

# Análisis de medidas repetidas

Maria del Carmen Ruiz de Villa

*Departament d'Estadística i Investigació Operativa. Universitat de Barcelona. Barcelona. España.*

**En este trabajo se realiza una exposición de los métodos estadísticos que se utilizan más frecuentemente en el análisis de medidas repetidas, haciendo énfasis en el tipo de datos y las limitaciones de cada método. El objetivo principal consiste en dar una visión comparativa de los distintos métodos desde un punto de vista práctico en el entorno de aplicaciones médicas. Para cada una de las aproximaciones se plantea una breve introducción matemática, una discusión de las condiciones de aplicación que permiten tener una idea del tipo de diseños en los que se puede utilizar y, finalmente, el análisis de unos datos concretos con una breve discusión de resultados.**

*Palabras clave:* Medidas repetidas. Análisis longitudinal. MANOVA. Modelos mixtos.

### Repeated measures analysis

**In this chapter we present the most frequent statistical methods used on the repeated measures analysis, putting emphasis on the different types of data and the limitation of each method. The main objective consists in giving a comparison of the different methods from a practical point of view and with a medical perspective. For each approach we give a brief mathematical introduction, a discussion about the conditions for its application suggesting what kind of data are appropriated and, finally, it is illustrated by the study of the analysis of some data sets including a brief discussion of the results.**

*Key words:* Repeated measures. Longitudinal analysis. MANOVA. Mixed models.

### Introducción

En medicina clínica es frecuente diseñar ensayos en los que se dispone de varias observaciones del mismo sujeto y se pueden tener además varios grupos de sujetos. Una de las situaciones más frecuentes en estos ensayos consiste en el análisis de datos en los que se dispone de la evolución de alguna variable a lo largo del tiempo para cada individuo (peso de pacientes, crecimiento tumoral, etc.) en diferentes situaciones experimentales (diferentes tratamientos, diferentes dietas, distinguiendo por sexos, etc.) y el objetivo del estudio consiste en realizar la comparación de los grupos mediante la comparación de las curvas obtenidas. En otras situaciones las observaciones repetidas sobre un mismo individuo vienen ocasionadas por el hecho de que se sabe que la variabilidad entre los individuos que se utilizan para realizar un determinado ensayo es muy grande. En este caso, con una observación por individuo será difícil determinar, entre las diferencias entre grupos que se observan, qué componente es debido a los tratamientos administrados y qué componente es debido a la propia variabilidad de la

muestra analizada. En estas situaciones suele ser interesante obtener varias medidas del mismo individuo para poder tener una aproximación (en términos estadísticos nos referiremos a control) a esta variabilidad individual.

El concepto de medidas repetidas es un término que engloba las situaciones comentadas anteriormente. En términos generales, se utiliza en aquellas situaciones en las que la variable respuesta en cada unidad experimental se mide en múltiples ocasiones y, posiblemente, en condiciones experimentales distintas. Su ámbito de aplicación es variado, incluyendo la biología, medicina, educación, psicología, econometría, etc.

Como se ha indicado anteriormente, este tipo de análisis suelen plantearse en estudios que son longitudinales por naturaleza, es decir, cuando las medidas se realizan a lo largo del tiempo. En estos casos para cada individuo se dispone de una curva que indica el comportamiento de la variable obtenida a lo largo del tiempo y, por lo tanto, el interés se centra en el modelo matemático que expresa la respuesta como una función del tiempo. Esta función, en general, será no-lineal. En el mejor de los casos, podremos proponer una expresión determinada para esta función teniendo en cuenta las características del proceso que estamos observando. Sin embargo, en muchos casos deberemos considerar modelos más generales (lineal, cuadrático, etc.).

En determinados estudios experimentales, las medidas repetidas corresponderán a la asignación de distintos tratamientos a los sujetos del estudio. En estos casos a cada individuo se le administran todos los tratamientos que se desean estudiar dejando un período prudencial de tiempo entre la administración de dos tratamientos consecutivos para evitar una posible interacción entre ellos. En el análisis de este tipo de situaciones el valor del tiempo en que se realiza cada medida es irrelevante, y el objetivo se centrará en la comparación de los tratamientos teniendo en cuenta que se dispone de una medida de la variabilidad individual al disponer de varias medidas sobre el mismo sujeto. Un caso particular de este tipo de análisis son los diseños *cross-over*. A modo de resumen, la tabla 1 indica las ventajas y los inconvenientes de este tipo de análisis en su aplicación en estudios clínicos.

Independientemente del diseño experimental con medidas repetidas que se utilice en una situación concreta, una de las características particulares común a todos ellos es la existencia de una estructura de correlaciones muy particular entre los datos: las observaciones de un mismo individuo están correlacionadas, mientras que las observaciones de individuos distintos están incorrelacionadas. Este hecho plantea una serie de problemas desde el punto de vista de análisis estadístico cuya solución, teniendo además en cuenta el tipo de cuestiones a resolver, da lugar a distintas aproximaciones al problema. En este trabajo se van a indicar los métodos más habituales para abordar este tipo de problemas en función del diseño particular y de los objetivos del estudio, haciendo hincapié en la adecuación de cada método a cada situación concreta.

Correspondencia: Dra. M.C. Ruiz de Villa.  
Departament d'Estadística. Universitat de Barcelona.  
Facultat Biologia-Principal.  
Avda. Diagonal, 645.  
08028 Barcelona. España.  
Correo electrónico: mruiz\_de\_villa@ub.edu

TABLA 1

**Ventajas e inconvenientes en el análisis de medidas repetidas**

<b>Ventajas del análisis</b>
Permiten evaluar eficientemente los cambios entre unidades experimentales y, dentro de éstas, a lo largo del tiempo
Dada la correlación positiva que existe entre las observaciones de una misma unidad experimental, las comparaciones entre los tratamientos asignados a cada unidad experimental se harán con más precisión
<b>Desventajas del análisis</b>
Dada la correlación positiva que existe entre las observaciones de una misma unidad experimental, las comparaciones entre los tratamientos asignados a cada unidad experimental se harán con más precisión
Las comparaciones entre tratamientos asignados a grupos de individuos (con el mismo tratamiento a lo largo del tiempo) se realizan de forma menos eficiente
Usualmente aparece una heterogeneidad entre las medidas repetidas que se suele reflejar en una heterocedasticidad, es decir, una varianza distinta a lo largo del tiempo. Este problema es característico en las curvas de crecimiento en las que en el tiempo inicial las curvas de las unidades experimentales parten prácticamente de un mismo punto, mientras que a medida que avanza el tiempo se van dispersando cada vez más. La solución a este problema requiere una cierta habilidad para encontrar el modelo de heterocedasticidad adecuado

En el ámbito de la investigación médica, este tipo de técnicas proporciona una herramienta de análisis de datos muy interesante. Por ejemplo, se ha aplicado a la evaluación de curvas de glucemia<sup>1</sup> y al estudio de la concentración de fármacos en sangre<sup>2</sup>. El estudio de curvas de crecimiento es otra de las aplicaciones muy frecuentes<sup>3</sup>. Otros ejemplos de aplicación serían el estudio del efecto de determinados tratamientos en el grosor de la córnea<sup>4</sup> o la evaluación de un programa de vacunas de la hepatitis B<sup>5</sup>. Desde un punto de vista técnico, y dada la importancia creciente de estas técnicas en medicina, han aparecido distintas revisiones que, desde distintos puntos de vista, analizan los principales métodos estadísticos de forma descriptiva<sup>6</sup> o bien con una cierta base estadística<sup>7-10</sup>.

TABLA 2

**Conceptos relacionados en distintos tipos de diseños y técnicas de análisis de medidas repetidas**

<b>Datos longitudinales</b>
En los que se estudia la variación de la respuesta a lo largo del tiempo. En ellos, fundamentalmente, interesa conocer el modelo que relaciona la respuesta con el tiempo
<b>Medidas repetidas</b>
Se asignan repetidamente una serie de tratamientos a cada individuo, de forma que cada unidad experimental se considera control de sí misma. Este nombre ha pasado a designar de forma general los estudios en que se observa un individuo de forma repetida, independientemente de la forma del experimento
<b>Curvas de crecimiento</b>
La relación que se desea estudiar es la de la respuesta en función de la dosis o bien en función del tiempo siendo ésta creciente
<b>Datos agrupados (multilevel)</b>
Son usuales cuando se estudian poblaciones con una estructura jerárquica (ciudad, hospital, etc.). En estos estudios el objetivo es saber en qué medida las características de un determinado nivel afectan al valor de la variable respuesta
<b>Datos multivariantes</b>
Se habla de datos multivariantes cuando se dispone del valor de varias variables sobre una misma unidad experimental, independientemente de cómo son estas variables y del tipo de correlación que exista entre ellas. En este caso una observación multivariante de un individuo sería el conjunto de medidas repetidas obtenidas en el mismo
<b>Datos correlacionados</b>
Este nombre incluye todo tipo de datos en los que se observa una cierta estructura de correlación

A pesar de que existe ya abundante información al respecto, no se encuentra una exposición de las distintas aproximaciones al análisis estadístico en estos ensayos en la que se haga más énfasis en el tipo de resultados que se pueden obtener con un ejemplo numérico de cada uno de los métodos que se pueden aplicar. Esta revisión se ha planteado con el objetivo de suplir este vacío.

*Definiciones básicas*

De cara a una mejor comprensión de los métodos de análisis en los diseños de medidas repetidas es preciso definir algunos conceptos estadísticos que son característicos de estos métodos. Se suele referir como covariables a aquellas variables o factores que pueden afectar a la respuesta. Estas covariables pueden clasificarse en dos categorías:

- Internas (*inner, within*): varían en las observaciones de una misma unidad experimental. Por ejemplo, serían el tiempo, los tratamientos en un diseño *cross-over*, etc.
- Externas (*outer, between*): se mantienen constantes para todas las observaciones de una misma unidad experimental. Por ejemplo, el sexo, los tratamientos cuando se administran a grupos de unidades experimentales, etc.

Dependiendo del número de medidas repetidas distinguiremos entre:

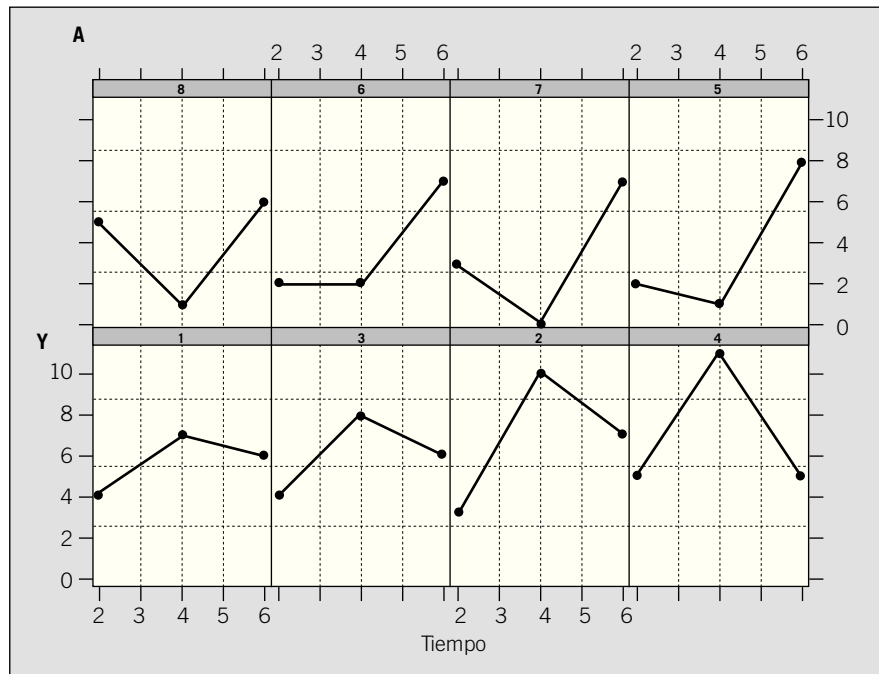
- Diseños equilibrados: con un número fijo de medidas por unidad experimental, tomadas a los mismos instantes de tiempo para todas ellas y sin valores faltantes.
- Diseños no equilibrados: con un número variable de medidas en cada unidad experimental, realizadas a distintos instantes de tiempo.

Debido a la multitud de aplicaciones que tiene esta metodología y a que existen diferentes aproximaciones estadísticas para analizar los datos, suele suceder que un mismo concepto aparezca en la bibliografía con distintos nombres. En particular, una primera aproximación al tema puede llevarnos a cierta confusión, tanto en lo que se refiere a la estructura de los datos y al diseño de los estudios como al mismo procedimiento estadístico de análisis de los mismos. Con el objetivo de facilitar la comprensión de los distintos conceptos, en la tabla 2 se definen los utilizados más frecuentemente. A lo largo de esta revisión, trataremos los diferentes métodos para el análisis de datos longitudinales. El estudio de los diseños *cross-sectional* queda fuera de los objetivos de este trabajo.

TABLA 3

**Resultado de un ensayo clínico donde se determina el estado de cada individuo a lo largo de tres instantes de tiempo. La asignación al grupo control y al grupo tratado se ha realizado al azar. Los datos son ficticios y se proponen a modo de ejemplo**

	Individuo	t1	t2	t3
Control	1	4	7	6
	2	3	10	7
	3	4	8	6
	4	5	11	5
Tratamiento	1	2	1	8
	2	2	2	7
	3	3	0	7
	4	5	1	6



Figs. 1a y b. Representación de los datos de la tabla 3. A: Por individuos. B: por grupos.

**Aproximación univariante**

Supongamos que queremos estudiar el efecto de un tratamiento sobre una determinada enfermedad, y que para ello realizamos un ensayo clínico controlado en el que comparamos un grupo de individuos que han recibido el tratamiento respecto a otro grupo control (tabla 3). Cada individuo se ha observado en 3 instantes de tiempo que supondremos iguales para todos los individuos. Una primera aproximación al análisis de estos datos consiste en realizar una representación gráfica. Ésta puede hacerse individualmente (fig. 1a) o bien con un gráfico para cada grupo (fig. 1b).

*Utilización del análisis de la varianza (ANOVA)*

En el ejemplo de la figura 1 encontramos una situación típica de medidas repetidas en las que se pueden resaltar algunas características:

- El comportamiento de ambos grupos a lo largo del tiempo no es igual. A pesar de que en ambos grupos la función se puede expresar mediante un polinomio de segundo grado, el hecho de que las curvas tengan una orientación opuesta indica que algunos coeficientes de los polinomios tendrán signo opuesto.
- El comportamiento de todos los individuos de un mismo grupo es similar, aunque se pueden observar diferencias entre ellos respecto al comportamiento medio del grupo.
- En algunos instantes de tiempo la variabilidad es mayor que en otros.

Una primera aproximación al problema consiste en plantear un diseño de análisis de la varianza que refleje de alguna forma la correlación que existe entre las observaciones de un mismo individuo. Un modelo adecuado sería:

$$y_{ijk} = \mu + \alpha_i + \pi_{j(i)} + \beta_k + \alpha\beta_{ik} + \beta\pi_{k(j)} + \epsilon_{ijk}$$

donde  $\alpha_i$  corresponde al efecto del grupo (con  $i = 1,2$  correspondiendo a control y tratamiento, respectivamente),  $\pi_{j(i)}$  al efecto del individuo  $j$  ( $j = 1,2,3,4$ ) del grupo  $i$  (indicamos

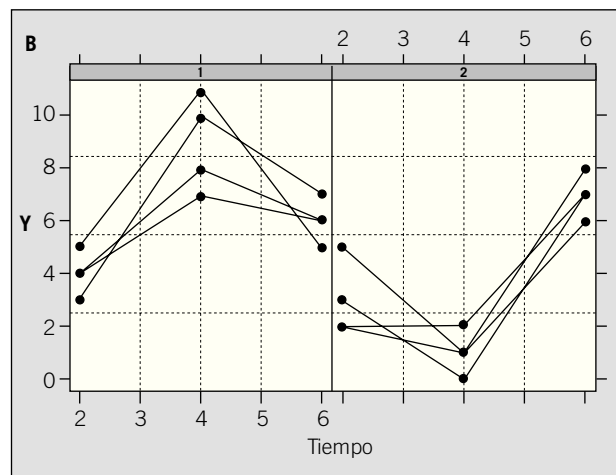


Fig. 1b.

entre paréntesis que cada individuo está exclusivamente –está anidado– en un grupo  $i$ ),  $\beta_k$  el efecto del tiempo ( $k = 1,2,3$ ) y los términos  $\alpha\beta_{ik}$  y  $\beta\pi_{k(j)}$  a las interacciones entre factores. Este tipo de diseño (conocido también como diseño *split-plot*) ha sido muy utilizado en el análisis de este tipo de datos a pesar de que son necesarios una serie de requisitos de difícil cumplimiento, sobre todo en el ámbito de los estudios clínicos. Estos requisitos son:

1. Homogeneidad de las matrices de covarianzas en cada grupo. Es decir la variabilidad de los datos en cada instante de tiempo tiene que ser la misma en todos los grupos y también la dependencia entre observaciones de un mismo individuo.
2. La matriz de covarianzas debe presentar la propiedad de simetría compuesta (*compound symmetry*), es decir la correlación entre observaciones alejadas en el tiempo, para un mismo individuo, ha de ser la misma independientemente de la distancia entre las observaciones.

TABLA 4

**Tabla ANOVA correspondiente al análisis de los datos de la tabla 3**

Efecto		Valor F	Valor de p
Grupo		64	0,000
Tiempo		10,8	0,002
Grupo x tiempo		26,8	0,000
	Efecto lineal	2,182	0,190
	Efecto cuadrático	56,889	0,000

3. El diseño debe ser equilibrado.

4. El número de individuos en cada grupo ha de ser el mismo.

La condición 2 es muy estricta, ya que requiere que exista la misma dependencia tanto entre dos observaciones seguidas en el tiempo como entre dos observaciones alejadas. En el caso particular de nuestro ejemplo se cumplen las condiciones para la aplicación del método univariante, por lo que procedemos a realizar el análisis correspondiente. La tabla 4 contiene los resultados del análisis de la varianza (tabla ANOVA). A partir de estos resultados podemos concluir que hay diferencias entre los grupos, entre los tiempos y además la interacción entre el tiempo y el grupo también es significativa. Esta interacción se interpreta como un comportamiento distinto, en cada grupo, de la evolución de la variable medida a lo largo del tiempo. Este comportamiento viene expresado en términos de una función matemática. En este caso se ha estudiado si podemos admitir una función lineal o una cuadrática. Tal como se ve en la tabla 4 es la componente cuadrática la que es significativa (como ya se intuía en la fig. 1).

#### Utilización incorrecta del análisis

Frente a unos datos del tipo del ejemplo anterior, se podría pensar que un análisis válido y sencillo sería realizar una comparación de los grupos en cada instante (por ejemplo, una t de Student en caso de 2 muestras). Esta solución se observa frecuentemente en trabajos de investigación médica con una estructura de datos similar a la de este ejemplo. Sin embargo, los resultados obtenidos mediante esta aproximación no se pueden considerar como válidos debido a las siguientes razones:

TABLA 5

**Estudio de la frecuencia cardíaca en 4 tratamientos. Cada individuo ha sido observado en siete instantes de tiempo distintos**

Tratamiento	Individuo	t1	t2	t3	t4	t5	t6	t7
1	1	73	74	74	78	79	81	79
1	2	78	77	76	79	84	88	86
1	3	73	73	74	78	81	81	79
1	4	75	72	73	74	79	82	80
1	5	75	74	73	78	82	84	82
2	6	72	74	76	80	81	83	81
2	7	77	77	78	81	85	86	84
2	8	72	74	78	80	84	86	84
2	9	77	80	78	83	85	86	84
2	10	74	75	75	80	84	85	83
3	11	79	80	81	83	83	83	82
3	12	73	76	74	81	79	82	81
3	13	68	68	74	75	76	78	77
3	14	75	77	80	81	85	89	88
3	15	72	74	81	79	82	83	82
4	16	78	82	83	87	91	92	87
4	17	78	81	81	85	89	92	87
4	18	78	80	81	82	87	87	82
4	19	71	75	76	80	85	85	80
4	20	80	83	86	86	90	90	85

– No se tiene en cuenta la correlación de los datos para un mismo individuo a lo largo del tiempo.

– El error asociado total sería mucho mayor que el que se obtiene cuando se realiza una sola prueba (es decir, estamos frente a un problema de comparaciones múltiples).

– La forma de las curvas en cada grupo puede ser distinta (como en nuestro ejemplo) y eso no queda reflejado en el análisis tiempo a tiempo.

– Normalmente el número de individuos utilizados en cada grupo es pequeño. Si se realiza un ANOVA, este problema no es grave ya que para realizar los contrastes se utiliza la información de todos los datos. Sin embargo, al realizar las comparaciones a cada instante de tiempo sólo se utilizan los datos de ese tiempo con lo que el análisis acostumbra tener una potencia (capacidad para encontrar diferencias entre grupos) muy baja.

#### Aproximación multivariante

En la mayoría de las situaciones experimentales relacionados con este tipo de diseño de datos no se cumplen las condiciones de aplicación del análisis anterior. En estos casos será necesario plantear nuevas aproximaciones que permitan resolver el problema de forma eficaz. Una solución consiste en utilizar las ventajas que ofrece el análisis multivariante de la varianza (MANOVA). Esta técnica no precisa de ningún requisito para la forma de la matriz de covarianzas (aunque sí es necesario que las matrices sean iguales entre grupos).

El MANOVA considera que todas las medidas repetidas de un mismo sujeto no son más que una observación multivariante y, por lo tanto, la comparación de grupos se realiza mediante la versión multivariante del análisis de la varianza. Este análisis será apropiado cuando se pretenda detectar diferencias entre grupos a partir del conjunto de medidas repetidas, pero no interesen tanto las medidas individuales a cada instante de tiempo.

Desde un punto de vista de estructura de datos, y de forma similar al modelo anterior, los instantes de tiempo en los que se realizan las observaciones deben ser iguales para todos los individuos y se deben tener observaciones completas. A modo de ejemplo, la tabla 5 contiene un conjunto de datos acerca de la frecuencia cardíaca (latidos por minuto) medidos en 20 varones asignados al azar a cada uno de 4 tratamientos (5 personas por tratamiento) y medidos cada 5 minutos durante 7 intervalos de tiempo. El conjunto de estas medidas no cumple las condiciones requeridas en el análisis univariante comentadas anteriormente. Por lo tanto aplicaremos el método multivariante.

#### Análisis de perfiles

En el MANOVA, la forma explícita de realizar la comparación de grupos es mediante el llamado análisis de perfiles. Este análisis estudia la forma de las curvas asociadas a cada sujeto permitiendo profundizar en las causas de las diferencias observadas en las observaciones y obtener resultados similares al estudio univariante. Así se podrá estudiar:

– La comparación de grupos, analizando la coincidencia de los perfiles de las curvas de los individuos de cada grupo.

– El efecto del tiempo, es decir, la equivalencia de las respuestas a partir de la constancia de los perfiles.

– La interacción entre el tiempo y el grupo, a partir del estudio del paralelismo de los perfiles.

La tabla 6 contiene los resultados de aplicar el análisis MANOVA a los datos del ejemplo de la tabla 5. Para cada uno de los

TABLA 6

**Tabla MANOVA resultado del análisis de los datos de la tabla 5. Se indican los valores de las distintas aproximaciones al test**

Fuente	Suma de cuadrados	gl	Media cuadrática	F	Valor de p
Trat	615,736	3	205,245	4,570	0,017
Efecto	Valor	F	Gl hipótesis	Gl del error	Valor de p
Tiempo					
Traza de Pillai	0,995	470,983	5,000	12,000	0,000
Lambda de Wilks	0,005	470,983	5,000	12,000	0,000
Traza de Hotelling	196,243	470,983	5,000	12,000	0,000
Raíz mayor de Roy	196,243	470,983	5,000	12,000	0,000
Tiempo x trat					
Traza de Pillai	1,910	4,908	15,000	42,000	0,000
Lambda de Wilks	0,003	15,737	15,000	33,528	0,000
Traza de Hotelling	78,109	55,544	15,000	32,000	0,000
Raíz mayor de Roy	75,678	211,897	5,000	14,000	0,000

factores analizados aparecen 4 valores de test con su nivel de significación, y se observa que hay diferencias entre tratamientos, entre tiempos y en la interacción. De estos resultados cabe destacar el análisis de la interacción. Como la interacción entre el tiempo y el tratamiento es significativa, podemos concluir que los perfiles no son paralelos y, por lo tanto, el comportamiento de los individuos de los 4 grupos es distinto.

### Modelos mixtos

Los 2 métodos anteriores no tienen en cuenta alguna de las características típicas de los datos con medidas repetidas. Estas características se ponían de manifiesto en el ejemplo de la figura 1 y se presentan frecuentemente en medicina. Así, por un lado se observa que existe una curva muy clara en la evolución de las observaciones de ambos grupos a lo largo del tiempo (en el ejemplo comentado, los resultados del gráfico sugieren un polinomio de segundo grado), siendo del mismo tipo para ambos grupos, pero seguramente con coeficientes distintos (la orientación del polinomio es distinta en cada grupo). Por otro lado, dentro de cada grupo se observa una gran variabilidad en las curvas individuales, lo que indica que un modelo que no tenga en cuenta este problema tendrá una variabilidad experimental muy grande, con lo cual no resultará eficaz para realizar el análisis. La aproximación a partir de los modelos mixtos es idónea para hacer frente a estos problemas ya que permite, por un lado, establecer la expresión funcional que explica la evolución de la medida observada a lo largo del tiempo y, por el otro, que los parámetros de esta función puedan variar de sujeto a sujeto. Así, podemos tener modelos lineales, que suponen que la relación se puede expresar mediante la ecuación de una recta o de un polinomio, o bien modelos no lineales que incluyen cualquier otro tipo de función. La decisión sobre el tipo de modelo puede ser un inconveniente en la utilización de estos modelos, ya que no siempre es fácil decidir cuál es la función que mejor se ajusta a nuestros datos. A pesar de este inconveniente, los modelos mixtos presentan muchas ventajas frente a los modelos anteriores por varias razones:

– Se basan en plantear un modelo para un individuo típico,  $i$ , al que se han realizado  $p_i$  observaciones a instantes de tiempo  $t_{i1}, t_{i2}, \dots, t_{ip_i}$ . En este modelo no es necesario tener los mismos instantes de tiempo para cada individuo ni observaciones completas. Tampoco es necesario que los grupos tengan el mismo número de unidades experimentales.

– El modelo individual, con parámetros propios, consiste en una función (lineal o no lineal, aunque en nuestro

caso, para simplificar, consideraremos que es lineal) cuyos coeficientes se expresan como una suma de 2 componentes: una componente poblacional, fija para todos los individuos y que representa el valor medio del parámetro dentro del grupo, y una componente aleatoria, distinta para cada individuo, que expresa la parte del parámetro que es propia de cada sujeto. Con ello los modelos mixtos recogen aquellas situaciones en las que los perfiles de los individuos de un mismo grupo puedan diferir entre sí.

– Las matrices de covarianzas no tienen por qué tener una estructura concreta y además permiten dar un modelo para explicar la diferente variabilidad de los datos a lo largo del tiempo y para identificar la correlación entre las observaciones a lo largo del tiempo.

### Modelo

En nuestro ejemplo, a partir de los gráficos de cada grupo (fig. 1), se puede concluir que un modelo apropiado podría ser un polinomio de segundo grado. El modelo mixto sería:

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + (\beta_2 + b_{2i})t_{ij}^2 + \varepsilon_{ij}$$

en el que  $y_{ij}$  es la respuesta del individuo  $i$  a tiempo  $t_{ij}$ ;  $\beta_0$ ,  $\beta_1$  y  $\beta_2$  corresponden a los parámetros poblacionales y  $b_{0i}$ ,  $b_{1i}$  y  $b_{2i}$  a la componente aleatoria de los mismos. El efecto del grupo se incorpora como un elemento más del modelo.

Desde un punto de vista teórico, se supone que  $b_{0i}$ ,  $b_{1i}$  y  $b_{2i}$  son variables aleatorias de media 0 y matriz de covarianzas desconocida  $\Psi$ , y que los errores siguen una ley normal con una matriz de covarianzas  $\Sigma_i$ , con una estructura similar para todos los individuos diferenciándose solamente en la dimensión pues, tal como hemos indicado, cada individuo puede tener un número distinto de observaciones. En esta última matriz se recogen las varianzas a cada instante de tiempo y las covarianzas (correlaciones) entre los tiempos. Aparte de esta matriz, el hecho de que todas las observaciones de un individuo compartan los mismos valores de  $b_{0i}$ ,  $b_{1i}$  y  $b_{2i}$  es otra forma de expresar la correlación de los datos.

De forma general se puede expresar el modelo mixto como:

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i$$

siendo  $y_i$  el vector de  $p_i$  observaciones repetidas del individuo  $i$ ,  $X_i$  la matriz de diseño de los efectos fijos,  $Z_i$  la matriz de diseño de los efectos aleatorios,  $b_i$  el vector de efectos aleatorios  $N(0, \Psi)$  y  $\varepsilon_i$  los residuos,  $N(0, \Sigma_i)$  siendo estos dos últimos independientes.

*Análisis estadístico*

El análisis estadístico de los modelos mixtos se basa en deducir y estimar el modelo que, con un mínimo número de parámetros, se ajusta a los datos experimentales. Para esto se sigue una serie de pasos sucesivos:

1. Se estima el modelo propuesto considerando que:  
 – Todos los parámetros tienen una componente fija y aleatoria.  
 – La matriz  $\Sigma_i$  tiene la estructura más sencilla posible, consistente en un valor de varianza igual a todos los tiempos ( $\sigma^2$ ) e incorrelación de los datos. Es decir  $\Sigma_i = \sigma^2 I$ .  
 Con este modelo los parámetros a estimar son: la componente fija de los parámetros ( $\beta_0, \beta_1$  y  $\beta_2$  en nuestro ejemplo), la estimación de los elementos de la matriz  $\Psi$  (con lo que será posible conocer la estimación de las componentes aleatorias) y la estimación de  $\sigma^2$ .

2. A partir de los resultados obtenidos se reajusta el modelo siguiendo los siguientes criterios:  
 – Si la varianza de la componente aleatoria de algún parámetro está cercana a cero, se elimina la componente aleatoria consiguiendo un modelo con menos términos y, por lo tanto, más fácil de interpretar.  
 – Mediante el análisis de un gráfico de residuos del modelo se cambia la matriz  $\Sigma_i$  dándole una estructura concreta. Este paso puede presentar alguna dificultad y requiere una cierta experiencia.

3. Cada vez que se considera un nuevo modelo, a partir de la modificación de un modelo previo, es necesario comprobar la validez de la modificación mediante una prueba estadística que certifique la mejora del ajuste con el nuevo modelo. En caso de que el test estadístico indique una mejora significativa, se aceptará la modificación. En este proceso iterativo de mejora del modelo, un punto importante es el estudio de la influencia de las covariables que identifican a los grupos de individuos y que se introducen en el modelo como un factor más. En caso de que el factor sea significativo, podremos confirmar la diferencia entre las curvas experimentales de los distintos grupos.

*Ejemplo de modelo mixto*

A modo de ilustración, consideraremos un ejemplo clásico analizado por Little y Rubin<sup>11</sup> y que corresponde al seguimiento del proceso de crecimiento en 11 niñas y 16 niños. A cada niño se le midió la distancia desde el centro de la pituitaria a la fisura maxilar, a las edades de 8, 10, 12 y 14 años. En este caso, se considera que la relación entre la medida y el tiempo es lineal. Los análisis han sido realizados mediante el paquete estadístico S-Plus<sup>12</sup>. La figura 2 muestra, para cada sexo, la adecuación de un modelo lineal para explicar la relación entre la distancia y la edad. También se observa la presencia de algún niño con un comportamiento errático y que podría ser un candidato a ser excluido, aunque nosotros lo dejaremos en el análisis ya que la exclusión de observaciones debe hacerse en colaboración con el experimentador.

A continuación se detallan los pasos seguidos para obtener el modelo definitivo con los resultados parciales que se han ido obteniendo y que justifican el proceso que se ha seguido. A lo largo del análisis se van a ir planteando distintos modelos que se estimarán y que nos permitirán determinar si se cumplen las hipótesis de igualdad de grupos. El primer ajuste realizado ha sido a un modelo mixto completo, es decir suponiendo que todos los parámetros del modelo tienen parte fija y aleatoria, con la matriz de varianzas y covarianzas de los efectos aleatorios sin una estructura particular que permita simplificar el modelo y con la matriz de varianzas y covarianzas de los errores igual a  $\sigma^2 I$ . Este primer modelo (que llamaremos modelo I) es, por lo tanto:

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})x_{ij} + \epsilon_{ij}$$

donde  $y_{ij}$  es la distancia del niño  $i$  medida a la edad  $j$ ,  $x_{ij}$  la edad  $j$  del niño  $i$ . La estimación de los parámetros que caracterizan este modelo son:

1. Para los efectos aleatorios, la estimación de las varianzas de los 2 parámetros y de la correlaciones entre ambos es suficiente para conocer la matriz  $\Psi$ :

$$\sigma^2(b_{0i}) = 5,415, \sigma^2(b_{1i}) = 0,0513, \text{corr}(b_{0i}, b_{1i}) = -0,609$$

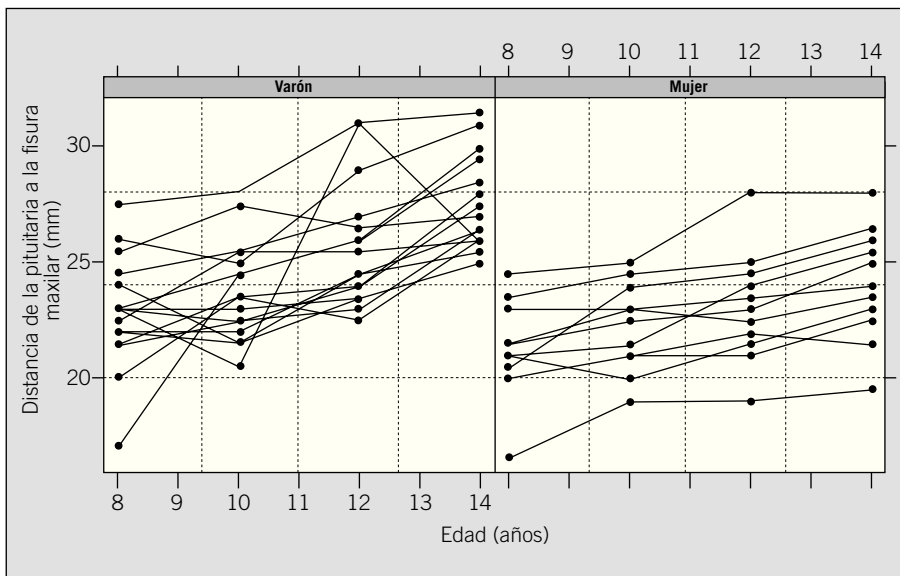


Fig. 2. Representación por sexos de los datos del ejemplo de la tabla 5.

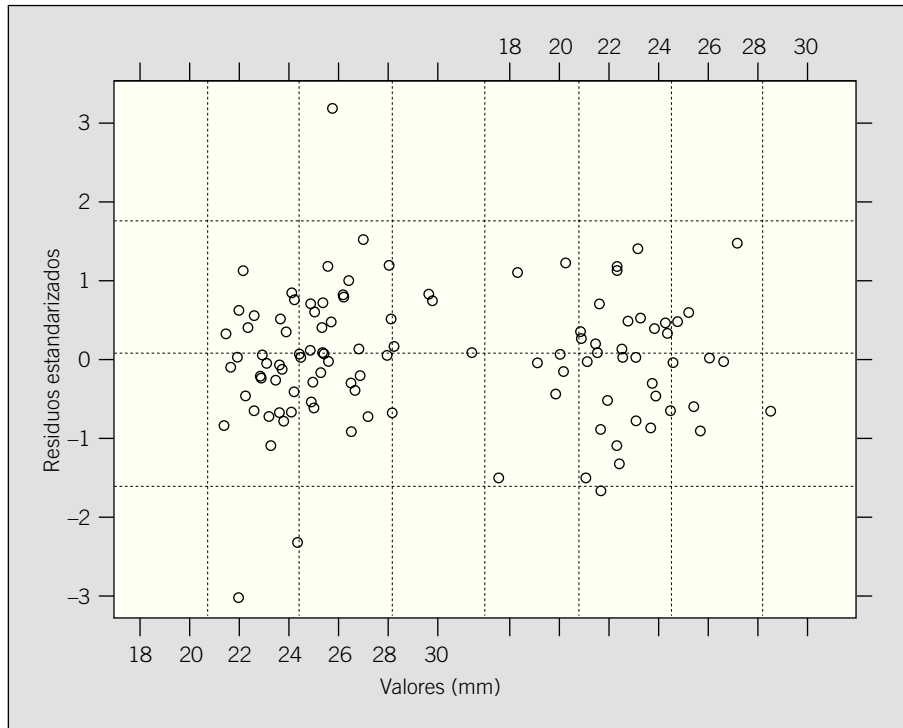


Fig. 3. Gráfico de residuos de modelo I correspondiente al análisis mediante la técnica de modelos mixtos (véase texto).

2. El valor de  $\sigma^2$  es 1,716.

3. Para los parámetros fijos:  $\beta_0 = 16,7611$ ,  $\beta_1 = 0,662$ , ambos significativos. El hecho de que  $\beta_1$  sea significativo indica que hay diferencias a lo largo de la edad.

El gráfico de residuos (fig. 3) muestra una mayor variabilidad en los niños que en las niñas, con lo que se sugiere una estructura para la matriz  $\Sigma_i$  que incluya esta heterocedasticidad. El modelo propuesto consiste en una matriz en la que la varianza de los residuos toma dos valores: uno para las observaciones de los niños y otro para la de las niñas. A continuación se realiza la estimación de los parámetros anteriores:  $-\sigma^2$  (niños) = 2,707 y  $\sigma^2$  (niñas) = 0,4418, y se ajusta un nuevo modelo (modelo II) corrigiendo el anterior a partir de las estimaciones de las varianzas obtenidas.

La comparación entre ambos modelos es significativa con un mejor ajuste para el segundo. El gráfico de residuos (que no se indica) muestra que en el modelo corregido los residuos ya no presentan heterocedasticidad. Como no se detecta la presencia de ninguna estructura de correlaciones para los residuos, se decide no analizar con más detalle la estructura de la matriz  $\Sigma_i$ .

Debido a que en la figura 2 se observaba que las rectas de los niños parecían distintas de las de las niñas, se continúa el estudio incluyendo la covariable Sex en el modelo inicial. La manera de incluirla consiste en definir el nuevo modelo (modelo III) con dos términos más, uno que recoge la influencia del sexo en las curvas y otro que recoge la interacción entre la edad y el sexo:

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})x_{ij} + \beta_2 \text{Sex} + \beta_3 \text{Sex} \cdot x_{ij} + \varepsilon_{ij}$$

El análisis de la significación de los dos términos añadidos es equivalente al estudio del efecto grupo y de la interacción de las aproximaciones univariante y multivariante descritas previamente. Este nuevo modelo ajusta significativamente mejor que el anterior (el valor del estadístico de razón de ve-

rosimilitud es 11,75 con un valor de p de 0,0028). Como resultado del ajuste la estimación de los parámetros fijos de este nuevo modelo es:

$\beta_0 = 16,857$ ,  $\beta_1 = 0,516$ ,  $\beta_2 = 0,632$  y  $\beta_3 = -0,152$ , siendo  $\beta_2$  no significativo.

Este hecho sugiere suprimir este término del modelo y ajustar un cuarto modelo (modelo IV). El nuevo modelo presenta un ajuste equivalente al tercero (el test de razón de verosimilitud es 0,5337 con un valor de p de 0,5021). Como que es más sencillo que el modelo III, ya que tiene un parámetro menos, se decide escoger este último como válido. Al haber excluido el término  $\beta_2 \text{Sex}$ , se deduce que las rectas de los niños y las niñas tienen aproximadamente la misma ordenada en origen; es decir, no hay diferencias al inicio del ensayo entre ambos grupos, aunque la existencia de un término de interacción  $\text{Sex} \cdot x_{ij}$  indica que la evolución de ambos grupos a lo largo de las distintas edades ha sido distinta.

El siguiente paso consiste en analizar si la parte del modelo propuesta para los efectos aleatorios puede ser simplificada. Recordemos que esto implica analizar con detalle la estructura de la matriz  $\Psi$ . Debido a que en los resultados del primer modelo se observaba una correlación baja (-0,609) vamos a considerar el caso en que sea nula. Para ello se ajusta un quinto modelo (modelo V) bajo esta suposición. La comparación con el anterior mediante el test de razón de verosimilitud da un estadístico de 0,4505 con un valor de p de 0,5021. Estos resultados muestran que no hay diferencias significativas entre ambos modelos, siendo el último más simple y, por lo tanto, el modelo definitivo escogido. Por lo tanto, el modelo que se ajusta mejor a las observaciones de los niños y las niñas es:

$$y_{ij} = (16,857 + b_{0i}) + (0,516 + b_{1i})x_{ij} - 0,152 \text{Sex} \cdot x_{ij} + \varepsilon_{ij}$$

$$\sigma^2(b_{0i}) = 5,415, \sigma^2(b_{1i}) = 0,0513, \text{corr}(b_{0i}, b_{1i}) = 0$$

La utilización de los modelos mixtos aporta un buen conocimiento del tipo de estructura estadística subyacente a los datos aunque, tal como se observa en este ejemplo, su aplicación requiere conocer adecuadamente el método estadístico y tener una cierta experiencia en su aplicación.

### Conclusiones

A lo largo de este trabajo, y mediante la discusión de los ejemplos, se ha intentado dar una visión comparativa de la forma de abordar el análisis de medidas repetidas en función del tipo de datos disponibles. Tal como se ha indicado, la aproximación univariante tiene como principal característica su facilidad de aplicación al poderse utilizar un software estándar para el análisis y, sobre todo, la facilidad de interpretación por la simplicidad de la técnica empleada. El principal inconveniente reside en la restricción que supone en la matriz de covarianzas, lo que delimita seriamente su campo de aplicación a casos muy concretos. Como primera alternativa, el método multivariante no tiene tantas restricciones en cuanto a la matriz de covarianzas, pero requiere que los datos sean equilibrados y completos, aspecto difícil de conseguir en los estudios médicos donde el seguimiento de los pacientes se realiza mediante visitas que no siempre se realizan en la fechas programadas. En este caso también se dispone de software estándar para realizar los análisis aunque la interpretación de los resultados ya no sea tan intuitiva como en el caso anterior.

Los modelos mixtos aportan una potencia de análisis que no permitan los métodos anteriormente indicados. Esta mejora en la capacidad de obtener el máximo de información de los datos va, no obstante, unida a una complejidad del método que determina que un usuario con conocimientos bási-

cos de estadística no pueda realizar el análisis de forma eficiente. En este caso es necesario acudir a un especialista para que, utilizando un software específico, pueda obtener el mejor modelo y nos ayude a explicar el comportamiento de los datos obtenidos.

### REFERENCIAS BIBLIOGRÁFICAS

1. Sievenpipe JL, Arnason JT, Leiter LA, Vuskan V. Variable effects of American ginseng: a batch of American ginseng (*Panax quinquefolius* L.) with a depressed ginsenoside profile does not affect postprandial glycaemia. *Eur J Clin Nutr* 2003;57:243-8.
2. Kinon BJ, Stauffer VL, McGuire HC, Kaiser CJ, Dickson RA, Kennedy JS. The effects of antipsychotic drug treatment on prolactin concentration in elderly patients. *J Am Med Dir Assoc* 2003;4:189-94.
3. Hanna SE, Law MC, Rosebaum PL, King GA, Walter SD, Russell DJ. Development of hand function among children with cerebral palsy: growth curve analysis for ages 16 to 70 months. *Dev Med Child Neurol* 2003;45:448-55.
4. Pall BR, Barr JT, Szczotka LB, Mitchell GL, Gleason W. Corneal thickness results in the Menicon Z 30-day continuous wear and ACUVUE 7 day extended wear contact lens clinical trial. *Eye Contact Lens* 2003;29:10-3.
5. Renard D, Bruckers L, Molenberghs G, Vellinga A, Van Damme P. Repeated measures models to evaluate a hepatitis B vaccination programme. *Stat Med* 2001;20:951-63.
6. Hennen J. Statistical methods for longitudinal research on bipolar disorders. *Bipolar Disord* 2003;5:156-68.
7. Goldstein H, Browne W, Rasbash J. Multilevel modelling of medical data. *Statistics in Med* 2002;21:3291-315.
8. Sithole JS, Jones PW. Repeated measures models for prescribing change. *Stat Med* 2002;21:571-87.
9. Park T, Lee YJ. Covariance models for nested repeated measures data: analysis of ovarian steroid secretion data. *Stat Med* 2002;21:143-64.
10. Littell RC, Pendergast J, Natarajan R. Modelling covariance structure in the analysis of repeated measures data. *Stat Med* 2000;19:1793-819.
11. Little RAJ, Rubin DB. *Statistical analysis with missing data*. New York: John Wiley, 1997.
12. Pinheiro J, Bates D. *Mixed models in S and S-Plus*. Statistics and computing series. New York: Springer Verlag, 2000.