

## 14A. PRUEBAS ESTADÍSTICAS PARA GRUPOS APAREADOS

En los estudios de **grupos apareados**, también llamados diseños de muestras, medidas o datos *intrasujetos*, *relacionados*, *agrupados* o *dependientes*, a diferencia de los estudios de grupos independientes, los elementos de los dos grupos son los mismos, aunque en circunstancias distintas (Figura 14A.1). La circunstancia diferenciadora más frecuente suele ser el tiempo, como analizar una variable (p. ej. *glucemia*) antes y después de administrar un tratamiento (p. ej. un nuevo antidiabético). Son los diseños pre-post. Además del tiempo pueden darse con otras circunstancias diferenciadoras como aplicar un colirio a un ojo y el placebo al otro, o bien un tratamiento tópico en un brazo y otro distinto o placebo en el otro brazo, etc.

**Figura 14A.1.** Diseño de estudios de grupos independientes y de grupos apareados.

Estudios o diseños de...			
...grupos o muestras o datos o medidas...			
...independientes o entresujetos		...dependientes o intrasujetos o apareados o relacionados	
Obesidad Sí/No	Glucosa (mg/dl)	Glucosa inicial	Glucosa final
Sí	132	132	133
Sí	158	158	144
.	.	.	.
No	95	95	101
No	130	130	110

A todos los sujetos del estudio se les ha sometido a un tratamiento dietético y farmacológico durante 6 meses con el fin de reducir su peso. Se obtiene el peso al finalizar el tratamiento, que llamaremos *Peso final*, y supuesta estable la talla, un *Imc* al final del tratamiento que llamaremos *Imc final*. El archivo *FactoresIntraF.sav* contiene las nuevas variables *Peso final*, *Imc final*, *Diferencia de Imc*, *Obesidad final* y *Nivel de masa corporal final*. Se trata de un diseño de grupos apareados pre-post con un solo grupo.

El análisis se realiza con pruebas estadísticas que permiten estudiar el cambio que se produce en la respuesta al pasar de un estado inicial a una respuesta final. Tienen la gran ventaja de eliminar la variabilidad entre sujetos, pero al no existir un grupo control disminuye la certeza de atribuir el cambio observado en la respuesta a la intervención, ya que entre ambos momentos pueden haber ocurrido otras circunstancias, además de la exposición, que sean las responsables del cambio observado.

La eficacia del tratamiento contra la obesidad se puede analizar con el estudio de la relación entre las variables inicial y final con las pruebas estadísticas para grupos apareados que dependen de la escala de medida de las variables utilizadas (Tabla 14A.1).

**Tabla 14A.1.** Pruebas estadísticas para datos apareados

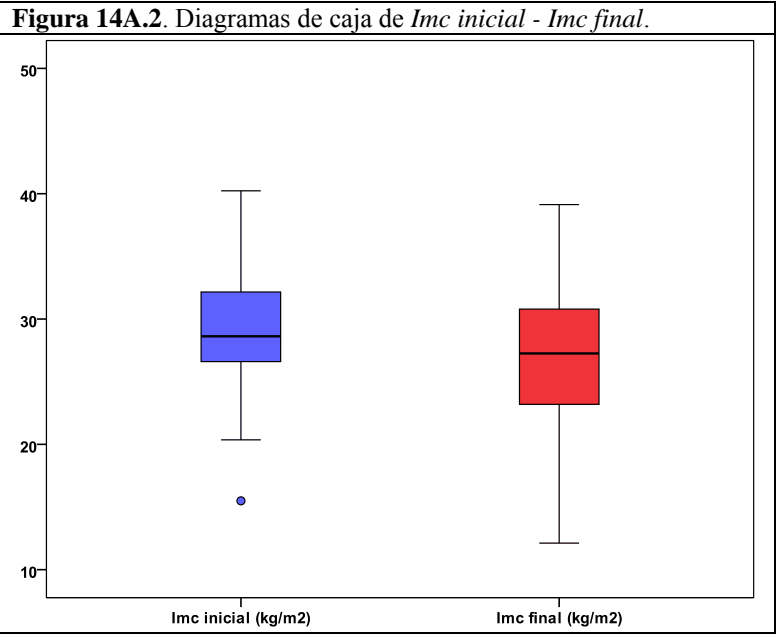
Variable respuesta	Variable exposición		
	Binaria <b>Obesidad inicial</b>	Politómica <b>Nivel masa corporal inicial</b>	Cuantitativa <b>Imc inicial</b>
Binaria <b>Obesidad final</b>	<i>Prueba de McNemar</i>		
Politómica <b>Nivel masa corporal final</b>		<i>Prueba de McNemar-Brooker</i>	
Cuantitativa <b>Imc final</b>			<i>Prueba de t de Student para grupos apareados</i> <i>Prueba no paramétrica</i> <i>T de Wilcoxon</i>

### RELACIÓN ENTRE DOS VARIABLES CUANTITATIVAS APAREADAS: *Imc* inicial y final

En la Tabla 14A.2 se observa que la media de *Imc final* es menor que la *inicial*. El diagrama de caja en la Figura 14A.2 muestra que el *Imc final* es menor que el *inicial* lo que también sugiere eficacia del tratamiento.

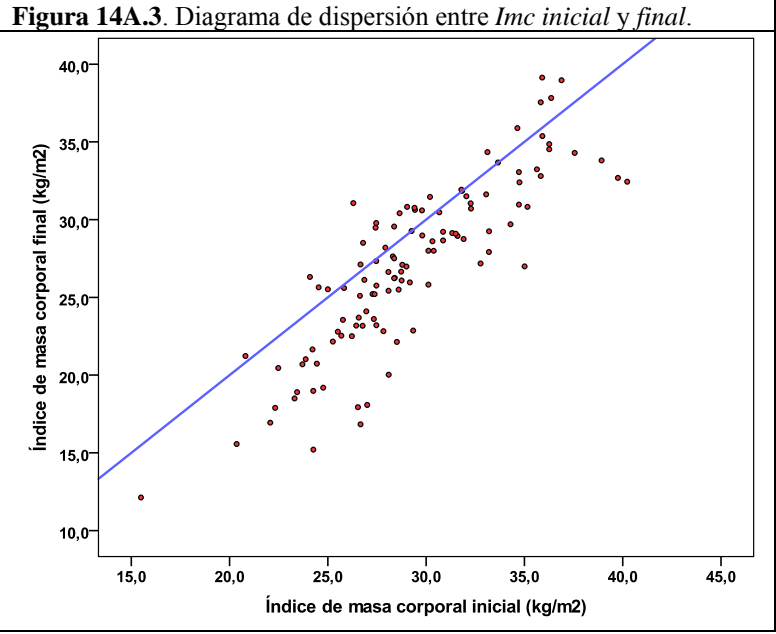
**Tabla 14A.2.** Comparación de medias *Imc inicial* - *Imc final*.

Media de <i>Imc inicial</i>	29,3 kg/m <sup>2</sup>
Media de <i>Imc final</i>	27,1 kg/m <sup>2</sup>



La Figura 14A.3 representa el diagrama de dispersión de *Imc inicial* y *final*. Se ha representado la línea diagonal media en azul. Es la función  $y = x$ ; en este caso  $Imc\ final = Imc\ inicial$ . Muestra de forma clara la lógica de la prueba *t de Student para grupos relacionados* para valorar la eficacia del tratamiento contra la obesidad y el cambio en el *Imc final* respecto al *inicial*.

La lineal diagonal media representa la ausencia de efecto: el *Imc inicial* es igual al *final*. La distancia vertical entre cada punto a la diagonal es el cambio de *Imc* en  $kg/m^2$  en cada sujeto a consecuencia del tratamiento. Es la variable *Diferencia de Imc*. La nube de puntos sugiere un efecto beneficioso del tratamiento, porque la mayor parte de los puntos están por debajo de la diagonal: el *Imc final* es predominantemente menor que el *inicial*, es decir, el tratamiento antiobesidad disminuye el *Imc*.



En la Figura 14A.4 se muestran las pruebas estadísticas utilizadas para la comparación de dos variables cuantitativas apareadas (pruebas *t de Student-Fisher para grupos apareados* y prueba no paramétrica *T de Wilcoxon*) y las condiciones de aplicación que deben evaluarse (Normalidad de la variable *Diferencia de Imc*).

▪ **Prueba t de Student-Fisher para muestras relacionadas**

Es la prueba para analizar la relación estadística entre dos variables cuantitativas apareadas, que es distinta de la prueba t de Student para grupos independientes descrita en la unidad 10. Comprueba si *Imc final* es menor que *Imc inicial* a partir de la variable *Diferencia de Imc* y valorar cuanto se aleja de cero su media.

**Figura 14A.4.** Comparación de dos variables cuantitativas apareadas: *Imc inicial* – *Imc final*.

Comparación dos variables cuantitativas apareadas <i>Imc inicial</i> – <i>Imc final</i>	
<b>Normalidad</b> (de la variable <i>Diferencia de Imc</i> )	
Pruebas: Kolmogorov-Smirnov y Shapiro-Wilks Gráficos: de Normalidad, de Caja, Histograma.	
Sí	No
<b>Prueba t de Student-Fisher para grupos apareados</b>	<b>Prueba no paramétrica T de Wilcoxon</b>

- Si *Imc inicial* e *Imc final* son iguales ( $H_0: Imc\ final - Imc = 0$ ) la diferencia será nula (*Diferencia de Imc* = 0), el estadístico de contraste de la t de Student valdrá cero y la prueba será no significativa ( $p = 1,000$ ).
- A medida que los valores de *Imc inicial* e *Imc final* se van haciendo más diferentes, el valor de la diferencia aumentará, al igual que el valor del estadístico de la t de Student y la prueba se hará más significativa.

En la Tabla 14A.3 se ofrece el resultado de la prueba t de Student para grupos apareados ( $p < 0,001$ ). Se puede concluir que el tratamiento contra la obesidad disminuye estadísticamente el *Imc*. Además informa de la magnitud del efecto: la media de las diferencias de *Imc* es 2,2 (IC95%: 1,7 – 2,8) kg/m<sup>2</sup> (Tabla 14A.3). El *Imc final* es, en promedio, 2,2 kg/m<sup>2</sup> y al menos 1,7 kg/m<sup>2</sup> (extremo más desfavorable del IC96%) menor que el inicial. La significación estadística también se puede deducir del IC95%: al no incluir la hipótesis nula (el cero) la diferencia es estadísticamente significativa al menos al 5% ( $p < 0,05$ ).

**Tabla 10A.3.** Pruebas estadísticas utilizadas en la comparación *Imc inicial* – *Imc final*.

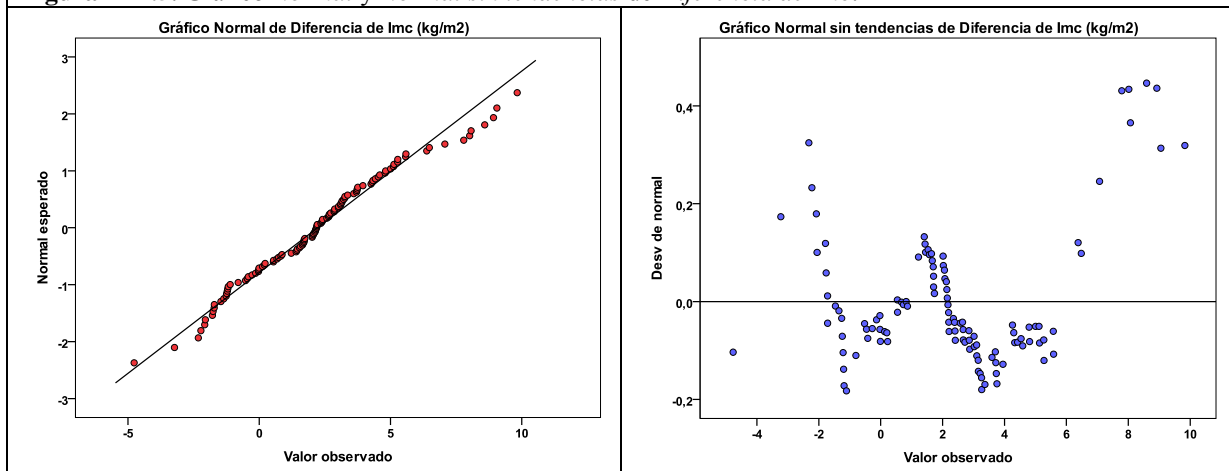
Prueba	Estadístico	gl	p	Media de las diferencias (IC95%)
Normalidad de Kolmogorov-Smirnov	0,071	112	<b>0,200</b>	
Normalidad de Shapiro-Wilks	0,982	112	0,124	
<b>t Student para grupos apareados</b>	t = 8,337	111	<b>&lt;0,001</b>	<b>2,2 (1,7 – 2,8) kg/m<sup>2</sup></b>
No paramétrica T de Wilcoxon	Z = -6,898		<0,001	

Esta prueba sirve para evaluar el efecto de una exposición, en este caso el tratamiento antiobesidad, sobre una respuesta, en este caso el cambio de *Imc*. Pero para valorar el grado de asociación lineal entre el *Imc inicial* y *final* se utiliza el *coeficiente de correlación de Pearson* entre ambas variables. En nuestra hipótesis esta prueba carece de interés porque lo que nos interesa es valorar el efecto del tratamiento sobre el *Imc* independientemente de que exista asociación entre ambas variables.

▪ **Pruebas y Gráficos de Normalidad**

Para poder aplicar la t de Student para grupos apareados se requiere que la *Diferencia de Imc* siga una distribución normal. La inspección de gráficos (Figura 14A.5) no contradice que siga la normalidad, y además la prueba de normalidad de Kolmogorov-Smirnov con la corrección de Lilliefors, adecuada para muestras > 30, resulta no significativa,  $p = 0,200$  (Tabla 14A.3), por lo que se puede asumir la normalidad.

**Figura 14A.5.** Gráfico Normal y Normal sin tendencias de *Diferencia de Imc*.



▪ **Prueba no paramétrica T de Wilcoxon**

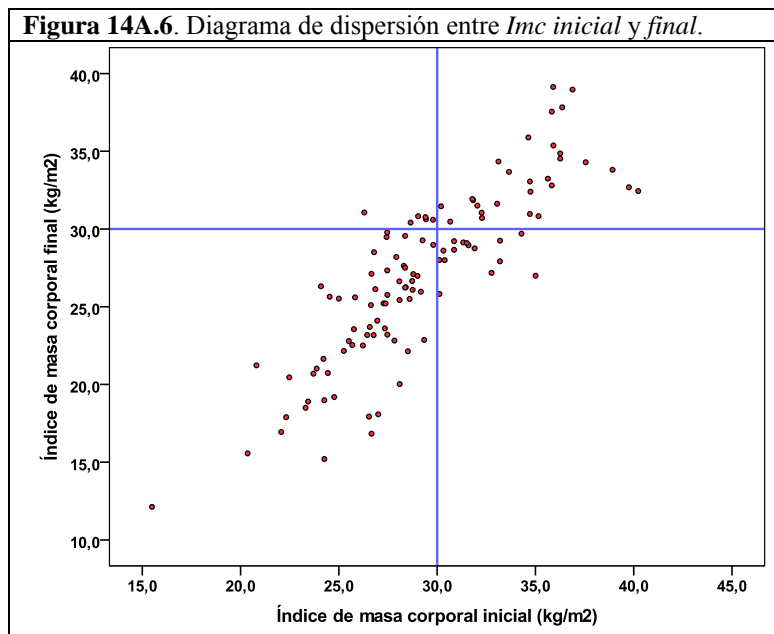
También llamada *prueba de los rangos con signo de Wilcoxon*, es la prueba no paramétrica para comparar dos variables cuantitativas apareadas utilizada cuando no se cumple el supuesto de normalidad. En esta caso también es significativa,  $p < 0,001$  (Tabla 14A.3).

### Videotutorial 14A1Apareados1.avi

Con el archivo *FactoresIntraF.sav* se muestra como se evalúa la relación entre las variables cuantitativas apareadas *Imc inicial* y *final*. Primero se obtiene la estadística descriptiva y el diagrama de caja de *Imc inicial* y *final* y los gráficos y pruebas de normalidad de *Diferencia de Imc* con el cuadro *Explorar*. Después se obtiene la **prueba t de Student-Fisher para grupos apareados** con el cuadro *Prueba t para muestras relacionadas* y la **prueba no paramétrica T de Wilcoxon** con el cuadro *Pruebas para dos muestras relacionadas*.

### RELACIÓN ENTRE DOS VARIABLES BINARIAS APAREADAS: Obesidad Sí/No inicial y final

La Figura 14A.6 representa el diagrama de dispersión de *Imc inicial* y *final*, con las líneas horizontal y vertical en azul referenciadas a  $Imc = 30 \text{ kg/m}^2$ . Muestra de forma clara la lógica de la prueba estadística de McNemar para grupos apareados, dirigida a valorar la relación entre el tratamiento contra la obesidad y el cambio en el estado de *Obesidad Sí/No*.



Estas líneas horizontal y vertical dividen a los casos en dos grupos según sean obesos al final (puntos por encima de la horizontal) o al inicio (puntos a la derecha de la vertical) o no sean obesos al final (puntos por debajo de la línea horizontal) o al inicio (puntos a la izquierda de la vertical). Estas dos líneas dividen el gráfico de dispersión en cuatro cuadrantes que aportan una información similar a la tabla de contingencia de *Obesidad inicial* y *Obesidad final* mostrada en la Tabla 14A.4.

**Tabla 14A.4.** Tabla de contingencia de *Obesidad inicial* y *final*.

		Obesidad final		Total
		No	Sí	
Obesidad inicial	Sí	15	28	43
	No	63	6	69
Total		78	34	112

- El cuadrante superior izquierdo contiene 6 puntos correspondientes a los no obesos antes del tratamiento y obesos después del tratamiento (recuadro inferior derecho de la Tabla 14A.4). En estos sujetos el tratamiento ha sido perjudicial.
- El cuadrante inferior derecho contiene 15 puntos, que son los 15 obesos antes del tratamiento que dejan de serlo después del tratamiento (recuadro superior izquierdo de la Tabla 14A.4). En estos casos el tratamiento es beneficioso. Los datos de estos 21 casos sugieren un efecto beneficioso del tratamiento: 15 pacientes mejoran y 6 empeoran.
- Cuadrante superior derecho: 28 obesos antes del tratamiento que continúan siéndolo después de él (recuadro superior derecho de la Tabla 14A.4). El tratamiento tiene efecto neutro sobre ellos.
- Cuadrante inferior izquierdo: 63 no obesos antes y después del tratamiento (recuadro inferior izquierdo de la Tabla 14A.4). El tratamiento tiene efecto neutro sobre ellos.

▪ **Prueba de simetría exacta de McNemar**

Es la prueba utilizada para la comparación de dos variables binarias de grupos apareados. La lógica del análisis de esta prueba consiste en comparar las casillas de la Tabla 14A.4 que indican cambios de respuesta en ambos sentidos: los 6 que pasan de no obesos a obesos y los 15 que pasan de obesos a no obesos. Los 91 casos restantes no intervienen en el análisis estadístico, y esta es la mayor debilidad de esta prueba. El nivel de significación exacto según la distribución binomial de la prueba de McNemar es  $p = 0,078$  (Tabla 14A.5). Se puede concluir que el tratamiento no es eficaz: no se puede descartar la hipótesis de que ha habido el mismo número de cambios de obesos a no obesos que de no obesos a obesos al aplicar el tratamiento ( $p = 0,078$ ).

El grado de asociación entre las variables *Obesidad inicial* y *Obesidad final* se mide mediante la **prueba de independencia de chi-cuadrado**. Es importante no confundirla con la prueba de simetría exacta de McNemar que es la prueba de homogeneidad para datos apareados. En nuestra hipótesis esta prueba de independencia de  $\chi^2$  carece de interés porque lo que nos interesa es valorar el efecto del tratamiento sobre la obesidad independientemente de que exista asociación entre la *Obesidad inicial* y *final*.

**Tabla 10A.5.** Pruebas estadísticas utilizadas en la comparación *Obesidad categórica inicial – final*.

Prueba	p
Mc Nemar ( <i>Obesidad inicial – final</i> )	0,078
Mc Nemar-Brooker ( <i>Nivel de Obesidad inicial – final</i> )	0,001

**RELACIÓN ENTRE DOS VARIABLES POLITÓMICAS APAREADAS:  
Nivel de masa corporal inicial y final**

Cuando se trata de comparar dos variables politómicas apareadas, *Nivel de masa corporal inicial* (antes del tratamiento) y *final* (tras finalizar el tratamiento), tendríamos una tabla de contingencia de 3x3 (Tabla 14A.6) y la lógica del análisis consiste en comparar las casillas que indican cambios de respuesta en ambos sentidos mediante una variante de la prueba de McNemar llamada **prueba de McNemar-Brooker**. Verifica si una tabla cuadrada es simétrica. El resultado es significativa ( $p=0,001$ ) (Tabla 14A.5). Para su cálculo utiliza la distribución de chi-cuadrado, por lo que requiere el supuesto de que la suma de los datos discordantes de la tabla sea  $\geq 10$ . En este caso se cumple, serían:  $0 + 15 + 18 + 6 + 3 + 0 \geq 10$ .

**Tabla 14A.6.** Tabla de contingencia de *Nivel de masa corporal inicial y final*.

		Nivel de masa corporal final			Total
		Normal	Sobrepeso	Obesidad	
Nivel de masa corporal inicial	Obesidad	0	15	28	43
	Sobrepeso	18	27	6	51
	Normal	15	3	0	18
Total		33	45	34	112

**Videotutorial 14A2Apareados3.avi**

Con el archivo *FactoresIntraF.sav* se muestra como se evalúa la relación entre dos variables categóricas apareadas. Primero se obtiene el gráfico de dispersión entre *Imc inicial* y *final* con el cuadro *Diagrama de dispersión simple* mostrando las líneas oblicua, horizontal y vertical. Después se obtiene la **prueba de McNemar** entre *Obesidad inicial* y *final* y la **prueba de McNemar-Brooker** entre *Nivel de masa corporal inicial* y *final* con el cuadro *Tabla de contingencia*.