

**INSTITUTO DE EDUCACIÓN TÉCNICA PROFESIONAL
DE ROLDANILLO, VALLE - INTEP**



**GUÍA PARA LA PRESENTACIÓN Y ANÁLISIS DE
DATOS DE LOS PROYECTOS DE
INVESTIGACIÓN EN LOS CICLOS TÉCNICO,
TECNOLÓGICO Y PROFESIONAL.**

**GESTIÓN DE INVESTIGACIÓN
OCTUBRE, 2009**



CONTENIDO

	Pág.
1. PRESENTACIÓN DE DATOS DE INVESTIGACIÓN - ANÁLISIS DESCRIPTIVO.	4
2. MEDIDAS DE TENDENCIA CENTRAL	8
3. DISEÑO EXPERIMENTAL O MUESTREO.	9
4. PLAN DE EJECUCIÓN DEL EXPERIMENTO CONSIDERACIONES	10
5. ERRORES MÁS FRECUENTES EN UN EXPERIMENTO	10
6. QUÉ SE DEBE TENER PRESENTE	10
7. DISEÑO EXPERIMENTAL	12
8. ANÁLISIS DE VARIANZA.	12
9. HIPÓTESIS	14
10. DISEÑOS EXPERIMENTALES BÁSICOS	14
a. DISEÑO COMPLETAMENTE AL AZAR	14
b. DISEÑO DE BLOQUES AL AZAR.	18



c. CARACTERÍSTICAS.	18
11. PRUEBAS DE COMPARACIÓN MÚLTIPLE	24
11.1 PRUEBA DE COMPARACIÓN MÚLTIPLE	24



1. PRESENTACIÓN DE DATOS DE INVESTIGACIÓN - ANÁLISIS DESCRIPTIVO

- Cuadros de resúmenes
- Representación Gráfica
 - De línea
 - Pictogramas
 - De barras
 - De cajas y alambres
 - Circulares
- Tablas de frecuencia: método que permite arreglar y ordenar los datos en intervalos.

Ejemplo 1.1

En un ensayo de agroquímicos para el control del pasador del tomate, se observa el número de frutos picados después de la aplicación de un tratamiento específico para su control; en el ensayo se tomaron 40 muestras de 10 frutos cada una y se determinaron los frutos con 0, 1, 2, 3, 4, 5 y más 6 larvas por fruto.

Datos del ensayo: 4, 2, 3, 5, 4, 2, 1, 6, 3, 6, 4, 5, 6, 6, 5, 4, 3, 5, 5, 3, 1, 6, 4, 3, 2, 3, 6, 2, 3, 2, 5, 5, 3, 5, 1, 2, 1, 4, 3

Tabla de distribución de frecuencias:

Clase	ni	hi	Ni	Hi
0	111	27,8%	111	27,8%
1	60	15,0%	171	42,8%
2	42	10,5%	213	53,3%
3	71	17,8%	284	71,0%
4	41	10,3%	325	81,3%
5	33	8,3%	358	89,5%
7	42	10,5%	400	100,0%
Σ	400	1,000		

Interpretación:

Frecuencia (ni) = 111: 111 de los 400 no presentan perforaciones ni larvas en los frutos



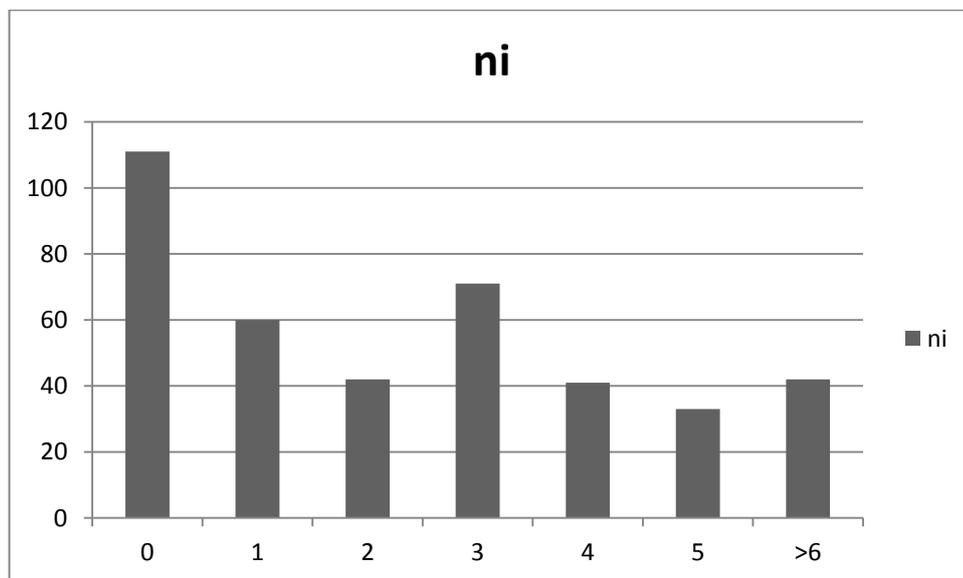
Porcentaje acumulado (H_1) = 0.278: el 27.8% de las muestras no presentan daño por el perforador

Porcentaje (h_1) = 0.103: el 10.3% de las muestras presentan 4 larvas por fruto del pasador

Frecuencia acumulada (N_3) = 284: de los 400 frutos, 284 tienen 3 o menos larvas por fruto del pasador

Porcentaje acumulado (H_5) = 0.895: el 89.5% de los 400 frutos tienen 5 o menos larvas por fruto del pasador

Histograma de Frecuencia.





19.0 22.2 24.8 23.6 24.0 20.9 23. 21.9 25.8 19.6 25.3
26.0
23.0 23.4 25.3 22.8 24.0 19.8 19.1 23.4 21.9 24.9 22.6
18.1

Número de Clases: $m=5$

Longitud del intervalo: C_1

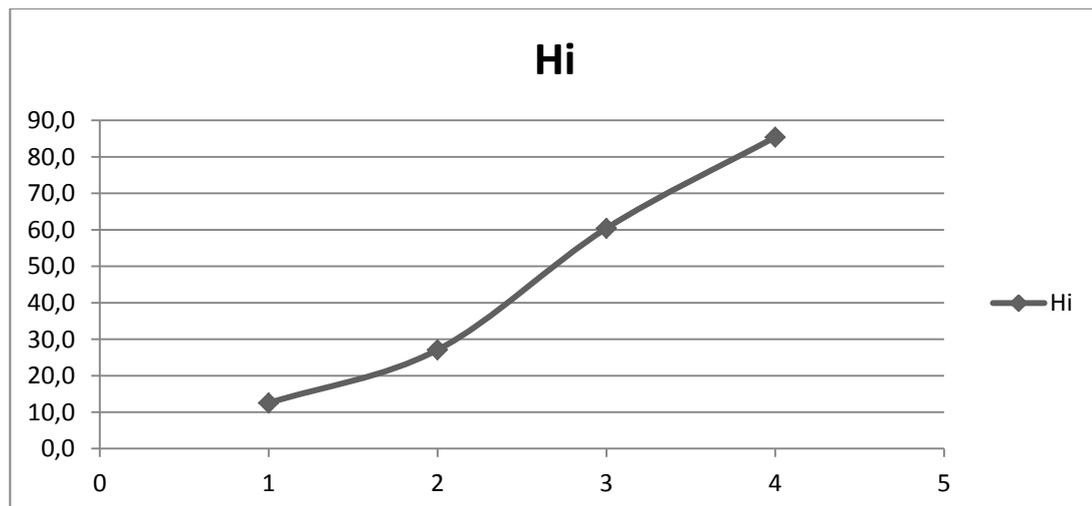
$C_1 = (\text{Valor Máximo} - \text{Valor Mínimo})/m$

$$C_1 = (26-18)/5 = 1.6$$

Tabla de Frecuencias

Intervalo de Clase	Marca de Clase	Frecuencia n_i	Porcentaje h_i	Frecuencia acumulada N_i	Porcentaje acumulado H_i
17,5 - 19,3	18,4	6	0,125	6	12,5
19,3 - 21,1	20,2	7	0,146	13	27,1
21,1 - 22,9	22	16	0,333	29	60,4
22,9 - 24,7	23,8	12	0,250	41	85,4
24,7 - 26,5	25,6	7	0,146	48	1,0
Σ		48	1,000		

Función de distribución acumulativa





2. MEDIDAS DE TENDENCIA CENTRAL

1. Media aritmética: $\bar{X} = \sum X_i/n$
2. La Mediana: Valor por debajo del cual se encuentran el 50% de los datos, no es sensible a valores extremos, Su cálculo no tiene en cuenta todos los datos.
3. Medidas de posición: indican la posición relativa de una calificación
 - Percentil 25 = Primer cuartil
 - Percentil 50 = Segundo cuartil = Mediana
 - Percentil 75 = Tercer cuartil
4. Cuartiles
 - Q25: El 25% de los datos están por debajo de este valor
 - Q50: El 50% de los datos están por debajo de este valor
 - Q75: El 75% de los datos están por debajo de este valor
5. Medidas de Dispersión: Sirven para describir la variabilidad de los datos, cuantificando su dispersión respecto de una medida de tendencia central.
 - Rango:
 $R = \text{Valor Máximo} - \text{Valor Mínimo}$
 - Rango Intercuartílico:
 $RI = Q_{75} - Q_{25}$
 - Varianza y Desviación Estándar: miden la dispersión de los datos con respecto a la media aritmética.

Varianza:

$$S^2 = \sum (x_1 - x_2)/(n-1)$$

Desviación Estándar:

$$S = \sqrt{S^2}$$

Coefficiente de Variación:

$$cv = s/x*100$$

Permite comparar la variabilidad para dos conjuntos de datos ya que éste es un indicador a dimensional.



3. DISEÑO EXPERIMENTAL O MUESTREO

Se debe especificar si el estudio o investigación involucra un diseño experimental en cuyo caso se deben plantear varias opciones y seleccionar el más conveniente.

Si el estudio involucra una técnica de muestreo, se debe definir la estrategia para la selección de las unidades.

Unidad experimental y repeticiones

UNIDAD EXPERIMENTAL. Es la mínima cantidad de material experimental, al cual se le aplica un tratamiento.

REPETICIONES. Indica el número de unidades experimentales a las cuales se les va a aplicar un tratamiento específico.

¿CUANTAS REPETICIONES?

- Grado de precisión requerido
- Heterogeneidad del material
- Número de tratamientos
- El diseño experimental
- Tamaño de parcela

Funciones

- Estimación del error experimental
- Mejorar la precisión del experimento
- Ampliar el área de inferencia

Datos a ser colectados – Situaciones:

- Recolección insuficiente de datos
- Recolección de datos en exceso

Consideraciones Generales

- Contar con los instrumentos adecuados
- Evitar preferencias por un tratamiento
- Supervisar la toma de datos
- No centrar el interés en una variable
- Anotar observaciones ordenadamente



- Muestrear la unidad si es necesario
- Visitar el campo experimental

Análisis Estadístico:

Antes de realizar el experimento se debe tener en cuenta que tipo de análisis estadístico se realizará, así como los posibles resultados a los que se puede apuntar.

4. PLAN DE EJECUCIÓN DEL EXPERIMENTO CONSIDERACIONES

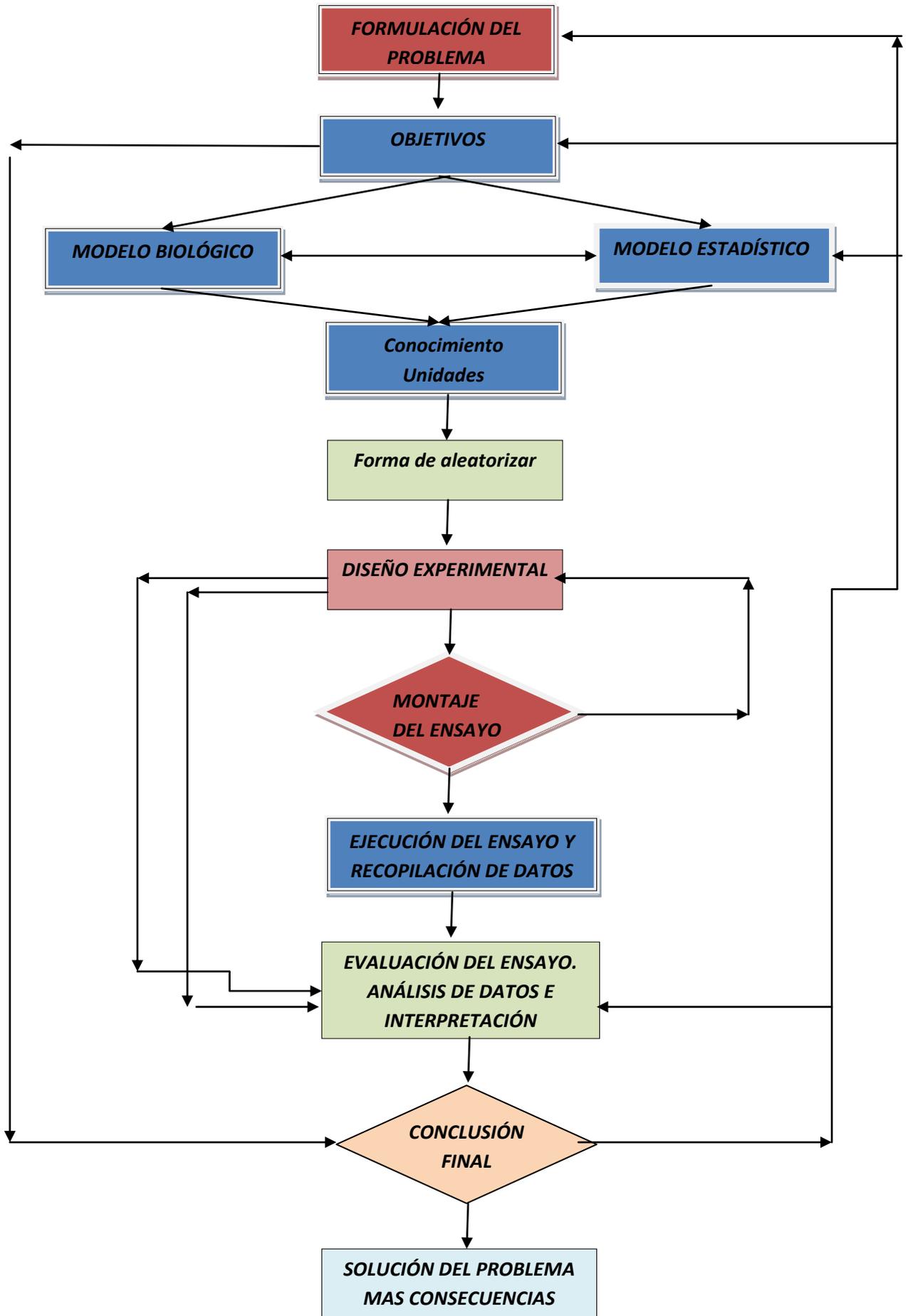
- Realizarlo de acuerdo a lo planeado en el anteproyecto
- Tener uniformidad durante la realización
- Aplicar los tratamientos en las parcelas previamente indicadas
- Hacer una buena planeación de las labores de campo y prácticas

5. ERRORES MÁS FRECUENTES EN UN EXPERIMENTO

- Incorrecta definición del problema
- Objetivos muy ambiciosos
- Gran número de tratamientos
- Material experimental poco homogéneo
- Diseño experimental inadecuado
- Tamaño y forma de las parcelas
- Interés en una sola variable
- Mala asignación de los tratamientos
- No estandarizar las unidades de medida
- Archivar y tomar de manera desordenada los datos del experimento.

6. QUÉ SE DEBE TENER PRESENTE

- Dedicar el suficiente tiempo a la realización de una buena planeación.
- Solo se obtienen buenos resultados de un experimento si cumple con los objetivos y si la ejecución es cuidadosa.
- El conocimiento de la estadística es una herramienta valiosa desde la planeación hasta la obtención de los resultados
- La investigación termina cuando se dan a conocer los resultados.



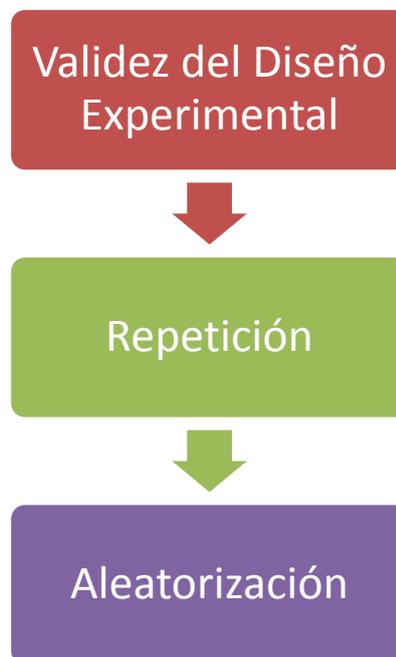


7. DISEÑO EXPERIMENTAL

El diseño experimental como parte del proceso de planeación está relacionado con la forma de asignar los tratamientos a las unidades experimentales a partir de unos factores especificados.

Objetivo del Diseño Experimental

Suministrar el máximo de información relevante a un problema bajo estudio y aun mínimo costo.



8. ANÁLISIS DE VARIANZA.

Particiona la varianza total en variación debido a factores controlados y en variación debido a factores no controlados, con el objetivo de inferir acerca de la no igualdad de los efectos



ocasionados por los



Objetivos

- Descomponer la variabilidad total de la característica de interés.
- Detectar diferencias entre medias de tratamientos

Estructura de la tabla de análisis de varianza

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
A	glA	SCA	CMA	FA	Pr(FA)
B	glB	SCB	CMB	FB	Pr(FB)
•	•	•	•	•	•
•	•	•	•	•	•
•	•	•	•	•	•
Error	glError	SCError	CMEror		
TOTAL	glTotal	SCTotal			



9. HIPÓTESIS

Hipótesis de investigación: Los tratamientos presentan efectos diferentes sobre la variable respuesta

Hipótesis Nula: El efecto de los tratamientos sobre la variable respuesta es el mismo

Hipótesis Alternativa: Al menos uno de los tratamientos presenta un efecto sobre la variable respuesta, diferente a los demás

10. DISEÑOS EXPERIMENTALES BÁSICOS

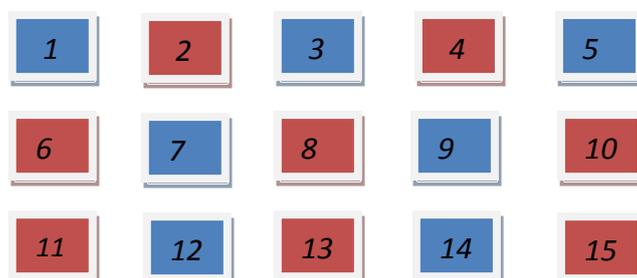
a. DISEÑO COMPLETAMENTE AL AZAR

Características

- Exige unidades experimentales homogéneas
- El número de tratamientos y repeticiones solamente se limita por las unidades experimentales disponibles.
- El análisis estadístico es sencillo aún si el número de repeticiones por tratamiento es diferente

Asignación de tratamientos

Un experimento con cinco tratamientos cada uno tres repeticiones para un total de 15 unidades experimentales.





Tratamiento	Rep. 1	Rep. 2	Rep. 3
A	a	a	a
B	b	b	b
C	c	c	c
D	d	d	d
E	e	e	e

d	d	b	e	c
c	b	a	c	a
d	b	e	a	e

Modelo Estadístico

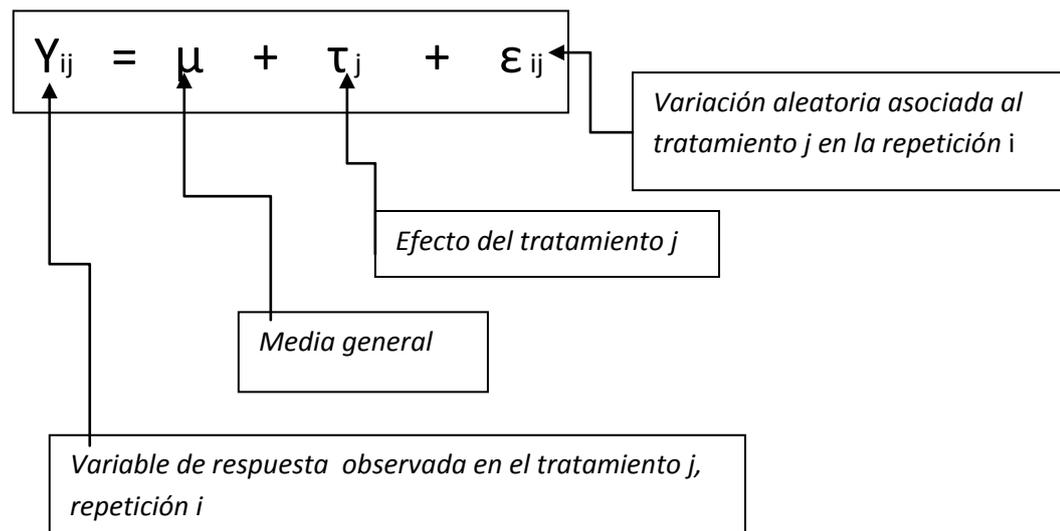




Tabla de Análisis de Varianza

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
Tratamiento	t-1	SC tratamiento	CM tratamiento	F tratamiento	Proba >F
Error	t(r-1)	SC error	CM error		
Total	rt-1	SC total			

Organización de los Datos

		Tratamientos					
		1	2	3	...	t	TOTAL
Repeticiones	1	y_{11}	y_{12}	y_{13}	...	y_{1j}	$y_{1\cdot}$
	2	y_{21}	y_{22}	y_{23}	...	y_{2j}	$y_{2\cdot}$
	3	y_{31}	y_{32}	y_{33}	...	y_{3j}	$y_{3\cdot}$
		
		
		
	r	y_{r1}	y_{r2}	y_{r3}	...	y_{rj}	$y_{r\cdot}$
TOTAL		$y_{\cdot 1}$	$y_{\cdot 2}$	$y_{\cdot 3}$...	$y_{\cdot j}$	$y_{\cdot \cdot}$

Cálculos necesarios para la tabla de análisis de varianza

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
Tratamiento	t-1	$\sum y_{\cdot j}^2 - FC$	SCtrat/t-1	CMtrat/CM error	Proba >F
Error	t(r-1)	Diferencia	SCerror/t-1		
Total	rt-1	$\sum \sum y_{ij}^2 - FC$			

Factor de corrección



$$FC = y..^2/rt$$

Ejemplo: En un ensayo se determinó el efecto de cuatro productos en el control del gusano pasador en el cultivo de lulo; se tomaron 100 frutos como unidad experimental, sobre los cuales se observaron, registrándose el número de frutos afectados por larvas. Los datos obtenidos para los cuatro productos y un testigo se presentan a continuación:

PRODUCTO	REPETICIÓN 1	REPETECIÓN 2	REPETICIÓN 3
Testigo	10	15	12
B	7	5	3
C	7	5	5
D	3	5	7
E	2	5	7

Hipótesis planteadas:

- Existen diferencias por el efecto del tratamiento entre el testigo y los cuatro productos de prueba.
- Existen diferencias entre los cuatro productos de prueba para el control de pasador del cultivo del lulo.

Organización de los datos:

	Testigo	B	C	D	E	Yi.
rep 1	10	7	7	3	2	29
rep 2	15	5	5	5	5	35
rep 3	12	3	5	7	7	34
Y.j	37	15	17	15	14	98

Pasos para el análisis de varianza

Paso 1: Calcular el factor de corrección

$$FC = (\sum\sum y_{ij})^2/r*t = (98)^2/5*3 = 640.27$$

Paso 2: Suma de cuadrados total

$$SCT = \sum\sum y_{ij}^2 - FC = (10^2+15^2+12^2+7^2+5^2+3^2+7^2+5^2+5^2+3^2+5^2+7^2+2^2+5^2+7^2) - 640.27 = 171.73$$

Paso 3: Suma de cuadrados de tratamientos



$$SC_{\text{trata.}} = \sum Y.j/r - FC = (37^2+15^2+17^2+15^2+14^2)/3 - 640.27 = 127.73$$

Tabla de Análisis de Varianza:

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
Tratamiento	t-1 = 4	127.73	31.93	7.26	3.48
Error	t(r-1) = 10	44.00	4.40		
Total	rt-1 = 14	171.73			

b. DISEÑO DE BLOQUES AL AZAR.

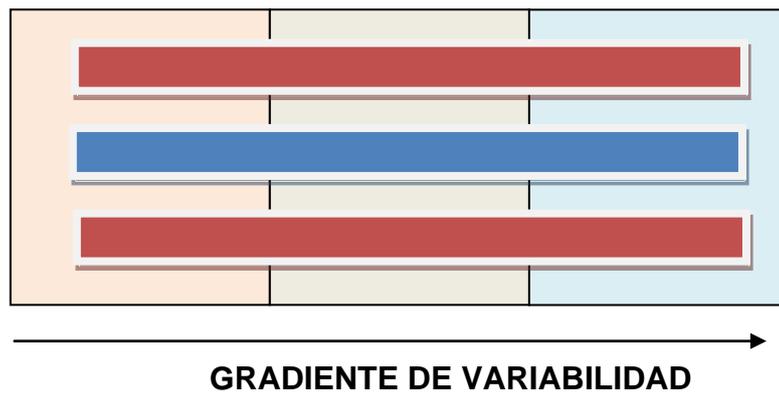
c. CARACTERÍSTICAS

- Permite el uso de unidades experimentales heterogéneas pero posibles de agrupar en bloques homogéneos.
- Su uso eficiente requiere que las unidades experimentales que integran el bloque, sean lo más homogéneas posibles (variación aleatoria) dejando la mayor heterogeneidad (variación sistemática) entre bloques.
- Cada bloque debe contener todos los tratamientos, los cuales se asignan al azar en forma independiente para cada bloque; en este diseño el término bloque es sinónimo de repetición.
- La mayor homogeneidad se consigue con un número bajo de tratamientos menos de diez tratamientos por bloque.
- El modelo asociado al diseño supone aditividad en el efecto de bloque y de tratamientos.
- No es necesario que los bloques queden físicamente juntos, ni que las parcelas que los conforman estén adyacentes; lo importante es que las parcelas que conforman el bloque tengan las mismas condiciones.
- Si se puede determinar un gradiente definido, los bloques se deben ubicar en el sentido contrario a este.
- La variabilidad entre bloques se elimina para comparar los tratamientos, reduciendo así el error experimental

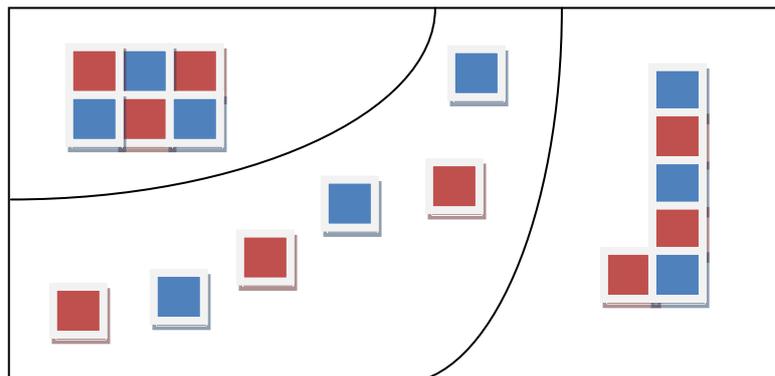
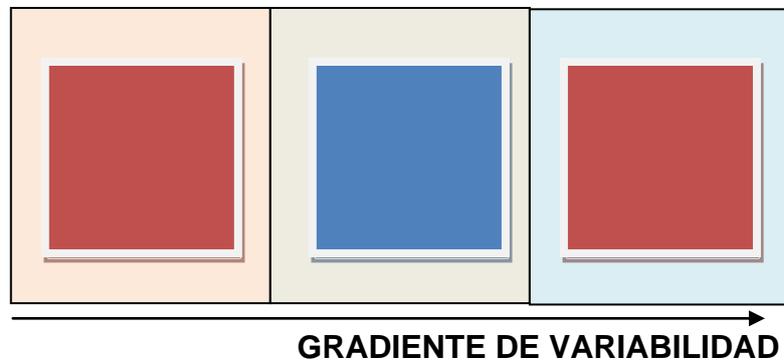


- El análisis es sencillo si el número de repeticiones es el mismo para todos los tratamientos
- Todos los tratamientos se comparan con la misma precisión.

Bloqueo inadecuado

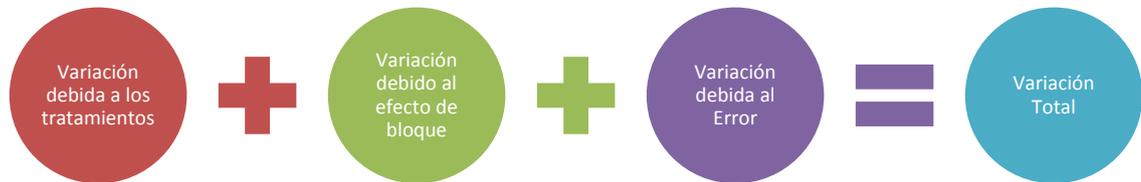


Bloqueo adecuado



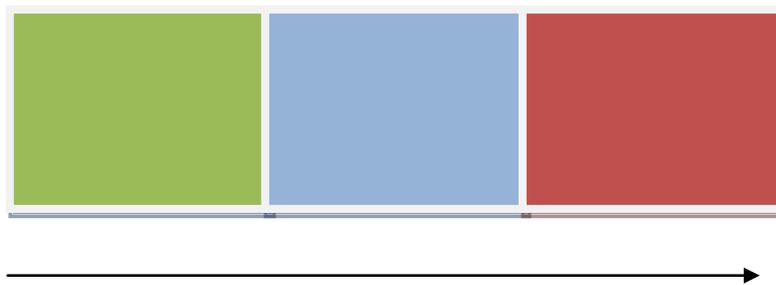


Variación Total.

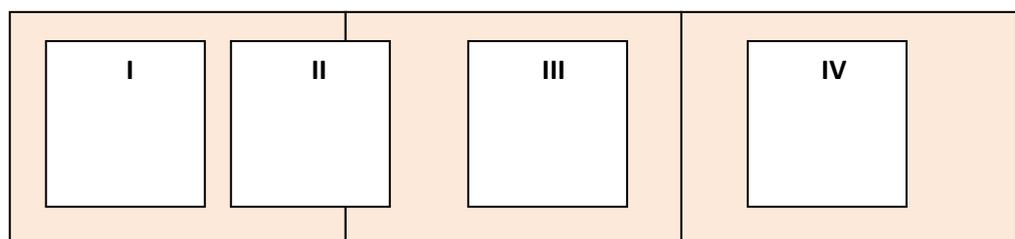


- **Asignación de Tratamientos.**

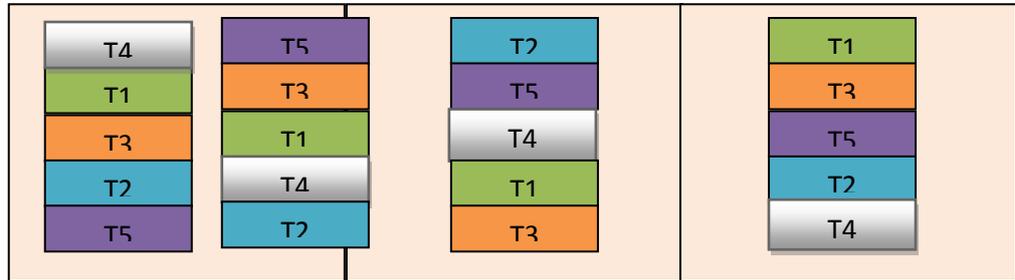
Se va a realizar un experimento con 5 tratamientos y 4 repeticiones con el siguiente gradiente de fertilidad del lote.



PASO 1: Divida el campo en partes iguales c/u representando un bloque.



PASO 2: SE divide cada bloque en cinco parcelas iguales c/u y se asignan al azar los tratamientos a las cinco parcelas en cada bloque uno por uno.



Modelo Estadístico

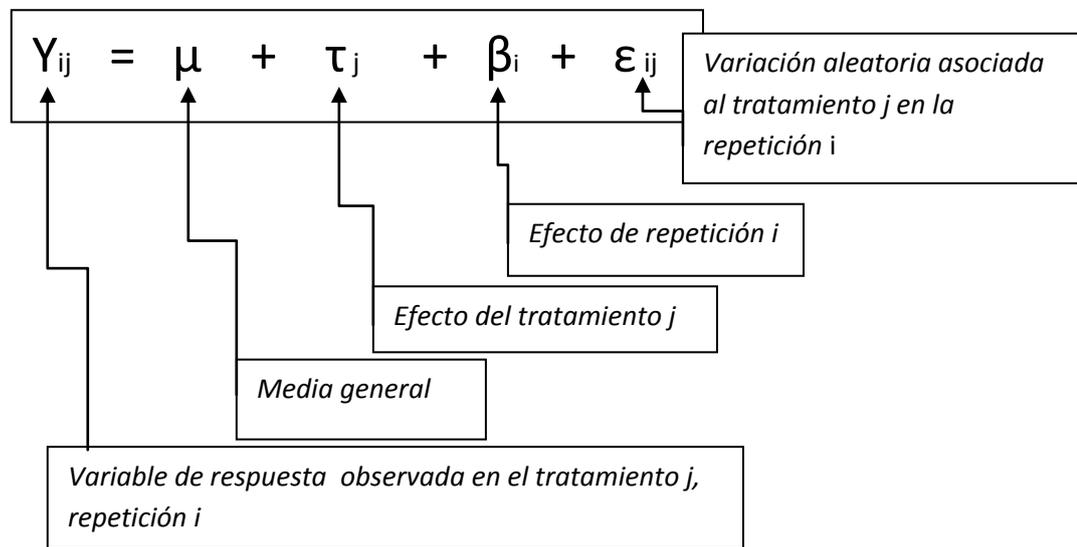


Tabla de Análisis de Varianza

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
Bloque	b-1	SC bloque	CM bloque	F bloque	Proba >F
Tratamiento	t-1	SC tratamiento	CM tratamiento	F tratamiento	Proba >F
Error	(t-1)(b-1)	SC error	CM error		
Total	bt-1	SC total			



Organización de los Datos

		Tratamientos					
		1	2	3	...	t	TOTAL
Bloques	1	y_{11}	y_{12}	y_{13}	...	y_{1j}	$y_{1\cdot}$
	2	y_{21}	y_{22}	y_{23}	...	y_{2j}	$y_{2\cdot}$
	3	y_{31}	y_{32}	y_{33}	...	y_{3j}	$y_{3\cdot}$
		
		
		
	r	y_{r1}	y_{r2}	y_{r3}	...	y_{rj}	$y_{r\cdot}$
TOTAL		$y_{\cdot 1}$	$y_{\cdot 2}$	$y_{\cdot 3}$...	$y_{\cdot j}$	$y_{\cdot \cdot}$

Cálculos necesarios para la tabla de análisis de varianza

FUENTES DE VARIACIÓN	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADOS MEDIOS	RAZÓN DE F.	PROBABILIDAD DE F.
Bloque	b-1	$\sum y_{i\cdot}^2 - FC$	$SC_{\text{bloque}}/b-1$	$CM_{\text{bloque}}/CM_{\text{error}}$	Proba >F
Tratamiento	t-1	$\sum y_{\cdot j}^2 - FC$	$SC_{\text{trat}}/t-1$	$CM_{\text{trat}}/CM_{\text{error}}$	Proba >F
Error	(t-1)(b-1)	Diferencia	$SC_{\text{error}}/t-1$		
Total	bt-1	$\sum \sum y_{ij}^2 - FC$			

Factor de corrección

$$FC = y_{\cdot \cdot}^2 / rt$$

Ejemplo. Con el fin de evaluar el control de malezas se probaron tres métodos y tomaron datos a los 10 y 20 días después de la germinación del cultivo.



Tratamiento	Bloque	% control 10 ddg	% control 20 ddg
A	1	12	25
B	1	12	25
C	1	18	26
A	2	11	21
B	2	10	22
C	2	16	22
A	3	11	26
B	3	10	26
C	3	16	28

Hipótesis Planteada: hay diferencias en el efecto de los tres métodos de control de malezas a los 10 y 20 días después de la aplicación.

Organización de los Datos:

	A	B	C	Yi.
Bloque 1	18	14	13	45
Bloque 2	20	17	15	52
Bloque 3	16	13	12	41
Y.J	54	44	40	138

Pasos para construir la tabla de análisis de varianza

Paso 1: Factor de corrección $FC = y..^2/rt$

$$FC = 138^2/3*3 = 2116$$

Paso 2: Suma de cuadrados total

$$SCT = \sum \sum y_{ij}^2 - FC = (18^2+20^2+16^2+14^2+17^2+13^2+13^2+15^2+12^2) - 2116 = 56$$

Paso 3: Suma de cuadrados de bloques

$$SC_{\text{bloques}} = \sum Y_{i.}/t - FC = (45^2+52^2+41^2)/3 - 2116 = 20.66$$

Paso 4: Suma de cuadrados de tratamientos

$$SC_{\text{trata.}} = \sum Y_{.j}/r - FC = (54^2+44^2+40^2)/3 - 2116 = 34.66$$

Paso 5 Suma de cuadrados del error



$$SC_{\text{error}} = SCT - SC_{\text{bloques}} - SC_{\text{trata.}} = 56.00 - 20.66 - 34.66 = 0.68$$

Tabla de análisis de varianza

Fuente de variación	G. L.	S.C	C.M	FC
BLOQUE	2	20.66	10.33	60.67
TRATAMIENTO	2	34.66	17.33	101.94
ERROR	4	0.68	0.17	
TOTAL	8	56.00		

Se compara el valor de Fc debe ser comparado con las tablas para 2 y 4 grados de libertad y

$\alpha = 0.05$. De la tabla 2 de valores de F se tiene:

F tabulada = 6.94

Condición de F calculada > F tabulada (60.67 > 6.94; 101.94 > 6.94), entonces se acepta la hipótesis de investigación, concluyendo que el efecto de los tres métodos de control de malezas, es estadísticamente diferente.

11. PRUEBAS DE COMPARACIÓN MÚLTIPLE

En el análisis de varianza la estadística de F se utiliza para probar que la hipótesis de efecto diferente de los tratamientos es cierta y que no son iguales. Cuando está hipótesis se acepta, se debe continuar el análisis estadístico con el fin de determinar cuáles de los tratamientos son los mejores. Para ello se utiliza una forma de análisis llamada prueba de comparación múltiple.

11.1 PRUEBA DE COMPARACIÓN MÚLTIPLE



La selección de la prueba depende de:

- Naturaleza de los tratamientos
- Comparación por pares o por grupos
- Número de repeticiones
- Comparación lógica desde el punto de vista biológico.

Tipos de prueba de comparación múltiple

- Diferencia mínima significativa LSD
- Prueba de Duncan
- Prueba de Tukey
- Prueba de Dunnett



	Nombre	Cargo	Firma	Fecha
Elaborado	Helbert Salinas Abadía	Coordinador CIPS		2009-10-20
Revisado	Ana Beatriz Jaramillo Abadía	Coordinadora Grupo Interno de Trabajo de Planeación		2009-10-20
Aprobado	Mandina Quizza Tómich	Rectora		2009-10-20