

Interpretación de los resultados del contraste de medias (t de Student)

Es muy conveniente que a cada tabla de resultados le acompañe una explicación clara.

Si va a haber varias tablas semejantes, la explicación se pone después de la primera tabla

Esta explicación (que podría ser mucho más breve) tiene estas ventajas:

1. Muestra que el que ha hecho los análisis los entiende,
2. Le sirve para otras ocasiones (su tesis o su trabajo es *su libro de texto*).
3. Facilita pedir o buscar aclaraciones de lo que no se entiende bien.

Asignatura: Psicología ¹	Primer curso N = 94	Último curso N = 43			
<i>Variables</i>	M (σ)	M (σ)	$M_{\text{últ.}} - M_{\text{prim.}}$	t (p)	d (%)
Enfoque profundo	50.84 (6.20)	53.80 (6.99)	2.96	2.52 (.013)	.46 (68 %)
Enfoque superficial	46.15 (6.78)	39.60 (6.16)	-4.56	-4.567 (.000)	.84 (80%)
Autorregulación	77.53 (9.99)	79.23 (9.05)	2.62	3.621 (.010)	.48 (68%)

Explicación de la tabla

Muy brevemente: M y σ = *media y desviación típica*; t = *t de Student*, p = *probabilidad de ocurrencia asociada al valor de t* y d *tamaño del efecto*, pero caben explicaciones más o menos amplias. Una orientación: la explicación puede ser lo suficientemente amplia como para que el sujeto que la escribe lo tenga todo claro él mismo (o un posible lector), sin utilizar expresiones que no sabría explicar.

1. En la primera columna tenemos la denominación de las variables (puede ponerse abreviada) en las que se van a comparar los dos grupos (cursos primero y último de Psicología). En este caso se trata de variables relacionadas con el estudio, medidas con tests conocidos. Se trata de *muestras independientes* (sujetos físicamente distintos).
2. En las dos columnas siguientes tenemos los datos descriptivos de los dos grupos (media y desviación típica).
3. En la cuarta columna tenemos la diferencia entre las dos medias (en este caso media del último curso *menos* media del primer curso; en otros planteamientos podría ser la media de grupo experimental menos la media del grupo de control).
4. En la quinta columna siguiente está el valor obtenido de la t de Student (para *muestras independientes*) junto con la probabilidad (p) asociada al valor obtenido de t (es decir, la significación estadística). Este valor de p es la probabilidad de obtener *una diferencia distinta de cero* en el caso de que ambas muestras pertenezcan a la misma población sin más diferencias que las puramente aleatorias (*no es* la probabilidad de obtener *la misma diferencia* en comparaciones semejantes).

Como nuestro nivel de confianza es p = .05, la diferencia es estadísticamente significativa siempre que este valor de p sea igual o inferior a .05. En este caso está puesto el valor exacto de la probabilidad, bastaría poner p < .05 o p > .05 si ese es nuestro nivel de confianza como es la práctica habitual.

¹ Datos tomados de la tesis doctoral de Juan Carlos Torre Puente, *La autoeficacia, la autorregulación y los enfoques de aprendizaje en estudiantes universitarios*, Madrid, Universidad Pontificia Comillas, 2006.

También es usual utilizar tres valores de referencia o probabilidades de error para afirmar la diferencia: .05 (5%), .01 (1%) y .001 (1/1000), pero nos basta que la probabilidad sea igual o inferior a .05.

El valor de t asociado a su probabilidad (p) de ocurrencia nos dice con qué *seguridad* podemos afirmar la diferencia; en el caso de *enfoque profundo* una diferencia distinta de cero la obtendríamos el 1.3% de las veces si hiciéramos esta comparación en múltiples ocasiones ($p < .05$).

En este caso las diferencias estadísticamente significativas las tenemos en las tres variables; en enfoque profundo (más alta en el último curso), en enfoque superficial (más baja en el último curso) y en autorregulación (más alta en el último curso).

Lo que *no nos dice* el valor de t es si la diferencia es grande o pequeña, simplemente nos remite a una probabilidad. Con muestras grandes se obtienen con facilidad diferencias estadísticamente significativas que pueden ser muy pequeñas y triviales. Un resultado *estadísticamente significativo* no quiere decir que sea *relevante* o *importante* (expresiones como *significativo* o *significancia* se prestan a interpretaciones equívocas, no es lo mismo significación *estadística* que significación *conceptual* o *relevancia práctica*).

Si el valor de t nos remite a una probabilidad mayor de .05 ($p > .05$) pero tampoco es una probabilidad muy grande (por ejemplo $p = .06$) y las muestras comparadas son más bien pequeñas, no podemos afirmar la diferencia (con nuestro nivel de confianza) pero sí se puede proponer, *como hipótesis*, que comparando muestras mayores es probable obtener una diferencia estadísticamente significativa, sobre todo si la *magnitud* de la diferencia (*tamaño del efecto*) es *apreciable*.

5. Por último tenemos el *tamaño del efecto* (de Cohen, 1988, el más utilizado) simbolizado con la letra d, y debajo del tamaño del efecto el porcentaje de sujetos del grupo con media inferior superados por el sujeto medio del grupo con media más alta (no siempre se pone este porcentaje pero es conveniente hacerlo). El signo del tamaño del efecto es indiferente (se puede poner o no poner, es el mismo de la diferencia entre las dos medias) pero hay que tenerlo en cuenta en la interpretación.

Con muestras de tamaño no muy desigual el tamaño del efecto se puede *estimar* con esta fórmula. $d = 2t(\sqrt{(N_1+N_2)-2})$.

Para valorar la magnitud del tamaño del efecto (de la diferencia entre las medias) suelen tenerse en cuenta las *orientaciones* de Cohen (1988):

d = .20	diferencia pequeña
d = .50	diferencia moderada
d = .80	diferencia grande

También pueden valorarse estas diferencias *en términos relativos* (cuáles son mayores y menores entre las que se encuentran el mismo planteamiento).

6. Junto al tamaño del efecto tenemos la proporción (o porcentaje) de sujetos del grupo con media más baja superados por el *alumno medio* del grupo con media más alta (que en su propio grupo superaría al 50 %). Por ejemplo, el alumno medio de primer curso supera al 80% de los alumnos del último curso en *estudio superficial*.

Como el *tamaño del efecto* (d) equivale a una *puntuación típica* (z) este porcentaje o proporción es la que en la distribución normal *cae* por debajo de una puntuación típica igual al tamaño del efecto. Esta proporción se busca en las tablas de la distribución normal o en programas de Internet. Estas proporciones son las que pueden justificar las valoraciones del tamaño del efecto.

Dos análisis que nos llevarían a la misma conclusión:

1) Calcular la *correlación entre cada variable y la pertenencia a uno u otro grupo* (1 ó 0). Los coeficientes de correlación (con valores extremos de 0 y ± 1) también expresan una magnitud. Los dos cálculos son intercambiables, incluso hay fórmulas para convertir la t en r y la r en t. El contraste de medias (t de Student) es para muchos más fácilmente comprensible.

2) Calcular los *intervalos de confianza de las medias: límites probables superior e inferior* en las poblaciones representadas por estos sujetos. Si el límite inferior del grupo con media mayor es superior al límite superior del grupo con media más baja, la diferencia entre las medias es estadísticamente significativa. En ocasiones puede ser un buen complemento al contraste de medias convencional que además se presta a representaciones gráficas muy claras; aun así haría falta calcular el tamaño del efecto.