

# Agrupamiento de ensayos Análisis combinados



JUNTA DE ANDALUCÍA

Consejería de Agricultura y Pesca

**AGRUPAMIENTO DE ENSAYOS**  
**ANÁLISIS COMBINADO**

AGRUPAMIENTO DE ENSAYOS. ANALISIS COMBINADO

© *Edita:* Junta de Andalucía. *Consejería de Agricultura y Pesca*

*Publica:* Secretaría General Técnica. Servicios de Publicaciones y Divulgación

© *de los textos:* Rodrigo Morillo-Velarde Perez-Barquero

*Colección:* POLÍTICA, ECONOMÍA Y SOCIOLOGÍA AGRARIA

*Serie:* Estadística

*Depósito Legal:* SE-3095-08

*ISBN:* 978-84-8474-241-8

*Fotocomposición e impresión:* J. de Haro Artes Gráficas, S.L. Parque Ind. P.I.S.A,  
Mairena del Aljarafe • Sevilla

**AGRUPAMIENTO DE ENSAYO**  
**ANÁLISIS COMBINADO**

**Rodrigo Morillo-Velarde Perez-Barquero**

**2008**



## ÍNDICE

<b>1 . INTRODUCCION .....</b>	<b>9</b>
1.1 Diseño. Tratamientos, repeticiones y localidades .....	14
<b>2. ANALISIS PRELIMINARES .....</b>	<b>21</b>
2.1 Analisis de los ensayos individuales .....	21
2.2 Primera estimación de resultados .....	30
2.3 Requisitos para agrupar ensayos .....	35
<b>3. ANÁLISIS COMBINADO DE LOCALIDADES.</b>	
<b>ANÁLISIS PRINCIPAL .....</b>	<b>47</b>
3.1 Analisis convencional .....	49
3.2 Análisis como parcela dividida .....	56
3.3 Agrupar ensayos de distinto tamaño .....	60
3.4 Agrupar ensayos con dos factores .....	61
<b>4. INTERACCION TRATAMIENTO POR AMBIENTE .....</b>	<b>63</b>
4.1 Partición de la interacción tratamiento x ambiente.....	67
4.2 Análisis de regresión lineal. Metodo de Eberhart y Russel .....	73
4.3 Algunos métodos no paramétricos .....	77
4.4 Análisis multivariante. Método AMMI.....	82
<b>5. ANALISIS COMBINADO DE LOCALIDADES Y AÑOS .....</b>	<b>87</b>
5.1 Análisis en años .....	87
5.2 Análisis combinado de localidades y años Diseños jerarquizados .....	91
5.3 Ensayos repetidos en varias localidades y años con más de un factor .....	109
5.4 Englobe de efectos .....	111
<b>6. BIBLIOGRAFIA .....</b>	<b>115</b>



## I. INTRODUCCION

El procedimiento habitual en la ciencia para conocer un hecho determinado es investigarlo. Es decir, estudiarlo para saber a que se debe, que causas son las que lo producen y qué lo puede explicar. El estudio implica generalmente la realización de experimentos para obtener datos y el análisis de estos. Como indicaron no hace mucho tiempo Steel y Torrie (1985) .-"el nuevo conocimiento se obtiene del cuidadoso análisis e interpretación de los datos".-

En los experimentos, diferentes factores que se supone pueden explicar las causas se prueban en forma de distintos tratamientos, analizándose su efecto. Pero, nunca el resultado total se explica totalmente por la variación de los tratamientos que se estudian (factores que son controlados por el experimentador), ya que existen factores no controlados (ajenos a los tratamientos ensayados, por ejemplo los factores ambientales que varían con el emplazamiento del experimento ) que se escapan a su control e influyen también en los resultados del experimento.

Cuando se hace solo un ensayo, ocurre a veces que el efecto de los factores no controlados es mas importante que el efecto de los que se ensayan, impidiendo obtener conclusiones. Sea cual sea el area de experimentación, esta claro que con un experimento en un único sitio el investigador puede evaluar el efecto del factor que controla pero no los ambientales. Si se quieren medir los efectos de los factores no controlados, generalmente como parte de la investigación, el único sistema es repetir el experimento en diferentes condiciones.

Un conjunto de experimentos (llamado también una serie), se puede hacer extensiva a varios lugares (sitios o condiciones), a lo largo de varios años o ambas cosas. Recibe el nombre genérico de experimentos en serie y su análisis conjunto, análisis combinado.

En todas las areas del conocimiento es frecuente hacer varios experimentos en lugares y a veces en ocasiones diferentes, con diferentes objeti-



vos y por diferentes razones: Siempre hay que considerar las razones para repetir los experimentos, porque ello afectará al diseño experimental. Algunos ejemplos de razones para repetir ensayos son las siguientes:

- Conseguir variación de los factores externos que están en estudio (relacionar el crecimiento de las plantas y la longitud del día, analizar el comportamiento de neumáticos en diferentes condiciones climáticas...). Se hacen generalmente con el objetivo de identificar una relación en forma de función entre la respuesta del tratamiento y algunos factores ambientales de modo que se pueda predecir esa respuesta (es lo que se llaman modelos).
- Generalizar una investigación en ciencias sociales, medicina y otras (precisión de una técnica en diferentes laboratorios, ensayos clínicos de evaluación de una técnica quirúrgica en varios hospitales, encuestas de un colegio a una ciudad ...). Ensayos comparativos de pruebas de determinación y validación de técnicas (llamados "ring tests").
- Identificar los mejores tratamientos en unas determinadas condiciones o localidades, pudiendo identificar el intervalo de adaptación espacial. Especialmente cuando se quiere hacer una recomendación que sea de utilidad para una extensión grande es poco eficaz hacer solo un experimento.
- Caracterizar un tratamiento o una práctica novedosa determinada respecto a su efecto a largo plazo.
- En el caso más común de una investigación, recordamos que el método científico consiste en proponer una hipótesis, hacer un experimento para testar la hipótesis, posiblemente modificar la hipótesis, hacer un nuevo experimento y así secuencialmente. Al final, es útil un análisis combinado de todos los experimentos realizados.

Dentro de todas las áreas del conocimiento, posiblemente es en investigación en agricultura donde es más frecuente hacer más de un ensayo dentro de una investigación (de variedades, nutrientes, número óptimo de plantas, fechas de siembra, otras prácticas agrícolas...) para evaluar los efectos de los incontrolables factores ambientales (suelo, clima, incidencia de parásitos...) sobre los factores en estudio. Los ensayos multilocalidades son la principal herramienta para conocer la respuesta varietal a diferentes ambientes y por ello lo son de la mejora genética. Se puede encontrar bibliografía específica sobre ensayos de adaptación (para medir la adaptabilidad geográfica de un tratamiento), ensayos de evaluación preliminar (para descartar tratamientos que no interesan) y ensayos básicos que permiten obtener información para resolver problemas prácticos y hacer recomendaciones (los experimentos se repiten en varias localidades en investigación aplicada con el propósito de obtener información útil para una zona amplia). En agri-

cultura, son necesarios porque el rendimiento de un cultivo depende del genotipo (variedad) o del tratamiento, del ambiente en el cual crece y de la interacción entre ambos, genotipo y ambiente. El estudio de la interacción entre el tratamiento y el ambiente es fundamental en estas investigaciones. La principal ventaja de los ensayos repetidos en agricultura es que suelen tener todos una similar precisión.

Los primeros trabajos dentro de este área se remontan a fechas muy tempranas. Yates y Cochran en 1938, fueron los pioneros en analizar el problema de combinar grupos de experimentos en el contexto de la agricultura.

Si interesa evaluar únicamente sitios, el planteamiento adecuado es hacer una comparación múltiple por análisis de la varianza de una vía (conocido por ANOVA, de sus siglas en inglés), donde los grupos son los diferentes sitios y es lo que se puede comparar. Un ejemplo sencillo se muestra a continuación:

*Ejemplo A. "Una compañía eléctrica quiere comparar los gastos de electricidad de 3 ciudades (sitios) con diferente ubicación. Para ello selecciona una muestra al azar de 20 hogares y mide el consumo de un determinado periodo" (Adaptado de Coakes y Steed, 1999)*

Se trata de un diseño de un solo factor con grupos diferentes (las ciudades) porque los hogares están en ciudades diferentes. La primera parte de la tabla ANOVA sería:

Factor de variación	Grados de libertad
Entre ciudades	2
Dentro de ciudades (residual o error)	57
Total	59

Independientemente de que haya o no diferencia significativa entre ciudades, no se trata realmente de un estudio multilocalidad y no se debe confundir, ya que no se comparan hogares y desde luego tampoco la relación entre hogares, (suponiendo que fueran iguales por tamaño, por ejemplo) y las ciudades, que es uno de los principales objetivos del análisis combinado.

Un paso más es considerar cada localidad como una repetición, se tendrían todos los tratamientos en cada localidad una sola vez, es decir, un experimento simple en el que los bloques o repeticiones serían las localidades. Este experimento se estudiaría por análisis de la varianza de dos vías, donde las fuentes principales de variación son los tratamientos, las repeti-

ciones (en este caso las localidades) y la interacción entre ambas, llamada error experimental o mas modernamente residual y se indica en la tabla 1.1

**Tabla 1. 1 Analisis de la varianza simple de t tratamientos en s sitios.**

Factor de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F
Sitios	s-1			
Tratamientos	t-1			
Error experimental	(s-1)*(t-1)			
Total	s*t -1			

El cuadrado medio se obtiene de la forma usual, dividiendo la suma de cuadrados por los grados de libertad y la prueba F de sitios y tratamientos a su vez dividiendo cada cuadrado medio por el cuadrado medio del error (para mas detalles del analisis de la varianza consultar Tejedor,1999). Consecuentemente, los errores estandar, se emplean como en otros métodos familiares.

Un análisis mas completo de t tratamientos repetidos r veces en s sitios se analiza mediante un ANOVA de dos vias, el cual proporciona información de la significacion de los efectos de los sitios, los tratamientos, la interaccion entre ambos y del error, que son todas las fuentes de variación. El análisis de la varianza se muestra en la tabla 1.2

**Tabla 1.2 Analisis de la varianza simple de t tratamientos repetidos r veces en s sitios.**

Factor de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F
Sitios	s-1			
Tratamientos	t-1			
Sitios x Tratamientos	(s-1)*(t-1)			
Error experimental	(r-1)*s*t			
Total	r*s*t -1			

Un ejemplo simple de este planteamiento es el siguiente:

*Ejemplo B. " Un distribuidor de cosméticos quiere determinar cual entre tres tipos de tiendas (pequeña en el barrio, mediana en centro comercial y gran superficie) y en qué localización (centro o suburbios) es mas eficaz en vender sus productos. Se eligen cuatro tiendas al azar para cada una de las seis situaciones y se consideran las ventas en un determinado periodo" (Adaptado de Coakes y Steed, 1999)*

Se trata de un diseño factorial 3 x 2. La primera parte de la tabla ANOVA sería

Factor de variación	Grados de libertad
Localización	1
Tipo de tienda	2
Tienda x Localización	2
Residual (o error)	18
Total	23

En este diseño ya es posible analizar la interacción entre tratamientos y localidad, es decir ver si hay alguna relación entre tratamientos (tipos de tienda) y su localización. Realmente este es el verdadero objeto de este experimento.

Pero para el análisis adecuado (que se verá más adelante) falta un factor de variación, las repeticiones dentro de las localidades. Esta componente es muy importante, ya que constituye un error que servirá para testar en la prueba F las distintas localidades.

Según Pearce y colaboradores (1988), el diseño de una serie de experimentos completos repetidos en varias localidades tiene dos fases:

1. Diseñar los experimentos en cada sitio.
2. Elegir los sitios (ambientes o localidades).

El diseño y análisis de un experimento no plantea especiales problemas. La elección de localidades hay que hacerla con algunas precauciones como se verá en el siguiente apartado.

Los principales problemas se presentan al analizar en conjunto varios experimentos son los posibles cambios en la varianza de la variable en respuesta a las distintas condiciones ambientales y si hay observaciones repetidas de una unidad experimental, alta correlación entre los errores de las

medidas repetidas. De ahí que es necesario verificar el cumplimiento de unos determinados requisitos.

Hacer varios ensayos es necesario en muchos casos y su planteamiento teórico no es especialmente difícil (el práctico suele ser más complicado). El análisis conjunto de todos los experimentos es importante y debe hacerse con precaución. Hay que considerar, por ejemplo, que muchos paquetes estadísticos (entre ellos el programa SPSS, que se utilizará en los ejemplos) no consideran algunas fuentes de variación como terminos de error, por lo que los test F en el ANOVA que muestran en las salidas son incorrectos.

## **1.1 Diseño. Tratamientos, repeticiones y localidades**

### Tratamientos

El objetivo de todo experimento es comparar tratamientos, que se asignan aleatoriamente a unas unidades experimentales. Si tratamientos es lo que se prueba de acuerdo con la hipótesis de partida (fertilizantes en agricultura, fármacos en farmacia...), las unidades experimentales es dónde se prueba (parcelas de terreno, ratones...). La variación existente entre unidades experimentales que han recibido el mismo tratamiento se llama error experimental e interesa que sea lo menor posible.

Los tratamientos pueden ser cuantitativos o cualitativos y su definición y número dependerá de los objetivos del experimento. Cuando un factor es cualitativo (la localidad siempre lo es) y otro es cuantitativo (por ejemplo dosis) la interpretación de la interacción puede ser complicada, en este caso es mejor hacer curvas de respuesta para comparar los tratamientos (Pearce, 2004). Si unos tratamientos tienen relación con los otros del mismo experimento se dice que están estructurados (por ejemplo dosis crecientes y su comparación se debe hacer por contrastes). Si no tienen relación unos con otros, se dice que no están estructurados (por ejemplo distintos fármacos) y se emplean las comparaciones múltiples de medias (Urquhart, 1981).

En ensayos con más de un factor, los tratamientos son combinaciones de factores y niveles (factor es todo lo que tenga un efecto y nivel es el estado de un factor). Si un factor se prueba con todos los niveles de otro se dice que tiene una estructura factorial. Esta combinación de tratamientos tiene grandes ventajas.

En el caso que nos ocupa, en la repetición de ensayos en varias localidades con  $t$  tratamientos repetidos  $r$ , veces en  $s$  localidades (o como siempre sitios) tenemos dos factores  $t$  y  $s$  en un experimento factorial, donde un factor se prueba con todos los niveles de otro. Como todos los experimentos factoriales, la intención es estudiar la interacción entre los factores tratamiento y localidad, en definitiva, saber si el mejor tratamiento depende de la localidad o que tratamiento es el más adecuado a cada localidad (o condición de ensayo).

Los factores (o los tratamientos) pueden ser fijos (cuando son mas importantes los niveles que el propio factor, por ejemplo variedad A,B...) o aleatorios (importa mas el factor, eligiéndose el tratamiento al azar entre los posibles, por ejemplo de ocho fármacos posibles se eligen dos para ensayar).

Con factores aleatorios no se puede hacer comparación de medias (este es uno de los errores mas comunes en el uso de programas estadísticos), hay que hacer componentes de la varianza. Recordamos que la interacción con un factor aleatorio es aleatoria. El hecho de ser factor fijo o aleatorio afecta a como construir la prueba F para cada factor de variación.

Un modelo con factores fijos y aleatorios se denomina modelo mixto. En un modelo mixto las componentes de la varianza difieren de los esperados para un modelo aleatorio. Las reglas para calcular los cuadrados medios esperados en un modelo mixto se pueden ver en Schultz (1955).

En los análisis combinados, generalmente los tratamientos suelen ser factores fijos en tanto los años se consideran como factores aleatorios y las localidades pueden ser fijas o aleatorias. De ahí que la interacción sea aleatoria. La situación se complica cuando en un mismo experimento intervienen factores fijos y aleatorios

Si los tratamientos deben estar asignados aleatoriamente (al azar) a las unidades experimentales, dentro de cada una de las repeticiones (réplicas o bloques), cada ensayo debe estar sorteado. No se debe repetir el mismo ensayo una y otra vez. Aunque los ensayos tengan el mismo tamaño y estructura hay que sortear cada ensayo independientemente (repetir el mismo sorteo es también otro error muy frecuente).

### Repeticiones

Para conocer la variabilidad que existe y el error experimental los tratamientos ensayados deben estar repetidos. El número de repeticiones depende de los recursos, de la variabilidad del material que se ensaya, del número de tratamientos y de la importancia de las comparaciones. Si el número de repeticiones aumenta, la precisión del ensayo aumenta (el coeficiente de variación disminuye) y los intervalos de confianza se reducen. Lo mínimo es que haya dos repeticiones, lo usual son cuatro y entre cuatro y ocho proporcionan un grado de precisión razonable. Para un calculo preciso del número de repeticiones necesarias es necesario conocer por adelantado la distribución (al menos el coeficiente de variación) y la diferencia verdadera que se desea conocer. El procedimiento está descrito en muchos libros de estadística, aunque Cochran y Cox (1957) proporcionan unas tablas adecuadas para determinar el número de repeticiones en función de la significación y probabilidad deseada. Se ha aconsejado que el número de repeti-

ciones mínimo debe asegurar entre 10 y 20 grados de libertad en el error en un ensayo individual (Mead, 1988), lo que hace necesario replantear el ANOVA del experimento antes de realizarlo.

Se ha sugerido también (y ha llegado a hacerse en algún caso) tomar las repeticiones de todas las localidades sumadas para una mejor respuesta de los tratamientos, pero lo que se consigue es sumar los efectos de las repeticiones en las localidades y de la interacción tratamiento por localidad. No se puede del análisis combinado. Lo ideal es que cada experimento de una serie tenga el mismo diseño, con igual número de tratamientos y repeticiones.

#### Ambientes, sitios o localidades

Las diferentes ubicaciones de los ensayos realizados en una serie y que pueden consistir, por ejemplo en diferentes suelos, presión de enfermedades, insectos o malas hierbas y para las cuales se desea información, han recibido indistintamente el nombre de ambientes (generalmente en ensayos de biología), sitios (en la literatura anglosajona) o localidades cuando hay una diferencia geográfica. En biología o agronomía, el término ambiente implica un conjunto de condiciones climáticas, de suelo, bióticas (plagas y enfermedades) o de manejo en un ensayo individual llevado a cabo en una determinada localidad y en un determinado año. Un ambiente identifica una combinación particular de localidad y año. En este texto nos referiremos indistintamente a cualquiera de los tres sinónimos.

Desde las primeras referencias, se ha insistido en que los lugares para hacer una serie de ensayos deben ser elegidos siempre que sea posible de forma aleatoria. Es decir, los sitios deben ser una muestra aleatoria de todos los sitios de la región (o condiciones posibles de experimentación), incluso la validez del análisis combinado se ha condicionado a la elección al azar de las localidades. No solo no es cierto que no sea válido, sino que es poco práctico (también influye el coste) y que cuando la elección de la localidad es totalmente al azar ocurren situaciones muy complicadas. Las localidades deben ser representativas, ya que en caso contrario los resultados tienen poco valor práctico. Parece evidente que para aumentar la representatividad es necesario aumentar en lo posible el número de ensayos. Cuando interesan las componentes de la varianza, lo más apropiado es elegir las localidades al azar.

En 1988, Mead indicó con gran sensatez, que repetir los experimentos totalmente al azar en una serie de localidades no es un uso eficiente de los recursos, ya que la variación de error experimental total (suma de todos los errores experimentales individuales) puede ser enorme. Aunque en el caso de no elegirlos al azar, los resultados deben ser interpretados estrictamente

para los sitios seleccionados y las medias de los tratamientos pueda representar menos a estos, propuso elegir las localidades muy cuidadosamente para evaluar la característica de la localidad, haciendo un esfuerzo en caracterizar las diferencias entre poblaciones. Las localidades deben estar bien definidas y deben ser suficientemente representativas.

El número óptimo de localidades depende de los objetivos del experimento, de la precisión de los mismos, de la variabilidad ambiental, del número de tratamientos en ensayo y del número de repeticiones en cada ensayo. Dagnelie (1981) propuso emplear el mismo método de cálculo para el número óptimo de localidades en una serie, que para el número de repeticiones en un experimento. También el número de sitios depende de la extensión de la región donde se van a situar los ensayos. Annicchiarico (2002) propone que no deberían ser menos de seis o siete. Como referencia, un proyecto de investigación nunca debería tener menos de 5 a 10 ensayos.

En relación con el procedimiento para elegir las localidades hay dos estrategias en función de los objetivos del estudio:

- 1.- Si el objetivo es hacer una recomendación concreta, primero hay que elegir los sitios y entre ellos seleccionar una muestra representativa. Los sitios deben ser elegidos estrictamente al azar.
- 2.- Si el objetivo es estudiar la interacción entre los tratamientos y el ambiente es mejor elegir algunos pocos sitios que representen cada ambiente y que no sean demasiados los ambientes diferentes, ya que el estudio de la interacción se complica.

En este punto surge muchas veces una duda (manifestada a menudo por estudiantes e investigadores): Supuesto que los recursos son limitados, en general a más sitios menos repeticiones en cada uno, entonces ¿es preferible más localidades con menos repeticiones o más repeticiones en cada localidad y menos localidades?.

En 1951, Sprague y Federer dieron una respuesta. Concluyeron que en un análisis de una serie de ensayos, en general, un aumento en el número de localidades y años y una disminución en el número de repeticiones produce un mejor resultado en un programa de investigación, pero dependiendo siempre de los objetivos de la investigación. Si se quiere testar una hipótesis, donde hay que tener una gran precisión, es necesario tener más repeticiones en cada ensayo a costa de menos localidades, pero si el objetivo es tener información para hacer una recomendación general en una región, donde se requiere menos precisión, se pueden tener menos repeticiones en cada ensayo y más localidades. En definitiva lo importante, de nuevo, son los objetivos del estudio.



Dagnelie (1981) con el método similar al número de repeticiones por ensayo, escribió que para un análisis de  $s$  sitios y  $r$  repeticiones y  $t$  tratamientos, la varianza de la diferencia entre dos medias de  $s \cdot r$  observaciones sería:

$$\sigma^2 (\text{diferencia}) = 2 (\sigma^2_{ts} / s + (\sigma^2_{tr} + \sigma^2) / s \cdot r)$$

Donde,  $\sigma^2_{ts}$  es la componente de la varianza relativa a tratamientos x sitios,  $\sigma^2_{tr}$  la componente de la varianza relativa a tratamientos x repeticiones y  $\sigma^2$  la varianza del error. Las diferencias entre localidades suelen ser mayores que las diferencias entre repeticiones, entonces siempre que  $\sigma^2_{ts}$  no sea nula, la varianza total es mínima cuando  $r$  es mínimo y  $s$  es máximo. De ello se desprende que en la práctica el número óptimo de repeticiones por ensayo debería ser de dos (este autor dirigió varios programas de investigación sobre agricultura en Africa con muchas localidades y tan sólo dos repeticiones por localidad). Proponía que a no ser que la variación entre repeticiones sea mucho mayor que la variación entre localidades, no es muy recomendable aumentar el número de repeticiones (se considera un despilfarro de recursos), ya que la varianza de la media de un tratamiento esta mas afectada por el numero de localidades que por el número de repeticiones. Sólo en el caso de que la varianza de las repeticiones sea mayor que la varianza de las localidades es mejor aumentar las repeticiones.

Algunos años más tarde Mead (1988) analizando los grados de libertad propuso lo contrario (la reducción del número de repeticiones por ensayo y el aumento del número de ensayos, provoca un aumento en los grados de libertad de la interacción tratamiento x localidad y la potencia de las comparaciones).

Para este autor es mejor hacer pocos experimentos en sitios bien elegidos (definidos adecuadamente por sus características) con gran precisión en cada experimento, que hacer un gran número de experimentos poco precisos cubriendo todas las variaciones de esta característica. Muchos sitios al azar es irrelevante en un programa de investigación. Es mejor caracterizar los sitios (no elegirlos al azar) y hacer un buen experimento en cada lugar. De esta forma la informacion generada es más útil. Hoy es la idea mas seguida.

El diseño también influye. Hacer 2 o 3 repeticiones con el empleo de un diseño eficaz, como bloques incompletos o diseño en filas y columnas, es adecuado en la mayor parte de las situaciones (según Cochran y Cox, 1957).

El diseño mas común, de  $t$  tratamientos (fijos), con  $r$  repeticiones (aleatorio) repetidos en  $s$  sitios (aleatorio) tiene como factores de variación en el análisis combinado: sitios, repeticiones dentro de sitios, tratamientos, interacción tratamientos x sitios e interacción tratamientos x repeticiones dentro

de sitios (error experimental o residual), al aumentar los sitios la interpretación de los resultados depende más de la interacción tratamientos x sitios que de tratamientos x repeticiones dentro de sitios (lo que llamamos error).

Como se ha indicado anteriormente, se hacen experimentos repetidos o en condiciones diferentes para saber si los tratamientos responden de igual manera en distintas condiciones. Esto es posible verlo con la significación de la interacción. Si la interacción tratamiento x localidad no es significativa, los tratamientos responden de igual manera en todos los sitios (la ordenación sería la misma) y el estudio concluiría con el análisis de los factores principales, tratamientos y localidades. Si es significativa, quiere decir que unos tratamientos van mejor en unas condiciones y otros en otras condiciones. Hay que considerar que el efecto de los tratamientos da una información relevante solo cuando no existe interacción entre tratamientos y ambientes. Con una interacción significativa las diferencias entre tratamientos pueden variar ampliamente entre ambientes o localidades. Dicho de otra forma, las medias describen adecuadamente las localidades o los tratamientos si la interacción no es significativa. Con una interacción significativa los efectos principales deben examinarse con precaución.

Se puede esperar una gran interacción cuando hay grandes diferencias entre los tratamientos que se ensayan o cuando hay una variación amplia entre los distintos ambientes donde se han realizado los experimentos. Hoy hay disponible mucha información sobre diferentes aspectos de la interacción entre tratamientos y ambientes. Algunas monografías recientes son las de Gausch (1992), Prabhakaran y Jain (1994) y Basford y Tukey (2000).

La interpretación de esta interacción, a veces puede ser complicada, haciéndose necesario estudiar la estabilidad de los resultados a través de los diferentes ambientes. Son los estudios de adaptabilidad y estabilidad de rendimientos. Información específica sobre estos análisis se pueden encontrar en Gausch y Zobel (1996), Piepho (1998) y Annicchiarico (2002).

A la hora de agrupar ensayos es preferible que estos sean iguales, ya que presenta grandes ventajas a la hora de aplicar ciertas pruebas estadísticas (por ejemplo el ANOVA es especialmente robusto para la violación de homogeneidad de varianza y normalidad cuando todos los grupos tienen el mismo tamaño) y para hacer un análisis ponderado dando a cada sitio un peso en proporción inversa al error medio cuadrado (hay que modificarlo si cada sitio tiene diferente número de repeticiones).

Siempre que aparezcan los mismos tratamientos, se puede hacer un análisis combinado, aun cuando los diferentes ensayos difieran en tamaño, estructura o incluso en diseño estadístico (por ejemplo agrupar un ensayo de bloques completos al azar con un cuadrado latino y con un látice) aunque, en este caso, el cálculo del cuadrado medio del error combinado requie-

re cuidado (si tienen mas o menos la misma precision los diferentes ensayos se puede ponderar por el grado de libertad, para mas detalles ver el libro de Cochran y Cox, 1955).

Otras posibilidades es hacer diseños sin repeticiones o diseños aumentados (ver Federer, 1951 o Kempton y Gleeson, 1997). En estos diseños se trata de comparar varios testigos que se repiten para obtener una estimación de error, con nuevos tratamientos en cada localidad. Las localidades deben estar proximas y se recomiendan especialmente cuando hay un número de tratamientos muy elevado. Si los ensayos sin repeticiones se hacen conjuntamente en sitios diferentes de una misma localidad, pueden ser considerados bloques o repeticiones de un mismo experimento en bloques completos ala azar realizado en la misma localidad.

En resumen un investigador que esté interesado en la respuesta de sus tratamientos a diferentes condiciones, intentará conocer que factores ambientales hacen aumentar o disminuir la respuesta, verá que tratamientos son mas estables, que tratamientos van bien en que lugares, ...tendra que conocer la interacción entre tratamientos y sitios.

## 2. ANALISIS PRELIMINARES

Antes de iniciar un analisis combinado es necesario hacer algunos análisis previos, es lo que se entiende por análisis preliminares. Básicamente tienen como objetivo determinar que experimentos de pueden agrupar, es decir cuales deben ser incluidos en el analisis combinado. Estos estudios suelen ser los mismos en todos los casos y son tres:

- 1.- Analizar los ensayos individuales. Es necesario completar cada experimento, analizar cada ensayo y rechazar los que no son válidos. Previamente es conveniente tener un criterio de validez de ensayos (por ejemplo un coeficiente de variación máximo...).
- 2.- Hacer una primera estimación de si las diferencias entre tratamientos son iguales en todos los experimentos y de la interacción entre tratamientos y localidades. Esto permitirá determinar si es necesario separar los experimentos en dos o mas grupos homogéneos para una misma o diferente recomendación.
- 3.- Verificar que se cumplen los requisitos necesarios. Independencia, normalidad y homogeneidad de las varianzas de los errores experimentales (comparar los errores experimentales para ver que son homogéneos).

Finalmente los ensayos se agrupan físicamente.

### 2.1.- Analisis de los ensayos individuales.

Cada ensayo se tiene que analizar e interpretar por separado. El análisis más simple es el análisis de la varianza (para una información detallada del ANOVA, consultar Tejedor, 1994). Para un experimento simple en bloques completos al azar con  $t$  tratamientos sorteados aleatoriamente en  $r$  repeticiones o bloques, el ANOVA tiene las fuentes de variación señaladas en la tabla 2.1

**Tabla 2.1 Análisis habitual de la varianza de un ensayo en bloques completos al azar con  $t$  tratamientos  $t$   $r$  repeticiones.**

Factor de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F
Repeticiones	$r-1$	SC1	CM1	CM1/CM3
Tratamientos	$t-1$	SC2	CM2	CM2/CM3
Repeticiones x Tratamientos (error)	$(r-1)(t-1)$	SC3	CM3	

Un primer paso para interpretar los resultados, es analizar el reparto proporcional de los cuadrados medios (las varianzas). Si la varianza del error (CM3) es proporcionalmente la más grande de las tres, la razón F de los tratamientos será pequeña, es decir indicará pocas diferencias entre tratamientos (realmente podemos decir que no es que no las haya, sino que no hemos sido capaces de encontrarlas). La validez de ensayos se suele comprobar bien por la varianza del error o por el coeficiente de variación (verejemplo). Es de esperar que la varianza del error varíe en los diferentes experimentos.

El siguiente paso es ver la significación de los tratamientos. Si la probabilidad asociada ( $p$  en la tabla de resultados viene en una columna junto al valor de F) es menor que 0.05 se considera que las diferencias entre tratamientos son estadísticamente significativas. Si es significativa, el siguiente paso consiste en determinar que tratamientos son diferentes de los otros y depende de que estos sean aleatorios o fijos. Si son aleatorios, solo es posible hacer un estudio de componentes de la varianza. Si los tratamientos son fijos, el procedimiento habitual como ya se ha indicado antes es hacer una comparación de medias.

Preferentemente debe elegirse comparaciones planeadas "a priori" y debe dejarse constancia de ello en el protocolo. En el caso de hacer comparaciones no planeadas, o "a posteriori", si se hacen pocas comparaciones no suele haber problema. Al aumentar el número de comparaciones aumenta el riesgo de cometer error de tipo I (rechazar la hipótesis nula de que no hay diferencias cuando es verdadera, más fácil, es equivalente a no dar por bueno un tratamiento que es eficaz). Para evitarlo hay que reducir las comparaciones que se quieren hacer o bien, para reducir el riesgo de error, bajar el nivel de alfa de 0.05 a 0.01, entonces las pruebas se hacen más conservadoras y es más difícil encontrar diferencias, aun cuando sean reales. Hay que tener cuidado con las comparaciones a posteriori pues se puede caer fácilmente en un error conocido ("efecto sugerido por los datos" o "razona-

miento posterior a los hechos”). Las medias de los diferentes tratamientos se pueden comparar también mediante contrastes, lo que se verá mas adelante.

Para ilustrar todo lo anterior veamos el siguiente ejemplo

### EJEMPLO I.

*“Se ensayan ocho variedades de remolacha azucarera en seis localidades diferentes, con el objetivo de conocer si la mas productiva lo es en todos los sitios o cada uno requiere una variedad diferente. En cada localidad se hace un experimento en bloques completos al azar (RCBD), sembrando las ocho variedades con cuatro repeticiones. Las localidades difieren geográficamente y en la cantidad de agua que recibe el cultivo, desde secano (sin riego), riegos de apoyo (solo uno o dos riegos) o riego durante el desarrollo del cultivo (entre dos y cuatro riegos)”.*

Como ejemplo del primer análisis se estudia la localidad 4. En SPSS se construye la base de datos en el editor con las variables: repetición (escala, de 1 a 4), tratamiento (nominal, nombres de las variedades de remolacha) y rendimiento (escala, valor en kg/ha de azúcar). El archivo tiene 32 casos. Las variables repetición y variedad se consideran factores fijos. Para obtener el resultado desde el menú se señala (SPSS, versión 8.0):

Statistics

General Lineal Model

GLM – General Factorial

- Dependent variable: Rendimiento
- Fixed Factor(s): Repetic, variedad

Model....

- Custom  
Model: Repetic, variedad
  - Sum of squares: Type III
  - Include intercept in model
- Continue

Post Hoc.....

- Post Hoc test for: variedad
- Tukey´s B

Continue

OK

En la salida (Output –SPSS Viewer) se obtienen las tablas del análisis de la varianza y de la comparación de medias:

**Tabla 2.2 Análisis de la varianza del experimento con ocho variedades y cuatro repeticiones de la localidad 4.**

**Test of Between-Subjects Effects**

Dependent Variable: Rendimiento en azúcar (t/ha)

Source	Type III Sum of Squares	df	Mean Square	F	Sig
Corrected Model	33,09 <sup>a</sup> 1	10	3,309	50,758	,000
Intercept	6052,75	1	6052,75	92842,91	,000
VARIEDAD	32,957	7	4,708	72,218	,000
REPETEC	,133	3	4,448E-C	,682	,573
Error	1,369	21		6,519E-C	
Total	6087,21	32			
Corrected total	34,460	31			

a.R Squared= 9,60 (Adjusted R Squared = ,941)

En el reparto de varianzas, el cuadrado medio de los tratamientos ocupa la mayor parte, siendo el efecto de las repeticiones y del error mucho menores en comparación. El estudio es muy adecuado por este análisis (el valor de r cuadrado es elevado y la significación también) y las diferencias entre variedades en esta localidad son estadísticamente muy significativas.

**Tabla 2.3 Comparación de medias para la variable rendimiento en azúcar del experimento anterior según el test de Tukey HSD a p=0.05**

**Rendimiento en azúcar (t/ha)**

**Tukey<sup>a b</sup>**

variedad	N	Subset			
		1	2	3	4
Alto	4	12.4500			
Pleno	4	12.7500	12.7500		
Hilma	4		13.1000		
Nita	4		13.2500		
Marina	4			13.9250	
Emerita4	4		14.3750		
Arista	4			14.4250	
Armida	4				15.7500

Means for groups in homogeneous subsets are displayed.  
based on Type III Sum of Squares

The error term is MEan Square(Error) = 6,519E-02

a. Uses Harmonic Mean Sample Size = 4,000

b. Alpha = ,05

Se puede apreciar como en esta localidad la variedad Armida es significativamente más productiva que el resto, seguida de las variedades Arista, Emerita y Marina. La variedad Alto es la que ha tenido un menor rendimiento en este ensayo sin ser estadísticamente diferente de la variedad Pleno.

La presentación mas común de estos resultados es la que se muestra en la tabla 2.4 donde se incluyen algunos parámetros estadísticos, necesarios para conocer la calidad del ensayo como el coeficiente de variación o el error estándar.

**Tabla 2.4 Presentación habitual de resultados del ensayo anterior (localidad 4).**

Variedad	Rendimiento (t/ha)	
Alto	12.45	d
Arista	14.42	b
Armida	15.75	a
Emerita	14.37	b
Hilma	13.10	c
Marina	13.92	b
Nita	13.25	c
Pleno	12.75	cd
Media	13.75	
C.V (1)	1.85	
F (signif.p)	72.21 (** 0.000)	
E.E (2)	0.18	

Letras diferentes indican diferencias significativas según el test de Tukey (HSD) a  $p = 0.05$

(T) = Testigos

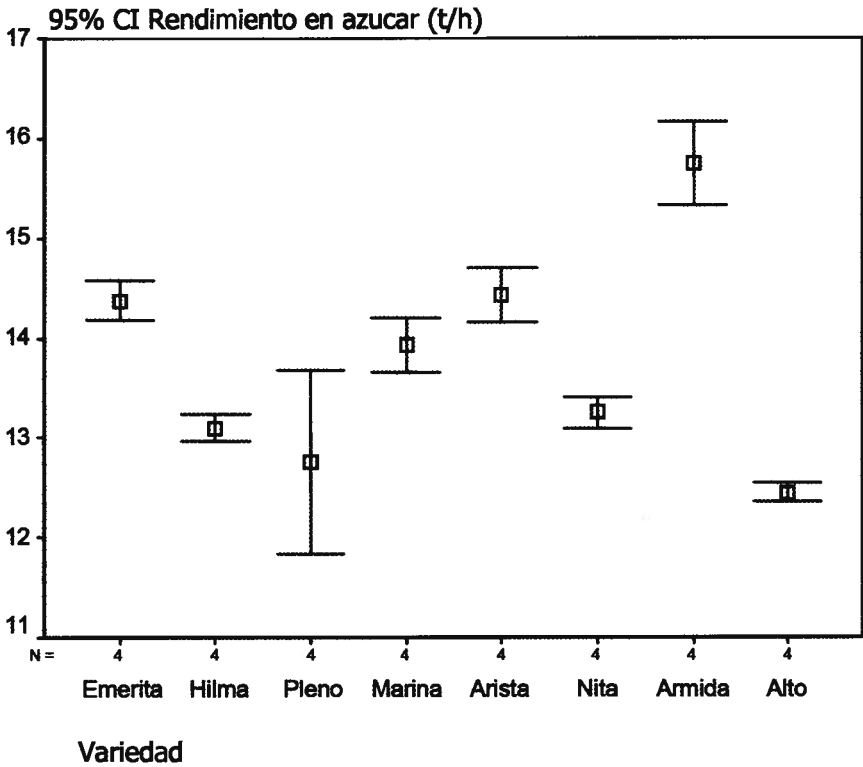
(1) C.V. = Coeficiente de Variación. Es la relación  $C.V = (Cme)^{1/2} / m$ , donde Cme es el Cuadrado Medio del error y m es la media general de los tratamientos. Curiosamente el SPSS (la versión 8) en la estadística descriptiva, no muestra el CV. En algunos protocolos se exige un CV mínimo para que un ensayo sea válido. Es frecuente el valor 10% (Bowman y Watson, 1997). Es un criterio útil para considerar si un ensayo se agrupa o no con el resto.



(2) E.E = Error estándar de la diferencia entre medias de dos tratamientos. Es la relación  $EE = (2 \cdot Cme / n)^{1/2}$ , donde  $n = n^{\circ}$  rep ..... Y el SPSS tampoco lo muestra. Este valor multiplicado por talfa, GL error da el valor crítico de comparación del test LSD (...que el SPSS tampoco lo muestra).

Los paquetes estadísticos actuales, entre ellos el programa SPSS, tienen una gran utilidad que son los gráficos. En las figuras 1 y 2 se muestran los más frecuentes. En la figura 1, el de barras de error (simple error bar) con el intervalo de confianza de la media (intuitivamente y como aproximación, dos variedades cuyos intervalos de confianza no se solapan son estadísticamente diferentes).

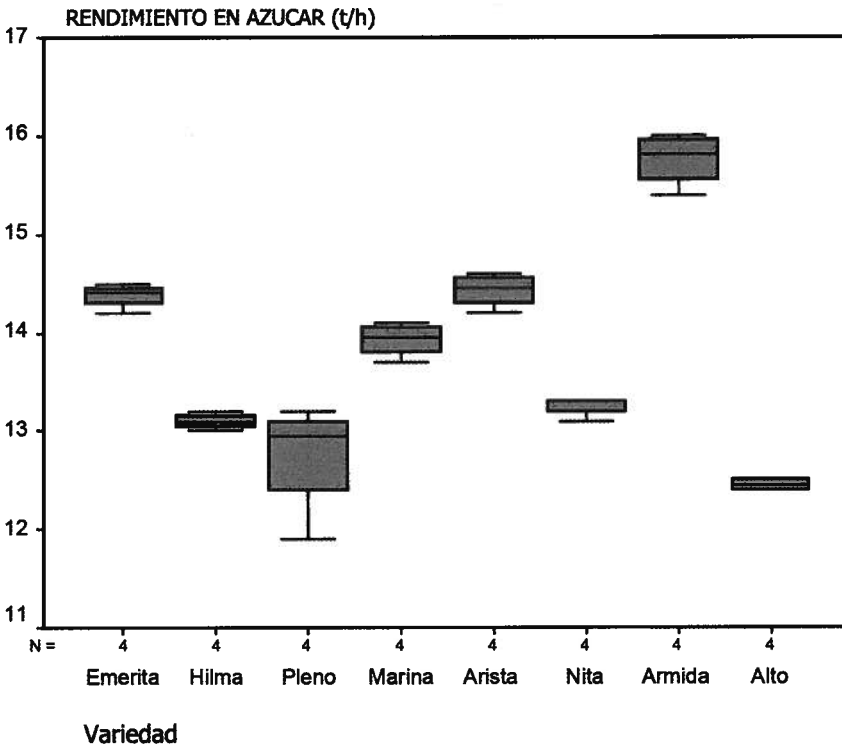
**Figura 1. Gráfico de Intervalos de confianza de la media de rendimientos de ocho variedades en la localidad 4.**



En la figura 2, se presenta el gráfico de cajas (boxplot) de los rendimientos medios de las ocho variedades en la localidad cuatro. Recordamos

que en estos gráficos, muy útiles y cada vez mas empleados, las líneas finales de los segmentos horizontales representan el valor mas alto y mas bajo de la distribución de los datos de cada variedad, la línea interior del recuadro la mediana y los extremos superiores e inferiores del mismo, el cuartil superior y el inferior. En nuestro ejemplo podemos apreciar que la variedad Alto tiene los valores de las cuatro repeticiones mas parecidos y la variedad Pleno es la que los tiene mas diferencia entre el valor mas alto y el mas bajo.

**Figura 2. Gráfico de cajas del rendimiento medio de las ocho variedades del experimento anterior.**



La sintaxis (se obtiene con el comando PASTE) resume las instrucciones que se han dado y en el análisis estadístico anterior han sido:

```

UNIANOVA
rendim BY variedad repetit
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
    
```

```
/POSTHOC = variedad ( BTUKEY )  
/EMMEANS = TABLES(variedad)  
/PRINT = DESCRIPTIVE  
/CRITERIA = ALPHA(.05)  
/DESIGN = variedad repetit .
```

Si los tratamientos hubiesen sido aleatorios, imaginemos, por ejemplo, que las ocho variedades han sido elegidas por sorteo al azar de entre todas las del catálogo o registro nacional, tras el análisis de la varianza, sin el cálculo de la razón F, se hace un estudio de componente de la varianza. El procedimiento es el siguiente:

```
Statistics  
  General Lineal Model  
    Variance component  
      • Dependent variable: Rendimiento  
      • Random Factor(s): Repetit, variedad  
Model....  
  • Custom  
    Model: Repetit, variedad  
  • Include intercept in model  
    Continue  
Options.....  
  • ANOVA  
  • Sum of Squares: Type I  
  • Display: Sums of squares  
    Expected mean squares  
    Continue  
    OK
```

El resultado del análisis de la varianza y la estimación de la varianza de cada componente se presenta en las tablas 2.5 y 2.6. La varianza estimada para las repeticiones es de 0.0025 (a veces tanto el método ANOVA como el MINQUE puede dar valores negativos cuando el verdadero valor está próximo a cero), para las variedades de 1.16 y para el error de 0.065.

**Tabla 2.5 Análisis de la varianza realizado por el procedimiento de componentes de la varianza.**

**ANOVA**

Source	Type I Sum of Squares	df	Mean Square
Corrected Model	33,091	10	3,309
Intercept	6052,750	1	6052,750
REPETIC	,133	3	4,448E-02
VARIEDAD	32,957	7	4,708
Error	1,369	21	6,519E-02
Total	6087,210	32	
Corrected Total	34,460	31	

Dependent Variable: RENDIM

**Tabla 2.6 Estimación de la varianza de las componentes repetición, variedad y error.**

**Variance Estimates**

Component	Estimate
Var (REPETIC)	-2,59E-03
Var (AVARIEDAD)	1.161
Var (Error)	6.519E-02

Dependent Variable: RENDIM

Method: ANOVA (Type I Su of Squares)

La corriente de control o sintaxis en este caso es:

```

VARCOMP
nota BY a b
/RANDOM = a b
/METHOD = SSTYPE (1)
/PRINT = EMS VARCOMP
rendim BY repetit variedad
/RANDOM = repetit variedad
/METHOD = SSTYPE (1)
/PRINT = SS
/PRINT = EMS
/DESIGN = repetit variedad
/INTERCEPT = INCLUDE.
    
```

## 2.2.- Primera estimación del resultado de los tratamientos en las distintas localidades.

Una vez analizados los ensayos individuales (hasta ahora se tienen seis archivos individuales) y descartado los que son poco precisos, es conveniente confeccionar una tabla de resultados con todos ellos juntos. Continuando con el ejemplo de las seis localidades (y supuesto que todos ellos se han aceptado por su CV) los resultados se muestran en la tabla 2.7. Junto a esta tabla, es útil incluir otra tabla donde se especifiquen los cuadrados medios de las fuentes de variación con su significación (ver tabla 2.8).

**Tabla 2.7 Resultados los seis experimentos con ocho variedades de remolacha ordenados por la producción media. Producción de azúcar en t/ha. Loc = Localidad.**

tratamiento	Loc1	Loc2	Loc3	Loc4	Loc5	Loc6
Alto	13.05	12.90	13.40	12.45	15.17	15.78
Arista	13.94	13.69	13.61	14.42	16.08	16.10
Armida	13.52	13.64	14.08	15.75	16.47	16.40
Emerita	14.01	13.89	14.19	14.37	14.10	14.40
Hilma	13.05	13.51	13.51	13.10	12.17	14.10
Marina	13.41	13.58	13.98	13.92	14.90	15.70
Nita	13.28	13.40	13.29	13.25	12.73	13.09
Pleno	13.00	13.80	14.06	12.75	13.08	13.40
Media	13.41	13.55	13.77	13.75	14.34	14.87

**Tabla 2.8. Analisis de la varianza de los ensayos individuales del ejemplo I. Cuadrados medios y significación.**

Fuente de variación	Grados libertad	Loc1	Loc2	Loc3	Loc4	Loc5	Loc6
Repetición	3	0.18	0.08	0.23	0.04	0.62	0.47
variedad	7	0.62NS	0.37NS	0.49NS	4.70**	10.01**	6.59**
Error	21	0.33	0.32	0.57	0.06	2.47	0.27

NS = No significativo, \* =  $p < 0.05$ , \*\* =  $p < 0.01$

Los resultados del análisis de la varianza muestran que hay tres localidades, las tres primeras, donde no hay diferencias estadísticamente significativas entre variedades, en tanto en las tres siguientes si las hay. Posiblemente en ambientes menos productivos las variedades mejores, que parecen ser Armida y Arista (ver tabla 9 en localidades 5 y 6), no puedan manifestar su potencial de rendimiento.

Existe un error bastante comun, que consiste en considerar en el estudio solamente los ensayos que han resultado significativos, despreciando el resto. Recordamos que los ensayos no son buenos o malos por encontrar diferencias significativas entre tratamientos, ya que esta depende de la técnica experimental en gran medida (se puede mejorar el encontrara diferencias entre tratamientos aumentando el número de repeticiones, de tratamientos, haciendo bloques o con medidad intermedias, haciendo análisis de la covarianza). En nuestro caso de no considerar las localidades 1, 2 y 3, se perderá eficacia para conocer la interacción entre variedades y localidades, en definitiva para el estudio global.

La observación de las tablas anteriores debe dar una primera idea de si las diferencias entre tratamientos son iguales en todos los experimentos o no, (dicho de otra forma, si las respuestas de los tratamientos son consistentes de un lugar a otro o hay variación de respuesta) y también de la posible interaccion entre tratamientos y localidades (que unos tratamientos vayan mejor en unas localidades que otras). Esto permitirá determinar si es necesario separar los experimentos en dos o mas grupos homogeneos para hacer una recomendación. Esto a veces no es fácil, especialmente cuando en número de tratamientos y de localidades es elevado. Algunas medidas para facilitar ver las diferencias entre tratamientos en las diferentes localidades son las siguientes:

- a) Es frecuente poner **testigos** para reagrupar. Se eligen uno o dos tratamientos comunes a todos los experimentos y la variable respuesta se transforma en porcentaje respecto a el o ellos. Los valores pasan a ser "relativos respecto a testigos" en lugar de valores absolutos. El uso de testigos tiene básicamente dos ventajas: proporciona una mejor base de comparación, se puede conocer el comportamiento de los tratamientos respecto a los ya conocidos y ayuda a interpretar los resultados (son especialmente necesarios para manejar la impredecible variacion entre años). Las características que deben tener los testigos son ser tratamientos conocidos y empleados, sobre todo a la hora de ensayar nuevas tesis. Pueden llamarse también tratamiento control (si no ha recibido ningún tratamiento) o tratamiento estándar (especialmente si sus efectos ya son conocidos). A veces se repiten varias veces en el experimento para aumentar el conocimiento y siempre deben ser

parte del propio experimento. Como alternativa, se puede emplear la media del campo para relativizar los resultados, aunque tiene como inconveniente las limitaciones conocidas de la media.

Por definición la media de testigos tiene el valor 100 y se suele llamar índice productivo. La significación no varía con los resultados transformados en relativos respecto a testigos, aunque si lo hacen los cuadrados medios y lógicamente, la media. Transformando los valores de la tabla 9, en relativos respecto a las variedades testigo Emérita y Marina (supongamos dos variedades muy conocidas y empleadas), obtenemos la tabla 2.9. Se puede apreciar, por ejemplo, como en la localidades mas productivas las mejores variedades superan a los testigos en casi un 10% en tanto las menos productivas llegan a tener un rendimiento casi un 15% inferior a los testigos.

**Tabla 2.9 Resultados de seis localidades en valores relativos respecto a los testigos Emerita y Marina**

tratamiento	Loc1	Loc2	Loc3	Loc4	Loc5	Loc6
Alto	95.20	93.95	95.22	88.04	104.65	104.90
Arista	101.69	99.70	96.67	102.01	110.97	106.97
Armida	98.65	99.36	100.05	111.38	113.62	109.00
Emerita (T)	102.22	101.22	100.81	101.66	97.24	95.68
Hilma	95.24	98.41	95.96	92.64	83.96	93.68
Marina (T)	97.86	98.96	99.34	98.47	102.75	104.31
Nita	96.91	97.59	94.42	93.70	87.79	86.99
Pleno	94.82	100.50	99.87	90.16	90.25	89.03
Cme(1)	17.74	17.53	28.98	3.26	117.62	12.28

(1) Cuadrado Medio del error

- b) **Análisis de residuos.** Si en el análisis de regresión se entiende por residuos las diferencias entre los valores reales u observados y los predichos por el modelo, para un mismo valor de la variable independiente, en análisis multilocalidades entendemos por residuos los valores que quedan tras quitar a cada unidad experimental (por ejemplo a una parcela...), la media del tratamiento correspondiente, la media de la localidad y sumar la media general. Intuitivamente se puede decir que es lo que aporta "el terreno". En estudios de agronomía son muy útiles para hacer mapas de fertilidad de suelo. De forma esquemática, su cálculo se expone en la tabla 2.10

**Tabla 2.10. Tabla descriptiva para el cálculo de los residuos**

	Loc1	Loc2	Loc3...	
Trat1	R1	R2	R3	Media Trat1
Trat2				
Trat3...				
.....	Media loc1			Media general

$R1 = X1 - \text{Media tratamiento 1} - \text{Media localidad 1} + \text{Media general}$ .

Si los valores de los residuos de un mismo tratamiento en distintas localidades (es decir R1, R2 y R3...) son iguales, existe un comportamiento o efecto similar de este tratamiento en todas las localidades. Podemos completar la tabla con las desviaciones típicas y los rangos de los residuos de cada tratamiento lo que facilitará las comparaciones.

Los residuos se pueden calcular con el programa SPSS, dentro del procedimiento:

- General Lineal Model
- GLM – General Factorial
- Save
- Residuals
- unstandardized

Creándose una variable nueva: res\_1. Lógicamente debe hacerse en un archivo de datos de todos los tratamientos y todas las localidades (como la tabla 2.11). Con los datos de nuestro ejemplo, los residuos de las ocho variedades en las seis localidades se muestran en la tabla 2.11



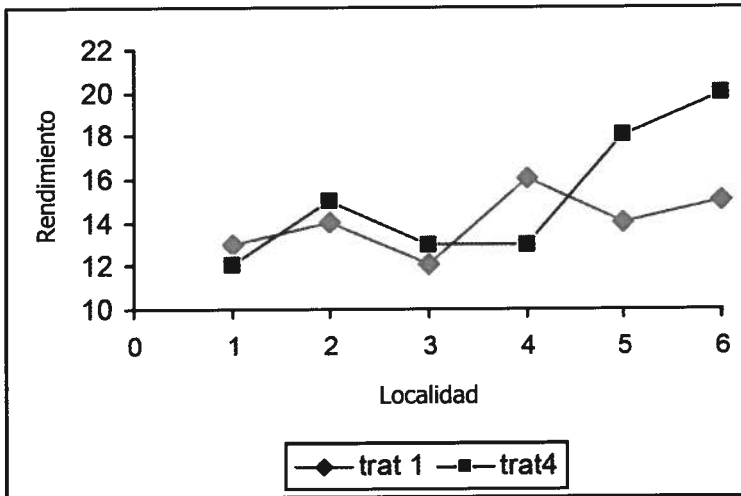
**Tabla 2.11 Tabla de residuos de las ocho variedades en seis localidades. DE = desviación estándar de los residuos de cada variedad. Entre paréntesis el rango de los mismos residuos.**

tratamiento	Loc1	Loc2	Loc3	Loc4	Loc5	Loc6	DE (Rango)
Alto	-0.20	-0.05	-0.21	-1.15	+0.99	+1.06	0.86 (2.2)
Arista	-0.16	-0.55	-0.85	-0.02	+1.05	+0.54	0.69 (1.9)
Armida	-0.92	-0.94	-0.71	+0.97	+1.10	+0.50	0.96 (2.0)
Emerita (T)	+0.39	+0.13	+0.21	+0.41	-0.45	-0.68	0.45 (1.0)
Hilma	+0.35	+0.69	+0.45	+0.06	-1.46	-0.06	0.76 (2.1)
Marina (T)	-0.30	-0.27	-0.09	-0.13	+0.26	+0.53	0.32 (0.8)
Nita	+0.65	+0.62	+0.30	+0.27	-0.83	-1.01	0.73 (1.6)
Pleno	+0.19	+0.85	+0.89	-0.40	-0.66	-0.87	0.76 (1.7)

Se puede ver como las variedades Alto, Arista e Hilma son las más variables de una localidad a otra. Los residuos tienen más utilidades, por ejemplo, se pueden usar como covariables en el análisis de la varianza en el caso que se necesite un análisis más detallado o preciso (que hubiera pasado si se hubiese eliminado el efecto "terreno", si todas las parcelas hubiesen estado en las mismas condiciones)

- c) **Representaciones gráficas.** Si por los datos de las tablas, se observa que un tratamiento va especialmente bien en una localidad, es útil destacarlo en un gráfico de doble eje, donde en el eje de abscisas se sitúan los resultados de todos los tratamientos en la localidad 1 (por ejemplo) y en el de ordenadas las medias del resto de las localidades. Si los tratamientos son pocos, e interesa alguna comparación en especial (por ejemplo con un testigo) es muy útil un gráfico como el indicado en la figura 3 (pueden incluir también las barras de error en cada punto). En él se aprecia como en la localidad 4 tiene un rendimiento más alto el tratamiento 1, en tanto en las localidades 5 y 6 es el tratamiento 4 el de mayor rendimiento.

**Figura 3. Gráfico de rendimientos comparativos entre el tratamiento 1 (T) y el tratamiento en las seis localidades.**



### 2.3.- Requisitos para agrupar los ensayos.

El análisis combinado, como cualquier análisis de la varianza, es un método paramétrico que necesita cumplir una serie de requisitos para poder hacerse.

Estos son: Independencia, normalidad y homogeneidad de las varianzas.

**1.- Independencia.** Puede ocurrir que los resultados de los tratamientos estén relacionados entre una localidad y otra, es decir que exista correlación entre localidades. Si se hace un análisis de la varianza, con datos que estén correlacionados, dependiendo de la intensidad de la correlación, ocurre que las pruebas estadísticas de significación suelen dar mas significacion (un valor de  $p$  mas bajo) de la que realmente existe, es decir se declaran como diferentes efectos que no son tales (se comete error de tipo I). Afortunadamente, esto no es muy frecuente, se suele dar solo cuando las medidas se toman en las mismas parcelas a lo largo del tiempo, por ejemplo en ensayos con cultivos perennes, donde unas medidas dependen de las anteriores.

Se puede comprobar si hay independencia con la matriz de correlación entre ensayos (ies la tabla 9!), que es de la forma:

	Loc1	loc2	loc3	...
Trat1	...	....	...	
Trat2	...	...	...	
Trat3	...	...	...	

El coeficiente de correlación  $r(\text{loc1,loc2,loc3...})$  entre localidades, para los diferentes valores de los tratamientos, indicará si existe relación entre localidades o no. Aunque es difícil, si existe relación solamente entre dos localidades, lo mejor es quitar una de ellas. Si existe una relación general hay que hacer un análisis de medidas repetidas. En el caso del ejemplo 1, el valor de  $r$  con las seis localidades solo es significativo entre las localidades 1 y 4 ( $r=0.70$ ,  $p=0.05$  \*,  $n=8$ ) por lo que es deseable eliminar una de ellas.

En ensayos con árboles, donde es frecuente que el error experimental en un año dependa del de otros, se suele trabajar con el rendimiento total en todos los años, haciendo para cada parcela la regresión lineal del rendimiento sobre años para probar los efectos de los tratamientos.

**2.- Normalidad.** Existen procedimientos muy conocidos tanto de pruebas estadísticas como de gráficos para verificar la normalidad dentro de cada tratamiento. Entre los primeros hay que señalar el test de Kolmogorov – Smirnov y entre los segundos los gráficos de tallos y hojas y los de cajas. La principal limitación para el empleo de pruebas de asimetría y curtosis o el test de chi cuadrado es el tamaño reducido de las muestras, en nuestro caso el número de datos disponible para cada tratamiento. El programa SPSS tiene una de sus mejores utilidades en las pruebas de normalidad, disponiendo de test estadísticos y de gráficos para verificarla. El procedimiento, a partir del archivo de todas las localidades es el siguiente:

```
Statistics
  Sumarize
    Explore
      Dependent list: Rendimiento
      Factor list: Variedad
        • Boxplot (1)
        • Steam-and- leaf (2)
        • Histograma
      - Normality plots with tests (3)
      Continue
```

(1) Gráfico de cajas (ya visto)

(2) Gráfico de tallos y hojas

(3) Proporciona el test de Kormogorov – Smirnov con la significación de Lilliefors para probar la normalidad. El estadístico de Shapiro – Wilk lo calcula en el caso de muestras con menos de 50 observaciones, lo mas usual en nuestro caso. también proporciona un gráfico interesante muy fácil de interpretar, el “normal Q – Q’ plot” donde aparecen los valores observados frente a los predichos si la distridución fuese perfectamente normal. La salida del test de Kormogorov – Smirnov, para los datos de nuestro ejemplo 1 se muestra en la tabla 2.12

**Tabla 2.12 Test de normalidad para las ocho variedades del ejemplo I.**

Tests of Normality

variedad		Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Rendimiento en azucar (t/ha)	Año	,281	6	,152	,852	6	,202
	Arista	,242	6	,200*	,779	6	,044
	Armida	,242	6	,200*	,808	6	,076
	Emerita	,186	6	,200*	,933	6	,554
	Hilma	,218	6	,200*	,956	6	,748
	Marina	,287	6	,135	,879	6	,310
	Nita	,292	6	,119	,846	6	,178
	Pleno	,203	6	,200*	,938	6	,592

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Se puede apreciar que la única variedad que no cumple la normalidad es Arista ( $p = 0.04 < 0.05$ ). Hay que señalar que el análisis de la varianza es bastante robusto frente a la falta de normalidad. Se permite (no causa error de tipo I) que si se comparan cinco grupos uno de ellos no sea normal. En el caso de que mas de un grupo (o distribución de datos en un tratamiento) no sean normales, hay dos posibilidades: aplicar un test no paramétrico (por ejemplo el test de Friedman) o la alternativa mas frecuente: transformar la variable dependiente.

Existe otra prueba de normalidad, el test de Anderson – Darling que es una modificación del anterior dando más peso a las colas. Es más sensible que este pero tiene la desventaja de que los valores críticos deben ser calculados para cada distribución. Es el test que emplea el programa estadístico **Irristat** y se puede encontrar su desarrollo en Stephen (1974).

**3.- Homocedasticidad.** Teóricamente el análisis combinado de la varianza es válido solo si los términos de error de todos las localidades son homogéneos.

Es muy importante que los errores experimentales de los ensayos que se quieren agrupar, sean iguales. Si las varianzas de los errores difieren entre

ensayos, las varianzas son heterogéneas, la prueba F de la interacción tratamiento por localidad no es estrictamente válida (aunque una forma ajustada puede valer como aproximación, ver Yates y Cochran, 1938 ó bibliografía especializada). Para comprobar si se cumple este requisito se han propuesto algunas pruebas y gráficos. Los test no los hacen rutinariamente los paquetes estadísticos y los más usados son:

**1.- Test de Hartley (H)** (también conocido por el test de Fmax). Es una prueba muy simple y rápida para comprobar la homogeneidad de varianzas cuando las muestras son de igual tamaño. Se ha supuesto que se han tomado al azar de diferentes poblaciones y es la relación entre la varianza ( $s^2$ ) mayor y la varianza menor de todas las varianzas de los tratamientos que se comparan (ver Upton y Cook, 2002).

Con los datos del ejemplo 1 (ver tabla 2.11)  $H = s^2(\text{mayor}) / s^2(\text{menor}) = 117.62 / 3.26 = 36.07$ . Para seis cuadrados medios del error o varianzas independientes, cada uno con siete grados de libertad (ya que se comparan ocho tratamientos), el valor crítico de comparación, a  $p=0.05$ , es 10.8 (ver tabla 31 de Biometric tables for statisticians, Vol 1, 3ª edición 1966, o también se puede encontrar en Petersen 1994 u otros textos de estadística). Como 36.07 es mayor que 10.8, concluimos que las varianzas no son homogéneas. Las seis localidades no se deben agrupar. Lo adecuado es eliminar las localidades con varianzas más extremas, tanto por altas como por bajas.

Parece apropiado, en nuestro caso, anular la localidad cuatro por tener una varianza del error demasiado baja frente al resto. Anulando esta localidad y repitiendo el proceso, ahora  $H = 117.62/12.28 = 9.57$ . Al ser el valor inferior a 9.70, valor crítico de comparación con 5 localidades, concluimos que las cinco localidades que hemos dejado si son homogéneas. Esta prueba no está disponible en el programa SPSS (al menos en la versión 8.0).

**2.- Test de Barlett (B)**. Este test se emplea para verificar también la igualdad de varianzas de muestras de poblaciones distribuidas normalmente y está basado en los logaritmos naturales de las varianzas. Su mayor inconveniente es que es muy sensible a la no normalidad de los datos. Para hacer este test el procedimiento es el siguiente: (ver Snedecor y Cochran, 1980 o Petersen, 1994):

Si

GL = Grados de Libertad del error en un ensayo individual

CME = Cuadrado medio del error de una localidad y

A = número de localidades

$M = GL ( a (\ln \underline{CME}) - \sum \ln CMe )$  donde  $\underline{CME} = \sum CMe / a$

y  $C = 1 + (a + 1) / 3 a GL$

La relación M/C es el estadístico (B) para testar la hipótesis de igualdad de varianzas y se distribuye como  $\chi^2$  con  $a - 1$  grados de libertad. Los datos para calcular este test se muestran en la tabla 2.13:

**Tabla 2.13. Cálculos para el test de Barlett**

Localidad	CMe	Ln CMe
1	17.74	2.87
2	17.53	2.86
3	28.98	3.36
4	3.26	1.18
5	117.62	4.76
6	12.28	2.50
total	197.41	17.53

Con GL del error=21 ,,  $a=6$  ,,  $CME=197.41 / 6 = 32.90$  ,,  $Ln CME=3.49$   
 $M = 21 \times (6 \times (3.49) - 17.53) = 71.61$

$C = 1 + 7/3 \times 6 \times 21 = 1.018$  ,,  $X^2 = M/C = 70.34$  Como el valor tabulado es inferior que el valor calculado ( $(X^2)_{0.05, 21} = 32.66 < 70.34$ ) podemos concluir que las varianzas de las seis localidades son heterogeneas. Esta prueba tampoco está disponible en el paquete estadístico SPSS.

**3.-Test de Levene.** Con este relativamente reciente test para comprobar la homogeneidad de las varianzas, se evita el problema del anterior ya que es menos sensible a la no normalidad de los datos. Por el contrario, tiene una limitación y es que puede ser afectado por valores erráticos (son observaciones muy diferentes del conjunto de datos, se suelen considerar así los que se separan mas de 2,5 veces la desviación estándar de la media y se llaman en inglés "outliers", Upton y Cook, 2002). Estos autores proponen para paliar este problema, emplear el test de Brown-Forsythe, que no es mas que trabajar con los valores modificados. En vez de usar  $Z_{jk}$  (valor de  $y - media$ ) emplear  $Z'_{jk} = (valor de y - mediana)$ . Lamentablemente mientras el test de Levene está en el programa SPSS, el test de Brown - Forsythe no esta disponible en SPSS y es tedioso hacerlo a mano. Es preferible revisar cuidadosamente los datos y evitar los outliers, ya que la mayor parte de las veces son debidos a errores de transcripción.

Al hacer el test de Levene, el valor L (levене Statistic) debe ser bajo. Un valor asociado de p (en significación) menor que 0.05 indica que las varianzas no son homogéneas, por lo que, como siempre, se rechazaría la hipótesis nula. En SPSS se puede hacer por explore (siguiendo el camino indicado

en las pruebas de normalidad) o bien por análisis de la varianza de una vía (oneway ANOVA). Este último es el siguiente:

```

Statistics
  Compare mean
  One Way ANOVA
    Dependent list: Rendimiento
    Factor : Variedad
Options:
  • Homogeneity – of - variance
  Continue
  OK
    
```

A partir de archivo conjunto de todas las variedades en todas las localidades, el resultado para la variable en rendimiento relativo se indica en la tabla 2.14.

**Tabla 2.14. Resultados de aplicar el test de Levene al archivo con las seis localidades.**

**Test of Homogeneity of Variance s**

RENDPORC

Levene Statistic	df1	df2	Sig.
8.909	5	42	,000

La probabilidad asociada a la prueba es menor que 0.05, por lo que se puede concluir como en las pruebas anteriores, que las varianzas son heterogéneas.

Si se quita la localidad 4 (simplemente se pueden ocultar los casos en el archivo de datos), el valor de L ha disminuido y la significación es mayor que 0.05, por lo que se puede concluir que ahora las varianzas son homogéneas (ver tabla 2.15).

**Tabla 2.15. Resultados de aplicar el test de Levene al archivo con cinco localidades.**

**Test of Homogeneity of Variance s**

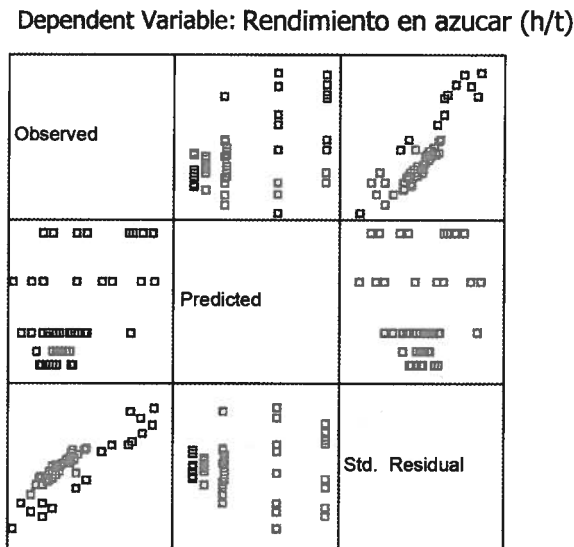
RENDPORC

Levene Statistic	df1	df2	Sig.
771	4	35	,475

**4.- Procedimientos gráfico en SPSS.**

- a) Gráficos de cajas. Este gráfico (ver figura 2) realizado a partir del archivo general, donde se sustituyen las variedades por las localidades, permite apreciar (de forma aproximada) si las varianzas de las localidades que se comparan son comparables.
- b) Gráfico de residuos. En el procedimiento de modelos generales lineales, existe la opción gráficos residuales, que produce un gráfico múltiple entre valores predichos, observados y residuos. El gráfico de los residuos estandarizados y los valores predichos (en el centro de los tres gráficos de la parte inferior) dá una idea del comportamiento de las varianzas (es el mismo sistema al que se emplea para verificar los requisitos de la regresión). En nuestro caso hay una tendencia clara a incrementar la variabilidad cuando la variable independiente aumenta (Ver figura 4 ).

**Figura 4. Gráficos del procedimiento General Lineal Model**



Model: Intercept + LOCALIDA

El procedimiento y las instrucciones son las siguientes:

General Lineal Model

GLM – General Factorial

Model : Hacer sólo con localidad

Options:

- Residual plot

Continue



Si no se cumple la condición de homogeneidad de varianzas y se hace el análisis combinado, se producen demasiados resultados significativos. Para resolver el problema de la heterocedasticidad hay varias posibilidades:

1) Eliminar las localidades que tienen muy alta o baja varianza del error

Esta opción es interesante en el caso de ser una o dos las localidades que se eliminan (a veces hay que eliminar tantas que invalidan el estudio multi-localidad). Es lo que se ha visto en el apartado anterior. Si es necesario considerar todas las localidades (puede perder sentido el programa de investigación si se eliminan), la opción es dividir las localidades en grupos con varianzas homogéneas y el análisis combinado se hace para cada grupo. Esta opción tiene la limitación de hacer recomendaciones solamente sobre un área. Si se hacen comparaciones sobre un subconjunto de lugares, la interacción tratamientos por localidades debe ser separada en tratamientos por subconjuntos y tratamientos por localidades dentro de subconjuntos. Si el cuadrado medio de la primera es mayor que la de la segunda, es recomendable analizar por separado los resultados para cada subconjunto (para más detalles ver Snedecor y Cochran 1967).

Alternativamente, Petersen (1994), propuso partir desde la suma de cuadrados en una serie de componentes ortogonales, llegando a obtener la interacción componente por localidad de cada componente. Se prueba cada componente de tratamiento con su propia interacción con la localidad. Debido a los grados de libertad del denominador, es un test bastante conservador. Este procedimiento se verá detenidamente más adelante.

2) Aumentar el nivel de riesgo que se asume (alfa)

Tiene un inconveniente, que aumentan los intervalos de confianza y hace que la prueba sea más estricta, de ahí que se emplee poco. Esta instrucción se modifica en SPSS de la manera siguiente:

General Lineal Model

GLM – General factorial

Options:

- Significance level:

Continue

Las varianzas de los errores, menudo difieren considerablemente de un sitio a otro. Si esto ocurre y se procede con el análisis combinado como si las varianzas fueran homogéneas, cuando se utiliza en el denominador de la prueba F el término de error combinado, la heterogeneidad tiende a invalidar la prueba para comparaciones en las que entra la interacción de tratamientos por localidad. El resultado es que la prueba F puede producir dema-

siados resultados significativos, sin que lo sean en realidad. Cochran y Cox en 1957 y mas tarde Steel y Torrie en 1986, propusieron una solución prudente: Se trata de comparar el valor de F calculado, con el valor de F tabulado para  $(t - 1)$  y  $n$  grados de libertad, donde  $t$  es el número de tratamientos y  $n$  el número de grados de libertad del error en un solo ensayo (en vez de compararlo con el valor de F tabulado para los grados de libertad de la interacion y el error combinado). Al ser esta modificación conservadora, si la prueba es significativa hay pocas dudas según estos autores. Si es negativa no sirve.

### 3) Hacer un agrupamiento ponderado.

Cuando las varianzas entre ambientes diferentes son heterogéneas, también se ha propuesto emplear medias ponderadas. Estas medias pueden ser inversamente proporcionales a la varianza del error o al cuadrado medio entre localidades. Dan un menor error estándar que la media aritmética y un valor crítico de comparación (LSD) mas alto (por lo que es mas difícil encontrar diferencias que sean significativas, Bernardo, 1992). Con el ejemplo, se pueden ponderar los resultados creando una nueva variable dividiendo el rendimiento en azúcar por el cuadrado medio del error de cada ensayo. Se hace fácilmente en SPSS.

1.- Crear una nueva variable :CME (cuadrado medio del error para cada ensayo)

2.- En Data

Weight Cases:

- Weight cases by : CME (la nueva variable)
- OK

Los resultados de los análisis parciales de localidades y variedades (sin considerar las repeticiones dentro de cada localidad), con las medias aritméticas y las medias ponderadas por la varianza del error, muestran que las variedades dejan de ser significativas (en el primer caso  $F = 4.60$  con  $p = 0.01$  y en el segundo  $F = 1.76$  con  $p = 0.127$ ) bajando el cuadrado medio del error de 28.16 a 0.899. Para lo que interesa de reducir la homocedasticidad, el test de Levene en el primer caso es de  $L = 8.90$  con  $p = 0.000$  y en el segundo  $L = 15.57$  con  $p = 0.000$  también. La transformacion ha tenido efecto sobre los resultados y no ha sido util para evitar la homocedasticidad.

Algun autor, en un trabajo reciente (Hühn, 1997), ha recomendado (lo mismo que Yates y Cochran en 1938), emplear para comparar los tratamientos siempre las medias aritméticas, antes que las medias ponderadas, incluso si las varianzas de los errores son heterogéneas. Otro autor, tambien recientemente (Yndgaard, 2002), recomienda ponderar las medias por el valor recíproco de la varianza del error, cuando hay una amplia variación de

las varianzas de los errores en los diferentes ensayos. Esta medida, indica este autor, elimina la necesidad de no considerar algunos ensayos, lo que siempre es una medida subjetiva. No se descarta que en alguna ocasión las medias ponderadas sean útiles para eliminar la heterogeneidad de las varianzas, pero hay que considerar que también tiene otras consecuencias.

#### 4) Transformar la variable.

Suele ser bastante usual la heterogeneidad de los errores experimentales. Los errores experimentales raramente son homogéneos en un conjunto de ensayos repetidos en varias localidades. Casi siempre están influidos por circunstancias específicas y tienden a ser menores en ambientes de bajo rendimiento. Para evitar la heterogeneidad de los errores los resultados se deben transformar para producir varianzas que sean homogéneas.

La transformación de datos tiene la ventaja de facilitar el que se cumplan los requisitos y la dificultad de interpretar los resultados. Cualquier análisis

estadístico (incluyendo, por supuesto, el análisis combinado) se debe hacer siempre con los datos transformados. Para facilitar la comprensión de los resultados, se suelen presentar estos con los valores reales y destacando que el análisis se ha hecho con los datos transformados, indicando siempre el tipo de transformación realizado. Existe mucha información sobre transformaciones (algunas referencias, por ejemplo, Bartlett, 1947, Little y Hills, 1978, Steel y Torrie, 1985, Gumpertz, 1995, Norman y Steiner, 1996...)

Análizando los resultados de la variable respuesta, hay que pensar en transformar los datos si están muy sesgados, si la distribución es poco normal, si los valores más altos o más bajos son los más abundantes (para todo esto es necesario representar gráficamente los datos), si las medidas están hechas en escalas arbitrarias (generalmente no hay que transformar si se mide en kg, m, ha...), si se comparan localidades con tamaños diferentes (caso de que falten tratamientos en algunas localidades...) o si las varianzas de errores son heterogéneas.

La transformación más adecuada suele depender del tipo de dato. Como orientación en la tabla 2.16 se indican algunas transformaciones recomendadas.

**Tabla 2.16. Transformaciones recomendadas a partir del tipo de dato.**

Tipo de dato	Transformación recomendada
Enteros pequeños (3,4,12...) ó porcentajes extremos (0-20% ó 80-100%)	Raiz cuadrada (x)
Enteros muy pequeños con ceros (0.5,1,0.8,...todos<10)	Raiz cuadrada (x+1/2) ó log(x+1/2). Añadir 1 si hay ceros
Enteros positivos en un gran intervalo (47,837,21,...todos>10)	log (x)
Medidas de escala con los valores bajos mucho mas abundantes. (1,1,2,1,2,1,1...5,9 en escala 1-10)	1/x Si los mas numerosos son los mas altos: - 1/x
Porcentajes con un intervalo amplio de valores (2%,75%...)	arcsen(raiz cuadrada(x))

La transformación mas adecuada se puede calcular de una forma mas precisa por la relación entre medias y varianzas de las variables (ver Norman y Steiner, 1996). Este procedimiento se emplea cuando al menos son tres las localidades a comparar. Es necesario revisar los resultados tras las transformaciones, pues se sabe que con algunas de ellas (por ejemplo con la transformación logaritmica en ciertos datos) la interacción entre tratamientos y localidades deja de ser significativa. En este caso hay que elegir otra transformación o procedimiento.

En SPSS, la transformación se puede hacer desde:

- 1.- La ventana de LEVENE:Transformado:Raiz...o algunas más.
- 2.- Barra de tareas inicial:  
Transform

Compute....

- Target variable (crea una nueva variable)
- Numeric expression... (a partir de la variable existente que se desee)

OK

En este punto se tienen varios ensayos válidos, cumpliéndose todos los requisitos para agruparlos. Para crear el archivo con todos los ensayos:

- 1.- Se crea una nueva variable (localidad, sitio...)
- 2.- A partir del archivo de datos de la primera localidad: Copiar  
+ **pegar** uno a uno cada ensayo sobre el primero.

### 3.- ANALISIS COMBINADO DE LOCALIDADES ANALISIS PRINCIPAL

El análisis combinado es un análisis de la varianza donde las fuentes de variación principales son los tratamientos, las localidades y los años. Básicamente existen dos modelos de análisis de la varianza, que difieren en número de factores, solo localidades o años y localidades y años. En el segundo caso existen también a su vez varios modelos en función de que los factores localidades y años estén cruzados o anidados.

El primer modelo incluye tres factores principales: los tratamientos, las localidades (o ambientes) y los bloques (o repeticiones) dentro de las localidades. Este modelo se usa en ensayos que no se repiten en el tiempo y tratamientos con localidades están cruzados, es decir todos los tratamientos están en todas las localidades. El segundo modelo incluye cuatro factores: los tratamientos, las localidades, los años (u otro factor de tiempo) cruzado con las localidades y los bloques dentro de las localidades y años. En otro caso este modelo es idéntico al anterior salvo que el factor tiempo esta anidado o jerarquizado al de localidades. Se usa particularmente cuando la comparación de años difiere entre localidades. Puede existir otra posibilidad que es el caso de que se consideren las localidades anidadas a los años.

En este capítulo se analizará el primer modelo, el analisis combinado de tratamientos y localidades, los otros modelos, se describen en el capítulo cinco.

El análisis combinado de una serie de experimentos también se llama análisis principal. Es una forma de análisis factorial donde los factores de variación son los tratamientos, las localidades y la interacción tratamiento por localidad. Este sería el modelo que se puede llamar "simple", que es insuficiente. Para el modelo completo, falta un nuevo factor de variación, las repeticiones dentro de las localidades, que constituyen un nuevo error que se emplea para probar la significación de las localidades (salvo en el caso de que ambos factores sean aleatorios, como se verá mas adelante). Esta par-

particular estructura con tratamientos, repeticiones y localidades ha hecho que se propongan dos alternativas para el análisis de la varianza: La primera tiene como factores las localidades, las repeticiones dentro de las localidades, los tratamientos, la interacción tratamiento por localidad y el error o residual. Esta estructura de análisis de la varianza la llamaremos convencional. Se podría llamar un modelo parcialmente jerarquizado de ANOVA. La segunda tiene como factores de variación las repeticiones, las localidades, la interacción entre repeticiones y localidades (se suele llamar error a), los tratamientos, la interacción entre tratamientos y localidades y el error general (llamado error b). Esta segunda estructura se diferencia de la primera en la descomposición del factor repeticiones dentro de localidades (en el primer caso juntos) en repeticiones y repeticiones por localidades y se conoce como parcela dividida.

El modelo general es válido y es el siguiente: Si varios tratamientos son evaluados en varios ambientes, el rendimiento del tratamiento  $t$  en la repetición  $r$  del ambiente  $j$  puede ser expresado de forma aditiva como:

$$Y_{jtr} = \mu + T_t + \theta_j + (T\theta)_{tj} + e_{tjr}$$

Donde:  $Y_{jtr}$  = media observada del tratamiento  $t$  en la localidad  $j$  y la repetición  $r$

$\mu$  = gran media

$T_t$  = efecto del tratamiento  $t$

$\theta_j$  = efecto de la localidad  $j$

$(T\theta)_{tj}$  = interacción del tratamiento  $t$  con la localidad  $j$

$e_{tjr}$  = media de los errores de las  $r$  parcelas que han recibido el tratamiento  $t$  en la localidad  $j$

Para este modelo se asume que las medias de todos los sumandos son  $N(0, \sigma^2)$ , normalmente distribuidas con media cero y varianza  $\sigma^2$  y se supone que los errores aleatorios medios,  $e_{tjr}$ , siguen una distribución normal  $N(0, \sigma^2 / r)$  y son independientes, donde  $\sigma^2$  es la varianza del error dentro de los ambientes que se suponen homogéneos (según se ha comprobado antes) y  $r$  es el número de repeticiones en una localidad.

Bajo las condiciones anteriores, los cuadrados medios esperados de los dos errores son para el error general o combinado:  $\sigma^2_1$  y para el de repeticiones dentro de localidades:  $\sigma^2_1 + t \sigma^2_2$ , siendo  $\sigma^2_2$  la varianza de las localidades. El error combinado (Cme) es la suma de los errores de todas las localidades (recordamos que en cada localidad, es la interacción entre tratamientos y repeticiones) y se emplea siempre como denominador en la prueba F de la interacción tratamientos por localidad. La interacción entre

tratamientos y localidades puede ser significativa o no. Conocer esto es uno de los principales objetivos del análisis combinado.

En el análisis combinado es interesante observar el reparto de varianzas en el ANOVA. Si el cuadrado medio de la interacción entre tratamientos y localidades, es muy superior al cuadrado medio de la interacción entre tratamientos y repeticiones (lo que es el error combinado) indica que las diferencias entre tratamientos no son iguales de un ensayo a otro y los experimentos deberían agruparse en grupos más homogéneos (dando lugar a recomendaciones diferentes). Es decir, existe interacción entre tratamientos y ensayos.

Si no existe interacción tratamiento por localidad, esta no es significativa, el estudio concluye con las pruebas de significación de los factores principales localidades y tratamientos. En este caso para saber si las diferencias entre medias de tratamientos son significativas o no, en el denominador para la prueba F se puede emplear el Cme o la suma del cuadrado medio de la interacción y del cuadrado medio del error (esta última es una opción más conservadora, según propuestas de Cochran y Cox, 1957 y Tejedor, 1999).

Asumiendo que las repeticiones son siempre al azar, el efecto principal de las localidades se prueba siempre con el error repeticiones dentro de localidades.

### **3.1.- ANÁLISIS CONVENCIONAL**

En el caso más común de una serie de experimentos iguales en bloques completos al azar (RCBD) realizados en distintas localidades y en idénticas condiciones experimentales, en función de que los tratamientos y las localidades sean fijas o aleatorias, el análisis de la varianza es el indicado en la tabla 3.1 (propuesta por McIntosh, 1983).



**Tabla 3.1. Valores de F según que los factores tratamientos y localidades sean fijos o aleatorios (McIntosh, 1983).**

Factor de variación	Grados de libertad	Cuadrado Medio	F1	F2	F3
Entre Loc	I-1	CM1	(CM1+/CM2) C8M2+(CM4)(*)	CM1/CM2	CM1/CM2
Rep. dentro de loc	I(r-1)	CM''			
Trat	t-1	CM3	CM3/CM4	CM3/CM4	CM3/CM5
Loc x trat	(t-1)(I -1)	CM4	CM4/CM5	CM4/CM5	CM4/CM5
Error combinado	I(r-1)(t-1)	CM5			
Total		Irt-1			

F1 = Si Localidades y tratamientos factores aleatorios

F2= Si Localidades aleatorias y tratamientos factor fijo

F3= Si Localidades y tratamientos factores fijos

No se considera el caso en que las localidades sean fijas y los tratamientos aleatorios por carecer de sentido real.

En el caso de que ambos factores tratamientos y localidades sean aleatorios, (F1\*), la prueba para medir el efecto de las localidades no es exacta. Satterthwaite (1946) proporcionó una aproximación, usando N'1 y N'2 grados de libertad en el numerador y denominador, donde

$$N'1 = (CM1+CM5)^2 / (CM1^2 / (I-1) + CM5^2 / I (r-1)(t-1))$$

$$N'2 = (CM2 + CM4)^2 / (CM2^2 / I(r-1) + CM4^2 / (I -1)(t-1))$$

Con I localidades, r repeticiones y t tratamientos.

Antes de este calculo tedioso, McIntosh (1985) propuso para ahorrar tiempo y esfuerzo, testar la significacion de las localidades usando los grados de libertad de las localidades para el numerador y los grados de libertad de repeticiones en localidades o la interacción d etratamientos por localidades en el denominador, cualquiera de estas que sea la mas pequeña. Este es un test muy conservador y si F es significativo no hay que calcular N'1 y N'2. Si F no es significativo, hay que hacerlo ya que usando N'1 y N'2 es posible que sea significativo.

Generalmente los tratamientos y las localidades se consideran efectos fijos (quizás un poco gratuitamente, según indican Cochran y Cox, aunque los tratamientos, casi siempre, son fijos) en tanto los años se consideran,

logicamente, siempre al azar. Un modelo con efectos fijos y aleatorios se denomina mixto.

La estimación de las componentes de la varianza, para el caso de efectos aleatorios se puede consultar en Anacchiaricco (2002) y Dagnelie (1975).

En el cálculo de la prueba F para los tratamientos se puede usar como denominador el cuadrado medio de la interacción tratamiento por localidad o el cuadrado medio del error como se ha indicado. Si la interacción ha resultado significativa, es preferible emplear el primero, aunque como el cuadrado medio del error usualmente tiene más grados de libertad que el cuadrado medio de la interacción, suele ser un mejor estimador. El cálculo de las pruebas F, cuando se emplea el programa SPSS, es mejor hacerlo a mano. Este programa (como todos) considera adecuadamente los denominadores en las pruebas F, según sean los factores fijos o aleatorios, pero no distingue los términos de error diferentes del error general.

Para hacer el análisis de la varianza mostrado en la tabla 3.2 a partir del archivo de datos general (con las variables: repetición – variedad – localidad – rendimiento) desde el menú principal, en la barra de tareas se señala:

#### Statistics

##### General Lineal Model

##### GLM – General Factorial

- Dependent variable: Rendimiento
- Fixed Factor(s): Repetic, variedad

1.- Model....(Nota: El modelo se construye en el orden que aparece en la columna de factores de variación del Anova combinado)

- Custom
- Model: localidad, repeticion x localidad, variedad, variedad x localidad.

Sum of squares: Type III

Include intercept in model

Continue

2.- Plots.....(Nota: En el caso de que salga significativa la interacción entre tratamientos y localidades, la forma mas fácil de verlo es gráficamente)

- Horizontal axis: localidad
- Separate lines: Variedad

Add

Continue

3.- Post Hoc.....

- Post Hoc test for: variedad
- Tukey´s B

Continue

4.- Options.....

- Display means for: localidad x variedad

(Nota: Esta tabla de localidad por variedad es útil para ver la interacción)

Continue

OK

Con los datos del ejemplo I (cinco localidades), considerando todos los factores fijos, los resultados del análisis de la varianza se indican en la tabla 3.2.

**Tabla 3.2. Analisis de la varianza combinado de 8 variedades repetidas 4 veces en cinco localidades del ejemplo 1. Rendimiento en porcentaje respecto a los testigos Emerita y Marina.**

Tests of Between-Subjects Effects

Dependent Variable: PERCENT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6187,120 <sup>a</sup>	54	114,576	2,950	,000
Intercept	1549665,0	1	1549665,0	39904,177	,000
LOCALIDA	39,172	4	9,793	,252	,908
REPETIC * LOCALIDA	228,753	15	15,250	,393	,978
VARIEDAD	2618,752	7	374,107	9,633	,000
VARIEDAD * LOCALIDA	3300,443	28	117,873	3,035	,000
Error	4077,639	105	38,835		
Total	1559929,7	160			
Corrected Total	10264,759	159			

a. R Squared = ,603 (Adjusted R Squared = ,398)

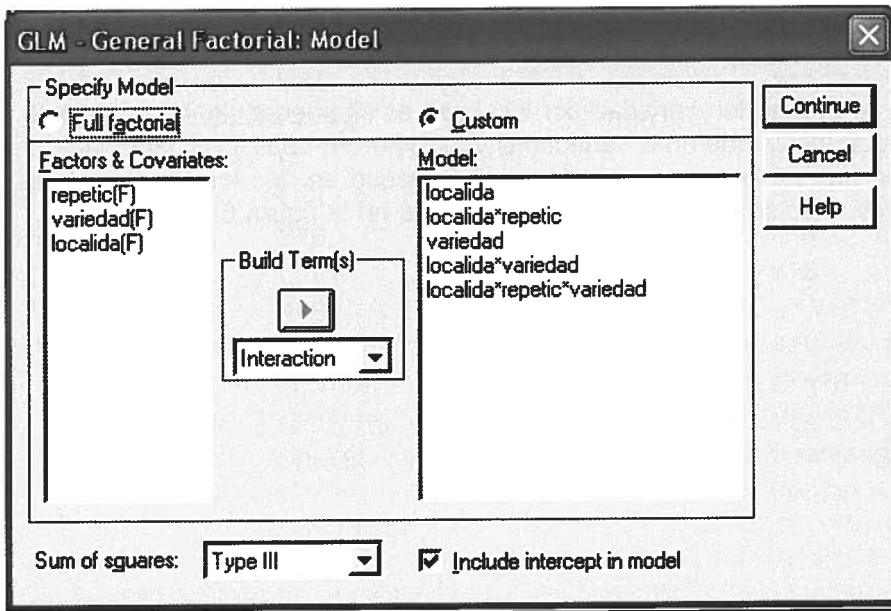
Observando la distribución de la suma de cuadrados, el efecto principal de los tratamientos explica el **25,5 %** ( $2618 \cdot 100 / 10264$ ) del total de la suma de cuadrados, mientras las diferencias entre localidades tan solo el **0,38 %** ( $39,17 \cdot 100 / 10264$ ) y la interacción contribuye en un **32,1 %** ( $3300 \cdot 100 / 10264$ ). La mayor suma de cuadrados de variedades que de localidades indica que las primeras son muy diferentes y las principales causantes, junto con la interacción, de la variación en el rendimiento de azúcar.

Acorde con las instrucciones señaladas en la tabla 3.2 el valor de F de la interacción variedad por localidad es  $117,87 / 38,83 = 3,035$  es significativo y se indica correctamente en la tabla 3.3. Para las variedades, (caso F3) el valor de F es  $374,10 / 38,83 = 9,633$  es significativo y también se muestra

correctamente en la tabla 20. Para las localidades, el valor de F es  $9,79/15,25 = 0,64$  no es significativo y **diferente** del expuesto en la tabla (0,252). Como se ha indicado anteriormente, el programa no reconoce el factor `repetic * localida` como un error y divide el cuadrado medio de las localidades por el cuadrado medio general. Hay que tener una especial precaución con esto.

Para evitar la posibilidad de equivocación, una alternativa es hacer el modelo completo, es decir, incluir el término de error también en el modelo, con lo que al no tener término de error para comparar, la columna de la F está vacía. Hay que tener la precaución de identificar correctamente en este caso lo que es el error. Ver figura 5 para la construcción del modelo completo y la tabla 3.3 para el análisis de la varianza.

**Figura 5. Factores de variación del modelo completo (sin concesión para el error).**



**Tabla 3.3. Análisis de la varianza del modelo completo (sin error).**

**Tests of Between-Subjects Effects**

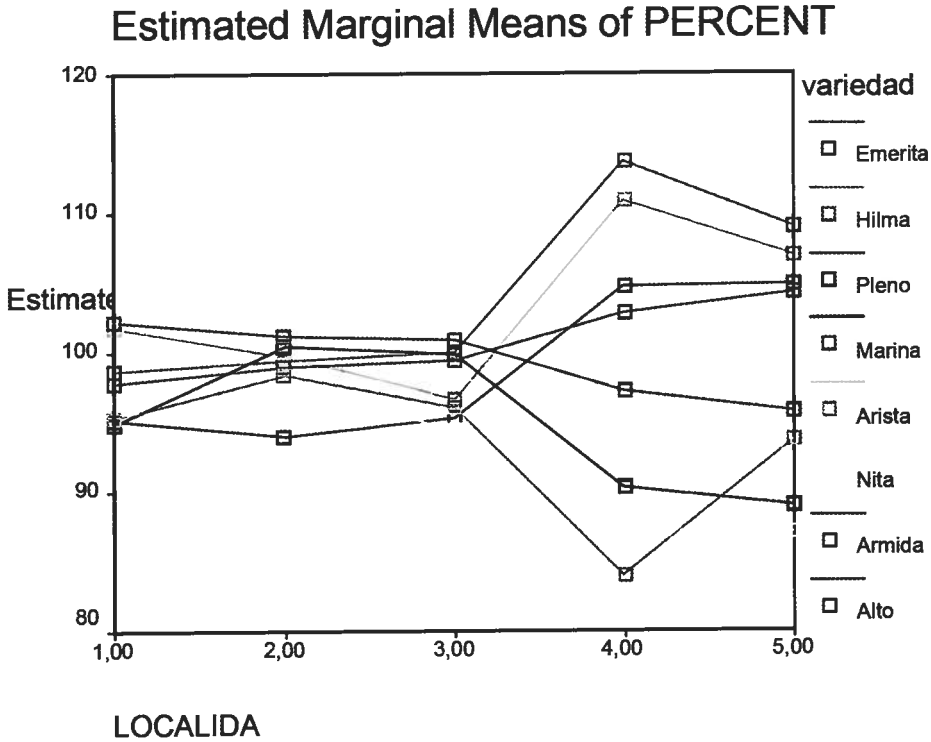
Dependent Variable: PERCENT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10264,759 <sup>a</sup>	159	64,558	,	,
Intercept	1549665,0	1	1549665,0	,	,
LOCALIDA	39,172	4	9,793	,	,
REPETIC * LOCALIDA	228,753	15	15,250	,	,
VARIEDAD	2618,752	7	374,107	,	,
VARIEDAD * LOCALIDA	3300,443	28	117,873	,	,
REPETIC * VARIEDAD * LOCALIDA	4077,639	105	38,835	,	,
Error	,000	0	,		
Total	1559929,7	160			
Corrected Total	10264,759	159			

a. R Squared = 1,000 (Adjusted R Squared = ,)

La interaccion variedad por localidad es altamente significativa ( $p < 0.01$ ) lo que indica que unas variedades van mejor en unas localidades que otras, o lo que es lo mismo, que la mejor variedad en una localidad no lo es en otras. Esto se puede ver de forma gráfica en la figura 6.

Figura 6. Valores medios de las 8 variedades en valores relativos respecto a testigos en las 5 localidades.



Se puede apreciar como en las localidades 1, 2 y 3 las diferencias en rendimiento entre las variedades es pequeña (realmente no existen diferencias significativas), en tanto en las localidades 4 y 5 se producen importantes diferencias. En estas dos localidades, las variedades Armida y Arista son las mas productivas en tanto Hilma es poco aconsejable para la localidad 4 y Nita y Pleno no lo son para la localidad 5.

Si la interacción entre tratamientos y localidades fuese no significativa, el cuadrado medio de la interacción puede partirse en diferentes componentes, comparando los tratamientos mas relevantes o significativos con las localidades y analizando su significación. Esta comparación se puede hacer mediante los contrastes ortogonales. Basicamente consiste en subdividir las sumas de cuadrados y grados de libertad del analisis de la varianza. Si en el analisis que nos ocupa ( tabla 3.3), las variedades tienen 7 grados de libertad, cada variedad tiene un contraste de 1 grado de libertad. Si la interacción variedad por localidad tiene 28 grados de libertad, entonces cada inter-

acción particular, variedad a x localidad, variedad b x localidad... tiene 7 grados de libertad. Para el cálculo de las F correspondientes, las componentes de la interacción se prueban con el error variedad x localidad. Cuando una interacción variedad a x localidad es significativa indica que el comportamiento individual de la variedad a depende de la localidad. Los aspectos prácticos de esta técnica que se verán mas adelante.

A efectos de presentar los resultados, si existe interacción tratamientos por localidad se deben hacer tablas de dos vías, si la interacción no es significativa es mas útil hacer tablas de una via con los medias efectos principales.

Aunque tiene un valor relativo la información del resultado medio de las variedades, ya que su comportamiento como se ha visto depende de la localidad, a efectos de ver la presentación y dado que el factor variedades ha resultado también altamente significativo ( $p < 0,01$ ), el siguiente paso sería comparar las medias para saber cuales son las de rendimiento medio mas elevado. Los resultados se muestran en la tabla 3.4. Esta tabla sería el resultado del análisis combinado en el caso de que la interacción hubiese resultado no significativa (en el caso del ejemplo I, donde la interaccion es significativa, el analisis combinado de los datos de todas las localidades tiene poco significado).

**Tabla 3.4. Comparación de medias para la variable rendimiento relativo según el test de Tukey HSD a  $p=0.05$**

**PERCENT**

Tukey B<sup>a,b</sup>

variedad	N	Subset			
		1	2	3	4
Nita	20	92,7452			
Hilma	20	93,4559	93,4559		
Pleno	20	94,9004	94,9004	94,9004	
Alto	20		98,7877	98,7877	98,7877
Emerita	20			99,4368	99,4368
Marina	20				100,6498
Arista	20				103,2018
Armida	20				104,1381

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 38,835.

a. Uses Harmonic Mean Sample Size = 20,000.

b. Alpha = ,05.

La corriente de control del análisis combinado con las gráficas y comparación de medias de variedades es la indicada a continuación:

UNIANOVA

percent BY repetit variedad localida

/METHOD = SSTYPE(3)

/INTERCEPT = INCLUDE

/POSTHOC = variedad ( BTUKEY )

/PLOT = PROFILE( localida\*variedad )

/EMMEANS = TABLES(localida)

/PRINT = HOMOGENEITY

/CRITERIA = ALPHA(.05)

/DESIGN = localida localida\*repetit variedad localida\*variedad .

El paso siguiente de la interacción tratamiento por localidad significativa se presenta en el siguiente capítulo.

### 3.2.- ANÁLISIS COMO PARCELA DIVIDIDA (SPLIT – PLOT en inglés)

Este conocido diseño de experimentos se emplea cuando algunas combinaciones de factores son difíciles de hacer y especialmente cuando el tamaño de las unidades experimentales donde se aplican los dos factores es diferente. Un ejemplo en agronomía es el estudio de la interacción entre tipos de labores y fechas de siembra (donde el primer factor requiere parcelas grandes en tanto el segundo no), en ciencias de la educación, donde un factor puede requerir una clase completa en tanto el otro tan solo un grupo de alumnos...

El factor que se estudia en parcelas grandes se llama factor principal en tanto el que se estudia en parcelas más pequeñas se denomina factor secundario. El procedimiento de diseño empieza porque las repeticiones se dividen en las parcelas principales y las parcelas principales se dividen a su vez en parcelas secundarias más pequeñas, siempre adjudicando los tratamientos a las parcelas al azar.

El análisis de la varianza se divide en dos partes, una para el estudio de las parcelas principales dentro de las repeticiones (donde en nuestro análisis estarían las localidades) y otra parte las subparcelas dentro de las parcelas principales (donde en nuestro análisis estarían los tratamientos). El análisis correcto requiere precauciones y en general las parcelas principales son estimadas con menor precisión (están dentro de las repeticiones y usualmente tienen menos grados de libertad para la estimación del error) que las secundarias (están dentro de las principales, son más pequeñas y más eficaces en el control de la variación ambiental, que en definitiva es lo que interesa). La interacción entre localidades (la parcela principal) y los tratamientos (la secundaria), se encuentra en el análisis de la subparcela.



El estudio multilocalidades se puede abordar y ha sido frecuentemente estudiado desde el diseño de parcela dividida. El análisis estadístico es mas complejo.

**Tabla 3.5. Análisis de la varianza de un diseño de parcela dividida. Factor principal A, factor secundario B y r repeticiones.**

Factor de variación	Grados de libertad	Cuadrado Medio
Repetición	$r-1$	CM1
Factor A	$a-1$	CM2
Repetición x factor A	$a-1$	CM3
	Error A	
Factor B	$b-1$	CM3
A x B	$(a-1)(b-1)$	CM4
Repetición x Factor B + repetición x AB	$a(r-1)(b-1) =$	CM5
	Error B.	
Total	$r*a*b-1$	

En el análisis combinado de tratamientos y localidades, el factor A serían las localidades y el factor B los tratamientos. La suma de cuadrados de localidades, tratamientos y de la interacción tratamiento x localidad son iguales que en el modelo convencional. La diferencia de este análisis con el convencional, es que aparece en este diseño el efecto de las repeticiones por separado, lo que puede ser de interés en algunas ocasiones. El factor repeticiones x localidad del primero es igual a la suma de repeticiones mas repeticiones x localidades del diseño en parcela dividida. Hay que tener precaución pues los grados de libertad son diferentes. La construcción del modelo en SPSS es similar al descrito para el análisis convencional exceptuando que en el modelo las fuentes de variación son: repeticiones, localidades, repeticiones x localidades, tratamientos y tratamientos x localidades.

Se muestra a continuación el enfoque convencional y en parcela dividida de un experimento en serie de cuatro variedades de cebada con dos repeticiones repetido en tres localidades, considerados efectos fijos (datos de Steel y Torrie, 1981).

**Tabla 3.6. Análisis de la varianza convencional de cuatro variedades de cebada repetido en tres localidades.**

## Tests of Between-Subjects Effects

Dependent Variable: RENDIMIE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	902,083 <sup>a</sup>	14	64,435	12,923	,000
Intercept	16590,042	1	16590,042	3327,251	,000
LOCALIDA	651,583	2	325,792	65,340	,000
REP * LOCALIDA	32,625	3	10,875	2,181	,160
TRATAM	109,125	3	36,375	7,295	,009
TRATAM *LOCALIDA	108,750	6	18,125	3,635	,041
Error	44,875	9	4,986		
Total	17537,000	24			
Corrected Total	946,958	23			

a. R Squared = ,953 (Adjusted R Squared = ,879)

**Tabla 3.7. Análisis de la varianza en parcela dividida de cuatro variedades de cebada repetido en tres localidades.**

## Tests of Between-Subjects Effects

Dependent Variable: RENDIMIE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	902,083 <sup>a</sup>	14	64,435	12,923	,000
Intercept	16590,042	1	16590,042	3327,251	,000
REP	9,375	1	9,375	1,880	,202
LOCALIDA	651,583	2	325,792	65,340	,000
REP * LOCALIDA	23,250	2	11,625	2,331	,153
TRATAM	109,125	3	36,375	7,295	,009
TRATAM *LOCALIDA	108,750	6	18,125	3,635	,041
Error	44,875	9	4,986		
Total	17537,000	24			
Corrected Total	946,958	23			

a. R Squared = ,953 (Adjusted R Squared = ,879)

Se puede ver lo antes indicado. La suma de cuadrados de repeticiones por localidad en el primer caso (32,62) es igual a la suma de repeticiones (9,37) mas repeticiones por localidad (23,25) en el segundo caso. El efecto de la interacción tratamientos x localidad es significativo ( $p=0,041 < 0,05$ ), el de los tratamientos, también ( $p=0,009 < 0,005$ ) y el de las localidades,

también ( $F = 325,79 / 10,87 = 29,97$  ,,  $p=0.000$ ). Recordamos que el valor de F para las localidades es necesario calcularlo a mano.

Antes que emplear los valores medios de cada tratamiento y localidad, el análisis combinado es preferible hacerlo con los valores de las parcelas o unidades experimentales, especialmente en el caso de emplear un diseño de bloques completos al azar (RCBD). Los valores de las medias son preferibles solo en los casos siguientes:

- diseños en bloques incompletos o láttices
- cuando los ensayos en las distintas localidades tienen diseños diferentes
- cuando los ensayos en las distintas localidades tienen distinto número de repeticiones.

Para convertir el resultado de un ANOVA hecho con las medias (ver tabla 2.9) en uno hecho con valores de parcelas, la suma de cuadrados y el cuadrado medio del efecto debe ser multiplicado por un factor r igual a la media armónica del número de las repeticiones en los experimentos (Cochran y Cox, 1957). Para p ensayos sería:

$r = p / \sum (1/r_j)$  ,, donde  $r_j$  es el número de repeticiones del ensayo j. Para un número constante de repeticiones coincide con la media aritmética.

El cuadrado medio (CM) del error combinado se puede estimar a partir de los CM de los errores de los ensayos individuales, ponderandolos con el número de repeticiones. Si hay el mismo número de repeticiones en todos los ensayos coincide con la media aritmética.

### **3.3.- AGRUPAR ENSAYOS DE DISTINTO TAMAÑO.**

Se pueden agrupar ensayos con solo algunos tratamientos comunes, es decir de distinto tamaño. Obviamente se obtiene información solo de los tratamientos comunes y ocurre igual que cuando se analiza un experimento simple, donde faltan tratamientos en algunas repeticiones, los grados de libertad del error disminuyen en una unidad por cada parcela perdida y en consecuencia el cuadrado medio del error aumenta (con lo que se pierde precisión en el experimento) y el valor de F de los tratamientos disminuye (hasta que deja de ser significativo, con lo que puede perder sentido el ensayo). Una opción conocida, es calcular el valor de la parcela perdida en función de los resultados de las otras parcelas, para lo que se han desarrollado fórmulas. Para más información, el lector puede consultar los textos de Snedecor y Cochran, Steel y Torrie, Petersen y otros. Actualmente, esta opción no se considera una buena alternativa para sustituir una parcela o dato perdido. Se estima que para poder agrupar una serie de ensayos hay que tener al menos un 75% de los tratamientos comunes.

En el ejemplo I si en vez de tener ocho tratamientos en cinco localidades, se tuvieran sólo seis tratamientos comunes (con el resto bien variedades nuevas una que interese probar en una localidad determinada o bien variedades perdidas), los grados de libertad del error combinado bajan de 105 (ver tabla 3.2) a 85, el cuadrado medio del error sube de 38,8 a 42 y la F de la interacción variedad por localidad, baja de 3.03 ( $p=0,000$ ) a 2,33 ( $p=0,03$ ), casi deja de ser significativa la interacción.

### 3.4.- AGRUPAR ENSAYOS CON DOS FACTORES

En el caso de que interese agrupar una serie de ensayos con dos factores, A y B, considerados fijos, repetidos r veces en l sitios o localidades, los análisis de la varianza con las pruebas F que hay que hacer en función de que las localidades sean fijas o aleatorias y que el diseño empleado sea factorial o parcela dividida, son los indicados en las tablas 3.8 y 3.9.

**Tabla 3.8. Analisis de la varianza de la agrupacion de l ensayos factoriales, con los factores A y B fijos.**

Factor de variación	Grados de libertad	Cuadro Medio	F1	F2
loc	$l-1$	CM1	CM1/CM2	CM1/CM2
Rep / loc	$l(r-1)$	CM2		
Tratam				
A	$a-1$	CM3	CM3/CM7	CM3/CM7
B	$b-1$	CM4	CM4/CM8	CM4/CM8
A x B	$(a-1)(b-1)$	CM6	CM6/CM9	CM6/CM10
Trat x loc				
A x Loc	$(b-1)(l-1)$	CM7	CM7/CM10	CM7/CM10
B x Loc	$(b-1)(l-1)$	CM8	CM8/CM10	CM8/CM10
A x B x Loc	$(a-1)(b-1)(l-1)$	CM9	CM9/CM10	CM9/CM10
Error	$al(r-1)(l-1)$	CM10		
Total	$ablr-1$			

F1 = Si Localidades aleatorios

F2= Si Localidades fijo

**Tabla 3.9. Analisis de la varianza de la agrupacion de l ensayos en parce-  
la dividida, con los factores A y B fijos, A factor principal y B factor secun-  
dario.**

Factor de variación	Grados de libertad	Cuadro Medio	F1	F2
loc	l-1	CM1	CM1/CM2	CM1/CM2
Rep / loc	l(r-1)	CM2		
A	a-1	CM3	CM3/CM4	CM3/CM5
A X Loc	(a-1)(l-1)	CM4	CM4/CM5	CM4/CM5
Error a	(a-1)(r-1)	CM5		CM4/CM8
B	b-1	CM6	CM6/CM7	CM6/CM10
B x Loc	(b-1)(r-1)	CM7	CM7/CM10	CM7/CM10
A x B	(a-1)(b-1)	CM8	CM8/CM9	CM8/CM10
A x B x Loc	(a-1)(b-1)(l-1)	CM9	CM9/CM10	CM9/CM10
Error b	al(r-1)(l-1)	CM10		
Total	abl-1			

F1 = Si Localidades aleatorios

F2= Si Localidades fijo

## 4. INTERACCION TRATAMIENTO POR AMBIENTE

Como se indicó en el primer capítulo, el principal objetivo de repetir un experimento en ambientes diferentes es conocer si todos los tratamientos se comportan igual en las diferentes condiciones o por el contrario si el resultado depende de ellas. En este caso se dice que existe una interacción tratamiento por ambiente y se habrá comprobado con la significación de este factor de variación en el análisis de la varianza. La interacción se puede definir como el efecto diferencial de un factor (tratamientos o ambientes) sobre el otro y surge cuando una variación ambiental tiene distinto efecto sobre los tratamientos o a la inversa, cuando un tratamiento responde de manera diferente a diversos ambientes.

Los primeros autores que usaron la ecuación lineal (lo que hoy conocemos por modelo general lineal) para describir los efectos conjuntos de los tratamientos y ambientes fueron Sprague y Federer en 1951. Estos autores señalaron que el resultado de varios ensayos repetidos depende de los tratamientos, del ambiente y de la interacción entre tratamiento y ambiente. El modelo extendido a localidades y años fue realizado más tarde por Miller y colaboradores en 1959. Trabajando con variedades, su propósito fue comparar la magnitud relativa de la componente de variedades por localidades y variedades por años con la componente de la varianza del error.

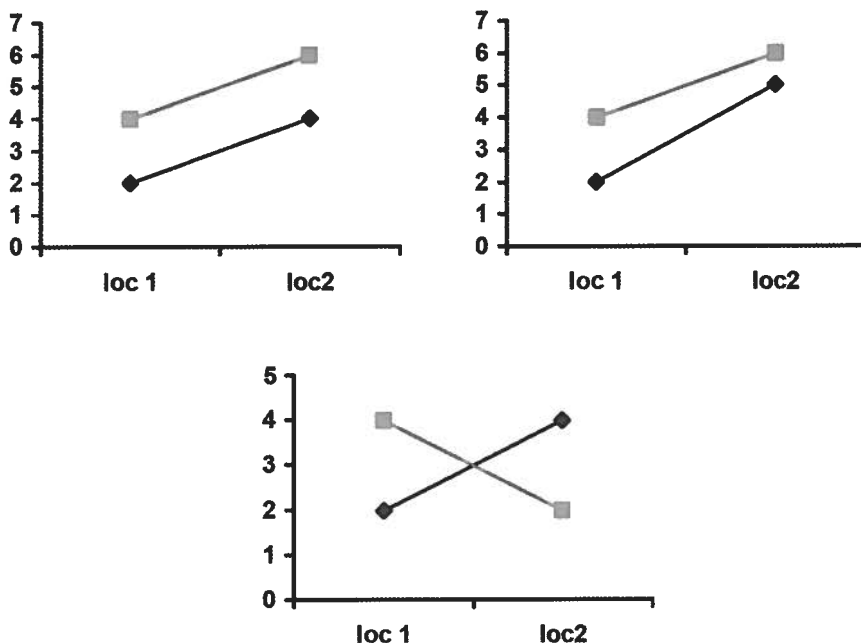
En el modelo completo aditivo (tratamientos + ambientes + tratamientos x ambientes) el error residual es en gran parte debido a la interacción. Hoy existen algunos modelos mas completos que incluyen terminos no lineales para describir la interacción (por ejemplo, modelos AMMI o de componentes principales).

La interacción fue inicialmente un término usado en el contexto de un experimento factorial, para describir los casos donde el efecto combinado de dos variables no es la suma simple de los efectos por separado. A veces los tratamientos se refuerzan entre si o se anulan o bien son efectos más grandes o más pequeños de los que cabría esperar si los efectos individuales se sumaran, como han definido Martin y Luna del Castillo en 1999.

La presencia de la interacción entre tratamientos y localidades se puede identificar en el análisis de la varianza o visualizar gráficamente en una simple figura (ver figura 7). Pueden haber diferencias en la cantidad de la respuesta o en la dirección. La interacción tratamiento por localidad se considera de tipo cuantitativo si las ordenaciones de los tratamientos no cambian al pasar de un ambiente a otro (figura 7,a y b), esta interacción es menos importante que la cualitativa, que es cuando cambian las ordenaciones al cambiar la localidad (fig 7, c).

En el análisis combinado, la interacción puede ser significativa o no significativa. Si la interacción no es significativa, esto indica que el comportamiento relativo de los tratamientos es similar en todas las localidades, las medias en los diferentes ambientes son indicadores adecuados y el análisis combinado termina con el estudio de los efectos principales. En este caso es las interacciones en la variación del error experimental (Tejedor 2003).

**Fig 7. Interpretación gráfica de la interacción. A= No interacción, B= Interacción moderada. (Hay diferencia en la cantidad de respuesta, los tratamientos rojo y azul no cambian el orden) C= Alta interacción (los tratamientos rojo y azul cambian de orden).**



Si la interacción es significativa, quiere decir que el comportamiento relativo entre tratamientos no es el mismo al cambiar de ambiente. Esto hace que la evaluación de los tratamientos mediante sus medias de los diferentes ambientes sea ineficiente. Se hace necesario conocer con más detalle tratamientos y ambientes para elegir el más adecuado para cada ambiente.

Hay que tener la mayor información posible sobre ambientes específicos. Si interesa, la interacción tratamiento por ambiente se puede reducir de dos maneras: estratificando los ambientes o seleccionando los tratamientos que interaccionen menos con el ambiente.

1).- Sitios más homogéneos. Se pueden agrupar los sitios en grupos homogéneos donde la interacción tratamiento x ambiente no sea significativa y a continuación identificar uno o varios tratamientos idóneos para cada grupo. Se puede reducir la interacción tratamiento x localidad dividiendo las regiones en subregiones. Se ha encontrado que la disminución en la varianza de la interacción es proporcional con el número de subgrupos que se hacen.

2).- Tratamientos más homogéneos. Se puede tratar de identificar uno o mas tratamientos cuya contribución a la interacción sea alta. En el gráfico de la interacción (tipo figura 6) se ven tratamientos que fluctúan mucho de una localidad a otra. Estos son los que contribuyen mas a la interacción (como se verá más adelante, son los que tienen un rango estrecho de adaptabilidad) y los menos apropiados para una recomendación general. A veces es precisamente esto lo que interesa, encontrar tratamientos adecuados a ambientes determinados.

Se han desarrollado numerosos modelos estadísticos para evaluar la interacción. Los trabajos donde se han estudiado en mayor profundidad ha sido entre variedades y ambientes por ser necesarios en mejora vegetal, pero como se ha señalado, se pueden usar igualmente para estudiar la interacción ente cualquier clase de tratamientos y diferentes ambientes (Crossa, 1990; Crossa y colaboradores, 2001).

Como han definido recientemente Cubero y Flores (1994) en relación con la interacción entre variedades (genotipos) y ambientes, la interacción entre ambos puede ser interpretada solo en parte como una respuesta diferencial a diferentes ambientes (o variaciones ambientales como sequía, enfermedades, ...), pero generalmente tan solo una pequeña parte de la interacción variedad (o tratamiento) por localidad puede ser atribuida a características ambientales conocidas, la mayor parte es una cantidad inexplicable en todos los análisis estadísticos y conocida como "error experimental".

La interacción variedad y ambiente se relaciona con la idea de resultados homogéneos o estables. Los investigadores en genética coinciden en gene-



ral en que es importante obtener variedades **estables** (lo que podemos hacer extensivo a cualquier tipo de tratamientos), aunque no hay acuerdo en definir lo que es estabilidad. Hay quien prefiere el término **adaptación** cuando se trata de una variación en el espacio y el término estabilidad cuando se trata de poca variación durante varios años. La capacidad de algunos genotipos para comportarse optimamente en un amplio intervalo de ambientes es lo que Finlay y Wilkinson (1963) definieron como adaptabilidad general y la capacidad de comportarse optimamente en algunos ambientes específicos, como estabilidad específica. En general se entiende por un tratamiento estable que muestra un rendimiento constante en diferentes ambientes, es decir los que muestran una varianza mínima frente a estos.

Los análisis anteriores se han sugerido en las siguientes condiciones (Annicchiarico, 2002)

- Un análisis de adaptación está realmente justificado si la componente de la varianza de la interacción tratamientos x localidades es significativa y moderadamente más grande que la componente de la varianza de los tratamientos (superior en un 30 – 35 %).
- La mejora para estabilidad de rendimientos está justificada cuando la varianza de la interacción tratamientos x localidades es mucho más grande que la varianza de los tratamientos (del orden de 150 – 200 %).
- La selección para un rendimiento estable requiere un amplio número de ambientes de selección.

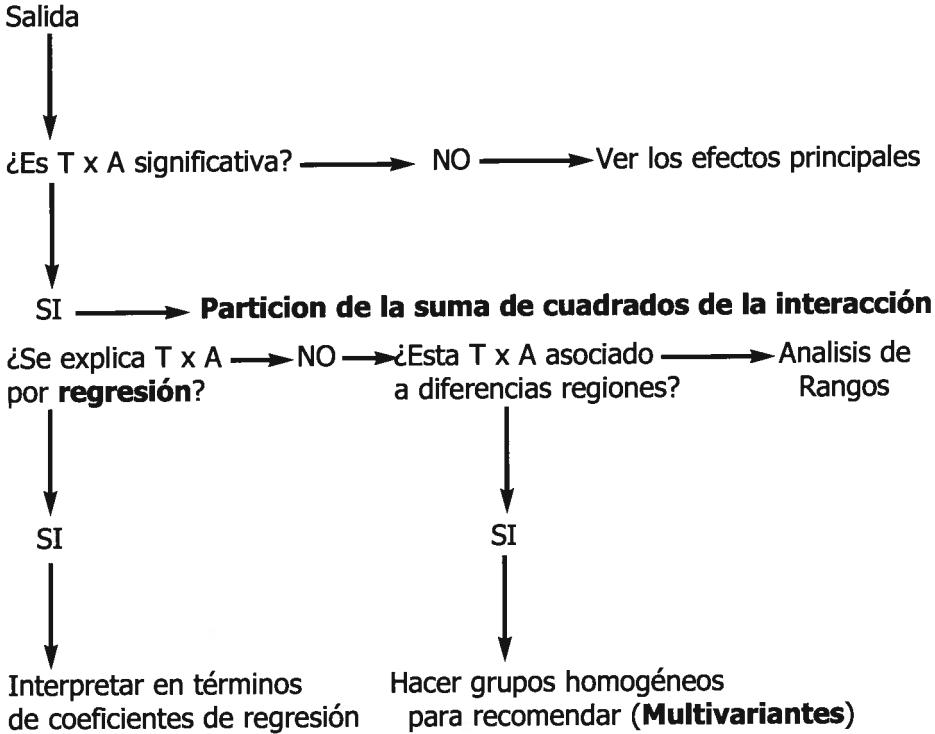
Con los datos del ejemplo 1, la variedad Pleno con una varianza de 33,84 es la más estable, en tanto las variedades Armida y Emerita con una varianza de 69,27 serían las menos estables.

Para estimar la estabilidad de las variedades se han propuesto numerosos métodos, una descripción detallada se puede encontrar en la publicación de estos autores (Cubero y Flores, 1994). Los métodos pueden ser univariantes o multivariantes, entre ellos, en resumen:

1. Métodos paramétricos univariantes
  - Partición de la suma de cuadrados de la interacción.
  - Análisis de regresión
2. Métodos no paramétricos univariantes
  - Posición y desviación (Ketata y colaboradores, 1989)
  - Estratificado de ordenaciones (Fox y colaboradores, 1990)
3. Métodos multivariantes
  - Métodos de ordenación o representación (análisis de componentes principales, de coordenadas principales y análisis factorial, análisis AMMI, análisis de correspondencia, análisis canónico)

- Métodos de clasificación (análisis de grupos y análisis discriminante).

En esta publicación se analizan los métodos univariantes, señalados en los apartados 1 y 2, siguiendo el diagrama de flujo señalado a continuación (donde TxA es la interacción tratamiento por ambiente, parcialmente adaptado de Kempton y Fox, 1997).



#### 4.1.- Particion de la suma de cuadrados de la interacción

##### Tratamiento x Localidad

La suma de cuadrados de la interacción puede ser descompuesta en partes que ayuden a describir el comportamiento de los diferentes tratamientos. Se han propuesto diferentes métodos para caracterizar la contribución de los tratamientos a la suma de cuadrados de la interacción.

Salmon (1951) propuso hacer un análisis de la varianza con todos los tratamientos tomados de dos en dos para estimar la varianza de la interacción de todas las combinaciones posibles tomadas de dos en dos. La media de la

varianza obtenida para cada tratamiento se emplea como un índice de la contribución de ese tratamiento a la interacción total tratamiento x ambiente (TXA) y permite, por lo tanto, conocer mejor el comportamiento de los tratamientos.

Un método útil y más detallado para tener más información sobre la naturaleza de la interacción es partir la suma de cuadrados de esta en un conjunto de contrastes ortogonales. Un contraste es un conjunto de números, cuya suma es cero, que indican una relación entre los tratamientos. Si se ha

identificado el tratamiento que más fluctúa de una localidad a otra o que es más relevante, se puede analizar la significación de este respecto a los demás y la interacción de este y la del resto con la localidad. Un ejemplo descriptivo (de Gómez y Gómez, 1986) es el siguiente:

Se diseña experimento con 7 tratamientos y 3 repeticiones en bloques completos al azar, que se repite en 9 localidades diferentes, la tabla del análisis de la varianza combinado es la siguiente (tabla 4.1):

**Tabla 4.1. Análisis de la varianza combinado de 7 tratamientos con 3 réplicas repetido en 9 localidades.**

Factor de variación	Grados de Libertad	Cuadro Medio	F
Localidades	8		
Rep dentro de loc	18		
Tratamientos	6		
Tratam x Local.	48		
Error combinado	108		
Total	188		

Supongamos que la interacción tratamiento por localidad resulta significativa y se observa que un tratamiento, por ejemplo el tratamiento 2, fluctúa mucho de una localidad a otra. Se prueba descomponer la suma de cuadrados de la interacción con un contraste ortogonal, tratamiento 2 contra el resto de tratamientos. Lo llamaremos E1 y sería de la forma (-1, 6, -1,-1,-1,-1,-1). E2 sería el resto, comparación entre los demás tratamientos. La nueva tabla de la la varianza que resulta, donde se incluyen los contrastes sería la siguiente (tabla 4.2).

**Tabla 4.2. Análisis de la varianza combinado de 7 tratamientos con 3 réplicas repetido en 9 localidades con dos contrastes E1 (tratam 2 vs otros tratam) y E2 (entre otros tratamientos).**

Factor de variación	Grados de Libertad	Cuadro Medio	F
Localidades	8		
Rep dentro de loc	18		
Tratamientos	6		
E1	1		
E2	5		
Tratam x Local.	48		
E1 x L	8		
E2 x L	40		
Error combinado	108		
Rep(Loc) x E1	18		
Rep(Loc) x E2	90		
Total 188	188		

Si E1 (tratam 2 contra el resto) resulta significativo y E1 x L (contraste por localidad) no significativo en la prueba F, hemos podido conocer que el tratamiento 2 es diferente estadísticamente del resto, pero que su interacción con el ambiente no lo es.

Los aspectos prácticos de esta técnica se detallan con el siguiente ejemplo:

#### EJEMPLO 2.

*Se prueban en ratones tres tratamientos (A1,A2 A3) con seis repeticiones en un diseño en bloques completos al azar en dos ambientes (B1,B2). A1 es un tratamiento natural y A2, A3 son tratamientos sintéticos a dos dosis diferentes. Se evalúa el peso. (Adaptado de Camussi y colaboradores, 1995).*

El análisis de la varianza combinado del factor A (tratamientos) y del factor B (ambientes), tomado como un factorial simple para simplificar el anova se muestra en la tabla 30. Al ser la interacción significativa ( $F=12,2$ ,  $p=0,000$ ) se prueba descomponer los tratamientos en un contraste. El contraste lógico a probar es  $(2 -1 -1)$ , es decir tratamiento natural frente a los tratamientos sintéticos. El procedimiento en SPSS es el indicado en el apartado de análisis convencional.

Una vez señalado el modelo (GLM) y las opciones, antes de pulsar el OK final, hay que pulsar el comando " Paste". En la rutina de salida:

unianova

- 1.- rendimie by bloque tratam
- 2.- /method = sstype(3)
- 3.- /intercept = include
- 4.- /criteria = alpha(.05)
- 5.- /DESIGN = bloque tratam

Se quitan las 3 sentencias 2, 3, 4 y se escriben las nuevas 2 y 3 :

- 2.- /LMATRIX "comp de tratam"
- 3.- tratam 0 -1 -1 -1 1 1 1 ;

A continuación se le da a run. "comp de tratam" significa comparación de tratamientos y la serie de números es el contraste que se quiere probar. En este caso la rutina de salida sería:

UNIANOVA

```
peso BY repetici tratam
/LMATRIX "comp de tratam"
tratam 2 2 -1 -1 -1 -1;
/DESIGN = repetici tratam .
```

El archivo de datos tendría las variables repetición (repetici), A, B, peso (como dependiente) y una nueva que se ha creado a efectos de hacer el contraste llamada tratam (tratamientos, combinación de A y B, con los valores A1, A1, A2, A2, A3, A3). El contraste tratamiento natural frente a sintéticos es 2 2 -1 -1 -1 -1 y lo llamamos L1. Al resultar L1 significativo, hay diferencia entre el tratamiento natural y los sintéticos. Los contrastes hay que hacerlos de uno en uno (la salida se indica en la tabla 4.4) e incorporarlos a la tabla del ANOVA (ver tabla 4.3).

**Tabla 4.3. Análisis de la varianza del ejemplo 2. Tres tratamientos (factor A) con seis repeticiones y dos ambientes (factor B). Análisis factorial.**

Tests of Between-Subjects Effects  
 Dependent Variable: PESO

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	849,000	5	169,800	11,320	,000
Intercept	15129,000	1	15129,000	1008,600	,000
A	402,000	2	201,000	13,400	,000
L1	220,500	1	220,500	19,433	,000
B	81,000	1	81,000	5,400	,027
A * B	366,000	2	183,000	12,200	,000
L1*B	112,500	1	112,500	7,5	,054
Error	450,000	30	15,000		
Total	16428,000	36			
Corrected Total	1299,000	35			

a R Squared = ,654 (Adjusted R Squared = ,596)

**Tabla 4.4. Resultados en SPSS del contraste propuesto. Contrast Results (K Matrix)**

		Dependent Variable
Contrast		PESO
L1	Contrast Estimate	-21,000
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-21,000
	Std. Error	4,764
	Sig.	,000
	95% Confidence Interval for Difference	Lower Bound -30,811
		Upper Bound -11,189

a Based on the user-specified contrast coefficients (L') matrix: comp de tratam  
 Test Results

Dependent Variable: PESO

Source	Sum of Squares	df	Mean Square	F	Sig.
Contrast	220,500	1	220,500	19,433	,000
Error	283,667	25	11,347		

Si los tratamientos son cuantitativos es muy interesante partir la suma de cuadrados de la interacción en un conjunto de contrastes polinómicos. Los polinomios ortogonales son contrastes en forma de ecuaciones lineales, cuadráticas, cúbicas... que indican las relaciones entre la variable dependiente y la independiente (tratamientos). En la interacción se obtienen las parciales  $L \times t$  lineal,  $L \times t$  cuadrático ..., donde L es la localidad y t el tratamiento cuantitativo.

Supuesto un experimento con cinco dosis crecientes con tres repeticiones en dos ambientes diferentes (por ejemplo, frío y calor), el análisis de la varianza combinado es el mostrado en la tabla 4.5.

La ejecución en SPSS es más simple que el anterior. Basta con señalar en GLM el botón "contrast", en "Factors" marcar la variable cuantitativa, en "Change contrast" poner el desplegable "polynomial" y dar a "change". La salida de SPSS es una tabla con información de cada contraste, lineal, cuadrático, cúbico y de orden 4. Entre la información se incluye la significación y la estimación. Una vez elegido el contraste más apropiado, la ecuación del mismo se obtiene por regresión.

**Tabla 4.5. Análisis de la varianza combinado de 5 tratamientos (dosis,D) con 3 repeticiones en 2 ambientes. Contrastes polinómicos.**

Factor de variación	Grados de Libertad	Cuadro Medio	F
Ambientes	1		
Rep dentro de Ambientes	4		
Dosis	4		
D Lineal	1		
D cuadrático	1		
D residual	2		
Amb x Dosis	4		
Amb x D lineal	1		
Amb x D cuadr.	1		
Amb x D resid.	2		
Error combinado	16		
Total	29		

#### **4.2.- Analisis de regresión lineal. Metodo de Eberhart y Russel (análisis de estabilidad)**

En 1938 Yates y Cochran emplearon la media de todos las diferentes variedades en ensayo en una localidad como "índice ambiental" de esa localidad. El comportamiento de una determinada variedad se estudiaba por medio de la regresión lineal de las medias de la variedad en cada ambiente en relación con esos índices ambientales y demostraron que la regresión lineal de los rendimientos de todas las variedades es una parte importante de la interacción en un grupo de ensayos. Estos autores fueron los pioneros en el uso de técnicas de regresión para identificar los tratamientos mas "estables" (y como ocurre a menudo, el método fue largamente ignorado por más de 20 años). En este método, las pendientes indicarían mayor o menor estabilidad de cada variedad y se podrían comparar entre ellas con un test de  $\chi^2$ .

La regresión lineal proporciona un buen modelo de una dimensión para analizar la estabilidad de los tratamientos. La regresión simple de la media de un tratamiento sobre la media de los ambientes proporciona una ecuación por cada tratamiento. En la ecuación, la pendiente da una medida de la estabilidad del tratamiento. Este método de regresión exige que todos los tratamientos esten en todos los ambientes.

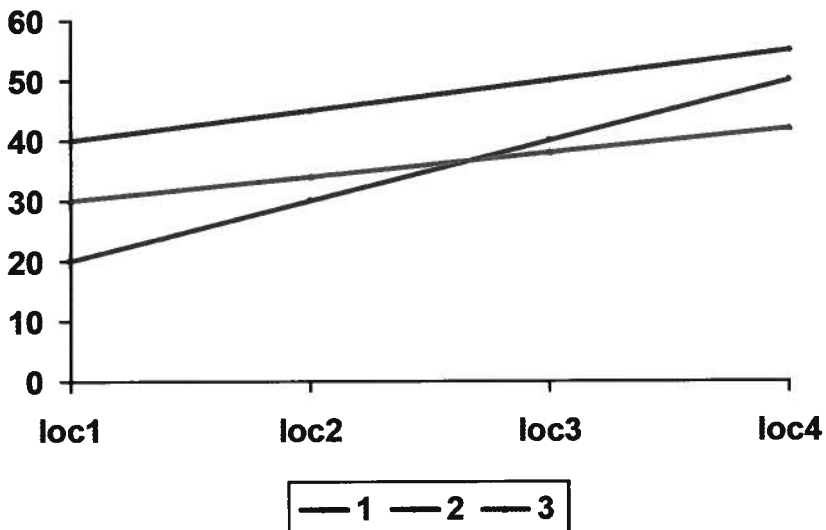


Años más tarde, Finlay y Wilkinson (en 1963) usaron este método para analizar la adaptación de un conjunto de variedades de cebada ensayadas en varias localidades. En definitiva se trataba de analizar la pendiente de la recta de regresión del rendimiento de cada variedad sobre el rendimiento medio de la localidad. Para cada tratamiento la ecuación tiene la forma :

$R.V = a + b*(R.M.L)$  ,, donde R.V = Rendimiento Variedad y R.M.L = Rendimiento medio de la localidad.

Existen muchas alternativas al rendimiento medio de la localidad, en definitiva para definir los diferentes ambientes, como variables climáticas (lluvia media, radiación solar...), variables demográficas, económicas.... En el modelo más conocido, se obtendrían respuestas como las indicadas en la figura 8, donde la línea 1 (azul) tendría una pendiente ligeramente inferior a 1 y tendría una adaptación horizontal, la línea 2 (verde) con una pendiente elevada, por ejemplo de valor 2, tendría un adaptación específica para ambientes más productivos (sería muy sensible al cambio de ambientes) y la línea 3 (roja) con una pendiente muy baja, próxima a 0, sería un tratamiento estable, recomendado para ambientes poco productivos.

**Figura 8. Respuestas de tres tratamientos a diferentes ambientes. (loc1-loc2-loc3-loc4) rendimiento medio de la localidades. (Loc = Localidades)**



Al medirse la estabilidad por el coeficiente de regresión, los tratamientos con coeficientes de regresión mayor que uno ( $b > 1$ ) indicarían un comportamiento mejor a condiciones favorables con respecto al rendimiento medio del sitio, mientras que pendientes próximas a cero ( $b = 0$ ) sugerirían ninguna variabilidad del tratamiento como respuesta al cambio del ambiente. Con los datos del ejemplo I (ver tabla 4.6), la variedad Emérita es adecuada para su uso en ambientes de alta producción (tipo línea verde), en tanto la variedad Pleno es la que responde menos al cambio de ambiente (tipo línea roja) aunque ambas no son significativas.

**Tabla 4.6. Valores de regresión de las diferentes variedades sobre la media de cada localidad. (Análisis realizado por el programa Irristat)**

tratamiento	media	b	E.E.	R2	Nº en la figura 9
Alto	13,79	0,37	0.46	0,86**	1
Arista	14,64	0,43	0.44	0,81**	2
Armida	14,97	0,33	0.67	0,7**	3
Emerita	14,16	1,68*	0.14	0,37NS	4
Hilma	13,24	0,16	0.57	0,23	5
Marina	14,24	0,62	0.02	0,99**	6
Nita	13,17	-1,51*	0.16	0,29NS	7
Pleno	13,34	-0,04	0.45	0,02NS	8

Notas:

b es la pendiente de la regresión de la media de cada variedad sobre las localidades. (\*)

Indica que la pendiente es significativamente diferente de la pendiente 1.

EE es el error estándar

Tres años después, Eberhart y Russell (1966), continuaron con el desarrollo de esta técnica. Consideraron que un estudio de las pendientes no es suficiente para determinar la estabilidad, debería entrar en el estudio también la media de cada variedad. Así las variedades con pendientes cercanas a 1 y alta media se considerarían bien adaptadas a todos los ambientes.

Consecuentemente, un tratamiento se consideraría más adaptado o estable si tiene un alto rendimiento y una baja fluctuación de rendimientos en los diferentes ambientes. Estos autores propusieron partir la interacción tratamiento por ambiente en dos partes; una sería la respuesta de los diferentes tratamientos a los diferentes índices ambientales (heterogeneidad de los coeficientes de regresión o regresión de tratamientos por ambientes) y otra

parte las desviaciones de las regresiones. El análisis de la varianza propuesto por Eberhart y Russell es:

**Tabla 4.7. Análisis de la varianza y regresión conjunta de t tratamientos, repetidos r veces en L ambientes (Método de Eberhart y Russell).**

Factor de variación	Grados de Libertad	Cuadrado Medio	F
Ambientes	L-1	CM1	CM1/CM2
Rep dentro de Ambientes	L*(r-1)	CM2	
Tratamientos	t-1	CM3	CM3/CM4
Tratam x Ambientes	(t-1)*(L-1)	CM4	CM4/CM7
Heterogeneidad	t-1	CM5	CM5/CM6
Desviación regresión	(t-1)*(L-2)	CM6	CM6/CM7
Error combinado	L*(t-1)*(r-1)	CM7	
Total	L*t*r-1		

**Tabla 4.8. Análisis de la varianza de la regresión de las ocho variedades en seis ambientes (datos de valores medios de las localidades del ejemplo 1 analizados con el programa Irristat).**

Factor de variación	Grados de Libertad	Suma de cuadrados	Cuadrado Medio	F
Localidades	7	18.94	2.70	
Tratamientos	5	12.20	2.44	
Tratam x localidades	35	20.91	0.59	
Regres trat x loc	7	12.10	1.72	5.49
Desviación regresión	28	8.81	0.31	(p=0.00)
Total	47	52,06		

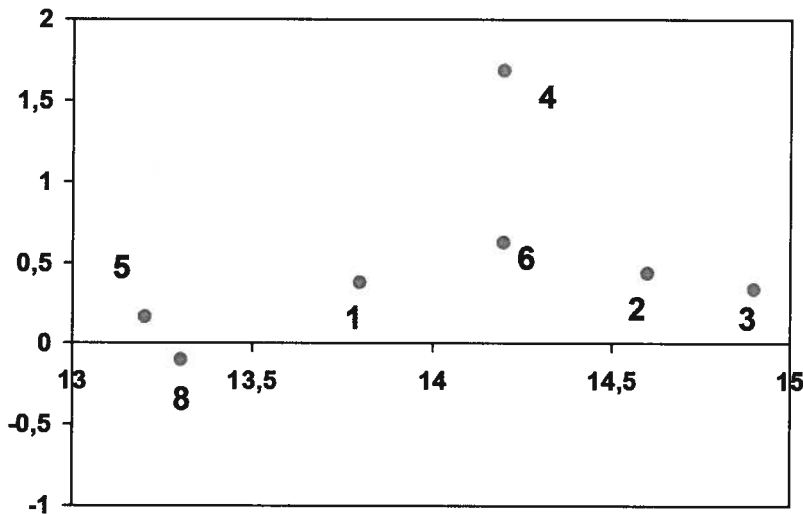
Para calcular las sumas de cuadrados existen programas (p.e. Irristat), subrutinas (especialmente para el programa SAS) o se pueden hacer a mano. Los detalles de cálculo se pueden consultar en la publicación de Cubero y Flores (1994). En la tabla 4.8 se muestra la salida del análisis de la varianza de los datos medios de las ocho variedades en seis localidades

con la significación de la regresión de los tratamientos por las localidades. Se puede apreciar como este método es capaz de explicar en este caso el 57.8 % de la suma de cuadrados de la interacción ( $12.10/20.91 \times 100$ )

Los resultados de rendimiento de cada tratamiento junto con su pendiente (coeficiente de regresión) se pueden representar en forma gráfica en lo que sería un diagrama de estabilidad (ver figura 9).

Pinthus en 1973 propuso emplear el coeficiente de determinación ( $R^2$ ) obtenido en el análisis de regresión como estimador de la estabilidad de una variedad. Según la tabla 4.6, las variedades más estables serían Marina y Alto. Esta propuesta tiene la ventaja de que es independiente de las unidades de medida.

**Figura 9. Diagrama de estabilidad de las variedades del ejemplo I. Rendimiento y pendiente de regresión. Las variedades 4 y 6 serían las más adaptadas a todos los ambientes.**



(nota: Se ha incluido la variedad 8 por su proximidad a cero. No se ha considerado la variedad 7, porque una pendiente negativa no tiene sentido. No es lógico que en un ambiente más productivo una variedad produzca menos. Es el efecto de un ejemplo)

Hoy existen métodos más completos (y más complejos) para estudiar la interacción tratamientos x ambiente basados también en la regresión, como la regresión factorial (FR, Van Eeuwijk y colaboradores, 1996) o la regresión PLS (Partial Least Squares, Aastveit y Martens, 1986).

### **4.3. Algunos métodos no paramétricos y otros métodos sencillos.**

Existen algunos métodos simples que en general se basan en el orden que obtienen los tratamientos en los diferentes experimentos. Son métodos que tienen algunas ventajas además de su sencillez, como son: no hay que tener en cuenta las suposiciones sobre las distribuciones (normalidad y homogeneidad de varianzas), son menos sensibles a los errores de medida y a la falta de observaciones y no está muy afectado por los resultados de ambientes extremos. Como inconveniente, estas técnicas son más descriptivas que predictivas por lo que no proporcionan mucha más información que la que da el análisis de la varianza. Algunas de estas técnicas son las siguientes:

#### *3.1 método de la posición y desviación (Ketata y colaboradores, 1989)*

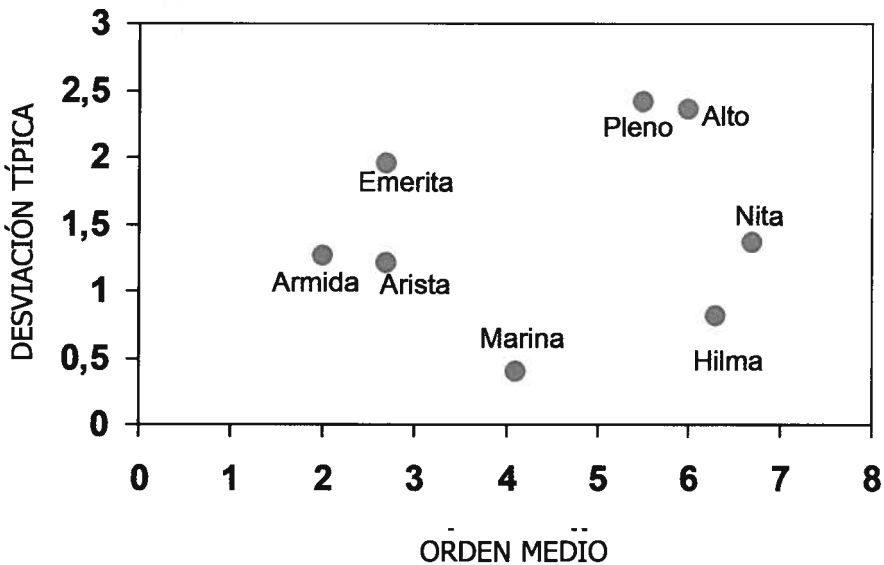
Estos autores propusieron usar la posición media y la desviación típica de cada variedad para seleccionar aquellas que tienen un rendimiento alto y una desviación típica pequeña (más estables). Los autores afirmaron que para un elevado número de tratamientos ambas variables no están correlacionadas, por lo que es posible seleccionar independientemente para alto rendimiento y baja variabilidad ambiental, lo que denominaron un comportamiento consistente. Con los datos del ejemplo I, el orden medio y desviación de las ocho variedades se muestra en la tabla 4.9. Se puede ver que se producen cambios en las ordenaciones de las variedades entre las distintas localidades lo que indica una interacción sustancial entre las variedades ensayadas y las localidades donde se han realizado los experimentos.

Las mejores variedades, las que ocupan de media los mejores puestos son Armida, Arista y Emérita. Su estabilidad es parecida. De las variedades que ocupan los últimos puestos, Hilma y Nita, la primera es la más estable (lo que es el peor resultado posible, ser casi siempre la última). Los resultados se pueden representar gráficamente lo que facilita su comprensión (ver figura 10). Se puede ver como entre las variedades Arista y Emérita que ocupan un orden parecido, Arista es más estable en los rendimientos.

**Tabla 4.9. Ordenaciones de las ocho variedades en cada uno de los ensayos realizados en relación con el rendimiento (posición en el ensayo). Orden medio y desviación típica.**

tratamiento	Loc1	Loc2	Loc3	Loc4	Loc5	Loc6	Orden medio	Desviación típica
Alto	7	8	7	8	3	3	6,0	2,36
Arista	2	3	5	2	2	2	2,7	1,21
Armida	3	4	2	1	1	1	2,0	1,26
Emerita	1	1	1	3	5	5	2,7	1,96
Hilma	6	6	6	6	8	6	6,3	0,81
Marina	4	5	4	4	4	4	4,1	0,40
Nita	5	7	8	5	7	8	6,7	1,36
Pleno	8	2	3	7	6	7	5,5	2,42

**Figura 10. Orden medio en los ensayos y desviación típica de las ocho variedades del ejemplo I.**



### *3.2. Método de ordenación estratificado. Analisis simple de Rangos (Fox y colaboradores, 1997)*

Si no hay cambio en la ordenación de los tratamientos en los distintos ambientes, los tratamientos con medias superiores pueden ser recomendados para todos los ambientes. La ordenación estratificada mide la proporción de ambientes donde un tratamiento esta situado en la tercera parte de las entradas mas altas, medias o bajas. Un tratamiento situado en el primer tercio superior en todos los ambientes estaría bien adaptado. La idea de los autores es que los cambios de ordenación en los tratamientos al cambiar el ambiente, indica la presencia de interacción tratamientos x ambiente y la falta de estabilidad para el carácter estudiado.

La semejanza de dos ordenaciones (para comparar dos grupos de ambientes diferentes) se puede medir por el coeficiente de correlación de rangos de Spearman.

Con este método, bastante restrictivo, prácticamente sólo las variedades Arista y Armida se podrían recomendar para emplear en toda la región (supuesto que las seis localidades representaran una determinada región).

### **3.3 Rendimiento relativo (RR, Yau y Hamblin, 1994)**

Parecido al orden se puede hacer con el rendimiento relativo. Es un método sencillo. El rendimiento relativo (RR) de un tratamiento es igual a  $R_{ij} / R_{.j} \times 100$  ,, donde  $R_{ij}$  es el rendimiento del tratamiento  $i$  en la localidad  $j$  y  $R_{.j}$  es el rendimiento medio de la localidad  $j$ . La desviación estándar de los RR de cada tratamiento es la medida de la estabilidad (como siempre, los tratamientos más estables son los que poseen menor desviación estandar). Este método tiene la ventaja de considerar igual a todos los ambientes, es decir no favorece a los ambientes más productivos.

### **3.4 Coeficiente de variación (Francis y Kannenberg, 1978).**

Un método simple para medir la estabilidad es el propuesto por estos autores. Se basa en representar gráficamente el coeficiente de variación (CV, desviación estándar dividida por la media) en el eje de abcisas y el rendimiento medio en el de ordenadas. La media de ambas variables da lugar a cuatro cuadrantes, con el origen en el CV y el rendimiento medio, clasificándose los tratamientos en alto rendimiento y baja variación, alto rendimiento y alta variación, bajo rendimiento y baja variación y bajo rendimiento y alta variación. Es un método más interesante para agrupar tratamientos que para caracterizar uno determinado individualmente (Cubero y Flores,1994).

Numerosos autores que han desarrollado métodos útiles para relacionar tratamientos y ambientes diferentes coinciden en que distintos métodos pueden dar resultados diferentes. Como señalan los últimos autores mencionados, esto puede ser consecuencia de la gran complejidad de los tratamientos en su interacción con el ambiente. Una recomendación es emplear siempre más de un método y recomendar en función de los resultados y del conocimiento que se tenga de los diferentes tratamientos.

En relación con los métodos multivariantes, existen técnicas de ordenación como el análisis de componentes principales (ACP, Kroonenberg y Basford, 1989; Cubero y Flores, 1994 incluyendo subrutinas para SAS), el análisis factorial, análisis de correspondencias o canónico, análisis AMMI (Additive main Effect and Multiplicative Interaction (Kempton, 1984; Gausch, 1988) en este método se parte la interacción tratamiento x ambiente en componentes multiplicativas y se basa en el análisis de la varianza y el ACP.

Se describe la interacción en más de una dimensión y suele ofrecer mejores oportunidades de interpretarla que el análisis de la varianza o la regresión sobre la media) y técnicas de clasificación como el análisis discriminante o el análisis de grupos (Cluster analysis, Sokal y Sneath, 1963). Este último método empleado habitualmente en taxonomía, se suele emplear frecuentemente para medir el grado de similaridad entre tratamientos. Se pueden encontrar referencias a estos métodos en Fox y colaboradores (1997) o Vargas y colaboradores (2001).

## **GRUPOS DE LOCALIDADES O SUBREGIONES.**

Si la interacción no se explica por regresión, un siguiente paso es hacer grupos homogéneos que tengan una misma recomendación dentro de ellos. Cuando un grupo de localidades es similar en términos de respuesta de los tratamientos, las localidades similares pueden agruparse. Cada grupo identifica un área uniforme donde la interacción tratamiento x localidad no es significativa. Estos grupos reciben también el nombre de subregiones, subzonas, subáreas o macroambientes (Annicchiarico, 2002). El factor subregión o subgrupo es siempre fijo.

Para analizar su efecto, el procedimiento básico es similar a la partición de la suma de cuadrados ya vista, con partición de efectos entre subregiones (sr) y entre localidades dentro de las subregiones  $l(sr)$ . Las fuentes de variación serían tratamientos (con  $t-1$  grados de libertad), localidades  $(l-1)$ , subregiones  $(sr-1)$ , localidades (subregiones) con  $(l-sr)$  grados de libertad, interacción entre tratamientos y subregiones  $(t-1)(sr-1)$ , interacción entre tratamientos y localidades(subregiones) con  $(t-1)(l-sr-1)$  grados de libertad como las interacciones de mayor interés.



Algunos otros métodos para determinar subgrupos (agrupar las localidades) son los siguientes:

- Regresión lineal conjunta (JRL, De Lacy y colaboradores, 1996)
- AMMI (Kempton, 1984, Zobel y colaboradores, 1988)
- Regresión factorial (Kang y Gorman, 1989, Piepho, 1998)
- Análisis de grupos (Pattern analysis, De Lacy y colaboradores 1996; Lin y Butler, 1988)

#### 4.4. Análisis multivariante. Método AMMI

De todos los métodos multivariantes, el método AMMI suele ser el más útil por dos razones: en general es el que explica mayor porcentaje de la suma de cuadrados y la representación gráfica que proporciona (biplot) donde se sitúan tratamientos y localidades es fácil de interpretar.

Esta técnica, relativamente reciente, permite identificar los tratamientos más adaptados a ciertos ambientes mejor que la regresión. El modelo se basa en la incorporación simultánea de efectos principales aditivos (tratamientos y localidades) y la partición de la interacción tratamiento x localidad mediante la extracción de componentes principales (Zobel y colaboradores, 1988). Por ello el procedimiento se denomina AMMI (Additive Main effects and Multiplicative

Interaction). Se ha mostrado que el análisis AMMI es más efectivo con la interacción que el análisis convencional de dos vías porque contiene menos grados de libertad y una menor suma de cuadrados (Zobel y colaboradores, 1988). Gauch (1992) señaló que un análisis AMMI con 2 repeticiones es igual de seguro que un análisis convencional con de 3 a 6 repeticiones.

Lin y colaboradores (1986) comentaron que dado que la respuesta de los tratamientos suele ser más multivariante que univariante, estos últimos métodos suelen ser preferibles. En este modelo los efectos principales se analizan por un análisis de la varianza y la interacción se analiza por componentes principales. El número óptimo de ejes de interacción de componentes principales (IPCA) para ser empleados en el modelo en orden de obtener la estimación más segura, de forma predictiva o postdistiva de la variable respuesta. Para el sistema postdistivo, el más empleado, se emplea un test F para evaluar la significación de cada IPCA.

La ecuación de este modelo es la siguiente:

$$Y_{ge} = \mu + \alpha_g + \beta_e + \sum_{n=1}^N \lambda_n * \varphi_{gn} * \delta_{en} + \theta_{ge} + \varepsilon_{ger}$$

Donde :  $Y_{ge}$  es el resultado del tratamiento  $g$  en el ambiente  $e$ ,  $\mu$  es la media general,  $\alpha_g$  es la media de los tratamientos,  $\beta_e$  es la media de los ambientes,  $\lambda_n$  es el valor principal de la componente principal  $n$ ,  $\varphi_{gn}$  y  $\delta_{en}$  son los valores de tratamiento y ambiente correspondientes a la componente principal  $n$ ,  $\theta_{ge}$  es el residual y  $\varepsilon_{ger}$  es el error aleatorio, correspondiente a la diferencia entre  $Y_{ge}$  y una simple observación para una repetición  $r$ . La parte aditiva del modelo AMMI ( $\mu, \alpha_g, \beta_e$ ) se estima por un análisis de la varianza y la multiplicativa ( $\lambda_n * \varphi_{gn} * \delta_{en}$ ) por un análisis de componentes principales. Si la mayor parte de la suma de cuadrados de la interacción se explica en el eje de la primera componente, se emplea un modelo de AMMI reducido solo con este primer eje. Este es el caso que se seguirá a continuación.

El AMMI tiene asociado una representación gráfica bidimensional que es una herramienta muy útil para interpretar la interacción. Cada tratamiento y cada localidad está representado por su rendimiento en el eje  $x$  y el peso correspondiente de la primera componente principal del análisis de componentes principales (IPCA1) en el eje  $y$  (ver figura 11). La interpretación es sencilla. Los tratamientos con valor próximo 0 de PCA tienen una adaptación horizontal en tanto los que tengan valores de PCA elevados tienen una adaptación específica a determinados ambientes. Los que tienen PCA positivos están adaptados a localidades con PCA positivos y los que tengan PCA negativos a las localidades con PCA negativos. A medida que el segmento entre el tratamiento y la localidad sea menor más adecuado será para esta localidad.

Aplicación del método AMMI a los datos del ensayo I con el programa Irristat

**Tabla 4.10 Analisis de la varianza**

Factor de variación	Grados de Libertad	Suma de cuadrados	Cuadrado Medio	F (probab.)
Localidades	5	12.20	2.44	
Tratamientos	7	18.94	2.70	
Tratam x localidades	35	20.91	0.59	
AMMI componente1	11	15.90	1.44	6.91 (0.000)
AMMI componente 2	9	3.36	0.37	3.93 (0.018)
AMMI componente 3	7	1.01	0.14	1.81 (0.210)
AMMI componente 4	5	0.55	0.11	4.11 (0.137)
Residual Trat x Loc	3	0.08		
Total	47	52,06		

La tabla del analisis de la varianza del AMMI, tabla 4.10, muestra que el 23.4% de la suma de cuadrados total es atribuible al ambiente, el 36,3% a los tratamientos y la mayor parte, el 40,1%, a la interacción entre tratamientos y ambientes. La magnitud de la suma de cuadrados de la interacción es mayor que la de los tratamientos y localidades lo que indica que hay una sustancial variación en la respuesta de los tratamientos a las localidades.

Los resultados del análisis AMMI (tabla 4.10), muestran también que el eje de la primera componente principal captura el 76% de la suma de cuadrados de la interacción y el 31,4% de los grados de libertad del error de la interacción.

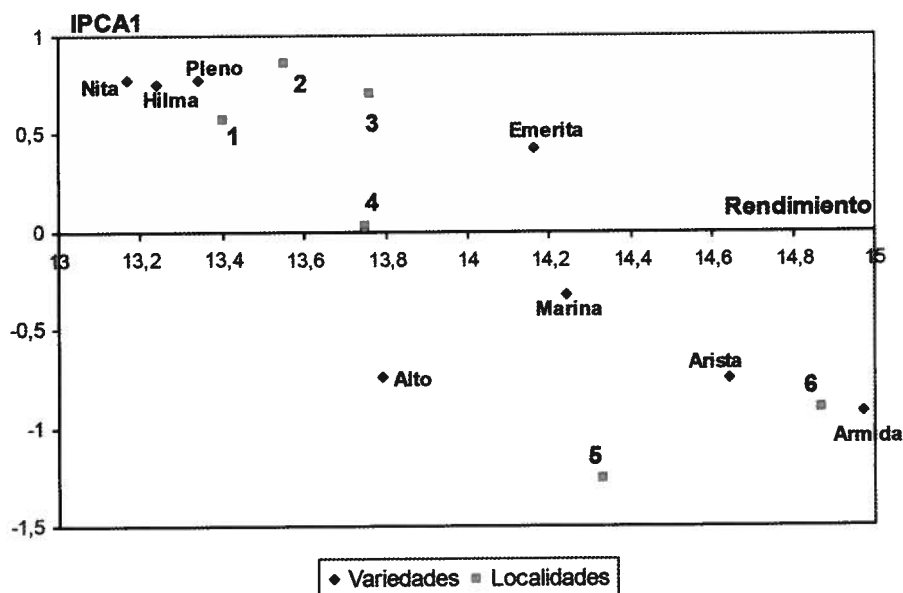
Comparativamente el método de regresión capturaba o explicaba tan solo el 57,8% de la suma de cuadrados de la interacción (tabla 4.8). Las dos primeras componentes capturan ya el 92% de la suma de cuadrados de la interacción. El cuadrado medio de las dos primeras componentes principales es significativo a  $p = 0.05$ .

Otra opción que suele ser el modelo más seguro de AMMI para predecir es el que emplea las dos componentes principales. Esto suele ocurrir a menudo y fue lo recomendado por Gausch y Zobel en 1996, pero es más difícil de interpretar gráficamente por lo que en nuestro caso utilizaremos el gráfico bidimensional de rendimiento y solo la primera componente principal.

Los valores de las componentes principales 1 (IPCA1) y las medias del rendimiento (means, kg azúcar/ha) necesarias para el gráfico (biplot) según la salida del programa Irristat son:

ITEM\$	IPCA1	Means
Alto	-0.737	13.79
Arista	-0.747	14.64
Armida	-0.916	14.97
Emerita	0.422	14.16
Hilma	0.752	13.24
Marina	-0.320	14.24
Nita	0.775	13.17
Pleno	0.772	13.34
Loc1 =A	0.572	13.40
Loc2 = B	0.857	13.55
Loc 3= C	0.707	13.76
Loc 4 = D	0.032	13.75
Loc 5 = E	-1.267	14.33
Loc 6 = F	-0.903	14.87

Figura 11. Representación gráfica del resultado del análisis AMMI (Componente principal 1 y rendimiento). Datos del ejemplo I



Las variedades muy próximas al eje de abscisas, como Marina, son variedades muy estables, que cambian poco su orden al cambiarlas de ambiente. Las variedades del cuadrante superior, Nita, Hilma y Pleno, están adaptadas a ambientes específicos, en este caso los menos productivos (localidades 1 y 2) y no van bien en los ambientes del cuadrante inferior. Las variedades del cuadrante inferior, Arista y Armida, estarían adaptadas específicamente a ambientes más productivos (localidad 6). En estos dos cuadrantes, como se ha señalado, a medida que el segmento que une cada tratamiento y cada localidad es más corto más adaptado está el tratamiento a esa localidad.

## 5. ANALISIS COMBINADO DE LOCALIDADES Y AÑOS.

### 5.1 Análisis en años.

A veces se realiza un experimento en la misma localidad durante varios años con objeto de evaluar el efecto del tiempo en unos determinados tratamientos. El analisis combinado de tratamientos y años es similar al caso anterior, cuando se combinan localidades en un mismo año, con la diferencia de sustituir años por localidades. Los primeros pasos son iguales en ambos casos. Las diferencias son que el factor años es siempre aleatorio (a la hora de interpretar los resultados, hay que tener la precaución de recordar que una muestra de varios años sucesivos no es representativa de años futuros) y que la variabilidad inducida por el tiempo en las unidades experimentales suele ser mayor que los cambios producidos por los tratamientos (Mead, 1988). Hay que señalar que algunos autores, en algunos planteamientos, han considerado los años como un factor fijo.

El análisis de la varianza tiene los siguientes factores de variación: Años, repeticiones dentro de años, tratamientos, interacción entre tratamientos y años y error combinado. Como ocurría con las localidades, el factor de variación que más interesa es la interacción, es decir la variación de comportamiento de los tratamientos en el tiempo. Si en el reparto de varianzas la interacción es pequeña en comparación con el efecto de los tratamientos (y normalmente no es significativa en este caso), el orden de los tratamientos a lo largo de los años es de esperar que sea constante (los tratamientos son estables) y la interacción su ele ser ignorada.

Para la prueba F de la interacción su cuadrado medio se divide por el cuadrado medio del error experimental. Si no es significativo, este denominador se emplea también para al cálculo de la F de los tratamientos. Si las diferencias entre los tratamientos resultan significativas, estos se comparan a continuación.

Si la interacción tratamientos por años es significativa hay que examinar que parte de la varianza total ocupa. Si es relativamente grande y el ranking (u orden) de los tratamientos cambia en los años (esto se puede ver en los agrupamientos de los años por separado), hay que examinar la naturaleza de la interacción por los métodos indicados en el apartado anterior, como por ejemplo hacer la partición de la suma de cuadrados basado en los contrastes apropiados. Siempre hay que recordar que la interpretación de la interacción año por tratamiento suele ser difícil por la complejidad de factores que determinan el efecto año.

Una práctica que simplifica la comprensión es comparar los años de dos en dos, analizar la diferencia de resultados y el comportamiento de los tratamientos en los años, así existirán tratamientos con buen resultado el primer año, el segundo o en ambos. Estos tratamientos serán los más consistentes.

El problema más frecuente de los experimentos similares, repetidos en el mismo sitio durante varios años, es que los rendimientos de la misma parcela en años sucesivos suelen estar correlacionados. Al no ser el error experimental de un año independiente del de los anteriores, el análisis de la varianza no es válido. Snedecor y Cochran en 1967 propusieron comparar los rendimientos globales de los tratamientos en todos los años para evitar lo anterior. De esta forma se proporciona un error válido para probar los efectos globales de los tratamientos. La limitación que tiene esta simplificación es que no se conoce el efecto de la interacción (que, recordamos, es lo que más interesa). El efecto de los años se puede conocer parcialmente probando la linealidad. Se puede completar lo anterior calculando para cada parcela la regresión lineal del rendimiento sobre los años para conocer la posible mejora de los resultados de año a año.

Otra alternativa para resolver el problema de las correlaciones entre años de las observaciones de una misma unidad experimental es hacer un modelo de parcela dividida donde los años sean las submuestras.

El problema asociado con la realización de un experimento en años consecutivos, de que estos años no son nunca una muestra aleatoria de todos los años posibles. Esto se puede reducir aumentando el número de años. Los experimentos repetidos a lo largo de un elevado número de años, reciben el nombre de experimentos a largo plazo ("Long-term experiments").

#### Experimentos a largo plazo.

En agronomía se usan generalmente ensayos repetidos en el tiempo, en la misma localidad, para estudiar el efecto a largo plazo de unos tratamientos. Por ejemplo cuando interesa conocer el efecto del uso de técnicas agrícolas, como la fertilización sobre propiedades del suelo, del control de insectos.

tos o malas hierbas sobre las diferentes poblaciones de especies...etc. El mismo ensayo se suele repetir durante un gran número de años, con los mismos tratamientos aplicados a las parcelas experimentales en el primer año. El sorteo se hace el primer año. Un ejemplo de diseño de ensayos a largo plazo es el siguiente:

Bloque I

Tratamiento I	Tratamiento 2	Tratamiento ...
Año 1 .....	Año 1 .....	Año 1 .....
Año 2 .....	Año 2 .....	Año 2 .....
Año 3 .....	Año 3 .....	Año 3 .....
.....	.....	.....
Media a1	b1	c1

Bloque II

Tratamiento I	Tratamiento 2	Tratamiento ...
Año 1 .....	Año 1 .....	Año 1 .....
Año 2 .....	Año 2 .....	Año 2 .....
Año 3 .....	Año 3 .....	Año 3 .....
.....	.....	.....
Media a2	b2	c2

Bloque III

Tratamiento I	Tratamiento 2	Tratamiento ...
Año 1 .....	Año 1 .....	Año 1 .....
Año 2 .....	Año 2 .....	Año 2 .....
Año 3 .....	Año 3 .....	Año 3 .....
.....	.....	.....
Media a2	b2	c2

Estos experimentos se analizan estadísticamente como parcela dividida (split-plot) en el tiempo, donde, en las repeticiones o bloques, las parcelas principales son los tratamientos (siempre sorteados) y dentro de ellas, las parcelas secundarias son los años (también, siempre sorteados). El análisis de la varianza correspondiente se muestra en la tabla 5.1.



**Tabla 5.1. Análisis de la varianza de un experimento de r repeticiones, con t tratamientos y a años diseñado en parcela dividida (Split-Plot) en el tiempo. Tratamientos factor principal.**

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio
Repeticiones	r-1	CM1	CM1/CM3
Tratamientos	t-1	CM2	CM2/CM3
rep x tratam (error a)	(r-1) (t-1)	CM3	-
Años	a-1	CM4	CM4/CM5 (1)
años x tratam	(a-1) (t-1)	CM5	CM5/CM6 (1)
Error b	diferencia	CM6 -	
Total	r*t*a-1		

(1)El efecto del año se prueba con la interacción tratamientos x años y la interacción a x t se prueba con el error b (McIntosh, 1983)

A diferencia del análisis multilocalidades, el error a, en vez de ser debido a la variación de bloques (o repeticiones) dentro de localidades, es debido a la interacción entre repeticiones y años. Si lo que cambia de año a año es el cultivo (ensayos de rotación de cultivos), el análisis de la varianza sería igual sustituyendo años por cultivo.

En ensayos donde los tratamientos están aplicados a las mismas parcelas año a año, los análisis de estabilidad son muy útiles y apropiados (Raun y colaboradores, 1993; Guertal y colaboradores, 1994).

El diseño en parcela dividida mostrado anteriormente se puede modificar, considerando los años como factor principal y tratamientos como secundario.

Este diseño es el más empleado en análisis de ensayos con cultivos perennes. Su tabla de análisis de la varianza sería el indicado en la tabla 5.2. En este caso la prueba F de los tratamientos es diferente del caso anterior. años diseñado en parcela dividida (Split-Plot) en el tiempo. Años factor principal.

**Tabla 5.2. Análisis de la varianza de un experimento de r repeticiones, con t tratamientos y a años diseñado en parcela dividida (Split-Plot) en el tiempo. Años factor principal.**

Fuente de variación	Grados de libertad	Suma de cuadrados
repeticiones	r-1	
Años	a-1	
Rep x Años (Error a)	(r-1)*(a-1)	
tratamientos	t-1	
tratamientos x años	(t-1)*(a-1)	
Trat*bloque + trat*bloque*año (error b)	Diferencia	
Total	r*t*a-1	

## 5.2 Análisis combinado de localidades y años

El caso más frecuente es cuando se repite un experimento en varias localidades y en varios años, bien con las mismas o diferentes localidades cada año. Existen varias posibilidades de abordar el problema.

En primer lugar, la elección del diseño depende de la duración de la investigación. Si se contemplan pocos años, es muy importante considerar la variación espacial. Si el experimento va a durar muchos años, es importante hacer una calibración, medir para predecir la variación de una parcela a otra en el mismo año. Haciéndolo en el primer año, se podrá usar como covariable para reducir la variación futura de las observaciones.

El resultado de un tratamiento en una determinada parcela viene definido por el modelo general:

$$Y_{ijk} = \mu + T_k + T_i + \theta_j + (T\theta)_{ij} + (TTT)_{ki} + (TTT\theta)_{kj} + (TT\theta)_{kij} + e_{ijk}$$

Donde:  $Y_{ijk}$  = media observada del tratamiento i en la localidad j y el año k

$\mu$  = gran media

$T_k$  = efecto del año k

$T_i$  = efecto del tratamiento i

$\theta_j$  = efecto de la localidad j

$(T\theta)_{ij}$ ,  $(TTT)_{ki}$ ,  $(TT\theta)_{kij}$  = interacciones dobles de los tratamientos, localidades y años.

$(TTT\theta)_{kj}$  = interacción triple

$e_{ijk}$  = media de los errores de las parcelas que han recibido el tratamiento i en la localidad j en el año k.

Se asume que las medias de todos los sumandos son  $N(0, \sigma^2)$ , normalmente distribuidas con media cero y varianza  $\sigma^2$  variable según el sumando. Existen cuatro modelos combinados de localidades y años en función de que las localidades cambien o se repitan y de si la relación entre localidades y años es cruzada o anidada.

**Modelo 1.** Es un análisis combinado sobre localidades y años con el mismo experimento repetido en las mismas localidades en varios años. Es el caso de experimentos en cultivos perennes y una extensión de experimentos repetidos en los años cuando existen varias localidades, es decir, un experimento repetido en las mismas localidades durante varios años. El análisis combinado de un experimento en bloques completos al azar de  $t$  tratamientos repetidos  $r$  veces en  $l$  localidades y  $a$  años con el mismo sorteo repetido año a año es el siguiente (tabla 5.3)

**Tabla 5.3. Analisis de la varianza de un experimento de  $t$  tratamientos con  $r$  repeticiones, con  $l$  localidades y  $a$  años. Todos los años hay el mismo sorteo en las mismas localidades.**

Fuente de variación	Grados de libertad	Suma de cuadrados
Localidades	$l-1$	
Rep (Localidades)	$(r-1)*l$	
Años	$a-1$	
Local x año	$(l-1)*(a-1)$	
Rep x año	$(loc) l*(r-1)*(a-1)$	
tratamientos	$t-1$	
Tratam x local	$(t-1)*(l-1)$	
Tratam x año	$(t-1)*(a-1)$	
Tratam x loc x año	$(t-1)*(l-1)*(a-1)$	
Error combinado $l*a*$	$(r-1)*(t-1)$	
Total	$T*r*l*a-1$	

En este modelo las localidades suelen ser fijas, empleándose generalmente fincas de centros públicos de experimentación agraria. La prueba  $F$  es la indicada en el caso 4 de los modelos que se muestran a continuación.

**Modelo 2.** En este modelo las localidades son las mismas cada año pero se sortean. Los años (u otro factor de tiempo) están cruzados con las localidades y los bloques se consideran dentro de las localidades y años. Es el modelo combinado más frecuente.

El modelo es una ampliación del modelo general antes indicado. La respuesta observada del tratamiento  $i$  en la localidad  $j$  en el año  $k$  y en la repetición  $r$  es:

$$R_{ijk_r} = m + G_i + L_j + Y_k + B_r(L_j Y_k) + G_{Lij} + GY_{jk} + LY_{jk} + GLY_{ijk} + e_{ijk_r}$$

Donde  $m$  es la media general,  $G$  es el tratamiento,  $L$  la localidad,  $Y$  el año,  $B$  el efecto de los bloques (o repeticiones) y  $e_{ijk_r}$  el error combinado.  $B_r(L_j Y_k)$  es el sumando añadido al modelo general.

Bajo este supuesto, un experimento de  $t$  tratamientos con  $r$  repeticiones en bloques completos al azar, repetido en  $l$  localidades durante  $a$  años, el análisis de la varianza en función de que los tratamientos y localidades sean fijos o aleatorias es el siguiente (ver tabla 5.4).

**Tabla 5.4. Valores de F según que los factores tratamientos y localidades sean fijos o aleatorios. Años siempre aleatorio. (En base a McIntosh, 1983).**

Factor de variación	Grados de libertad	Cuad. Medio	F1	F2	F3
Años	a-1	CM1	(1 + 8) / (3 + 6)	CM1/CM3	CM2/CM4
Localidades	l-1	CM2	(1)(2)	CM2/CM3	CM3/CM4
Años x Local	(a-1)(l-1)	CM3	(2+8)/(3+7) (2)	CM3/CM4	
Rep/años x Local	a*I*(r-1)	CM4	(3+9)/(4+8)	CM1/CM4	
Tratamientos	t-1	CM5	(5+8)/(6+7) (2)	5+8/6+7(2)	CM7/CM9
Tratam x años	(t-1)(a -1)	CM6	CM6/CM8	CM6/CM8	CM8/CM9
Tratam xLocal	(t-1)(l-1)	CM7	CM7/CM8	CM7/CM8	
Tratam x años x loc	(t-1)(l-1)(a-1)	CM8	CM8/CM9	CM8/CM9	
Error combinado	a*I*(t-1)(r-1)	CM9		CM5/CM9 CM6/CM9	
Total	a*I*r*t-1				

(1) Quiere decir (CM1+CM8)/(CM3+CM6)

(2) Los test señalados son factibles pero no absolutamente ciertos. Ver el texto.

F1 = Si Localidades y tratamientos son factores aleatorios. Años aleatorio  
 F2= Si Localidades son aleatorias y tratamientos factor fijo. Años aleatorio  
 F3= Si Localidades y tratamientos factores fijos. Años aleatorio

En el caso de que localidades y tratamientos sean factores aleatorios, no existe un test exacto para evaluar los efectos de las localidades. Una aproximación fue propuesta por Satterwaithe (1946) con los grados de libertad  $n_1$  y  $n_2$  que deberían usarse.

El test F para probar la significación es el siguiente

$$F_{n_1, n_2} = (CM_5 + CM_8) / (CM_6 + CM_7).$$

$$n_1 = (CM_1 + CM_5)^2 / (CM_1)^2 / (l-1) + (CM_5)^2 / l * (r-1)(t-1)$$

$$n_2 = (CM_2 + CM_4)^2 / (CM_2)^2 / l * (r-1) + (CM_4)^2 / (l-1)(t-1)$$

Como se ha indicado anteriormente, el valor de F calculado se compara con el valor de F tabulado para los valores de grados de libertad de  $n_1$  y  $n_2$  (Las tablas de la distribución F se pueden encontrar en muchos libros de estadística).

Antes de calcular estos grados de libertad aproximados  $n_1$  y  $n_2$ , se puede ganar tiempo comprobando la significación de las localidades usando los grados de libertad de las localidades como numerador y los grados de libertad de repeticiones en años o de la interacción tratamientos por localidades (el que sea más pequeño) en el denominador. Este es un test aproximado más conservador y si esta nueva F es significativa, no es necesario calcular  $n_1$  y  $n_2$ .

Kempthorne (1952) sugirió los cuadrados medios del error que deberían usarse como mas apropiados siendo una combinación de varios cuadrados medios. Para este modelo 2 son:

$$CM_{\text{error}} = CM_5 + CM_6 - CM_8 \text{ con los grados de libertad}$$

asociados n:

$$n = (CM_5 + CM_6 - CM_8)^2 / (CM_5^2 / (t-1)(l-1) + CM_6^2 / (t-1)(a-1) + CM_8^2 / (t * (l-1)(a-1)))$$

La razón que adujo es que otros efectos aleatorios pueden contribuir también al termino de error en el test F. Este test F es preferible al propuesto por Cochran y Cox (1957) porque el término de error puede ser rápidamente adoptado en formulas para el calculo de los errores estándar y la comparación de tratamientos.

Para calcular el reparto de varianzas, útil en el caso de factores aleatorios, la varianza estimada de los efectos principales y de las interacciones son las siguientes:

	Si tratamientos y localidades son aleatorios	Si tratamientos fijo y localidad aleatoria
Efecto de los tratamientos ( $\sigma^2_t$ )	(CM5+CM8-CM7-CM6)/ r*I*a	
Interacción tratamientos x localidades ( $\sigma^2_{tl}$ )	(CM7-CM8)/ r*a	(CM7-CM8)/ r*a
Interacción tratamientos x años ( $\sigma^2_{ta}$ )	(CM6-CM8)/ r*I	CM6-CM8)/ r*I
Interacción tratamientos x localidad x año ( $\sigma^2_{tla}$ )	(CM8-CM9)/ r	(CM8-CM9)/ r
Error experimental ( $\sigma^2_e$ )	CM9	CM9

**Modelo 3.** Este modelo es idéntico al anterior salvo que el factor tiempo esta anidado o jerarquizado al de localidades (ver diseños anidados más adelante). Es decir, para cada localidad hay años diferentes. Se usa cuando la comparación de años difiere entre localidades.

La respuesta observada del tratamiento i en la localidad j en el año k y en la repetición r es:

$$R_{ijk_r} = m + G_i + L_j + Y_k(L_j) + Br(Y_k L_j) + GL_{ij} + GY_{jk}(L_j) + e_{ijk_r}$$

Donde m es la media general, G es el tratamiento, L la localidad, Y el año, B el efecto de los bloques (o repeticiones) y  $e_{ijk_r}$  el error combinado. Respecto al modelo dos, no aparecen los sumando años ( $Y_k$ ) e interacción triple.  $Y_k(L_j)$  años jerarquizado a localidades y  $GY_{jk}(L_j)$  tratamientos x años jerarquizados a localidades son los sumandos añadidos al modelo general.

Bajo este supuesto, un experimento de t tratamientos con r repeticiones en bloques completos al azar, repetido en l localidades durante a años, con los años jerarquizados a las localidades, el análisis de la varianza en función de que los tratamientos y localidades sean fijas o aleatorias es el siguiente (ver tabla 5.5).

**Tabla 5.5. Valores de F según que los factores tratamientos y localidades sean fijos o aleatorios. Años siempre aleatorio. (En base a Anicchiarico, 2002).**

Factor de variación	Grados de libertad	Cuad. med.	F1	F2	F3
Tratamientos	t-1	CM1	CM1/CM5	CM1/CM5	CM1/CM6
Localidades	l-1	CM2	(2)	CM2/CM3	CM2/CM3
años(Localidades)	l*(a-1)	CM3	CM3/CM6	CM3/CM7	CM3/CM7
Rep/años x Local	a*l*(r-1)	CM4			
Tratam x Local	(t-1)(l-1)	CM5	CM5/CM6	CM5/CM6	CM5/CM6
Tratam x año(Local)	l*(t-1)(a-1)	CM6	CM6/CM7	CM6/CM7	CM6/CM7
Error combinado	t*a*l*(r-1)	CM7			

(2) Ver tabla 5.4

F1 = Si Localidades y tratamientos son factores aleatorios. Años aleatorio

F2= Si Localidades son aleatorias y tratamientos factor fijo. Años aleatorio

F3= Si Localidades y tratamientos factores fijos. Años aleatorio

En relación con el análisis jerarquizado, las localidades son los grupos y años(localidades) es la comparación entre subgrupos (los años) dentro de los grupos.

En este modelo, cuando los factores son aleatorios, la varianza estimada de los efectos principales y de las interacciones son las siguientes:

	Si tratamientos y localidades son aleatorios	Si tratamientos fijo y localidad aleatoria
Efecto de los tratamientos ( $\sigma^2_t$ )	$(M1-CM5)/r*l*a*$	-
Interacción tratamientos x localidades ( $\sigma^2_{tl}$ )	$(CM5-CM6)/ r*a$	$(CM5-CM6)/ r*a$
Interacción tratamientos x años ( $\sigma^2_{tla}$ )	$(CM6-CM7)/ r$	$CM6-CM7)/ r$
Error experimental ( $\sigma^2_e$ )	CM7	CM7

Modelo 4. Este modelo se puede considerar el inverso del anterior, donde los años son el factor grupos y las localidades jerarquizadas a los años los subgrupos. Las localidades son diferentes (o nuevas) cada año, es decir una determinada localidad tiene diferentes años que otra. La ecuación del modelo es similar a la anterior cambiando años (Y) por localidades (L).

El análisis de la varianza en función de que tratamientos y localidades sean fijas o aleatorias se indica en la tabla 5.6. La estimación de las varianzas es la misma del modelo anterior.

**Tabla 5.6. Valores de F según que los factores tratamientos y localidades sean fijos o aleatorios. Años siempre aleatorio. (Annicchiarico, 2002).**

Factor de variación	Grados de libertad	Cuad. med.	F1	F2	F3
Tratamientos	t-1	CM1	CM1/CM5	CM1/CM5	CM1/CM6
Localidades	l-1	CM2	(2)	CM2/CM3	CM2/CM3
localidades(años)	a*(l-1)	CM3	CM3/CM6	CM3/CM7	CM3/CM7
Rep/años x Local	a*l*(r-1)	CM4			
Tratam x Local	(t-1)(l-1)	CM5	CM5/CM6	CM5/CM6	CM5/CM6
Tratam x Local (año)	a(t-1)(l-1)	CM6	CM6/CM7	CM6/CM7	CM6/CM7
Error combinado	t*a*l*(r-1)	CM7			

(2) Ver tabla 5.4

F1 = Si Localidades y tratamientos son factores aleatorios. Años aleatorio

F2= Si Localidades son aleatorias y tratamientos factor fijo. Años aleatorio

F3= Si Localidades y tratamientos factores fijos. Años aleatorio

Este diseño dificulta el análisis de la respuesta de los tratamientos a las localidades por ello no se suele usar.

Un test F aproximado para las localidades en el caso de que los tratamientos sean aleatorios, sea como sea el factor localidad, en el modelo 2 se puede hacer usando el término de la interacción localidad x año como término de error lo que da un cuadrado medio más grande que usando la interacción tratamiento x localidad.

Con el año como factor aleatorio, el término de error apropiado para testar la interacción tratamiento x localidad en todos los experimentos multi-  
nacionales es siempre la interacción triple (tratamiento x localidad x año) cuando el año está cruzado con la localidad (modelo 2) y la interacción tratamiento x año cuando el factor años está jerarquizado en las localidades (modelo 3).



La mayor diferencia entre los modelos 2 y 3 es la incapacidad de este último modelo para separar la variación debida al efecto año a través de todas las localidades (Annicchiarico, 2002).

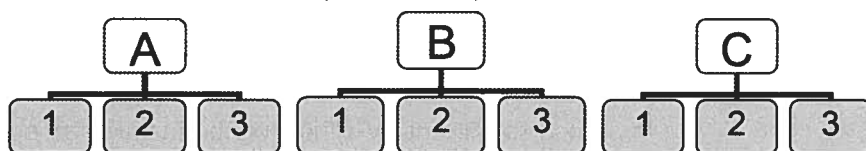
En ensayos en los que los tratamientos se repiten en varias localidades y años, la interacción que se consideraba en el capítulo anterior de interacción entre tratamientos y localidades se divide realmente en tres partes: tratamientos x localidades, tratamientos x años y tratamientos x localidades x años. Si de las tres interacciones es la más importante hay que dividir las localidades en subregiones para minimizar esta interacción. Cuando tratamientos x año o tratamientos x localidades x año son los términos dominantes no hay una solución clara. La interacción triple es difícil de interpretar.

### **Diseños Jerarquizados** (anidados o en inglés "nested")

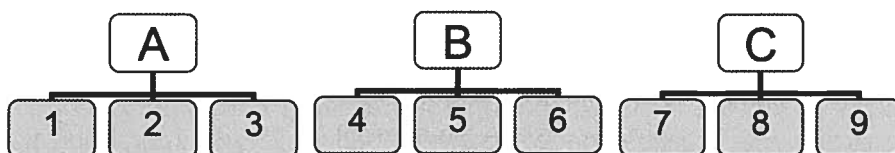
Si en un ensayo con dos factores, algunos niveles de un factor están ligados a ciertos niveles de otro se trata de un modelo jerárquico. El factor principal se denomina grupos y el secundario jerarquizado al principal subgrupos. Cada conjunto de a grupos (que es el nivel principal de clasificación) está dividido en b subgrupos jerarquizados al anterior. Hay n medidas en cada subgrupo. Puede haber más de un factor anidado y los factores pueden ser fijos o aleatorios.

En el diagrama adjunto se pueden apreciar las diferencias entre un diseño cruzado (convencional), apartado 1 y uno jerarquizado, apartado 2. En el supuesto de localidades y años, el factor principal (A-B-C) podrían ser tres años y el secundario (1-2-3-...) tres localidades iguales en el diseño cruzado o nueve localidades diferentes en el diseño anidado.

#### 1.- ANALISIS CRUZADO (Convencional)



#### 2.- ANALISIS JERARQUIZADO



Si los niveles del factor b pueden ser renumerados arbitrariamente son jerarquizados a otro a. Este diseño es frecuentes en ensayos de laboratorio, en procesos industriales o en estudios sobre métodos analíticos.

El análisis de la varianza de un modelo jerarquizado es el indicado en la tabla 5.7

**Tabla 5.7. Análisis de la varianza de un diseño con dos factores, grupos con a elementos y subgrupos dentro de grupos con b elementos y en cada uno n repeticiones.**

Factor de variación	Grados de Libertad	Suma de cuadrados	Cuadrado Medio	F
Entre grupos	a-1		CM1	F1
Entre subgrupos dentro de grupos	a(b-1)		CM2	F2
Dentro de subgrupos (Error)	a*b*(n-1)		CM3	
Total	a*b*n-1			

Para probar la significación de la variación entre grupos (F1)

$$F1 = CM1/CM2$$

Para probar la significación de los subgrupos dentro de grupos (F2)

$$F2 = CM2/CM3$$

En un modelo con factores aleatorios es adecuado estimar las varianzas de los grupos, de los subgrupos y del error, así como el porcentaje de variación total asignable a cada fuente de variación. Si el modelo es de factores fijos, los diferentes niveles del factor principal a se pueden comparar usando el término de error CM3.

En el programa SPSS para especificar que se trata de un diseño anidado hay que usar el editor de sintaxis (PASTE), cambiando la última fila, donde hay que incluir la sentencia factor b (factor a) y finalizar con run.

- /DESIGN = factor b. **factor b (factor a)**  
RUN

#### EJEMPLO 4.

*"Se prueba un procedimiento en tres laboratorios diferentes y en cada uno de ellos por cuatro operarios. Cada operario hace dos medidas. La varia-*

ble respuesta es longitud en unidades micrométricas.(Adaptado de Sokal y Rohlf, 1977)”

Factor a = 3 laboratorios.

Factor b = 4 operarios por laboratorio. Aleatorio (El operario de un laboratorio es diferente del de otro)

	Laboratorio 1				Laboratorio 2				Laboratorio 3			
Operario	1	2	3	4	1	2	3	4	1	2	3	4

Ejecución en SPSS (v 8.0)

UNIANOVA

longitud BY laboratorio operario

/RANDOM = operario

/METHOD = SSTYPE(3)

/INTERCEPT = INCLUDE

/CRITERIA = ALPHA(.05)

/DESIGN = laboratorio operario (laboratorio). (Esto se pone nuevo)

**Tabla 5.8. resultados del análisis análisis de la varianza del ejemplo de diseño jerarquizado 4.**

**Tests of Between-Subjects Effects**

Dependent Variable: LONGITUD

Source		Type III Sum of Squares	df	Mean Squares	F	Sig.
Intercept	Hypothesis	106560,027	1	106560,027	557,362	,000
	Error	1720,678	9	191,186a		
LABORATORIO	Hypothesis	665,676	2	332,838	1,741	,230
	Error	1720,678	9	191,18a6		
OPERAR(LAB)	Hypothesis	1720,678	9	191,186	146,878	,000
	Error	15,620	12	1,302b		

a. MS(OPERAR(LAB))

b. MS(Error)

Como se puede apreciar el la tabla 5.8, no existe evidencia de una componente significativa de la varianza entre laboratorios, pero existe una componente añadida de la varianza altamente significativa entre operarios.

El cuadrado medio del error es = 1.30. Al ser aleatorio, se analizan las varianzas.

Entre subgrupos dentro de grupos =  $(191.19 - 1.30) / n = 94.94$

Entre grupos =  $(332.84 - 191.19) / n*b = 17.71$

Como interesan solo las magnitudes relativas, los componentes de la varianza se pueden poner como porcentaje de la suma de sus varianzas:

$$1.30 + 94.94 + 1.71 = 113.95$$

$$\text{El error representa} = 1.3/113.95 * 100 = 1.14 \%$$

$$\text{Los operarios representan} = 94.94/113.95 * 100 = 83.3 \%$$

Los laboratorios representan =  $17.71/113.95 * 100 = 15.54 \%$ . Este estudio se conoce como "reparto de varianzas".

En un análisis multianual, para poder conocer las interacciones de los tratamientos x año, tratamientos x localidades o tratamientos x año x localidades, es necesario que un mismo grupo de tratamientos se haya probado en la misma localidad durante más de un año. El análisis estadístico se complica sobremedida cuando intervienen ensayos diferentes en cada año, cuando los ensayos tienen diferentes repeticiones o cuando en las localidades hay tratamientos diferentes. Esto último es bastante frecuente, incluir en cada localidad un tratamiento "local" de referencia (para información en estos casos, consultar la bibliografía especializada como Cochran y Cox, 1957).

Se puede detectar una menor diferencia significativa entre tratamientos a medida que la varianza de los tratamientos decrece. Es decir, dos tratamientos que difieran en una menor cantidad se considerarán estadísticamente diferentes. En este sentido, se ha demostrado que a medida que el número de años, localidades o repeticiones de un experimento aumenta, la varianza estimada de los tratamientos es menor. Esto es especialmente importante si se pretende comparar con precisión los tratamientos en un experimento multilocalidad o multianual. Se ha observado que cuando el número de localidades aumenta, el efecto de aumentar el número de años para reducir la varianza de los tratamientos es menor. Esto tiene una gran relevancia, dado que uno de los posibles objetivos de hacer ensayos repetidos en localidades y años es determinar el número de localidades idóneas para hacer una red de ensayos para futuras recomendaciones.

Para un determinado número de repeticiones, también es posible determinar el número óptimo de localidades y años, siempre que se conozca (por ejemplo, por estudios o trabajos previos) la componente de la varianza de las interacciones tratamiento x año y tratamiento x localidad.

El valor de F de la interacción triple se obtiene siempre dividiendo su cuadrado medio por el cuadrado medio del error combinado. Si esta interacción triple es significativa, la práctica común es tratar cada experimento como único efecto localidad x año.

Si las interacciones dobles tratamiento x año o tratamiento x localidad no son significativas (generalmente se dividen por la interacción triple, pues las localidades suelen ser aleatorias) la otra interacción años x localidad se puede usar para testar los tratamientos.

Si la interacción tratamiento x localidad resulta no significativa y la interacción tratamiento x año resulta significativa, se divide en cuadrado medio de los tratamientos por el cuadrado medio de tratamientos x años si el valor de F es igual de un año a otro.

Cuando la interacción tratamientos x años es significativa y es el cuadrado medio más alto, es evidente que las diferencias entre años (¡el clima!) son el factor más importante en determinar el efecto de los tratamientos y será difícil predecir los efectos de los tratamientos en el futuro. En este caso se suelen comparar los resultados de los tratamientos de cada año en una tabla.

El efecto de los años se puede estudiar midiendo las tendencias de los resultados de los diferentes tratamientos en los años. Se pueden hacer contrastes con las medias de años, igual que se hicieron contrastes con las medias de los tratamientos. Todas las interacciones se pueden partir en componentes tal y como se describió en el capítulo anterior.

Recordamos que este cálculo hay que hacerlo siempre a mano.

#### **EJEMPLO 5.**

*"Se ensayan tres tratamientos de laboreo del suelo diferentes (C1,C2,C3) que producen distinto grado de compactación del terreno, en un diseño en bloques completos al azar con 4 repeticiones. Se evalúa el porcentaje de plantas de habas nacidas respecto a las sembradas. El ensayo se repite en las mismas 4 localidades durante 5 años. Los tratamientos y las localidades se consideran fijos.*

El primer análisis es de un ensayo simple. En la tabla 5.9 se puede ver que en esta localidad del primer año, las diferencias entre tratamientos es altamente significativa ( $p = 0.015 < 0.05$ ). La comparación de medias de tratamientos, según el test de Tukey (HSD), ver tabla 5.10, indica que sólo los tratamientos 1 y 3 son estadísticamente diferentes, siendo C1 el tratamiento de laboreo que proporciona un mayor porcentaje de nascencia de plantas.

**Tabla 5.9 Análisis de la varianza de un simple ensayo en una localidad. 3 tratamientos y 4 repeticiones en SPSS.****Tests of Between-Subjects Effects**

Dependent Variable: PLANTAS

Source	Type III Sum of Squares	df	Mean Squares	F	Sig.
Corrected Model	659,833	5	131,967	4,307	,052
Intercept	56856,333	1	56856,333	1855,692	,000
REPETICI	93,667	3	31,222	1,019	,448
TRATAMIE	566,167	2	283,083	9,239	,015
Error	183,833	6	30,639		
Total	57700,000	12			
Corrected Total	843,667	11			

a R Squared = ,782 (Adjusted R Squared = ,601)

**Tabla 5.10 Comparación de medias de los tratamientos del ejemplo 5 según el test de Tukey B a  $p= 0.05$ . Variable Plantas.**

	N	Subset	
TRATAMI		1	2
3,00	4	60,0000	
2,00	4	69,7500	69,7500
1,00	4		76,7500

Means for groups in homogeneous subsets are displayed. Based on Type III Sum of Squares

The error term is Mean Square(Error) = 30,639.

a Uses Harmonic Mean Sample Size = 4,000.

b Alpha = ,05.

En el análisis combinado del primer año (ver tabla 5.11), resulta la interacción entre tratamientos y localidades no significativa ( $p= 0,113$ ). Por lo que se puede concluir que en este primer año no ha existido un tratamiento idóneo para una determinada localidad. El tratamiento 1 (C1) con una nascencia de 71.93 %, ha resultado significativamente el mejor en todas las localidades de este primer año. (Ver Tabla 5.12). Los cuadrados medios (CM) de todos los factores de variación, se incluyen en la tabla 5.13, necesaria para la comparación entre años. Igual se procedería con los años 2º, 3º, 4º y 5º.

El procedimiento en SPSS de analizar cada año por separado puede ser tedioso. Una alternativa es partir el archivo, organizando la salida para cada

grupo definido por la variable año. Para partir el archivo en SPSS las instrucciones son:

Data

Split file

Organize output by group

• Group based on: Años

OK

Los resultados del análisis combinado anuales se muestran en la tabla 5.12

**Tabla 5.11. Análisis de la varianza combinado de los experimentos de las 4 localidades del primer año. 3 tratamientos, 4 repeticiones y 4 localidades. Análisis en SPSS.**

**Tests of Between-Subjects Effects**

Dependent Variable: PLANTAS

Source	Type III Sum of Squares	df	Mean Squares	F	Sig.
Corrected Model	3077,583	23	133,808	4,105	,001
Intercept	213600,083	1	213600,083	6552,708	,000
LOCALIDA	1543,417	3	514,472	15,783	,000
REPETICI * LOCALIDA	291,167	12	24,264	,744	,697
TRATAMIE	861,292	2	430,646	13,211	,000
TRATAMIE * LOCALIDA	381,708	6	63,618	1,952	,113
Error	782,333	24	32,597		
Total	217460,000	48			
Corrected Total	3859,917	47			

a R Squared = ,797 (Adjusted R Squared = ,603)

**Tabla 5.12 Comparación de medias de los tratamientos del ejemplo 5 en el primer año, según el test de Tukey B a p= 0.05. Variable Plantas.**

	N	Subset		
TRATAMI		1	2	3
3,00	16	61,5625		
2,00	16		66,6250	
1,00	16			71,9375

Means for groups in homogeneous subsets are displayed. Based on Type III Sum of Squares

The error term is Mean Square(Error) = 32,597.

a Uses Harmonic Mean Sample Size = 16,000.

b Alpha = ,05.

**Tabla 5.13. Análisis de la varianza combinado anual del ejemplo II. (3 tratam, 4 rep, 4 loc y 5 años). CM = Cuadrado Medio.**

Fuente de variación	Grados de libertad	CM año1	CM año2	CM año3	CM año4	CM año5
Localidad	3	514,472**	109,500 NS	195,576 NS	226,632 *	26,688NS
Rep/local	12	24,264	34,792	46,007	34,243	54,424
tratam	2	430,646 **	1807,938**	1101,813**	1263,896 **	1311,812**
Trata x loc	6	63,618 NS	28,104 NS	32,035 NS	42,424 NS	30,646
Error com	24	32,597	61,479	81,174	35,764	30,299

NS = No significativo

\* = significativo a  $p=0.05$

\*\* = significativo a  $p= 0.01$

En el análisis año a año, ver tabla 5.13, se puede apreciar que la interacción tratamiento x localidad no es significativa en ningún año. Las diferencias entre tratamientos es altamente significativa y el efecto de la localidad es o no significativo según el año. Analizando el reparto de varianzas, la mayor parte de la variación es debida a los tratamientos (entre el 90% del 5º año y el 40% del 1º año), seguida de las localidades ( de media han representado el 16% de la variación) y de la interacción. El error ha ocupado aproximadamente sólo el 3% de la variación. Un análisis similar se puede hacer entre tratamientos y localidades.

**Tabla 5.14. Resultados medios en 4 localidades de 3 tratamientos de laboreo probados en 5 años consecutivos (porcentaje de plantas de habas nacidas).**

Tratamiento	Año 1	Año 2	Año 3	Año 4	Año 5
C3	61.2 c	54.5 c	57.6 c	56.2 c	58.7 c
C2	66.6 b	64.6 b	58.6 b	63.0 b	65.0 b
C1	71.9 a	75.8 a 7	2.5 a	73.8 a	76.1 a
Media	66.7	65.0	62.9	64.3	66.4
E.E.	4.03	5.54	6.37	4.22	3.89

E.E = Error estándar de la diferencia entre dos medias

Letras diferentes indican diferencias significativas según el test de Tukey(HSB) a  $p=0.05$



En relación con el comportamiento de los tratamientos, el tratamiento C1 es el que produce significativamente una mayor nascencia de plantas todos los años. En el caso de que algún año, el orden hubiese cambiado, sería necesario estudiar detenidamente que condiciones han sido particularmente diferentes ese año.

El resultado del análisis combinado multianual de localidades y años se muestra en la tabla 5.15.

**Tabla 5.15. Análisis de la varianza combinado completo del ejemplo 5. (3 tratam, 4 rep, 4 loc y 5 años). Ejecución en SPSS.**

Tests of Between-Subjects Effects

Dependent Variable: PLANTAS.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	18967,643	119	159,392	3,276	,000
Intercept	1012733,379	1	1012733,379	20812,609	,000
AÑO	486,092	4	121,523	2,497	,046
LOCALIDA	566,113	3	188,704	3,878	,011
LOCALIDA * AÑO	2654,685	12	221,224	4,546	,000
REPETICI *	2402,014	60	40,034	,823	,798
LOCALIDA * AÑO					
TRATAMIE	10815,959	2	5407,980	111,139	,000
TRATAMIE * AÑO	790,009	8	98,751	2,029	,048
TRATAMIE *	319,634	6	53,272	1,095	,369
LOCALIDA					
TRATAMIE *	846,230	24	35,260	,725	,818
LOCALIDA * AÑO					
Error	5839,153	120	48,660		
Total	1041799,000	240			
Corrected Total	24806,796	239			

a R Squared = ,765 (Adjusted R Squared = ,531)

Se puede observar por la significación del modelo completo y el valor del  $r^2$  que el análisis de la varianza es muy apropiado para el análisis de resultados de la comparación de los tres tratamientos de laboreo en las diferentes localidades y años. La interacción triple de tratamiento por localidad y año es no significativa, por lo que a continuación, se analizan las interaccio-

nes dobles. La interacción tratamiento x localidad no es significativa (en casi todas las localidades es mejor el tratamiento C1, por lo que no hay interacción). La interacción tratamiento x año está próxima a la no significación (se ha analizado el comportamiento de los tratamientos en los años en la tabla 5.14) y la interacción localidad x año es muy significativa. Esto quiere decir que las distintas localidades han dado resultados diferentes según los años (posiblemente las condiciones climáticas del año variables según la localidad, han podido afectar a la nascencia de las plantas). Para un estudio detallado, en caso que interese, se podría hacer una tabla intermedia (tipo 5.14) de localidades y años.

Al no ser significativas las dos primeras interacciones es conveniente emplear la tercera como denominador para testar los tratamientos, así tendríamos que la F de los tratamientos sería  $5407,98 / 221,22 = 24.44$  (\*\*\*) significativo también pero valor más conservador que el calculado en el modelo general (111.13 \*\*\*).

En general se pretende que las conclusiones obtenidas de un estudio en el que se implican varias localidades y años, sean de aplicación en una region amplia y para varios años. Para ello es importante que los resultados obtenidos en condiciones bien definidas y conocidas, por ejemplo en estaciones de experimentación, sean verificados siempre en otras condiciones. Si se pretende hacer una recomendación general es necesario realizar los ensayos en diferentes localidades cada año, cambiar las localidades para proporcionar condiciones lo más diversas posibles.

Los experimentos en los que intervienen localidades y años se pueden diseñar y analizar estadísticamente también como ensayos en parcela dividida (split-plot) en el tiempo, donde, en los años son las repeticiones, las parcelas principales son las localidades y dentro de ellas, las parcelas secundarias son los tratamientos (repetidos, logicamente). Las localidades en los años pueden ser las mismas (sorteadas) o diferentes. Los tratamientos dentro de las localidades deben estar siempre sorteados.

El análisis de la varianza correspondiente a ambos casos con un diseño en parcela dividida (con los correspondientes términos de error a y error b) ha sido propuesto en el programa MSTATC (1988). Las diferencias se pueden ver en las tablas 5.16 y 5.17.

**Tabla 5.16 Análisis de la varianza combinado de un experimento realizado en varias localidades (l) y años (a) con las mismas localidades cada año (sorteadas).**

Fuente de variación	Grados de libertad	Suma de cuadrados
Años	a-1	
Localidades	l-1	
Localidades x años	(a-1)*(l-1)	
Repet dentro de loc x años	a*(r-1) = error a	
tratamientos	t-1	
tratamientos x años	(t-1)*(a-1)	
Tratam x Local	(t-1)*(l-1)	
tratamx loc x añosx	(l-1)*(a-1)*(t-1)	
Repet combinadas en a*I*t	l*a*(r-1)*(t-1)= error b	
total	a*I*r*t-1	

**Tabla 5.17. Análisis de la varianza combinado de un experimento realizado en varias localidades (l) y años (a) con diferentes localidades cada año.**

Fuente de variación	Grados de libertad	Suma de cuadrados
Años	a-1	
Localidades en años	a*(l-1)	
Repet dentro de localidades x años	a*(r-1) = error a	
tratamientos	t-1	
tratamientos x años	(t-1)*(a-1)	
Local x tratam (en años)	a*(l-1)*(t-1)	
Repet x tratam (en años)	a*(r-1)*(t-1) = error b	
total	a*I*r*t-1	

Dependiendo del planteamiento del experimento, existen numerosas posibilidades de combinar tratamientos, localidades y años. Entre ellas un diseño en parcela sub -subdividida (split-split-plot), donde la parcela principal es la repetición, dentro de esta la secundaria la localidad y a su vez dentro de esta, la terciaria, es el año dentro de los cuales se encuentran sorteados los tratamientos. Por ejemplo, con r repeticiones, l localidades, a años

y tratamientos, todos considerados fijos excepto el año, el análisis de la varianza sería el indicado en la tabla 5.18. Cada factor de variación se testa con el error situado inmediatamente debajo (como un diseño en bloques divididos).

**Tabla 5.18. Análisis de la varianza combinado de un experimento realizado en varias localidades (l) y años (a) en diseño split-split-plot.**

Fuente de variación	Grados de libertad	Suma de cuadrados
Repeticiones	$r-1$	
Localidades	$l-1$	
Repet en localidades = error a	$(r-1)*(l-1)$	
tratamientos	$t-1$	
Tratamientos*local	$(t-1)*(l-1)$	
Repet en Localidades* tratam=error b	$l*(r-1)*(t-1)$	
Años	$a-1$	
Años*local	$(a-1)*(l-1)$	
Tratam*Años	$(t-1)*(a-1)$	
Local*Tratam*Años	$(l-1)*(t-1)*(a-1)$	
Repet en interac triple = Error c	Por diferencia	
total	$a*l*r*t-1$	

### 5.3 Ensayos repetidos en varias localidades y años con más de un factor

Si los tratamientos, a su vez, son combinaciones de varios factores, por ejemplo A y B, en el caso de un experimento con estos dos factores fijos en parcela dividida, repetidos  $r$  veces en  $l$  localidades e  $y$  años, se puede emplear el modelo de análisis de la varianza para el análisis combinado indicado en la tabla 5.19.

**Tabla 5.19. Analisis de la varianza de un experimento diseñado en parcela dividida combinado sobre localidades y años. A y B factores fijos. Adaptado de McIntosh, 1983.**

Fuente de variación	Grados de libertad	Suma de Cuadr.	F1	F2
Años	$y-1$	CM1	CM1/CM3	CM1/CM4
localidades	$l-1$	CM2	CM2/CM3	CM2/CM3
años x localidades	$(y-1)(l-1)$	CM3	CM3/CM4	CM3/CM4
bloque/loc x años	$y^*l*(r-1)$	CM4		
Factor A	$a-1$	CM5	$(CM5+CM8)/$ $(CM6+CM7)$	CM5/CM6
Factor A x años	$(a-1)(y-1)$	CM6	CM6/CM8	CM6/CM9
Factor Ax localidad	$(a-1)(l-1)$	CM7	CM7/CM8	CM7/CM8
Factor A x año x loc	$(a-1)(y-1)(l-1)$	CM8	CM8/CM9	CM8/CM9
Error a	$y^*l*(a-1)(r-1)$	CM9		
Factor B	$(b-1)$	CM10	$(CM19+CM13)/$ $(CM11+CM12)$	CM10/CM11
Factor B x años	$(b-1)(y-1)$	CM11	CM11/CM13	CM11/CM18
Factor B x loc	$(b-1)(l-1)$	CM12	CM12/CM13	CM12/CM13
Factor B x año x Loc	$(b-1)(y-1)(l-1)$	CM13	CM13/CM18	CM13/CM18
Factor A x Factor B	$(a-1)(b-1)$	CM14	$(CM14+CM17)/$ $(CM15+CM16)$	CM14/CM15
A x B X años	$(a-1)(b-1)$	CM15	CM15/CM17	CM15/CM18
		$(y-1)$		
A x B X Loc	$(a-1)(b-1)(l-1)$	CM16	CM16/CM17	CM16/CM17
A x B X año x loc	$(a-1)(b-1)$	CM17	CM17/CM18	CM17/CM18
	$(y-1)(l-1)$			
Error combinado =	$y^*l*a*(b-1)$	CM18		
Error b	$(r-1)$			

F1 = Si años y localidades son al azar

F2 = Si años son al azar y localidades son fijas

Como referencia, las fuentes de variación en diseños combinados generales que implican varias localidades y años y dos factores A y B, en función de diferentes enfoques (convencional o parcela dividida) y variación o no de las localidades serían las indicadas en la tabla 5.20

**Tabla 5.20 Factores de variación en los análisis de la varianza combinados de localidades y años con dos factores A y B. (Y años, L localidades, R repeticiones)**

Nueva Localidad cada año	Mismas Localidades pero sorteadas cada año	Mismas Localidades y mismo sorteo cada año	Nueva Localidad cada año Parcela dividida	Mismas Localidades pero sorteadas cada año Parcela dividida	Mismas Localidades y mismo sorteo cada año Parcela dividida
Y	Y	L	Y	Y	L
L(Y)	L	R(L)	L(Y)	L	R(L)
R(LY)	Y*L	Y	R(LY)	Y*L	Y
A	R(LY)	L*Y	A	R(YL)	L*Y
Y*A	A	R*Y(L)	Y*A	A R*	L(Y)
L*A(Y)	Y*A	A	L*A(Y)	Y*A	A
B	L*A	L*A	ERROR a	L*A	L*A
Y*B	Y*L*A	Y*A	B	Y*L*A	Y*A
L*B(Y)	B	L*Y*A	Y*B	ERROR a	L*Y*A
A*B	Y*B	B	L*B(Y)	B	ERROR a
Y*A*B	L*B	L*B	A*B	Y*B	B
L*A*B(Y)	Y*L*B	Y*B	Y*A*B	L*B	L*B
ERROR	A*B	L*Y*B	L*A*B(Y)	Y*L*B	Y*B
	Y*A*B	A*B	ERROR	A*B	L*Y*B
	L*A*B	L*A*B		Y*A*B	A*B
	Y*L*A*B	Y*A*B		L*A*B	L*A*B
	ERROR	L*Y*A*B		Y*L*A*B	Y*A*B
		ERROR		ERROR	L*Y*A*B
					ERROR

### 5.4 Englobe de efectos

Cuando el valor de F calculado para una determinada fuente de variación se aproxima a uno, los cuadrados medios, el del numerador y el del denominador son cifras parecidas y podría no tener sentido mantener los dos factores en el modelo. Se podrían englobar en uno solo. Esta técnica se denomina englobe de efectos. Con ello se obtiene una mejor estima de la variación del factor que actúa como denominador (generalmente el error en los factores principales).

Esta técnica es especialmente útil cuando hay muchas fuentes de variación, bien factores principales, interacciones simples o dobles.

El problema es decidir si la F es diferente de 1 y no cometer un error de tipo II. En relación con esta técnica hay opiniones diferentes. Unas a favor de mantener todos los términos del modelo y no englobarlos unos con otros aunque no sean significativos y otras a favor de englobar siempre que haya sido significativo el efecto a un nivel de significación decidido de antemano. Se ha aconsejado el englobe de efectos solo si la prueba F no es significativa al nivel del 25% (Bozovich y colaboradores, 1956).

Para englobar los efectos se empieza por las interacciones de orden superior hasta llegar a los factores principales. Estos no se suelen incluir. Para hacer el englobe se suman las sumas de cuadrados y los grados de libertad de los efectos que se engloban, con los nuevos valores se obtiene el correspondiente cuadrado medio.

Como ejemplo supongamos un experimento de 7 tratamientos con 4 bloques repetido en 3 años (ver tabla 5.1). El análisis de la varianza se muestra en la tabla 5.21

**Tabla 5.21. Análisis de la varianza del experimento de 7 tratamientos con 4 bloques, repetido en 3 años.**

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F calculada	Sign.
Bloques	3				
Tratamientos	6				
Bloquesx tratam	18				
Años	2				
Años x tratam	12	4.72	0.393	1.03 >	0.25
Error	42	16.00	0.381		
Total	83	8.976	0.102		

De la tabla anterior, la interacción años x tratamiento se puede englobar con el error para dar una nueva estima de este efecto. Los nuevos grados de libertad del error serían  $12+42=54$  y la nueva suma de cuadrados:  $4.72+16.00= 20.72$  por lo que la nueva estimación del cuadrado medio del error sera  $20.72/54 = 0.383$ . La tabla nueva será ahora:

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F calculada	Sign.
Bloques	3				
Tratamientos	6				
Bloques x tratam	18	7.61	0.423	1.10	>0.25
Año	3				
Error	54	20.72	0.383		

Ahora la interacción bloque x tratamiento puede englobarse con el nuevo error, con lo que quedaría:

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio	F calculada	Sign.
Bloques	3				
Tratamientos	6				
Año	3				
Error	72	28.33	0.393		

La varianza del error pasa de 0.381 a 0.383 y a 0.393, con lo que la prueba es más conservadora al tener mayor número de grados de libertad en el denominador.





## 6. BIBLIOGRAFIA

- H.Aastveit y H.Martens (1986). "ANOVA interactions interpreted by partial least squares regression". *Biometrics* nº 42, pág. 829-844.
- P.Annicchiarico (2002). "Genotype by environment interactions". *FAO paper* nº 174. Ed: FAO. Rome. Pág 132.
- K.E. Basford y J.W. Tukey (2000). "Graphical analysis of multiresponse data: illustrated with a plant breeding trial". Ed: Chapman & Hall /CRC Press. Boca Ratón. EE.UU.
- M.S.Bartlett (1947). "The use of transformations". *Biometrics*, nº 3. pág. 39 – 52.
- R.Bernardo (1992). "Weighted vs unweighted mean performance of varieties across environments". *Crops Science* nº 32. pág. 490 – 492.
- D.T.Bowman y C.E.Watson (1997). "Measures of validity in cultivar performance trials". *Agronomy Journal* nº 89, pág. 860 – 866.
- H.Bozivich, T.A.Bancroft y H.O.Hartley (1956). "Power of analysis of variance test procedures for incompletely specified models". *Annals of mathematical statistics*, vol 27, pág. 1017 – 1043.
- A.Camussi, F.Möller, E. Ottaviano y M.Sari Gorla (1995). "Metodi statistici per la sperimentazione biologica". Zanichelli Editore s.p.a. Bologna. pág. 479.
- S.J.Coakes y L.G.Steed (1999). "SPSS analysis without anguish". Ed. John Wiley and Sons. Australia.Ltd. 283 pág.
- W.G.Cochran y G.M.Cox (1957). "Experimental Designs". Ed. John Wiley and Sons, Inc. pág. 661.
- J.Crossa.(1990). "Statistical analyses of multilocation trials". *Advances in agronomy*. Nº 44, pág. 55-85.

- J.I.Cubero y F.Flores. (1994). "Metodos estadísticos para el estudio de la estabilidad varietal en ensayos agrícolas". Ed: Junta de Andalucía. serie: Monografías. 12/94. 176 pág.
- P.Dagnelie (1975). "Theorie et méthodes statistiques ". Vol 2. Ed: Les presses agronomiques de Gembloux, Belgica.
- P.Dagnelie (1981). "Principes d' experimentation". Ed: Les presses agronomiques de Gembloux, Belgica. A.S.B.L. pág. 182.
- I.H. DeLacy, K.E. Basford, M. Cooper, J.K. Bull y C.G. McLaren (1996). "Analysis of multienvironment data". In Plant adaptation and crop improvement, pág. 39-124. Ed: Cooper and Hammer. Wallingford. Reino Unido.
- S.A.Eberhart y W.A.Russell (1966). " Stability parameters for comparing varieties". Crops Science nº 6 pág. 36-40
- W.T. Federer (1956). "Augmented designs". Hawaiian Planters Record LV(2) pág. 191-208.
- K.W.Finlay y G.N.Wilkinson (1963). " The analysis of adaptation in a plant breeding programme". Australian Journal Agricultural research. nº 14. pág. 742-754.
- P.N.Fox, J.Crossa y I. Romagosa. (1997) "Multi-environment testing and genotype x environment interaction". Cap 8. Statistical methods for plant varieties evaluation. Ed. CIHEAM y Chapman & Hall, pág. 117-138.
- P.N.Fox, B.Skovmand, B.K.Thompson, H.J.Braun y R.Cormier.(1990). "Yield and adaptation of hexaploid spring triticale". Euphytica nº 47, pág. 57-64.
- T.R.Francis y L.W.Kannenber (1978). " Yield stability studies in short-season maize.A descriptive method fot grouping genotypes". Canada Journal of Plant Science, nº 58, pág. 1029-1034.
- H.G.Gausch (1992). "Statistical analysis of regional yield trials:AMMI analysis of factorial design". Ed. Elsevier. Amsterdam.
- H.G.Gausch y R.W. Zobel (1996). "AMMI analysis of yield in trials". In Genotype by environment interaction pág. 85-122. Ed:CRC Press. Boca Raton EEUU.
- K.A.Gomez y A.A. Gomez (1984). "Statistical procedures for agricultural research". Ed. John Wiley and Sons,Inc. pág. 680.
- M.L.Gumpertz (1995). "Data transformation". Special Report. North Carolina State Univ. Biological and cultural tests. pág. 5.
- E.A.Guertal, W.R.Raun, R.L.Westerman y R.K. Boman (1994). "Application of stability analysis for single-site, long term experiments". Agronomy Journal nº 86, pág. 1016-1019.

- M.Hühn (1997). "Weighted mean are unnecessary in cultivar performance trials". *Crop Science* nº 37. págs. 1745 – 1750.
- IRRI STAT. Software desarrollado por el IRRI (International Rice Research Institute). Se puede descargar de [www.irri.org/](http://www.irri.org/). IRRI biometrics Unit. DAPO Box 7777, Metro Manila Philippines.
- M.S. Kang y D.P. Gorman (1989). "Genotype x environment interaction in maize". *Agronomy Journal*. nº 81 págs. 662-664.
- R.A. Kempton (1984). "The use of biplots in interpreting variety by environment interactions". *Journal of Agricultural Science*. Nº 103, págs. 123-135.
- O. Kempthorne (1952). "The design and analysis of experiments". Ed: J.Wiley and Sons. N. York.
- R.A.Kempton y A.C.Gleeson. (1997). "Unreplicated trials". En *Statistical methods for plant varieties evaluation*. Ed. CIHEAM y Chapman & H.Ketata, S.K.Yan y M.Nachit.(1989)."Relative consistency performance across environments".*International Symposium on Physiology and Breeding of winter cereals for stressed mediterranean environment*. Hall, págs. 86-100.
- P.M.Kroonenberg y K.Basford. (1989). "An investigation of multi-attribute genotype response across environments using three-mode principal component analysis". *Euphytica*, nº 44, págs. 109-123.
- C.S. Lin, M.R.Binns y L.P.Lefkovitch. (1986). "Stability analysis: where do we stand?". *Crop Science*, nº 26, págs. 894-900.
- C.S. Lin y G. Butler (1988). "A data base approach for selecting locations for regional trials". *Canadian Journal of plant science*. Nº 68 págs. 651- 659.
- T.M.Little y F.J.Hills (1978). "Agricultural experimentation. Design and analysis". Ed: John Wiley and Sons, Inc, N.York. págs. 630.
- A.Martin y J de D. Luna del castillo. (1999)."Bioestadística para ciencias de la salud". Ed. Norma/ Capitel ediciones. 622 págs.
- M.S.McIntosh (1983)." Analysis of combined experiments". *Agronomy Journal* nº 75. págs. 153-155.
- R.Mead (1988)." The design of experiments. Statistical principles and practical applications". Cambridge University Press. Cambridge. págs. 620.
- A.D. Miller, J.C.Williams y H.F. Robinson (1959). "Variety – environment interactions and their implications on testing methods". *Agronomy Journal* nº 51, págs. 132 – 135.
- MSTAT- C. User´s guide to MSTAT-C. (1988). "A software program for the design, management and analysis of agronomic research experiments". Ed. Michigan State University.

- N.Norman y S.Steiner (1996). "Bioestadística". Ed. Mosby / Doyma libros. pág. 235.
- S.C.Pearce (2005). "The factorial field experiment". Expl.Agric. Cambridge Univ.Press . volume 41, pág.109 – 120.
- S.C.Pearce, G.M.Clarke, G.V.Dyke y R.E. Kempson (1988). "Manual of crop experimentation". Ed. Charles Griffin & Co. Ltd. Oxford University press. N.York, pág. 358.
- R.G.Petersen (1994). "Agricultural field experiments. Design and analysis". Ed. Marcel dekker, N.York, pág. 409.
- H.P. Piepho (1998). "Methods for comparing the yield stability of cropping systems- a overview". Journal of Agron. Crop Science. Nº 180, pág. 193-213.
- M.J.Pinthus. (1973). "Estimate of genotypic value: A proposed method". Euphytica nº 22, pág. 121-123.
- V.T. Prabhakaran y J.P.Jain (1994). " Statistical techniques for studying genotype – environment interactions". Ed: South Asia Publishing. New Delhi.
- W.R.Raun, H.J.Barreto y R.L.Westerman (1993). "Use of stability analysis for long-term soil fertility experiments". Agronomy Journal nº 85, pág. 159-167.
- S.C.Salmon. (1951). "Analysis of variance and longtime variety tests of wheat". Agronomy Journal nº 43, pág. 562-570.
- R.R Sokal y F.J Rohlf. (1977). "Principios y métodos estadísticos en la investigación biológica". Ed. H.Blume. Madrid. Pág 415.
- R.R.Sokal y P.A.H.Sneath (1963). "Principes of numerical taxonomy". Ed. W.H.Freeman and Company.
- F.E.Satterthwaite (1946). "An approximate distribution of estimates of variance components. Biometric Bulletin nº 2, pág. 110-114.
- G.W.Snedecor y W.G.Cochran (1967). "Statistical methods". Ed. Iowa State University press, pág. 703.
- E.F.Schultz (1955). "Rules of thumb for determining expectations of mean squares in analysis of variance". Biometrics 25, pág... 427-430.
- SPSS (1999). Base 10.0 Application guide. SPSS Inc. 233 South Wacker Drive 11 th floor. Chicago .IL. 60606-6307. [http:// www.spss.com](http://www.spss.com)
- G.F.Sprague y W.T.Federer (1951). "A comparison of variance components in corn yield trials". Agronomy Journal nº 43, pág. 535-541.
- R.G.D.Steel y J.H.Torrie (1985). "Bioestadística. Principios y procedimientos". Ed: McGraw-Hill, S.A. pág. 621.

- M.A.Stephen (1974). "EDF statistic for goodness of fit and some comparisons". Journal of American Statistic Assoc. Vol 96. pág. 730-737.
- F.Tejedor (1999). "Análisis de la varianza". Ed: La Muralla / Hespérides. Pág 132.
- G.Upton y I. Cook (2002). "Dictionary of Statistics". Ed: Oxford University Press, N.York. pág. 420.
- N.S.Urquhart (1981). "Anatomy of a study". Journal article 839. New Mexico State Univ. Las Cruces.and HortScience. Vol 16 (5). Pág 621-627
- F.A.Van Eeuwijk, J.B.Denis y M.S. Kang.(1996). "Incorporating additional information on genotype and environment in models for two way GXE tables". En G X E interaction.Ed. CRC. Press. Boca Ratón. Florida.
- M.Vargas, J.Crossa, F Van Eeuwijk, K.D.Sayre y M. Reynolds (2001). "Interpreting treatment x environment interaction in agronomy trials.Agronomy Journal nº 93. pág. 949-960. F.Yates y W.G.Cochran (1938). "The analysis of groups of experiments". Journal of Agric Science. Nº 28 (4) pág. 556 – 580.
- S.K.Yau y J.Hamblin (1994). "Relative yield as a measure of entry performance in variable environments". Crop Science nº 34, pág. 831-817
- F.Yndgaard. (2002). "Variety trials in sugarbeet. Methodology and design". Ed. IIRB.vol 4. pág. 3-12.
- R.W. Zobel, M.J. Wright y H.G. Gausch (1988). "Statistical analysis of a yield trial". Agronomy Journal. Nº 80 pág. 388-393.

