Serie Didáctica Nro. 21

Facultad de Ciencias Forestales
UNIVERSIDAD NACIONAL DE SANTIAGO DEL ESTERO

CÁTEDRA DE
ESTADÍSTICA O. F.

ANÁLISIS DE LA VARIANCIA EN EXPERIMENTOS FACTORIALES

Cátedra de Estadística Forestal

Prof. Titular **Celia G. de BENITEZ**
Prof. Asoc. **Marta G. PECE**
J.T.P. **Margarita J. de GALINDEZ**

Febrero 2010

INDICE

A nuestros alumnos

Estos apuntes han sido preparados para facilitar la comprensión de los experimentos factoriales y su procesamiento mediante la utilización de tres sofwares estadísticos: INFOSTAT, SPSS y SAS. Es nuestro deseo, que sea éste un aporte para el aprendizaje y análisis de estos experimentes. El que presentamos a continuación es una versión actualizada y aumentada de la Serie Didáctica Nº 21 de la Facultad de Ciencias Forestales, titulada "Análisis de la Variancia en Experimentos Factoriales" y que fuera publicada por primera vez en el año 2006.

Ing. Celia Gaillard de Benítez
Dra. Marta Graciela Pece
MSc Margarita Juárez de Galíndez
Febrero de 2010

EXPERIMENTOS FACTORIALES

Introducción

Muchas veces, en la práctica forestal es de interés conocer la influencia de dos o más factores sobre una variable respuesta. Por ejemplo en el estudio de comportamientos de varios clones de álamos podría ser oportuno estudiar simultáneamente la influencia del distanciamiento sobre la variable respuesta, por ejemplo, crecimiento en volumen por hectárea y año. En casos como el mencionado lo adecuado es realizar un experimento factorial: esto significa que cada tratamiento estará definido por la combinación de los factores: clon y distanciamientos probados.

Por lo tanto, se puede definir a los experimentos factoriales como aquellos en los que se comparan o estudian simultáneamente dos o más factores principales, incluyendo los diferentes niveles o modalidades de cada uno.

El Anova en experimentos factoriales constituye una técnica estadística para analizar el efecto de dos ó más variables independientes (factores) sobre una variable respuesta. Hasta el momento se ha estudiado el efecto de un factor sobre la variable respuesta, pero en muchas situaciones prácticas es necesario investigar el efecto de varios factores.

Como en estos experimentos los tratamientos se forman combinando cada nivel de un factor con cada uno de los niveles del otro (o de los otros, si hubiere más de dos), este tipo de experimento permite además evaluar los efectos de las interacciones. Se dice que entre dos factores hay interacción si los efectos de un nivel de un factor dependen de los niveles del otro. Dicho con otras palabras la respuesta de un factor es influenciada en forma diferenciada por los niveles del otro.

La existencia de interacciones indica que los efectos de los factores sobre la respuesta no son aditivos y por tanto no pueden separarse los efectos de los factores.

Conceptos básicos

Factores son características que involucra a dos o más modalidades, variantes o **niveles** diferentes y pueden ser:

a) **Cualitativos**: Son aquellos en los cuales los niveles definen o expresan una modalidad particular de las características del factor; cada nivel tiene un interés intrínseco o independiente de los otros niveles. Estos factores responden a las características de las variables cualitativas.

Ej: Diferentes métodos de riego (manto, surco, aspersión).

Ej: Variedades de un tratamiento cultural: método de poda, de raleo, forma de aplicación de productos terapéuticos, etc.
EJ: Variedad de una determinada especie incluye V_1, V_2 y V_3

Factor = Variedad
Niveles = V_1, V_2, V_3

b) **Cuantitativos**: Son aquellos cuyos valores corresponden a cantidades numéricas, es decir valores inherentes a una variable cuantitativa.
Ej: Supongamos que en una experiencia se prueba fertilizar con diferentes dosis de nitrógeno N: 0-10-20-30 Kg/ha.

Factor = Nitrógeno (N)
Niveles = N_0, N_1, N_2, N_3 que corresponden a las dosis 0-10-20 y 30

Ej: Dosis creciente de un fertilizante medida en kg del elemento por hectárea
Ej: Diferentes dosis de un producto terapéutico
Ej: Concentración de diferentes drogas o reactivos
Ej: Diferentes Tº de aplicación de tratamientos, etc.

Para simbolizar a los factores se ha generalizado el uso de la letra mayúscula vinculada con el nombre del factor y esa letra (que puede ser mayúscula o minúscula) con un subíndice numérico para los niveles.

Ej: Fertilizante nitrogenado con 3 niveles: **Factor Nitrógeno: N; Niveles: n_1, n_2, n_3.**
Si se incluye al control, sin nitrógeno, se acostumbra designarlo con **n_0**.
Si se opta por la letra mayúscula para representar a los niveles del factor:

Ej: Fertilizante nitrogenado con 3 niveles. **Factor Nitrógeno: N; Niveles: N_1, N_2, N_3, N_0.**

También se puede utilizar una letra mayúscula para el factor y otras letras para los niveles que reemplazan los nombres.

Ej.: **FactorVariedades**: **V; Niveles**: **A, B, C.**

En un experimento factorial los **tratamientos surgen de la combinación de los niveles de un factor con los niveles de los otros factores.** Por ejemplo si se combinan 3 dosis de Nitrógeno con 2 variedades (A y B), los **tratamientos** resultantes son 3 x 2 = 6 y se pueden designar como sigue: **N_1A, N_1B, N_2A, N_2B, N_3A, N_3B**.
Los experimentos factoriales se pueden aplicar a diferentes diseños: completamente aleatorizados, bloques, cuadrados latinos.
<u>Los Experimentos Factoriales completos</u> incluyen, por razones de balanceo, a todas las combinaciones posibles entre los distintos niveles del factor involucrado en el experimento.
Supongamos el factorial más simple: 2 factores A y B y cada uno de ellos con dos niveles "a_1 y a_2 ; b_1 y b_2"
Las combinaciones posibles son 4. Si se consideran a las combinaciones como tratamientos, éstos son 4 y se identifican como: a_1b_1, a_1b_2, a_2b_1 , a_2b_2.. En la tabla 1 se aprecia la "estructura" de los tratamientos.

Tabla 1: Tratamientos en un factorial 2 x 2.

		Niveles deA	
		a_1	**a_2**
Niveles de B	**b_1**	a_1 b_1	a_2 b_1
	b_2	a_1 b_2	a_2 b_2

La expresión experimento factorial 2x2: indica que el 1º factor tiene 2 niveles y el 2º. Si se desea efectuar un diseño Experimental para este factorial 2x2 es decir con k = 4 tratamientos o combinaciones y r = 5 repeticiones, se puede utilizar cualquiera de los diseños básicos : Completamente aleatorizado (C.A.), Bloques al Azar (B.AS) o Cuadrado Latino (C.L.).
De igual modo, la expresión experimento factorial 2x3 indica que el 1º factor tiene dos niveles y el 2º tres. En este caso el número de combinaciones (tratamientos) es 2x3 = 6 los que se identifican por la simbología que figura en las celdas de la tabla 2.

Tabla 2: Tratamientos de un factorial 2 x 3.

Niveles del Factor B	**Niveles del Factor A**	
	a_1	**a_2**
b_1	a_1 b_1	a_2 b_1
b_2	a_1 b_2	a_2 b_2
b_3	a_1 b_3	a_2 b_3

A medida que aumenta el número de factores y/o los niveles, aumenta sensiblemente el número de tratamientos y con ello la dificultad de elegir el diseño adecuado, particularmente cuando se trata de un experimento a campo.
Si es un factorial 3x2x3 es decir de 18 tratamientos se descarta el C.L. Si se piensa en B.A. este número de tratamientos es algo elevado a lo que se podría agregar que si la especie con la que se está trabajando exige parcelas grandes entonces el tamaño de los bloques tiene que ser también grande con la consiguiente dificultad de encontrar sectores de terreno homogéneos para ubicar los bloques.
El problema subsiste, si se piensa utilizar C.A., pues el experimento ocupa una superficie grande lo que dificulta la homogeneidad de las unidades experimentales. En esas condiciones quizá se deba pensar en el uso de otros diseños.
Los experimentos factoriales proporcionan en general una información más completa que los experimentos comunes, pues posibilita el estudio de factores principales, las combinaciones de todos los niveles y la interacción de los factores.
En los experimentos factoriales algunos autores hablan de "estructura de tratamientos" indicando con esto que los tratamientos se forman por combinaciones de factores.

Interacción:
Es el efecto recíproco entre 2 o más factores, o la modificación de efecto de un factor por la acción de otro u otros. El estudio de la interacción entre los factores es una de las características importantes en los experimentos factoriales.
La posibilidad de estudios en forma conjunta de dos o más factores con sus correspondientes niveles, hace a los factoriales muy útiles para investigaciones exploratorias y como un paso previo para concentrar posteriormente la atención en los aspectos que puedan ser de mayor interés, de acuerdo a las conclusiones generales que proporcionan estos experimentos.

Ventajas y Desventajas:
- Posibilita el estudio simultáneo de dos o más factores
- Permite estudiar la posible interacción entre los factores intervinientes, y consecuentemente con ello el efecto o comportamiento de cada factor en los diferentes niveles del otro factor.

- Son más eficientes que los experimentos simples, donde se estudia un solo factor. Proporcionan además resultados generales que los hacen útiles en experimentos exploratorios.
- Como se incluyen todas las combinaciones posibles de los diferentes niveles, proporcionan habitualmente un número elevado de grados de libertad para el error experimental, con la consiguiente ventaja que esto significa.

Como contrapartida de lo anterior, a medida que se incrementa el número de factores y niveles se hace mayor el número de tratamientos, aumentando cuando se trata de experimentos a campo la superficie requerida para todo el experimento y en particular para cada repetición. Con todo ello aumenta la dificultad de adaptar el diseño más adecuado al terreno ó al material experimental y se eleva significativamente el costo de cada repetición. Esta circunstancia obliga algunas veces, a recurrir a otro diseño (bloques incompletos) o adoptar sistemas "factoriales" en confundido", cuyo análisis y planeamiento es más dificultoso, además de la pérdida de información sobre algunas interacciones.

A pesar de que no todas las combinaciones entre los diferentes niveles son de interés para el investigador, en estos experimentos no pueden ser excluidos por razones de balanceo que exige el análisis.

<u>Efectos Principales. Efectos simples. Interacción</u>

Analizando los resultados supuestos del Experimento factorial más simple (2x2) vamos a determinar los efectos principales y simples de los factores intervinientes, el efecto de la interacción y la medida de estos efectos.

Supongamos que uno de los factores es fertilización con Nitrógeno (N) con 2 niveles N_0 y N_1 (ausencia y presencia); el otro factor es fertilización con fósforo (P) con los mismos niveles P_0 y P_1. En el cuadro se indican las combinaciones entre los niveles N y P, y entre paréntesis los supuestos rendimientos obtenidos con dichas combinaciones.

Tabla Nº 3: Resultados de un factorial 2 x 2 (cifras entre paréntesis)

Factor P	Factor N		Totales del factor P
	N_0	N_1	
P_0	$N_0P_0(6)$	$N_1P_0(4)$	$P_0 = 10$
P_1	$N_0P_1(2)$	$N_1P_1(10)$	$P_1 = 12$
Totales factor N	$N_0 = 8$	$N_1 = 14$	22

<u>Efecto Principal de N</u>

Es la diferencia entre los tratamientos que tienen nitrógeno (N_1) menos los que no lo tienen (N_0): $N_1 – N_0= 14 – 8 = 6$

<u>Efecto principal de P</u> = Total P_1-Total P_0= 12-10 = 2

Efectos Simples

Es la diferencia entre dos niveles de un factor, a un mismo nivel del otro.

Efecto simple del N a un mismo nivel P

Efecto simple de N en $P_0 = N_1P_0-N_0P_0 = 4-6 = -2$

Efecto simple de N en $P_1 = N_1P_1-N_0P_1 = 10-2 = 8$

De igual forma el efecto simple de P en $N_0 = N_0P_1 -N_0P_0 = 2-6 = - 4$

Efecto simple de P en $N_1 = N_1P_1-N_1P_0 = 10-4 = 6$

Es fácil advertir que la suma de los efectos simples de un elemento es igual a su efecto principal. Comprobémoslo para el N:
Efecto de N en P_0 + Efecto de N en P_1 = $(N_1P_0-N_0P_0) + (N_1P_1-N_0P_1) = -2+8 = 6$ = Efecto Principal

Efecto de la Interacción

Observemos los efectos simples de N a los dos niveles de P

Efecto de N en $P_1 = N_1P_1-N_0P_1 = 10-2 = 8$
Efecto de N en $P_0 = N_1P_0-N_0P_0 = 4-6 = -2$
Se advierte que el comportamiento del N en presencia del fósforo es diferente: en presencia del fósforo (P_1) el rendimiento aumenta cuando se agrega nitrógeno: $N_1P_1-N_1P_0 = 10-2 = 8$
En cambio en ausencia de fósforo (P_0) el rendimiento disminuye cuando se agrega Nitrógeno:$N_1P_0-N_0P_0 = 4-6 = -2$
Esto muestra que, en este ejemplo, hay interacción entre los dos elementos. La medida de esta interacción la da la diferencia entre esos valores, es decir la diferencia del comportamiento del N en presencia de P_1 ($N_1P_1-N_0P_1$) menos la diferencia del N en presencia de P_0 ($N_1P_0-N_0P_0$).
Efecto de la interacción NP = $(N_1P_1-N_0P_1)-(N_1P_0-N_0P_0) = N_1P_1+ N_0P_0-N_1P_0-N_0P_1 = 10+6-4-2 = 10$
Resumiendo: la interacción NP = Efecto N en P_1- Efecto N en P_0
Si se consideran las diferencias en sentido contrario, el resultado debe ser el mismo.
Efecto de esta interacción, que llamaremos PN es el efecto del P en N_1 - efecto de P en N_0 =
= $(N_1P_1 - N_1P_0) - (N_0P_1-N_0P_0) = = N_1P_1+ N_0P_0 - N_1P_0 - N_0P_1 = 10 + 6 – 4 – 2 = 10$
Interacción NP =diferencia de efectos = 8-(-2)=10

Gráficamente

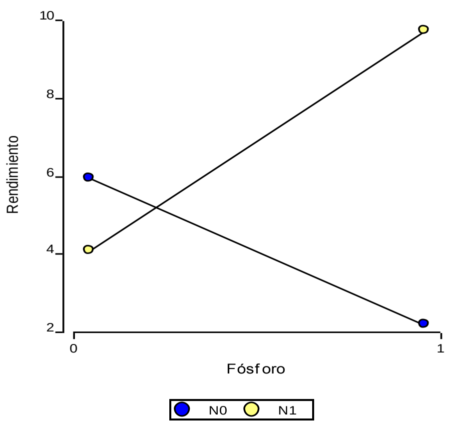


Gráfico 1: Interacción NP = 10 en los datos de la Tabla 3. (P en abcisas)

Se advierte en el gráfico que los segmentos de la recta que representan a N_1 y N_0 se cruzan. Este cambio de dirección es la manifestación de la interacción NP, la medida de la misma es la diferencia de los efectos simples de N.
Observense los efecto simples de P para los dos niveles de P_0: Efecto P en $N_1= N_1P_1-N_1P_0$= 10-4=6 = Efecto P en $N_0= N_0P_1-N_0P_0$=2-6=-4
Se nota aquí que el comportamiento del P en presencia o ausencia de N es diferente: en presencia de nitrógeno(N_1) el rendimiento es mayor o aumenta cuando se agrega fósforo: $N_1P_1-N_1P_0 = 10-4 = 6$. En ausencia de N en cambio el rendimiento es menor cuando se agrega P: $N_0P_1-N_0P_0 = 2-6 = -4$

Se advierte también de esta forma que la existencia de interacción entre los factores y como en el caso anterior, se puede medir con la diferencia de los efectos simples. Llamaremos a esta interacción PN
Efecto de la interacción PN=(N_0P_1-N_1P_0)-(N_0P_1-N_0P_0)=N_1P_1+ N_0P_0-N_1P_0-N_0P_1=10+6-4-2=10
Interacción NP = diferencia de efectos = 6-(-4)=10. Adviértase que la interacción NP =PN =10

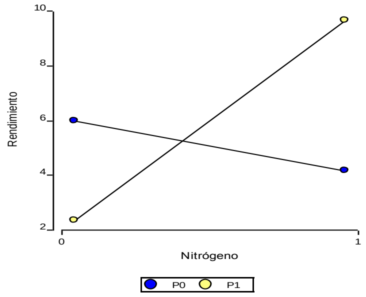


Gráfico 2: Interacción NP = 10 en los datos de la Tabla 3. (N en abcisas)

También en este grafico la interacción se manifiesta con el cruce de segmentos o cambio de dirección de los mismos; lo medido es = a la interacción NP
La medida de la interacción NP o PN se puede hallar restando del efecto combinado de N y P, los efectos de esos elementos aplicados solos o independientemente, es decir, el efecto de cada uno de ellos en ausencia del otro:

Efecto combinado N y P = N_1P_1-N_0P_0=10-6=4
Efecto independiente de N = N_1P_1-N_0P_0=4-6=-2
Efecto independiente de P = N_0P_1-N_0P_0=2-6=-4

Luego, el efecto de la interacción es NP = 4-(-2-4)=10

Ausencia de interacción

Supongamos que los resultados, es decir, los totales de tratamiento hubiera sido el siguiente:

Tabla Nº 4: Otros resultados de un factorial 2 x 2 (cifras entre paréntesis)

Factor P	Factor N		Totales del factor P
	N_0	N_1	
P_0	N_0P_0(2)	N_1P_0(7)	P_0 = 9
P_1	N_0P_1(4)	N_1P_1(9)	P_1 = 13
Totales factor N	N_0 = 8	N_1 = 14	22

Calculemos los efectos simples de N en presencia y ausencia de P, pero con ellos calcular la interacción NP:

Efecto simple de N en P_1= N_1P_1 -- N_0P_1= 9-4 = 5
Efecto simple de N en P_0= N_1P_0 -- N_0P_0= 7-2 = 5

Se advierte que no hay diferencias entre los efectos simples, lo que indica que no hay interacción entre N y P. El efecto de la interacción es, como sabemos, la diferencia de estos efectos simples:
Efecto interacción NP = (N_1P_1 - N_0P_1) - (N_1P_0 - N_0P_0)= N_1P_1 - N_0P_1- N_1P_0 + N_0P_0 = 9 – 4 -7 + 2= 0

Gráficamente se manifiesta de la siguiente manera: Interacción NP = 5-5 = 0

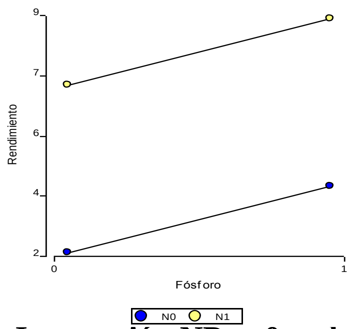


Gráfico 3: Interacción NP = 0 en los datos de la Tabla 4

Se advierte que los segmentos de recta que representan a N_1 y N_0 son paralelos; esta es la evidencia gráfica de la ausencia de interacción.

Otra forma de interacción

No necesariamente los segmentos de recta que representan los efectos simples deben intersectarse cuando existe interacción entre los factores. En las dos situaciones que presentamos a continuación hay interacción entre los dos factores y la manifestación grafica es de otro tipo:

Supongamos las tablas de los tratamientos como el cuadro siguiente:

Tabla Nº 5: Otros resultados de un factorial 2 x 2 (cifras entre paréntesis)

Factor P	Factor N		Totales del factor P
	N_0	N_1	
P_0	$N_0P_0(5)$	$N_1P_0(6)$	$P_0 = 9$
P_1	$N_0P_1(2)$	$N_1P_1(10)$	$P_1 = 13$
Totales factor N	$N_0 = 7$	$N_1 = 16$	23

Los efectos simples de N en presencia o ausencia de P son:

Efecto simple de N en $P_1 = N_1P_1 - N_0P_1 = 10 - 2 = 8$

Efecto simple de N en $P_0 = N_1P_0 - N_0P_0 = 6 - 5 = 1$

La diferencia entre estos efectos simple es la medida de la interacción NP = 8-1 = 7

Efecto de la interacción = $(N_1P_1 - N_0P_1) - (N_1P_0 - N_0P_0) = N_1P_1 - N_0P_1 - N_1P_0 + N_0P_0 = 10 - 2 - 6 + 5 = 7$

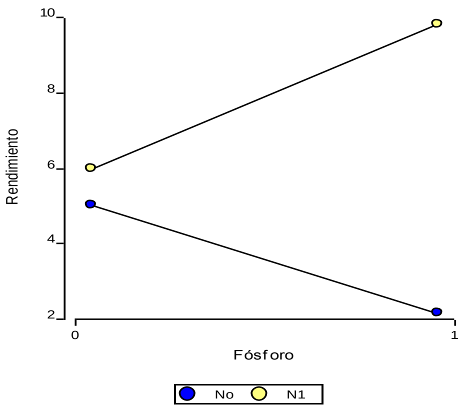


Gráfico 4: Interacción NP = 7 en los datos de la Tabla 5

En el gráfico se nota que cuando se agrega nitrógeno(N_1) el rendimiento aumenta en presencia de (P_1), ocurre lo contrario cuando no se agrega nitrógeno(N_0), el rendimiento disminuye en presencia de fósforo(P).Los segmentos que representa a N_1 y N_0 no se cortan pero sus direcciones son diferentes; la pendiente de N_1 es positiva, en cambio la de N_0 es negativo.

Supongamos ahora los totales de tratamiento que se incluyen en el cuadro siguiente, ver tabla 6:

Tabla Nº 6: Otros resultados de un factorial 2 x 2 (cifras entre paréntesis)

Factor P	Factor N		Totales del factor P
	N_0	N_1	
P_0	$N_0P_0(2)$	$N_1P_0(4)$	$P_0 = 6$
P_1	$N_0P_1(2)$	$N_1P_1(10)$	$P_1 = 12$
Totales factor N	$N_0 = 4$	$N_1 = 14$	18

Efecto simple de N en P_1 = $N_1P_1 - N_0P_1$= 10-2 = 8
Efecto simple de N en P_0 = $N_1P_0 - N_0P_0$ = 4 - 2 = 2
La diferencia entre estos efectos simples es la medida de la interacción P = 8 -2 = 6
Efecto de la interacción NP = $(N_1P_1 - N_0P_1)- (N_1P_0 - N_0P_0) = N_1P_1 - N_0P_1- N_1P_0 + N_0P_0$ = 10 – 2 – 4 + 2 = 6 Interacción NP = 8-2 = 6

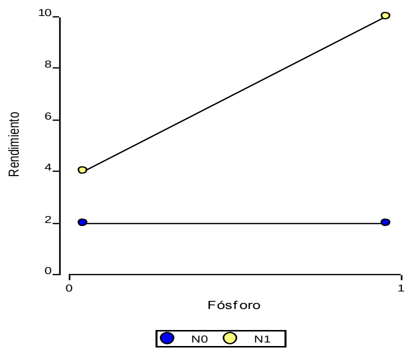


Gráfico 5: Interacción NP = 6 en los datos de la Tabla 6

En el grafico anterior se advierte que cuando se agrega nitrógeno (N_1) el rendimiento aumenta en presencia de fósforo (P_1), no ocurre lo mismo cuando no se agrega nitrógeno (N_0), el rendimiento permanece constante en presencia de fósforo (P_1). Los segmentos que representan a N_1 y N_0 no se cortan pero sus direcciones son diferentes.

<u>Modelo estadístico</u>: Las observaciones pueden describirse mediante un modelo estadístico lineal.

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

i = 1,2......,a
j = 1,2......,b
k =1,2…..,n

donde

μ_i = media general
τ_i = efecto del iésimo nivel del factor A
β_j = efecto del j-ésimo nivel del factor B
$(\tau\beta)_{ij}$ = efecto de la interacción entre τ_i y β_j
ε_{ijk} = Componente del error aleatorio

Los supuestos sobre la componente aleatoria ε_{ijk} son: los errores son independientes e idénticamente distribuidos: $\varepsilon_{ijk} \sim N(0,\sigma)$. Esta simbología indica que la distribución debe ser normal con media cero y variancia común.
Ambos factores son fijos y los efectos de los tratamientos se definen como desviaciones de la media general y deben cumplir la restricción:

$$\sum \tau_i = 0 \qquad \sum \beta_j = 0$$

Se supone que los efectos de la interacción son fijos y se definen como: $\sum (\tau\beta)_{ij} = 0$

Hay un total de "***abn***" observaciones: porque los tratamientos son "***ab***" y se realizan "***n***" réplicas.
Tanto el factor A como el factor B tienen el mismo interés. La finalidad consiste en probar hipótesis de mismo efecto de tratamientos en renglón
Ho= $\tau_1 = \tau_2 = \ldots = \tau_a = 0$
H_1 = al menos un $\tau i \neq 0$

Y de la igualdad de los tratamientos de columna
H_0: $\beta_1 = \beta_2 = \ldots = \beta_b = 0$
H_1: al menos un $\beta_i \neq 0$

También es importante probar si los tratamientos $(\tau\beta)_{ij}$ de las celdas interaccionan
$H_0 = (\tau\beta)_{ij} = 0$ para todo *i,j*
H_1 al menos uno $(\tau\beta)_{ij} \neq 0$

Sea el siguiente ejemplo donde se presentan ***a*** niveles de un factor A y ***b*** niveles de un factor B dispuestos en un diseño factorial, cada repetición del experimento contiene todas las combinaciones "***ab***" (es decir hay ***a x b*** tratamientos) con "***n***" réplicas en cada uno de los tratamientos.

Ejemplo: Del libro de Douglas Montgomery "Diseño y Análisis de Experimentos", 1993.

Un Ingeniero diseñó una batería para su uso en un dispositivo que será sometido a ciertas variaciones extremas de temperatura. El único parámetro de diseño que el puede seleccionar es la cubierta de la batería y tiene 3 alternativas (factor A, ***a*** = 3). Cuando el dispositivo se manufactura y se envía a campo, el Ingeniero no tiene control sobre los extremos de la Tº a que será expuesto el dispositivo y sabe por experiencia que es probable que la Tº influya en la duración efectiva de la batería. Sin embargo es posible controlar la Tº en el laboratorio de desarrollo de productos para los fines del ensayo.
El Ing. decide probar los 3 niveles de cubierta a tres niveles de Tº (15º, 70º y 125º F) consistentes en el entorno de la cubierta final (factor B, ***b*** = 3).
Se prueban 4 baterías (***n*** = 4) en cada combinación (material de cubierta, Tº), y las 3 x 3 x 4 = 36 pruebas se ejecutan en orden determinado al azar.
El Ing. quiere contestar las siguientes preguntas:

1)- ¿Qué efecto tienen el tipo de material y la Tº sobre la duración de la batería?
2)- ¿Existe un material que de como resultado una duración uniforme sin importar la Tº?
Esta pregunta es importante, o por el contrario: ¿Existe la posibilidad de hallar un material que sea muy afectado por la Tº?.
De ser así el Ing. ¿puede hacer que la batería sea robusta a la variación de Tº en el campo?

Este es un diseño con dos factores
La observación $\boldsymbol{Y_{ijk}}$ se encuentra en el i- iésimo nivel del factor A(i =1,…,a)
j- ésimo nivel de factor B (i =1,…,b)
k-iésimo rep. de la combinaciones(ij) k = (1,…,n)

Tabla 7: duración en horas de baterías con 3 cubiertas distintas y sometidas a 3 temperaturas de trabajo

		Tº Factor B (grados F)						
		15	***Yij•***	**70**	***Yij•***	**125**	***Yij•***	**Y*i*••**
Factor A: tipo de material	**1**	130-155 70-180	535	34-40 80-75	229	20-70 85-58	233	997
	2	150-188 159-106	603	136-122 106-115	479	25-70 58-45	198	1280
	3	138-110 168-160	576	174-120 150-139	583	96-104 82-60	342	1501
	$\bar{Y}_{\bullet j \bullet}$	1714		1291		773		**Y••• =** =3778

$Y_{\bullet\bullet\bullet}$ es el total general de todas las observaciones
$Y_{i\bullet\bullet}$ es el total de las observaciones bajo el i-ésimo nivel del factor A
$Y_{\bullet j \bullet}$ es el total delas observaciones bajo el j-ésimo nivel del factor B y
$Y_{ij\bullet}$ es el total de las observaciones de la ij-ésima celda,

Se define $\bar{Y}_{i\bullet\bullet}$ $\bar{Y}_{\bullet j \bullet}$ $\bar{Y}_{ij\bullet}$ $\bar{Y}_{\bullet\bullet\bullet}$ promedio de renglón, columna , celda y general

$$Y_{i\bullet\bullet} = \sum_{j}^{b}\sum_{k}^{n} Y_{ijk} \qquad \bar{Y}_{i\bullet\bullet} = \frac{Y_{i\bullet\bullet}}{bn} \qquad i=1,\dots, a$$

$$Y_{\bullet j \bullet} = \sum_{i}^{a}\sum_{k}^{n} Y_{ijk} \qquad \bar{Y}_{\bullet j \bullet} = \frac{Y_{\bullet j \bullet}}{an} \qquad i= 1,\dots, b$$

$$Y_{ij\bullet} = \sum_{k=1}^{n} Y_{ijk} \qquad \bar{Y}_{ij\bullet} = \frac{Y_{ij\bullet}}{n} \qquad i=1,\dots, a\ ;\ j= 1,\dots, b$$

$$Y_{\bullet\bullet\bullet} = \sum_{i=1}^{i=a}\sum_{j=1}^{j=b}\sum_{k=1}^{k=n} Y_{ijk} \qquad \bar{Y}_{\bullet\bullet\bullet} = \frac{Y_{\bullet\bullet\bullet}}{abn} \qquad i =1,\dots,a;\quad j = 1,\dots,b;\quad k = 1,\dots,n$$

Descomposición

$$\mathbf{SCT = SC_A + SC_B + SC_{AB} + SC_{EE}}$$

Los Grados de Libertad Asociados a cada suma de cuadrados se presentan en la tabla 8.

Tabla 8: descomposición de los grados de libertad en un experimento factorial

Causas	G de L
A	a-1
B	b-1
Interacción AB	(a-1)(b-1)
ERROR	ab(n-1)
TOTAL	abn-1

Como se calculan las sumas de cuadrados:

$$\Rightarrow \ \boldsymbol{SCT} = \sum\sum\sum Y_{ijk}^2 - \frac{Y_{\bullet\bullet\bullet}^2}{abn} \qquad\qquad \mathbf{TC} = \frac{Y_{\bullet\bullet\bullet}^2}{abn}$$

Suma de Cuadrados de efectos principales

$$\Rightarrow \ \boldsymbol{SC}_{\mathbf{A}} = \sum \frac{Y_{i\bullet\bullet}^2}{bn} - TC$$

$$\Rightarrow \ \mathbf{SC_B} = \sum \frac{Y_{\bullet j\bullet}^2}{an} - TC$$

es conveniente obtener SC_{AB} en 2 etapas. Primero se calcula la suma de cuadrados de los totales de cada uno de los ***ab*** tratamientos (combinaciones de los dos factores: ver celdas de la tabla 7), conocido como la suma de cuadrados debido a los "subtotales" o también "combinaciones".

$$\mathbf{SC_{Subtotal}} = \mathbf{SC_{Combinaciones}} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{Y}_{ij\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{Y_{ij\bullet}^2}{n} - TC$$

Esta suma de cuadrados contiene a la SC_A y SC_B en la segunda etapa.

$$\Rightarrow \ \mathbf{SC_{AB} = SC_{Subtotal} - SC_A - SC_B}$$

$$\Rightarrow \ \mathbf{SC_E = SCT - SC_{AB} - SC_A - SC_B}$$

$$\Rightarrow \ \mathbf{SC_E = SCT - SC_{Subtotal}}$$

Ejemplo:

$$Y_{\bullet\bullet\bullet} = 3.778$$

$$\mathbf{SCTotal} = \left(130^2 + 155^2 + \ldots\ldots + 82^2 + 60^2\right) - \frac{\left(3778\right)^2}{3\times3\times4} = 77.351{,}89$$

$$\mathbf{SC_A = SC_{Material}} = \frac{997^2 + 1280^2 + 1501^2}{3\times4} - \frac{\left(3778\right)^2}{3\times3\times4} = 10.637{,}39$$

$$\mathbf{SC_B = SC_{T^o}} = \frac{1714^2 + 1291^2 + 773^2}{3\times4} - \frac{\left(3778\right)^2}{3\times3\times4} = 37.020{,}39$$

$$\mathbf{SC_{Combinación}} = \frac{535^2 + 603^2 + \ldots + 342^2}{4} - \frac{3778^2}{3\times3\times4} = 56.979{,}39$$

La suma de cuadrados de las combinaciones incluye a las sumas de cuadrados de A, de B y de AB.

$$\mathbf{SC_{Combinación} = SCA_{Material} + SCB_{T^o} + SC_{(AxB)}}$$

Por lo tanto:

$$\mathbf{SC_{\ Interacción} = SC_{Combinación} - SCA_{Material} - SCB_{T^o} =}$$
$$\mathbf{= 56.979,39 - 10.637,39 - 37.020,39 = 9.321,61}$$

Por diferencia se obtiene la suma de cuadrados del error:

$$\mathbf{SC_{EE} = 77.351,89 - 10.637,39 - 37.020,39 - 9.321,61 = 20.372,50}$$

Los grados de libertad son:

- Grados de Libertad de A = a – 1 = 3 – 1 = 2
- Grados de Libertad de B = b – 1 = 3 – 1 = 2
- Grados de Libertad de la Interacción AB = G. de Libertad de cada celda (ab-1) menos los g. de libertad de los 2 efectos principales A y B.
  ab-1- (a-1) - (b-1) = (a-1)(b-1) = 2 x 2 = 4
- Dentro de cada una de las celdas hay ***n-1*** grados de Libertad entre las ***n*** réplicas, por lo tanto hay

  ab(n-1) G. libertad error = 3 x 3 x 3 = 27

Las esperanzas de los cuadrados medios son:

E(CMA)= $E\left(\frac{SC_A}{a-1}\right) = \sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$ ***;*** ***E(CMB)=*** $E\left(\frac{SC_B}{b-1}\right) = \sigma^2 + \frac{an\sum \beta_i^2}{b-1}$

E(CMAB)= $E\left(\frac{SC_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n\sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$ ***;*** ***E(CME)=*** $E\left(\frac{SCE}{ab(n-1)}\right) = \sigma^2$

Si las hipótesis nulas, que consisten en proponer que no hay efecto de tratamientos de renglón, columna e interacción son verdaderas, entonces CMA, CMB, CMAB y CME son estimadores σ^2.
Sin embargo, si por ej. existen diferencia entre materiales (filas), entonces CMA será mayor CME.
En forma similar, si hay efecto de tratamiento de columna (Tº) o de la interacción, la media de los cuadrados correspondientes serán mayores que CME.
Por lo tanto, para probar el significado de ambos efectos principales así como su interacción, simplemente deben dividirse las medias de los cuadrados correspondientes entre la media de Cuadrados del Error.
Si el modelo es adecuado y ε_{ij} son independientes, con distribuciones normales y con variancias constantes $\sigma^2 \Rightarrow$ las razones que se presentan más abajo se distribuyen como F con los grados de libertad que se indican.

$$\frac{CMA}{CME} \approx F\left(a-1; ab(n-1)\right)$$
$$\frac{CMB}{CME} \approx F\left(b-1; ab(n-1)\right)$$
$$\frac{CMAB}{CME} \approx F\left((a-1)(b-1); ab(n-1)\right)$$

Tabla 9: Análisis de la variancia de los datos de la tabla 7.

F. de Variación	SC	G de Lib.	CM	F	F_{tabla}
Total	77.351,89	35			
Tipo de	10.637,39	2	5.318,69	7,05*	$F_{(2,27)\ 0,.05}$=3.35
Material	37.020,39	2	18.510,19	24,53*	$F_{(4,27)\ 0,.05}$=2.73
Temperatura	9.321,61	4	2.330,40	3.09*	
Interacción	20.372,50	27	754,54		
Error					

Como auxiliar en la interpretación de los resultados de este experimento resulto útil la construcción de un gráfico de las respuestas promedio (Gráfico 6). El hecho de que las rectas no sean paralelas indica una interacción significativa.

Tabla 10: Medias por tratamiento de los datos de la tabla 7

Y	Tº		
Material	**15º**	**70º**	**125º**
1	535/4=133,75	229/4= 57,25	233/4=58,25
2	603/4=150,75	479/4=119,75	198/4=49,5
3	576/4=144	583/4=145,75	342/4=85,5

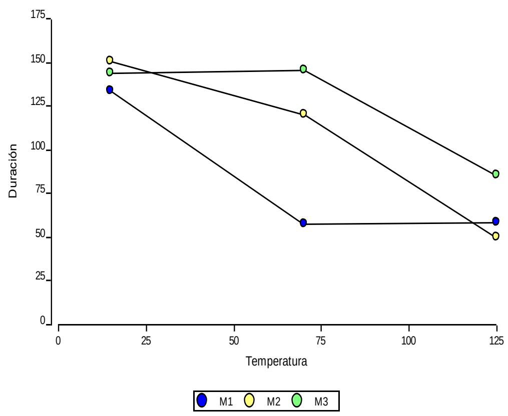


Gráfico 6. Promedios de duración de los diferentes materiales en función de la Tº

Comparaciones Múltiples

Si el análisis de las Variancias indica que hay diferencia en el nivel medio de renglones o columna, resulta de interés llevar a cabo comparaciones medias individuales de renglones o columna para descubrir diferencias específicas. Pero cuando la interacción es significativa, las diferencias en los medias de un factor (por ej A) pueden ser ocultadas por la interacción AB. El enfoque consiste en fijar el factor B en un nivel específico, y aplicar la prueba de intervalos múltiples de Duncan a las medias del factor A en ese nivel. Por ejemplo si se desea detectar diferencias en el nivel medio de los tres tipos de material, como la interacción es significativa, las comparaciones deben realizase en un solo nivel de la temperatura por separado. Por ejemplo, el gráfico 6 muestra que las diferencias son más marcadas en el nivel 2 (70ºC). Los promedios de los tres tipos de material en el nivel de Tº 70º, organizados en orden ascendente son:

$\overline{Y}_{12} = 57{,}25$ **Material tipo 1**

$\overline{Y}_{22} = 119{,}75$ **Material tipo 2**

$\overline{Y}_{32} = 145{,}75$ **Material tipo 3**

Se supone que el mejor estimador de la Variancia del error es MSE obtenido de la tabla de Análisis de la Variancia. Además, se utiliza la suposición de que la variancia del Error Experimental es la misma en todas las combinaciones de tratamientos.

El error estándar de estos promedios o medias de tratamientos es:

$$S\overline{Y}_{12} = \sqrt{\frac{CME}{n}} = \sqrt{\frac{754,54}{4}} = 13,73$$

Ya que cada promedio se calcula mediante n = 4 observaciones

Duncan $\begin{cases} \Delta_{0,05}\,(2,27)=2,91 \\ \Delta_{0,05}\,(3,27)=3,06 \end{cases}$

Intervalos mínimos significativos entre medias de materiales para un nivel de Tº.

D_2= 2,91 x 13,73= 39,97

D_3= 3,06 x 13,73 = 42,02

Para la Tº = 70º, las diferencias observadas entre los promedios de materiales son:

Diferencia Material 3 – Material 1 = 145,75 – 57,25 = 88,5 > 42,02 (D_3)
Diferencia Material 3 – Material 2 = 145,75 – 119,75 = 26 < 39,97 (D_2)
Diferencia Material 2 – Material 1 = 119,75 – 57,25 = 62,5 > 42,02 (D_3)

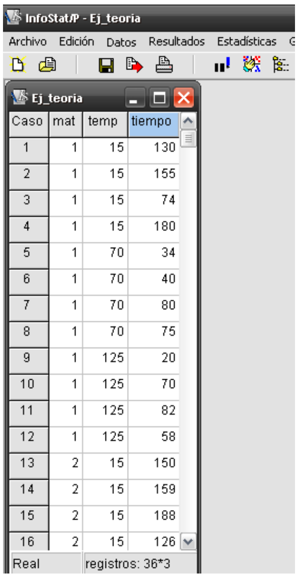
El análisis indica que al nivel de temperatura de 70ºC, la duración media de las baterías con cubiertas de los materiales 2 y 3 es la misma y mientras que a esa temperatura la duración media con el material 1 es significativamente menor que la de los elementos con cubierta de materiales 2 y 3.
Cuando la interacción es significativa, el investigador puede comparar las medias de todas las celdas para determinar en cuales hay diferencias significativas. En este análisis las diferencias entre las celdas incluyen tanto los efectos principales como el efecto de interacción. Para completar el análisis de este ejemplo se debe efectuar el análisis que se hizo para 70º , para los otros niveles de Tº.

RESOLUCIÓN DEL EJEMPLO ANTERIOR EMPLEANDO DIFERENTES SOFTWARES

DESARROLLO CON INFOSTAT (V.2008).

Creación del archivo de datos:

Se introducen en una primera columna, los tipos de material: 1, 2, 3 ; en una segunda columna las temperaturas: 15, 70, 125, y en una tercera columna los valores de la variable, obteniéndose la siguiente formato de archivo (Figura 1).



Caso	mat	temp	tiempo
1	1	15	130
2	1	15	155
3	1	15	74
4	1	15	180
5	1	70	34
6	1	70	40
7	1	70	80
8	1	70	75
9	1	125	20
10	1	125	70
11	1	125	82
12	1	125	58
13	2	15	150
14	2	15	159
15	2	15	188
16	2	15	126

Figura 1: **Creación del archivo de datos para análisis del experimento factorial.**

Luego se crea una nueva columna que estará formada por la combinación de cada nivel de material por cada nivel de temperatura. Esto se puede hacer con el menú de INFOSTAT seleccionar: **datos, cruzar categorías** y aparece este cuadro de diálogo. Luego se marca materia y temperatura y con la flecha se pasa al cuadro de la derecha, criterios de clasificación, como lo indica la siguiente figura (Figura 2)

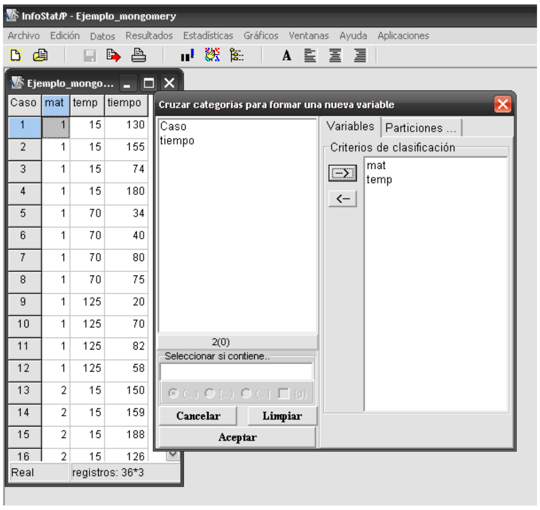


Figura 2: **Cruce de categorías con los niveles de material y temperatura.**

Al aceptar queda creada la columna mat_temp. Se puede cambiar ese nombre, haciendo doble clic en el mismo y colocando trat Figuras 3 y 4.

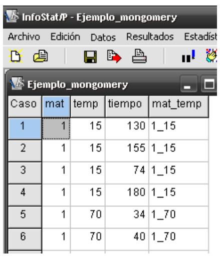


Figura 3: **Variable creada con Cruce de categorías**

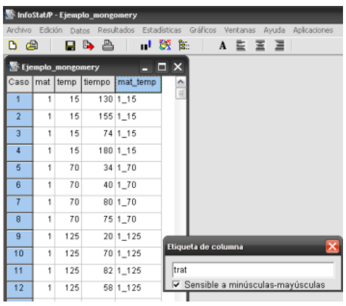


Figura 4: **Cambio de nombre**

De esta forma el archivo queda con 4 variables: material, temperatura, tiempo y tratamiento. Una vez completo el archivo se realiza el Anova. Para ello se procede de la siguiente manera seleccionando el menú**: Estadística_Análisis de la varianza. A**parece la pantalla que se encuentra en la Figura 5 y se marca y pasa con las flechas correspondientes:

- La variable tiempo al cuadro variable dependiente y
- mat y temp al cuadro variables de clasificación

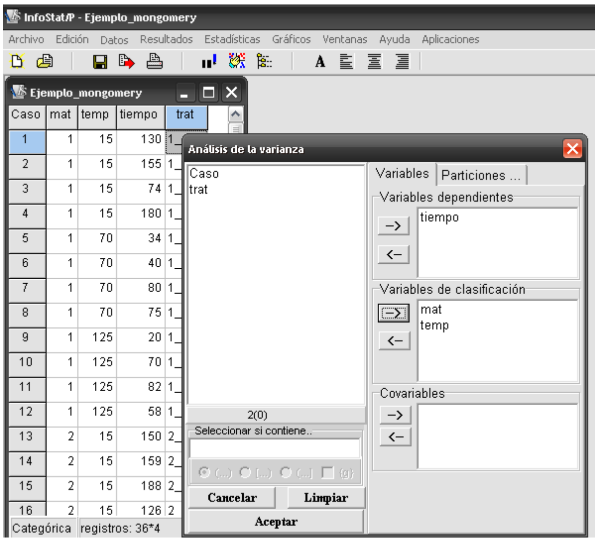


Figura 5: Selección de las variables y de los factores.

Se acepta y se muestra un nuevo cuadro (Figura 6).

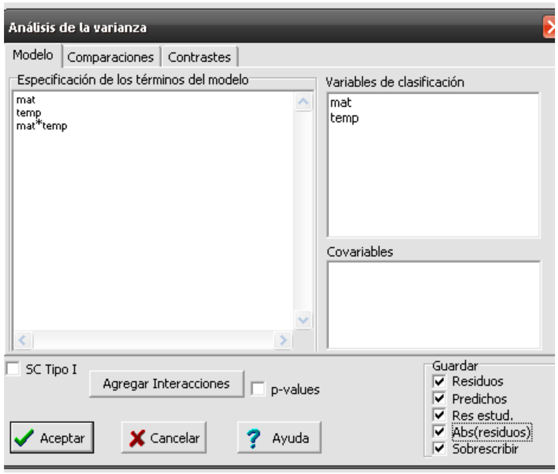
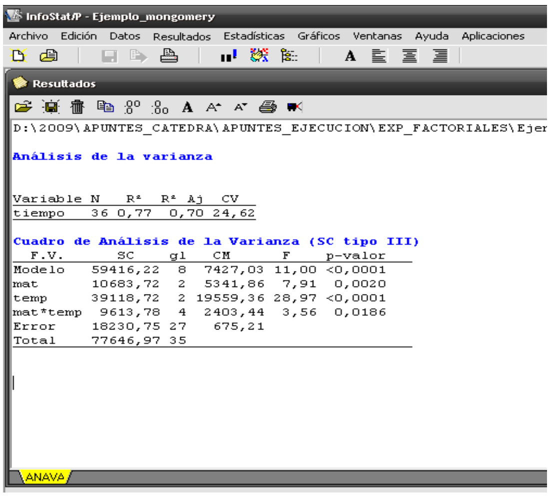


Figura 6. Indicaciones para realizar el ANOVA del experimento factorial y grabar los residuos para prueba de supuestos.

Para incorporar las interacciones al modelo, se hace clic en **agregar interacciones** y en el cuadro de especificación del modelo, que ya tiene incorporado **mat y temp** aparece automáticamente una nueva línea **mat*temp** (que es la interacción de los factores).En el cuadro **guardar** se marcan mediante un clic: Residuos, Predichos, Res estud.; Abs(residuos) y Sobrescribir ya esta marcado por default. Figura 6.

Al **aceptar** aparece por un lado una hoja de resultados (Salida 1) en la que se encuentran los resultados del Análisis de la _Variancia del experimento factorial.


D:\2009\APUNTES_CATEDRA\APUNTES_EJECUCION\EXP_FACTORIALES\Ejer

Análisis de la varianza

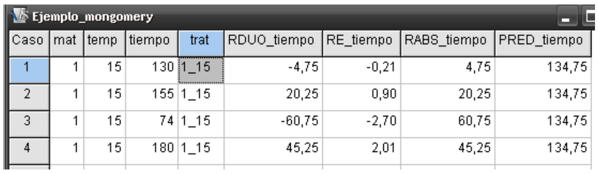
Variable	N	R²	R² Aj	CV
tiempo	36	0,77	0,70	24,62

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	59416,22	8	7427,03	11,00	<0,0001
mat	10683,72	2	5341,86	7,91	0,0020
temp	39118,72	2	19559,36	28,97	<0,0001
mat*temp	9613,78	4	2403,44	3,56	0,0186
Error	18230,75	27	675,21		
Total	77646,97	35			

Salida 1: **Resultados del ANOVA**

Si se observa la tabla de datos se ve que se crearon nuevas columnas (Figura 7) correspondiente a las variables, residuos y predichos (RDUO, RABS, RE)



Caso	mat	temp	tiempo	trat	RDUO_tiempo	RE_tiempo	RABS_tiempo	PRED_tiempo
1	1	15	130	1_15	-4,75	-0,21	4,75	134,75
2	1	15	155	1_15	20,25	0,90	20,25	134,75
3	1	15	74	1_15	-60,75	-2,70	60,75	134,75
4	1	15	180	1_15	45,25	2,01	45,25	134,75

Figura 7: Archivo de datos luego de la creación de variables con la opción guardar

Con estas nuevas variables se realizan las prueba de los supuestos necesarios para que el ANOVA sea válido:

<u>Prueba de Normalidad</u> de Shapiro Wilks. En esta prueba se trabaja con los residuos del ANOVA. Para realizarla se utiliza el menú **Estadísticas_Inferencia basada en una muestra, Prueba de Normalidad (Shapiro-Wilks modificado) (**Figura **8).**

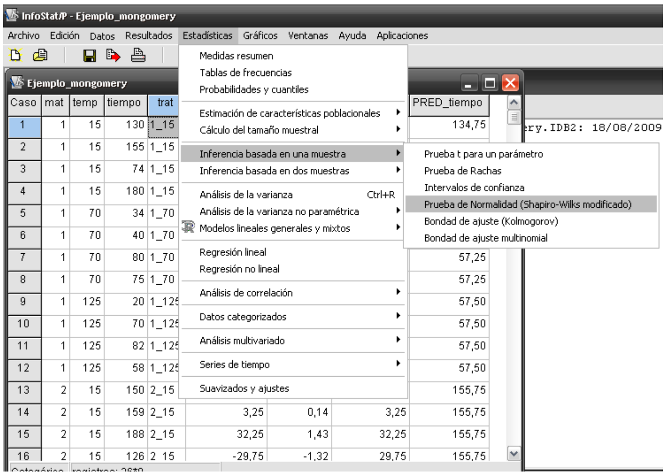


Figura 8: Selección de la prueba de normalidad

Al aceptar aparece el siguiente cuadro (Figura 9). Para seleccionar la variable con la que se realizará la prueba de normalidad, se marca RDUO_tiempo y se la pasa al cuadro de variables.

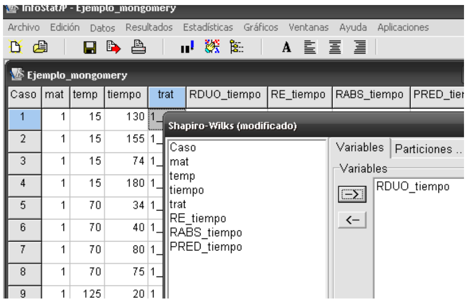
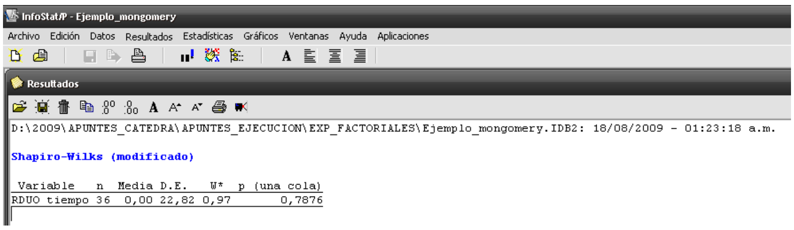


Figura 9. Selección de los residuos para la prueba de normalidad

Al aceptar obtenemos una hoja con los resultados de la prueba (Salida 2)
Las hipótesis puestas en juego son:

$H_0 : \varepsilon_{ij}$ Tienen distribución normal

$H_1 : \varepsilon_{ij}$ No tienen distribución normal



D:\2009\APUNTES_CATEDRA\APUNTES_EJECUCION\EXP_FACTORIALES\Ejemplo_mongomery.IDB2: 18/08/2009 - 01:23:18 a.m.

Shapiro-Wilks (modificado)

Variable	n	Media	D.E.	W*	p (una cola)
RDUO tiempo	36	0,00	22,82	0,97	0,7876

Salida 2. Resultados del test de normalidad de residuos.

Con un valor alto (p = 0,7876) no se rechaza la hipótesis nula. No hay evidencias suficientes para decir que los residuos no tienen distribución normal. También se puede realizar una prueba gráfica con un **qq-plot** para la variable RDUO_tiempo.

Para ello se selecciona la opción **Gráficos, qq-plot**, se acepta, se marca y se pasa la variable RDUO_tiempo al cuadro de variable y luego se selecciona la Distribución Normal del siguiente cuadro (Figura 10).

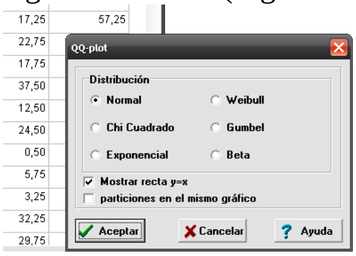
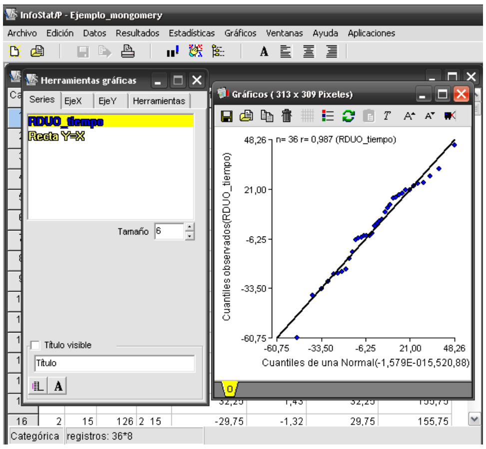


Figura 10. Selección de la distribución normal para la realización del qq-plot

Al aceptar se obtiene el siguiente gráfico (Salida 3). La aproximación de los puntos a la recta sugiere que la distribución de los residuos es normal.



Salida 3 . Qq-plot de los residuos del ANOVA

<u>**Prueba de Homogeneidad**</u> de Levene. Esta prueba consiste en un ANOVA con una causa de variación (Tratamientos) de los **valores absolutos de los residuos.** Seleccionar **Estadísticas-Análisis de la varianza** considerando como variable dependiente a RABS_tiempo y como variable de clasificación Tratamientos (Figura 11).

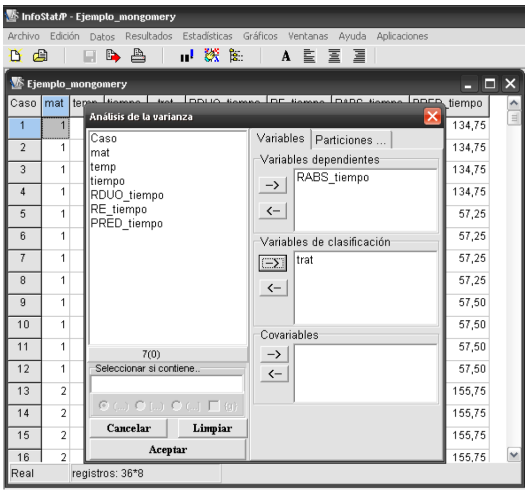
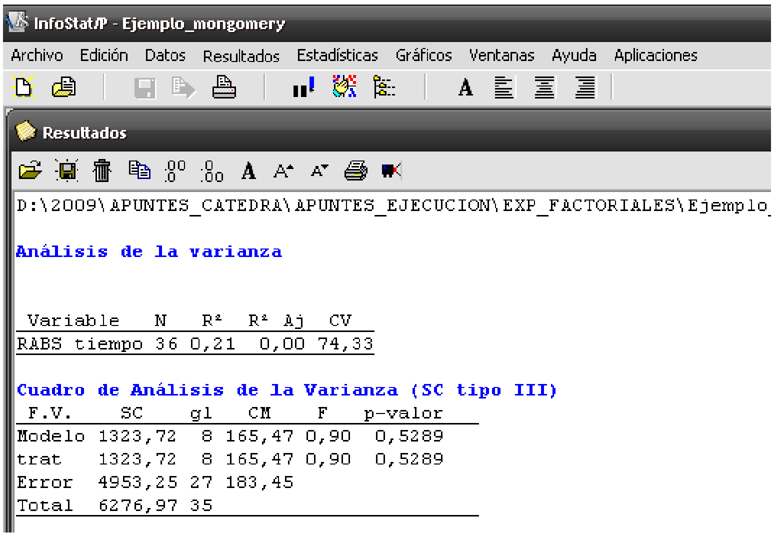


Figura 11. Selección de variables para la prueba de homogeneidad de variancias de Levene.

Al aceptar se presentan los resultados de la prueba en la Salida 4.



D:\2009\APUNTES_CATEDRA\APUNTES_EJECUCION\EXP_FACTORIALES\Ejemplo_

Análisis de la varianza

Variable	N	R²	R² Aj	CV
RABS tiempo	36	0,21	0,00	74,33

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	1323,72	8	165,47	0,90	0,5289
trat	1323,72	8	165,47	0,90	0,5289
Error	4953,25	27	183,45		
Total	6276,97	35			

Salida 4. Resultados de la prueba de Levene

Las hipótesis puestas en juego son:

$$H_0 : \sigma_i^2 = \sigma^2 \, (hom\,eneidad\,)$$

$$H_1 : al\ menos\ una\ \sigma_i^2 difiere\ de\ otra$$

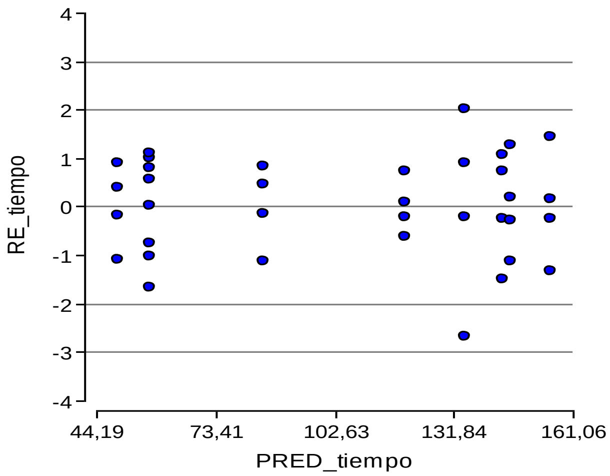
Con un valor de p = 0,5289, no se rechaza la hipótesis nula. No hay evidencias suficientes para decir que los residuos no son homogéneos.

También se podría inspeccionar gráficamente la homogeneidad de variancias mediante un **gráfico de dispersión**, colocando en el eje de las ordenadas los residuos estudentizados (RE_tiempo) y en el eje de las abscisas los valores estimados de la variable (PRED_tiempo). La gráfica no debe manifestar ninguna tendencia y ni los residuos estudentizados ser mayores o menores a 3 y -3.

Seleccionar Gráficos_Dispersión: Marcar RE_tiempo y con flecha pasar la variable al eje Y; marcar PRED_tiempo y con flecha pasar la variable al Eje X (Figura 12).

Figura 12. Selección de variables para el gráfico de dispersión de los residuos (RE_tiempo) en función de los valores estimados o predichos (PRED_tiempo).

Al aceptar se obtiene el gráfico de la salida 5, en el que se observa una distribución homogénea de los residuos, ya que los mismos no muestran patrón alguno de dispersión con valores comprendidos entre +3 y -3.



Salida 5: Residuos estudentizados en función de predichos.

Una vez realizadas las pruebas de cumplimiento de los supuestos, se pueden considerar válidos los resultados del análisis de la varianza que se presentaron en la salida 1. En este ejemplo, se examinará primero la interacción para saber si ésta es estadísticamente significativa.

1) Prueba de hipótesis sobre el efecto de la interacción de los factores: Material y Temperatura

Es importante probar si los factores de renglón y columna interaccionan $(\tau\beta)_{ij}$

$$\mathbf{H_0: (\tau\beta)_{ij} = 0 \quad para\ todo\ i,j}$$

$$\mathbf{H_1: al\ menos\ uno\ (\tau\beta)_{ij} \neq 0}$$

Con un valor p tan bajo (0,0186) se rechaza la hipótesis nula y se concluye que el efecto de la interacción material-temperatura es diferente de cero, es decir que el tiempo promedio de duración depende de la combinación Temperatura-Material, por lo que es conveniente realizar pruebas de diferencias entre temperaturas para un mismo material. Para ello se utilizará la prueba de Diferencia Límite Significativa. Ésta se realiza efectuando el ANOVA de la manera ya indicada, seleccionando en la solapa "comparaciones" e indicando que la prueba se efectúe entre medias de Tratamientos (Figura 13).

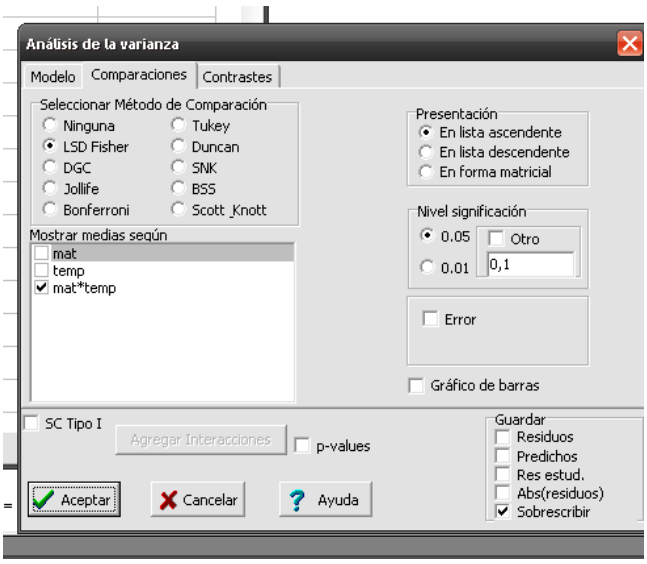
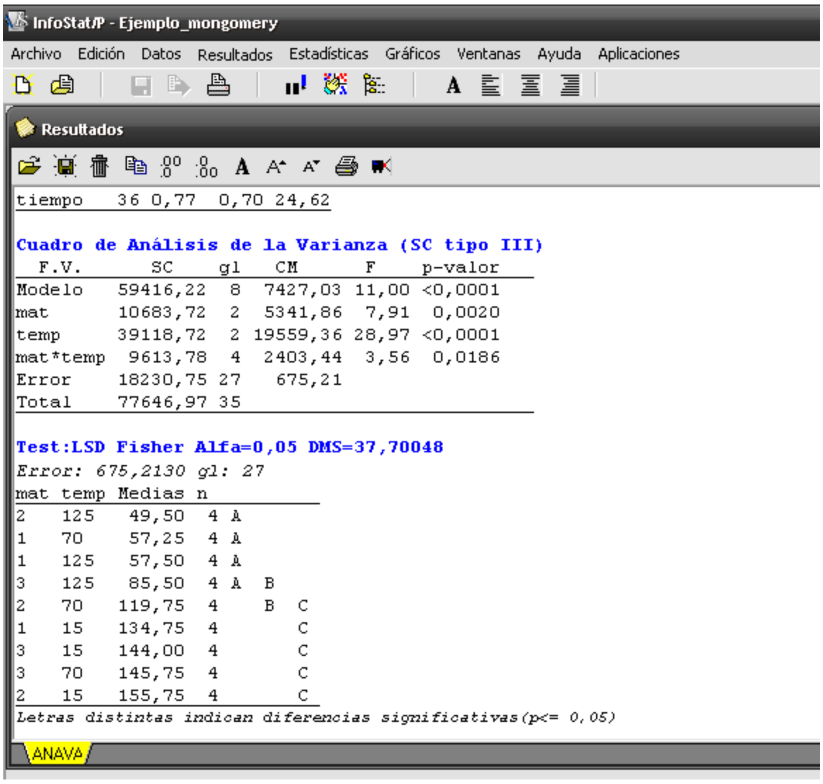


Figura 13: Selección de prueba de medias en caso de interacción significativa

Al aceptar se obtiene la siguiente salida (Salida 6).



tiempo 36 0,77 0,70 24,62

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	59416,22	8	7427,03	11,00	<0,0001
mat	10683,72	2	5341,86	7,91	0,0020
temp	39118,72	2	19559,36	28,97	<0,0001
mat*temp	9613,78	4	2403,44	3,56	0,0186
Error	18230,75	27	675,21		
Total	77646,97	35			

Test:LSD Fisher Alfa=0,05 DMS=37,70048

Error: 675,2130 gl: 27

mat	temp	Medias	n			
2	125	49,50	4	A		
1	70	57,25	4	A		
1	125	57,50	4	A		
3	125	85,50	4	A	B	
2	70	119,75	4		B	C
1	15	134,75	4			C
3	15	144,00	4			C
3	70	145,75	4			C
2	15	155,75	4			C

Letras distintas indican diferencias significativas(p<= 0,05)

Salida 6. Diferencia entre medias de tratamientos

La interpretación se debe hacer para cada material:

Mat. 1: la duración es mayor a 15º y difiere significativamente de 70 y 125.

Mat. 2: con 15 º es la combinación que da más duración de batería y difiere significativamente si se la usa a 125º (menor duración).

Mat. 3: a 15 y 70º duran más que a 125º.

Para recomendar un material debería preguntarse a qué temperatura trabajará la batería porque estos dos factores interactúan.

El ANOVA muestra diferencias en materiales y temperaturas pero, al dar la interacción significativa es conveniente primero analizar el comportamiento de los niveles de un factor en un nivel fijo del otro lo que ya se realizó (Salida 6). Para observar gráfiamente el comportamiento se puede realizar un gráfico de las medias del tiempo de duración de las baterías en función de los dos factores, de la siguiente manera: Se cliquea gráfico_Diagrama de puntos (Figura 14)

Figura 14: **Diagrama de puntos**.

Se cliquea en particiones y se coloca la variable materiales (mat)(Figura 15)

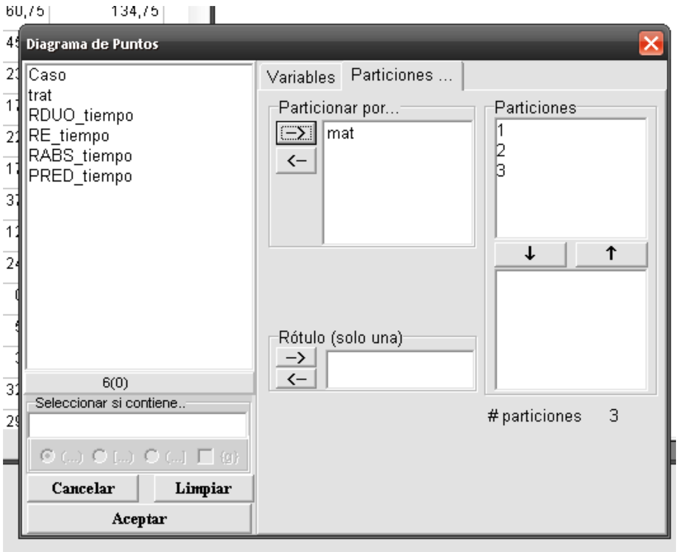


Figura 15: **Diagrama de puntos, partición por materiales.**

Al aceptar se obtiene la Figura 16, se marca media; en el cuadro de medidas de confianza marcamos ninguna, luego se marca tratar el eje x como categórico y particiones en el mismo gráfico.

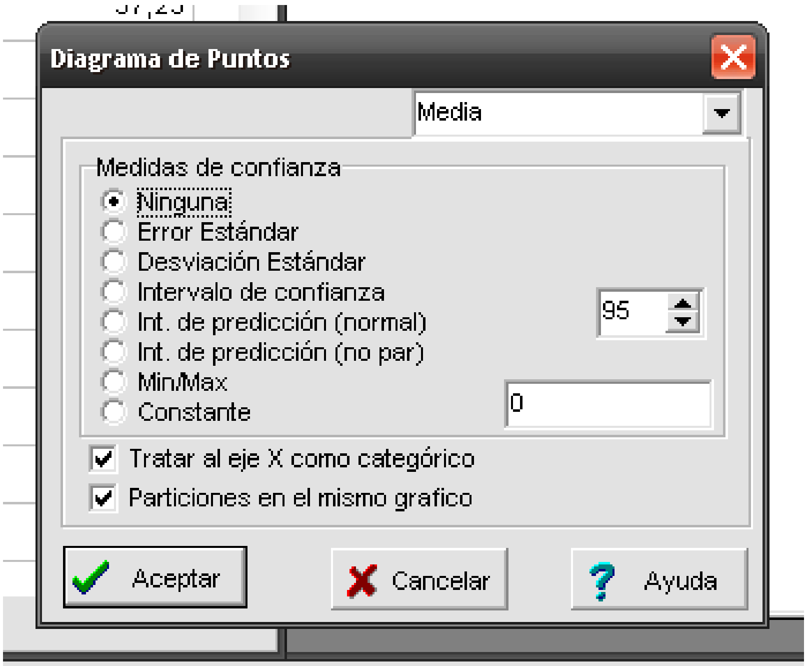
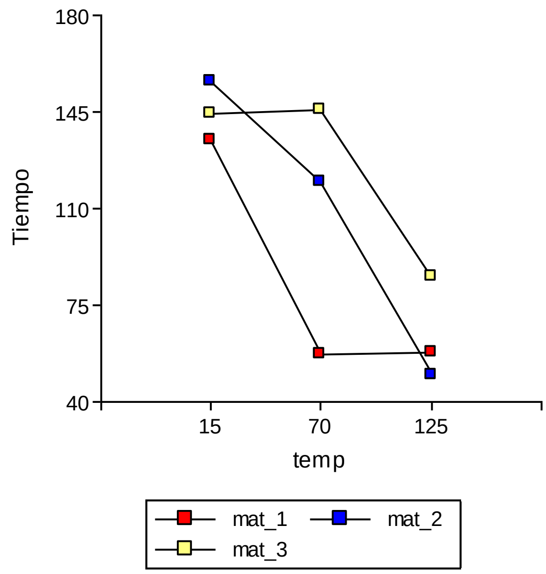


Figura 16: **Instrucciones para el diagrama de puntos.**

Al aceptar se obtiene la Salida 7. La salida 7 es de gran ayuda en esta interpretación: se ve que todos los materiales son más durables a 15º pero que 1 y 2 son más sensibles al aumento de temperatura.



Salida 7. Promedio de tiempo, según temperatura para cada material

Muchas veces ocurre que la interacción puede enmascarar la diferencia entre los efectos principales. Sin embargo, no sucede en este ejemplo.

Si se examina la significación de los efectos principales según los resultados del ANOVA:

Prueba de hipótesis sobre el efecto del factor renglón: Material

$$\mathbf{H_0 = \tau_1 = \tau_2 = ... = \tau_a = 0}$$

$$\mathbf{H_1 = \text{al menos un } \tau_i \neq 0}$$

Con un valor p tan bajo p = 0,0020 se rechaza la hipótesis nula y se concluye que el efecto del material es diferente de cero, es decir que el tiempo promedio de duración es diferentes entre al menos dos materiales.

Prueba de hipótesis sobre el efecto del factor columna: Temperatura

$$\mathbf{H_0: \beta_1 = \beta_2 = ... = \beta_b = 0}$$

$$\mathbf{H_1: \text{al menos un } \beta_i \neq 0}$$

Con un valor p tan bajo p < 0,0001 se rechaza la hipótesis nula y se concluye que el efecto de la temperatura en el tiempo de duración de las baterías es diferente de cero, es decir que el tiempo promedio de duración es diferente entre al menos dos temperaturas, pero no se las prueba por haber dado significativa la interacción.

2. DESARROLLO CON SPSS Ver. 15.0

En SPSS cuando se pide archivo **nuevo_Datos** aparecen dos pestañas en la parte inferior: Vista de variables y Vista de datos- Se hace clic en la pestaña Vista de variables se poder poner el nombre, ancho y tipo de variables (figura 17):

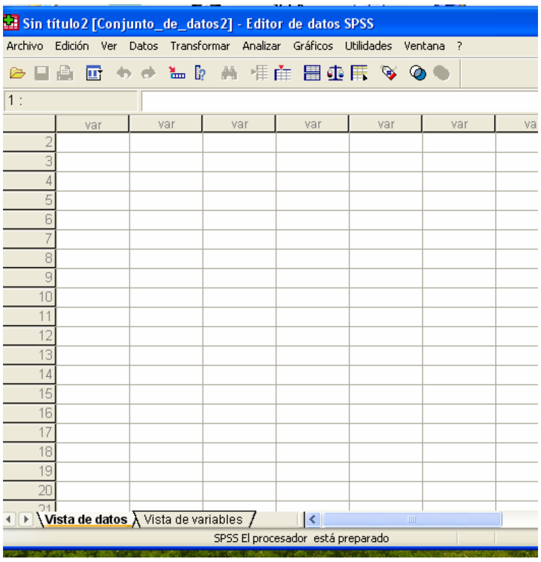


Figura 17: Tabla para crear el archivo de datos

Una vez que se completa los cuadros queda la información como se muestra en la Figura 18.

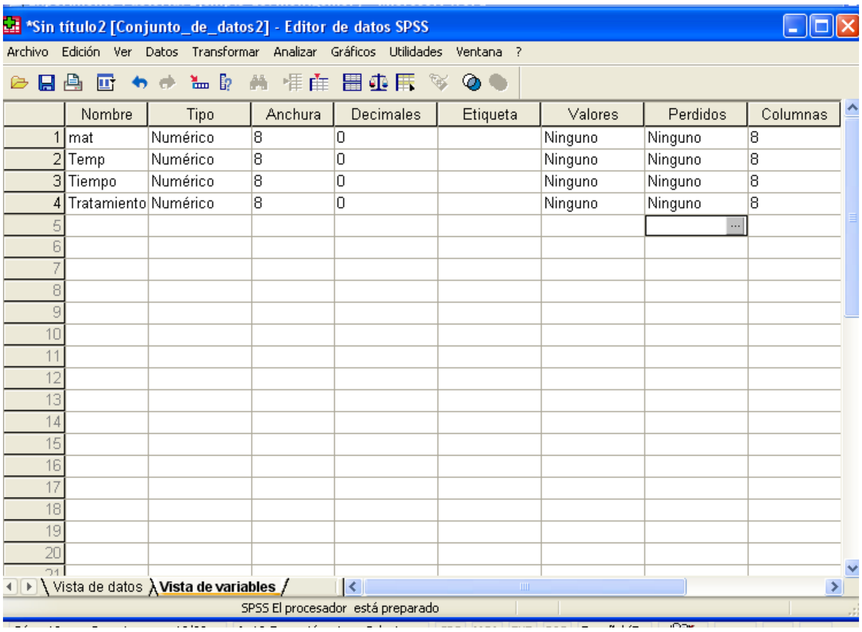


Figura 18: Información de las variables del archivo.

Para introducir los datos, se debe hacer ahora clic en la pestaña **Vista de datos** y se observa que tenemos la planilla con el nombre de las variables en las columnas (Figura 19).

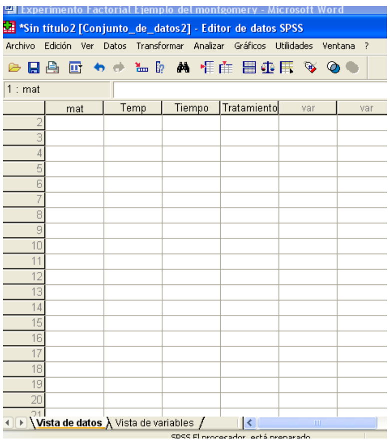
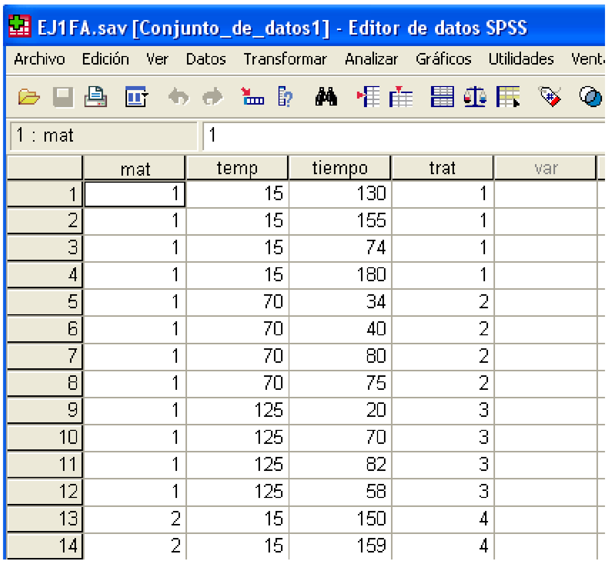


Figura 19. **Planilla para introducir los datos**

La planilla esta lista para que se incorporen los datos. Se debe notar que hemos creado nosotros la columna tratamientos y debemos designar un número para cada combinación de niveles (recordar que en INFOSTAT hay una sentencia que completa la columna). En la figura 20 se muestran los datos cargados



	mat	temp	tiempo	trat	var
1	1	15	130	1	
2	1	15	155	1	
3	1	15	74	1	
4	1	15	180	1	
5	1	70	34	2	
6	1	70	40	2	
7	1	70	80	2	
8	1	70	75	2	
9	1	125	20	3	
10	1	125	70	3	
11	1	125	82	3	
12	1	125	58	3	
13	2	15	150	4	
14	2	15	159	4	

.

Figura 20.Planilla con datos cargados.

Se procede a realizar un análisis exploratorio para ver el comportamiento de cada uno de los factores utilizando **boxplot** o de la combinación de los niveles de las mismas que serían los **box-plot múltiples.** Para obtener un **box-plot** para cada uno de los factores se marca:

Gráfico_Cuadros de diálogos antiguos_Diagrama de cajas (Figura 21).

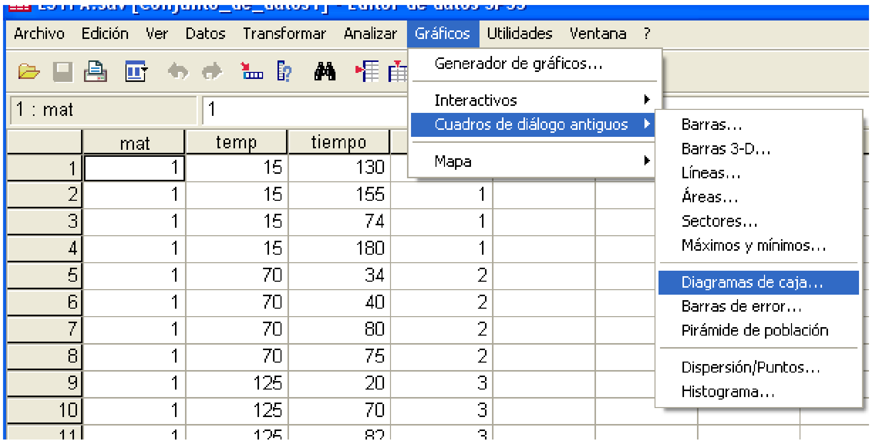


Figura 21: **Caminos para obtener gráfico de Cajas.**

Al hacer click en diagrama de caja, aparece el cuadro que se muestra en la figura 22.

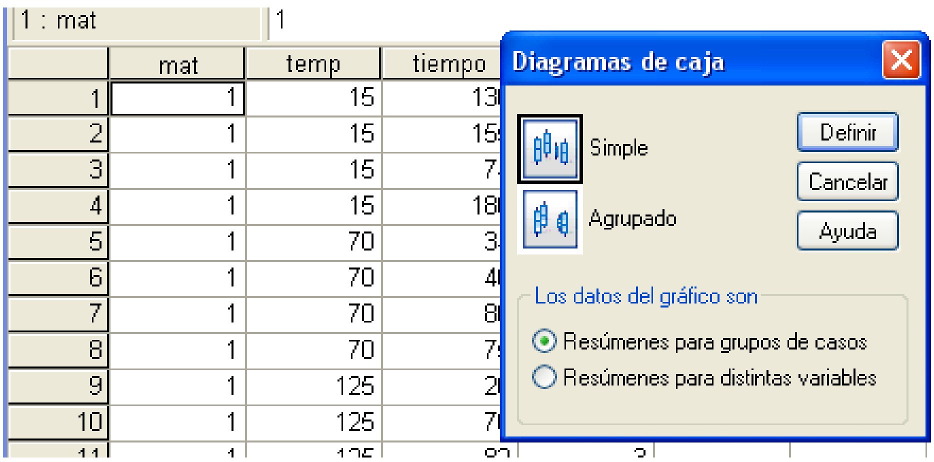


Figura 22: **Opciones a marcar para Box Plot**

Se marca simple y se hace clic en definir, aparece la pantalla que se muestra en la figura 23 a.

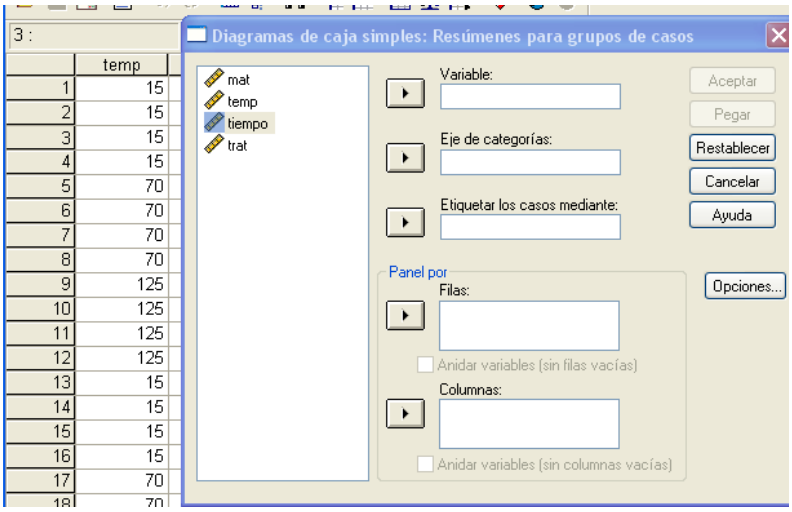


Figura 23 a: Como se indican las variables.

Se completa el **cuadro de variable**, señalando tiempo y pasándolo con flecha derecha. De igual forma se pasa la variable mat al cuadro **eje de categoría** (Fig.23 b).

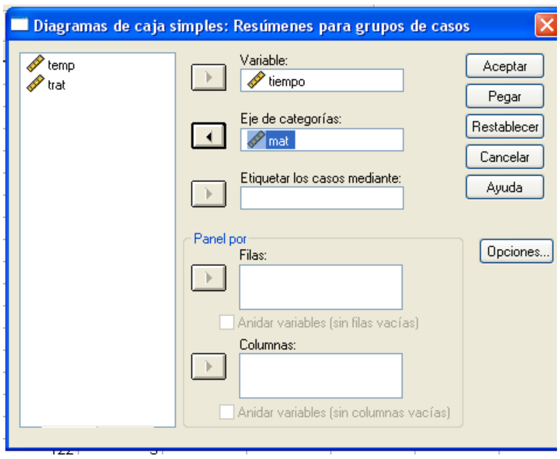
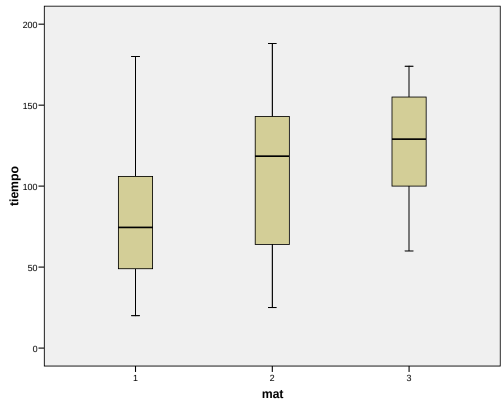


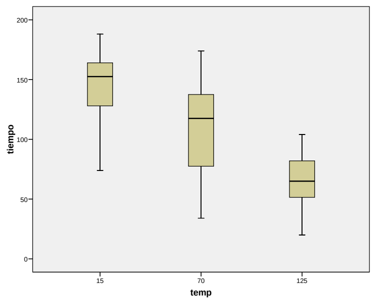
Figura 23 b. Cuadro completo de Box-plot Tiempo para cada material

Al aceptar aparece el gráfico correspondiente a tiempo según material (Salida 8).



Salida 8. Box-plot de la variable tiempo según material.

Se procede igual manera para obtener el box-plot correspondiente a la variable tiempo según temperatura, el gráfico puede verse en la Salida 9.



Salida 9. Box-plot de la variable tiempo para cada temperatura.

Para observar el comportamiento del tiempo de duración de las pilas en función de la temperatura y el material se realiza un boxplot múltiple de la siguiente manera: Seleccionar **Gráficos_Diagrama**

de caja, aparece la pantalla que se muestra en la figura 24; se señala **agrupados_definir y** en pantalla aparece el cuadro que se visualiza en la figura 25.

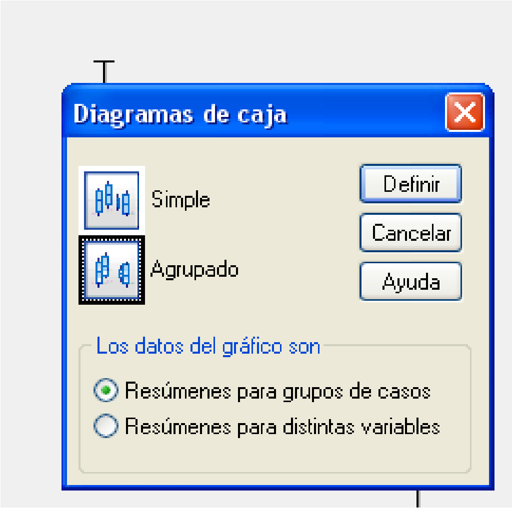
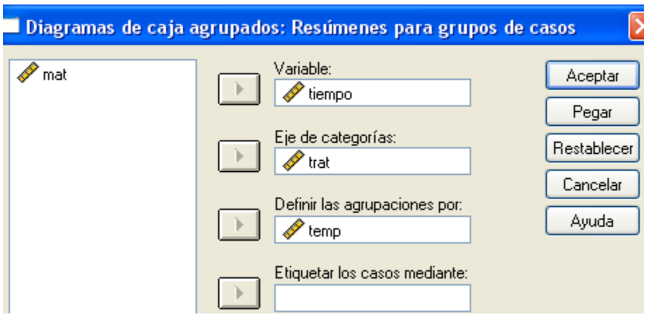


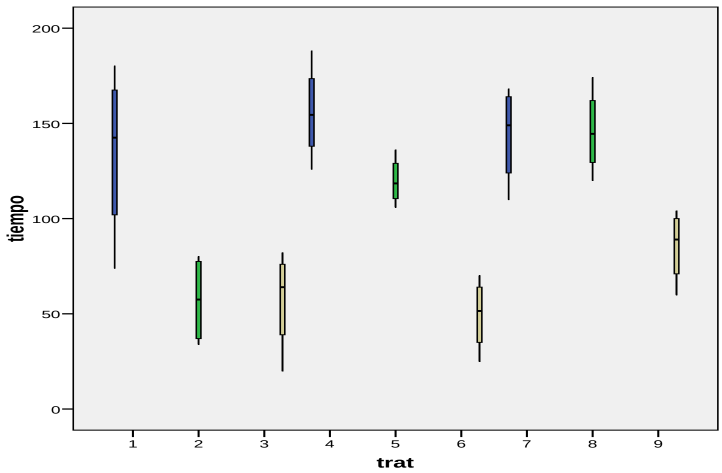
Figura 24. **Pantalla correspondiente al box-plot**

Se marca y pasa la variable tiempo al campo **variable**; tratamiento al campo **eje de categorías** y la variable temperatura al campo **definir las agrupaciones por (**Figura25).

Figura 25: **Cuadro de diagrama de boxplot múltiple completo**



Al aceptar aparece el gráfico como lo muestra la salida 10.



Salida 10. **Box-plot de tiempo para cada tratamiento (combinación de niveles de temperatura-material**

Para realizar el análisis de la variancia vamos al menú Analizar_Modelo lineal general_univariante (figura 26).

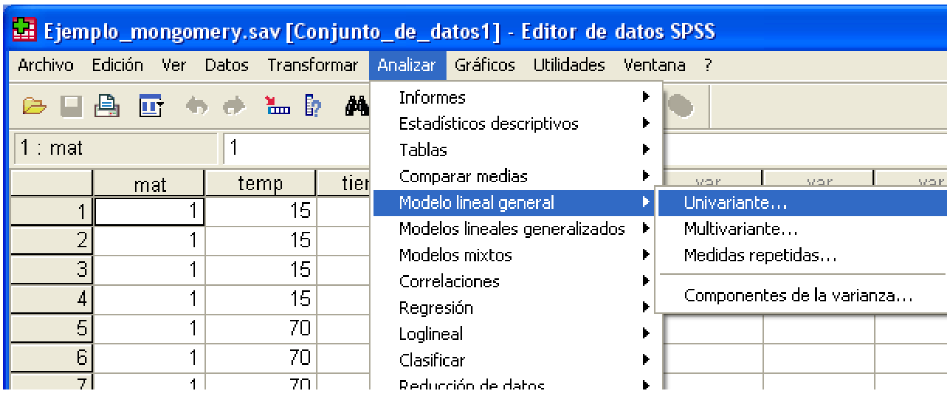


Figura 26: **Pantalla para indicar el análisis de la varianza**

Al hacer clic, aparece la pantalla que se muestra en la Figura 27, se marca y pasa la variable tiempo al cuadro dependiente y al cuadro de Factores fijos los factores: material y temperatura.

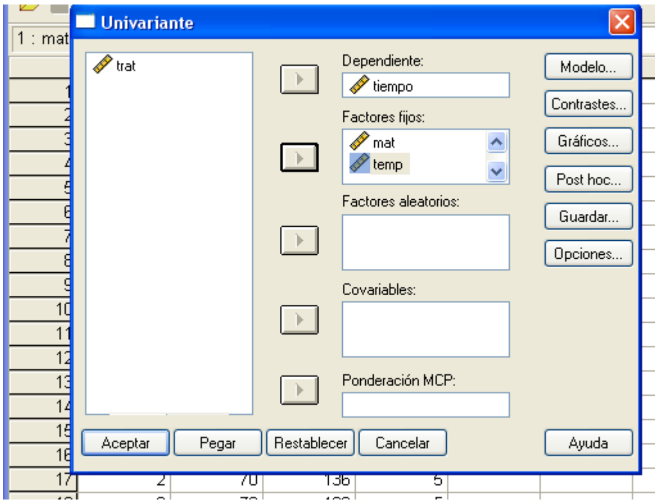


Figura 27: Ubicación de la variable y factores en los cuadros

Hacer clic en Modelo para indicar que además de los factores debemos colocar en el Análisis la interacción de ellos (Figura 28).

Al marcar personalizado, se resaltan los colores en el cuadro Factores y covariables, en los que se nota que al lado el nombre de los factores (Figura 28) aparece (F) indicando que son fijos.

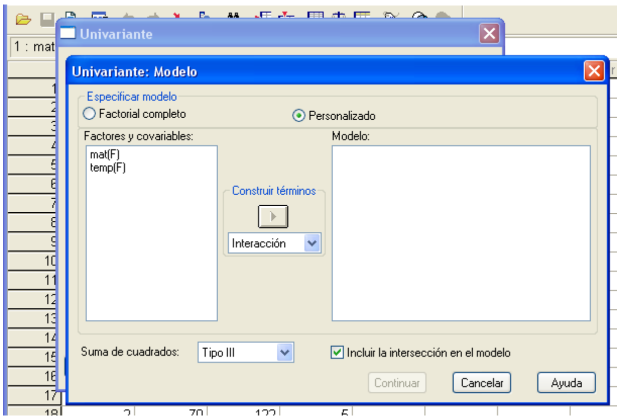


Figura 28: **Pantalla para especificar un factorial**

Se debe marcar **incluir interacción en el modelo** y en la parte central debe estar colocado interacción. Con el mousse se marca cada uno de los factores por separados y se los pasa con la flecha al cuadro titulado modelo, luego se marcan los dos juntos y se los pasa con la flecha y aparece la interacción en el cuadro titulado Modelo; se marca continuar (figura 29).

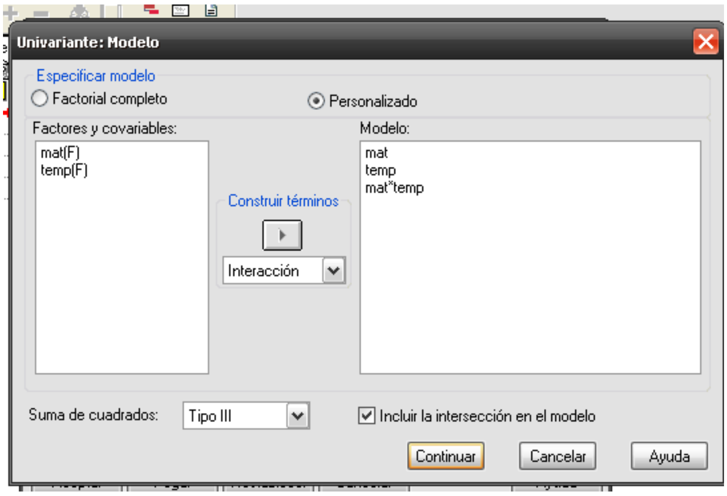
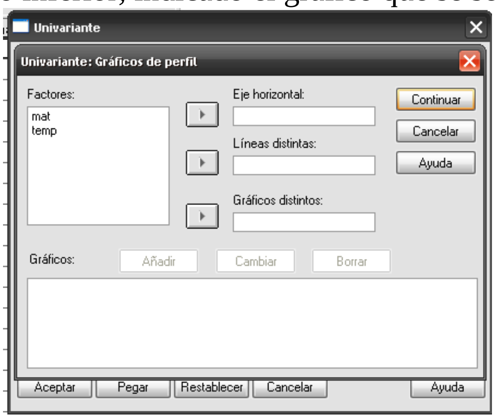


Figura 29: Pantalla con indicación de interacción

Al marcar continuar, se vuelve a la figura 27 y al hacer un clic en **gráfico** aparece la siguiente pantalla (figura 30a). Se pasa material (mat) al eje horizontal; temperatura (temp) a líneas distintas y se habilita en la parte inferior añadir (figura 30b), al hacer clic aparece en el cuadro inferior, indicado el gráfico que se solicitó es mat*temp (figura 30c).



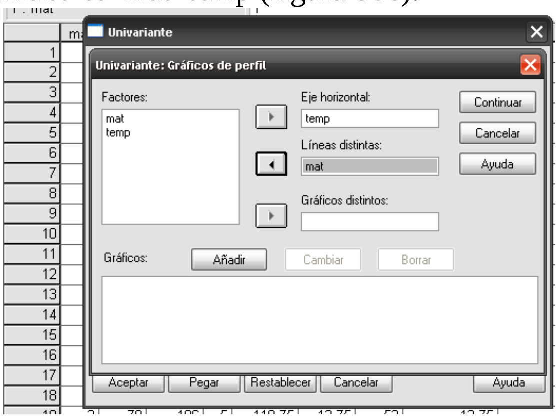


Figura 30 a **Figura 30 b**

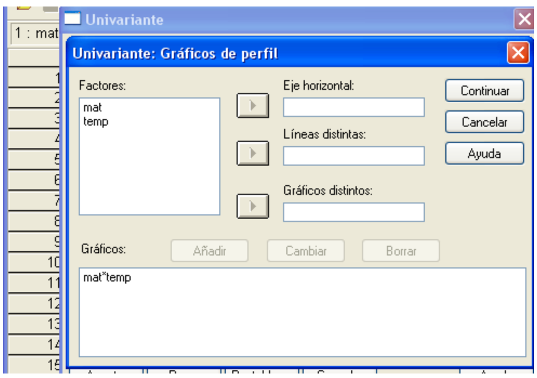
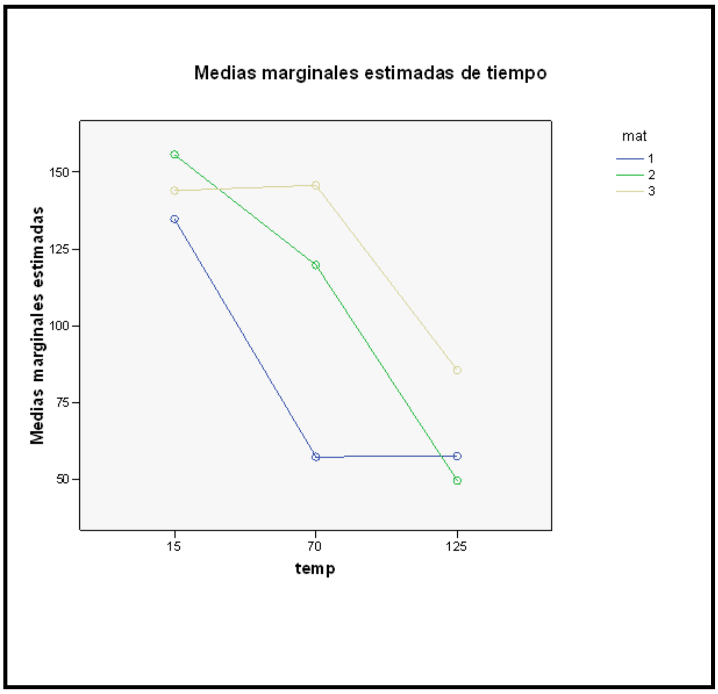


Figura 30 c

Figura 30 (a,b,c): Secuencias de pantallas para especificar un gráfico de tiempo para los diferentes materiales según temperatura

El gráfico de tiempo por temperatura y materialse muestra en la salida 11.



Salida 11. Tiempo en función de temperatura y tipo de material

Se hace clic en continuar y se vuelve a la figura 27 y se marca **guardar**. Esta opción permite guardar los residuos y los valores estimados para realizar luego las diferentes pruebas de supuestos gráfica y analíticamente.
Luego de hacer clic en guardar aparece la siguiente pantalla (figura 31).

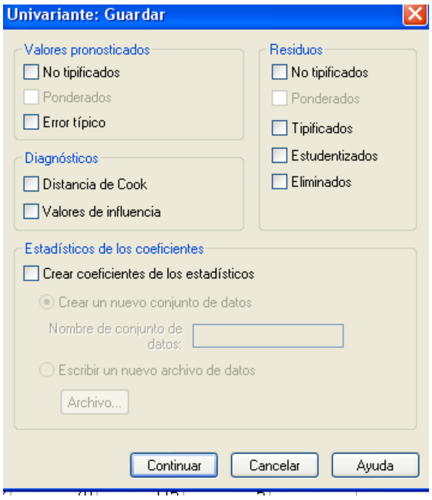


Figura 31: Pantalla para indicar los valores pronósticados y los residuos que se deben guardar

Para ello se tilda, en el cuadro de **Valores pronosticados**: No tipificados, en el cuadro de **Residuos**, no tipificados y tipificados (estandarizados), lo que permite posteriormente que el programa cree dichas columnas en el archivo. Continuar
Por último nuevamente aparece la figura 27, hacemos clic en Opciones y se visualiza la pantalla siguiente (figura 33).

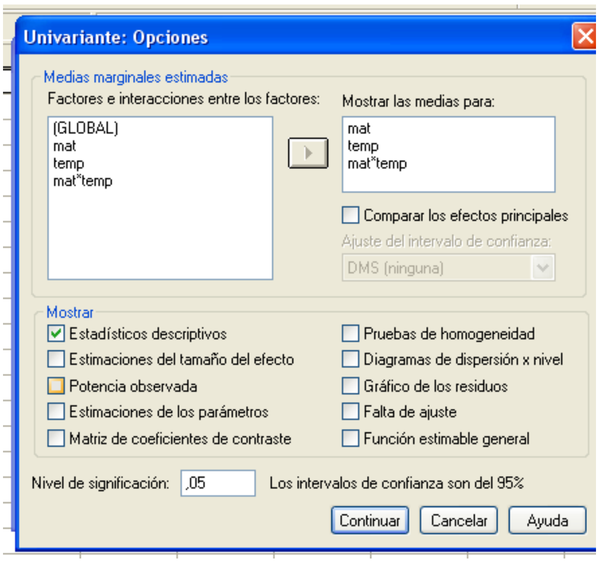


Figura 33: Pantalla para solicitar estadísticos descriptivos para factores e interacciones.

Se pasa de la forma acostumbrada, los factores y la interacción al cuadro ”mostrar las medias para: mat, temp y mat*temp y en el cuadro Mostrar se marca estadísticos descriptivos, se deja nivel de significación en 0,05 o puede cambiarse a 0,01 como se desee.
Continuar, continuar y aparecen en la hoja de resultados los siguientes valores (Salida 12a).

Análisis de varianza univariante

Factores inter-sujetos

Estadísticos descriptivos

Variable dependiente: tiempo

		N
mat	**1**	**12**
	2	**12**
	3	**12**
temp	**15**	**12**
	70	**12**
	125	**12**

Salida 12 a. Media, desviación estándar y número de repeticiones en la que se presenta para cada combinación material-temperatura, y también los mismos estadísticos para material y para temperatura.

mat	temp	Media	Desv. típ.	N
1	15	134,75	45,353	4
	70	57,25	23,599	4
	125	57,50	26,851	4
	Total	83,17	48,589	12
2	15	155,75	25,617	4
	70	119,75	12,659	4
	125	49,50	19,261	4
	Total	108,33	49,472	12
3	15	144,00	25,974	4
	70	145,75	22,544	4
	125	85,50	19,279	4
	Total	125,08	35,766	12
Total	15	144,83	31,694	12
	70	107,58	42,883	12
	125	64,17	25,672	12
	Total	105,53	47,101	36

En la **Salida 12 b,** Aparece el análisis de la variancia con valores de Sumas de cuadrados y F iguales a los valores obtenidos con el software anterior, por lo que se realizan las mismas pruebas y test de hipótesis realizados anteriormente y que no es necesario repetirlo.

Pruebas de los efectos inter-sujetos

Variable dependiente: tiempo

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo corregido	59416,222(a)	8	7427,028	11,000	,000
Intersección	400900,028	1	400900,028	593,739	,000
mat	10683,722	2	5341,861	7,911	,002
temp	39118,722	2	19559,361	28,968	,000
mat * temp	9613,778	4	2403,444	3,560	,019
Error	18230,750	27	675,213		
Total	478547,000	36			
Total corregida	77646,972	35			

a R cuadrado = ,765 (R cuadrado corregida = ,696)

Salida 12 b. Análisis de la variancia

La salida que figura abajo **Salida 12 c,** corresponde a la salida de la opción en donde se colocó estadística para factores y para interacción

Medias marginales estimadas

1. mat

Variable dependiente: tiempo

mat	Media	Error típ.	Intervalo de confianza al 95%. Límite inferior	Límite superior
1	83,167	7,501	67,776	98,558
2	108,333	7,501	92,942	123,724
3	125,083	7,501	109,692	140,474

2. temp

Variable dependiente: tiempo

temp	Media	Error típ.	Intervalo de confianza al 95%. Límite inferior	Límite superior
15	144,833	7,501	129,442	160,224
70	107,583	7,501	92,192	122,974
125	64,167	7,501	48,776	79,558

3. mat * temp

Variable dependiente: tiempo

mat	temp	Media	Error típ.	Intervalo de confianza al 95%. Límite inferior	Límite superior
1	15	134,750	12,992	108,092	161,408
	70	57,250	12,992	30,592	83,908
	125	57,500	12,992	30,842	84,158
2	15	155,750	12,992	129,092	182,408
	70	119,750	12,992	93,092	146,408
	125	49,500	12,992	22,842	76,158
3	15	144,000	12,992	117,342	170,658
	70	145,750	12,992	119,092	172,408
	125	85,500	12,992	58,842	112,158

Salida 12 c. Estadísticos para factores e interacción.

Para que el análisis de la variancia sea válido, se deben probar los supuestos
1.- Prueba de Normalidad, con la Prueba de Shapiro Wilks-
Seleccionar**: Analizar_Estadísticos descriptivos_explorar** (figura 34).

Figura 34: Pasos para realizar prueba de normalidad de residuos

Al confirmar la selección aparece la pantalla de abajo (Figura 35). Se pasa la variable residuos (RES_1), en el cuadro **Mostrar** se marca **Gráficos** y aparece la pantalla (Figura36), en el cuadro **diagrama de caja** se marca **ninguna**, en el cuadro **Descriptivos**, se desmarcan tallos y hojas e Histogramas. Se marca **gráficos con prueba de normalidad**
Continuar y aceptar. Los resultados de la prueba se encuentran en la salida 13

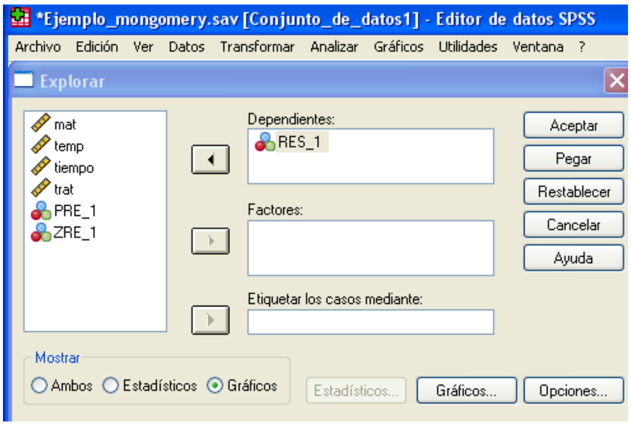


Figura 35 :Selección de variable para prueba de normalidad

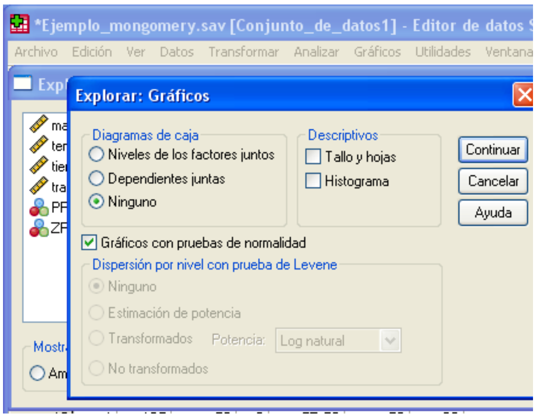


Figura 36: Selección del test de normalidad

Salida 13. Prueba de normalidad gráfica y analíticamente

Explorar

Resumen del procesamiento de los casos

	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
Residuo para tiempo	36	100,0%	0	,0%	36	100,0%

Pruebas de normalidad

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Residuo para tiempo	,106	36	,200(*)	,976	36	,612

* Este es un límite inferior de la significación verdadera.
a Corrección de la significación de Lilliefors

Como p=0,612 No se rechaza la hipótesis que los residuos siguen una distribución normal. La figura 37 muestra gráficamente la normalidad (obtenida en la salida 13).

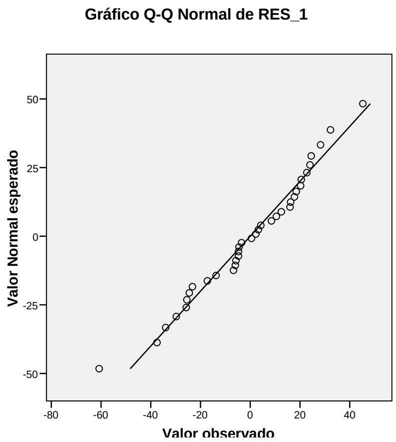
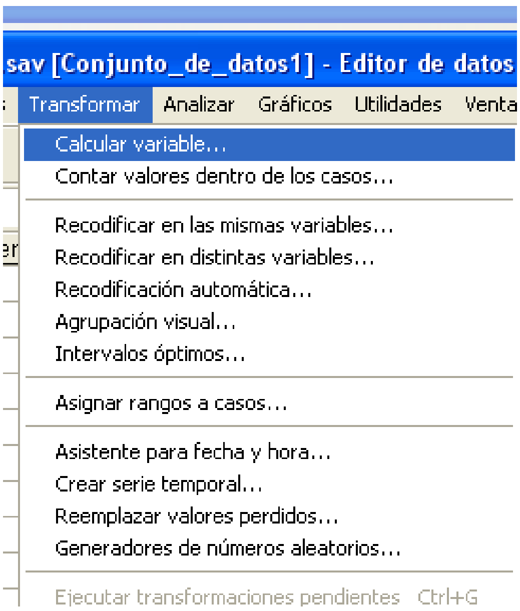


Figura 37: Q-Q plot de distribución normal para residuos

2.-Prueba de homogeneidad

La prueba de homogeneidad de Levene consiste en realizar un ANOVA con los valores absolutos de los residuos considerando como factor de clasificación a los tratamientos.



El SPSS no guarda el valor absoluto de los residuos por lo que se los debe obtener. Para ello se procede de la siguiente manera: se selcciona del menú principal **Transformar_Calcular variable(**Figura 38) y al hacer clic aparece la siguiente pantalla (Figura 39).

Figura 38: Transformar variable mediante cálculo

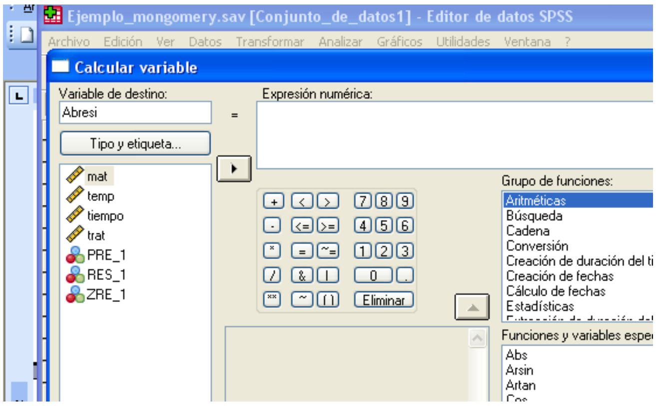


Figura 39: Pantalla a completar para obtener la variable transformada

En variable destino se coloca la variable a crear que es el residuo en valor absoluto, a la que se denominará **absresi** . Pasamos el cursor a **Expresión numérica** buscamos en **Grupo de funciones** las **Aritméticas** y en **Funciones y variables especiales** se selecciona Abs, se aplica doble clic y aparece la función seleccionada en expresiones numéricas de la siguiente manera ABS(?). Con el cursor se va al cuadro **Tipo y etiqueta** y con un doble clic en la variable RESI_1 se obtiene en el cuadro de Expresiones numéricas: ABS(RESI_1) se aceptar, y la variable creada se visualiza en el archivo de datos la variable creada (Figura 40).

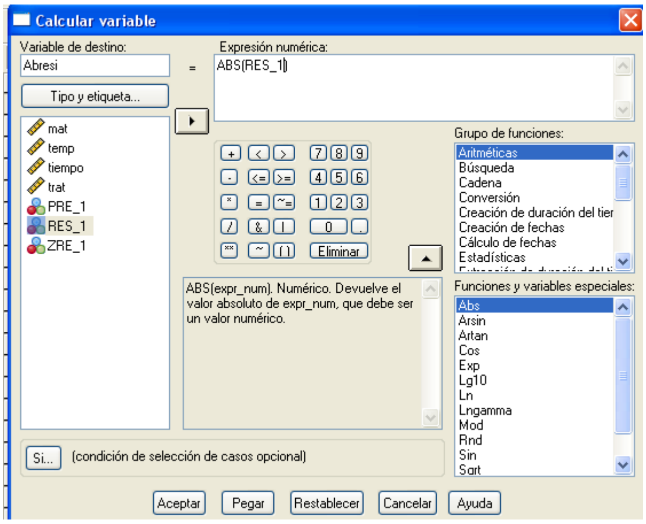
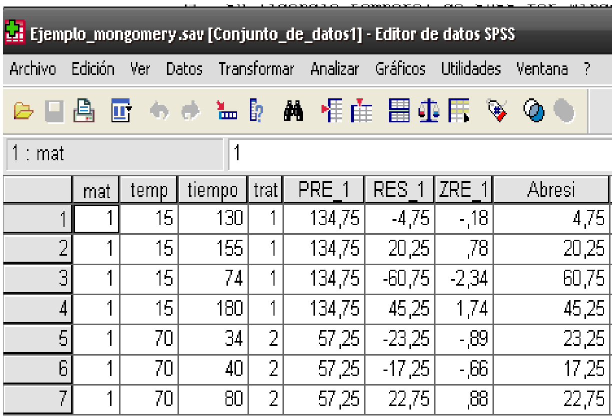


Figura 40: Pantalla con cuadros completados para obtener la variable transformada.

La variable creada, aparece como una nueva columna en la hoja de datos (figura 41).



	mat	temp	tiempo	trat	PRE_1	RES_1	ZRE_1	Abresi
1	1	15	130	1	134,75	-4,75	-,18	4,75
2	1	15	155	1	134,75	20,25	,78	20,25
3	1	15	74	1	134,75	-60,75	-2,34	60,75
4	1	15	180	1	134,75	45,25	1,74	45,25
5	1	70	34	2	57,25	-23,25	-,89	23,25
6	1	70	40	2	57,25	-17,25	-,66	17,25
7	1	70	80	2	57,25	22,75	,88	22,75

Figura 41: Pantalla con la nueva variable transformada.

Se está ahora en condiciones de realizar el análisis de la varianza. Se selecciona en el menú principal: **Analizar_Modelo Lineal General_Univariante;** Marcar **Abresi** y con la flecha pasar a al cuadro titulado dependiente; **Marcar tratamiento** y con la flecha pasar a **Factores Fijos** (figura 42).

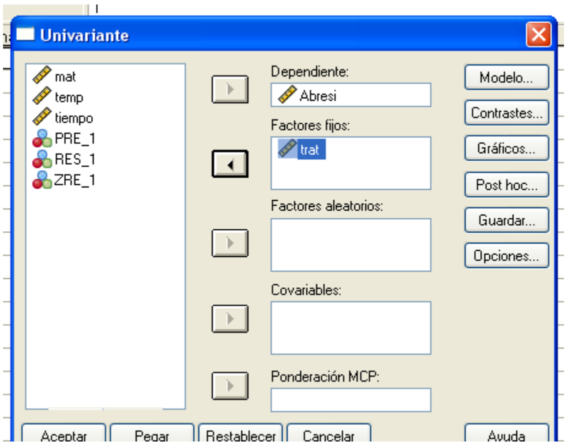


Figura 42: Pantalla con los cuadros completos para realizar la prueba de Levene

Al **aceptar** , se obtiene el análisis de la varianza que aparece en la salida 14 y se observa la significación para la variable tratamiento, el valor de p debe ser mayor a 0,05 para aceptar la hipótesis de homogeneidad de variancias.

Tests of Between-Subjects Effects

Dependent Variable: Abresi

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1323,722[a]	8	165,465	,902	,529
Intercept	11953,778	1	11953,778	65,160	,000
trat	1323,722	8	165,465	,902	,529
Error	4953,250	27	183,454		
Total	18230,750	36			
Corrected Total	6276,972	35			

a. R Squared = ,211 (Adjusted R Squared = -,023)

Salida 14: Análisis de la variancia correspondiente a la prueba de homogeneidad de Levene

Con un p = 0,529 no se rechaza la hipótesis de que existe homogeneidad de varianza entre los diferentes tratamientos.

La prueba gráfica de homogeneidad consiste en graficar los residuos estandarizados en función de los valores estimados de la variable dependiente, no debiendo los mismos mostrar un patrón determinado. Para ello: del menú principal se selecciona

Gráfico_Cuadro de diálogos antiguos (Figura 43)_Dispersión/Puntos…,

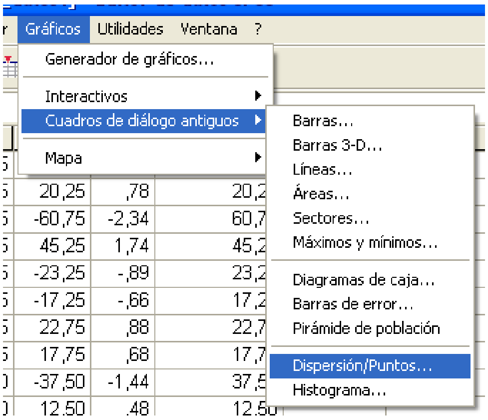


Figura 43:Pasos para gráfico de dispersión

Se hace clic y aparece la pantalla **(figura 44)** en la que se marca **Dispersión simple_Definir**

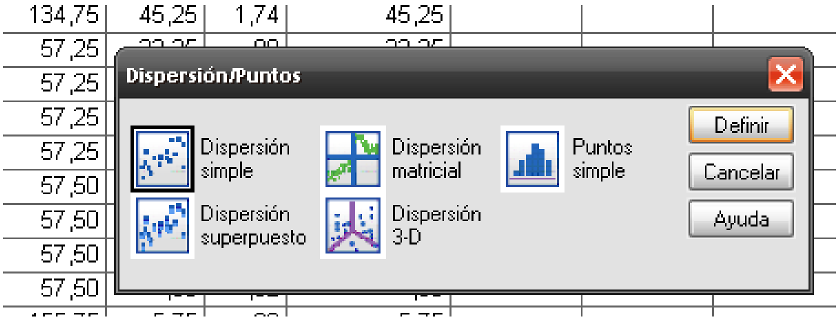


Figura 44: Pantalla para definir grafico de dispersión simple

Luego de realizar clic se obtiene la pantalla siguiente (figura 45).

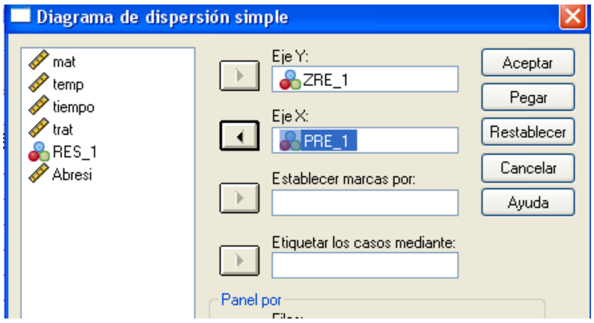
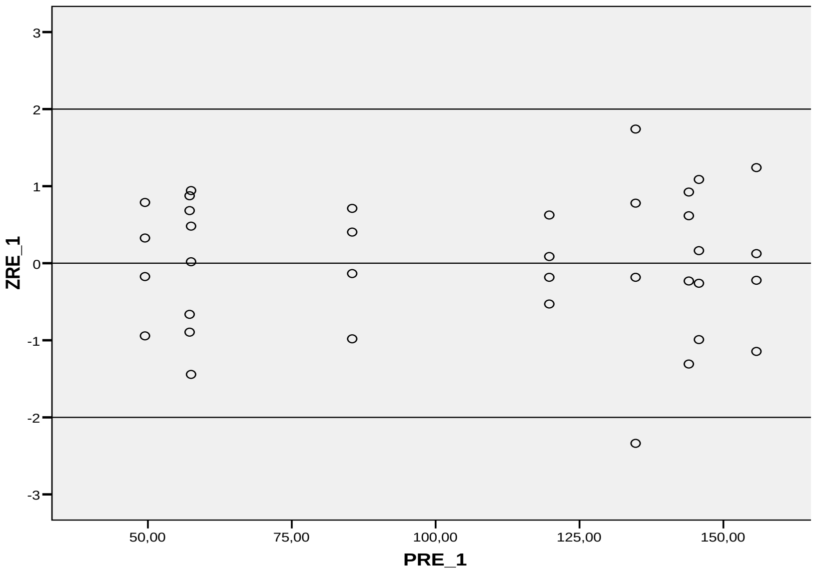


Figura 45: Pantalla con las variables en los ejes y y x del gráfico de dispersión.

Se coloca ZRE_1 en el eje Y y PRE_1 en el eje de las X. Al aceptar se obtiene la salida 15



Salida 15. Residuos estandarizados versus variable estimada

Las salidas 13, 14 y 15 demuestran que los residuos cumplen con los supuestos de normalidad y homogeneidad de variancias. Es por ello que se consideran válidos los resultados de la salida 12 y se realizan las siguientes conclusiones.

Conclusiónes

La interacción es significativa (p = 0.019; tabla de anova salida 12b). Existen diferencias de las medias de los tiempos para las diferentes temperaturas en cada nivel del factor material o viceversa. Existen diferencias entre las medias de los tiempos para las diferentes temperaturas (p= 0.0001; tabla de anova salida 12b.) Existen diferencias entre las medias de los tiempos para los diferentes materiales (p= 0.0020; tabla de anova salida 12 b)

Se debe considerar la primera conclusión, los estadísticos deben construirse a mano:
Comparación entre:

Material 1 a las diferentes temperaturas

$$\overline{y}_{11.} = 134,75$$
$$\overline{y}_{12..} = 57,25$$
$$\overline{y}_{13.} = 57,50$$

$$DLS_{0,05} = t_{27\ 0,05}\sqrt{\frac{CME}{n}} = 2,052\sqrt{\frac{675,213}{4}} = 2,052 * 12,992 = 26,66$$

Material 1. hay diferencias significativas entre las medias del tiempo de duración entre la temperatura de 15 º y las temperaturas de 70º y 125º
No hay diferencias entre las medias del tiempo de duración entre las temperatura 70º La batería no debe ser expuesta a temperatura de 125º

$$d_1 = \overline{y}_{11.} - \overline{y}_{12.} = 134,75 - 57,25 = 77,5*$$
$$d_2 = \overline{y}_{11..} - \overline{y}_{13.} = 134,75 - 57,50 = 57,25*$$
$$d_3 = \overline{y}_{12-} - \overline{y}_{13.} = 57,25 - 57,50 = -0,25$$

Material 2

$$d_3 = \overline{y}_{21.} - \overline{y}_{22.} = 155,75 - 119,75 = 36*$$
$$d_2 = \overline{y}_{21..} - \overline{y}_{23.} = 155,75 - 49,50 = 106,25*$$
$$d_4 = \overline{y}_{22-} - \overline{y}_{23.} = 119,75 - 47,50 = 70,25*$$

Para el material 2, se aconseja 70º pues la batería dura más tiempo

Material 3

$$d_5 = \overline{y}_{31.} - \overline{y}_{32.} = 144,00 - 145,75 = -1,75$$
$$d_6 = \overline{y}_{31..} - \overline{y}_{33.} = 144,00 - 85,5 = 58,5*$$
$$d_7 = \overline{y}_{32-} - \overline{y}_{3.} = 145,75 - 85,50 = 60,25*$$

Para el material 3, se aconseja 125º pues la batería dura más tiempo

Si se debe elegir material, se recomienda el material 3, pues es el que más tiempo de duración alcanza a la máxima temperatura.

3.-Desarrollo con SAS.

Es necesario realizar el pequeño programa que figura a continuación en el que también se introducen los datos, se efectúa el análisis de la variancia, las pruebas de supuestos y las comparaciones entre medias.

```
data factorial;
input mat temp tiempo;
cards;
        1 15 130 1
        1 15 74  1
        1 15 155 1
        1 15 180 1
        2 15 150 2
        2 15 188 2
        2 15 159 2
        2 15 126 2
        3 15 138 3
        3 15 168 3
        3 15 110 3
        3 15 160 3
        1 70 34  4
        1 70 40  4
        1 70 80  4
        1 70 75  4
        2 70 136 5
        2 70 122 5
        2 70 106 5
        2 70 115 5
        3 70 174 6
        3 70 120 6
        3 70 150 6
        3 70 139 6
        1 125 20 7
        1 125 70 7
        1 125 82 7
        1 125 58 7
        2 125 25 8
        2 125 70 8
        2 125 58 8
        2 125 45 8
        3 125 96 9
        3 125 104 9
        3 125 82 9
        3 125 60 9
             ;
proc print data= factorial;
run; proc glm data= factorial;
class mat temp;
model tiempo=mat temp mat*temp;
LSMEANS mat*temp/TDIFF;
output out=pp p=dpred r=resid;
proc univariate data=pp plot normal;
var resid;
proc plot data=pp vpercent=50; plot resid*dpred /vref=0 box;
data pp; set pp;
z=ABS(resid);
proc anova data=pp;
class trat;
```

```
model z=trat;
run;
```

La salida es la que se muestra a continuación:

Salida 16: correspondiente al programa detallado arriba

```
Salida
                                   Sistema SAS

                             Obs    mat    temp    tiempo    trat

                               1     1      15      130        1
                               2     1      15       74        1
                               3     1      15      155        1
                               4     1      15      180        1
                               5     2      15      150        2
                               6     2      15      188        2
                               7     2      15      159        2
                               8     2      15      126        2
                               9     3      15      138        3
                              10     3      15      168        3
                              11     3      15      110        3
                              12     3      15      160        3
                              13     1      70       34        4
                              14     1      70       40        4
                              15     1      70       80        4
                              16     1      70       75        4
                              17     2      70      136        5
                              18     2      70      122        5
                              19     2      70      106        5
                              20     2      70      115        5
                              21     3      70      174        6
                              22     3      70      120        6
                              23     3      70      150        6
                              24     3      70      139        6
                              25     1     125       20        7
                              26     1     125       70        7
                              27     1     125       82        7
                              28     1     125       58        7
                              29     2     125       25        8
                              30     2     125       70        8
                              31     2     125       58        8
                              32     2     125       45        8
                              33     3     125       96        9
                              34     3     125      104        9
                              35     3     125       82        9
                              36     3     125       60        9

                                    Sistema SAS

                                 Procedimiento GLM

                            Información de nivel de clase

                              Clase         Niveles    Valores

                              mat                 3    1 2 3

                              temp                3    15 70 125


                        Número de observaciones leídas          36
                        Número de observaciones usadas          36

                                    Sistema SAS

                                 Procedimiento GLM

Variable dependiente: tiempo

                                                 Suma de      Cuadrado de
        Fuente                      DF          cuadrados        la media      F-Valor     Pr > F
```

Modelo	8	59416.22222	7427.02778	11.00	<.0001
Error	27	18230.75000	675.21296		
Total correcto	35	77646.97222			

R-cuadrado	Coef Var	Raiz MSE	tiempo Media
0.765210	24.62372	25.98486	105.5278

Fuente	DF	Tipo I SS	Cuadrado de la media	F-Valor	Pr > F
mat	2	10683.72222	5341.86111	7.91	0.0020
temp	2	39118.72222	19559.36111	28.97	<.0001
mat*temp	4	9613.77778	2403.44444	3.56	0.0186

Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F
mat	2	10683.72222	5341.86111	7.91	0.0020
temp	2	39118.72222	19559.36111	28.97	<.0001
mat*temp	4	9613.77778	2403.44444	3.56	0.0186

Salida 16 (continuación)

PRUEBA DE SUPUESTOS

Sistema SAS
Procedimiento UNIVARIATE
Variable: **resid**
Momentos

N	36	Pesos de la suma	36
Media	0	Observaciones de la suma	0
Desviación típica	22.8227643	Varianza	520.878571
Asimetría	-0.4635911	Kurtosis	0.09663605
Suma de cuadrados no corregidos	18230.75	Suma de cuadrados corregidos	18230.75
Coeficiente de variación	.	Media de error estándar	3.80379405

Medidas estadísticas básicas

Localización		Variabilidad	
Media	0.000000	Desviación típica	22.82276
Mediana	1.375000	Varianza	520.87857
Moda	.	Rango	106.00000
		Rango intercuantil	33.62500

Tests para posición: Mu0=0

Test	-Estadístico-		-----P-valor------	
T de Student	t	0	Pr > |t|	1.0000
Signo	M	1	Pr >= |M|	0.8679
Puntuación con signo	S	3.5	Pr >= |S|	0.9571

Tests para normalidad

Test	--Estadístico--		-----P-valor------	
Shapiro-Wilk	#11 X	0.976057	Pr < W	0.6117
Kolmogorov-Smirnov	D	0.10593	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.054415	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.340337	Pr > A-Sq	>0.2500

```
                    Cuantiles (Definición 5)

                    Cuantil        Estimador

                    100% Máx          45.250
                    99%               45.250
                    95%               32.250
                    90%               24.500
                    75% Q3            18.125
                    50% Mediana        1.375
                    25% Q1           -15.500

                           Sistema SAS

                    Procedimiento UNIVARIATE
                        Variable:  resid

                    Cuantiles (Definición 5)

                    Cuantil        Estimador

                    10%              -29.750
                    5%               -37.500
                    1%               -60.750
                    0% Mín           -60.750


                     Observaciones extremas
          ------Inferior------        ------Superior------

           Valor   Observación         Valor   Observación

          -60.75             2         24.00            10
          -37.50            25         24.50            27
          -34.00            11         28.25            21
          -29.75             8         32.25             6
          -25.75            22         45.25             4


          Stem Hoja                    #            de caja
             4 5                       1               |
             3 2                       1               |
             2 003448                  6               |
             1 026688                  6            +-----+
             0 02348                   5            *--+--*
            -0 7665544                 7            |     |
            -1 74                      2            +-----+
            -2 6643                    4               |
            -3 840                     3               |
            -4                                         |
            -5                                         |
            -6 1                       1               |
               ----+----+----+----+
       Multiplicar Stem.Leaf por 10**+1
```

Salida 16 (continuación)

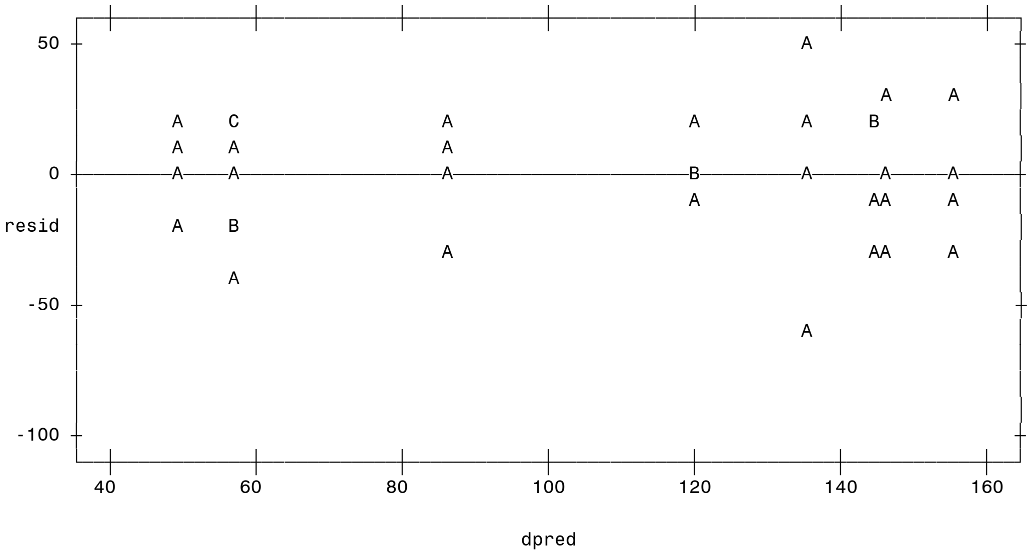
```
                                 Sistema SAS

                        Procedimiento UNIVARIATE
                            Variable:  resid

                         Trazado de probablidad normal
            45+                                             +++*
              |                                          +++*
              |                                      *****+* *
              |                                  *****+
              |                               ***++
              |                           *****
              |                       +***+
              |                   *+****
              |               *+*+
              |           ++++
              |++++
           -65+    *
               +----+----+----+----+----+----+----+----+----+----+
                   -2        -1         0        +1        +2
```

PRUEBA DE HOMOGENEIDAD DE VARIANCIAS

```
                                 Sistema SAS           1
                 Trazado def resid*dpred.  Leyenda: A = 1 obs, B = 2 obs, etc.
```



TEST DE LEVENE

```
                                 Sistema SAS
                             Procedimiento ANOVA
                         Información de nivel de clase

                  Clase        Niveles    Valores

                  trat               9    1 2 3 4 5 6 7 8 9


                  Número de observaciones leídas         36
                  Número de observaciones usadas         36

                                 Sistema SAS

                             Procedimiento ANOVA

Variable dependiente: z
```

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Modelo	8	1323.722222	165.465278	0.90	0.5289
Error	27	4953.250000	183.453704		
Total correcto	35	6276.972222			

R-cuadrado	Coef Var	Raiz MSE	z Media
0.210885	74.32962	13.54451	18.22222

Fuente	DF	Anova SS	Cuadrado de la media	F-Valor	Pr > F
trat	8	1323.722222	165.465278	0.90	0.5289

Salida 16 (continuación)

PRUEBA DE LA DIFERENCIA MINIMA SIGNIFICATIVA (LSD)

Sistema SAS

Procedimiento GLM
Medias de cuadrados mínimos

mat	temp	tiempo LSMEAN	Número LSMEAN
1	15	134.750000	1
1	70	57.250000	2
1	125	57.500000	3
2	15	155.750000	4
2	70	119.750000	5
2	125	49.500000	6
3	15	144.000000	7
3	70	145.750000	8
3	125	85.500000	9

Aquí tenemos el primer valor es t y el de abajo es el valor de probabilidad para la H_0 planteada. Por ej. para un p=0.0002 se rechaza la igualdad de medias entre el tratamiento 1 (mat 1 temp 15) y el 2 Mat 1 temp 70).

Medias de cuadrados mínimos para el efecto mat*temp
t para H0: MediaLS(i)=MediaLS(j) / Pr > |t|

Variable dependiente: tiempo

i/j	1	2	3	4	5	6	7	8	9
1		4.2179	4.204294	-1.14291	0.816368	4.63969	-0.50343	-0.59867	2.680408
		0.0002	0.0003	0.2631	0.4214	<.0001	0.6187	0.5544	0.0124
2	-4.2179		-0.01361	-5.36082	-3.40153	0.42179	-4.72133	-4.81657	-1.53749
	0.0002		0.9892	<.0001	0.0021	0.6765	<.0001	<.0001	0.1358
3	-4.20429	0.013606		-5.34721	-3.38793	0.435396	-4.70772	-4.80296	-1.52389
	0.0003	0.9892		<.0001	0.0022	0.6667	<.0001	<.0001	0.1392
4	1.142915	5.360815	5.347209		1.959283	5.782605	0.639488	0.544245	3.823323
	0.2631	<.0001	<.0001		0.0605	<.0001	0.5279	0.5907	0.0007
5	-0.81637	3.401533	3.387926	-1.95928		3.823323	-1.31979	-1.41504	1.86404
	0.4214	0.0021	0.0022	0.0605		0.0007	0.1980	0.1685	0.0732
6	-4.63969	-0.42179	-0.4354	-5.78261	-3.82332		-5.14312	-5.23836	-1.95928
	<.0001	0.6765	0.6667	<.0001	0.0007		<.0001	<.0001	0.0605
7	0.503427	4.721327	4.707721	-0.63949	1.319795	5.143117		-0.09524	3.183834
	0.6187	<.0001	<.0001	0.5279	0.1980	<.0001		0.9248	0.0036
8	0.59867	4.81657	4.802964	-0.54425	1.415038	5.23836	0.095243		3.279077
	0.5544	<.0001	<.0001	0.5907	0.1685	<.0001	0.9248		0.0029
9	-2.68041	1.537493	1.523887	-3.82332	-1.86404	1.959283	-3.18383	-3.27908	
	0.0124	0.1358	0.1392	0.0007	0.0732	0.0605	0.0036	0.0029	

NOTA: Para asegurar un nivel de protección completo, sólo se deben usar probabilidades asociadas con comparaciones preplanificadas .

BIBLIOGRAFÍA

Balzarini M.G., Gonzalez L., Tablada M., Casanoves F., Di Rienzo J.A., Robledo C.W. (2008). *Manual del Usuario INFOSTAT.* Editorial Brujas, Córdoba, Argentina.

Little, R.; Freund, R.; Spector P.1993.SAS System for linear models. Third edition.329p.

Montgomery, Douglas.1993.Diseño y Análisis de Experimentos. Grupo editorial Ibero-Americana. México.DF. 589p.

Perez, César.2001. Técnicas estadísticas con SPSS. Prentice Hall.571p.

Robles C. Y Gaillard de Benítez, Celia. 1988. Estadística experimental agrícola y forestal. Apuntes de Cátedra.