

TEMA 10

COMPARAR MEDIAS

Los procedimientos incluidos en el menú **Comparar medias** permiten el cálculo de medias y otros estadísticos, así como la comparación de medias para diferentes tipos de variables, mediante las pruebas *t* de Student y ANOVA.

MEDIAS

Mediante el procedimiento **Medias** pueden calcularse medias y otros estadísticos para variables dependientes dentro de las categorías de una o varias variables independientes. Además puede realizarse el análisis de la varianza para variables cualitativas de varias categorías. Por lo tanto las variables dependientes o variables resultado serán cuantitativas y las variables independientes, categóricas.

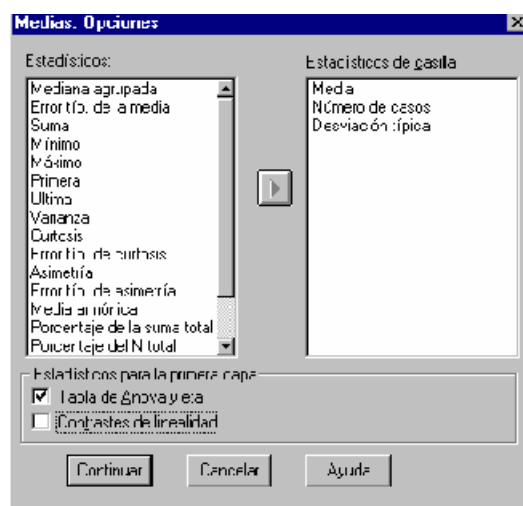
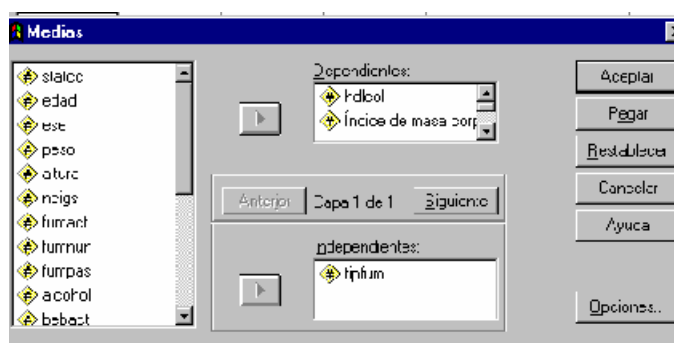
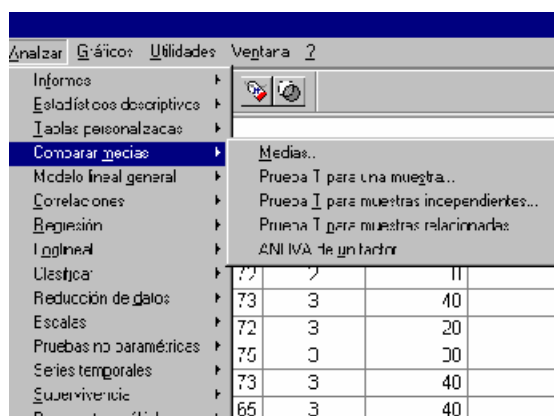
Pueden seleccionarse varias variables dependientes, y también varias independientes de forma que los resultados irán mostrando los valores de las variables resultado para cada variable independiente por separado.

Se puede realizar una estratificación por más de una variable independiente, seleccionando éstas e incluyéndolas en diferentes capas. De este modo el análisis se realizará estratificando por esas variables en el orden en que las hayamos puesto en las capas.

OPCIONES

El subcuadro de diálogo Medias: Opciones nos muestra la información adicional que podemos pedir en este procedimiento; se trata de una serie de estadísticos cuyo orden de aparición será el mismo en el que quedan en el cuadro Estadísticos de casilla.

Además puede obtenerse una tabla de análisis de la varianza si la variable independiente tiene más de dos categorías, y parámetros de asociación (η y η^2). Otra opción es la de presentar también parámetros de linealidad como son R y R^2 .



PRUEBAS T DE STUDENT

La prueba *t de Student* es utilizada para la estimación de medias y proporciones en variables cuantitativas y para la comparación de medias y proporciones en distintas poblaciones. El procedimiento **Comparar medias** ofrece varias opciones en función de la procedencia de los datos que queremos comparar.

PRUEBA T PARA MUESTRAS INDEPENDIENTES

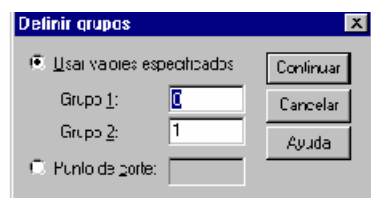
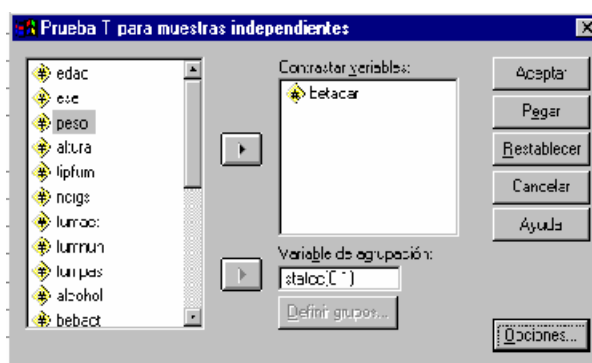
La *t de Student* para datos independientes se usa para comparar medias de variables en dos grupos de casos que son independientes entre sí, de forma que los sujetos de cada grupo deberían haber sido elegidos de forma aleatoria. En caso contrario elegiríamos la opción de *t de Student* para datos apareados o dependientes.

Al hacer la comparación de medias hay que tener en cuenta otro factor, que es la igualdad

o no de las varianzas en los dos grupos, pues los resultados no serán iguales en ambos casos. Si asumimos igualdad de varianzas podremos calcular un estimador único y más estable de la varianza poblacional a partir de las varianzas muestrales. Si las varianzas no son iguales no podremos hacer esto. El procedimiento Comparar medias realiza automáticamente un test de hipótesis para la igualdad de las varianzas (test de Levene).

En el cuadro de diálogo podemos seleccionar las variables que queremos analizar, y en función de qué vamos a comparar los grupos, mediante la variable de agrupación, que puede ser cualitativa o cuantitativa. Una vez elegida esta variable hemos de definir los grupos de comparación pulsando el botón *Definir grupos*.

Aquí tenemos dos opciones. Pueden elegirse dos valores especificados para comparar las medias (en general, dos categorías en variables cualitativas), pero también puede establecerse un punto de corte para variables cuantitativas que divida esa variable en dos grupos, por encima y por debajo de ese valor.



Estadísticos de grupo

	STATCC	N	Media	Desviación tp.	Error tp. de la media
BETACAR	0	700	,3726	,4011	,015
	1	635	,2849	,2862	,011

Prueba de muestras independientes

		Pruebas de Levene para la igualdad de varianzas		Pruebas T para la igualdad de medias						
		F	Sig.	t		Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% intervalo de confianza para la diferencia	
				l	gl				inferior	superior
BETACAR	Se han asumido varianzas iguales	13,945	,000	4,656	1333	,000	,038	,019	,051	,1246
	No se han asumido varianzas iguales			4,744	1224,394	,000	,030	,010	,051	,1209

Los resultados que SPSS muestra vienen en dos tablas: la primera, *Estadísticos de grupo*, incluye el tamaño muestral, la media, la desviación típica y el error estándar de la media. En la segunda se efectúa la prueba t.

En primer lugar se muestra el resultado para la prueba de Levene de igualdad de varianzas. Si éste es estadísticamente significativo, asumiremos que las varianzas no son iguales y entonces tomaremos los datos de la comparación de medias de la fila inferior (*No se han asumido varianzas iguales*). Si el test de Levene no es significativo, asumiremos igualdad de varianzas y nos quedaremos con los datos de la fila superior.

Los datos que el procedimiento nos muestra son:

- *t*: estadístico utilizado para el contraste de hipótesis.
- *gl*: número de grados de libertad del estadístico t. Cuando se asumen varianzas iguales, *gl* es igual a $(n_1 + n_2 - 2)$, pero cuando las varianzas no son iguales hay que calcularlo de otras formas. La más usada es la aproximación de Satterthwaite, que SPSS calcula. Este valor será menor que el anterior porque se pierde precisión con la desigualdad de las varianzas.
- *Sig. (bilateral)*: valor p de significación estadística obtenido para el contraste de hipótesis.
- *Diferencia de medias*: estimador puntual de la diferencia de medias.
- *Error típ. de la diferencia*: valor del error estándar de la diferencia de medias (obtenido a partir de los errores estándar de la tabla superior).
- *Intervalo de confianza para la media*: límites inferior y superior del intervalo de confianza que valora la precisión de la estimación que estamos realizando para la diferencia de medias.

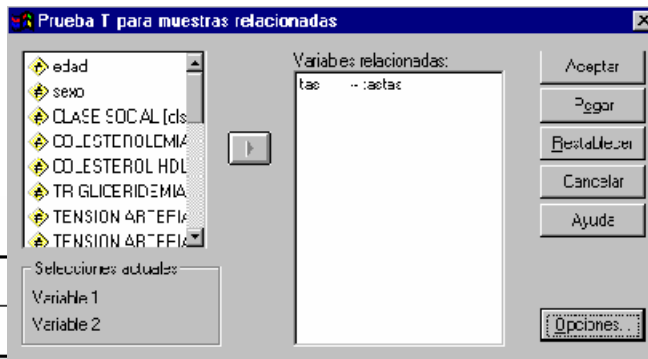
PRUEBA T PARA MUESTRAS RELACIONADAS

En este caso las muestras no son independientes, sino que están relacionadas. Lo más habitual es que se trate de variables mediadas en los mismos casos en dos momentos temporales diferentes, o que sean casos apareados en función de otras variables.

En el cuadro de diálogo hay que seleccionar el par de variables cuyas medias vamos a comparar, y pasarlo al campo de la derecha *Variables relacionadas*.

Estadísticos de muestras relacionadas

Par		Media	N	Desviación tp.
Par 1	ANTESITO	140,11	73	19,54
	POSTITO	131,03	73	23,05



Prueba de muestras relacionadas

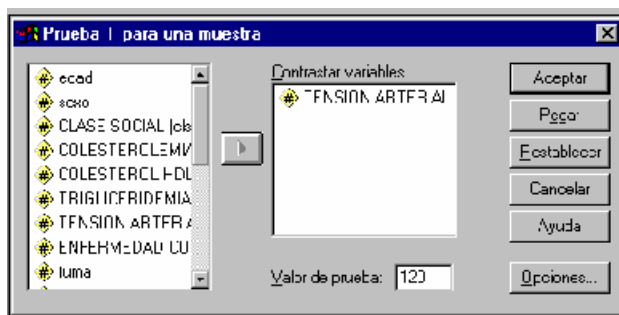
		Diferencias relacionadas				t	gl	Sig. (bilateral)	
		Media	Desviación tp.	Error típ. de la media	95% Intervalo de confianza para la diferencia				
					Inferior				Superior
Par 1	ANTESITO POSTITO	9,09	11,41	1,33	6,37	11,01	3,637	63	,000

Igual que en el caso anterior primero aparece una tabla con los estadísticos de las muestras y después la prueba *t de Student*. En esta ocasión se tratan los datos como parejas de datos, y se halla la media de las diferencias entre cada pareja, así como su desviación típica para calcular el error estándar de la media. El resto de los valores son equivalentes a los del caso anterior.

PRUEBA T PARA UNA MUESTRA

La prueba *t* para una muestra efectúa un contraste de hipótesis para comprobar si la media de una variable difiere de forma significativa de un valor que nosotros mismos seleccionamos.

En este caso hay que seleccionar la variable elegida y el valor que queremos contrastar. Al pulsar Aceptar se efectuará el contraste de hipótesis.



Estadísticos para una muestra

	N	Media	Desviación tp.	Error típ. de la media
TENSION ARTERIAL SISTOLICA	73	140,11	19,54	2,33

Prueba para una muestra

	Valor de prueba = 120					
	t	gl	Sig. (bilateral)	Diferencia de medias	95% Intervalo de confianza para la diferencia	
					Inferior	Superior
TENSION ARTERIAL SISTOLICA	3,614	63	,000	20,11	15,46	24,77

ANOVA DE UN FACTOR

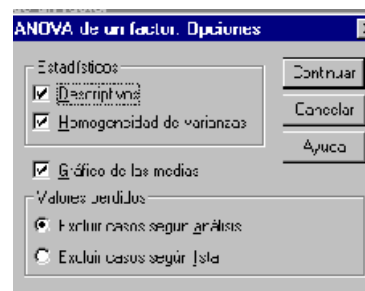
Este procedimiento es el equivalente a la *t de Student* para más de dos muestras, es decir, realiza una comparación de medias en variables cuantitativas para más de dos grupos de casos. Se usa para contrastar la hipótesis de que varias medias son iguales.

En el correspondiente cuadro de diálogo seleccionaremos la variable cuantitativa que vamos a analizar así como la variable cualitativa de más de dos categorías (*Factor*). Además tenemos varios botones que nos ofrecen distintas posibilidades.



OPCIONES

Podemos pedir los estadísticos descriptivos para cada grupo de casos, así como la prueba de Levene para la homogeneidad de las varianzas. También puede presentarse un gráfico con las medias para cada grupo.



CONTRASTES

Además pueden efectuarse contrastes individuales entre varios de los valores que puede adquirir la variable factor. Esto se hace asignando coeficientes a cada uno de los valores posibles. Pueden contrastarse varios grupos con uno solo, o varios con varios, en función de los valores que se les dé a los coeficientes. La suma de estos coeficientes debe ser 0.



Descriptivos

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
					Límite inferior	Límite superior		
Nunca fumador	212	5,423	1,017	,089	5,28E	5,560	2,97	8,67
Ex. Fumador	319	5,634	1,149	,064	5,507	5,760	2,73	9,27
Fumador actual	507	5,463	1,177	,061	5,364	5,561	2,77	9,44
Total	1033	5,507	1,110	,035	5,44C	5,575	2,27	9,49

ANOVA

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	7,603	2	3,802	3,09E	,046
Intra-grupos	12E4,543	1030	1,228		
Total	1272,146	1032			

Coefficientes de los contrastes

Contraste	TIPFUM		
	Nunca fumador	Exfumador	Fumador actual
1	-.5	0	-.5

Prueba de homogeneidad de varianzas

TOTCOL

Estadístico de Levene	gl1	gl2	Sig.
1,514	2	1000	,221

Pruebas para los contrastes

		Contraste	Valor del contraste	Error típico	t	gl	Sig. (bilateral)
TOTCOL	Asumiendo igualdad de varianzas	1	,1909	,077	2,483	1000	,013
	No asumiendo	1	,1909	,077	2,470	579,103	,014