Diseño de Bloques Completamente al Azar (DBCA)

Luis Fernando Delgado Muñoz - Ing. Agroindustrial, M.Sc

lfdelgadom@unal.edu.co

Universidad Nacional de Colombia Facultad de Ingeniería y Administración Departamento de Ingenieria

UNIVERSIDAD NACIONAL DE COLOMBIA

Contenido

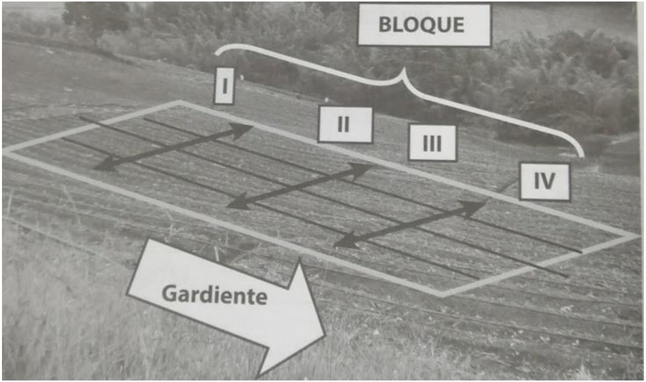
Diseño de Bloques Completamente al Azar (DBCA)

La idea surge cuando, al tratar de realizar el proceso de aleatorización, no es posible hacerse de forma completa si no restringida.

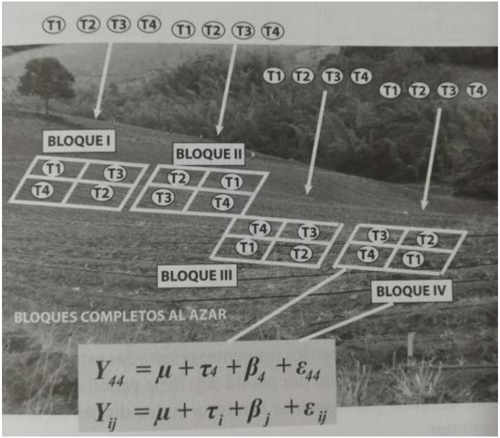
Factor Bloque

Son las Variables adicionales al factor de interés que se incorporan de manera explícita en un experimento comparativo para lograr mayor precisión en la comparación.

- Se dispone de un lote experimental en una zona de ladera con pendiente de 10%.
- Gradiente de fertilidad en el sentido de la pendiente.
- Bloqueo perpendicular al gradiente.



Cuatro balotas (1 por tratamiento), se lanzan 1 vez en cada bloque, de tal manera que una y solo una vaya a cada parcela.



Los bloques no necesariamente deben tienen que estar adyacentes

Ejemplo

Una empresa de alimentos está estudiando el efecto de 4 temperaturas sobre el rendimiento del almidón de yuca (en gramos). Cada evaluación tarda aproximadamente una hora y media, por lo que solamente se pueden realizar 4 evaluaciones por día.

Temperatura(°C)	Día 1	Día 2	Día 3	Día 4
30	14.0	14.1	14.5	14.0
35	13.9	13.8	14.2	14.0
40	14.1	14.2	14.4	13.9
45	13.0	13.1	13.5	13.2

La idea general consiste en realizar una repetición completa del experimento en cada uno de los niveles del factor de control (factor bloque) y se realiza la aleatorización de los tratamientos dentro de cada bloque.

El modelo

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad i = 1, \cdots, t \qquad j = 1, \cdots, b$$

Y_{ij} Es la observación de la variable de respuesta debida al i − esimo tratamiento en el j − esimo bloque.

μ Es la media global.

τ_i Es el efecto del i − esimo tratamiento **sobre** la variable de respuesta.

β_j Es el efecto del j − esimo bloque **sobre** la variable de respuesta.

ε_{ij} Es el error aleatorio debida al i − esimo tratamiento en el j − esimo bloque.

El modelo

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad i = 1, \cdots, t \qquad j = 1, \cdots, b$$

Y_{ij} Es la observación de la variable de respuesta debida al i − esimo tratamiento en el j − esimo bloque.

μ Es la media global.

τ_i Es el efecto del i − esimo tratamiento **sobre** la variable de respuesta.

β_j Es el efecto del j − esimo bloque **sobre** la variable de respuesta.

ε_{ij} Es el error aleatorio debida al i − esimo tratamiento en el j − esimo bloque.

El modelo

Supuestos

1. Correcta especificación del modelo $E(\varepsilon_{ij}) = 0$.
2. Homogeneidad de Varianza (Homocedasticidad) $V(\varepsilon_{ij}) = \sigma^2$
3. Independencia $Cov(\varepsilon_{ij}, \varepsilon_{i'j'}) = 0$
4. Normalidad $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Hipótesis

1. Sobre los Tratamientos

$$H_o : \tau_1 = \tau_2 = \cdots = \tau_t = 0$$
$$H_a : \tau_i \neq 0 \text{ para algún } i = 1, 2, \cdots, t.$$

2. Sobre los Bloques

$$H_o : \beta_1 = \beta_2 = \cdots = \beta_b = 0$$
$$H_a : \beta_j \neq 0 \text{ para algún } j = 1, 2, \cdots, b.$$

Análisis de Varianza

En este caso,

$$SC_T = SC_\tau + SC_\beta + SC_\varepsilon$$

Variabilidad = Total   Variabilidad debida a los tratamientos   + Variabilidad debida a los bloques   + Variabilidad debida al error aleatorio

Tabla del ANOVA

$$H_o : \tau_1 = \tau_2 = \cdots = \tau_t = 0$$

Fuente de Variabilidad	Grados de Libertad	Suma de Cuadrados	Cuadrados Medios	F_c	Valor p
Tratamientos	$t-1$	$b\sum_{i=1}^{t}(\bar{Y}_{i.}-\bar{Y}_{..})^2$	$CM_\tau = \frac{SC_\tau}{t-1}$	$\frac{CM_\tau}{CM_\varepsilon}$	$\Pr(F > F_c)$
Bloque	$b-1$	$t\sum_{j=1}^{b}(\bar{Y}_{.j}-\bar{Y}_{..})^2$	$CM_\beta = \frac{SC_\beta}{b-1}$	$\frac{CM_\beta}{CM_\varepsilon}$	$\Pr(F > F_c)$
Error	$(t-1)(b-1)$	$\sum_{i=1}^{t}\sum_{j=1}^{b}(Y_{ij}-\bar{Y}_{i.}-\bar{Y}_{.j}+\bar{Y}_{..})^2$	$CM_\varepsilon = \frac{SC_\varepsilon}{(t-1)(b-1)}$		
Total	$tb-1$	$\sum_{i=1}^{t}\sum_{j=1}^{b}(Y_{ij}-\bar{Y}_{..})^2$			

Verificación de supuestos

La verificación de los supuestos puede realizarse fácilmente mediante los residuales

$$e_{ij} = Y_{ij} - \hat{Y}_{ij}$$

donde Y_{ij} es el valor observado y $\hat{Y}_{ij}$ es el valor estimado con el modelo.

Ejemplo

Temperatura(ºC)	Día 1	Día 2	Día 3	Día 4	Total Yi.	Media $\bar{Y}i.$
30	14	14.1	14.5	14	56.6	14.15
35	13.9	13.8	14.2	14	55.9	13.975
40	14.1	14.2	14.4	13.9	56.6	14.15
45	13	13.1	13.5	13.2	52.8	13.2
Total Y.j	55	55.2	56.6	55.1	**Y.. = 221.9**	
Media $\bar{Y}.j$	13.75	13.8	14.15	13.775		**$\bar{Y}..$ = 13.86**

Suma de cuadrados de la media o factor de corrección

$$FC = \frac{(Y..)^2}{r \text{ x } t}$$

$$FC = \frac{(221.9)^2}{4 \text{ x } 4} = 3077.47$$

Suma de cuadrados total corregida

$$SCT(c) = \sum_{i=1}^{4}\sum_{j=1}^{4} Y_{IJ}^2 - FC$$

$$SCT(c) = (14^2 + 14.1^2 + \ldots\ldots\ldots\ldots 13.5^2 + 13.2^2) - 3077.47$$

$$\mathrm{SCT(c)} = 3$$

Suma cuadrados de los tratamientos

$$SCTR = \frac{\sum_{i=1}^{4} Y_{i.}^{2}}{B} - FC$$

$$SCTR = \frac{(56.6^2 + 55.9^2 + 56.6^2 + 52.8^{2)}}{4} - 3077.47$$

$$SCTR = 2.47$$

Suma cuadrados de los bloques (días)

$$SCB = \frac{\sum_{i=1}^{4} Y_{i.}^{2}}{Trat} - FC$$

$$SCB = \frac{(55^2 + 55.2^2 + 56.6^2 + 55.1^{2)}}{4} - 3077.47$$

$$SCB = 0.43$$

Suma cuadrados del error experi

$$SCEE = SCT(c) - SCTR - SCB$$

$$SCEE = 3 - 2.47 - 0.43$$

$$SCEE = 0.1$$

Análisis de varianza para la variable "Rendimiento de yuca" de acuerdo con el DBCA

Fuente de variabilidad	Grados de libertad	Sumas de cuadrados	Cuadrados medios	Fc	Ft
Tratamientos	3	2.47	0.823	74.1	3.86
Bloques	3	0.43	0.143	12.9	3.86
Error	9	0.1	0.011		
Total (corregido)	15	3			

$$Fc(Bloques) = 14 \; > Ft^{4,9}_{0.05} = 3.86$$

GL (TTO y Bloque) = 3

GL error = 9

$\boxed{3.86}$

No se acepta la hipótesis nula y se llega a la conclusión de que el criterio de bloqueo fue acertado.

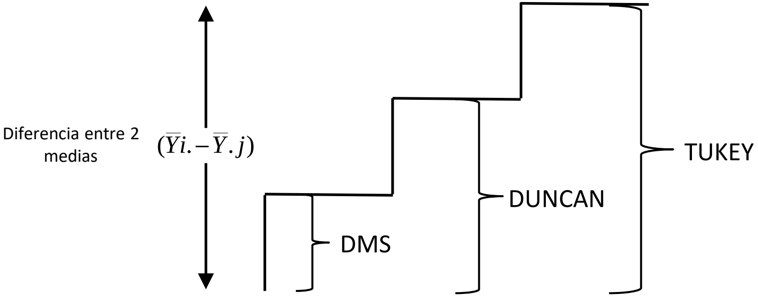
Conclusión

Al remover de la SC de cuadrados del error experimental la variación debida al bloqueo, se mejora la precisión del experimento, Siempre que el sacrificio en grados de libertad del error se compense con una suficiente reducción de la SCEE

PRUEBAS POST ANOVA

Pruebas post ANOVA

La estadística ofrece un menú de pruebas post anova desde las “muy liberales” como la DMS, hasta las “muy conservadoras” como la prueba de Tukey. Algunas son de rango “único” como la DMS, Tukey y Dunnett, otras de rango “multiple” como la de Duncan y la de Student-Newman –Keuls.



Prueba de la diferencia mínima significativa DMS

La DMS (LSD en ingles)se define como la diferencia mínima que debe existir entre dos medias poblacionales $(\mu_i - \mu_j)$ para considerarlas estadísticamente diferentes. Para que esto ocurra la diferencia entre ellos debe ser o igual que la DMS.

$$DMS = t_{\frac{\alpha}{2}} \sqrt{\frac{2CME}{r}}$$

$t_{\frac{\alpha}{2}}$ Valor en la distribución " t", se busca con los grados de libertad del error, en una tabla de dos colas.

$$DMS_{a,b} = t_{\frac{\alpha}{2}} \sqrt{\frac{CME}{ra} + \frac{CME}{rb}}$$

Diferente número de repeticiones por tratamiento (no balanceado)

DMS para tratamientos: experimento "Rendimiento de yuca"

CME = 0.01
r= Bloques = 4
GLE = 9
$t_{\frac{\alpha}{2}}$ = buscado en la tabla "t" = 2.262 con un nivel de significancia del 5%

$$DMS = 2.262\sqrt{\frac{2(0.01)}{4}}$$

$$DMS = 0.159\ g$$

Se ordenan los promedios en forma descendente

$$30ºC = 14.15$$
$$40ºC = 14.15$$
$$35ºC = 13.97$$
$$45ºC = 13.2$$

$$Numero \text{ de comparaciones} = \frac{t^2 - t}{2}$$

$$Numero \text{ de comparaciones} = \frac{4^2 - 4}{2}$$

$$Numero \text{ de comparaciones} = 6$$

$$DMS = 0.159\ g$$

$$1 - 3 = 14.15 - 14.15$$
$$= 0\ NS$$
$$1 - 2 = 14.15 - 13.97$$
$$= 0.53 *$$
$$1 - 4 = 1.3 *$$
$$3 - 2 = 0.53 *$$
$$3 - 4 = 1.3 *$$
$$2 - 4 = 0.77 *$$

$30ºC = 14.15\ a$
$40ºC = 14.15\ a$
$35ºC = 13.97\ b$
$45ºC = 13.2\ c$

Prueba de Tukey tratamientos: experimento "Rendimiento de yuca"

La prueba Tukey, también denominada Diferencia Honesta Significativa (DHS) es una prueba de diferencia única como la DMS. Con los grados de libertad del error y el numero de tratamientos del experimento. Se busca en la tabla Tukey el valor de la amplitud estudentizada. Se calcula el valor DHS mediante la siguiente expresión:

$$DHS = Q_{\alpha}\sqrt{\frac{CMEE}{r}}$$

Q_{α} = buscado en la tabla "Tukey" = 4.41 con un nivel de significancia del 5%

$$DHS = 4.41\sqrt{\frac{0.01}{4}}$$

$$DHS = 0.220$$

$$30ºC = 14.15\ a$$
$$40ºC = 14.15\ a$$
$$35ºC = 13.97\ a$$
$$45ºC = 13.2\ b$$

Prueba de DUNCAN tratamientos: experimento "Rendimiento de yuca"

Error estandar de la media.

$$EE = \sqrt{\frac{CME}{r}}$$

$$EE = \sqrt{\frac{0.01}{4}}$$

$$EE = 0.05$$

Valores Duncan para comparar medias, con 9 GLE y 4 tratamientos. Ver Tabla (Duncan).

Grados de libertad del error	# de Tratamientos			
	2	3	4	5
1				
2				
.				
.				
.				
9	3.199	3.339	3.42	3.47

Rangos críticos de Duncan

Se efectúa el producto entre el error estándar de la media y los rangos críticos de Duncan

	# de Tratamientos		
Grados de libertad del error	**2**	**3**	**4**
9	3.199	3.339	3.42
Error estandar de la media	0.05	0.05	0.05
EE*Rangos Duncan	**0.160**	**0.167**	**0.171**

$30ºC = 14.15\ a$
$40ºC = 14.15\ a$
$35ºC = 13.97\ a$
$45ºC = 13.2\ b$

MUCHAS GRACIAS