

### 3. ARTÍCULOS DE APLICACIÓN

#### ANÁLISIS DE SERIES TEMPORALES DE PACIENTES CON TRASTORNOS DE LA CONDUCTA ALIMENTARIA

María Isabel Parra Arévalo

Departamento de Matemáticas

Universidad de Extremadura

##### 1. Introducción

Este trabajo es consecuencia del estudio de algunos problemas planteados desde la Unidad de Psiquiatría del Complejo Hospitalario Infanta Cristina de Badajoz. Se trataba de modelar y caracterizar la sintomatología de distintas alteraciones psiquiátricas, así como valorar la acción terapéutica, para poder disponer de medios objetivos de diagnóstico y predicción de la evolución clínica.

Nos centramos en el estudio de los trastornos de la conducta alimentaria, cuyos principales diagnósticos son la anorexia nerviosa (AN) y la bulimia nerviosa (BN). Tienen una gran relevancia socio-sanitaria, por su peculiaridad, gravedad, interés social, complejidad y dificultad de diagnóstico y tratamiento. Además, pueden llegar a constituir cuadros de extrema gravedad si no se tratan precozmente, lo que ha provocado una gran alarma social. Aunque la enfermedad no es nueva, el interés por estos temas es reciente.

El objetivo de este trabajo es poner de manifiesto la utilidad de la metodología estadística, especialmente el análisis de series temporales, en el diagnóstico y pronóstico de tales trastornos.

##### 2. Tratamiento estadístico del problema

Los datos disponibles provienen de pacientes diagnosticados de algún trastorno de la conducta alimentaria según los criterios establecidos en el Diagnostic and Statistical Manual of Mental Disorders (véase [1]) y de sujetos control. Todos los pacientes fueron tratados en el mismo contexto, recibiendo igual tratamiento farmacológico, programa dietético y psicoterapia de orientación cognitivo-conductual. Los pacientes ingresan en el hospital cuando su bajo peso hace peligrar su salud y salen cuando su peso corporal, estabilización analítica y evolución lo permiten.

La variable fundamental de análisis elegida ha

sido el peso diario, por tratarse de una variable física que se puede medir de forma precisa, objetiva y resulta ser suficientemente representativa de la enfermedad. Para los pacientes hospitalizados, las medidas fueron tomadas todos los días, por la mañana en ayunas, aproximadamente a la misma hora, por el personal sanitario, con una báscula de precisión. Los sujetos con tratamiento ambulatorio y de control, llevaron a cabo ellos mismos idéntico protocolo. En consecuencia, podrían estar sujetos a mayor error.

En la Figura 1 mostramos, de forma resumida, las medidas terapéuticas a desarrollar por los profesionales, según sea la evolución del peso de un paciente con anorexia nerviosa a lo largo de un ingreso hospitalario.

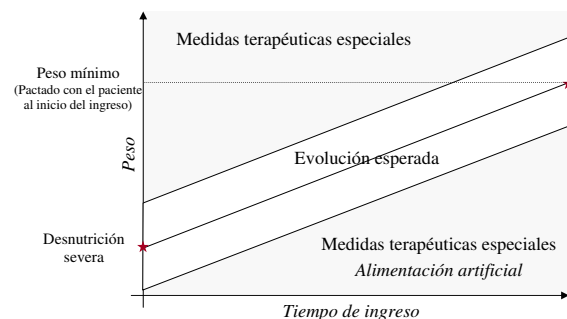


Figura 1: Medidas terapéuticas generales.

Trabajar con datos reales implica una serie de limitaciones generales, acrecentadas en nuestro caso por cuestiones éticas. Se trata pues de series temporales cortas, ruidosas, medidas, a veces, a intervalos irregulares y con comportamiento no estacionario. Por este motivo hemos de utilizar la metodología con cierta precaución.

Con el objeto de realizar estudios comparativos, trabajaremos con valores del peso normalizados. Es decir

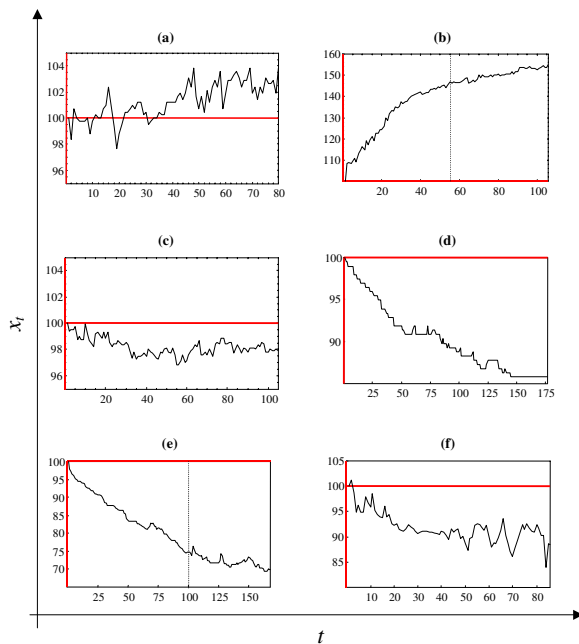
$$x_t^i = \frac{w_t^i}{w_1^i} \times 100,$$

donde  $w_t^i$  representa el peso registrado para el sujeto  $i$ -ésimo en el  $t$ -ésimo día de seguimiento siendo  $w_1^i$  su peso inicial. En consecuencia, disponemos de las series

$$\{x_t^i\}_{t=1}^{N_i}, \quad i = 1, 2, \dots, N_c,$$

donde  $N_i$  denota la longitud de la serie correspondiente al  $i$ -ésimo paciente, siendo  $N_c$  el número total de casos analizados.

Como ilustración, en la Figura 2 representamos una serie temporal típica de cada uno de los distintos patrones de comportamiento que hemos encontrado en el total de series estudiadas.



**Figura 2:** Ejemplos de patrones de comportamiento en pacientes con trastornos de la conducta alimentaria: (a) paciente diagnosticado de AN, con alimentación oral, (b) paciente diagnosticado de AN, con alimentación artificial durante el primer periodo de su ingreso hospitalario (hasta la línea vertical) y por vía oral a continuación, (c) sujeto sano de control, (d) sujeto diagnosticado de trastorno facticio bajo dieta de adelgazamiento, (e) paciente con antecedentes de AN durante su adolescencia, bajo dieta de adelgazamiento por excesivo peso, (f) paciente diagnosticado de BN con tratamiento ambulatorio.

Cabe señalar que en todas las secuencias relativas a pacientes se aprecian ciertas tendencias o cambios de nivel a lo largo del tiempo. Sólo las series de los sujetos control, sin ningún tipo de restricción

en su alimentación, resultan ser aproximadamente constantes en media. En los casos de AN las series son ascendentes, con una pendiente tanto más pronunciada cuanto mayor es el control al que está sometido el paciente. Cuando el paciente recibe alimentación artificial, a través de sonda nasogástrica, se aprecia un fuerte cambio de nivel que se amortigua cuando pasa a la alimentación oral. Las secuencias de sujetos control sometidos a dieta estricta presentan una evolución simétrica al caso de alimentación artificial. Finalmente, los casos de BN se asemejan a la imagen simétrica de AN, pero con fluctuaciones relativamente mayores.

Con objeto de solucionar una de las principales limitaciones de nuestro estudio: el reducido número de series temporales disponibles, hemos desarrollado dos algoritmos para obtener series temporales de hipotéticos sujetos, con características similares a las reales (véase [2]).

En ambos casos, comenzamos realizando una diferencia de orden 1, para eliminar la tendencia. Las series transformadas son estacionarias.

El primer algoritmo está basado en la función de distribución empírica de los incrementos de peso diario. Simulamos una secuencia de números aleatorios distribuidos según dicha función, fijamos un peso de ingreso, y deshacemos la transformación, obteniendo series temporales similares a las originales.

El segundo está basado en reglas lógicas:  $R_1, \dots, R_N$ , donde  $N = \max\{N_1, \dots, N_{N_s}\}$ , siendo  $N_s$  el número de series temporales disponibles en el grupo correspondiente (AN o control). Cada regla,  $R_t$ , se construye utilizando la información de las series cuya longitud es mayor o igual que  $t$ . Tratan de modelar la relación que existe entre los incrementos de peso en  $t-1$  y los de  $t$ . Como caso particular,  $R_1$  modela la relación entre la longitud del ingreso y el incremento de peso para el primer día. En este caso, el programa es determinista, y podrá proporcionarnos únicamente  $N$  trayectorias, es decir, obtendremos únicamente una secuencia de datos para cada tamaño.

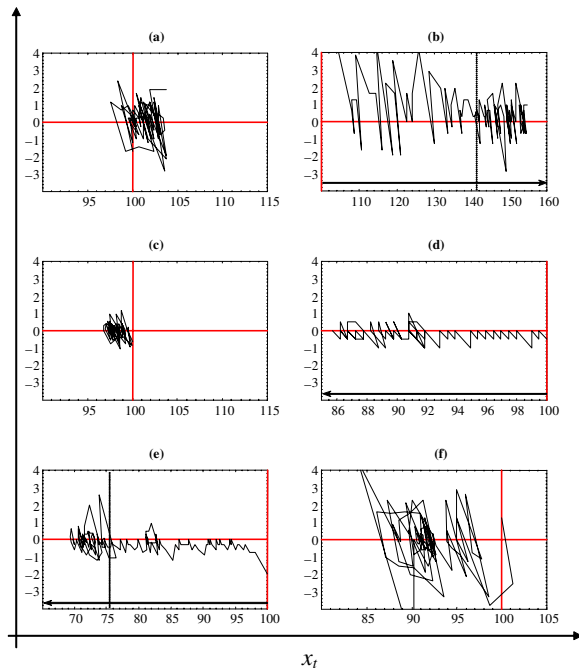
En la búsqueda de un modelo, como primera aproximación, teniendo en cuenta el protocolo de tratamiento de la enfermedad, podríamos asumir un modelo lineal. Sin embargo, puesto que el estudio de los residuos indica que aún queda información en ellos que podríamos reproducir a través de un

modelo AR(1), hemos considerado como modelo:

$$x_{t+1} = \underbrace{\mu + m(t+1)}_{\text{modelo lineal}} + \underbrace{\phi(x_t - (\mu + mt))}_{\text{fluctuaciones sobre la recta}} + \underbrace{\varepsilon_t}_{\text{residuos}}$$

Se trata de un modelo efectivo y parsimonioso<sup>1</sup>. Su sencillez permite interpretar sus coeficientes para aplicaciones prácticas,  $\mu$  y  $m$  son similares a la ordenada en el origen y la pendiente (incremento de peso diario) del modelo lineal, mientras que  $\phi$  representa las fluctuaciones sobre dicho modelo lineal. Su interpretación en la práctica clínica no resulta ser trivial ya que al aparecer multiplicado por  $(x_t - (\mu + mt))$  introduce cierta confusión.

El espacio de fases o espacio de estados constituye una manera muy eficaz de observar gráficamente toda la dinámica de un sistema de una sola vez, ya que contiene todos los aspectos de la dinámica del sistema en sus ejes. Para nuestro problema es suficiente considerar un plano. Basta representar el peso normalizado,  $x_t$ , frente a la variación de peso,  $x'_t = x_t - x_{t-1}$ , para apreciar con gran claridad los distintos patrones de evolución (véase Figura 3).



**Figura 3:** *Espacio de fases bidimensional donde se representa geoméricamente la evolución de las series temporales mostradas en la Figura 2. Las líneas verticales discontinuas marcan cambios significativos en la alimentación del sujeto y las flechas horizontales el sentido de la evolución.*

Como primera aproximación, podríamos distinguir dos tipos de comportamiento claramente diferenciados. Los sujetos sometidos a un control estricto, casos (b), (d) y (e) con trayectorias que recorren un amplio intervalo de valores sobre el eje horizontal y sólo llegan a enredarse cuando se relaja el control, línea vertical en casos (b) y (e). La principal diferencia entre estos casos es el sentido de la evolución. En los casos de AN con alimentación por sonda nasogástrica, (b), la trayectoria se mueve de izquierda a derecha, y para los controles a dieta, casos (d) y (e), de derecha a izquierda. Por el contrario, los pacientes con trastornos de la conducta alimentaria que reciben alimentación por vía oral, así como los controles sanos, muestran trayectorias que se enredan sobre sí mismas una y otra vez. En este caso, la principal diferencia que puede apreciarse a simple vista, entre pacientes y controles, se basa en el tamaño de la región del plano que ocupan. Esta región es mucho menor en el caso de los controles.

### 3. Caracterización de las series temporales

Cabe preguntarse, si sería posible analizar tales diagramas desde el punto de vista de la teoría del caos. Si pudiéramos determinar el exponente de Lyapunov máximo ( $\lambda$ ) o la entropía de Kolmogorov-Sinai por unidad de tiempo ( $h_{KS}$ ) de la secuencia de datos, conoceríamos, sin incertidumbre, si la serie está generada por una ley determinista ( $\lambda, h_{KS} < \infty$ ) o estocástica ( $\lambda, h_{KS} \rightarrow \infty$ ). En la práctica los problemas para calcular estas medidas, que se definen como medias en el tiempo infinito, y en el límite para una resolución arbitrariamente fina, son evidentes, teniendo en cuenta que disponemos de un número finito (y en nuestro caso muy limitado) de rango de escalas. Para salvar estas limitaciones, hacemos uso del exponente de Lyapunov de tamaño finito,  $\lambda(\varepsilon)$ , como una variante del exponente de Lyapunov, la  $(\varepsilon, \tau)$ -entropía por unidad de tiempo,  $h(\varepsilon, \tau)$ , como una generalización de la entropía de Kolmogorov-Sinai, y la redundancia  $r(\varepsilon)$ , que mide la cantidad de incertidumbre en observaciones futuras que puede ser eliminada con el conocimiento del pasado (para más información véase [4]).

Cuando estimamos la entropía y la redundancia de tamaño finito por unidad de tiempo en el problema que nos ocupa, encontramos que se trata de

<sup>1</sup>Hemos elegido aquél, entre los modelos válidos, que minimiza el Akaike Information Criterion [3].

secuencias deterministas, puesto que la primera es constante y la segunda proporcional a  $-\log \varepsilon$ , para una región del rango de escalas (para más detalles véase [5]).

Existen diferentes vías para cuantificar la autosimilitud de objetos geométricos a partir de la dimensión (el lector interesado puede consultar [6], [7] ó [8] para más información al respecto). Por supuesto, en todos los casos, se requiere que la definición coincida con la noción usual de dimensión, cuando sea evaluada sobre objetos no-fractales.

Supongamos que el objeto está cubierto por un conjunto de  $n(\varepsilon)$  hipercubos de lado  $\varepsilon$ , y sea  $p_i$  la probabilidad de que un punto se encuentre en el  $i$ -ésimo hipercubo. La definición general de dimensión de orden  $q$  viene dada por la expresión:

$$d^{(q)} = \frac{1}{q-1} \lim_{\varepsilon \rightarrow 0} \left( \frac{\log \sum_{i=1}^{n(\varepsilon)} p_i^q}{\log \varepsilon} \right).$$

(véase, [9]). Los casos particulares  $d^{(0)}$ ,  $d^{(1)}$  y  $d^{(2)}$  corresponden a la dimensión de capacidad ( $d_C$ ), la dimensión de información ( $d_I$ ), y la dimensión de correlación ( $d_G$ ), respectivamente.

La dimensión de capacidad nos proporciona información acerca de cómo cambia el número de hipercubos necesario para cubrir el objeto cuando su lado  $\varepsilon$  decrece. Puede calcularse a partir de la expresión:

$$d_C = \lim_{\varepsilon \rightarrow 0} \left( \frac{\log n(\varepsilon)}{\log 1/\varepsilon} \right).$$

La dimensión de información se obtiene como:

$$d_I = \lim_{\varepsilon \rightarrow 0} \left( \frac{\sum_{i=1}^{n(\varepsilon)} p_i \log p_i}{\log \varepsilon} \right) = \lim_{\varepsilon \rightarrow 0} \left( -\frac{I(\varepsilon)}{\log \varepsilon} \right),$$

donde  $-I(\varepsilon)$  es la pérdida de información o entropía. Esta dimensión mide la variación en el exponente de escalado de la entropía con  $\varepsilon$ .

Finalmente, la dimensión de correlación se define como:

$$d_G = \lim_{\varepsilon \rightarrow 0} \left( \frac{\log \sum_{i=1}^{n(\varepsilon)} p_i^2}{\log \varepsilon} \right).$$

La dimensión fractal ha sido propuesta como una medida de gran utilidad para la caracterización de series temporales fisiológicas (véase [10],

[11] ó [12]). Uno de los principales problemas de esta aproximación es la dificultad para obtener series temporales estacionarias, suficientemente largas para determinar la dimensión fractal “real” (véase por ejemplo [13], [14] ó [15]). Nosotros defendemos la postura de que es posible hacer estimaciones a partir de conjuntos de datos pequeños (véase [2]). Utilizando series temporales de diferentes longitudes, mostramos que existe una relación monótona entre la estimación de la dimensión fractal y el número de datos<sup>2</sup>, por lo que puede permitírnos realizar comparaciones entre las estimaciones obtenidas para distintas secuencias de datos, aunque éstas sean cortas.

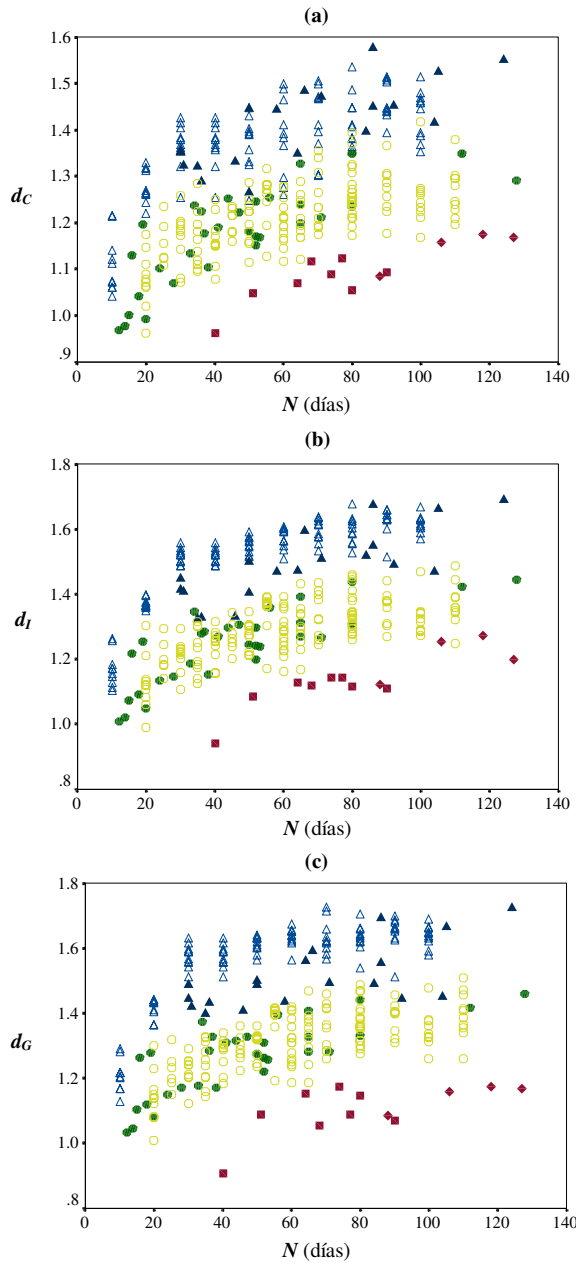
Desde que comenzamos a trabajar con la idea de que la dimensión fractal podía servir para caracterizar la evolución de las series temporales del peso en pacientes con trastornos de la conducta alimetaria, los resultados fueron muy halagüeños (véase [17] ó [18]). Sin embargo, hay que ser conscientes de las serias limitaciones a las que nos enfrentamos, por lo que tratamos de ser muy cautelosos a la hora de estimar la dimensión fractal. Para estar más seguros de que la metodología empleada es la adecuada, hemos realizado una serie de comprobaciones.

En primer lugar, hemos estimado la dimensión a lo largo del tiempo,  $t$ , para cada una de las series temporales correspondientes a pacientes con trastornos de la conducta alimentaria, y para series simuladas mediante un modelo lineal y para otras simuladas a partir de secuencias de números aleatorios. En todos los casos, las estimaciones de la dimensión fractal correspondientes a series de pacientes se mantienen acotadas inferiormente por las estimaciones relativas a las series lineales simuladas y superiormente por las de las series aleatorias simuladas, siempre que el número de datos considerados sea mayor que 30.

En segundo lugar, hemos realizado estimaciones a partir de distintas definiciones de la dimensión ( $d_C$ ,  $d_I$  y  $d_G$ ) para estudiar la robustez de los resultados.

En la Figura 4 hemos representado los resultados obtenidos, en relación a las estimaciones de la dimensión fractal, tanto para las series proporcionadas por sujetos reales como para las simuladas.

<sup>2</sup>Otros autores llegaron a la misma conclusión, para estudiar las diferencias en la complejidad de señales de electroencefalogramas para series desde 128 a 2048 datos [16].



**Figura 4:** Resultados obtenidos en la estimación de las dimensiones de capacidad (a), información (b) y correlación (c), para las series temporales de pacientes con AN ( $\bullet$ ), series simuladas entrenadas con las anteriores ( $\circ$ ), sujetos de control ( $\blacktriangle$ ), series simuladas entrenadas con las anteriores ( $\triangle$ ), pacientes con otro tipo de patología sujetos a una dieta estricta ( $\blacksquare$ ), y pacientes con anorexia nerviosa recibiendo alimentación a través de sonda nasogástrica ( $\blacklozenge$ ), representados frente a la longitud de la serie temporal.

Como es lógico, cuando estimamos la dimensión

de los posibles atractores en el plano de fases, encontramos que en todos los casos estudiados los valores se encuentran entre 1 y 2. Para los casos correspondientes a sujetos con un control estricto de su alimentación, los valores se encuentran próximos a la unidad. Este hecho confirma que dichas series puedan ser aproximadas por un modelo lineal. Por otra parte, sería lógico pensar que un sujeto con una alimentación sana tuviera un peso constante, lo que en el diagrama de fases quedaría representado por un punto, cuya dimensión es cero. Sin embargo, en la realidad, ocurre que está teniendo pequeñas variaciones de peso, causadas por múltiples factores, provocando que su comportamiento parezca aleatorio. En estos casos, es cuando las estimaciones de la dimensión toman los valores mayores, puesto que las trayectorias fásicas tienden a llenar completamente una pequeña región del espacio de fases. Finalmente, las series correspondientes a pacientes con trastornos de la conducta alimentaria muestran comportamientos intermedios a los dos anteriores. Dentro de este grupo, las estimaciones de la dimensión toman valores tanto mayores cuanto más irregular ha sido el comportamiento del paciente durante su seguimiento.

Se aprecia una clara diferencia entre los valores de la dimensión estimados para los distintos grupos de series, independientemente de la definición utilizada ( $d_C$ ,  $d_I$  ó  $d_G$ ). En la banda inferior se sitúan los correspondientes a las series de sujetos con un control estricto (dieta o alimentación artificial por sonda nasogástrica). En una banda intermedia aparecen los correspondientes a los distintos ingresos de pacientes con trastornos de la conducta alimentaria. Finalmente, en la banda superior se sitúan los relativos a las series de sujetos control. La clara diferencia entre estos grupos nos hace conjeturar si la dimensión fractal podría ser una medida objetiva de apoyo en el tratamiento de la enfermedad, al permitir caracterizar la evolución del peso durante el seguimiento.

Finalmente, siguiendo la estrategia que describíamos en la sección anterior, hemos estimado la dimensión para series temporales simuladas tanto a partir de las disponibles para sujetos con AN como a partir de las de sujetos control. Los resultados, muestran que los valores obtenidos se mueven en el mismo rango que los de las series originales correspondientes.

#### 4. Conclusiones

Hemos realizado un análisis exhaustivo de las series temporales del peso diario de pacientes diagnosticados de trastornos de la conducta alimentaria, comparándolas con las de sujetos control, persiguiendo el objetivo de encontrar resultados que sean útiles en la práctica clínica. Para conseguirlo, hemos utilizado desde herramientas estadísticas tradicionales hasta medidas de sistemas dinámicos no lineales.

Como primera aproximación, hemos utilizado herramientas estadísticas tradicionales de modelación. El mejor modelo lineal que hemos encontrado para representarlas es un  $AR(1)$  con constante y el tiempo como variable explicativa.

Para solucionar el limitado número de casos disponibles, hemos diseñado una estrategia de simulación que permite obtener un número ilimitado de secuencias con propiedades similares a las originales.

Dando un enfoque original al análisis, hemos estudiado la evolución del peso en el diagrama de fases, caracterizando dicha evolución a partir de la dimensión fractal del atractor. Las estimaciones de la dimensión fractal para los diagramas de fase de sujetos con un control estricto de su alimentación toman los valores más pequeños, próximos a la unidad. Justo lo contrario que ocurre para los sujetos sanos, sin ningún tipo de forzamiento respecto a su ingesta, ya que toman los valores más altos. Los pacientes con trastornos de la conducta alimentaria presentan valores intermedios de la dimensión, tanto mayores cuanto más irregular es su comportamiento. Finalmente, para las series temporales simuladas, obtenemos resultados comparables a los del grupo que sirvió de base para la simulación.

Para comprobar la consistencia y robustez de los resultados obtenidos en la estimación de la dimensión fractal estimamos su valor utilizando distintas definiciones ( $d_C$ ,  $d_I$  y  $d_G$ ), encontrando resultados comparables, y calculamos su valor a lo largo del tiempo obteniendo que a partir de un tamaño  $\geq 30$  podemos caracterizar el comportamiento de los pacientes con trastornos de la conducta alimentaria a partir de su dimensión fractal.

Como resultado de la investigación, conjeturamos que la dimensión box-counting resulta ser una medida de gran utilidad para caracterizar la evo-

lución de pacientes con trastornos de la conducta alimentaria, ya que puede ayudar al terapeuta a adoptar estrategias de tratamiento.

#### Referencias

- [1] APA (1994). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*. American Psychiatric Press: Washington.
- [2] Cuadros F., Parra M.I., Salcedo M.S. y Gutiérrez J.R. (2007). Weight time series of anorexic inpatients. *International Journal of Bifurcation & Chaos*, enviado para publicación.
- [3] Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**: 716-723.
- [4] Cohen A. y Procaccia I. (1985). Computing the Kolmogorov entropy from time signals of dissipative and conservative dynamical systems, *Physical Review A*, **31**: 1872-1882.
- [5] Parra M.I., Cuadros F., Salcedo M.S. y Gutiérrez J.R. (2000). ¿Es caótica la evolución del peso en pacientes con anaorexia nerviosa? *No-Linear 2000*.
- [6] Baker G.L. y Gollub J.P. (1990). *Chaotic Dynamics: an introduction*. Cambridge: New York.
- [7] Schroeder M. (1991). *Fractal, Chaos, Power Laws*. Freeman: New York.
- [8] Eke A., Herman P., Kocsis L y Kozak L.R. (2002). Fractal characterization of complexity in temporal physiological signals, *Physiological Measurement*, **23**: 1-38.
- [9] Atmanspacher H., Scheingraber H. y Voges W. (1988). Global scaling properties of chaotic attractor reconstructed from experimental data, *Physical Review A*, **34**: 11314-1322.
- [10] Elbert T., Ray W.J., Kowalik Z.J., Skinner J.E., Graf K.E. y Birbaumer N. (1994). Chaos and physiology: deterministic chaos in excitable cell assemblies, *Physical Review*, **74**: 1-47.
- [11] Kim Y.M., Kriehle K.K., Kim C.M., Reed J. y Rae-Grant A.D (1995). Differentiation of alpha coma from awake alpha by nonlinear dy-

- namics of electroencephalography, *Electroencephalography and Clinical Neurophysiology*, **98**: 35-41.
- [12] Lutzenberger W., Elbert T., Birbaumer N., Ray W.J. y Schupp H.T. (1992). The scalp distribution of fractal dimension of the EEG and its variation with mental task, *Brain Topography*, **5**: 27-34.
- [13] Eckmann J.P. y Ruelle D. (1993). Fundamental limitations for estimating dimension and Lyapunov exponents in dynamical systems, *Physica D*, **56**: 185-187.
- [14] Procaccia I. (1988). Complex or just complicated, *Nature*, **333**: 489-499.
- [15] Smith L.A. (1988). Intrinsic limits on dimension calculation, *Physics Letters A*, **133**: 283-288.
- [16] Preißl H., Lutzenberger W., Pulvermüller F. y Birbaumer N. (1997). Fractal dimensions of short time series in humans, *Neuroscience Letters*, **225**: 77-80.
- [17] Parra M.I., Cuadros F., Salcedo M.S. y Gutiérrez J.R. (1998). Chaotic dynamics of weight time series in anorexic inpatients. *XXth IUPAP International Conference on Statistical Physics*.
- [18] Parra M.I., Cuadros F., Salcedo M.S. y Gutiérrez J.R. (1999). Caracterización del caos en series temporales de interés psicológico. En: Montenegro R., Montero G. y Winter G. (Eds.), *Actas del XVI y VI CMA*, 1261-1268.