

Comparación de medias de grupos independientes y grupos relacionados

[10.1] ¿Cómo estudiar este tema?

[10.2] Grupos independientes

[10.3] Comparación de las medias de dos grupos independientes

[10.4] Comparación de las medias de más de dos grupos independientes

[10.5] Grupos relacionados

[10.6] Comparación paramétrica de dos grupos relacionados

[10.7] Comparación no paramétrica de dos grupos relacionados

10

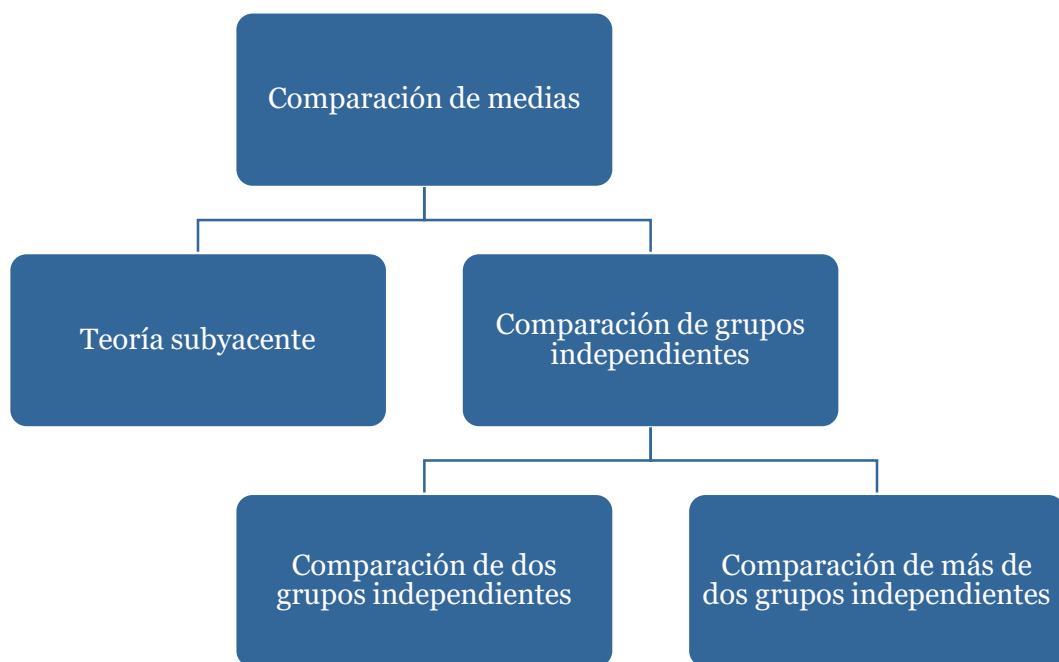
T E M A

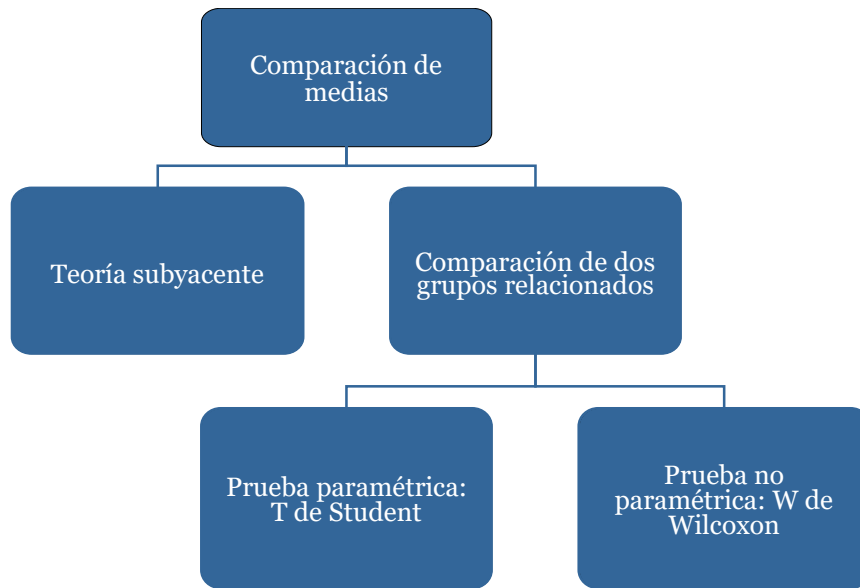
Ideas clave

10.1. ¿Cómo estudiar este tema?

En este tema veremos la comparación de las medias de grupos independientes. Por un lado, presentaremos la teoría subyacente de esta comparación de grupos para después pasar a ver cómo utilizar el complemento de Excel EZAnalyze para comparar las medias de diferentes grupos en el caso paramétrico y para el caso no paramétrico usaremos SPSS.

Para estudiar este tema **deberás comprender las Ideas clave** expuestas en este documento y que han sido elaboradas por el profesor de la asignatura. Estas ideas se van a complementar con lecturas y otros documentos para que puedas ampliar los conocimientos sobre el mismo.





10.2. Grupos independientes

Tipos de variables

Existen diferentes distinciones de tipos de variables. Atendiendo al tipo de valores que pueden tomar se pueden distinguir dos grandes bloques de variables:

1. Variables cuantitativas o numéricas: si sus valores son numéricos. Este tipo de variables se divide a su vez en otros dos tipos:

» Continuas: si puede tomar cualquier valor (en otras palabras si puede contener decimales).

Ejemplos:

- Peso (81,7).
- Altura (1.73).
- Puntuación en un examen (8.75).

» Discretas: si solo pueden tomar valores enteros (es decir, sin decimales).

Ejemplos:

- Número de hijos.
- Edad.
- Veces que sale el 6 al tirar un dado.

2. Variables cualitativas: si sus valores no se puede asociar naturalmente a un número. Dentro de este tipo de variables existen dos tipos:

» Nominales: si sus valores no se pueden ordenar.

Ejemplos:

- Sexo.
- Nacionalidad.
- Ideas políticas.

» Ordinales: si se pueden ordenar pero no hay la misma distancia entre las diferentes categorías.

Ejemplos:

- Intensidad de dolor (leve, moderado, alto).
- Notas (suspense, aprobado, notable, sobresaliente).
- Grado de satisfacción (bajo, normal, alto).

Si en lugar de atender al tipo de valores que puede tomar, tenemos en cuenta el papel que juega en la **investigación** se tiene la siguiente clasificación:

» Variables **independientes**: es una variable que actúa como estímulo para provocar cambios. Es la «causa».

» Variables **dependientes**: es la variable donde se observan los cambios, es el objeto de interés de la investigación.

» Variables **intervenientes**: son variables que afectan a la dependiente, pero que no producen cambios que interesen para la investigación. También existen dentro de este grupo dos categorías:

- Variables de control: si se incluyen en el estudio.
- Variables extrañas: si no se incluyen en el estudio.

Tipos de grupos

Atendiendo a las características de los sujetos que los componen, los grupos se pueden diferenciar en:

- » **Grupos independientes:** si los sujetos se distribuyen de manera aleatoria en los grupos. El problema de estos grupos es que el no tener en cuenta un determinado factor puede afectar a los resultados obtenidos en el estudio.
- » **Grupos relacionados:** si se mide previamente un factor que se considera que puede afectar a los resultados y se tienen en cuenta dichos resultados a la hora de realizar los grupos para hacer grupos homogéneos.

Pruebas

Las pruebas que vamos a ver en este tema son variadas y dependerán de una serie de condicionantes.

1. Si tenemos datos de **dos grupos independientes**, como por ejemplo, rubios y morenos o chicos y chicas.
 - » Si la muestra cuenta con al menos 30 individuos, se cumplen los supuestos de normalidad, igualdad de varianzas e independencia de observaciones y además las variables a observar son cuantitativas:

- **T de Student**

Ejemplo. Se dispone de dos grupos de alumnos, chicas (40) y chicos (35) y se desea saber si las medias de notas son similares para ambos grupos.

- » Si no se verifica alguna de las condiciones:

- **U de Mann-Whitney**

Ejemplo. Se desea saber si la intensidad de dolor medias es igual en rubios (40) y morenos (20) dentro de nuestra clase.

2. Si tenemos datos de **más de dos grupos independientes**, como por ejemplo, rubios, morenos, pelirrojos y castaños.

- » Si la muestra cuenta con al menos 30 individuos, se cumplen los supuestos de normalidad, igualdad de varianzas e independencia de observaciones y además las variables a observar son cuantitativas:

- **ANOVA: F de Snedecor**

Ejemplo. Se desea saber si las notas medias de rubios (35), morenos (37) y castaños (30) son iguales dentro de nuestra clase.

- » Si no se verifica alguna de las condiciones:

- **H DE Kruskal-Wallis**

Ejemplo. Se desea conocer si el número de hijos (0-1, 2-3, >3) es igual en profesores con ojos azules (20), con ojos marrones (30) y con ojos verdes (50).

10.3. Comparación de las medias de dos grupos independientes

A continuación pasaremos a enumerar las diferentes pruebas que utilizaremos para comparar las medias de dos grupos de sujetos independientes.

Comparación de dos grupos independientes: T de Student (paramétrica)

La prueba T de Student es un estadístico paramétrico que se utiliza cuando se desea comparar las medias de dos grupos independientes de sujetos que deben contar con las siguientes características:

- » Tamaño grande, es decir, por lo menos 30 sujetos.
- » Que sean de naturaleza cuantitativa.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Y la hipótesis alternativa:

Sí existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Los resultados de la prueba T de Student van a venir representados por dos tablas distintas y un gráfico.

Una primera tabla en la que se nos muestra la media y la desviación típica de cada grupo por separado.

- » Media de cada grupo.
- » Número de casos por cada grupo.

Una segunda tabla en la que nos muestra:

- » La diferencia entre las medias.
- » El valor del estadístico T.
- » Eta squared (que no nos interesa).
- » El P-valor que indica si hay o no significatividad.

Un gráfico con las medias de cada grupo: en el que se observa la diferencia entre las medias de ambos grupos.

Ejemplo de comparación de dos grupos independientes (caso paramétrico)

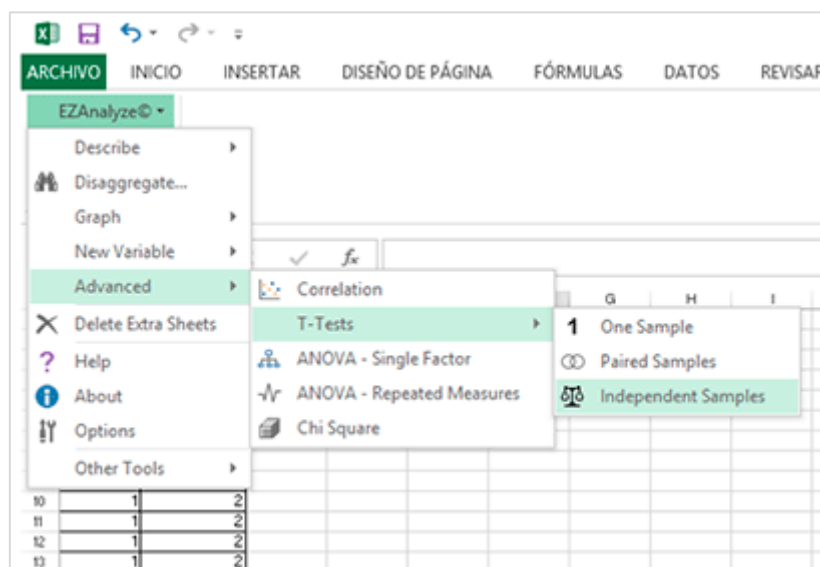
Tenemos los datos, expresados en miles, de las visitas a la página web de dos periódicos (30 datos de cada periódico) en la hoja de Excel que se observa en la siguiente figura.

Podemos observar que son dos grupos independientes claramente. Ahora veamos cómo medir si existen diferencias significativas y, si las hay, si estas son significativas o debidas al azar.

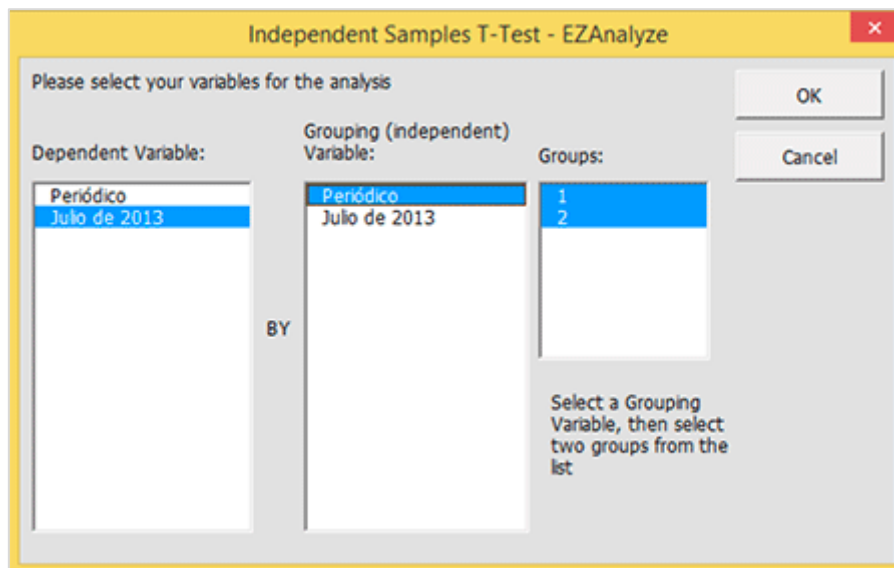
	A	B	C
1	Periódico	Julio de 2013	
2	1	1	
3	1	1	
4	1	1	
5	1	1	
6	1	1	
7	3	1	
8	3	1	
9	3	1	
10	3	1	
11	3	1	
12	3	1	
13	3	1	
14	3	1	
15	3	1	
16	3	1	
17	3	1	
18	3	1	
19	3	1	
20	3	1	
21	3	1	
22	3	1	
23	2	1	
24	1	2	
25	1	2	
26	1	2	
27	1	2	
28	1	2	
29	1	2	
30	1	2	
31	1	2	
32	1	2	
33	1	2	
34	1	2	

Para ello seguimos la siguiente ruta:

Complementos > EZAnalyze > Advanced > T-Test > Independent Samples. Podemos ver la ruta elegida en la siguiente figura.



A través de estas opciones obtenemos el siguiente cuadro, en el que elegimos como variable dependiente julio del 2013 y como independiente el periódico.

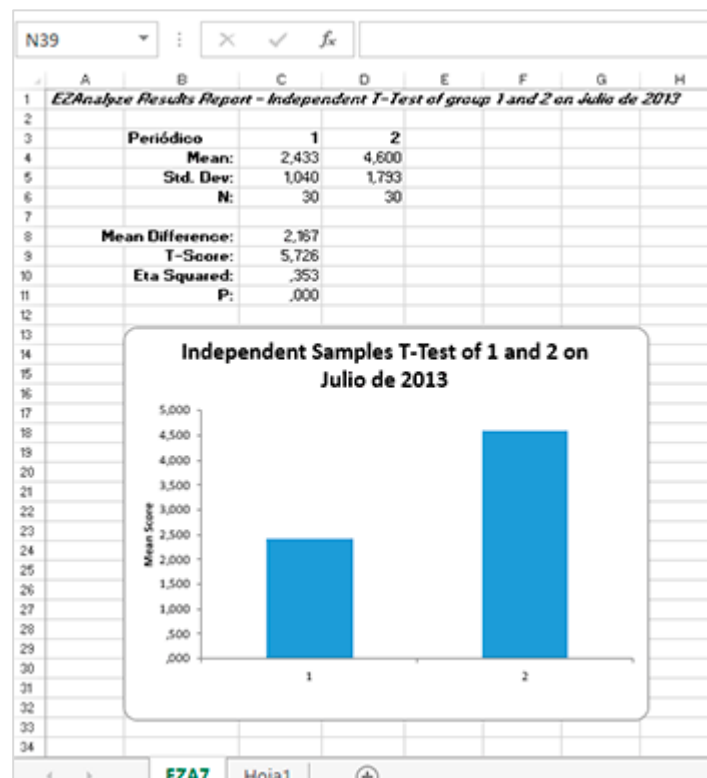


Pulsamos OK y obtenemos la siguiente información que vemos en la figura.

En esta hoja podemos ver los siguientes datos:

Una primera tabla en la que se nos muestra la media y la desviación típica de cada grupo por separado.

- » Media del periódico 1: 2.433
- » Media del periódico 2: 4.600.
- » Número de casos por cada caso: 30.



Una segunda tabla en la que nos muestra:

- » La diferencia entre las medias: 2.167 significa que la media de ventas es superior en el periódico 2 en 2167 ejemplares mayor que la media de ventas del periódico 1.
- » El valor del estadístico T: 5.726.
- » Eta squared (que no nos interesa).
- » El P-valor que indica si hay o no significatividad: $P=0$, por lo tanto la diferencia entre las medias es significativa.

Un gráfico con las medias de cada grupo en el que se observa la diferencia entre ambas.

Comparación de dos grupos independientes: U de Mann-Whitney (no paramétrica)

La prueba U de Mann-Whitney es un estadístico no paramétrico que se utiliza cuando se desea comparar las medias de dos grupos independientes de sujetos cuando al menos una de las siguientes características se cumple:

- » Tamaño pequeño, es decir, menos de 30 sujetos.
- » Cuando la variable dependiente es ordinal.
- » Cuando la variable dependiente no se distribuye de forma normal.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Y la **hipótesis alternativa**:

Sí existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Los resultados de la prueba U de Mann-Whitney van a venir dados por un resumen en el que nos dirá si debemos o no rechazar la hipótesis nula, es decir, que nos dirá si debemos o no admitir que las diferencias entre los dos grupos son significativas.

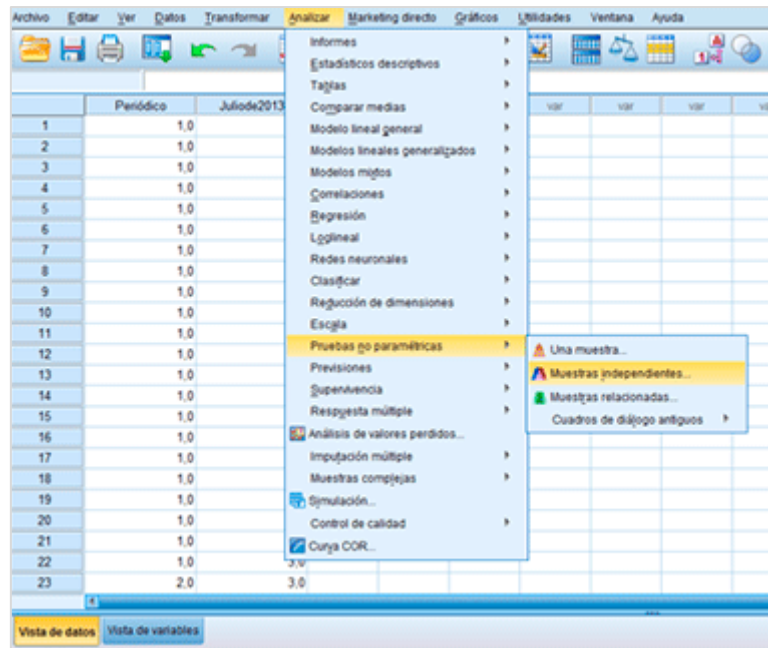
Ejemplo de comparación de dos grupos independientes (caso no paramétrico)

Tenemos los datos (23 para el periódico 1 y 21 para el periódico 2), expresados en miles, de las ventas de dos periódicos en la hoja de PSPP que se observa en la siguiente figura:

	Antes	Después	var
1	1,0	2,0	
2	1,0	2,0	
3	2,0	4,0	
4	3,0	7,0	
5	2,0	4,0	
6	3,0	7,0	
7	2,0	5,0	
8	3,0	7,0	
9	2,0	5,0	
10	3,0	8,0	
11	2,0	4,0	
12	3,0	8,0	
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			

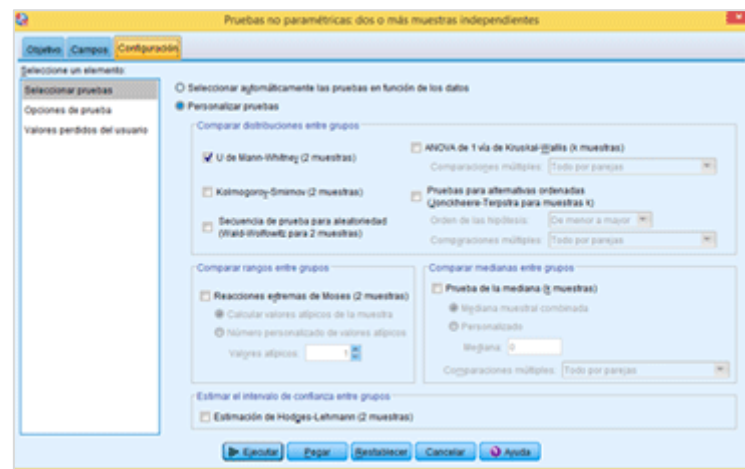
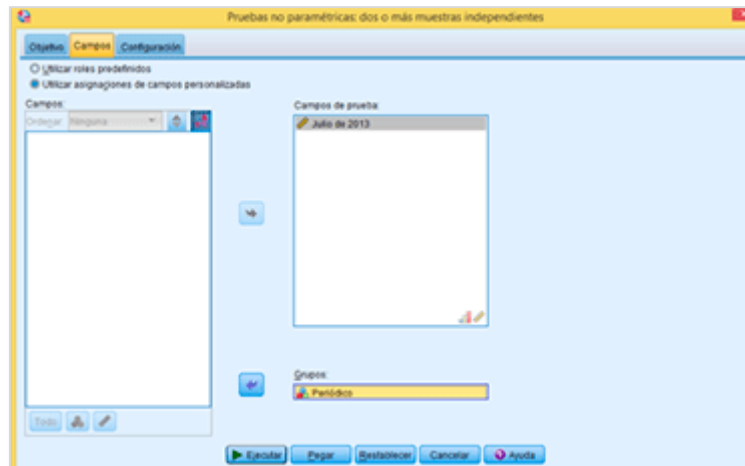
Entonces vemos que son dos grupos independientes claramente pero tenemos pocas mediciones por lo que estamos en el caso no paramétrico. Ahora veamos cómo medir si existen diferencias significativas y si las hay si estas son significativas o debidas al azar.

Para ello seguimos la siguiente ruta: Analizar > Pruebas no paramétricas > Muestras independientes. En la siguiente figura vemos la ruta elegida.

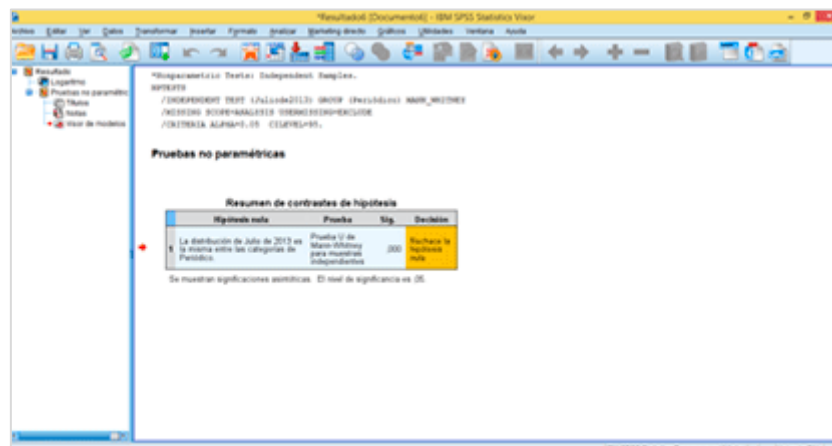


Nos aparecerá un nuevo cuadro con tres pestañas, las cuales se muestran en las siguientes figuras.





Pulsamos ejecutar y obtenemos lo que se observa en la siguiente figura.



Nos está diciendo que rechazemos la hipótesis nula y, por lo tanto, las diferencias con el nuevo editor jefe son significativas. Además, sin más que echar una ojeada se observa como las ventas son mayores con el nuevo editor por lo que el cambio ha sido positivo.

10.4. Comparación de las medias de más de dos grupos independientes

Comparación de más de dos grupos independientes: ANOVA F de Snedecor (paramétrica)

La prueba F de Snedecor es un estadístico paramétrico que se utiliza cuando se desea comparar las medias de más de dos grupos independientes de sujetos que deben contar con las siguientes características:

- » Tamaño grande, es decir, por lo menos 30 sujetos.
- » Que sean de naturaleza cuantitativa.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Y la **hipótesis alternativa**:

Sí existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Los resultados de la prueba F de Snedecor van a venir representados por cuatro tablas distintas y un gráfico:

Una primera tabla en la que se nos muestra la media total de todo el conjunto.

Una segunda tabla en la que nos muestra por cada grupo.

- » Media.
- » Desviación típica.
- » Número de valores.

Una tercera tabla (la tabla ANOVA) en la que nos muestra:

- » Toda la información ANOVA.
- » Eta squared (que no nos interesa).
- » El P-valor que indica si hay o no significatividad.

Un aviso en el que nos dice si el estudio de ANOVA ha determinado que al menos dos de las medias presentan diferencias significativas.

Una cuarta tabla en la que nos muestra las diferencias de las medias con sus correspondientes P valores.

Un gráfico con las medias de cada grupo: en el que se observa la diferencia entre las medias de ambos grupos.

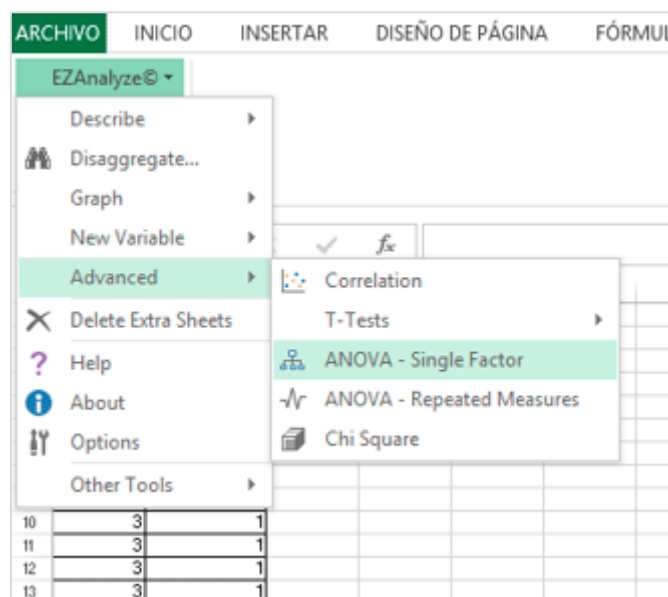
Ejemplo de comparación de tres grupos independientes (caso paramétrico)

Tenemos los datos, expresados en miles, de las ventas de 3 periódicos durante el mes de julio de 2013 (disponemos de al menos 30 observaciones por cada uno de ellos) en la hoja de Excel que se observa en la siguiente figura:

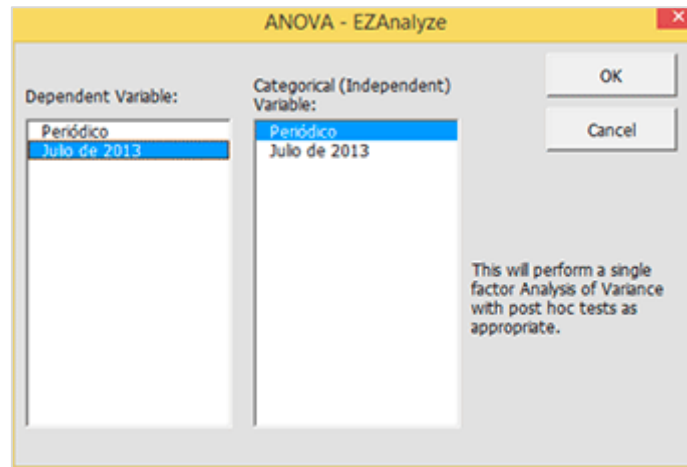
Vemos que son tres grupos independientes claramente. Ahora veamos cómo medir si existen diferencias significativas y, si las hay, si estas son significativas o debidas al azar.

	A	B
1	Periódico	Julio de 2013
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	3	1
8	3	1
9	3	1
10	3	1
11	3	1
12	3	1
13	3	1
14	3	1
15	3	1
16	3	1
17	3	1
18	3	1
19	3	1
20	3	1
21	3	1
22	3	1
23	2	1
24	1	2
25	1	2
26	1	2
27	1	2
28	1	2
29	1	2
30	1	2
31	1	2
32	1	2
33	1	2
34	1	2

Para ello seguimos la siguiente ruta: Complementos > EZAnalyze > Advanced > ANOVA-Single Factor.

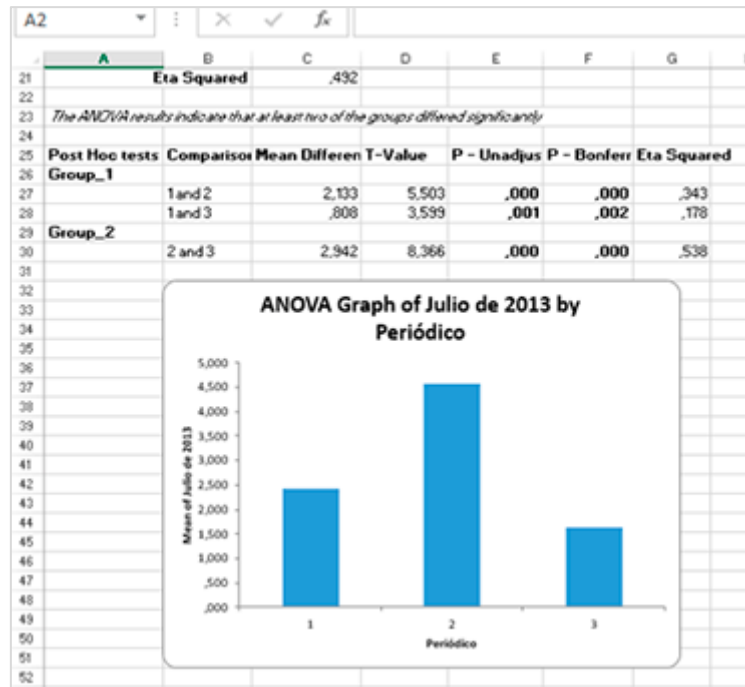


Obtenemos el siguiente cuadro, en el que elegimos como dependiente las ventas en julio de 2013 y como independiente el periódico.



Pulsamos OK y obtenemos una hoja (que dividiremos en dos imágenes para una mejor visualización) con la siguiente información.

EZAnalyze©				
Comandos de menú				
A2				
	A	B	C	D
2				
3	0 cases were removed due to missing data			
4				
5	Grand Mean	2,848		
6	N	92		
7				
8	Periódico(group)	N	Group Mean	Std Deviation
9	1	30	2,433	1,040
10	2	30	4,567	1,851
11	3	32	1,625	,707
12				
13				
14	ANOVA Table			
15	Source of Variance	SS	DF	MS
16	Between Groups	141,636	2,000	70,818
17	Within Groups	146,233	89,000	1,643
18	Total	287,870		
19				
20		P	,000	
21	Eta Squared	,492		
22				
23	The ANOVA results indicate that at least two of the groups differed significantly			
24				
25				
EZAnalyze Hoja1				



En esta hoja observamos:

Una primera tabla en la que se nos muestra la media total de todo.

» 2.484 y 92 valores.

Una segunda tabla en la que nos muestra:

» Periódico 1

- Media: 2.433.
- Desviación típica: 1.040.
- Número de valores: 30.

» Grupo 2

- Media: 4.567.
- Desviación típica: 1.851.
- Número de valores: 30.

» Grupo 3

- Media: 1.625.
- Desviación típica: 0.707.
- Número de valores: 32.

Una tercera tabla (la tabla ANOVA) en la que nos muestra:

- » Toda la información ANOVA.
- » Eta squared (que no nos interesa): 0.492.
- » El P-valor que indica si hay o no significatividad: $P=0$, por lo tanto la diferencia entre las medias es significativa.

Un aviso en el que nos dice que el estudio de ANOVA ha determinado que al menos dos de las medias presentan diferencias significativas.

Una cuarta tabla en la que nos muestra las diferencias de las medias con sus correspondientes P valores.

- » Diferencias entre periódico 1 y 2: 2.133 y $P=0$ por lo tanto significativa.
- » Diferencias entre periódico 1 y 3: 0.808 y $P=0.001$ por lo tanto significativa.
- » Diferencias entre periódico 2 y 3: 2.942 y $P=0$ por lo tanto significativa.

Un gráfico con las medias de cada grupo: en el que se observa la diferencia entre las medias de ambos grupos en el que vemos que el periódico vende que más es el 2, el que menos el 3 y el 1 queda en medio.

Comparación de más de dos Grupos independientes: H de Kruskal Wallis (no Paramétrica)

La prueba H de Kruskal Wallis la usaremos cuando tengamos que comparar más de dos grupos o, en otras palabras, cuando nuestra variable independiente tenga más de dos categorías y al menos una de las siguientes características se cumple:

- » Tamaño pequeño, es decir, menos de 30 sujetos.
- » Cuando la variable dependiente es ordinal aunque tenga al menos 30 sujetos.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

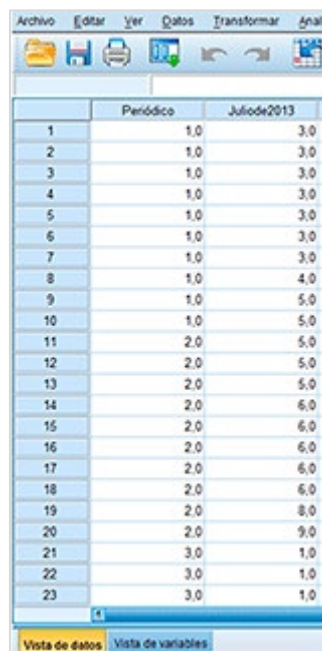
Y la hipótesis alternativa:

Sí existen diferencias significativas en el rendimiento académico en estadística (variable dependiente) de los periodistas en función del número de idiomas (variable independiente) que dominen.

Los resultados de la prueba van a venir representados por un resumen. En este resumen se nos señala si debemos o no rechazar la hipótesis nula, la cual nos dice que las diferencias entre los grupos no son significativas.

Ejemplo de comparación de tres grupos independientes (caso no paramétrico)

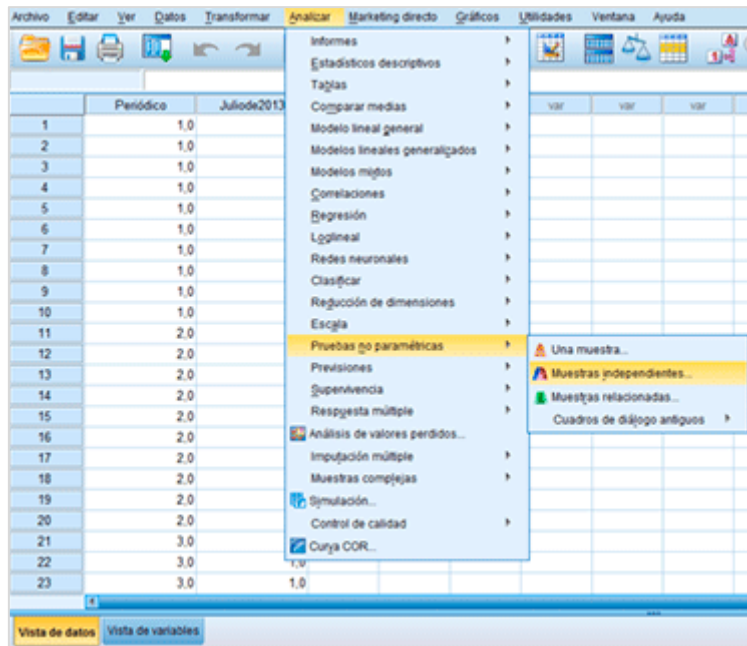
Tenemos los datos, expresados en miles, de las ventas de 3 periódicos durante el mes de julio de 2013 (disponemos de 10 observaciones por cada uno de ellos) en la hoja de PSPP que se observa en la siguiente figura:



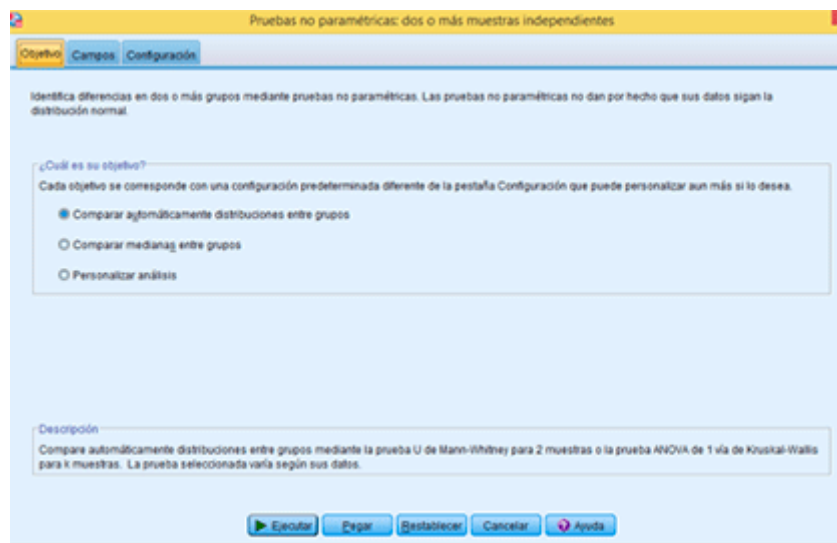
	Periódico	Julio de 2013
1	1.0	3.0
2	1.0	3.0
3	1.0	3.0
4	1.0	3.0
5	1.0	3.0
6	1.0	3.0
7	1.0	3.0
8	1.0	4.0
9	1.0	5.0
10	1.0	5.0
11	2.0	5.0
12	2.0	5.0
13	2.0	5.0
14	2.0	6.0
15	2.0	6.0
16	2.0	6.0
17	2.0	6.0
18	2.0	6.0
19	2.0	8.0
20	2.0	9.0
21	3.0	1.0
22	3.0	1.0
23	3.0	1.0

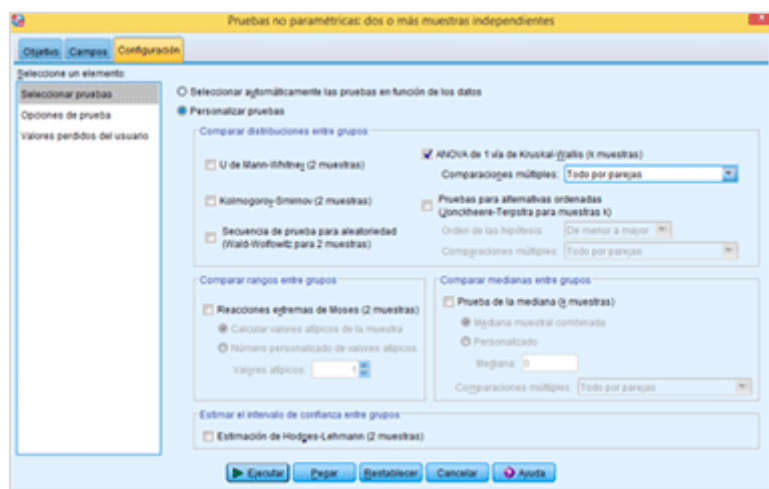
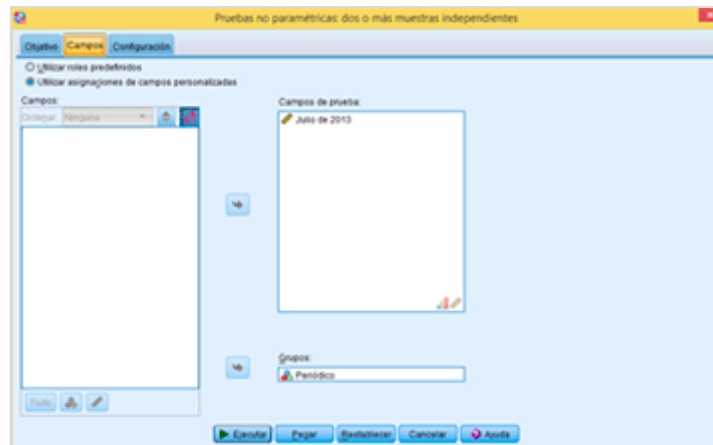
Podemos ver que son tres grupos independientes claramente. Ahora veamos cómo medir si existen diferencias significativas y, si las hay, si estas son significativas o debidas al azar.

Para ello seguimos la siguiente ruta: Analizar > Pruebas no paramétricas > Muestras independientes



Nos aparecerá un nuevo cuadro con tres pestañas, las cuales se muestran en las siguientes figuras.





Pulsamos ejecutar y obtenemos lo que se observa en la siguiente figura:



Tras todo este proceso, nos indica que rechazamos la hipótesis nula y, por lo tanto, las diferencias en las ventas de los 3 periódicos son significativas. Además, observamos dónde las ventas son mayores (periódico 2) y dónde menores (periódico 3).

10.5. Grupos relacionados

Recordamos que los grupos relacionados son aquellos en los que se mide previamente un factor que se considera puede afectar a los resultados y se tiene en cuenta dichos resultados a la hora de realizar los grupos para hacerlos homogéneos.

Pruebas

Las pruebas que vamos a ver en este tema son referentes a dos grupos relacionados, como por ejemplo, el mismo grupo antes y después de cierto tratamiento.

- » Si la muestra cuenta con al menos 30 individuos, se cumplen los supuestos de normalidad, igualdad de varianzas e independencia de observaciones y además las variables a observar son cuantitativas:

- **T de Student**

Ejemplo. Se dispone de un grupo de 40 alumnos, del cual se han tomado datos antes y después de un tratamiento corrector de lectura y se desea ver si la media del coeficiente de lectura es igual en ambos casos o por el contrario varía con el tratamiento.

- » Si no se verifica alguna de las condiciones:

- **W de Wilcoxon**

Ejemplo. Se dispone de un grupo de alumnos con una cierta enfermedad que son tratados con un medicamento y se desea saber si la intensidad de dolor es igual a un cierto nivel como probó cierto autor.

10.6. Comparación paramétrica de dos grupos relacionados

Comparación de dos grupos relacionados: T de Student (paramétrica)

La prueba T de Student es un estadístico paramétrico que se utiliza cuando se desea comparar las medias de dos grupos relacionados de sujetos, que suelen ser el pretest y el posttest, y que deben contar con las siguientes características:

- » Tamaño grande, es decir, por lo menos 30 sujetos.
- » Que sean de naturaleza cuantitativa.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas entre las puntuaciones pretest y posttest del número de faltas de ortografía.

Y la hipótesis alternativa:

Sí existen diferencias significativas entre las puntuaciones pretest y posttest del número de faltas de ortografía.

Los resultados de la prueba T de Student van a venir representados por dos tablas distintas y un gráfico:

- » En la primera vamos a ver el conjunto de estadísticos de muestras relacionadas de las variables pretest y posttest. En ella veremos los resultados medios obtenidos antes y después del tratamiento.
- » Una segunda tabla en la que nos muestra:
 - El número de pares con información.
 - La diferencia entre las medias.
 - SE o diff (que no nos interesa).
 - Eta squared (que no nos interesa).
 - El valor del estadístico T.
 - El P-valor que indica si hay o no significatividad.
- » Un gráfico con las medias de cada grupo.

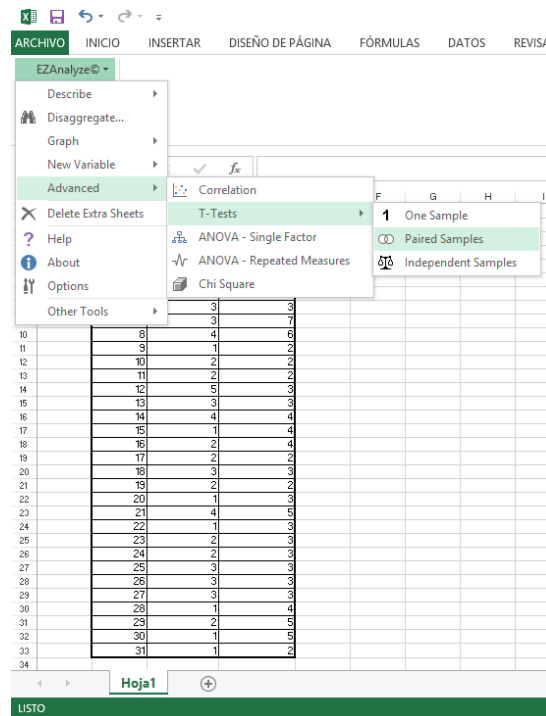
Ejemplo de comparación de dos grupos relacionados (caso paramétrico)

Tenemos los datos, expresados en miles, de las visitas a la página web de nuestro periódico en el mes de julio durante dos años consecutivos (31 datos de cada año) en la hoja de Excel que se observa en la siguiente figura:

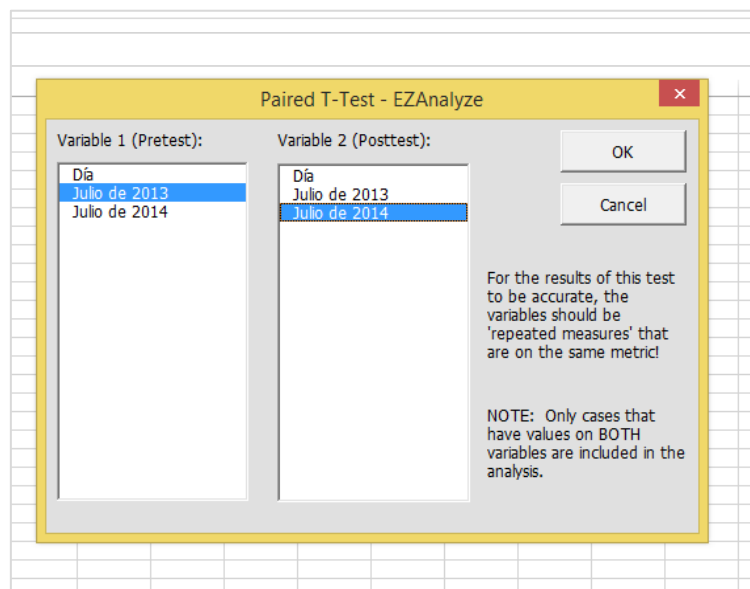
Día	Julio de 2013	Julio de 2014
1	1	4
2	2	4
3	2	4
4	2	3
5	2	3
6	3	7
7	4	6
8	1	2
9	2	2
10	2	2
11	5	3
12	3	3
13	4	4
14	1	4
15	2	4
16	2	2
17	3	3
18	2	2
19	1	3
20	4	5
21	1	3
22	2	3
23	2	3
24	3	3
25	3	3
26	3	3
27	1	4
28	2	5
29	1	5
30	1	2
31		

El día 1 de enero de 2014 se decidió poner una tira cómica interactiva en la página web y queremos ver si esto ha influido en el número de visitas. Entonces vemos que son dos grupos relacionados claramente. Ahora veamos cómo medir si existen diferencias significativas y si las hay si estas son significativas o debidas al azar.

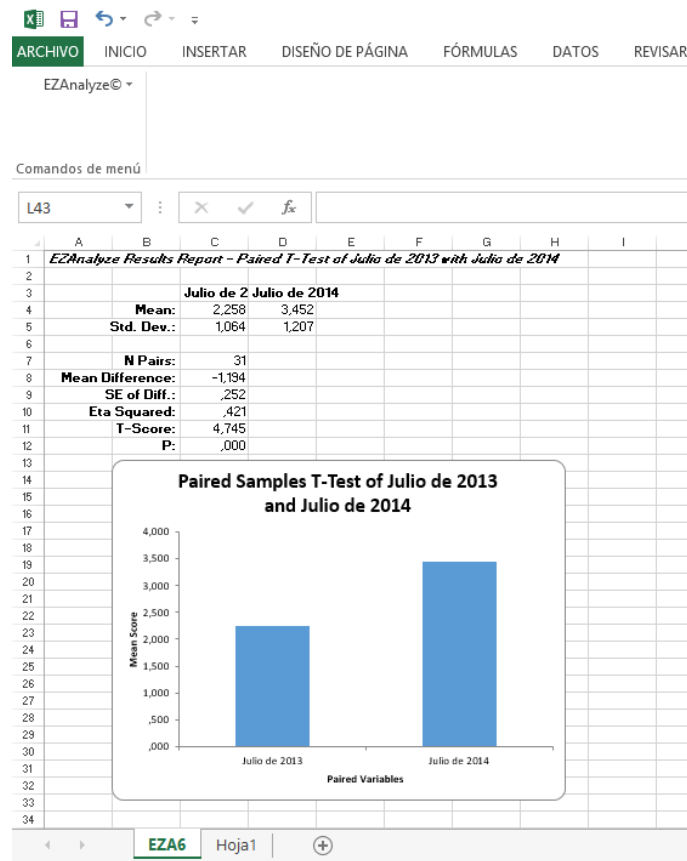
Para ello seguimos la siguiente ruta: Complementos > EZAnalyze > Advanced > T-Test > Paried Samples



Obtenemos el siguiente cuadro, en el cual elegimos como pretest julio del 2013 y como posttest julio del 2014.



Pulsamos OK y obtenemos una hoja con la siguiente información.



De esta hoja extraemos una serie de datos:

- » Una primera tabla en la que nos muestra la media y la desviación típica de cada grupo por separado.
 - Media julio 2013: 2.258.
 - Media julio 2014: 3452.
- » Una segunda tabla en la que nos muestra:
 - El número de pares con información: 31.
 - La diferencia entre las medias: -1194 significa que en 2014 la media fue mayor en 1194 puntos.
 - SE o diff (que no nos interesa).
 - Eta squared (que no nos interesa).
 - El valor del estadístico T: 4.745.
 - El P-valor que indica si hay o no significatividad: $P=0$, por lo tanto la diferencia entre las medias es significativa.

- » Un gráfico con las medias de cada grupo, en el que se observa la diferencia entre las medias de ambos.

10.7. Comparación no paramétrica de dos grupos relacionados

Comparación de dos grupos relacionados: W de Wilcoxon

La prueba W de Wilcoxon es un estadístico no paramétrico que se utiliza cuando se desea comparar los resultados obtenidos por dos grupos pequeños de sujetos y relacionados (normalmente se comparan los resultados de los mismos sujetos en un pretest y un posttest). Por lo tanto, son grupos con las siguientes características:

- » Tamaño pequeño, es decir, menos de 30 sujetos.
- » Que no sean de naturaleza cuantitativa.

La **hipótesis nula** es de la siguiente manera:

No existen diferencias significativas entre las puntuaciones pretest y posttest de la felicidad de los futbolistas al publicar una noticia sobre su ciudad natal.

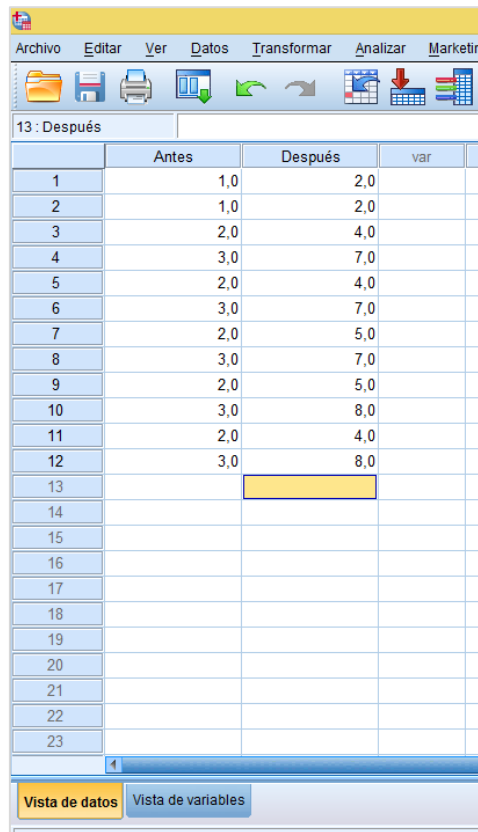
Y la **hipótesis alternativa**:

Sí existen diferencias significativas entre las puntuaciones pretest y posttest de la felicidad de los futbolistas al publicar una noticia sobre su ciudad natal.

Los resultados de la prueba W de Wilcoxon van a venir representados en un cuadro resumen que nos indica si debemos aceptar la hipótesis nula, es decir, que las diferencias no son significativas, o por el contrario, si debemos rechazarla.

Ejemplo de comparación de dos grupos relacionados (caso no paramétrico)

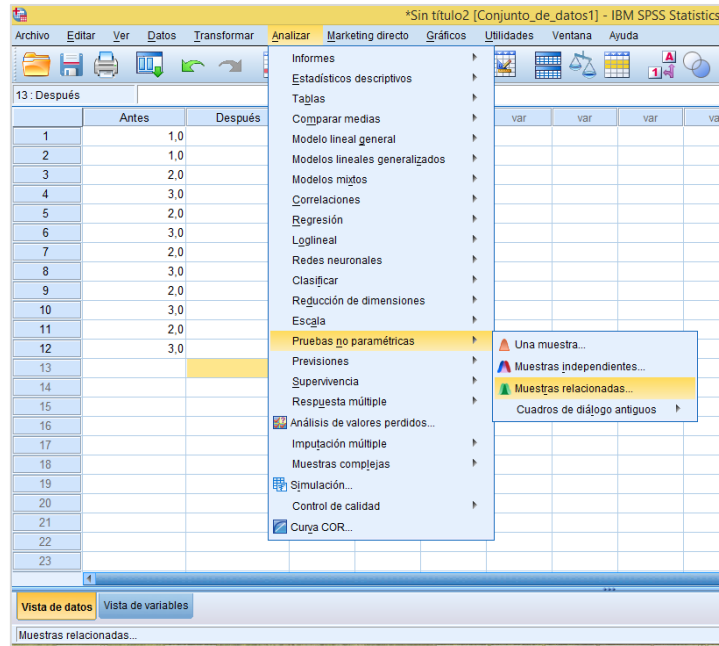
Tenemos los datos, expresados en miles, de las ventas mensuales de nuestro periódico antes y después de cambiar de editor (12 mediciones) en la hoja de PSPP que se observa en la siguiente figura:



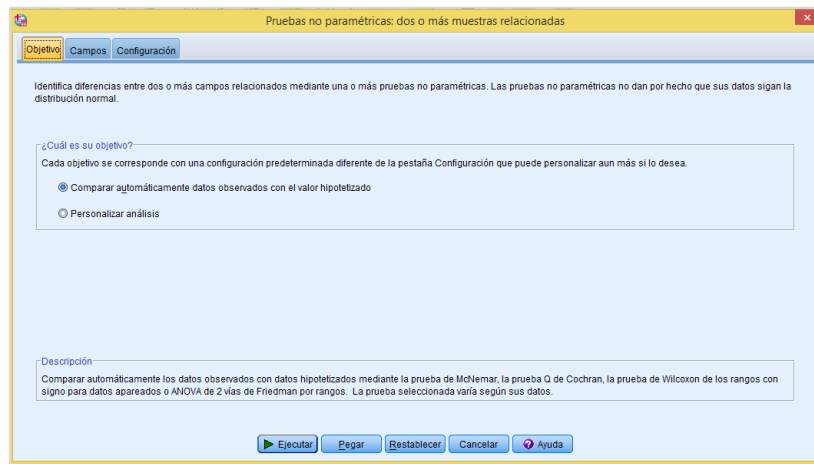
	Antes	Después	var
1	1,0	2,0	
2	1,0	2,0	
3	2,0	4,0	
4	3,0	7,0	
5	2,0	4,0	
6	3,0	7,0	
7	2,0	5,0	
8	3,0	7,0	
9	2,0	5,0	
10	3,0	8,0	
11	2,0	4,0	
12	3,0	8,0	
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			

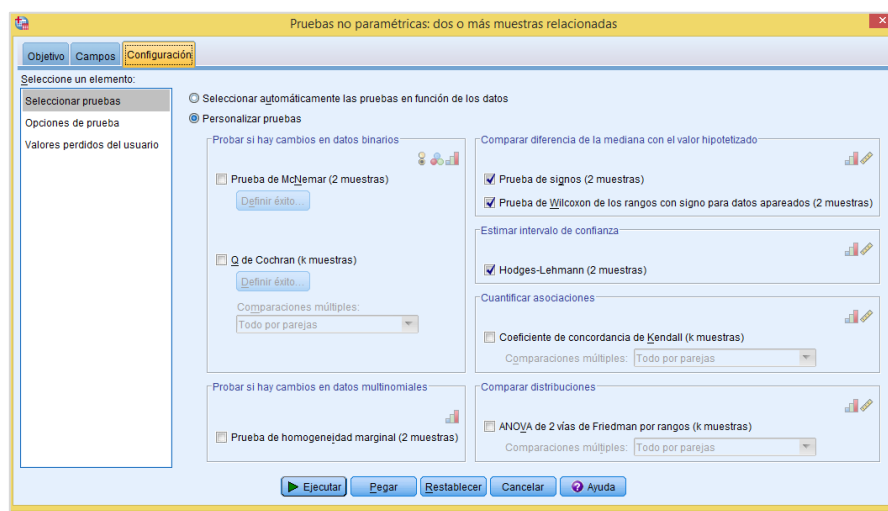
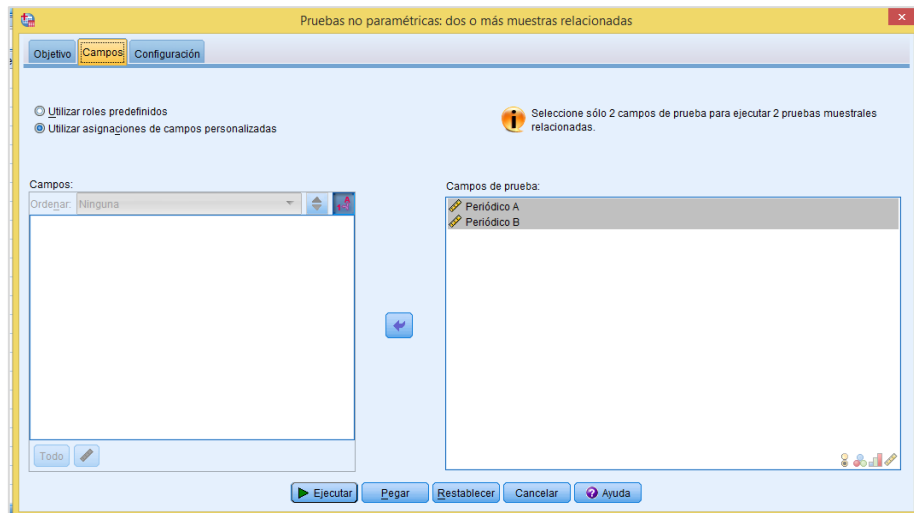
El día 1 de enero de 2014 se cambia el editor jefe y queremos ver si esto ha influido en el número de ventas. Observamos que son dos grupos relacionados claramente pero únicamente tenemos 12 mediciones, por lo que estamos en el caso no paramétrico. Ahora veamos cómo medir si existen diferencias significativas y, si las hay, si estas son significativas o debidas al azar.

Para ello seguimos la siguiente ruta: Analizar > Pruebas no paramétricas > Muestras Relacionadas

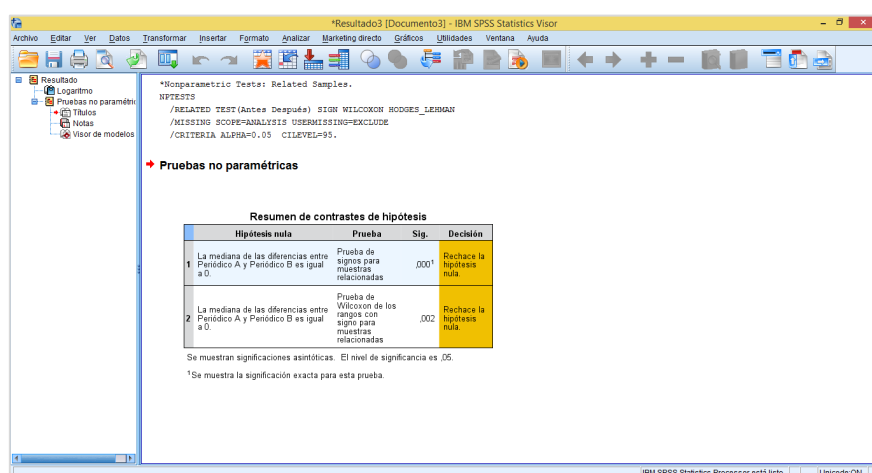


Nos aparecerá un nuevo cuadro con tres pestañas, las cuales se muestran en las siguientes figuras.





Pulsamos Ejecutar y obtenemos lo que se observa en la siguiente figura:



El cuadro resumen nos indica que rechazamos la hipótesis nula y, por lo tanto, las diferencias con el nuevo editor jefe son significativas. Además, podemos ver que las ventas son mayores con el nuevo editor.

Lo + recomendado

No dejes de leer...

Bioestadística amigable

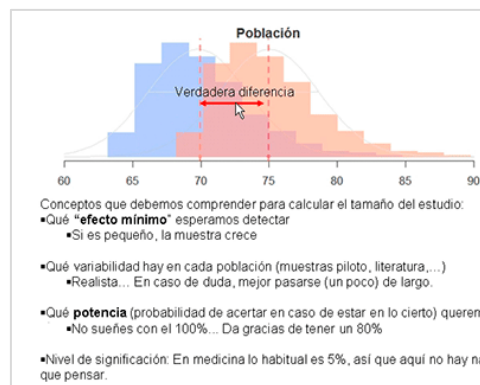
Martínez, M.A., Sánchez-Villegas, A. y Faulín, F. J. (2006). *Bioestadística amigable*. Madrid: Díaz de Santos.

En este libro se introduce al lector en el uso de Excel y SPSS.

No dejes de ver...

Cálculo del tamaño para la comparación de medios

Este vídeo supone un acercamiento al alumno hacia la comparación de grupos y para ello se explicará cómo calcular el tamaño para poder relacionar 2 grupos.

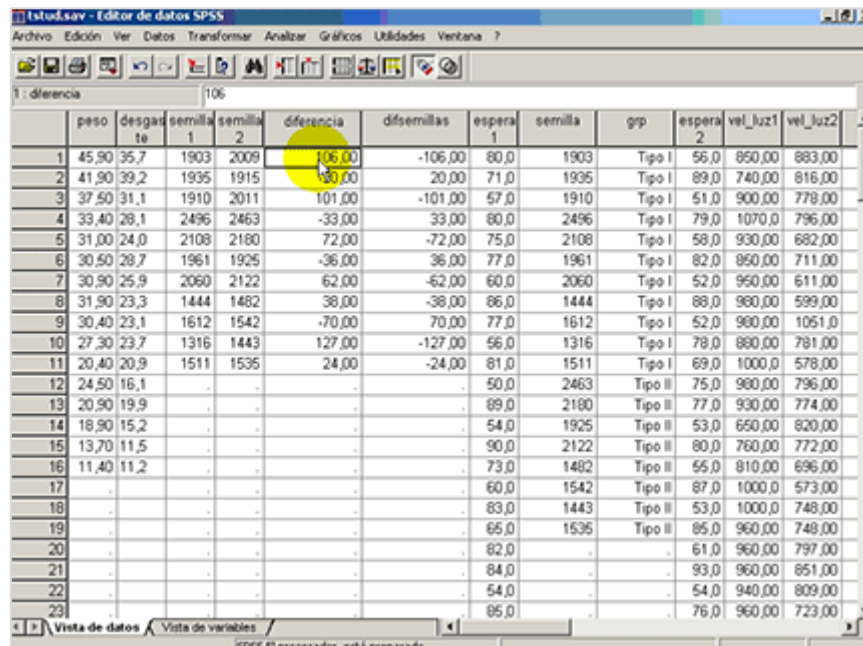


Accede al vídeo a través del aula virtual o desde la siguiente dirección web:

<http://www.bioestadistica.uma.es/baron/wordpress/master/calculo-de-tamano-de-muestras-para-comparar-2-grupos>

Comparar medias con SPSS

Este vídeo muestra los comandos de SPSS para comparar 2 grupos independientes.

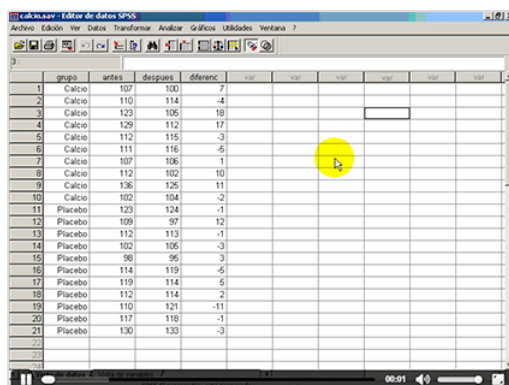


	peso	desgaste	semilla 1	semilla 2	diferencia	dísemillas	espera 1	semilla	g/p	espera 2	vel_luz1	vel_luz2
1	45,90	35,7	1903	2009	-106,00	-106,00	80,0	1903	Tipo I	56,0	850,00	883,00
2	41,90	39,2	1935	1915	-20,00	20,00	71,0	1935	Tipo I	89,0	740,00	816,00
3	37,50	31,1	1910	2011	-101,00	-101,00	57,0	1910	Tipo I	51,0	900,00	778,00
4	33,40	28,1	2496	2463	-33,00	33,00	80,0	2496	Tipo I	79,0	1070,0	796,00
5	31,00	24,0	2108	2180	-72,00	-72,00	75,0	2108	Tipo I	58,0	930,00	682,00
6	30,50	28,7	1961	1925	-36,00	36,00	77,0	1961	Tipo I	82,0	850,00	711,00
7	30,90	25,9	2060	2122	-62,00	62,00	60,0	2060	Tipo I	52,0	950,00	611,00
8	31,90	23,3	1444	1482	-38,00	-38,00	86,0	1444	Tipo I	88,0	980,00	599,00
9	30,40	23,1	1612	1542	-70,00	70,00	77,0	1612	Tipo I	52,0	980,00	1051,0
10	27,30	23,7	1316	1443	-127,00	-127,00	56,0	1316	Tipo I	78,0	880,00	781,00
11	20,40	20,9	1511	1535	-24,00	-24,00	81,0	1511	Tipo I	69,0	1000,0	578,00
12	24,50	16,1	-	-	-	-	50,0	2463	Tipo II	75,0	980,00	796,00
13	20,90	19,9	-	-	-	-	89,0	2180	Tipo II	77,0	930,00	774,00
14	18,90	15,2	-	-	-	-	54,0	1925	Tipo II	53,0	650,00	820,00
15	13,70	11,5	-	-	-	-	90,0	2122	Tipo II	80,0	760,00	772,00
16	11,40	11,2	-	-	-	-	73,0	1482	Tipo II	55,0	810,00	696,00
17	-	-	-	-	-	-	60,0	1542	Tipo II	87,0	1000,0	573,00
18	-	-	-	-	-	-	83,0	1443	Tipo II	53,0	1000,0	748,00
19	-	-	-	-	-	-	65,0	1535	Tipo II	85,0	960,00	748,00
20	-	-	-	-	-	-	82,0	-	-	61,0	960,00	797,00
21	-	-	-	-	-	-	84,0	-	-	93,0	960,00	851,00
22	-	-	-	-	-	-	54,0	-	-	54,0	940,00	809,00
23	-	-	-	-	-	-	85,0	-	-	76,0	960,00	723,00

Accede al vídeo a través del aula virtual o desde la siguiente dirección web:

http://www.bioestadistica.uma.es/baron/apuntes/ficheros/prac_doc_o4a.mp4

Comparación de grupos relacionados



	grupo	antes	después	diferencia	var1	var2	var3	var4	var5
1	Calcio	107	100	7					
2	Calcio	110	114	-4					
3	Calcio	123	106	18					
4	Calcio	129	112	17					
5	Calcio	112	115	-3					
6	Calcio	111	116	-5					
7	Calcio	107	106	1					
8	Calcio	112	102	10					
9	Calcio	136	125	11					
10	Calcio	102	104	-2					
11	Placebo	123	124	-1					
12	Placebo	109	97	12					
13	Placebo	112	113	-1					
14	Placebo	102	106	-3					
15	Placebo	98	96	3					
16	Placebo	114	119	-5					
17	Placebo	119	114	5					
18	Placebo	112	114	-2					
19	Placebo	110	121	-11					
20	Placebo	117	118	-1					
21	Placebo	130	133	-3					

Este vídeo explica los comandos para comparar dos grupos relacionados utilizando SPSS.

Accede al vídeo a través del aula virtual o desde la siguiente dirección web:

http://www.bioestadistica.uma.es/baron/apuntes/ficheros/prac_doc_o4b.mp4

+ Información

Webgrafía

Las siguientes páginas webs tienen explicaciones y ejemplos sobre la comparación de grupos independientes.

Prueba ANOVA: comparación de las medias de tres o más grupos



Accede a la página web a través del aula virtual o desde la siguiente dirección:

<http://epidemiologiamolecular.com/prueba-anova-comparacion-medias-grupos/>

Métodos no paramétricos para la comparación de dos muestras



Accede a la página web a través del aula virtual o desde la siguiente dirección:

<https://www.fisterra.com/mbe/investiga/noParametricos/noParametricos.asp>

Diferencia de medias poblacionales

DIFERENCIA DE MEDIAS POBLACIONALES $\mu_1 - \mu_2$

En ocasiones interesa definir un intervalo de valores tal que permita establecer cuáles son los valores mínimo y máximo aceptables para la diferencia entre las medias de dos poblaciones. Pueden darse dos situaciones según las muestras sean o no independientes; siendo en ambos casos condición necesaria que las poblaciones de origen sean normales o aproximadamente normales:

- **MUESTRAS INDEPENDIENTES**

Si puede suponerse que las varianzas de ambas poblaciones son iguales, el intervalo de confianza para la diferencia de medias poblacionales está centrado en la diferencia de las medias muestrales, siendo sus límites superior e inferior:

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$t_{\alpha/2}$ es el valor crítico correspondiente al grado de confianza $1-\alpha$ de la distribución t de Student con $n_1 + n_2 - 2$ grados de libertad y $S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$ es una estimación de la desviación típica común a ambas poblaciones obtenida a partir de las varianzas de las dos muestras. En la práctica si n_1 y n_2 son moderadamente grandes, el valor crítico $t_{\alpha/2}$ se aproxima, como ya se ha visto anteriormente, a los valores de la distribución normal.

Accede a la página web a través del aula virtual o desde la siguiente dirección:

http://www.ub.edu/aplica_infor/spss/cap4-3.htm

Comparación de grupos relacionados

En esta página web encontrarás la teoría y aplicación para la comparación de medias de grupos tanto independientes como relacionados.

Accede a la página web a través del aula virtual o desde la siguiente dirección:

http://www.mate.unlp.edu.ar/practicas/55_8_30052010143027.pdf

Bibliografía

Guàrdia, J. y Però, M. (2001). *Esquemas de estadística: aplicaciones en intervención ambiental*. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.

Salafranca, L., Solanas, A. y Núñez, M. I. (2000). *Estadística aplicada con SPSS y Statgraphics*. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.

Test

1. La prueba para comparar dos grupos independientes de 30 sujetos cada uno y variables cuantitativas es:
 - A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.

2. La prueba para comparar tres grupos independientes de 30 sujetos cada uno y variables cuantitativas es:
 - A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.

3. La prueba para comparar dos grupos independientes uno de 30 sujetos y el otro de 28 y variables cuantitativas es:
 - A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.

4. La prueba para comparar dos grupos independientes de 28 sujetos cada uno y variables cuantitativas es:
 - A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.

5. La prueba para comparar tres grupos independientes de 28 sujetos cada uno y variables cuantitativas es:
 - A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.

- 6.** La prueba para comparar dos grupos relacionados de 60 sujetos cada uno y variables cualitativas es:
- A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. W de Wilcoxon.
- 7.** La prueba para comparar dos grupos relacionados de 35 sujetos cada uno y variables cuantitativas es:
- A. F de Snedecor.
 - B. T de Student.
 - C. W de Wilcoxon.
 - D. H de Kruskal Wallis.
- 8.** La prueba para comparar dos grupos relacionados de 90 sujetos cada uno y variables cuantitativas es:
- A. F de Snedecor.
 - B. T de Student.
 - C. W de Wilcoxon.
 - D. H de Kruskal Wallis.
- 9.** La prueba para comparar dos grupos relacionados de 450 sujetos cada uno y variables cuantitativas es:
- A. W de Wilcoxon.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. H de Kruskal Wallis.
- 10.** La prueba para comparar dos grupos independientes de 10 sujetos cada uno y variables cualitativas es:
- A. F de Snedecor.
 - B. T de Student.
 - C. U de Mann-Whitney.
 - D. W de Wilcoxon.