UNIVERSIDAD AUTÓNOMA AGRARIA

ANTONIO NARRO

DIVISIÓN DE CIENCIA ANIMAL

Diseños Experimentales Especiales para la Investigación Agropecuaria

POR:

Job Hernández Alejandro

MONOGRAFÍA

Presentada como Requisito Parcial para

Obtener el Titulo de:

Ingeniero Agrónomo Zootecnista

Buenavista, Saltillo, Coahuila, México.

Marzo del 2001

UNIVERSIDAD AUTÓNOMA AGRARIA

ANTONIO NARRO

DIVISION DE CIENCIA ANIMAL

Diseños Experimentales Especiales para la Investigación Agropecuaria

Por:

Job Hernández Alejandro

Que somete a la consideración del H. Jurado examinador como requisito parcial para obtener el titulo de:

Ingeniero Agrónomo Zootecnista

APROBADA

Asesor principal

____

MC. Jaime M Rodríguez Del Ángel

Asesor | Asesor

______________________________ ______________________________

ING. José Rodolfo Peña Oranday | ING. Jesús Macias Hernández

El Coordinador de la División de Ciencia Animal

____

ING. José Rodolfo Peña Oranday

Buenavista, Saltillo, Coahuila Marzo del 2001

CONTENIDO

INTRODUCCIÓN

Durante nuestra preparación como profesionales en zootecnia, se contempla como materias básicas la Estadística y los Diseños Experimentales, cátedras que nos inician en el proceso de evaluación de respuestas mediante el método científico, lo cual incluye por supuesto, la utilización de la estadística descriptiva y la probabilidad como herramientas prioritarias. No obstante, las limitantes de tiempo, la magnitud de la carga académica y la falta de equipo entre otras, no dan espacio para la capacitación en técnicas estadísticas especificas utilizadas en el campo laboral que ahora nos ocupa. Actualmente los sistemas computacionales independientemente de facilitarnos los cálculos que implica la utilización de la estadística, nos permiten contar con un gran volumen de información actualizada, la cual si bien es cierto, no esta aplicada directamente a los aspectos agropecuarios gran parte de ella puede ser utilizada mediante inferencia, en esta área.

La presente recopilación de información sobre temas estadísticos de uso agropecuario.Diseño en Parcelas Divididas, Experimentos Factoriales con un tratamiento extra y Experimentos Combinados, nace de la necesidad de ampliar el conocimiento sobre las diferentes formas en que el estudio de un fenómeno de interés puede ser observado, analizado y evaluado, con el propósito de obtener un conocimiento más completo, por otra parte los temas antes mencionados por ser específicos, no son de uso cotidiano y generalmente no son contemplados en los programas analíticos de las materias tradicionales.

A continuación y sin pretender establecer un tratado sobre cada uno de los temas que aquí se mencionan, describiremos brevemente la utilización de los mismos en el campo de la experimentación agrícola y pecuaria.

Diseño Parcelas Divididas. Este diseño contempla los principios de los experimentos factoriales y puede ser incorporado a los diseños básicos; Completamente al Azar, Bloques al Azar o Cuadro Latino. Normalmente es deseable en los siguientes casos.

Cuando los tratamientos relacionados con los niveles de uno o más de los factores necesitan mayores cantidades de material experimental en una unidad experimental que los tratamientos de otros factores.

Otra forma común de utilización se refiere al hecho de considerar la evaluación de un factor como generalizada y la de un segundo como información especifica y con mayor precisión.

También el diseño puede usarse si en una segunda evaluación de un experimento se pretende incorporar un factor adicional con el propósito de aumentar su alcance.

A partir de la información anterior, se puede saber que normalmente existirán diferencias mayores entre los niveles de ciertos factores que entre los niveles de otros.

Experimentos Factoriales con un Tratamiento Extra o Testigo. Cuando se desea evaluar el efecto de dos o mas factores sobre la unidad experimental, los tratamientos son resultantes de la combinación de los niveles dentro de factores, por otra parte es común, el incluir dentro del estudio uno o varios tratamientos testigo, que se asume son representativos del área de influencia donde se desea hacer inferencia. Lo anterior se presta a confusiones, debido a que el tratamiento testigo normalmente es creado y manejado como un tratamiento mas dentro del estudio y difícilmente representa una situación real del área de influencia. Por esta razón se propone que los testigos sean evaluados fuera del estudio general y que la información de respuesta se obtenga de unidades experimentales semejantes, localizadas en el área donde se pretende hacer inferencia y que además sean manejadas de forma tradicional. Todo lo anterior trae consigo una modificación en el modelo estadístico y en la forma de análisis, lo cual se aborda en este tema.

Experimentos combinados. Este tipo de experimentos son comúnmente utilizados con tratamientos simples o factoriales, pero cuando lo mismos son evaluados en diferentes localidades y durante varias épocas o años. Lo anterior permite ampliar las recomendaciones practicas sobre los tratamientos estudiados, como consecuencia de las repeticiones de los experimentos en las localidades y en las estaciones.
El numero de localidades para muestrear adecuadamente una área geográfica en particular en una extensión considerable, es una función de la variabilidad del área, similarmente, cuando las fluctuaciones ambientales de estación a estación son extremas, obliga a que mas estaciones o años sean incluidos en el estudio para muestrear adecuadamente la longitud del ciclo ambiental en una localidad en particular.
Con esta estratificación de localidades y estaciones el investigador podrá hacer mejor inferencia sobre el efecto de tratamientos para una región específica con una gran seguridad de que una diferencia representara los efectos reales y no justamente la variabilidad ocasional.

Por todo lo anterior y deseando que la presente recopilación y ordenamiento de temas aquí expuestos, sirva como una experiencia útil en el ejercicio profesional, se plantean los siguientes objetivos:

- Teniendo en cuenta que los ejemplos contenidos en los temas se refieren a aspectos agropecuarios, creemos se tendrá un conocimiento más acorde respecto a la aplicación de las técnicas estadísticas y su interpretación.

- Considerando lo expresado en el Articulo 85°, IV. Del Reglamento para Alumnos de Licenciatura de la U A A A N, esta monografía sirva como requisito parcial para obtener el Titulo de Ingeniero Agrónomo Zootecnista.

PARCELAS DIVIDIDAS

INTRODUCCIÓN

La designación de parcelas divididas proviene del hecho de que una unidad experimental, tratada bajo algún estímulo, dentro del contexto de algún diseño básico, es particionada para dar cabida a la aplicación de otro factor diferente, que también se desea estudiar. Esto implica que las parcelas divididas se trabajan en arreglos factoriales. Así, un factor es aplicado en forma aleatoria en la parcela mayor o principal, y otro se aleatoriza en sus niveles respectivos, dentro de la parcela mencionada. Esto provoca que la variabilidad entre las unidades experimentales tratadas por los factores no sea homogénea, ya que mientras un factor es aplicado a las parcelas proporcionalmente mayores, otro es estudiado dentro de esas parcelas. Comúnmente, la minimización de la variabilidad para el factor que se aplica dentro de la parcela permite una mejor estimación de su efecto, sacrificando precisión en el estudio del factor aplicado a la parcela principal.

Lo anterior es de importancia práctica en las ciencias agropecuarias, puesto que en ocasiones se tienen estudios acerca del efecto de un factor en general, pero es de interés la respuesta del mismo y otro estímulo, tanto en forma adicional como combinatoria.

El arreglo de campo consiste en asignar a las unidades experimentales los niveles del factor que se considere de menor importancia, todo esto dentro de los principios de un diseño básico, que corresponda a las características de las unidades, y posteriormente aleatorizar dentro de las parcelas mayores los niveles del factor que deseamos sea estudiado con mayor precisión.

* Un arreglo de tratamientos en donde el factor en su parcela tiene mayor importancia (se maneja con menor error).
* Tiene dos errores específicos, para parcela mayor y subparcela.
* Primero se sortea el factor menos importante y este define el diseño, posteriormente se sortea el factor importante como subparcela.
* Las pruebas de comparación de medias se basan en varianzas especificas.

OTROS CONCEPTOS SOBRE PARCELAS DIVIDIDAS

Los diseños de parcelas divididas se usan frecuentemente en experimentos factoriales, tales diseños pueden incorporar uno o más de los diseños completamente aleatorio, de bloques completos aleatorizados o de cuadro latino. El principio básico es éste; las parcelas completas o unidades completas, a las cuales se le aplican niveles de uno o más factores, se dividen en subparcelas o subunidades a las cuales se les aplican niveles de uno o más factores adicionales, de este modo, cada unidad completa se convierte en un bloque para los tratamientos de subunidades.

El diseño de parcelas divididas es deseable en las siguientes situaciones:

Puede usarse cuando los tratamientos relacionados con los niveles de uno o más de los factores necesitan mayores cantidades de material experimental en una unidad experimental que los tratamientos de otros factores. Esto es común en experimentación sobre el campo, el laboratorio, industrial y social, por ejemplo, en un experimento sobre el campo uno de los factores puede ser métodos de preparación del suelo o aplicación de un fertilizante, factores que necesitan ambos, por lo general parcelas o unidades experimentales grandes.

El diseño puede usarse si va incorporarse en un experimento un factor adicional para aumentar su alcance. Por ejemplo, supongamos que el objeto principal de un experimento es comparar los efectos de varios funguicidas como protectores contra infección por una enfermedad. Para aumentar el alcance del experimento, se incluyen en varias variedades de las cuales se sabe que difieren en su resistencia a la enfermedad.

A partir de la información anterior, se puede saber que pueden esperarse diferencias mayores entre los niveles de ciertos factores que entre los niveles de otros. En este caso, las combinaciones de los tratamientos para los factores donde se esperan diferencias grandes pueden asignarse aleatoriamente a las unidades completas, simplemente por comodidad.

El diseño se usa cuando se desea mayor precisión para comparaciones entre ciertos factores, que para otras. Esta es esencialmente la misma que la tercera situación, pero las razones pueden ser diferentes.

EJEMPLO PRACTICO

Los siguientes datos se refieren al número de raíces en acodos de vid en cuatro tiempos de medición (variable endógena), los mismos fueron expuestos al efecto de un factor generalizado que se denomino longitud de vara, el segundo factor y más importante se aplico a la plantación, y se refiere a la dosis de enraizador de un mismo producto. Los resultados son los siguientes:

		Longitud 8 cm				Longitud 12 cm				Longitud 16 cm			
Dosis	**Días**	**7**	**14**	**21**	**28**	**7**	**14**	**21**	**28**	**7**	**14**	**21**	**28**
0.5		3	4	8	10	3	6	12	14	5	7	12	12
1.0		4	3	6	10	3	6	12	10	5	6	13	10
1.5		4	5	9	12	4	10	14	14	6	8	15	15
2.0		6	6	9	13	5	11	15	16	8	10	18	18

PARCELAS DIVIDIDAS EN UN DISEÑO BLOQUES AL AZAR.

Se utilizo este diseño debido a que se observo una variable endógena (tiempo) propia de los muestreos utilizados, es decir se tiene un control local y tenemos que bloquear y no se hizo transformación de datos debido que están reportados en cm.

En el diseño de bloques al azar, los bloques (B) son conjuntos de unidades experimentales dispuestas o seleccionadas con anterioridad a la asignación de tratamientos, control local, de tal manera que la variabilidad existente es minimizada dentro de los bloques y maximizada entre los mismos.

MODELO:

El modelo estadístico para un diseño bloques al azar en parcelas divididas es el siguiente:

$$Y_{ijk} = \mu + B_k + \alpha_i + \in_{ik} + \beta_j + \alpha\beta_{ij} + \delta_{ijk}$$

i = 1,2,3, ..., a (parcela mayor)

j = 1,2,3, ..., b (parcela menor)

k = 1,2,3, ..., r (bloques o repetición)

$\in_{ij}$ = (error experimental parcela mayor)

δ = (error experimental parcela menor)

Parcela mayor Longitud (cm)	Parcela menor Dosis	B1	B2	B3	B4
a1	b1	y111	y112	y113	y114
	b2	y121	y122	y123	y124
	b3	y131	y132	y133	y134
	b4	y141	y142	y143	y144
a2	b1	y211	y212	y213	y214
	b2	y221	y222	y223	y224
	b3	y231	y232	y233	y234
	b4	y241	y242	y243	y244
a3	b1	y311	y312	y313	y314
	b2	y321	y322	y323	y324
	b3	y331	y332	y333	y334
	b4	y341	y342	y343	y344

Concentración de datos.

Parcela mayor longitud	Parcela menor Dosis	B1 7	B2 14	B3 21	B4 28	**yij.**
8 cm	0.5	3	4	8	10	**25**
	1.0	4	3	6	10	**23**
	1.5	4	5	9	12	**30**
	2.0	<u>6</u>	<u>6</u>	<u>9</u>	<u>13</u>	**34**
	yi.k	**17**	**18**	**32**	**45**	
12 cm	0.5	3	6	12	14	**35**
	1.0	3	6	12	10	**31**
	1.5	4	10	14	14	**42**
	2.0	<u>5</u>	<u>11</u>	<u>15</u>	<u>16</u>	**47**
	yi.k	**15**	**33**	**53**	**54**	
16 cm	0.5	5	7	12	12	**36**
	1.0	5	6	13	10	**34**
	2.0	6	8	15	15	**44**

	2.5	8	10	18	18	**54**
	yi.k	**24**	**31**	**58**	**55**	
	y..k	**56**	**82**	**143**	**154**	

Donde:

i = 3

j = 4

k = 4

Cuadro de concentración para sumas de cuadrados simples.

	b1	b2	b3	b4	**Yi..**
Aa1	25	23	30	34	112
Aa2	35	31	42	41	155
Aa3	36	34	44	54	168
Y.j.	96	88	116	135	435 **Y...**

Análisis de varianza

F.V.	g.l.	Sc	CM	Fc	$F\alpha$ .05 .01
Bloques	(r-1) = 3	559.895	183.631	23.678 **	4.53 9.78
Longitud (A)	(a-1) = 2	107.375	53.687	6.812 *	5.14 10.92
E exp. Longitud	(a-1)(r-1) =6	47.291	7.881		
Dosis (B)	(b-1) = 3	111.229	37.076	39.192 **	2.96 4.60
Longitud x Dosis (AB)	(a-1)/b-1) = 6	7.458	1.243	1.313 NS	2.46 3.56
E exp. Dosis	a(b-1)(r-1) = 27	25.564	0.946		
Total	Abr-1 = 47	858.812			

De acuerdo a los resultados obtenidos en el análisis de varianza podemos observar alta significancía entre los bloques (tratamientos) y el factor B (dosis), por otro lado observamos que el factor A (longitud) es significante, sin embargo, la interacción AB se ve que es no significante.

Como en ocasiones anteriores, la significación de los efectos está en función del contraste entre las estadísticas de prueba y los valores tabulares de Snedecor, También es necesario considerar los ajustes de medias.

$$Sc_{longitud} = \sum_{i=1}^{a} \frac{Yi..^2}{br} - \frac{Y...^2}{abr}$$

$$Sc_{longitud} = \frac{112.00^2 + 155.00^2 + 168.00^2}{4x4} - \frac{435.00^2}{3x4x4}$$

$$Sc_{longitud} = 107.375$$

$$Sc_{dosis} = \sum_{j=1}^{b} \frac{Y.j.^2}{ar} - \frac{Y...^2}{abr}$$

$$Sc_{dosis} = \frac{96.00^2 + 88.00^2 + 116.00^2 + 135.00^2}{3x4} - \frac{435.00^2}{3x4x4}$$

$$Sc_{dosis} = 111.229$$

$$Sc_{bloque} = \sum_{k=1}^{r} \frac{Y..k^2}{ab} - \frac{Y...^2}{abr}$$

$$Sc_{bloque} = \frac{56.00^2 + 82.00^2 + 143.00^2 + 154.00^2}{3x4} - \frac{435.00^2}{3x4x4}$$

$$Sc_{bloque} = 559.895$$

$$ScEPM_{longitud} = \sum_{i=1}^{a} \sum_{k=1}^{r} \frac{Yi.k^2}{b} - \frac{Y...^2}{abr} - (ScA + Scbloque)$$

$$ScEPM_{longitud} = \frac{17.00^2 + .. + 55.00^2}{4} - \frac{435.00^2}{abr} - (107.375 + 559.895)$$

$$ScEPM_{longiutd} = 47.291$$

$$ScAxB_{longitudxdosis} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{Yij.^2}{r} - \frac{Y...^2}{abr} - (ScA + ScB)$$

$$ScAxB_{longitudxdosis} = \frac{25.00^2 + .. + 54.00^2}{4} - \frac{435.00^2}{3x4x4} - (107.375 + 111.229)$$

$$ScAxB_{longitudxdosis} = 7.458$$

$$Sc_{Total} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{r} Yijk^2 - \frac{Y...^2}{abr}$$

$$Sc_{Total} = 3.0^2 + .. + 18.00^2 - \frac{435.00^2}{3x4x4}$$

$$Sc_{Total} = 858.812$$

$$ScEpm_{dosis} = Sc_{Total} - (ScA + ScB + Scbloques + ScEPM + ScAxB)$$
$$ScEpm_{dosis} = 858.912 - (107.375 + 111.229 + 559.895 + 47.291 + 7.458)$$
$$ScEpm_{dosis} = 25.564$$

AJUSTE PARA COMPARACIÓN DE MEDIAS.

Es importante considerar alguna prueba estadística que nos permita estratificar en función de los objetivos, el comportamiento de los diferentes efectos. Sin embargo, es necesario tomar en cuenta que existen dos cuadrados medios de error, por lo que es indispensable efectuar los siguientes ajustes para la comparación de medias.

Efectos	Totales	Medias	Ajuste
Parcela Mayor	Yi..	$\bar{Y}.../br$	$\delta\bar{x} = \sqrt{\frac{Ea}{rb}} = \sqrt{\frac{7.881}{16}} = 0.701$
Ai = Aj	1.- 112.00	7.00 b	
	2.- 155.00	9.68 a	
	3.- 168.00	10.50 a	
Parcela Menor	Y.j.	Y.j./ar	$\delta\bar{x} = \sqrt{\frac{Eb}{ra}} = \sqrt{\frac{0.946}{12}} = 0.280$
Bi = Bj	1.- 96.00	8.00 c	
	2.- 88.00	7.33 c	
	3.-116.00	9.66 b	
	4.-135.00	11.25 a	
Entre tratamientos	Aij.	Yij./r	
AiBi = AjBj	1.- 25.00	6.25 EF	
	2.- 23.00	5.75 F	
	3.- 30.00	7.50 DE	
	4.- 34.00	8.50 CD	
	5.- 35.00	8.75 CD	
	6.- 31.00	7.75 CD	
	7.- 42.00	10.50 B	
	8.- 47.00	11.75 B	
	9.- 36.00	9.00 C	
	10.- 34.00	8.50 CD	
	11.- 44.00	11.00 B	

12.- 54.00 13.50 A

$$\delta\bar{x} = \sqrt{\frac{(b-1)Eb + Ea}{rb}} = \sqrt{\frac{(4-1)0.946 + 7.881}{4x4}} = 0.818$$

POLINOMIOS ORTOGONALES (DOSIS)

Conceptos

1.- Un polinomio no puede ser mayor de t – 1, pero no necesariamente deberá agotar esa magnitud.

2.- Los tratamientos deberán ser variables cuantitativas y preferentemente equidistantes.

Con base a lo anterior, el modelo para polinomios ortogonales será:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_{t-1} x^{t-1} + \in_i$$

96	88	116	135	$\sum_{i=1}^{t} C_{ij} y_{i.}$	$r\sum_{i=1}^{t} C_{ij}^{2}$	$Sc = \frac{\left(\sum_{i=1}^{t} C_{ij} y_{i.}\right)^2}{r\sum_{i=1}^{t} C_{ij}^{2}}$	$CM = \frac{ScEf}{1}$	$Fc = \frac{CMEf}{CM_{EEXP} = 0.946}$
-3	-1	1	3	145	240	87.604	87.604	90.604**
1	-1	-1	1	27	48	15.187	15.187	16.187**
-1	3	-3	1	-45	240	8.437	8.437	8.434*

DEFINICIÓN DE TENDENCIA DEL POLINOMIO ORTOGONAL.

En primer lugar, se debe definir el grado máximo del polinomio (t –1), en este caso es **cubica**. Luego describir los coeficientes independientes, para el cálculo de suma de cuadrados (tabla de coeficientes ortogonales).

MODELO DE SUPERFICIE:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3$$

$$\hat{y}_i = \alpha_0 + \alpha_1 \mu_1 P_1 + \alpha_2 \mu_2 P_2 + \alpha_3 \mu_3 P_3$$

$$\alpha_0 = \frac{y...}{abr} = \frac{435.00}{48} = 9.062$$

$$\alpha_1 = \frac{\sum_{i=1}^{t} C_{1ij} y_{i.}}{\sum_{i=1}^{t} C_{1ij}^{2}} = \frac{145.00}{240.00} = 0.604$$

$$\alpha_2 = \frac{\sum_{i=1}^{t} C_{2ij} y_{i.}}{\sum_{i=1}^{t} C_{2ij}^{\ 2}} = \frac{27.00}{48.00} = 0.562$$

$$\alpha_3 = \frac{\sum_{i=1}^{t} C_{3ij} y_{i.}}{\sum_{i=1}^{t} C_{3ij}^{\ 2}} = \frac{-45.00}{240.00} = -0.187$$

TRANSFORMADORES LOGARITMICOS VALORES TABULARES

$$\mu_1 = 2$$
$$\mu_2 = 1$$
$$\mu_3 = 10/3 = 3.333$$

$$P_1 = \frac{x_i - \bar{x}}{d}$$

$\bar{x} =$ Es una media de los estímulos.

d = Espaciamiento que hay entre los estímulos.

$$\bar{x} = \frac{0.5 + 1.0 + 1.5 + 2.0}{4} = 1.25$$

d = 0.5

$$P_1 = \frac{x_i - 1.25}{4}$$

$$P_1 = 2x_i - 2.5$$

$$P_2 = P_1^2 - \left(\frac{t^2 - 1}{12}\right)$$

$$P_2 = (2x_i - 1.25)^2 - \left(\frac{4^2 - 1}{12}\right)$$

$$P_2 = 4x_i^{\ 2} - 10x_i + 6.25 - 1.25$$

$$P_2 = 4x_i^{\ 2} - 10x_i + 5$$

$$P_3 = P_1^3 - \left(\frac{3t^2 - 7}{20}\right) P_1$$

$$P_3 = 8x_i^{\ 3} - 30x_i^{\ 2} + 37.5x_i - 15.625 - \left(\frac{3(4)^2 - 7}{20}\right)(2x_i - 2.5)$$

$$P_3 = 8x_i^3 - 30x_i^2 + 37.5x_i - 15.625 - 2.05(2x_i - 2.5)$$
$$p_3 = 8x_i^3 - 30x_i^2 + 37.5x_i - 15.625 - 4.1x_i + 5.125$$
$$p_3 = 8x_i^3 - 30x_i^2 + 33.4x_i - 10.5$$

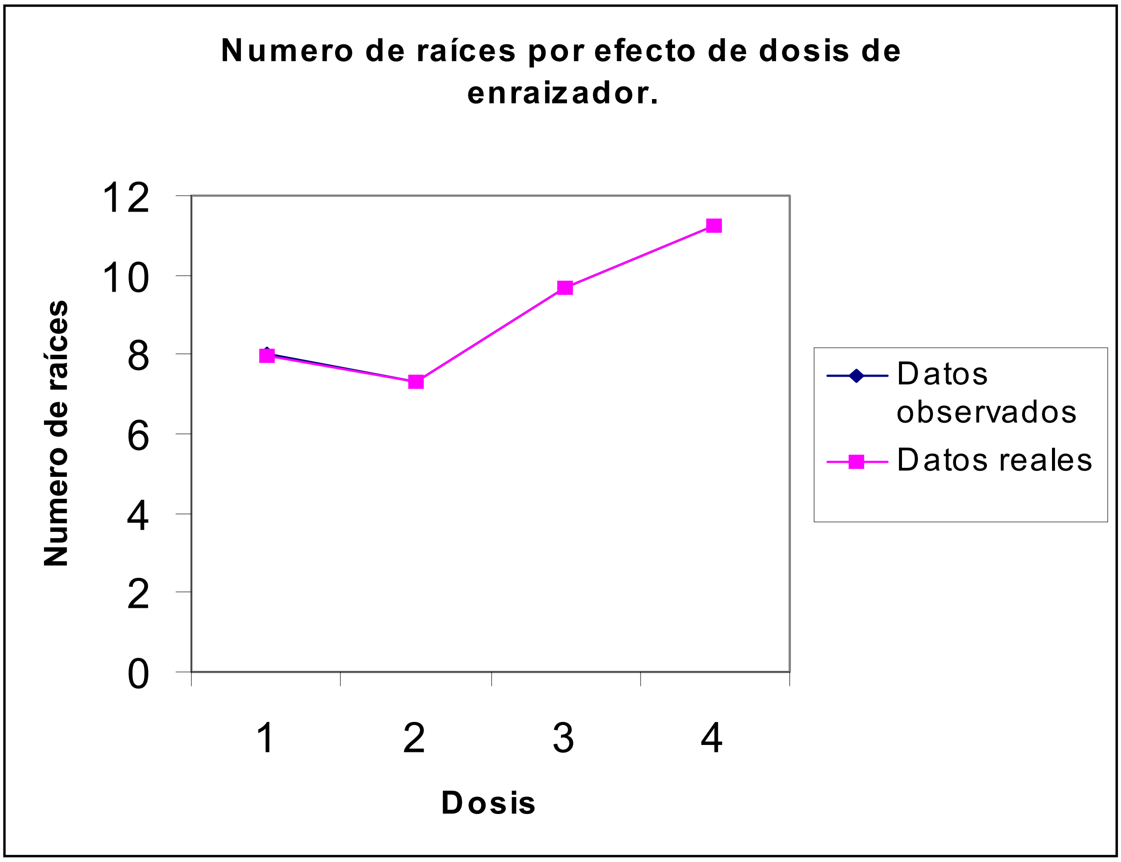
$$\hat{Y}_i = 9.062 + (0.604)(2)(2x_i - 2.5) + (0.562)(1)(4x_i^2 - 10x_i + 5) + (-0.187)(3.333)$$
$$(8x_i^3 - 30x_i^2 + 33.4x_i - 10.5)$$

$$\hat{Y}_i = 9.062 + 2.416x_i - 3.02 + 2.248x_i^2 - 5.62x_i + 2.81 - 4.986x_i^3 + 18.698x_i^2 - 20.817x_i + 6.544$$
$$\hat{Y}_i = 15.396 - 24.021x_i + 20.946x_i^2 - 4.986x_i^3$$

Esta expresión, independientemente de describir la tendencia de respuesta, permite predecir valores no considerados para las variables (X y Y) dentro de los rangos estudiados.

La importancia práctica de este tipo de expresiones radica en que la relación optima entre las variables relacionadas, en función de los objetivos, es descubierta fácilmente.

DOSIS	DATOS OBSERVADOS $\bar{Y}.j.$	DATOS REALES $\hat{Y}.j.$
0.5	8.00	7.998
1.0	7.33	7.335
1.5	9.66	9.665
2.5	11.25	11.250



POLINOMIOS ORTOGONALES (TIEMPO)

Con base a lo anterior, el modelo para polinomios ortogonales será:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_{t-1} x^{t-1} + \in_i$$

56	82	143	154	$\sum_{i=1}^{t} C_{ij} y_{i.}$	$r\sum_{i=1}^{t} C_{ij}^2$	$Sc = \frac{\left(\sum_{i=1}^{t} C_{ij} y_{i.}\right)^2}{r\sum_{i=1}^{t} C_{ij}^2}$	$CM = \frac{ScEf}{1}$	$Fc = \frac{CMEf}{CM_{EEXP} = 0.946}$
-3	-1	1	3	355	240	525.104	525.104	555.078 **
1	-1	-1	1	-15	48	4.687	4.687	4.955 *
-1	3	-3	1	-85	240	30.104	30.104	31.822 **

En primer lugar, se debe definir el grado máximo del polinomio (t –1), en este caso es **cubica**. Luego describir los coeficientes independientes, para el cálculo de suma de cuadrados (tabla de coeficientes ortogonales).

MODELO DE SUPERFICIE:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3$$

$$\hat{y}_i = \alpha_0 + \alpha_1 \mu_1 P_1 + \alpha_2 \mu_2 P_2 + \alpha_3 \mu_3 P_3$$

$$\alpha_0 = \frac{y...}{abr} = \frac{435.00}{48} = 9.062$$

$$\alpha_1 = \frac{\sum_{i=1}^{t} C_{1ij} y_{i.}}{\sum_{i=1}^{t} C_{1ij}{}^2} = \frac{355.00}{240.00} = 1.479$$

$$\alpha_2 = \frac{\sum_{i=1}^{t} C_{2ij} y_{i.}}{\sum_{i=1}^{t} C_{2ij}{}^2} = \frac{-15.00}{48.00} = -0.312$$

$$\alpha_3 = \frac{\sum_{i=1}^{t} C_{3ij} y_{i.}}{\sum_{i=1}^{t} C_{3ij}{}^2} = \frac{-85.00}{240.00} = -0.354$$

TRANSFORMADORES LOGARITMICOS VALORES TABULARES.

$\mu_1 = 2$

$\mu_2 = 1$

$\mu_3 = 10/3 = 3.333$

$$P_1 = \frac{x_i - \bar{x}}{d}$$

$\bar{x} =$ Es una media de los estímulos.

d = Espaciamiento que hay entre los estímulos.

$$\bar{x} = \frac{7 + 14 + 21 + 28}{4} = 17.5$$

d = 7

$$P_1 = \frac{x_i - 17.5}{7}$$

$$P_1 = 0.142 x_i - 2.5$$

$$P_2 = P_1^2 - \left(\frac{t^2 - 1}{12} \right)$$

$$P_2 = (0.142 x_i - 2.5)^2 - \left(\frac{4^2 - 1}{12} \right)$$

$$P_2 = 0.0201 x_i^{\,2} - 0.710 x_i + 6.25 - 1.25$$

$$P_2 = 0.0201x_i^{\,2} - 0.710x_i + 5$$

$$P_3 = P_1^3 - \left(\frac{3t^2 - 7}{20}\right)P_1$$

$$P_3 = 0.0028x_i^{\,3} - 0.1510x_i^{\,2} + 2.6625x_i - 15.625 - \left(\frac{3(4)^2 - 7}{20}\right)(0.142x_i - 2.5)$$

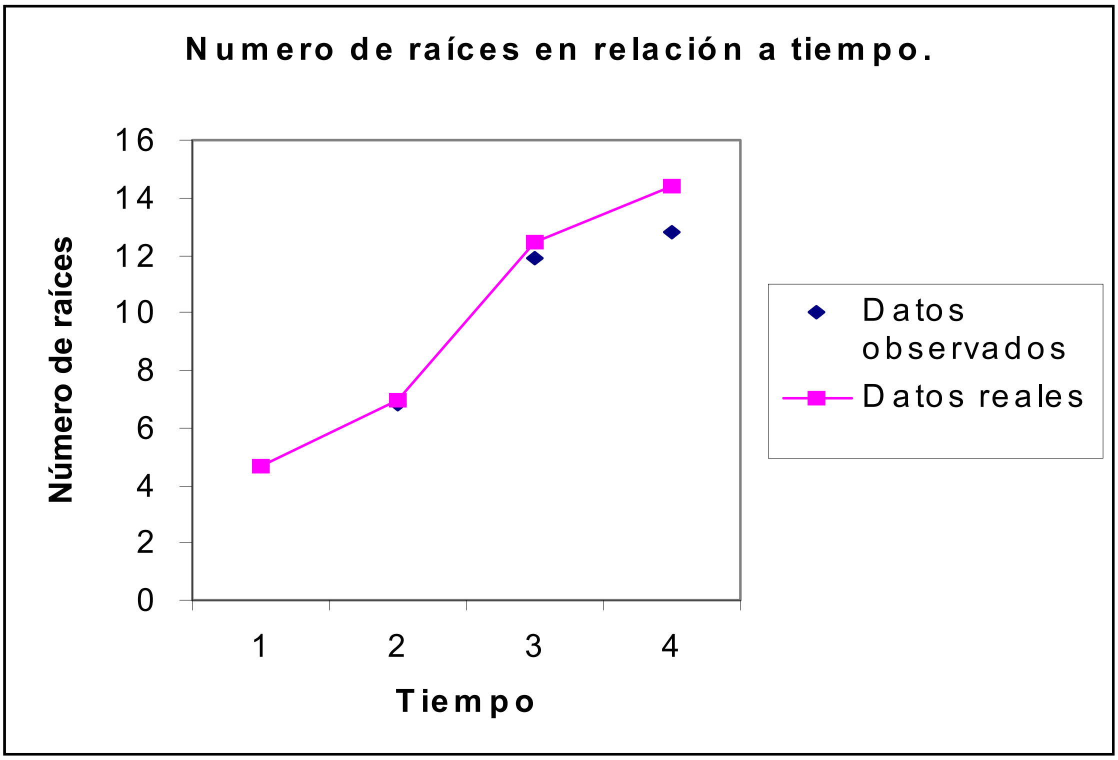
$$P_3 = 0.0028x_i^{\,3} - 0.1510x_i^2 + 2.6625x_i - 15.625 - 2.05(0.142x_i - 2.5)$$
$$p_3 = 0.0028x_i^3 - 0.1510x_i^2 + 2.6625x_i - 15.625 - 0.2911x_i + 5.125$$
$$p_3 = 0.0028x_i^3 - 0.1510x_i^2 + 2.3714x_i - 10.5$$

$$\hat{Y}_i = 9.062 + (1.479)(2)(0.142x_i - 2.5) + (-0.312)(1)(0.0201x_i^{\,2} - 0.710x_i + 5) + (-0.354)(3.333)$$
$$(0.0028x_i^{\,3} - 0.1510x_i^{\,2} + 2.3714x_i - 10.5)$$

$$\hat{Y}_i = 9.062 + 0.420x_i - 7.395 - 0.0062x_i^{\,2} + 0.2215x_i - 1.56 - 0.0033x_i^3 + 0.178x_i^2 - 2.797x_i + 12.388$$

$$\hat{Y}_i = 12.495 - 2.1555x_i + 0.1718x_i^2 - 0.0033x_i^3$$

TIEMPO	DATOS OBSERVADOS $\bar{Y}.j.$	DATOS REALES $\hat{Y}.j.$
7	4.666	4.692
14	6.833	6.935
21	11.916	12.432
28	12.833	14.390



POLINOMIOS ORTOGONALES (LONGITUD)

Con base a lo anterior, el modelo para polinomios ortogonales será:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_{t-1} x^{t-1} + \in_i$$

112	155	168	$\sum_{i=1}^{t} C_{ij} y_{i.}$	$r\sum_{i=1}^{t} C_{ij}^{\ 2}$	$Sc = \frac{\left(\sum_{i=1}^{t} C_{ij} y_{i.}\right)^2}{r\sum_{i=1}^{t} C_{ij}^{\ 2}}$	$CM = \frac{ScEf}{1}$	$Fc = \frac{CMEf}{CM_{EEXP} = 0.946}$
-1	0	1	56	32	98.000	98.000	103.594 **
1	-2	1	30	96	9.375	9.375	9.910 **

En primer lugar, se debe definir el grado máximo del polinomio (t –1), en este caso es **cuadratica**. Luego describir los coeficientes independientes, para el cálculo de suma de cuadrados (tabla de coeficientes ortogonales).

MODELO DE SUPERFICIE:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$$

$$\hat{y}_i = \alpha_0 + \alpha_1 \mu_1 P_1 + \alpha_2 \mu_2 P_2$$

$$\alpha_0 = \frac{y...}{abr} = \frac{435.00}{48} = 9.062$$

$$\alpha_1 = \frac{\sum_{i=1}^{t} C_{1ij} y_{i.}}{\sum_{i=1}^{t} C_{1ij}^{\ 2}} = \frac{56.00}{32.00} = 1.75$$

$$\alpha_2 = \frac{\sum_{i=1}^{t} C_{2ij} y_{i.}}{\sum_{i=1}^{t} C_{2ij}^{\ 2}} = \frac{30.00}{96.00} = 0.312$$

TRANSFORMADORES LOGARITMICOS VALORES TABULARES.

$\mu_1 = 2$

$\mu_2 = 1$

$$P_1 = \frac{x_i - \bar{x}}{d}$$

$\bar{x} =$ Es una media de los estímulos.

d = Espaciamiento que hay entre los estímulos.

$$\bar{x} = \frac{8+12+16}{3} = 12$$

d = 4

$$P_1 = \frac{x_i - 12}{4}$$

$$P_1 = 0.25x_i - 3$$

$$P_2 = P_1^2 - \left(\frac{t^2 - 1}{12}\right)$$

$$P_2 = (0.25x_i - 3)^2 - \left(\frac{4^2 - 1}{12}\right)$$

$$P_2 = 0.0625x_i^{\ 2} - 1.5x_i + 9.0 - 1.25$$

$$P_2 = 0.0625x_i^{\ 2} - 1.5x_i + 7.75$$

$$\hat{Y}_i = 9.062 + (1.175)(1)(0.25x_i - 3) + (0.312)(3)(0.0625x_i^{\ 2} - 1.5x_i + 7.75)$$

$$\hat{Y}_i = 9.062 + 0.4375x_i - 5.25 + 0.0585x_i^{\ 2} - 1.404x_i + 7.254$$

$\hat{Y}_i = 11.066 - 0.9665x_i + 0.0585x_i^2$

LONGITUD	DATOS OBSERVADOS $\bar{Y}.j.$	DATOS REALES $\hat{Y}.j.$
8	7.000	7.078
12	9.687	9.720
16	10.500	10.578

Conclusiones. El análisis de varianza indica diferencia altamente significativa para el factor A (longitud) y para el factor B (dosis). La prueba de medias indica que la dosis 4 (2.0) de enraizador provoca la formación de mayor número de raíces y por lo tanto mayor altura de la vareta. El tiempo fue un factor que también tuvo influencia en el número de raíces, es decir a medida que transcurrió el tiempo se incremento el número de raíces.

LITERATURA REVISADA

Castillo P.J, J.G.Arias. 1998. Estadística inferencial básica. Grupo editorial Iberoamérica. México.

Cochran. G. William. 1980. Diseños experimentales. Editorial Trillas. México

Infante G. S. 1997. Métodos estadísticos. Editorial Trillas. México.

Kreyszig Erwin. 1979. Introducción a la estadística matemática. Editorial Limusa. México

Montgomery D. C.1991.Diseño y análisis de experimentos. Grupo editorial Iberoamérica. México.

Reyes C. P. 1978 Diseños de experimentos agrícolas. Editorial Trillas. México.

Rodríguez del A. J. 1991. Métodos de investigación pecuaria. Editorial Trillas. México.

Snedecor W. George, W. G. Cochran. 1979 métodos estadísticos. Editorial Continental. México.

Steel G.D Robert, J. H. Torrie.1981. 2ª. Principles and procedures of statistics a biometrical approach. 2ª. Ed. Editorial Mc Graw-Hill. USA.

EXPERIMENTOS FACTORIALES CON UN TRATAMIENTO EXTRA

INTRODUCCION

Los experimentos factoriales son arreglos de tratamientos que permiten aplicar de una sola vez una serie de estímulos o tratamientos que se consideran intervengan en la respuesta dada por unidad experimental.

Los experimentos factoriales son de importancia práctica, ya que permiten el estudio de un estimulo tal y su respuesta combinatoria respecto de otras condiciones generadas por la interacción con otros factores, dando así información más completa, aun cuando los efectos interaccionados no sean significativos. (Rodríguez, 1991)

La experimentación factorial puede ser adecuada en trabajos de exploración, donde el objeto es determinar rápidamente los efectos independientes, de cierto número de factores dentro de un intervalo específico, en investigaciones de las interacciones entre los efectos de varios factores, en experimentos diseñados para poder llegar a recomendaciones que deben aplicarse a una gran variedades de condiciones. (Cocharn, cox, 1983)

Se pueden realizar experimentos factoriales con diferentes diseños, como completamente al azar ó bloques al azar, análisis de covarianza, parcelas divididas etc. En este caso en particular hablaremos de experimentos factoriales con un tratamiento extra, considerando que este tema por lo general no esta incluido en los textos tradicionales.

En la investigación agropecuaria tradicionalmente dentro del planteamiento general del estudio se considera un tratamiento testigo, el cual asumimos es representativo tradicional o cero, de la respuesta que hasta ese momento se tiene en el área de influencia. Lo anterior puede ser confuso debido a que comúnmente los testigos son creados y manejados por las mismas personas que plantean el estudio y puede no ser reflejo de la situación real. Por esta razón aquí se propone que los testigos sean avaluados fuera del trabajo general y que la información de los mismos se obtenga de unidades experimentales semejantes, pero que estén localizadas en el área donde se pretende hacer inferencia y que sean manejadas en la forma tradicional. Cabe aclarar que él numero de repeticiones en el testigo no necesariamente tendrá que ser el mismo que se planteo en el estudio general, lo cual permite tener una mayor precisión en la evaluación original.

ARREGLO FACTORIAL CON TRES FACTORES Y UN TRATAMIENTO EXTRA.

El modelo estadístico de un arreglo factorial con tres factores y un tratamiento extra, en un diseño de bloques al azar es el siguiente:

$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \lambda_k + \alpha\lambda_{ik} + \beta\lambda_{jk} + \alpha\beta\lambda_{ijk} + B_l + to + \varepsilon_{ijkl}$

i = 1,2,3,... a

j = 1,2,3,... b

k = 1,2,3,.... c

l = 1,2,3,......r (numero de bloques),* el testigo puede tener q repeticiones.

Los efectos producto de la combinación de los factores, en este caso tres, dan como resultado que la partición de los tratamientos sea en siete elementos. El ANVA quedará de la siguiente manera.

F.V	Gl	Sc*	CM	FC
Test / Fact	1		ScT / glT	CMT/CMEE
Bloques (B)	r – 1		ScB /glB	CMB/CMEE
Factor A	a – 1		Sc A / gl A	CMA/CMEE
Factor B	b – 1		SC *B* /gl *B*	CM*B*/CMEE
Interaccion AxB	(a – 1)(b – 1)		ScAxB/glAxB	CMAB/CME
Factor C	c - 1		Sc C / gl C	CMC /CMEE
Interaccion AxC	(a – 1)(c – 1)		ScAxC/glAxC	CMAC/CMEE
Interaccion BxC	(b – 1)(c – 1)		ScBxC/glBxC	CMBC/CMEE

AxBxC Error experimental total	(a-1)(b-1)(c-1) (abc – 1)(r-1) abcr – 1		ScABC/glABC ScEE / glEE	CMABC/CMEE

* Las formulas de sumas de cuadrados se definen en el ejemplo que se presenta a continuación.

EJEMPLO PRACTICO CON TRES FACTORES Y UN TRATAMIENTO EXTRA

La digestibilidad porcentual de cuatro forrajes fue evaluada con rumiantes fistulados, utilizando tres formas físicas de presentación (trozo, molido y rolado) y en tres tiempos después de la ingesta (3, 6 y 9), la edad de los animales fue variable por lo que se utilizó un control local (bloques),

Por otro lado existe un testigo para digestibilidad general con los siguientes datos, 22.3, 21.4, 23.2, 17.9, 21.3, 25.5, 26.1, 24.5.

			Edades		
Forrajes	Proceso	horas	I	II	III
		3	12.2	12.5	14.1
	trozo	6	13.9	15.1	15.7
		9	15.2	16.3	16.5
trigo		3	22.1	19.1	17.2
	molido	6	23.2	23.2	19.6
		9	24.5	25.9	25.8
		3	24.2	20.1	18.4
	rolado	6	24.5	24.1	21.6
		9	25.7	26.3	27.4
		3	13.5	13.4	14.7
	trozo	6	14.2	14.1	15.6
		9	16.1	15.9	16.8
Avena		3	21.2	18.5	17.4
	Molido	6	22.9	22.6	20.1
		9	24.1	25.6	25.6

		3	23.9	19.9	18.6
	rolado	6	23.9	23.6	21.4
		9	24.5	26.9	26.8

Forraje	proceso	horas	I	II	III
		3	12.9	12.6	13.6
	Trozo	6	13.4	13.5	14.2
		9	15.4	15.4	15.8
Cebada		3	19.8	17.3	16.8
	molido	6	21.0	21.0	19.3
		9	23.1	21.2	23.2
		3	21.3	18.5	17.5
	rolado	6	22.6	22.3	21.5
		9	24.5	23.4	24.2
		3	14.1	15.1	15.9
	trozo	6	15.2	17.1	17.1
		9	16.9	18.7	19.1
Kochia		3	23.0	22.3	21.6
	molido	6	25.6	25.3	25.3
		9	28.7	28.1	27.9
		3	24.5	24.1	23.6
	rolado	6	26.1	26.4	26.3
		9	29.1	29.1	30.1

Debido a la forma en que se evaluó la respuesta (porcentaje), los datos no cumplen cabalmente con los supuestos del análisis de varianza, es decir no se distribuyen normalmente de menos infinito a infinito, por lo que es necesario hacer una transformación, que en este caso parece ser que logaritmo es la mas adecuada por las características porcentuales.

Datos transformados a logaritmo.

Forrajes	proceso	horas	I	II	III	yijk..
		3	1.086	1.096	1.149	**3.331**
	trozo	6	1.143	1.178	1.195	**3.516**
		9	1.181	1.212	1.217	**3.610**
Trigo		3	1.344	1.281	1.235	**3.860**
	Molido	6	1.365	1.365	1.292	**4.022**
		9	1.389	1.413	1.411	**4.213**
		3	1.383	1.303	1.264	**3.950**
	rolado	6	1.389	1.382	1.334	**4.105**
		9	1.409	1.419	1.437	**4.265**
		3	1.130	1.127	1.167	**3.424**
	trozo	6	1.152	1.149	1.193	**3.494**
		9	1.206	1.201	1.225	**3.632**
avena		3	1.326	1.267	1.240	**3.833**
	molido	6	1.359	1.354	1.303	**4.016**
		9	1.382	1.408	1.408	**4.198**
		3	1.378	1.298	1.269	**3.945**
	rolado	6	1.378	1.372	1.330	**4.080**
		9	1.389	1.429	1.428	**4.246**
		3	1.110	1.100	1.133	**3.343**
	trozo	6	1.127	1.130	1.152	**3.409**
		9	1.187	1.187	1.198	**3.572**
		3	1.296	1.238	1.225	**3.759**
cebada	molido	6	1.322	1.322	1.285	**3.929**
		9	1.363	1.326	1.365	**4.054**
		3	1.328	1.267	1.243	**3.838**
	rolado	6	1.354	1.348	1.332	**4.034**
		9	1.389	1.369	1.383	**4.141**
		3	1.149	1.178	1.201	**3.528**
	trozo	6	1.181	1.232	1.232	**3.645**
		9	1.227	1.271	1.281	**3.779**

kochia		3	1.361	1.348	1.334	**4.043**
	molido	6	1.408	1.403	1.403	**4.214**
		9	1.457	1.448	1.445	**4.350**
		3	1.389	1.382	1.372	**4.143**
	rolado	6	1.416	1.421	1.419	**4.256**
		9	1.463	1.463	1.478	**4.404**
Y...l			**46.916**	**46.687**		**46.578**
Y.... **140.181**						

Yij..

	a1	a2	a3	a4	$\mathbf{Y_{.j..}}$
b_1	10.457	10.550	10.324	10.952	**42.283**
b_2	12.095	12.047	11.742	12.607	**48.491**
b_3	12.320	12.271	12.013	12.803	**49.407**
$\mathbf{Y_{i...}}$	**34.872**	**34.868**	**34.079**	**36.362**	**140.181**
					Y…

Yi.k.

	a_1	a_2	a_3	a_4	$\mathbf{Y_{..k}}$
c_1	11.141	11.202	10.940	11.714	**44.997**
c_2	11.643	11.590	11.372	12.115	**46.720**
c_3	12.088	12.076	11.767	12.533	**48.464**
$\mathbf{Y_{i...}}$	**34.872**	**34.868**	**34.079**	**36.362**	**140.181**

Y.jk.

	b_1	b_2	b_3	$\mathbf{Y_{..k}}$
c_1	13.626	15.495	15.876	**44.997**
c_2	14.064	16.181	16.475	**46.720**
c_3	14.593	16.815	17.056	**48.464**
$\mathbf{Y_{.j..}}$	**42.283**	**48.491**	**49.407**	**140.181**

Testigo 1.348 + 1.330 + 1.365 + 1.252 + 1.328 + 1.406 + 1.416 + 1.389 = 10.834

ANÁLISIS DE VARIANZA CON UN TRATAMIENTO EXTRA.

F.V	G.L	SC	CM	FC	Fα
Test &fact	1	0.023	0.023	23.000	Fα1 y 70gl 3.980**
Factor A	3	0.101	0.033	33.000	Fα3 y 70gl 2.740**
Factor B	2	0.834	0.417	417.000	Fα2 y 70gl 3.130**
I A x B	6	0.003	0.0005	0.500	Fα6 y 70gl 2.230NS
Factor C	2	0.166	0.083	830.000	Fα2 y 70gl 3.313**
I A x C	6	0.001	0.0001	0.100	Fα6 y 70gl 2.230NS
I B x C	4	0.004	0.001	1.000	Fα4 y 70gl 2.500NS
I AxBxC	12	0.002	0.00016	0.160	Fα12y70gl 1.900NS
Bloques	2	0.001	0.0005	0.500	Fα2 y 70gl 3.130NS
Tratram	35	1.111	0.0317	31.700	Fα35y70gl **
Error (EE)	70	0.076	0.001		
Total	108	1.211			

La relación testigo-factorial es altamente significativo (p< 0.01)

$$SCto = \frac{to2}{q} + \sum_{l=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{r} \frac{y_{ijkl}^2}{abcr} - \frac{(y\ldots + to)^2}{abcr + q}$$

$$= \frac{10.834^2}{8} + \frac{140.181^2}{4x3x3x3} - \frac{(140.181 + 10.834)^2}{(108) + 8}$$

$$= 0.023$$

Cmto = Scto ÷ Glto = 0.023 ÷ 1 = 0.023

Fcto = Cmto ÷CMEE = 0.023 ÷0.001

$$Sctot = \sum_{l=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum^{r} y_{ijkl}^2 + \sum_{t=1}^{q} ti^2 - \frac{(y\ldots + to)^2}{abcr + q} =$$

$$= (1.086^2 + \ldots + 1.478^2) + (1.348^2 + ..+1.416^2) - \frac{(140.181+10.834)^2}{108 + 8} = 1.211$$

Factor A = Forraje

Factor B = Tipo de procesamiento

Factor C = tiempo despues de la ingesta.

A X B = interaccion forraje x tipo de proceso

A X C = interaccion forraje x tiempo despues de la ingesta.

B X C = interaccion tipo de procesamiento x tiempo despues de la ingesta

A X B X C = interaccion forraje x tipo de proceso x tiempo despues de la ingesta

Bloque = Edad.

To = efecto del testigo

$$Sc\ A = \sum_{I=1}^{a} \frac{y_{i}...^2}{bcr} - \frac{y^2...}{abcr} = \frac{(34.872^2 + \ldots 36.362^2)}{3x3x3} - \frac{140.181^2}{108} = 0.101$$

$$Sc\ B = \sum_{j=1}^{b} \frac{y_{.j}..^2}{acr} - \frac{y^2...}{abcr} = \frac{(42.283^2 + \ldots 36.362^2)}{4x3x3} - \frac{140.181^2}{108} = 0.834$$

$$ScAxB = \sum_{I=1}^{a}\sum_{j=1}^{b} \frac{y_{ij}..^2}{cr} - \frac{y^2...}{abcr} - (SCA + SCB) =$$

$$= \frac{(10.457^2 + \ldots + 12.803^2)}{3x3} - \frac{140.181^2}{108} - (.101 + .834) = 0.003$$

$$Sc\ C = \sum_{k=1}^{c} \frac{y_{..k.}^2}{abr} - \frac{y^2...}{abcr} = \frac{(44.997^2 + \ldots + 48.464^2)}{4x3x3} - \frac{180.141^2}{108} = 0.166$$

$$ScAxC = \sum_{I=1}^{a}\sum_{k=1}^{c} \frac{y_{i.j.}^2}{br} - \frac{y^2...}{abcr} - (SCA + SCC) =$$

$$= \frac{(11.141^2 + \ldots + 12.533^2)}{3x3} - \frac{140.181^2}{108} - (0.101 + 0.166) = .001$$

$$ScBxC = \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{.jk.}^2}{ar} - \frac{y^2...}{abcr} - (SCB + SCC) =$$

$$= \frac{(13.626^2 + . .. + 17.056^2)}{4 \times 3} - \frac{140.181^2}{108} - (0.834 + 0.166) = .004$$

$$ScABC = \sum_{l=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{ijk.}^2}{r} - \frac{y^2...}{abcr} - (SCA + SCB + SCC + SCAB + SCAC + SCBC)$$

$$= \frac{(3.331^2 + . \quad .+4.404^2)}{3} - \frac{140.181^2}{108} - (.101+.384+.166+.003+.001 + .004)$$

$$= 0.002$$

$$Sc\mathbf{B} = \sum_{l=1}^{r} \frac{y_{ij..}^2}{abc} - \frac{y^2...}{abcr} = \frac{46.916^2 + 46.687^2 + 46.578^2}{4x3x3} - \frac{140.181^2}{108} = .001$$

$$Sctot = \sum_{l=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum^{r} y_{ijkl}^2 - \frac{y^2...}{abcr} = 1.086^2 + \ldots + 1.478^2 - \frac{140.181^2}{108} = 1.1679$$

Sctrat = (SCA + SCB + SCC + SCAB + SCAC + SCBC + SCAXBXC) =

= (0.101 +0.834 + 0.003 + .166 + .001 + .004 + .002) = 1.111

PRUEBA DE MEDIAS ANTES DE CONSIDERAR EL TESTIGO.

Prueba de medias para el factor A

		K	T	A	C
		1.345	1.29155	1.29140	1.262
C	1.262	0.866*	0.02955*	0.0294*	0
A	1.29140	0.0546*	0.00015 NS	0	
T	1.29155	0.0544*	0		
K	1.345	0			

Tukey $q\alpha\ \sigma x$

$$\sigma x = \sqrt{\frac{CMEE}{bcr}} = \sqrt{\frac{0.00078}{27}} = 0.00537$$

Tukey q(0.05) 4 y 70 gl (3.7317)(0.00537) = 0.0200

trigo	1.29155b
avena	1.29140b
cebada	1.262 c
Kochia	1.346 a

Prueba de medias del factor B

	T	M	R
	1.372	1.346	1.174
T 1.174	0.198**	0.172**	0
M 1.346	0.026**	0	
R 1.372	0		

Tukey qα σx

$$\sigma x = \sqrt{\frac{CMEE}{acr}} = \sqrt{\frac{0.00078}{36}} = 0.00465$$

Tukey q(0.05) 3 y 70 gl (3.3934)(0.00465) = 0.0157

Trozo	1.174c
Molido	1.346b
Rolado	1.372a

Prueba de medias para el factor C

	9	6	3
	1.346	1.297	1.249
3 1.249	0.097 *	0.048 *	0
6 1.297	0.049 *	0	
9 1.346	0		

Tukey qα σx

$$\sigma x = \frac{CMEE}{\sqrt{abr}} = \frac{0.00078}{\sqrt{36}} = 0.00465$$

Tukey q(0.05) 3 y 70 gl (3.3934)(0.00465) = 0.0157

Tiempo (horas)

3	1.249c
6	1.297b
9	1.346a

CONCLUSIONES

Se encontró diferencia significativa ($p<0.05$) para la digestibilidad porcentual de los forrajes evaluados con rumiantes fistulados que fueron trigo, avena, cebada y kochia. Se encontró diferencia significativa para la digestibidad porcentual ($p> 0.05$) en el segundo factor que fue el tipo de procesamiento del forraje (trozo, molido y rolado) también se encontró diferencia significativa ($p< 0.05$) en los tres tiempos de ingesta.

No se encontró diferencia significativa ($p> 0.05$) para la interacción forraje-tipo de proceso, forraje-tiempos después de la ingesta, así como tampoco se encontró diferencia significativa para la interacción forraje-tipo de proceso-horas después de la ingesta.

No se encontró diferencia significativa ($p> 0.05$) en la digestibilidad general para el control local de los animales que fue la edad.

Al realiza la prueba de medias de tukey con un nivel de significancía de 0.05 se encontró que la digestibilidad porcentual fue mayor en la Kochia, en el trigo y avena la digestibilidad fue similar y la cebada fue la que obtuvo el porcentaje mas bajo de digestibilidad.

En la prueba de medias para el factor tipo de procesamiento, la digestibilidad fue diferente en los tres procesos siendo mejor en el forraje rolado. Seguido del molido y por ultimo en trozo.

En la prueba de medias para los tiempos después de la ingesta, se encontró que en la hora 9 se obtuvo el mayor porcentaje de digestilibidad. Esto es lógico ya que el forraje estuvo mayor tiempo en el rumen, por lo que su degradación fue mayor.

Por lo que en forma general podemos concluir que la Kochia con el tipo de procesamiento rolado es la que obtiene mayor digestibilidad

Digestibilidad general de 4 forrajes diferentes y con 3 tipos de procesamiento a 3 tiempos después de la ingesta. Datos originales porcentaje.

Pasto Factor A	Media Real %	Tipo de proceso	Media real %	Horas después de la ingesta	Media real %
Trigo	19.56b	Trozo	14.92 c	3	17.74 c
Avena	19.56b	Molido	22.18 b	6	19.81 b
Cebada	18.28c	rolado	23.55 a	9	22.18 a
kochia	22.18a				

Tratamientos con letras iguales son no significativos(P< 0.05)

PRUEBA DE MEDIAS INCORPORANDO EL TESTIGO

Prueba de medias para el factor A

Tuquey = q $\alpha.\sigma\overline{x}$

$$\sigma x = \sqrt{\frac{CMEE}{bcr}} = \sqrt{\frac{0.001}{27}} = 0.0060$$

$$tc = \frac{\overline{to} - \overline{y}i...}{sd} \approx t\alpha/2\ (EE) \qquad t\alpha\ 0.05/2 = 1.994$$

sd = $\sqrt{s^2}$ (1/q + 1/r) = $\sqrt{0.001}$ (1/8 + 1/27) 0.0127

tc to&trigo = 1.354 – 1.29155 ÷ 0.012 = 5.25 * b

tc to&avena = 1.354 – 1.29140 ÷ 0.012 = 5.216 * b

tc to&cebada = 1.354 – 1.26200 ÷ 0.012 = 7.660 * b

tc to& Kochia = 1.354 – 1.34600 ÷ 0.012 = 0.666 NS a

La digestibilidad del testigo es similar a la digestibilidad de la Kochia, pero diferente del trigo, avena cebada y trigo ($p < 0.05$)

Prueba de medias para el factor B incorporando el testigo.

Sd = .0123 $t\alpha$ 0.05/2 = 1.994

Tc to&trozo = 1.354 – 1.174 ÷ 0.0123 = 14.634 * b

Tc to&molido = 1.354 – 1.346 ÷ 0.0123 = 0.650 NS a

Tc to& rolado = 1.354 – 1.372 ÷ 0.0123 = 1.463 NS a

La digestibilidad de la forma de presentación molido y rolado es similar al la digestibilidad general del testigo, pero la presentación en trozo es diferente a la del testigo. ($p < 0.05$)

Prueba de medias para el Factor C contra el testigo.

Sd = 0.123 $t\alpha$ 0.05/2 = 1.994

Tc to&3 = 1.354 – 1.249 ÷ 0.0123 = 8.536* b

Tc to&6 = 1.354 – 1.297 ÷ 0.0123 = 4.634* b

Tc to&9 = 1.354 – 1.346 ÷ 0.0123 = 0.650 NSa

La digestibilidad presentada en la hora 9 resulta ser similar a la digestibilidad general del testigo.

LITERATURA REVISADA

Castillo P.J, J.G.Arias. 1998. Estadística inferencial básica. Grupo editorial Iberoamérica. México.

Cochran. G. William. 1980. Diseños experimentales. Editorial Trillas. México

Infante G. S. 1997. Métodos estadísticos. Editorial Trillas. México.

Kreyszig Erwin. 1979. Introducción a la estadística matemática. Editorial Limusa. México

Montgomery D. C.1991.Diseño y análisis de experimentos. Grupo editorial Iberoamérica. México.

Reyes C. P. 1978 Diseños de experimentos agrícolas. Editorial Trillas. México.

Rodríguez del A. J. 1991. Métodos de investigación pecuaria. Editorial Trillas. México.

Snedecor W. George, W. G. Cochran. 1979 métodos estadísticos. Editorial Continental. México.

Steel G.D Robert, J. H. Torrie.1981. 2ª. Principles and procedures of statistics a biometrical approach. 2ª. Ed. Editorial Mc Graw-Hill. USA.

EXPERIMENTOS COMBINADOS

INTRODUCCION

Los experimentos agrícolas, tales como pruebas de variedades, experimentos con semillas como tratamientos, determinaciones de prácticas culturales en las investigaciones de campo son usualmente repetidas en varias localidades para un número de años. Esto es necesario porque los efectos de muchas variables o factores que varían considerablemente de localidad a localidad así como de año a año (variables endógenas).

Con este tipo de experimentación agrícola es posible determinar los efectos ambientales (años) y localidades sobre las variedades o tratamientos estudiados como una consecuencia de las repeticiones de los experimentos en las localidades y estaciones, y es también posible hacer más amplias recomendaciones aplicables.

De cualquier manera, en este tipo de experimentación el investigador debe considerar la representatividad de los efectos experimentales en la población. En experimentos que involucran repeticiones de ambas localidades y tiempo (años), es deseable que las muestras para localidades sean tan representativas como sea posible del área donde se aplicaran los resultados. La muestra de años debe ser representativa de la longitud del ciclo de la estación en la localidad en particular. Como una regla, el investigador tiene poco control sobre la selección de las localidades e inseguridad de que las muestras de las estaciones son típicas.

El número de localidades necesarias para muestrear adecuadamente un área geográfica en particular en una extensión considerable, es una función de la variedad del área. Así, una variable requiere el uso de más localidades en campo. Similarmente, cuando las fluctuaciones ambientales de estación a estación son extremas, esto aparenta que más estaciones deban ser incluidas en orden para muestrear adecuadamente la longitud del ciclo ambiental en una localidad en particular.

Bajo condiciones de alta variabilidad de localidad y estación es extremadamente difícil distinguir entre variedades cuando el promedio esta sobre todas las localidades. Para superar esta dificultad, existen dos soluciones, las cuales no son necesariamente y mutuamente exclusivas son aparentes. Una es usar más repeticiones en cada localidad, técnica más refinada de parcelas, y adicionar pruebas de sitios. La otra es agrupar o estatificar las localidades en base a su respuesta promedio y su respuesta de variabilidad.

Un problema igualmente grave es la alta variabilidad año con año en algunas localidades. Aquí otra vez la estratificación puede ser de utilidad. Por este medio las localidades son agrupadas dentro de estos con una variable ambiental y esto da uniformidad ambiental.

Otro acceso, donde se detalla los records del clima son disponibles, para analizar el clima por los años en los cuales las pruebas se condujeron y compararon con el promedio sobre la longitud del ciclo de los años. Esto permitirá medir de cuantas muestras en particular dentro de años derivará de la población de años.

Con esta estratificación de localidades y estaciones el investigador podrá hacer mejor inferencia sobre la variedad para una región especifica con una gran seguridad que una diferencia representara los efectos reales y no justamente la variabilidad ocasional.

La repetición de tales experimentos permite al investigador con una colección de pruebas comprender variedades similares o tratamientos los cuales requieren un análisis combinado así como una sumarización. El análisis apropiado de experimentos combinados es complejo; requiere la aplicación de procedimientos diferentes de aquellos que se aplican a los experimentos simples. Una de estas diferencias es la introducción del concepto de interacción de variables.

Los agrónomos reconocen que las pruebas de rendimiento son comúnmente hechas para dos propósitos. Uno es el de proveer una prueba de significancia para una variedad especifica o un experimento con fertilización. La otra es la proveer información como una base para recomendaciones generales para un área geográfica. Para el primer propósito, una amplia repetición en cada localidad es deseable. Menos repeticiones por localidad son necesarias donde el objetivo es hacer una recomendación generalizada para un área en particular. De cualquier manera en esta situación, el investigador podrá confiar más de los resultados cuando al menos dos repeticiones estén disponibles en cada localidad.

Una nueva aleatorización de variaciones o tratamientos deberán ser efectuados para cada experimento. El investigador arriesga la introducción de una parcialidad en los datos cuando la misma aleatorización es usada para cada una de las localidades y años, o cuando las variaciones o tratamientos tienen la misma posición relativa en las diferentes repeticiones del mismo experimento. Consecutivamente las variaciones son aptas para ser correlacionadas.

APLICACIÓN DE LA PRUEBA A UNA VARIEDAD DE CEBADA

El desarrollo de variedades superiores de cultivos requieren de muchas pruebas de rendimiento hechas en varias localidades y en sucesivas estaciones para hacer recomendaciones para regiones o estados. Las variedades no siempre responden igual en rendimiento cuando crecen en diferentes localidades o en diferentes años, esto hace necesario evaluar grupos con pruebas individuales de rendimiento. Las pruebas individuales en cada una de las localidades pueden ser designadas como un bloques completos al azar, cuadro latino, parcelas divididas o bloques incompletos.

Los datos de rendimiento para 5 variedades de cebada colectados y reportados por un experimento de bloques completos al azar en cuatro localidades en Minnesota por un período de dos años serán usados para: a).- ilustrar el método de combinación de pruebas individuales de rendimiento dentro de un análisis y b).- Hacer inferencias apropiadas de los datos. Cada experimento consistió de 3 repeticiones. Las mismas cinco variedades fueron evaluadas en Minnesota en la University Farm, Waseca, Crookston y Grand Rapids, para los años de 1932 y 1935. el rendimiento en buschels por acre para cada una de las parcelas de cada variedad se muestra en la tabla siguiente.

Antes de analizar los datos, es importante recordar que una suposición fundamental para el análisis de varianza es que el error experimental debe tener una varianza común. Por lo tanto, antes de proceder con los cálculos de estos experimentos combinados, es necesario determinar si estos datos, tienen las varianzas homogéneas, normalidad y los demás supuestos. Consecutivamente será apropiado proceder con el análisis de los datos combinados.

Las preguntas que surgen de estos experimentos y que pueden ser contestadas son las siguientes:

a) Si hay diferencias significativas entre el rendimiento de las variedades como un promedio de todas las localidades y años.

b) Si hay diferencias significativas en las respuestas entre las diferentes variedades que crecen en las diferentes localidades, por ejemplo si la relación entre variedades es la misma en todas las localidades.

c) Si las variedades responden de manera diferente en los deferentes años.

d) Si las variedades responden de manera diferente en ciertas localidades y en ciertos años.

e) Si la diferencia parietal es de suficiente magnitud para que una variedad pueda esperar servir en grandes pociones ó el área completa incluida en la prueba sobre un periodo de años.

El análisis de los datos de rendimiento de la tabla siguiente ilustran la aplicación del análisis de varianza en un esfuerzo para obtener respuestas a las preguntas planteadas.

Tabla 1.- Rendimientos de 5 variedades de cebada repetidas 3 veces en
Cada una de las 4 localidades en 1932 y 1935.

	Número repeticiones								Total para
	I	II	III	TOTAL	I	II	III	TOTAL	ambos años
	University Farm 1932				University Farm 1935				

Manchuria	19.7	31.4	29.6	80.7	45.5	50.3	60.0	155.8	236.5	
Glabron	28.6	38.3	43.5	110.4	47.5	41.1	49.4	138.0	248.4	
Velvet	20.3	27.5	32.6	80.4	54.2	52.3	64.5	171.0	251.4	
WISC #38	27.9	40.0	48.1	114.0	82.2	53.1	74.7	190.0	304.0	
Peatland	22.3	30.8	31.1	84.2	47.4	57.8	50.5	155.7	239.9	
Total	118.8		188.0	182.9	469.7	256.8	254.6	299.1	810.5	1280.2
		Waseca 1932					Waseca 1935			
Manchuria	40.8	29.4	30.2	100.4	53.9	58.8	47.7	180.4	280.8	
Glabron	44.4	34.9	33.9	113.2	63.7	61.1	52.2	177.0	290.2	
Velvet	44.6	41.4	26.2	112.2	53.9	59.1	58.4	169.4	281.6	
WISC #38	39.8	39.2	29.1	108.1	74.2	75.6	67.0	216.8	324.9	
Peatland	71.5	47.6	55.4	174.5	51.1	47.3	45.0	143.4	317.9	
Total	241.1		192.5	174.8	608.4	296.8	301.9	268.3	867.0	1475.4
		Crookston 1932					Crookston 1935			
Manchuria	34.7	29.1	35.1	98.9	42.1	47.1	30.8	120.0	218.9	
Glabron	28.8	28.7	21.0	78.5	38.8	29.4	30.5	98.7	177.2	
Velvet	29.8	38.4	28.0	96.2	42.1	40.0	39.8	121.9	218.1	
WISC#38	27.7	27.6	20.4	75.7	44.3	43.5	47.7	135.5	211.2	
Peatland	43.0	32.7	32.0	107.7	53.9	51.8	50.3	156.0	263.7	
Total	184.0		158.5	136.5	457.0	221.2	211.8	199.1	832.1	1089.1
		Grand Rapids 1932					Grand Rapids 1935			
Manchuria	20.2	30.2	16.0	66.4	28.6	28.5	32.7	85.8	152.2	
Glabron	13.2	20.5	9.6	43.3	21.4	18.7	24.1	64.2	107.5	
Velvet	24.5	41.6	30.6	96.7	20.7	26.8	30.4	77.9	177.6	
WISC # 38		19.0	18.4	24.6	62.0	20.7	23.6	30.9	75.2	137.2
Peatland	27.6	30.0	22.7	80.3	32.6	40.0	34.2	106.8	187.1	
Total	104.5		140.0	103.5	348.7	122.0	135.8	152.3	409.9	758.6
Total de 4	628.4		657.7	597.7	1883.8		896.8	903.9	918.8	2719.5
	4603.3									
Secciones										

Análisis de pruebas dentro de componentes.

El análisis de pruebas combinadas de rendimiento de un cultivo anual es esencialmente una extensión del análisis de varianza aplicado a un experimento simple de bloques completos al azar. De cualquier manera, esto también involucra el principio del análisis de parcela dividida en el que se requieren dos errores. Un error se necesita para pruebas de localidades, años, y la interacción de localidades por años, otra es necesaria para pruebas de variedades y todas las interacciones que involucran variedades.

Los factores e interacciones juntos con sus correspondientes grados de libertad para este experimento pueden ser representados como sigue:

Fuentes de Variación		Grados de Libertad
Entre experimentos		7
Localidades		3
Años		1
Localidades por años		3
Dentro de experimentos (error(a))		16
Repeticiones / Experimentos		23
Variedades		4
Variedades por localidades		12
Variedades por años		4
Variedades por localidades por años		12
Repeticiones por localidades	8	
Repeticiones por variedades por localidades	24	
Repeticiones por localidades por años	8	Error (b) 64
Repeticiones por localidades por años	24	
Total		119

Habrá un total de 119 grados de libertad para las pruebas combinadas donde hay 120 parcelas individuales en las 8 pruebas. Los grados de libertad para los efectos principales serán 1 menos (n-1) que el número de localidades (4-1=3), años (2-1=1) y variedades (5-1=4), representativamente. Los grados de libertad para las interacciones se obtienen multiplicando los grados de libertado por los efectos principales involucrados. Por ejemplo variedades por localidades será 4x3=12. para la segunda interacción variedades por localidades por año, los grados de libertad serán 4x3x1=12. los cálculos de sumas de cuadrados y cuadrados medios se hacen de acuerdo con lo anteriormente planteado.

Deberá notarse que las repeticiones son numeradas como 1, 2 y 3 en la tabla anterior de cualquier manera esto no indica que una repetición particular en una localidad especifica tenga una relación directa con su correspondiente repetición en alguna otra localidad. Por ejemplo, la repetición en Waseca no lleva una relación con la repetición en Grand Rapids, o en alguna otra localidad.

Porque esta carece de asociación entre repeticiones y localidades, el análisis de varianza de los experimentos combinados está basado en el principio del experimento de parcelas divididas. La primer parte representa una sola clasificación de análisis como “entre experimentos” y “ dentro de experimentos” comenzando con los 8 experimentos bajo consideración. Subsecuentemente, entre experimento esta subdividido dentro de las partes componentes como localidades, años y localidades por años. La segunda parte del análisis considera los efectos varietales y sus interacciones con localidades y años.

a).- Cálculo de las sumas de cuadrados:

El factor de corrección (FC) es calculado del gran total en la tabla 1, como sigue:

FC = $(4603.3)^2/120=176,586.42$

El primer paso es obtener la suma de cuadrados (SC) para la variabilidad total de los experimentos combinados de la siguiente manera:

SCtotal (tabla 1):

$= ((19.7) + (28.6)^2 +...+(30.9)^2+(34.2)^2)-(4603.3)^2/120$

$=200,897.35 – 176,586.42 = 24,292.93$ para 119 g.l.

Los cálculos para las otras sumas de cuadrados se calculan de la siguiente manera:

SC experimentos (tabla 1):

$$=\frac{(496.7)^2 + (608.4)^2+...+(409.9)^2}{15} - (4603.3)^2/120$$

$=193,158.47 – 176,586.42 = 16,572.05$ para 7 g.l.

La tabla 2 debe ser presentada para incluir los datos apropiados que se necesitan para calcular la suma de cuadrados para localidades, años y para localidades por años. Son el gran total en cada uno de los 8 experimentos de la tabla 1.

Tabla 2.- Total de rendimiento para localidades y años.

Año	Localidad				Total
	U. Farm	Waseca	Crookston	Gm Rpds	
1932	469.7	608.4	457.0	348.7	1883.8
1935	810.5	867.0	632.1	409.9	2719.5
Total	1280.2	1475.4	1089.1	758.6	4603.3

SC Localidades (tabla 2):

$$= \frac{(1280.2)^2+(1475.4)^2 +...+(578.6)^2}{30} - (4603.3)^2/120$$

=185,991.00 – 176,586.42=9,324.58 para 3 g.l.

SC años (tabla 2):

$$= \frac{(1883.8)^2+(2719.5)^2}{15} - (4603.3)2/120$$

=193,158.47 –176,586.42 = 16,572.05 = 7 g.l.

- Localidades = - 9,324.58 = -3 g.l.
- Años = - 5,819.96 = -1 g.l.

1, 427.51 =3 g.l.

está claro que los cálculos para la interacción y años de los 8 totales individuales en el cuerpo de la tabla 2 fueron cuadrados y sumados. La suma de estos totales cuadrados se dividió por 15 y se ordenaron para poner la suma de cuadrados en una parcela básica individual. En otras palabras, cada total comprendió de 15 parcelas totales, 5 variedades repetidas 3 veces. De esta es sustraído el factor de corrección. Esto representa 7 grados de libertad. De este total es sustraída la suma de cuadrados para localidades (9324.58. para 3 g.l.) y años (5819.96 para 1 g.l.), con una diferencia de 1427.51 (3 g.l.) para la interacción.

Así, en términos generales, la suma de cuadrados para la interacción de dos factores es la diferencia después de restar de las sumas de cuadrados para dos variables, representando por los totales marginales, del total de suma de cuadrados en el cuerpo de la tabla 2.

El siguiente paso es calcular la suma de cuadrados para repeticiones sobre experimentos (tabla 1):

$$\frac{(118.8)^2+(241.1)^2+\dots+(152.3)^2}{5} - (4603.3)^2/120$$

=194,864.59 – 176,586.42 = 18,278.14 para 23 g.l.

La suma de cuadrados para el error (a), dentro de experimentos representa 2 g.l. para repeticiones en cada uno de los 8 experimentos (2 x 8 =16).

18,278.14 – 9,324.58 – 5,819.96 – 1,427.51 = 1,706.12 para 16 g.l.

Los datos apropiados para cálculos de las sumas de cuadrados para variedades, así como para las interacciones de dos factores con dos variedades, se presentan en las tablas 3 y 4.

Los totales para variedades y localidades se muestran en la tabla 3.

Estos datos son tomados de la columna de la derecha de la tabla 1.

Tabla 3.- Totales de rendimientos agrupados para variedades y localidades.

	Localidad				
Variedad	U. Farm	Waseca	Crookston	Grn. Rpds.	Total
Manchuria	236.5	260.8	218.9	152.2	868.4
Glabron	248.4	290.2	177.2	107.5	823.3
Velvet	251.4	281.6	218.1	174.6	925.7
WISC # 38	304.0	324.9	211.2	137.2	977.3
Peatland	239.9	317.9	263.7	187.1	1008.6
Total	1280.2	1475.4	1089.1	758.6	4603.3

En la tabla 4 los totales son presentados para variedades y años los cuales son obtenidos sumando los rendimientos de cada variedad en cada una de las 4 localidades.

Tabla 4.- Total de rendimiento agrupado para variedades y años.

Variedad	Año		Total
	1932	1935	
Manchuria	346.4	522.0	868.4
Glabron	345.4	477.9	823.3
Velvet	385.5	540.2	925.7
WISC # 38	359.8	617.5	977.3
Peatland	446.7	561.9	1008.6
Total	1883.8	2719.5	4603.3

SC Variedades (tabla 3):

$$= \frac{(868.4)^2+(823.3)^2+...+(1008.6)^2}{24} - (4603.3)^2/120$$

= 177,552.13 – 176,586.42 = 965.71 para 4 g.l.

SC Variedades x localidades (tabla 14.3):

$$= \frac{(236.5)^2+(248.4)^2+...+(187.1)^2}{6} - (4603.3)^2/120$$

=188,170.12 – 176,586.42 = 11,583.70 =9 g.l.

-variedades = -965.71 = -4 g.l.

- localidades = - 9,324.58 = -3 g.l.

1,293.41 = 12 g.l.

SC Variedades x años (tabla 4):

$$= \frac{(346.4)^2+(345.4)^2 +...+ (561.9)^2}{12} - (4603.3)^2/120$$

= 183,885.63 – 176,586.42 =7,299.21 = 9 g.l.

variedades = - 965.71 = -4 g.l.

Años = - 5,819.58 = -3 g.l.

513.54 = 4 g.l.

La comparación para la interacción de tres factores variedades por localidades por años se calcula de los totales del cuerpo de la tabla 5. Estos totales son la suma de las tres repeticiones en cada una de las variedades en cada una de las localidades para cada uno de los años en la tabla 1.

Tabla 5 .- Total de rendimiento para variedades, localidades y años.

		Variedades					
Localidades	años	Manchuria	Glabron	Velvet	Wisc. # 38	Peatland	Total.
U.Farm	1932	80.7	110.4	80.4	114.0	84.2	469.7
Waseca	1932	100.4	113.2	112.2	108.1	174.5	608.4
Crookston	1932	98.4	78.5	96.2	75.7	107.7	457.0
Grand Rapids	1932	66.4	43.3	96.7	62.0	80.3	348.7
U. Farm	1935	155.8	138.0	171.0	190.0	155.7	810.5
Waseca	1935	160.4	177.0	169.4	216.8	143.4	867.0
Crookston	1935	120.0	98.7	121.9	135.5	156.0	632.1
Grand Rapis	1935	85.8	64.2	77.9	75.2	106.8	409.9
Total		868.4	823.3	925.7	977.3	1008.6	4603.3

SC variedades por localidades por años (tabla 5):

$$= \frac{(80.7)^2+(100.4)^2+...+(106.8)^2}{3} - (4603.3)^2/120$$

=197,951.68 – 176,586.42 = 21,365.26 = 39 g.l.

-variedades =965.71 = -4 g.l.

- localidades =9,324.58 = -3 g.l.

- años = -5,819.96 = -1

- variedades por localidades = - 1, 293.41 = -12 g.l.

- variedades por años = - 513.54 = - 4 g.l.

- localidades por años = - 1, 427.51 = - 3 g.l.

2,020.55 = 12 g.l.

De lo anterior, está claro que la segunda interacción de tres factores es obtenida de la resta de la suma de cuadrados para tres variables (21365.26) las sumas de cuadrados para los tres efectos principales y sus respectivas interacciones (variedades =965.71, localidades =9324.58, años = 5819.96, variedades por localidades = 1293.41, variedades por años = 513.54, localidades por años = 1427.51) para obtener 2020.55.

El análisis de varianza completo está preparado y es presentado en la tabla 6. La suma de cuadrados para el error (b) para 64 g.l. es obtenido por diferencia de la siguiente manera:

Total – (repeticiones / experimentos + variedades) + (variedades por localidades)+(variedades por años)+(variedades por localidades por año)).

Por sustitución, la suma de cuadrados para el error (b) es determinada como sigue:

24,292.93 – (18,278.17 + 965.71 + 1,293.41 + 513.54 + 2, 020.55)

=1,221.55 para 64 g.l.

Tabla 6.- Análisis de varianza de experimentos combinados en la tabla 1.

Fuentes de variación	G.L.	S.C.	C.M.
Experimentos	7	16,572.05	
Localidades	3	9,324.58	3,108.19
Años	1	5,819.96	5,819.96
Localidades por años	3	1,427.51	475.84
Dentro de Experimentos Error (a)	16	1,706.12	106.63
Repeticiones sobre experimentos	23	18,278.17	
Variedades	4	965.71	241.63
Variedades por localidades	12	1,293.41	107.78
Variedades por años	4	513.54	128.39
Variedades por localidades por años	12	2,020.55	168.38
Error (b)	64	1,221.55	19.09
Total	119	24,292.93	

b).- Comparación de sumas de cuadrados simples vs experimentos combinados.

En este punto se informa lo relativo al análisis de varianza combinado en la tabla 6 para los análisis de varianza de los 8 experimentos individuales en bloques completos al azar. La suma de cuadrados para el total, repeticiones, variedades, y el segundo error (b) para los 8 experimentos son sumados en la tabla 7.

Tabla 7.- Suma de cuadrados calculados de los experimentos por separado.

Prueba	Años	SC Total	SC Rep.	SC Var.	SC error (b)
U. Farm	1932	867.30	450.10	375.61	41.59
U. Farm	1935	1,031.71	251.62	506.36	273.73
Waseca	1932	1,907.14	471.40	1,196.33	239.41
Waseca	1935	1,203.56	131.15	993.84	78.57
Crookston	1932	487.87	80.83	252.62	154.42
Crookston	1935	807.84	49.21	595.82	162.81
G. Rapids	1932	905.56	179.69	536.17	189.70
G. Rapids	1935	509.91	92.13	336.46	81.32
Total		7,720.89	1,706.13	4,793.21	1,221.55

La suma de cuadrados total (7,728.89) en la tabla 7 es para 112 g.l. para esto debe ser adicionado la correspondiente suma de cuadrados para los componentes de los 7 g.l. para los experimentos que nos darán 119 g.l. para el análisis combinado. Esto será de la siguiente manera:

Tabla 7	7,720.89	= 112 g.l.
Localidades	+ 9, 324.58	= + 3 g.l.
Años	+ 5,819.96	= + 1 g.l.
Localidades por años	+ 1,427.51	= + 3 g.l.

	24,292.94	= 119 g.l.

Así, el total anterior (24,292.94) corresponde al de la tabla 6.

La suma de cuadrados para “dentro de experimentos” (1,706.12) en la tabla 6 comprende el total de las sumas de cuadrados por repeticiones en cada uno de los 8 experimentos (1,706.13) como se indica en la tabla 7.

Del total de la suma de cuadrados para variedades (4,793.21) debe ser restado de la sumas de cuadrados para variedades por localidades (1,293.41), variedades por años (513.54), y variedades por localidades por años (2,020.55) para obtener el valor de 965.71 en la tabla 6 específicamente es los siguiente:

4,793.21 – 1,293.41 – 513.54 – 2,020.55 = 965.71

esto puede ser notado en la tabla 7 que la suma de cuadrados del error (b) para las 8 pruebas totales es 1,221.55. Esto de acuerdo con las sumas de cuadrados para el error (b) de la tabla 6. Así, el error (b) en la tabla 6 es simplemente un total de la sumas de cuadrados para el error individual de los análisis de varianza de las pruebas individuales.

Comparación de los grados de libertad del análisis de varianza combinado de los 8 experimentos separados se muestra en la tabla 8.

De la tabla 8, la analogía entre el análisis de varianza separado para cada experimento y el análisis combinado en la tabla 6 es claro. Los 16 g.l. para el análisis combinado en la tabla 8 representan el error (a) en la tabla 6. Los 32 grados de libertad para el análisis combinado en la tabla 8 representan los siguientes componentes en la tabla 6: variedades = 4 g.l; Variedades por localidades = 12 g.l. variedades por años = 4 g.l. variedades por localidades por años = 12 g.l. Esto podría aparentar que los 64 g.l. para el error (b) en la tabla 6 consisten del producto de los 8 g.l. para el error de cada análisis de varianza en los 8 experimentos 8 x 8 = 64. Los 112 g.l. para el total en la tabla 8 representan los totales respectivos para “dentro de experimentos”. Cuando los 7 g.l. “entre experimentos” (estaciones = 3 años = 1 y localidades por años = 3) son sumados con aquellos “entre experimentos” (112 g.l.) los 119 g.l. son obtenidos.

Tabla 8.- Comparación de grados de las 8 pruebas separadas con el análisis de varianza combinado en la tabla 6.

Fuente de variación	Número de experimentos		Grados de libertad Análisis Separado		Análisis Combinado
Repeticiones	8	x	2	=	16
Variedades	8	x	4	=	32
Error	8	x	8	=	64
Total	8	x	14	=	112
Experimentos	8		-	=	7
Gran Total	-		-		119

Es importante recordar que, cuando un experimento se repite en un número de localidades y sobre varios años debe ser conservado el mismo tamaño y forma de las parcelas, de cualquier manera, la aleatorización debe hacerse por separado para cada experimento.

Interpretación de datos:

La interpretación de datos de un experimento combinado es más complicado que aquellos experimentos simples de magnitud relativa de los varios tipos de interacciones de variedad por ambiente se han estudiado en variedades de algodón como pruebas de rendimiento, con particular referencia a las implicaciones de estas interacciones en la evaluación de tales tipos de experimentos. Se han hecho estudios también en varias interacciones y efectos de variedad en experimentos de tabaco conducidos sobre un período de años en varias localidades. Las interacciones son particularmente importantes en experimentos combinados.

a).- Significado de las interacciones.

Antes de completar la interpretación hecha para las pruebas de rendimiento en cebada presentadas en la tabla 1, es aconsejable presentar una discusión general sobre la interpretación del significado de las interacciones involucradas en las pruebas de rendimientos combinadas con las variedades. De acuerdo con esto, las interpretaciones que se harán son las siguientes:

1.- la interacción de localidades y años indica que la cebada (como un promedio de todas las variedades) rindió relativamente mejor en algunas localidades en algunos años que en otros años. Puesto que las diferencias varietales no son parte de esta comparación esta interacción no tiene relación para las recomendaciones con respecto a las variedades. Así, las comparaciones de localidades y años son de interés solamente en la medida en que es demostrada la magnitud de rendimiento de las diferentes variedades en las diferentes localidades y en sucesivas estaciones.

2.- La interacción de variedades y años indica que algunas variedades, como un promedio de todas las localidades, rindieron mejor en algunos años que en otros que fueron menos productivas en los diferentes años. Al menos el promedio de diferencias entre variedades sobre todos los años es mayor que la interacción de variedades por año, una recomendación no puede hacerse para alguna variedad para todas las estaciones. De cualquier manera, cuando los cuadrados medios para variedades es significativamente mayor que el cuadrado medio para variedades por año, algunas variedades pueden considerarse superiores a otras.

3.- la interacción de variedades y localidades significan que algunas variedades son más productivas en promedio que otras para todas las localidades. Cuando el cuadrado medio para variedades excede significativamente la interacción de variedades y localidades, hay evidencia que la ejecución varietal generalmente fue bastante consistente para demostrar que, como un promedio de todos los años, algunas variedades fueron superiores en todas las localidades. De otra manera una prueba no significativa podría indicar que las recomendaciones de las variedades deberían ser hechas cada una por separado. Más, las estaciones en las cuales se hicieron los experimentos son consideradas como una muestra al azar de los años futuros. Esto es solamente cuando el número de localidades y años se considera una adecuada prueba de todos los posibles lugares y años que las predicciones se han hechos para todos los lugares en la región o estado y futuros años.

La relación entre las diferentes variedades y las condiciones ambientales, en las diferentes localidades y sus diferentes aspectos sobre el rendimiento de cebada, se investigó con más detalle mediante un análisis de regresión de Yates y Conchran. Torrie y Dickson presentan una discusión detallada del procedimiento para obtener información adicional sobre las variedades de cebada para varios factores en diferentes localidades y en diferentes años. Ambas referencias citadas se refieren al uso de varias interacciones en la interpretación de deducciones generales para extensas áreas geográficas.

La referencia esta hecha también para los trabajos de Weiss, et al, los cuales interpretan el análisis de varianza de experimentos combinados que involucran variedades datos de plantación, y años sobre las bases de los componentes de varianza estimados y discutidos Crump. Igualmente, Jonson y Keepin discuten la interpretación de las interacciones de un experimento combinado. En ambos casos están haciendo la evaluación, la varianza de cada contribución para el total de la variabilidad en el análisis de varianza es estimado y relacionado para el apropiado error experimental.

De cualquier manera para ilustrar el método para la determinación de significación entre ciertas variedades en ciertas localidades, el procedimiento apropiado se presentara en detalle. Tal interacción de dos factores involucran la diferencia entre dos diferencias (Fuerte diferencia). El error estándar de una diferencia entre dos diferencias es aplicable a la interacción de variedades por localidades, será como sigue:

$$\sqrt{\frac{19.09 \times 2 \times 2}{6}} = 3.57 \text{ bu.}$$

La única diferencia en esta fórmula, de que para el error estándar de una diferencia entre dos efectos principales, es aquel que una $\sqrt{2}$ adicional es incluido.

La prueba de comparación de medias (DMS) se aplica para la determinación de significancia de ciertas interacciones o “fuertes diferencias”. Como se indico previamente, esta prueba es apropiada para pruebas de comparación de medias decididas antes que los datos combinados son analizados. Para este propósito, el valor de t = 2.00 para 64 g.l., es usando para determinar la DMS apropiada, que es 2.00 x 3.57 = 7.14 Bu.

University Farm			Grand Rapids	
Wisc #38	50.7 bu.		Wisc # 38	22.9 bu.
Peatland	40.0 “		Peatland	31.2 “
	----------------			-----------------
	10.7 bu.			–8.3 bu.

La pregunta que involucra si la diferencia entre estas dos diferencias, 10.7 –8.3 es significativa. La diferencia entre dos diferencias se realiza como sigue:

(Wisc. # 38 - Peatland) – (Wisc. # 38 – Peatland)

(50.7 – 40.0) - (22.9 - 31.2)

10.7 - 19.0 bu.

Por lo tanto puede concluirse que los rendimientos de Wisc. # 38 y Peatland fueron diferentes en U. Farm y en Grand Rapids, la habilidad relativa de rendimientos de dos variedades no fue el mismo en esas dos localidades como un promedio para los años 1932 y 1935.

b.- Componentes de varianza y pruebas de significación:

Como previamente se discutió la aplicación completa del análisis de varianza para experimentos combinados presentan situaciones que requieren la selección de un error apropiado en la interpretación de los resultados. La consideración de componentes de varianza será de utilidad para el establecimiento de un error en términos apropiados para tipos específicos de inferencias.

Como una base para amplias inferencias, el investigador asume que las localidades y estaciones (años) constituye una prueba representativa de la población de localidades y estaciones para lo cual los resultados son aplicables. Por lo tanto, parece apropiado apreciar que las localidades y estaciones son variables y aleatorias. Las variedades, de otra manera, son tomadas como variables fijas.

Tabla 9.- Varianzas en el Análisis de Varianza de Experimentos de Variedades Hechos en Diferentes Localidades en Varios Años.

		Modelos y Varianzas			
		Fijos	Mezclados		
Fuentes De variación	Cuadrados Medios	Repeticiones Variedades Años y Localidades Fijas	Repeticiones y años aleatorizados: Localidades y Variedades fijas	Repeticiones y Localidades Aleatorizadas: Años y Variedades fijas	Repeticiones Años y Localidades aleatorizadas Variedades fijas
Repeticiones En experimentos	R				
Variedades	V	$\sigma^2+ryl(V)$	$\sigma^2+rl\ \sigma^2VY+ryl(V)$	$\sigma^2+ry\ \sigma^2VL+ryl(V)$	$\sigma^2+r\sigma^2Vyl+ry\sigma^2VL+rl\ \sigma^2VY+ryl(V)$
Variedades Por años	VY	$\sigma^2+rl(VY)$	$\sigma^2+rl\ \sigma^2VY$	$\sigma^2+rVYL+rl(VY)$	$\sigma^2+r\ \sigma^2VYL+rl\ \sigma^2VY$
Variedades Por localidades	VL	$\sigma^2+ry(VL)$	$\sigma^2+r\ \sigma^2VYL+ry(VL)$	$\sigma^2+ry\ \sigma^2VL$	$\sigma^2+r\ \sigma^2VYL+ry\ \sigma^2VL$
Variedades por Años por Localidades	VYL	$\sigma^2+r(VYL)$	$\sigma^2+\ r\sigma^2VYL$	$\sigma^2+r\ \sigma^2VYL$	$\sigma^2+r\ \sigma^2VYL$
Error		σ^2	σ^2	σ^2	σ^2

Es aparente, que el punto de vista de los análisis de varianza, que un modelo mezcla, es discutido en otra sección, es usualmente apropiado para rendimiento en experimentos combinados con cultivos anuales.

Las partes importantes del análisis de varianza, en términos de componentes de varianza como una ayuda haciendo apropiadas pruebas de significancia son indicados en la tabla 9.

Con respecto a la tabla 9, sobre la suposición de los efectos fijos para todas las variables, esta claro que el término error (σ^2)es el denominador apropiado para la F para probar variedades, variedades por años, Variedades por localidades y variedades por años por localidades. En este tipo de pruebas se asume que los experimentos comprenden esencialmente el entero de las poblaciones finitas de las tres variables y sus interacciones. Por lo tanto, las inferencias hechas para las variedades, localidades y años las cuales comprenden los experimentos. Esto aparenta que las recomendaciones no pueden hacerse de este tipo de pruebas de hipótesis aplicables a grandes áreas geográficas o para futuras estaciones. Obviamente tales inferencias podrían ser un valor limitado para el fitomejorador o el agrónomo. El investigador, en lugar de eso generalmente se interesa en amplias inferencias con respuesta apreciables a las variedades.

La habilidad para predecir los rendimientos de las variedades sobre áreas grandes y para futuros años es esencial para extender los experimentos de campo para rendimiento en variedades como se indico previamente el modelo más apropiado para este propósito es el modelo mezclado donde las repeticiones, localidades y años son considerados para producir efectos aleatorios, con variedades como variables fijas.

Esto es que el principio general para la formación de un F es para elegir dos cálculos de cuadrados medios los cuales difieren en sus valores esperados por solamente un término, el término que involucra el efecto recién probado. Cuando este principio es aplicado, es aparente que pruebas apropiadas de significancia son indicadas como sigue:

a).- la interacción de 3 factores de variedades por localidades por años es probada en contra del error como

$\sigma^2 + r\ \sigma\ VYL\ ^2 / \sigma$

b).- la interacción de variedades por localidades es probada en contra de la interacción de 3 factores como:

$\sigma^2 + r\ \sigma\ VYL^2 + r\ y\ \sigma\ VL\ ^2 / \sigma^2 + r\ \sigma\ VYL^2$

c).- la interacción también es probada en contra de la interacción de variedades por localidades por años, como:

$\sigma^2 + r\ \sigma\ VYL^2 + rl\ \sigma\ \ VY^2 / \sigma^2 + r\ \sigma\ VYL^2$

d).- La prueba de significación para variedades no es aparente de la relación de los componentes de varianza. Esto es evidente de la tabla 9 que ninguna de las otras líneas en la tabla de análisis de varianza podrían servir como un denominar apropiado para una prueba F del cuadrado medio de la variedad. Esta apropiada (prueba F´) para esta relación, la cual es como sigue:

$$F' = \frac{{}^{s}V^2 + {}^{s}VLY^2}{{}^{s}VL^2 + {}^{s}VY^2}$$

Donde la S^2 representa los valores para los cuadrados medios respectivos en el análisis de varianza. En el uso de las tablas de F sera necesario determinar los apropiados grados de libertad asociados con la prueba F´. Estos han sido desarrollados por Satterthwaite y son calculos como sigue:

$$F1' = \frac{({}^{s}V^2 + {}^{s}VLY^2)^2}{{}^{s}V^4 \qquad {}^{s}VLY^4}$$

----------- +---------

fV   fVLY

Y

$$F2' = \frac{({}^{s}VL^{2}+{}^{s}VY^{2})^{2}}{\dfrac{{}^{s}VL^{4}}{fVL}+\dfrac{{}^{s}VY^{4}}{fVY}}$$

Donde f1 y f2 se aproximan el número de g.l. para los cuadrados medios correspondientes.

Debería aparentar que no es frecuente que el investigador este interesado en una prueba F para variedades puesto que indica si ó no son diferentes significativamente las variedades como un promedio de todas las localidades y años.

e).- Las pruebas de significancia para estaciones y años, como efectos aleatorios son probados en contra de la interacción de estaciones y años (en la tabla 6).

Las inferencias tanto como las pruebas de significancia para algunas de las hipótesis importantes en un análisis de varianza de experimentos combinados son determinados por el cuadrado medio apropiado (tabla 9).

a).- La prueba de F para variedades, indica si las diferencias entre variedades son consistentes sobre todo los años y localidades.

b).- La prueba de F de V/Vy determina las consistencias de variedades sobre futuros años, en estas localidades especificas.

c).- La consistencia de las diferencias de variedades sobre todos los lugares en estos años es determinada por la prueba de F de V/VL.

d).- La respuesta de variedades que en una manera diferencial sobre todos los años es determinada por F=VY/VL.

e).- La respuesta de variedades en una manera diferencial sobre todo de las localidades es determinada por F=VL/VLY.

f).- La respuesta diferencial de variedades y estaciones en los años especificados es indicada por VL/E, donde E es el error.

El modelo adecuado discutido es también apropiado para evaluación de funguicidas, insecticidas, o herbicidas. En tales situaciones el investigador quiere hacer inferencias sobre un rango del desarrollo de variedades en un área especifica donde las recomendaciones son hechas. En esta situación los efectos de los funguicidas, insecticidas y herbicidas son considerados como fijos en los efectos de variedad que son aleatorios.

c).- Pruebas de variedad en una localidad:
una sola seria de experimentos puede ser repetidas en varias alocalidades en una estación. En experimentos con fertilización, por ejemplo, podría desearse determinar la respuesta de un cultivo a diferentes cantidades de un fertilizante a un número de localidades. El análisis completo de tales datos involucra una evaluación de todos los experimentos individuales en un análisis.

Para hacer inferencias, se desea probar los efectos de tratamiento para la población infinita de localidades en la cual las localidades en los experimentos son consideradas una muestra aleatoria. Los tratamientos con fertilizantes se asume que tienen efectos fijos.

Por lo tanto cuando uno asume que los efectos para tratamiento son fijos y los efectos para localidades (estaciones) y repeticiones son aleatorias la interpretación del análisis de varianza combinado es de las esperanzas de los cuadrados medios de la siguiente manera:

Fuente de variación	Grados de libertad	Varianzas
Repeticiones en estaciones (R)	s(r-1)	
Estaciones (S)	(s-1)	
Tratamientos (T)		σ^2+r σTS^2+rs (T)
S x T	(s-1)(t-1)	σ^2+ r $\sigma^2 S^2$
Error	s(r-1)(t-1)	σ^2

La hipótesis de los tratamientos son representativas de la misma población es probada de la siguiente manera:

$$F = \frac{\sigma^2 + r\,\sigma^2\ TS + rs\ (T)}{\sigma^2 + r\,\sigma^2\ TS}$$

Mientras la significación de la interacción de las localidades y tratamiento es probada como sigue:

$$F = \frac{\sigma^2 + r\,\sigma^2\ TS}{\sigma^2}$$

d).- Evaluación de la interacción de tratamientos por estaciones:

la interacción de tratamientos por estaciones (tabla 6) puede ser probada como sigue:

$$F = \frac{\sigma^2 + r\,\sigma TYS^2 + ry\,\sigma^2\,TS}{\sigma^2 + r\,\sigma\,TY\,S^2} = \frac{sTS^2}{sTYS^2} = \frac{107.78}{168.38}$$

Puesto que esta proporción es menos que la unidad, puede concluirse que la interacción de tratamientos por estaciones es N.S.

HOMOGENEIDAD DE LA VARIANZA DEL ERROR EXPERIMENTAL.

Una de las suposiciones del análisis de varianza, es aquel donde los factores experimentales tienen la misma varianza. La suposición de igual varianza se obtiene solamente cuando todos los experimentos se han conducidos de la misma manera, con la misma cantidad de control sobre las condiciones ambientales, y con material experimental de la misma variabilidad.

Esto reconoce generalmente que en experimentos de campo con cultivos, tierra y condiciones climáticas, estos grados de uniformidad rara vez son logrados. Porque de esta condición se podría esperar errores experimentales de experimentos individuales que serán diferentes de localidad a localidad y de estación a estación.

De la suposición que remarca el análisis de varianza, una sola asume que las varianzas del error experimental son las mismas en todos los experimentos y es lo más importantes.

Como se discutió brevemente los experimentos de campo son con frecuencia combinados en orden para estimar un error experimental para los datos reunidos. Este procedimiento es valido cuando todas las colecciones de experimentos razonablemente pueden ser esperados que tengan la misma varianza, esto es, sus varianzas son homogéneas. Bajo tales condiciones, las colecciones originales pueden ser consideradas posteriormente como muestras aleatorias de la misma población. Cuando las varianzas no son homogéneas no pueden ser consideradas como de la misma población. En tales casos, el error no es realmente aplicable para todas las colecciones originales. La importancia de esta situación fue referida en la sección 2 en la discusión de experimentos combinados hechos en varias localidades y años sucesivos.

Cuando una prueba falla para la homogeneidad de los cuadrados medios del error de los experimentos individuales en los análisis de varianza combinados, el efecto general valor de la F tabulada el cual será demasiado bajo cuando los datos no son homogéneos. Como un resultado, la significancia se obtendrá más frecuentemente que podría ser el caso. En términos de tipos de errores, las hipótesis nulas podrán ser rechazadas cuando esto sea verdadero. En otras palabras, el error tipo 1 puede ser que se cometa.

Algunos métodos son disponibles para probar la homogeneidad de varianza (1,2,10). La prueba descrita por Bartlett, es lo mejor que se conoce, es la que más se usa, y se prefiere porque es aplicable si cada prueba de varianza esta basada sobre el mismo o diferentes números de grados de libertad. Esta basada en una Chi-cuadrada que da una prueba aproximada de homogeneidad.

Ejemplo numérico y cálculos para la prueba de Bartlett

Los datos de las 8 pruebas de las variedades de cebada discutidas en la tabla 1 se usaran para ilustrar los cálculos necesarios para probar la homogeneidad de las varianzas del error de las 8 pruebas individuales. Un análisis de varianza de estas 8 pruebas por separado (tabla 7) proporcionan los cuadrados medios para el error respectivo de la siguiente manera:

Localidad	Año	Cuadrados medios

University Farm	1932	5.20
" "	1935	34.22
Waseca	1932	29.93
"	1935	9.82
Crookston	1932	19.30
"	1935	20.35
Grand Rapids	1932	23.71
" "	1935	10.16

En cada uno de los 8 experimentos, los g.l. para los cuadrados medios del error fueron 8. Esta es una evidencia de que la llave para calcular los g.l. para un experimento individual, es como sigue:

Repeticiones	2 g.l.
Variedades	4 g.l.
Error	8 g.l.

Total	14 g.l.

En esta prueba para homogeneidad de varianzas, Bartlett define a Chi-cuadrada como sigue:

$X^2 = 1/C\ (f_t \log e\ s^2_p – s(f_i \log e\ s^2_i))$ para K-1 grados de libertad.

Donde:

K = Número de cuadrados medias recién comparados.

C = Factor de corrección

$$= 1 + \frac{1}{3(k-1)} \left(S\left(\frac{1}{fi}\right) - \frac{1}{ft} \right) \text{----------}(2)$$

fi = g.l. asociados con cada cuadrado medio

ft = total de g.l. para los cuadrados medios individuales

s^2_p = cuadrados medios individuales

s^2_i = cuadrados medios combinados que se calcula de s (fisi2)

----

ft

En la tabla 10, los cálculos necesarios se presentan para ilustrar la aplicación de esta prueba. Bajo la hipótesis de muestras aleatorias de una población con una varianza común, los valores de la s^2_i de la tabla 10 se asumen que todos son estimados de la varianza de la misma población. Si esto es cierto después los valores s^2_i no mostraran alguna diferencia más que las esperadas en las muestras aleatorias de una población común.

Tabla 10.- Cálculos preeliminares para la determinación de la prueba X2 para homogeneidad de varianzas.

		Experimentos individuales				
Localidades y años de experimentos		G.L. (fi)	S.C (fi s^2_i)	C.M. (s^2_i)	(log s^2_i)	(fi loge s^2_i)
U. Farm	1932	8	41.59	5.20	1.6487	13.1896
“	1935	8	273.73	34.22	3.5328	28.2624
Waseca	1932	8	239.41	29.93	3.3989	27.1912
“	1935	8	78.41	9.82	2.2844	18.2752
Crookston	1932	8	154.42	19.30	2.9601	23.6808
“	1935	8	162.81	20.35	3.0131	24.1048
Grand Rapids	1932	8	189.70	23.71	3.1659	25.3272
“	1935	8	81.32	10.16	2.3184	18.5472
Total		64	1221.55	152.6	22.3223	178.5784

Para las situaciones donde los grados de libertad asociados con cada cuadrado medio (fi) son los mismos, como en este caso los totales para las dos últimas columnas de la tabla 10 pueden determinarse por un procedimiento corto. Esto es una evidencia que el total (1221.55) para (fi s^2_i) puede obtenerse multiplicando 152.69 x (=1221.52. igualmente, el total (178.5784) para (fi loge s^2_i) es igual a 22.3223 x 8 = 178.5784. de cualquier manera en casos donde los g.l. se asocian son cada cuadrado medio del error no son los mismos cada valor para (fi s^2_i) y (fi loge s^2_i) podrían ser necesitar ser calculados, con los valores individuales de cada una de las columnas sumadas.

Tamaño f_t = 64

$$S^2_p = \frac{S(fi\ s^2_i)}{f_t} = \frac{1221.55}{64} = 19.086$$

Este es el promedio de los cuadrados medios, 152.69/8 = 19.09 (vea los cuadrados medios para el error en la tabla 6) entonces,

Ft loge s^2_p = 64 x loge 19.086

= 64 x 2.9489 = 188.7302

Sustituyendo en la fórmula (2) para C

C= 1 + 1/21 ((1/8 + 1/8 +1/8 + 1/8 +1/8+ 1/8+ 1/8+ 1/8) – 1/64) = 1.0469

Desde que el factor de corrección (C) siempre será mayor que la unidad disminuirá la ji-cuadrada no corregida. Consecuentemente será necesario ajustar la ji-cuadrada ordinaria con el factor de corrección cuando ji-cuadrada ordinaria es significativa. Sustituyendo en la fórmula (1) para determinar los valores de la ji-cuadrada ajustada, de la siguiente manera:

X2 = 1/C (ft loge s^2_p – S (fi loge s^2_i))

= 1/1.0469 (188.7302 – 178.5784) = 9.70

Referencias para la tabla de ji-cuadrada (Snedecor y Cochran) para 7 g.l. (involucra 8 cuadrados medios individuales) muestra que la probabilidad de un valor mayor que ji-cuadrada = 9.70 esta entre p= 0.20 y 0.30. desde que el valor obtenido de ji-cuadrada es no significativo los datos no ofrecen una evidencia aceptable en contra de las hipótesis de las muestras aleatorias de una población con varianza común. Así, podrá concluirse que los cuadrados medios individuales s^2_i, pueden considerarse como podrá estimados de la misma varianza poblacional. Por lo tanto, el error para 8 experimentos de los datos de rendimientos en la tabla

6 son suficientemente homogéneos para permitir el análisis de varianza combinado, con los cálculos de un error generalizado (19.09) para comparación de todos los experimentos.

Cuando este procedimiento conduce a un valor altamente significativo de ji-cuadrada, extremadamente alto o bajo cuadrados medios del error podrán rechazarse hasta que los valores de ji-cuadrada obtenidos sean no significativos. Experimentos donde los cuadrados medios del error han sido rechazados no pueden ser incluidos en los análisis combinados. Cuando muchos experimentos han sido rechazados, usualmente es deseable agrupar los datos en dos o más colecciones, así que los cuadrados medios de cada colección serán homogéneos. Para este procedimiento, todos los datos se incluirán para hacer inferencias.

LITERATURA REVISADA

Castillo P.J, J.G.Arias. 1998. Estadística inferencial básica. Grupo editorial Iberoamérica. México.

Cochran. G. William. 1980. Diseños experimentales. Editorial Trillas. México

Infante G. S. 1997. Métodos estadísticos. Editorial Trillas. México.

Kreyszig Erwin. 1979. Introducción a la estadística matemática. Editorial Limusa. México

Montgomery D. C.1991.Diseño y análisis de experimentos. Grupo editorial Iberoamérica. México.

Reyes C. P. 1978 Diseños de experimentos agrícolas. Editorial Trillas. México.

Rodríguez del A. J. 1991. Métodos de investigación pecuaria. Editorial Trillas. México.

Snedecor W. George, W. G. Cochran. 1979 métodos estadísticos. Editorial Continental. México.

Steel G.D Robert, J. H. Torrie.1981. 2ª. Principles and procedures of statistics a biometrical approach. 2ª. Ed. Editorial Mc Graw-Hill. USA.