
ANÁLISIS DE LA VARIABILIDAD DE LOS SISTEMAS COMPLEJOS UTILIZANDO ENTROPÍA APROXIMADA
COMPLEX SYSTEMS VARIABILITY ANALYSIS USING APPROXIMATE ENTROPY

Eduardo Cuestas

Resumen:

Los sistemas biológicos son sistemas altamente complejos, tanto espacial como temporalmente. Los mismos están cimentados en una red dinámica interconectada marcadamente interdependiente, redundante y pleiotrópica. Las propiedades de un sistema son distintas de las de sus partes, aunque ellas dependen de la integridad del todo. Las propiedades sistémicas desaparecen cuando el sistema se rompe, mientras que las propiedades de sus componentes se mantienen. La enfermedad puede entenderse como una alteración funcional sistémica del organismo humano, y se presenta con una enorme variabilidad en los patrones de severidad, estabilidad y duración. Los sistemas biológicos complejos se caracterizan por presentar ritmos o ciclos medibles y objetivables, los ritmos anormales están asociados a la enfermedad y pueden estar involucrados en su patogénesis, este fenómeno se denomina "enfermedad dinámica". Los médicos hace más de 20 siglos que reconocen la asociación de las alteraciones de los ritmos fisiológicos con la presencia de enfermedad. La medición de los valores absolutos de los parámetros clínicos es una fuente altamente significativa y relevante de datos sobre el estado de los pacientes, sin embargo la medición de la variabilidad de estos ofrece una muy valiosa información adicional. El objetivo de esta revisión fue estudiar uno de los avances más recientes en la medición y caracterización de la variabilidad biológica posibilitados por el desarrollo de modelos matemáticos basados en la teoría del caos y de las dinámicas no lineales, como el cálculo de la entropía aproximada, que ofrece una herramienta para poder discernir el significado de las diferencias en la variabilidad de los signos biológicos en diferentes grupos de pacientes.

Palabras clave: sistemas complejos, entropía aproximada, variabilidad.

Abstract:

Biological systems are highly complex systems, both spatially and temporally. They are rooted in an interdependent, redundant and pleiotropic interconnected dynamic network. The properties of a system are different from those of their parts, and they depend on the integrity of the whole. The systemic properties vanish when the system breaks down, while the properties of its components are maintained. The disease can be understood as a systemic functional alteration of the human body, which present with a varying severity, stability and durability. Biological systems are characterized by measurable complex rhythms, abnormal rhythms are associated with disease and may be involved in its pathogenesis, they are been termed "dynamic disease." Physicians have long time recognized that alterations of physiological rhythms are associated with disease. Measuring absolute values of clinical parameters yields highly significant, clinically useful information, however evaluating clinical parameters the variability provides additionally useful clinical information. The aim of this review was to study one of the most recent advances in the measurement and characterization of biological variability made possible by the development of mathematical models based on chaos theory and nonlinear dynamics, as approximate entropy, has provided us with greater ability to discern meaningful distinctions between biological signals from clinically distinct groups of patients.

Keywords: complex systems, approximate entropy, variability

Servicio de Pediatría y Neonatología. Hospital Privado.

Cátedra de Clínica Pediátrica. Facultad de Ciencias Médicas. Universidad Nacional de Córdoba.

Cátedra de Metodología de la Investigación y Bioestadística. Facultad de Medicina. Universidad Católica de Córdoba.

Servicio de Pediatría y Neonatología. Hospital Privado. Naciones Unidas 346

X5016 Córdoba- Argentina Teléfono 0351-4688241

E-mail: ecueatas@hospitalprivadosa.com.ar

Financiación: propia.

Conflictos de intereses: ninguno.

Número de palabras: 2868

Enviado: 20 /01/ 2011

Aceptado 10/02/2011

En un esfuerzo por comprender los fenómenos complejos, se ha considerado a la teoría del caos como un posible modelo explicativo subyacente. Los parámetros asociados al caos son: las medidas de dimensión, las tasas de generación de información (entropía) y el espectro de Lyapunov. Para poder analizar estos parámetros en términos probabilísticos, se han desarrollado una serie de fórmulas utilizando teoremas ergódicos, basadas en los trabajos de Kolmogorov¹, Sinai² y Oseledets³. Como estos análisis requieren grandes series de datos que frecuentemente hacen muy difícil sino imposible su cálculo en términos prácticos, fue entonces Pincus⁴ quien dividió la teoría y el método para poder hacer factible el estudio estadístico de la regularidad, relacionando el índice de generación de la información con la entropía de Kolmogorov, aplicado a pequeñas series de datos clínicos originados de sistemas complejos "ruidosos" para demostrar la existencia o inexistencia de caos y no linealidad, creando una nueva herramienta para el análisis estadístico de series de datos denominada entropía aproximada (EnAp). La EnAp no es más que un estadístico de regularidad que cuantifica la impredecibilidad de las fluctuaciones en una serie de datos generada a través del tiempo, y aunque intuitivamente se podría suponer que la presencia de un patrón repetitivo de las fluctuaciones de una serie la hace más predecible que otra serie de datos en que ese patrón se encuentra ausente, la EnAp refleja la probabilidad de que un patrón "similar" de observaciones puede no ser seguido de observaciones adicionales similares. Una serie que contiene patrones de variabilidad con valores pequeños de EnAp es más predecible y menos caótica que una que arroja valores mayores⁵.

El principal objetivo de este resumen es presentar en forma concisa un resumen que incluya la definición de ApEn, su cálculo, interpretación y aplicación clínica, junto a una visión de sus alcances y limitaciones.

Elementos generales para el análisis de la variabilidad En general el análisis de la variabilidad implica conocer cuatro conceptos básicos que lo distinguen y lo hacen posible, estos son⁶.

Técnica de muestreo y tamaño muestral:

El análisis de los diferentes patrones de cambio o de la variabilidad de los parámetros clínicos o fisiológicos (desde ahora señales) se realiza sobre una serie de datos recolectados en forma continua o semicontinua a través del tiempo. El muestreo de estas señales puede realizarse entonces en intervalos continuos o discretos, que luego deben ser reconstruidos para evitar errores según el teorema de Nyquist, el cual enuncia que la frecuencia de muestreo debe ser por lo menos el doble que la mayor frecuencia de señales muestreadas.

Estacionalidad:

La estacionalidad requiere que las propiedades estadísticas de las señales, como la media y la desviación

típica, permanezcan iguales a través del periodo de relevamiento de los datos. La estacionalidad implica que los periodos de medición deben ser regulares y no cambiar en su periodicidad.

Control de los artefactos

El análisis de variabilidad debe hacerse sobre señales libres de artefactos, con el mínimo ruido posible. El ruido siempre significa una imprecisión en la técnica de medición, constituyendo entonces un error de medición importante. Existen muchas técnicas para controlar el ruido como por ejemplo el gráfico de Poincaré, que permite observar las diferencias entre datos puntuales consecutivos⁷.

Estandarización de la técnica

En deferencia al diseño experimental de Heisenberg⁸, se debe tener en cuenta que el proceso de medición puede alterar por sí mismo la variabilidad intrínseca. Uno de los componentes más importantes en la estandarización de las técnicas de medición es la duración de la misma, que debe ser siempre igual, para posibilitar análisis válidos.

ANÁLISIS DE LA ENTROPÍA APROXIMADA

Definición

La entropía es una medida de desorden o de aleatoriedad consagrada en la segunda ley de la termodinámica, la cual establece que la entropía de un sistema tiende hacia un máximo. En otras palabras, los estados tienden a evolucionar desde formas ordenadas aunque estadísticamente menos probables hacia configuraciones menos ordenadas pero estadísticamente más probables. Por ejemplo un anillo de humo (configuración ordenada) se difunde por el aire (configuración aleatoria), la situación espontánea inversa es estadísticamente muy poco probable hasta el punto de la imposibilidad absoluta.

En relación al análisis de series de datos temporales la EnAp provee una medida del grado de irregularidad o aleatoriedad dentro de una serie relativamente pequeña de datos.

Si se entiende EnAp como una medida de la complejidad de un sistema, los valores más pequeños indican una gran regularidad, mientras que los valores más grandes son indicadores de un mayor desorden⁹.

Cálculo

Para medir el grado de regularidad de una serie de datos de longitud N, la serie de datos debe ser evaluada según patrones de recurrencia. Esto se efectúa mediante la evaluación de secuencias de datos de longitud m, y se determina la probabilidad de que otras secuencias posteriores de la misma longitud sean similares, dentro de un intervalo de tolerancia r.

Entonces, para calcular la EnAp deben fijarse previamente dos parámetros m y r y una vez que se calcula la frecuencia de la ocurrencia de series repetidas, la medida de su prevalencia se resuelve como el logaritmo natural negativo de la probabilidad condicional.

Consecuentemente la EnAp mide la diferencia entre las frecuencias logarítmicas de patrones repetitivos de series de longitud m y series con patrones repetitivos de longitud $m+1$.

Como el algoritmo de cómputo de la EnAp ya se encuentra publicado en su máximo detalle⁴, aquí se enunciará un breve resumen del método de cálculo aplicado a una serie temporal de mediciones sucesivas de la una frecuencia (F). Entonces, dada una secuencia SN, consistente en N mediciones instantáneas de F , denominadas $F(1), F(2), \dots, F(N)$, se pueden tomar los valores de dos parámetros, m y r , para computar la EnAp (SN, m, r) de la secuencia. El parámetro m define la longitud de las subsecuencias y r especifica el criterio de similitud. Se denota una subsecuencia o patrón de m medidas de la F , comenzando por la medición i con SN, por el vector $Pm(i)$. Dos patrones $Pm(i)$ y $Pm(j)$ son similares si la diferencia entre cualquier par de mediciones correspondientes en los patrones es menor a r , por lo que sí:

$$|FC(i+k)-FC(j+k)| < r \text{ para } 0 \leq k < m$$

consideremos el conjunto de los vectores Pm de todos los patrones de longitud m , $Pm(1), Pm(2), \dots, Pm(N-m+1)$, dentro de SN, Se puede definir:

$$Cim(r) = \text{nim}(r)/N-m+1$$

donde $\text{nim}(r)$ es el número de patrones en Pm que son similares a $Pm(i)$, dada la similitud de criterio r , la cantidad $Cim(r)$ es la fracción de patrones de longitud m que remedan el patrón de la misma longitud que comienza en el intervalo i . Se puede calcular $Cim(r)$ para cada patrón en Pm , y es posible definir $Cim(r)$ como la media de esos valores de $Cim(r)$. La cantidad $Cm(r)$ expresa la prevalencia de patrones repetitivos de longitud m en SN. Finalmente, es posible definir la EnAp de SN, para un patrón de longitud m y un criterio de similaridad r , como:

$$ApEn(SN, m, r) = \ln [Cm(r)/ Cm+1(r)]$$

el logaritmo natural de la prevalencia relativa de los patrones repetidos de longitud m comparados con los de longitud $m+1$.

Así es que si encontramos patrones similares en una frecuencia (F) en una serie de mediciones, la EnAp estima la probabilidad logarítmica que el próximo intervalo después de cada patrón pueda diferir de éste. Valores pequeños de EnAp implican una mayor probabilidad de que patrones similares sean seguidos por un patrón adicional sucesivo similar, mientras que si la serie tiene secuencias de valores muy irregulares o variables, la ocurrencia de patrones similares no tendría valor predictivo para las mediciones sucesivas, y el valor de la EnAp sería entonces relativamente grande⁵.

Interpretación y aplicación clínica

La EnAp es una representación de la tasa de generación de nueva información dentro de un sistema biológico porque provee una medida del grado de irregularidad o de desorden dentro de una serie de señales. Como tal, debe entonces ser utilizada como una medida de la complejidad subyacente de cualquier sistema dinámico¹⁰.

El valor clínico del análisis de la complejidad de las variables fisiológicas tiene una potencialidad enorme, ya que la presencia de enfermedad implica siempre una pérdida de complejidad del sistema.

La EnAp ha sido utilizada ampliamente en estudios dinámicos de la frecuencia cardíaca. Su aplicación ha dado importantes aportes al conocimiento y detección de ciertas patologías, como por ejemplo que la EnAp de la frecuencia cardíaca se encuentra disminuida en lactantes con casi muerte súbita comparados a lactantes normales¹¹.

Como también la EnAp puede ser aplicada a series de señales ruidosas y poco extensas es de mucha utilidad para medir la variación de un parámetro en donde el muestreo del mismo es dificultoso, como por ejemplo cuando es impracticable efectuar frecuentes extracciones de sangre o hay relativamente pocos datos sin posibilidad de aumentar su número. Este hecho es particularmente aparente en la investigación de la variabilidad de los niveles plasmáticos de ciertas hormonas¹².

La EnAp también ha sido aplicada con éxito para investigar la variabilidad de parámetros neurológicos, respiratorios y más recientemente, al estudio de la temperatura corporal en diferentes circunstancias¹³⁻¹⁵.

Fortalezas y debilidades

La principal ventaja de la aplicación de EnAp a la investigación de señales biológicas es que puede ser calculada sobre series de datos relativamente pequeñas, tanto en el análisis teórico y básico como en la clínica. Pincus y Goldberg¹⁶ concluyeron que puede utilizarse como mínimo un set de $N = 50$ con una $m = 2$ y $r = 10-25\%$ de la desviación típica de los N valores.

El cálculo de la EnAp no es independiente de la secuencia de los datos, por lo que requiere que la conformación de los vectores represente exactamente el orden de los datos, consecutivamente, tal cual aparecieron al ser medidos. Es fundamental que este orden sea respetado también en la fase de análisis de los datos.

La debilidad más significativa de este método es que en el análisis de las series de datos muy inestables o ruidosas, comprometen la interpretación de los resultados obtenidos¹⁷.

Un sesgo inherente al cálculo de la EnAp se produce debido a que el algoritmo cuenta secuencias similares a una secuencia dada de longitud m , incluyendo en el conteo la secuencia la misma secuencia dada, para evitar el logaritmo de 0 dentro de los cálculos. Como resultado de esto, la EnAp puede ser sensible al tamaño de la serie de datos, arrojando valores inapropiadamente bajos, cuando ésta es pequeña⁹.

Por último, la EnAp evalúa la regularidad de una serie solo

en una escala, por lo que no es de utilidad en la investigación de fenómenos multiescales¹⁸.

Conclusión

El análisis de variabilidad representa una nueva forma de evaluar y tratar a los pacientes individualmente, teniendo como base líneas de investigación analítica y epidemiológica sobre grandes grupos de personas. Como hemos visto existe una extensa bibliografía que prueba el valor de las investigaciones sobre variabilidad biológica para proveer información útil a fin de comprender la fisiopatología de las enfermedades y a la vez facilitar la aplicación clínica del análisis de variabilidad para establecer diagnósticos y pronósticos. Sin duda las investigaciones futuras en este campo mejorarán la calidad del cuidado y de los resultados obtenidos para combatir las dolencias de nuestros enfermos¹⁹⁻²¹.

Bibliografía

1. Kolmogorov AN. *Dokl Akad Nauk SSSR* 1968;119:861-864.
2. Sinai Ya G. *Dokl Akad Nauk SSSR* 1959;124:768-771.
3. Oseledets VI. *Trna Moscow Math Soc* 1968;19:197-231.
4. Pincus SM. Approximate entropy as a measure of system complexity. *Proc Nat Acad Sci USA* 1991;88:2297-2301.
5. Pincus SM. Approximate entropy (ApEn) as a complexity measure. *Chaos* 1995;5:110-117.
6. Pincus SM, Singer BH. Randomness and degrees of irregularity. *Proc Nat Acad Sci USA* 1995;93:2083-2088.
7. Cunningham S, Symon AG, McIntosh N. The practical management of artifacts on computerized physiological data. *Int J Clin Monit Comput* 1994;11:211-216.
8. Heisemberg W. Über den anschaulichen Inhalt der quantentheoretischen Kinetik und Mechanik. *Zeitschrift für Physik* 1927;43:172-198.
9. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol* 2000;278:2039-2049.
10. Pincus SM. Assessing serial irregularity and its implications for health. *Ann NY Acad Sci* 2001;954:245-267.
11. Pincus SM, Cummings TR, Haddad GG. Heart rate control in normal and aborted –SIDS infants. *Am J Physiol* 1993;264:638-646.
12. Veldman RG, Frolich M, Pincus SM, Veldhuis JD, Roelfsema F. Patients with Cushing's disease secrete adrenocorticotropin and cortisol jointly more asynchronous than healthy subjects. *J Clin Endocrinol Metab* 1998;83:688-692.
13. Engoren M. Approximate entropy of respiratory rate and tidal volume during weaning from mechanical ventilation. *Crit Care Med* 1998;26:1817-1823.
14. Burnes J, Quigg M, Zanelli S, Goodkin HP. Clinical Severity, Rather Than Body Temperature, During the Rewarming Phase of Therapeutic Hypothermia Affect Quantitative EEG in Neonates With Hypoxic Ischemic Encephalopathy. *J Clin Neurophysiol*. 2011 Jan 10. [Epub ahead of print]
15. Varela M, Jimenez L, Farina R. Complexity analysis of the temperature curve: new information from body temperature. *Eur J Appl Physiol* 2003;89:230-237.
16. Pincus SM, Goldberger AL. Physiological time-series analysis: what do regularity quantify? *Am J Physiol* 1994;643-656.
17. Ryan SM, Goldberger AL, Pincus SM, Mietus J, Lipsitz LA. Gender and age-related differences in heart rate dynamics: are women more complex than men? *Am Coll Cardiol* 1994;24:1700-1707.
18. Costa M, Goldberger AL, Peng CK. Multiscale entropy to analysis of complex physiologic time series. *Phys Rev Lett* 2002;89:68-102.
19. Pan YH, Wang YH, Liang SF, Lee KT. Fast computation of sample entropy and approximate entropy in biomedicine. *Comput Methods Programs Biomed*. 2011 Jan 3. [Epub ahead of print].
20. Deffeyes JE, Harbourne RT, Stuber WA, Stergiou N. Approximate entropy used to assess sitting postural sway of infants with developmental delay. *Infant Behav Dev*. 2010 Dec 1. [Epub ahead of print].
21. Jovic A, Bogunovic N. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artif Intell Med*. 2010 Oct 25. [Epub ahead of print].