

Práctica ANOVA 1 via (factor de variación)

Generalización del tema de contrastes de medias para dos grupos manteniendo el riesgo tipo I al nivel deseado.

Ejemplo:

En un experimento clínico se han ensayado 5 tratamientos diferentes correspondientes a 5 principios activos distintos para el tratamiento de una enfermedad, que tienen como efecto secundario el producir una cierta sensación de fatiga. Para ello se dispone de 50 individuos sanos voluntarios, asignados en 5 grupos de igual tamaño al azar, a los que se le administran los 5 principios activos. Como respuesta se utiliza un índice de la disminución en la actividad de los individuos basado en diversas medidas obtenidas sobre cada paciente. Se pretende comprobar si los 5 tratamientos son igualmente efectivos.

Los datos son los siguientes:

Tratamiento	Actividad									
T1	10	12	8	10	6	13	9	10	9	9
T2	11	18	12	15	13	8	15	16	9	13
T3	7	14	10	11	9	10	9	11	7	9
T4	12	9	11	10	7	8	13	14	10	11
T5	7	6	10	7	7	5	6	7	9	6

OBJETIVO:

Comprobar si la variabilidad de la actividad de los individuos se puede explicar por los distintos tratamientos suministrados (niveles del factor) a los individuos.

Resultados:

A)-Una vez introducidos los datos de forma conveniente en el programa SPSS, procederemos a una inspección previa de los mismos con el fin de obtener una descripción estadística de la actividad de los individuos bajo cada tratamiento.

De una **exploración** de la situación obtenemos:

Explorar

☐ VAR00003
☐ VAR00004

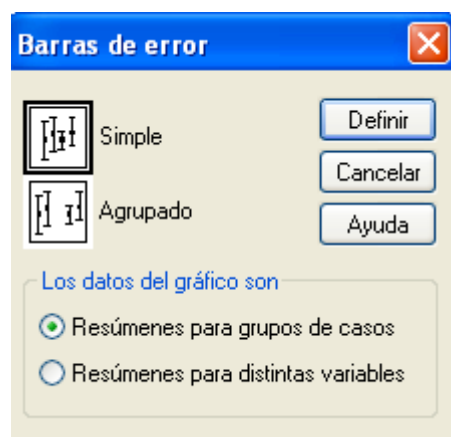
☒ Ambos
 ☐ Estadísticos
 ☐ Gráficos

Descriptivos

Tratamiento			Estadístico
t1	Media		9,6000
	Intervalo de confianza para la media al 95%	Límite inferior	8,2014
		Límite superior	10,9986
t2	Media		13,0000
	Intervalo de confianza para la media al 95%	Límite inferior	10,7631
		Límite superior	15,2369
t3	Media		8,7000
	Intervalo de confianza para la media al 95%	Límite inferior	7,1528
		Límite superior	10,2472
t4	Media		10,5000
	Intervalo de confianza para la media al 95%	Límite inferior	8,9455
		Límite superior	12,0545
t5	Media		7,0000
	Intervalo de confianza para la media al 95%	Límite inferior	5,9336
		Límite superior	8,0664

Extracto de la tabla obtenida

Podemos además dibujar los intervalos de confianza obtenidos en el apartado anterior con el fin de visualizar mejor los resultados: Para ello en la opción de gráficos seleccionamos barras de error:



A la vista de los resultados podemos hacernos una idea de las posibles diferencias entre los grupos en estudio aunque, eso sí, con la problemática de la *t* de Student, es decir sin control del riesgo de rechazo indebido de la hipótesis nula.

Para solucionar estos problemas realizaremos un análisis de la varianza.

B) Requisitos previos a la realización del Anova

NORMALIDAD (ver gráficos box-plot) Parece ser que los grupos no se alejan demasiado de la normalidad

HOMOSCEDASTICIDAD : test de Levene: La prueba de homogeneidad de varianzas tiene una significación de $p = 0,252$ de lo que se desprende que podemos considerar que los 5 grupos son homoscedásticos

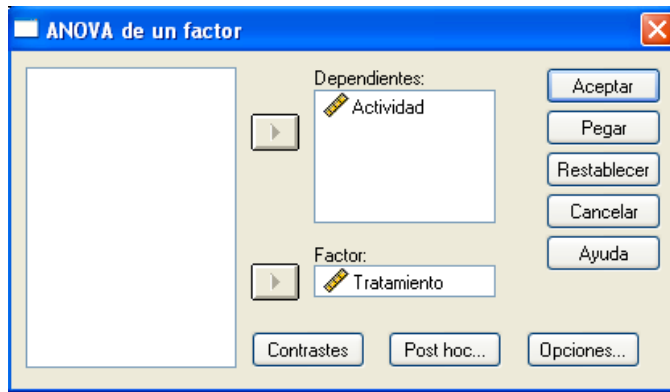
:

Prueba de homogeneidad de la varianza

Estadístico de Levene	gl1	gl2	Sig.
1,392	4	45	,252

Estamos pues en condiciones de aplicar la técnica de ANOVA de 1 vía:

Para ello vamos al menú **Analizar—comparar medias---Anova un factor:**



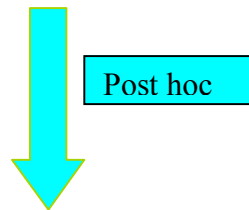
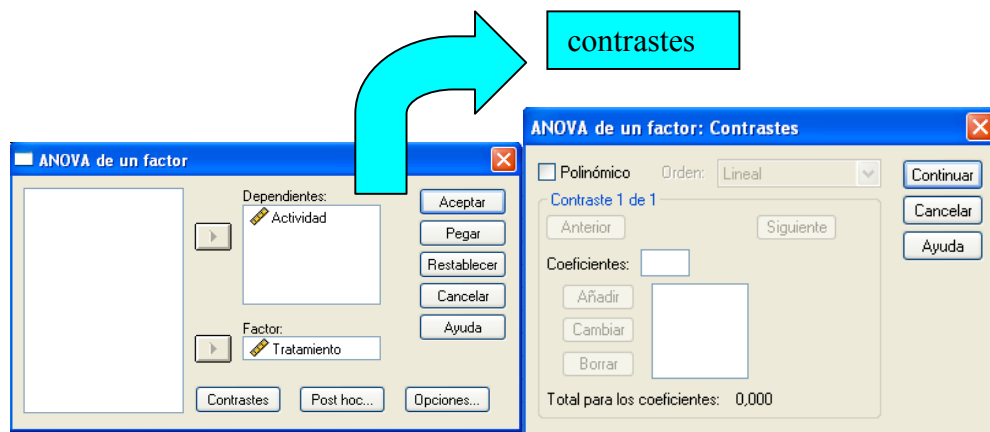
Seleccionamos la variable dependiente y el factor de variación y el **botón aceptar** nos arroja los resultados del ANOVA:

ANOVA

actividad					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	184,920	4	46,230	9,329	,000
Intra-grupos	223,000	45	4,956		
Total	407,920	49			

Lo que nos indica que rechazamos la hipótesis nula de igualdad de medias con resultados altamente significativos. Por tanto nos gustaría saber entre que grupos se ha producido la diferencia. Para ello, la opción **post hoc**: nos permite seleccionar el tipo de contraste que deseamos realizar para poner de manifiesto las diferencias detectadas a nivel global con el ANOVA.

Cuando no queremos hacer todas las comparaciones por pareja, o todas las combinaciones posibles sino que antes de realizar el Anova tuviésemos planeados aquellos contrastes que realmente nos interesan para el estudio el **botón de contrastes** nos permite diseñar nuestros propios contrastes de forma más específica asignando coeficientes positivos y negativos para diferenciar a los grupos que creemos que tienen comportamientos medios distintos. La suma de dichos coeficientes ha de ser igual a cero



ANOVA de un factor: Comparaciones múltiples post hoc

Asumiendo varianzas iguales

<input type="checkbox"/> DMS	<input type="checkbox"/> S-N-K	<input type="checkbox"/> Waller-Duncan
<input type="checkbox"/> Bonferroni	<input type="checkbox"/> Tukey	Tasa de errores tipo I/tipo II: 100
<input type="checkbox"/> Sidak	<input type="checkbox"/> Tukey-b	<input type="checkbox"/> Dunnett
<input type="checkbox"/> Scheffe	<input type="checkbox"/> Duncan	Categoría de control: Ultima
<input type="checkbox"/> R-E-G-W F	<input type="checkbox"/> GT2 de Hochberg	<u>Contraste</u>
<input type="checkbox"/> R-E-G-W Q	<input type="checkbox"/> Gabriel	<input checked="" type="radio"/> Bilateral <input type="radio"/> < Control <input type="radio"/> > Control

No asumiendo varianzas iguales

<input type="checkbox"/> T2 de Tamhane	<input type="checkbox"/> T3 de Dunnett	<input type="checkbox"/> Games-Howell	<input type="checkbox"/> C de Dunnett
--	--	---------------------------------------	---------------------------------------

Nivel de significación: .05

Continuar Cancelar Ayuda

Práctica ANOVA 2 vías (factores de variación)

Ejemplo:

Se está investigando cual es el efecto de tres tipos de abono sobre dos tipos de suelo. Se espera que el efecto de los distintos abonos se manifieste de forma diferente dependiendo del tipo de suelo. Para el presente estudio tomaremos dos tipos de suelo, ácido y alcalino y tres tipos de abono que denotaremos con A B y C. Tenemos así dos factores (suelo y abono) con dos y tres niveles respectivamente, que resultan en 6 combinaciones. Tomaremos un diseño factorial con dos factores y tres replicas en cada una de las combinaciones de los niveles de los dos factores. La respuesta es un índice de abundancia de una determinada especie tras la aplicación de los distintos abonos. Los resultados se muestran en la tabla siguiente.

	A	B	C
Ácido	8	10	8
	4	8	6
	0	6	4
Alcalino	14	4	15
	10	2	12
	6	0	9

OBJETIVO:

Comprobar si la abundancia de la especie analizada se puede explicar por los distintos tratamientos suministrados (abonos y tipos de suelo) así como a la interacción de ambos tratamientos.

Resultados:

A)-Una vez introducidos los datos de forma conveniente en el programa SPSS, procederemos a una inspección previa de los mismos con el fin de obtener una descripción estadística de la abundancia de la especie bajo cada tratamiento y tipo de suelo.

Descriptivos

Suelo				Estadístico	Error típ.
VAR00001	ácido	Media		6,0000	1,00000
		Intervalo de confianza para la media al 95%	Límite inferior	3,6940	
			Límite superior	8,3060	
	alcalino	Media		8,0000	1,77169
		Intervalo de confianza para la media al 95%	Límite inferior	3,9145	
			Límite superior	12,0855	

Prueba de homogeneidad de la varianza

	Estadístico de Levene	gl1	gl2	Sig.
VAR00001 Basándose en la media	4,692	1	16	,046

Descriptivos

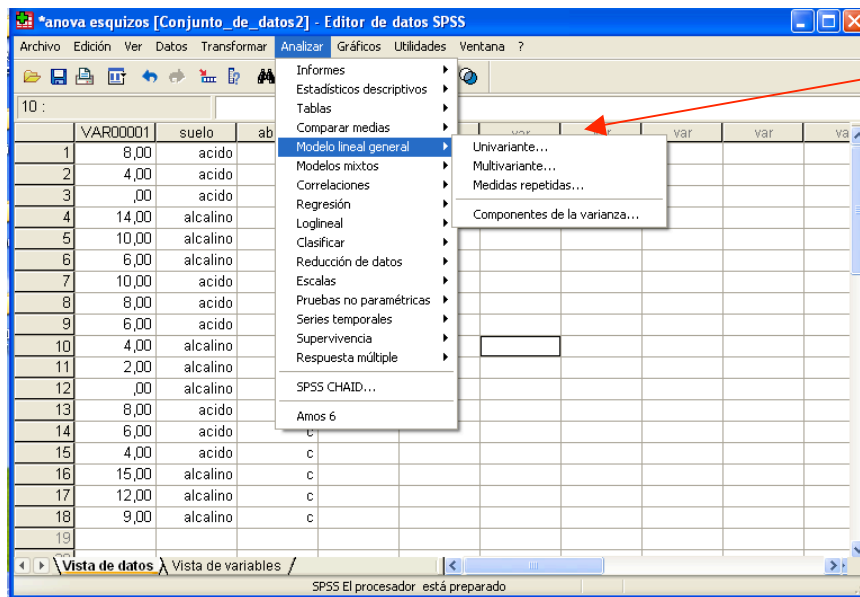
abono				Estadístico	Error típ.
VAR00001	a	Media		7,0000	1,98326
		Intervalo de confianza para la media al 95%	Límite inferior	1,9019	
			Límite superior	12,0981	
	b	Media		5,0000	1,52753
		Intervalo de confianza para la media al 95%	Límite inferior	1,0734	
			Límite superior	8,9266	
	c	Media		9,0000	1,63299
		Intervalo de confianza para la media al 95%	Límite inferior	4,8023	
			Límite superior	13,1977	

Prueba de homogeneidad de la varianza

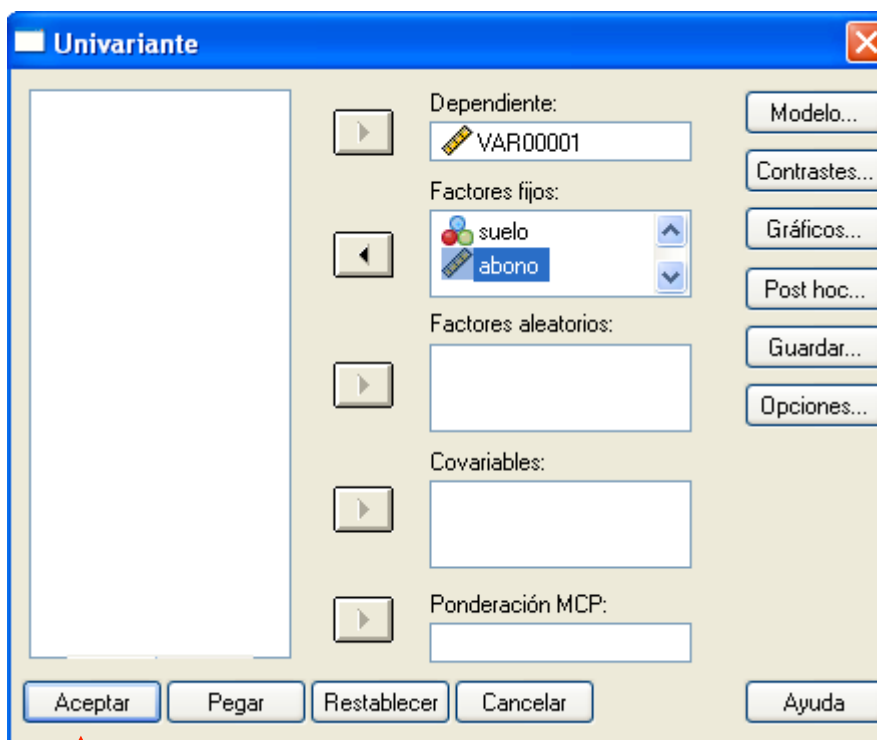
	Estadístico de Levene	gl1	gl2	Sig.
VAR00001 Basándose en la media	,168	2	15	,847

v

Una vez explorados los datos y analizados los requisitos previos para el análisis (normalidad y homoscedasticidad) procedemos a la realización del análisis de la varianza de dos vías con interacción



Seleccionamos de entre las variables del listado nuestra variable dependiente y los factores que queremos analizar: El programa permite diferenciar si los factores son fijos o son aleatorios.



Si seleccionamos el botón aceptar obtenemos la siguiente información:

Factores inter-sujetos

		Etiqueta del valor	N
suelo	1	acido	9
	2	alcalino	9
abono	1,00	a	6
	2,00	b	6
	3,00	c	6

Una primera tabla donde nos detalla para cada factor sus niveles y respectivos tamaños muestrales.

A continuación aparece, en el visor de resultados, una tabla resumen para el Anova donde se obtiene información para los contrastes de las distintas fuentes o factores de variación y para la interacción. Como ya sabemos el análisis de esta tabla depende totalmente del resultado del contraste para la interacción.

Además de los tres contrastes planteados en este análisis podemos efectuar un contraste para ver si el modelo completo (modelo corregido) explica de forma significativa una parte de la variabilidad de la variable dependiente. El R cuadrado de la base de la tabla nos informa del porcentaje de la variabilidad de la variable dependiente explicada por el modelo completo. El término intersección hace referencia a la constante del modelo.

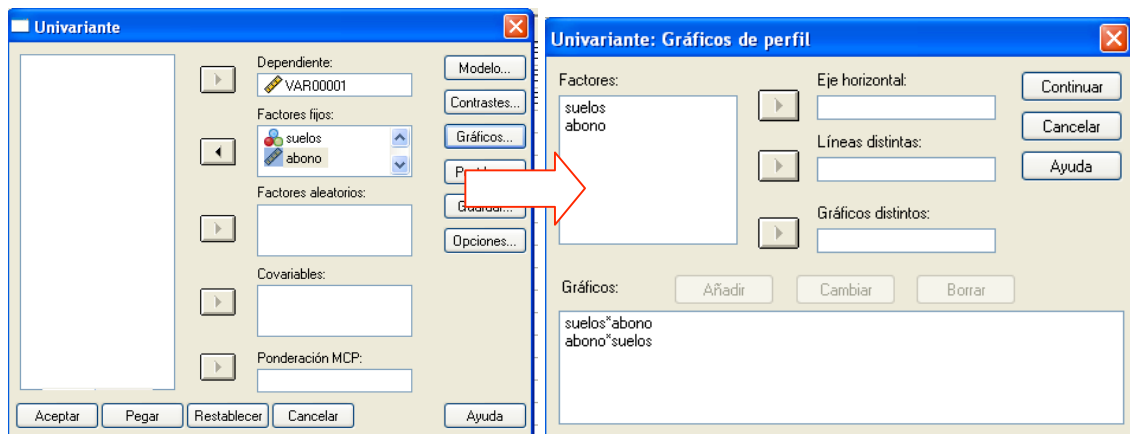
Pruebas de los efectos inter-sujetos

Variable dependiente: VAR00001

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo corregido	210,000(a)	5	42,000	4,755	,013
Intersección	882,000	1	882,000	99,849	,000
suelo	18,000	1	18,000	2,038	,179
abono	48,000	2	24,000	2,717	,106
suelo * abono	144,000	2	72,000	8,151	,006
Error	106,000	12	8,833		
Total	1198,000	18			
Total corregida	316,000	17			

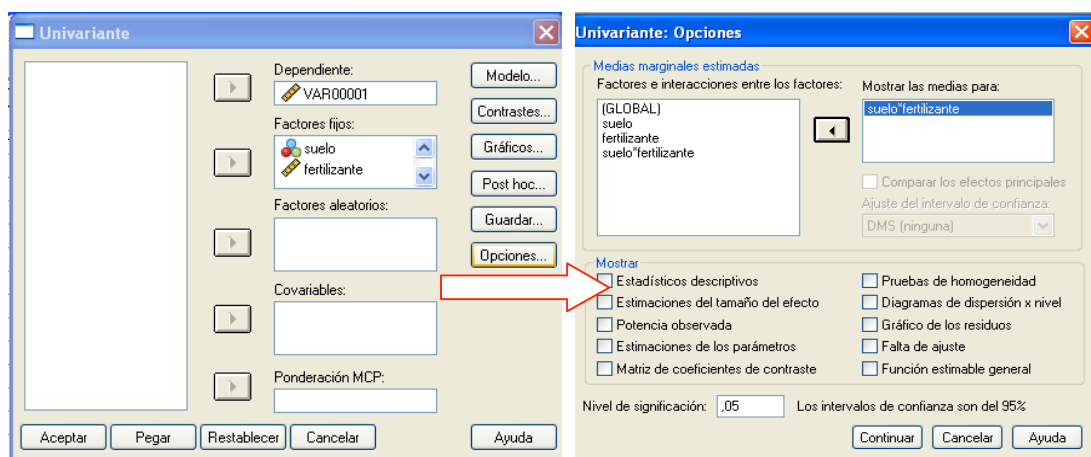
a R cuadrado = ,665 (R cuadrado corregida = ,525)

Si la interacción resulta significativa los gráficos de medias pueden resultar valiosos. para analizar lo que está ocurriendo en la realidad. Para obtenerlos utilizamos el botón de gráficos del menú principal del análisis:

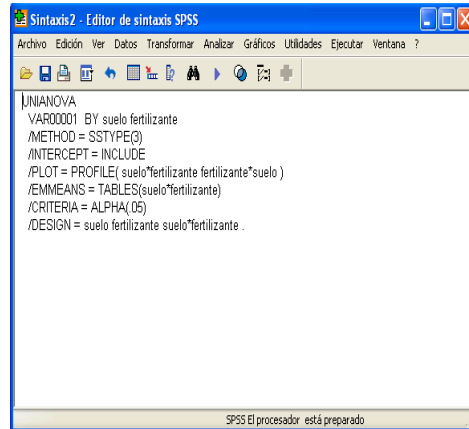
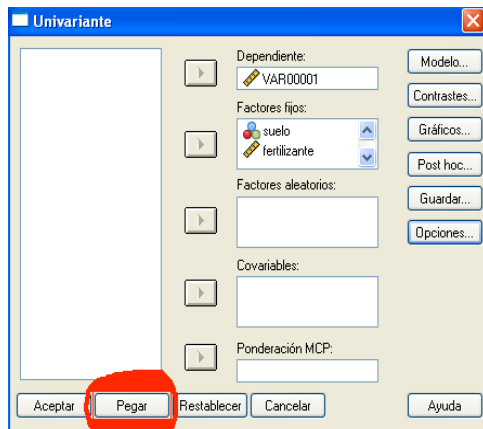


Así se observa cual es el significado del efecto interacción, pero una interpretación correcta del efecto requiere realizar comparaciones por separado para los niveles de un factor en cada uno de los niveles del otro factor, desafortunadamente el procedimiento univariante no permite de forma directa realizar estas comparaciones por lo que es necesario **utilizar la sintaxis de spss**:

1ª En el botón **opciones** seleccionamos mostrar las medias para l interacción y pular el botón continuar:



2º En el botón **pegar** hay que modificar la sintaxis de la siguiente forma:
 Modificar en la línea “**EMMEANS=TABLES(SUELOS*FERTILIZANTES)**” **añadiendo lo siguiente: COMPARE(fertilizante) ADJ (BONFERRONI)**



Lo que nos muestra en primer lugar es la tabla de medias para cada celdilla de la tabla (combinación de los niveles de los factores)

Variable dependiente: VAR00001

suelo	fertilizante	Intervalo de confianza al 95%.			
		Media	Error típ.	Límite inferior	Límite superior
ácido	a	4,000	1,716	,261	7,739
	b	8,000	1,716	4,261	11,739
	c	6,000	1,716	2,261	9,739
alcalino	a	10,000	1,716	6,261	13,739
	b	2,000	1,716	-1,739	5,739
	c	12,000	1,716	8,261	15,739

A continuación nos aparecen los contrastes (comparaciones por pares) entre fertilizantes para cada nivel por separado del factor suelo. Es decir la tabla donde podemos analizar realmente la interacción.

Comparaciones por pares

suelo	(I) fertilizante	(J) fertilizante	Diferencia entre medias (I-J)	Error típ.	Significación(a)	Intervalo de confianza al 95 % para la diferencia(a)	
						Límite inferior	Límite superior
ácido	a	b	-4,000	2,427	,376	-10,745	2,745
		c	-2,000	2,427	1,000	-8,745	4,745
	b	a	4,000	2,427	,376	-2,745	10,745
		c	2,000	2,427	1,000	-4,745	8,745
	c	a	2,000	2,427	1,000	-4,745	8,745
		b	-2,000	2,427	1,000	-8,745	4,745
alcalino	a	b	8,000(*)	2,427	,019	1,255	14,745
		c	-2,000	2,427	1,000	-8,745	4,745
	b	a	-8,000(*)	2,427	,019	-14,745	-1,255
		c	-10,000(*)	2,427	,004	-16,745	-3,255
	c	a	2,000	2,427	1,000	-4,745	8,745
		b	10,000(*)	2,427	,004	3,255	16,745

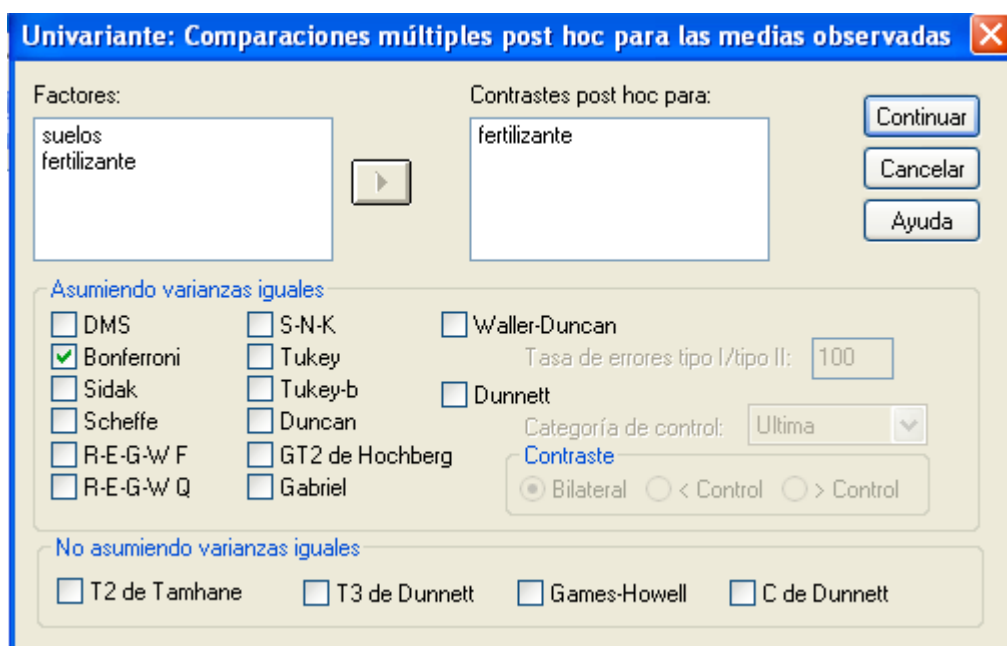
El Botón **modelo**: permite definir o bien el modelo factorial completo o bien un modelo en el que sólo se consideren los factores sin la interacción o si tuviésemos más factores las interacciones de orden que nos fueran de interés. Podría ser de utilidad para el caso en que la interacción no hubiese resultado significativa y deseáramos eliminarla del análisis.



En la **opción contrastes** es posible realizar contrastes planificados de antemano para los distintos factores del análisis de interés cuando la interacción no es significativa o cuando tenemos diseños sin interacción.

En la opción **post hoc** igual que en el análisis de la varianza de un factor es posible realizar contrastes no planificados de antemano para los distintos factores del análisis de interés cuando la interacción no es significativa o cuando tenemos diseños sin interacción.

Si en el **botón modelo** seleccionamos únicamente los dos factores de variación sin la interacción podemos luego seleccionar siempre que tenga sentido los contrastes post hoc para cada factor por separado. En nuestro caso como sólo fertilizante tiene más de dos niveles elegimos a este factor:



Pruebas post hoc

fertilizante

Comparaciones múltiples

Bonferroni

					Intervalo de confianza al 95%.	
(I) fertilizante	(J) fertilizante	Diferencia entre medias (I-J)	Error típ.	Significación	Límite inferior	Límite superior
a	b	2,0000	2,43975	1,000	-4,6306	8,6306
	c	-2,0000	2,43975	1,000	-8,6306	4,6306
b	a	-2,0000	2,43975	1,000	-8,6306	4,6306
	c	-4,0000	2,43975	,370	-10,6306	2,6306
c	a	2,0000	2,43975	1,000	-4,6306	8,6306
	b	4,0000	2,43975	,370	-2,6306	10,6306