



Universitat d'Alacant
Universidad de Alicante

***Métodos estadísticos
avanzados: Modelos de
Análisis de la varianza y
covarianza***

***Materiales para la
docencia no presencial***

Introducción

Este documento pretende guiarle de forma estructurada y secuencial a lo largo de un curso de introducción a algunos métodos de análisis estadístico multivariante con variable respuesta cuantitativa, particularmente de los procedimientos conocidos como Análisis de la varianza y covarianza.

Dispondrá de 4 sesiones teórico-prácticas en las que se prevee que realizando las actividades previstas debe alcanzar los objetivos formulados para la sesión. Se contempla revisión de conceptos teóricos con cuestiones, y teórico-prácticos con actividades de análisis de datos.

Las sesiones requieren disponer del cuadernillo de trabajo: Analisis de datos continuos_ANOVA-ANCOVA.pdf que puede Vd. obtener en la dirección:

<http://hdl.handle.net/10045/113344>

En el anexo dispone Vd. de las bases de datos necesarias para la realización de las actividades prácticas propuestas. Todas las actividades están pensadas para ser resueltas con SPSS® y deberá convertir los datos a este formato.

El autor le desea un aprendizaje fructífero

Andreu Nolasco



Contenidos

Sesión 0. Análisis de la varianza (ANOVA) de un factor

Sesión 1. Análisis de la varianza de dos o más factores

Sesión 2. Análisis de la covarianza (ANCOVA)

Sesión 3. Análisis de la varianza de medidas repetidas

Sesión 4. Práctica de revisión de Análisis de la varianza (ANOVA) y de la covarianza (ANCOVA)

Sesión 5. Práctica de revisión de Análisis de la varianza de medidas repetidas (ANOVAMR)



Sesión 0

Análisis de la varianza (ANOVA) de un factor



Sesión 0.- Análisis de la varianza (ANOVA) de una vía

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

- Identificar situaciones de investigación en las que pueda ser utilizado un ANOVA de una vía o factor
- Identificar el modelo de ANOVA con un único factor
- Realizar un ANOVA de un factor e interpretar sus resultados

Actividades: Las actividades previstas son las siguientes

- Revisión de los contenidos teóricos
- Realización de las actividades de autoaprendizaje resueltas y comprobación de resultados
- Realización opcional de actividades de autoaprendizaje para practicar

Contenidos teóricos: Material teórico pgs. 9-18 de la monografía Modelos de Análisis de la Varianza y Covarianza (<http://hdl.handle.net/10045/113344>)

Actividades de autoaprendizaje: Ver actividades de autoaprendizaje sesión 0

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Cuadernillo de prácticas, práctica 2
- Bases de datos: DATOS EJEMPLO TEXTO.sav



Sesión 0.- Actividades de autoaprendizaje (resueltas)

0.1. Diga si los siguientes enunciados son verdaderos (V) o falsos (F)

- i. El ANOVA de un factor fijo es un procedimiento cuyo objetivo principal es para comparar las varianzas de una variable cuantitativa según las categorías de un factor
- ii. En la mayoría de estudios epidemiológicos observacionales es posible utilizar métodos de ANOVA basados en un diseño equilibrado o balanceado
- iii. En ANOVA identificamos como efecto de un factor a la existencia de diferencias en las medias de la variable respuesta según sus categorías
- iv. En ANOVA, podemos comprobar las diferencias entre cualesquiera par de categorías del factor utilizando la prueba t de student, sin necesidad de ningún tipo de corrección
- v. En caso de comparar las medias entre todas las parejas de categorías del factor, es conveniente realizar una corrección a través de pruebas 'honestas' de comparación
- vi. En el caso de la cuestión v, el método de Bonferroni es uno de los posibles para corregir la significación de las diferencias
- vii. Si el requerimiento de homogeneidad de varianzas no se cumple, el método de Scheffé es el preferido en el caso de la cuestión v
- viii. El requerimiento de la normalidad de la variable respuesta en cada subgrupo generado por las categorías del factor puede ser comprobado a través de la normalidad de los residuos del modelo de ANOVA
- ix. El método Tipo III de descomposición de la variabilidad para la obtención de sumas de cuadrados necesarias en ANOVA se utiliza normalmente para cualquier modelo equilibrado o desequilibrado sin casillas vacías
- x. En general, los estadísticos para los contrastes de hipótesis en ANOVA suelen seguir la distribución F de Snedecor

0.2. Resuelva las siguientes aplicaciones: PRÁCTICA 2 del cuadernillo 'PRACTICAS_CURSO_DATOS_CONTINUOS.pdf'

A partir del archivo DATOS EJEMPLOS TEXTO.SAV. Queremos evaluar el efecto de la edad (categorizada en grupos) sobre el ácido úrico.

Abra el archivo DATOS EJEMPLOS TEXTO.SAV. En él encontrará las variables

EDAD: En años	CALCIO: mg/100ml
PESO: En Kg.	URICO: mg/litro
COLEST: mg/100ml	ALCOHOL: 1 'bajo' 2'moderado/alto'
ALBUMINA: mg/100ml	EDADREC: 1 '≤30' 2 '30-40' 3 '>40'

- i. Inspeccione los estadísticos descriptivos de ácido úrico según grupos de edad
- ii. Realice un ANOVA para comprobar si existen diferencias en las medias de ácido úrico según grupos de edad. Utilice el procedimiento ANOVA de un factor del comando MEDIAS del SPSS.
- iii.. Compruebe entre qué grupos de edad existen diferencias con los tests de TUKEY y de SCHEFFÉ.



- iv. Resuelva el problema a través de comparaciones múltiples con la corrección de BONFERRONI
- v. Compruebe los requerimientos del ANOVA (normalidad y homogeneidad de varianzas)
- vi. Dada la falta de homogeneidad de varianzas, realice transformaciones logarítmica y de raíz cuadrada y vuelva a resolver el ANOVA y la comprobación de requerimientos

0.3. Ejercicio práctico propuesto (para practicar)

Repita los puntos anteriores para las variables PESO, COLEST, CALCIO y ALBUMINA. ¿En cuáles de estas variables hay diferencia de medias según grupos de edad? ¿Entre qué grupos de edad? ¿Se cumplen los requerimientos?



Sesión 0.- Soluciones a las actividades de autoaprendizaje

0.1 Cuestiones

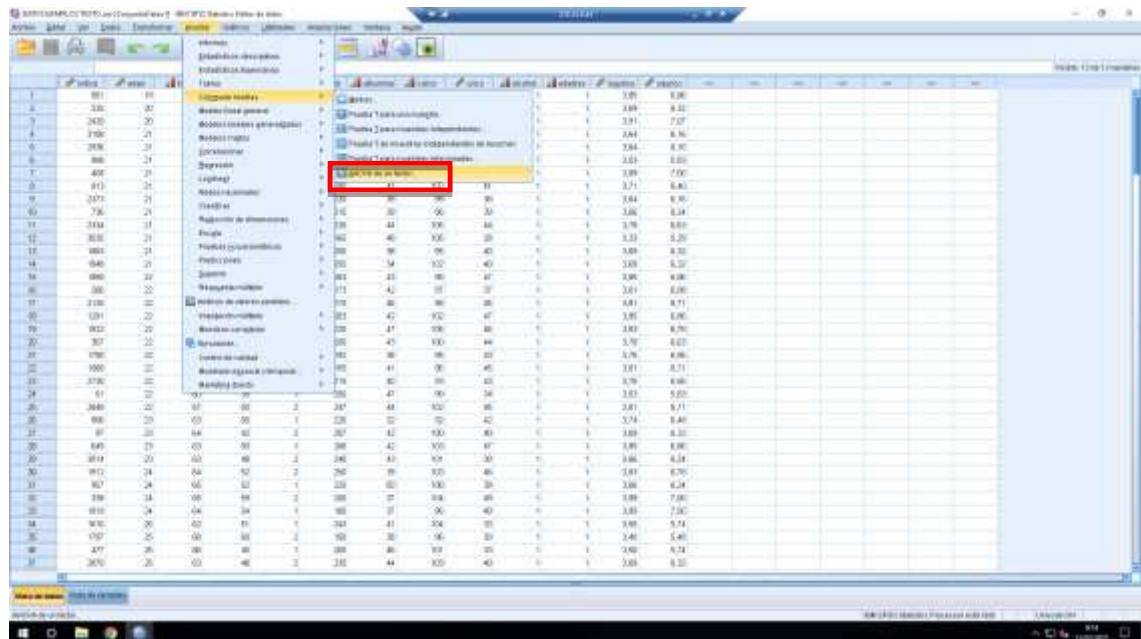
Cuestión	i	ii	iii	iv	v	vi	vii	viii	ix	x
Respuesta	F	F	V	F	V	V	F	V	V	V



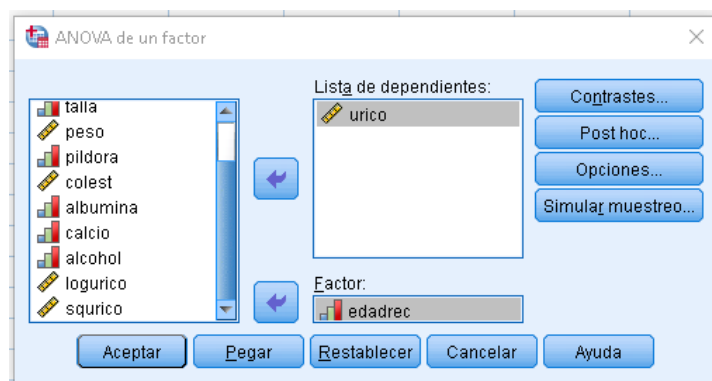
0.2 Aplicaciones

Consideraciones previas. Cómo aplicar el procedimiento ANOVA de un factor con SPSS

Aunque en las sesiones siguientes de verá que para todos los análisis de ANOVA, ANCOVA y ANOVA de medidas repetidas se utilizará el procedimiento 'Modelo lineal general' del SPSS, el ANOVA de una vía o factor puede ser también resuelto (de manera un poca más sintética) a través del procedimiento Comparar Medias:



REC y
ra ello



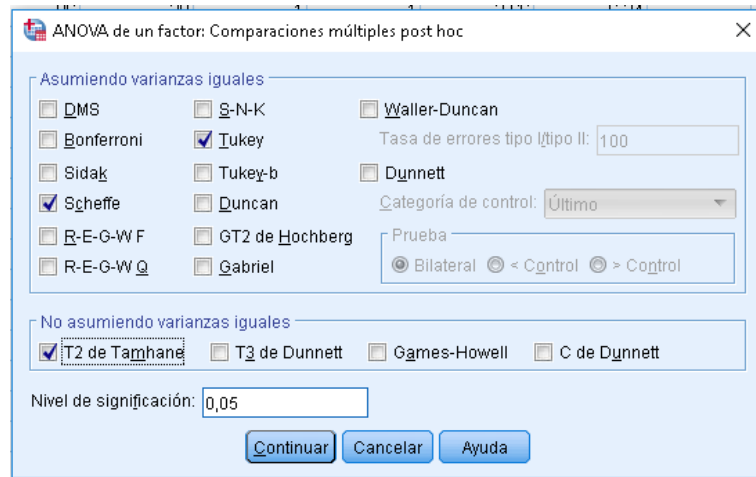
usando URICO como variable dependiente/respuesta y EDADREC como factor/variable explicativa.

A continuación definiremos un análisis básico, con todos los elementos para realizar, para ir contestando con los resultados las preguntas del ejercicio. Revisemos que definir en cada pestaña:

CONTRASTES: No se utiliza en este procedimiento. Se utilizará en ANOVA de medidas repetidas

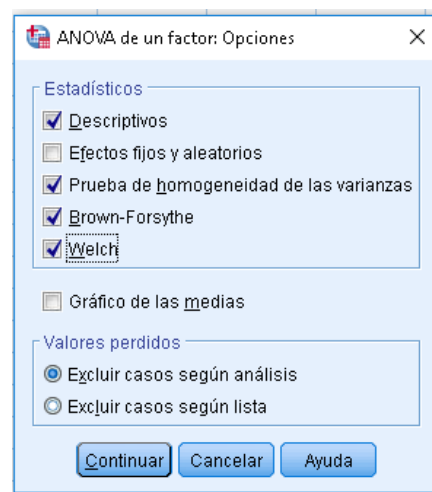
POSTHOC: Se pueden seleccionar pruebas para las comparaciones múltiples de las medias según categorías de la variable explicativa (2 a 2):





Se ha seleccionado las pruebas de Tukey (n,s iguales) y Scheffé (n,s diferentes) para varianzas iguales (son las pruebas de mayor potencia y más utilizadas) y T2 de Tamhane para varianzas desiguales (es la prueba más conservadora)

OPCIONES: En esta pestaña se pueden seleccionar diversos elementos. Elegimos los estadísticos descriptivos, la prueba de homogeneidad de varianzas y los estadísticos de Brown-Forsythe y Welch para testar la hipótesis nula de igualdad de medias si las varianzas fueran distintas



Resolución del ejercicio

i. Inspeccione los estadísticos descriptivos de ácido úrico según grupos de edad

De la salida obtenida con las selecciones anteriores tomamos la tabla de estadísticos descriptivos:

Descriptivos								
urico								
	N	Media	Desv. Desviación	Desv. Error	95% del intervalo de confianza para la media		Mínimo	Máximo
					Límite inferior	Límite superior		
Edad<30	74	46,59	10,794	1,255	44,09	49,10	24	84
30<=Edad<40	54	44,28	10,267	1,397	41,48	47,08	22	66
40<=Edad	59	58,37	13,538	1,762	54,84	61,90	25	98
Total	187	49,64	13,006	,951	47,77	51,52	22	98

Observamos como el ácido úrico sube en media a partir de 40 años. También es más variable en este grupo.

ii. Realice un ANOVA para comprobar si existen diferencias en las medias de ácido úrico según grupos de edad. Utilice el procedimiento ANOVA de un factor del comando MEDIAS del SPSS.

En la salida de SPSS obtenemos:

Prueba de homogeneidad de varianzas					
		Estadístico de Levene	gl1	gl2	Sig.
urico	Se basa en la media	4,350	2	184	,014
	Se basa en la mediana	3,870	2	184	,023
	Se basa en la mediana y con gl ajustado	3,870	2	181,312	,023
	Se basa en la media recortada	4,445	2	184	,013

ANOVA					
urico					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	6738,527	2	3369,263	25,076	,000
Dentro de grupos	24722,468	184	134,361		
Total	31460,995	186			

Pruebas robustas de igualdad de medias					
urico					
	Estadístico ^a	gl1	gl2	Sig.	
Welch	21,274	2	116,274	,000	
Brown-Forsythe	24,881	2	164,719	,000	

a. F distribuida de forma asintótica

La prueba de Levene de comparación de varianzas lleva a concluir varianzas diferentes ($p=0,014$). Por ello, la significación de la tabla ANOVA es sólo orientativa, ya que rechazamos el requerimiento de homogeneidad de varianzas. Sin embargo las pruebas robustas, que no requieren homogeneidad de varianzas (Welch y Brown-Forsythe) son significativas ($p<0,001$) y permiten rechazar la hipótesis



de igualdad de medias, concluyendo que las medias de ácido úrico son significativamente diferentes entre los grupos de edad

iii.. Compruebe entre qué grupos de edad existen diferencias con los tests de TUKEY y de SCHEFFÉ.

Usando la salida correspondiente al POST-HOC, comparaciones múltiples,

Comparaciones múltiples

Variable dependiente: urico

				Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
		(I) edadrec	(J) edadrec				Límite inferior	Límite superior
HSD Tukey	Edad<30	30<=Edad<40		2,317	2,075	,505	-2,59	7,22
			40<=Edad	-11,778	2,023	,000	-16,56	-7,00
	30<=Edad<40	Edad<30		-2,317	2,075	,505	-7,22	2,59
			40<=Edad	-14,095	2,183	,000	-19,25	-8,94
	40<=Edad	Edad<30		11,778	2,023	,000	7,00	16,56
			30<=Edad<40	14,095	2,183	,000	8,94	19,25
Scheffe	Edad<30	30<=Edad<40		2,317	2,075	,537	-2,80	7,44
			40<=Edad	-11,778	2,023	,000	-16,77	-6,79
	30<=Edad<40	Edad<30		-2,317	2,075	,537	-7,44	2,80
			40<=Edad	-14,095	2,183	,000	-19,48	-8,71
	40<=Edad	Edad<30		11,778	2,023	,000	6,79	16,77
			30<=Edad<40	14,095	2,183	,000	8,71	19,48
Tamhane	Edad<30	30<=Edad<40		2,317	1,878	,525	-2,23	6,87
			40<=Edad	-11,778	2,164	,000	-17,02	-6,53
	30<=Edad<40	Edad<30		-2,317	1,878	,525	-6,87	2,23
			40<=Edad	-14,095	2,249	,000	-19,55	-8,64
	40<=Edad	Edad<30		11,778	2,164	,000	6,53	17,02
			30<=Edad<40	14,095	2,249	,000	8,64	19,55

*. La diferencia de medias es significativa en el nivel 0.05.

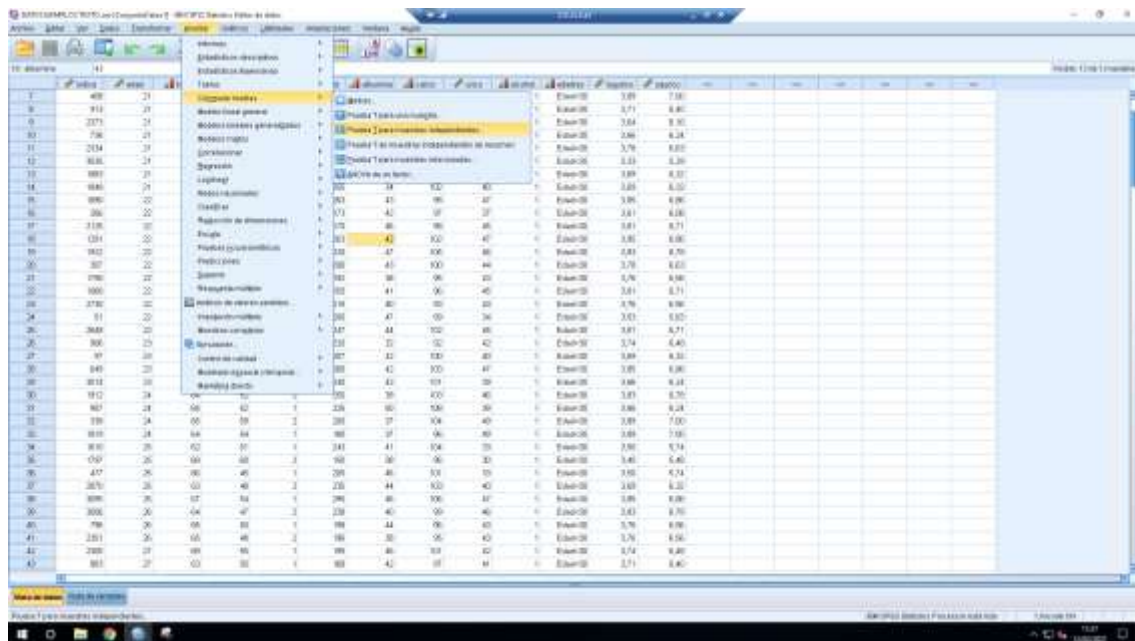
Tanto Tukey, Scheffé y Tamhane detectan que las diferencias significativas se producen entre el grupo de >40 años y el resto de grupos. Sin embargo, aunque coincidan en este caso, Tukey y Scheffé requieren homogeneidad de varianzas (su resultado sería sólo orientativo, no riguroso), que hemos visto que no se da, mientras que Tamhane no la requiere (resultado correcto).

iv. Resuelva el problema a través de comparaciones múltiples con la corrección de BONFERRONI

Lo que se pretende con este ejercicio es realizar las comparaciones 2 a 2 con la prueba t de student para dos grupos, pero aplicando la corrección de Bonferroni, que dice que si deseamos un nivel de significación 0,05, tendremos que resolver cada comparación con nivel $0,05/n^{\circ}$ de comparaciones. Con 3 grupos en número de comparaciones es 3 (G1 con G2, G1 con G3 y G2 con G3) y el nivel a aplicar para cada comparación será $0,05/3=0,017$. Veamos:

Utilizando la prueba t para grupos independientes:





La variable de prueba es el ácido úrico y vamos a compararlo en primer lugar entre los grupos 1 y 2 de edadrec (eran <30 y 30-40 años).
Obtenemos los siguientes resultados:

Estadísticas de grupo					
	edadrec	N	Media	Desv. Desviación	Desv. Error promedio
urico	Edad<30	74	46,59	10,794	1,255
	30<=Edad<40	54	44,28	10,267	1,397

Prueba de muestras independientes									
		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia
									Inferior Superior
urico	Se asumen varianzas iguales	,191	,663	1,224	126	,223	2,317	1,893	-1,429 6,063
	No se asumen varianzas iguales			1,234	117,487	,220	2,317	1,878	-1,402 6,036

Se aceptan varianzas iguales y no hay diferencia significativa en las medias de ácido úrico entre estos grupos ($p=0,223$)

Repetimos para las otras dos comparaciones:



<30 años con >=40 años

Estadísticas de grupo					
	edadrec	N	Media	Desv. Desviación	Desv. Error promedio
urico	Edad<30	74	46,59	10,794	1,255
	40<=Edad	59	58,37	13,538	1,762

Prueba de muestras independientes									
		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia Inferior Superior
urico	Se asumen varianzas iguales	7,162	.003	-5,584	131	.000	-11,778	2,109	-15,951 -7,605
	No se asumen varianzas iguales			-5,444	109,374	.000	-11,778	2,164	-16,066 -7,490

Se rechazan varianzas iguales y sí que hay diferencia significativa $p > 0,001$ y por tanto $p < 0,017$ (valor de Bonferroni)

30-40 años con >=40 años

Estadísticas de grupo					
	edadrec	N	Media	Desv. Desviación	Desv. Error promedio
urico	30<=Edad<40	54	44,28	10,267	1,397
	40<=Edad	59	58,37	13,538	1,762

Prueba de muestras independientes									
		Prueba de Levene de igualdad de varianzas		prueba t para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia Inferior Superior
urico	Se asumen varianzas iguales	5,262	.024	-6,192	111	.000	-14,095	2,276	-18,606 -9,584
	No se asumen varianzas iguales			-6,267	107,391	.000	-14,095	2,249	-18,553 -9,637

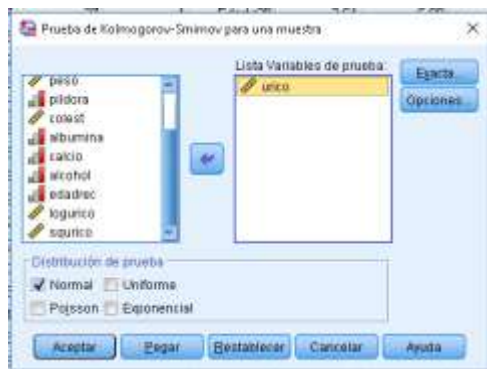
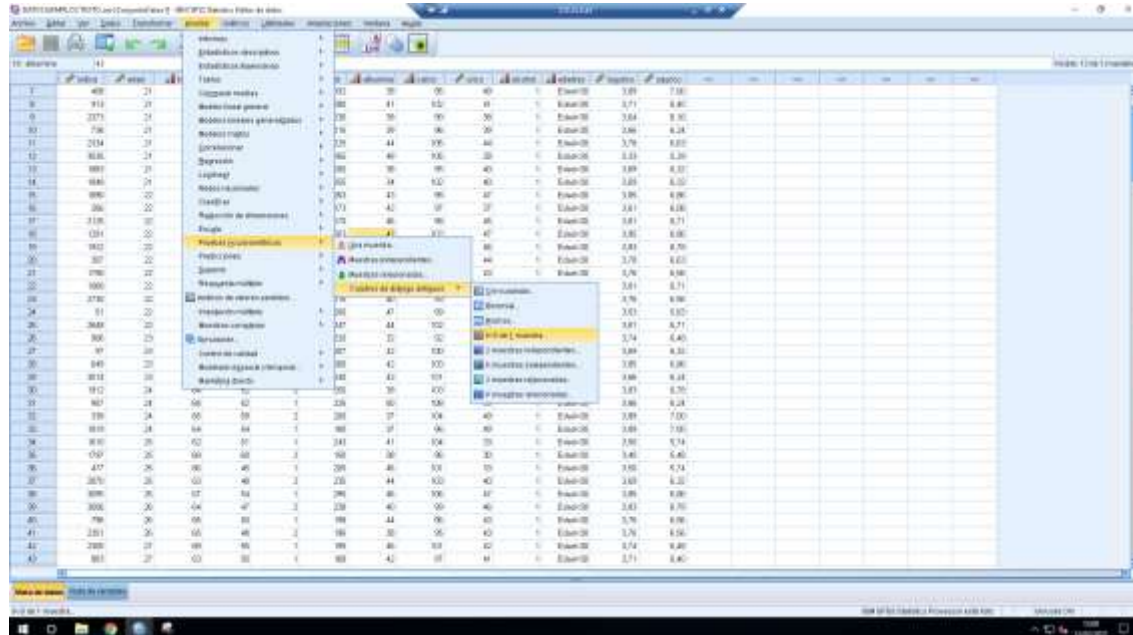
También se rechazan varianzas iguales y el resultado para las medias es como la comparación anterior, difieren significativamente $p < 0,001$, y por tanto $p < 0,017$



v. Compruebe los requerimientos del ANOVA (normalidad y homogeneidad de varianzas)

En i. fue comprobado que las varianzas no podrían suponerse iguales

Respecto a la normalidad, realizamos una prueba de Kolmogorov-Smirnov a la variable úrico:



Prueba de Kolmogorov-Smirnov para una muestra

		úrico
Parámetros normales ^{a,b}	N	187
	Media	49,64
	Desv.	13,006
	Desviación	
Máximas diferencias extremas	Absoluto	,113
	Positivo	,113
	Negativo	-,059
Estadístico de prueba		,113
Sig. asintótica(bilateral)		,000 ^c

- a. La distribución de prueba es normal.
 b. Se calcula a partir de datos.
 c. Corrección de significación de Lilliefors.

Resultando que se rechaza la normalidad con $p < 0,001$

vi. Dada la falta de homogeneidad de varianzas y normalidad, realice transformaciones logarítmica y de raíz cuadrada y vuelva a resolver el ANOVA y la comprobación de requerimientos

Vamos a repetir el análisis de comparación de medias (ANOVA) pero con el logaritmo neperiano del ácido úrico, LOGURICO en la base de datos.

Tendremos:



Descriptivos

logurico

	N	Media	Desv. Desviación	Desv. Error	95% del intervalo de confianza para la media		Mínimo	Máximo
					Límite inferior	Límite superior		
Edad<30	74	3,8173	,21880	,02543	3,7666	3,8680	3,18	4,43
30<=Edad<40	54	3,7643	,23203	,03158	3,7010	3,8276	3,09	4,19
40<=Edad	59	4,0387	,24566	,03198	3,9747	4,1027	3,22	4,58
Total	187	3,8718	,25753	,01883	3,8347	3,9090	3,09	4,58

Prueba de homogeneidad de varianzas

	Estadístico de Levene	gl1	gl2	Sig.
logurico				
Se basa en la media	1,172	2	184	,312
Se basa en la mediana	,852	2	184	,428
Se basa en la mediana y con gl ajustado	,852	2	182,480	,428
Se basa en la media recortada	1,141	2	184	,322

ANOVA

logurico

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	2,488	2	1,244	23,238	,000
Dentro de grupos	9,848	184	,054		
Total	12,336	186			

Pruebas robustas de igualdad de medias

logurico

	Estadístico ^a	gl1	gl2	Sig.
Welch	21,468	2	115,996	,000
Brown-Forsythe	22,921	2	171,825	,000

a. F distribuida de forma asintótica

Comparaciones múltiples

Variable dependiente: logurico

	(I) edadrec	(J) edadrec	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
HSD Tukey	Edad<30	30<=Edad<40	,05300	,04141	,408	-,0448	,1508
		40<=Edad	-,22141*	,04038	,000	-,3168	-,1260
	30<=Edad<40	Edad<30	-,05300	,04141	,408	-,1508	,0448
		40<=Edad	-,27441*	,04357	,000	-,3774	-,1715
	40<=Edad	Edad<30	,22141*	,04038	,000	,1260	,3168
		30<=Edad<40	,27441*	,04357	,000	,1715	,3774
Scheffe	Edad<30	30<=Edad<40	,05300	,04141	,442	-,0492	,1552
		40<=Edad	-,22141*	,04038	,000	-,3211	-,1218
	30<=Edad<40	Edad<30	-,05300	,04141	,442	-,1552	,0492
		40<=Edad	-,27441*	,04357	,000	-,3819	-,1669
	40<=Edad	Edad<30	,22141*	,04038	,000	,1218	,3211
		30<=Edad<40	,27441*	,04357	,000	,1669	,3819
Tamhane	Edad<30	30<=Edad<40	,05300	,04055	,476	-,0453	,1513
		40<=Edad	-,22141*	,04086	,000	-,3204	-,1224
	30<=Edad<40	Edad<30	-,05300	,04055	,476	-,1513	,0453
		40<=Edad	-,27441*	,04494	,000	-,3834	-,1655
	40<=Edad	Edad<30	,22141*	,04086	,000	,1224	,3204
		30<=Edad<40	,27441*	,04494	,000	,1655	,3834

*. La diferencia de medias es significativa en el nivel 0.05.

Resultados similares a los obtenidos con URICO, pero con la diferencia que no se puede rechazar la homogeneidad de varianzas ($p=0,312$).



Si realizamos la prueba de normalidad, tenemos que:

Prueba de Kolmogorov-Smirnov para una muestra

		logurico
N		187
Parámetros normales ^{a,b}	Media	3,8718
	Desv. Desviación	,25753
Máximas diferencias extremas	Absoluto	,063
	Positivo	,063
	Negativo	-,049
Estadístico de prueba		,063
Sig. asintótica(bilateral)		,064 ^c

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

No pudiendo rechazar tampoco la normalidad

Por tanto para esta variable podemos suponer ambos requerimientos.

Sesión 1

Análisis de la varianza de dos o más factores



Sesión 1.- Análisis de la varianza (ANOVA)

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

Identificar situaciones de investigación en las que deba ser utilizado un ANOVA
Identificar los modelos de ANOVA con varios factores
Realizar un ANOVA e interpretar sus resultados
Identificar las relaciones/asociaciones entre las variables explicativas y respuesta
Identificar los efectos de confusión e interacción entre las variables del ANOVA

Actividades: Las actividades previstas son las siguientes

Revisión de los contenidos teóricos
Realización de las actividades de autoaprendizaje resueltas y comprobación de resultados
Realización opcional de actividades de autoaprendizaje para practicar

Contenidos teóricos: Material teórico pgs. 20-29 de la monografía Modelos de Análisis de la Varianza y Covarianza, <http://hdl.handle.net/10045/113344>

Actividades de autoaprendizaje: Ver actividades de autoaprendizaje sesión 1

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Cuadernillo de prácticas, prácticas 3 y 4
- Bases de datos: DATOS EJEMPLO TEXTO.sav; ANOVAS.sav



Sesión 1.- Actividades de autoaprendizaje (resueltas)

1.1. Diga si los siguientes enunciados son verdaderos (V) o falsos (F)

- i. El ANOVA de dos factores fijos es un procedimiento para comparar las medias de una variable cuantitativa según las categorías de cada uno de los factores y de la combinación de categorías de los factores
- ii. En la mayoría de estudios epidemiológicos observacionales es posible utilizar métodos de ANOVA basados en un diseño equilibrado o balanceado
- iii. En ANOVA identificamos como efecto de un factor a la existencia de diferencias en las medias de la variable respuesta según sus categorías
- iv. En ANOVA de dos factores sólo puede haber efecto de confusión cuando las diferencias entre niveles de un factor no son las mismas según las categorías del otro factor
- v. En ANOVA de dos factores, ambos con dos categorías, si graficamos las medias de la variable respuesta según categorías de uno de los factores y separando las rectas que las unen según las categorías del otro factor, el modelo de rectas paralelas separadas, que a su vez son paralelas al eje de abscisas, representa una situación de efecto de un único factor
- vi. En el caso de la cuestión v, el modelo de interacción se representaría como dos rectas paralelas con inclinación (no paralelas al eje de abscisas)
- vii. El requerimiento de la normalidad de la variable respuesta en cada subgrupo generado por los cruces de las categorías de los factores puede ser comprobado a través de la normalidad de los residuos del modelo de ANOVA
- viii. El método Tipo III de descomposición de la variabilidad para la obtención de sumas de cuadrados necesarias en ANOVA se utiliza normalmente para cualquier modelo equilibrado o desequilibrado sin casillas vacías
- ix. La existencia de interacción establece un efecto conjunto de los factores, que interactúan para producir comportamientos específicos y diferenciados según sus niveles conjuntos. La interacción es un efecto de rango superior a la existencia de diferencias de primer orden (para uno u otro de los factores)
- x. En general, los estadísticos para los contrastes de hipótesis en ANOVA suelen seguir la distribución F de Snedecor

1.2. Resuelva las siguientes aplicaciones: PRÁCTICA 3 del cuadernillo 'Prácticas_curso_datos_continuos.pdf'

A partir del archivo DATOS EJEMPLOS TEXTO.SAV. Queremos evaluar el efecto de la edad (categorizada en grupos) y el alcohol sobre el ácido úrico. Abra el archivo DATOS EJEMPLOS TEXTO.SAV. En él encontrará las variables

EDAD: En años	CALCIO: mg/100ml
PESO: En Kg.	URICO: mg/litro
COLEST: mg/100ml	ALCOHOL: 1 'bajo' 2'moderado/alto'
ALBUMINA: mg/100ml	EDADREC: 1 '≤30' 2 '30-40' 3 '>40'

- i. Inspeccione los estadísticos descriptivos de ácido úrico según grupos de edad y consumo de alcohol. ¿Qué le sugieren estos resultados?



- ii. Realice un ANOVA de dos factores para comprobar los efectos existentes. ¿Hay interacción? ¿Hay efecto de la edad? ¿Hay efecto del alcohol?
- iii. Realice comparaciones múltiples con los tests de TUKEY y SCHEFFÉ
- iv. Calcule intervalos de confianza al 95% para la diferencia de medias de ácido úrico entre grupos de edad. Idem entre grupos de alcohol (intervalos ajustados por comparaciones múltiples)
- v. Compruebe los requerimientos del modelo
- vi. Puesto que la normalidad de los residuos es rechazable, repita el ANOVA anterior con transformaciones de la variable respuesta (logaritmo y raíz cuadrada)

1.3. Ejercicio práctico propuesto (para practicar)

Repita los puntos i a iv del punto 1.2 anterior tomando como variables respuesta PESO y COLEST



Sesión 1.- Soluciones a las actividades de autoaprendizaje

1.1 Cuestiones

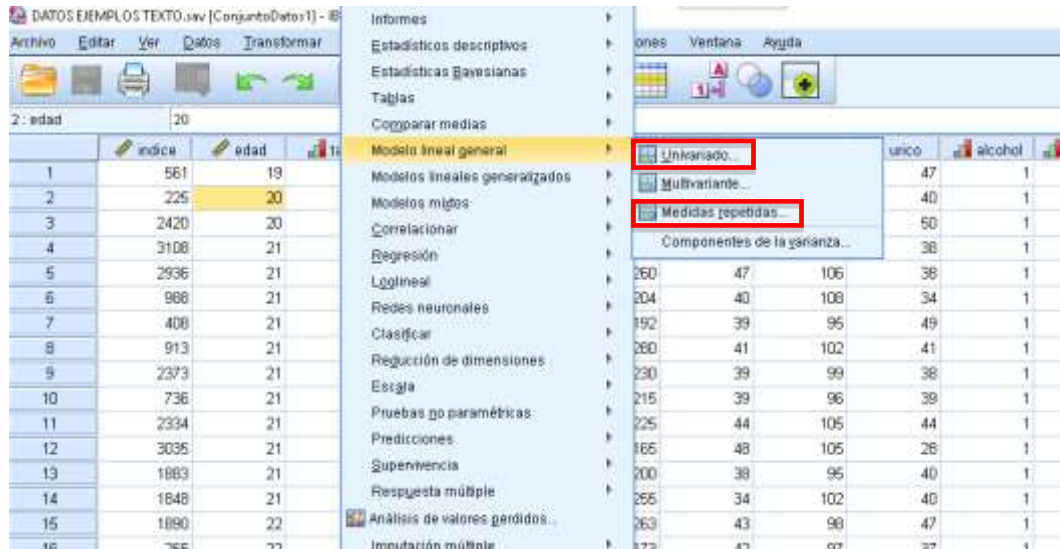
Cuestión	i	ii	iii	iv	v	vi	vii	viii	ix	x
Respuesta	V	F	V	F	V	F	V	V	V	V



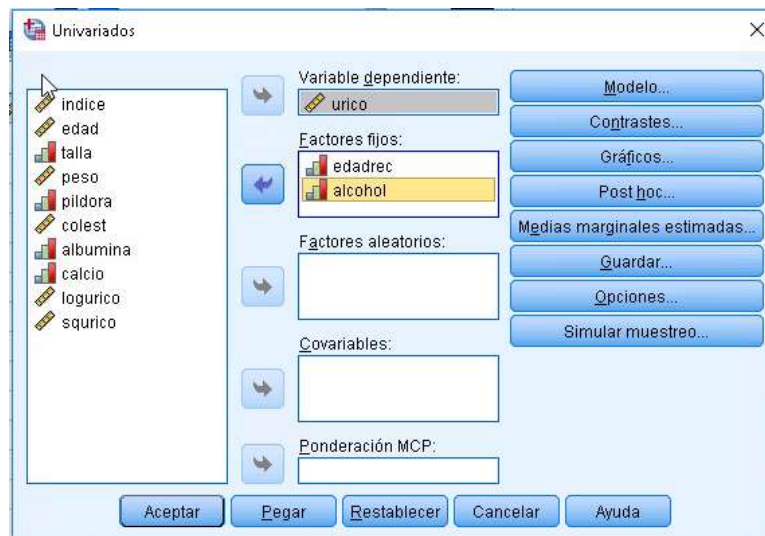
1.2 Aplicaciones

Consideraciones previas. Cómo aplicar el procedimiento ANOVA con SPSS

Para todos los análisis de ANOVA, ANCOVA y ANOVA de medidas repetidas se utilizará el procedimiento 'Modelo lineal general' del SPSS, usando las pestañas/procedimientos 'Univariado' (caso de ANOVA Y ANCOVA) y 'Medidas repetidas' (caso de ANOVA de medidas repetidas):



Dado que la variable respuesta en continua (URICO) y las explicativas son categóricas (EDADREC y ALCOHOL), resolveremos las preguntas de todo el ejercicio con ANOVA de factores fijos. Para ello entramos en **Univariado** y definimos:



usando URICO como variable dependiente/respuesta y EDADREC y ALCOHOL como factores fijos/explicativas.

A continuación definiremos un análisis básico, con todos los elementos para realizar, para ir contestando con los resultados las preguntas del ejercicio. Revisemos que definir en cada pestaña:

MODELO: Permite seleccionar qué efectos incluirá nuestro modelo:

Por defecto figura el modelo factorial completo. Este modelo incluye todos los efectos posibles que, en nuestro caso serán el de ALCOHOL, el de EDADREC y el de la INTERACCION entre ambas. Puede utilizarse la opción **Construir términos** para definir el modelo que queramos, por ejemplo, el modelo factorial completo seleccionando nosotros los términos sería:

pasando la interacción como ALCOHOL*EDADREC, lo cual se consigue seleccionando ambas juntas y pulsando el pulsador intermedio que indicará Tipo interacción.

Otro aspecto de esta pantalla en que aparece la **Suma de cuadrados: Tipo III**. Esta no debe ser modificada, ya que en ANOVA es la descomposición de variabilidad recomendada

CONTRASTES: No se utiliza en este procedimiento. Se utilizará en ANOVA de medidas repetidas



GRÁFICOS: Podemos seleccionar gráficos de las medias según categorías o combinación de categorías de hasta 3 variables. En las imágenes siguientes vemos la selección de las variables y, tras pulsar el botón **Añadir** como pasan a la lista de gráficos pedida. En este caso también se ha pedido las barras de error (intervalo de confianza 95% para las medias):

POSTHOC: Se pueden seleccionar pruebas para las comparaciones múltiples de las medias según categorías de las variables explicativas. Para ANOVA multivariante (2 o más factores) sólo se pueden seleccionar las pruebas con homogeneidad de varianzas:

MEDIAS MARGINALES ESTIMADAS: Esta opción sólo se aplica sin el modelo que estamos estimando no es el factorial completo, es decir, tiene una reducción de efectos. En el modelo factorial completo las medias estimadas coinciden con las observadas, ya que el modelo reproduce las medias exactamente

GUARDAR: Pueden seleccionarse diversas variables resultado para añadir las a la base de datos. En este caso se seleccionan los residuos para ver su normalidad:

Univariados: Guardar

Valores pronosticados

- ☐ No estandarizados
- ☐ Ponderados
- ☐ Error estándar

Residuos

- ☐ No estandarizados
- ☐ Ponderados
- ☒ Estandarizados
- ☐ Método de Student
- ☐ Eliminados

Diagnósticos

- ☐ Distancia de Cook
- ☐ Valores de influencia

Estadísticos de los coeficientes

- ☒ Crear estadísticos de los coeficientes

Tipo

- ☒ Estadísticas estándar
- ☐ Estadísticas coherentes con heterocedasticidad

Requiere la selección de estimaciones de parámetro con errores estándar robustos en Opciones.

Destino

- ☒ Crear un nuevo conjunto de datos
- ☐ Escribir un nuevo archivo de datos

Nombre de conjunto de datos:

Archivo...

Continuar Cancelar Ayuda

OPCIONES: En esta pestaña se pueden seleccionar diversos elementos. Elegimos los estadísticos descriptivos y la prueba de homogeneidad de varianzas para el modelo propuesto

Univariados: Opciones

Visualización

- ☒ Estadísticos descriptivos
- ☐ Estimaciones del tamaño del efecto
- ☐ Potencia observada
- ☐ Estimaciones de los parámetros
- ☐ Matriz de coeficientes del contraste
- ☒ Pruebas de homogeneidad
- ☐ Gráfico de dispersión versus gráfico de niveles
- ☐ Gráfico de residuos
- ☐ Falta de ajuste
- ☐ Función estimable general

Pruebas de heterocedasticidad

- ☐ Test de Breusch-Pagan modificado
- ☐ Prueba F
- ☐ Test de Breusch-Pagan
- ☐ Prueba de White

Modelo...

Modelo...

Estimaciones de parámetro con errores estándar robustos

- ☐ HC0
- ☐ HC1
- ☐ HC2
- ☐ HC3
- ☒ HC4

Nivel de significación: 0,05 Los intervalos de confianza son 95,0 %

Continuar Cancelar Ayuda

SIMULAR MUESTREO (BOOTSTRAPPING): Activando esta opción se estimarían por simulación ciertos intervalos de confianza de parámetros del modelo (en caso de haberlos solicitado).

Resolución del ejercicio

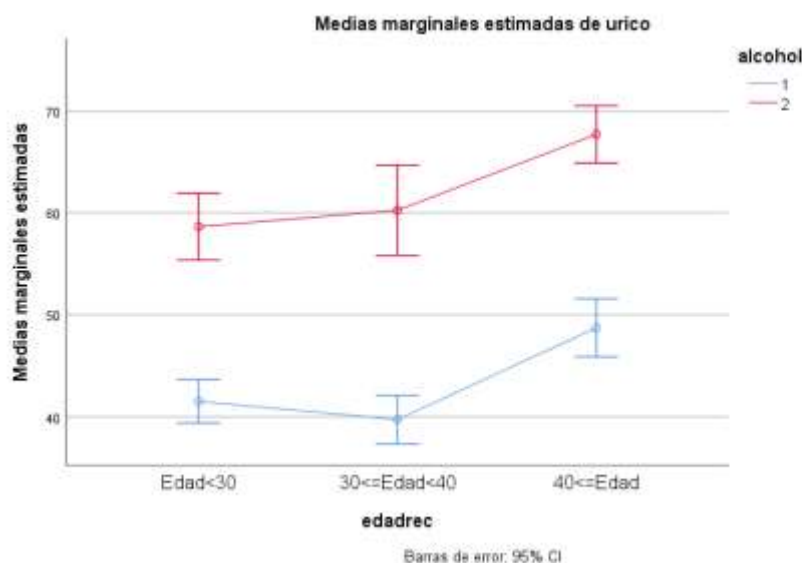
i. Utilizando directamente el procedimiento para realizar un ANOVA, inspeccione los estadísticos descriptivos de ácido úrico según grupos de edad y consumo de alcohol. ¿Qué le sugieren estos resultados?

De la salida obtenida con las selecciones anteriores tomamos la tabla de estadísticos descriptivos y los gráficos:

Estadísticos descriptivos

Variable dependiente: urico

edadrec	alcohol	Media	Desv. Desviación	N
Edad<30	1	41,50	5,599	52
	2	58,64	10,617	22
	Total	46,59	10,794	74
30<=Edad<40	1	39,71	5,944	42
	2	60,25	4,288	12
	Total	44,28	10,267	54
40<=Edad	1	48,72	11,934	29
	2	67,70	6,778	30
	Total	58,37	13,538	59
Total	1	42,59	8,364	123
	2	63,19	8,972	64
	Total	49,64	13,006	187



Los resultados sugieren un posible efecto de la edad (incremento en ambos grupos de alcohol a partir de 40 años), un efecto del alcohol (siempre por encima en el grupo 2, 'consumo alto') y la ausencia de interacción ya que es un modelo de rectas paralelas y los efectos de cada variable son los mismos en cualquier

categoría de la otra variable (misma distancia de alcohol 1 a alcohol 2 para cualquier edad y misma tendencia en edad en ambos grupos de alcohol)

ii. Realice un ANOVA de dos factores para comprobar los efectos existentes. ¿Hay interacción? ¿Hay efecto de la edad? ¿Hay efecto del alcohol?

Tomaremos la tabla de evaluación de efectos del ANOVA:

Pruebas de efectos inter-sujetos

Variable dependiente: urico

Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	20523,989 ^a	5	4104,798	67,932	,000
Intersección	418069,431	1	418069,431	6918,765	,000
edadrec	2472,711	2	1236,355	20,461	,000
alcohol	13390,636	1	13390,636	221,606	,000
edadrec * alcohol	70,148	2	35,074	,580	,561
Error	10937,005	181	60,425		
Total	492285,000	187			
Total corregido	31460,995	186			

a. R al cuadrado = ,652 (R al cuadrado ajustada = ,643)

Como se observa el efecto de interacción no es significativo ($p=0,561$), mientras que si lo son los efectos de la edad y el alcohol (ambos $p<0,001$). Esto nos lleva a afirmar la existencia de diferencias significativas en las medias de úrico según alcohol y según edad, pero diferencias constantes, sin interacción que supondría que las diferencias no serían constantes.

El valor $R^2=0,652$ representa el porcentaje de variabilidad de úrico explicada por el modelo completo (edad+alcohol+interacción)

iii. Realice comparaciones múltiples con los tests de TUKEY y SCHEFFÉ

iv. Calcule intervalos de confianza al 95% para la diferencia de medias de ácido úrico entre grupos de edad. Idem entre grupos de alcohol (intervalos ajustados por comparaciones múltiples)

Comparaciones múltiples

Variable dependiente: urico

	(I) edadrec	(J) edadrec	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
						Límite inferior	Límite superior
HSD Tukey	Edad<30	30<=Edad<40	2,32	1,391	,221	-,97	5,60
		40<=Edad	-11,78 [*]	1,357	,000	-14,98	-8,57
	30<=Edad<40	Edad<30	-2,32	1,391	,221	-5,60	,97
		40<=Edad	-14,10 [*]	1,464	,000	-17,55	-10,64
	40<=Edad	Edad<30	11,78 [*]	1,357	,000	8,57	14,98
		30<=Edad<40	14,10 [*]	1,464	,000	10,64	17,55
Scheffe	Edad<30	30<=Edad<40	2,32	1,391	,253	-1,12	5,75
		40<=Edad	-11,78 [*]	1,357	,000	-15,13	-8,43
	30<=Edad<40	Edad<30	-2,32	1,391	,253	-5,75	1,12
		40<=Edad	-14,10 [*]	1,464	,000	-17,71	-10,48
	40<=Edad	Edad<30	11,78 [*]	1,357	,000	8,43	15,13
		30<=Edad<40	14,10 [*]	1,464	,000	10,48	17,71

Se basa en las medias observadas.

El término de error es la media cuadrática(Error) = 60,425.

*. La diferencia de medias es significativa en el nivel ,05.



Los apartados iii y iv se resuelven conjuntamente. Teniendo en cuenta que el alcohol solo tiene 2 categorías, no es necesario aplicar correcciones a la comparación de medias, y el intervalo de la diferencia se haría con la prueba t univariante. No es así en el caso de la edad que tiene 3 categorías y en ese caso si que hay comparaciones múltiples. La tabla siguiente presenta los resultados con Scheffé y Tukey, que son bastante similares:

Subconjuntos homogéneos

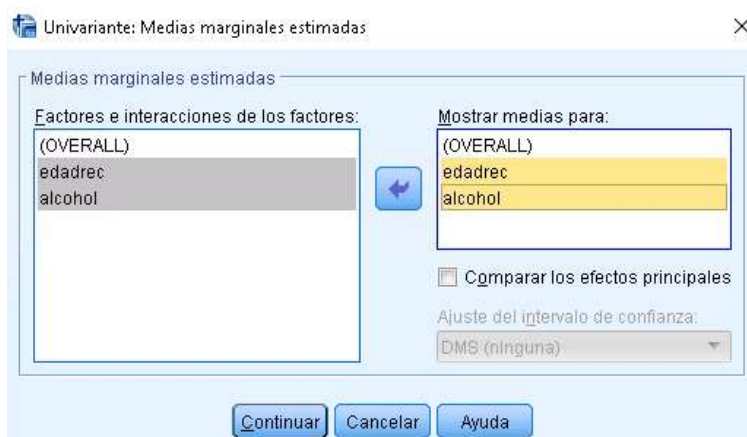
urico			Subconjunto	
	edadrec	N	1	2
HSD Tukey ^{a,b,c}	30<=Edad<40	54	44,28	
	Edad<30	74	46,59	
	40<=Edad	59		58,37
	Sig.		,228	1,000
Scheffe ^{a,b,c}	30<=Edad<40	54	44,28	
	Edad<30	74	46,59	
	40<=Edad	59		58,37
	Sig.		,259	1,000

Se visualizan las medias para los grupos en los subconjuntos homogéneos.

Se basa en las medias observadas.

Reflexiones: modificación del modelo y nuevos resultados

A la vista de los resultados obtenidos hasta el momento, el modelo adecuado es un modelo de efectos principales de edad y alcohol, sin interacción. Procede pues modificar la selección del modelo y repetir el análisis, particularmente para los requerimientos:



procediendo también a pedir las **MEDIAS MARGINALES ESTIMADAS**, ya que, al haber modificado el modelo y quitada la interacción no coincidirán con las observadas, siendo estimadas por el modelo con efectos edad y alcohol (sin interacción):

Resultados para el nuevo modelo

La tabla del ANOVA permite comprobar la significación de los efectos de edad y alcohol ($p < 0,001$). El R^2 es 0,650, prácticamente igual al del modelo anterior con interacción (al no ser ésta significativa) que era 0,652

Pruebas de efectos inter-sujetos

Variable dependiente: urico

Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	20453,841 ^a	3	6817,947	113,352	,000
Intersección	457292,759	1	457292,759	7602,744	,000
edadrec	2600,271	2	1300,135	21,615	,000
alcohol	13715,314	1	13715,314	228,025	,000
Error	11007,154	183	60,148		
Total	492285,000	187			
Total corregido	31460,995	186			

a. R al cuadrado = ,650 (R al cuadrado ajustada = ,644)

Las medias marginales estimadas son:

1. Media global

Variable dependiente: urico

Media	Desv. Error	Intervalo de confianza al 95%	
		Límite inferior	Límite superior
52,679	,604	51,487	53,871

2. edadrec

Variable dependiente: urico

edadrec	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
Edad<30	50,370	,936	48,524	52,216
30<=Edad<40	49,451	1,110	47,262	51,641
40<=Edad	58,215	1,010	56,223	60,207

3. alcohol

Variable dependiente: urico

alcohol	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	43,366	,711	41,963	44,770
2	61,991	,992	60,033	63,949

que, como se puede observar, no coinciden exactamente con las medias reales observadas; éstas son las estimadas por el modelo.



v. Compruebe los requerimientos del modelo

Los requerimientos deben ser comprobados sobre el modelo final adoptado, en este caso el modelo con efectos edad y alcohol, sin interacción.

La prueba de Levene de homogeneidad de varianzas y prueba de Kolmogorov-Smirnov sobre los residuos del modelo son:

Prueba de igualdad de Levene de varianzas de error^a

Variable dependiente: urico

F	gl1	gl2	Sig.
2,605	5	181	,027

Prueba la hipótesis nula de que la varianza de error de la variable dependiente es igual entre grupos.

a. Diseño : Intersección + edadrec + alcohol

Prueba de Kolmogorov-Smirnov para una muestra

		Residuo estandarizado para urico
N		187
Parámetros normales ^{a,b}	Media	,0000
	Desv. Desviación	,99190
Máximas diferencias extremas	Absoluto	,102
	Positivo	,102
	Negativo	-,071
Estadístico de prueba		,102
Sig. asintótica(bilateral)		,000 ^c

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

lo cual sugiere que ni la varianza puede considerarse homogénea ni los residuos normales (ambas pruebas son significativas).

vi. Puesto que la normalidad de los residuos es rechazable, repita el ANOVA anterior con transformaciones de la variable respuesta (logaritmo y raíz cuadrada)

Vamos a repetir el análisis de la significación de efectos para la variable SQRICO (raíz cuadrada del urico).

Pruebas de efectos inter-sujetos

Variable dependiente: squico

Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	98,525 ^a	3	32,842	112,648	,000
Intersección	8539,037	1	8539,037	29289,236	,000
alcohol	66,537	1	66,537	228,226	,000
edadrec	12,158	2	6,079	20,851	,000
Error	53,352	183	,292		
Total	9283,000	187			
Total corregido	151,877	186			

a. R al cuadrado = ,649 (R al cuadrado ajustada = ,643)

observando que se mantiene el efecto significativo de alcohol y de edad y, como vemos a continuación, se pueden aceptar los requerimientos de homogeneidad de varianzas ($p=0,143$) y de normalidad (p exacta= $0,07$)

Prueba de igualdad de Levene de varianzas de error^a

Variable dependiente: squico

F	gl1	gl2	Sig.
1,675	5	181	,143

Prueba la hipótesis nula de que la varianza de error de la variable dependiente es igual entre grupos.

a. Diseño : Intersección + alcohol + edadrec

Prueba de Kolmogorov-Smirnov para una muestra

		Residuo estandarizado para squico
N		187
Parámetros normales ^{a,b}	Media	,0000
	Desv. Desviación	,99190
Máximas diferencias extremas	Absoluto	,094
	Positivo	,094
	Negativo	-,072
Estadístico de prueba		,094
Sig. asintótica(bilateral)		,000 ^c
Significación exacta (bilateral)		,070
Probabilidad en el punto		,000

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

Esto refuerza todos los resultados previos obtenidos con la variable original urico.

Sesión 2

Análisis de la Covarianza



Sesión 2.- Análisis de la covarianza (ANCOVA)

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

Identificar situaciones de investigación en las que deba ser utilizado un ANCOVA
Identificar los modelos de ANCOVA con varios factores y covariables
Realizar un ANCOVA e interpretar sus resultados
Identificar las relaciones/asociaciones entre las variables explicativas y respuesta
Identificar los efectos de confusión e interacción entre las variables del ANCOVA

Actividades: Las actividades previstas son las siguientes

Revisión de los contenidos teóricos
Realización de las actividades de autoaprendizaje resueltas y comprobación de resultados
Realización opcional de actividades de autoaprendizaje para practicar

Contenidos teóricos: Material teórico pgs. 29-41 de la monografía Modelos de Análisis de la Varianza y Covarianza, <http://hdl.handle.net/10045/113344>

Actividades de autoaprendizaje: Ver actividades de autoaprendizaje sesión 2

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Cuadernillo de prácticas, prácticas 6 y 7
- Bases de datos: DATOS EJEMPLO TEXTO.sav; ANOVAS.sav



Sesión 2.- Actividades de autoaprendizaje (resueltas)

2.1. Diga si los siguientes enunciados son verdaderos (V) o falsos (F)

Para todas las cuestiones se supondrá que estamos interesados en averiguar el efecto de un factor (variable cualitativa/categorica) y una covariable (variable cuantitativa) sobre una variable respuesta cuantitativa

- i. El ANCOVA es un procedimiento que se puede considerar una generalización del ANOVA y de la regresión lineal ya que permite introducir variables explicativas cualitativas y cuantitativas
- ii. El diagrama de dispersión de la variable respuesta como función de la covariable, separando por categorías del factor es un instrumento descriptivo adecuado para identificar los posibles efectos existentes entre las variables
- iii. Una inspección gráfica del diagrama de dispersión de la variable respuesta en función de la variable explicativa cuantitativa que produce rectas paralelas separadas para las categorías de la variable explicativa categorica sugiere una posible interacción entre las variables.
- iv. Si el modelo resultante es un modelo de regresión lineal, con efecto tanto de la variable categorica como de la cuantitativa sobre la variable respuesta, en el que los coeficientes de las variables explicativas coinciden con el que tendrían en un análisis simple, podemos decir que no hay confusión al evaluar el efecto sobre la variable respuesta
- v. El modelo de interacción puede ser representado en forma de modelos de regresión, bien como dos modelos distintos para cada categoría del factor o bien como un modelo de regresión en el que se introduzca un término de interacción como producto del factor por la covariable
- vi. Si no hay interacción y hay efecto de factor y de covariable, el ANCOVA permite estimar los valores de las medias de la variable respuesta en los niveles del factor ajustadas por la covariable, lo cual tiene particular interés si las medias de la covariable son las mismas en cada categoría del factor
- vii. En el ANCOVA, a los requerimientos descritos para ANOVA y regresión (normalidad, homogeneidad de varianzas e independencia) se une la necesidad de que la covariable sea también normal
- viii. Si en un ANCOVA se detecta un efecto de interacción deja de tener sentido indagar si existe efecto de confusión
- ix. En el ANCOVA, la descomposición de la suma de cuadrados de tipo I, según la cual se evalúa el incremento de variabilidad explicada cuando cada uno de los efectos es añadido en el modelo a los efectos precedentes, es preferible al método de tipo III
- x. Si el efecto del factor sobre la respuesta es el mismo tanto si se incluye la covariable en el modelo como si no se incluye es que ésta no confunde el efecto del factor

2.2. Resuelva las siguientes aplicaciones: PRÁCTICA 6 del cuadernillo 'Prácticas_curso_datos_continuos.pdf'

A partir del archivo DATOS EJEMPLOS TEXTO.SAV. Queremos evaluar el efecto de la edad (categorizada en grupos) y el alcohol sobre el ácido úrico. Abra el archivo DATOS EJEMPLOS TEXTO.SAV. En él encontrará las variables

EDAD: En años	CALCIO: mg/100ml
PESO: En Kg.	URICO: mg/litro
COLEST: mg/100ml	ALCOHOL: 1 'bajo' 2'moderado/alto'
ALBUMINA: mg/100ml	EDADREC: 1 '≤30' 2 '30-40' 3 '>40'



Abra el archivo DATOS EJEMPLO TEXTO.SAV.

Utilizando como variable respuesta el PESO y explicativas ALCOHOL y EDAD (continua), construya un gráfico de dispersión que le permita describir la distribución de pesos y edades según alcohol. ¿Qué sugiere el gráfico?

Utilizando el procedimiento UNIVARIANTE del comando MODELO LINEAL GENERAL del SPSS, produzca un ANCOVA que le permita discutir la existencia de interacción.

Construya los modelos de regresión lineal simple para el peso como función de la edad en cada uno de los grupos de alcohol. Evalúe los efectos de la edad en cada grupo. Compárelos

Reproduzca con un modelo de regresión que incluya un término de interacción la detección de la interacción significativa

Repita 1 a 3 para la variable respuesta URICO.

Dado que no hay interacción significativa, discuta el efecto de confusión, si lo hay, que introduce la edad sobre el efecto del alcohol y viceversa

Calcule las medias de ácido úrico para los grupos de consumo de alcohol ajustadas por edad (en la edad promedio). Calcule sus intervalos de confianza al 95%.

Compruebe los requerimientos del ANCOVA para el modelo ácido úrico función de edad y alcohol

2.3 Ejercicio práctico propuesto (para practicar)

Repita un ANCOVA para evaluar los efectos existentes para las variables respuesta CALCIO y ALBUMINA y explicativas la edad y el alcohol

Sesión 2.- Soluciones a las actividades de autoaprendizaje

2.1 Cuestiones

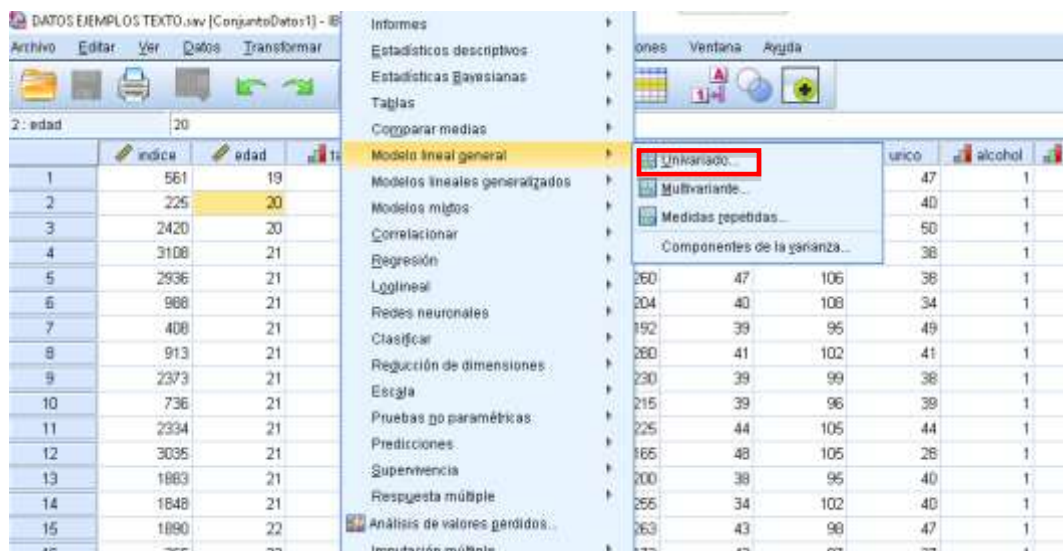
Cuestión	i	ii	iii	iv	v	vi	vii	viii	ix	x
Respuesta	V	V	F	V	V	F	F	V	V	V



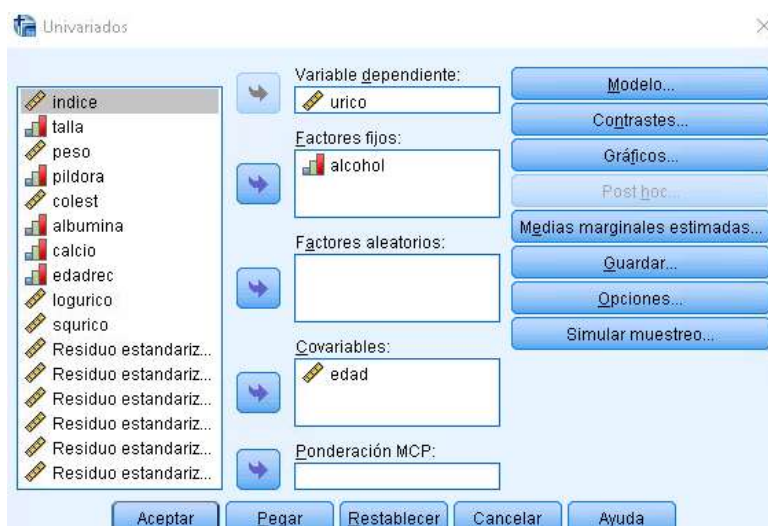
2.2 Aplicaciones

Consideraciones previas. Cómo aplicar el procedimiento ANCOVA con SPSS

Como ya se dijo en la sesión anterior, para el ANCOVA se utilizará el procedimiento 'Modelo lineal general' del SPSS, **univariado**:



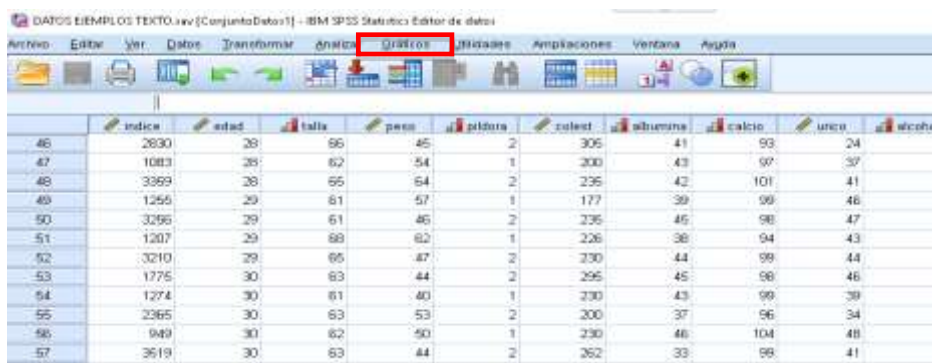
Tomaremos como ejemplo de variable respuesta continua URICO y las explicativas serán una categórica (factor, ALCOHOL) y la otra cuantitativa (covariable, EDAD), tras entrar en **Univariado** definimos:



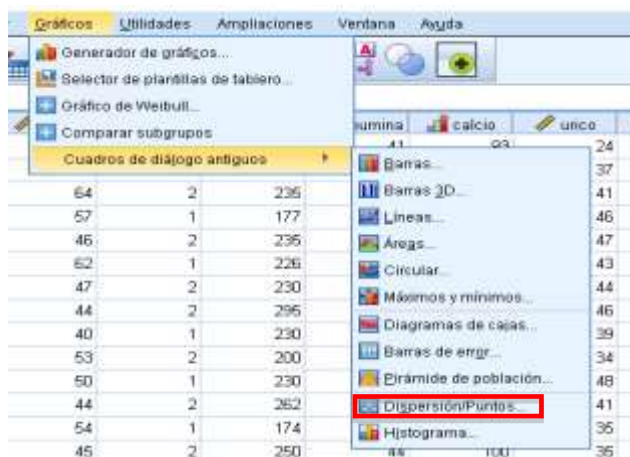
usando URICO como variable dependiente/respuesta, ALCOHOL como factor y EDAD como covariable.

A continuación definiremos como construir gráficos descriptivos orientativos y un análisis básico, con todos los elementos para realizar, para ir contestando con los resultados las preguntas del ejercicio. Revisemos que definir en cada pestaña:

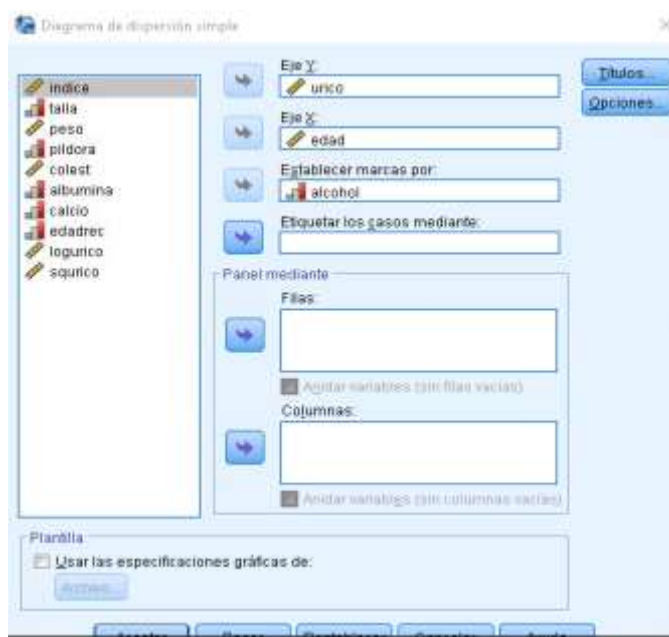
GRÁFICOS: La opción de **Gráficos** incluida como pestaña en el procedimiento **Modelo lineal general** no es útil para el ANCOVA. Los gráficos de dispersión de interés pueden ser contruidos con SPSS en la pestaña general de **Gráficos del programa**:



	indice	edad	talla	peso	pildora	colest	albumina	calcio	unico	alcohol
46	2630	28	56	45	2	305	41	93	24	1
47	1083	28	62	54	1	200	43	90	37	1
48	3359	28	65	64	2	235	42	101	41	1
49	1255	29	61	57	1	177	39	99	46	1
50	3255	29	61	46	2	235	45	98	47	1
51	1207	29	68	62	1	226	36	94	43	1
52	3210	29	65	47	2	230	44	99	44	1
53	1775	30	63	44	2	255	45	98	46	1
54	1274	30	61	40	1	230	43	99	39	1
55	2365	30	63	53	2	200	37	96	34	1
56	949	30	62	50	1	230	46	104	48	1
57	3519	30	63	44	2	262	33	98	41	1

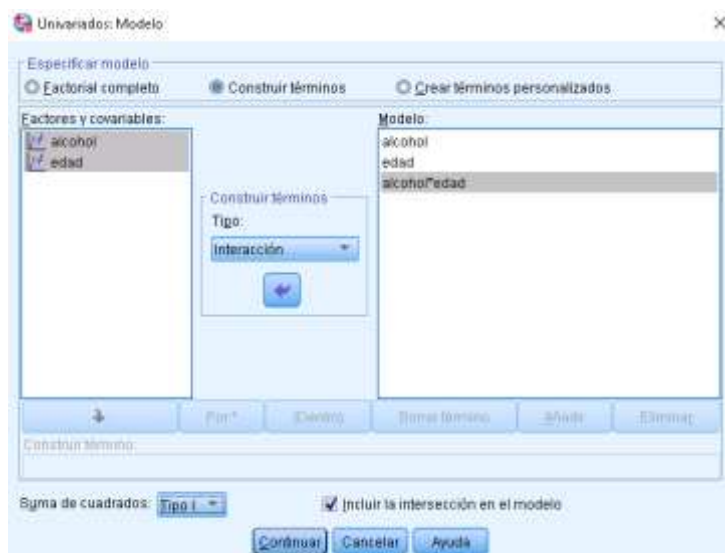


Seleccionaremos el diagrama de **Dispersión/Puntos**, con opción **Dispersión simple** y pasamos a definirlo:



poniendo la variable respuesta en el eje Y, la covariable en el eje X y separando los casos según alcohol. Posteriormente, en las soluciones del ejercicio, veremos como manejar e interpretar el resultado.

MODELO: Permite seleccionar qué efectos incluirá nuestro modelo:



Por defecto figura el modelo factorial completo. Este modelo incluye todos los efectos posibles que, en nuestro caso serán el de ALCOHOL, el de EDAD y el de la INTERACCION entre ambas. A diferencia del ANOVA, en el ANCOVA se recomienda utilizar la opción **Construir términos**, incluyendo inicialmente todos los términos posibles (los efectos principales y la interacción), combinada con la **Suma de cuadrados** Tipo I. Luego, en la resolución del ejercicio, veremos qué nos permite evaluar esta manera de construcción.

La interacción se pasa como ALCOHOL*EDADREC, y recordemos

que se consigue seleccionando ambas juntas y pulsando el pulsador intermedio que indicará Tipo interacción.

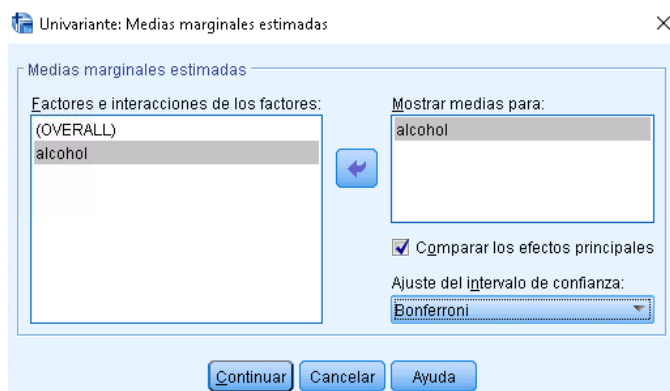
CONTRASTES: No se utiliza en este procedimiento. Se utilizará en ANOVA de medidas repetidas

POSTHOC: Al igual que en el ANOVA Se pueden seleccionar pruebas para las comparaciones

múltiples de las medias según categorías de las variables explicativas categóricas. Sólo se pueden seleccionar las pruebas con homogeneidad de varianzas asumida.

En este caso la única variable categórica es ALCOHOL y, al tener sólo 2 categorías, no se activa la pestaña **Posthoc** correspondiente

MEDIAS MARGINALES ESTIMADAS: En el caso del ANCOVA esta opción permite estimar las medias de la variable respuesta (URICO) en cada categoría del factor (ALCOHOL) pero para un valor común de



la covariable (EDAD) que es en la media (de edad).

Si elegimos **Comparar los efectos principales**, con método **Bonferroni**, nos producirá los intervalos de confianza de las medias estimadas y para la diferencia entre las medias con corrección de Bonferroni.

Univariados: Guardar

Valores pronosticados:

- ☐ No estandarizados
- ☐ Ponderados
- ☐ Error estándar

Residuos:

- ☐ No estandarizados
- ☐ Ponderados
- ☒ Estandarizados
- ☐ Método de Student
- ☐ Eliminados

Diagnósticos:

- ☐ Distancia de Cook
- ☐ Valores de influencia

Estadísticos de los coeficientes:

- ☒ Crear estadísticos de los coeficientes

Tipo:

- ☒ Estadísticas estándar
- ☐ Estadísticas coherentes con heterocedasticidad

Requiere la selección de estimaciones de parámetro con errores estándar robustos en Opciones.

Destino:

- ☒ Crear un nuevo conjunto de datos
- ☐ Escribir un nuevo archivo de datos

Nombre de conjunto de datos:

Archivo...

GUARDAR: Pueden seleccionarse diversas variables resultado para añadirlas a la base de datos. En este caso se seleccionan los residuos para ver su normalidad:

Univariados: Opciones

Visualización:

- ☒ Estadísticos descriptivos
- ☐ Estimaciones del tamaño del efecto
- ☐ Potencia observada
- ☐ Estimaciones de los parámetros
- ☐ Matriz de coeficientes del contraste
- ☒ Pruebas de homogeneidad
- ☐ Gráfico de dispersión versus gráfico de niveles
- ☐ Gráfico de residuos
- ☐ Falta de ajuste
- ☐ Función estimable general

Pruebas de heterocedasticidad:

- ☐ Test de Breusch-Pagan modificado
- ☒ Test de Breusch-Pagan
- ☐ Prueba F
- ☐ Prueba de White

Modelo...

Estimaciones de parámetro con errores estándar robustos:

- ☒ HC0
- ☐ HC1
- ☐ HC2
- ☐ HC3
- ☐ HC4

Nivel de significación: Los intervalos de confianza son 95,0 %

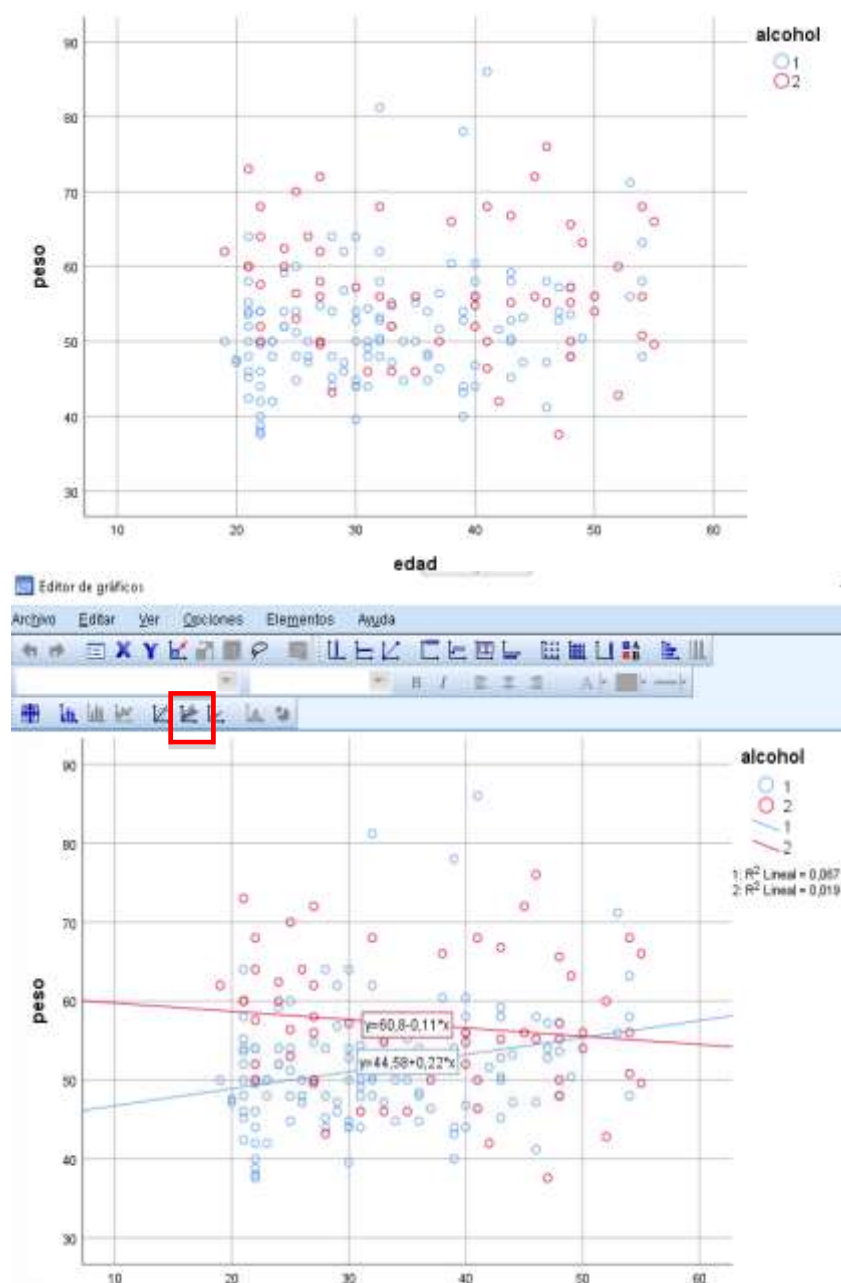
OPCIONES: En esta pestaña se pueden seleccionar diversos elementos. Elegimos los estadísticos descriptivos y la prueba de homogeneidad de varianzas para el modelo propuesto

SIMULAR MUESTREO (BOOTSTRAPPING): Activando esta opción se estimarían por simulación ciertos intervalos de confianza de parámetros del modelo (en caso de haberlos solicitado).

Resolución del ejercicio

Utilizando como variable respuesta el PESO y explicativas ALCOHOL y EDAD (continua), construya un gráfico de dispersión que le permita describir la distribución de pesos y edades según alcohol. ¿Qué sugiere el gráfico?

Construimos el gráfico en la forma descrita anteriormente, produciendo la siguiente salida:



En la pantalla de resultados pulsamos dos veces en el gráfico y se abre en una pantalla de edición de gráficos.

Si seleccionamos el pulsador en con recuadro rojo en la imagen obtendremos las líneas de ajuste de regresión para el PESO en función de la EDAD, y, como se ve, separando según consumo de ALCOHOL.

El gráfico sugiere varias posibilidades, siendo las más probables

- Interacción, al tratarse de rectas que se cruzan
- No efecto de ninguna variable, pudiendo ser rectas planas, sin diferencias por alcohol

Por otra parte, los estadísticos descriptivos de las variables, obtenidos por OPCIONES son:

Estadísticos descriptivos

Variable dependiente: peso

alcohol	Media	Desv. Desviación	N
1	51,57	7,761	124
2	56,96	8,468	62
Total	53,37	8,377	186

sin ofrecer información descriptiva para la edad. Estos estadísticos no aportan nada relevante. El gráfico, sin embargo, si nos ha aportado mucha información

Utilizando el procedimiento UNIVARIANTE del comando MODELO LINEAL GENERAL del SPSS, produzca un ANCOVA que le permita discutir la existencia de interacción.

La tabla de evaluación de los efectos es:

Pruebas de efectos inter-sujetos

Variable dependiente: peso

Origen	Tipo I de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	1780,237 ^a	3	593,412	9,642	,000
Intersección	529792,249	1	529792,249	8607,837	,000
alcohol	1199,527	1	1199,527	19,489	,000
edad	124,245	1	124,245	2,019	,157
alcohol * edad	456,465	1	456,465	7,416	,007
Error	11201,674	182	61,548		
Total	542774,160	186			
Total corregido	12981,911	185			

a. R al cuadrado = ,137 (R al cuadrado ajustada = ,123)

Al haber sido construida con suma de cuadrados Tipo I, lo que podemos afirmar es que:

- El ALCOHOL sólo tiene efecto significativo ($p < 0,001$)
- La EDAD no añade capacidad explicativa significativa al ALCOHOL ($p = 0,157$)
- La interacción ALCOHOL*EDAD si que añade capacidad explicativa significativa a los anteriores ($p = 0,007$)

Concluyendo así la existencia de interacción significativa.

Construya los modelos de regresión lineal simple para el peso como función de la edad en cada uno de los grupos de alcohol. Evalúe los efectos de la edad en cada grupo. Compárelos

Dado que existe interacción significativa, ahora hay que explicarla y caracterizarla. La interacción significa que el efecto de la EDAD no es el mismo en cada grupo de ALCOHOL (rectas diferentes) o que las diferencias entre los grupos de ALCOHOL dependen de la EDAD.

Una forma de describir la interacción es construir los modelos de regresión separando por grupos del factor (ALCOHOL). Usando el modelo de regresión lineal con variable respuesta el PESO y explicativa la EDAD, obtenemos:

Resumen del modelo



alcohol	Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	1	,259 ^a	,067	,059	7,526
2	1	,138 ^a	,019	,003	8,456

a. Predictores: (Constante), edad

ANOVA ^a							
alcohol	Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	1	Regresión	496,942	1	496,942	8,773	,004 ^b
		Residuo	6910,975	122	56,647		
		Total	7407,917	123			
2	1	Regresión	83,769	1	83,769	1,171	,283 ^b
		Residuo	4290,699	60	71,512		
		Total	4374,467	61			

a. Variable dependiente: peso

b. Predictores: (Constante), edad

Coeficientes ^a								
alcohol	Modelo		Coeficientes no estandarizados		Coeficientes estandarizados		95,0% intervalo de confianza para B	
			B	Desv. Error	Beta	t	Sig.	Límite inferior / Límite superior
1	1	(Constante)	44,584	2,455		18,160	,000	39,723 / 49,444
		edad	,216	,073	,259	2,962	,004	,072 / ,360
2	1	(Constante)	60,801	3,707		16,402	,000	53,386 / 68,217
		edad	-,105	,097	-,138	-1,082	,283	-,299 / ,089

a. Variable dependiente: peso

Pudiendo comprobar que la EDAD sólo presenta efecto significativo sobre el PESO en los no consumidores de ALCOHOL (alcohol=1), mientras que no es significativo (incluso es negativo) en el grupo de ALCOHOL=2. Disponemos además de las estimaciones del efecto (coeficientes B) y de sus correspondientes intervalos de confianza.

La evaluación de las diferencias entre grupos de ALCOHOL es difícil de realizar, ya que éstas serían diferentes a cada EDAD. Con la salida de las **Medias marginales estimadas** del **Modelo lineal general** disponemos de información para una edad concreta, para la media de edad:

Estimaciones

Variable dependiente: peso

alcohol	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	51,872 ^a	,712	50,466	53,277
2	57,251 ^a	1,027	55,225	59,277

a. Las covariables que aparecen en el modelo se evalúan en los valores siguientes: edad = 33,77.

Comparaciones por parejas

Variable dependiente: peso

(I) alcohol	(J) alcohol	Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b	
					Límite inferior	Límite superior
1	2	-5,379	1,250	,000	-7,845	-2,914
2	1	5,379	1,250	,000	2,914	7,845

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.



b. Ajuste para varias comparaciones: Bonferroni.

Estos resultados estiman una diferencia en las medias de PESO de 5,379 kg (significativa, $p < 0,001$) entre los grupos de ALCOHOL cuando la edad es la edad media de 33,77 años

Reproduzca con un modelo de regresión que incluya un término de interacción la detección de la interacción significativa

Otra forma de comprobar la existencia de interacción sería utilizando un modelo de regresión con un término de interacción creado como producto de las variables explicativas ALCOHOL y EDAD, lo cual en este caso es relativamente viable al ser el ALCOHOL dicotómica (equivalente a dummy), pero es fácil imaginar que con variables categóricas de múltiples categorías el procedimiento sería tedioso.

A partir de crear una variable INTERACCION = ALCOHOL*EDAD (*=producto)

construimos los modelos de regresión:

$$\begin{aligned}\text{PESO} &= \text{ALCOHOL} + \text{EDAD} \\ \text{PESO} &= \text{ALCOHOL} + \text{EDAD} + \text{INTERACCIÓN}\end{aligned}$$

evaluando el R^2 y F de cambio de uno a otro modelo:

Resumen del modelo									
Modelo	R	Estadísticos de cambio			Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F
		R cuadrado	R cuadrado ajustado	Error estándar de la estimación					
1	,319 ^a	,102	,092	7,982	,102	10,390	2	183	,000
2	,370 ^b	,137	,123	7,845	,035	7,416	1	182	,007

a. Predictores: (Constante), edad, alcohol

b. Predictores: (Constante), edad, alcohol, INTERACCION

ANOVA ^a						
Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	1323,772	2	661,886	10,390	,000 ^b
	Residuo	11658,139	183	63,706		
	Total	12981,911	185			
2	Regresión	1780,237	3	593,412	9,642	,000 ^c
	Residuo	11201,674	182	61,548		
	Total	12981,911	185			

a. Variable dependiente: peso

b. Predictores: (Constante), edad, alcohol

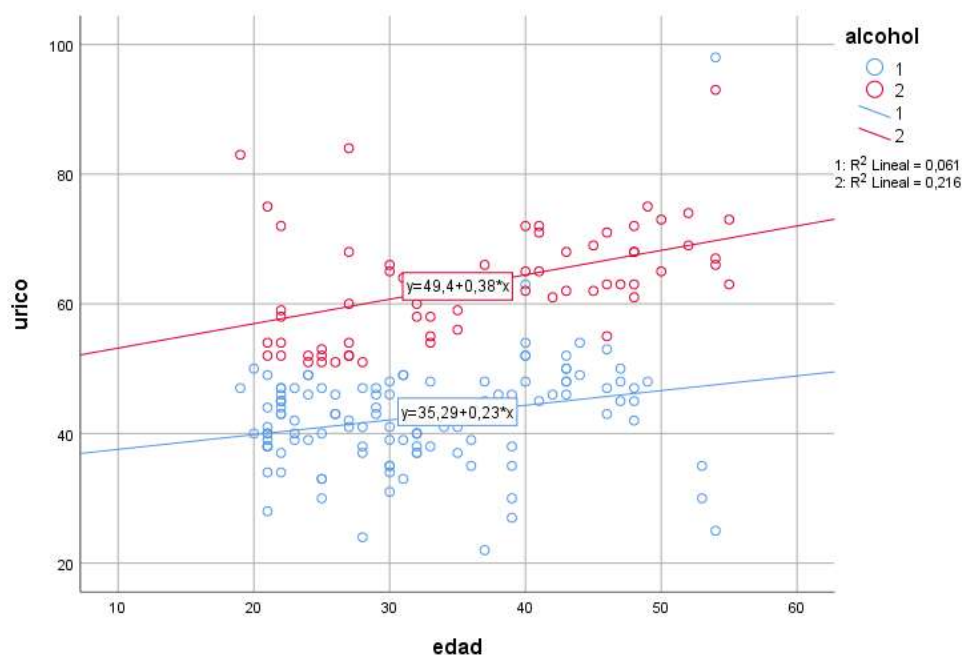
c. Predictores: (Constante), edad, alcohol, INTERACCION

pudiendo observar que el modelo 2, que añade la INTERACCION, añade capacidad explicativa significativa ($p=0,007$) al modelo 1, ALCOHOL + EDAD

Repita 1 a 3 para la variable respuesta URICO

El gráfico de dispersión obtenido es:





lo cual sugiere un posible efecto de la edad y del alcohol, probablemente sin interacción (modelo de rectas paralelas)

El resultado del ANCOVA para evaluar los efectos es:

Pruebas de efectos inter-sujetos

Variable dependiente: urico

Origen	Tipo I de suma de cuadrados	gl	Media cuadrática	F	Sig.
Modelo corregido	19473,155 ^a	3	6491,052	99,089	,000
Intersección	460824,005	1	460824,005	7034,695	,000
alcohol	17853,570	1	17853,570	272,543	,000
edad	1520,449	1	1520,449	23,210	,000
alcohol * edad	99,136	1	99,136	1,513	,220
Error	11987,840	183	65,507		
Total	492285,000	187			
Total corregido	31460,995	186			

a. R al cuadrado = ,619 (R al cuadrado ajustada = ,613)

Con efecto de interacción que no añade capacidad explicativa significativa al ALCOHOL y la EDAD ($p=0,220$)

Dado que no hay interacción significativa, discuta el efecto de confusión, si lo hay, que introduce la edad sobre el efecto del alcohol y viceversa

Para comprobar la existencia o no de confusión hay que hacer los análisis de regresión detallados, con efectos simples para ALCOHOL y para EDAD, y luego un modelo conjunto, para ver si hay cambios importantes. Usando el modelo de regresión lineal:

Modelo ALCOHOL

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados ^a		Sig.	95,0% intervalo de confianza para B	
		B	Desv. Error	Beta	t		Límite inferior	Límite superior
1	(Constante)	21,999	1,882		11,691	,000	18,287	25,712
	alcohol	20,594	1,322	,753	15,580	,000	17,986	23,202

a. Variable dependiente: urico

Modelo EDAD

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados ^a		Sig.	95,0% intervalo de confianza para B	
		B	Desv. Error	Beta	t		Límite inferior	Límite superior
1	(Constante)	33,434	3,113		10,740	,000	27,292	39,576
	edad	,481	,089	,371	5,431	,000	,306	,655

a. Variable dependiente: urico

Modelo ALCOHOL + EDAD**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	,785 ^a	,616	,612	8,105

a. Predictores: (Constante), edad, alcohol

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	19374,019	2	9687,009	147,465	,000 ^b
	Residuo	12086,976	184	65,690		
	Total	31460,995	186			

a. Variable dependiente: urico

b. Predictores: (Constante), edad, alcohol

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados ^a		Sig.	95,0% intervalo de confianza para B	
		B	Desv. Error	Beta	t		Límite inferior	Límite superior
1	(Constante)	13,883	2,451		5,663	,000	9,047	18,719
	alcohol	19,324	1,277	,707	15,135	,000	16,805	21,843
	edad	,291	,061	,225	4,811	,000	,172	,411

a. Variable dependiente: urico

Podemos observar que el modelo ALCOHOL+EDAD incluye ambas variables como significativas, tiene un $R^2 = 0,616$, y es en conjunto significativo.



Si observamos los cambios en los coeficientes B e intervalos de confianza, desde los modelos simples al modelo conjunto, vemos que el ALCOHOL apenas modifica su coeficiente (de 20,594 a 19,324) y que la EDAD modifica más su coeficiente (de 0,481 a 0,291), aunque los intervalos de confianza al 95% se solapan en todos los casos. Podríamos decir que el ALCOHOL induce mayor confusión sobre la EDAD que lo contrario

Calcule las medias de ácido úrico para los grupos de consumo de alcohol ajustadas por edad (en la edad promedio). Calcule sus intervalos de confianza al 95%.

Se trata de obtener con el **Modelo lineal general** las medias marginales estimadas para el modelo ALCOHOL + EDAD, que es el modelo definitivo obtenido. Para ello, en el procedimiento **Univariado** se define un modelo sin interacción (sólo efectos ALCOHOL y EDAD) y se piden las **Medias marginales estimadas** para alcohol:

Estimaciones				
Variable dependiente: urico				
alcohol	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	43,028 ^a	,736	41,575	44,481
2	62,352 ^a	1,028	60,324	64,380

a. Las covariables que aparecen en el modelo se evalúan en los valores siguientes: edad = 33,71.

Comparaciones por parejas						
Variable dependiente: urico						
(I) alcohol	(J) alcohol	Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b	
					Límite inferior	Límite superior
1	2	-19,324 [*]	1,277	,000	-21,843	-16,805
2	1	19,324 [*]	1,277	,000	16,805	21,843

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.

b. Ajuste para varias comparaciones: Bonferroni.

Podemos ver que las medias son 43,028 y 62,352 para los grupos ALCOHOL 1 y 2 respectivamente, evaluadas a la EDAD media de 33,71 años (también disponemos de los intervalos de confianza). La segunda tabla presenta las diferencias entre las medias que, al no haber interacción son las mismas a cualquier edad y coinciden con el coeficiente B estimado por regresión lineal en el modelo conjunto ALCOHOL+EDAD (ver resultado del punto vi)

Compruebe los requerimientos del ANCOVA para el modelo ácido úrico función de edad y alcohol

Para ello en el modelo final ALCOHOL+EDAD se ha pedido la **Prueba de homogeneidad de varianzas** (en **Opciones**) y los **Residuos estandarizados** en **Guardar**. Con los residuos se ha aplicado una prueba de Kolmogorov-Smirnov de normalidad:

Prueba de igualdad de Levene de varianzas de error^a

Variable dependiente: urico

F	gl1	gl2	Sig.
,567	1	185	,453

Prueba la hipótesis nula de que la varianza de error de la variable dependiente es igual entre grupos.

a. Diseño : Intersección + alcohol + edad

Prueba de Kolmogorov-Smirnov para una muestra

		Residuo estandarizado para urico
N		187
Parámetros normales ^{a,b}	Media	,0000
	Desv. Desviación	,99461
Máximas diferencias extremas	Absoluto	,089
	Positivo	,089
	Negativo	-,057
Estadístico de prueba		,089
Sig. asintótica(bilateral)		,001 ^c
Significación exacta (bilateral)		,099
Probabilidad en el punto		,000

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

Como podemos observar, ninguna de las dos pruebas es significativa, por lo que los supuestos de homogeneidad de varianzas y normalidad son aceptables para este modelo

Sesión 3

Análisis de la Varianza de medidas repetidas



Sesión 3.- Análisis de la varianza de medidas repetidas (ANOVAMR)

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

Identificar situaciones de investigación en las que deba ser utilizado un ANOVAMR

Identificar los modelos de ANOVAMR con factores y covariables

Realizar un ANOVAMR e interpretar sus resultados

Identificar las relaciones/asociaciones entre las variables explicativas y respuesta

Identificar los efectos de confusión e interacción entre las variables del ANOVAMR

Actividades: Las actividades previstas son las siguientes

Revisión de los contenidos teóricos

Realización de las actividades de autoaprendizaje resueltas y comprobación de resultados

Realización opcional de actividades de autoaprendizaje para practicar

Contenidos teóricos: Material teórico pgs. 42-49 de la monografía Modelos de Análisis de la Varianza y Covarianza, <http://hdl.handle.net/10045/113344>

Actividades de autoaprendizaje: Ver actividades de autoaprendizaje sesión 3

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Cuadernillo de prácticas, prácticas 8 y 9
- Bases de datos: MEDREP.sav; Apareado.sav



Sesión 3.- Actividades de autoaprendizaje (resueltas)

3.1. Diga si los siguientes enunciados son verdaderos (V) o falsos (F)

- i. El ANOVAMR es un procedimiento indicado cuando el diseño de los grupos de análisis es apareado
- ii. El nombre de análisis de la varianza de medidas repetidas (ANOVAMR) responde al hecho de que debemos tener variables repetidas sobre sujetos distintos para su correcta aplicación
- iii. Si queremos comprobar si un grupo de individuos presenta cambios en las medias de cierta variable cuantitativa entre diferentes momentos del tiempo, la hipótesis nula a contrastar es que no hay diferencias en las medias
- iv. En el ANOVAMR las variables originales deben ser transformadas en contrastes ortonormalizados que son variables independientes (ortogonales) y tienen la propiedad de ser normalizados (la suma de los cuadrados de sus coeficientes es 1)
- v. Si nuestro análisis quiere comprobar las diferencias en un variable en varios momentos del tiempo, el número de contrastes ortonormalizados es igual al de momentos a estudio (o niveles del factor momento)
- vi. En el contraste de tipo 'simple' se compara cada momento del tiempo con un momento de referencia
- vii. En el contraste de tipo 'Helmert' se compara cada momento del tiempo con la media de los anteriores
- viii. En el ANOVAMR, a los requerimientos de normalidad y homogeneidad de varianzas se une la necesidad de que las covarianzas entre las variables contraste sean nulas
- ix. La prueba de esfericidad de Mauchly contrasta la hipótesis nula de que las varianzas son constantes y las covarianzas entre las variables originales son nulas
- x. Entre los estadísticos univariados que se pueden utilizar para contrastar la igualdad de medias entre diferentes momentos del tiempo, es de Huynh-Feldt es el más conservador

3.2. Resuelva las siguientes aplicaciones: PRÁCTICA 9 del cuadernillo 'Prácticas_curso_datos_continuos.pdf'

A partir del archivo MEDREP.sav. Los datos que almacena se refieren al tratamiento de la hipertensión idiopática (causa desconocida). Veinte pacientes recibieron, secuencialmente, los tratamientos:

c = Periodo de control

a = Propanolol + Fenoxibenzamina

b = Propanolol + Fenoxibenzamina + Hidroclorotiazida

La variable grupo = sexo, 1 'Hombres' 2 'mujeres'

Realice un ANOVA de medidas repetidas para comprobar si hay diferencias en las medias de tensión arterial entre los tratamientos

Compruebe entre que tratamientos existen diferencias significativas

Compruebe si las medias presentan tendencia lineal o cuadrática



Realice pruebas no paramétricas para comprobar las diferencias entre tratamientos

El efecto del tratamiento sobre las medias de tensión arterial podría interaccionar con el sexo (ser distinto en hombres y en mujeres). Compruebe esta posibilidad introduciendo la variable grupo como un factor intersujetos. ¿Hay interacción significativa? ¿Hay efecto del factor?

3.3 Ejercicio práctico propuesto (para practicar)

Usando la base de datos Apareado.sav, repita un ANOVA de medidas repetidas para evaluar los cambios en la variable úrico y si existe interacción con la variable 'dieta' (es decir si los cambios son los mismos o no) según el tipo de dieta: 1 Tipo A, 2 Tipo B, que representa el tipo de dieta seguido por estos sujetos.



Sesión 3.- Soluciones a las actividades de autoaprendizaje

3.1 Cuestiones

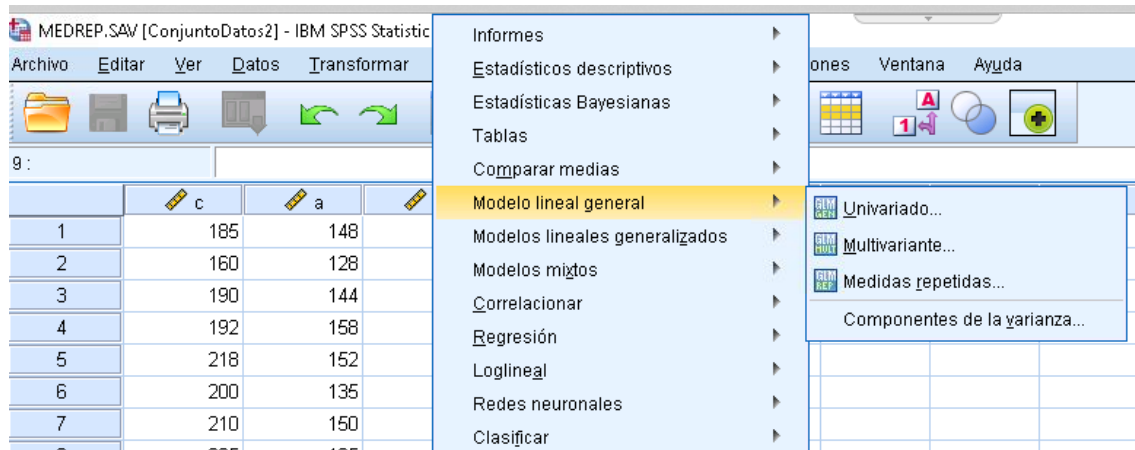
Cuestión	i	ii	iii	iv	v	vi	vii	viii	ix	x
Respuesta	V	F	V	V	F	V	F	V	V	F



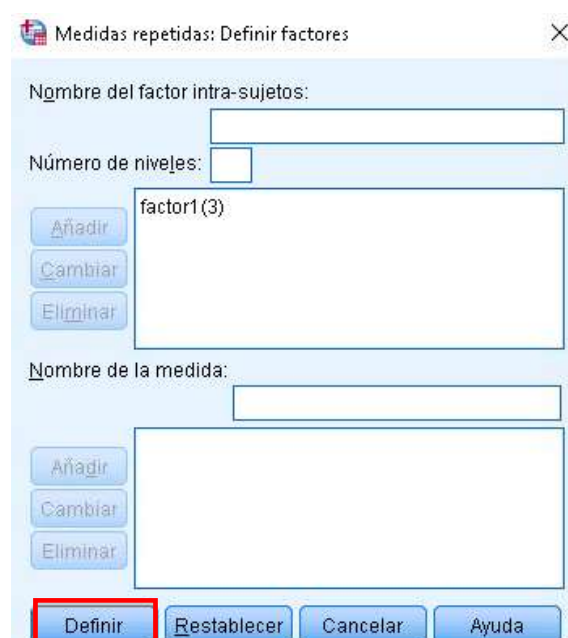
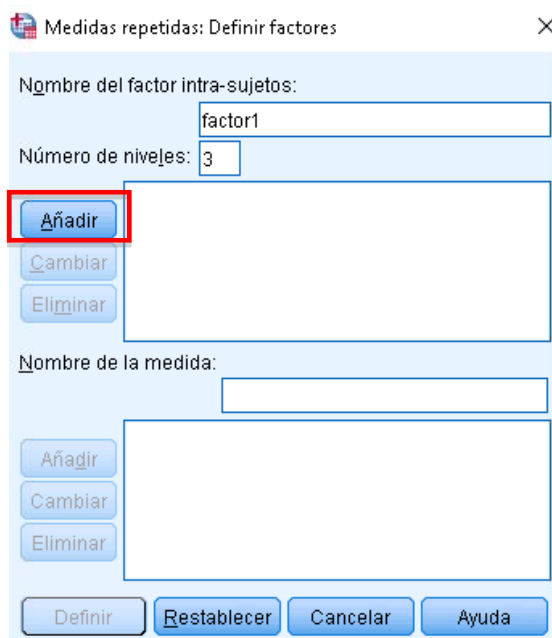
3.2 Aplicaciones

Consideraciones previas. Cómo aplicar el procedimiento ANOVA de medidas repetidas (ANOVAMR) con SPSS

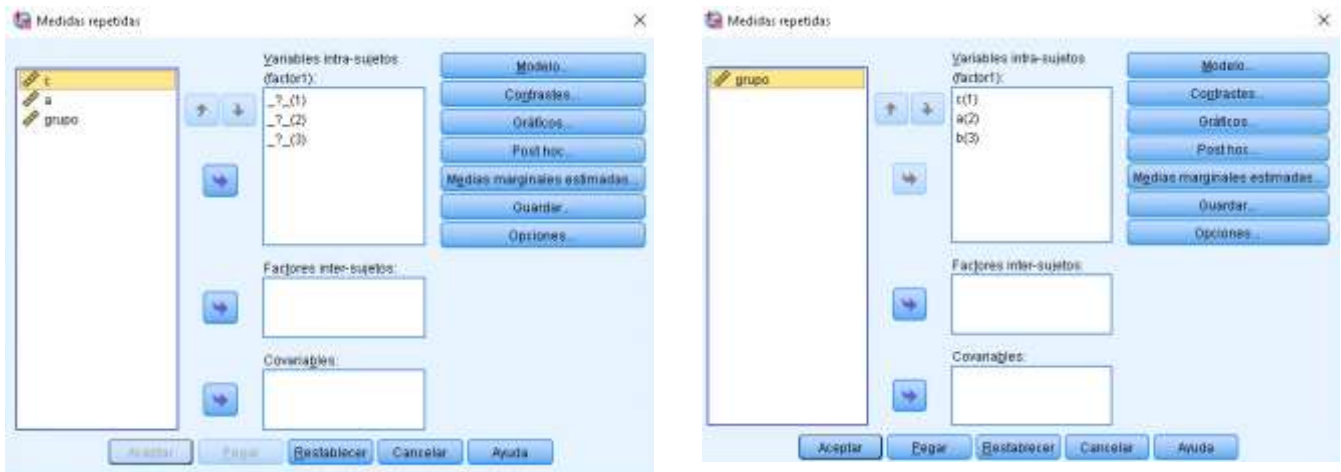
Abra el archivo MEDREP.sav. Se utilizará el procedimiento 'Modelo lineal general' del SPSS, **Medidas repetidas**:



Al entrar en el procedimiento aparece una primera pantalla de identificación del FACTOR intra-sujetos que es la variable que vamos a comparar entre los 3 momentos a estudio (corresponden a los 3 tratamientos distintos secuenciados en el tiempo), la tensión arterial sistólica TAS. Se define 3 niveles y se la da a Añadir, pasando a continuación a Definir el procedimiento

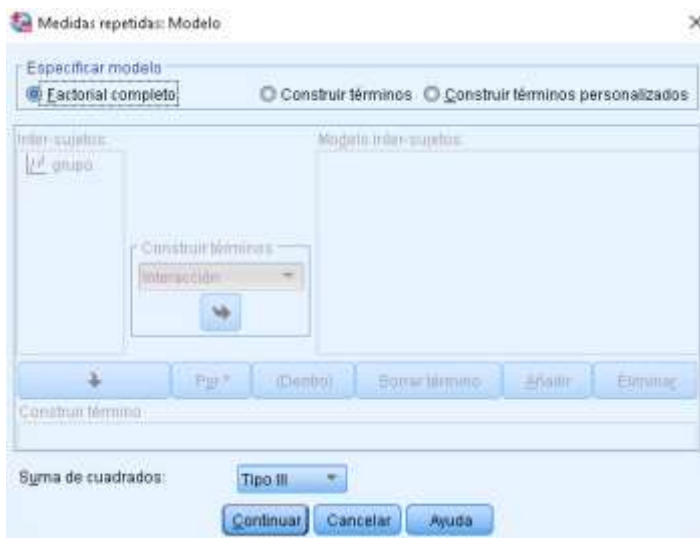


En las siguientes pantallas seleccionamos las variables en el orden que queramos comparar, en este caso c, a, b, y las pasamos a la pantalla de variables intrasujetos



A continuación pasamos a describir el uso de las pestañas del procedimiento

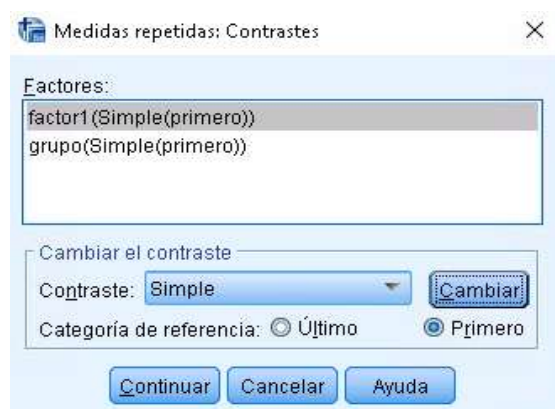
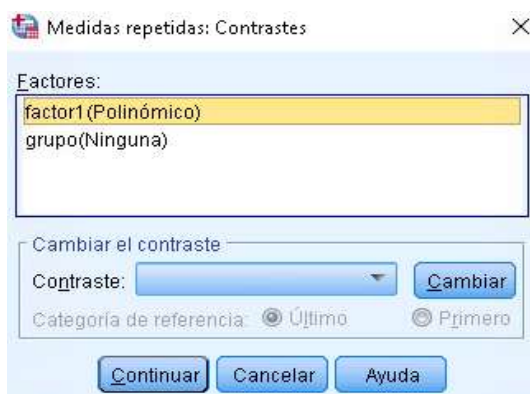
MODELO: Permite seleccionar qué efectos incluirá nuestro modelo:



Por defecto figura el modelo factorial completo. En ANOVAMR este modelo incluye todos los efectos posibles que son el de las diferencias entre momentos (tratamientos) (**factor intra-sujetos**. TAS), el de las diferencias entre categorías de la variable GRUPO (sexo) que es un factor no apareado, no son los mismos sujetos en cada sexo, y que se denomina **factor inter-sujetos**, y el de la interacción entre ambos. De existir, el efecto de interacción debe ser interpretado como que las diferencias en el factor intra-sujetos (TAS) entre momentos no son las mismas por categorías del factor inter-sujetos (GRUPO), es decir que no son las mismas en hombres que en mujeres. Si no hubiera interacción se interpretaría como que hombres y mujeres presentan las mismas diferencias entre momentos.

Nótese que la suma de cuadrados es Tipo III.

CONTRASTES: En esta pestaña podemos definir la forma de construir las variables de contraste (tipos **polinómico, diferencias, Helmert, repetido, simple y desviación**):



Por defecto los factores intra-sujetos aparecen con contraste polinómico y los inter-sujetos sin definición. Para simplificar interpretaciones es mejor utilizar contrastes sencillos, como el **simple**, que tiene interpretación similar a las dummies definidas en regresión. Con la pestaña contraste seleccionamos simple y tomamos marcamos la categoría de referencia primera, así comparará la TAS del momento 3 y 2 contra el 1 y el grupo 2 contra el 1 (en esta variable no era necesario definir contraste simple, ya que solo tiene dos categorías, lo hemos hecho a modo de ejemplo).

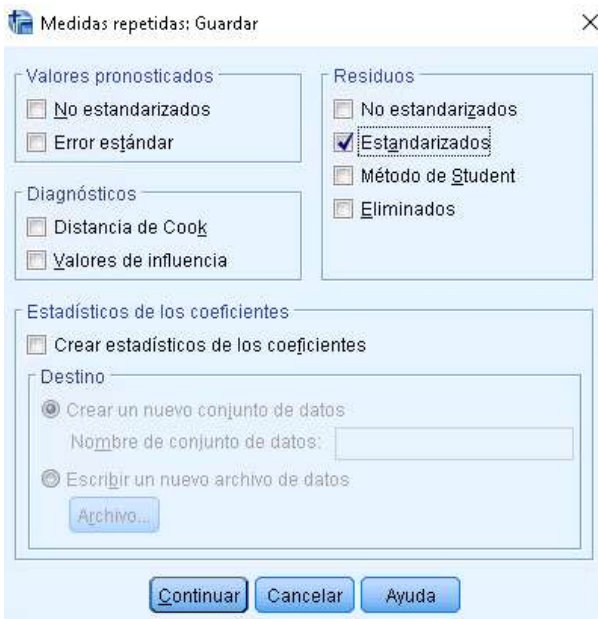
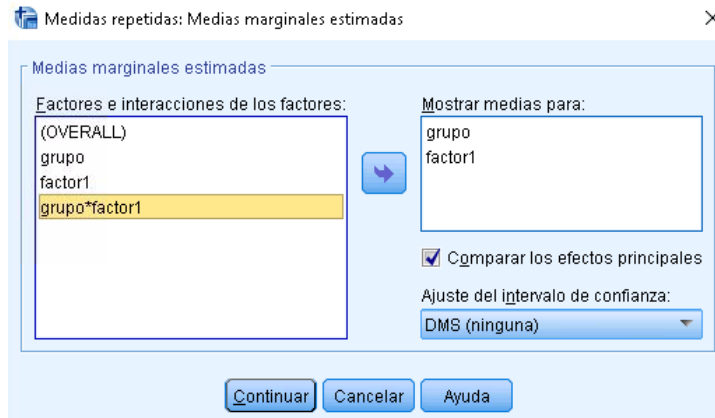
GRÁFICOS: Es conveniente visualizar información descriptiva para hacernos una idea del comportamiento de las variables.

En este caso se ha elegido el factor1 (recordemos que es el factor intra-sujetos, o sea la TAS en los 3 momentos) en el eje horizontal, mostrando líneas separadas según las categorías del factor inter-sujetos, o sea el grupo (1 hombres 2 mujeres).

Se han pedido también los intervalos de confianza para las medias en cada momento

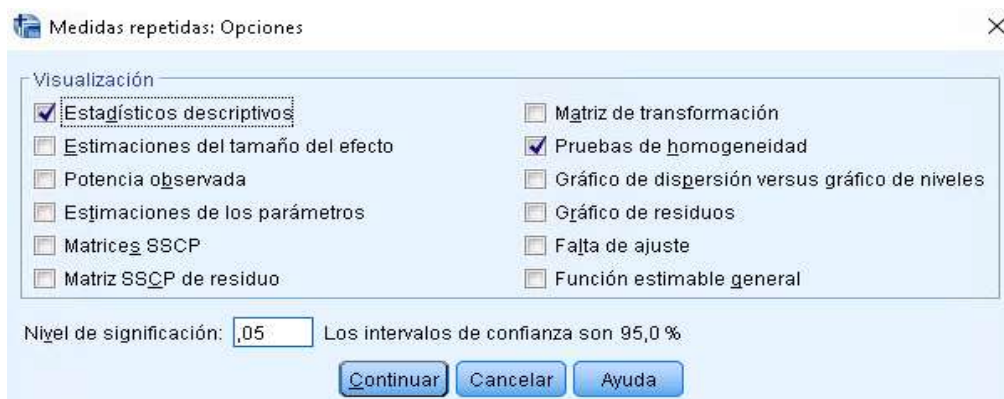
POSTHOC: Al igual que en el ANOVA Se pueden seleccionar pruebas para las comparaciones múltiples de las medias según categorías de los factores inter-sujetos, en este caso la variable independiente GRUPO. Hemos seleccionado comparaciones con Scheffé, pero no las realizará, ya que la variable grupo sólo tiene 2 categorías

MEDIAS MARGINALES ESTIMADAS: Opción más interesante que el posthoc, ya que permite realizar las comparaciones entre momentos del factor, o entre categorías de grupo para las medias estimadas por el modelo, pudiendo corregir las comparaciones por Bonferroni, si elegimos **Comparar los efectos principales**, con método **Bonferroni**, nos producirá los intervalos de confianza de las medias estimadas y para la diferencia entre las medias con corrección de Bonferroni.



GUARDAR: Pueden seleccionarse diversas variables resultado para añadirlas a la base de datos. En este caso se seleccionan los residuos para ver su normalidad. Sin embargo, hay que tener en cuenta que el ANOVAMR no tiene exactamente los mismos requerimientos de los modelos anteriores. Así, si la suposición de homogeneidad de varianzas y covarianzas nulas es aceptable, los **estadísticos univariados** son aceptados como correctos. Los **estadísticos multivariados** requieren normalidad multivariante, de difícil comprobación. Por tanto, los residuos seleccionados no serán de gran utilidad.

OPCIONES: En esta pestaña se pueden seleccionar diversos elementos. Elegimos los estadísticos descriptivos y la prueba de homogeneidad de varianzas para el modelo propuesto.



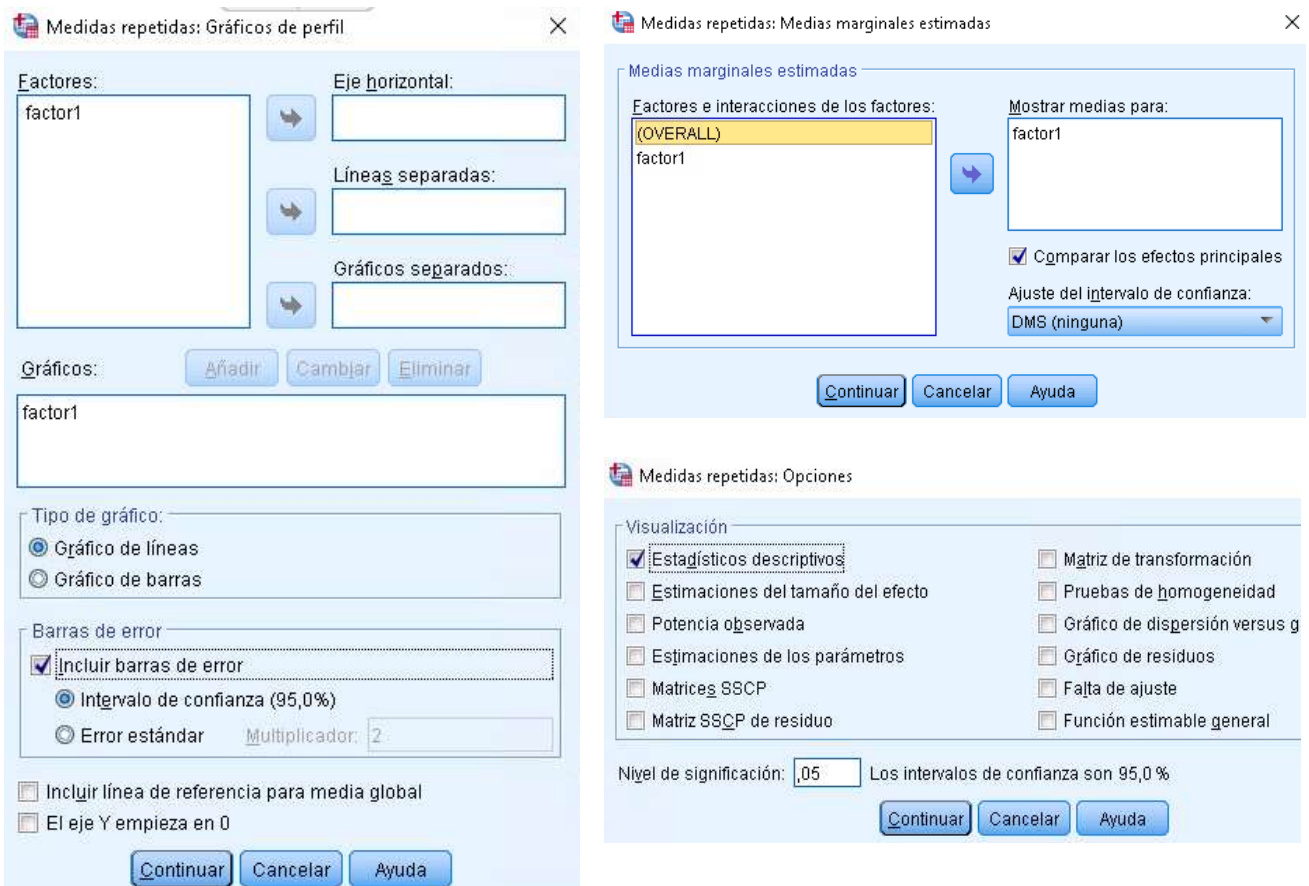
Resolución del ejercicio

Realice un ANOVA de medidas repetidas para comprobar si hay diferencias en las medias de tensión arterial entre los tratamientos

Para empezar realizaremos un ANOVAMR sólo para comprobar las diferencias en el factor intra-sujetos, o sea diferencias en TAS entre los tratamientos/momentos, sin introducir la variable inter-sujetos, GRUPO. En **CONTRASTES**, elegimos el tipo **simple**, con categoría de referencia la **primera** (el tratamiento c, primer tratamiento)



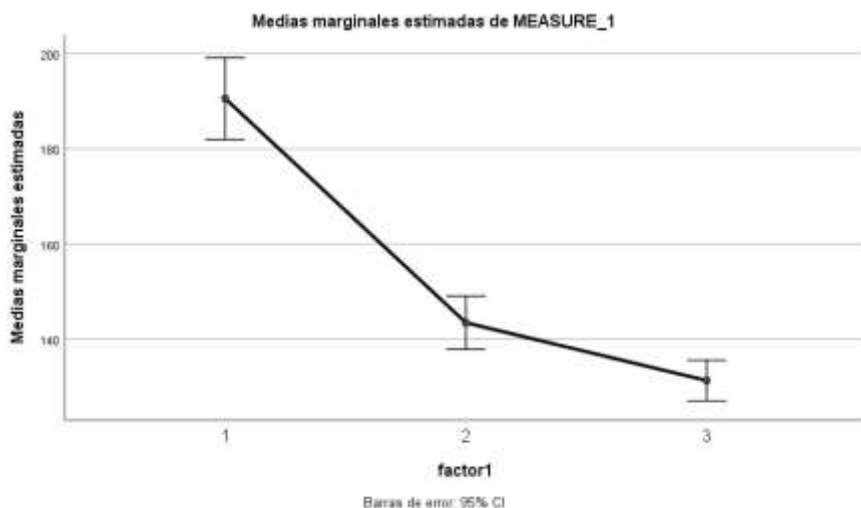
Pedimos un **GRÁFICO** de las medias en los 3 momentos, las **MEDIAS MARGINALES ESTIMADAS**, con comparación entre los grupos ajustada por Bonferroni, y los **descriptivos** en **OPCIONES**. La homogeneidad de varianzas no es necesaria, ya que la comprobaremos, junto con la de covarianzas nulas con la prueba de Mauchly que sale por defecto siempre:



Veamos los resultados, empezando por los descriptivos y los gráficos:

Estadísticos descriptivos

	Media	Desv. Desviación	N
c	190,60	18,317	20
a	143,45	11,958	20
b	131,30	9,102	20



Se observa una disminución importante del tratamiento c al a, y continúa disminuyendo al pasar al b.

En segundo lugar veremos si podemos afirmar la existencia de diferencias significativas entre los 3 tratamientos:

Pruebas multivariante^a

Efecto		Valor	F	gl de hipótesis	gl de error	Sig.
factor1	Traza de Pillai	,932	122,449 ^b	2,000	18,000	,000
	Lambda de Wilks	,068	122,449 ^b	2,000	18,000	,000
	Traza de Hotelling	13,605	122,449 ^b	2,000	18,000	,000
	Raíz mayor de Roy	13,605	122,449 ^b	2,000	18,000	,000

a. Diseño : Intersección

Diseño intra-sujetos: factor1

b. Estadístico exacto

La tabla de estadísticos multivariantes detecta diferencias muy significativas ($p < 0,001$) entre tratamientos (recordemos que los tratamientos son identificados como **factor1** en el procedimiento, es el factor intra-sujetos, en esta tabla todos los estadísticos coinciden siempre en ANOVAMR. Se debe suponer normalidad multivariante, algo difícil de asegurar.

Prueba de esfericidad de Mauchly^a

Medida: MEASURE_1

Efecto intra-sujetos	W de Mauchly	Aprox. Chi-cuadrado	gl	Sig.	Épsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Límite inferior
factor1	,885	2,190	2	,335	,897	,985	,500

Prueba la hipótesis nula de que la matriz de covarianzas de error de las variables dependientes con transformación ortonormalizada es proporcional a una matriz de identidad.

a. Diseño : Intersección

Diseño intra-sujetos: factor1

b. Se puede utilizar para ajustar los grados de libertad para las pruebas promedio de significación. Las pruebas corregidas se visualizan en la tabla de pruebas de efectos intra-sujetos.

La prueba de esfericidad de Mauchly da un resultado no significativo ($p=0,335$), lo cual debe ser interpretado como que podemos aceptar el requerimiento de homogeneidad de varianzas y covarianzas nulas (esfericidad), lo cual da soporte a las pruebas siguientes:

Pruebas de efectos intra-sujetos

Medida: MEASURE_1

Origen		Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
factor1	Esfericidad asumida	39248,233	2	19624,117	156,474	,000
	Greenhouse-Geisser	39248,233	1,794	21872,089	156,474	,000
	Huynh-Feldt	39248,233	1,970	19926,530	156,474	,000
	Límite inferior	39248,233	1,000	39248,233	156,474	,000
Error(factor1)	Esfericidad asumida	4765,767	38	125,415		
	Greenhouse-Geisser	4765,767	34,094	139,781		
	Huynh-Feldt	4765,767	37,423	127,348		
	Límite inferior	4765,767	19,000	250,830		

Las pruebas **intra-sujetos** contrastan la existencia o no de diferencias entre los tratamientos y son aceptables al aceptarse el requerimiento de homogeneidad de varianzas y covarianzas nulas. Aunque parecen iguales, dando como resultado la existencia de diferencias, la significación no es exactamente la misma (los grados de libertad, gl, no son exactamente iguales). Como aquí se puede asumir la esfericidad, podemos leer la primera prueba de Esfericidad asumida, $p<0,001$.

Se concluye la existencia de diferencias significativas entre los tratamientos. La pregunta ahora sería ¿entre qué tratamientos?

Pruebas de contrastes intra-sujetos

Medida: MEASURE_1

Origen		Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
factor1	Nivel 2 versus nivel 1	44462,450	1	44462,450	142,977	,000
	Nivel 3 versus nivel 1	70329,800	1	70329,800	257,956	,000
Error(factor1)	Nivel 2 versus nivel 1	5908,550	19	310,976		
	Nivel 3 versus nivel 1	5180,200	19	272,642		

Esta parte de la salida corresponde a la elección del **tipo simple** en la pestaña de **CONTRASTES**. Se tomó como referencia el tratamiento primero c (nivel 1), y se comprueba que el tratamiento a (nivel 2) y b (nivel 3) difieren significativamente del c ($p<0,001$). Pero no sabemos si el tratamiento b difiere del a.

Compruebe entre que tratamientos existen diferencias significativas

Los resultados de las medias marginales estimadas permiten realizar comparaciones múltiples entre tratamientos, ajustando por Bonferroni

Medias marginales estimadas

Estimaciones

Medida: MEASURE_1

factor1	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	190,600	4,096	182,027	199,173
2	143,450	2,674	137,853	149,047
3	131,300	2,035	127,040	135,560

Comparaciones por parejas

Medida: MEASURE_1

(I) factor1	(J) factor1	Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b	
					Límite inferior	Límite superior
1	2	47,150	3,943	,000	38,897	55,403
	3	59,300	3,692	,000	51,572	67,028
2	1	-47,150	3,943	,000	-55,403	-38,897
	3	12,150	2,906	,001	6,068	18,232
3	1	-59,300	3,692	,000	-67,028	-51,572
	2	-12,150	2,906	,001	-18,232	-6,068

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.

b. Ajuste para varias comparaciones: menor diferencia significativa (equivalente a sin ajustes).

Concluyendo que hay diferencia significativa en las medias de TAS entre cualesquiera tratamientos.

Otra forma de llegar a esta conclusión hubiera sido cambiando la definición del tipo de **CONTRASTE**, eligiendo, por ejemplo, **tipo simple** con referencia la **última** categoría (tratamiento b):

Pruebas de contrastes intra-sujetos

Medida: MEASURE_1

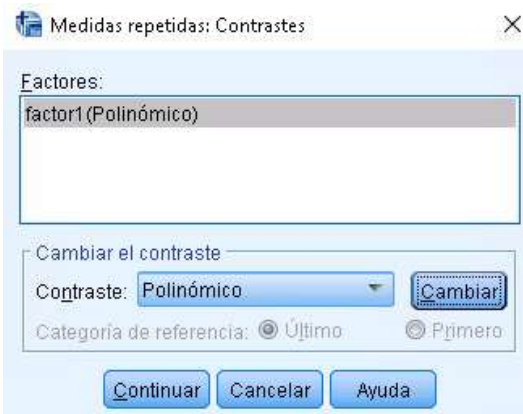
Origen	factor1	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
factor1	Nivel 1 versus nivel 3	70329,800	1	70329,800	257,956	,000
	Nivel 2 versus nivel 3	2952,450	1	2952,450	17,483	,001
Error(factor1)	Nivel 1 versus nivel 3	5180,200	19	272,642		
	Nivel 2 versus nivel 3	3208,550	19	168,871		

que nos hubiera dado las diferencias entre el Nivel 2 (tratamiento a) y Nivel 3 (tratamiento b) que nos faltaba

Compruebe si las medias presentan tendencia lineal o cuadrática



Para ello deberíamos elegir la opción **Polinómico** en **CONTRASTES**:



Pruebas de contrastes intra-sujetos

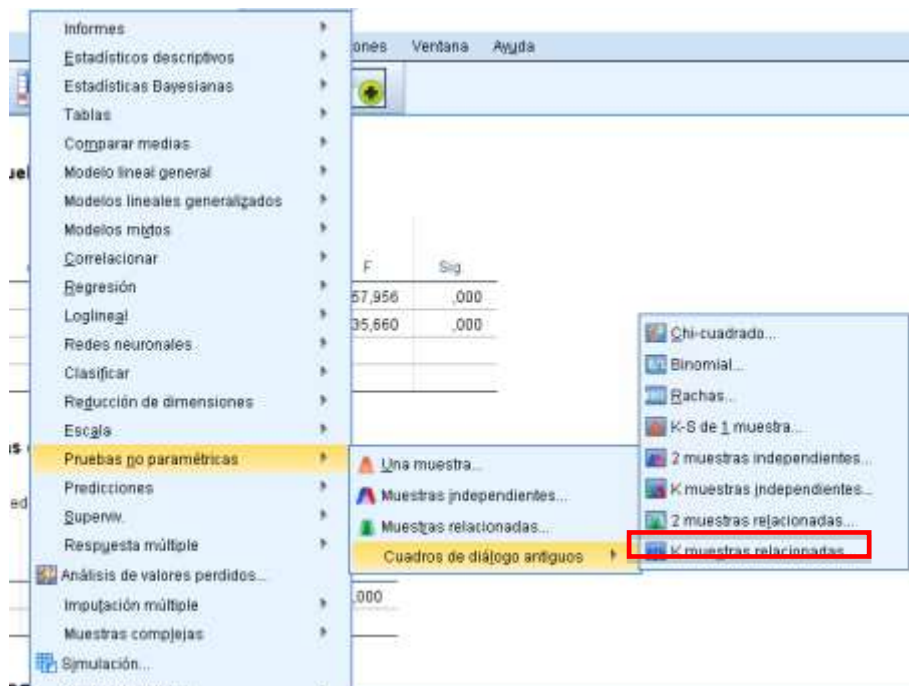
Medida: MEASURE_1

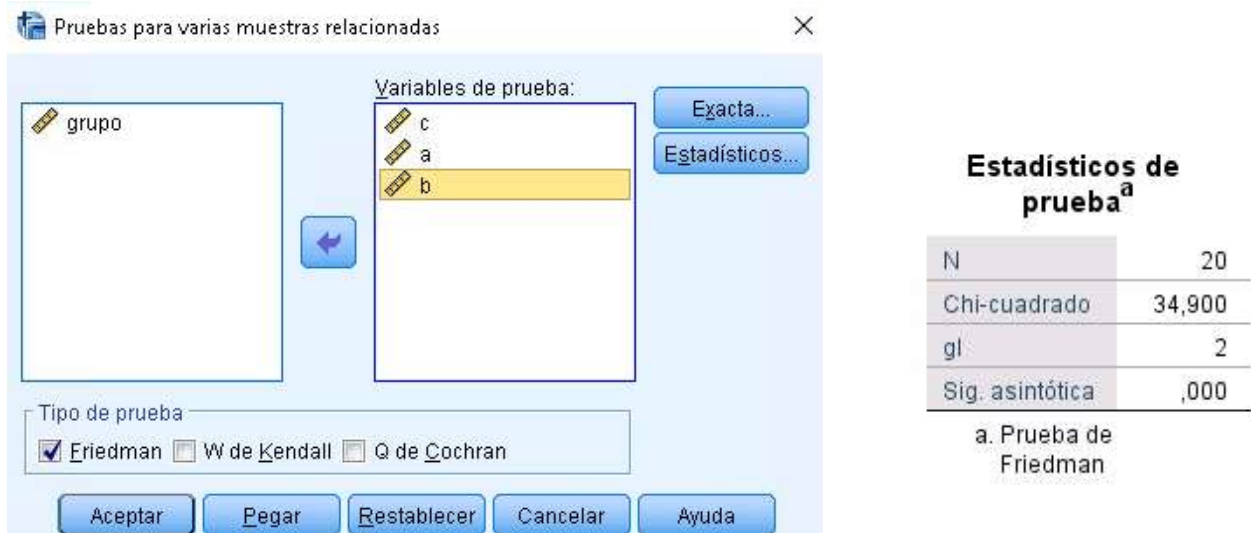
Origen	factor1	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
factor1	Lineal	35164,900	1	35164,900	257,956	,000
	Cuadrático	4083,333	1	4083,333	35,660	,000
Error(factor1)	Lineal	2590,100	19	136,321		
	Cuadrático	2175,667	19	114,509		

que sólo modifica en la salida la tabla de Pruebas de contrastes intra-sujetos, donde vemos que aparece el efecto lineal y el efecto cuadrático (polinomio de grado dos). Hay que recordar que el grado máximo del polinomio a ajustar es el del número de momentos/puntos (en este caso 3) menos 1. Ambos efectos son significativos ($p < 0,001$), lo cual nos lleva a quedarnos con el polinomio de mayor grado, es decir el cuadrático, ya que la componente lineal es significativa y también la cuadrática. Si solo hubiera sido significativa la lineal diríamos que la tendencia es lineal.

Realice pruebas no paramétricas para comprobar las diferencias entre tratamientos

A veces los requerimientos no se cumplen (prueba de Mauchly significativa) y es conveniente pensar en hacer comparaciones no paramétricas para reforzar los resultados. La comparación entre varios grupos apareados se puede resolver con la **prueba de Friedman**, y las comparaciones 2 a 2 con la **prueba de Wilcoxon por rangos** (con corrección de Bonferroni cuando hay 3 o más de comparaciones). La prueba de Friedman se obtendría como:





Estadísticos de prueba^a

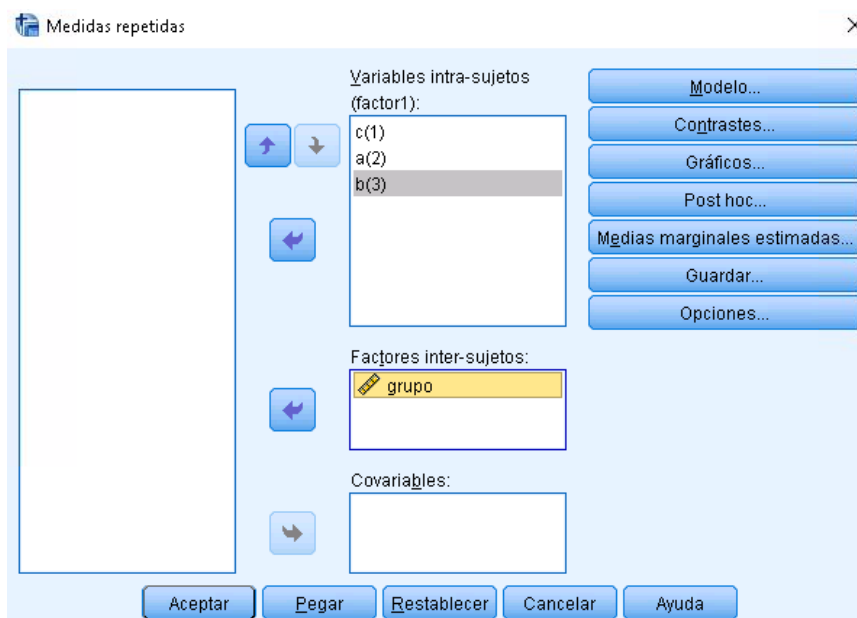
N	20
Chi-cuadrado	34,900
gl	2
Sig. asintótica	,000

a. Prueba de Friedman

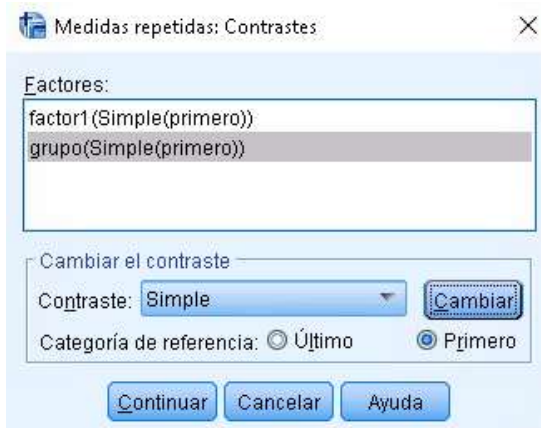
y podemos observar que la prueba de Friedman detecta diferencias significativas ($p < 0,001$) en las medianas de los 3 tratamientos.

El efecto del tratamiento sobre las medias de tensión arterial podría interaccionar con el sexo (ser distinto en hombres y en mujeres). Compruebe esta posibilidad introduciendo la variable grupo como un factor intersujetos. ¿Hay interacción significativa? ¿Hay efecto del factor?

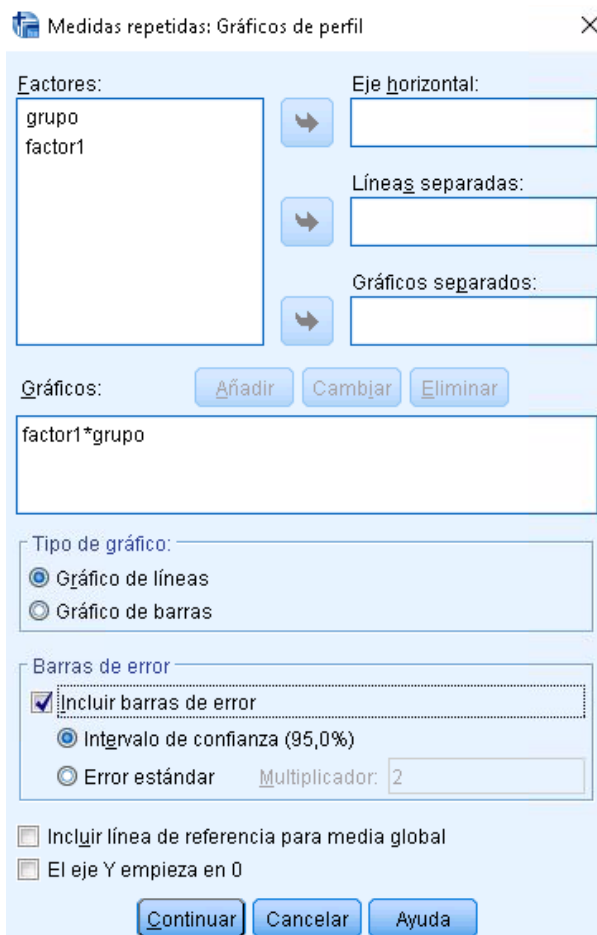
Realizaremos en ANOVAMR con 1 **factor intra-sujetos** (tratamientos c, a, b) y un factor **inter-sujetos** (variable GRUPO, 1 hombre, 2 mujer).



El Modelo será por defecto el Factorial completo, que incluirá todos los efectos. Tratamiento, Grupo e Interacción



Mantendremos la definición anterior de **CONTRASTES** en el **tipo simple**, con referencia la **primera** categoría de cada variable

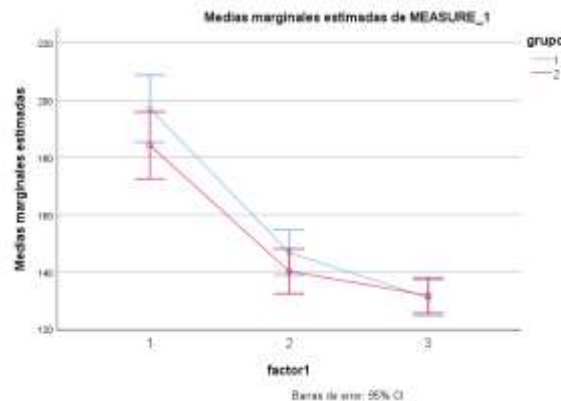


En **GRÁFICOS** seleccionamos el tratamiento (FACTOR1) en el eje horizontal y líneas separadas por GRUPO. No ayudará a interpretar el posible efecto de interacción (si lo hubiera). También podemos pedir los **Descriptivos** de la pestaña de **OPCIONES**

Puesto que vamos a indagar primero el posible efecto de interacción no pedimos nada más. Si observamos los resultados descriptivos y de gráficos en primer lugar, tenemos:

Estadísticos descriptivos

	grupo	Media	Desv. Desviación	N
c	1	197,00	18,403	10
	2	184,20	16,692	10
	Total	190,60	18,317	20
a	1	146,70	11,935	10
	2	140,20	11,660	10
	Total	143,45	11,958	20
b	1	130,90	10,588	10
	2	131,70	7,903	10
	Total	131,30	9,102	20



Podemos observar que, aunque empiezan de niveles ligeramente diferentes hombres y mujeres, al final acaban igual, con un comportamiento similar, la impresión es de rectas paralelas, y sugiere la no existencia de interacción. Veámoslo:

Prueba de esfericidad de Mauchly^a

Medida: MEASURE_1

Efecto intra-sujetos	W de Mauchly	Aprox. Chi-cuadrado	gl	Sig.	Épsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Límite inferior
factor1	,862	2,526	2	,283	,879	1,000	,500

Prueba la hipótesis nula de que la matriz de covarianzas de error de las variables dependientes con transformación ortonormalizada es proporcional a una matriz de identidad.

a. Diseño : Intersección + grupo

Diseño intra-sujetos: factor1

b. Se puede utilizar para ajustar los grados de libertad para las pruebas promedio de significación. Las pruebas corregidas se visualizan en la tabla de pruebas de efectos intra-sujetos.

Pruebas de efectos intra-sujetos

Medida: MEASURE_1

Origen		Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.
factor1	Esfericidad asumida	39248,233	2	19624,117	164,198	,000
	Greenhouse-Geisser	39248,233	1,757	22333,872	164,198	,000
	Huynh-Feldt	39248,233	2,000	19624,117	164,198	,000
	Límite inferior	39248,233	1,000	39248,233	164,198	,000
factor1 * grupo	Esfericidad asumida	463,233	2	231,617	1,938	,159
	Greenhouse-Geisser	463,233	1,757	263,599	1,938	,165
	Huynh-Feldt	463,233	2,000	231,617	1,938	,159
	Límite inferior	463,233	1,000	463,233	1,938	,181
Error(factor1)	Esfericidad asumida	4302,533	36	119,515		
	Greenhouse-Geisser	4302,533	31,632	136,018		
	Huynh-Feldt	4302,533	36,000	119,515		
	Límite inferior	4302,533	18,000	239,030		

La prueba de esfericidad de Mauchly no es significativa y las Pruebas de efectos intra-sujetos incluyen un efecto de factor1 (tratamiento) que es significativo ($p < 0,001$) y un efecto de interacción, que no es significativo con ninguna de las pruebas disponibles (aunque aquí será adecuada la de Esfericidad asumida ya que Mauchly no es significativa).



Concluimos por tanto la no existencia de interacción con el grupo y los resultados de los apartados anteriores son los adecuados.

Caso de haber existido interacción significativa, lo cual se interpretaría como que las diferencias (o su ausencia) entre tratamientos no serían las mismas en hombres y en mujeres, la opción que permitiría explicar mejor los comportamientos sería replicar los análisis anteriores (de i a iv) pero por separado para hombres y mujeres y así ver en que se traduciría la interacción.



Sesión 4

Práctica de revisión de Análisis de la varianza (ANOVA) y de la covarianza (ANCOVA)



Sesión 4.- Práctica de revisión de Análisis de la varianza (ANOVA) y de la covarianza (ANCOVA)

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

Aplicar un análisis de la varianza sobre un conjunto de datos, identificando los modelos ANOVA y ANCOVA pertinentes

Interpretar los resultados obtenidos en términos de los efectos y relaciones entre las variables implicadas y las preguntas de interés sobre las mismas

Actividades: Las actividades previstas son las siguientes

Revisión de lo aprendido en las sesiones 1 y 2

Realización de la actividad obligatoria propuesta

Elaboración de un documento en formato de word o pdf en el que se contesten las preguntas planteadas en la actividad (Extensión máxima 5 páginas)

Contenidos teóricos: Material teórico pgs. 20-41 de la monografía Modelos de Análisis de la Varianza y Covarianza. <http://hdl.handle.net/10045/113344>

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Sesiones 1 y 2
- Base de datos: PAS.sav



Sesión 4.- Actividad obligatoria

Actividad práctica

La base PAS.SAV contiene datos obtenidos sobre 32 sujetos de las siguientes variables:

PAS: Presión arterial sistólica	QUETELET: Índice de masa corporal
EDAD: En años	TABACO: 0 'No fumador' 1 'Fumador'
NEDUC: Nivel educativo 1 'Sin estudios' 2 'EGB' 3 'Secundaria' 4 'Superiores'	
ECG: 0 'Sin antecedentes de anormalidad' 1 'Con antecedentes'	

1. Queremos averiguar si el tabaco es factor de riesgo para la tensión arterial, ajustado por edad. Para ello:

Construya y discuta el diagrama de dispersión de la PAS como función de la edad según tabaco. ¿Qué le sugiere el diagrama?

Realice un ANCOVA que le permita discutir la existencia o no de interacción entre edad , tabaco y pas

Compruebe y valore si existe confusión por parte de la edad sobre el efecto del tabaco

¿Cuánto vale el efecto del tabaco sobre la pas?

Para el mejor modelo obtenido, compruebe sus requerimientos

2. Realice un ANOVA para comprobar los efectos existentes entre tabaco y nivel educativo sobre la pas.

¿Tiene efecto significativo el nivel educativo? Dé el valor de la p

¿Hay interacción significativa? Dé el valor de p

¿Tiene efecto el tabaco? Dé el valor de p

Compruebe los requerimientos del modelo. Dé el valor de la p de homogeneidad de varianzas y de la prueba de normalidad sobre los residuos

3. Realice un ANCOVA con variable respuesta la pas y explicativas la edad, quetelet, tabaco y nivel educativo.

Teniendo en cuenta que solo tienen interés las interacciones entre tabaco y el resto de variables (cualitativa o cuantitativas), compruebe si existe alguna de estas interacciones significativa

Si ninguna de las interacciones es significativa, obtenga el modelo óptimo para explicar la pas ¿qué variables tienen efecto significativo? ¿qué variables inducen confusión sobre el tabaco?



Sesión 5

Práctica de revisión de Análisis de la varianza de medidas repetidas (ANOVAMR)



Sesión 5. - Práctica de revisión de Análisis de la varianza de medidas repetidas (ANOVAMR)

Objetivos: Al finalizar las actividades previstas el alumno debe ser capaz de

Aplicar un análisis de la varianza sobre un conjunto de datos, identificando los modelos ANOVAMR pertinentes

Interpretar los resultados obtenidos en términos de los efectos y relaciones entre las variables implicadas y las preguntas de interés sobre las mismas

Actividades: Las actividades previstas son las siguientes

Revisión de lo aprendido en la sesión 3

Realización de la actividad obligatoria propuesta

Elaboración de un documento en formato de word o pdf en el que se contesten las preguntas planteadas en la actividad (Extensión máxima 5 páginas)

Contenidos teóricos: Material teórico pgs. 41-50 de la monografía Modelos de Análisis de la Varianza y Covarianza, <http://hdl.handle.net/10045/113344>

Materiales:

- Modelos de Análisis de la Varianza y Covarianza
- Sesión 3
- Base de datos: APAREADO.sav



Sesión 5.- Actividad obligatoria

Actividad práctica

La base APAREADO.sav contiene datos obtenidos sobre 20 sujetos de las siguientes variables:

URICO1: Nivel ac. úrico al inicio de un tratamiento consistente en una dieta
URICO2: Nivel ac. úrico 3 meses después del inicio del tratamiento
URICO3: Nivel ac. úrico 6 meses después del inicio del tratamiento
DIETA: 1 'No hacen dieta' 2 'Hacen dieta'

Queremos averiguar si se puede afirmar significativamente que la dieta ha sido efectiva, disminuyendo el ácido úrico. Realice las siguientes aplicaciones con el procedimiento de ANOVA de medidas repetidas

A través del gráfico de las medias y de los estadísticos descriptivos que le permite el procedimiento sugiera que efectos podrían estar produciéndose (diferencias en las medias, efecto de interacción)

Compruebe la existencia o no de interacción significativa a través de las Pruebas univariadas de efectos intra-sujetos. ¿Puede asumirse la esfericidad? ¿Qué prueba univariada elegiría de las 4 disponibles para discutir los efectos existentes? ¿Porqué? ¿Se puede afirmar que hay interacción significativa? ¿Porqué?

Descartada la interacción, repita el análisis anterior pero sin introducir la variable dieta. ¿Podemos afirmar que existe una disminución significativa del ácido úrico a lo largo del tratamiento?

Utilizando los CONTRASTES adecuados, compruebe si hay diferencia significativa en las medias de ácido úrico desde el inicio hasta los 3 meses y hasta los 6 meses, y también si hay diferencia significativa de los 3 a los 6 meses

Averigüe si las medias de ácido úrico presentan una tendencia lineal o cuadrática

Aunque los resultados no permiten concluir la efectividad de la dieta, ¿qué le sugieren a Vd.? ¿Porqué no podemos concluir significativamente la efectividad de la dieta?



Anexo

Bases de datos



DATOS EJEMPLO TEXTO.sav

Caso	edad	peso	colest	albumina	calcio	urico	alcohol	edadrec
1	19	50	158	41	99	47	1	1
2	20	47	210	39	95	40	1	1
3	20	48	192	38	93	50	1	1
4	21	54	245	34	106	38	1	1
5	21	48	260	47	106	38	1	1
6	21	54	204	40	108	34	1	1
7	21	58	192	39	95	49	1	1
8	21	55	280	41	102	41	1	1
9	21	45	230	39	99	38	1	1
10	21	64	215	39	96	39	1	1
11	21	46	225	44	105	44	1	1
12	21	50	165	48	105	28	1	1
13	21	42	200	38	95	40	1	1
14	21	52	255	34	102	40	1	1
15	22	54	263	43	98	47	1	1
16	22	44	173	42	97	37	1	1
17	22	42	170	46	98	45	1	1
18	22	46	263	42	102	47	1	1
19	22	38	220	47	105	46	1	1
20	22	50	200	43	100	44	1	1
21	22	39	192	38	95	43	1	1
22	22	40	155	41	96	45	1	1
23	22	54	215	40	93	43	1	1
24	22	38	200	47	99	34	1	1
25	22	50	247	44	102	45	1	1
26	23	50	220	32	92	42	1	1
27	23	42	207	42	100	40	1	1
28	23	50	266	42	103	47	1	1
29	23	48	240	43	101	39	1	1
30	24	52	250	39	103	46	1	1
31	24	52	225	50	108	39	1	1
32	24	59	200	37	104	49	1	1
33	24	54	180	37	96	49	1	1
34	25	51	243	41	104	33	1	1
35	25	60	150	38	96	30	1	1
36	25	45	205	46	101	33	1	1
37	25	48	235	44	103	40	1	1
38	25	54	295	46	106	47	1	1
39	26	47	238	40	99	46	1	1



40	26	50	198	44	96	43	1	1
41	26	48	196	38	95	43	1	1
42	27	55	195	46	101	42	1	1
43	27	50	168	42	97	41	1	1
44	28	44	250	44	105	38	1	1
45	28	48	175	48	100	47	1	1
46	28	45	305	41	93	24	1	1
47	28	54	200	43	97	37	1	1
48	28	64	235	42	101	41	1	1
49	29	57	177	39	99	46	1	1
50	29	46	235	45	98	47	1	1
51	29	62	226	38	94	43	1	1
52	29	47	230	44	99	44	1	1
53	30	44	295	45	98	46	1	2
54	30	40	230	43	99	39	1	2
55	30	53	200	37	96	34	1	2
56	30	50	230	46	104	48	1	2
57	30	44	262	33	99	41	1	2
58	30	54	174	40	95	35	1	2
59	30	45	250	44	100	35	1	2
60	30	64	217	35	95	31	1	2
61	31	50	250	43	98	39	1	2
62	31	48	237	34	91	49	1	2
63	31	44	280	44	99	49	1	2
64	31	49	238	37	96	33	1	2
65	31	54	218	38	95	42	1	2
66	32	53	185	39	103	37	1	2
67	32	81	235	38	99	37	1	2
68	32	62	262	37	99	43	1	2
69	32	50	160	41	97	40	1	2
70	32	50	189	40	94	40	1	2
71	32	48	260	43	107	38	1	2
72	32	58	240	45	108	42	1	2
73	32	53	180	32	95	40	1	2
74	33	47	260	38	99	38	1	2
75	33	52	203	44	101	48	1	2
76	33	55	222	40	104	42	1	2
77	34	45	197	37	93	44	1	2
78	34	50	245	36	95	41	1	2
79	35	50	223	40	100	37	1	2
80	35	55	254	39	107	41	1	2
81	36	54	247	34	90	44	1	2
82	36	48	175	46	103	39	1	2



83	36	45	215	43	104	42	1	2
84	36	48	270	43	98	35	1	2
85	37	56	237		105	45	1	2
86	37	46	270	42	100	48	1	2
87	37	52	230	36	91	22	1	2
88	38	60	275	38	94	46	1	2
89	39	54	210	40	95	46	1	2
90	39	43	198	44	90	38	1	2
91	39	78	260	40	108	42	1	2
92	39	53	180	39	94	30	1	2
93	39	40	210	45	91	27	1	2
94	39	44	235	41	99	35	1	2
95	40	44	196	39	97	52	1	3
96	40	60	305	39	99	48	1	3
97	40	58	170	45	100	63	1	3
98	40	56	272	41	91	54	1	3
99	40	47	290	42	99	52	1	3
100	41	86	255	43	105	45	1	3
101	42	52	210	40	100	46	1	3
102	43	58	250	36	98	52	1	3
103	43	53	230	42	98	48	1	3
104	43	50	285	45	105	50	1	3
105	43	45	200	40	93	46	1	3
106	43	50	280	45	106	48	1	3
107	43	59	276	41	105	50	1	3
108	44	47	253	43	94	54	1	3
109	44	53	242	47	104	49	1	3
110	46	58	320	40	101	47	1	3
111	46	47	190	38	95	43	1	3
112	46	41	230	43	102	53	1	3
113	47	54	297	42	100	45	1	3
114	47	57	255	41	100	50	1	3
115	47	53	257	39	96	48	1	3
116	48	57	216	40	96	47	1	3
117	48	54	248	42	102	42	1	3
118	48	48	235	36	97	45	1	3
119	49	50	255	41	102	48	1	3
120	53	71	227	39		35	1	3
121	53	56	220	40	107	30	1	3
122	54	48	227		86	25	1	3
123	54	58	320	39			1	3
124	54	63	305	42	103	98	1	3
125	19	62	255	45	105	83	2	1



126	21	73	246	42	101	52	2	1
127	21	60	208	38	98	54	2	1
128	21	60	220	47	102	75	2	1
129	22	58	200	43	98	54	2	1
130	22	64	150	35		72	2	1
131	22	68	290	37	98	59	2	1
132	22	50	247	42	104	52	2	1
133	22	52	175	44	106	58	2	1
134	24	60	195	49	106	52	2	1
135	24	62	240	42	102	51	2	1
136	25	53	330	48	101	53	2	1
137	25	70	175	39	93	51	2	1
138	25	56	230	38	101	52	2	1
139	26	64	240	48	103	51	2	1
140	27	58	172	43	98	60	2	1
141	27	72	317	37	98	84	2	1
142	27	50	185	36	94	54	2	1
143	27	50	200	40	96	52	2	1
144	27	56	250	36	98	68	2	1
145	27	62	280	42	103	52	2	1
146	28	43	260	48	106	51	2	1
147	30	57	218	45	107	65	2	2
148	30		232	38	100	66	2	2
149	31	46	290	41	111	64	2	2
150	32	68	225	37	90	60	2	2
151	32	56	217	44	106	58	2	2
152	33	46	225	47	100	54	2	2
153	33	55	263	41	106	55	2	2
154	33	52	215	40	95	58	2	2
155	35	46	200	40	91	59	2	2
156	35	56	265	39	105	56	2	2
157	37	50	220	45	99	66	2	2
158	38	66	275	44	102	62	2	2
159	40	56	276	46	100	65	2	3
160	40	55	315	37	96	72	2	3
161	40	52	300	40	106	62	2	3
162	41	46	320	44	111	71	2	3
163	41	50	306	45	98	72	2	3
164	41	68	324	40	99	65	2	3
165	42	42	240	41	101	61	2	3
166	43	67	210	40	100	62	2	3
167	43	55	335	44	105	68	2	3
168	45	72	160	38	97	69	2	3



169	45	56	263	45	107	62	2	3
170	46		250	41	101	63	2	3
171	46	55	257	40	90	71	2	3
172	46	76	265	41	108	55	2	3
173	47	38	257	41	103	63	2	3
174	48	48	300	39	94	61	2	3
175	48	57	225	40	100	72	2	3
176	48	66	306	44	100	68	2	3
177	48	50	195	41	93	63	2	3
178	48	55	338	57	100	68	2	3
179	49	63	217	36	106	75	2	3
180	50	54	295	43	105	73	2	3
181	50	56	390	46	97	65	2	3
182	52	60	250	42	97	69	2	3
183	52	43	265	46	104	74	2	3
184	54	51	260	44	106	67	2	3
185	54	56	245	39	104	66	2	3
186	54	68	220	35	88	93	2	3
187	55	50	275	40	98	63	2	3
188	55	66	298	36	100	73	2	3
Total	N	188	186	188	186	185	187	188



ANOVAS.sav

	sexo	edad	alcohol	talla	peso	glucosa	colestot	triglice	quetelet	calculos	Nivel de obesidad
1	1	47	2	173	69,0	1	263	39	28,05	1	1
2	1	67	1	168	66,0	2	163	86	23,38	1	1
3	1	67	2	154	69,0	2	304	355	34,09	1	2
4	1	23	2	165	67,0	1	205	68	24,61	1	1
5	1	25	2	178	76,0	1	221	55	23,99	1	1
6	1	58	1	176	90,0	1	268	144	29,05	1	2
7	1	20	2	167	52,0	1	216	49	18,65	1	1
8	1	56	2	174	83,0	1	199	79	32,41	1	2
9	2	34	1	162	62,0	1	235	70	23,62	1	1
10	2	65	1	149	70,0	2	237	101	31,53	1	3
11	2	65	1	149	77,0	1	305	140	34,68	1	3
12	2	70	1	146	72,0	1	235	148	33,78	1	3
13	1	41	1	166	65,0	1	159	59	23,59	1	1
14	2	57	1	151	68,0	1	344	126	29,82	1	2
15	2	65	1	156	70,0	2	206	148	28,76	1	2
16	2	53	1	157	80,0	1	242	139	32,46	2	3
17	2	20	1	159	60,0	1	163	58	23,73	1	1
18	1	27	1	177	67,0	1	161	128	21,39	1	1
19	1	41	2	142	65,0	1	269	213	37,24	1	3
20	2	57	2	164	89,0	1	322	84	38,09	1	3
21	2	62	1	159	82,0	1	266	85	32,44	2	3
22	1	46	2	193	101,0	1	156	146	32,11	1	2
23	1	67	1	168	72,0	1	278	80	25,51	2	2
24	2	40	1	160	74,0	1	254	55	28,91	1	2
25	1	46	1	161	84,0	1	347	252	32,41	1	3
26	1	41	1	178	99,0	1	239	191	31,25	1	3
27	2	39	1	149	52,0	1	218	98	23,42	1	1
28	2	59	1	158	75,0	1	222	90	30,04	1	3
29	1	27	2	171	65,0	1	195	75	22,23	1	1
30	2	20	1	159	50,0	1	180	96	19,78	1	1
31	2	65	2	143	64,0	2	311	123	36,30	1	3
32	1	34	2	164	65,0	1	309	51	24,17	1	1
33	1	50	2	163	79,0	1	256	187	34,73	1	2
34	2	45	1	164	65,0	1	251	107	24,17	1	1
35	2	50	1	145	58,0				27,59	1	2
36	2	28	1	149	58,0	1	216	79	26,12	1	2
37	1	66	2	172	70,0	1	250	83	28,66	1	1
38	2	57	1	158	82,0	2	217	125	32,85	1	3



39	2	65	1	150	68,0	1	268	170	30,22	2	3
40	1	55	2	167	74,0	1	228	100	31,53	1	2
41	1	71	1	164	86,0	1	167	97	31,98	1	3
42	2	62	1	149	64,0	1	218	54	28,83	1	2
43	2	39	1	156	64,0	1	182	55	26,30	1	2
44	2	28	1	162	61,0	1	160	37	23,24	1	1
45	1	59	1	179	94,0	1	290	120	29,34	2	2
46	1	46	1	170	79,0	2	314	414	27,34	1	2
47	1	72	2	174	83,0	1	299	174	32,41	1	2
48	2	45	1	157	75,0	1	245	95	30,43	1	3
49	1	35	2	173	88,0	1	346	508	29,40	1	2
50	2	26	1	162	57,0	1	212	84	21,72	1	1
51	2	45	2	156	66,0	1	230	83	32,12	1	2
52	2	61	1	158	69,0	1	197	116	27,64	1	2
53	1	55	1	166	75,0	2	212	534	27,22	1	2
54	2	38	2	159	63,0	1	186	38	29,92	1	1
55	1	31	2	183	98,0	1	194	48	29,26	1	2
56	2	67	1	150	68,0	1	302	93	30,22	1	3
57	1	55	2	172	84,0	1	262	209	33,39	1	2
58	2	53	1	154	73,0	1	228	164	30,78	2	3
59	1	45	1	170	90,0	1	209	98	31,14	1	3
60	2	48	1	162	78,0	1	167	100	29,72	1	2
61	2	38	1	163	63,0	1	185	39	23,71	1	1
62	1	60	2	168	69,0	1	188	109	29,45	1	1
63	1	66	2	160	65,0	1	209	130	30,39	1	2
64	1	25	1	167	68,0	1	226	90	24,38	1	1
65	1	54	2	174	86,0	2	180	92	33,41	2	2
66	2	28	1	160	55,0	1	251	145	21,48	1	1
67	2	65	1	160	70,0	2	326	208	27,34	1	2
68	2	60	1	156	85,0	1	195	162	34,93	1	3
69	2	26	1	170	73,0	1	156	64	25,26	1	2
70	2	70	1	152	56,0	1	241	46	24,24	1	1
71	2	64	1	149	54,0	1	294	107	24,32	2	1
72	2	26	1	155	62,0	1	198	76	25,81	1	2
73	2	56	1	160	80,0	1	209	85	31,25	1	3
74	2	23	1	152	58,0	1	239	80	25,10	1	2
75	2	60	1	161	72,0	1	234	40	27,78	1	2
76	2	56	1	159	91,0	1	241	106	36,00	1	3
77	1	38	2	166	64,0	1	234	418	28,23	1	1
78	1	54	1	167	65,0	1	188	66	23,31	1	1
79	2	51	1	154	70,0	1	266	99	29,52	1	2
80	2	43	1	157	62,0	1	227	83	25,15	1	2
81	1	56	2	172	80,0	2	232	529	32,04	2	2



82	1	25	1	177	67,0	1	156	190	21,39	1	1
83	2	48	1	163	75,0	1	261	160	28,23	1	2
84	1	29	1	165	77,0	1	231	58	28,28	1	2
85	1	65	1	165	66,0	1	202	54	24,24	1	1
86	1	64	1	170	85,0	1	225	83	29,41	2	2
87	1	49	2	168	91,0	1	306	174	37,24	1	3
88	1	57	2	162	59,0	1	355	191	27,48	1	1
89	2	60	1	151	58,0	2	310	119	25,44	1	2
90	1	45	1	169	83,0	1	232	125	29,06	1	2
91	1	69	2	163	65,0	1	169	41	29,46	1	1
92	1	29	2	174	79,0	1	344	319	26,09	1	2
93	2	25	1	169	55,0	1	173	31	19,26	1	1
94	2	59	1	162	76,0	1	301	107	28,96	2	2
95	2	68	1	157	50,0	1	203	74	20,28	1	1
96	1	68	1	161	72,0	1	193	97	27,78	1	2
97	2	70	1	148	63,0	1	188	69	28,76	1	2
98	2	64	1	158	79,0	1	263	134	31,65	1	3
99	1	63	2	161	78,0	1	196	117	35,09	1	3
100	1	24	1	159	65,0	1	187	58	25,71	1	2
101	1	48	2	180	86,0	1	180	96	31,54	1	2
102	1	29	2	183	86,0	1	201	94	25,68	1	2
103	2	47	1	153	66,0	2	206	409	28,19	1	2
104	1	36	1	177	68,0	1	214	72	21,71	1	1
105	1	32	1	174	88,0	1	263	179	29,07	1	2
106	2	25	1	165	55,0	1	160	41	20,20	1	1
107	2	60	2	155	70,0	1	298	73	34,14	1	2
108	2	36	2	158	53,0	1	236	34	26,23	1	1
109	1	65	2	167	72,0	1	237	61	30,82	1	2
110	1	60	2	165	79,0	2	337	145	34,02	1	2
111	1	69	1	163	81,0	2	259	151	30,49	1	3
112	2	74	1	153	63,0	1	313	114	26,91	1	2
113	1	55	2	162	70,0	1	219	86	31,67	1	2
114	2	34	2	152	44,0	1	233	32	19,04	1	1
115	2	63	1	147	55,0	2	254	345	25,45	2	2
116	2	58	1	157	67,0	1	316	152	27,18	1	2
117	1	30	1	173	81,0	1	203	46	27,06	1	2
118	2	22	1	161	50,0	1	162	46	19,29	1	1
119	1	66	1	162	77,0	1	214	90	29,34	1	2
120	2	54	2	158	72,0	1	289	167	33,84	1	2
121	2	57	1	155	73,0	1	250	43	30,39	1	3
122	2	57	1	153	82,0	1	290	208	35,03	1	3
123	1	70	2	149	51,0	1	172	64	27,97	1	1
124	2	70	1	151	65,0	1	247	118	28,51	1	2



125	2	71	1	151	70,0	1	209	82	30,70	1	3
126	2	75	1	149	59,0	2	235	307	26,58	2	2
127	2	62	2	158	53,0	1	233	64	26,23	1	1
128	1	67	1	162	78,0	1	248	167	29,72	1	2
129	1	62	1	168	70,0	2	225	154	24,80	1	1
130	2	63	1	162	75,0	1	366	298	28,58	1	2
131	1	61	1	165	80,0	2	210	105	29,38	1	2
132	2	57	1	157	74,0	1	236	91	30,02	1	3
133	2	60	2	150	65,0	1	188	267	33,89	1	2
134	1	60	1	173	83,0	1	200	79	27,73	1	2
135	2	69	1	157	60,0	1	314	131	24,34	1	1
136	1	58	1	158	61,0	1	251	196	24,44	1	1
137	1	57	1	162	69,0	1	244	72	26,29	1	2
138	2	31	1	166	58,0	1	164	65	21,05	1	1
139	2	73	1	148	68,5	1	233	47	31,27	1	3
140	1	57	2	169	79,0	1	137	171	32,66	1	2
141	1	54	2	163	67,0	1	258	150	30,22	1	2
142	1	21	2	175	89,0	1	216	110	29,06	1	2
143	1	48	2	155	51,0	1	256	105	26,23	1	1
144	2	50	2	147	63,0	1	232	66	34,15	1	2
145	2	58	1	164	71,5	1	259	132	26,58	1	2
146	2	47	2	158	60,0	1	193	42	29,03	1	1
147	1	43	2	178	87,0	1	257	139	32,46	1	2
148	2	47	1	155	63,0	1	245	53	26,22	1	2
149	1	42	2	172	70,0	1	204	77	28,66	1	1
150	2	52	1	160	58,0	1	151	37	22,66	1	1
151	2	38	1	169	69,0	1	167	61	24,16	1	1
152	2	41	1	156	59,0	1	166	55	24,24	1	1
153	2	41	1	161	60,0	1	188	58	23,15	1	1
154	1	35	2	169	77,0	1	260	180	26,96	1	2
155	1	35	1	177	70,0	1	195	66	22,34	1	1
156	1	40	2	175	78,0	1	181	191	30,47	1	2
157	2	39	2	156	67,0	1	317	111	32,53	1	2
158	1	34	1	162	81,0	1	158	184	30,86	1	3
159	2	42	1	144	54,0	1	211	53	26,04	1	2
160	2	33	1	165	69,0	1	188	96	25,34	1	2
161	2	32	1	154	53,0	1	271	56	22,35	1	1
162	2	30	1	157	68,0	1	164	65	27,59	1	2
163	2	29	2	165	65,0	1	221	85	23,88	1	1
164	2	68	1	160	72,0	1	229	180	28,13	2	2
165	1	31	2	171	90,0	1	224	442	30,78	1	3
166	1	30	1	167	76,0	1	302	138	27,25	1	2
167	2	27	1	158	51,0	1	172	60	20,43	1	1



168	2	25	1	161	64,0	1	145	37	24,69	1	1
169	1	22	1	179	75,0	1	169	77	23,41	1	1
170	2	25	1	163	66,0	1	244	133	24,84	1	1
171	2	24	1	149	74,0	1	137	186	33,33	1	3
172	1	44	2	171	65,0	1	219	79	27,23	2	1
173	1	23	1	181	90,0	1	226	62	27,47	1	2
174	2	24	1	154	63,5	1	146	75	26,78	1	2
175	1	54	1	150	60,0	1	174	40	26,67	1	2
176	1	48	2	168	93,0	1	263	139	37,95	1	3
177	2	21	1	173	67,0	1	137	72	22,39	1	1
178	2	54	1	152	63,0	1	244	70	27,27	1	2
179	2	52	1	153	84,0	1	303	411	35,88	2	3
180	2	45	1	163	73,0	1	305	73	27,48	1	2
181	1	39	1	169	92,0	1	154	208	32,21	1	3
182	2	40	1	158	74,0	1	206	60	29,64	1	2
183	1	29	2	173	70,0	1	197	63	23,39	1	1
184	2	27	1	160	86,5	1	173	77	33,79	1	3
185	2	26	1	156	55,0	1	159	48	22,60	1	1
186	1	58	2	165	78,0	1	237	187	33,65	1	2
187	1	54	1	166	77,0	1	227	104	27,94	1	2
188	1	47	1	172	72,0	1	244	112	24,34	1	1
189	1	39	2	154	64,0	1	311	129	31,99	1	2
190	2	39	1	163	60,0	1	200	50	22,58	1	1
191	2	26	1	156	46,0	1	133	45	18,90	1	1
192	2	31	1	154	50,0	1	252	56	21,08	1	1
193	2	65	1	160	71,0	1	193	58	27,73	1	2
194	2	66	1	154	65,0	1	268	120	27,41	1	2
195	1	60	2	161	60,0	1	247	79	28,15	1	1
196	1	25	1	189	82,0	1	160	71	22,96	1	1
197	1	50	2	164	55,5	1	220	96	25,64	1	1
198	1	50	1	162	74,0	1	190	77	28,20	1	2
199	2	29	1	148	61,0	1	164	57	27,85	2	2
200	2	35	1	164	56,5	1	226	48	21,01	1	1
Total	N	200	200	200	200	200	199	199	199	200	200



MEDREP.sav

Resúmenes de casos					
	c	a	b	grupo	edad
1	185	148	132	1	64
2	160	128	120	1	58
3	190	144	118	1	72
4	192	158	115	1	74
5	218	152	148	1	75
6	200	135	134	1	68
7	210	150	128	1	69
8	225	165	140	1	75
9	190	155	138	1	67
10	200	132	136	1	69
11	210	140	135	2	70
12	180	165	144	2	63
13	180	140	115	2	63
14	176	135	126	2	60
15	200	130	140	2	70
16	188	122	130	2	69
17	200	150	132	2	70
18	178	145	135	2	60
19	180	140	130	2	68
20	150	135	130	2	60
Total	N	20	20	20	20



APAREADO.sav

Resúmenes de casos^a

	urico1	urico2	urico3	dieta
1	6,5	5,9	7,2	1
2	5,0	5,5	6,2	1
3	6,4	6,4	7,0	1
4	7,2	7,8	6,4	1
5	9,2	7,1	6,5	1
6	8,0	7,0	7,0	1
7	7,5	5,0	5,1	1
8	9,2	6,5	5,1	1
9	6,9	6,9	5,5	1
10	7,8	6,0	5,9	1
11	7,0	4,0	4,2	2
12	6,5	6,3	6,2	2
13	7,0	6,5	6,0	2
14	9,0	6,0	5,0	2
15	9,0	5,0	4,0	2
16	8,4	6,0	5,0	2
17	9,3	7,0	5,0	2
18	6,0	6,0	5,0	2
19	7,0	5,0	5,0	2
20	5,0	5,0	5,0	2
Total	N	20	20	20



PAS.sav

	pas	quetelet	edad	tabaco	neduc
1	135	28,8	45	0	4
2	122	32,5	41	0	3
3	130	31,0	49	0	3
4	148	37,7	52	0	1
5	146	29,8	54	1	1
6	129	27,9	47	1	1
7	162	36,7	60	1	2
8	160	36,1	48	1	2
9	144	23,7	44	1	2
10	180	46,4	64	1	3
11	166	38,8	59	1	2
12	138	40,3	51	1	1
13	152	41,1	64	0	3
14	138	36,7	56	0	3
15	140	35,6	54	1	2
16	134	30,0	50	1	1
17	145	33,6	49	1	3
18	142	30,2	46	1	4
19	135	31,7	57	0	1
20	142	34,0	56	0	3
21	150	36,3	56	1	1
22	144	37,5	58	0	4
23	137	32,9	53	0	1
24	132	32,1	50	0	3
25	149	33,0	54	1	1
26	132	30,2	48	1	1
27	120	27,9	43	0	2
28	126	29,6	43	1	2
29	161	38,0	63	0	2
30	170	41,3	63	1	2
31	152	39,6	62	0	2
32	164	40,1	65	0	3
Total N	32	32	32	32	32

