

Diseño de bloques al azar



Diseño de bloques al azar

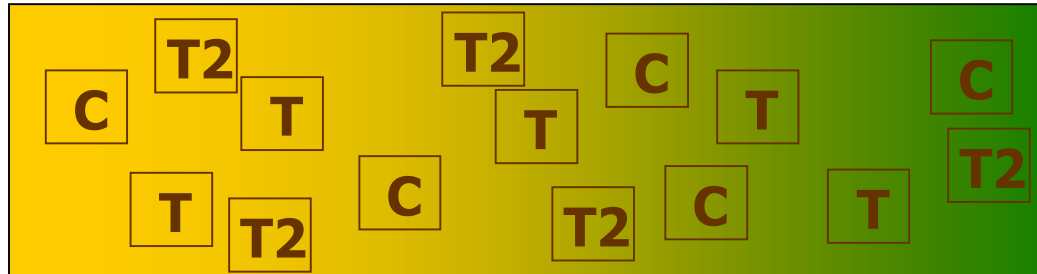
¿Qué es un bloque?

Es una o varias unidades experimentales homogéneas:

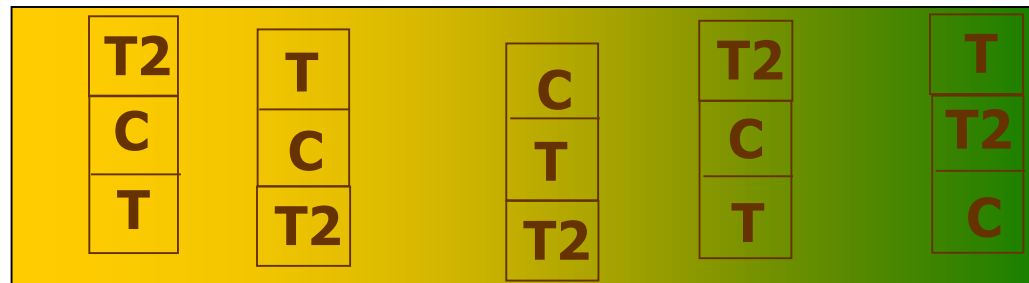
- ▣ Una parcela que se divide en subparcelas
- ▣ Una camada compuesta por varios individuos
- ▣ Un mismo individuo que se puede tratar por más de un método o proporcionar más de un dato

Y dentro de cada bloque la asignación de los tratamientos a las u.e. es aleatoria

¿Mejora la productividad de una plantación de trigo si se aplica un fungicida para el control de la roya?



Diseño completamente aleatorizado (DCA)



Unidad experimental:

Réplicas:

Diseño de bloques al azar (DBA)

¿Con qué criterio bloquear?

- ❑ Utilizando cualquier factor que afecte la variable respuesta y que varíe entre las unidades experimentales
- ❑ No debería interesar efectuar comparaciones entre niveles del factor bloque
- ❑ El bloque debe ser internamente homogéneo
- ❑ El criterio debe ser tal que minimice la variabilidad entre las subparcelas dentro de un mismo bloque (control del error)
- ❑ La prueba t para muestras pareadas es equivalente a DBA con 2 tratamientos

Buenos criterios para bloquear

- ❑ Proximidad (parcelas adyacentes)
- ❑ Características físicas (edad, peso)
- ❑ Características genéticas (parientes, camadas, un mismo individuo)
- ❑ Tiempo

DBA

Consiste en descomponer la variabilidad de la variable respuesta en:

- ▣ Variabilidad explicada o debida a los tratamientos (variabilidad **entre** tratamientos)
- ▣ Variabilidad explicada o debida a las parcelas (variabilidad **entre** bloques)
- ▣ Variabilidad no explicada por los tratamientos o variabilidad aleatoria o residual (variabilidad **dentro**)

El modelo estadístico es:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

- ▣ donde μ es la media general o media de la población
- ▣ α_i es el efecto del tratamiento i y que es común a todos los individuos que recibieron ese tratamiento
- ▣ β_j es el efecto del bloque j y que es común a todos los tratamientos que se aplicaron en ese bloque
- ▣ ε_{ij} es el residuo o error aleatorio (dentro)

Hipótesis en DBA

- A** ☐ $H_0: \mu_1 = \mu_2 = \mu_3 = \mu$
es decir no hay efecto de los tratamientos sobre la variable respuesta
- ☐ $H_1: \text{alguna } \mu_i \neq \mu$
es decir al menos un tratamiento tiene efecto sobre la variable respuesta
- B** ☐ $H_0: \sigma^2_{\text{bloques}} = 0$
es decir no hay variabilidad debida a los bloques
- ☐ $H_1: \sigma^2_{\text{bloques}} \neq 0$
es decir al hay variabilidad debida a los bloques

Tabla de ANOVA

Fuente de Variación	SC	GL	CM	F
A Entre Tratamientos	$b\sum(\bar{y}_i - \bar{y})^2$	a-1	$\frac{SC_A}{GL_A}$	$\frac{CM_A}{CMd}$
B Entre Bloques	$a\sum(\bar{y}_j - \bar{y})^2$	b-1	$\frac{SC_B}{GL_B}$	$\frac{CM_B}{CMd}$
Dentro de tratamientos	Por diferencia	(a-1)x (b-1)	$\frac{SCd}{GLd}$	
Total	$\sum(y_{ij} - \bar{y})^2$	n-1		

Volvamos al ejemplo

Bloque		Control	Fungicida Dosis baja	Fungicida Dosis alta	Media Bloque
$a=3$ $b=5$ $n_i=1$ $n=15$	1	2.83	2.95	3.52	3.10
	2	2.62	2.88	3.47	2.99
	3	3.44	3.55	4.20	3.73
	4	3.05	3.08	3.98	3.37
	5	2.44	2.54	3.60	2.86
Media Trat		2.876	3.000	3.754	3.21

Fuente de Variación	SC	GL	CM	F
A Entre Tratamientos	$b\Sigma(\bar{y}_i - \bar{\bar{y}})^2$	a-1	$\frac{SC_A}{GL_A}$	$\frac{CM_A}{CMd}$
B Entre Bloques	$a\Sigma(\bar{y}_j - \bar{\bar{y}})^2$	b-1	$\frac{SC_B}{GL_B}$	$\frac{CM_B}{CMd}$
Dentro de tratamientos	Por diferencia	(a-1)x (b-1)	$\frac{SCd}{GLd}$	
Total	$\Sigma(y_{ij} - \bar{\bar{y}})^2$	n-1		

Bloque		Control	Fungicida Dosis baja	Fungicida Dosis alta	Media Bloque
a=3	1	2.83	2.95	3.52	3.10
b=5	2	2.62	2.88	3.47	2.99
n _i =1	3	3.44	3.55	4.20	3.73
n=15	4	3.05	3.08	3.98	3.37
	5	2.44	2.54	3.60	2.86
Media Trat		2.876	3.000	3.754	3.21

En Infostat:

▣ Ingreso de datos



InfoStat/P - clase5

Archivo Edición Datos Resultados Estadístico

InfoStat clase5

caso	Tratamiento	Parcela	Rinde
1	control	1.00	2.83
2	control	2.00	2.62
3	control	3.00	3.44
4	control	4.00	3.05
5	control	5.00	2.44
6	fung baja	1.00	2.95
7	fung baja	2.00	2.88
8	fung baja	3.00	3.55
9	fung baja	4.00	3.08
10	fung baja	5.00	2.54
11	fung alta	1.00	3.52
12	fung alta	2.00	3.47

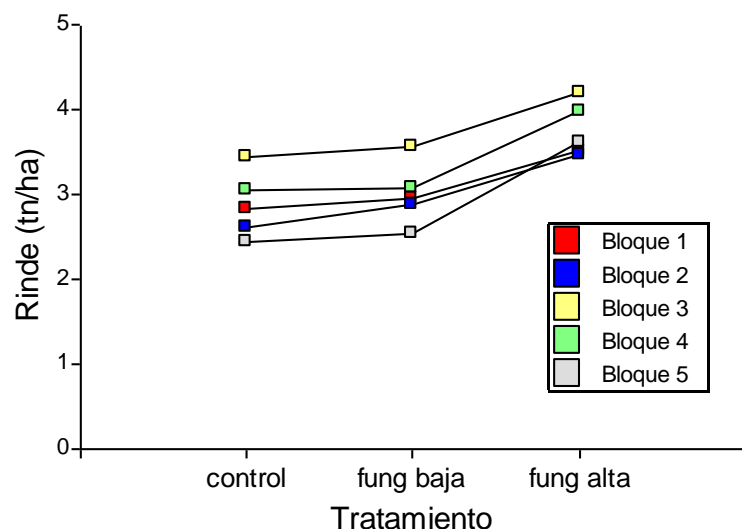
Anova:

Cuadro de Análisis de la Varianza (SC tipo III)

	F.V.	SC	gl	CM	F	p-valor
A	Modelo	3.69	6	0.62	42.60	<0.0001
	Tratamiento	2.26	2	1.13	78.10	<0.0001
B	Parcela	1.44	4	0.36	24.85	0.0001
	Error	0.12	8	0.01		
	Total	3.81	14			

Obsérvese que

SCA + SCB + SCd = SC total
GLA + GLB + GLd = GL total



Supuestos del modelo

- ❑ Los b bloques deben ser extraídos al **azar** y deben ser **independientes** entre sí.
- ❑ Los tratamientos deben ser asignados **al azar** dentro de cada bloque
- ❑ Los errores deben estar **normalmente** distribuidos
- ❑ Los bloques deben responder en forma paralela (**aditividad-paralelismo**)
- ❑ Los tratamientos deben tener la misma variabilidad (**homocedasticidad**)

¿Cómo detectamos si hay normalidad de los errores?

- ❑ **Gráficamente: QQ plot**
- ❑ **Analíticamente: Prueba de Shapiro Wilks**

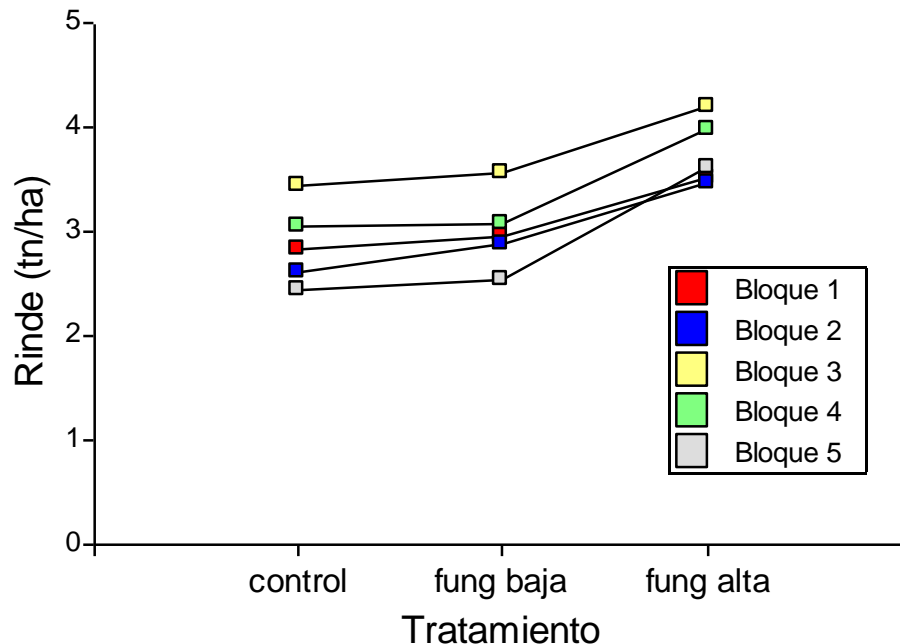
¿Cómo detectamos si hay paralelismo entre bloques?

□ **Gráficamente: Gráfico de perfiles**

En Infostat => Gráfico de puntos. Variable a graficar: rinde. Criterios de clasificación: Tratamiento. Partición: Bloque (Particiones en el mismo gráfico)

□ **Analíticamente: Prueba de no aditividad de Tukey**

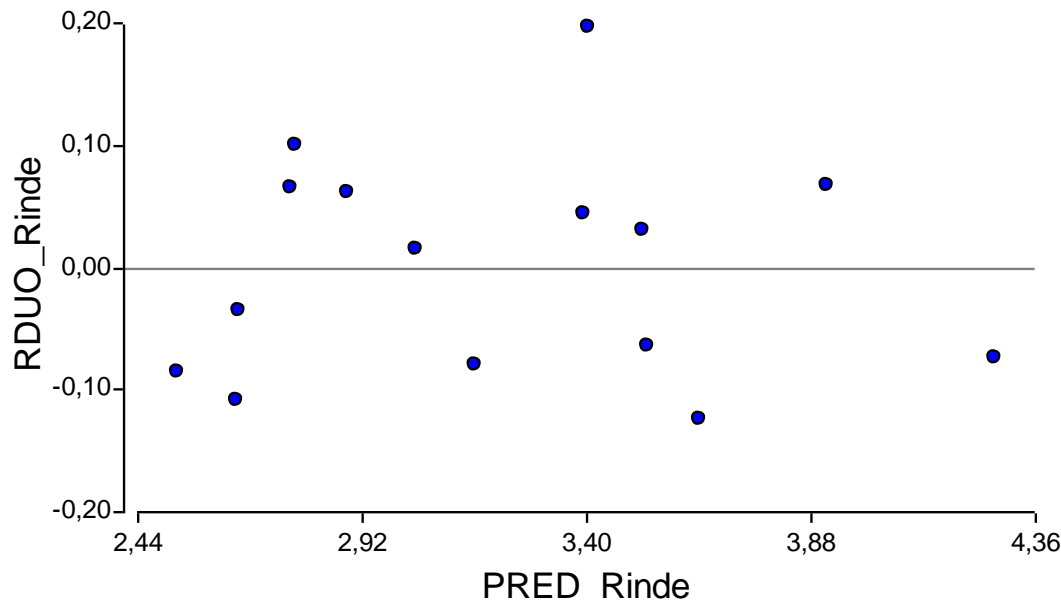
Gráfico de paralelismo



¿Cómo detectamos si hay heterocedasticidad?

- Se calculan los residuos y se grafican vs los valores esperados o predichos
- Se espera encontrar una distribución al azar y con variabilidad constante

Gráfico de dispersión



Comparaciones múltiples

- ▣ Tienen sentido solo para comparar tratamientos
- ▣ Es irrelevante comparar bloques
- ▣ El procedimiento es el mismo que en anova de un factor