Análisis de varianza para verificar *tendencias*

©Pedro Morales Vallejo
Universidad Pontificia Comillas, Madrid
Facultad de Ciencias Humanas y Sociales
(última revisión 28 de febrero de 2009)

índice

1. Introducción

El análisis de varianza para muestras *independientes* o para muestras *relacionadas* nos dice si entre las medias de varias muestras hay diferencias significativas, pero no nos dice si en las medias se observa una *tendencia a crecer o decrecer.* Es más, al comparar varios grupos podemos encontrarnos con una razón F no significativa que nos dice que no hay diferencias (todas las medias proceden de la misma población) cuando por simple observación de los datos vemos que hay una clara tendencia en las medias, que van siendo progresivamente mayores o menores. Lo mismo puede suceder con las medias de un mismo grupo obtenidas en ocasiones sucesivas: entre las medias puede no haber diferencias significativas, pero sí puede haber una tendencia clara y *superior a lo aleatorio* a aumentar o disminuir.

Siempre que podamos *ordenar* los grupos con algún criterio (edad, curso, ocasiones sucesivas, etc.) podemos comprobar si se da una *tendencia lineal (linear trend)* a aumentar o disminuir; también podemos comprobar otras tendencias, pero ahora nos fijamos en las tendencias lineales. Las categorías de clasificación son por lo tanto *cuantitativas,* de manera que sea posible establecer un *orden* numérico.

El análisis de varianza convencional responde a una hipótesis muy genérica (si los grupos pertenecen o no a la misma población), pero podemos hacer además preguntas más específicas como si se observa o no una *tendencia a aumentar o disminuir* superior a lo meramente aleatorio. A esta pregunta podemos responder también mediante el análisis de varianza apropiado[1].

2. Verificación de tendencias en muestras independientes

2.1. Análisis de varianza

Vamos a verlo con un ejemplo[2]. Tenemos cinco grupos *ordenados de menos a más* con un criterio claro, como puede ser la edad. En este ejemplo (datos ficticios, tabla 1) tenemos un total de 20 sujetos (N = 20) con 4 sujetos en cada grupo (k o número de grupos = 5).

Tanto en este caso como con muestras relacionadas, la solución tiene dos pasos:

1º Se resuelve el análisis de varianza convencional (en este caso se trata de muestras independientes),

2º Con los datos obtenidos se pasa a un segundo análisis de varianza específico para verificar *tendencias* (entre los grupos puede no haber diferencias significativas, pero la *tendencia* a aumentar o disminuir sí puede ser significativa).

---

[1] En el anexo VII mencionamos dos métodos *no paramétricos* para verificar tendencias en muestras independientes (Jonckheere) y relacionadas (Page). Existe al menos otro método no paramétrico para verificar tendencias cuando n = 1 (prueba de Mann) que puede ser de especial interés para analizar terapias individuales (por ejemplo en psicología clínica y educación especial).

[2] El procedimiento que vamos a exponer puede verse también explicado y con ejemplos resueltos en bastantes textos no elementales (como Guilford y Fruchter, 1973; Rosenthal, 1987; Kirk, 1995). A veces se advierten aparentes discrepancias metodológicas en los distintos autores, simplemente porque unos prefieren utilizar como dato la *media* de cada grupo y otros el *total* de las puntuaciones de cada grupo.

	1º	2º	3º	4º	5º
N = 20	14	15	16	17	18
n = 4	14	15	16	17	18
k = 5	11	12	13	14	15
	10	11	12	13	14
Media =	12.25	13.25	14.25	15.25	16.25
Desviación (σ_n)=	1.785	1.785	1.785	1.785	1.785
Total =	49	53	57	61	65

Tabla 1

Si hacemos un análisis de varianza convencional nos encontramos con estos resultados (tabla 2):

origen	Suma de Cuadrados	Grados de Libertad	Cuadrados Medios	F
entre los grupos	40	k - 1 = 5 - 1 = 4	40/4 = 10	2.35, p > .05
dentro de los grupos	63.75	N - k = 20 - 5 = 15	63.75/15 = 4.25	

Tabla 2

De esta tabla solamente vamos a utilizar los *Cuadrados Medios dentro de los grupos* (que es el *término del error*) por lo que podemos calcularlos directamente si únicamente nos interesa la *tendencia* y no nos interesa comprobar si las medias difieren significativamente entre sí:

$$CM_{dentro} = \frac{\Sigma n \sigma_{\mathbf{n}}^2}{N - k} \qquad [1]$$

En cada grupo multiplicamos el número de sujetos (de cada grupo, k = 4 en este ejemplo) por su varianza (calculada dividiendo por N, no por N-1), sumamos estos productos y los dividimos por los grados de libertad *dentro* de los grupos (número total de sujetos menos número de grupos). Si hemos calculado las desviaciones dividiendo por N-1 (σ_{n-1}) el denominador de [1] será N.

En nuestro ejemplo, como los grupos son del mismo tamaño, podemos simplificar el numerador ($n\Sigma\sigma^2$ en vez de $\Sigma n\sigma^2$):

$$\textit{Cuadrados Medios dentro} = \frac{4(1.785^2 + 1.785^2 + 1.785^2 + 1.785^2 + 1.785^2)}{20 - 4} = 4.25$$

Como en casos semejantes (muestras independientes), si calculamos las varianzas de los grupos dividiendo por N-1, los *Cuadrados Medios dentro* serán $\Sigma n\sigma^2_{n-1}/N$, y si los grupos son de idéntico tamaño $n\Sigma\sigma^2_{n-1}/N$.

Vemos en la tabla 2 que la razón F no es estadísticamente significativa; para 4 y 15 grados de libertad nos haría falta un valor de F = 3.06 para rechazar (no aceptar) la Hipótesis

Nula de no diferencia y el que hemos encontrado es F = 2.35; nuestra primera conclusión es que las diferencias entre los grupos están dentro de lo normal.

Este resultado contradice aparentemente lo que vemos en los datos (ejemplo ficticio). Observamos que las medias van aumentando progresivamente, cada grupo tiene una media mayor que el anterior. Es más, si calculamos la correlación entre el total (o la media) de cada grupo y el *número de orden* del grupo (que podríamos substituir por la edad media del grupo, por ejemplo) nos encontramos con una relación perfecta de r = 1 (en este caso N = 5 y grados de libertad igual a N - 2 = 3).

Lo que sucede es que en el análisis de varianza que hemos hecho *se prescinde del orden* y la respuesta que obtenemos es muy genérica (*¿hay diferencias entre los grupos*?) y realmente no corresponde a nuestra pregunta de interés: si se da o no se da una *tendencia*, en este caso a aumentar progresivamente. Si ordenamos los grupos de manera diferente, el valor de F será el mismo, pues es independiente del orden.

La respuesta a nuestra pregunta (hay o no hay una determinada *tendencia*) es hacer un análisis de varianza que tiene en cuenta la varianza debida a la hipotética tendencia a aumentar.

El procedimiento lo explicamos a partir de los datos de la tabla 3.

	Grupo 1º	Grupo 2º	Grupo 3º	Grupo 4º	Grupo 5º	
Totales T =	49	53	57	61	65	
pesos λ =	-2	-1	0	+1	+2	$\Sigma\lambda^2 = 10$
Tλ =	-98	-53	0	+ 61	+ 130	$\Sigma T\lambda = L = 40$

Tabla 3

Qué hemos hecho:

1º Tomamos como dato inicial los totales de cada grupo (tomados de la tabla 1). Si los grupos son de distinto tamaño, el total de cada grupo será igual a la media del grupo multiplicada por la *media armónica del número de sujetos* (lo aclaramos más en el apartado 2.3).

2º Asignamos a cada grupo un *peso* (simbolizado por la letra griega *lambda,* λ) que expresa nuestra predicción. Nuestra predicción es que las medias van *de menos a más,* por eso los pesos asignados se corresponden con la tendencia que queremos contrastar. *La condición para asignar estos pesos es que su suma sea igual a 0.*

Si nuestra predicción no hubiera sido una tendencia lineal (crecimiento o decrecimiento progresivo) sino una tendencia *cuadrática* (tendencia primero a aumentar y luego a disminuir, o al revés) nuestros pesos hubieran sido +2, -1, -2, -1, +2. Estas tendencias cuadráticas tienen la forma de ∩ o ∪.

Los pesos (λ) para estos contrastes según las predicciones (lineales, cuadráticas, cúbicas) y según el número de grupos (o número de ocasiones si se trata del mismo grupo) se pueden encontrar en muchos textos (en el anexo ponemos los pesos para los casos más frecuentes (tendencias lineales y cuadráticas, entre tres y seis grupos u ocasiones).

3º Sumamos los pesos elevados previamente al cuadrado:

$$\Sigma\lambda^2 = (-2)^2 + (-1)^2 + (+1)^2 + (+2)^2 = 10$$

4º Multiplicamos cada Total por su peso λ (Tλ)

5º Sumamos estos valores de Tλ; designamos esta suma (ΣTλ) con el símbolo L:

$$L = -98 - 53 + 61 + 130 = 40$$

6º Calculamos la *Suma de Cuadrados* correspondiente a la tendencia lineal mediante esta fórmula:

$$SC_{tendencia} = \frac{L^2}{(n)(\Sigma\lambda^2)} \qquad [2] \qquad = \frac{40^2}{(4)(10)} = \frac{1600}{40} = 40$$

Cuadrados Medios (o varianza correspondiente a la tendencia):

$$CM_{tendencia} = \frac{SC_{tendencia}}{\text{grados de lib.}[=1]} \qquad [3] \qquad = 40$$

En estos contrastes los grados de libertad de la tendencia son siempre igual a 1, por lo que la Suma de Cuadrados es igual a los Cuadrados Medios.

7º Calculamos la razón F dividiendo los Cuadrados Medios de la tendencia lineal por el término del error (Cuadrados Medios *dentro* de los grupos) calculado antes (tabla 2):

$$F = \frac{40}{4.25} = 9.41$$

En las tablas de la razón F encontramos que para un grado de libertad en el numerador y 15 en el denominador (N-k = 20 - 5), este valor corresponde a $p < .01$ (ó $p < .001$ si consultamos tablas más completas). Podemos afirmar con mucha seguridad que la tendencia a aumentar que observamos dista de ser casual.

2.2. Cálculo de coeficientes de correlación

A partir de los datos ya calculados podemos calcular dos coeficientes de correlación (r de Pearson) entre la variable independiente (como puede ser la edad, o el número asignado a cada grupo) y la variable dependiente (la que hemos medido); estos coeficientes son una estimación del *tamaño del efecto* que nos permite apreciar y *valorar la magnitud de la tendencia* (el coeficiente de correlación tiene pleno sentido tratándose precisamente de una tendencia linear).

1º Tomando el *grupo* como unidad (Rosenthal, 1987:141):

$$r = \sqrt{\frac{SC_{tendencia}}{SC_{entre}}} \qquad [4] \qquad = \sqrt{\frac{40}{40}} = 1$$

Llegaremos al mismo resultado si calculamos el coeficiente de correlación entre la media (o el total) de cada grupo y la numeración *(número de orden)* que corresponde a cada grupo (el *sujeto* o unidad es aquí el grupo, tenemos tantos *sujetos* como grupos). Es como si calculáramos la correlación con todos los sujetos, pero asignando a cada sujeto, como puntuación individual, la media (o total) de su grupo.

2º Tomando el *sujeto* como unidad:

$$r = \sqrt{\frac{F}{F + \text{grados de libertad } \textit{dentro}}} \qquad [5] \qquad = \sqrt{\frac{9.41}{9.41+15}} = .62$$

$$\text{o también} \quad r = \sqrt{\frac{SC_{\text{tendencia}}}{SC_{\text{tendencia}} + SC_{\text{dentro}}}} \qquad [6] \qquad = \sqrt{\frac{40}{40+63.75}} = .62$$

En los denominadores tenemos en [5] los grados de libertad *dentro de los grupos* o en [6] la suma de cuadrados *dentro de los grupos;* quizás con más propiedad habría que decir grados de libertad o suma de cuadrados del *término del error* (tratándose de varias muestras independientes el término del error es la varianza *dentro* de los grupos).

Cuando tomamos el sujeto individual como unidad los coeficientes de correlación son menores que cuando utilizamos los totales o medias de cada grupo (en éste y en cualquier otro caso), sencillamente porque no prescindimos de las diferencias individuales dentro de cada grupo.

Los dos coeficientes dan información distinta. Cuando nos referimos, por ejemplo, a la correlación entre edad y cualquier otra variable, nos referimos habitualmente a la correlación utilizando a *cada sujeto como unidad;* el utilizar la media o el total del grupo (como en [4]) también es útil porque nos *avisa* sobre la tendencia de una manera muy clara. Sin embargo estas correlaciones tomando el grupo como unidad son una estimación pobre de la correlación entre el dato individual y el grupo de pertenencia (edad en este caso): aquí hemos bajado de una perfecta relación (r = 1) a una relación moderada de r = .62 Como criterio general, tomaremos al sujeto como unidad utilizando las fórmulas [5] ó [6].

2.3. Cuando los grupos son de distinto tamaño

Cuando los grupos son de distinto tamaño el procedimiento es el mismo con dos modificaciones:

1º el valor de n (número de sujetos en cada grupo), pasa a ser la *media armónica* de n:

$$\bar{n}_{\text{armónica}} = \frac{k}{\Sigma(1/k)} \qquad [7]$$

Vamos a suponer que en el ejemplo anterior los valores de n para los cinco grupos son 4, 5, 4, 6, y 5; la *media armónica* de *n* sería en este caso:

$$\bar{n}_{\text{armónica}} = \frac{5}{\frac{1}{4}+\frac{1}{5}+\frac{1}{4}+\frac{1}{6}+\frac{1}{5}} = \frac{5}{1.0667} = 4.69$$

2º El valor de T pasa a ser la media de cada grupo (M) multiplicada por la media armónica del número de sujetos:

$$T = (M)(\bar{n}_{\text{armónica}}) \qquad [8]$$

2.4. Contrastes posteriores entre medias

Si nos interesa podemos comparar las medias de los grupos de dos en dos, o también podemos comparar una media con la media combinada de otros grupos; podemos comparar tanto medias como *medias de medias* si queremos unir dos grupos o más, teniendo en cuenta, al aplicar la fórmula, que si unimos dos grupos o más cambiará el número de sujetos.

El contraste de medias lo hacemos mediante la t de Student con esta fórmula[3]:

$$t = \frac{M_1 - M_2}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)CM_{\text{dentro}}}} \qquad [9]$$

Los grados de libertad para consultar las tablas son los de los Cuadrados Medios *dentro* de los grupos *(término del error)*, 15 en este ejemplo.

Lo vemos con dos ejemplos:

a) Comparamos las medias de dos grupos, el 5º y el 3º:

$$t = \frac{16.25-14.25}{\sqrt{\left(\frac{1}{4}+\frac{1}{4}\right)4.25}} = 1.372 \ (p > .05)$$

b) Comparamos la media de un grupo, el 5º, con la media combinada de los grupos 1º, 2º y 3º.

La media de los tres primeros grupos, como tienen el mismo número de sujetos, es igual a la *media de las medias* = 13.25; el número de sujetos de este *nuevo* grupo es n = 4+4+4 = 12:

$$t = \frac{16.25-13.25}{\sqrt{\left(\frac{1}{12}+\frac{1}{4}\right)4.25}} = 2.52 \text{ (con 15 de libertad: } p < .05)$$

[3] Es la misma fórmula de los contrastes posteriores de Fisher (LSD, *Least Statistical Difference*).

De manera semejante podemos hacer otros contrastes de interés, aunque en este planteamiento lo que suele interesar fundamentalmente es verificar si se da o no se da una determinada tendencia: comprobar si la varianza debida a la tendencia a aumentar es superior a la varianza del *término del error* (la varianza dentro de los grupos, que expresa la *variabilidad normal).*

3. Verificación de tendencias en muestras relacionadas

Cuando se trata de los *mismos sujetos* medidos en la misma variable en *ocasiones sucesivas* tenemos un planteamiento semejante: podemos comprobar si una determinada tendencia es estadísticamente significativa (superior a lo que podríamos encontrar por azar).

Tenemos dos procedimientos para llegar a una respuesta:

1º Un análisis de varianza semejante al anterior;
2º Un contraste de medias utilizando unas puntuaciones individuales de *tendencia al cambio* que podemos calcular para cada sujeto.

Este segundo procedimiento es independiente del análisis de varianza propiamente dicho pero completa la información; responde a estas preguntas ¿El cambio o evolución observado en nuestra muestra, se aparta significativamente de una media de cambio = 0? ¿Cuál es la *magnitud* del cambio?

3.1. Análisis de varianza

Vamos a suponer que los mismos cuatro sujetos han sido medidos en tres ocasiones sucesivas (tabla 4).

	1ª ocasión	2ª ocasión	3ª ocasión	Total	Media
N = 12	10	12	14	36	12
n = 4	9	11	13	33	11
k = 3	5	3	7	15	5
	4	2	6	12	4
Total =	28	28	40		
Media =	7	7	10		
Desviación =	2.55	4.527	3.535		

Tabla 4

Los únicos datos que necesitamos de la tabla 4 son los *totales* de las *filas* y de las *columnas*, aunque también hemos puesto en la tabla las medias (si sólo tenemos las medias, los totales son igual a la media por número de sujetos).

Lo primero que vamos a hacer es un *análisis de varianza convencional* para *muestras relacionadas* (porque tenemos al mismo sujeto en cada fila); para calcular las sumas de cuadrados de filas y columnas son suficientes o los totales o las medias.

Por lo que respecta a las *filas* (los sujetos) lo más cómodo suele ser utilizar los totales. En el caso de las *columnas* (las ocasiones en que han sido medidos los sujetos) las medias y desviaciones son en sí mismas informativas, aunque también podemos hacer el análisis de varianza a partir de los totales.

Los resultados del análisis de varianza para muestras relacionadas están en la tabla 5.

origen	Sumas de Cuadrados	Grados de libertad	Cuadrados Medios	F
sujetos *(filas)*	150	4 -1 = 3	150/3 = 50	37.59, $p < .01$
ocasiones *(columnas)*	24	3 -1 = 2	24/2 = 12	9.02, $p<.01$
error *(fil. x col)*	8	(f-1)(c-1) = 6	8/6 = 1.33	
total	182	N -1 = 11		

Tabla 5

La varianza debida a que los sujetos *(filas)* son distintos es estadísticamente significativa, pero este resultado no nos interesa especialmente. La varianza debida a las ocasiones también es estadísticamente significativa; es decir, hay diferencias significativas entre las ocasiones, pero lo que *no nos dice* este resultado es si la *tendencia a aumentar progresivamente* es estadísticamente significativa, que es precisamente lo que queremos verificar. Es verdad que al observar las medias podemos ver que la media que corresponde a la tercera ocasión es la media mayor, pero con los mismos datos, si cambiamos de orden las columnas, tendríamos el mismo resultado (F estadísticamente significativa). Si nos limitamos a este análisis de varianza no verificamos si esta *tendencia a aumentar* de ocasión a ocasión es superior a lo que podemos esperar por azar.

Para comprobar si la *tendencia lineal* es estadísticamente significativa hacemos un análisis de varianza semejante al que hemos visto con muestras independientes. Necesitamos calcular los Cuadrados Medios (varianza) de esta tendencia para dividirlo por la varianza del término del error que ya hemos calculado en el análisis de varianza precedente (tabla 5). Los datos los disponemos tal como figuran en la tabla 6.

ocasiones	1ª	2ª	3ª	
Totales	28	28	40	
λ	-1	0	+1	$\Sigma\lambda^2 = 2$
$T\lambda$	-28	0	+ 40	$L = \Sigma T\lambda = 12$

Tabla 6

Los coeficientes λ corresponden a una tendencia lineal (crecimiento o decrecimiento progresivo).

La Suma de Cuadrados de esta tendencia la calculamos como antes [2]:

$$CM_{\text{tendencia}} = \frac{L^2}{(n)(\Sigma\lambda^2)} = \frac{12^2}{(4)(\ 2)} = 18$$

Como los grados de libertad son siempre = 1, los Cuadros Medios coinciden con la Suma de Cuadrados. Ahora calculamos la razón F utilizando el término del error calculado en el análisis de varianza precedente (tabla 5):

$$F = \frac{18}{1.33} = 13.53$$

Para grados de libertad = 1 (numerador) y 6 (denominador) tenemos que $p < .05$

No solamente hemos probado que hay diferencias entre las ocasiones (primer análisis de varianza) sino que además se da una *tendencia lineal* a aumentar progresivamente de ocasión a ocasión.

En este ejemplo el resultado era de esperar, pues vemos que la media en la tercera ocasión es superior a las medias de las ocasiones anteriores, pero no siempre los datos son tan claros. Y también puede suceder que en el primer análisis de varianza no tengamos resultados significativos y sí lo tengamos al verificar de manera específica la tendencia.

Si nuestra hipótesis o predicción fuera que la tendencia es *cuadrática* (ir primero *de menos a más* y después *de más a menos* o viceversa) el procedimiento es el mismo; lo que cambian son los pesos λ por los que multiplicamos los totales, que son los que expresan una tendencia *cuadrática* . En este caso los cálculos serían los que figuran en la tabla 7.

ocasiones	1ª	2ª	3ª	
Totales	28	28	40	
λ	-1	+2	-1	$\Sigma\lambda^2 = 6$
$T\lambda$	-28	+ 56	- 40	$L = \Sigma T\lambda = -12$

Tabla 7

Los Cuadrados Medios de la *tendencia cuadrática* son = $\frac{-12^2}{(4)(6)} = .50$

Y la razón F es igual a F = $\frac{.50}{1.33} = .37$

En este caso, puesto solamente como ejemplo, el resultado es obviamente no significativo (el denominador es superior al numerador): hay diferencias entre las columnas (ocasiones) tal como vimos en el primer análisis de varianza para muestras relacionadas, pero lo que no hay es una *tendencia cuadrática,* sino lineal en este caso. Con los mismos datos podemos verificar las dos tendencias.

3.2. Contraste de medias

Como análisis complementario podemos *verificar la magnitud del cambio*. El proceso que vamos a seguir es el siguiente:

1º A cada sujeto le calculamos una *puntuación individual de tendencia (de cambio individual).*

2º Calculamos la media y desviación típica de estas puntuaciones de *tendencia;*

3º Hacemos un contraste de medias comparando la media de la muestra con la media de una hipotética población cuya media fuera 0; se trata del contraste de la media de una muestra con la media de una población de cambio cero.

Ya hemos advertido antes que este procedimiento tiene una ventaja adicional: al disponer de una puntuación individual de *tendencia a progresar* (si la hipótesis es de tendencia lineal), podemos además comprobar *relaciones posibles entre esta tendencia y otras variables* que conozcamos de los sujetos (edad, sexo o cualquier otra).

Vamos a ver el procedimiento con los datos de la tabla 4.

1º Para calcular la *puntuación en tendencia* de cada sujeto, multiplicamos sus puntuaciones por los coeficientes λ oportunos; en este caso -1, 0 y +1 y sumamos a cada sujeto su nuevo total, que lo es de su tendencia al cambio progresivo. El nuevo cuadro de datos está en la tabla 8. En este caso se trata simplemente de restar la primera puntuación de la tercera (pero no sería lo mismo si tuviéramos más de tres ocasiones porque los valores de λ serían distintos).

2º De estas puntuaciones de *tendencia* calculamos la *media y la desviación típica* (también calculadas en la tabla 8).

1ª ocasión(-1)	2ª ocasión (0)	3ª ocasión (+1)	nuevo total
(10)(-1) = -10	(12)(0) = 0	(14)(+1) = 14	4
(9)(-1) = -9	(11)(0)= 0	(13)(+1) = 13	4
(5)(-1) = - 5	(3)(0) = 0	(7)(+1) = +7	2
(4)(-1) = -4	(2)(0) = 0	(6)(+1) = +6	2
		Media =	3
		Desviación =	1

Tabla 8

3º Ahora calculamos la t de Student con la fórmula habitual para comparar la media de una muestra con la de una población hipotética de media $\mu = 0$:

$$t = \frac{M - 0}{\sqrt{\frac{\sigma^2}{N-1}}} = \qquad [10] \qquad \frac{3-0}{\sqrt{\frac{1^2}{4-1}}} = 5.19, \text{ con grados de libertad} = 3,\ p < .05$$

Nuestra muestra (con una media en cambio = 3) no pertenece a una población con una media de *cero* en cambio. Los resultados de ambos métodos, análisis de varianza y t de Student, no dan idénticos resultados aunque suelen ser similares y suelen llevar a las mismas conclusiones[4].

3.3. 'Magnitud del cambio' (*tamaño del efecto*)

En este caso (con tres ocasiones el cambio equivale a la diferencia entre la 3º ocasión y la 1ª) podríamos además calcular el *tamaño del efecto* habitual: diferencia entre las dos medias (ocasiones 1ª y 3ª) dividida por la desviación típica combinada de las dos, o la de la 3ª ocasión (siguiendo a Glass, como si se tratara de un post-test).

En este ejemplo *la media en cambio* es 3 y la desviación del post-test es 3.535, por lo que el *tamaño del efecto* es 3/3.535 = .848; podemos valorar este cambio como grande.

[4] Las diferencias pueden verse en Rosenthal (1987:166).

3.4. Análisis correlacionales: relación entre cambio individual y otras variables

La ventaja de disponer de una puntuación individual de *tendencia* (lineal, que es quizás lo más común, o cuadrática) es que nos permite *comprobar relaciones* entre la tendencia (o cambio) individual y otras características de los sujetos. También podemos hacer lo mismo cuando las ocasiones son dos (contraste de medias entre *antes* y *después*).

El hecho de que la diferencia entre la media de la primera ocasión y la media de la última ocasión sea estadísticamente significativa quiere decir que el cambio en el grupo es superior a lo que podemos esperar por azar; al grupo *se ha movido*, ha evolucionado, pero no nos dice nada sobre cada sujeto en particular. Unos sujetos han podido cambiar más, otros menos, otros nada e incluso alguno ha podido cambiar en dirección opuesta a la de la mayoría. Si disponemos de otros datos de los sujetos podemos verificar la relación entre otras variables y el cambio. Este tipo de análisis puede aportar gran riqueza informativa a cualquier investigación.

Por ejemplo, en un estudio sobre la eficacia de un tratamiento de la dislexia (Benito, 1999) los sujetos fueron medidos en tres ocasiones distintas en diversas variables que sirven de diagnóstico de la dislexia. Los resultados muestran una tendencia estadísticamente significativa (análisis de varianza, los sujetos van mejorando de una vez a otra), y el cambio final (*tamaño del efecto*) es grande. Pero *además* se disponía de otros datos de los sujetos; calculando las correlaciones entre estos datos y el cambio (*mejora*) en dislexia, se observa que esta mejora en dislexia es independiente del nivel socio-económico de la familia o número de hermanos y en cambio está relacionada con variables tales (entre otras) como grado de asistencia de los padres a las sesiones de evaluación y asistencia regular de los niños a clase.

Naturalmente para poder hacer estos análisis hay que haber previsto qué información adicional se necesita o puede ser conveniente.

4. Referencias bibliográficas

BENITO PEREGRINA, MANUELA de (1999). *Aproximación al concepto de dislexia: un estudio sobre las características de los alumnos disléxicos en un entorno bilingüe*. Tesis doctoral. Madrid: Universidad Pontificia Comillas.

ESCOTET, MIGUEL A., (1980). *Diseño multivariado en psicología y educación.* Barcelona: Ceac.

GUILFORD, J. P. Y FRUCHTER, B., (1984). *Estadística aplicada a la psicología y la educación,* México: McGraw-Hill. [En inglés: *Fundamental Statistics in Psychology and Education,* 1973. New York: McGraw-Hill].

KIRK, ROGER E., (1995). *Experimental Design, Procedures for the Behavioral Sciences.* Boston: Brooks/Cole.

ROSENTHAL, ROBERT, (1987). *Judgment Studies, Design, analysis and meta-analysis.* Cambridge: Cambridge University Press.

TEJEDOR, FRANCISCO JAVIER, (1984). *Análisis de varianza aplicada a la investigación en pedagogía y psicología.* Madrid: Anaya

Anexo. Tabla de los pesos (λ) aplicables para verificar tendencias

En la tabla 9 están los pesos (λ) aplicables para verificar tendencias (lineares o cuadráticas)[5] cuando las muestras (columnas, muestras independientes o muestras relacionadas) están ordenadas (tablas más completas pueden verse en los autores citados).

tendencia	*número de columnas*					
	1	2	3	4	5	6
3 Linear	- 1	0	+ 1			
Cuadrática	+1	- 2	+ 1			
4 Linear	- 3	- 1	+ 1	+ 3		
Cuadrática	+1	- 1	- 1	+ 1		
5 Linear	- 2	- 1	0	+ 1	+ 2	
Cuadrática	+2	- 1	- 2	- 1	+ 2	
6 Linear	- 5	- 3	- 1	+ 1	+ 3	+ 5
Cuadrática	+5	- 1	- 4	- 4	- 1	+ 5

Tabla 9

[5] Estas tablas se encuentran en numerosos textos (por ejemplo Guilford y Fruchter; 1973; Escotet, 1980; Tejedor, 1984; Rosenthal, 1987; Kirk, 1995), y en Internet, por ejemplo (hasta 10 medidas o columnas) en LANE, DAVID M. HyperStat Online Statistics Textbook, http://davidmlane.com/hyperstat/index.html, nº 12. Las tendencias *cúbicas* (dos cambios de dirección) pueden comprobarse a partir de cuatro ocasiones, suelen ser de interés con menor frecuencia que las otras dos.