VIII. ARREGLOS FACTORIALES

CARACTERISTICAS.

En investigación agrícola y pecuaria, es conocido que diversas variables afectan los resultados de un experimento. El ejemplo clásico es la investigación del rendimiento de un cultivo bajo diferentes dosis (niveles) de Nitrógeno, Fósforo y Potasio, estos últimos (los nutrimentos: N,P,K), son las variables o mejor como se conocen: factores. En caso de un ensayo de nutrición se tiene como factores: proteína, el factor fuente de energía, el factor raza, el factor concentración, el factor temperatura. En industria tendríamos por ejemplo la producción de un químico que contemple como factores a: temperaturas, presiones y diluciones.

Un "factor" es entonces una clase de tratamiento y en arreglos factoriales, cualquier factor corresponde a varios tratamientos. El término "nivel" hace referencia a un valor de un factor. También se refiere a la cantidad o un estado de un factor. Si comparan cinco clases de forrajes y tres tipos de manejos, el experimento se denomina un factorial 5 x 3, con cinco niveles del factor forraje y 3 niveles del factor manejo. Así se tienen los niveles alto y bajo del factor proteína, los niveles animales y vegetal del factor fuente; los niveles 15°C, 20°C y 30°C del factor temperatura, etc.

Por lo anterior, se observa que la aplicabilidad y uso de los arreglos factoriales comprende numerosas ciencias. Particularmente, en el área agrícola, son de gran utilidad cada vez que se genera ó produce un material genético nuevo dígase un híbrido, variedad ó clon. Este nuevo material genético, cuando se entrega al agricultor debe ir acompañado de lo que se denomina el paquete tecnológico, es decir, se debe indicar cual es la densidad y época de siembra oportuna, las cantidades adecuadas de elementos mayores: nitrógeno, fósforo y potasio, etc.

Los experimentos o arreglos factoriales, no constituyen un nuevo tipo de diseño experimental, corresponden como su nombre lo indica a una forma particular de combinar un conjunto de tratamientos y no a la forma o procedimiento de como se asignan los tratamientos a las unidades experimentales. Razón por la cual un arreglo factorial puede conducirse en cualquiera de los diseños hasta ahora estudiados: diseño completamente al azar, diseño bloques completos al azar y diseño cuadrado latino.

VENTAJAS.

Las ventajas de la experimentación factorial naturalmente dependen de los propósitos del experimento en particular, en general para grandes experimentos factoriales el error estandar por unidad se incrementa. Usualmente este incremento puede conservarse pequeño usando el artificio conocido como confusión. Algunas de las ventajas son:

1. Son recomendables si se desea investigar la (s) interacción (s) entre varios factores, igualmente son de gran valor en experimentos exploratorios, puesto que se pueden estudiar varios factores conjuntamente.

2. Son convenientes si se desea llegar a recomendaciones aplicables a una gran variedad de condiciones.

3. En los experimentos factoriales cada factor (el efecto) se estima con la misma precisión que se obtendría si todo el experimento se dedica un factor específico.

4. Es altamente eficiente porque cualquier observación suministra información acerca de todos los factores incluídos en el experimento.

5. Algunas veces se puede introducir otro factor dentro del experimento que no es por sí solo de interés, pero puede formar las bases para recomendaciones sanas acerca de los otros factores, cuando éllos no son independientes.

DESVENTAJAS.

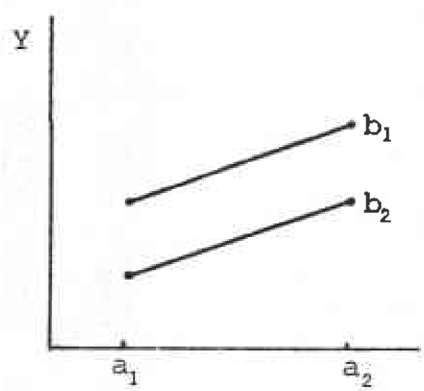
1. Los resultados pueden ser difíciles de interpretar, sobre todo las interacciones entre los factores. Sin embargo, conviene anotar que no es un defecto del factorial , sino consecuencia de la complejidad del fenómeno en estudio.

2. El aumento de factores o bien de niveles puede conducir a un aumento considerable del total de tratamientos, lo cual conlleva a tamaños de replicaciones grandes. Sin embargo, puede remediarse el problema con la técnica de confusión.

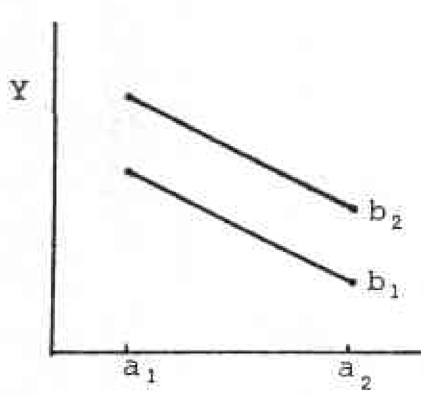
INTERACCION: Por interacción entre dos factores en estudio, por ejemplo A y B, se entenderá al hecho de que la diferencia en respuesta, a los niveles de un factor, sea igual ó diferente a los distintos niveles del otro factor.

Una interacción no significativa indica que los dos factores en estudio son independientes entre sí, es decir, los efectos de un factor son los mismos para todos los niveles de los otros factores, y la variación que se observe será debida al azar y no a la condición de los factores como tal. Cuando los factores no son independientes, los datos requieren de un estudio más detallado, para saber cuál es la causa de la interacción y probablemente de la ejecución de los experimentos adicionales en el futuro.

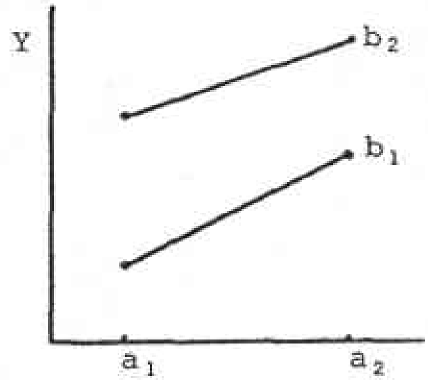
Las siguientes figuras ilustran la existencia ó no de la interacción entre los factores A y B.



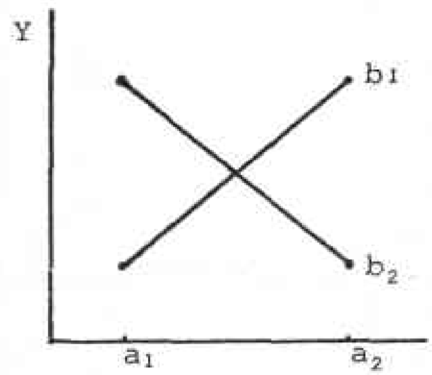
No Interacción



No Interacción



Interacción: Escala



Interacción Dirección

NOTACION.

Los sistemas de notación para los experimentos factoriales en general han sido similares, aunque en ciertos libros difieren un poco. Se seguirá la notación propuesta por Yates, donde las letras mayúsculas se usan para referirse a factores: por ejemplo A - B - N - P - D y las minúsculas correspondientes para denotar los niveles de los factores. Considérese que se tienen dos factores cada uno a dos niveles así:

Factores		Niveles
A	→	a_o, a_1
B	→	b_o, b_1

Las cuatro combinaciones de tratamientos puede representarse por a_ob_o, a_ob_1, a_1b_o, a_1b_1, se acostumbra también las siguientes notaciones para las combinaciones:

a_ob_o	equivalente	a	(1)	equivalente	a	(0,0)
a_ob_1	"	"	b	"	"	(0,1)
a_1b_o	"	"	a	"	"	(1,0)
a_1b_1	"	"	ab	"	"	(1,1)

La extensión para más de dos factores, y dos niveles para cada factor por ejemplo: 3 factores A, B, C se tendrían 8 combinaciones de tratamientos, así:

$a_ob_oc_o$	equivalente	a	(1)	equivalente	a	(000)
$a_1b_oc_o$	"	"	a	"	"	(100)
$a_ob_1c_o$	"	"	b	"	"	(010)
$a_1b_1c_o$	"	"	ab	"	"	(110)
$a_ob_oc_1$	"	"	c	"	"	(001)
$a_1b_oc_1$	"	"	ac	"	"	(101)
$a_ob_1c_1$	"	"	bc	"	"	(011)
$a_1b_1c_1$	"	"	abc	"	"	(111)

ESTIMACION.

Caso de dos factores:

Considérense los factores A y B cada uno con dos niveles. Se podría estimar A así:

Efecto de A al nivel b_o : $a_1b_o - a_ob_o$

Efecto de A al nivel b_1 : $a_1b_1 - a_ob_1$

Como aproximación de A podríamos considerar el promedio de los efectos de A al nivel b_o y A al nivel b_1 así:

$$A = 1/2 \{(a_1b_1 - a_ob_1) + (a_1b_o - a_ob_o)\}$$

$$A = 1/2 \{(ab - b) + (a - 1)\}$$

Simbólicamente puede escribirse: $A = 1/2\ (a - 1)\ (b+ 1)$

en forma similar el efecto de B sería:

Efecto de B al nivel a_o: $a_ob_1 - a_ob_o$

Efecto de B al nivel a_1: $a_1b_1 - a_1b_o$

entonces:

$$B = 1/2 \{(a_1b_1 - a_1b_o) + (a_ob_1 - a_ob_o)\}$$

$$B = 1/2 \{(ab - a) + (b - 1)\}$$

Simbólicamente se tiene:

$$B = 1/2\ (a + 1)\ (b - 1)$$

Como una medida de la interacción podemos considerar la diferencia entre el efecto de A a los dos niveles de B:

$$AB = 1/2 \{(a_1b_1 - a_ob_1) - (a_1b_o - a_ob_o)\}$$

$$AB = 1/2 \{(ab - b) - (a - 1)\}$$

Simbólicamente:

$$AB = 1/2\ (a - 1)\ (b - 1)$$

Puede comprobarse fácilmente que daría el mismo resultado, tomar la diferencia de los efectos de B a los dos niveles de A:

$$AB = 1/2 \{ (a_1 b_1 - a_1 b_o) - (a_o b_1 - a_o b_o) \}$$

$$AB = 1/2 \{ (ab - a) - (b - 1) \}$$

Caso de tres factores:

Supóngase 3 factores a dos niveles cada uno:

Factor		Niveles
A	$\rightarrow$	a_o, a_1
B	$\rightarrow$	b_o, b_1
C	$\rightarrow$	c_o, c_1

Se tendrían ocho combinaciones de tratamientos:

$a_o b_o c_o$	equivalente	a	(1)
$a_1 b_o c_o$	"	"	a
$a_o b_1 c_o$	"	"	b
$a_1 b_1 c_o$	"	"	ab
$a_o b_o c_1$	"	"	c
$a_1 b_o c_1$	"	"	ac
$a_o b_1 c_1$	"	"	bc
$a_1 b_1 c_1$	"	"	abc

Para estimar el efecto de incrementar el factor a_o a a_1 podríamos generar la siguiente tabla:

Nivel B	Nivel C.	Efecto de A
b_o	c_o	$a_1b_oc_o - a_ob_oc_o$
b_1	c_o	$a_1b_1c_o - a_ob_1c_o$
b_o	c_1	$a_1b_oc_1 - a_ob_oc_1$
b_1	c_1	$a_1b_1c_1 - a_ob_1c_1$
Efecto medio de A		1/4 de la suma

$$A = 1/4 \{ (a_1b_oc_o - a_ob_oc_o) + (a_1b_1c_o - a_ob_1c_o) + (a_1b_oc_1 - a_ob_oc_1) + (a_1b_1c_1 - a_ob_1c_1) \}$$

En la otra notación sería:

$$A = 1/4 \{ (a - 1) + (ab - b) + (ac - c) + (abc - bc)\}$$

Simbólicamente puede escribirse.

$$A = 1/4 \; (a - 1) \; (b + 1) \; (c + 1)$$

Para estimar la interacción considérese AB:

Efecto de A, para b_o; $:1/2 \{ (a_1b_oc_o - a_ob_oc_o) + (a_1b_oc_1 - a_ob_oc_1)\}$

media de c_o y c_1.

Efecto de A, para b_1; $: 1/2 \{ (a_1b_1c_o - a_ob_1c_o) + (a_1b_1c_1 - a_ob_1c_1) \}$

media de c_o y c_1.

La diferencia promedio entre estas cantidades sería la interacción AB.

$$AB = 1/4\ \{(a_1b_1c_o - a_ob_1c_o)+(a_1b_1c_1 - a_ob_1c_1)-(a_1b_oc_o - a_ob_oc_o) - (a_1b_oc_1 - a_ob_oc_1)\}$$

$$AB = 1/4\ (ab - b + abc - bc - a + 1 - ac + c)$$

Simbólicamente:

$$AB = 1/4\ (a - 1)\ (b-1)(c+1)$$

Para la interacción triple consideremos ABC:

Interacción AB a nivel c_o:

$$1/2\ \{(a_1b_1c_o - a_ob_1c_o) - (a_1b_oc_o - a_ob_oc_o)\}$$

Interacción AB a nivel c_1:

$$1/2\ \{(a_1b_1c_1 - a_ob_1c_1) - (a_1b_oc_1 - a_ob_oc_1)\}$$

El promedio de estos valores sería la interacción AB antes mencionada. En caso de que estas dos interacciones sean diferentes, la media entre las diferencias de ellas correspondería ala interacción ABC.

$$ABC = 1/4\ \{(a_1b_1c_1 - a_ob_1c_1 - a_1b_oc_1 + a_ob_oc_1) - (a_1b_1c_o - a_ob_1c_o - a_1b_oc_o + a_ob_oc_o)\}$$

En la notación alterna sería:

$$ABC = 1/4\ (abc - bc - ac + c - ab + b + a - 1)$$

Y simbólicamente

ABC = 1/4 (a-1) (b-1) (c-1)

Los efectos e interacciones pueden resumirse en la siguiente matriz donde M es la media general:

	(1)	a	b	ab	c	ac	bc	abc
8M	1	1	1	1	1	1	1	1
4A	-1	1	-1	1	-1	1	-1	1
4B	-1	-1	1	1	-1	-1	1	1
4AB	1	-1	-1	1	1	-1	-1	1
4C	-1	-1	-1	-1	1	1	1	1
4AC	1	-1	1	-1	-1	1	-1	1
4BC	1	1	-1	-1	-1	-1	1	1
4ABC	-1	1	1	-1	1	-1	-1	1

Las hileras de la tabla dan los efectos factoriales en términos de los promedios de tratamientos, mientras que las columnas dan las medidas de los tratamientos en términos de los efectos factoriales. Por ejemplo: a - (1) = A - AB - AC + ABC., que es el simple efecto de A cuando B y C están al primer nivel.

Por niveles es de interés en el siguiente sentido. La contribución de cualquier efecto o la suma de cuadrados para tratamientos es $\{\quad\}^2/8r$, donde { } denota el efecto factorial total, donde r es el número de replicaciones, por ejemplo:

$\{A\} = -\{1\} + \{a\} - \{b\} + \{ab\} - \{c\} + \{ac\} - \{bc\} + \{abc\}$

La matriz es fâcil de construîr: se ha adoptado de orden estandar para los sîmbolos de tratamientos y efectos y los nûmeros para A, B y C serân + 1 si la letra mayûscula correspondiente estâ presente y -1 si no lo estâ. Los nûmeros para las combinaciones de letras, por ejemplo AB son el producto de los correspondientes para A y B.

Caso general:

En el caso general con n factores, cada uno a dos niveles se tendrâ:

$$X = \frac{1}{2^{n-1}} (a \pm 1)(b \pm 1)(c \pm 1) \ldots$$

Donde el signo en cada parêntesis es positivo, si la letra mayûscula correspondiente no estâ presente en X y negativa si lo estâ.

En el caso general de n factores cada uno a dos niveles (sistema 2) por ejemplo A-B-C-D... las combinaciones de tratamientos serîan:

1	d	e	de	
a	ad	ae	ade	
b	bd	be	bde	
ab	abd	abe	abde	
c	cd	ce	cde	
ac	acd	ace	acde	
bc	bcd	bce	bcde	etc...
abc	abcd	abce	abcde	

Se tendrá el siguiente número de efectos e interacciones.

	Número
Efectos principales	n
Interacciones dobles	$\binom{n}{2}$
Interacciones triples	$\binom{n}{3}$
Interacciones cuádruples	$\binom{n}{4}$
.	.
.	.
.	.
Total	$2^n - 1$

Los efectos e interacciones estarían dados por:

$$A = \frac{1}{2^{n-1}} (a-1)(b+1)\ (c+1)\ (d+1) \ldots$$

$$AB= \frac{1}{2^{n-1}}(a-1)\ (b-1)\ (c+1)\ (d+1)\ (e+1) \ldots$$

$$BDE = \frac{1}{2^{n-1}}\ (a+1)\ (b-1)\ (c+1)\ (d-1)\ (e-1) \ldots$$

Como ejercicio al estudiante, se deja la construcción de la matriz de coeficientes para la estimación de los efectos e interacciones en un 2^4 y un 2^3, siguiendo las reglas dadas en las secciones anteriores.

CONTRASTES

Un contraste o una comparación entre tratamientos como se definió, está dado por la siguiente ecuación:

$$Q = \Sigma c_i \tau_i \text{ con } \Sigma c_i = 0$$

donde los τ_i son tratamientos en comparación y los c_i son números, en la mayoría de los casos los c_i son enteros pero bien puede ocurrir que sean fracciones.

El estimador para $\Sigma c_i \tau_i$ será:

$$\Sigma c_i \hat{\tau}_i \text{ con } \Sigma c_i = 0$$

donde $\hat{\tau}_i$ es el estimador de τ_i, es decir, el promedio de cada tratamiento en cuestión.

La suma de cuadrados atribuíble a dicha comparación está definida por:

$$\frac{r \; (\Sigma c_i \hat{\tau}_i)^2}{\Sigma c_i^2}$$

r es el número de repeticiones.

Considérese dos comparaciones o contrastes así:

$Q_1 = \Sigma c_i \tau_i$ y $Q_2 = \Sigma d_i \tau_i$. Las comparaciones se dice que son ortogonales si.

$$\Sigma c_i d_i = 0$$

ésto es, la suma de los productos de los coeficientes es cero, por ejemplo.

$$Q_1 = \tau_1 - \tau_2 \quad Q_2 = \tau_1 + \tau_2 - 2\tau_3$$

entonces: $c_1 = 1,\ c_2 = -1,\ c_3 = 0;\quad d_1 = 1,\ d_2 = 1,\ d_3 = -2$

$$\Sigma c_i d_i = \{1 \times 1\} + \{(-1) \times 1\} + \{0 \times (-2)\} = 0$$

La varianza de un contraste está dada por: $\frac{\sigma^2}{r} \; \Sigma c_i^2$

El concepto de contraste o comparación, unido a los de estimación anteriormente señalados, permiten la ejecución de las pruebas de hipótesis mediante la tabla del análisis de varianza.

MODELO ESTADISTICO.

El modelo estadístico para un arreglo factorial 2^2 ejecutado en el diseño completamente al azar es:

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ijk}$$

donde:

Y_{ijk} = variable aleatoria observable

μ = media general

A_i = efecto del i-ésimo nivel del factor A

B_j = efecto del j-ésimo nivel del factor B

AB_{ij} = efecto de la interacción

$\varepsilon_{ijk} \sim N(0,\sigma^2)$, no correlacionada

Si fuera en bloques completos al azar el modelo sería:

$$Y_{ijk} = \mu + R_k + A_i + B_j + AB_{ij} + \varepsilon_{ijk}$$

donde:

R_k = Efecto de la k-ésima replicación

Los otros efectos como se definieron en el parágrafo precedente.

EJEMPLO DE UN FACTORIAL 2^2. Como de costumbre se ilustrará la metodología de análisis mediante un ejemplo númerico. En este caso se quizo estudiar el efecto de dos fuentes de energía (papa y melaza) y un anabólico (con y sin) sobre la utilización de úrea en ceba de bovinos con ensilaje de maíz. Los datos son cortesía del Dr. Nobel Jiménez, corresponden a incrementos de peso. El diseño fue completamente al azar en un arreglo 2^2 y 5 animales por combinación de energía y anabólico. Denote a E: e_0 = papa, e_1 = melaza; A: a_0 = sin anabólico, a_1 = con anabólico.

E:	e_0 = Papa		e_1 = Melaza		
	a_0	a_1	a_0	a_1	
A:	Sin Anab.	Con Anab.	Sin Anab.	Con Anab.	
	32	43	24	29	
	28	15	28	27	
	22	32	13	36	
	30	28	36	29	
	31	34	14	37	
Total	143	152	115	158	568
PROM.	28.60	30.40	23.00	31.60	
TRAT.	1	a	e	ea	

El análisis de varianza es:

TABLA DEL ANALISIS DE VARIANZA

F.V.	G.L.	S.C.	C.M.	F
Trat.	3	217.2	72.4	1.23
Error	16	939.6	58.7	
Total	10	1156.8		

El objetivo es dividir los G.L. y S.C. de tratamientos en: Energía, anabólico y su interacción para tal propósito se prosigue así:

$E = 1/2\ (e - 1)\ (a + 1) = 1/2\ (ae + e - a - 1)$

$E = 1/2\ (31.60 + 23.00 - 30.40 - 28.60) = -2.20$

$A = 1/2\ (a - 1)\ (e + 1) = 1/2\ (ae + a - e - 1)$

$A = 1/2\ (31.60 + 30.40 - 23.00 - 28.60) = 5.20$

$AE = 1/2\ (e - 1)\ (a - 1) = 1/2\ (ea + 1 - a - e)$

$AE = 1/2\ (31.60 + 28.60 - 30.40 - 23.00) = 3.40$

Cada uno de estos efectos es un contraste, por consiguiente se pueden calcular sus sumas de cuadrados. En el caso particular de energía se tiene:

$\hat{\tau}_1 = 31.60;\ \hat{\tau}_2 = 23.00;\ \hat{\tau}_3 = 30.40;\ \hat{\tau}_4 = 28.60$

$C_1 = 1/2;\ C_2 = 1/2;\ C_3 = -1/2;\ C_4 = -1/2$

$\Sigma C_i = 0; \qquad \Sigma C_i^2 = 1$

De forma similar se plantean los efectos de anabólico y la interacción, las sumas de cuadrados son:

$$\text{S.C. Contraste} = r(\Sigma c_i \hat{\tau}_i)^2/\Sigma c_i^2$$

$$\text{S.C.(E)} = 5(-2.20)^2/1 = 24.20$$

$$\text{S.C.(A)} = 5(5.20)^2/1 = 135.20$$

$$\text{S.C.(EA)} = 5(3.40)^2/1 = 57.80$$

Se observa que la suma de las tres sumas de cuadrados es igual a la S.C. de tratamientos. Existe la metodología alterna a la anterior para obtener cada suma de cuadrados en tal caso se prosigue:

	Papa	Melaza	Total
Sin	143	115	258
Con	152	158	310
Tot.	295	273	568

$$\text{F.C.} = 568^2/20 = 16131.20$$

$$\text{S.C. Energía} = \frac{295^2 + 273^2}{10} - FC = 24.20$$

$$\text{S.C. Anabol} = \frac{258^2 + 310^2}{10} - FC = 135.20$$

$$\text{S.C.E x A} = \frac{143^2 + \ldots + 158^2}{5} - FC - \text{S.C.(E)} - \text{S.C.(A)}$$

$$\text{S.C.E x A}' = 57.80$$

Finalmente la tabla del análisis de variancia quedaría:

TABLA DEL ANALISIS DE VARIANCIA

F.V.	G.L.	S.C.	C.M.	Fc	
Energía	1	24.20	24.2	0.41	n.s.
Anabol.	1	135.20	135.2	2.30	n.s.
E x A	1	57.80	57.8	0.98	n.s.
Error	16	939.60	58.7		
Total	19	1156.80			

FACTORIAL A TRES NIVELES: Hasta aquí, se ha estudiado lo que constituye el sistema 2^n, se observó su estimación, el modelo y finalmente se hizo la ilustración con un ejemplo numérico. Análogamente existen los sistemas 3^n, p^n donde 3 y p son los niveles de n factores, ambos sistemas son de gran utilidad y aplicación tanto en las ciencias pecuarias como agrícolas. No nos extenderemos en estos sistemas tan detenidamente como se hizo con el 2^n, ésto no quiere decir que no sean importantes, sólo se presentará un ejemplo numérico de un 3^n y del p^n se observará en otro capítulo también mediante un ejemplo numérico.

Los datos que se expresan a continuación, corresponden a producción del Pasto Braquiaria en terrazas media de los Llanos Orientales. El propósito es el de observar la respuesta del pasto a la aplicación de nitrógeno (0-50-100 kg/ha) y Cal (0-2-4 Ton/ha). El diseño empleado fue de bloques completos al azar 9 tratamientos y 3 replicaciones. Los nueve tratamientos corresponden a un arreglo factorial 3 x 3 ó 3^2.

REPETICIONES

TRAT.	N	CAL	I	II	III	TOTAL
1	0	0	5.711	4.770	6.814	17.295
2	0	2	6.944	5.938	8.372	21.254
3	0	4	7.983	5.938	5.030	18.951
4	50	0	5.938	6.555	6.328	18.821
5	50	2	6.685	4.965	8.112	19.762
6	50	4	6.133	5.614	5.646	17.393
7	100	0	7.204	5.841	6.750	19.795
8	100	2	8.567	7.139	8.177	23.883
9	100	4	6.263	5.322	5.809	17.394
TOTAL			61.428	52.082	61.038	174.548

		CAL 0	CAL 2	CAL 4	TOTAL
	0	17.295	21.254	18.951	57.500
NIT	50	18.821	19.762	17.393	55.976
	100	19.795	23.883	17.394	61.072
TOTAL		55.911	64.899	53.738	174.548

Con las tablas anteriores se procede al cálculo de las sumas de cuadrados así:

$$FC = \frac{174.548^2}{27} = 1128.408$$

$$S.C.\ REP = \frac{61.428^2 + \ldots + 61.038^2}{9} - FC = 6.212$$

$$S.C.\ NIT = \frac{57.500^2 + \ldots + 61.072^2}{9} - FC = 1.520$$

$$S.C.\ CAL = \frac{55.911^2 + \ldots + 53.738^2}{9} - FC = 7.781$$

$$S.C.\ N \times C \quad \frac{17.295^2 + \ldots + 17.394^2}{3} - FC - SC\ NIT - S.C.\ CAL = 2.979$$

$$S.C.\ TOTAL = 5.711^2 + \ldots + 5.809^2 - FC = 29.767$$

$$S.C.\ Error = S.C.\ Total - S.C.\ Rep - S.C.\ Nit - S.C.\ Cal - S.C.\ N \times C$$

$$S.C.\ Error = 29.974 - 6.212 - 1.520 - 7.781 = 11.275$$

Se establece la tabla del análisis de variancia:

TABLA DEL ANALISIS DE VARIANCIA

F.V.	G.L.	S.C.	C.M.	F
Rep.	2	6.212	3.106	4.41 *
Nitrog.	2	1.520	0.760	1.08 NS
Cal	2	7.781	3.890	5.52 *
Nit x Cal	4	2.979	0.745	1.06 NS
Error	16	11.275	0.7047	
Total	26	29.767		

Se concluye que hubo efecto de la cal sobre la producción de forraje. Se procede a calcular los promedios para cada nivel de cal y se realiza una prueba de comparación múltiple o bien una descomposición de la variación de la cal en efectos: lineal y cuadrático.

Promedios:

		CAL			
		0	2	4	$\bar{X}$
	0	5.765	7.085	6.317	6.389
NIT	50	6.274	6.587	6.798	6.220
	100	6.598	7.961	5.798	6.786
	$\bar{X}$	6.212	7.211	5.971	6.465

FACTORIAL 2 x 3 x 3. Los datos que se expresan a continuación corresponden a aumentos de peso diario en gazapos, expresados en gramos. El diseño experimental fue de completamente al azar, en arreglo factorial 2 x 3 x 3 y dos repeticiones (promedio de 5 animales). Los datos son cortesía del programa de especies menores del ICA.

Energía (E): 3100 y 3200 calorías

Proteína (P): 12, 15 y 18 porcentaje en la dieta

Tiempo de suministro (T): 80,100 y 120 días

No.	E	P	T	I	II	TOTAL
1.	3100	12	80	13,16	10,97	24.13
2.	3100	12	100	12,72	9,55	22,27
3.	3100	12	120	11,35	8.67	20,02
4.	3100	15	80	16.13	17,30	33,43
5.	3100	15	100	13,75	14,56	28,31
6.	3100	15	120	11,96	13,93	25,89
7.	3100	18	80	15,25	14,65	29,90
8.	3100	18	100	13,75	12,86	26,61
9.	3100	18	120	12,10	11,11	23,21
10.	3200	12	80	12,52	14,84	27,36
11.	3200	12	100	9.76	13,15	22,91
12.	3200	12	120	8.13	10.84	18,97
13.	3200	15	80	16,18	19,37	35,55
14.	3200	15	100	15,14	18,77	33,91
15.	3200	15	120	15,07	17,35	32,42

No.	E.	P	T	I	II	TOTAL
16.	3200	18	80	17,74	20,54	38,28
17.	3200	18	100	16,71	17,12	33,83
18.	3200	18	120	15,73	17,22	32,95

Para el cálculo de las sumas de cuadrados, se procede como sigue:

$$FC = \frac{509{,}95^2}{36} = 7223{,}58$$

		Proteína			
		12	15	18	$\overline{X}$
Energía	3100	66,42	87,63	79,72	233,77
	3200	69,24	101,88	105,06	276,18
		135,66	189,51	184,78	509,95

$$S.C.\ Ener. = \frac{233{,}77^2 + 276{,}18^2}{18} - FC$$

$$S.C.\ Ener. = 49{,}96$$

$$S.C.\ Prot. = \frac{135{,}66^2 + 189{,}51^2 + 184{,}78^2}{12} - FC$$

$$S.C.\ Prot. = 148{,}19$$

$$S.C.\ E \times P = \frac{66{,}41^2 + \ldots + 105{,}06^2}{6} - FC - S.C.\ Ener. - S.C.\ Prot.$$

$$S.C.\ E \times P = 21{,}13$$

		Tiempo			
		80	100	120	$\bar{X}$
Energía	3100	87,46	77,19	69,12	233,77
	3200	101,19	90,65	84,34	276,18
		188,65	167,84	153,46	509,95

$$S.C.\ Tiem. = \frac{188,65^2 + 167,84^2 + 153,46^2}{12} - FC$$

S.C. Tiem. = 52,17

$$S.C.\ Ener.\ x\ Tiem. = \frac{87,46^2 + \ldots + 84,34^2}{6} - FC - SC\ Ener. - SC\ Tiemp.$$

S.C. E x T = 0,15

		Proteína			
		12	15	18	
Tiempo	80	51,49	68,98	68,18	188,65
	100	45,18	62,22	60,44	167,84
	120	38,99	58,31	56,16	153,46
		135,66	189,51	184,78	

$$S.C.\ T x P = \frac{51,49^2 + \ldots + 56,16^2}{4} - FC = SCT - SCP$$

S.C T x P = 0,49

$$S.C.\ ExPX\ T = \frac{24,13^2 + \ldots + 32,95^2}{2} - FC - SCE - SCP$$

$$- SCT - SC\ ExP - SCExT - SCPxT$$

S.C. ExPXT = 5,67

S.C. TOTAL = $13{,}16^2 + \ldots + 17{,}22^2 -$ FC

S.C. TOTAL = 324,29

S.C. Error = S.C. TOTAL - S.C. Resto

S.C. Error = 46,53

La tabla del análisis de variancia quedaría como sigue:

TABLA DEL ANALISIS DE VARIANCIA

F. de V.	G.L.	S.C.	C.M.	Fc
Ener. (E)	1	49,96	49,96	19,33 **
Prot. (P)	2	148,19	74,10	28,66 **
E x P	2	21,13	10,57	4,09 *
Tiempo (T)	2	52,17	26,09	10,09 **
E x T	2	0,15	0,07	0,03
P x T	4	0,49	0,12	0,05
E x P x T	4	5,67	1,42	0,55
Error	18	46,53	2,59	
Total	35	324,29		

Las tablas de promedios para cada factor e interacciones son:

		Proteína			
		12	15	18	$\bar{X}$
	300	11,07	14,61	13,29	12,99
Ener.	3200	11,54	16,98	17,51	15,34
	$\bar{X}$	11,31	15,79	15,40	

		Tiempo			
		80	100	120	$\bar{X}$
	3100	14,58	12,87	11,52	12,99
Ener.	3200	16,87	15,11	14,06	15,34
	$\bar{X}$	15,72	13,99	12,79	

		Tiempo			
		80	100	120	$\bar{X}$
	12	12,87	11,30	9,75	11,31
Prot.	15	17,25	15,56	14,58	15,79
	18	17,05	15,11	14,04	15,40
	$\bar{X}$	15,72	13,99	12,79	

De la tabla del análisis de variancia se obtiene como resultado importante, la significancia de la interacción Energía x Proteina, la cual se explicará a continuación, mediante una prueba de comparación múltiple, como Duncan al nivel del 5%, se tiene:

$$S_{\bar{X}} = \sqrt{CME/r} = \sqrt{2,59/6} = 0,6570$$

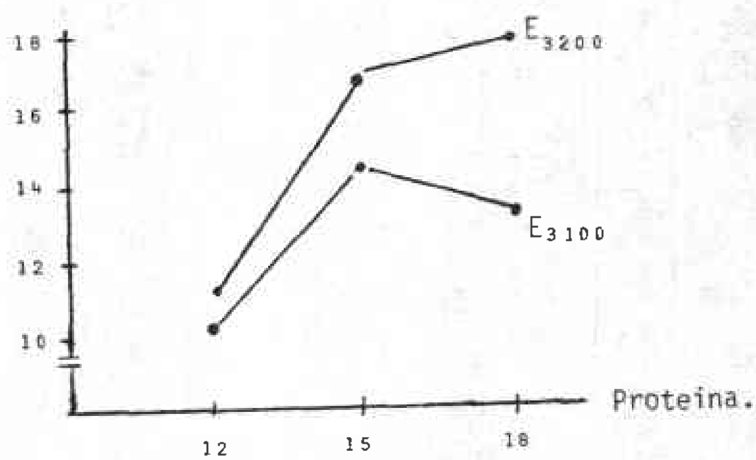
g.l. error = 18

AES (5%) =	2,97	3,12	3,21	3,27	3,32
ALS (5%) =	1,95	2,05	2,11	2,15	2,18

		Proteina		
		12	15	18
Ener.	31 00	11,07[d]	14,61[b]	13,29[bc]
	32 00	11,54[cd]	16,98[a]	17,51[a]

Esta prueba no tiene mucho sentido, puesto que los niveles son cuantitativos se hace solamente por ilustracióna académica.

Gráficamente, la interacción Energía x Proteina, se expresa como sigue:



la interacción es del tipo de escala y dirección simultáneamente.

Restaría por interpretar el tiempo, el cual se postergará cuando tengamos el conocimiento de los polinomios ortogonales y así evaluar el tipo o forma de respuesta: lineal y/o cuadrática, aunque aparentemente, los promedios parecen indicar, un tipo de respuesta lineal únicamente.

FACTORIAL 2^3. Los valores que se expresan a continuación, corresponden a toneladas de maíz por hectárea, en un estudio donde se buscaba evaluar la respuesta del maíz a las aplicaciones de nitrógeno, fósforo y la densidad de población. El diseño experimental usado fue de bloques completos al azar en un arreglo factorial 2^3 y 5 bloques.

	TRATAMIENTO (n) N	(p) P_2O_5	(d) D.P.	c.t.	REPLICACIONES I	II	III	IV	V	Total	$\bar{Y}$
1.	30	40	20	(1)	3.31	1.43	1.31	2.38	1.91	10.34	2.068
2.	30	40	50	d	4.41	1.63	1.09	2.61	2.53	12.27	2.454
3.	30	80	20	p	3.55	1.74	1.55	2.44	2.46	11.74	2.348
4.	30	80	50	pd	4.57	2.16	1.45	2.45	2.57	13.20	2.640
5.	60	40	20	n	3.35	1.43	1.06	2.29	2.23	10.36	2.072
6.	60	40	50	nd	4.45	1.78	1.12	1.99	2.82	12.16	2.432
7.	60	80	20	np	3.37	1.75	1.50	2.39	2.28	11.29	2.258
8.	60	80	50	npd	5.16	2.19	1.44	2.32	2.92	14.03	2.806
				TOTAL	32.17	14.11	10.52	18.87	19.72	95.39	2.385

$$FC = \frac{95.39^2}{40} = 227{,}4813$$

$$SC\ TOT = 3.31^2 + 4.41^2 + \ldots + 2.92^2 - FC = 39{,}082$$

$$SC\ REP = \frac{32.17^2 + \ldots + 19.72^2}{8} - FC = 33{,}722$$

$$SC\ TRAT. = \frac{10.34^2 + \ldots + 14.03^2}{5} - FC = 2{,}326$$

$$SC\ ERROR = SC\ TOT - (SC\ REP + SC\ TRAT) = 3{,}034$$

La tabla del análisis de variancia quedaría como sigue:

F. de V.	G.L.	S.C.	C.M.	F
Rept	4	33,722	8,431	77,79 **
Trat.	7	2,326	0,332	3,07 *
Error	28	3,034	0,108	
TOTAL	39	39,082		

El valor de las tablas para F es: $F(7,28,\ 0{,}05) = 2{,}36$

Por lo tanto existen diferencias significativas entre tratamientos; se procederá entonces a descomponer los 7 grados de libertad de tratamientos, en los efectos de nitrógeno, fósforo, densidad de población y sus respectivas interacciones. Se hará uso del desarrollo analítico presentado en la Sección de estimación y del concepto de contraste, así:

c.t.	(1)	n	p	np	d	nd	pd	npd
$\bar{Y}$	2,068	2,072	2,348	2,258	2,454	2,432	2,640	2,806

$$\hat{N} = \tfrac{1}{4}(n-1)(p+1)(d+1) = \tfrac{1}{4}(npd+np+nd+n-pd-p-d-(1))$$

$$= \tfrac{1}{4}(2{,}806+2{,}258+2{,}432+2{,}072-2{,}640-2{,}348-2{,}454-2{,}068) = 0{,}0145$$

$$\hat{P} = \tfrac{1}{4}(n+1)(p-1)(d+1) = \tfrac{1}{4}(npd+np+pd+p-nd-n-d-(1))$$

$$\hat{P} = \tfrac{1}{4}(2{,}806+2{,}258+2{,}640+2{,}348-2{,}432-2{,}072-2{,}454-2{,}068) = 0{,}2565$$

$\hat{NP} = {}^1/_4(n-1)\ (p-1)\ (d+1) = {}^1/_4(npd + np + d + (1) - nd - n - pd - p)$

$\hat{NP} = {}^1/_4(2{,}806 + 2{,}258 + 2{,}454 + 2{,}068 - 2{,}432 - 2{,}072 - 2{,}640 - 2{,}348) = 0{,}0235$

$\hat{D} = {}^1/_4(n+1)\ (p+1)\ (d-1) = {}^1/_4(d + nd + pd + npd - (1) - n - p - np)$

$\hat{D} = {}^1/_4(2{,}454 + 2{,}432 + 2{,}640 + 2{,}806 - 2{,}068 - 2{,}072 - 2{,}348 - 2{,}258) = 0{,}3965$

$\hat{ND} = {}^1/_4(n-1)\ (p+1)\ (d-1) = {}^1/_4(npd + nd + p + (1) - np - n - pd - d)$

$\hat{ND} = {}^1/_4(2{,}806 + 2{,}432 + 2{,}348 + 2{,}068 - 2{,}258 - 2{,}072 - 2{,}640 - 2{,}454) = 0{,}0575$

$\hat{PD} = {}^1/_4(n+1)\ (p-1)\ (d-1) = {}^1/_4(npd + n + pd + (1) - np - nd - p - d)$

$\hat{PD} = {}^1/_4(2{,}806 + 2{,}072 + 2{,}640 + 2{,}068 - 2{,}258 - 2{,}432 - 2{,}348 - 2{,}454) = 0{,}0235$

$\hat{NPD} = {}^1/_4(n-1)\ (p-1)\ (d-1) = {}^1/_4(npd + n + p + d - (1) - np - nd - pd)$

$\hat{NPD} = {}^1/_4(2{,}806 + 2{,}072 + 2{,}348 + 2{,}454 - 2{,}068 - 2{,}258 - 2{,}432 - 2{,}640) = 0{,}0705$

Obsérvese: $C_1 = {}^1/_4 \quad C_2 = {}^1/_4 \quad C_3 = {}^1/_4 \quad C_4 = {}^1/_4$

$C_5 = -{}^1/_4 \quad C_6 = -{}^1/_4 \quad C_7 = -{}^1/_4 \quad C_8 = -{}^1/_4$

$\Sigma C_i = 0 \quad ; \quad \Sigma C_i^2 = {}^1/_2$

$$SC(\hat{N}) = \frac{r(\hat{N})^2}{\Sigma C_i^2} = \frac{5 \times (0{,}0145)^2}{{}^1/_2} = 0{,}0022$$

$$SC(\hat{P}) = \frac{r(\hat{P})^2}{\Sigma C_i^2} = \frac{5 \times (0{,}2565)^2}{{}^1/_2} = 0{,}6579$$

$$SC(\widehat{NP}) = \frac{r(\widehat{NP})^2}{\Sigma C_i^2} = \frac{5 \times (0{,}0235)^2}{^1/_2} = 0{,}0055$$

$$SC(\widehat{D}) = \frac{r(\widehat{D})^2}{\Sigma C_i^2} = \frac{5(0{,}3965)^2}{^1/_2} = 1{,}5721$$

$$SC(\widehat{ND}) = \frac{r(\widehat{ND})^2}{\Sigma C_i^2} = \frac{5(0{,}0575)^2}{^1/_2} = 0{,}0331$$

$$SC(\widehat{PD}) = \frac{r(\widehat{PD})^2}{\Sigma C_i^2} = \frac{5(0{,}0235)^2}{^1/_2} = 0{,}0055$$

$$SC(\widehat{NPD}) = \frac{r(\widehat{NPD})^2}{\Sigma C_i^2} = \frac{5(0{,}0705)^2}{^1/_2} = 0{,}0497$$

En consecuencia, la tabla del análisis de variancia y considerando los resultados anteriores sería:

F. de V.	G.L.	S.C.	C.M.	F
REPT.	4	33,7220	8,431	77,79**
N	1	0,0022	0,0022	0,02
P	1	0,6579	0,6579	6,09 *
NP	1	0,0055	0,0055	0,05
D	1	1,5721	1,5721	14,56**
ND	1	0,0331	0,0331	0,31
PD	1	0,0055	0,0055	0,05
NPD	1	0,0497	0,0497	0,46
ERROR	28	3,0340	0,1080	
TOTAL	39	39,0820		

El valor de F(1,28, 0,05) = 4.20

Restaría por analizar los promedios de los efectos que fueron significativos así:

$\overline{P}_{40}$ = 2,2565 $\overline{P}_{80}$ = 2,5130

$\overline{D}_{20}$ = 2,1865 $\overline{D}_{50}$ = 2,5830

Puesto que son sólo dos promedios en comparación, la prueba de F es suficiente para discriminar entre los dos "tratamientos" en cuestión. En este caso particular la dosis de 80 kilos de P_2O_5 por hectárea y una densidad de 50.000 plantas son las mejores.

EJERCICIOS

1. Los datos que se dan a continuación corresponden a producción de maíz en kilos por parcela. El estudio corresponde al efecto de asociar maíz con fríjol. El diseño fue de bloques al azar con arreglo factorial 3 x 2, se estudiaron tres variedades de maíz y dos de fríjol.

No.	MAIZ	FRIJOL	I	II	III
1.	D. CABALLO	VIBORAL	0,41	0,52	0,40
2.	D. CABALLO	RADICAL	0,28	0,26	0,32
3.	ICA-V-402	VIBORAL	0,37	0,22	0,41
4.	ICA-V-402	RADICAL	0,26	0,26	0,32
5.	ICA-V-453	VIBORAL	0,40	0,46	0,41
6.	ICA-V-453	RADICAL	0,11	0,18	0,18

2. En experimentos de fertilización de pasto braquiaria se obtuvieron los siguientes rendimientos en toneladas de forraje por hectárea en dos cortes diferentes. El diseño experimental es factorial 3 x 3 y 3 replicaciones en bloques completos al azar.

		1er. Corte		
P_2O_5	CAL	I	II	III
0	0	6,5	7,0	6,8
0	2	6,4	6,6	6,9
0	4	7,0	7,5	6,9
50	0	8,0	8,1	8,4
50	2	8,5	8,6	8,8
50	4	9,0	9,5	10,0
100	0	10,4	9,8	9,8
100	2	11,3	12,0	11,7
100	4	9,8	10,0	9,5

2do. C o r t e

P_2O_5	CAL	I	II	III
0	0	9,5	8,6	9,0
0	2	8,9	8,5	9,5
0	4	9,9	9,0	8,9
50	0	10,0	9,6	8,9
50	2	9,9	10,0	10,1
50	4	10,2	9,9	10,5
100	0	9,8	9,9	10,2
100	2	11,0	11,6	10,9
100	4	13,2	12,8	10,9

3. Los datos que se dan a continuación corresponden a producción de leche, corregida según edad del animal. Se busca evaluar el efecto de la raza Holstein en dos cargas diferentes. El diseño experimental fue completamente al azar, en arreglo factorial 2 x 2 y seis observaciones por combinación de tratamientos. Los datos son a los 15 y 30 días de iniciado el experimento.

	Cebú x Criollo		Cebú x Criollo X Holstein	
Obs.	3 animales por ha.	5 animales por ha.	3 animales por ha.	5 animales por ha.
		15 dias		
1.	168,7	134,9	145,5	102,5
2.	92,2	107,3	144,5	113,0
3.	156,9	148,6	81,1	120,7
4.	162,6	188,9	101,9	122,7
5.	140,3	120,6	108,1	81,0
6.	138,2	130,9	153,7	111,2

30 días

1.	198,4	164,7	172,2	132,4
2.	119.1	136,6	166,5	141,6
3.	166,3	168,7	111,5	166,9
4.	183,3	181,4	153,2	150,9
5.	167,8	191,7	132,1	128,8
6.	180,6	169,0	154,8	123,8

a. En cada caso ejecute tabla de análisis de variancia

b. Si la interacción es significativa, interprétela y realice prueba de comparación múltiple.

4. Se realizaron dos experimentos, con el propósito de evaluar la respuesta de maíz (ICA H-211) a las aplicaciones del nitrógeno, fósforo y potasio. El diseño experimental fue de bloques completos al azar en arreglo factorial 2 x 2 x 2 y cuatro replicaciones. Los datos corresponden a toneladas de maíz por hectárea y son cortesía del Dr. Jaime Lugo, del Distrito del Guamo, Tolima.

Localidad 1

No.	N	P	K	I	II	III	IV
1.	50	20	20	2,60	3,60	2,75	4,45
2.	50	20	40	3,95	3,65	2,60	2,90
3.	50	40	20	3,45	3,53	2,15	2,40
4.	50	40	40	2,75	2,31	1,85	3,36
5.	70	20	20	3,73	2,66	3,15	2,66
6.	70	20	40	3,95	4,15	2,35	2,93
7.	70	40	20	3,95	3,89	3,50	3,31
8.	70	40	40	4,50	3,75	4,02	3,25

Localidad 2

No.	N	P	K	I	II	III	IV
1.	50	20	20	2,80	2,90	2,25	3,30
2.	50	20	40	2,77	2,26	2,41	2,57
3.	50	40	20	3,45	2,00	2,33	2,21
4.	50	40	40	2,09	2,76	2,75	2,70
5.	70	20	20	3,02	3,19	3,15	3,40
6.	70	20	40	1,87	3,35	2,55	3,19
7.	70	40	20	2,35	2,55	2,55	2,46
8.	70	40	40	2,60	2,45	2,17	2,95

a. En cada caso ejecute tabla de análisis de variancia.

b. Interprete los resultados en cada caso.

5. Se realizó un ensayo con el propósito de evaluar la respuesta del híbrido de sorgo SORGHICA-NH301 a las aplicaciones de Nitrógeno y Potasio en suelos del Ariari. Los valores que se observan corresponden a toneladas de sorgo por hectárea. El diseño fue de bloques completos al azar en arreglo factorial 4 x 4.

Tratamiento	N	P_2O_5	I	II	III	IV	Total
1.	0	0	2,47	2,30	2,47	2,43	9,67
2.	0	30	2,50	2,60	2,73	2,50	10,33
3.	0	60	2,73	2,77	2,75	3,08	11,33
4.	0	90	2,50	2,43	2,25	2,40	9,58
5.	25	0	3,10	3,30	3,03	3,23	12,66
6.	25	30	2,98	3,02	3,00	3,00	12,00
7.	25	60	3,50	3,33	3,50	3,33	13,66
8.	25	90	3,37	3,58	3,50	3,63	14,08
9.	50	0	3,57	3,33	3,43	3,08	13,41
10.	50	30	4,37	3,67	3,63	3,83	15,50
11.	50	60	3,60	4,16	4,06	3,33	15,15
12.	50	90	3,83	3,33	3,16	3,50	13,82
13.	75	0	3,90	4,26	4,00	4,33	16,49
14.	75	30	4,33	4,33	4,50	4,33	17,49
15.	75	60	4,23	4,41	4,10	3,92	16,66
16.	75	90	4,40	4,83	4,50	5,10	17,23
			55,38	55,65	54,61	53,42	219.06

a. Ejecute Tabla de análisis de varianza

b. Interprete resultados y estime promedios.

BIBLIOGRAFIA

1. CALZADA J. 1970. Métodos para la Investigación. Editorial Jurídica, S.A. Lima, Perú.

2. COCHRAN, W.G. y G.M. COX. 1957. Experimental designs. John Wiley, Inc. and edition. New York.

3. KEMPTHORNE O. 1952. Design and analysis of experiments. Capítulo 13.

4. SNEDECOR G.W. 1970. Métodos Estadísticos. Trad. al Español.