAS A, B, C Y D.
- SE CONSIDERARON LOS SIGUIENTES FACTORES DE BLOQUEO:
  - FILA: TIPO DE VEHÍCULO
  - COLUMNA: CONDUCTOR
  - LETRA GRIEGA: TIPO DE CARRETERA
- CON UN ARREGLO EN FORMA DE CUADRADO GRECO LATINO, EL NÚMERO DE MILLAS POR GALÓN RESULTÓ SER:

	Conductor 1	**Conductor 2**	**Conductor 3**	**Conductor 4**
Vehículo 1	B_γ 19	A_β 16	D_δ 16	C_α 14
Vehículo 2	A_δ 15	B_α 18	C_γ 11	D_β 15
Vehículo 3	D_α 14	C_δ 11	B_β 21	A_γ 16
Vehículo 4	C_β 16	D_γ 16	A_α 15	B_δ 23

- ANALIZAR A UN NIVEL DE SIGNIFICACIÓN DEL 5% SI EXISTE DIFERENCIA SIGNIFICATIVA ENTRE LOS CUATRO TIPOS DE GASOLINA.

SOLUCIÓN

- PRIMERO SE OBTIENEN LOS TOTALES DE LOS RENGLONES Y COLUMNAS:

					Total
	B_γ 19	A_β 16	D_δ 16	C_α 14	65
	A_δ 15	B_α 18	C_γ 11	D_β 15	59
	D_α 14	C_δ 11	B_β 21	A_γ 16	62
	C_β 16	D_γ 16	A_α 15	B_δ 23	70
Total	64	61	63	68	256

- DESPUÉS SE OBTIENEN LOS TOTALES CORRESPONDIENTES A CADA LETRA LATINA Y CADA LETRA GRIEGA:

Total A: $15 + 16 + 15 + 16 = 62$

Total B: $19 + 18 + 21 + 23 = 81$

Total C: $16 + 11 + 11 + 14 = 52$

Total D: $14 + 16 + 16 + 15 = 61$

Total α: $14 + 18 + 15 + 14 = 61$

Total β: $16 + 16 + 21 + 15 = 68$

Total γ: $19 + 16 + 11 + 16 = 62$

Total δ: $15 + 11 + 16 + 23 = 65$

SOLUCIÓN

- AHORA SE CALCULAN LAS VARIACIONES CORRESPONDIENTES:
  - RENGLONES: $\frac{65^2}{4}+\frac{59^2}{4}+\frac{62^2}{4}+\frac{70^2}{4}-\frac{256^2}{16}=4112{,}50-4096=16{,}50$
  - COLUMNAS: $\frac{64^2}{4}+\frac{61^2}{4}+\frac{63^2}{4}+\frac{68^2}{4}-\frac{256^2}{16}=4102{,}50-4096=6{,}50$
  - GASOLINAS (*A, B, C, D*): $\frac{62^2}{4}+\frac{81^2}{4}+\frac{52^2}{4}+\frac{61^2}{4}-\frac{256^2}{16}=4207{,}50-4096=111{,}50$
  - CARRETERAS (Α, Β, Γ, Δ): $\frac{61^2}{4}+\frac{68^2}{4}+\frac{62^2}{4}+\frac{65^2}{4}-\frac{256^2}{16}=4103{,}50-4096=7{,}50$
- LA VARIACIÓN TOTAL ES:

$$19^2+16^2+16^2+\cdots+15^2+23^2-\frac{256^2}{16}=4244-4096=148{,}00$$

- LA VARIACIÓN DEBIDA AL ERROR ES:

$$148{,}00-16{,}50-6{,}50-111{,}50-7{,}50=6{,}00$$

SOLUCIÓN

- LOS RESULTADOS SE MUESTRAN EN LA SIGUIENTE TABLA DE ANÁLISIS DE VARIANZA. EL NÚMERO TOTAL DE GRADOS DE LIBERTAD ES $N^2 - 1$, YA QUE SE TRATA DE UN CUADRADO DE $N \times N$. CADA UNO DE LOS RENGLONES, DE LAS COLUMNAS, DE LAS LETRAS LATINAS Y DE LAS LETRAS GRIEGAS TIENE $N - 1$ GRADOS DE LIBERTAD. POR LO TANTO, LOS GRADOS DE LIBERTAD PARA EL ERROR SON $N^2 - 1 - 4(N - 1) = (N - 1)(N - 3)$. EN ESTE CASO, $N = 4$

Variación	Grados de libertad	Cuadrado medio	F
Renglones (automóviles), $16{,}50$	3	$5{,}5$	$5{,}5\ /2\ = 2{,}75$
Columnas (conductores), $6{,}50$	3	$2{,}167$	$2{,}167/2\ = 1{,}08$
Gasolinas (A, B, C, D), $111{,}50$	3	$37{,}167$	$37{,}167/2\ = 18{,}6$
Carreteras (α, β, γ, δ), $7{,}50$	3	$2{,}5$	$2{,}5/2\ = 1{,}25$
Error, $6{,}00$	3	2	
Total, $148{,}00$	15		

- SE TIENE QUE $F_{0.05,3,3} = 9{,}28\ (F_{(\alpha;(N-1);(N-1)*(N-3))})$. POR LO TANTO, LA HIPÓTESIS DE QUE LAS GASOLINAS SON IGUALES PUEDE RECHAZARSE A UN NIVEL DE SIGNIFICANCIA DE 0,05