



Los modelos multinivel o la importancia de la jerarquía

E. Sánchez-Cantalejo / R. Ocaña-Riola
Escuela Andaluza de Salud Pública (Granada)

Correspondencia: Emilio Sánchez-Cantalejo Ramírez. Escuela Andaluza de Salud Pública. Campus Universitario de Cartuja, Apdo. de Correos 2070. 18080 Granada.
E-mail:esc@easp.es

Recibido: 21 de julio de 1998
Aceptado: 24 de diciembre de 1998

(Multilevel models or the importance of ranking)

Resumen

Una parte importante de la investigación sanitaria proporciona bases de datos en las que se puede establecer alguna estructura jerárquica. Así, los individuos estudiados, unidades muestrales de nivel 1, pueden pertenecer a grupos o unidades mayores, por ejemplo, la zona geográfica donde viven, el médico que los atiende, el hospital que los asiste, etc., las unidades de nivel 2. La posible mayor homogeneidad entre individuos de un mismo grupo respecto a individuos de distintos grupos invalidaría la hipótesis de independencia necesaria para poder utilizar los modelos tradicionales de regresión. Los modelos multinivel, también denominados modelos jerárquicos, permiten solventar esta dificultad al distinguir los distintos niveles jerárquicos de las predictoras, separando la variabilidad de los individuos objeto de estudio de la de los grupos a los que pertenecen. Aunque estos modelos se desarrollaron en la pasada década, especialmente en el campo de la educación, en los últimos años se ha puesto en evidencia su aplicabilidad en la investigación sanitaria. Este trabajo describe algunos modelos multinivel sencillos, discute sus ventajas sobre los métodos clásicos y presenta algunos ejemplos de aplicación tanto en la investigación epidemiológica como en la de servicios sanitarios.

Palabras clave: Análisis Multinivel; Modelos Estadísticos; Análisis de Regresión; Análisis Multivariante; Análisis Estadístico.

Summary

Many researchers in Public Health have data bases with a hierarchical structure. The studied patients (level 1) can be nested in groups, i.e., district, doctor, hospital, etc. (level 2). It is possible that patients in the same group be similar, so traditional regression models can not be used because the hypothesis of independent observations is not satisfied. A Multilevel Analysis, using hierarchical models, can be a solution for this problem; these models take into account the distribution of the data at different levels to estimate two types of variability: one due to individuals in the study and another due to the groups in which patients are nested. These types of models were applied in education in the last decade, however they have been recently applied in Health Research. This paper is a review about multilevel analysis. A discussion about hierarchical models versus traditional regression models is presented and some applications in Epidemiology and Health Research are showed.

While the health of manual men and women was almost always poorer than that of non-manual, it is clear that types of living area do make a difference.

Blaxter M. *Health and Lifestyles*. 1990.

abundante investigación empírica sobre la interacción entre características individuales y las variables que describen distintos contextos.

Por otra parte, en las dos últimas décadas ha habido un interés creciente en la evaluación de la calidad del producto tanto en instituciones privadas como públicas^{1,2}. Esto es especialmente cierto en el campo de la educación, sobre todo en Gran Bretaña, donde se ha investigado con el fin de evaluar la influencia que, aparte de las características individuales de los alumnos, tienen tanto los profesores como los centros escolares en el progreso de los estudiantes. Así, ¿progresan de la misma forma los alumnos, independientemente de los profesores que los enseñan?, ¿el rendimiento de los alumnos se ve afectado por ciertas características de los colegios: público/privado, religioso/laico, etc.?

Introducción

Durante mucho tiempo, los científicos sociales se han interesado por las relaciones entre el individuo y el medio social en el que éste se desarrolla, con el objetivo de evaluar la influencia del contexto sobre el comportamiento individual. La hipótesis global es que el comportamiento de los individuos se ve influenciado por su contexto; esto ha acarreado

En la investigación sanitaria tienen también larga tradición los estudios acerca de la variabilidad, tanto de indicadores de salud^{3,4,5} como de consumo de recursos^{6,7,8}, entre los individuos de distintos grupos o zonas geográficas, de los que se derivan relaciones más o menos evidentes entre la salud de los individuos y la zona donde habitan o entre el tratamiento recibido por los pacientes y las características del médico y/o servicio sanitario donde son atendidos.

En este tipo de estudios podríamos disponer de información tanto acerca de los individuos como de sus contextos, es decir, de información a nivel micro y a nivel macro. Por tanto, podemos hablar de distintas jerarquías de la información disponible: por una parte, el nivel micro o individual, el nivel 1 y, por otra, el nivel 2 o nivel macro, el contexto o grupo al que pertenece el individuo.

El hecho de establecer esta jerarquía entre las distintas variables, es decir, el reconocimiento de que los distintos individuos pertenecen a distintos grupos no es gratuito pues tiene consecuencias importantes a la hora de analizar los datos. No cuesta reconocer que, en general, los individuos pertenecientes a un mismo contexto tenderán a ser más similares en su comportamiento entre sí que respecto a los que pertenezcan a distintos contextos. Así, las personas que viven en la misma área sanitaria podrían tener hábitos de vida más parecidos entre sí que respecto a personas de otra área distinta por el hecho de haber distintas culturas y/o políticas de promoción de salud en las respectivas áreas; el hecho de que una mujer sea sometida a una histerectomía puede depender de la práctica médica más o menos conservadora del servicio donde es atendida. Esta similitud entre los individuos dentro de los grupos establece una estructura de correlación intracontextual que impide el cumplimiento de la hipótesis de independencia sobre la que están basados los modelos de regresión tradicionales e invalida por tanto sus métodos de estimación, lo que se traduce en estimaciones incorrectas de los errores estándar. En los últimos años ha habido un considerable esfuerzo, en especial en la investigación educativa, para adaptar esta estructura jerárquica de los datos al marco de los modelos lineales generalizados; el resultado han sido los llamados modelos multinivel o modelos jerárquicos.

Es un hecho frecuente que técnicas estadísticas desarrolladas para resolver distintos problemas en ámbitos específicos del saber humano sean de gran utilidad para el estudio de otros problemas en el campo de la investigación sanitaria, si bien la transferencia tecnológica al campo de la salud no sea tan rápida como sería deseable⁹. Este trabajo trata de presentar a los investigadores, tanto en el campo de la salud pública como en el de los servicios sanitarios, estos nuevos modelos.

El análisis tradicional y sus consecuencias

Para hacer más comprensible la exposición posterior consideremos la situación concreta de un gestor sanitario interesado en conocer la relación entre el grado de satisfacción de sus pacientes y la experiencia de sus profesionales, medida ésta en número de años como profesional. Como quiera que el grado de satisfacción, que consideraremos como variable continua, también va a depender de algunas características del paciente, habrá que tenerlas en cuenta; por motivos de sencillez vamos a considerar sólo una característica del paciente, el grado de salud percibida y representaremos por Y a la satisfacción y por X a la salud percibida de los pacientes; Z indicará la experiencia de los profesionales.

Una posible estrategia de análisis sería no distinguir entre los distintos médicos, es decir, que considerar al conjunto de pacientes como un todo y proponer el modelo lineal

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

donde y_i es el grado de satisfacción del paciente i y x_i es la salud percibida de ese mismo paciente.

El coeficiente β_1 mide el cambio promedio en satisfacción de los pacientes por unidad de aumento de la salud percibida; si consideramos a esta variable centrada, es decir, si restamos a cada valor de la salud percibida su valor medio en todos los pacientes, el coeficiente β_0 se puede interpretar como el grado medio de satisfacción de un paciente con una salud percibida media. En el modelo de regresión clásico aparece un solo término de error, el e_i , que describe la variabilidad de la satisfacción de los pacientes de todos los médicos.

Si, como se acaba de hacer, se elige trabajar sólo a nivel del paciente ignorando las características del contexto se puede caer en la falacia atomista, proponiendo las mismas asociaciones encontradas a nivel individual como relaciones a nivel contextual. Si no se tiene en cuenta el contexto, en este caso el médico que atiende al paciente, el modelo lineal entre la satisfacción y la salud percibida por los pacientes puede representarse mediante una sola recta, la correspondiente a la estimación del modelo anterior.

Para distinguir cómo las distintas experiencias de los profesionales afectan a esa relación se podría considerar un modelo en el que se incluyera la experiencia del médico, la variable Z por lo que el modelo anterior quedaría de la forma

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_j + e_{ij}$$

donde ahora y_{ij} es la satisfacción del paciente i atendido por el médico j y z_j es la experiencia de tal médico;

β_2 es el cambio promedio en satisfacción entre dos pacientes con la misma salud percibida pero atendidos por dos médicos que se diferencian en un año de experiencia. Obsérvese que la representación gráfica de este nuevo modelo más general sería un conjunto de rectas ligando la satisfacción con la salud percibida, una recta para cada uno de los médicos, pero todas las rectas paralelas, pues la pendiente o efecto de la salud percibida sobre la satisfacción es β_1 , que no depende de j .

Si introducimos un término de interacción entre la experiencia del profesional y la salud del paciente, posibilitaremos que el coeficiente de la salud percibida dependa de la experiencia del profesional; en efecto, si añadimos un término de interacción $\beta_3 x_{ij} z_j$ modelo anterior se puede escribir de la forma

$$y_{ij} = \beta_0 + (\beta_1 + \beta_3 z_j) x_{ij} + \beta_2 z_j + e_{ij}$$

y ahora el coeficiente de la salud percibida $\beta_1 + \beta_3 z_j$ sí depende de la experiencia del médico. La representación gráfica de este nuevo modelo sería un conjunto de rectas, una para cada médico, pero sin la restricción de que deban ser paralelas, pues la pendiente de cada una de ellas ahora depende de la experiencia del profesional que atiende al paciente.

Ninguno de los modelos propuestos hasta ahora distingue la estructura jerárquica, ya que mientras la salud percibida pertenece al nivel 1, el paciente, la experiencia del profesional que lo atiende pertenece al contexto, es decir, el médico o nivel 2. Ante esta dificultad, los investigadores han utilizado en ocasiones la estrategia de analizar los datos a nivel contextual; las variables individuales se agregan a nivel de contexto, lo que para nuestro ejemplo equivaldría a estudiar la relación entre la satisfacción media de los enfermos atendidos por los distintos médicos en función de la salud percibida media de los enfermos atendidos y la experiencia del profesional que los atiende. Obsérvese que ahora no se habla de salud y satisfacción de los pacientes sino de la media de estas variables en los distintos grupos de pacientes que atienden los distintos médicos; por tanto las tres variables pertenecen al nivel contextual y no hay jerarquía entre ellas. Sin embargo, se corren grandes riesgos si de este tipo de análisis, llamado ecológico, se pretenden deducir relaciones para los individuos; prueba de ello es la llamada falacia ecológica^{10,11}. En efecto, los resultados contextuales no reproducen necesariamente los del nivel individual.

Aparte de consideraciones sobre eficiencia estadística, los modelos anteriores consideran a los efectos de las variables de contexto como parámetros fijos lo que hace que las inferencias sean exclusivas para los contextos, que en nuestro ejemplo serían los médicos, concretos muestreados y no para la población de profesionales de la que los estudiados proceden.

Modelos multinivel

Con el ánimo de superar esta dificultad, vamos a considerar el modelo

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}$$

donde el término β_{0j} es el grado de satisfacción promedio de los pacientes atendidos por el médico j y con un nivel medio de salud percibida, supuesta ésta centrada. Pues bien, pensamos ahora que los médicos estudiados son una muestra de una determinada población de profesionales y que la relación entre las dos variables del nivel 1, el paciente, es la misma para todos los médicos salvo que, por término medio, los pacientes de unos médicos están más satisfechos con su atención que los pacientes de otros médicos. Esta situación se puede modelar permitiendo que la satisfacción media de cada médico se exprese, no como un parámetro fijo, sino aleatorio y constituido por dos componentes, uno fijo, que puede ser una constante o depender de la experiencia del médico, y otro aleatorio, es decir,

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}$$

donde z_j es la experiencia del médico j , variable que también consideraremos centrada. El componente aleatorio u_{0j} , que expresa la variabilidad en el grado medio de satisfacción entre los distintos médicos, juega de ahora en adelante un papel fundamental. Sustituyendo esta última expresión en la anterior y reordenando tenemos que

$$y_{ij} = \gamma_{00} + \beta_1 x_{ij} + \gamma_{01} z_j + (e_{ij} + u_{0j})$$

modelo que tiene un componente fijo y otro componente aleatorio $e_{ij} + u_{0j}$. En esta nueva situación, γ_{00} no es más que la satisfacción media de los enfermos con una salud percibida media y que son atendidos por un profesional de experiencia media. Este coeficiente γ_{00} se podría interpretar como una satisfacción de referencia, alrededor de la cual van a variar tanto los pacientes como los médicos; γ_{01} es el cambio promedio de satisfacción en los enfermos por cada año más de experiencia de su médico en igualdad de salud percibida; β_1 es el cambio promedio en satisfacción por cada unidad de aumento de salud percibida, y este cambio es el mismo para todos los médicos.

Este modelo presenta la gran novedad de que en el componente aleatorio figuran dos términos: el e_{ij} o variabilidad en la satisfacción entre los pacientes del médico j y u_{0j} que mide la variabilidad entre los médicos en términos de grado medio de satisfacción de sus pacientes. Por tanto, se combinan, y esto es lo importante, en un mismo modelo un error a nivel micro o in-

dividual y un error a nivel macro o de contexto; este hecho justifica el apellido de multinivel o jerárquico para estos modelos.

Si aceptamos que tales errores se distribuyen de forma independiente siguiendo distribuciones normales con medias cero y varianzas σ_e^2 y σ_{u0}^2 , fijadas una experiencia profesional y una salud percibida concretas, la varianza de la satisfacción de los pacientes se puede expresar como la suma de la varianza entre los pacientes dentro de cada médico, σ_e^2 , más la varianza entre los distintos médicos, σ_{u0}^2 ; este modelo se conoce como modelo de los componentes de la varianza (*variance component models*); es el modelo multinivel más simple pues solo contempla dos términos de error.

El cociente

$$\frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}$$

se conoce con el nombre de correlación intra-contextos y mide qué parte de la variabilidad de la satisfacción ($\sigma_e^2 + \sigma_{u0}^2$), es achacable a la variabilidad entre los contextos. Cuanto mayor sea esta correlación más variabilidad hay entre los distintos médicos y más correlación existe entre los pacientes de un mismo médico, por lo que más inapropiada sería la utilización de los métodos clásicos de regresión. Evidentemente, si esta correlación fuese cero, el modelo anterior tan sólo tendría un término de error y, como consecuencia, el modelo multinivel se reduce al clásico.

Son varios los métodos de estimación propuestos para estos modelos; así, Goldstein¹² propone un procedimiento iterativo basado en el método de los mínimos cuadrados que es equivalente al propuesto por Longford¹³. Otra aproximación a la estimación de los parámetros del modelo es la que Diggle et al¹⁴ establecen en el marco de los estudios longitudinales que, como luego se verá, también en ellos está implicada una estructura jerárquica. Sea cual sea el procedimiento utilizado, éste proporciona una estimación tanto de los efectos fijos del modelo como de los aleatorios, así como de sus varianzas. También se puede obtener una lejanía (*deviance*) para el modelo ajustado, medida que cuantifica la similitud entre las respuestas observadas y las predichas por el modelo; esta lejanía permite evaluar la significación estadística, tanto de los parámetros fijos como de los aleatorios del modelo ajustado. Recientemente se está trabajando con mucho éxito en el tratamiento de estos modelos desde el punto de vista bayesiano, mediante la utilización de procedimientos MCMC (*Markov Chain Monte Carlo*)^{15,16}.

En el modelo lineal clásico el coeficiente de determinación es un indicador de la parte de variabilidad de la variable resultado que se puede explicar en función de las predictoras; de alguna manera, este coeficiente de determinación cuantifica la cantidad de in-

formación que aportan las predictoras a la hora de disminuir la incertidumbre sobre la variable resultado. Para el modelo de componentes de la varianza que considera predictoras a dos niveles, se han definido dos coeficientes de determinación¹⁷, uno para el nivel 1 y otro para el nivel 2, que indican las variabilidades explicadas por las variables individuales y de contexto, respectivamente.

Sixma HJ et al.¹⁸ describen, en un estudio sobre 13.014 pacientes pertenecientes a 161 médicos, una aplicación de este modelo para medir, entre otras, la relación entre la satisfacción de los pacientes en términos de accesibilidad y características tanto individuales de los pacientes: la edad, género, salud percibida, etc., como del médico que los atienden: experiencia, género, etc.; también consideran variables a nivel de contexto resultantes de agregar variables individuales, por ejemplo, edad media del grupo que atiende el médico, porcentaje de mujeres, etc. Aquí los efectos de las predictoras se consideran fijos por lo que, salvo la diferencia de haber más predictoras, es decir, efectos fijos, la estructura de error es la misma.

Del modelo ajustado que presentan se deduce que, las variables asociadas con la satisfacción, en términos de accesibilidad, son la edad, el tener problemas psicológicos y el número de enfermedades crónicas. Como características del contexto, tan solo encuentran asociación con el hecho de poder ver al médico en el mismo día en que lo solicitan y con la cantidad media de problemas de los grupos de pacientes atendidos por los distintos médicos. La varianza estimada para el nivel del paciente es 4.070, mucho mayor que la del nivel del médico que es 0.282; por tanto, el coeficiente de correlación intra-médicos es $0,282/(0,282+4,070) = 0,065$. Es decir, controlando por las variables anteriormente citadas, de la variabilidad de la satisfacción de los pacientes, un 6,5% es achacable a la variabilidad entre los médicos, el nivel 2, y el resto, el 93,5% es debido a la variabilidad a nivel de los pacientes, nivel 1, dentro de los médicos. La variabilidad entre los médicos, aunque pequeña, es estadísticamente distinta de cero, lo que se puede interpretar diciendo que, aunque se controle por las variables anteriores, todavía hay características referentes a los médicos que los hacen distintos en términos de satisfacción de sus pacientes.

Otros modelos multinivel

Como se ha dicho antes, el modelo de componentes de la varianza establece que, aunque hay un efecto del médico, el cambio en la satisfacción de los pacientes por unidad de aumento de su salud percibida es la misma

para todos los médicos; sin embargo, esta es una restricción que en ocasiones puede ser demasiado fuerte. Para liberarse de tal restricción vamos a permitir que la citada relación pueda cambiar, considerando el coeficiente de la salud percibida también como aleatorio, es decir,

$$\beta_{ij} = \gamma_{i0} + \gamma_{i1}Z_j + u_{ij}$$

donde u_{ij} , es otro término de error contextual, del médico, que posibilita el hecho que médicos distintos tengan efectos distintos en la relación entre satisfacción y salud percibida. Sustituyendo esta última expresión obtenemos el nuevo modelo

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{11}X_{ij}Z_j + \gamma_{01}Z_j + (\epsilon_{ij} + X_{ij}u_{1j} + u_{0j})$$

denominado modelo de coeficientes aleatorios (*random coefficient models*), que presenta la novedad de implicar a tres componentes de error, dos de los cuales, el u_{0j} y el u_{1j} , son a nivel del médico y el otro, el ϵ_{ij} , a nivel del paciente. Lo que caracteriza a este nuevo modelo de coeficientes aleatorios es la presencia del nuevo término de error u_{ij} . En este caso, la variabilidad total es ahora

$$(\sigma_{u1}^2 X_{ij}^2 + \sigma_{u0}^2 + 2\sigma_{u01} X_{ij} + \sigma_{\epsilon}^2)$$

donde σ_{u1}^2 mide la variabilidad entre los coeficientes o efectos de la salud percibida sobre la satisfacción para los distintos médicos. Por otra parte, σ_{u01} es la covarianza entre los dos errores al nivel 2, el médico; tal covarianza mide pues la relación entre la pendiente y la media de satisfacción para los distintos médicos.

Considerando las variables predictoras centradas, los coeficientes fijos del modelo se pueden interpretar como sigue: γ_{00} es satisfacción media general, γ_{11} el efecto de la variable contextual Z, la experiencia del médico, sobre la variable individual, salud percibida; es decir, el efecto del médico depende de la salud percibida del paciente, o lo que es igual, este coeficiente cuantifica la interacción entre los niveles 1 y 2, el paciente y su médico. Evidentemente, si este coeficiente es cero ello significa que no existe interacción entre los dos niveles, es decir, el efecto de la salud percibida no depende del grado de experiencia del médico. En ese caso, los coeficientes γ_{10} y γ_{01} miden los efectos de la salud percibida y la experiencia del médico, respectivamente.

El control del gasto sanitario es un reto muy importante para todos los actores implicados en los servicios de salud, los ciudadanos en general. Los cambios demográficos y sociales y la universalización del sistema de salud en nuestro país, han tenido como consecuencia un aumento muy importante en el consumo de recursos: número de visitas al médico, pruebas complementarias, hospitalizaciones, etc.; que este consu-

mo viene condicionado tanto por el médico como por el paciente es una obviedad. Tiene cierta tradición la investigación relativa a las características de los pacientes que puedan estar más asociadas a un mayor número de visitas al médico; sin embargo, el papel del médico ha sido menos estudiado, por lo que es pertinente la siguiente pregunta: a igualdad de perfil de usuarios, ¿hay médicos que ven más a sus pacientes que otros o, por el contrario, los médicos se comportan de forma parecida, por lo que el número de visitas de los pacientes está determinado solamente por las características de éstos?

Los datos que se analizan a continuación son parte de un proyecto de investigación multicéntrico, cuyo objetivo es evaluar el comportamiento de un sistema de clasificación de pacientes, los Grupos de Atención Ambulatoria, diseñado por la universidad John Hopkins para identificar tipos de pacientes con un consumo de recursos similar, en el entorno de la atención primaria. A efectos de ejemplo, como variable respuesta hemos elegido el número de visitas y como predictoras, el género (0 = hombre, 1 = mujer) y la edad, en años, de cada paciente. Ya que el número de visitas sigue una distribución sesgada a la derecha, se ha transformado logarítmicamente esta variable, lo que tiene consecuencias a la hora de interpretar los coeficientes del modelo, como luego veremos. En primer lugar hemos ajustado un modelo nulo (Modelo A), con sólo tres parámetros a estimar: el término constante, y las varianzas tanto de los pacientes como de los médicos (tabla 1). Con este modelo se formula que la respuesta, el logaritmo de las visitas, se puede explicar mediante un valor que es igual para todos los pacientes, más una parte aleatoria que representa la heterogeneidad entre los pacientes, en cuanto al número de visitas, y otra correspondiente a los médicos. Según la tabla 1, la estimación del valor común, la constante, es 1.162, interpretable como la media del logaritmo de las visitas; dicho de otra forma, el número medio de visitas por paciente es $\exp(1.162) = 3,2$ visitas; de la variabilidad observada ($0,893 + 0,121$) = 1.014, la mayor parte un 88% se debe a los pacientes y el 12% restante es debida a que unos médicos generan más visitas que otros.

El modelo ajustado no contempla las características de los pacientes, por lo que no hay evidencias de que la variabilidad observada entre los médicos no sea más que un artefacto debido al distinto perfil de pacientes que atienden los médicos. Para responder a esta cuestión hemos introducido, en el modelo B, dos características de los pacientes: el género y la edad, esta última centrada en su valor medio; ahora el término constante se interpreta como el logaritmo del número de visitas de un paciente «tipo»: hombre de 62 años; según el modelo ajustado, no hay evidencia de diferencias entre hombres y mujeres, mientras que las edades mayores están asociadas a una mayor frecuentación; el descenso

Tabla 1. Ajuste de tres modelos multinivel

	Modelo A		Modelo B		Modelo C	
	Estimación	Error estándar	Estimación	Error estándar	Estimación	Error estándar
<i>Parámetros fijos</i>						
Constante	1,162	0,079	1,109	0,099	1,102	0,099
Género: mujer			0,034	0,044	0,035	0,044
Edad-62			0,008	0,001	0,008	0,001
<i>Parámetros aleatorios</i>						
σ_e^2	0,893	0,028	0,881	0,027	0,878	0,027
σ_{u0}^2	0,121	0,040	0,089	0,030	0,085	0,032
σ_{u1}^2					1,3E-5	1,6E-5
σ_{u01}^2					1,0E-4	4,9E-4
Lejanía	5778,27	2097 g,l,	5743,61	2095 g,l,	5743,56	2093 g,l,

en lejanía entre los modelos A y B indica la «importancia» de las características de los pacientes en relación al número de visitas al médico. La introducción de estas variables del paciente conlleva una reducción de la variabilidad entre los pacientes, aunque ciertamente muy pequeña: de 0,893 a 0,881; la variabilidad entre los médicos pasa de 0,121 en el modelo anterior a 0,089 para el actual, lo que refleja que una parte de la variabilidad entre los médicos era debida a que atienden distintos perfiles de pacientes; cuando estos se tienen en cuenta, los médicos tienen un comportamiento más homogéneo, aunque todavía existe variabilidad estadísticamente significativa.

El modelo B postula que, por ejemplo, el efecto de la edad es el mismo para todos los médicos, lo que no tiene porqué ser cierto. El modelo C, de coeficientes aleatorios, permite que este efecto varíe entre los distintos médicos, por lo que tendremos otro componente aleatorio en relación a este efecto, más otro que mide la covarianza entre el término constante y la edad; una estimación de la variabilidad del referido efecto entre los distintos médicos es 1,3E-5 y la covarianza es 1,0E-4. La significación estadística de los parámetros fijos se puede evaluar mediante el estadístico de Wald, estimación dividida entre su error estándar, pero para el caso de los parámetros aleatorios es más pertinente la utilización de la diferencia en lejanías comparándola con una chi-cuadrado cuyos grados de libertad son la diferencia de éstos entre los dos modelos; para nuestro ejemplo, $5743,61 - 5743,56 = 0,05$, no significativa para una chi-cuadrado con $2095 - 2093 = 2$ grados de libertad; la correlación entre los componentes aleatorios de la constante y el efecto de la edad es muy baja, 0,03. Ya que el modelo C no explica mejor los datos que el B y éste es más parsimonioso, será el de elección; sin embargo, la variabilidad residual a ambos niveles todavía es significativa, lo que indica la existencia de características, tanto de los pacientes como de los mé-

dicos, no tenidas en cuenta en este análisis y que pueden explicar, al menos en parte, tales variabilidades. En definitiva, controlando por el género de los pacientes, el efecto de la edad de los pacientes es el mismo para todos los médicos.

Los modelos propuestos hasta ahora están pensados para una respuesta continua; sin embargo, en investigación sanitaria son frecuentes las variables respuesta dicotómicas. Por ejemplo, consideremos el hecho de realizar o no realizar una angiografía a un paciente que ha sufrido un infarto agudo de miocardio; es posible que tal hecho pueda depender de la situación clínica y características del paciente, pero también puede depender de otras circunstancias ajenas al paciente y relacionadas con, por ejemplo, la «cultura» y medios del hospital donde el paciente es atendido. En este caso, la respuesta, realizar o no la angiografía, es una variable dicotómica, por lo que los modelos anteriores no son aplicables a esta nueva situación. Pues bien, el modelo logístico tradicional también se puede extender al caso de datos jerárquicos permitiendo que alguno o algunos de sus parámetros sean aleatorios. Gatsonis¹⁹ describe una aplicación del modelo logístico multinivel para estudiar la utilización de la angiografía en ancianos con infarto en función de variables del paciente: edad, género, raza y un índice de comorbilidad y también en función de la «disponibilidad» de la angiografía en cada uno de los estados de USA.

Para el caso de una respuesta categórica con más de dos posibles resultados se han propuesto modelos multinivel de regresión logística policotómica; Daniels et al²⁰ lo ejemplifican con los datos del estudio anterior pero considerando tres alternativas para la variable resultado: no angiografía, sí angiografía y, por último, angiografía coronaria y revascularización.

También tienen mucho interés este tipo de modelos en análisis de supervivencia. En efecto, cuando se contrastan las experiencias de supervivencia entre gru-

pos de pacientes atendidos en distintos hospitales o pertenecientes a distintos ámbitos geográficos, es frecuente observar una gran variabilidad de la supervivencia de los pacientes²¹; los modelos multinivel permiten evaluar la posible heterogeneidad de los hospitales en términos de la supervivencia de sus pacientes²². Sastry²³ utiliza los modelos multinivel de riesgos proporcionales para analizar la mortalidad infantil en Brasil sobre una muestra elegida mediante un muestreo por *clusters* multietápico: primero se tomó una muestra de unidades administrativas (comunidades), dentro de las cuales se eligieron familias y dentro de estas a los niños.

Extensiones a otros escenarios

En las situaciones discutidas hasta ahora se han considerado sólo dos niveles de jerarquía, pero estos pueden ser tres o más, no habiendo gran dificultad teórica en extender los modelos multinivel; por ejemplo, la satisfacción de los pacientes puede cambiar con sus propias características pero también con las de los profesionales que los atienden, los servicios donde estos trabajan y el área sanitaria donde están localizados los servicios.

En la investigación sanitaria son frecuentes los estudios en los que no se cumple la hipótesis de independencia entre los valores de la variable resultado, necesaria para poder utilizar los modelos de regresión comunes; así, en los ensayos clínicos cruzados (*cross-over*) o en los estudios de seguimiento con medidas repetidas, a un mismo individuo se le mide la variable resultado en distintas ocasiones; como quiera que estos diseños son más interesantes desde el punto de vista de la interpretabilidad de los resultados, son de un interés creciente. En epidemiología genética se presenta una situación similar cuando, al poder pertenecer dos o más individuos de la muestra a una misma familia, estos van a compartir una carga genética que va en contra de la hipótesis de independencia. Pues bien, todo esto se puede ver desde el punto de vista jerárquico, pensando en las distintas medidas sobre un individuo, o sobre los de una misma familia, como el primer nivel y en el individuo o familia, respectivamente, como el segundo nivel. Snijders²⁴ muestra la aplicación de distintos modelos multinivel a un estudio en el que se midió en 245 individuos un índice de neurosis en cinco ocasiones, durante el período 1970-1986, con el objetivo de estudiar la evolución de tal índice a lo largo del tiempo, utilizando tanto variables predictoras fijas, género y nivel educativo, como dependientes del tiempo.

Esta situación es similar, desde el punto de vista conceptual, al caso de disponer de varias variables respuesta. Siguiendo con el ejemplo anterior, la satisfacción del paciente con la atención recibida tiene que ver,

tanto con el componente de accesibilidad a los servicios, como con otros posibles: la humanidad del trato y/o la calidad de la información recibida, etc. Por tanto, los modelos multinivel pueden estudiar simultáneamente distintos componentes de un mismo fenómeno, posibilitando conocer las interrelaciones entre las distintas respuestas. Así, podemos responder a preguntas del tipo: a) ¿es igual o no la relación entre cada uno de los componentes de la satisfacción y la edad del paciente?; b) ¿los más satisfechos en una componente están, en general, más satisfechos también en las restantes componentes? Además de esta información de gran interés a nivel individual, los modelos multinivel nos podrían informar acerca de si, controlando por las características de los pacientes atendidos por cada profesional, el médico más accesible, tiene mejor trato con los pacientes y/o informa mejor a sus pacientes²⁵.

Por último, Goldstein¹² ha propuesto modelos multinivel de clasificación cruzada que no exigen una estructura estrictamente jerárquica; por ejemplo, los individuos, el nivel 1, pueden ser clasificados según el sitio donde residen y el lugar de trabajo, sin que entre estas dos clasificaciones distintas exista una jerarquía; en definitiva, una unidad de un determinado nivel puede pertenecer a una o más de una del nivel inmediatamente superior. Langford²⁶ utiliza un modelo multinivel de clasificación cruzada para estudiar la mortalidad en los distritos de Inglaterra y Gales.

Aquí no se agota la flexibilidad de estos modelos, pues las posibilidades que presenta el hecho de poder modelar el término de error, son de mucho interés, tanto conceptual como estadístico; así, se puede tratar con modelos lineales en los que la condición de homocedasticidad no se cumple, o bien, podemos modelar la sobredispersión que en regresión logística, más que una excepción, es casi una regla.

Conclusiones

Los modelos multinivel son una respuesta a la necesidad de analizar la relación entre los individuos y el medio en donde se desenvuelven; poder separar el papel de cada uno de los componentes de la compleja estructura implicada puede llevar a un mejor conocimiento de la realidad socio-sanitaria para así poder intervenir de forma más eficiente.

Por otra parte, la introducción de un nuevo procedimiento estadístico en la investigación aplicada, en particular en ciencias de la salud, es un proceso que se desarrolla en distintas etapas: después de una fase de exaltación sigue un período de reflexión crítica que lleva a una evaluación realista de lo que el nuevo método puede aportar, quedando finalmente incorporado al armamentarium metodológico²⁷. Para los modelos multilevel parece que

ha pasado la primera fase, estando en la actualidad en la etapa de discusión crítica, tanto de aspectos puramente estadísticos, como de los problemas del marco teórico que su aplicación plantea en la investigación sanitaria.

En resumen, los modelos multinivel ofrecen distintas ventajas respecto a los modelos tradicionales: dan una versión más realista ya que modelan cada nivel de jerarquía; no requieren la hipótesis de independencia entre las medidas de la variable resultado y también dan estimaciones más precisas. Un precio a pagar es la mayor complejidad tanto del marco teórico como del modelo propuesto para analizar los datos, lo que conlleva una mayor dificultad en la comunicación de los resultados. De cualquier forma, la gran frecuencia con que se encuentran estructuras jerárquicas en los datos que

provienen de los estudios epidemiológicos, de la investigación de servicios de salud, etc., está demandando una mayor utilización de estos modelos, lo que puede contribuir a una revitalización en la investigación de los determinantes colectivos y sociales de la salud de los individuos^{28,29,30,31}.

Agradecimientos

A nuestras compañeras de la Escuela Andaluza de Salud Pública: Carmen Martínez, que nos ayudó con su crítica y Alexandra Prados, investigadora principal del proyecto sobre Grupos de Atención Ambulatoria (FIS 96/0576), que nos permitió la utilización de parte de la base de datos.

Bibliografía

- Box G. Statistics and quality improvement. *J Royal Stat Soc, Series A* 1994;157:209-29.
- Goldstein H, Spiegelhalter D. League tables and their limitations: statistical issues in comparasions of institutional performance (with discussion). *J Royal Statis Soc, Series A* 1996;159:385-443.
- Buck CW. Prenatal and perinatal causes of early death and defect. In Clark DW, MacMahon B (eds). *Preventive and Community Medicine* 2d ed. Little Brown and Co. Boston MA, 1981.
- Blaxter M. *Health and Lifestyles*. Londres, Routledgr, 1990.
- Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wayes. *BMJ* 1989;299:886-9.
- Gronick M. Medicare patients: regional differences in length of hospital stay, 1969-1971. *Soc Sec Bull* 1975;40:22-41.
- Wennberg JE, Freeman JL, Culp WJ. Are hospital services rationed in New Haven or over-utilised in Boston?. *Lancet* 1987;337:1185-9.
- McPherson K, Wennberg JE, Hovind OE, Clifford P. Small-area variation in the use of common surgical procedures: an international comparasion of New England, England and Norway. *N Engl J Med* 1982;307:1310-4.
- Altman DG, Goodman SN. Transfer of technology from statistical journal to biomedical literature. Past trends and future predictions. *JAMA* 1994;272:129-32.
- Morgenstern H. Ecologic studies in epidemiology. *Annu Rev Public Health* 1995;16:61-81.
- Robinson WS. Ecological correlations and the behavior of individuals. *Am Sociol Rev* 1950;15:351-7.
- Goldstein H. *Multilevel statistical models*, 2.^a ed. Arnold, 1995.
- Longford NT. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 1987;74:817-27.
- Diggle PJ, Liang KY, Zeger SL. *Analysis of longitudinal data*. Oxford Univ Press, Londres, 1994.
- Clayton D. Generalized linear mixed models. Gilks WR, Richardson S, Spiegelhalter DJ (eds). *Markov Chain Monte Carlo in Practice*. Londres: Chapman and Hall, 1996.
- Richardson S. Développements récents de la biostatistique. *Rev Epidem et Santé Publ* 1996;44:482-93.
- Snijders TAB, Bosker RJ. Model variance in two-levels models. *Sociol Methods Research* 1994;22:342-63.
- Sixma HJ, Spreeuwenberg PMM, van der Pasch MAA. Patient satisfaction in the general practitioner: a two-level analysis. *Medical Care* 1998;36:212-29.
- Gatsonis CA, Epstein AM, Newhouse JP, Normand SL, McNeil BJ. Variations in the utilization of coronary angiography for elderly patients with an acute myocardial infarction: an analysis using hierarchical logistic regression. *Medical Care* 1995;33:625-42.
- Daniels MJ, Gatsonis CA. Hierarchical polytomous regression models with application to health services research. *Stat Med* 1997;16:2311-25.
- Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J. Survival of cancer patients in Europe: The EUROCARE study. *IARC Sci Publ No 132*. IARC, Lyon, 1995.
- Stangl DK, Greenhouse JB. Assesing placebo response using bayesian hierarchical survival models. *Lifetime Data Anal* 1998;4:5-28.
- Sastry N. A multilevel hazards model for hierarchically clustered data: model estimation and an application to the study of child survival in northeast Brazil. *RAND Working Paper Series 95-15*, 1995.
- Snijders T. Analysis of longitudinal data using the hierarchical linear model. *Quality and Quantity* 1996;30:405-26.
- Duncan C, Jones K, Moon G. Health-related behaviour in context: a multilevel modelling approach. *Soc Sci Med* 1996;42:817-30.
- Langford IH, Bentham G. Regional variations in mortality rates in England and Wales: an analysis using multi-level modelling. *Soc Sci Med* 1996;42:897-908.
- Mason WM. Comment. *J. Educational and Behavioral Statistics* 1995;20:221-7.
- Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health* 1998;88:216-22.
- Korff M von, Koepsell T, Curry S, Diehr P. Multi-level analysis in epidemiologic research on health behaviors and outcomes. *Am J Epidemiol* 1992;135:1077-82.
- O'Campo P, Gielen AC, Faden RR, Xue X, Kass N, Wang RR. Violence by male partners against women during the childbearing year: a contextual analysis. *Am J Public Health* 1995;85:1092-97.
- Duncan C, Jones K, Moon G. Do places matter? A multi-level analysis of regional variations in health-related behaviour in Britain. *Soc Sci Med*. 1993;37:725-33.