

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO

ESCUELA DE POSGRADO

MAESTRÍA EN ESTADÍSTICA



TRABAJO INDIVIDUAL 2

DISEÑO DE EXPERIMENTOS

AUTOR:

KEVIN HEBERTH HAQUEHUA APAZA

DOCENTE:

Mtro. ARTURO ZUÑIGA BLANCO

CUSCO - PERÚ

2025

ACTIVIDAD 1

TRABAJO DE INVESTIGACIÓN

Investigar

Investigar sobre uno de estos tres temas:

1. **DCA con submuestreo**
2. **Diseño Cuadrado Latino**
3. **Diseño Greco Latino**

Presente un informe académico de alguno de los temas elegidos, este informe debe tener fundamentos teóricos, aplicaciones y un ejercicio resuelto paso a paso usando las sumas de cuadrados (descomponiendo la variabilidad total), y corroborado en el R; Verificar los supuestos del modelo en el R, (utilizar transformaciones si fuera necesario) Así mismo si es pertinente usar pruebas de comparación múltiple.

Componentes mínimo del informe

- a) Presente el modelo aditivo lineal y explique sus componentes según el enunciado de la pregunta.
- b) Aplicaciones.
- c) Enunciado del ejemplo (cite o link)
- d) Anova (paso a paso)
- e) comprobación con el R
- f) Comprobación de supuestos, (Transformar si es necesario)
- g) Comparaciones multiples

Indicaciones finales

- a) El documento debe utilizar un lenguaje académico.
- b) Subir el informe final-documento en pdf.
- c) Subir el Script y data comprimido (se descontaran puntos si no lo hace).
- d) Subir tu informe antes de las 23:59 pm (18-09-2025)

1.1 SOLUCIÓN

Para realizar la resolución de esta activad se esta usando el editor de texto \LaTeX en el cual se hizo las debidas configuraciones en el preámbulo para obtener resultados en formato APA, me decidí por buscar el caso del diseño cuadrado latino.

1.1.1 *Diseño en cuadrados latinos*

Introducción

En este modelo se consideran simultáneamente dos variables de bloque (Universidad de Granada, Área de Bioestadística, sf). Por ejemplo considera un experimento en el que se quiere estudiar el efecto de distintos tipos de semilla en el rendimiento del trigo y se considera que en el rendimiento puede influir los tipos de abonos e insecticidas empleados. Se puede trabajar con un diseño completo aleatorizado donde el factor es el tipo de semilla y las variables de bloques son los tipos de abono e insecticida.

Pero su inconveniente es el de requerir excesivas unidades para su realización, para estos tres factores se tengan K_1 , K_2 y K_3 niveles en cada uno de los factores de lo que requeriría $K_1 \times K_2 \times K_3$ unidades experimentales por lo que pueden resultar muy costos. Entonces se recurre a un tipo especial de diseño en bloques incompletos aleatorizados en el que se selecciona una parte del diseño completo de tal forma que se puedan estimar los efectos de interés.

Uno de estos bloques es el modelo en cuadrado latino el cual requiere el mismo número de niveles para los tres factores.

Descripción del modelo

En estos diseños el número de niveles del factor principal tiene que coincidir con el número de niveles de las dos variables de bloque o factores secundarios y además hay que suponer que no existe interacción entre ninguna pareja de factores.

Suponiendo que el número de niveles de cada uno de los factores es K . El diseño cuadrado latino utiliza K^2 bloques, cada bloque corresponde a una de las posibles combinaciones de niveles de los dos factores de control. En cada bloque se aplica un solo tratamiento de manera que cada tratamiento debe aparecer con cada uno de los niveles de los dos factores de control.

Para el caso de abonos donde se tienen para la semilla, abono e insecticidad con $K = 4$ niveles se muestra la siguiente tabla:

	Abonos			
Insecticidas	a1	a2	a3	a4
i1	s4	s3	s1	s2
i2	s2	s1	s3	s4
i3	s3	s4	s2	s1
i4	s1	s2	s4	s3

De donde un diseño cuadrado latino tiene las siguientes características:

1. Se controlan tres fuentes de variabilidad, un factor principal y dos factores de bloque.
2. Cada uno de los factores tiene el mismo número de niveles, K .
3. Cada nivel del factor principal aparece una sola vez en cada fila y una vez en cada columna.
4. No hay interacción entre los factores.

Planteamiento del modelo

En un diseño en cuadrado latino intervienen los siguientes factores: un factor principal y dos factores secundarios o variables de bloque. Se supone que no existe interacción entre estos

tres factores, teniendo un modelo aditivo.

Si consideramos que los tres factores son de efectos fijos, el modelo estadístico para este diseño es:

$$y_{ijh} = \mu + \tau_i + \beta_j + \gamma_h + \epsilon_{ijh}$$

En el que $i = j = h = (1, 2, \dots, K)$ y donde:

- y_{ijh} : Representa la observación correspondiente de la i -ésima fila, j -ésima columna y h -ésima letra latina.
- μ : Es la media global.
- τ_i : Es el efecto producido por el i -ésimo nivel del factor fila. Los cuales están sujetos a la restricción $\sum \tau_i = 0$
- β_j : Es el efecto producido por el j -ésimo nivel del factor columna. Dichos efectos están sujetos a la restricción $\sum \beta_j = 0$
- γ_h : Es el efecto producido por la h -ésima letra latina. Dichos efectos están sujetos a la restricción $\sum \gamma_h = 0$
- ϵ_{ijh} : Son los errores que siguen una distribución $N \sim (0, \sigma)$

Ejemplo

Se ha realizado el experimento aleatorio correspondiente y se designó por las letras (A, B, C, D) a los tratamientos. Así, el cuadrado latino resultante junto con las observaciones obtenidas, dan lugar a la siguiente tabla

Abonos	Insecticidas				$y_{i..}$	$y_{i..}^2$
	i1	i2	i3	i4		
a1	C 7	D 8	B 4	A 3	22	484
a2	B 15	A 16	C 18	D 23	72	5184
a3	D 18	C 12	A 12	B 10	52	2704
a4	A 14	B 13	D 16	C 14	57	3249
$y_{.j.}$	54	49	50	50	203	11621
$y_{.j.}^2$	2916	2401	2500	2500	10317	
$\sum y_{ij(.)}^2$	794	633	740	834	3001	

Asi mismo la suma de cuadrados para el análisis de la varianza

Letra latina	Observaciones				$y_{..h}$	$y_{..h}^2$
A	3	16	12	14	45	2025
B	4	15	10	13	42	1764
C	7	18	12	14	51	2601
D	8	23	18	16	65	4225
					203	10615

$$SCT = \sum_{i=1}^4 \sum_{j=1}^4 y_{ij.}^2 - \frac{y_{...}^2}{K^2} = 3001 - \frac{203^2}{4^2} = 425,4375$$

$$SCF = \frac{1}{K} \sum_{i=1}^4 y_{i..}^2 - \frac{y_{...}^2}{K^2} = \frac{11621}{4} - \frac{203^2}{4^2} = 329,6875$$

$$SCC = \frac{1}{K} \sum_{j=1}^4 y_{.j.}^2 - \frac{y_{...}^2}{K^2} = \frac{10317}{4} - \frac{203^2}{4^2} = 3,6875$$

$$SCL = \frac{1}{K} \sum_{h=1}^4 y_{..h}^2 - \frac{y_{...}^2}{K^2} = \frac{10615}{4} - \frac{203^2}{4^2} = 78,1875$$

Y la suma de cuadrados del error se obtiene por diferencia

$$SCR = SCT - SCF - SCC - SCL = 13,875$$

En el que la tabla ANOVA seria el siguiente

Fuentes de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F_{exp}	% Explicado
E. fila (Abonos)	329.6875	3	109.89583	47.523	77.49 %
E. column. (Insect.)	3.6875	3	1.22917	0.532	0.87 %
E. trat. (Semillas)	78.1875	3	26.06250	11.270	18.38 %
Residual	13.8750	6	2.31250		
TOTAL	425.4375	15			96.74 %

Realizando el contraste al 5 % y tomamos el F_{cal} con el F_{teo} . Se concluye que los efectos de los abonos y semillas son significativas, pero no de los insecticidas.

Por última veamos la aplicación de los resultados en RStudio.

```
library(readxl)
semilla <- read_excel("data_semillas.xlsx")

# Realizamos el modelo
ajuste <- aov(y ~ tratamiento + fila + columna, data = semilla)

# Para mostrar la tabla ANOVA con anova()
anova(ajuste)

## Analysis of Variance Table
##
## Response: y
##          Df Sum Sq Mean Sq F value    Pr(>F)
## tratamiento  3   78.19   26.063  11.2703 0.0070517 **
## fila         3 329.69  109.896  47.5225 0.0001419 ***
## columna      3    3.69    1.229   0.5315 0.6771834
## Residuals    6   13.88    2.313
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

BIBLIOGRAFÍA

Universidad de Granada, Área de Bioestadística (s.f.). Capítulo 7: Diseños en cuadrados latinos.

<https://wpd.ugr.es/~bioestad/wp-content/uploads/Latinos.pdf>. Material docente, Guía de R.