Diseño en Bloques Completamente Aleatorizados

Prueba de Aditividad de Tukey

Note que en el diseño de bloques, uno de los supuestos es la Aditividad entre el Bloque y el Factor. Para ello, existe una prueba propuesta por Tukey(1949).

Para detectar la no aditividad gráficamente, se realiza un gráfico de dispersión entre los residuales y los valores predichos. Si se observa una tendencia cuadrática en el gráfico da indicios de presencia de NO aditividad transformable, esto es, existe NO aditividad que se puede remover al aplicar una transformación. Para la determinación de la NO aditividad, considere el siguiente modelo para el diseño de bloques completos al azar con interacción:

$$Y_{ij} = \mu + \tau_i + \beta_j + \gamma\tau_i\beta_j + \varepsilon_{ij}$$

El valor de γ es estimado a partir de:

$$\hat{\gamma} = \frac{\sum_i \sum_j (Y_{i\,.} - \bar{Y}_{.\,.})(Y_{.\,j} - \bar{Y}_{.\,.}) Y_{ij}}{\sum_i (Y_{i\,.} - \bar{Y}_{.\,.})^2 \sum_j (Y_{.\,j} - \bar{Y}_{.\,.})^2}$$

...

Además, la SS_T se puede descomponer como:

$$SS_T = SSA + SSB + SSAB^* + SS_{Res}$$

donde $SSAB^*$ representa la suma de cuadrados de la **interacción** entre el **Factor** (A) y el **bloque** (B) y SS_{Res}: El restante de la suma de cuadrados, dado por:

$$SS_{Res} = SS_T - SSA - SSB - SSAB^*$$

$SSAB^*$ se halla a partir de un estimador para:

$$\sum_i \sum_j \gamma^2 \tau_i^2 \beta_j^2$$

obtenido a partir de:

$$SSAB^* = \frac{\left[\sum_i \sum_j (Y_{i\,.} - \bar{Y}_{.\,.})(Y_{.\,j} - \bar{Y}_{.\,.}) Y_{ij}\right]^2}{\sum_i (Y_{i\,.} - \bar{Y}_{.\,.})^2 \sum_j (Y_{.\,j} - \bar{Y}_{.\,.})^2}$$

...

La hipótesis de aditividad se traduce en probar:

$$H_0: \gamma = 0 \text{ vs. } H_1: \gamma \neq 0$$

Bajo H_0, el siguiente Estadístico de Prueba:

$$F^* = \frac{SSAB^*/1}{SS_{Res}/(ab-a-b)}$$

tiene un distribución $F_{1,\, ab-a-b}$.
Se rechaza H_0 si $F_{cal} > F_{\alpha,1,ab-a-b}$ y con el VP de la prueba, se rechaza H_0 si $VP < \alpha$.
En $\mathbf{R}$ se utiliza la función **tukey.add.test** de la librería: **library(asbio)**. Para el caso de estudio:

```
library(asbio)
tukey.add.test(y, agente, rollo)
```

Resultados que arroja la función:

```
Tukey's one df test for additivity
F = 0.1060423   Denom df = 11    p-value = 0.7508062
```

Como el $VP = 0.75 > \alpha = 0.05$ no se rechaza H_0, y por tanto se concluye que existe aditividad entre el Agente químico y el Rollo de tela a un nivel de significancia del $5\,\%$.

Diseño de Cuadrados Latinos

Anotaciones

- Se utiliza cuando en un experimento de un factor se desea determinar si hay diferencias entre los tratamientos controlando por **DOS fuentes de variabilidad externa**, con el fin de **aislar** y **eliminar** del término de error la variación atribuida a las dos fuentes.
- Se conoce como un diseño de factor con **DOS** restricciones de aleatorización.
- Se denomina Cuadrados Latinos ya que para denotar los tratamientos se utilizan letras **latinas**: $A, B, C, D, E, \ldots$
- Los niveles tanto de los tratamientos (factor) como de los bloques es el mismo.
- En total este diseño cuenta con **CUATRO** fuentes de variabilidad que pueden afectar la respuesta de interés:
  - Los tratamientos (**Letra Latina**)
  - El factor de bloque **I** (**Fila**)
  - El factor de bloque **II** (**Columna**)
  - El error Aleatorio.
- Los Cuadrados Latinos tienen la particularidad que en cada **Fila** y en cada **Columna** debe aplicarse todos los **Tratamientos**.
- El cuadrado latino se debe escoger antes de realizar el experimento.

Diseño de Cuadrados Latinos

Ejemplos de Cuadrados Latinos

- $\mathbf{p = 3}$ Tratamientos- Cuadrado Latino estándar

Fila/Columna	1	2	3
1	A	B	C
2	B	C	A
3	C	A	B

- $\mathbf{p = 4}$ Tratamientos- Cuadrado Latino estándar

Fila/Columna	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

...

- $\mathbf{p = 5}$ Tratamientos- Cuadrado Latino estándar

Fila/Columna	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

- **IMPORTANTE**
  En general un cuadrado latino para $\mathbf{p}$-tratamientos es un cuadrado $\mathbf{p} \times \mathbf{p}$, $\mathbf{p}$-filas $\mathbf{p}$-columnas, donde cada una de las celdas resultante, $\mathbf{p^2}$, contiene una de las $\mathbf{p}$-letras y cada letra ocurre una sola vez en cada fila y columna.

Modelo Estadístico

$$\boxed{\mathbf{Y_{ijk}} = \mu + \alpha_{\mathbf{i}} + \tau_{\mathbf{j}} + \beta_{\mathbf{k}} + \varepsilon_{\mathbf{ijk}}, \quad \mathbf{i},\ \mathbf{j},\ \mathbf{k} = \mathbf{1},\ \mathbf{2},\ \ldots,\ \mathbf{p}}$$

$\mathbf{i}$: Fila, $\mathbf{j}$: Tratamiento, $\mathbf{k}$: Columna; $\alpha_{\mathbf{i}}$: Efecto de la Fila; $\tau_{\mathbf{j}}$: Efecto del factor; $\beta_{\mathbf{k}}$: Efecto de la Columna.

El modelo anterior es un **modelo completamente aditivo** (**NO** hay interacción entre Filas-Columnas-Tratamientos).

Supuesto: $\varepsilon_{\mathbf{ijk}} \sim \mathbf{NID}(\mathbf{0},\ \sigma^{\mathbf{2}})$.

Descomposición de la Suma de Cuadrados:

$$\begin{aligned}
\mathbf{SST} &= \mathbf{SSF} + \mathbf{SSCol} + \mathbf{SSTrat} + \mathbf{SSE} \\
(\mathbf{gl})\mathbf{p^2} - \mathbf{1} &= (\mathbf{p}-\mathbf{1}) + (\mathbf{p}-\mathbf{1}) + (\mathbf{p}-\mathbf{1}) + (\mathbf{p}-\mathbf{2})(\mathbf{p}-\mathbf{1})
\end{aligned}$$

Hipótesis Central

$$\mathbf{H_0}: \ \mu_1 = \mu_2 = \ldots = \mu_p \quad \mathbf{H_1} ::; \mu_j \neq \mu \text{ p.a } \mathbf{j}$$

equivalente a:

$$\mathbf{H_0}: \ \tau_1 = \tau_2 = \ldots = \tau_p = \mathbf{0} \quad \mathbf{H_1}: \ \tau_j \neq \mathbf{0} \text{ p.a} \mathbf{j}.$$

Estadístico de Prueba

$$\mathbf{F_0} = \frac{\mathbf{MSTrat}}{\mathbf{MSE}} \sim_{\text{Bajo}\mathbf{H_0}} \mathbf{F}_{(\mathbf{p}-\mathbf{1}),\,(\mathbf{p}-\mathbf{2})(\mathbf{p}-\mathbf{1})}.$$

Diseño de Cuadrados Latinos

Anova

Fuente de Variación	SS	gl	MS	F_0	VP
Tratamientos	$SSTrat = \frac{1}{p}\sum_{j=1}^{p} Y_{.j.}^2 - \frac{Y_{...}^2}{N}$	$p-1$	MSTrat	$\frac{MSTrat}{MSE}$	
Fila	$SSF = \frac{1}{p}\sum_{i=1}^{p} Y_{i..}^2 - \frac{Y_{...}^2}{N}$	$p-1$	MSF	$\frac{MSF}{MSE}$	
Columna	$SSC = \frac{1}{p}\sum_{k=1}^{p} Y_{..k}^2 - \frac{Y_{...}^2}{N}$	$p-1$	MSC	$\frac{MSC}{MSE}$	
Error	SSE	$(p-1)(p-2)$	MSE		
Total	$SST = \sum\sum\sum Y_{ijk}^2 - \frac{Y_{...}^2}{N}$	p^2-1			

Residuales:

$$e_{ijk} = Y_{ijk} - \hat{Y}_{ijk} = Y_{ijk} - \bar{Y}_{i..} - \bar{Y}_{.j.} - \bar{Y}_{..k} + 2\bar{Y}_{...}$$

Replicación de los Cuadrados Latinos

Una de las desventajas de los cuadrados latinos con p-pequeño es el número reducido de grados de libertad para el error. Una solución aconsejable es replicarlos, para incrementar los grados de libertad del error.

Hay tres formas de replicarlos:

- Usando las mismas filas y columnas en cada réplica.
- Usando las mismas filas pero diferentes columnas en cada réplica (o al revés).
- Usando diferentes filas y diferentes columnas.

Diseño de Cuadrados Latinos

CASO 1:

$\mathbf{N = np^2}$ $\mathbf{n}$: Número de réplicas.

Cada fila y cada columna cuenta con $\mathbf{np}$-observaciones, igual los tratamientos.

Parte de la ANOVA

Fuente de Variación	Grados Libertad
Tratamientos	$p-1$
Fila	$p-1$
Columna	$p-1$
Réplica	$n-1$
Error	$(p-1)[n(p+1)-3]$
Total	np^2

NOTA: Los otros casos se pueden consultar en el texto de Montgomery.

Diseño de Cuadrado Grecolatino

Estudio de Caso 4: Producción de Maíz

En un experimento que involucra diferentes variedades de maíz se consideró que el **espaciamiento del maíz** podría influir en la producción, así que **cinco métodos de espaciamiento** fueron estudiados para **cinco variedades de maíz** y **cinco localizaciones** al igual que **cinco tratamientos** de fertilizantes. Un diseño en cuadrado grecolatino fue elegido para el estudio y las producciones de maíz por **Ha** se presentan a continuación:

		Localización				
		L1	L2	L3	L4	L5
Variedad	V1	G2 C 5.65	G4 D 7.68	G1 E 8.75	G3 B 4.32	G5 A 5.27
	V2	G5 B 3.79	G3 C 8.35	G2 D 4.98	G1 A 5.94	G4 E 7.50
	V3	G3 E 8.12	G2 A 6.27	G4 B 4.22	G5 D 7.29	G1 C 4.71
	V4	G4 A 7.93	G1 B 4.77	G5 C 6.92	G2 E 8.48	G3 D 6.5
	V5	G1 D 4.85	G5 E 8.88	G3 A 8.45	G4 C 4.49	G2 B 4.88

Diseño de Cuadrado Grecolatino

Anotaciones

- Este diseño resulta cuando se **sobrepone** en un cuadrado latino un segundo cuadrado latino, donde los tratamientos se denotan por **letras griegas**
- **Interés:**
  Controlar **TRES** fuentes de variabilidad externa con el fin de determinar si hay o no diferencias entre los tratamientos.
- Si se garantiza la propiedad de que cada letra **griega** aparece una y sólo una vez con cada letra **latina**, entonces los dos cuadrados latinos se dice que son **ortogonales** y el diseño obtenido es llamado **cuadrado grecolatino**.
- Se conoce con el nombre de **Bloqueo** en **TRES** direcciones.
- En este tipo de diseño las unidades experimentales se agrupan en **TRES** formas diferentes.
- Cada letra griega debe aparecer un vez y sólo una en cada fila y en cada columna. Así las letras forman un cuadrado latino con respecto a filas y columnas.

Diseño de Cuadrado Grecolatino

Ejemplo de un Cuadrado Grecolatino

$\mathbf{p = 4}$ Tratamientos- Cuadrado Latino Estándar

Fila/Columna	1	2	3	4
1	A α	B β	C γ	D δ
2	B δ	A γ	D β	C α
3	C β	D α	A δ	B γ
4	D γ	C δ	B α	A β

Modelo Estadístico

$$\boxed{\mathbf{Y_{ijkl}} = \mu + \theta_{\mathbf{i}} + \tau_{\mathbf{j}} + \omega_{\mathbf{k}} + \psi_{\mathbf{l}} + \varepsilon_{\mathbf{ijkl}}}$$

Supuestos: $\varepsilon_{\mathbf{ijkl}} \sim \mathbf{N}(\mathbf{0}, \sigma^2)$ y aditividad entre todos los efectos considerados.
Efectos: Fila ($\theta_{\mathbf{i}}$), Letra latina- Tratamientos ($\tau_{\mathbf{j}}$), Letra Griega ($\omega_{\mathbf{k}}$), columna ($\psi_{\mathbf{l}}$).

ANOVA de un cuadrado Grecolatino

Fuente de Variación	Grados Libertad	Sumas de Cuadrados
Letra griega	$p-1$	$\frac{1}{p}\sum_{k=1}^{p} Y_{..k.}^2 - \frac{Y_{....}^2}{N}$
Letra latina	$p-1$	$\frac{1}{p}\sum_{j=1}^{p} Y_{.j..}^2 - \frac{Y_{....}^2}{N}$
Fila	$p-1$	$\frac{1}{p}\sum_{i=1}^{p} Y_{i...}^2 - \frac{Y_{....}^2}{N}$
Columna	$p-1$	$\frac{1}{p}\sum_{l=1}^{p} Y_{...l}^2 - \frac{Y_{....}^2}{N}$
Error	$(p-1)(p-3)$	Por diferencia
Total	p^2-1	$\sum_{i,j,k,l}^{p} Y_{ijkl}^2 - \frac{Y_{....}^2}{N}$

$\mathbf{N = p^2}$.

Diseño de Cuadrado Grecolatino

Taller 3 Estudio de Caso 4: Producción de Maíz

En un experimento que involucra diferentes variedades de maíz se consideró que el **espaciamiento del maíz** podría influir en la producción, así que **cinco métodos de espaciamiento** fueron estudiados para **cinco variedades de maíz** y **cinco localizaciones** al igual que **cinco tratamientos** de fertilizantes. Un diseño en cuadrado grecolatino fue elegido para el estudio y las producciones de maíz por **Ha** se presentan a continuación:

		Localización				
		L1	L2	L3	L4	L5
Variedad	V1	G2 C 5.65	G4 D 7.68	G1 E 8.75	G3 B 4.32	G5 A 5.27
	V2	G5 B 3.79	G3 C 8.35	G2 D 4.98	G1 A 5.94	G4 E 7.50
	V3	G3 E 8.12	G2 A 6.27	G4 B 4.22	G5 D 7.29	G1 C 4.71
	V4	G4 A 7.93	G1 B 4.77	G5 C 6.92	G2 E 8.48	G3 D 6.5
	V5	G1 D 4.85	G5 E 8.88	G3 A 8.45	G4 C 4.49	G2 B 4.88

Letra latina: Tipo de fertilizante, letra Griega: Gj: Espaciamiento. ¿Hay diferencia significativa entre las producciones bajo los diferentes fertilizantes?.

1. Identifique el nombre del diseño experimental utilizado, describa la forma como se garantizaría que los datos obtenidos cumplen con los requerimientos del diseño.
2. Escriba el modelo estadístico, identificando claramente cada una de sus componentes para el problema bajo estudio, indicando los respectivos supuestos.

Diseño de Cuadrado Grecolatino

Taller 3 Continuación

3. Formule la hipótesis que se quiere probar de acuerdo al problema.
4. Realice el análisis estadístico apropiado a estos datos y obtenga las respectivas conclusiones. (Verificación de los supuestos, probar la significancia del factor, y determinar si es necesario o no considerar cada uno de los bloques involucrados, determine las respectivas agrupaciones entre los tratamientos.)

Problema Adicional: Se diseñó un experimento para estudiar el rendimiento de **cuatro** detergentes. Las siguientes lecturas de blancura se obtuvieron con un equipo especial diseñado para **12 cargas** de lavado distribuídas en **tres** modelos de lavadoras:

	Lavadora 1	Lavadora 2	Lavadora 3
Det. A	45	43	51
Det. B	47	44	52
Det. C	50	49	57
Det. D	42	37	49

1. Identifique el nombre del diseño experimental utilizado, describa la forma como se garantizaría que los datos obtenidos cumplen con los requerimientos del diseño.
2. Escriba el modelo estadístico, identificando claramente cada una de sus componentes para el problema bajo estudio, indicando los respectivos supuestos.
3. Formule la hipótesis que se quiere probar de acuerdo al problema.
4. Realice el análisis estadístico apropiado a estos datos y obtenga las respectivas conclusiones. (Verificación de los supuestos, probar la significancia del factor, y determinar si es necesario o no considerar el tipo de lavadora, determine las respectivas agrupaciones entre los tratamientos.)
5. Conteste los incisos anteriores sin tener en cuenta el efecto del tipo de lavadora. Hay diferencias en las conclusiones obtenidas y aquellas que se obtuvieron en 4.? Explique. Con cuáles conclusiones se queda? Argumente.

Diseño de Cuadrado Grecolatino

Anotaciones con respecto a los cuadrados Latinos

1. Note que si $\mathbf{p}=\mathbf{4}$ se pueden construir un total de $576=4!*(4-1)!\times\#CLE$ donde $\#CLE$ es el número de cuadrados latinos estándar, en este caso 4.
2. Los cuadrados latinos estándar son aquellos donde la primera fila y la primera columna están ordenados en orden alfabético, por ejemplo los cuatro cuadrados latinos estándar en el caso $p=4$ son.

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

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

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

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

3. Con respecto a la selección aleatoria del cuadrado latino: Como dado un cuadrado latino cualquier intercambio de columnas o de filas produce un nuevo cuadrado Latino. Se usa como estrategia de selección y aleatorización la siguiente:
   i. Se construye el cuadrado latino estándar más sencillo.
   ii. Se aleatoriza el orden de las filas (o columnas) y después se aleatoriza el orden de las columnas (o filas).
   iii. Por último, los tratamientos a comparar se asignan en forma aleatoria a las letras latinas.

Diseño de Cuadrado Grecolatino

Cuadrados Latinos Aleatorios

Por ejemplo:

CUADRADO ESTÁNDAR INICIAL					Aleatorización FILA					Aleatorización COLUMNA			
Aleatorización					sample(1:4)			2 1 4 3				[1] 2 4 3 1	
										1	2	3	4
A	B	C	D		B	C	D	A	1	C	A	D	B
B	C	D	A		A	B	C	D	2	B	D	C	A
C	D	A	B		D	A	B	C	3	A	C	B	D
D	A	B	C		C	D	A	B	4	D	B	A	C

Nota:

- El cuadrado latino tiene dos restricciones a la aleatorización que se deben a los dos factores de bloque, lo cual implica que a la hora de correr el experimento no hay ningún margen de aleatorización. Es decir, se puede correr por columna o por renglón según convenga.
- No es correcto hacer todas las pruebas de un tratamiento, luego todas las de otro, y así sucesivamente, puesto que se puede introducir ruido adicional debido a factores no controlables que cambian con el tiempo.

Diseño de Cuadrado Grecolatino

Continuación

En **R**:

```
library(agricolae)
Tratamiento <- c("A","B","C","D")
outdesign     <- design.lsd(Tratamiento,serie=2,seed=23)
lsd  <- outdesign$book
print(outdesign$sketch)
     [,1] [,2] [,3] [,4]
[1,] "C"  "A"  "B"  "D"
[2,] "B"  "D"  "A"  "C"
[3,] "A"  "C"  "D"  "B"
[4,] "D"  "B"  "C"  "A"
```

Diseño de Cuadrado Grecolatino

Caso del Cuadrado Grecolatino

Para el caso de un cuadrado Grecolatino, se pueden generar aleatoriamente dos cuadrados latinos y luego se superponen. En **R**:

```
T1<-c("A","B","C","D")
T2<-c("G1","G2","G3","G4")
outdesign <- design.graeco(T1,T2,serie=1)
plots <-as.numeric(graeco[,1])
print(outdesign$sketch)
#     [,1]   [,2]   [,3]   [,4]
#[1,] "A G1" "C G2" "D G4" "B G3"
#[2,] "C G4" "A G3" "B G1" "D G2"
#[3,] "D G3" "B G4" "A G2" "C G1"
#[4,] "B G2" "D G1" "C G3" "A G4"
```