Revista Investigación Educativa - N.º 25 - 1995

ESTUDIO METODOLÓGICO

ANÁLISIS SECUENCIAL DE DATOS OBSERVACIONALES EN INVESTIGACIÓN EDUCATIVA (I): PERSPECTIVA BIVARIANTE

por
Juan Carlos Tójar y José Serrano
Métodos de Investigación y Diagnóstico en Educación
Universidad de Málaga

RESUMEN

El análisis secuencial de datos, procedentes de la observación de contextos particulares, en los que se producen comportamientos en interacción, ha experimentado en los últimos años un gran auge en la ciencias sociales. En el ámbito de la Psicología y de la Educación, esta nueva perspectiva supone un aumento de las posibilidades en el tratamiento de los datos categoriales secuenciales y, en definitiva, un enriquecimiento en las técnicas de análisis con las que aproximarnos a la realidad educativa.

En este trabajo se reúnen y sistematizan las técnicas clásicas de análisis secuencial de datos observacionales para facilitar su manejo en diseños de investigación educativa. Se estudian formalmente las condiciones del número necesario de observaciones y el control del error tipo I, ofreciendo dos tablas que facilitan decisiones a la hora de planificar la investigación.

Palabras clave: Observación; Análisis secuencial de datos; Procesos de Markov; Técnica de retardos; Tamaño de la muestra; Significación estadística.

ABSTRACT

Sequential data analysis derived from the observation of particular interactive contexts has undergone a step forward in the field of social sciences over the last few years. This new approach entails an increase of possibilities in the analysis of categorial sequential data in the psychological and educational fields and, therefore, an enrichment of the analysis techniques which could be used to approach the educational reality.

Throughout this report the traditional techniques of sequential analysis of observational data are gathered and systematized in order to facilitate its use in educational research design. The conditions of sample size and the control of the type I error are studied by means of two tables which provide a number of decisions when planning a research.

Key words: Observation; Sequential data analysis; Markov processes; Lag method; Sample size; Statistical significance.

1. INTRODUCCIÓN

Las técnicas de análisis secuencial de datos han supuesto una revolución en el tratamiento de la información recogida mediante *observación sistemática.* Los datos recogidos en esquemas de codificación o *sistemas de categorías*, empleados como instrumentos de investigación en el aula (v.g. las amplias recopilaciones de Simon y Boyer, 1967 y 1970), se encontraban limitados a análisis no secuenciales, que informaban poco más que de la proporción de ocurrencia de determinado comportamiento con respecto al resto.

El comportamiento de un organismo o de un grupo que interactúa en un contexto particular puede ser descrito a partir de las frecuencias de los estados por los que pasa. Así por ejemplo, observando y registrando el desarrollo de un debate en un aula, y atendiendo únicamente al parámetro *frecuencia*, es posible determinar quién o quiénes han participado más asiduamente y quiénes no han participado tanto.

Sean cual sean los comportamientos considerados, como es obvio, se desarrollan en el tiempo y una descripción de los mismos se hace tanto más ajustada a la realidad en tanto que incluye, además de la frecuencia, al menos el *orden* en que dichos estados o comportamientos se producen. Si en el ejemplo simple de la asamblea, además de tener en cuenta la *cantidad de participación* de cada sujeto, se registra el orden en que las diferentes personas van tomando la palabra, la idea que se obtiene, a partir del análisis de los registros, de la forma en que se ha ido desarrollando la asamblea es más completa.

Si se une a ellos la inclusión del parámetro *duración* la descripción de la sucesión de los diferentes estados o comportamientos puede completarse aún más y la información registrada ganar en calidad. La consideración o no de este último parámetro se puede realizar en función del nivel de respuesta o del objeto de estudio de interés, aunque a veces puede verse limitada por el grado de instrumentación disponible en la investigación.

En pocas palabras, tomar en cuenta la perspectiva secuencial en el registro y en el análisis de los datos permite poner de manifiesto la *natural secuencialidad* de cualquier sistema en interacción.

El análisis secuencial, denominación utilizada para englobar un conjunto de diferentes técnicas aplicables a datos secuenciales categóricos, se ha desarrollado de forma espectacular en las dos últimas décadas (Bakeman, 1991:14). En España, destacan los recientes trabajos de Anguera (1991; 1993) y Tójar (1993; 1994).

Los objetivos del análisis secuencial se pueden resumir en *descubrir patrones estocásticos en los datos* y *evaluar el efecto de variables contextuales y explicativas en la estructura secuencial* (Gottman y Roy, 1989: 19). Analizando secuencias se pretende, además de la identificación de las situaciones o comportamientos más probables en un contexto dado, la descripción, por un lado, y la explicación y predicción, por otro, de los estados por los que pasa una interacción. La descripción se puede realizar mediante cadenas de secuencias de situaciones previamente categorizadas, y en cada instante es posible predecir qué situaciones son inhibidas y/o excitadas con respecto a otra en concreto (criterio).

Bakeman y Dabbs (1976) establecieron un doble criterio, ya clásico, según el cual pueden recogerse los datos en investigación observacional. Considerando el criterio *ocurrencia*, los datos pueden ser clasificados en *secuenciales* o *concurrentes*. Y mediante lo que estos autores denominaron criterio de *base*, en datos de *evento* o de *tiempo*. Combinando ambos criterios aparecen cuatro diferentes tipos de datos como puede verse en la tabla 1.

Tabla 1
BAKEMAN Y DABBS (1976), BAKEMAN (1978)

	OCURRENCIA	
BASE	Secuenciales	Concurrentes
Evento	I	II
Tiempo	III	IV

Según este doble criterio los datos tipo I se considerarían secuencialmente, el observador anota su orden pero no su duración (cuidando la mutua exclusividad entre categorías). Los datos tipo II, se producen cuando el sistema de categorías no es mutuamente exclusivo y aunque el codificador registre el orden en que ocurren los eventos, en el registro se muestran concurrencias de categorías.

Para los datos tipo III y IV se tiene siempre en cuenta la duración de cada evento. La diferencia se sitúa en que mientras que en el tipo III no aparecen concurrencias, en los datos tipo IV sí (como consecuencia del uso de unas categorías no mutuamente excluyentes).

Los diferentes tipos de datos pueden transformarse entre sí. Los de tipo IV y tipo III se pueden transformar en tipo II y tipo I respectivamente si se deja de considerar la duración de las categorías en el registro. Los de tipo II y tipo IV se pueden transformar respectivamente en tipo I y tipo III, aunque esto se traduzca en aumentar el número de categorías y la complejidad del registro (Anguera, 1983). De cara a eliminar complejidad en el análisis y a encontrar regularidades y patrones con mayor facilidad lo mejor es trabajar en lo posible con datos tipo III, si se precisa considerar el criterio tiempo, o con los datos tipo I, si no interesa considerar la duración de las categorías.

La clasificación de Bakeman y Dabbs (1976) condiciona la manera de hacer los registros (aun teniendo en cuenta las posibles transformaciones entre los mismos), para que a partir de un análisis adecuado de los mismos, sea posible identificar patrones concurrentes o secuenciales. Si se compara esta clasificación con todas las posibilidades formas de registrar datos observacionales aparecidas en los últimos años (Quera, 1991) se observa que dicho esquema se ha quedado pequeño.

Con el objeto de dar cabida a más posibilidades de registro, evitando ceñirse a normas demasiado estrictas, facilitando asimismo el intercambio, la comunicación entre investigadores y el desarrollo de programas informáticos de análisis de datos Bakeman y Quera (1992), han propuesto un nuevo estándar denominado SDIS (*Sequential Data Interchange Standard*).

SDIS es un lenguaje de datos secuenciales, con normas específicas sintácticas y de puntuación, basado en una reconsideración de las técnicas más habituales de recoger datos observacionales. Una exposición completa así como la sintaxis del SDIS pueden ser consultadas en el referido trabajo de Bakeman y Quera. A modo ilustrativo en este trabajo simplemente se van a presentar las clases fundamentales del nuevo estándar destacando en cada caso el registro del que derivan cada una de ellas:

1) Secuencias de eventos o ESD (*Event Sequential Data*). Consisten en series de códigos que representan unidades de conducta mutuamente exclusivas y en los que no se ha registrado el tiempo. Es el formato más simple que representa un registro activado por transiciones y coincide plenamente con los datos tipo I del estándar de la tabla 1. Se pueden subdividir en ESD repetibles y ESD no repetibles en función de que se permita recoger en la secuencia códigos adyacentes iguales o no.

2) Secuencias de estados o SSD (*State Sequential Data*). Se obtienen a partir de series de códigos que representan unidades de conducta en las que el tiempo ha sido considerado. Este formato está previsto para registro activados mediante transiciones en los que se ha recogido la duración de cada ocurrencia o del inicio de la misma. Los datos tipo III e incluso los de tipo IV en algunas ocasiones pueden considerarse como SSD.

3) Secuencias mixtas de eventos y estados o TSD (*Timed Sequential Data*). Este es el formato más complejo al combinar eventos y/o estados que no han de ser necesariamente ni mutuamente excluyentes ni exhaustivos. Se produce a partir de un registro activado por transiciones en los que se recoge además el tiempo (duración, tiempo de inicio o de finalización de cada categoría). Las secuencias TSD pueden transformarse en una o varias SSD.

4) Secuencias de intervalos o ISD (*Interval Sequential Data*). Una secuencia ISD está formada por intervalos de tiempo constante y puede representar las ocurrencias de categorías recogidas mediante los RAUT A, B y C (tres modalidades de registro activado por unidades de Tiempo, ver Quera, 1991).

En este trabajo se van a revisar y sistematizar las dos técnicas clásicas para detectar patrones secuenciales, las *cadenas de Markov* y la *técnica de retardos*. Los modelos *log-lineales* también pueden ser utilizados para realizar el análisis secuencial desde una perspectiva multivariante (Bakeman, Adamson y Strisik, 1989), sin embargo realizar una revisión de su aplicación superaría los límites establecidos para la confección de este trabajo, por lo que se ha optado por realizarla aparte.

2. CADENAS DE MARKOV

Desde que en 1952, G. Miller (cit. en Gottman y Roy, 1990) introdujera los procesos de Markov en psicología, muchos autores los han utilizado como modelos simples para las secuencias de categorías conductuales. Las cadenas de Markov permiten el tratamiento de los datos observacionales, a partir del análisis de las probabilidades de transición, con el fin de detectar patrones secuenciales entre pares de conductas o situaciones categorizadas a lo largo del tiempo (Gottman y Notarius, 1978).

Los requisitos para que una secuencia de categorías pueda ser considerada como un proceso de Markov se encuentran destacadas en un trabajo de Quera (1991).

En primer lugar, la probabilidad de que una determinada categoría (o estado en la terminología markoviana) ocurra tras una transición depende únicamente del estado anterior que ha finalizado y no de los anteriores a ella. Las probabilidades de transición pues son independientes de la historia previa a cada una de ellas e incluso del tiempo transcurrido desde el inicio de la secuencia de categorías. En segundo lugar, *las duraciones de ocurrencia de cada categoría se distribuyen de forma exponencial* (Quera, 1991: 285), hecho que se produce sólo si los momentos en los que aparecen las transiciones son puntos aleatorios de un continuo temporal (la probabilidad de que finalice una categoría iniciada es independiente también del tiempo transcurrido desde su comienzo).

Supóngase un proceso que genera una serie de elementos completamente aleatoria. Este proceso de Poisson, en el que un cierto evento aparece aleatoriamente con una tasa constante en los momentos $t_0, t_1, t_2, ...t_m$, se basa en los siguientes supuestos:

1) En el intervalo comprendido entre t y $t + d$ (para d entero y mayor o igual que 1), los eventos suceden con una tasa de ocurrencia constante λ_d.
2) La probabilidad de que ocurran dos eventos o más en el intervalo entre t y $t + 1$ es nula.
3) El hecho de que el evento ocurra en el momento t es independiente de lo ocurrido antes de t.

Cuando estas condiciones se cumplen, la probabilidad de que el evento ocurra *n* veces desde el inicio de la sesión, esto es, en el intervalo [0, *t*], viene dada por la ley de Poisson cuya expresión es la siguiente:

$$P(k_t = n) = \frac{(\lambda_t)^n}{n!} e^{-\lambda_t} \qquad [1]$$

donde k_t es la variable aleatoria discreta número de eventos que ocurren en un intervalo de duración *t*.

Si se denomina *X* a la variable continua formada por los tiempos transcurridos entre dos eventos sucesivos, la función de densidad de *X* es:

$$f(X) = \lambda e^{-\lambda x} \qquad [2]$$

Cuya esperanza matemática y variancia valen respectivamente:

$$E(X) = \frac{1}{\lambda} \qquad V(X) = \frac{1}{\lambda^2} \qquad [3]$$

Lo que significa que la duración media esperada entre eventos sucesivos será igual al inverso de la tasa de dichos eventos.

Este modelo es verosimil cuando «*a) Todas las categorías son eventos y el tiempo transcurrido entre dos de ellos es independiente de qué categorías se trata; en ese caso todos los tiempos entre eventos sucesivos tienen una única función de densidad exponencial. O bien, b) todas las categorías son estados y las propiedades estadísticas de las duraciones de ocurrencia de categorías distintas son sin embargo idénticas; en este caso todas las duraciones de ocurrencia tienen una única función de densidad exponencial*» (Quera, 1991: 287).

Como conclusión queda que si una serie de transiciones entre categorías estado sigue un proceso de Poisson, entonces es también un proceso de Markov (lo contrario no es necesariamente cierto, puesto que es posible que las duraciones de ocurrencia tengan una diferente distribución exponencial con lo que no seguirían un proceso de Poisson).

Para el cálculo de las cadenas de Markov es preciso partir de estados diádicos (Gottman y Bakeman, 1979). Supóngase un sistema de categorías con n códigos. Dichos códigos pueden representarse en un *sistema* con *k* estados. El sistema puede ser descrito mediante transiciones de estados diádicos y discretos (perfectamente compatible con la mutua exclusividad de las categorías).

Las frecuencias correspondientes a la transición entre estados pueden representarse en una matriz de transición *F* de *mxm* en el tiempo igual a *t*+1:

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & f_{1,3} & \cdots & f_{1,m\text{-}1} & f_{1,m} \\ f_{2,1} & f_{2,2} & f_{2,3} & \cdots & f_{2,m\text{-}1} & f_{2,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ f_{m\text{-}1,1} & f_{m\text{-}1,2} & f_{m\text{-}1,3} & \cdots & f_{m\text{-}1,m\text{-}1} & f_{m\text{-}1,m} \\ f_{m,1} & f_{m,2} & f_{m,3} & \cdots & f_{m,m\text{-}1} & f_{m,m} \end{bmatrix} \quad [4]$$

Y de igual forma se puede construir la matriz de probabilidades de transición de primer orden (de t a t+1):

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \cdots & p_{1,m\text{-}1} & p_{1,m} \\ p_{2,1} & p_{2,2} & p_{2,3} & \cdots & p_{2,m\text{-}1} & p_{2,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{m\text{-}1,1} & p_{m\text{-}1,2} & p_{m\text{-}1,3} & \cdots & p_{m\text{-}1,m\text{-}1} & p_{m\text{-}1,m} \\ p_{m,1} & p_{m,2} & p_{m,3} & \cdots & p_{m,m\text{-}1} & p_{m,m} \end{bmatrix} \quad [5]$$

p_{ij} es la probabilidad de transición del estado i en el tiempo t al estado j en el tiempo t+1, que se calcula así:

$$p_{ij} = \frac{f_{ij}}{\sum_{j=1}^{m} f_{ij}} \quad [6]$$

donde f_i es la frecuencia del estado i.
f_{ij} es la frecuencia de transición de i a j, o número de veces que j sigue a i.

Para detectar las transiciones significativas es posible calcular diversos índices de conexión secuencial:

a) Una *prueba binomial*, propuesta inicialmente por Sackett (1974, cit. en Gottman y Roy,1990), fue adaptada del contraste paramétrico z entre dos proporciones. Más tarde fue modificada por Gottman (1980) y Allison y Liker (1982). Consiste en un estadístico z asintóticamente normal, que puede ser comparado con la distribución normal estándar para una muestra lo suficientemente grande (v. g. $n > 20$). La expresión que aquí se presenta es la ofrecida en el texto de Gottman y Roy (1990).

$$z^* = \frac{p(B_{t+k}/A) - p(B)}{\sqrt{\dfrac{p(B)[1-p(B)][1-p(A)]}{(n-k)p(A)}}} \qquad [7]$$

donde A y B son las dos categorías (estado) de un sistema.
$p(B)$ y $p(A)$ son las probabilidades de ocurrencia de cada categoría.
$p(B_{t+k}/A)$ es la probabilidad de transición de la categoría A a la categoría B en un proceso de orden k.
n es el número total de ocurrencias de todas las categorías.
k es el orden (en el tiempo) del modelo de Markov $(t+k)$.

b) La *prueba* χ^2 de Pearson (Gottman y Notarius, 1978), aplicada a la matriz de frecuencias de transición de orden k: El resultado que se obtenga ha de ser comparado con un valor crítico de la distribución teórica de χ^2 de Pearson con $(m-1)^2$ grados de libertad, donde m es el número de códigos o categorías del sistema.

$$\chi^2 = \sum_{j=1}^{m} \sum_{i=1}^{m} \frac{\left(f_{ij} - \hat{f}_{ij}\right)^2}{\hat{f}_{ij}} \qquad [8]$$

donde $\hat{f}_{ij}$ son las frecuencias de transición del estado i al j en el modelo de orden k.
$\hat{f}_{ij}$ son las frecuencias esperadas bajo el supuesto de independencia.

c) La *razón de verosimilitud* ($LR\chi^2$: *Likelihood Ratio*). Utilizando la misma notación que en la anterior χ^2, se aplica a la matriz de frecuencias de transición de orden k: El resultado que se obtenga se compara también con un valor crítico de la distribución teórica de χ^2 de Pearson con $(m-1)^2$ grados de libertad, donde m es el número de categorías del sistema.

$$LR\chi^2 = 2\sum_{j=1}^{m} \sum_{i=1}^{m} f_{ij} \log\left(\frac{f_{ij}}{\hat{f}_{ij}}\right) \qquad [9]$$

d) La *transformación logit* (Fienberg, 1980; Allison y Liker, 1982), es una medida no influida por los marginales (de las filas) totales, que permite comparar las contingencias secuenciales de dos grupos. Su fórmula es la siguiente:

$$z = \frac{\beta_1 - \beta_2}{\sqrt{\sum_{i=1}^{m} \left(\frac{1}{f_i}\right)}} \qquad [10]$$

El estadístico β se define como:

$$\beta = \text{logit}\,[p(B_{t+k} = 1/\, A = 1)] - \text{logit}\,[p(B_{t+k} = 1/\, A = 0)] \qquad [11]$$

donde el *logit* de una proporción cualquiera es igual al logaritmo neperiano del cociente entre esa proporción y su complementaria, esto es:

$$logit\,(p) = ln\left(\frac{p}{1 - p}\right) \qquad [12]$$

3. LA TÉCNICA DE RETARDOS

La técnica de retardos (*lag method*) propuesta por Sackett (1979) precisa de la elección de una categoría como criterio. En pocas palabras se puede decir que se cuenta el número de veces que cada una de las restantes categorías del sistema sigue (o precede) al criterio de forma consecutiva (primer retardo), con una categoría intermedia (segundo retardo), con dos categorías intermedias (tercer retardo), y así sucesivamente hasta el máximo retardo (*max lag*) de interés (ver figura 1).

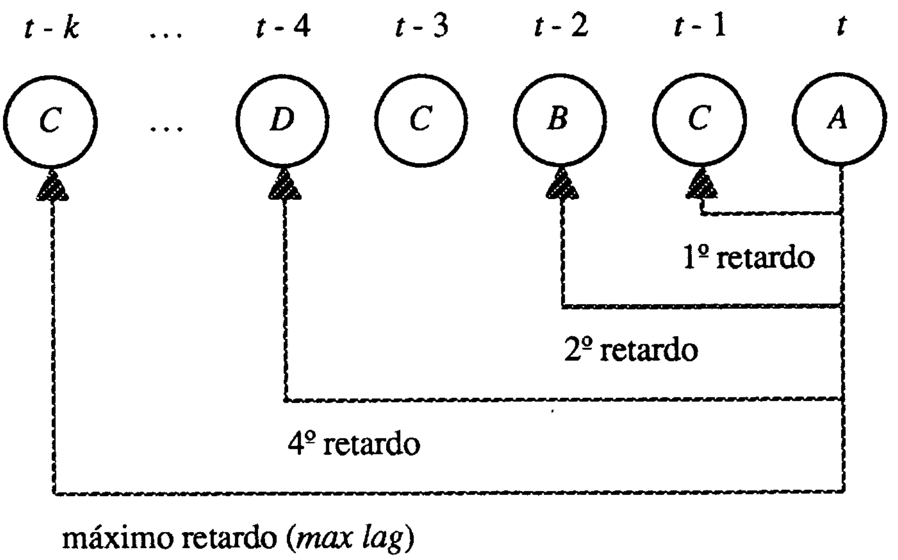

Figura 1
RETARDOS EN UNA CADENA TOMANDO COMO CRITERIO LA CATEGORÍA A

Usando la nomenclatura de Bakeman (1978) se puede decir que a partir de datos tipo I (ver tabla 1), se calcularían los denominados *retardos de evento*, mientras que si se utilizan los datos tipo III se hablaría de *retardos de tiempo*.

Mediante la técnica de retardos es posible además contrastar las probabilidades de retardo frente a una hipótesis nula de no dependencia entre situaciones que aparecen secuencialmente en el tiempo. Si la hipótesis nula es aceptada, en base a un determinado nivel de significación, una categoría *B* ocurre con la misma probabilidad si no se tiene en cuenta que se ha considerado apareada con el criterio *A*. La diferencia entre las probabilidades de retardo observadas (condicionales) y esperadas (incondicionales) pueden contrastarse mediante la siguiente prueba de *z*:

$$z = \frac{p(B_{t+k}/A) - p(B)}{\sqrt{\frac{p(B)[1-p(B)]}{(n-k)p(A)}}} \qquad [13]$$

donde *A* y *B* son las únicas dos categorías de un sistema.
$p(B)$ y $p(A)$ son las probabilidades (incondicionales) esperadas de ocurrencia de cada categoría.
$p(B_{t+k}/A)$ es la probabilidad de que *B* ocurra *condicionada* a que *A* haya sucedido hace *k* retardos.
n es el número total de ocurrencias de todas las categorías.
k es el número de unidades (de evento o de tiempo) que separa la categoría *B* del criterio *A*.
$\sqrt{\frac{p(B)[1-p(B)]}{(n-k)p(A)}}$ es el error estándar o la raíz cuadrada de la variancia esperada.

El verdadero valor de $p(B)$ se puede estimar mediante:

$$p(B) \pm \left| z_{\alpha/2} \sqrt{\frac{p(B)[1-p(B)]}{(n-k)p(A)}} \right| \qquad [14]$$

Si la operación se repite para cada retardo y categoría se obtiene una banda de confianza. Si la probabilidad observada no excede de las bandas de confianza se concluye que la aparición de la secuencia en cuestión es aleatoria, mientras que si excede por encima, o por debajo, de los límites fijados se puede hablar de dependencia excitatoria (positiva), o inhibitoria (negativa) según el caso.

Allison y Liker (1982) introdujeron una pequeña modificación en los cálculos anteriores. Estos autores basaron su modificación en que el error estándar de la diferencia entre lo esperado y lo observado sería correcto si la probabilidad de la categoría *B* (la que acompaña al criterio *A*) fuese el *valor real*, en lugar de uno estimado a partir de la muestra de observaciones.

Para corregir este defecto, estos autores proponen introducir el producto [Picture] en el cálculo del error estándar:

$$EE = \sqrt{\frac{p(B)[1-p(B)][1-p(A)]}{(n-k)\,p(A)}} \quad [15]$$

Como resultado aparece una prueba z menos restrictiva (presentada anteriormente en la ecuación [7]):

$$z^* = \frac{p(B_{t+k}/A) - p(B)}{\sqrt{\frac{p(B)[1-p(B)][1-p(A)]}{(n-k)\,p(A)}}} \quad [7]$$

Esta prueba había sido propuesta anteriormente por Gottman (1980). Este autor opina que no es necesario aplicar z^*. El estadístico propuesto por Sackett es asintóticamente correcto, ya que cuando n aumenta la estimación muestral de $p(B)$ mejora al parámetro poblacional. Por tanto, z^* no supone una crítica seria al estadístico de Sackett (Gottman y Roy, 1990).

Cuando es imposible suponer que las muestras de secuencias obtenidas siguen una distribución binomial, incluso si el sistema continúa teniendo sólo dos categorías, debido a que el muestreo sea exhaustivo, no queda otra alternativa que acogerse a la *ley de probabilidad hipergeométrica.* Wampold y Margolin (1982) propusieron un estadístico basado en la distribución hipergeométrica mediante una expresión que incluye la z de Sackett (1979):

$$z_H = \frac{z}{\sqrt{\frac{(n - f_{ij})}{(n - 1)}}} \quad [16]$$

donde f_{ij} es el número de casos en que el estado j puede ocurrir después de cada estado i.

z_H se encuentra relacionado también con la prueba χ^2 de Pearson (expresión [8]) mediante la siguiente fórmula (Wampold y Margolin, 1982: 759):

$$z_H^2 = \left(\frac{n-1}{n}\right)\chi^2 \quad [17]$$

En general, z_H es también menos restrictivo que la z propuesta por Sackett (1979).

Como mantiene el propio Sackett (1979: 642) la técnica de retardos puede ser aplicada en cualquier situación de investigación en que las categorías sean medidas en una secuencia ordenada de eventos o de tiempo. Un gran número de autores coinciden en la preferencia de utilizar la técnica de los retardos de Sackett frente a las cadenas de Markov (Gottman y Notarius, 1978; Anguera, 1983; Bakeman y Gottman, 1986).

En primer lugar, la técnica de retardos es más parsimoniosa y produce más información que las cadenas de Markov. Mediante la primera se pueden obtener medidas de contingencia entre conductas lejanas tanto en orden, para eventos secuenciales, como en el tiempo, si se tiene en cuenta la duración de los eventos. Incluso facilita la obtención de medidas directas de ciclicidad para una única situación categorizada (autocontingencia) o relaciones de fase entre varias (contingencia cruzada).

En segundo lugar, es preciso señalar que la suposición de que un estado, o una situación categorizada, de un sistema en el tiempo t+1 dependa única y exclusivamente del estado que se produce en el tiempo *t* resulta limitada. Esta suposición, en la base de las cadenas markovianas, supone además contrastar la misma hipótesis respecto de la información que se va ganando para cada estado, lo que dificulta la obtención de patrones y otros fenómenos identificables mediante la técnica de Sackett (Anguera, 1983).

Una razón práctica también se alía con el análisis secuencial propuesto en la técnica de retardos ya que resulta más manejable y más apropiado para el análisis de datos categorizados en tiempo discreto y los programas y paquetes estadísticos lo hacen actualmente más factible (Gottman y Notarius, 1978).

4. TAMAÑO DE LA MUESTRA Y ERROR TIPO I

En el momento de diseñar el plan de una investigación observacional que va a aplicar un análisis secuencial de retardos, es preciso tener en cuenta dos cuestiones. La primera de ellas hace referencia al número de datos, o secuencias (de evento o de tiempo), que como mínimo es necesario disponer. El motivo que fundamenta lo anterior es justificar la significación de la prueba *z* utilizada en base a una aproximación de la distribución binomial (o hipergeométrica) a la normal. La segunda es el problema que se produce al trabajar con una técnica de análisis univariada. A medida que el número de pruebas estadísticas aumenta el error tipo I (riesgo *a*) asociado a cada una de ellas se dispara peligrosamente, por lo que es necesario establecer un plan previo de control de dicho error.

4.1. Determinación del tamaño (*n*) de la muestra de observaciones

Siegel (1956) sugirió que la variancia de la distribución binomial (*npq*) debería de valer al menos 9 para hablar de una aproximación a la normal. Adaptando esta regla deducida empíricamente a la perspectiva secuencial Bakeman y Gottman

(1986: 178) propusieron utilizar la siguiente expresión para determinar el número suficiente de secuencias:

$$n_s = \frac{9}{p(1-p)} + L - 1 + I \qquad [18]$$

donde L es la longitud de la secuencia.
I es el número de interrupciones en la secuencia.

La única dificultad se encuentra en determinar el valor de p. Como señalan Bakeman y Gottman (1986), dos tipos de aspectos afectan al valor de p: Si las categorías son equiprobables o no y sí se admite, o no, que los *códigos adyacentes* puedan ser iguales (ESD repetibles o ESD no repetibles según el estándar de Bakeman y Quera, 1992).

En primer lugar, se va a suponer que las *categorías son equiprobables* e independientes. Esta suposición es verosímil en estudios exploratorios, e incluso en confirmatorios cuando no hay evidencia de que las probabilidades incondicionales (no secuenciales) sean diferentes. La suposición de que las categorías sean equiprobables no es posible cuando por ejemplo existan categorías con una elevada *saliencia*, independientemente de que sea más o menos relevante el *núcleo categorial*. Cuando hay equiprobabilidad p está en función inversa con el número de secuencias posibles:

$$p = \frac{1}{m^L} \qquad [19]$$

donde m^L es el número de secuencias de longitud L para un sistema de categorías con m códigos, esto es, el número de variaciones con repetición de m elementos tomados de L en L.

Sustituyendo el valor de p obtenido en [19] en la expresión [18] se obtiene el *número de secuencias para admitir la aproximación normal de la binomial cuando las categorías son equiprobables y son admisibles códigos adyacentes iguales* (ESD repetibles).

$$n_s = \frac{9}{\frac{1}{m^L}\left(1 - \frac{1}{m^L}\right)} + L - 1 + I = \frac{9m^L}{\left(1 - \frac{1}{m^L}\right)} + L - 1 + I = \frac{m^L\left(9m^L\right)}{m^L\left(1 - \frac{1}{m^L}\right)} + L - 1 + I =$$

$$n_s = \frac{9m^{2L}}{m^L - 1} + L - 1 + I \qquad [20]$$

Si los códigos adyacentes deben ser diferentes, el valor de p está en función inversa de las posibles secuencias de longitud L:

$$p = \frac{1}{V_m^L} = \frac{(m-L)!}{m!} \qquad [21]$$

Y sustituyendo el valor de p en la expresión general [18] se obtiene el *número de secuencias para admitir la aproximación normal de la binomial cuando las categorías son equiprobables y no son admisibles códigos adyacentes iguales* (ESD no repetibles).

$$n_s^* = \frac{9}{\frac{1}{V_m^L}\left(1 - \frac{1}{V_m^L}\right)} + L - 1 + I = \frac{9V_m^L}{\left(1 - \frac{(m-L)!}{m!}\right)} + L - 1 + I =$$

$$n_s^* = \frac{9\, V_m^L\; m!}{m!\left(1 - \frac{(m-L)!}{m!}\right)} + L - 1 + I = \frac{9\, V_m^L\; m!}{m! - (m-L)!} + L - 1 + I \qquad [22]$$

Para el caso común de que la longitud de la secuencia sea L = 2, se puede construir la tabla 2:

Tabla 2
NÚMERO SUFICIENTE DE SECUENCIAS EN UNA INVESTIGACIÓN OBSERVACIONAL (L=2) CON M CATEGORÍAS EQUIPROBABLES

m	ns	n^*s
2	37	49
3	66	93
4	118	154
5	190	235
6	280	334
7	388	451
8	513	586
9	658	739
10	810	910

En segundo lugar, si se sospecha que las categorías no poseen probabilidades iguales de ocurrencia, y teniendo en cuenta la relación inversa entre el tamaño de la muestra y la probabilidad de las categorías (Cooley y Mao, 1981), Bakeman y Gottman (1986), propusieron aplicar la expresión [18], utilizando la secuencia de códigos de interés para la investigación menos frecuente, argumentando que si se tienen suficientes datos para esta secuencia es de esperar que se posean más que suficientes para el resto. Por ejemplo, supongamos que en un sistema de categorías la secuencia (L = 2) menos frecuente es laAB, con una distancia entre ellas en el tiempo de $t+k$. El número de secuencias necesarias vendría dado por la expresión:

$$n_d = \frac{9}{p(B_{t+k}/A)[1 - p(B_{t+k}/A)]} + L - 1 + l \qquad [23]$$

3.2. Control del error tipo I

El problema se deriva del uso de una técnica univariante para realizar múltiples contrastes entre variables. En efecto, si se realizan c comparaciones con un riesgo α fijado *a priori* para cada una de ellas, el riesgo del conjunto de las c comparaciones es superior. Por ejemplo, para T comparaciones con α = 0.05, la probabilidad de tomar la decisión correcta en todas las T pruebas, aplicando la ley multiplicativa de las probabilidades:

$$0.95 \text{ x } 0.95 \text{ x...x } 0.95 = 0.95\text{T} = (1-\alpha)^T \qquad [24]$$

Para controlar el error tipo I es preciso partir de un valor α menor, que tenga en cuenta el número de pruebas que se van a realizar. Esto se consigue con el procedimiento de Bonferroni cuya expresión es la siguiente:

$$\alpha_\beta = \frac{\alpha}{T} \qquad [25]$$

Donde α es el nivel máximo de error tipo uno que se desea admitir, y t el número de total de pruebas que se van ha realizar simultáneamente. T depende del número de categorías (m) que tenga el sistema y de la longitud de la secuencia (L):

$$T = V_m^L \qquad [26]$$

Así, para α = 0.05 (bilateral) y L = 2, es posible construir la tabla 3:

Tabla 3
ERROR TIPO I CORREGIDO Y PUNTUACIÓN ZC ASOCIADA SEGÚN LA DISTRIBUCIÓN NORMAL ESTANDARIZADA PARA T PRUEBAS Y M CATEGORÍAS

m	T	α_β	z_c
2	2	0,025	2,24
3	6	0,0083	2,64
4	12	0,0041	2,86
5	20	0,0025	3,01
6	30	0,0016	3,12
7	42	0,0012	3,22
8	56	0,00089	3,295
9	72	0,00069	3,355
10	90	0,00055	3,40

REFERENCIAS

ALLISON, P. D. y LIKER, J. K. (1982): Analyzing Sequential Categorial data on Dyadic Interaction: A Comment on Gottman. *Psychological Bulletin,* 91 (2), 393-403.

ANGUERA, M. T. (1983): *Manual de Prácticas de Observación.* México: Trillas.

ANGUERA, M. T. (1991): (Ed.) *Metodología observacional en investigación psicológica. Vol. I. Fundamentación (1).* Barcelona: PPU.

ANGUERA, M. T. (1993): (Ed.) *Metodología observacional en investigación psicológica. Vol. II. Fundamentación (2).* Barcelona: PPU.

BAKEMAN, R. (1978): Untangling streams of behavior: Sequential analyses of observation data. En G. P. Sackett (Ed.) *Observing Behavior: Data collection and analysis methods. Vol. II.* (pp. 63-78). Baltimore: University of Park Press.

BAKEMAN, R. (1991b): Prologue. En M. T. Anguera (Ed.): *Metodología observacional en investigación psicológica.* (pp. 13-17). Barcelona: PPU.

BAKEMAN, R., ADAMSON, L. B. y STRISIK, P. (1989): Lags and Logs: Statistical Aproaches to Interaction. En M. H. Bornstein y J. Bruner (Eds.): *Interaction in Human Development* (pp. 241-260). Hillsdale, N. J.: Erlbaum.

BAKEMAN, R. y DABBS, J. M. (1976): Social interaction observed: Some approaches to the analysis of behavior streams. *Journal of Social and Personality Behavior,* 2, 335-345.

BAKEMAN, R. y GOTTMAN, J. M. (1986): *Observing Interaction: An Introducing to Sequential Analysis.* Cambridge: Cambridge University Press. (Trad. al castellano con el título *Observación de la interacción: Introducción al análisis secuencial,* Madrid: Morata, 1989).

BAKEMAN, R. y QUERA, V. (1992): SDIS: A Sequential Data Interchange Standard. *Behavior Research Methods, Instruments, & Computers,* 24 (4), 554-559.

COOLEY, W. W. y MAO, B. (1981): The sample of classroom time observed. *Journal of Classroom Interaction,* 17 (1), 31-36.

FIENBERG, S. E. (1980): *The analysis of cross-classified categorial data* (2nd Ed.) Cambridge: MIT Press.

GOTTMAN, J. M. (1980): On analyzing for sequential connection and assessing interobsever reliability for the sequential analysis of observational data. *Behavioral Assessment,* 2, 361-368.

GOTTMAN, J. M. y BAKEMAN, R. (1979): The sequential analysis of observational data. En M. E. Lamb, S. J. Suomi y G. R. Stephenson (Eds.): *Social interaction analysis: Methodological issuses.* Madison: University of Wisconsin Press.

GOTTMAN, J. M. y NOTARIUS, C. (1978): Sequential Analysis of Observational Data Using Markov Chains. En T. R. Kratochwill (Ed.): *Single Subject research: Strategies for Evaluating Change.* (pp. 237-285) London: Academic Press, Inc.

GOTTMAN, J. M. y ROY, A. K. (1990): *Sequential analysis. Aguide for behavioral researchers.* Cambridge: Cambridge University Press.

QUERA, V. (1991): Muestreo y Registro Observacional. En M. T. ANGUERA (Ed.): *Metodología observacional en la investigación psicológica.* Vol I. Fundamentación (1). (pp. 241-329). Barcelona: PPU.

SACKETT, G. P. (1978): Measurement in observational research. En G. P. Sackett (Ed.): *Observing behavior. Vol. 2: Data collection and analysis methods.* (pp. 25-43): Baltimore: University Park Press.

SACKETT, G. P. (1979): The Lag Sequential Analysis of Contingency and Cyclicity in Behavioral Interaction Research. En J. D. Osofsky (Ed.): *Handbook of Infant Development.* (pp. 623-649) New York: Wiley.

SIEGEL, S. (1970): *Estadística no paramétrica.* México: Trillas. (Versión original de 1956).

SIMON, A. y BOYER, E. G. (1967). (Eds.): *Mirrors for behavior: An anthology of Classroom Observation Instruments.* Philadelphia: Research for Better Schools, Inc.

SIMON, A. y BOYER, E. G. (1970). (Eds.): *Mirrors for behavior II: An anthology of Classroom Observation Instruments.* 2 vol. Philadelphia: Classroom Interaction Newsletter & Research for Better Schools, Inc.

TÓJAR, J. C. (1993): *Concordancia del registro observacional en datos secuenciales. Investigación aplicada en el contexto del aula.* Málaga: Secretariado de Publicaciones de la Universidad de Málaga.

TÓJAR, J. C. (1994): *Concordancia en los registros de observación. Calidad de la investigación educativa en Metodología Observacional.* Barcelona: PPU.

WAMPOLD, B. E. y MARGOLIN, G. (1982): Nonparametric strategies to test the independence of behavioral states in sequential data. *Psychological Bulletin,* 92 (3), 217-225.