

# **Medidas de entropía en el procesado de señales biológicas: robustez y caracterización frente a pérdida de muestras y longitud de los registros**

**EVA MARÍA CIRUGEDA ROLDAN**

**EDITORIAL  
UNIVERSITAT POLITÈCNICA DE VALÈNCIA**



---

# Medidas de entropía en el procesado de señales biológicas: robustez y caracterización frente a pérdida de muestras y longitud de los registros

---

Julio 2014

**AUTOR:**

Eva María Cirugeda Roldán

**DIRECTOR:**

David Cuesta Frau



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

*Colección Tesis Doctorales*

© Eva María Cirugeda Roldan

© 2014, de la presente edición: Editorial Universitat Politècnica de València  
Telf.: 963 877 012 / [www.lalibreria.upv.es](http://www.lalibreria.upv.es)

ISBN: 978-84-9048-254-4 (version CD-Rom)

Queda prohibida la reproducción, distribución, comercialización, transformación, y en general, cualquier otra forma de explotación, por cualquier procedimiento, de todo o parte de los contenidos de esta obra sin autorización expresa y por escrito de sus autores.

El único límite para nuestra comprensión  
del mañana serán nuestras dudas del  
presente. Duda siempre de ti  
mismo, hasta que los datos  
no dejen lugar a dudas.

Franklin D. Roosevelt (1882–1945)  
Louis Pasteur (1822–1895)



A mi Abuelo



# *Agradecimientos*

A David Cuesta, por valorar mi solicitud y darme la oportunidad y el soporte para desarrollar esta larga andadura que es la tesis doctoral. A mis compañeros de laboratorio, que más que menos hemos compartido buenos momentos dentro y fuera de él.

A la gente que estos cuatro años han formado parte de mi vida de una u otra forma, gente de todo ámbito y parecer, polifacética y no tanto, de paso o para quedarse. Gracias por haber estado ahí porque me habeis enseñado mucho. El tiempo, dueño de nuestras vidas, borrará recuerdos, anécdotas y personas pero otras perdurarán en mi memoria y corazón para siempre.

A una persona muy especial en mi vida, que no ansía protagonismo, que prefiere pasar desapercibida, pero no podía evitar dedicarle unas líneas, por su inagotable apoyo, por su incansable lucha para conmigo, porque me ha hecho reír cuando sólo había tristeza, porque me daba aliento cuando todo era desasosiego, porque me hacía ver siempre la luz al final del túnel. Porque eres especial aunque tú no lo creas!

A mi familia. La que de verdad importa. La que siempre esta ahí para levantarte tras una caída. La que no importa cuantas veces te equivoques. La que te ama incondicionalmente. La que te valora como nadie lo hará nunca. La que te apoya, te ayuda, te enseña y te hace reír y llorar. Mi madre, la persona con mayor corazón que he conocido en mi vida, porque como ella no hay dos, porque si la perfección existe, ella es su máxima representante. A mi hermano, el niño de mis ojos y ahora mi hombrecito, porque es lo mejor que tengo en mi vida, porque me llena de orgullo cada día. Y a mi abuela, porque él se fue y ella ha sido fuerte por mí muchas veces, porque si de alguien he aprendido que no podemos dejar que el dolor nos consuma ha sido de ella. Porque siempre has estado ahí para mí, para tu niña aunque ya toda un mujer.

Y tras esta parrafada, me emociono pensando que no tengo palabras para él. Porque nada que escriba en este momento será suficiente. Porque una promesa se cumple hasta el final. A mi abuelo Antonio. Porque sé que dónde quiera que estes, si me estas viendo, tienes que ser el hombre más feliz del firmamento. La estrella más brillante. Irradiando orgullo, dicha y satisfacción por cada poro de tu piel. A partir de hoy dejarás de decir, mi nieta va a ser doctora. Hoy ya puedes decir que tienes una nieta doctora.



# RESUMEN

Las medidas de complejidad son un conjunto de métodos estadísticos que permiten estimar la regularidad de un sistema. Estos métodos se basan en técnicas de análisis no lineal de forma que se pueda caracterizar un señal sin hacer asunciones implícitas de estacionariedad o ergodicidad de la misma.

Estos métodos se están aplicando ampliamente sobre señales biológicas debido a la naturaleza de las mismas. Las señales biológicas se caracterizan por ser irregulares, no lineales y variables en el tiempo, de forma que los métodos tradicionales de análisis lineal no consiguen caracterizar su comportamiento completamente.

Estas medidas funcionan muy bien en la práctica, ya que consiguen extraer información de las señales que de otra forma no es posible. Entre otras capacidades, consiguen diferenciar estados patológicos, precedir la aparición de un ataque epiléptico o distinguir entre estados del sueño. Pero su aplicación presenta cierta controversia, ya que carecen de una caracterización que indique al usuario qué medida aplicar en función de las características del registro, cómo debe ser aplicada o incluso cómo interpretar los resultados obtenidos.

En este trabajo se ha propuesto abordar una caracterización de algunas de las medidas de complejidad de uso más común. Se muestra una caracterización de la entropía aproximada (ApEn), la entropía muestral (SampEn), la entropía en múltiples escalas o multiescala (MSE), el análisis de fluctuaciones sin tendencias (DFA), la entropía cuadrada de Rényi (QSE) y el coeficiente de entropía muestral (CosEn), ante situaciones en las que las señales han perdido muestras o cuya longitud es limitada.

La pérdida de muestras es algo muy común en la actualidad, dónde la mayoría de los registros se hacen de forma ambulatoria y el espacio de almacenamiento es limitado (compresión de datos) o la transmisión se hace de forma inalámbrica, donde el canal puede presentar condiciones inestables o interferencias que causen la pérdida de muestras, bien de forma uniforme o aleatoria. La longitud limitada de los registros puede deberse, entre otras posibilidades, a que la toma de datos se ha realizado de forma manual o ésta resulta incómoda para el paciente.

Se presenta una caracterización paramétrica de las medidas para las señales de longitud reducidas y se proponen dos métodos de optimización no supervisada para el análisis de registros de corta duración con QSE o CosEn.

Este trabajo muestra cómo las medidas de entropía consideradas, presentan un comportamiento similar ante una misma situación, conservando las capacidad de separabilidad entre clases, independientemente del registro biológico analizado, siempre y cuando la medida se use de forma correcta.

SampEn se ha erigido como la medida más estable y de mayor aplicabilidad en registros de duración media ( $300 < N < 5000$ ) cuando las señales pierden muestras, tanto de forma aleatoria, como uniforme mateniendo coeficientes de correlación cruzados por encima de 0.8 hasta un 70 % de pérdidas. Si las señales presentan desviaciones estándar altas o gran variabilidad, se recomienda la aplicación de MSE ya que introduce un suavizado y decorrelación de los patrones.

En señales de corta duración ( $100 < N < 300$ ) se recomienda el uso de DFA, ya que permite una caracterización de la complejidad de forma estable y robusta aunque con un coste computacional alto y la necesidad de realizar una inspección visual para determinar el número de coeficientes de escalado necesarios.

Finalmente, en señales de muy corta duración ( $N < 100$ ) se recomienda el uso de CosEn. Se han conseguido segmentar señales de HTA en humanos de apenas 55 muestras, algo muy novedoso, con mejores estadísticos que QSE.

# ABSTRACT

Complexity measures are a set of statistical methods that enable the estimation of system regularity. These methods are based on nonlinear signal processing techniques that allow signal characterization without any underlying assumption on stationarity or ergodicity.

These nonlinear processing techniques are being widely used over biological signals due to their nature. Biological signals are characterized by an irregular, nonlinear and time-varying behavior. This chaotic behavior makes traditional linear methods fail as they can not fully characterize it.

Regularity measures work well in practice, as they can unveil information from the signals which otherwise is not possible. They can discern between healthy and pathologic states, predict the occurrence of an epileptic seizure or classify between sleep stages, among others. Even though their application is useful, it presents some controversy, as there is no previous characterization on how to apply them or even how the results should be interpreted.

This thesis proposes a deep characterization of some of the most common complexity measures. A characterization of Approximate Entropy (ApEn), Sample Entropy (SampEn), MultiScale Entropy (MSE), Detrended Fluctuation Analysis (DFA), Rényi Square Entropy (QSE) and the Coefficient of Sample Entropy (Co-sEn) is driven against situations where biological signals suffered sample loss or present a limited length.

Sample loss is actually quite common where most of the signal transduction and recording is made somehow in an ambulatory manner and the storage capabilities are limited (data compression) or signals are wirelessly transmitted over channels that can present unstable conditions or interference that can cause sample loss. The limited length of the recordings can be due to a manual annotation or cumbersome and uncomfortable recording techniques.

A parametric characterization of short length entropy measures is performed and two unsupervised optimization techniques for the analysis of short record length by means of QSE and CosEn is proposed.

This thesis shows a similar behavior of the considered complexity measures against a similar scenario while preserving the segmentation capabilities between classes with independence to the biological recording on analysis, as long as their application is done in the correct manner.

SampEn has proven to be the most stable measure and with wider applicability in medium length records ( $300 < N < 5000$ ) against both, random and uniform, sample loss. Cross correlation coefficient for SampEn relay over 0.8 up tu a 70 % sample loss. Wether biological records show high variability MSE is recommended as the coarse grained series introduce signal soften and template decorrelation.

In short length signals ( $100 < N < 300$ ) DFA is recommended. DFA allows complexity characterization in a robust and reliable way. It has a high computational cost and a visual inspection for scaling region delimitation should be performed.

Finally in very short length records ( $N < 100$ ) Cosen is recommended. Human hipertensive records with less than 55 samples have been segmented with good statistical validity and better results than QSE.

# RESUM

Les mesures de complexitat són un conjunt de mètodes estadístics que permeten valorar la regularitat d'un sistema. Aquests mètodes es basen en tècniques d'anàlisi no lineal de manera que es pugui caracteritzar una senyal sense fer suposicions implícites d'estacionarietat o ergodicitat de la mateixa.

Aquests mètodes estan àmpliament aplicant-se sobre senyals biològiques a causa de la naturalesa de les mateixes. Les senyals biològiques es caracteritzen per ser irregulars, no lineals i variables en el temps, de manera que els mètodes tradicionals d'anàlisi lineal no aconseguen caracteritzar el seu comportament completament.

Aquestes mesures funcionen molt bé en la pràctica, ja que aconseguen extreure informació de les senyals que d'una altra forma no és possible. Entre altres capacitats, permeten diferenciar estats patològics, predir l'aparició d'un atac epilèptic o distingir entre estats del somni. Però la seva aplicació presenta certa controvèrsia, ja que manquen d'una caracterització que indiqui a l'usuari què mesura aplicar en funció de les característiques del registre, com ha de ser aplicada o fins i tot com interpretar els resultats obtinguts.

En aquest treball s'ha proposat oferir una caracterització d'algunes de les mesures de complexitat d'ús més comú davant situacions quotidianes no valorades amb anterioritat. Aquest treball mostra una caracterització de l'entropia aproximada (ApEn), l'entropia mostral (SampEn), l'entropia en múltiples escales (MSE), l'anàlisi de fluctuacions sense tendències (DFA), l'entropia quadrada de Rényi (QSE) i el coeficient d'entropia mostral (CosEn), davant situacions en les quals les senyals han perdut mostres o la longitud de les quals és limitada. La pèrdua de mostres és alguna cosa molt comuna en l'actualitat, on la majoria dels registres es fan de forma ambulatoria i l'espai d'emmagatzematge és limitat (compressió de dades) o la transmissió es fa sense fils on el canal pot presentar condicions inestables o interferències que causin la pèrdua de mostres bé de forma uniforme o aleatòria. La longitud limitada dels registres pot ocórrer, entre altres possibilitats, al fet que la presa de dades s'ha realitzat de forma manual o aquesta resulta incòmoda per al pacient.

Aquest treball mostra com les mesures d'entropia considerades, presenten un comportament similar davant una mateixa situació conservant la capacitat de separabilitat entre classes, independentment del registre biològic analitzat, sempre que la mesura s'utilitzi de forma correcta. Es presenta una caracterització paramètrica de les mesures per a les senyals de longitud reduïda i es proposen dos mètodes d'optimització no supervisada per a l'anàlisi de registre de curta durada amb QSE o CosEn.

SampEn s'ha erigit com la mesura més estable i de major aplicabilitat en registres de durada mitjana ( $300 < N < 5000$ ) quan els senyals perden mostres tant de forma aleatòria com uniforme, mantenint coeficients de correlació creuats per sobre de 0.8 fins a un 70% de pèrdues. Si les senyals presenten desviacions estàndard altes o gran variabilitat, es recomana l'aplicació de MSE ja que introdueix un suavitzat i decorrelació dels patrons.

En senyals de curta durada ( $100 < N < 300$ ) es recomana l'ús de DFA, ja que permet una caracterització de la complexitat de forma estable i robusta encara que amb un cost computacional alt i la necessitat de realitzar una inspecció visual per determinar el nombre de coeficients d'escalat necessaris.

Finalment en senyals de molt curta durada es recomana l'ús de CosEn, ja que s'han aconseguit segmentar senyals de HTA en humans de només 55 mostres, cosa molt nova, amb millors estadístics que QSE.

# Índice general

Índice general	I
Índice de figuras	V
Índice de tablas	XIII
1 Introducción	1
2 Estado del arte	5
3 Objetivos	9
4 Métodos	13
4.1 Entropía aproximada. . . . .	13
4.2 Entropía muestral . . . . .	16
4.3 Entropía multiescala . . . . .	18
4.4 Análisis de fluctuaciones sin tendencias . . . . .	19
4.4.1 Estimación del punto de cross-over . . . . .	22
4.5 Entropía cuadrada de Rényi . . . . .	23
4.6 Coeficiente de entropía muestral . . . . .	26
4.7 Métodos de Evaluación . . . . .	27
4.7.1 Test de Segmentación . . . . .	27
4.7.2 Test de Estabilidad. . . . .	28

5	Conjunto experimental	31
5.1	Base de datos de electroencefalogramas . . . . .	31
5.2	Base de datos de registros de hipertensión arterial . . . . .	33
5.3	Bases de datos de registros de glucemia . . . . .	35
5.3.1	Registros de glucemia en pacientes críticos . . . . .	35
5.3.2	Registros de glucemia en pacientes ambulatorios . . . . .	36
5.3.3	Registros de glucemia de pacientes post-operatorios . . . . .	37
5.4	Base de datos de electrogramas de fibrilación auricular . . . . .	38
5.5	Base de datos de registros RR . . . . .	39
5.6	Base de datos de dinámicas de movilidad en pacientes con enfermedades neurodegenerativas . . . . .	40
6	Experimentos	43
6.1	Pérdida de muestras . . . . .	43
6.2	Señales de corta duración . . . . .	45
6.2.1	Parametrización . . . . .	45
6.2.2	Entropías específicas . . . . .	46
6.3	Evaluación . . . . .	46
7	Resultados	49
7.1	Entropía aproximada (ApEn) . . . . .	50
7.1.1	Pérdida aleatoria de muestras . . . . .	50
7.1.2	Pérdida uniforme de muestras . . . . .	54
7.1.3	Señales de corta duración . . . . .	58
7.2	Entropía muestral (SampEn) . . . . .	62
7.2.1	Pérdida aleatoria de muestras . . . . .	62
7.2.2	Pérdida uniforme de muestras . . . . .	65
7.2.3	Señales de corta duración . . . . .	68
7.3	Entropía Multiescala (MSE) . . . . .	70
7.3.1	Pérdida aleatoria de muestras . . . . .	70
7.3.2	Pérdida uniforme de muestras . . . . .	80
7.3.3	Señales de corta duración . . . . .	90
7.4	Análisis de fluctuaciones sin tendencias (DFA) . . . . .	92
7.4.1	Pérdida aleatoria de muestras . . . . .	92
7.4.2	Pérdida uniforme de muestras . . . . .	95

7.4.3 Señales de corta duración . . . . .	98
7.5 Entropía cuadrada de Rényi (QSE) . . . . .	105
7.5.1 Hipertensión Arterial . . . . .	106
7.6 Coeficiente de entropía muestral (CosEn) . . . . .	107
7.6.1 Hipertensión Arterial . . . . .	108
7.6.2 Glucemia Postoperatoria . . . . .	109
8 Discusión . . . . .	111
8.1 Pérdida de muestras . . . . .	112
8.1.1 Pérdida de muestras aleatoria . . . . .	112
8.1.2 pérdida de muestras uniforme . . . . .	115
8.2 Señales de corta duración. . . . .	117
9 Conclusión . . . . .	121
10 Líneas futuras . . . . .	125
11 Aportaciones de la tesis . . . . .	127
11.1 Contribuciones a congresos . . . . .	127
11.2 Artículos en revistas . . . . .	128
Bibliografía . . . . .	129
Índice de Abreviaturas . . . . .	143



# Índice de figuras

4.1. Estimación de QSE para una señal de hipertensión arterial. (a) Umbral $r$ necesario en función del número de patrones coincidentes $M$ . (b) Cálculo de QSE en función de $M$ . (c) Estimación $M_{opt}$ : curva del área bajo la ROC (AUROC) en función de $M$ en azul y su versión filtrada en rojo. (d) Máximos detectados en función de $M$ y región de filtrado de mediana en rojo. (e) Distribución de $M_{opt}(k)$	26
5.1. Representación temporal (EEG) y espectral (PSD) para una señal de cada set. Sets A a E respectivamente. . . . .	32
5.2. Variables fisiológicas contenidas en la base de datos de hipertensión arterial (HTA). (a) Presión arterial sistólica (SBP), (b) presión arterial diastólica (DBP), (c) presión arterial media (MBP) y (d) frecuencia cardíaca ( $F_C$ ). En rojo se presentan los sujetos con hipertensión severa (patológicos) y en azul los sujetos de control. La línea continua presenta la señal media de cada variable y las discontinuas representan los intervalos de confianza al 95%. . . . .	34
5.3. Niveles de glucemia medios para (a) diabéticos (DM) en azul y no diabéticos (NDM) en rojo y (b) supervivientes (SV) en azul y no supervivientes (NSV) en rojo. La línea sólida presenta el valor medio y las discontinuas los intervalos de confianza al 95%. . . . .	35
5.4. Niveles de glucemia medios para los sujetos de control (CTRL) en azul. En rojo los sujetos con síndrome metabólico (MS). En negro sujetos con diabetes tipo II (DM). . . . .	36
5.5. Histograma del número de señales en función de las longitudes. . .	37
5.6. Señal ejemplo de un electrograma para cada grado de fraccionamiento (a) C0, (b) C1, (c) C2 y (d) C3. . . . .	39

5.7. Registros RR de un sujeto de cada tipo de medicación antes del tratamiento (Pre) en azul y durante el tratamiento (Post) con (a) encainida, (b) frecaínida y (c) moricicina en rojo . . . . .	40
5.8. Definición de los términos (a) <i>Stance</i> y (b) <i>Stride</i> . . . . .	41
5.9. Tiempo de apoyo del pie derecho (TAP). Se muestra una señal ejemplo de cada grupo (a) control, (b) esclerosis lateral amiotrófica, (c) Huntington y (d) Párkinson. . . . .	41
6.1. Señal ECG real ampliada y saturada de forma artificial para simular el efecto de una saturación en un registro . . . . .	44
6.2. Se muestra en la fila superior el proceso aplicado para la eliminación de muestras de forma aleatoria de un 60 % y en la inferior para el algoritmo de pérdidas de muestras uniforme con un 40 % de pérdidas. La primera figura representa una señal RR sobre la cual se eliminan muestras, la segunda columna representa la señal con las muestras eliminadas y en la última vemos finalmente la señal a analizar, donde al eliminar las muestras ésta ha visto su longitud disminuída. . . . .	45
7.1. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con muestras eliminadas en función del ratio de pérdidas. . . . .	50
7.2. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul tratados con encainida. (b) Coeficiente de correlación cruzada de ApEn para cada grupo en función del porcentaje de muestras perdidas. . . . .	52
7.3. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EGM no recomendada para ablación (NRA) en azul o recomendada para ablación (RA) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) de la regularidad en función del ratio de muestras perdidas. . . . .	54

7.4. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con pérdidas en función del ratio de muestras eliminadas. . . . .	55
7.5. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre registros RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul tratados con encainida. (b) Coeficiente de correlación cruzada de ApEn para cada grupo en función del porcentaje de muestras perdidas. . . . .	56
7.6. (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul o recomendadas para ablación (RA) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) de la regularidad en función del ratio de pérdidas. . . . .	58
7.7. (a) Comportamiento de SampEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con pérdidas en función del ratio de muestras eliminadas. . . . .	62
7.8. Comportamiento de SampEn cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul, tratados con encainida. (a) Evolución de SampEn y (b) Coeficiente de correlación cruzada entre la serie original y sus versiones con pérdida de muestras. . . .	64
7.9. (a) Comportamiento de SampEn cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul y recomendadas para ablación (RA) en rojo y (b) $C_{xy}$ entre la regularidad de la serie original y sus versiones con pérdidas para distintos niveles de fraccionamiento. . .	65
7.10. Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) o patológica (PT). (a) Evolución de SampEn en función del ratio de pérdidas, (b) coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la serie original y la de las series con pérdidas. . . . .	66

7.11. Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul. (a) Evolución en función del porcentaje de muestras eliminadas de forma uniforme, (b) coeficiente de correlación cruzada entre la regularidad de la señal original y sus versiones con perdidas. . . . .	67
7.12. (a) Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal EGM no recomendados para ablación (NRA) en azul y recomendados para ablación (RA) en rojo, y (b) coeficiente de correlación cruzada entre la regularidad de la serie original y las series con pérdida de muestras. . . . .	68
7.13. Comportamiento de MSE cuando se eliminan muestras de forma aleatoria sobre una señal EEG control (CT) o patológica (PT). (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10 . . . . .	73
7.14. Comportamiento de MSE sobre registros RR obtenidos en pacientes antes (Pre) y después (Post) del tratamiento con encainida cuando se da pérdida de muestras aleatoria. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10. . . . .	76
7.15. Comportamiento de MSE cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul, y recomendadas para ablación (RA) en rojo, para distintas escalas de MSE. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10. . . . .	80
7.16. Comportamiento de MSE cuando se eliminan muestras de forma uniforme sobre una señal EEG control (CT) en azul o patológica (PT) en rojo. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10 . . . . .	83
7.17. Comportamiento de MSE sobre registros RR obtenidos en pacientes tratados con encainida, antes (Pre) y después (Post) del tratamiento cuando se han perdido muestras de forma uniforme. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10 . . . . .	86
7.18. Comportamiento de MSE cuando se eliminan muestras de forma uniforme sobre una EGM registrados en regiones no recomendadas para ablación (NRA) en azul y recomendadas para ablación (RA) en rojo, para distintas escalas de MSE. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10 . . . . .	89

7.19. (a) Comportamiento de DFA ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la señal original y de las series con pérdidas en función del ratio de muestras eliminadas. . . . .	92
7.20. Comportamiento de DFA cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos antes (Pre) en rojo y después (Post) en azul, del tratamiento con encainida. (a) Evolución en función del ratio de muestras eliminadas y (d) coeficiente de correlación cruzada de la complejidad de cada grupo en las series con pérdidas con respecto a la complejidad de las series originales. . . .	93
7.21. (a) Comportamiento de DFA cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en rojo y recomendadas para ablación (RA) en azul. (b) Coeficiente de correlación cruzada entre la señal original y las señales con pérdida de muestras en función del porcentaje de muestras eliminadas. . . . .	94
7.22. (a) Comportamiento de DFA ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre las regularidades de las series orginales y las series con pérdidas. . . . .	95
7.23. (a) Comportamiento de DFA cuando se eliminan muestras de forma uniforme sobre una señal RR antes (PRE) en azul y después (POST) en rojo tratados del tratamiento con encainida. (b) coeficiente de correlación cruzada de la entropía de cada grupo en las series con pérdida de muestras respecto a la compejidad de la serie original. . .	96
7.24. (a) Comportamiento de DFA cuando se eliminan muestras de forma uniforme sobre EGM registrados en regiones no recomendadas para ablación (NRA) en rojo y recomendadas para ablación (RA) en azul. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la señal original y la de las versiones con pérdidas en función del porcentaje de muestras eliminadas. . . . .	98

7.25. Función de fluctuación ( $F(L)$ ) media para sujetos no diabéticos (NDM) en escala logarítmica y distintos órdenes de integración de DFA: en azul orden 1, en negro orden 2 y en rojo orden 3. La línea sólida identifica $F(n)$ y la punteada la regresión lineal. (a) ventanas no solapadas (b) ventanas deslizantes (c) dos coeficientes: la línea punteada identifica la regresión lineal en la primera parte y la línea discontinua en la segunda parte. . . . .	101
7.26. Función de fluctuación ( $F(L)$ ) media para sujetos sanos (NDM) en escala logarítmica y distintos órdenes de integración de DFA. En azul orden 1, en negro orden 2 y en rojo orden 3. La línea sólida denota $F(n)$ y la punteada la regresión lineal. (a) Ventanas no solapadas, (b) ventanas deslizantes y (c) dos coeficientes: la línea punteada representa la regresión lineal en la primera parte y la línea discontinua en la segunda parte. . . . .	103
7.27. Función de fluctuación ( $F(L)$ ) media para sujetos de control (CT) en escala logarítmica y distintos órdenes de integración de DFA. En azul orden 1, en rojo orden 2 y en negro orden 3. La línea sólida denota $F(n)$ y la punteada la regresión lineal. (a) ventanas no solapadas (b) ventanas deslizantes. . . . .	105
7.28. Cálculo de QSE sobre señales HTA. (a) Evolución del umbral, $r$ , con el número de patrones coincidentes, $M$ , para una señal. (b) Evolución de QSE en función de $M$ para una señal de cada grupo, control (CT) en rojo y patológica (PT) en azul. (c) Evolución del área bajo la curva ROC (AUC) de cada variable de la base de datos, presión sistólica (SBP), diastólica (DBP), media (MBP) y frecuencia cardíaca (FC) en función de $M$ . (d) AUC de SBP en azul y su versión filtrada ( $AUC_F$ ) en rojo. . . . .	106
7.29. Curva del área bajo la ROC (AUC) en función del número de coincidencias ( <i>matches</i> , $M$ ) de CosEn para las cuatro variables consideradas: presión sistólica (SBP), diastólica (DBP), media (PAM) y frecuencia cardíaca (FC) filtrada para distintas longitudes del filtro de media móvil, (a) $K = 11$ , (b) $K = 51$ y (c) $K = 101$ . . . . .	108
7.30. Distribución del número de <i>matches</i> óptimo, $M$ , en azul, para una segmentación máxima en función de la longitud de las señales $N$ con distintos tipos de longitud de la ventana de Hamming para el filtro basado en el periodograma de Welch, (a) $K = 51$ , (b) $K = 101$ y (c) $K = 151$ . En rojo se presenta el número máximo de <i>matches</i> . . . . .	110

7.31. Capacidades estadísticas de CosEn evaluando series de glucemia en sujetos con control pre-operatorio o post-operatorio en pacientes que han desarrollado complicaciones post-operatorias. Estimación con un filtrado basado en el periodograma de Welch con una ventana Hamming de 51 muestras. (a) Intervalos de confianza (CI) y (b) probabilidad de separación estadística (UT) en función de la longitud  $N$  de la señal analizada. . . . . 110



# Índice de tablas

5.1. Resumen de las bases de datos consideradas. Se muestra para cada base de datos, el grupo considerado de control y los patológicos, el número de registros en cada grupo (NR) y la longitud de éstos ( $N$ ).	42
7.1. Caracterización estadística de ApEn sobre señales EEG de sujetos control (CT) o patológicos (PT) que han sufrido pérdida de muestras de forma aleatoria. Se muestra el estudio de normalidad (JB), estadísticos de segmentación (UT), intervalos de confianza (CI) al 95 % y coeficiente de correlación cruzada ( $C_{xy}$ ).	51
7.2. Caracterización estadística de ApEn sobre registros RR de sujetos pre-tratamiento (PRE) y post-tratamiento (POST) cuando las pérdidas se producen de forma aleatoria. Se muestra el test de normalidad (JB) y de segmentación (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).	53
7.3. Caracterización de ApEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se pierden de forma aleatoria. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).	54
7.4. Caracterización de ApEn sobre señales EEG de control (CT) y patológicas (PT). Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).	56

7.5. Caracterización de ApEn para registros RR de pacientes antes (Pre) y después (Post) del tratamiento con encainida. Los registros han sufrido una pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	57
7.6. Caracterización de ApEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las pérdidas se producen de forma uniforme. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	57
7.7. Combinación paramétrica que permite la segmentación entre sujetos supervivientes (SV) y no supervivientes (NSV) en función de los niveles de glucemia. Se muestran también los intervalos de confianza al 95 % (CI) y los test de normalidad (JB) y separabilidad (UT). .	59
7.8. Combinación paramétrica que permite la segmentación entre distintas clases. Se muestran también los intervalos de confianza al 95 % (CI) y los test de normalidad (JB) y separabilidad (UT). En negrita se identifica dónde no hay segmentación entre clases. $UT_1$ : segmentación entre NDM y DM, $UT_2$ : segmentación entre NDM y MS y $UT_3$ : segmentación entre DM y MS. . . . .	60
7.9. Caracterización de ApEn en función de los parámetros ( $m, r$ ) para los grupos considerados en la base de datos de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA). $UT_1, UT_2, UT_3$ indican la probabilidad estadística de segmentación entre LSA, PD y HD con respecto a CT.	61
7.10. Caracterización de SampEn sobre registros EEG de control (CT) y patológicos (PT) que han perdido muestras de forma aleatoria. Se muestra en función del porcentaje de pérdidas en la señal la normalidad (JB), la separabilidad de los grupos (UT), los intervalos de confianza al 95 % (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).	63
7.11. Caracterización de SampEn sobre registros RR en pacientes antes (Pre) y después (Post) del tratamiento con encainida que han sufrido pérdidas de forma aleatoria. Se muestra, en función del porcentaje de pérdidas, la normalidad (JB), la separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	63

- 
- 7.12. Caracterización de SampEn sobre los EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando la pérdida de muestras se produce de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). 64
- 7.13. Caracterización de SampEn en señales EEG control (CT) y patológicas (PT) con pérdida de muestras de forma uniforme. Se muestra: test de normalidad (JB) y separabilidad (UT), intervalos de confianza al 95 % (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . 65
- 7.14. Caracterización de SampEn en registros RR antes (Pre) y después (Post) del tratamiento con encainida y pérdida de muestras de forma uniforme. Se muestra: test de normalidad (JB) y separabilidad (UT), intervalos de confianza al 95 % (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . . 66
- 7.15. Caracterización de SampEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se pierden de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . 67
- 7.16. Caracterización de SampEn sobre registros de glucemia ambulatoria. Para cada combinación paramétrica se dan los intervalos de confianza (CI), test de normalidad (JB) y separabilidad (UT). En negrita se identifican las combinaciones sin separación estadística entre clases.  $UT_1$ : segmentación entre NDM y DM,  $UT_2$ : segmentación entre NDM y MS y  $UT_3$ : segmentación entre DM y MS. . . 69
- 7.17. Caracterización de SampEn en función de los parámetros ( $m, r$ ) para los registros de dinámicas de movilidad, control (CT), sujetos con Parkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA).  $UT_1, UT_2, UT_3$  indican la probabilidad estadística de segmentación entre LSA, PD y HD con respecto a CT. . . . . 70
- 7.18. Caracterización de MSE 2 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . 70
- 7.19. Caracterización de MSE 3 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . 71

7.20. Caracterización de MSE 4 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	71
7.21. Caracterización de MSE 5 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	71
7.22. Caracterización de MSE 6 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	71
7.23. Caracterización de MSE 7 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	72
7.24. Caracterización de MSE 8 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	72
7.25. Caracterización de MSE 9 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	72
7.26. Caracterización de MSE 10 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . .	72
7.27. Caracterización de MSE 2 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	74
7.28. Caracterización de MSE 3 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	74

---

7.29. Caracterización de MSE 4 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	74
7.30. Caracterización de MSE 5 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	75
7.31. Caracterización de MSE 6 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	75
7.32. Caracterización de MSE 7 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	75
7.33. Caracterización de MSE 8 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	75
7.34. Caracterización de MSE 9 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	76
7.35. Caracterización de MSE 10 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	76

7.36. Caracterización de MSE 2 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	77
7.37. Caracterización de MSE 3 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	77
7.38. Caracterización de MSE 4 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	78
7.39. Caracterización de MSE 5 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	78
7.40. Caracterización de MSE 6 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	78
7.41. Caracterización de MSE 7 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	78
7.42. Caracterización de MSE 8 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	79

---

7.43. Caracterización de MSE 9 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	79
7.44. Caracterización de MSE 10 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	79
7.45. Caracterización de MSE 2 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	80
7.46. Caracterización de MSE 3 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	81
7.47. Caracterización de MSE 4 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	81
7.48. Caracterización de MSE 5 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	81
7.49. Caracterización de MSE 6 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	81
7.50. Caracterización de MSE 7 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	82
7.51. Caracterización de MSE 8 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	82

7.52. Caracterización de MSE 9 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) . . . . .	82
7.53. Caracterización de MSE 10 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) . . . . .	82
7.54. Caracterización de MSE 2 con pérdida de muestras de forma uniforme en registros RR antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	84
7.55. Caracterización de MSE 3 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	84
7.56. Caracterización de MSE 4 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	84
7.57. Caracterización de MSE 5 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	84
7.58. Caracterización de MSE 6 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	85
7.59. Caracterización de MSE 7 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	85

---

7.60. Caracterización de MSE 8 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	85
7.61. Caracterización de MSE 9 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	85
7.62. Caracterización de MSE 10 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	86
7.63. Caracterización de MSE 2 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	87
7.64. Caracterización de MSE 3 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	87
7.65. Caracterización de MSE 4 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	87
7.66. Caracterización de MSE 5 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	88

7.67. Caracterización de MSE 6 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	88
7.68. Caracterización de MSE 7 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	88
7.69. Caracterización de MSE 8 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	88
7.70. Caracterización de MSE 9 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	89
7.71. Caracterización de MSE 10 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ). . . . .	89
7.72. Estudio estadístico de la variabilidad de MSE 2 en función de los parámetros ( $m, r$ ) para los grupos de sujetos DM, MS y NDM. En negrita se identifican las distribuciones donde no se obtiene separación entre clases. $UT_1$ : Segmentación estadística entre medianas DM y NDM, $UT_2$ : entre MS y NDM y $UT_3$ : entre DM y MS. . . .	90
7.73. Estudio estadístico de la variabilidad de MSE 3 en función de los parámetros ( $m, r$ ) para los grupos de sujetos DM, MS y NDM. En negrita se remarcan las distribuciones que no segmentan los grupos	91
7.74. Caracterización estadística de MSE 2 en función de los parámetros ( $m, r$ ) para los grupos considerados en los registros de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA). $UT_1$ , $UT_2$ y $UT_3$ muestran la probabilidad de segmentación entre LSA, PD y HD respecto a CT.	91

7.75. Caracterización estadística de MSE 3 en función de los parámetros $(m, r)$ para los grupos considerados en los registros de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA). $UT_1$ , $UT_2$ y $UT_3$ muestran la probabilidad de segmentación entre LSA, PD y HD respecto a CT.	92
7.76. Caracterización de DFA para señales EEG de control (CT) y patológicos (PT) cuando se pierden muestras de forma aleatoria. Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada $(C_{xy})$ .	93
7.77. Caracterización de DFA sobre registros RR antes (Pre) y después (Post) del tratamiento con encainida cuando se pierden muestras de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada $(C_{xy})$ .	94
7.78. Caracterización de DFA sobre EGM registrados en regiones no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se eliminan de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada $(C_{xy})$ .	95
7.79. Caracterización de DFA para señales EEG de control (CT) y patológicos (PT) con pérdidas de forma uniforme. Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada $(C_{xy})$ .	96
7.80. Caracterización de DFA sobre registros RR antes (PRE) y después (POST) del tratamiento con encainida cuando las muestras se eliminan de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada $(C_{xy})$ .	97
7.81. Caracterización de DFA sobre EGM registrados en regiones no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando la pérdida de muestras es de forma uniforme. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada $(C_{xy})$ .	97
7.82. Coeficientes de escalado $(\mu \pm \sigma)$ y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y las dos clasificaciones consideradas: supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM). En negrita se indica dónde existe separación.	99

7.83. Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y las dos clasificaciones consideradas: supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM) utilizando ventanas deslizantes. En negrita se remarca dónde existe segmentación entre grupos. . . . .	99
7.84. Coeficientes de escalado a corto ( $\alpha_{sr}$ ) y largo plazo ( $\alpha_{lr}$ ) ( $\mu \pm \sigma$ ), localización del punto de <i>cross-over</i> (CO) y estadísticos de separabilidad (UT) para DFA-2 en supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM). En negrita se identifica dónde existe segmentación entre grupos. . . . .	100
7.85. Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad ( $UT_i$ ) para distintos órdenes de integración de DFA (DFA- $l$ ) y los distintos grupos considerados, diabéticos (DM), síndrome metabólico (MS) y no diabéticos (NDM). Separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM . . . . .	101
7.86. Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos $l$ para los distintos grupos considerados, diabéticos (DM), síndrome metabólico (MS) y no diabético (NDM), se indica la separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM	102
7.87. Coeficientes de escalado a corto ( $\alpha_{sr}$ ) y largo plazo ( $\alpha_{lr}$ ) ( $\mu \pm 2\sigma$ ), y localización del punto de <i>cross-over</i> (CO) para $l = 2$ en registros de sujetos diabéticos (DM), con síndrome metabólico (MS) y no diabéticos (NDM). Se indica la separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM. . . . .	103
7.88. Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y los distintos grupos de registros considerados. Registros de control (CT), de Párkinson (PD), Huntinton (HD) y esclerosis lateral amiotrófica (LSA). $UT_1$ , $UT_2$ y $UT_3$ son los estadísticos de segmentación entre LSA, PD y HD con respecto a CT. . . . .	104
7.89. Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y los diferentes grupos de registros considerados en el algoritmo de DFA estimado con ventanas deslizantes. Registros de control (CT), de Párkinson (PD), Huntinton (HD) y esclerosis lateral amiotrófica (LSA). $UT_1$ , $UT_2$ y $UT_3$ son los estadísticos de segmentación entre LSA, PD y HD con respecto a CT. . . . .	105

- 
- 7.90. Caracterización de QSE sobre registros de hipertensión arterial con 4 variables: presión arterial sistólica (SBP), diastólica (DBP), media (MBP) y frecuencia cardíaca (FC). Se muestra el número de patrones coincidentes óptimo  $M_{opt}$  para cada grupo de señales de la base de datos de entrenamiento. Estadísticos QSE calculados con  $M_{opt}$  en la base de datos completas, intervalos de confianza al 95% (CI) y separabilidad (UT) . . . . . 107
- 7.91. Caracterización estadística de CosEn ( $m=2$ ) sobre señales de hipertensión arterial (HTA) en sujetos control (CT) y patológicos (PT). Se muestra test de normalidad (JB), intervalos de confianza y test de separabilidad (UT) de cada variable considerada, presión sistólica (SBP), diastólica (DBP), media (PAM) y frecuencia cardíaca (FC). . . . . 109



# Abreviaturas

## Acrónimos de las medidas de regularidad

ApEn	Entropía Aproximada (Approximate Entropy)
CosEn	Coefficiente de Entropía Muestral (Coefficient of Sample Entropy)
DFA	Análisis de fluctuaciones sin tendencias (Detrended Fluctuation Analysis)
MSE	Entropía multiescala o en escalas múltiples (MultiScale Entropy)
RSE	Entropía cuadrada de Rényi (Rényi Square Entropy)
SampEn	Entropía muestral (Sample Entropy)

## Parámetros de las medidas de regularidad

$m$	Longitud de los patrones, dimensión embebida
$r$	Umbral del filtro
$\alpha$	Coefficiente de escalado de DFA
$l$	Orden de DFA (polinomio de integración)
$S$	Escala de MSE
$M$	Número de patrones coincidentes ( <i>matches</i> )

## Acrónimos de señales

ADN	Ácido Desoxirribonucleico
AF	Fibrilación auricular
DBP	Presión arterial diastólica
ECG	Electrocardiograma
EEG	Electroencefalograma
EKG	Electrograma
Fc	Frecuencia cardíaca
HRV	Variabilidad del ritmo cardíaco
HTA	Hipertensión arterial
MBP	Presión arterial media
SBP	Presión arterial sistólica

## Acrónimos de clasificación

CT	registro de control
DM	registro diabético
HD	registro enfermedad de Huntington
LSA	registro esclerosis lateral amiotrófica
MS	registro síndrome metabólico
NDM	registro no diabético
NSV	registro en sujeto no superviviente
POST	registro posterior a un evento
PRE	registro anterior a un evento
PT	registro patológico
PD	registro enfermedad de Parkinson
SV	registro en sujeto superviviente

## Unidades de medida

bpm	latidos por minuto
mmHg	milímetros de mercurio
mmol/l	miligramos de mol por litro
s	segundos
ms	milisegundos

## Otros acrónimos

$C_{xy}$	Coefficiente de correlación cruzada
CI	Intervalos de confianza al 95 %
JB	Test de hipótesis de Jarque–Bera
UT	<i>Mann–Whitney U Test</i>
$\mu$	Media de una secuencia numérica
$\sigma$	Desviación estándar de una secuencia numérica

# Capítulo 1

## Introducción

Las señales biológicas se caracterizan por presentar un comportamiento que cambia a lo largo de tiempo, es decir, son señales no estacionarias, que a su vez y debido al múltiple número de entradas de las que está compuesto su sistema de generación, presentan un comportamiento no lineal.

Generalmente estas señales se han venido analizando y caracterizando mediante técnicas temporales, frecuenciales o de tiempo–frecuencia, siendo en su mayoría técnicas de análisis lineal y asumiendo en mayor o menor grado una existente estacionareidad y/o ergodicidad que no tiene por qué existir [1, 2, 3, 4]. Por ejemplo, durante la fibrilación auricular, la actividad eléctrica de la aurícula presenta un gran nivel de desorganización [5, 6].

Para satisfacer la demanda de análisis de dichos sistemas y poder llevar a cabo su completa caracterización, se ha dispuesto del estudio del caos para ver si hay alguna ley interna que permita clasificar a los distintos grupos de señales.

El estudio del caos nace como alternativa a los métodos tradicionales para clasificar sistemas dinámicos. Las técnicas típicas para el estudio del caos son, entre otras, las medidas de dimensión [7, 8], de entropía [9, 10, 11] y el espectro de Lyapunov [12].

Las medidas de dimensión necesitan de un gran número de muestras antes de converger a un resultado robusto. Se establece que entre  $10^d$  y  $30^d$  muestras son necesarias para cubrir un atractor  $d$ -dimensional [9, 10]. Las señales biológicas, generalmente, no mantienen sus características temporales para un gran número de muestras, de forma que no se puede obtener una medida de la dimensión robusta. El espectro de Lyapunov se sustenta sobre los trabajos previos de Kolmogorov–Sinai [13], métodos que asumen cierta ergodicidad de las señales y poco ruido,

circunstancias que no se dan en la realidad. Finalmente, aunque algunas medidas de entropía se sustentan también sobre principios de ergodicidad, Pincus en 1991 [9, 14] describió por primera vez una familia de estadísticos que no requieren de una alta dimensionalidad y pueden caracterizar y diferenciar entre los distintos sistemas existentes. Richman y Moorman [10] proponen la entropía muestral (SampEn), como mejora a la entropía aproximada (ApEn) al reducir el sesgo en el análisis de señales de longitud corta ( $N < 1000$ ). Costa [15] propone la entropía multiescala (MSE) para el análisis de señales de larga duración ( $N \geq 10.000$ ).

Estas medidas de entropía, entre otras, funcionan muy bien en la práctica, extrayendo información de las señales que no es posible detectar por los métodos clásicos lineales [9, 10, 16]. Sin embargo, no son muy adecuados para el análisis de conjuntos cortos y contaminados por factores extrínsecos a su sistema de generación, algo frecuente en las señales biomédicas [8, 14, 17, 18, 19, 20, 21]. Por ejemplo ApEn presenta un sesgo dependiente de la longitud que SampEn pretende corregir aunque no lo consigue erradicar completamente [10]. Por otro lado, en análisis de fluctuaciones sin tendencias (DFA) o la dimensión de correlación (CD) necesitan de secuencias muy largas para ofrecer resultados robustos [8, 21] que generalmente no pueden obtenerse en señales biológicas. Estas situaciones, secuencias de longitud reducida, artefactos y pérdida de muestras o tomas inválidas en los registros no han sido previamente caracterizadas ni valoradas en otros trabajos salvo para algunos casos o señales.

Kim en [22] realiza una caracterización de ApEn, SampEn, DFA y representación de Poincaré sobre una señal de variabilidad del ritmo cardíaco (HRV) que se deriva de registros RR que han perdido un segmento de duración determinada. Govindan [19] realiza una caracterización de DFA ante pérdida de muestras tanto uniforme como aleatoria sobre una señal HRV. Lake [17] presenta un muy breve estudio del coeficiente de correlación cruzada cuando las series de ritmo cardíaco (HR) han perdido muestras de forma aleatoria. Esto lleva a plantearse, ¿Presenta la entropía un mismo comportamiento independientemente de la señal biológica que caracterizan?, ¿Presentan las entropías el mismo comportamiento independientemente de la longitud de las señales biológicas?, ¿Influye el tipo de pérdidas sobre la entropía?, entre otra muchas preguntas.

En este trabajo se ha propuesto buscar la respuesta a las preguntas planteadas en el párrafo anterior, mediante la caracterización de algunas de las medidas de entropía más comunes frente a la pérdida de muestras y la ausencia de ellas en distintos registros biológicos.

Como medidas de entropía a caracterizar se han seleccionado ApEn [9], SampEn [10], MSE [15], DFA [8], entropía cuadrada de Rényi (QSE) [20] y el coeficiente de entropía muestral (CosEn) [23]. ApEn, SampEn, MSE y DFA han sido ampliamente utilizadas pero no han sido caracterizadas frente a la pérdida de muestras salvo en los casos particulares previamente mencionados. En este trabajo se ha de-

---

cido caracterizarlas frente a pérdida de muestras tanto de forma aleatoria como uniforme, sobre distintos tipos de registros biológicos como son los electroencefalogramas (EEG), los registros RR y los electrogramas (EGM). En segundo lugar, QSE y CosEn han sido estudiadas sobre registros HRV y HR de corta duración (apenas 12 muestras) con resultados satisfactorios [20, 23]. Por ello se han seleccionado registros de hipertensión arterial en humanos y registros de glucemia, ambos inferiores a 100 muestras de duración. Finalmente se pretende caracterizar otros tipos de señales, donde bien el muestreo es uniforme (registros de glucemia) o no uniforme (dinámicas de movilidad) pero cuya longitud apenas alcanza las 300 muestras, y ver la influencia que presentan los parámetros de cada una de las medidas propuestas (ApEn, SampEn, MSE y DFA) en su correcta identificación.

Esta tesis doctoral se estructura de la siguiente forma: en el capítulo 2 se presenta un estado del arte de los estudios de complejidad aplicados a señales biológicas en la actualidad. En el capítulo 3 se presentan los objetivos fundamentales de la tesis. Los capítulos 4 y 5 presentan los métodos y el conjunto experimental utilizados a lo largo de este trabajo, respectivamente. En el capítulo 6 se presentan los experimentos para la caracterización de las medidas de complejidad llevados a cabo y en el 7 los resultados obtenidos en cada uno de ellos. El capítulo 8 presenta la discusión de los resultados y el 9 desarrolla la conclusión, dejando abiertas posibilidades de investigación en el capítulo 10 con las líneas futuras. Finalmente, las aportaciones de la tesis, expresadas en forma de artículos publicados en revistas científicas o contribuciones a congresos, se exponen en el capítulo 11.



## Capítulo 2

# Estado del arte

Las primeras medidas de complejidad para cuantificar la cantidad de información en sistemas caóticos se basan en identificaciones y cambios en los comportamientos de los denominados atractores extraños [24, 25].

Los atractores extraños identifican a los sistemas que evolucionan después de un tiempo suficientemente largo en los cuales las trayectorias de variación deben permanecer próximas incluso si son ligeramente perturbadas y se considera que tienen estructura en todas las escalas, es decir, estructura fractal [25].

Para cuantificar dichos atractores en medicina y fisiología se utilizaron técnicas de análisis fractal [24, 25] o los exponentes de Lyapunov [12]. El problema de estas técnicas radica en el coste computacional asociado a los algoritmos, además de necesitar un gran número de muestras para poder ofrecer resultados robustos. El número de muestras necesario oscila entre  $10^d$  y  $30^d$  para un atractor de dimensión  $d$  y su coste computacional crece de forma exponencial con la dimensión [12, 26].

Como forma de reducir el coste computacional asociado al cálculo de dimensiones, Grassberger y Procaccia desarrollan la entropía de Kolmogorov–Sinai (KS) [13]. La entropía KS resulta ser muy sensible al ruido, incluso cuando éste es muy bajo. Su aplicación no supervisada puede inducir a resultados erróneos ya que termina evaluando el ruido presente y no el propio sistema complejo [14].

Hasta ese momento las entropías servían para cuantificar la cantidad de información generado por un sistema, a partir de [9] van a servir para caracterizar la complejidad del mismo y poder revelar anomalías allí dónde los estadísticos tradicionales fallan, ya que de otra forma no serían aparentes.

De nuevo se genera un problema, no sólo el coste computacional, sino la robustez frente artefactos o ruido. Es por ello que en 1991 Pincus propone una familia de

estadísticos denominado Entropía Aproximada (ApEn) [9, 14]. ApEn se trata de una modificación de KS que realiza comparaciones temporales de la señal a escalas mayores a la mayoría del ruido de forma que solo se ve comprometida en ratios señal-a-ruido inferiores a 3 dB [14].

ApEn presenta una mejora en cuanto al coste computacional del algoritmo, puesto que con apenas 50 muestras puede dar resultados bastante aproximados a los teóricos. Se ve poco afectada por el ruido, siempre y cuando el umbral seleccionado sea superior a la mayor parte de éste [14].

Con el tiempo se descubre que ApEn es dependiente de la longitud de la señal y presenta un sesgo bastante notable cuando ésta es corta. Por ello Richmann y Moorman presentan SampEn [10]. SampEn nace como una mejora de ApEn en la cual ya no se usan patrones para las comparaciones sino que se eliminan los denominados *automatches*, es decir, las similitudes con las mismas secuencias, consiguiendo reducir el sesgo introducido por ApEn.

Estas dos medidas son las más populares y comunmente usadas pero con el tiempo se han ido introduciendo modificaciones y definiendo nuevas medidas en función del aumento de las necesidades de los estudios.

Por ejemplo, para el estudio de secuencias de corta duración se ha desarrollado el coeficiente de entropía muestral (CosEn) [23] o la entropía cuadrada de Rényi (QSE) [20]. CosEn deriva de la entropía QSE [11] y ambas se pueden obtener a partir de SampEn. El problema de ApEn y SampEn radica en que no son medidas absolutas sino comparativas y que su resultado depende de los parámetros de entrada del sistema:  $(m, r)$ . Es aquí cuando se decide rescatar la entropía de Rényi y relacionarla con SampEn dando lugar a la QSE, de forma que se elimine la dependencia con  $r$  y aun así, se mantenga la capacidad comparativa de la medida [20]. CosEn aparece más tarde, cuando se considera que los estadísticos de primer orden pueden contener información relevante para la complejidad [23].

Para estudios en los que SampEn o ApEn no dan resultados suficientemente válidos, nace también la entropía Multiescala (MSE). Esta medida trata de valorar las diferencias no solo en variaciones temporales sino en escalas mayores. De nuevo al igual que CosEn o QSE, se estima a partir de SampEn sobre señales de mayor longitud [15, 27].

Se siguen manteniendo técnicas de análisis de dimensión fractal como el análisis de fluctuaciones (DFA) [8] o la entropía multifractal [28], pero ya no son necesarias longitudes tan largas. Ambas medidas son usadas en sistemas de variación temporal lenta como son, por ejemplo, los niveles de glucosa en sangre.

No obstante hoy en día se siguen utilizando un gran número de medidas de complejidad. Es una técnica poco explorada. Entre otras medidas de complejidad aparte de las nombradas con anterioridad se puede destacar: la complejidad de Lempel-

---

Ziv (LZC) [29, 30], ApEn y SampEn cruzadas [10], la entropía borrosa [31], la entropía espectral [32], la entropía delta [33] la información mutua [34], dimensión de correlación [35] o el uso de Wavelets [36], entre otras.

En el ámbito de aplicación médica o biomédica, estas medidas se han venido utilizando de forma constante para caracterizar la regularidad de las señales y así caracterizar el estado patológico de un sujeto, para predecir la posible aparición de eventos patológicos como pueden ser ataques epilépticos o distinguir entre etapas del sueño.

ApEn se ha utilizado para caracterizar el estado patológico de neonatos a partir del análisis de la variabilidad del ritmo cardíaco (HRV) [14] o del ritmo cardíaco (HR) [10, 37, 38] o registros RR [18], para identificar la existencia de sufrimiento fetal mediante el análisis de señales HR [39] y para la detección automática no supervisada de ataques epilépticos en señales electroencefalográficas (EEG) [40] entre otras.

SampEn se ha utilizado para caracterizar o diferenciar entre estados sanos o patológicos [10, 17] o determinar si hay sufrimiento fetal [39] sobre señales HR, para estudiar las variaciones de HRV durante apnea obstructiva del sueño [41]. Para la caracterización de la fibrilación auricular mediante el análisis del ECG [42, 43] o predecir el final de un evento de fibrilación auricular [44].

DFA ha sido utilizada para caracterizar las dinámicas de variación del EEG [45] en situaciones de apnea del sueño [3], en sujetos con Alzheimer [46], con traumatismos cerebrales [47] o el estudio del sueño [48, 49]. Se ha utilizado también en genómica en el estudio de las secuencias de ADN [50], y para evaluar los efectos de la postura, el ejercicio y la edad o el bloqueo autonómico en la presión sanguínea y la HRV [51, 52], para la caracterización de HR [8, 19, 53]. Ha sido utilizada también en el estudio de la diabetes a través de los niveles de glucemia [54, 55, 56, 57, 58] y en electromiogramas [59] entre otras.

MSE se aplica sobre señales HR para determinar si hay sufrimiento fetal [39] o valorar condiciones patológicas [15, 60], en la caracterización del ECG durante el embarazo [61]. Se ha aplicado para caracterizar las dinámicas de variación del EEG en autismo [62], Alzheimer [63] o esquizofrenia [64].

La dimensión de correlación se ha utilizado en el análisis de HRV en pacientes con cardiomiopatía dilatada [65], para diferenciar entre sujetos patológicos y sanos evaluando el ECG [7], para evaluar la dependencia entre el desarrollo neuronal y la complejidad en EEG de neonatos [66], en el estudio de sueño [67] o la caracterización de las dinámicas de variación del EEG en sujetos con migrañas nocturnas espontáneas [68] y en la presión sanguínea de animales [69].

La complejidad de Lempel–Ziv ha sido utilizada para caracterizar la presión arterial en condiciones patológicas [70] en la caracterización de EEG [71]. Y la entropía

multifractal se ha considerado para el estudio neuronal [72, 73], para la evaluación de patologías y la edad en las alteraciones del HR [74], entre otras muchas aplicaciones.

# Capítulo 3

## Objetivos

Las métricas de entropía son métodos de análisis no lineal que cuantifican la capacidad de generar información por parte de un sistema [9]. Estas medidas son ampliamente utilizadas en el contexto del procesamiento de señales biomédicas para, entre otras cosas, identificar patologías, clasificar sistemas, o incluso predecir la ocurrencia de eventos. Estos métodos ofrecen excelentes prestaciones en la extracción de información de las señales que de otra forma no sería posible obtener [10, 14].

Sin embargo, las prestaciones de estas medidas se degradan significativamente cuando las señales a caracterizar son ruidosas, presentan artefactos, han perdido muestras o su longitud es muy reducida, entre otras anomalías. Por ejemplo, el electroencefalograma puede estar contaminado por otras señales biomédicas como el oculograma, ruido procedente del dispositivo de medida o incluso de los electrodos, enmascarando la señal original [75]. El electrocardiograma suele aparecer contaminado por ruido muscular, ruido del dispositivo de medida, o por la respiración, entre otros factores [75]. Algunas de estas interferencias pueden ser reducidas mediante técnicas clásicas de preprocesado de señales [75], pero esto no siempre es posible o las distorsiones no pueden ser eliminadas completamente.

Por tanto, la cuestión que surge es cómo se comportan las medidas de entropía cuando existen factores anómalos en las señales que no pueden ser evitados o corregidos. Para poder contestar a esta pregunta fundamental es necesaria una caracterización completa y particular de estas medidas en cada caso.

Específicamente, en esta tesis se aborda dicha caracterización en el caso concreto de pérdida o insuficiente número de muestras. Por ejemplo, cuando ocurre una saturación del dispositivo de medida durante un cierto intervalo de tiempo o cuando se realiza una compresión con pérdidas [76, 77], si una interferencia radio externa

impide la correcta recepción de una señal inalámbrica de forma completa, o si un canal de transmisión inestable provoca que se pierdan muestras de forma aleatoria [78, 79, 80, 81], esto puede afectar al cálculo de las medidas de entropía. En otros casos las señales biológicas pueden presentar muy pocas muestras. Esto podría ser debido a que la toma de datos es manual, el registro es incómodo para el paciente, o por las características específicas del registro su duración es muy limitada, lo cual también puede distorsionar los resultados de la estimación de entropía.

Esta tesis por tanto tiene como objetivo principal el proporcionar las pautas necesarias para saber en qué condiciones y bajo qué supuestos se pueden aplicar ciertas medidas de estimación de entropía en este contexto, de forma que se realice una correcta interpretación de los resultados cuando los registros presenten pérdida o ausencia de muestras. Para ello se han fijado los siguientes aspectos a abordar:

- (i) Caracterización de medidas de entropía sobre señales que han sufrido una pérdida de muestras de forma aleatoria.
- (ii) Caracterización de medidas de entropía sobre señales que han sufrido una pérdida de muestras de forma uniforme.
- (iii) Caracterización de medidas de entropía en señales de duración limitada.

Como se ha descrito en los capítulos de introducción y estado del arte, son muchas las medidas de complejidad existentes y sería inabordable su completa caracterización en un único trabajo. Por ello, se han seleccionado aquellas medidas de entropía más representativas y utilizadas en este campo: la Entropía Aproximada (ApEn) [9], la Entropía Muestral (SampEn) [10], la Entropía en Múltiples Escalas (MSE) [15], el Análisis de Fluctuaciones sin Tendencias (DFA) [8], la Entropía Cuadrada de Rényi (QSE) [20] y el Coeficiente de Entropía Muestral (CosEn) [23]. Las medidas ApEn, SampEn, MSE y DFA han sido seleccionadas por ser las de uso más común sobre señales biológicas. Todas ellas pueden presentar problemas cuando los registros son de duración limitada, asociados propiamente al algoritmo de cálculo, como puede ser el caso de DFA o MSE [8, 15] o bien por errores de interpretación en los resultados, como sucede con ApEn y SampEn [10, 37, 38]. Para caracterizar estas limitaciones se ha realizado un estudio paramétrico de dichas medidas sobre señales de longitud aproximada de 300 muestras. Por otro lado, CosEn y QSE han sido seleccionadas por tratarse de medidas de reciente proposición para el análisis de registros de muy corta duración con resultados satisfactorios sobre señales HR con apenas 12 latidos [20, 23, 82].

Otra cuestión planteada en la caracterización a abordar en la presente tesis concierne al tipo o propiedades de los registros sobre los que se aplican dichas medidas de entropía. Es decir, se debe estudiar si presenta el mismo comportamiento una medida de entropía sobre distintos tipos de registros biológicos, y si no es así, determinar el por qué y las consecuencias prácticas. Con este objetivo se ha seleccionado un amplio y variado conjunto de registros fisiológicos. Se usarán señales

---

EEG, RR y EGM para abordar los puntos (i) y (ii) expuestos anteriormente, y señales de dinámicas de movilidad, glucemia e hipertensión arterial para el (iii). Estas señales cuantifican algún proceso fisiológico, pero a su vez presentan distintas características en términos de componentes frecuenciales, ancho de banda, frecuencias de muestreo y dinámicas de variación.



# Capítulo 4

## Métodos

En este capítulo se presentan y describen las medidas de complejidad y/o regularidad objetivo de la caracterización de la tesis.

Las medidas de complejidad son una familia de estadísticos que permiten evaluar la similitud de una serie temporal a lo largo del tiempo sin asumir propiedades de estacionariedad o lineales de la misma como en el caso de autocorrelaciones. También miden la capacidad de predicción de las señales en función de muestras anteriores.

Entre el gran número de medidas de complejidad disponibles, en este trabajo se ha decidido caracterizar la entropía aproximada (ApEn), la entropía muestral (SampEn), la entropía multiescala (MSE), el análisis de fluctuaciones (DFA), la entropía cuadrada de Rényi (QSE) y el coeficiente de entropía muestral (CosEn).

### 4.1 Entropía aproximada

La entropía aproximada o más conocida por su denominación en inglés, *Approximate Entropy (ApEn)*, se trata de una familia de estadísticos desarrollada para cuantificar el concepto de complejidad variante en un sistema. Su principal uso es el de detectar anomalías en los registros de larga duración que no son de otra forma aparentes [9].

ApEn se define como el ratio de verosimilitud entre la probabilidad condicional de que secuencias o patrones de longitud determinada considerados similares o próximos continúen siéndolo cuando la longitud de dichas secuencias se aumenta en una muestra. Representa la capacidad media de que aparezcan patrones similares y se caracteriza por ser capaz de distinguir una gran variedad de sistemas, entre ellos,

sistemas deterministas con baja dimensionalidad, periódicos y con periodicidad múltiple, sistemas caóticos, estocásticos o una mezcla de ellos [9, 14].

Se ha utilizado en infinidad de aplicaciones por ser considerada una medida muy estable frente a artefactos numéricos y/o estimaciones erróneas, siempre y cuando sean infrecuentes aunque sus valores sean extremos. En [14] se utiliza para analizar señales de ritmo cardíaco para diferenciar entre neonatos sanos o enfermos siendo capaz de diferenciar características que pasarían desapercibidas pero que indican un principio de enfermedad o para identificar si existe o no sufrimiento fetal [39]. Para el análisis de señales de variabilidad del ritmo cardíaco y registros RR [18]. Detección de eventos epilépticos en señales EEG [40] y para establecer el balanceo natural en niños con desarrollo tardío [83].

El cálculo de ApEn para una secuencia temporal  $u(i)$  de longitud  $N$  muestras responde al siguiente algoritmo [14]:

**1. Definir los subvectores:**

Se definen los subvectores de longitud  $m$  de acuerdo a:

$$x(i) = [u(i) u(i + 1) \dots u(i + m - 1)]^T \quad i = 1, 2, \dots, N - m \quad (4.1)$$

**2. Definir la medida de similitud:**

Para establecer si los patrones o subvectores definidos en (4.1) son próximos es necesario establecer una medida de similitud o distancia. En este caso dicha distancia responde a:

$$\begin{aligned} d(i, j) &= d(x(i), x(j)) \\ &= \max_k \{|x(i + k - 1) - x(j + k - 1)|\} \quad k = 1, 2 \dots m \end{aligned} \quad (4.2)$$

**3. Calcular la regularidad:**

La regularidad o frecuencia de aparición de los patrones definidos en (4.1) identificados como similares a uno dado se define como  $C_i^m(r)$ . Se consideran patrones similares aquellos en los que la distancia de similitud definida en (4.2) es inferior a un valor dado. Dicho valor se denota como  $r$  y representa, en esencia, un nivel de filtrado.

$$C_i^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} (d(i, j) \leq r) \quad (4.3)$$

**4. Calcular las probabilidades condicionales:**

La probabilidad condicional,  $\Phi^m(r)$ , de que dos patrones de longitud  $m$  sean similares se calcula de acuerdo a:

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln \{C_i^m(r)\} \quad (4.4)$$

### 5. Obtener ApEn:

ApEn se obtiene como el ratio de verosimilitud entre la probabilidad condicional definida en (4.4) para subvectores de longitud  $m$  y  $m + 1$  muestras respectivamente, de acuerdo a:

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \quad (4.5)$$

La cota superior de ApEn para una serie temporal de duración finita y longitud  $N$  es  $\log(N - m)$ . Se trata de una medida no negativa e invariante frente a traslaciones o escalados. No se puede ver como una medida absoluta puesto que para un mismo sistema, al cambiar los valores de  $m$  y  $r$  su valor cambia de forma notable y además si el ruido del sistema en grande, es decir, para ratios de señal a ruido inferiores a 3 dB, la validez de ApEn puede verse seriamente comprometida [9, 14].

Señales con alto grado de regularidad, en las cuales se puede reconocer al menos un patrón que se repite obtendrá probabilidades condicionales en (4.4) proximas a 1 de forma que sus logaritmos tenderán a 0. Esto se traduce en valores de ApEn bajos, al contrario de lo que sucede con señales que presentan un comportamiento aparentemente aleatorio donde dichas probabilidades seran próximas a 0 y en consecuencia ApEn obtendrá valores altos [38].

Aunque ApEn ha sido utilizada para una gran variedad de aplicaciones, no se ha establecido una forma única de definir los parámetros  $m$  y  $r$ . Si  $r$  se selecciona demasiado pequeño, las probabilidades condicionales que se obtendrían resultarían en estimaciones burdas respecto de las reales además de numéricamente inestables, mientras que si se selecciona demasiado grande la información útil puede pasar desapercibida [9, 14, 38]. Por ello se establece que para evitar una contribución significativa del ruido,  $r$  debe ser superior a la mayor parte de éste. En [14] se han obtenido buenos resultados con un valor de  $r$  al menos 3 veces la amplitud media del ruido.

En [9, 38] se establece que con  $m = 2$  y  $0,10 \leq r/\sigma_u \leq 0,25$  (siendo  $\sigma_u$  la desviación estándar de la serie a analizar  $u(i)$ ) se pueden obtener aproximaciones de ApEn para señales de un mínimo de 1000 muestras con una validez estadística razonable. En [14] se dice que con  $m = \{2, 3\}$  se asegura que las probabilidades condicionales definidas en (4.4) sean estimadas de forma razonable, siempre y cuando la longitud de la serie original  $u(i)$  sea de al menos  $10^m$  y preferiblemente  $30^m$  muestras (generalmente entre 75 y 5000 muestras [37]).

Aunque se trata de un estadístico sesgado, puede ser asintóticamente insesgado para muchos procesos dependiendo de la longitud  $N$  de la señal a analizar. Este

sesgo indica una mayor regularidad y/o similaridad en la señal de la que realmente existe al proporcionar valores inferiores a los reales [10, 38]. Además del sesgo, los errores de ApEn dependen de la desviación estándar de la señal y de la frecuencia de muestreo. El remuestreo introduce a su vez un sesgo en la medida que depende de la frecuencia a la cual se remuestrea [18].

Finalmente, si lo que se desea es comparar sistemas, para que la comparación sea consistente se deben utilizar los mismos valores para  $m$ ,  $r$  y  $N$  [9].

## 4.2 Entropía muestral

La entropía muestral conocida como SampEn (acrónimo de *Sample Entropy*), al igual que la entropía aproximada, se trata de una familia de estadísticos desarrollados y caracterizados para reducir el sesgo asociado a ApEn. Al no tener en cuenta un patrón referencia se impide que cualquier secuencia se compare consigo misma [10].

SampEn se define como el logaritmo natural negativo de la probabilidad condicional de que dos secuencias que siendo similares para  $m$  puntos lo sean también para  $m + 1$  puntos, siendo estas dos secuencias siempre distintas [10]. Se trata de un algoritmo más simple y con menos coste computacional que el de ApEn.

Al contrario que ApEn, SampEn es muy independiente de la longitud de la señal a analizar  $N$  y en consecuencia se ha utilizado para analizar registros de apenas 100 muestras. Presenta una mayor consistencia en circunstancias donde ApEn no lo hace [10, 41]. SampEn se calcula de forma similar a ApEn pero presenta algunas modificaciones. De acuerdo a Richmann y Moorman [10], el cálculo de SampEn para una secuencia  $u(i)$  de longitud  $N$  muestras, se define como:

$$SampEn(N, m, r) = -\log \frac{A^m(r)}{B^m(r)} \quad (4.6)$$

donde  $B^m(r)$  representa la probabilidad de que dos secuencias coincidan para  $m$  muestras con una tolerancia  $r$ , y  $A^m(r)$  lo hagan para  $m + 1$  muestras. En consecuencia SampEn es la probabilidad condicional de que dos secuencias que coinciden en  $m$  puntos con una tolerancia  $r$  lo hagan también para  $m + 1$ . Estas probabilidades se definen de acuerdo a:

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \quad (4.7)$$

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \quad (4.8)$$

$$A_i^m(r) = \frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} (d(i, j) < r) \quad (4.9)$$

$$B_i^m(r) = \frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} (d(i, j) < r) \quad (4.10)$$

donde  $d(i, j)$  es la distancia entre dos subvectores de  $m$  muestras en el caso de  $B_i^m(r)$  y  $m+1$  muestras para  $A_i^m(r)$  según se definió para ApEn en (4.2) y los subvectores se construyen de acuerdo a (4.1). Los parámetros  $m$  y  $r$  son los mismos que los definidos para ApEn en la sección 7.1.

SampEn es asintóticamente insesgada y presenta una cota máxima  $SampEn \leq \log(N-m) + \log(N-m+1) - \log(2)$ . Tiene un mayor ajuste que ApEn a las estadísticas de la teoría de números aleatorios, cuyas características probabilísticas son conocidas, en un ancho rango de condiciones operativas. Presenta un sesgo residual para secuencias  $u(i)$  de muy corta duración, en gran parte asociado a que los subvectores no son completamente independientes [10].

La precisión y confianza de las estimaciones de SampEn dependen del número de coincidencias. A mayor número de éstas, más robusta es la estimación obtenida. El número de similitudes o coincidencias puede ser aumentado al igual que en el caso de ApEn aumentando  $r$  o disminuyendo  $m$ . En el primer caso (aumenta  $r$ ) SampEn tiende asintóticamente a 0 para todos los procesos, mientras que en el segundo caso (disminuir  $m$ ) se puede perder información relevante de procesos fisiológicos que no son aparentes en esos niveles [17, 84]. Si el número de coincidencias es el adecuado, SampEn se puede asumir como procedente de una distribución de probabilidad normal [17].

Esta medida de regularidad ha sido usada en multitud de ocasiones para caracterizar distintos procesos biológicos como en señales de variabilidad del ritmo cardíaco (HRV) en neonatos [17, 84], en apnea del sueño [41], análisis de registros RR y QT en distintos estados [85] y fibrilación auricular [42], entre otros muchos.

### 4.3 Entropía multiescala

De acuerdo a [15], cuando la longitud  $N$  de la serie temporal  $u(i)$  a analizar es elevada, la entropía, bien sea ApEn o SampEn, se subestima y tiende a 0. Por ello se desarrolló la denominada entropía multiescala (MSE) [15]. Las típicas medidas de complejidad de escala única (e.g. SampEn o ApEn) pueden, en algunos casos, asignar valores de entropía superiores a los reales a determinadas series patológicas que se sabe presentan dinámicas menos complejas que las que se derivan de los sistemas sanos [15].

La idea básica de MSE radica en que determinados sistemas biológicos pueden exhibir comportamientos dinámicos que distan mucho tanto de una regularidad perfecta o de una completa aleatoriedad y que dichas dinámicas típicamente revelan estructuras en distintas escalas tanto temporales como espaciales [15].

El cálculo de MSE se basa en el algoritmo de SampEn descrito anteriormente en la sección 4.2 pero sobre series con distinto escalado temporal. Es decir, sea  $u(n)$  la serie temporal sujeta a análisis, se deben construir las denominadas *coarse grained series*, o series con mayor escalado temporal  $y_j^\tau$  de acuerdo a [15]:

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} u(i) \quad j = 1, 2, \dots, \lfloor N/\tau \rfloor \quad (4.11)$$

donde  $\tau$  se conoce como la escala, y  $\lfloor \cdot \rfloor$  indica la parte entera por defecto. Finalmente, MSE se obtiene al aplicar SampEn según se definió en (4.6) sobre la señal escalada  $y_j^\tau$  obtenida de acuerdo a (4.11). Estas series, introducen un suavizado en la señal y una decorrelación entre los subvectores. Los parámetros  $m$  y  $r$  son los mismos a los definidos para ApEn y/o SampEn, pero en este caso, la tolerancia  $r$  se ajusta proporcional a la serie original  $u(i)$  de forma que sea igual en todas las escalas. Esto implica que la desviación estándar disminuye al aumentar la escala [60].

La selección óptima de la escala depende de tres factores: (i) el tiempo de muestreo  $\tau_s$ , (ii) el tiempo de correlación  $\tau_c$  y (iii) los períodos de posibles oscilaciones no lineales  $\tau_p$ . El tiempo de correlación se puede obtener mediante la función de autocorrelación y  $\tau_p$  a partir de la detección de picos en espectro de Fourier [60]. MSE ha sido utilizada en la caracterización de distintos procesos biológicos como pueden ser el análisis de EEG [62, 63, 64] o series derivadas del ECG como RR o HRV [27, 39] entre otras.

## 4.4 Análisis de fluctuaciones sin tendencias

El análisis de fluctuaciones sin tendencia o también conocido como *Detrended Fluctuation Analysis* (DFA) es una medida de regularidad que analiza las fluctuaciones temporales o correlaciones de larga duración presentes en una señal que es aparentemente no estacionaria durante un período de tiempo relativamente corto. Evita la asunción de linealidad y baja dimensionalidad del caos [8, 45]. Fundamentalmente se utiliza para describir las propiedades estadísticas de segundo orden de la señal que probablemente reflejen procesos de larga duración y con memoria, correlaciones de larga duración, escalas fractales o evoluciones no estacionarias intrínsecas [47]. Fue desarrollada en primera instancia por Peng *et al.* [8] para analizar las tendencias no estacionarias presentes en una señal de ritmo cardíaco.

Este método de análisis de regularidad se caracteriza por ser capaz de evitar la estimación de correlaciones en la señal de larga duración consideradas como espúreas al ser artefactos de la no estacionariedad implícita en la señal [8].

El método consiste en integrar una serie temporal una vez se haya eliminado su tendencia central y analizar las posibles fluctuaciones existentes en las distintas escalas temporales. Tal como se describe en su configuración original [8], considerando una serie temporal  $u(i)$  de duración  $N$  muestras, el primer paso consiste en eliminar la tendencia central y proceder a su integración para eliminar los posibles artefactos extrínsecos a la señal:

$$y(n) = \sum_{k=1}^n u(k) - \bar{u} \quad (4.12)$$

donde  $\bar{u}$  representa la tendencia central de la serie  $u(i)$  o su media. En algunos trabajos, debido a la corta duración de los registros, no se realiza la integración de la serie arriba descrita lo que hace que se pierda la capacidad de poder interpretar los resultados de forma simple [45]. En [19] se propone normalizar (4.12) por la desviación estándar de la señal, de forma que sea más fácil realizar una combinación lineal de las distintas funciones de fluctuación.

En segundo lugar, se divide la señal en un número de segmentos  $S$ , disjuntos y no solapados de igual longitud  $s$  definidos como  $y_s(n)$ . A cada uno de los segmentos se le realiza un proceso de eliminación de tendencia mediante una función polinómica  $y_{ks}(n)$  de forma que se pueda calcular el error cuadrático de fluctuación para cada uno de ellos. Éste error se define de acuerdo a [8, 45]:

$$F_s^2(K) = \frac{1}{K} \sum_{n=(s-1)K+1}^{sK} (y_s(n) - y_{ks}(n))^2 \quad (4.13)$$

donde,  $F_s^2(K)$  denota el error cuadrático medio de fluctuación en el segmento  $s$  de longitud  $K$  muestras. Este error generalmente aumenta de forma directamente proporcional a  $K$  [8]. El analizar la serie sólo de forma local en distintos segmentos permite que el método sea insensible a correlaciones espúreas introducidas por tendencias externas de baja frecuencia [45]

La función polinómica utilizada para eliminar la tendencia dentro de cada segmento, determina el orden de DFA. Suele utilizarse de grado 1 aunque en algunos casos se llega a utilizar grado 3, no siendo recomendable utilizar mayores grados puesto que se puede llegar a eliminar la señal en sí. Toda función o tendencia existente en la señal por debajo del orden de la función polinómica no tiene efecto en las propiedades de escalado (coeficiente DFA), haciendo este método estable y así dotándole de la capacidad para estimar de forma bastante exacta las correlaciones en señales ruidosas integradas en tendencias con funciones polinomiales [86].

A continuación se estima el error cuadrático medio de fluctuación para una longitud de segmento dada:

$$F^2(K) = \frac{1}{S} \sum_{i=1}^S F_s^2(K) \quad (4.14)$$

El procedimiento descrito en las ecuaciones (4.13) y (4.14), es repetido para distintos valores de  $K$ . Estos valores representan las distintas escalas temporales en las que se analiza la señal. Una vez obtenidos los distintos valores de la función de fluctuación  $F(K)$  para todos los valores de  $K$  considerados, se debe realizar una regresión lineal en el dominio logarítmico, de forma que el coeficiente DFA,  $\alpha$ , de la señal  $u(i)$  responda a:

$$F(K) \propto K^\alpha \quad (4.15)$$

El coeficiente de escalado,  $\alpha$ , caracteriza el tipo de correlaciones existentes en la señal y en consecuencia es siempre no-negativo. Cuando  $\alpha \in (0, 0,5)$  se dice que las correlaciones de la señal describen una función alternante en el tiempo, es decir, la señal es anticorrelada de forma que valores altos y bajos en la serie temporal tienden a aparecer de forma alterna o que dichas correlaciones no siguen ningún tipo de función temporal [8, 86, 87] Si  $\alpha \in (0,5, 1)$  indica que la señal tiene correlaciones de larga duración y persistentes en el tiempo definidas por una función de correlación positiva creciente o decreciente en el tiempo, esto es, valores altos de la señal aparecen seguidos de valores altos y viceversa [8, 86, 87]. Los puntos umbral caracterizan a dos tipos de ruido,  $\alpha = 0,5$  indica ruido blanco

y se dice que la señal es incorrelada,  $\alpha = 1$  indica ruido  $1/f$ . Finalmente  $\alpha = 1,5$  determina ruido *Browniano* [8, 86, 87].

La secuencia de ventanas o longitudes de segmentos  $K$  en los que se divide la señal para su análisis puede ser realizado de distintas formas. Se establece como valor mínimo un longitud  $K$  superior al orden del polinomio utilizado para eliminar la tendencia local del segmento en dos muestras, esto es,  $K_{min} = \text{grad}\{y_{ks}\} + 2$  [19]. Como la longitud superior depende del autor, algunos proponen utilizar la longitud máxima de la señal  $N$ , o reducirla a  $N/4$  [19] o  $N/10$  [46] para garantizar al menos más de una estimación de la función en  $K_{max}$ . Para definir las longitudes de las escalas intermedias se han propuesto distintas posibilidades. Por ejemplo, en [45] se propone hacerlo mediante intervalos equidistantes en la escala logarítmica. En [48] los límites superior e inferior de  $K$  se establecen mediante inspección visual de señales seleccionadas de forma aleatoria.

El grado del polinomio de integración para la eliminación de la tendencia local,  $l$ , debe ser seleccionado de forma que verifique  $\alpha \leq l$  [88].

El coeficiente de escalado de DFA,  $\alpha$  se considera una medida de regularidad en base a que cuanto mayor es este valor más suaves son las variaciones existentes en la señal [8]. Es bastante robusta, ya que aplicar interpolación sobre la función de fluctuación tras su cómputo no afecta a la estimación final de  $\alpha$ . La media de coeficientes estimados sobre distintas señales es igual a la media del coeficiente estimado sobre la función media  $F(K)$  de dichas señales [45]. Se ha establecido también que consigue reportar resultados válidos incluso cuando se han eliminado hasta un 50 % de las muestras de la señal de forma aleatoria o hasta 50 minutos consecutivos en registros de ritmo cardíaco fetal de 8h de duración [19], de forma que se puedan analizar señales que no hayan sido necesariamente registradas de forma continua en el tiempo .

Este método presenta sus limitaciones. La primera es que se necesitan registros largos, en torno a 24h, para que su estimación sea suficientemente robusta ya que de otra forma  $F(K)$  presenta gran variabilidad o la existencia de un punto en el cual la función de fluctuación presenta un cambio en su pendiente (punto de *cross-over*) [86, 89, 90]. El punto de *cross-over* puede estar asociado a un cambio en las correlaciones temporales de mayor duración debido a cambios en la propiedades correlativas de las señal en distintas escalas temporales o bien de distintas tendencias existentes en la señal como efectos circadianos o geoestacionarios [51, 90]. Este punto de *cross-over* puede ser utilizado para determinar el orden de las tendencias polinómicas presentes en la señal de análisis [86]. Dependiendo de cómo sean las tendencias intrínsecas a la señal se puede necesitar el uso de más de un coeficiente de escalado  $\alpha$  para determinar el comportamiento de la señal [8, 86, 89, 90]. Estas influencias si son cíclicas o periódicas en alguna forma pueden ser reducidas mediante una descomposición SVD de la señal, donde los autovalores de mayor peso están asociados a dichas tendencias [89].

Para solventar el problema asociado a secuencias de corta duración o las fluctuaciones aparentes cuando se estima  $F(K)$  en ventanas de larga duración [50] propone el uso de ventanas deslizantes. Esta modificación implica que en lugar de analizar intervalos disjuntos de forma independiente una única vez, se analizan varias veces considerando solapamiento de las ventanas en las sucesivas iteraciones. En este caso las ecuaciones en (4.13) y (4.14) quedan modificadas de acuerdo a:

$$F^2(K) = \frac{1}{(N - K + 1)(K - 1)} \sum_{i=0}^{N-K} \left( \sum_{k=n}^{n+K-1} (y_s(n) - y_{ks}(n))^2 \right) \quad (4.16)$$

Con esto se consigue que en ventanas de mayor duración se posea un mayor número de estimaciones para el cómputo de  $F(K)$  y obtener mejor un comportamiento estadístico.

Este método de análisis de regularidad ha sido utilizado de forma satisfactoria entre otras aplicaciones, en el análisis del ritmo cardíaco [8], ritmo cardíaco fetal [19], análisis de secuencias de ADN [50], análisis de electroencefalogramas en sujetos con apnea del sueño [3], o Alzheimer [46] en el análisis de fluctuaciones propias del EEG [45], para estudiar las distintas etapas del sueño [48] e incluso para establecer la profundidad de la anestesia en un paciente [87].

En este trabajo se ha elegido normalizar la función perfil en (4.12) de acuerdo a [19]. El orden de polinomio de eliminación de la tendencia local en cada segmento se ha seleccionado de acuerdo a [88] de forma que sea superior al coeficiente DFA de la señal. Finalmente, la secuencia de longitudes se ha definido como  $K_{min} = \text{grad}\{y_{ks}\} + 2$ ,  $K_{max} = N/4$  según [19] y equidistante en escala logarítmica con 50 longitudes distintas de acuerdo a [45], eliminando, si existen, valores iguales en escala lineal [87]. La señal se ha analizado dos veces, iniciando el análisis cada vez desde un extremo, de forma que nunca se descarten muestras en el caso de que  $K$  no sea un divisor entero de  $N$  de acuerdo a [19, 91].

#### 4.4.1 Estimación del punto de cross-over

Para la estimación del punto de cross-over se ha generado un algoritmo basado en el criterio de Akaike y en la minimización del error respecto a la recta de regresión lineal, [90, 92].

En primer lugar, mediante el uso del criterio de Akaike se ha estimado el grado mínimo del polinomio de ajuste a  $F(K)$  con error cuadrático medio mínimo a cada lado de cada posible punto de *cross-over*. Se consideran puntos de *cross-over* todos los valores  $K$  excepto los dos primeros y dos últimos de forma que se garantice que cada región contiene 3 puntos y poder así aplicar el algoritmo de mínimos cuadrados en el cual se basa dicho criterio.

La relación que se establece entre  $\log F(K) \sim \log(K)$  debe ser lineal (representable mediante un polinomio de grado 1). Las escalas temporales  $K$  en las que el criterio de Akaike garantiza una regresión lineal (polinomio de grado 1), se seleccionan como posibles puntos de *cross-over*. A cada lado de éstos se realiza una regresión lineal y se establece como punto de *cross-over* el que presenta menor error de aproximación entre  $F(K)$  y su regresión lineal, similar a la propuesta de [90], donde se establecen los puntos de *cross-over* mediante la intersección entre las regresiones lineales realizadas desde ambos extremos.

Se ha implementado este algoritmo para la estimación del punto de *cross-over* de forma autónoma y no supervisada. En la mayoría de la literatura existente, como por ejemplo, [46, 52] se realiza una inspección visual para establecer los límites y determinar el valor del *cross-over*. Tal como se ha descrito este punto es de interés puesto que su localización puede estar relacionada con la patología [52], pudiendo dar una idea del grado de enfermedad de paciente que haciéndolo mediante inspección visual puede no ser suficientemente robusto.

## 4.5 Entropía cuadrada de Rényi

La entropía de Rényi [11] surgió como una generalización y caracterización de la entropía de Shannon [93]. Esta medida cuantifica la pérdida de información en un sistema dinámico.

Rényi establece la entropía generalizada de orden  $\alpha$  de una función de distribución de probabilidad discreta  $\varphi = (p_1, p_2, \dots, p_n)$  donde  $p_k = P(\xi = x_k)$ ,  $k = 1, 2, \dots, n$  como:

$$H_\alpha(\varphi) = \frac{1}{1-\alpha} \log_2 \left\{ \frac{\sum_{k=1}^n p_k^\alpha}{\sum_{k=1}^n p_k} \right\} \quad (4.17)$$

donde la entropía de Shannon quedaría definida en el límite cuando  $\alpha$  tiende a 1 [11]. Suponiendo que la distribución es completa ( $\sum_{k=1}^n p_k = 1$ ), la entropía de Rényi queda:

$$H_\alpha(\varphi) = \frac{1}{1-\alpha} \log_2 \left\{ \sum_{k=1}^n p_k^\alpha \right\} \quad \begin{matrix} \alpha > 0 \\ \alpha \neq 1 \end{matrix} \quad (4.18)$$

De acuerdo a sus propiedades, la entropía de Rényi para sistemas no caóticos tiende a 0, para sistemas caóticos será superior a 0, mientras que para sistemas aleatorios tenderá a infinito [94]. Con el tiempo el logaritmo en base 2 se ha reemplazado por el logaritmo natural [95, 96].

Lake [20], establece que si una señal, en este caso biológica, puede considerarse suficientemente estocástica como para modelarla mediante un proceso aleatorio, sobre ella se puede aplicar la entropía de Rényi. El argumento del teorema central del límite sugiere que dicha entropía refleja la complejidad del sistema fisiológico en el proceso de transducción de la señal. En este artículo, [20], se deriva la fórmula que relaciona la entropía cuadrada de Rényi (QSE) a través de SampEn. Esta relación supone el cálculo de densidades de probabilidad en lugar de probabilidades condicionales como en SampEn. Lake considera que dividiendo la región en la que se buscan las secuencias similares, necesarias para el cálculo de SampEn, por el volumen de dicha región  $(2r)^m$ , se transforman las probabilidades en densidades y así se puede hacer uso de la entropía de Rényi, de forma que:

$$QSE(x, m) = SampEn(x, m, r) + \log(2r) \quad (4.19)$$

En este caso se elimina la dependencia de la medida con  $r$  ya que ésta puede ser optimizada de forma individual para cada señal y así obtener unos valores de complejidad más robustos. El hecho de mantener el mismo número de similitudes  $M$  en el cómputo de SampEn conserva la comparabilidad de la medida entre distintas señales [20].

Esta medida se ha utilizado en diversas aplicaciones como medida del contenido de información en tiempo–frecuencia [97], separación ciega de fuentes [96] usando la entropía cuadrada de Rényi, para clasificar efectos del daño cerebral en el balanceo postural [98], en señales de ritmo cardíaco [20], y caracterización de registros de voz [94].

En esta tesis se ha desarrollado un algoritmo basado en QSE, teniendo en cuenta la relación con SampEn definida por Lake en [20] para el análisis de señales de corta duración (longitudes inferiores a 100 muestras) donde SampEn y ApEn se ven comprometidas. Este algoritmo se basa en mantener el mismo número de secuencias similares en todas las señales a caracterizar y calcular el umbral en función de éstas para obtener una máxima segmentación entre clases. Consta de los siguientes pasos:

### 1. Cómputo de umbrales y número de similitudes:

Primero es necesario definir una longitud de secuencia  $m$  para la cual se calcula la entropía final. Generalmente y debido a la corta duración de las señales se selecciona  $m < 3$  y generalmente aquella que maximice la función de autocorrelación entre segmentos [20].

Para cada señal  $u(i)$  de longitud  $N$ , se define que para SampEn, existe un número máximo de similitudes definido por:  $M_{max} = \frac{(N-m)(N-m-1)}{2}$ . Para cada posibilidad de ellas entre  $M = 1 : M_{max}$  se calcula el umbral  $r(M, u)$  en función de la distancia de similitud definida en (4.2). Se descartan aquellas

$M_i$  que no están presentes en todas las señales. La figura 4.1.a muestra  $r(M)$  para una señal aleatoria.

## 2. Cómputo de QSE:

Se calcula el valor inicial de QSE según (4.19) parametrizado en función de los  $M_i$  no descartados en el punto anterior, para cada una de las señales en la base de datos. La figura 4.1.b muestra el  $QSE(M)$  para una señal aleatoria de la base de datos.

## 3. Obtención $M_{opt}$

La idea principal es la segmentación entre clases. Para ello se busca maximizar la distancia de separación entre éstas y se sabe que ésta depende de  $M$ . Para alcanzar este objetivo, se utiliza el criterio de maximización del área bajo la curva ROC (AUROC). La estimación del valor  $M_i$  en el cuál se produce dicho máximo se obtiene mediante:

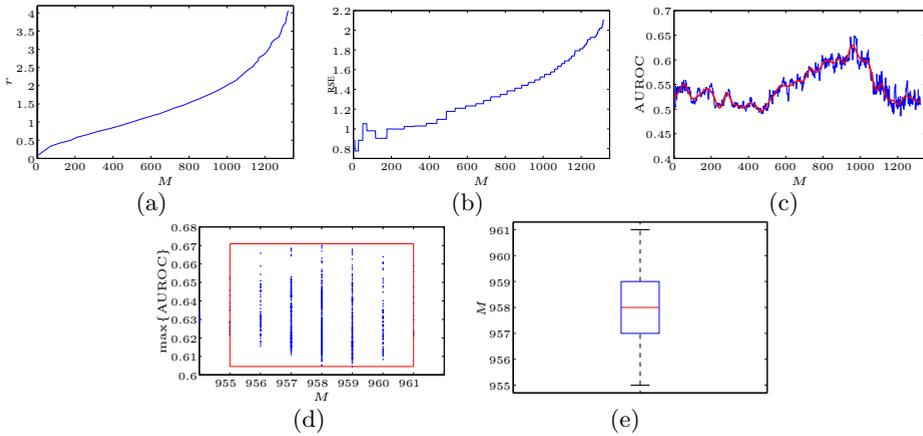
$$M_{opt} = \operatorname{argmax}\{\text{AUROC}(M)\} \quad (4.20)$$

Las curvas ROC (*Receiving Operating Characteristics*) presentan de forma gráfica la capacidad de un método de análisis no lineal para discriminar entre clases, generalmente entre sujetos sanos y sujetos que sufren algún tipo de patología. Representan la sensibilidad frente a la especificidad de clasificación del método. La sensibilidad indica la capacidad de clasificar correctamente a los sujetos (verdaderos positivos) mientras que la especificidad representa el porcentaje de sujetos adecuadamente identificados (verdaderos negativos). El AUROC indica de forma aproximada el rendimiento de método evaluado, un mayor valor de AUROC indica una mejor discriminación entre grupos mientras que cuando  $\text{AUROC} \rightarrow 0.5$ , el método es incapaz de separar o clasificar a los pacientes en sus respectivos grupos de forma correcta [99, 100, 101].

Para obtener  $M_{opt}$  de forma insesgada se ha aplicado un algoritmo de *cross-validation*. En bases de datos de menos de 100 señales se utilizará el *Leave-One-Out Bootstrap algorithm* (LOO), y si el número es superior se aplicará *K-fold Bootstrap algorithm*, de forma que no se introduzca un sesgo adicional asociado al método [102].

Para reducir el posible ruido existente en las  $\text{AUROC}(M)$  se aplica una descomposición Wavelet, donde la escala y nivel se seleccionan de forma automática de acuerdo a las características de las AUROC y se utiliza la familia Daubechies [103]. La figura 4.1.c muestra la  $\text{AUROC}(M)$  de una iteración de LOO en azul y su versión filtrada en rojo.

Sobre cada iteración del método se obtiene un  $M_{opt}(k)$  (ver figura 4.1.d). Para realizar una correcta estimación del máximo se aplica un filtrado de mediana sobre la serie de  $M_{opt}$  obtenidos definiendo una región en función



**Figura 4.1:** Estimación de QSE para una señal de hipertensión arterial. (a) Umbral  $r$  necesario en función del número de patrones coincidentes  $M$ . (b) Cálculo de QSE en función de  $M$ . (c) Estimación  $M_{opt}$ : curva del área bajo la ROC (AUROC) en función de  $M$  en azul y su versión filtrada en rojo. (d) Máximos detectados en función de  $M$  y región de filtrado de mediana en rojo. (e) Distribución de  $M_{opt}(k)$

de la mediana y la desviación mediana absoluta (MAD) y así eliminando los outliers de acuerdo a:

$$M_{opt}^j = \{M_{opt}(k) \in [\text{med}\{M_{opt}(k)\} \pm 1,5\text{MAD}\{M_{opt}(k)\}]\} \quad (4.21)$$

$$\text{MAD}\{x\} = \text{med}\{|x - \text{med}\{x\}|\} \quad (4.22)$$

donde  $\text{med}\{x\}$  denota el operador mediana. Sobre los  $M_{opt}^j$  que cumplen el criterio indicado en (4.21) se selecciona el óptimo de acuerdo a:

$$M_{opt} = \text{med}\{M_{opt}^j\} \quad (4.23)$$

## 4.6 Coeficiente de entropía muestral

El coeficiente de entropía muestral (CosEn), es utilizado y publicado por primera vez por Lake en [23]. Se trata de una modificación de SampEn para poder estimar la entropía de forma robusta en señales de muy corta duración y donde la información contenida en la media de la señal se considera relevante.

Para una señal  $u(i)$  de longitud  $N$ , CosEn se calcula de acuerdo a:

$$\text{CosEn}(u) = \text{SampEn} - \log(2r) - \log(\bar{u}) \quad (4.24)$$

Se observa como la expresión es muy similar a la definida en (4.19) y de hecho todo el algoritmo aplicado para el uso de esta medida es igual al aplicado para la entropía de Rényi dónde se ha cambiado simplemente la medida de complejidad utilizada, en lugar de QSE, se calcula CosEn. La principal diferencia con el resto de medidas de complejidad es que presenta generalmente valores negativos.

Esta medida ha permitido detectar de forma bastante robusta la fibrilación auricular con  $M = 5$  en apenas 12 latidos [23]. Se trata de una medida de reciente incorporación, por ello hay poco material al respecto. Un segundo artículo [82] la compara con otras medidas, pero también sobre señales RR. Nosotros hemos tratado de ampliar este hecho, aplicándola sobre otras señales de corta duración (hipertensión arterial y glucemia).

## 4.7 Métodos de Evaluación

### 4.7.1 Test de Segmentación

Para evaluar la capacidad de segmentación entre clases de las medidas de complejidad propuestas se han considerado distintas pruebas de hipótesis. En primer lugar se les ha aplicado el test de hipótesis de *Jarque-Bera* (JB) para determinar si las distribuciones de los valores de complejidad responden a un comportamiento normal o no [104]. En caso de que las medidas de complejidad respondan a un comportamiento normal ( $p > 0,05$ ) se considera aplicar el *Student T-Test* (TT2) y en caso contrario se recomienda aplicar el *Mann-Whitney U Test* (UT).

El *Student T-Test* (ver definición en 4.7.1) hace un estudio de las poblaciones en función de su media y desviación estándar y en función de su estadístico devuelve la probabilidad de que las dos distribuciones provengan de una distribución normal con igual valor medio. Por otro lado, el *Mann-Whitney U Test* se trata de un estadístico no paramétrico y el estudio lo hace en función de los valores medianos de la distribución. En ambos casos si la probabilidad devuelta por estos métodos,  $p$ , es inferior a 0.05 se considera que las distribuciones presentan medias/medianas distintas y por tanto los dos conjuntos son separables.

### **Student T-Test**

El Student T-Test realiza un t-test de hipótesis nula en cual considera que dos series numéricas proceden de distribuciones normales con igual media pero varianza desconocida.

Considerando dos series numéricas,  $x$  e  $y$ , de longitudes conocidas  $n_1$  y  $n_2$  respectivamente. El estadístico calculado por el *T-Test* responde a:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}} \quad (4.25)$$

donde  $\bar{x}$  y  $\bar{y}$  representan las medias de las secuencias numéricas  $x$  e  $y$ ,  $s_x^2$  y  $s_y^2$  las varianzas de dichas series. Este test devuelve la probabilidad  $p$  de que bajo la hipótesis nula se observe un valor más extremo al del estadístico.

### **Mann-Whitney U Test o Wilconson Rank Sum Test**

Se trata de la versión no paramétrica del *Student T-Test*. La hipótesis nula considera que ambas series numéricas presentan la misma mediana.

Considerando dos series numéricas,  $x$  e  $y$ , de longitudes conocidas  $n_1$  y  $n_2$  respectivamente. El estadístico calculado por el *U-Test* responde a:

$$\begin{aligned} U_1 &= n_1 n_2 + \frac{n_1(n_1 + 1)}{2} + R_1 \\ U_2 &= n_1 n_2 + \frac{n_2(n_2 + 1)}{2} + R_2 \\ U &= \text{mín}\{U_1, U_2\} \end{aligned} \quad (4.26)$$

donde  $R_1$  y  $R_2$  son las sumas de los rangos de las series  $x$  e  $y$  respectivamente.

### **4.7.2 Test de Estabilidad**

Para calcular la estabilidad o robustez de la medida en algunos casos se ha considerado el coeficiente de correlación cruzada (CC), el cual mide el grado de similitud entre dos series numéricas. Si este coeficiente es superior a 0.8, se considera que existe robustez y validez clínica en los resultados [105]. Se aplica a su vez un test de hipótesis nula donde se considera que las series numéricas no presentan correlación alguna, de forma que se considera que la correlación es significativa cuando  $p < 0,05$ .

***Coefficiente de correlación cruzada***

El coeficiente de correlación cruzada se define a partir de las matrices de covarianzas de dos secuencias  $x$  e  $y$  de longitud  $n$ , de acuerdo a:

$$R = \frac{C_{xy}}{\sqrt{C_x C_y}} \quad (4.27)$$

donde  $C_{xy}$ ,  $C_x$  y  $C_y$  representan las matrices de covarianza y autocovarianzas de las series  $x$  e  $y$  respectivamente, calculadas de acuerdo a:

$$\begin{aligned} C_{xy} &= \overline{(x - \bar{x}) * (y - \bar{y})} \\ &= \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \end{aligned} \quad (4.28)$$

con  $\bar{x}$  y  $\bar{y}$  siendo las medias de  $x$  e  $y$  respectivamente.



# Capítulo 5

## Conjunto experimental

A continuación se describen las distintas base de datos que contienen los registros utilizados para caracterizar las medidas presentadas en el capítulo 4. Se trata generalmente de señales biológicas con al menos dos grupos diferenciados.

Al final se muestra la tabla 5.1 que contiene un resumen de las bases de datos consideradas en esta tesis, con el número de señales en cada grupo y la longitud de los registros.

### 5.1 Base de datos de electroencefalogramas

El electroencefalograma (EEG) es el registro superficial de la actividad eléctrica del cerebro. Está compuesto por la actividad conjunta de millones de neuronas corticales hasta una profundidad de apenas algunos milímetros. El EEG se compone de distintos ritmos que dependen, entre otras cosas, del estado mental del sujeto como puede ser el estado de atención o dormir. Estos ritmos se caracterizan por su rango frecuencial y su amplitud relativa [75].

La amplitud del EEG está relacionada con el grado de sincronía en la interacción de la distintas neuronas corticales. El EEG generalmente presenta una amplitud desde unos pocos  $\mu\text{V}$  hasta  $100 \mu\text{V}$  y un contenido frecuencial comprendido desde  $0.5 \text{ Hz}$  hasta  $30\text{--}40 \text{ Hz}$  comprendido en 5 rangos frecuenciales [75].

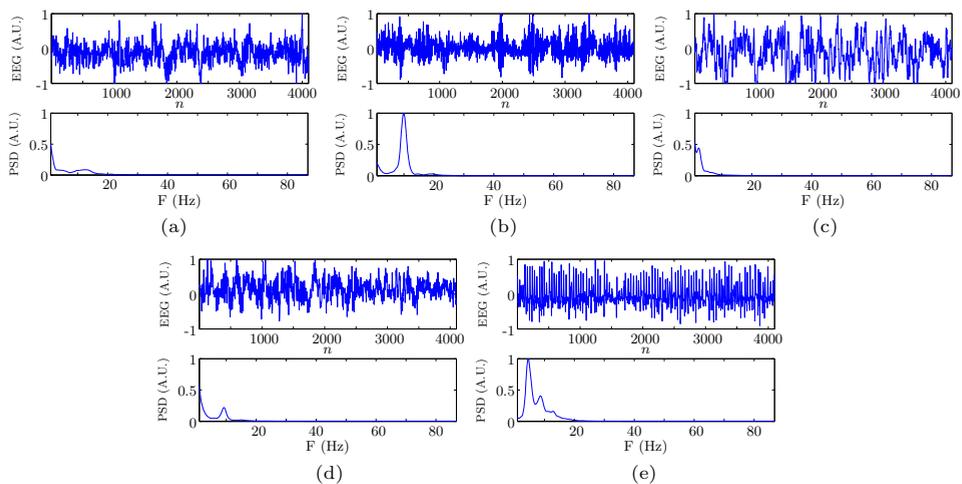
Los principales rangos frecuenciales del EEG son los siguientes [75]:

- **Delta:** localizado por debajo de  $4 \text{ Hz}$ , generalmente aparece durante el sueño profundo y tiene gran amplitud. No se suele observar en adultos despiertos y es indicativo de encefalopatía.

- **Theta:** se localiza entre 4 y 7 Hz, aparece durante estados de mareo y en ciertas etapas de sueño.
- **Alpha:** comprende el rango frecuencial entre 8 y 13 Hz. Es la componente frecuencial dominante en sujetos normales en estado de relax, despiertos y con ojos cerrados. Esta actividad desaparece cuando se abren los ojos. Su amplitud es máxima en las regiones occipitales.
- **Beta:** localizado en 14 y 30 Hz. Se tratan de ondas con variaciones frecuenciales rápidas y baja amplitud. Está asociado a una activación cortical que se observa durante ciertas etapas del sueño. Generalmente aparece en las regiones frontales y centrales del cortex cerebral.
- **Gamma:** ondas que comprenden el rango frecuencial superior a 30 Hz. Está relacionado con el procesado activo de información en el cortex.

Esta base de datos consta de señales electroencefalográficas (EEG), es de libre distribución y se puede descargar en <http://epileptologie-bonn.de>.

Esta base de datos está compuesta por 5 grupos o sets de señales. Cada uno contiene 100 señales de 23.6 s de duración adquiridas con una frecuencia de muestreo  $F_s = 173,61$  Hz (4096 muestras) [106].



**Figura 5.1:** Representación temporal (EEG) y espectral (PSD) para una señal de cada set. Sets A a E respectivamente.

Cada señal ha sido seleccionada y extraída de un registro EEG multicanal continuo sometido a inspección visual para eliminar artefactos como oculogramas o miogramas. Estos segmentos además han tenido que cumplir requisitos de estacionariedad descritos en [106].

Los set A y B proceden de registros EEG continuos de 5 sujetos sanos (control) mientras que los sets C, D y E proceden de registros EEG continuos de 5 sujetos epilépticos (patológicos) que han conseguido controlar sus ataques epilépticos tras la resección completa del hipocampo diagnosticado como zona epileptogénica. Los sets corresponden a:

- **Set A:** Control, despierto, en estado de relax y con los ojos abiertos.
- **Set B:** Control, despierto, en estado de relax y con los ojos cerrados.
- **Set C:** Patológico, EEG entre ataques epilépticos registrados en el hipocampo opuesto a la zona epileptogénica.
- **Set D:** Patológico, EEG de intervalos entre ataques epilépticos procedentes de la zona epileptogénica.
- **Set E:** Patológicos, contienen señal correspondiente a ataques epilépticos.

Las señales han sido prefiltradas con un filtro paso banda de frecuencias de corte  $[0.53, 40]$  Hz y 12 dB/oct. de caída. Se puede obtener más información de esta base de datos en [106] o en la url mencionada al inicio de la sección.

Para este trabajo se han inspeccionado visualmente todas las señales, en las que se ha observado que muchas de ellas presentan contenido espectral relevante por encima de 40 Hz. Se ha utilizado un filtro Butterworth paso bajo de orden 3 y frecuencia de corte de 40 Hz para eliminar este contenido espectral usando un algoritmo *Backward-Forward-IIR-filtering* para no introducir distorsión de fase, de forma que se ajustaran a la descripción del artículo [106].

La figura 5.1 muestra una señal tipo para cada set de esta base de datos tanto en su dominio temporal como espectral, en unidades normalizadas. Su densidad espectral de potencia (PSD) se ha obtenido mediante un periodograma de Welch con ventanas Hamming de 1 segundo de duración y un solapamiento de medio segundo.

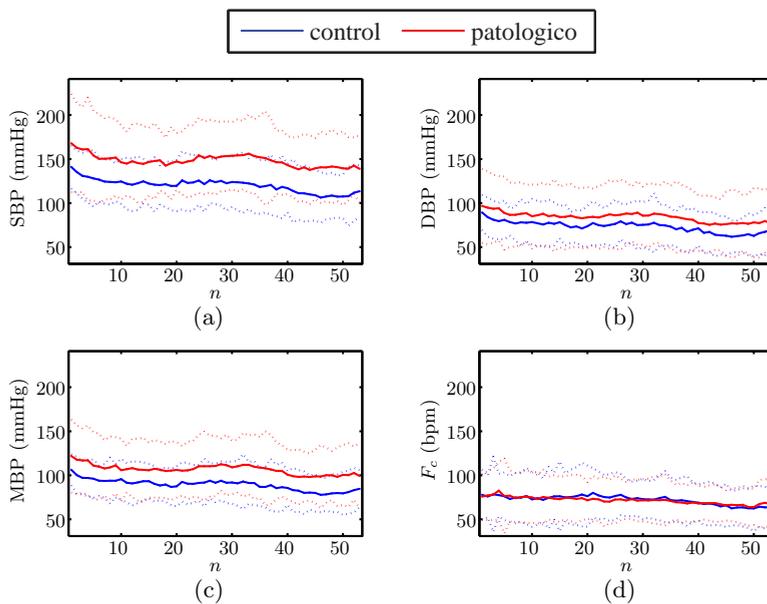
Para los posteriores análisis, estos sets se han reagrupado en controles (A y B) y patológicos (C, D y E).

## 5.2 Base de datos de registros de hipertensión arterial

Esta base de datos contiene registros de monitorización ambulatoria de presión arterial de 24h de duración, proporcionados por la unidad de hipertensión del Hospital Universitario de Móstoles (Madrid). Los registros han sido realizados mediante el uso del SpaceLabs 90207 [107]. Este dispositivo oscilométrico es no invasivo y automático, programado para obtener una medida de presión arterial

cada 20 minutos durante períodos diurnos y cada 30 minutos durante períodos nocturnas.

Los registros fueron realizados en días laborables. A los pacientes se les comunicó que siguieran con sus actividades normales, pero debían mantener el brazo extendido e inmóvil mientras se tomaba la muestra, es decir, mientras el hinchado del manguito tuviera lugar. Para poder ser clasificados como válidos, los registros deben presentar más del 80 % de medidas correctas en ambos períodos, diurno y nocturno, además de una medida de presión arterial cada hora. Los períodos diurno y nocturno fueron definidos de forma individual para cada paciente de acuerdo a sus hábitos.



**Figura 5.2:** Variables fisiológicas contenidas en la base de datos de hipertensión arterial (HTA). (a) Presión arterial sistólica (SBP), (b) presión arterial diastólica (DBP), (c) presión arterial media (MBP) y (d) frecuencia cardíaca ( $F_C$ ). En rojo se presentan los sujetos con hipertensión severa (patológicos) y en azul los sujetos de control. La línea continua presenta la señal media de cada variable y las discontinuas representan los intervalos de confianza al 95 %.

Esta base de datos contiene 61 registros, de los cuales 31 corresponden a pacientes ya diagnosticados de hipertensión severa, y tratados con tres o más fármacos que además son atendidos de forma regular en la unidad de hipertensión del hospital. Los otros 30 registros corresponden a pacientes con sospecha de hipertensión o con hipertensión arterial de reciente diagnóstico, así que para poder confirmar

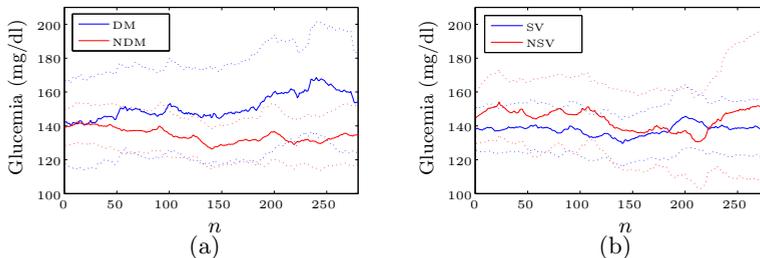
dicho diagnóstico no han sido tratados y son utilizados como grupo de control o referencia.

Cada registro contiene medidas de presión arterial sistólica, diastólica, media y frecuencia cardíaca. De acuerdo al protocolo de adquisición de las medidas, estos registros no están uniformemente muestreados y algunas medidas obtenidas no son válidas. Las medidas incorrectas han sido eliminadas, de forma que las señales presentan longitudes comprendidas entre un mínimo de 52 muestras y un máximo de 63.

## 5.3 Bases de datos de registros de glucemia

### 5.3.1 Registros de glucemia en pacientes críticos

Esta base de datos está compuesta por 42 pacientes en estado crítico e ingresados en la unidad de cuidados intensivos (UCI) del Hospital de Móstoles en Madrid entre noviembre 2007 y enero 2009, de los que finalmente se seleccionaron 38 [54]. El grupo de pacientes comprende sujetos mayores de edad y con una puntuación APACHE II [108] superior a 14 y con estancia en la UCI de al menos 72h. Los pacientes pueden ser clasificados como diabéticos y no diabéticos, o supervivientes y no supervivientes. Se consideran supervivientes aquellos que fueron dados de alta de la UCI.



**Figura 5.3:** Niveles de glucemia medios para (a) diabéticos (DM) en azul y no diabéticos (NDM) en rojo y (b) supervivientes (SV) en azul y no supervivientes (NSV) en rojo. La línea sólida presenta el valor medio y las discontinuas los intervalos de confianza al 95 %.

El nivel de glucosa en sangre ha sido obtenido de forma subcutánea a partir del fluido intersticial, mediante el uso de un sensor colocado en el tejido subcutáneo abdominal. Las medidas han sido tomadas cada 5 minutos durante al menos 48h mediante un sistema de monitorización continua (CGMS System Gold, Medtronic MiniMed). Este sistema era diariamente calibrado mediante 4 medidas superficiales obtenidas a partir de las yemas de los dedos. También se disponen de datos como la

dosificación de insulina administrada, el uso de medicación vasoactiva, excursiones, desviaciones estándar o MAGE [109] para cada paciente.

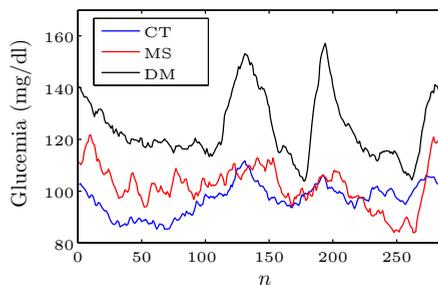
Las series finales comprenden 24h de registro continuo, cuya primera muestra comienza a las 8 a.m. del día posterior a su ingreso en la UCI. La figura 5.3 muestra los niveles de glucosa en sangre medios de las señales de la base de datos en función de la clasificación definida.

El uso de esta base de datos ha sido aprobada por el comite institucional de ética del hospital de Móstoles y todos los sujetos o sus tutores legales han dado su consentimiento para el estudio. Para más información sobre esta base de datos se puede consultar el artículo original [54].

### 5.3.2 Registros de glucemia en pacientes ambulatorios

Esta base de datos consta de 32 registros divididos en tres grupos distintos. El primer grupo consta de 17 señales y comprende sujetos de control (NDM). El segundo contiene 9 señales de sujetos con síndrome metabólico (MS) y el tercero consta de 6 señales de sujetos diabéticos (DM) de reciente comienzo, esto es, diagnosticados en un período inferior al año y sin tratamiento insulínico. Todos ellos pacientes del hospital universitario de Móstoles en Madrid.

Para determinar los sujetos con síndrome metabólico se ha considerado el criterio definido por el programa de educación nacional sobre colesterol y tratamiento adulto (NCEP-ATPII) definido en [110]. Los sujetos con diabetes mellitus (tipo II) se han definido a partir de dos medidas consecutivas donde el nivel de glucosa en plasma es superior a 126 mg/dl.



**Figura 5.4:** Niveles de glucemia medios para los sujetos de control (CTRL) en azul. En rojo los sujetos con síndrome metabólico (MS). En negro sujetos con diabetes tipo II (DM).

Para evitar factores externos como el sexo y posibles cambios hormonales, solo hombres han sido incluidos en el estudio. Las muestras se han obtenido mediante el mismo protocolo descrito en la sección 5.3.1 a partir de un sensor abdominal

subcutáneo que registraba los niveles de glucosa en el líquido intersticial cada 5 minutos durante al menos 48 horas. Los sujetos han desarrollado su vida normal durante todo el registro y debían hacerse 3 medidas de glucemia en sangre a partir de un pinchazo en la yema de los dedos para calibrar las medidas reales.

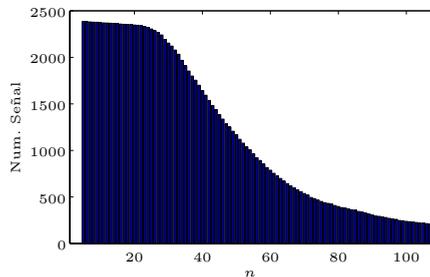
De la serie de 48 horas se ha extraído una serie de 24h para cada paciente (288 muestras) donde la primera muestra corresponde a la de las 6 a.m. del segundo día de registro. También se han obtenido medidas convencionales de glucemia. La figura 5.4, muestra como ejemplo el perfil glicémico medio para cada tipo de paciente, sujeto sano, con MS o con diabetes de tipo II (DM).

El uso de esta base de datos ha sido aprobado por el comité institucional de ética del hospital de Móstoles y todos los sujetos han dado su consentimiento para el estudio. Para más información, se puede consultar el artículo original [55].

### 5.3.3 Registros de glucemia de pacientes post-operatorios

Esta base de datos se ha obtenido en un hospital de Praga, República Checa, en pacientes sometidos a cirugía cardíaca.

Los sujetos de estudio presentan una edad entre 18 y 90 años, y son hemodinámicamente estables. Se esperaba que tuvieran un tratamiento postoperatorio en UCI durante al menos 2 días consecutivos. Los criterios de exclusión eran: alergia a la insulina, incapacidad mental, barreras lingüísticas y/o negativa del paciente.



**Figura 5.5:** Histograma del número de señales en función de las longitudes.

Los pacientes fueron asignados de forma aleatoria a dos grupos, uno con control pre y postoperatorio (PRE) y un segundo con control únicamente postoperatorio (POST) del nivel de glucosa en sangre. La clasificación fue llevada a cabo por el anestesista. El control sanguíneo del nivel de glucosa fue llevado a cabo mediante la administración de insulina por vía intravenosa.

La administración de insulina en el grupo PRE comenzó al inicio de la intervención cardíaca cuando el nivel de glucosa en sangre excedía los 6.1 mmol/l. En el grupo

POST ésta empieza en el momento de ingreso en la UCI tras la intervención quirúrgica o durante la misma si el nivel de glucosa excedía los 10 mmol/l. En ambos grupos el control se realiza hasta el final de su estancia en la UCI o cuando el paciente empezó a comer de forma normal.

El control del nivel de glucemia se realizó por parte de las enfermeras de la UCI de forma rutinaria. Los niveles de glucosa en sangre se obtuvieron mediante catéteres arteriales cuando fuera posible o mediante catéteres en venas centrales. Las medidas se obtuvieron mediante análisis espectrométricos de la sangre (86,9%; ABL 700, Radiometer Medical, Copenague, Dinamarca) o glucómetros (86,9%; ABL 700, Radiometer Medical, Copenague, Dinamarca).

La base de datos consta de 2383 registros, 1780 correspondientes a sujetos control (no diabéticos) y 604 a sujetos patológicos (diabéticos). Por otro lado, de estos 2383 registros, 1249 presentan control postoperatorio, mientras que 1134 presentan control pre y postoperatorio.

Las muestras fueron tomadas de forma manual, y no uniforme. Las señales comprenden longitudes entre 5 y 108 muestras. Para ilustrar esta variabilidad en la duración, la figura 5.5 muestra un diagrama de barras con la representación del número de señales disponibles para cada una de las posibles longitudes.

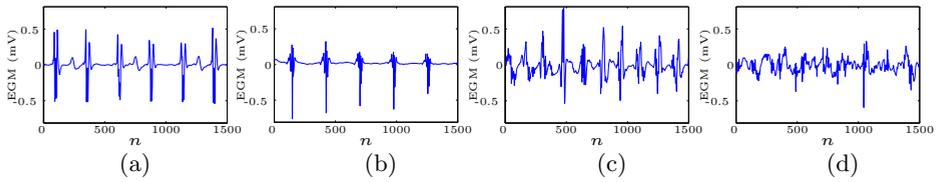
La mayor diferencia radica en que a los sujetos se les ha hecho un control del nivel de glucosa en sangre preoperatorio o solo postoperatorio para buscar así que técnica es la más conveniente a la hora de garantizar el mínimo número de complicaciones durante el postoperatorio.

## 5.4 Base de datos de electrogramas de fibrilación auricular

Los electrogramas (EGM) auriculares son señales que miden la actividad eléctrica de un conjunto de células del tejido cardíaco en la aurícula. Esta base de datos comprende señales EGM registrados durante episodios de fibrilación auricular (AF). Se han considerado cuatro clases o niveles de fraccionamiento de los EGM:

- **C0:** Electrogramas no fraccionados.
- **C1:** Electrogramas fraccionados con activaciones periódicas.
- **C2:** Electrogramas fraccionados con activaciones tanto periódicas como no periódicas.
- **C3:** Electrogramas de alta frecuencia con actividad continuada. No presentan activaciones de forma regular.

Estos cuatro niveles a su vez se reagrupan en dos grupos principales, esto es debido a que en la práctica los tipos C0 y C1 no se consideran responsables de la generación



**Figura 5.6:** Señal ejemplo de un electrograma para cada grado de fraccionamiento (a) C0, (b) C1, (c) C2 y (d) C3.

y mantenimiento de la fibrilación auricular y por tanto no se considera necesario aplicar una ablación del tejido en esas regiones [111, 112]. A su vez, las clases C2 y C3 se consideran responsables de la generación y mantenimiento de AF y por tanto las zonas donde se registran estos EGM deberían ser ablacionadas [111, 112]. A partir de este punto se hará referencia a estas dos clases como EGM recomendados para ablación (RA) o no recomendados para ablación (NRA).

Los 113 registros analizados han sido preseleccionados por expertos clínicos procedentes de una base de datos mayor. Estas series temporales han sido registradas durante el proceso de mapeo de AF en pacientes que han sido indicados para ablación por radiofrecuencia.

Las señales han sido registradas mediante el dispositivo CardioLab7000, Prucka Inc. a una frecuencia de muestreo de 997 Hz y remuestreadas a 1 KHz durante 1.5 s. Cada señal ha sido seleccionada de forma que se garantice un buen ratio señal a ruido y clasificada en uno de los cuatro niveles de fraccionamiento iniciales mencionados, por un experto electrofisiólogo.

Las 113 señales de la base de datos han quedado clasificadas en los cuatro grupos anteriores de la siguiente forma: 22 señales de nivel C0, 42 señales de nivel C1, 36 señales de nivel C2 y 13 señales de nivel C3. Un ejemplo gráfico de una señal de cada nivel de fraccionamiento se muestra en la figura 5.6.

## 5.5 Base de datos de registros RR

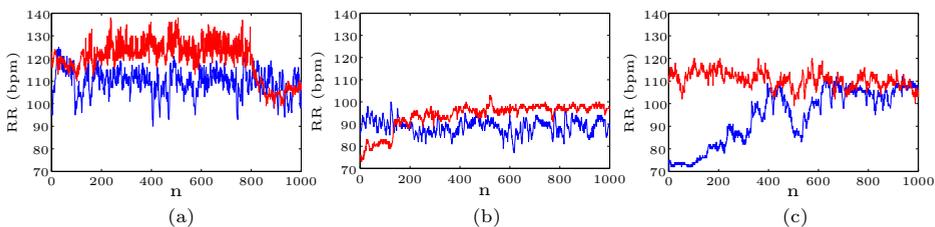
Esta base de datos se encuentra ampliamente descrita en [113] y es de libre distribución, pudiendo acceder a ella a través de [www.physionet.org/physiobank](http://www.physionet.org/physiobank) [114]. Forma parte del estudio para la supresión de las arritmias cardíacas (*Cardiac Arrhythmia Suppression Trial, CAST*). Este estudio trata de valorar la eficacia en la supresión de arritmias por parte de tres medicaciones distintas : encainida, flecainida y moricicina.

El estudio requería que los sujetos hubieran sufrido un infarto agudo de miocardio en los dos años anteriores y al menos presentasen 6 latidos ventriculares ectópicos

por hora en los registros holter pre-tratamiento. A los sujetos seleccionados se les aplicó de forma aleatoria una medicación de las tres disponibles.

De los 3549 pacientes iniciales solo 769 se han considerado en la base de datos en función de su potencialidad para obtener resultados viables en la evaluación y caracterización de los posibles cambios de la variabilidad del ritmo cardíaco tras el tratamiento.

De los 769 sujetos finales, 286 han sido tratados con encainida, 229 con flecainida y 294 con moricicina. La figura 5.7 muestra un segmento de señal de un sujeto pre y post-tratamiento con cada tipo de medicación.



**Figura 5.7:** Registros RR de un sujeto de cada tipo de medicación antes del tratamiento (Pre) en azul y durante el tratamiento (Post) con (a) encainida, (b) flecainida y (c) moricicina en rojo .

## 5.6 Base de datos de dinámicas de movilidad en pacientes con enfermedades neurodegenerativas

Las enfermedades neurodegenerativas generalmente afectan a la movilidad del paciente. Para entender mejor la fisiopatología de dichas enfermedades y mejorar la habilidad para medir las respuestas a posibles intervenciones terapéuticas, resulta de ayuda la cuantificación y caracterización de las dinámicas de movilidad.

Esta base de datos contiene registros de 15 pacientes con la enfermedad de Párkinson, 20 con la enfermedad de Huntington y 13 con esclerosis lateral amiotrófica, además de 16 sujetos sanos.

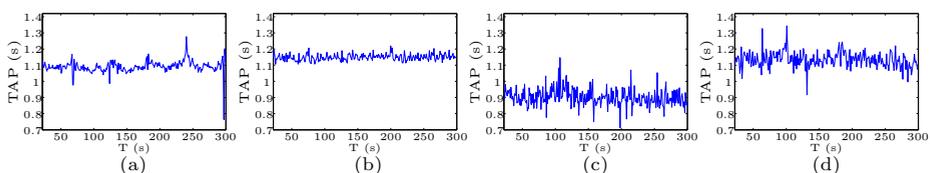
Las señales han sido registradas con sensores resistivos de presión, con los que se obtiene una señal proporcional a la fuerza con la que el sujeto pisa al caminar. A partir de estas señales se han derivado registros que miden el tiempo entre pasos.

Esta base de datos es de libre distribución y se puede obtener de forma gratuita en <http://www.physionet.org/physiobank/database/gaitnidd/>. Contiene distintos tipos de variables. De entre ellas se ha seleccionado la duración porcentual de un paso completo con la pierna derecha (*right stance interval*), con respecto



**Figura 5.8:** Definición de los términos (a) *Stance* y (b) *Stride*.

al tiempo entre dos pisadas sucesivas del pie derecho (*stride*). En la figura 5.8 se muestra a qué se refiere cada parte de la variable y qué cuantifica. La figura 5.9 muestra un ejemplo de una señal para cada enfermedad neurodegenerativa frente a un caso de control para la duración de un apoyo de pie completo.



**Figura 5.9:** Tiempo de apoyo del pie derecho (TAP). Se muestra una señal ejemplo de cada grupo (a) control, (b) esclerosis lateral amiotrófica, (c) Huntington y (d) Párkinson.

**Tabla 5.1:** Resumen de las bases de datos consideradas. Se muestra para cada base de datos, el grupo considerado de control y los patológicos, el número de registros en cada grupo (NR) y la longitud de éstos ( $N$ ).

Base de datos	grupos	clasificación	NR	$N$
Electroencefalogramas	control ojos abiertos	control	100	4096
	control ojos cerrados	control	100	4096
	epiléptico interictal	patológico	100	4096
	epiléptico ictal	patológico	100	4096
	epiléptico area opuesta	patológico	100	4096
Hipertensión	no diagnosticados	control	30	52 – 63
	diagnosticados	patológicos	31	52 – 63
Glucemia (crítica)	diabéticos	patológico	11	288
	no diabéticos	control	27	288
	superivivientes	control	28	288
	no supervivientes	patológico	10	288
Glucemia (ambulatoria)	no diabéticos	control	17	288
	síndrome metabólico	patológico	9	288
	diabéticos	patológico	6	288
Glucemia (postoperatoria)	no diabéticos	control	1780	5 – 108
	control preoperatorio	control	1134	5 – 108
	diabéticos	patológico	604	5 – 108
	control postoperatorio	patológico	1249	5 – 108
Electrogramas	nivel C0	control	22	1500
	nivel C1	control	42	1500
	nivel C2	patológico	36	1500
	nivel C3	patológico	13	1500
Registros RR	no tratados	control	602	10000
	tratados con Encainida	patológico	203	10000
	tratados con Flecainida	patológico	206	10000
	tratados con Moricicina	patológico	209	10000
Dinámicas de Movilidad	control	control	16	198
	Huntington	patológico	20	167
	Parkinson	patológico	15	203
	Esclerosis	patológico	13	122

# Capítulo 6

## Experimentos

### 6.1 Pérdida de muestras

Se considera una señal base  $u(n)$  de longitud  $N$  muestras a la cual se le aplica un algoritmo para eliminar muestras en un porcentaje proporcional a su longitud.

Se siguen dos algoritmos de eliminación de muestras:

- **Aleatorio:** este algoritmo para eliminar muestras de un registro sigue una distribución aleatoria. Se seleccionan un número de muestras proporcional a su longitud que se eliminan de la señal.
- **Uniforme:** este algoritmo para eliminar muestras de una señal sigue una distribución uniforme. Se establece el número de muestras a eliminar en función de un porcentaje proporcional a la longitud total de la señal. La muestra inicial a partir de la cual se eliminan las muestras se selecciona de acuerdo a una distribución aleatoria.

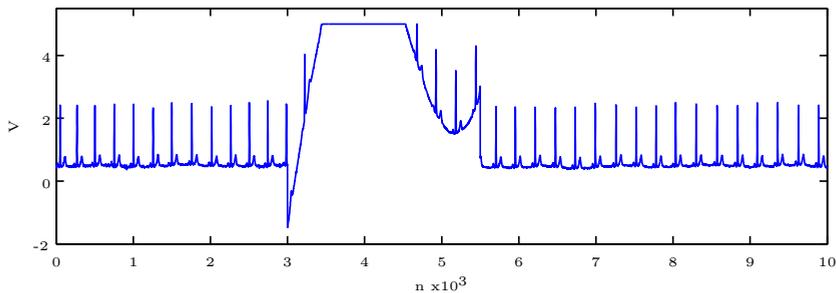
Debido a que estas simulaciones necesitan señales largas de forma que se cumplan los requisitos de las medidas de complejidad con respecto a  $m$ ,  $r$  y  $N$ , sólo se han considerado las señales que presentan al menos 1000 muestras. Estos algoritmos se han aplicado sobre las bases de datos de ECG ( $N = 4096$ ), EEG ( $N > 1500$ ), RR ( $N = 1000$ ).

Para las bases de datos de EEG y EGM al tener un número limitado de registros se han realizado 100 repeticiones del experimento donde la aleatoriedad reside en que la muestra que se elimina en cada repetición es distinta de la anterior. Para la base de datos RR puesto que se dispone de aproximadamente 28000 segmentos de

señal en cada grupo, se han seleccionado de forma aleatoria 500 pares para cada medicación.

En todas las medidas se han aplicado parámetros clásicos considerados como adecuados y propuestos por Pincus en [14],  $(m,r) = (2, 0.2)$ . Los ratios de eliminación se han considerado a intervalos de un 10 % hasta alcanzar el 90 % de muestras eliminadas. Las medidas de complejidad QSE y CosEn no se han considerado en este experimento puesto que están optimizadas para señales de corta duración  $N < 100$ , debido en gran parte, a su alto coste computacional.

Este experimento intenta simular condiciones reales de las condiciones de registro y transmisión de las señales biomédicas en la época actual. En primer lugar, los dispositivos de medida y adquisición de registros biomédicos pueden estar mal calibrados o debido a las condiciones del paciente durante el registro de la señal, como por ejemplo movimientos torácicos durante el registro de un ECG, pueden provocar saturaciones del dispositivo y obtener una señal como la mostrada en la figura 6.1. En esta figura se ha intentado simular el efecto de una saturación de un electrocardiógrafo común que presenta un nivel de saturación de 5v como puede ser el EDAN SE-1010. En este caso una vez limpia esta señal, sin línea de base, sin interferencia de 50 Hz y de ruido, se tendría un lapso temporal entre los instantes 3000 a 5500 en los cuales la señal experimenta una alteración de su línea de base provocando una saturación en el registro que resulta irrecuperable.

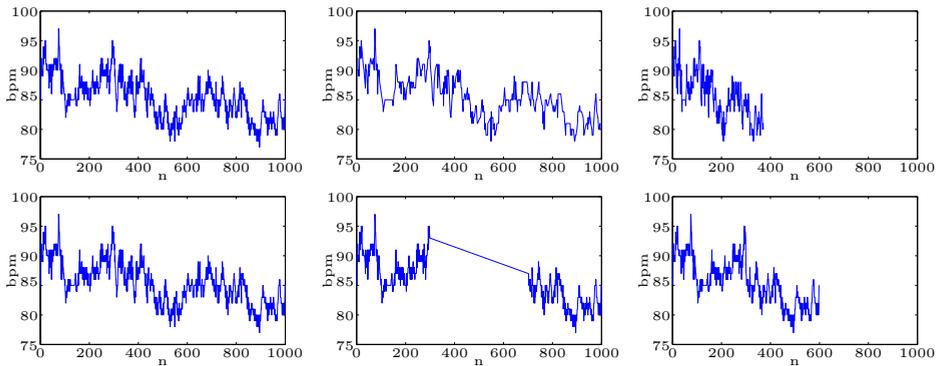


**Figura 6.1:** Señal ECG real ampliada y saturada de forma artificial para simular el efecto de una saturación en un registro

En segundo lugar, la telemedicina está cada vez más extendida y muchas de sus transmisiones requieren transmisión sin cable, bien sea por WiFi, Bluetooth, UMSST o GSM entre otros. Estas transmisiones pueden sufrir interrupciones en la conexión, pérdida de paquetes, exceso de ruido o interferencias, lo que puede provocar a su vez que se pierdan bien paquetes de forma aleatoria o bien una secuencia de muestras continua [78, 80, 105].

También algunas técnicas de procesado como puede ser detección de eventos, compresión de datos o especificaciones hardware pueden causar la pérdida de información que puede resultar relevante [77, 115].

La figura 6.2 muestra el proceso aplicado para eliminar las muestras tanto de forma aleatoria como uniforme a partir de una señal RR.



**Figura 6.2:** Se muestra en la fila superior el proceso aplicado para la eliminación de muestras de forma aleatoria de un 60% y en la inferior para el algoritmo de pérdidas de muestras uniforme con un 40% de pérdidas. La primera figura representa una señal RR sobre la cual se eliminan muestras, la segunda columna representa la señal con las muestras eliminadas y en la última vemos finalmente la señal a analizar, donde al eliminar las muestras ésta ha visto su longitud disminuída.

## 6.2 Señales de corta duración

### 6.2.1 Parametrización

En este experimento se ha considerado estudiar las variaciones de las medidas de complejidad en función de los posibles valores que pueden tomar sus parámetros, en señales de corta duración ( $N < 1000$ ).

En el caso de ApEn y SampEn, se ha considerado la siguiente parametrización:

- $m = \{1, 2, \dots, 10\}$
- $r = \{0.1:0.1:1\}$

Para MSE, puesto que se añade el parámetro de la escala se ha considerado hasta la décima escala en el caso de que sea factible, o escalas inferiores en su defecto. Finalmente en DFA no existe una parametrización tal cual, los parámetros de los

cuales depende DFA son la forma de aplicar el inventariado (disjunto o solapado), la forma de definir las (longitudes y número de escalas temporales) y el orden del polinomio utilizado para eliminar la tendencia. La longitud inicial de la secuencia de inventariado viene definida por el algoritmo dejando a elección la longitud de la ventana de máxima duración. Ésta se ha establecido en  $N/10$  siempre y cuando hubiera al menos 20 ventanas en la secuencia o en  $N/4$  o  $N$  de igual forma, siendo  $N$  la longitud de la señal a analizar. Se ha considerado el utilizar ventanas deslizantes y ver su influencia en el resultado final. El grado mínimo del polinomio se puede considerar también como predefinido atendiendo a [88] pero en este caso se ha decidido estudiar los tres primeros órdenes ya que no se recomienda usar un orden mayor a 3. Finalmente en DFA queda estudiar la caracterización del punto de *cross-over* y su posible influencia en el cómputo de los coeficientes.

Este experimento se ha realizado sobre las siguientes bases de datos:

- **Glucemia en pacientes críticos:** ApEn, SampEn, MSE y DFA
- **Glucemia ambulatoria:** ApEn, SampEn, MSE y DFA
- **Dinámicas de movilidad en pacientes con enfermedades neurodegenerativas:** ApEn, SampEn, MSE y DFA

En el caso de MSE se ha considerado siempre y cuando sea posible hasta la escala 10, no siendo este el caso de las series de glucemia en pacientes críticos ni ambulatorios puesto que solo disponen de 288 muestras. Esto implica que únicamente pueda analizarse hasta escala 3 (96 muestras/serie).

### 6.2.2 Entropías específicas

Para el análisis de señales de muy corta duración ( $N < 100$ ) se ha utilizado particularmente QSE y CosEn. Estas métricas están diseñadas para su uso sobre señales con pocas muestras y además su algoritmo tiene un coste computacional alto lo que no hace factible su uso sobre señales más largas. Con QSE y CosEn se han analizado las secuencias de hipertensión arterial y las señales de glucemia con control pre y post-operatorio.

## 6.3 Evaluación

En todas las simulaciones existe la posibilidad que según sea el valor de los parámetros del método seleccionado, no se obtenga un resultado válido de entropía.

Por ejemplo, en ApEn existe la posibilidad de obtener un valor 0. Ésto no quiere decir que su entropía sea nula, sino todo lo contrario, indica que no se han encon-

trado otros *matches* en ese registro aparte de la comparación con el patrón inicial, por ello los logaritmos son 0 al deberse a un *auto-match* [10].

Por otro lado, existe la posibilidad de que el número de señales en las cuales se pueda considerar una estimación fiable de la entropía sea muy limitado. Todos aquellos casos en los que el número de señales con estimación robusta de entropía sea inferior a 5 no serán considerados en la evaluación, ya que no se puede considerar que dicho grupo sea representativo de la población, ni apelar a la teoría de los grandes números .



# Capítulo 7

## Resultados

En este capítulo se mostrarán los resultados obtenidos en los distintos experimentos en función de cada medida descrita en la sección 4 y obtenidos sobre cada base de datos. Se muestran únicamente los resultados más significativos:

- Electroencefalogramas (EEG) de sujetos control (CT) y patológicos (PT).
- Registros RR de pacientes tratados con encainida antes (PRE) o después (POST) del tratamiento.
- Electrogramas (EGM) registrados en regiones recomendadas (RA) o no recomendadas para ablación (NRA).
- Dinámicas de movilidad con el pie derecho en pacientes en sujetos control (CT) y con patologías neurodegenerativas, esclerosis lateral amiotrófica (LSA), Parkinson (PD) o Huntinton (HD).
- Series de glucemia en pacientes críticos, sujetos diabéticos (DM) o no diabéticos (NDM).
- Series de glucemia ambulatoria en pacientes diabéticos (DM), no diabéticos (NDM) o con síndrome metabólico (MS).
- Series de hipertensión arterial (HTA) en pacientes de control (CT) y patológicos (PT).
- Series de glucemia en pacientes que han desarrollado complicaciones post-operativas en sujetos con tratamiento insulínico pre y post-operatorio (PRE) o sólo post-operatorio (POST).

## 7.1 Entropía aproximada (ApEn)

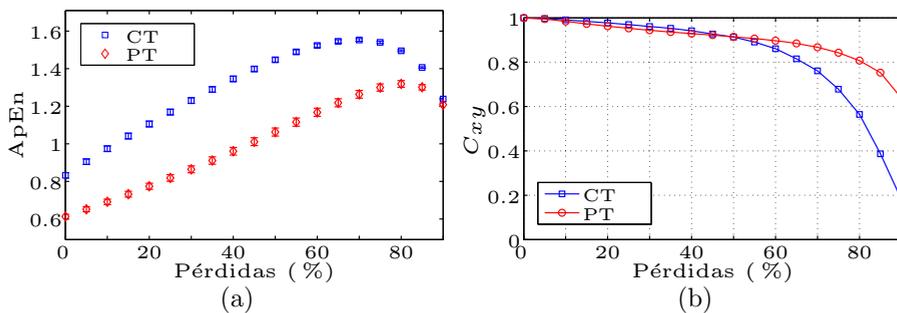
A continuación se muestran los resultados obtenidos para ApEn en los distintos experimentos realizados. Pérdida aleatoria y uniforme de muestras y parametrización en señales de corta duración.

### 7.1.1 Pérdida aleatoria de muestras

En esta sección se muestran los resultados relativos al experimento descrito en la sección 6.1 para ApEn sobre las series de electroencefalogramas (EEG), registros RR y electrogramas (EGM).

#### *Electroencefalogramas*

La figura 7.1.a muestra la evolución de ApEn en función del ratio de muestras perdidas y en 7.1.b el coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de las series con pérdidas y las originales para cada grupo. Los resultados mostrados en la figura 7.1.a sugieren una menor complejidad o mayor regularidad en sujetos patológicos (PT). Esto está sustentado en la literatura, donde en [14] se establece que ApEn tiene la capacidad de distinguir entre características imperceptibles y en consecuencia indicar un principio patológico.



**Figura 7.1:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con muestras eliminadas en función del ratio de pérdidas.

En la figura 7.1.a se observa cómo en primer lugar el valor de ApEn aumenta hasta alcanzar un máximo a partir del cual decrece. El efecto obtenido es en cierta forma esperado. En primer lugar al eliminar muestras de la señal de forma aleatoria, se está reduciendo su predictibilidad y por tanto haciéndola más aleatoria, de ahí el aumento en la complejidad. En segundo lugar, el punto de inflexión puede asociarse al sesgo que introduce el algoritmo de ApEn cuando la señal presenta

pocas muestras, y está asociado a los denominados *auto-matches* resultado de la comparación de patrones inherente al algoritmo de cálculo. Para ratios de eliminación del 60 % se dispone de señales de 164 muestras, algo limitado para ApEn (en [9] se recomiendan  $30^m$  muestras). Este punto de inflexión se produce antes en señales de control (CT) que en PT. Eso puede estar asociado a que la pérdida de muestras de forma aleatoria puede inducir cierta regularidad si la señal de origen es ya suficientemente aleatoria.

El valor de  $C_{xy}$  (ver figura 7.1.b) da una idea de la robustez de la medida frente al número de muestras perdidas en las señales. Shimizu *et. al* establece en [105] que dada una serie numérica, si su versión con pérdidas presenta un coeficiente de correlación cruzada superior a 0.8, ésta sigue presentado validez clínica. En este caso se observa que ApEn presentaría validez clínica en ambos grupos hasta un 60 % de pérdidas, porcentaje a partir del cual  $C_{xy} < 0,8$  para CT.

Los resultados gráficos se refuerzan con la caracterización estadística presentada en la tabla 7.1. En esta tabla se muestran los estudios de normalidad estadística de las series (JB test), los estudios de separabilidad (U Test), los intervalos de confianza al 95 % (CI) para cada grupo, definidos como  $\mu \pm 1,96\sigma$ , siendo  $\mu$  la media y  $\sigma$  la desviación estándar y  $C_{xy}$  para cada grupo.

**Tabla 7.1:** Caracterización estadística de ApEn sobre señales EEG de sujetos control (CT) o patológicos (PT) que han sufrido pérdida de muestras de forma aleatoria. Se muestra el estudio de normalidad (JB), estadísticos de segmentación (UT), intervalos de confianza (CI) al 95 % y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[0.822, 0.842]	[0.963, 0.985]	[1.219, 1.241]	[1.438, 1.456]	[1.548, 1.559]	[1.235, 1.242]
CI PT	[0.604, 0.621]	[0.681, 0.701]	[0.851, 0.877]	[1.048, 1.078]	[1.249, 1.278]	[1.201, 1.215]
$C_{xy}$ (p)	1.000(0.001)	0.991(0.001)	0.969(0.001)	0.944(0.001)	0.889(0.001)	0.508(0.001)

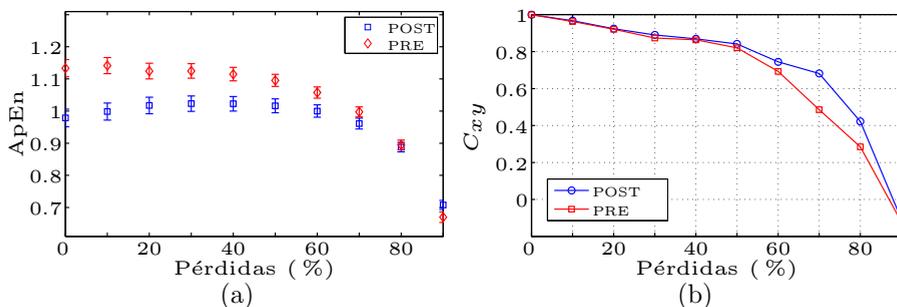
Con los resultados mostrados en la tabla 7.1 se comprueba cómo las distribuciones no son normales ( $p < 0,001$ ). La segmentación se mantiene a lo largo de todo el rango de pérdidas considerado ( $p < 0,001$ ) y los CI no se solapan en ningún caso, lo que aumenta la capacidad de segmentación y clasificación. Finalmente se observa como  $C_{xy}$  presenta validez clínica hasta alcanzar el 70 % con validez estadística ( $p < 0,001$ ).

### Señales RR

Estas señales, como se describieron en la sección 5.5 presentan un gran número de lo que a priori se considerarían “estimaciones erróneas” debido a latidos ectópicos. Generalmente, interesa eliminar o corregir esas estimaciones. En este caso no interesa su corrección puesto que el experimento inicial para el cual se registró la base de datos [113], pretendía estudiar si la medicación aplicada era adecuada o no para el tratamiento de las arritmias provocando un decremento en el número de latidos ectópicos.

En este caso los pacientes pre-tratamiento (PRE) presentan un gran número de *spikes* asociados a la presencia de latidos ectópicos ventriculares. Son más patológicos pero a su vez la señal presenta mayor aleatoriedad ya que estos no se producen con ningún tipo de cadencia. Por otro lado, los pacientes post-tratamiento (POST) han sido medicados para reducir la arritmia, con lo que se producirán menos latidos ectópicos y en consecuencia la señal será más regular.

Como se observa en la figura 7.2, PRE presenta menor regularidad que POST. Además se observa un comportamiento similar al obtenido con la base de datos EEG descrito en el punto anterior. ApEn aumenta su valor hasta un máximo a partir del cuál empieza a disminuir. Este punto se localiza en torno al 10 % para PRE y al 40 % para POST. De nuevo, el punto de inflexión ocurre en porcentajes de pérdida de muestras mayores para señales en origen más regulares. Cuando la longitud de la señal es limitada, ApEn se ve dominada por el efecto de los *auto-matches* y el decaimiento de la medida vuelve a ser exponencial (pérdidas > 70 %,  $N < 300$  muestras).



**Figura 7.2:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul tratados con encainida. (b) Coeficiente de correlación cruzada de ApEn para cada grupo en función del porcentaje de muestras perdidas.

Con el coeficiente de correlación cruzada se observa que la validez clínica ( $C_{xy} \leq 0,8$ ) se mantiene hasta alcanzar el 50 % de pérdidas en ambos grupos, permitiendo

por ejemplo, comunicaciones sobre canales bastante inestables en una transmisión GSM o WiFi.

La tabla 7.2 presenta una caracterización estadística de ApEn. En ella se muestra el test de normalidad (JB Test) donde se aprecia cómo en ningún caso las distribuciones pueden considerarse normales ( $p < 0,001$ ). Se dispone también del test de segmentación estadística. En cualquier caso existen diferencias estadísticas relevantes entre las medianas y por ello diferenciación entre grupos ( $p < 0,003$ ). Se dan los CI de ambos grupos, mostrando que no hay solapamiento hasta un porcentaje de muestras perdidas del 90 % y que existe correlación cruzada significativa en las series de regularidad en todos los casos ( $p < 0,001$ ), pero ésta solo tiene validez clínica hasta que se alcanza el 50 % de muestras perdidas.

**Tabla 7.2:** Caracterización estadística de ApEn sobre registros RR de sujetos pre-tratamiento (PRE) y post-tratamiento (POST) cuando las pérdidas se producen de forma aleatoria. Se muestra el test de normalidad (JB) y de segmentación (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

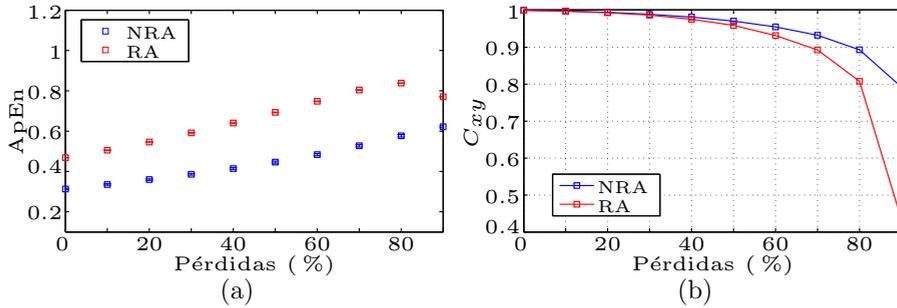
Pérdida	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.003	0.002
CI Pre	[0.951,1.006]	[0.972,1.025]	[0.999,1.047]	[0.995,1.038]	[0.944,0.978]	[0.693,0.722]
CI Post	[1.107,1.160]	[1.116,1.166]	[1.102,1.147]	[1.077,1.114]	[0.982,1.013]	[0.653,0.686]
$C_{xy}$ (p)	1.000 (0.001)	0.924 (0.001)	0.886 (0.001)	0.835 (0.001)	0.592 (0.001)	-0.138 (0.001)

### Electrogramas

A continuación se presentan los resultados obtenidos al eliminar muestras de forma aleatoria sobre señales EGM con distinto nivel de fraccionamiento. En primer lugar, se muestra la figura 7.3 donde se presenta tanto la evolución de ApEn como el coeficiente de correlación cruzada ( $C_{xy}$ ) para los valores de ApEn en función del porcentaje de pérdida de muestras.

El comportamiento que se observa en la figura 7.3 es similar al observado para las señales EEG y los registros RR. La complejidad es mayor para RA que para NRA. Esto es coherente con la literatura, ya que las señales RA mostradas en 5.6 son más irregulares y aleatorias que las señales NRA. El punto de inflexión se observa en torno al 80 % de pérdidas para RA (señal más irregular y  $N < 300$  muestras) mientras que en NRA pasa completamente desapercibido. Esto refrenda que cuanto más aleatoria es la señal, antes se aprecia el sesgo asociado a los *auto-matches*.

El  $C_{xy}$  muestra cómo para ambas clases, la validez clínica de ApEn se mantiene hasta alcanzar ratios de muestras perdidas del 70 %, lo cual indica que la regularidad de estas señales frente a una pérdida de muestras aleatoria se puede obtener de forma robusta.



**Figura 7.3:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EGM no recomendada para ablación (NRA) en azul o recomendada para ablación (RA) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) de la regularidad en función del ratio de muestras perdidas.

La tabla 7.3 presenta la caracterización estadística de ApEn en el análisis de señales EGM con pérdida de muestras aleatoria. Se muestra el test de normalidad (JB) y de separabilidad (UT), los intervalos de confianza (CI) de las dos clases principales y el  $C_{xy}$  entre los valores de entropía de las series originales con respecto a la entropía de sus versiones con pérdida de muestras. Los valores de los CI muestran cómo la segmentación es total, al no producirse solape entre las clases NRA y RA.

**Tabla 7.3:** Caracterización de ApEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se pierden de forma aleatoria. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.457	0.474	0.490	0.414	0.141	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.309, 0.316]	[0.332, 0.339]	[0.382, 0.390]	[0.441, 0.451]	[0.522, 0.533]	[0.617, 0.626]
CI RA	[0.465, 0.472]	[0.502, 0.510]	[0.587, 0.596]	[0.688, 0.698]	[0.799, 0.809]	[0.768, 0.773]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.995(0.001)	0.983(0.001)	0.957(0.001)	0.831(0.001)

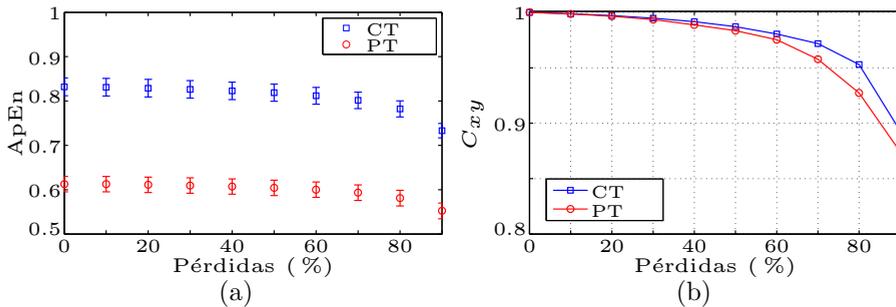
### 7.1.2 Pérdida uniforme de muestras

A continuación se incluyen los resultados relativos al experimento descrito en la sección 6.1 con diezmado uniforme para la entropía ApEn sobre las bases de datos de EEG, RR y EGM.

### Electroencefalogramas

En este punto se muestra el análisis de ApEn sobre señales EEG registradas sobre sujetos de control (CT) o sujetos patológicos (PT) donde se ha producido una pérdida de muestras de forma uniforme. En la figura 7.4 se observa el comportamiento de ApEn para sujetos CT y PT en función del ratio de pérdidas, y su coeficiente de correlación cruzada ( $C_{xy}$ ).

El comportamiento de ApEn para este tipo de pérdidas es distinto al mostrado en la sección 7.1.1 con pérdidas de tipo aleatorio. En este caso, la entropía se mantiene más o menos constante hasta alcanzar pérdidas del 60 % en CT y 70 % en PT aproximadamente, a partir de los cuales describe un comportamiento decreciente. Este comportamiento puede estar asociado al hecho de que al eliminar de forma uniforme un segmento de señal se están eliminando en similar proporción el número de coincidencias de longitud  $m$  que de longitud  $m + 1$  y por ello la entropía se ve mínimamente alterada. El punto en el cual empieza a decaer se puede asociar al punto de inflexión que se experimentaba con un algoritmo de pérdidas aleatorio, efecto que tiene el sesgo inherente al algoritmo de ApEn asociado a los *auto-matches* y a  $N$ .



**Figura 7.4:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con pérdidas en función del ratio de muestras eliminadas.

El valor de  $C_{xy}$  se mantiene por encima del umbral de validez clínica para todo el rango de pérdidas, algo que con un algoritmo aleatorio no sucedía. Esto indica que ApEn es más robusta ante unas pérdidas de forma uniforme que aleatoria.

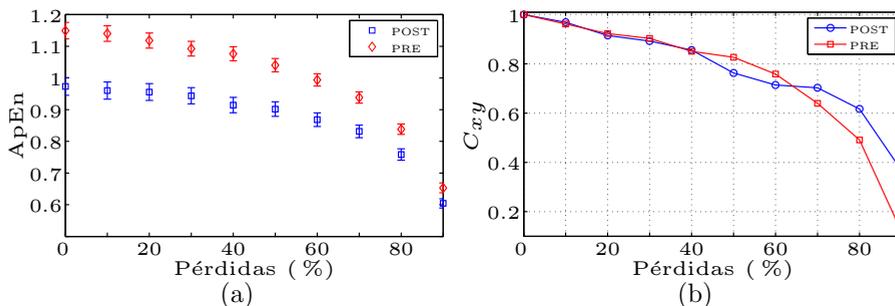
La tabla 7.4 presenta la caracterización estadística de ApEn donde se muestran los tests de normalidad (JB) y separabilidad (UT), junto con los intervalos de confianza (CI) al 95 % y  $C_{xy}$ . En ella se constata que las distribuciones no son normales ( $p < 0,001$ ), y las clases son completamente separables ( $p < 0,001$ ). Los intervalos de confianza no presentan solapamiento y  $C_{xy}$  indica que las series mantiene su validez clínica hasta un 90 % de pérdidas.

**Tabla 7.4:** Caracterización de ApEn sobre señales EEG de control (CT) y patológicas (PT). Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[0.812, 0.852]	[0.811, 0.851]	[0.807, 0.846]	[0.799, 0.838]	[0.783, 0.820]	[0.717, 0.750]
CI PT	[0.595, 0.630]	[0.595, 0.630]	[0.592, 0.627]	[0.587, 0.622]	[0.576, 0.611]	[0.535, 0.570]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.997(0.001)	0.993(0.001)	0.981(0.001)	0.933(0.001)

### Señales RR

En la figura 7.5 se muestra la evolución de ApEn para señales RR registradas en sujetos antes (PRE) y después (POST) del tratamiento de arritmia cardíaca con encainida en función del porcentaje de muestras eliminadas de la señal. Aunque la similitud no es tan evidente, se encuentran semejanzas con los resultados expuestos para EEG (sec. 7.1.2).



**Figura 7.5:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre registros RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul tratados con encainida. (b) Coeficiente de correlación cruzada de ApEn para cada grupo en función del porcentaje de muestras perdidas.

Para estas señales la zona de estabilidad se situaría antes del 10% de pérdidas para el grupo PRE y del 30% para el grupo POST. Se observa que los rangos de estabilidad de la medida se han reducido dando lugar a la aparición del efecto del sesgo de ApEn mucho antes, para longitudes de señal inferiores a 800 muestras. No obstante, también se observa cómo la robustez de la medida junto con su validez clínica se ven disminuídas, viéndose ésta comprometida a partir de un porcentaje de muestras perdidas superior al 40%.

La tabla 7.5 muestra las características estadísticas de ApEn sobre señales RR a las cuales se les ha aplicado un algoritmo de pérdida de muestras de forma uniforme. En ella se establecen distribuciones no normales ( $p < 0.001$ ), hay separabilidad

**Tabla 7.5:** Caracterización de ApEn para registros RR de pacientes antes (Pre) y después (Post) del tratamiento con encainida. Los registros han sufrido una pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.00	0.001	0.001	0.001
CI Post	[0.946,1.001]	[0.933,0.988]	[0.918,0.969]	[0.879,0.925]	[0.811,0.851]	[0.590,0.619]
CI Pre	[1.123,1.175]	[1.116,1.165]	[1.069,1.116]	[1.019,1.062]	[0.920,0.956]	[0.638,0.669]
$C_{xy}$ (p)	1 (0.001)	0.968 (0.001)	0.905 (0.001)	0.807 (0.001)	0.695 (0.001)	0.270 (0.001)

entre grupos ( $p < 0,001$ ), CI sin solapamiento y validez clínica hasta el 50 % de muestras perdidas ( $C_{xy} > 0.8$ ).

### *Electrogramas*

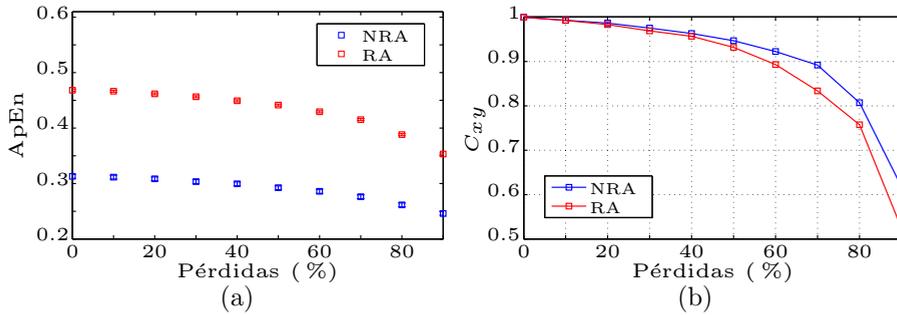
La tabla 7.6 muestra la caracterización estadística de ApEn sobre registros EGM AF. La distribuciones de ApEn se pueden considerar normales para cualquier valor de muestras perdidas ( $p > 0.05$ ), mientras que existen diferencias estadísticas en las medias de las clases ( $p < 0.05$ ).

**Tabla 7.6:** Caracterización de ApEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las pérdidas se producen de forma uniforme. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.457	0.475	0.394	0.257	0.141	0.139
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.309, 0.316]	[0.308, 0.315]	[0.300, 0.307]	[0.289, 0.296]	[0.273, 0.280]	[0.243, 0.249]
CI RA	[0.465, 0.472]	[0.463, 0.470]	[0.453, 0.460]	[0.438, 0.445]	[0.412, 0.419]	[0.349, 0.357]
$C_{xy}$ (p)	1.000(0.000)	1.000 (0.000)	0.998 (0.000)	0.987 (0.000)	0.952 (0.000)	0.789 (0.000)

La figura 7.6 muestra la evolución de la entropía para distintos porcentajes de muestras perdidas, y el coeficiente de correlación cruzada de la entropía de las series con pérdidas con respecto a las series originales.

Se observa el mismo comportamiento que en los casos anteriores. Inicialmente se mantiene más o menos estable en torno a 0.45 para RA y 0.3 para NRA, para luego empezar a disminuir (en torno al 30–40 % pérdidas). De nuevo se observa cómo señales más aleatorias presentan un decaimiento más temprano, para menores pérdidas. La medida es estable por debajo del 70 % de muestras perdidas y se puede considerar que las series aún albergan cierta validez clínica.



**Figura 7.6:** (a) Comportamiento de ApEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul o recomendadas para ablación (RA) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) de la regularidad en función del ratio de pérdidas.

### 7.1.3 Señales de corta duración

En esta sección se muestran los resultados de aplicar ApEn en señales de longitud inferior a 1000 muestras, que se consideran de corta duración. Debido a los problemas que presenta ApEn con longitudes inferiores a 100 muestras, se han considerado solo las señales de Glucemia tanto en pacientes críticos como ambulatorios ( $N = 288$ ) y las señales de dinámicas de movilidad ( $N \approx 300$ ).

#### *Glucemia en pacientes críticos*

Estas señales presentan dos posibles clasificaciones, bien entre pacientes diabéticos (DM) y no diabéticos (NDM) o bien entre supervivientes (SV) y no supervivientes (NSV) [54].

A estas señales se les aplicó un procedimiento paramétrico con  $m = \{1, 2, \dots, 10\}$  y  $r = \{0.1, 0.2, \dots, 1\}$  de forma que se pudieran encontrar las combinaciones para las cuales se obtenía separación estadística entre clases.

En la tabla 7.7 se muestran únicamente las combinaciones de parámetros que dan lugar a separación estadística entre SV y NSV. Las combinaciones que consiguen obtener algún tipo de segmentación entre SV y NSV presentan  $m$  bajo y  $r$  relativamente alto.

ApEn no sería una medida adecuada para procesar este tipo de señales ya que incluso con valores de  $r$  altos y  $m$  bajos no se encuentran el número suficientes de patrones para asegurar una distribución normal y en consecuencia un resultado estable de la entropía [38].

**Tabla 7.7:** Combinación paramétrica que permite la segmentación entre sujetos supervivientes (SV) y no supervivientes (NSV) en función de los niveles de glucemia. Se muestran también los intervalos de confianza al 95 % (CI) y los test de normalidad (JB) y separabilidad (UT).

(m, r)	CI SV	CI NSV	JB	UT
(1,0.5)	[0.112, 0.174]	[0.053, 0.132]	0.004	0.043
(1,0.6)	[0.092, 0.144]	[0.043, 0.107]	0.004	0.040
(1,0.8)	[0.063, 0.102]	[0.027, 0.077]	0.001	0.043
(1,0.9)	[0.055, 0.087]	[0.024, 0.065]	0.001	0.034
(2,0.8)	[0.063, 0.101]	[0.027, 0.076]	0.001	0.043
(2,0.9)	[0.054, 0.085]	[0.023, 0.064]	0.002	0.030
(2,1.0)	[0.047, 0.072]	[0.020, 0.055]	0.003	0.043
(3,0.9)	[0.050, 0.080]	[0.023, 0.063]	0.001	0.040
(3,1.0)	[0.044, 0.069]	[0.020, 0.056]	0.003	0.047

Por contra, no se pudo conseguir una sola combinación de parámetros que permitiera a ApEn discernir entre DM y NDM. Esto es algo que refuerza las conclusiones obtenidas en [54] donde la glucemia puede ser crítica para determinar la supervivencia o no de un paciente en la UVI pero la condición crítica de éstos no permite encontrar diferencias entre DM y NDM.

### *Glucemia ambulatoria*

En este punto se presenta el análisis de las señales de glucemia en pacientes ambulatorios con ApEn. Se dispone de tres grupos de sujetos: sujetos diabéticos (DM) con síndrome metabólico (MS) o no diabéticos (NDM).

Al igual que en caso anterior, glucemia en pacientes críticos se realiza un estudio paramétrico con  $m = \{1, 2, \dots, 5\}$  y  $r = \{0.1, 0.2, \dots, 1\}$  para encontrar qué combinaciones permitían encontrar diferencias estadísticas y en qué condiciones.

La tabla 7.8 muestra las combinaciones de parámetros que permiten hallar diferencias estadísticas entre al menos dos grupos, DM/NDM, DM/MS, MS/NDM. Todas las combinaciones de parámetros consideradas excepto  $\{(3, 0.1), (4, 0.1)\}$  consiguen segmentación entre algún grupo. Los resultados en la tabla se han limitado a aquellas combinaciones paramétricas que permiten discernir estadísticamente entre alguno de los grupos y aquellas combinaciones en las que  $r$  maximiza ApEn para cada  $m$ .

DM y NDM pueden ser separados estadísticamente prácticamente con cualquier combinación paramétrica que se considere excepto aquellas con  $r = 0.1$ . Esto quizá se asocie de nuevo a que las variaciones en glucemia son lentas y de mayor amplitud, por ello no se ven en escalas bajas. Las combinaciones que quizá interesen más son las que permiten segmentar MS de los demás grupos. MS se considera un estado intermedio en la evolución de la patología diabética, es decir, se considera

**Tabla 7.8:** Combinación paramétrica que permite la segmentación entre distintas clases. Se muestran también los intervalos de confianza al 95% (CI) y los test de normalidad (JB) y separabilidad (UT). En negrita se identifica dónde no hay segmentación entre clases.  $UT_1$ : segmentación entre NDM y DM,  $UT_2$ : segmentación entre NDM y MS y  $UT_3$ : segmentación entre DM y MS.

(m, r)	CI NDM	CI MS	CI DM	JB	$UT_1$	$UT_2$	$UT_3$
(1,0.1)	[1.126,1.415]	[1.045,1.307]	[0.646,1.071]	0.237	0.018	<b>0.190</b>	0.026
(2,0.1)	[0.561,0.710]	[0.659,0.759]	[0.451,0.587]	0.373	<b>0.102</b>	<b>0.084</b>	0.002
(2,0.2)	[0.566,0.755]	[0.522,0.722]	[0.344,0.523]	0.427	0.011	<b>0.551</b>	<b>0.066</b>
(3,0.2)	[0.436,0.552]	[0.428,0.567]	[0.279,0.393]	0.448	0.009	<b>0.976</b>	0.008
(4,0.2)	[0.339,0.397]	[0.345,0.411]	[0.236,0.313]	0.281	0.005	<b>0.633</b>	0.005
(4,0.3)	[0.326,0.423]	[0.300,0.437]	[0.197,0.281]	0.263	0.007	<b>0.905</b>	0.013
(5,0.2)	[0.219,0.264]	[0.250,0.310]	[0.210,0.250]	0.428	<b>0.697</b>	<b>0.074</b>	0.024
(5,0.3)	[0.279,0.344]	[0.264,0.363]	[0.174,0.235]	0.459	0.004	<b>0.905</b>	0.005

que cuando un sujeto se clasifica como MS tiene grandes probabilidades de, si no cambia sus hábitos de vida, terminar desarrollando una diabetes.

ApEn sí es capaz de distinguir entre un sujeto MS y DM en la mayoría de los casos, algo útil puesto que permite saber con apenas un registro de un día si un sujeto que se sospecha ha desarrollado la enfermedad lo ha hecho. Por contra, no es capaz de distinguir entre NDM y MS, lo que no permite establecer en qué punto patológico se encuentra.

Como se puede observar  $CI\ DM < CI\ MS < CI\ NDM$ , es decir, sujetos sanos presentan mayor complejidad que sujetos patológicos. Esto es algo que ya Churrua estableció en [55], lo que refuerza la idea de que sistemas más deteriorados (patológicos) presentan mayor regularidad.

Finalmente si se busca una optimización en cuanto a minimización de la variabilidad de la entropía para seleccionar  $m$  y segmentación entre al menos dos clases simultáneamente, se escoge la combinación (4,0.3) donde  $ApEn_{NDM} = 0.374 \pm 0.049$ ,  $ApEn_{MS} = 0.368 \pm 0.069$ ,  $ApEn_{DM} = 0.239 \pm 0.042$ .

Al contrario de los resultados obtenidos sobre las señales de glucemia en pacientes críticos, los parámetros de ApEn han sido considerados en la literatura como adecuados, además la distribución de ApEn es normal lo que indica de acuerdo a [37] que el número de patrones coincidentes es suficiente para producir un resultado de entropía estable.

### Dinámicas de movilidad

La base de datos de dinámicas de movilidad presenta 4 grupos de señales: registros de control (CT), registros de pacientes con Parkinson (PD), sujetos con Huntington (HD) y sujetos con esclerosis lateral amiotrófica (LSA), todas ellas enfermedades neurodegenerativas.

Al igual que para los casos de glucemia, se ha desarrollado un estudio paramétrico en función de  $m$  y  $r$ . En la tabla 7.9 se muestran las combinaciones de parámetros que permiten segmentación estadística entre CT y alguna de las patologías. Se muestran solo aquellas combinaciones donde  $r$  maximiza la entropía de cada variable para cada  $m$ , descartando aquellas combinaciones donde  $r$  no presentara un máximo absoluto. Esto limita a  $m \leq 5$ . Se observa que a medida que aumentamos  $m$  es necesario aumentar  $r$  para obtener valores de entropía similares, lo que se ajusta a lo descrito en la literatura [10, 38].

**Tabla 7.9:** Caracterización de ApEn en función de los parámetros ( $m, r$ ) para los grupos considerados en la base de datos de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA).  $UT_1, UT_2, UT_3$  indican la probabilidad estadística de segmentación entre LSA, PD y HD con respecto a CT.

(m, r)	CI LSA	CI PD	CI HD	CI CT	$UT_1$	$UT_2$	$UT_3$
(1,0.10)	[1.610,1.697]	[1.879,1.951]	[1.589,1.951]	[1.854,1.946]	0.000	<b>0.828</b>	0.014
(2,0.30)	[0.889,0.983]	[1.034,1.140]	[0.902,1.107]	[1.008,1.125]	0.002	<b>0.782</b>	<b>0.161</b>
(2,0.40)	[0.777,0.942]	[0.932,1.090]	[0.920,1.144]	[0.928,1.106]	0.010	<b>0.813</b>	<b>0.426</b>
(3,0.40)	[0.582,0.662]	[0.698,0.763]	[0.557,0.686]	[0.665,0.723]	0.010	<b>0.185</b>	0.022
(3,0.50)	[0.558,0.671]	[0.676,0.800]	[0.646,0.793]	[0.656,0.770]	0.006	<b>0.333</b>	<b>0.577</b>
(3,0.60)	[0.488,0.640]	[0.608,0.755]	[0.656,0.823]	[0.613,0.762]	0.018	<b>0.797</b>	0.040
(4,0.60)	[0.412,0.496]	[0.506,0.596]	[0.425,0.527]	[0.494,0.581]	0.004	<b>0.649</b>	0.018
(4,0.70)	[0.385,0.492]	[0.476,0.591]	[0.477,0.594]	[0.484,0.608]	0.005	<b>0.812</b>	<b>0.599</b>
(4,0.80)	[0.340,0.477]	[0.428,0.562]	[0.481,0.611]	[0.446,0.585]	0.041	<b>0.678</b>	<b>0.373</b>
(5,0.70)	[0.323,0.397]	[0.394,0.472]	[0.318,0.396]	[0.392,0.470]	0.012	<b>0.984</b>	0.004
(5,0.80)	[0.309,0.395]	[0.370,0.472]	[0.364,0.452]	[0.382,0.487]	0.018	<b>0.678</b>	<b>0.286</b>
(5,0.90)	[0.277,0.380]	[0.339,0.457]	[0.380,0.478]	[0.364,0.482]	0.020	<b>0.489</b>	<b>0.987</b>

Ninguna combinación paramétrica considerada consigue simultáneamente encontrar diferencias estadísticas entre cada patología y CT. En particular, ApEn es incapaz de establecer diferencias entre PD y CT con ninguna combinación. Apenas 5 combinaciones permiten segmentar al menos CT de LSA y HD simultáneamente.

Buscando los dos criterios básicos de optimización, maximización de entropía y minimización de variabilidad los estudios indican que para esta base de datos se debería utilizar (3,0.4), de forma que  $ApEn_{LSA} = 0.622 \pm 0.040$ ,  $ApEn_{PD} = 0.731 \pm 0.032$ ,  $ApEn_{HD} = 0.622 \pm 0.065$  y  $ApEn_{CT} = 0.694 \pm 0.029$ .

## 7.2 Entropía muestral (SampEn)

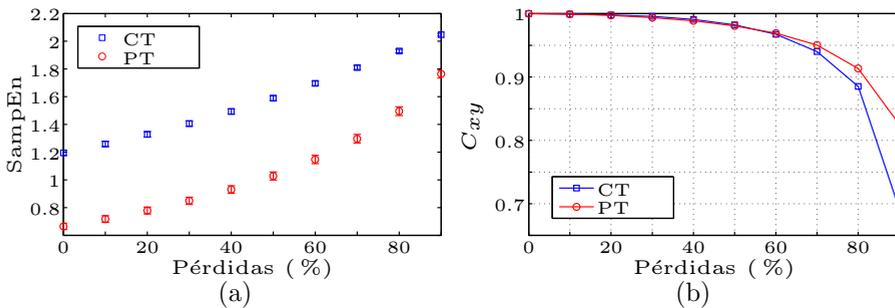
A continuación se muestran los resultados de SampEn en el procesado de EEG, EGM y registros RR ante la pérdida de muestras y en análisis sobre señales de corta duración.

### 7.2.1 Pérdida aleatoria de muestras

En esta sección se muestran los resultados relativos al experimento descrito en la sección 6.1 con pérdida aleatoria de muestras sobre señales EEG, RR y EGM.

#### *Electroencefalogramas*

La figura 7.7 muestra la evolución de SampEn con respecto al grado de pérdidas para las señales EEG de control (CT) y patológicas (PT) y el coeficiente de correlación cruzada ( $C_{xy}$ ).



**Figura 7.7:** (a) Comportamiento de SampEn ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de señal original y de las series con pérdidas en función del ratio de muestras eliminadas.

En ella se observa que la segmentación se mantiene para todo el rango de porcentaje de pérdidas considerado, con mayor complejidad asociada a CT. En este caso, al contrario que para ApEn no se observa un punto de inflexión. Esto es resultado de que SampEn no se calcula a partir de la comparativa de patrones y por ello no se da el fenómeno de *auto-matches*. La validez de la medida también se ve incrementada, de un 60% en el caso de ApEn hasta más de un 80% de muestras perdidas en ambas clases con SampEn.

La tabla 7.10 muestra las características estadísticas de SampEn sobre señales EEG. Se indica la validez clínica de las series hasta alcanzar un 70% de muestras eliminadas de forma aleatoria.

**Tabla 7.10:** Caracterización de SampEn sobre registros EEG de control (CT) y patológicos (PT) que han perdido muestras de forma aleatoria. Se muestra en función del porcentaje de pérdidas en la señal la normalidad (JB), la separabilidad de los grupos (UT), los intervalos de confianza al 95 % (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
U (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.176, 1.214]	[1.239, 1.278]	[1.387, 1.425]	[1.571, 1.607]	[1.795, 1.826]	[2.037, 2.058]
CI PT	[0.641, 0.688]	[0.694, 0.742]	[0.822, 0.875]	[0.999, 1.057]	[1.266, 1.328]	[1.735, 1.792]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.995(0.001)	0.982(0.001)	0.945(0.001)	0.758(0.001)

### Señales RR

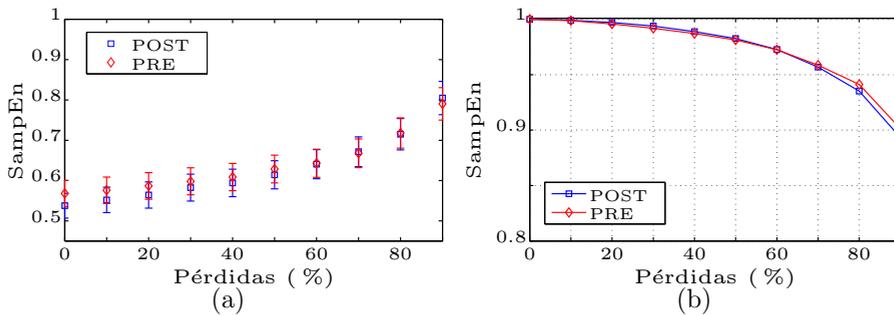
En la tabla 7.11 se muestra la caracterización de SampEn en función del ratio de pérdidas de forma aleatoria que se le ha aplicado a las señales RR. En este caso no existe segmentación, esto nos indicaría que la medicación no es eficiente, ya que no se pueden encontrar diferencias estadísticas entre las clases ( $p > 0.05$ ).

**Tabla 7.11:** Caracterización de SampEn sobre registros RR en pacientes antes (Pre) y después (Post) del tratamiento con encainida que han sufrido pérdidas de forma aleatoria. Se muestra, en función del porcentaje de pérdidas, la normalidad (JB), la separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.293	0.439	0.568	0.574	0.929	0.586
CI POST	[0.507,0.569]	[0.521,0.584]	[0.549,0.616]	[0.579,0.649]	[0.635,0.709]	[0.763,0.846]
CI PRE	[0.536,0.601]	[0.543,0.609]	[0.565,0.632]	[0.594,0.663]	[0.632,0.703]	[0.750,0.830]
$C_{xy}$ (p)	1 (0.001)	0.998 (0.001)	0.990 (0.001)	0.981 (0.001)	0.957(0.001)	0.895 (0.001)

La figura 7.8 muestra la evolución de SampEn en función del ratio de eliminación de muestras considerado. Sin tener en cuenta que las medidas se solapan, se observa como al igual que sucede en las señales EEG o para ApEn la medida experimenta un comportamiento creciente. Este comportamiento viene asociado a la aleatorización que se produce al eliminar muestras en los registros. En este caso, no se encuentra el punto de inflexión donde la medida cambia de tendencia y decrece. Esto se debe a que SampEn, como se ha comentado con anterioridad, no considera los *auto-matches* y elimina ese sesgo que indica mayor regularidad de la existente en la señal.

La validez clínica se mantiene ( $C_{xy} > 0.8$ ). Esto indica que la medida es estable en todo el rango considerado.



**Figura 7.8:** Comportamiento de SampEn cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul, tratados con encainida. (a) Evolución de SampEn y (b) Coeficiente de correlación cruzada entre la serie original y sus versiones con pérdida de muestras.

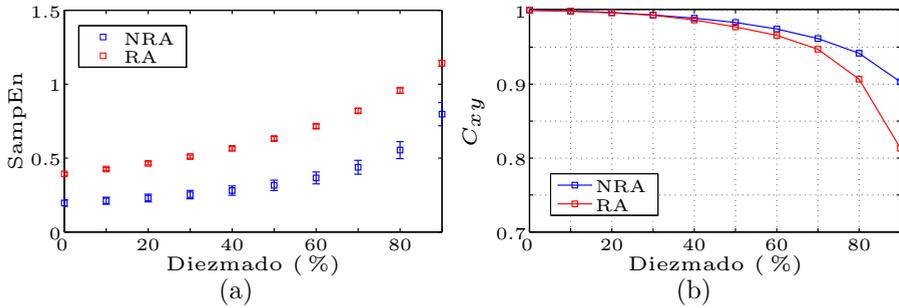
### Electrogramas

La tabla 7.12 muestra la caracterización estadística del comportamiento de SampEn para distintos valores de pérdidas. La distribución de SampEn es normal para todo el rango considerado y la segmentación entre clases completa. En este caso SampEn también es más estable que ApEn ya que para los mismos porcentajes de pérdida de muestras,  $C_{xy}$  presenta mayores valores.

**Tabla 7.12:** Caracterización de SampEn sobre los EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando la pérdida de muestras se produce de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.051	0.061	0.093	0.202	0.493	0.241
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.193, 0.199]	[0.208, 0.214]	[0.248, 0.256]	[0.310, 0.320]	[0.431, 0.444]	[0.786, 0.808]
CI RA	[0.366, 0.376]	[0.397, 0.407]	[0.476, 0.488]	[0.595, 0.609]	[0.801, 0.818]	[1.300, 1.321]
$C_{xy}$ (p)	1.000(0.001)	1.000 (0.001)	0.997 (0.001)	0.991 (0.001)	0.979 (0.001)	0.937 (0.001)

En la figura 7.9 se muestran tanto la evolución de SampEn con el número de muestras perdidas como el coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de las señales originales frente a la regularidad de sus versiones diezadas. Se puede observar cómo la segmentación entre EGM NRA y EGM RA es completa. Al igual que para señales EEG y RR, SampEn aumenta con el diezado y tampoco presenta el punto de cambio de tendencia que se observaba con ApEn.



**Figura 7.9:** (a) Comportamiento de SampEn cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul y recomendadas para ablación (RA) en rojo y (b)  $C_{xy}$  entre la regularidad de la serie original y sus versiones con pérdidas para distintos niveles de fraccionamiento.

## 7.2.2 Pérdida uniforme de muestras

En esta sección se muestran los resultados relativos al experimento descrito en la sección 6.1 para la entropía SampEn con pérdidas de forma uniforme sobre las bases de datos EEG, RR y EGM.

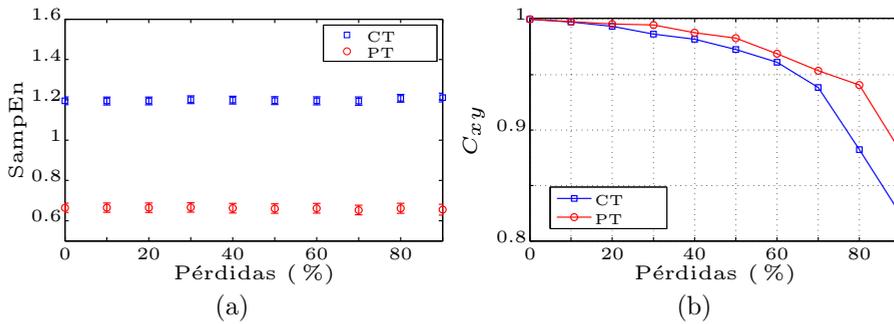
### *Electroencefalogramas*

En la tabla 7.13 se muestra la caracterización estadística de SampEn. La segmentación es completa para cualquier porcentaje de pérdidas ( $p = 0.001$ ) y los intervalos de confianza no se solapan. La medida es estable ( $C_{xy} > 0.8$ ) en todo el rango.

**Tabla 7.13:** Caracterización de SampEn en señales EEG control (CT) y patológicas (PT) con pérdida de muestras de forma uniforme. Se muestra: test de normalidad (JB) y separabilidad (UT), intervalos de confianza al 95 % (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.176, 1.214]	[1.175, 1.213]	[1.182, 1.220]	[1.176, 1.216]	[1.173, 1.214]	[1.190, 1.233]
CI PT	[0.641, 0.688]	[0.642, 0.688]	[0.643, 0.689]	[0.637, 0.684]	[0.629, 0.677]	[0.628, 0.682]
$C_{xy}$ (p)	1.000(0.001)	0.998(0.001)	0.991(0.001)	0.973(0.001)	0.946(0.001)	0.853(0.001)

La figura 7.10 muestra la evolución de SampEn cuando la señal experimenta una pérdida de muestras de forma uniforme. En este caso se observa que la medida se mantiene más o menos constante para todos los valores de pérdida de muestras considerados.



**Figura 7.10:** Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) o patológica (PT). (a) Evolución de SampEn en función del ratio de pérdidas, (b) coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la serie original y la de las series con pérdidas.

Al igual que ocurría en ApEn se presenta un comportamiento en el cuál la entropía se mantiene aproximadamente constante de forma independiente al porcentaje de pérdidas que se le aplica a la señal. Por otro lado se observa que no se obtiene ese punto de cambio de tendencia característico de ApEn tal y como se obtuvo con las pérdidas aleatorias.

### Señales RR

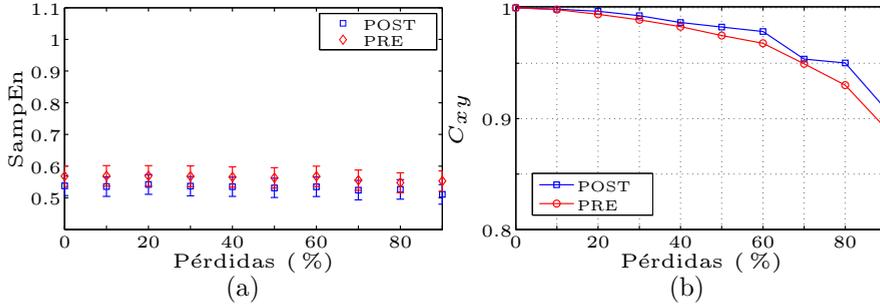
La tabla 7.14 muestra la caracterización estadística de SampEn sobre señales RR con pérdida de muestras de forma uniforme. De nuevo las distribuciones no son normales y no existe separación estadística entre clases, tal y como pasaba cuando se tenía pérdida de muestras de forma aleatoria. Se puede inferir que las clases no son separables independientemente del tipo de pérdidas sufrido.

**Tabla 7.14:** Caracterización de SampEn en registros RR antes (Pre) y después (Post) del tratamiento con encainida y pérdida de muestras de forma uniforme. Se muestra: test de normalidad (JB) y separabilidad (UT), intervalos de confianza al 95 % (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.293	0.251	0.365	0.270	0.237	0.066
CI Post	[0.507,0.569]	[0.505,0.566]	[0.506,0.567]	[0.501,0.562]	[0.493,0.555]	[0.480,0.542]
CI Pre	[0.536,0.601]	[0.537,0.602]	[0.535,0.601]	[0.531,0.595]	[0.524,0.588]	[0.521,0.585]
$C_{xy}$ (p)	1.000 (0.001)	0.999 (0.001)	0.991 (0.001)	0.979 (0.001)	0.952 (0.001)	0.900 (0.001)

En la figura 7.11 se presenta el comportamiento de SampEn en función del porcentaje de muestras que se han eliminado de los registros RR y el coeficiente de correlación cruzada de la medida de entropía. Como se ha observado para el caso

de los EEG, SampEn se mantiene constante para todo el rango de pérdidas con validez clínica.



**Figura 7.11:** Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal RR en sujetos pre-medicación (PRE) en rojo y post-medicación (POST) en azul. (a) Evolución en función del porcentaje de muestras eliminadas de forma uniforme, (b) coeficiente de correlación cruzada entre la regularidad de la señal original y sus versiones con pérdidas.

### Electrogramas

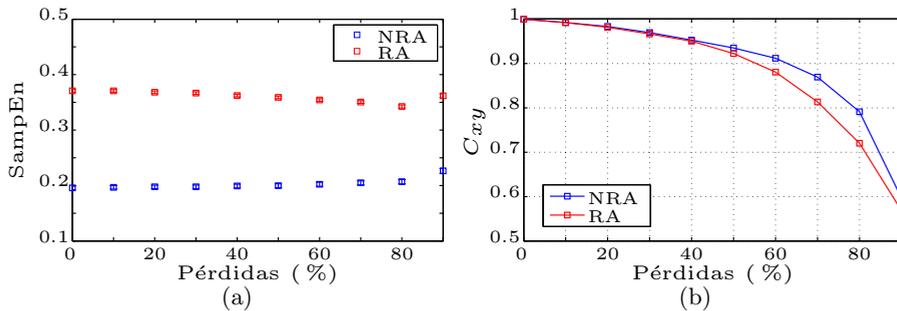
La tabla 7.15 muestra la caracterización estadística de la medida, en ella se reflejan distribuciones normales hasta alcanzar el 90 % de muestras perdidas, segmentación entre clases ( $p=0.001$ ) e intervalos de confianza sin solapamiento.

**Tabla 7.15:** Caracterización de SampEn sobre EGM registrados en áreas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se pierden de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.051	0.052	0.071	0.100	0.090	0.004
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.193, 0.199]	[0.194, 0.200]	[0.195, 0.201]	[0.196, 0.203]	[0.201, 0.209]	[0.222, 0.231]
CI RA	[0.366, 0.376]	[0.366, 0.376]	[0.362, 0.372]	[0.354, 0.364]	[0.346, 0.356]	[0.355, 0.369]
$C_{xy}$ (p)	1.000(0.001)	1.000 (0.001)	0.998 (0.001)	0.981 (0.001)	0.927 (0.001)	0.757 (0.001)

La figura 7.12 muestra la evolución de SampEn para registros de electrogramas (EGM) de fibrilación auricular registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) y el coeficiente de correlación cruzada ( $C_{xy}$ ) entre las entropías de las series originales y sus versiones con pérdidas.

Al igual que se ha observado en las señales anteriores (EEG y RR) la entropía se mantiene aproximadamente constante en un intervalo amplio (hasta 60–70 % de muestras perdidas) que es el rango donde existe validez clínica ( $C_{xy} > 0.8$ ).



**Figura 7.12:** (a) Comportamiento de SampEn cuando se eliminan muestras de forma uniforme sobre una señal EGM no recomendados para ablación (NRA) en azul y recomendados para ablación (RA) en rojo, y (b) coeficiente de correlación cruzada entre la regularidad de la serie original y las series con pérdida de muestras.

### 7.2.3 Señales de corta duración

En esta sección se muestran los resultados de aplicar ApEn en señales de longitud inferior a 1000 muestras, que se consideran como de corta duración. Debido a los problemas que presenta ApEn con longitudes inferiores a 100 muestras se han considerado solo las señales de glucemia tanto en pacientes críticos como ambulatorios ( $N = 288$ ) y las señales de dinámicas de movilidad ( $N \approx 300$ ).

#### *Glucemia en pacientes críticos*

Al igual que ocurría con ApEn, SampEn tampoco es capaz de encontrar diferencias significativas entre sujetos diabéticos (DM) respecto a sujetos no diabéticos (NDM). Tampoco entre sujetos supervivientes (SV) con respecto a sujetos no supervivientes (NSV).

Aún cuando SampEn es una medida más estable y coherente que ApEn, se observa que en sujetos en estado crítico no consigue valorar las posibles diferencias estadísticas que éstos presenten y es incapaz de encontrar ninguna diferencia significativa entre las clases.

#### *Glucemia ambulatoria*

El comportamiento que se obtenía con ApEn para estos registros se encuentra reproducido aquí (tabla 7.16). SampEn sí es capaz de segmentar entre NDM y DM, pero no lo es para NDM y MS.

Como diferencia, menos combinaciones paramétricas permiten obtener una segmentación entre MS y DM. En la tabla se ha decidido mostrar aquellas combina-

**Tabla 7.16:** Caracterización de SampEn sobre registros de glucemia ambulatoria. Para cada combinación paramétrica se dan los intervalos de confianza (CI), test de normalidad (JB) y separabilidad (UT). En negrita se identifican las combinaciones sin separación estadística entre clases. UT<sub>1</sub>: segmentación entre NDM y DM, UT<sub>2</sub>: segmentación entre NDM y MS y UT<sub>3</sub>: segmentación entre DM y MS.

(m, r)	CI NDM	CI MS	CI DM	JB	UT <sub>1</sub>	UT <sub>2</sub>	UT <sub>3</sub>
(1,0.2)	[1.936,2.194]	[1.899,2.121]	[1.733,1.983]	0.500	0.027	<b>0.571</b>	<b>0.066</b>
(2,0.1)	[1.127,1.580]	[0.937,1.306]	[0.471,0.939]	0.380	0.007	<b>0.152</b>	0.028
(3,0.1)	[0.920,1.392]	[0.798,1.081]	[0.418,0.702]	0.179	0.008	<b>0.283</b>	0.005
(4,0.1)	[0.841,1.277]	[0.710,0.949]	[0.354,0.634]	0.208	0.009	<b>0.233</b>	0.008
(5,0.1)	[0.867,1.509]	[0.639,0.903]	[0.306,0.630]	0.054	0.017	<b>0.136</b>	0.018

ciones paramétricas de acuerdo a las siguientes restricciones: (i)  $r$  debe garantizar SampEn máxima para cada  $m$  en cada variable. (ii) deben permitir segmentación entre al menos algún grupo de forma significativa.

Se observa la consistencia de la medida, donde el máximo de SampEn se produce siempre para los tres grupos considerados con la misma combinación paramétrica. Tenemos el comportamiento establecido en [55] donde la complejidad es menor a medida que se constata la patología diabética.

Considerando las restricciones establecidas, la combinación óptima de parámetros  $(m,r)$  es (2,0.1). Con esos parámetros se obtiene  $\text{SampEn}_{\text{NDM}} = 1.354 \pm 0.227$ ,  $\text{SampEn}_{\text{MS}} = 1.122 \pm 0.185$ ,  $\text{SampEn}_{\text{DM}} = 0.705 \pm 0.24$ , ya que determina  $r$  para SampEn máximo en  $m=2$  y segmentación en dos variables.

### *Dinámicas de movilidad*

La tabla 7.17 muestra los resultados de SampEn para las señales de dinámicas de movilidad. En ella se muestran los intervalos al 95 % y los estadísticos de separabilidad entre las señales de control (CT) y las señales patológicas que corresponden a sujetos con Párkinson (PD), Huntington (HD) o esclerosis lateral amiotrófica (LSA). Se muestran únicamente aquellas combinaciones paramétricas  $(m, r)$  que presentan segmentación estadística para aquel  $r$  que permite maximizar el valor de SampEn en cada  $m$ .

La primera diferencia que debemos notar es que SampEn obtiene peor segmentación que ApEn en estas señales pese a ser una medida más consistente y estable.

La consistencia se observa en que en todos los grupos el máximo valor de entropía se obtiene simultáneamente con los mismos parámetros en los cuatro grupos. En este caso solo se obtiene separación estadística entre CT y HD independientemente de la combinación de parámetros elegida.

**Tabla 7.17:** Caracterización de SampEn en función de los parámetros  $(m, r)$  para los registros de dinámicas de movilidad, control (CT), sujetos con Parkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA).  $UT_1, UT_2, UT_3$  indican la probabilidad estadística de segmentación entre LSA, PD y HD con respecto a CT.

$(m, r)$	CI LSA	CI PD	CI HD	CI CT	$UT_1$	$UT_2$	$UT_3$
(1,0.10)	[2.200,2.517]	[2.389,2.632]	[2.458,2.844]	[2.406,2.714]	<b>0.114</b>	<b>0.374</b>	0.043
(2,0.10)	[2.082,2.437]	[2.284,2.539]	[2.303,2.889]	[2.298,2.610]	<b>0.100</b>	<b>0.441</b>	0.015
(3,0.20)	[1.402,1.974]	[1.555,1.830]	[1.676,2.144]	[1.594,1.917]	<b>0.324</b>	<b>0.594</b>	0.047
(4,0.30)	[1.054,1.396]	[1.193,1.500]	[1.364,1.817]	[1.139,1.478]	<b>0.584</b>	<b>0.767</b>	0.013
(5,0.40)	[0.741,1.056]	[0.922,1.250]	[1.122,1.524]	[0.926,1.251]	<b>0.142</b>	<b>0.890</b>	0.037
(6,0.80)	[0.289,0.536]	[0.381,0.588]	[0.559,0.787]	[0.413,0.630]	<b>0.219</b>	<b>0.707</b>	0.037
(7,0.60)	[0.439,0.768]	[0.535,0.808]	[0.757,1.227]	[0.583,0.840]	<b>0.324</b>	<b>0.678</b>	0.043
(8,0.80)	[0.288,0.558]	[0.352,0.594]	[0.571,0.862]	[0.398,0.623]	<b>0.283</b>	<b>0.722</b>	0.034

### 7.3 Entropía Multiescala (MSE)

En esta sección se muestran los resultados obtenidos de los distintos registros con entropía multiescala (MSE). Esta entropía se suele aplicar a registros largos y se encuentra poco caracterizada con señales cortas. En cada caso se han considerado las escalas que garantizaran al menos el análisis de registros de 100 muestras.

#### 7.3.1 Pérdida aleatoria de muestras

##### *Electroencefalogramas*

De la tabla 7.18 a la tabla 7.26 se muestra la caracterización estadística de MSE sobre señales EEG a las que se les ha aplicado un algoritmo de pérdida de muestras de forma aleatoria. La escala 1 no se muestra por tratarse de SampEn, cuyos resultados se han presentado en la sección 7.2.1.

**Tabla 7.18:** Caracterización de MSE 2 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.688, 1.724]	[1.737, 1.772]	[1.828, 1.860]	[1.894, 1.919]	[1.904, 1.920]	[1.871, 1.893]
CI PT	[1.074, 1.136]	[1.133, 1.196]	[1.270, 1.335]	[1.437, 1.501]	[1.640, 1.699]	[1.835, 1.875]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.994(0.001)	0.970(0.001)	0.855(0.001)	0.375(0.001)

En todas las escalas se observa segmentación estadística de las clases hasta ratios del 70 %, aún cuando la tendencia en la complejidad varía en las dos clases. La

**Tabla 7.19:** Caracterización de MSE 3 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.886, 1.917]	[1.907, 1.936]	[1.928, 1.950]	[1.903, 1.918]	[1.826, 1.844]	[1.751, 1.781]
CI PT	[1.322, 1.388]	[1.376, 1.441]	[1.492, 1.556]	[1.622, 1.682]	[1.746, 1.796]	[1.774, 1.812]
$C_{xy}$ (p)	1.000(0.001)	0.998(0.001)	0.984(0.001)	0.911(0.001)	0.618(0.001)	0.064(0.001)

**Tabla 7.20:** Caracterización de MSE 4 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.947, 1.971]	[1.946, 1.968]	[1.915, 1.930]	[1.839, 1.856]	[1.740, 1.766]	[1.663, 1.702]
CI PT	[1.475, 1.540]	[1.520, 1.584]	[1.616, 1.676]	[1.710, 1.764]	[1.764, 1.809]	[1.695, 1.736]
$C_{xy}$ (p)	1.000(0.001)	0.996(0.001)	0.960(0.001)	0.790(0.001)	0.358(0.001)	-0.040(0.001)

**Tabla 7.21:** Caracterización de MSE 5 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.945, 1.964]	[1.926, 1.941]	[1.860, 1.876]	[1.766, 1.789]	[1.672, 1.705]	[1.593, 1.637]
CI PT	[1.574, 1.637]	[1.612, 1.674]	[1.687, 1.744]	[1.744, 1.794]	[1.748, 1.793]	[1.616, 1.660]
$C_{xy}$ (p)	1.000(0.001)	0.991(0.001)	0.913(0.001)	0.652(0.001)	0.209(0.001)	-0.052(0.001)

**Tabla 7.22:** Caracterización de MSE 6 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.911, 1.925]	[1.877, 1.894]	[1.799, 1.820]	[1.700, 1.730]	[1.623, 1.661]	[1.529, 1.578]
CI PT	[1.642, 1.702]	[1.671, 1.729]	[1.724, 1.778]	[1.747, 1.795]	[1.718, 1.765]	[1.540, 1.588]
$C_{xy}$ (p)	1.000(0.001)	0.984(0.001)	0.872(0.001)	0.569(0.001)	0.197(0.001)	-0.020(0.001)

**Tabla 7.23:** Caracterización de MSE 7 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.927
CI CT	[1.858, 1.878]	[1.822, 1.843]	[1.738, 1.764]	[1.644, 1.680]	[1.584, 1.625]	[1.473, 1.527]
CI PT	[1.683, 1.742]	[1.707, 1.763]	[1.739, 1.790]	[1.734, 1.783]	[1.682, 1.733]	[1.475, 1.526]
$C_{xy}$ (p)	1.000(0.001)	0.978(0.001)	0.848(0.001)	0.549(0.001)	0.249(0.001)	0.027(0.001)

**Tabla 7.24:** Caracterización de MSE 8 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.807, 1.831]	[1.771, 1.796]	[1.682, 1.713]	[1.603, 1.644]	[1.550, 1.595]	[1.423, 1.479]
CI PT	[1.710, 1.767]	[1.725, 1.780]	[1.736, 1.787]	[1.717, 1.768]	[1.644, 1.700]	[1.411, 1.463]
$C_{xy}$ (p)	1.000(0.001)	0.966(0.001)	0.818(0.001)	0.615(0.001)	0.379(0.001)	0.103(0.001)

**Tabla 7.25:** Caracterización de MSE 9 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

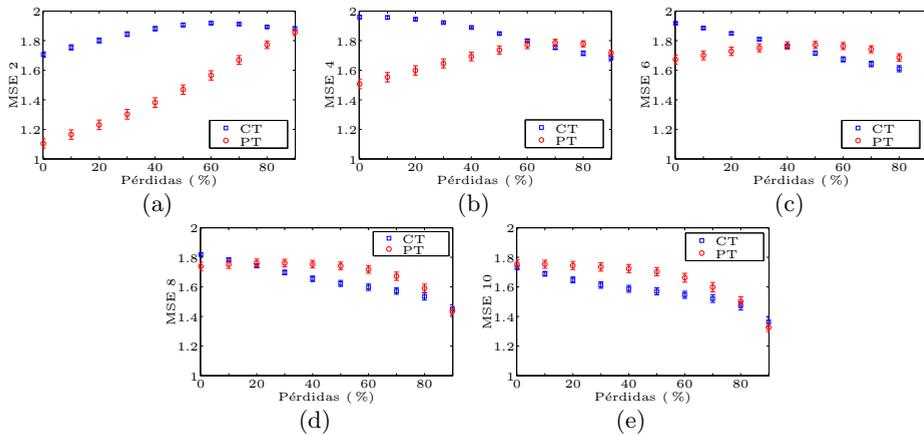
Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.764, 1.789]	[1.722, 1.750]	[1.632, 1.670]	[1.569, 1.614]	[1.522, 1.570]	[1.375, 1.436]
CI PT	[1.725, 1.779]	[1.732, 1.785]	[1.723, 1.776]	[1.697, 1.750]	[1.605, 1.665]	[1.353, 1.407]
$C_{xy}$ (p)	1.000(0.001)	0.960(0.001)	0.835(0.001)	0.665(0.001)	0.443(0.001)	0.126(0.001)

**Tabla 7.26:** Caracterización de MSE 10 cuando se eliminan muestras de una señal EEG de control (CT) o patológica (PT) de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza(CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.717, 1.744]	[1.672, 1.705]	[1.591, 1.635]	[1.546, 1.594]	[1.495, 1.547]	[1.332, 1.394]
CI PT	[1.727, 1.783]	[1.727, 1.781]	[1.707, 1.763]	[1.675, 1.731]	[1.566, 1.630]	[1.298, 1.354]
$C_{xy}$ PT (p)	1.000(0.001)	0.960(0.001)	0.835(0.001)	0.665(0.001)	0.443(0.001)	0.126(0.001)

validez clínica de la medida se ve comprometida antes (menores valores de los porcentajes de muestras perdidas) según aumenta la escala considerada.

En la figura 7.13 se presenta la evolución de MSE en función del ratio de muestras eliminadas para algunas de las escalas consideradas. Se observa cómo a medida que aumenta la escala la tendencia en la complejidad de la señal se invierte. Según aumenta la escala, el punto de cruce entre CT y PT se obtiene para menores ratios de muestras eliminadas. Este cambio de tendencia podría indicar que las señales de control a largo plazo son más regulares manteniendo comportamientos más similares y predictivos que las patológicas.



**Figura 7.13:** Comportamiento de MSE cuando se eliminan muestras de forma aleatoria sobre una señal EEG control (CT) o patológica (PT). (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10

De acuerdo a los resultados presentados, a mayor escalado menor estabilidad presenta la medida, pues  $C_{xy} < 0.8$  para menores valores de pérdidas, según se ha mostrado en las tablas 7.18–7.26. La variación que experimenta PT con respecto a la variación de CT en función del ratio de muestras perdidas es más similar a mayor escala. Se puede suponer que en señales más aleatorias (CT), la influencia de la escala de MSE se traducirá en una mayor regularización de la señal que en señales a priori más regulares.

No se debería considerar el uso de escalas superiores a 7 en ningún caso pues proporcionan resultados contrarios a la literatura indicando que CT presenta mayor regularidad que PT.

### Señales RR

Las tablas 7.27 a 7.35 presentan la caracterización estadística de MSE sobre señales RR registradas sobre sujetos tratados con encainida, antes (Pre) y después (Post) del tratamiento.

**Tabla 7.27:** Caracterización de MSE 2 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.031	0.465	0.147	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.948, 1.014]	[0.978, 1.044]	[1.053, 1.119]	[1.146, 1.213]	[1.257, 1.326]	[1.429, 1.504]
CI Pre	[1.112, 1.198]	[1.131, 1.216]	[1.183, 1.267]	[1.264, 1.347]	[1.357, 1.441]	[1.555, 1.645]
$C_{xy}$ (p)	1.000(0.001)	0.984(0.001)	0.948(0.001)	0.898(0.001)	0.811(0.001)	0.647(0.001)

**Tabla 7.28:** Caracterización de MSE 3 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.166	0.242	0.077	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.929, 0.993]	[0.979, 1.043]	[1.083, 1.148]	[1.178, 1.242]	[1.279, 1.344]	[1.450, 1.532]
CI Pre	[1.040, 1.115]	[1.069, 1.144]	[1.151, 1.224]	[1.245, 1.321]	[1.341, 1.417]	[1.543, 1.632]
$C_{xy}$ (p)	1.000(0.001)	0.980(0.001)	0.939(0.001)	0.901(0.001)	0.812(0.001)	0.558(0.001)

**Tabla 7.29:** Caracterización de MSE 4 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.055	0.215	0.045	0.001	0.001	0.001
UT (p)	0.001	0.001	0.002	0.001	0.014	0.097
CI Post	[0.945, 1.010]	[1.004, 1.067]	[1.107, 1.169]	[1.194, 1.256]	[1.289, 1.351]	[1.440, 1.524]
CI Pre	[1.043, 1.118]	[1.087, 1.161]	[1.168, 1.241]	[1.258, 1.332]	[1.329, 1.404]	[1.486, 1.577]
$C_{xy}$ (p)	1.000(0.001)	0.980(0.001)	0.940(0.001)	0.881(0.001)	0.771(0.001)	0.481(0.001)

La figura 7.14 muestra la evolución de MSE para distintas escalas en función del porcentaje de muestras perdidas en los registros RR.

En ningún caso se dispone de distribuciones normales pero sí de separación entre clases para la mayoría de las escalas y porcentajes de pérdida de muestras. La medida es estable en todas las escalas siempre que el porcentaje de muestras perdidas no sea superior al 50%.

**Tabla 7.30:** Caracterización de MSE 5 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.008	0.007	0.001	0.001	0.001	0.001
UT (p)	0.001	0.002	0.011	0.015	0.036	0.231
CI Post	[0.991, 1.057]	[1.040, 1.104]	[1.128, 1.189]	[1.202, 1.263]	[1.279, 1.341]	[1.454, 1.548]
CI Pre	[1.070, 1.143]	[1.102, 1.174]	[1.172, 1.243]	[1.242, 1.312]	[1.310, 1.381]	[1.486, 1.582]
$C_{xy}$ (p)	1.000(0.001)	0.972(0.001)	0.936(0.001)	0.874(0.001)	0.756(0.001)	0.472(0.001)

**Tabla 7.31:** Caracterización de MSE 6 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.002	0.001	0.002	0.454
CI Post	[1.002, 1.070]	[1.049, 1.114]	[1.129, 1.190]	[1.191, 1.252]	[1.254, 1.315]	[1.426, 1.522]
CI Pre	[1.087, 1.161]	[1.116, 1.189]	[1.181, 1.250]	[1.250, 1.321]	[1.307, 1.379]	[1.445, 1.543]
$C_{xy}$ (p)	1.000(0.001)	0.977(0.001)	0.931(0.001)	0.872(0.001)	0.690(0.001)	0.453(0.001)

**Tabla 7.32:** Caracterización de MSE 7 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.002	0.002	0.001	0.001	0.001
UT (p)	0.001	0.001	0.004	0.002	0.081	0.275
CI Post	[1.024, 1.093]	[1.058, 1.125]	[1.135, 1.198]	[1.186, 1.248]	[1.259, 1.323]	[1.369, 1.460]
CI Pre	[1.102, 1.175]	[1.131, 1.203]	[1.190, 1.260]	[1.241, 1.309]	[1.279, 1.349]	[1.401, 1.505]
$C_{xy}$ (p)	1.000(0.001)	0.976(0.001)	0.935(0.001)	0.869(0.001)	0.716(0.001)	0.1402(0.001)

**Tabla 7.33:** Caracterización de MSE 8 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

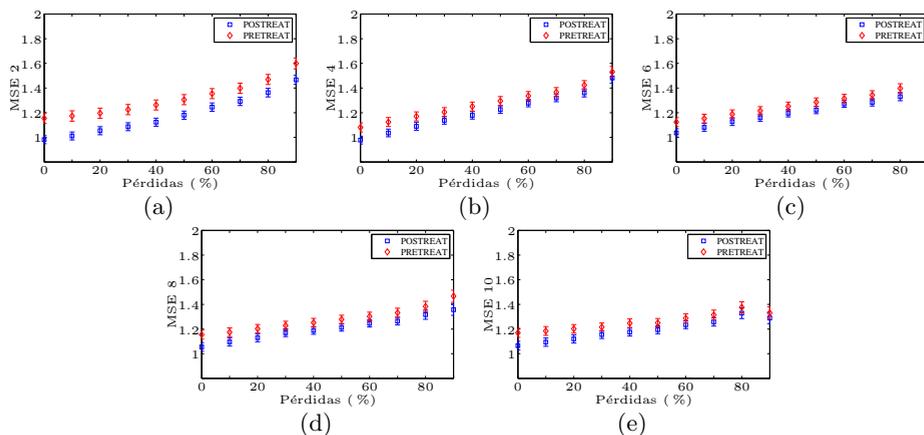
Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.105	0.001
UT (p)	0.001	0.001	0.003	0.001	0.001	0.001
CI Post	[1.019, 1.091]	[1.062, 1.130]	[1.138, 1.202]	[1.182, 1.244]	[1.232, 1.296]	[1.311, 1.405]
CI Pre	[1.117, 1.192]	[1.138, 1.210]	[1.194, 1.264]	[1.243, 1.313]	[1.295, 1.370]	[1.416, 1.517]
$C_{xy}$ (p)	1.000(0.001)	0.971(0.001)	0.931(0.001)	0.845(0.001)	0.679(0.001)	0.392(0.001)

**Tabla 7.34:** Caracterización de MSE 9 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.009	0.006	0.446	0.006
UT (p)	0.001	0.001	0.003	0.007	0.034	0.864
CI Post	[1.028, 1.099]	[1.066, 1.135]	[1.127, 1.191]	[1.172, 1.236]	[1.227, 1.292]	[1.290, 1.383]
CI Pre	[1.123, 1.197]	[1.146, 1.218]	[1.185, 1.255]	[1.219, 1.289]	[1.269, 1.342]	[1.296, 1.389]
$C_{xy}$ (p)	1.000(0.001)	0.969(0.001)	0.921(0.001)	0.853(0.001)	0.648(0.001)	0.351(0.001)

**Tabla 7.35:** Caracterización de MSE 10 con pérdida de muestras de forma aleatoria en registros RR antes (Pre) y después del tratamiento (Post) con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdida	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.037	0.001	0.002
UT (p)	0.001	0.001	0.002	0.006	0.013	0.203
CI Post	[1.030, 1.101]	[1.059, 1.128]	[1.123, 1.188]	[1.162, 1.228]	[1.224, 1.293]	[1.243, 1.338]
CI Pre	[1.132, 1.207]	[1.148, 1.220]	[1.182, 1.250]	[1.216, 1.286]	[1.276, 1.354]	[1.281, 1.381]
$C_{xy}$ (p)	1.000(0.001)	0.962(0.001)	0.917(0.001)	0.839(0.001)	0.622(0.001)	0.314(0.001)



**Figura 7.14:** Comportamiento de MSE sobre registros RR obtenidos en pacientes antes (Pre) y después (Post) del tratamiento con encainida cuando se da pérdida de muestras aleatoria. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10.

En primer lugar se observa que para escalas superiores a 1 (SampEn, sec. 7.2.1) esta medida al igual, que ApEn, (sec. 7.1.1) sí es capaz de separar de forma estadística las clases dando lugar a intervalos de confianza (CI) sin solapamiento. Se observa el mismo comportamiento que ocurría tanto para SampEn como para ApEn y para

EEG o EGM, la medida aumenta con el número de muestras perdidas. Presentan siempre mayor regularidad las señales Post que Pre. El aumento en la escala se traduce en una mayor similitud entre las regularidades de ambos grupos, aumenta la regularidad de Pre y disminuye la de Post, aunque se mantiene la separación estadística entre clases.

### Electrogramas

En este punto se muestran los resultados de aplicar pérdida de muestras de tipo aleatorio a los EGM para distintas escalas de MSE.

En las tablas 7.36 a 7.44 se muestra la caracterización estadística de la medida para las distintas escalas que se han considerado. En ella se denota la normalidad de la distribución en prácticamente cada caso (escala y valor de muestras perdidas) y la capacidad de segmentación entre clases.

**Tabla 7.36:** Caracterización de MSE 2 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.098	0.147	0.284	0.492	0.399	0.015
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.325, 0.336]	[0.350, 0.361]	[0.412, 0.426]	[0.507, 0.523]	[0.683, 0.704]	[1.131, 1.161]
CI RA	[0.640, 0.656]	[0.683, 0.700]	[0.791, 0.809]	[0.943, 0.962]	[1.186, 1.207]	[1.664, 1.686]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.997 (0.001)	0.990 (0.001)	0.970 (0.001)	0.897 (0.001)

**Tabla 7.37:** Caracterización de MSE 3 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.238	0.348	0.499	0.423	0.103	0.004
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.429, 0.443]	[0.462, 0.476]	[0.539, 0.556]	[0.654, 0.674]	[0.861, 0.886]	[1.300, 1.330]
CI RA	[0.838, 0.857]	[0.884, 0.904]	[0.998, 1.018]	[1.155, 1.176]	[1.405, 1.427]	[1.766, 1.787]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.996 (0.001)	0.985 (0.001)	0.963 (0.001)	0.846 (0.001)

En la figura 7.15 se muestra la evolución de la medida en función del ratio de muestras perdidas para algunas de las escalas consideradas. Se observa el mismo comportamiento que en las señales EEG, la complejidad de las señales menos regulares (EGM RA) empieza a disminuir mientras que la de las señales más regulares (EGM NRA) aumenta. Se observa también la existencia de un punto en el que se cruzan las complejidades, tal y como ocurría con las señales EEG. La variabilidad de NRA es superior a RA.

**Tabla 7.38:** Caracterización de MSE 4 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.452	0.489	0.437	0.215	0.034	0.003
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.528, 0.545]	[0.555, 0.573]	[0.643, 0.662]	[0.773, 0.796]	[0.994, 1.021]	[1.378, 1.408]
CI RA	[0.991, 1.012]	[1.029, 1.049]	[1.143, 1.164]	[1.306, 1.328]	[1.548, 1.569]	[1.790, 1.814]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.994 (0.001)	0.983 (0.001)	0.958 (0.001)	0.800 (0.001)

**Tabla 7.39:** Caracterización de MSE 5 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.409	0.323	0.156	0.048	0.010	0.028
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.664, 0.684]	[0.706, 0.727]	[0.810, 0.834]	[0.957, 0.984]	[1.169, 1.199]	[1.430, 1.461]
CI RA	[1.176, 1.197]	[1.227, 1.248]	[1.353, 1.374]	[1.516, 1.538]	[1.690, 1.711]	[1.713, 1.745]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.994 (0.001)	0.984 (0.001)	0.942 (0.001)	0.642 (0.001)

**Tabla 7.40:** Caracterización de MSE 6 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.409	0.323	0.156	0.048	0.010	0.028
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.664, 0.684]	[0.706, 0.727]	[0.810, 0.834]	[0.957, 0.984]	[1.169, 1.199]	[1.430, 1.461]
CI RA	[1.176, 1.197]	[1.227, 1.248]	[1.353, 1.374]	[1.516, 1.538]	[1.690, 1.711]	[1.713, 1.745]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.994 (0.001)	0.984 (0.001)	0.942 (0.001)	0.642 (0.001)

**Tabla 7.41:** Caracterización de MSE 7 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.246	0.206	0.090	0.029	0.007	0.219
UT (p)	0.001	0.001	0.001	0.001	0.001	NaN
CI NRA	[0.727, 0.749]	[0.771, 0.794]	[0.881, 0.906]	[1.027, 1.055]	[1.221, 1.252]	[1.420, 1.452]
CI RA	[1.248, 1.269]	[1.305, 1.327]	[1.433, 1.455]	[1.585, 1.606]	[1.724, 1.745]	[1.649, 1.683]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.994 (0.001)	0.983 (0.001)	0.937 (0.001)	0.633 (0.001)

**Tabla 7.42:** Caracterización de MSE 8 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.147	0.125	0.055	0.017	0.005	0.500
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.787, 0.811]	[0.828, 0.853]	[0.941, 0.968]	[1.090, 1.119]	[1.265, 1.296]	[1.391, 1.424]
CI RA	[1.333, 1.354]	[1.376, 1.397]	[1.502, 1.524]	[1.637, 1.658]	[1.738, 1.758]	[1.579, 1.617]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.995 (0.001)	0.980 (0.001)	0.921 (0.001)	0.596 (0.001)

**Tabla 7.43:** Caracterización de MSE 9 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

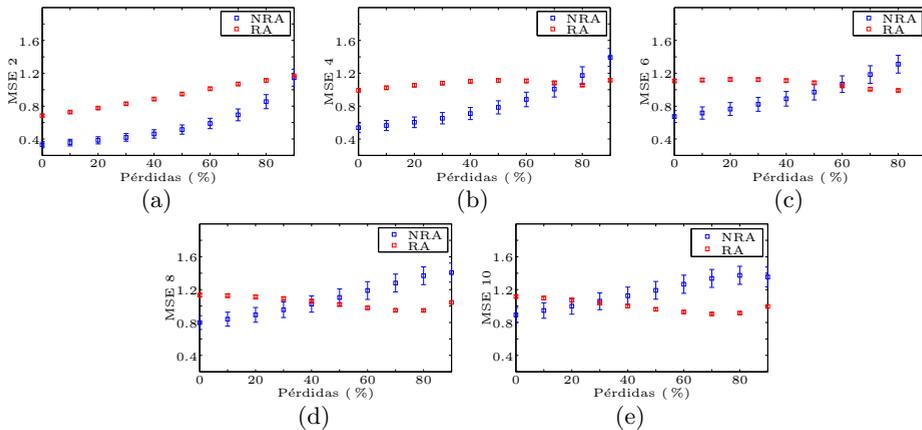
Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.140	0.082	0.036	0.012	0.005	0.500
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.835, 0.859]	[0.883, 0.908]	[0.997, 1.025]	[1.138, 1.168]	[1.301, 1.332]	[1.368, 1.402]
CI RA	[1.386, 1.408]	[1.439, 1.461]	[1.558, 1.579]	[1.675, 1.695]	[1.750, 1.771]	[1.550, 1.590]
$C_{xy}$ (p)	1.000(0.001)	0.997 (0.001)	0.992 (0.001)	0.971 (0.001)	0.896 (0.001)	0.610 (0.001)

**Tabla 7.44:** Caracterización de MSE 10 sobre EGM registrados en zonas no recomendadas para ablación (NRA) y recomendadas para ablación (RA) con pérdida de muestras aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.133	0.057	0.025	0.010	0.006	0.500
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.879, 0.904]	[0.932, 0.959]	[1.044, 1.073]	[1.176, 1.206]	[1.321, 1.352]	[1.337, 1.372]
CI RA	[1.451, 1.473]	[1.492, 1.513]	[1.601, 1.622]	[1.703, 1.723]	[1.749, 1.771]	[1.477, 1.518]
$C_{xy}$ (p)	1.000(0.001)	0.995 (0.001)	0.989 (0.001)	0.968 (0.001)	0.880 (0.001)	0.570 (0.001)

z

De nuevo no es recomendable usar escalas grandes pues ofrecen resultados de complejidad contradictorios con la literatura, otorgando mayor complejidad a señales más regulares. Escalas mayores a 4 no ofrecen resultados fiables para porcentaje de muestras perdidas altos.



**Figura 7.15:** Comportamiento de MSE cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en azul, y recomendadas para ablación (RA) en rojo, para distintas escalas de MSE. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10.

### 7.3.2 Pérdida uniforme de muestras

#### *Electroencefalogramas*

Las tablas 7.45 a 7.53 muestran la caracterización estadística de MSE para distintas escalas. Se trata de distribuciones no normales. Existen diferencias estadísticas entre las medianas de cada clase para cualquier escala y nivel de muestras eliminadas.

**Tabla 7.45:** Caracterización de MSE 2 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.688, 1.724]	[1.687, 1.723]	[1.694, 1.731]	[1.689, 1.727]	[1.687, 1.726]	[1.711, 1.753]
CI PT	[1.074, 1.136]	[1.075, 1.136]	[1.076, 1.139]	[1.069, 1.133]	[1.057, 1.124]	[1.056, 1.131]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.997(0.001)	0.992(0.001)	0.979(0.001)	0.946(0.001)

La figura 7.16 muestra el comportamiento de MSE en función del porcentaje de muestras que se han perdido para algunas de las escalas evaluadas. El comportamiento se ajusta a los casos anteriores de pérdidas uniformes, la complejidad de la señal se mantiene aproximadamente constante independientemente de las

**Tabla 7.46:** Caracterización de MSE 3 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) .

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.886, 1.917]	[1.885, 1.916]	[1.891, 1.921]	[1.889, 1.921]	[1.889, 1.923]	[1.912, 1.953]
CI PT	[1.322, 1.388]	[1.324, 1.389]	[1.326, 1.392]	[1.318, 1.387]	[1.303, 1.375]	[1.297, 1.378]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.997(0.001)	0.991(0.001)	0.975(0.001)	0.932(0.001)

**Tabla 7.47:** Caracterización de MSE 4 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) .

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.947, 1.971]	[1.947, 1.972]	[1.950, 1.975]	[1.947, 1.973]	[1.949, 1.977]	[1.968, 2.013]
CI PT	[1.475, 1.540]	[1.476, 1.541]	[1.477, 1.543]	[1.471, 1.538]	[1.454, 1.526]	[1.453, 1.534]
$C_{xy}$ (p)	1.000(0.001)	0.998(0.001)	0.995(0.001)	0.988(0.001)	0.967(0.001)	0.906(0.001)

**Tabla 7.48:** Caracterización de MSE 5 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) .

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.945, 1.964]	[1.944, 1.962]	[1.950, 1.969]	[1.948, 1.968]	[1.946, 1.971]	[1.953, 1.998]
CI PT	[1.574, 1.637]	[1.575, 1.637]	[1.579, 1.642]	[1.574, 1.639]	[1.554, 1.625]	[1.555, 1.640]
$C_{xy}$ (p)	1.000(0.001)	0.998(0.001)	0.993(0.001)	0.983(0.001)	0.953(0.001)	0.864(0.001)

**Tabla 7.49:** Caracterización de MSE 6 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ) .

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.911, 1.925]	[1.908, 1.924]	[1.912, 1.929]	[1.911, 1.929]	[1.915, 1.938]	[1.932, 1.976]
CI PT	[1.642, 1.702]	[1.643, 1.702]	[1.646, 1.707]	[1.639, 1.702]	[1.621, 1.690]	[1.630, 1.712]
$C_{xy}$ (p)	1.000(0.001)	0.993(0.001)	0.990(0.001)	0.974(0.001)	0.937(0.001)	0.807(0.001)

**Tabla 7.50:** Caracterización de MSE 7 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.858, 1.878]	[1.855, 1.875]	[1.859, 1.880]	[1.855, 1.878]	[1.851, 1.881]	[1.853, 1.909]
CI PT	[1.683, 1.742]	[1.685, 1.743]	[1.689, 1.748]	[1.681, 1.742]	[1.665, 1.733]	[1.684, 1.768]
$C_{xy}$ (p)	1.000(0.001)	0.989(0.001)	0.983(0.001)	0.965(0.001)	0.915(0.001)	0.747(0.001)

**Tabla 7.51:** Caracterización de MSE 8 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.807, 1.831]	[1.804, 1.828]	[1.809, 1.834]	[1.807, 1.834]	[1.806, 1.841]	[1.822, 1.884]
CI PT	[1.710, 1.767]	[1.711, 1.767]	[1.714, 1.771]	[1.710, 1.768]	[1.690, 1.755]	[1.718, 1.801]
$C_{xy}$ (p)	1.000(0.001)	0.987(0.001)	0.979(0.001)	0.961(0.001)	0.887(0.001)	0.719(0.001)

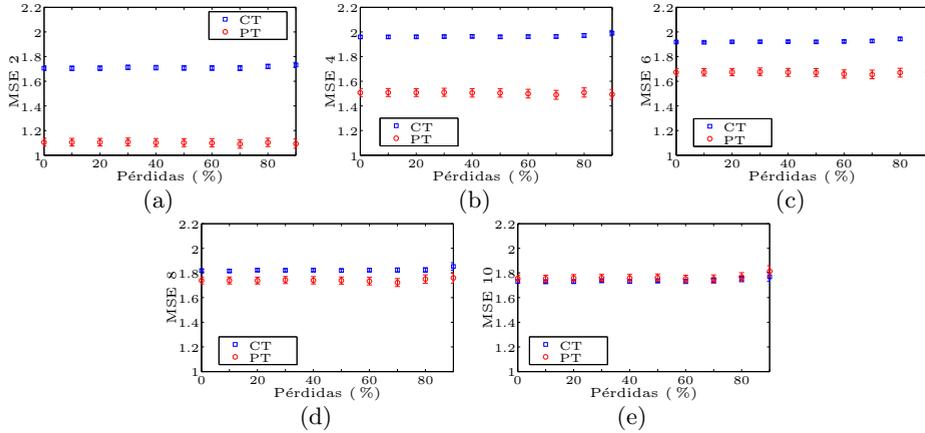
**Tabla 7.52:** Caracterización de MSE 9 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.764, 1.789]	[1.756, 1.781]	[1.765, 1.791]	[1.760, 1.790]	[1.757, 1.793]	[1.777, 1.841]
CI PT	[1.725, 1.779]	[1.723, 1.778]	[1.728, 1.784]	[1.724, 1.782]	[1.707, 1.772]	[1.732, 1.820]
$C_{xy}$ (p)	1.000(0.001)	0.981(0.001)	0.974(0.001)	0.954(0.001)	0.887(0.001)	0.669(0.001)

**Tabla 7.53:** Caracterización de MSE 10 cuando se eliminan muestras de un EEG de control (CT) o patológica (PT) de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[1.717, 1.744]	[1.715, 1.743]	[1.723, 1.751]	[1.721, 1.751]	[1.720, 1.758]	[1.732, 1.806]
CI PT	[1.727, 1.783]	[1.726, 1.782]	[1.732, 1.788]	[1.733, 1.792]	[1.719, 1.784]	[1.768, 1.859]
$C_{xy}$ (p)	1.000(0.001)	0.994(0.001)	0.982(0.001)	0.954(0.001)	0.884(0.001)	0.647(0.001)

muestras eliminadas. Se observa claramente el efecto que produce el escalado en la señal. Este aumenta progresivamente la complejidad de las señales patológicas, a priori, más regulares que las de control.



**Figura 7.16:** Comportamiento de MSE cuando se eliminan muestras de forma uniforme sobre una señal EEG control (CT) en azul o patológica (PT) en rojo. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10

A medida que aumenta la escala, la diferencia entre clases va disminuyendo de forma progresiva hasta prácticamente desaparecer en MSE-10, aunque estadísticamente se encuentren diferencias entre sus medianas en cualquier escala y para todos los valores de muestras eliminadas considerados.

### Señales RR

Las tablas 7.54 a 7.62 contienen la caracterización estadística de MSE para escalas de 2 a 10. En ella se aprecia validez y robustez de la medida al menos hasta alcanzar un 50 % de muestras perdidas en todas las escalas consideradas. La distribución de MSE no se corresponde con una normal pero existe separación estadística de las medianas de cada clase en todas las escalas y ante cualquier porcentaje de pérdida de muestras.

La figura 7.17 muestra la evolución de MSE en función del porcentaje de muestras perdidas. Como se ha observado en experimentos previos (ApEn, sec. 7.1.2 o SampEn, sec.7.2.2) los registros obtenidos sobre sujetos antes del tratamiento (Pre) presentan mayor complejidad que una vez el paciente ha sido tratado (Post). Esto sugiere que la medicación consigue reducir parte de los latidos ectópicos.

Como contrapunto a SampEn, MSE al igual que ApEn sí segmenta las clases en todo el rango muestras perdidas considerado y para todas las escalas. La medida

**Tabla 7.54:** Caracterización de MSE 2 con pérdida de muestras de forma uniforme en registros RR antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.002	0.002	0.004	0.002	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.932, 1.005]	[0.928, 1.001]	[0.945, 1.019]	[0.957, 1.036]	[0.970, 1.051]	[1.044, 1.141]
CI Pre	[1.127, 1.211]	[1.135, 1.221]	[1.145, 1.230]	[1.165, 1.249]	[1.188, 1.278]	[1.310, 1.419]
$C_{xy}$ (p)	1.000(0.001)	0.962(0.001)	0.907(0.001)	0.812(0.001)	0.743(0.001)	0.537(0.001)

**Tabla 7.55:** Caracterización de MSE 3 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.018	0.016	0.009	0.019	0.015	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.893, 0.962]	[0.900, 0.970]	[0.902, 0.972]	[0.917, 0.990]	[0.937, 1.012]	[1.037, 1.134]
CI Pre	[1.060, 1.134]	[1.061, 1.135]	[1.068, 1.144]	[1.104, 1.181]	[1.126, 1.209]	[1.274, 1.381]
$C_{xy}$ (p)	1.000(0.001)	0.968(0.001)	0.917(0.001)	0.864(0.001)	0.778(0.001)	0.585(0.001)

**Tabla 7.56:** Caracterización de MSE 4 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.005	0.003	0.003	0.003	0.008	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.901, 0.970]	[0.903, 0.972]	[0.920, 0.991]	[0.933, 1.007]	[0.964, 1.043]	[1.092, 1.200]
CI Pre	[1.066, 1.138]	[1.073, 1.145]	[1.086, 1.161]	[1.116, 1.189]	[1.148, 1.228]	[1.304, 1.413]
$C_{xy}$ (p)	1.000(0.001)	0.979(0.001)	0.931(0.001)	0.867(0.001)	0.779(0.001)	0.545(0.001)

**Tabla 7.57:** Caracterización de MSE 5 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.002	0.002	0.002	0.003	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.937, 1.007]	[0.941, 1.011]	[0.957, 1.030]	[0.967, 1.043]	[0.996, 1.076]	[1.145, 1.257]
CI Pre	[1.092, 1.162]	[1.098, 1.168]	[1.117, 1.190]	[1.153, 1.226]	[1.182, 1.263]	[1.367, 1.486]
$C_{xy}$ (p)	1.000(0.001)	0.982(0.001)	0.933(0.001)	0.873(0.001)	0.783(0.001)	0.531(0.001)

**Tabla 7.58:** Caracterización de MSE 6 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.002	0.002	0.071	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.944, 1.016]	[0.951, 1.023]	[0.962, 1.036]	[0.976, 1.054]	[1.013, 1.099]	[1.153, 1.271]
CI Pre	[1.126, 1.199]	[1.129, 1.201]	[1.138, 1.213]	[1.183, 1.259]	[1.214, 1.296]	[1.386, 1.508]
$C_{xy}$ (p)	1.000(0.001)	0.975(0.001)	0.941(0.001)	0.871(0.001)	0.782(0.001)	0.463(0.001)

**Tabla 7.59:** Caracterización de MSE 7 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.002	0.001	0.005	0.001	0.012	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.961, 1.033]	[0.962, 1.034]	[0.979, 1.052]	[0.992, 1.069]	[1.024, 1.108]	[1.112, 1.230]
CI Pre	[1.137, 1.210]	[1.140, 1.214]	[1.150, 1.224]	[1.202, 1.276]	[1.257, 1.342]	[1.348, 1.470]
$C_{xy}$ (p)	1.000(0.001)	0.979(0.001)	0.939(0.001)	0.871(0.001)	0.780(0.001)	0.471(0.001)

**Tabla 7.60:** Caracterización de MSE 8 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

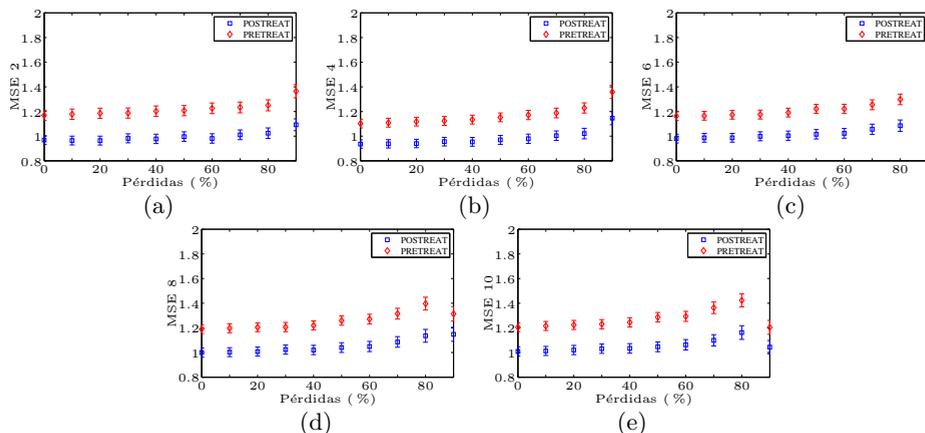
Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.007	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.963, 1.038]	[0.966, 1.040]	[0.986, 1.062]	[1.000, 1.081]	[1.042, 1.130]	[1.090, 1.206]
CI Pre	[1.153, 1.226]	[1.160, 1.234]	[1.168, 1.244]	[1.221, 1.298]	[1.271, 1.359]	[1.254, 1.374]
$C_{xy}$ (p)	1.000(0.001)	0.975(0.001)	0.933(0.001)	0.864(0.001)	0.763(0.001)	0.456(0.001)

**Tabla 7.61:** Caracterización de MSE 9 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.003	0.023	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.004
CI Post	[0.969, 1.043]	[0.975, 1.050]	[0.993, 1.069]	[0.998, 1.078]	[1.048, 1.138]	[1.075, 1.189]
CI Pre	[1.160, 1.234]	[1.169, 1.244]	[1.181, 1.257]	[1.229, 1.306]	[1.289, 1.380]	[1.195, 1.310]
$C_{xy}$ (p)	1.000(0.001)	0.981(0.001)	0.933(0.001)	0.861(0.001)	0.754(0.001)	0.386(0.001)

**Tabla 7.62:** Caracterización de MSE 10 con pérdida de muestras de forma uniforme en registros RR de sujetos antes (Pre) y después (Post) del tratamiento con encainida. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.012	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[1.168, 1.242]	[1.178, 1.253]	[1.191, 1.268]	[1.247, 1.326]	[1.315, 1.411]	[1.147, 1.262]
CI Pre	[0.972, 1.047]	[0.975, 1.050]	[0.992, 1.069]	[1.006, 1.087]	[1.053, 1.146]	[0.988, 1.098]
$C_{xy}$ (p)	1.000(0.001)	0.987(0.001)	0.939(0.001)	0.865(0.001)	0.762(0.001)	0.391(0.001)



**Figura 7.17:** Comportamiento de MSE sobre registros RR obtenidos en pacientes tratados con encainida, antes (Pre) y después (Post) del tratamiento cuando se han perdido muestras de forma uniforme. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10

se mantiene estable en todo el rango de pérdidas. La entropía varía de una escala a otra pero se mantiene en torno a 1.2 en Pre y 1 en Post y la distancia entre clases se mantiene relativamente estable, contrario al comportamiento en EEG.

En escalas altas, MSE 8, y porcentaje de pérdidas elevados (80–90 %), aparece un punto de inflexión. Éste se asemeja al encontrado en ApEn. SampEn no presenta el problema de los *auto-matches* pero presenta un sesgo para registros muy cortos ( $N < 100$ ).

**Electrogramas**

Las tablas 7.63 a 7.71 muestran la caracterización estadística de MSE para las escalas 2 a 10 sobre EGM registrados en regiones recomendadas para ablación (RA) y regiones no recomendadas para ablación (NRA) cuando se han eliminado muestras de forma uniforme.

**Tabla 7.63:** Caracterización de MSE 2 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.098	0.101	0.169	0.196	0.126	0.055
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.325, 0.336]	[0.327, 0.338]	[0.331, 0.342]	[0.335, 0.346]	[0.341, 0.353]	[0.389, 0.404]
CI RA	[0.640, 0.656]	[0.641, 0.657]	[0.635, 0.650]	[0.626, 0.641]	[0.618, 0.634]	[0.644, 0.666]
$C_{xy}$ (p)	1.000(0.001)	1.000 (0.001)	0.997 (0.001)	0.980 (0.001)	0.935 (0.001)	0.768 (0.001)

**Tabla 7.64:** Caracterización de MSE 3 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.238	0.262	0.317	0.296	0.138	0.053
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.429, 0.443]	[0.433, 0.447]	[0.439, 0.454]	[0.440, 0.456]	[0.456, 0.473]	[0.521, 0.541]
CI RA	[0.838, 0.857]	[0.839, 0.858]	[0.832, 0.851]	[0.824, 0.843]	[0.819, 0.840]	[0.884, 0.913]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.997 (0.001)	0.982 (0.001)	0.936 (0.001)	0.762 (0.001)

**Tabla 7.65:** Caracterización de MSE 4 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.452	0.426	0.463	0.372	0.142	0.123
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.528, 0.545]	[0.526, 0.543]	[0.532, 0.550]	[0.539, 0.558]	[0.556, 0.575]	[0.656, 0.683]
CI RA	[0.991, 1.012]	[0.983, 1.004]	[0.987, 1.008]	[0.977, 0.999]	[0.975, 0.998]	[1.062, 1.096]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.997 (0.001)	0.979 (0.001)	0.932 (0.001)	0.734 (0.001)

Estas tablas muestran en función del porcentaje de muestras perdidas, los intervalos de confianza (CI) de MSE para cada grupo, la normalidad (JB) y separabilidad (UT) y el coeficiente de correlación cruzada de la entropía de las series sin pérdidas con respecto a la entropía de las series con pérdidas.

**Tabla 7.66:** Caracterización de MSE 5 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.441	0.492	0.425	0.412	0.133	0.093
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.598, 0.617]	[0.602, 0.621]	[0.605, 0.625]	[0.613, 0.633]	[0.635, 0.657]	[0.772, 0.804]
CI RA	[1.097, 1.118]	[1.094, 1.116]	[1.091, 1.112]	[1.083, 1.106]	[1.092, 1.117]	[1.221, 1.258]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.997 (0.001)	0.980 (0.001)	0.933 (0.001)	0.739 (0.001)

**Tabla 7.67:** Caracterización de MSE 6 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.409	0.394	0.438	0.318	0.141	0.130
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.664, 0.684]	[0.669, 0.689]	[0.674, 0.695]	[0.684, 0.705]	[0.710, 0.733]	[0.881, 0.919]
CI RA	[1.176, 1.197]	[1.178, 1.199]	[1.174, 1.196]	[1.174, 1.198]	[1.185, 1.212]	[1.372, 1.419]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.997 (0.001)	0.981 (0.001)	0.926 (0.001)	0.722 (0.001)

**Tabla 7.68:** Caracterización de MSE 7 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.246	0.254	0.336	0.416	0.244	0.350
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.727, 0.749]	[0.732, 0.754]	[0.739, 0.761]	[0.750, 0.774]	[0.779, 0.804]	[1.023, 1.067]
CI RA	[1.248, 1.269]	[1.250, 1.271]	[1.252, 1.275]	[1.249, 1.274]	[1.269, 1.297]	[1.448, 1.495]
$C_{xy}$ (p)	1.000(0.001)	1.000 (0.001)	0.996 (0.001)	0.977 (0.001)	0.924 (0.001)	0.635 (0.001)

**Tabla 7.69:** Caracterización de MSE 8 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

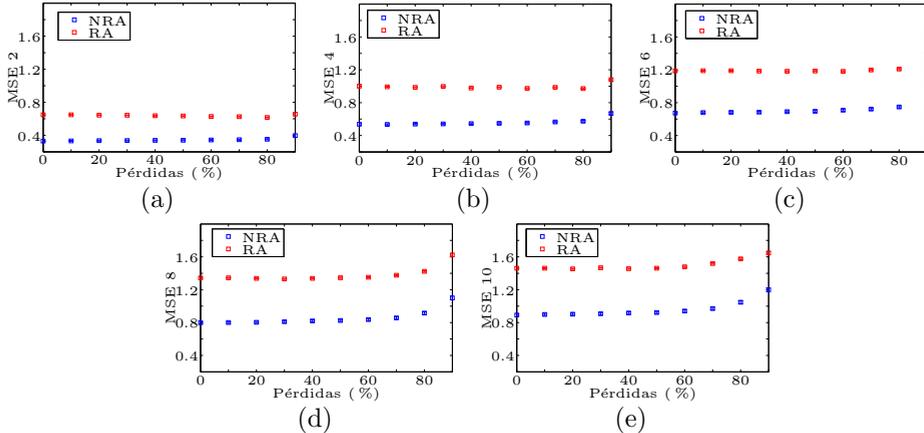
Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.147	0.204	0.142	0.165	0.143	0.237
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.787, 0.811]	[0.788, 0.812]	[0.798, 0.822]	[0.813, 0.838]	[0.844, 0.871]	[1.077, 1.122]
CI RA	[1.333, 1.354]	[1.333, 1.355]	[1.320, 1.342]	[1.334, 1.358]	[1.361, 1.391]	[1.596, 1.651]
$C_{xy}$ (p)	1.000(0.001)	0.998 (0.001)	0.994 (0.001)	0.979 (0.001)	0.921 (0.001)	0.679 (0.001)

**Tabla 7.70:** Caracterización de MSE 9 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.140	0.119	0.101	0.189	0.124	0.145
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.835, 0.859]	[0.845, 0.869]	[0.855, 0.880]	[0.864, 0.890]	[0.906, 0.935]	[1.151, 1.197]
CI RA	[1.386, 1.408]	[1.390, 1.411]	[1.393, 1.416]	[1.385, 1.410]	[1.424, 1.453]	[1.576, 1.628]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.995 (0.001)	0.976 (0.001)	0.925 (0.001)	0.567 (0.001)

**Tabla 7.71:** Caracterización de MSE 10 sobre EGM registrados en zonas recomendadas para ablación (RA) o no recomendadas para ablación (NRA) con pérdida de muestras de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.133	0.100	0.147	0.217	0.205	0.143
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI NRA	[0.879, 0.904]	[0.886, 0.912]	[0.894, 0.921]	[0.909, 0.936]	[0.956, 0.986]	[1.175, 1.223]
CI RA	[1.451, 1.473]	[1.450, 1.472]	[1.455, 1.479]	[1.448, 1.474]	[1.503, 1.534]	[1.623, 1.674]
$C_{xy}$ (p)	1.000(0.001)	0.996 (0.001)	0.995 (0.001)	0.976 (0.001)	0.925 (0.001)	0.525 (0.001)



**Figura 7.18:** Comportamiento de MSE cuando se eliminan muestras de forma uniforme sobre una EGM registrados en regiones no recomendadas para ablación (NRA) en azul y recomendadas para ablación (RA) en rojo, para distintas escalas de MSE. (a) MSE 2, (b) MSE 4, (c) MSE 6, (d) MSE 8 y (e) MSE 10

La figura 7.18 muestra, para las escalas pares, el comportamiento de la entropía en ambas clases de registros EGM en función del porcentaje de muestras que se han perdido de forma uniforme en los registros. En primer lugar se observa una

entropía constante hasta alcanzar pérdidas elevadas en escalas altas. Contrario a lo que se obtenía en EEG, la distancia entre clases aumenta con la escala, y la complejidad de ambas clases aumenta con la escala.

### 7.3.3 Señales de corta duración

Con restricciones asociadas a la longitud de las señales analizadas  $N < 300$ , no es aconsejable aplicar el análisis en escalas superiores a 3 ya que se estaría aplicando SampEn a registros de longitudes inferiores a 100 muestras.

#### *Glucemia en pacientes críticos*

Al igual que sucede con SampEn, en escalas más altas, tampoco se consiguen obtener diferencias significativas entre clases para ninguna de las combinaciones paramétricas consideradas.

#### *Glucemia ambulatoria*

De forma similar a SampEn, el máximo de la distribución se produce en la misma combinación de parámetros en las tres variables simultáneamente.

En los resultados presentados en la tabla 7.72, se observa que para una escala 2 tampoco es posible discernir entre sujetos NDM y con síndrome metabólico (MS), algo que resultaría de especial interés para considerar el grado de susceptibilidad del sujeto a desarrollar diabetes.

**Tabla 7.72:** Estudio estadístico de la variabilidad de MSE 2 en función de los parámetros  $(m, r)$  para los grupos de sujetos DM, MS y NDM. En negrita se identifican las distribuciones donde no se obtiene separación entre clases. UT<sub>1</sub>: Segmentación estadística entre medianas DM y NDM, UT<sub>2</sub>: entre MS y NDM y UT<sub>3</sub>: entre DM y MS.

$(m, r)$	CI NDM	CI MS	CI DM	JB Test	U Test <sub>1</sub>	U Test <sub>2</sub>	U Test <sub>3</sub>
(1,0.1)	[2.630,2.911]	[2.557,2.802]	[2.374,2.666]	0.500	0.032	<b>0.404</b>	<b>0.113</b>
(2,0.1)	[1.356,1.777]	[1.207,1.616]	[0.832,1.162]	0.457	0.009	<b>0.311</b>	0.012
(3,0.1)	[1.150,1.724]	[1.006,1.391]	[0.649,0.930]	0.008	0.006	<b>0.283</b>	0.018
(4,0.1)	[0.994,1.608]	[0.890,1.560]	[0.541,0.843]	0.030	0.007	<b>0.592</b>	0.008
(5,0.1)	[0.844,1.433]	[0.795,1.386]	[0.438,0.971]	0.269	<b>0.060</b>	<b>0.951</b>	0.027

NDM es separable de MS en escala 3 (tabla 7.73), pero mientras que se consigue discernir entre NDM y MS, no se consigue discernir entre DM y NDM, además de  $MSE_{MS} > MSE_{NDM}$ , contrario a la literatura [55].

En este caso, existen variables en las cuales MSE no es capaz de discernir entre sujetos no diabéticos (NDM) y diabéticos (DM), pero se dan en dimensiones  $m$  altas (resultados no incluidos).

**Tabla 7.73:** Estudio estadístico de la variabilidad de MSE 3 en función de los parámetros  $(m, r)$  para los grupos de sujetos DM, MS y NDM. En negrita se remarcán las distribuciones que no segmentan los grupos

$(m, r)$	CI NDM	CI MS	CI DM	JB Test	U Test <sub>1</sub>	U Test <sub>2</sub>	U Test <sub>3</sub>
(1,0.1)	[2.622,2.854]	[2.577,2.821]	[2.435,2.672]	0.500	0.047	<b>0.881</b>	<b>0.088</b>
(2,0.1)	[1.573,1.970]	[1.384,1.828]	[1.109,1.446]	0.500	0.004	<b>0.404</b>	0.036
(3,0.1)	[1.392,2.035]	[1.167,1.717]	[0.795,1.158]	0.035	0.003	<b>0.387</b>	0.026
(4,0.2)	[0.695,1.029]	[0.631,1.173]	[0.362,0.624]	0.145	0.009	<b>0.858</b>	<b>0.066</b>
(5,0.1)	[0.502,0.979]	[0.830,1.890]	[0.488,1.242]	0.081	<b>0.808</b>	0.045	<b>0.139</b>

Considerando una combinación de parámetros que permitan realizar al menos dos diferencias significativas, presenten distribuciones normales y mínima variabilidad en la entropía, se consideraría para ambas escalas (2,0.1).

### *Dinámicas de movilidad*

Las tablas 7.74 y 7.75 muestran las caracterizaciones estadísticas de MSE en escalas 2 y 3, respectivamente. Se muestran únicamente las combinaciones paramétricas que permiten establecer alguna segmentación entre la clase de control (CT) y alguna de las clases con registros patológicos.

**Tabla 7.74:** Caracterización estadística de MSE 2 en función de los parámetros  $(m, r)$  para los grupos considerados en los registros de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA). UT<sub>1</sub>, UT<sub>2</sub> y UT<sub>3</sub> muestran la probabilidad de segmentación entre LSA, PD y HD respecto a CT.

$(m, r)$	CI LSA	CI PD	CI HD	CI CT	UT <sub>1</sub>	UT <sub>2</sub>	UT <sub>3</sub>
(2,0.10)	[1.877,2.204]	[2.139,2.430]	[2.098,2.645]	[2.044,2.392]	<b>0.181</b>	<b>0.502</b>	0.032
(2,0.20)	[1.219,1.520]	[1.392,1.681]	[1.493,1.888]	[1.433,1.737]	0.041	<b>0.540</b>	<b>0.115</b>
(3,0.20)	[1.148,1.403]	[1.385,1.761]	[1.448,1.875]	[1.394,1.673]	0.013	<b>0.514</b>	<b>0.075</b>
(4,0.30)	[0.820,1.465]	[0.987,1.353]	[1.193,1.570]	[0.966,1.441]	<b>0.539</b>	<b>0.937</b>	0.029
(4,0.40)	[0.596,1.010]	[0.704,0.956]	[0.921,1.197]	[0.740,1.029]	<b>0.283</b>	<b>0.737</b>	0.018
(5,0.50)	[0.398,0.932]	[0.549,0.799]	[0.728,0.966]	[0.599,0.841]	<b>0.100</b>	<b>0.580</b>	0.043

En ellas se observan cómo los resultados empeoran considerablemente con respecto a los proporcionados por ApEn o SampEn ya que en contadas ocasiones se consigue ofrecer algún tipo de segmentación y sólo con respecto a una única variable. MSE no sería adecuado para analizar estas señales, ni ninguna de las consideradas con esta longitud.

**Tabla 7.75:** Caracterización estadística de MSE 3 en función de los parámetros  $(m, r)$  para los grupos considerados en los registros de dinámicas de movilidad: control (CT), Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA).  $UT_1$ ,  $UT_2$  y  $UT_3$  muestran la probabilidad de segmentación entre LSA, PD y HD respecto a CT.

$(m, r)$	CI LSA	CI PD	CI HD	CI CT	$UT_1$	$UT_2$	$UT_3$
(3,0.10)	[1.475,1.875]	[1.618,2.174]	[1.631,2.271]	[2.033,2.438]	0.001	<b>0.086</b>	<b>0.270</b>
(5,0.20)	[0.831,1.260]	[1.139,1.820]	[0.851,1.608]	[0.977,1.164]	<b>0.871</b>	0.045	<b>0.364</b>

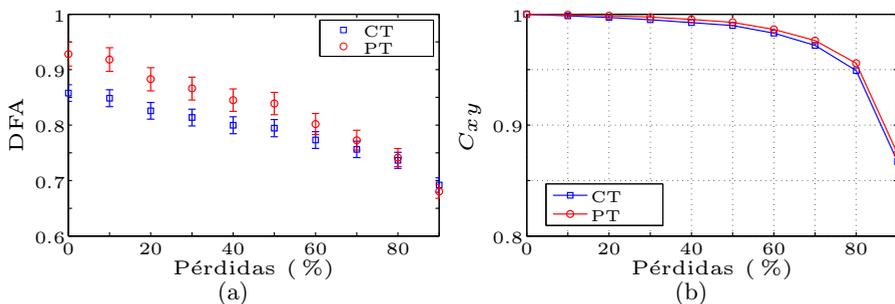
## 7.4 Análisis de fluctuaciones sin tendencias (DFA)

### 7.4.1 Pérdida aleatoria de muestras

Se ha aplicado el algoritmo clásico de DFA, descrito en [8]. La secuencia de enventanado de longitud máxima es de  $N/4$ , y orden de integración 1.

#### *Electroencefalogramas*

La figura 7.19 muestra la evolución del coeficiente de DFA en función del ratio de muestras perdidas que experimenta la señal, tanto para el grupo de control (CT) como el de registros patológicos (PT).



**Figura 7.19:** (a) Comportamiento de DFA ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma uniforme sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la señal original y de las series con pérdidas en función del ratio de muestras eliminadas.

La tabla 7.76 muestra la caracterización estadística de DFA, las distribuciones no se ajustan a una normal pero existe separación entre medianas y la estabilidad de la medida y validez clínica se mantiene a lo largo de todo el rango de pérdidas.

Teniendo en cuenta que esta medida representa la regularidad y no la complejidad, podemos establecer que el comportamiento es el mismo que en casos anteriores.

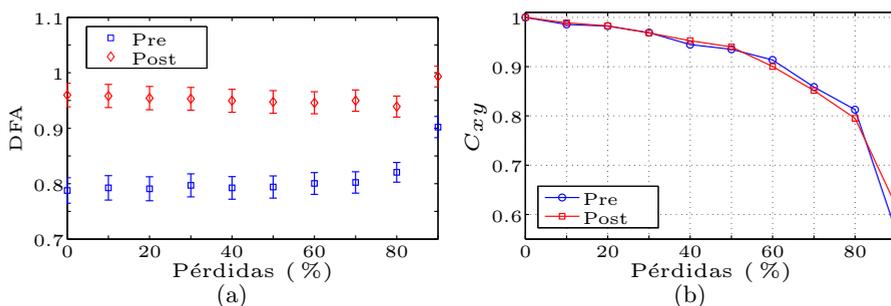
**Tabla 7.76:** Caracterización de DFA para señales EEG de control (CT) y patológicas (PT) cuando se pierden muestras de forma aleatoria. Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[0.843, 0.873]	[0.833, 0.864]	[0.799, 0.829]	[0.779, 0.810]	[0.741, 0.771]	[0.679, 0.705]
CI PT	[0.907, 0.949]	[0.897, 0.940]	[0.845, 0.887]	[0.819, 0.859]	[0.755, 0.791]	[0.668, 0.693]
$C_{xy}$ (p)	1.000(0.001)	0.999(0.001)	0.996(0.001)	0.989(0.001)	0.962(0.001)	0.828(0.001)

CT presenta menor regularidad (mayor complejidad) que PT, y la regularidad disminuye a medida que aumenta el número de muestras perdidas.

### Señales RR

La figura 7.20 muestra el coeficiente de DFA para señales RR provenientes de sujetos pre-medición (Pre) y post-medición (Post) tratados con encainida y su coeficiente de correlación cruzada.



**Figura 7.20:** Comportamiento de DFA cuando se eliminan muestras de forma aleatoria sobre una señal RR en sujetos antes (Pre) en rojo y después (Post) en azul, del tratamiento con encainida. (a) Evolución en función del ratio de muestras eliminadas y (d) coeficiente de correlación cruzada de la complejidad de cada grupo en las series con pérdidas con respecto a la complejidad de las series originales.

La tabla 7.77 muestra la caracterización estadística de DFA. En ella se determinan distribuciones no normales y separación estadística entre clases, junto con estabilidad y validez clínica de la medida hasta el 70 % de pérdidas.

Por un lado, se observa un comportamiento similar a los obtenidos para las mismas señales en capítulos anteriores, donde las señales Pre presentan menor regularidad que las señales Post. Por otro lado, el comportamiento en función del porcentaje de muestras perdidas no es coherente. Con una reducción de muestras aleatoria, la

**Tabla 7.77:** Caracterización de DFA sobre registros RR antes (Pre) y después (Post) del tratamiento con encainida cuando se pierden muestras de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

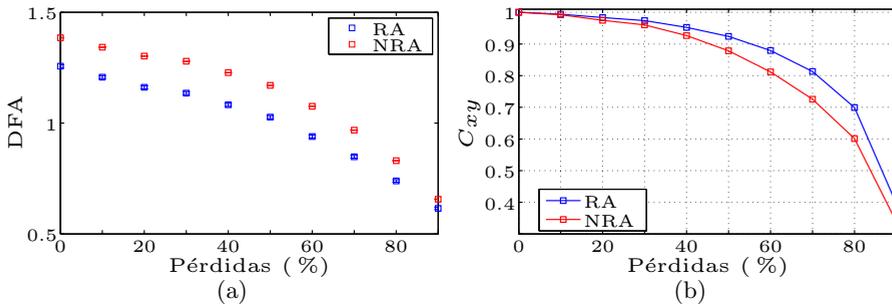
Eliminación	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI Post	[0.938,0.981]	[0.941,0.985]	[0.946,0.992]	[0.946,0.996]	[0.958,1.015]	[0.955,1.019]
CI Pre	[0.765,0.811]	[0.763,0.810]	[0.757,0.808]	[0.746,0.802]	[0.718,0.779]	[0.803,0.873]
$C_{xy}$ (p)	1.000 (0.001)	0.987 (0.001)	0.967 (0.001)	0.909 (0.001)	0.836 (0.001)	0.608 (0.001)

complejidad ha aumentado. En este caso se mantiene constante independientemente del diemzado. Este comportamiento sugiere que estos registros RR matienen sus propiedades de forma más estable que EEG.

Como prácticamente en cada estudio, las distribuciones no son normales, pero existe diferenciación estadística entre las medianas de las clases consideradas. La validez se mantiene hasta el 70 % de pérdidas.

### Electrogramas

La figura 7.21 muestra la evolución de DFA en función del ratio de muestras perdidas para cada clase. De nuevo se observa cómo la regularidad es mayor para señales registradas en áreas no recomendadas para ablación. Al igual que en casos anteriores, la regularidad de la señal disminuye con el porcentaje de pérdidas y la validez clínica está comprometida por encima del 60 % de muestras perdidas para las clases individuales.



**Figura 7.21:** (a) Comportamiento de DFA cuando se eliminan muestras de forma aleatoria sobre EGM registrados en áreas no recomendadas para ablación (NRA) en rojo y recomendadas para ablación (RA) en azul. (b) Coeficiente de correlación cruzada entre la señal original y las señales con pérdida de muestras en función del porcentaje de muestras eliminadas.

La tabla 7.78 muestra la caracterización estadística de DFA sobre señales EGM registradas en áreas recomendadas para ablación (RA) y no recomendadas para ablación (NRA). Como en casos anteriores, las distribuciones presentan normalidad excepto para el 90 % de pérdidas. La segmentación entre clases se mantiene en todo el rango y la validez clínica se mantiene en conjunto hasta el 70 % de muestras eliminadas.

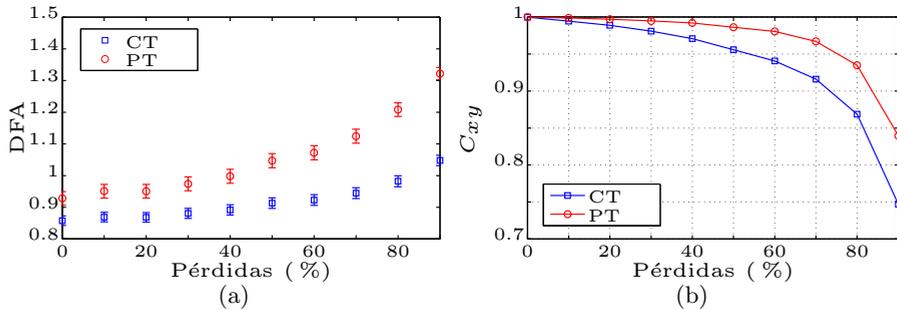
**Tabla 7.78:** Caracterización de DFA sobre EGM registrados en regiones no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando las muestras se eliminan de forma aleatoria. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.457	0.474	0.490	0.414	0.141	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI RA	[1.250, 1.263]	[1.200, 1.214]	[1.129, 1.142]	[1.020, 1.034]	[0.842, 0.854]	[0.610, 0.620]
CI NRA	[1.380, 1.390]	[1.337, 1.347]	[1.273, 1.284]	[1.164, 1.175]	[0.962, 0.974]	[0.651, 0.663]
$C_{xy}$ (p)	1.000(0.001)	0.996 (0.001)	0.979 (0.001)	0.932 (0.001)	0.829 (0.001)	0.436 (0.001)

## 7.4.2 Pérdida uniforme de muestras

### *Electroencefalogramas*

La figura 7.22 muestra la evolución del coeficiente de escalado en función del porcentaje de muestra eliminadas en las señales EEG de control (CT) y patológicas (PT).



**Figura 7.22:** (a) Comportamiento de DFA ( $\mu \pm 2\sigma$ ) cuando se eliminan muestras de forma aleatoria sobre una señal EEG de control (CT) en azul y patológica (PT) en rojo. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre las regularidades de las series originales y las series con pérdidas.

En las medidas de entropía se ha observado que ésta se mantenía más o menos constante. En este caso, la regularidad de la señal aumenta en función de la pérdida

de muestras. PT presenta a su vez mayor regularidad que CT, algo que se ha venido observando con anterioridad.

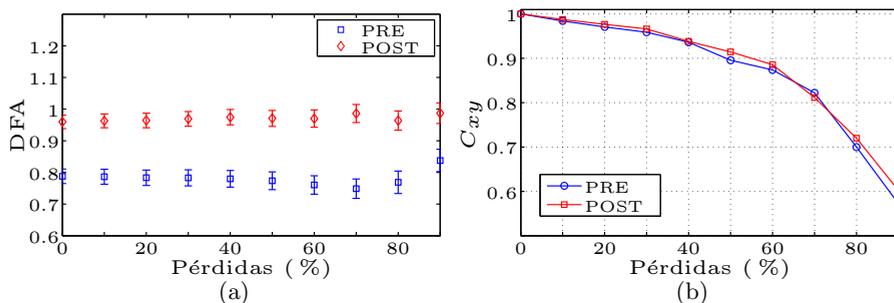
La tabla 7.79 presenta la caracterización estadística de DFA para algunos de los ratios considerados. DFA no presenta una distribución normal pero permite clasificar entre CT y PT ( $p=0.001$ ). Los intervalos de confianza no presentan solapamiento y la validez de la medida se mantiene hasta un 70 % de muestras perdidas.

**Tabla 7.79:** Caracterización de DFA para señales EEG de control (CT) y patológicos (PT) con pérdidas de forma uniforme. Se muestran los test estadísticos de normalidad (JB) y separabilidad (UT), los intervalos de confianza (CI) y el coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0 %	10 %	30 %	50 %	70 %	90 %
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI CT	[0.843, 0.873]	[0.854, 0.884]	[0.865, 0.896]	[0.897, 0.930]	[0.928, 0.962]	[1.032, 1.064]
CI PT	[0.907, 0.949]	[0.929, 0.972]	[0.952, 0.996]	[1.025, 1.070]	[1.102, 1.147]	[1.303, 1.341]
$C_{xy}$ (p)	1.000(0.001)	0.998(0.001)	0.991(0.001)	0.969(0.001)	0.926(0.001)	0.735(0.001)

### Señales RR

La figura 7.23 muestra la evolución del coeficiente de escalado en función del porcentaje de muestras eliminadas en los registros RR. Como se viene observando siempre que se ha producido una pérdida uniforme para SampEn o ApEn la regularidad ha permanecido más o menos constante dentro de unos límites. Aquí se observa lo mismo, y la medida se mantiene estable hasta pérdidas del 70 % ( $C_{xy} > 0.8$ ). Este comportamiento es opuesto al observado en EEG, donde la regularidad de la señal aumentaba con el número de muestras eliminadas en la señal.



**Figura 7.23:** (a) Comportamiento de DFA cuando se eliminan muestras de forma uniforme sobre una señal RR antes (PRE) en azul y después (POST) en rojo tratados del tratamiento con encanida. (b) coeficiente de correlación cruzada de la entropía de cada grupo en las series con pérdida de muestras respecto a la complejidad de la serie original.

La caracterización estadística de DFA se muestra en la tabla 7.80. Se constata la separabilidad de las clases ( $p=0.001$ ) y la no normalidad de las distribuciones. Los intervalos de confianza (CI) de las clases apenas sufren variación cuando se incrementa el número de muestras eliminadas y no presentan solapamiento.

**Tabla 7.80:** Caracterización de DFA sobre registros RR antes (PRE) y después (POST) del tratamiento con encainida cuando las muestras se eliminan de forma uniforme. Se muestran estadísticos de normalidad (JB) y separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.001	0.001	0.001	0.001	0.001	0.001
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI POST	[0.938,0.981]	[0.937,0.979]	[0.933,0.974]	[0.927,0.968]	[0.930,0.969]	[0.975,1.012]
CI PRE	[0.765,0.811]	[0.770,0.814]	[0.776,0.818]	[0.774,0.814]	[0.783,0.821]	[0.883,0.921]
$C_{xy}$ (p)	1.000 (0.001)	0.989 (0.001)	0.971 (0.001)	0.943 (0.001)	0.869 (0.001)	0.614 (0.001)

### Electrogramas

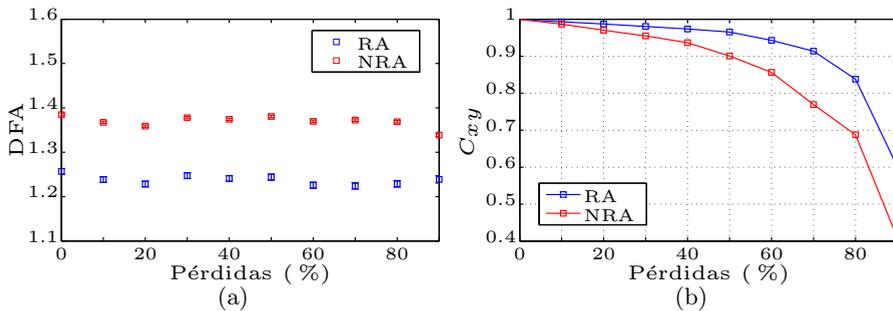
La tabla 7.81 muestra la caracterización estadística del coeficiente de escalado para señales EGM registradas en regiones indicadas para ablación (RA) o no recomendadas para ablación (NRA). Como se ha observado en otras medidas de entropía, las distribuciones generalmente son normales y las clases están separadas. Los intervalos de confianza no presentan solapamiento y la medida presenta validez clínica hasta un 70 % de pérdidas.

**Tabla 7.81:** Caracterización de DFA sobre EGM registrados en regiones no recomendadas para ablación (NRA) y recomendadas para ablación (RA) cuando la pérdida de muestras es de forma uniforme. Se muestran estadísticos de normalidad (JB), separabilidad (UT), intervalos de confianza (CI) y coeficiente de correlación cruzada ( $C_{xy}$ ).

Pérdidas	0%	10%	30%	50%	70%	90%
JB (p)	0.057	0.048	0.048	0.069	0.060	0.046
UT (p)	0.001	0.001	0.001	0.001	0.001	0.001
CI RA	[1.380, 1.390]	[1.362, 1.373]	[1.373, 1.383]	[1.375, 1.386]	[1.367, 1.378]	[1.332, 1.346]
CI NRA	[1.250, 1.263]	[1.232, 1.245]	[1.241, 1.254]	[1.237, 1.251]	[1.217, 1.231]	[1.230, 1.247]
$C_{xy}$ (p)	1.000(0.001)	0.999 (0.001)	0.996 (0.001)	0.983 (0.001)	0.938 (0.001)	0.708 (0.001)

La figura 7.24 muestra, por un lado, la evolución del coeficiente de escalado de DFA en función del ratio de muestras eliminadas en la señales EGM. Por otro lado, muestra el coeficiente de correlación cruzada entre la entropía de la series originales y las series con muestras eliminadas.

Se aprecia el mismo comportamiento que en casos anteriores, RA presenta menor regularidad que NRA y ésta se mantiene aproximadamente constante en todo el rango de pérdida de muestras.  $C_{xy}$  presenta validez clínica hasta el 60 % de pérdidas en ambos grupos.



**Figura 7.24:** (a) Comportamiento de DFA cuando se eliminan muestras de forma uniforme sobre EGM registrados en regiones no recomendadas para ablación (NRA) en rojo y recomendadas para ablación (RA) en azul. (b) Coeficiente de correlación cruzada ( $C_{xy}$ ) entre la regularidad de la señal original y la de las versiones con pérdidas en función del porcentaje de muestras eliminadas.

### 7.4.3 Señales de corta duración

En estas señales, se han considerado variaciones con respecto al algoritmo clásico de orden 1 y ventanas sin solapamiento. Se ha considerado aplicar distintos órdenes de integración ( $l$ ) hasta alcanzar la restricción impuesta por [88] donde  $\alpha \leq l$ . Se ha considerado el uso de ventana deslizante [50] y múltiples coeficientes de escalado [45] en caso necesario mediante el método expuesto para estimar el cross-over de forma autónoma no supervisada expuesto en la sección 4.

#### *Glucemia en pacientes críticos*

En el artículo de referencia [54], se consideró utilizar un algoritmo con orden de integración 1, secuencia de enventanado con longitudes que sólo consideran divisores enteros de la longitud del registro  $N$  y un único coeficiente de escalado. Se consiguieron diferencias entre pacientes supervivientes (SV) y no supervivientes (NSV) pero no se consiguió discernir sujetos diabéticos (DM) de no diabéticos (NDM).

A continuación se consideran las modificaciones propuestas con respecto al algoritmo clásico de DFA propuesto por Peng *et al* en [8]:

#### a) Orden de integración:

La primera variación considerada ha sido el uso de distintos órdenes de integración. Los resultados obtenidos y mostrados en la tabla 7.82 son similares a los obtenidos en el artículo de referencia [54], donde solo se consigue segmentación entre SV y NSV. Pero se observa cómo es necesario el uso de al menos un orden 2 para garantizar  $\alpha \leq l$ .

**Tabla 7.82:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y las dos clasificaciones consideradas: supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM). En negrita se indica dónde existe separación.

	DFA-1	DFA-2	DFA-3
NDM	$1.448 \pm 0.027$	$1.606 \pm 0.0306$	$1.688 \pm 0.033$
DM	$1.470 \pm 0.047$	$1.581 \pm 0.0573$	$1.608 \pm 0.062$
UT	0.674	0.682	0.221
SV	$1.535 \pm 0.053$	$1.698 \pm 0.062$	$1.757 \pm 0.054$
NSV	$1.426 \pm 0.023$	$1.563 \pm 0.027$	$1.632 \pm 0.034$
UT	<b>0.036</b>	<b>0.026</b>	0.062

En la figura 7.25.a se muestra la función de fluctuación media ( $\bar{F}(L)$ ) en escala logarítmica del grupo NDM para los tres órdenes de integración considerados. En ella se observa cómo al aumentar el grado, la curva presenta menor ordenada, es decir, existe menor error de fluctuación. Aunque el ajuste se podría considerar como bueno de acuerdo a [116],  $\alpha = 1.448$  y  $R^2 = 0.991$  para DFA-1,  $\alpha = 1.606$  y  $R^2 = 0.993$  para DFA-2 y  $\alpha = 1.688$  y  $R^2 = 0.997$  para DFA-3, se observan dos regiones de escalado. Para DFA-2 se podría hacer un ajuste con dos coeficientes que describiría mejor este comportamiento. Esto se analizará en el punto 3.

A partir de ahora sólo se considerará orden 1 y 2. Orden 1 a modo comparativo con [54] y orden 2 ya que con éste es suficiente para caracterizar este tipo de señales.

#### b) Ventanas deslizantes:

La segunda variación considerada fue el uso de ventanas deslizantes tal y como se hace en [50]. La tabla 7.83 muestra los coeficientes de escalado y los estadísticos de separabilidad entre grupos. Comparándola con la tabla 7.82, los resultados no presentan grandes diferencias y la segmentación entre clases tampoco mejora, incluso empeora para DFA-1.

**Tabla 7.83:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y las dos clasificaciones consideradas: supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM) utilizando ventanas deslizantes. En negrita se remarca dónde existe segmentación entre grupos.

	DFA-1	DFA-2
NDM	$1.427 \pm 0.029$	$1.596 \pm 0.032$
DM	$1.442 \pm 0.052$	$1.570 \pm 0.059$
UT	0.790	0.673
SV	$1.500 \pm 0.054$	$1.694 \pm 0.062$
NSV	$1.407 \pm 0.027$	$1.551 \pm 0.028$
UT	0.102	<b>0.023</b>

No obstante, en la figura 7.25.b se observa que el ajuste es muy similar al obtenido para ventanas no solapadas. La diferencia es que las curvas son más suaves, pero sigue apareciendo ese error de ajuste entre la regresión lineal y  $\log\{F(L)\}$ , de igual forma que aparecía en 7.25.a. Esto nos sugiere que estos sistemas necesitan al menos dos coeficientes de escalado para una completa caracterización.

c) **Múltiples coeficientes de escalado:**

En este punto se aplicó el algoritmo descrito en la sección 4 basado en el criterio de Akaike y la minimización del error de aproximación entre  $\log\{F(L)\}$  y la regresión lineal. Sólo se muestra orden 2 puesto que DFA-1 no cumple  $\alpha \leq l$ .

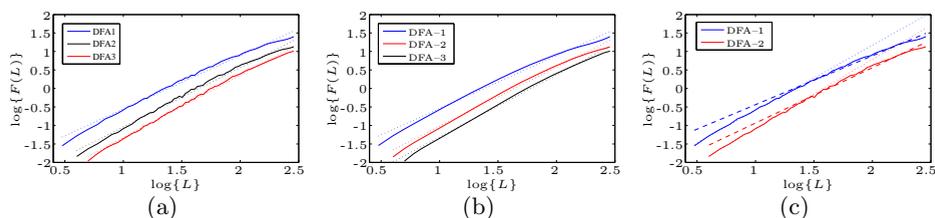
La tabla 7.84 muestra los coeficientes de escalado a corto ( $\alpha_{sr}$ ) y largo plazo ( $\alpha_{lr}$ ), puntos de *cross-over* (CO) para cada uno de los grupos y los estadísticos de separabilidad entre ellos.

**Tabla 7.84:** Coeficientes de escalado a corto ( $\alpha_{sr}$ ) y largo plazo ( $\alpha_{lr}$ ) ( $\mu \pm \sigma$ ), localización del punto de *cross-over* (CO) y estadísticos de separabilidad (UT) para DFA-2 en supervivientes (SV) y no supervivientes (NSV), diabéticos (DM) y no diabéticos (NDM). En negrita se identifica dónde existe segmentación entre grupos.

	$\alpha_{sr}$	$\alpha_{lr}$	CO
NDM	$1.867 \pm 0.134$	$1.498 \pm 0.074$	16
DM	$1.629 \pm 0.038$	$1.532 \pm 0.116$	41
UT	<b>0.046</b>	0.748	0.071
SV	$1.863 \pm 0.141$	$1.501 \pm 0.055$	11
NSV	$1.822 \pm 0.097$	$1.515 \pm 0.200$	34
UT	0.934	0.908	0.410

La figura 7.25.c muestra  $\log\{F(L)\}$  media de sujetos NDM y el ajuste con dos coeficientes. La tabla 7.84 muestra  $\alpha_{sr}$  y  $\alpha_{lr}$  con  $l=2$ . Calculando estos coeficientes a partir de  $\log\{F(n)\}$  media, tenemos que para NDM:  $\alpha_{lr} = 1.867 \pm 0.040$  con  $R^2 = 0.999$  y  $\alpha_{sr} = 1.498 \pm 0.040$  con  $R^2 = 0.992$ , mismos valores medios que los mostrados en la tabla 7.84 y de nuevo mejor ajuste que con un único coeficiente.

Adicionalmente, en este punto sí se observan diferencias significativas en el coeficiente de corto plazo entre sujetos DM y NDM cosa que antes no se obtenía. El punto de *cross-over* varía ligeramente con el grado. Los coeficientes por contra aumentan y en sujetos patológicos el punto de *cross-over* aparece más retrasado que en sujetos sanos. Este retraso puede indicar también la existencia de la patología. Por contra, no es útil para realizar la segmentación entre SV y NSV, ya que no consiguen encontrar diferencias ni a corto ni a largo plazo tal y como se muestra en la tabla 7.84. Aunque no se encuentren diferencias en los coeficientes de escalado, el punto de *cross-over* en NSV aparece retrasado con respecto al de SV.



**Figura 7.25:** Función de fluctuación ( $F(L)$ ) media para sujetos no diabéticos (NDM) en escala logarítmica y distintos órdenes de integración de DFA: en azul orden 1, en negro orden 2 y en rojo orden 3. La línea sólida identifica  $F(n)$  y la punteada la regresión lineal. (a) ventanas no solapadas (b) ventanas deslizantes (c) dos coeficientes: la línea punteada identifica la regresión lineal en la primera parte y la línea discontinua en la segunda parte.

### Glucemia ambulatoria

En el artículo de referencia [55] se consideró utilizar un algoritmo con orden de integración 1, secuencia de enventanados con longitudes que sólo consideran divisores enteros de la longitud total del registro  $N$  y un único coeficiente de escalado. En este punto, se han considerado variaciones en el cómputo de DFA como el orden de integración, el uso de ventanas deslizantes o múltiples coeficientes de escalado.

#### a) Orden de integración:

La tabla 7.85 muestra la caracterización estadística de  $\alpha$  en función del orden  $l$  del polinomio de integración. Al contrario de lo que sucedía en la base de datos de pacientes críticos, en ésta se halla segmentación prácticamente para cada orden de integración y entre los tres grupos considerados.

**Tabla 7.85:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad ( $UT_i$ ) para distintos órdenes de integración de DFA (DFA- $l$ ) y los distintos grupos considerados, diabéticos (DM), síndrome metabólico (MS) y no diabéticos (NDM). Separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM

	DFA-1	DFA-2	DFA-3
NDM	$1.199 \pm 0.018$	$1.314 \pm 0.025$	$1.410 \pm 0.033$
MS	$1.261 \pm 0.021$	$1.435 \pm 0.048$	$1.527 \pm 0.049$
DM	$1.363 \pm 0.018$	$1.677 \pm 0.044$	$1.795 \pm 0.062$
$UT_1$	<b>0.001</b>	<b>0.001</b>	<b>0.001</b>
$UT_2$	<b>0.037</b>	<b>0.032</b>	0.074
$UT_3$	<b>0.008</b>	<b>0.003</b>	<b>0.008</b>

Según [55] en función de la severidad de la patología del sujeto, mayor es  $\alpha$ . Se debería tener  $\alpha_{\text{NDM}} \leq \alpha_{\text{MS}} \leq \alpha_{\text{DM}}$ . Este comportamiento se obtiene para todo  $l$  considerado. El valor de  $l = 3$  no segmenta entre NDM y MS. Esto nos sugiere que  $l = 2$  es el orden adecuado para eliminar la tendencia en la señal base ya

que de acuerdo a [88],  $l \geq \alpha$ . Algo que ya se observa en las series de glucemia de pacientes críticos en 7.4.3.

La figura 7.26.a muestra  $\log(F(L)) \sim \log(L)$ . En ella se observa cómo la relación no es estrictamente lineal. Se observan dos regiones de escalado. Se sugiere el uso de dos coeficientes aún cuando, por ejemplo, para NDM,  $\alpha = 1.199 \pm 0.017$  con  $R^2 = 0.988$  para  $l = 1$ ,  $\alpha = 1.314 \pm 0.025$ ,  $R^2 = 0.979$  para  $l = 2$  y  $\alpha = 1.410 \pm 0.027$ ,  $R^2 = 0.980$  para  $l = 3$ . Estos valores han sido obtenidos sobre las  $F(L)$  medias, mientras que los mostrados en la tabla 7.85 han sido estimados sobre las  $F(L)$  individuales. DFA se puede considerar una medida robusta en términos estadísticos.

b) **Ventanas deslizantes:**

La tabla 7.86 muestra la caracterización estadística ( $\mu \pm \sigma$ ) de  $\alpha$  para cada grupo de la base de datos en función de  $l$ , estimados con el algoritmo de ventanas deslizantes. En ella se observa cómo la segmentación entre los tres grupos simultáneamente se obtiene con  $l=2$ , mientras que  $l=\{1,3\}$  no consiguen diferenciar entre NDM y MS. Por otro lado, de nuevo  $l=1$  tampoco cumple las restricciones impuestas en [88] ya que  $l < \alpha$ .

**Tabla 7.86:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos  $l$  para los distintos grupos considerados, diabéticos (DM), síndrome metabólico (MS) y no diabético (NDM), se indica la separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM

	DFA-1	DFA-2	DFA-3
NDM	$1.199 \pm 0.020$	$1.312 \pm 0.028$	$1.405 \pm 0.038$
MS	$1.256 \pm 0.020$	$1.437 \pm 0.052$	$1.535 \pm 0.054$
DM	$1.384 \pm 0.022$	$1.687 \pm 0.041$	$1.828 \pm 0.059$
UT <sub>1</sub>	<b>0.002</b>	<b>0.001</b>	<b>0.001</b>
UT <sub>2</sub>	0.084	<b>0.043</b>	0.084
UT <sub>3</sub>	<b>0.005</b>	<b>0.005</b>	<b>0.005</b>

En la figura 7.26.b se muestra  $\log(F(L)) \sim \log(L)$  media para NDM. Se observan curvas más suavizadas que con el algoritmo clásico, pero el ajuste no es mucho mejor que en el caso de ventanas no solapadas salvo para el caso  $l = 3$ . Los coeficientes de escalado para cada grado de DFA considerados son:  $\alpha = 1.198 \pm 0.017$  y  $R^2 = 0.988$  para  $l=1$ ,  $\alpha = 1.312 \pm 0.025$ ,  $R^2 = 0.979$  para  $l = 2$  y  $\alpha = 1.528 \pm 0.021$ ,  $R^2 = 0.988$  para  $l = 3$ .

Se observa la necesidad de un segundo coeficiente de escalado, ya que el comportamiento a simple vista no es lineal y se observa un punto de cambio de tendencia, más acusado a mayor orden de integración en torno a  $\log(L) = 1.5$ .

c) **Múltiples coeficientes de escalado:**

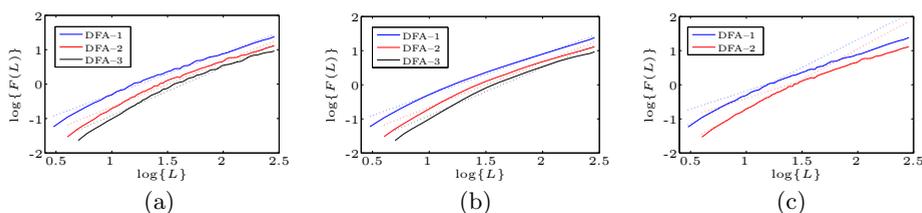
En los casos anteriores se ha visto cómo DFA-3 empeoraba, no conseguía segmentar entre los tres grupos simultáneamente y DFA-1 no cumplía  $l \leq \alpha$ . En este punto se ha considerado sólo DFA-2.

En la tabla 7.87 se muestra la caracterización de  $\alpha$  para los tres grupos con  $l=2$ . Existen diferencias estadísticas en  $\alpha_{sr}$  pero no en  $\alpha_{lr}$ , y también sobre el *cross-over*. Esto unido a los resultados en pacientes críticos sugiere que las diferencias patológicas se encuentran a corto plazo ( $\log(L) \leq 1.5$ ), y que a mayor grado de patología más tarde se localiza el *cross-over*. Estos resultados surgieron que existe una relación entre la escala temporal en la que aparece el *cross-over* y la activación completa del sistema glucorregulador.

**Tabla 7.87:** Coeficientes de escalado a corto ( $\alpha_{sr}$ ) y largo plazo ( $\alpha_{lr}$ ) ( $\mu \pm 2\sigma$ ), y localización del punto de *cross-over* (CO) para  $l = 2$  en registros de sujetos diabéticos (DM), con síndrome metabólico (MS) y no diabéticos (NDM). Se indica la separabilidad (UT) entre 1: DM-NDM, 2: NDM-MS y 3: MS-DM.

	$\alpha_{sr}$	$\alpha_{lr}$	CO
NDM	$1.6761 \pm 0.0622$	$1.0170 \pm 0.0442$	24
MS	$1.7427 \pm 0.0878$	$0.9702 \pm 0.1275$	51
DM	$1.8660 \pm 0.0723$	$1.0184 \pm 0.1239$	81
UT <sub>1</sub>	<b>0.005</b>	0.976	<b>0.001</b>
UT <sub>2</sub>	<b>0.048</b>	0.765	<b>0.034</b>
UT <sub>3</sub>	<b>0.039</b>	0.754	<b>0.015</b>

La figura 7.26.c muestra  $\log(F(L)) \sim \log(L)$  y las regresiones lineales (líneas punteadas) a cada lado del punto de *cross-over*. El *cross-over* indica un cambio en la tendencia o comportamiento de la señal sugiriendo la existencia de un proceso más regular al inicio (mayor  $\alpha$ ) y más irregular al final. Se observa cómo en función de  $l$  este punto aparece retrasado, y se ha comprobado que es así para las tres clases. Este punto también aparece retrasado en función del grado de patología, al igual que en las series de sujetos críticos.



**Figura 7.26:** Función de fluctuación ( $F(L)$ ) media para sujetos sanos (NDM) en escala logarítmica y distintos órdenes de integración de DFA. En azul orden 1, en negro orden 2 y en rojo orden 3. La línea sólida denota  $F(n)$  y la punteada la regresión lineal. (a) Ventanas no solapadas, (b) ventanas deslizantes y (c) dos coeficientes: la línea punteada representa la regresión lineal en la primera parte y la línea discontinua en la segunda parte.

### Dinámicas de movilidad

De igual forma que para las señales de glucemia, se ha considerado valorar variaciones en el algoritmo de DFA sobre las señales de dinámicas de movilidad. Estas variaciones consisten en el orden de integración, el uso de ventanas deslizantes y en caso necesario múltiples coeficientes de escalado.

#### a) Orden de integración:

La tabla 7.88, muestra  $\alpha$  para cada tipo de registro contenido en la base de datos de dinámicas de movilidad. En este caso  $\alpha < 1$  de forma que con  $l=1$ , sería suficiente para tener un resultado válido.

**Tabla 7.88:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y los distintos grupos de registros considerados. Registros de control (CT), de Párkinson (PD), Huntington (HD) y esclerosis lateral amiotrófica (LSA).  $UT_1$ ,  $UT_2$  y  $UT_3$  son los estadísticos de segmentación entre LSA, PD y HD con respecto a CT.

	DFA-1	DFA-2	DFA-3
LSA	$0.684 \pm 0.048$	$0.714 \pm 0.053$	$0.746 \pm 0.055$
PD	$0.692 \pm 0.033$	$0.689 \pm 0.032$	$0.703 \pm 0.029$
HD	$0.613 \pm 0.054$	$0.599 \pm 0.055$	$0.609 \pm 0.050$
CT	$0.668 \pm 0.019$	$0.651 \pm 0.018$	$0.673 \pm 0.020$
$UT_1$	0.228	0.100	0.091
$UT_2$	0.984	0.353	0.353
$UT_3$	<b>0.032</b>	<b>0.009</b>	<b>0.004</b>

En la figura 7.27.a se observa la influencia del orden en  $F(L)$  reflejada en la ordenada. Apenas existe variación de los  $\alpha$  en cada grupo (aprox. 4% superiores para cada orden).

El ajuste con un único coeficiente en DFA-1 es muy bueno,  $R^2 = 99.89\%$  en CT,  $99.79\%$  en HD y PD y  $99.29\%$  en LSA y puesto que cumple que  $\alpha < l$  y se observa una única region de escalado,  $l = 1$  sería suficiente [88, 116].

#### b) Ventanas deslizantes:

La tabla 7.89 muestra  $\alpha$  ( $\mu \pm \sigma$ ) para  $l$  y cada grupo. Ahora con  $l = 1$ , no se consigue separar ningún grupo pero para  $l = \{2,3\}$  los resultados de separabilidad se mantienen. Respecto al uso de ventanas disjuntas,  $\alpha$  generalmente disminuye.

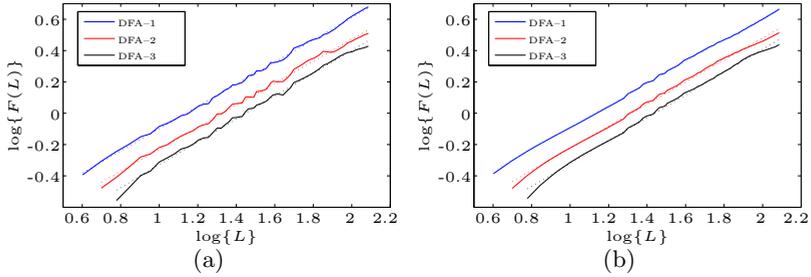
La figura 7.27.b presenta  $\log\{F(L)\} \sim \log(L)$ . La única diferencia apreciable son curvas menos ruidosas con respecto a 7.27.a.

#### c) Múltiples coeficientes de escalado:

**Tabla 7.89:** Coeficientes de escalado ( $\mu \pm \sigma$ ) y estadísticos de separabilidad (UT) para distintos órdenes de integración de DFA (DFA- $l$ ) y los diferentes grupos de registros considerados en el algoritmo de DFA estimado con ventanas deslizantes. Registros de control (CT), de Párkinson (PD), Huntinton (HD) y esclerosis lateral amiotrófica (LSA). UT<sub>1</sub>, UT<sub>2</sub> y UT<sub>3</sub> son los estadísticos de segmentación entre LSA, PD y HD con respecto a CT.

	DFA-1	DFA-2	DFA-3
LSA	$0.667 \pm 0.051$	$0.728 \pm 0.056$	$0.757 \pm 0.060$
PD	$0.675 \pm 0.027$	$0.680 \pm 0.028$	$0.695 \pm 0.026$
HD	$0.608 \pm 0.049$	$0.596 \pm 0.053$	$0.606 \pm 0.047$
CT	$0.660 \pm 0.024$	$0.651 \pm 0.023$	$0.672 \pm 0.024$
UT <sub>1</sub>	0.469	0.120	0.167
UT <sub>2</sub>	0.767	0.441	0.514
UT <sub>3</sub>	0.058	<b>0.015</b>	<b>0.008</b>

En la figura 7.27 se observa que no tiene ningún sentido aplicar un ajuste por doble coeficiente, ya que solo se aprecia una única región de escalado. La señal es completamente monofractal.



**Figura 7.27:** Función de fluctuación ( $F(L)$ ) media para sujetos de control (CT) en escala logarítmica y distintos órdenes de integración de DFA. En azul orden 1, en rojo orden 2 y en negro orden 3. La línea sólida denota  $F(n)$  y la punteada la regresión lineal. (a) ventanas no solapadas (b) ventanas deslizantes.

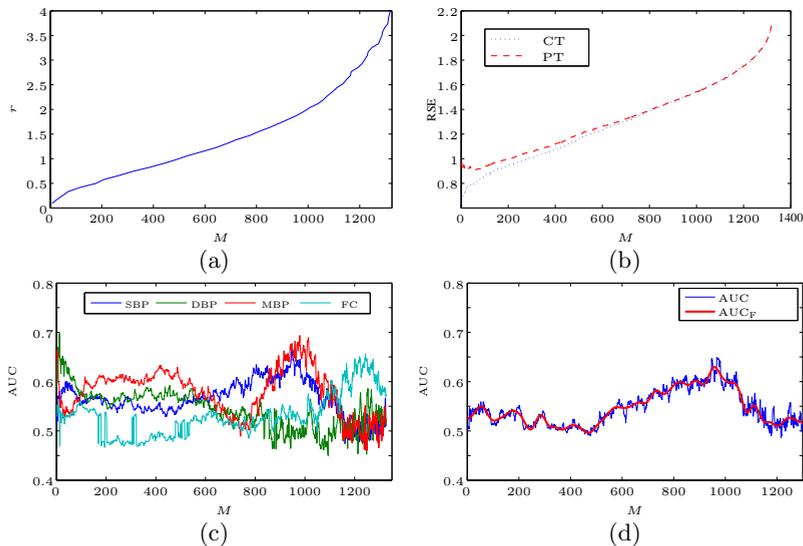
## 7.5 Entropía cuadrada de Rényi (QSE)

La entropía cuadrada de Rényi, también denominada por Lake en [23] como entropía muestral cuadrática, aplica un algoritmo basado en la selección óptima del umbral  $r$  para el cálculo de SampEn de forma que se optimice la segmentación entre clases. Está diseñado para señales de corta duración principalmente debido al coste computacional asociado, por ello aquí solo se ha aplicado en señales de hipertensión arterial (HTA) que contienen apenas 55 muestras.

### 7.5.1 Hipertensión Arterial

Este algoritmo trata de buscar el umbral óptimo para cada señal a partir de la selección del número de patrones coincidentes que ofrece una máxima segmentación de las clases. Para ello se necesita aplicar un algoritmo de búsqueda inverso. Se seleccionó  $m=2$  por convención en SampEn pero en [117] se ofrece un mayor número de resultados y comparativas con otras medidas y otros valores  $m$ .

Se propuso definir una función para buscar el número de patrones coincidentes ( $M$ ) en cada señal y a partir de éstos definir el umbral necesario para que se garantice ese valor de  $M$  en cada señal. Se seleccionó un 30 % del número de señales de cada grupo de la base de datos como grupo de entrenamiento ( $TRAIN$ ). La figura 7.28.a muestra la evolución del umbral  $r$  para el número de patrones coincidentes  $M$  en una señal de la base de datos. El segundo paso es calcular QSE para cada señal del grupo  $TRAIN$ , figura 7.28.b.



**Figura 7.28:** Cálculo de QSE sobre señales HTA. (a) Evolución del umbral,  $r$ , con el número de patrones coincidentes,  $M$ , para una señal. (b) Evolución de QSE en función de  $M$  para una señal de cada grupo, control (CT) en rojo y patológica (PT) en azul. (c) Evolución del área bajo la curva ROC (AUC) de cada variable de la base de datos, presión sistólica (SBP), diastólica (DBP), media (MBP) y frecuencia cardíaca (FC) en función de  $M$ . (d) AUC de SBP en azul y su versión filtrada ( $AUC_F$ ) en rojo.

Para buscar la máxima segmentación entre clases o clasificación óptima se optó por buscar aquel  $M$  que maximizara el área bajo la curva ROC (AUC). Para reducir la posibilidad de un sesgo asociado a una  $TRAIN$  pequeña, se optó por aplicar un algoritmo de *bootstrap*. Dado el limitado número de señales (9 por clase) se

aplicó *Leave-One-Out* (LOO) eliminando una señal de cada grupo en cada iteración, dando lugar a 18 iteraciones. La figura 7.28.c muestra la curva AUC media en cada variable de la base de datos. Por debajo de 0.5 no existe clasificación y cuanto más proxima a 1 es el AUC, mejor clasificación se obtendrá en las clases. Para reducir las rápidas variaciones existentes en la AUC, ésta es filtrada. La figura 7.28.d muestra la AUC filtrada para una de las variables de la base de datos. Y se selecciona como  $M_{opt}$  aquel para el cual se obtiene el máximo de  $AUC_F$ .

La tabla 7.90 muestra los valores  $M_{opt}$  para cada grupo y los intervalos de confianza (CI) junto con la probabilidad de segmentación entre clases (UT) obtenidas con dicho  $M_{opt}$ , para cada variable. Tradicionalmente se ha venido utilizando la presión arterial sistólica para determinar la condición de hipertensión arterial de un paciente, mientras que aquí se propone que con la selección de parámetros adecuada se pueden encontrar diferencias entre clases en la diastólica y en la media, no siendo útil la frecuencia cardíaca en esta situación.

**Tabla 7.90:** Caracterización de QSE sobre registros de hipertensión arterial con 4 variables: presión arterial sistólica (SBP), diastólica (DBP), media (MBP) y frecuencia cardíaca (FC). Se muestra el número de patrones coincidentes óptimo  $M_{opt}$  para cada grupo de señales de la base de datos de entrenamiento. Estadísticos QSE calculados con  $M_{opt}$  en la base de datos completas, intervalos de confianza al 95 % (CI) y separabilidad (UT)

VAR.	$M_{opt}$	CI CT	CI PT	UT
SBP	958	[1.470, 1.511]	[1.456, 1.488]	0.158
DBP	17	[0.559, 0.872]	[0.816, 1.047]	0.029
MBP	960	[1.495, 1.524]	[1.467, 1.504]	0.046
FC	1248	[1.820, 1.949]	[1.855, 1.937]	0.765

## 7.6 Coeficiente de entropía muestral (CosEn)

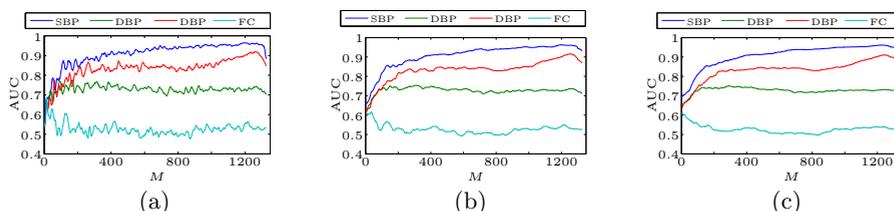
El coeficiente de entropía muestral (CosEn) es una modificación de la entropía cuadrada de Rényi, donde se considera que la señal media contiene información importante y que no debería obviarse ya que podría contribuir a una mejor caracterización de la entropía.

Esta medida se ha aplicado a señales de corta duración debido a su alto coste computacional. Se han considerados las bases de datos de registros de hipertensión arterial ( $N < 55$ ) y de glucemia post-operatoria ( $N < 108$ ) a las cuales difícilmente se les puede aplicar ninguna medida de complejidad de forma robusta.

### 7.6.1 Hipertensión Arterial

A continuación se muestran los resultados de aplicar CosEn a una serie de registros de hipertensión arterial, de longitud máxima 55 muestras. La metodología es similar a la expuesta en 7.5.1 salvo que en este caso se ha considerado que la señal presenta información relevante en su media que puede ayudar a mejorar los resultados estadísticos de clasificación entre grupos.

Teniendo en cuenta resultados previos y el limitado número de muestras de estas señales ( $N=55$ ), se ha considerado aplicar CosEn con  $m=2$ . El proceso es el mismo al presentado en la figura 7.28 salvo que el filtrado de las AUC se implementó mediante un filtro de media móvil de distinta longitud. La figura 7.29 presenta tres longitudes distintas del filtro de media móvil para las AUC. A mayor longitud del filtro, las series aparecen más suavizadas. Es necesario encontrar la longitud del filtro adecuada que resulte en una serie sin ruido, pero sin un suavizado excesivo. En este trabajo se seleccionó una longitud de filtro  $K=51$ .



**Figura 7.29:** Curva del área bajo la ROC (AUC) en función del número de coincidencias (*matches*,  $M$ ) de CosEn para las cuatro variables consideradas: presión sistólica (SBP), diastólica (DBP), media (PAM) y frecuencia cardíaca (FC) filtrada para distintas longitudes del filtro de media móvil, (a)  $K=11$ , (b)  $K=51$  y (c)  $K=101$ .

Tras aplicar el filtrado a las series AUC y seleccionar el máximo, se recomputó CosEn para ese valor de patrones coincidentes  $M$  obteniendo una segmentación en al menos tres variables de las consideradas. Como se preveía en un inicio, los estudios estadísticos revelan que la frecuencia cardíaca no es una variable que permita discernir entre sujetos con o sin hipertensión arterial. Estos resultados se muestran en la tabla 7.91. CosEn, al contrario que todas las medidas de entropía, resulta en valores negativos. Esto es resultado de considerar la media de la señal en la medida de entropía. La interpretación es la misma, mayores valores de entropía dan lugar a mayores complejidades. El considerar la media dentro de la medida de complejidad, incrementa la separación entre clases con respecto a RSE.

**Tabla 7.91:** Caracterización estadística de CosEn ( $m=2$ ) sobre señales de hipertensión arterial (HTA) en sujetos control (CT) y patológicos (PT). Se muestra test de normalidad (JB), intervalos de confianza y test de separabilidad (UT) de cada variable considerada, presión sistólica (SBP), diastólica (DBP), media (PAM) y frecuencia cardíaca (FC).

	SBP	DBP	PAM	FC
JB	0.052	0.165	0.023	0.868
CI CT	[-6.451, -4.130]	[-4.207, -4.130]	[-6.274, -6.230]	[-2.547, -2.356]
CI PT	[-6.669, -6.603]	[-4.393, -4.251]	[-6.467, -6.394]	[-2.596, -2.357]
UT	0.001	0.001	0.001	0.749

## 7.6.2 Glucemia Postoperatoria

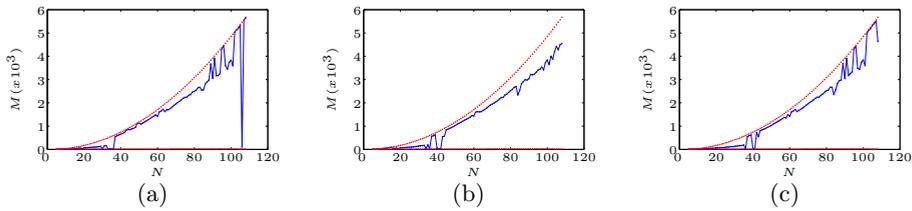
En este punto se incluye el estudio realizado sobre las señales de glucemia postoperatoria registradas en un hospital de la República Checa. Se han considerado dos grupos principales, el primero, a los cuales se les ha llevado un control glüquémico pre y postoperatorio y el segundo solo se ha considerado control postoperatorio. Se han evaluado la probabilidad de desarrollar complicaciones postoperatorias en función del control que se les ha aplicado.

Estas señales no contienen una longitud estándar, y van desde las 5 muestras hasta las 108 muestras. La distribución del número de señales en función del número de muestras se describe en la sección 5.

Se han realizado análisis para las distintas longitudes. En lugar de aplicar un filtro de media móvil se consideró aplicar un filtrado adaptativo en función del periodograma de Welch, con distintas longitudes de ventana de Hamming para el filtrado.

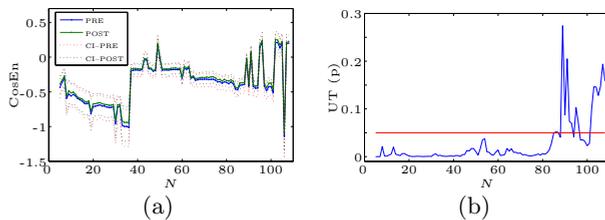
La figura 7.30 muestra la distribución del número de *matches* óptimo en función de la longitud de los registros. Se observa cómo para longitudes de filtrado más pequeñas, hay más oscilaciones. En longitudes pequeñas se tiende a valores menores y en longitudes de señal mayores, suele seleccionarse un valor próximo al máximo. Para un filtrado de hamming de 51 muestras se observa un mayor número de fluctuaciones en los *matches* elegidos como óptimos al igual que en longitudes de ventana grandes. La idea es buscar un filtrado que permita elegir el número óptimo que debería describir una función logarítmica como el número máximo (línea roja superior), sin variaciones bruscas y en punto más o menos intermedio de la región definida entre mínimo y máximo número de *matches*.

La figura 7.31.a muestra CosEn sobre estas señales, sus intervalos de confianza. La probabilidad de encontrar diferencias estadísticas entre las clases se muestra en la figura 7.31.b. Por debajo de 85 muestras en la señal se encuentran diferencias estadísticas entre las clases. En la figura previa, se observa que hay un pico en torno a esa longitud decreciente y empieza a presentar más oscilaciones en la curva del número de *matches*. Encontrando el filtrado óptimo adaptado a la AUC



**Figura 7.30:** Distribución del número de *matches* óptimo,  $M$ , en azul, para una segmentación máxima en función de la longitud de las señales  $N$  con distintos tipos de longitud de la ventana de Hamming para el filtro basado en el periodograma de Welch, (a)  $K = 51$ , (b)  $K = 101$  y (c)  $K = 151$ . En rojo se presenta el número máximo de *matches*.

para maximizar la segmentación se podrían encontrar diferencias estadísticas en todas las longitudes consideradas.



**Figura 7.31:** Capacidades estadísticas de CosEn evaluando series de glucemia en sujetos con control pre-operatorio o post-operatorio en pacientes que han desarrollado complicaciones post-operatorias. Estimación con un filtrado basado en el periodograma de Welch con una ventana Hamming de 51 muestras. (a) Intervalos de confianza (CI) y (b) probabilidad de separación estadística (UT) en función de la longitud  $N$  de la señal analizada.

# Capítulo 8

## Discusión

En este trabajo se ha realizado una caracterización de algunas de las medidas de entropía más comunes ante situaciones en las cuales la longitud de los registros sobre los cuales se han aplicado es limitado. En concreto, las medidas de entropía caracterizadas en este trabajo son ApEn, SampEn, MSE y DFA, como genéricas, o más específicas y adecuadas a longitudes extremadamente cortas como QSE o CosEn.

La limitada longitud de los registros que se ha considerado puede estar asociada a diversos factores: (i) baja frecuencia de muestreo, (ii) imposibilidad de registro de mayor duración de las señales, (iii) pérdida de muestras durante el registro o transmisión de las señales. En (i) tenemos el caso de señales de glucemia. Por tratarse de señales con variaciones lentas y baja frecuencia se toma una muestra cada 5 minutos, resultando en registros de 24h con apenas 300 muestras. En (ii) se tienen las señales de hipertensión arterial o glucemia post-operatoria. En el primer caso, el registro de la señal es muy incómodo para el paciente, de forma que se toma una muestra cada 30 minutos. Además es un proceso poco robusto y muchas veces la muestra tomada no es válida. En glucemia post-operatoria, la toma de datos se realiza también a mano de forma que apenas se dispone de 2 o 3 medidas diarias. Finalmente, en (iii) se tienen registros que se hacen de forma ambulatoria o remota y se necesita una transmisión de la información de la señal a un centro donde ser analizada. Estas transmisión suelen ser inalámbricas mediante Bluetooth, WiFi o incluso infrarrojos aunque también se hacen de forma cableada. Estos registros pueden sufrir distorsiones asociadas al canal, lo que provoca que haya interferencias y en consecuencia la pérdida de alguna muestra (pérdida aleatoria) o de un bloque de muestras (uniforme).

## 8.1 Pérdida de muestras

Como se ha visto en todas las medidas de complejidad analizadas, siempre y cuando la entropía haya sido estimada de forma robusta ( $N$  suficiente para garantizar un número adecuado de patrones coincidentes), la entropía asociada a los registros patológicos es inferior a la asociada a los registros de control. En registros EEG se ha observado que la entropía de CT es superior a PT, lo cual es acorde con [14, 38]. En estos estudios se observó que sobre registros RR, ApEn obtenía valores superiores en neonatos que sanos que en enfermos. Para señales EGM, el grupo de control sería NRA, puesto que se consideran señales más regulares y zonas sin alteraciones que no necesitan ser ablacionadas. En glucemia la entropía de SV es superior a NSV y de NDM a DM, al igual que en dinámicas de movilidad CT es superior a LSA, HD y PD.

Con los registros RR se obtiene que la entropía de POST es inferior a la de PRE, este resultado parece contrario a [14] que muestra cómo cuando un neonato pasa de estado enfermo a sano su entropía aumenta, pero no es justamente el caso de estos registros. En estos registros hay que considerar que las señales presentan un gran número de latidos ectópicos, lo que se traduce en un gran número de *outliers*. En [9] se establece que los *outliers* son artefactos poco frecuentes que resultan en una muy pequeña contribución al valor final de la entropía. En este caso los *outliers* son realmente la información a caracterizar en la señal, ya que se pretendía establecer si la medicación aplicada conseguía reducir el número de ectópicos, por ello que los resultados sean contrarios y PRE presente mayor entropía que POST, porque la variabilidad o aleatoriedad de la señal viene en los latidos ectópicos.

### 8.1.1 Pérdida de muestras aleatoria

El comportamiento de ApEn, SampEn y MSE cuando se eliminan muestras de una señal de forma aleatoria es el mismo: la entropía aumenta. En DFA, al tratarse de una medida de regularidad, el coeficiente de escalado disminuye excepto para los registros RR de fibrilación auricular, donde se mantiene aproximadamente constante. Este aumento de la entropía o disminución de la regularidad, puede estar asociado a la aleatorización que se introduce en la señal al eliminar muestras de forma no ordenada.

En ApEn se observa que la entropía aumenta hasta alcanzar un máximo a partir del cual decae. En las señales EEG ( $N = 4096$ ) este punto se encuentra en torno al 70 % en CT y en torno al 80 % en PT. Para señales RR ( $N = 1000$ ) el punto se encuentra en torno al 10 % en el grupo PRE y en torno al 40 % en el grupo POST. Para los registros EGM ( $N = 1500$ ) el punto de cambio de tendencia se encuentra en torno al 80 % para RA y no se aprecia en NRA. Pincus en [14, 38] ya establece que se necesitan registros con  $10^m \leq N \leq 30^m$ . Generalmente entre 75 y 5000 muestras pero preferiblemente  $N \geq 100$  [37]. En este experimento se consideró  $m$

= 2 puesto que en [14, 38] se hace esa recomendación si se desea estimar las probabilidades condicionales de forma razonable. Para poder tener una estimación razonable de ApEn se necesitaría  $100 \leq N \leq 900$ . Todas las clases presentan registros de más de 100 muestras en los puntos donde se produce el cambio de tendencia en ApEn pero en cualquier caso están por debajo de 900. Según [14], esto puede producir que las estimaciones de las probabilidades condicionales no se hayan estimado de forma suficientemente robusta, debido a que no existe un número de patrones coincidentes suficientes. Tal y como se detalla en [38], ApEn se calcula mediante un algoritmo basado en identificación de patrones, de forma que se garantice que los cálculos logarítmicos presenten una solución finita. Este método de cálculo implica que las probabilidades condicionales sean subestimadas, lo que se traduce en un sesgo de hasta el 20–30 % [38], lo que hace que ApEn sea muy dependiente de la longitud [10].

Excepto para los registros RR, el punto de inflexión de ApEn se produce en porcentajes de pérdidas altos, lo cual sugiere que el número de patrones coincidentes es muy limitado y el decremento en la entropía está asociado al sesgo descrito en [38]. Este punto de inflexión aparece más retrasado en las señales que a priori suelen ser más regulares (EEG PT y EGM NRA) y en consecuencia presentan un mayor número de patrones coincidentes para una misma longitud de registro. Los registros RR son un tanto especial, puesto que la información relativa se encuentra tal y como se ha dicho antes en los latidos ectópicos, lo que provoca que ya desde un principio los patrones coincidentes sean muy limitados. Debido a la gran variabilidad que presentan estas señales, los puntos de inflexión se observan en porcentajes de pérdidas muy bajos, 10 % en PRE y 40 % en POST. Pero presentan igual comportamiento que EEG o EGM donde en señales más regulares la influencia de este sesgo aparece en porcentajes de pérdidas más altos.

No obstante, se puede considerar en primer lugar que la entropía de EEG puede ser estimada mediante ApEn de forma válida siempre que el porcentaje de muestras perdidas sea inferior al 60 % en segmentos de  $N = 4096$  (la longitud residual de la señal debe ser superior a 1600 muestras). De esta forma se garantiza que existan un suficiente número de patrones coincidentes para que no aparezca el sesgo asociado a los *automatches*, identificado como el punto de inflexión, y se mantenga un coeficiente de correlación cruzada superior al 80 %. En registros EGM, inicialmente más regulares que EEG, se aceptan hasta pérdidas de muestras del 80 % en registros de 1500 muestras (registro residual de longitud  $N \geq 300$ ). Los EGM mantienen la correlación cruzada por encima del 80 % en cualquier caso, resultando en valores de ApEn en cierta forma robustos. Finalmente en registros RR con  $N = 1000$ , apenas podemos admitir una pérdida del 10 % (100 muestras). En [22] se realiza un estudio de entropía sobre señales de variabilidad del ritmo cardíaco (HRV) en el cual se ha perdido un segmento de señal en los registros RR. Aunque la forma en que se han perdido muestras no es la misma, en él se establece que sobre señales RR de 5 minutos registradas a 128 Hz ( $N = 640$ ) solo se pueden aceptar pérdidas

de hasta 20 s máximo, lo que representa un 7%, resultados coherentes con los obtenidos en este punto.

SampEn es una mejora de ApEn donde se elimina la dependencia con  $N$  al no considerar los *automatches* en el algoritmo de cálculo [10]. Esto se observa en que la complejidad es siempre creciente para los tres tipos de registros considerados y el punto de inflexión que se producía en ApEn desaparece. La medida es más estable que ApEn ya que todos los registros presentan un valor de correlación cruzada por encima de 0.8 hasta pérdidas del 80%.

En [17] se presenta un breve estudio que muestra el coeficiente de correlación cruzada entre los valores de entropía muestral estimados sobre series de ritmo cardíaco (HR) y sobre las mismas con pérdidas aleatorias, utilizando  $m = 3$  y  $r=0.25$ . Las series HR consideradas presentan 4096 intervalos RR, misma longitud que los registros EEG y superior a EGM y RR. Aunque los parámetros de análisis no son los mismos, se muestra que  $C_{xy} \geq 0.95$  hasta un porcentaje del 40% de pérdidas en las series. Los resultados presentados en esta tesis muestran como tanto para señales EEG de igual longitud como para registros RR y señales EGM, con pérdidas aleatorias del 40%,  $C_{xy} \geq 0.95$ . La validez clínica ( $C_{xy} \geq 0.8$ ) de SampEn en los registros HR utilizados en [17] se mantiene hasta alcanzar un 70% de muestras perdidas. Las series utilizadas en esta tesis, presentan validez hasta un mínimo del 80% de pérdidas.

Contrario a ApEn, SampEn no consigue segmentar los registros RR, esto sugiere que la medicación aplicada no es tan indicada como podía considerarse en un inicio.

Podemos sugerir que el incremento de la complejidad esta asociado a la aleatorización que este tipo de pérdidas introduce en las señales, y que SampEn es más estable y robusta que ApEn frente a la pérdida de muestras.

La entropía en multiples escalas o multiescala (MSE), analiza las series mediante el uso de SampEn en escalas temporales de mayor duración [15]. Para ello realiza una media de  $S$  muestras consecutivas, donde  $S$  es la escala en la que analiza la serie. De acuerdo a la literatura existente referente a las señales analizadas y a los resultados previos, se establece que sujetos sanos o de control presentan menor regularidad que sujetos patológicos [17, 41, 42, 84]. En los resultados presentados para MSE se observa que en función de la escala, el comportamiento en la complejidad de la series se invierte, llegando a tener por ejemplo en EEG, que a partir de la escala 8 CT presenta menor complejidad que PT. En EGM se observa el mismo comportamiento: las series RA reducen su complejidad por debajo de las series NRA. El porcentaje de muestras perdidas en el que se produce este cambio de comportamiento es menor a medida que aumenta la escala. Sin embargo las series RR mantienen el comportamiento en la complejidad, reduciendo la distancia de separación entre PRE y POST a medida que aumenta la escala.

De acuerdo a los artículos originales de Costa [15, 118], a mayores escalas la entropía de registros procedentes de sujetos enfermos, deben presentar menor complejidad que la de sujetos sanos. Los resultados obtenidos para EEG y EGM son contrarios a este hecho. El punto en el cual cambia la señal para EEG en, por ejemplo, escala 4, se produce en un 60 % de pérdidas. Las series analizadas son de 400 muestras de longitud. En la escala 6 se produce con registros de 300 muestras. SampEn presenta un sesgo residual para longitudes muy cortas [10] asociado a las correlaciones entre los subvectores, que puede ser la causa de este cambio en el comportamiento. En cualquier caso no se debe recomendar el uso de escalas superiores a 2 en estas señales.

Finalmente, DFA no presenta un mismo comportamiento para los tres tipos de registros considerados. Para registros EGM y EEG la regularidad dada por el coeficiente de escalado disminuye, resultado coherente con los obtenidos en SampEn y ApEn. Dicha regularidad tiende a la asociada al ruido blanco,  $\alpha = 0.5$ . Por ello se puede decir que una pérdida de muestras de forma aleatoria provoca una aleatorización en la señal. Por el contrario, el coeficiente de escalado de los registros RR se mantiene aproximadamente constante a pesar que la señal presente pérdida de muestras. Esto sugiere que el comportamiento a largo plazo de los registros RR no cambia, pero revisando un poco la literatura actual [19, 51, 119] se sugiere el uso de dos coeficientes para la completa caracterización del comportamiento fractal de la señal. No obstante Govindan en [19], establece que DFA ofrece resultados estables en señales de ritmo cardíaco en fetos con una pérdida de muestras de hasta el 50 %. En esta tesis se observa resultados estables en series RR hasta un 80 % de pérdidas ( $C_{xy} \leq 0.8$ ), hasta un 60 % en series EGM y en todo el rango considerado en EEG.

### 8.1.2 pérdida de muestras uniforme

El comportamiento de ApEn, SampEn y MSE es el mismo en las tres bases de datos consideradas cuando la pérdida de muestras se produce de forma uniforme. La hipótesis formulada es que al eliminar un segmento de señal se eliminan aproximadamente el mismo número de patrones coincidentes de longitud  $m$  que de longitud  $m + 1$ , de forma que la complejidad se ve mínimamente alterada. Esta hipótesis es confirmada con los resultados presentados en este trabajo.

En ApEn el comportamiento no es estable a lo largo de todo el rango de pérdidas. Al igual que sucedía cuando se eliminaban muestras de forma aleatoria, existe un porcentaje de pérdidas a partir del cual la complejidad comienza a disminuir. En señales EEG este punto se produce en torno al 60–70 % de pérdidas, en registros RR se encuentra en torno al 10 % y en señales EGM en torno al 30 %. En este caso, el punto de cambio de tendencia se produce antes que para pérdidas de forma aleatoria. Posiblemente esté asociado a que en este caso no se aleatoriza la señal,

y sea más fácil encontrar patrones coincidentes, lo que provoca un descenso en la entropía de forma que el sesgo producido por los *automatches* es más evidente.

SampEn por el contrario, no muestra ese punto de cambio de tendencia, de forma que la complejidad se mantiene constante para todo el rango de pérdidas, resultando en coeficientes de correlación cruzada mayores.

En [22] se hace un análisis exhaustivo de registros de variabilidad del ritmo cardíaco (HRV) en los cuales se han perdido intervalos del registro RR de los que derivan. Las señales analizadas presentaban una duración de 5 minutos a una frecuencia de muestreo de 128 Hz (640 muestras). En este trabajo, se utilizan tres técnicas de reconstrucción (*jackknifing*, *bootstrapping* e interpolación) para el segmento de señal RR que se ha perdido, pero como no reducen el error normalizado, se considera que la serie original sin reconstrucción es óptima para el análisis de HRV en el dominio temporal. En este trabajo [22] también se establece que SampEn es más estable y robusta que ApEn.

ApEn se considera robusta en situaciones en las cuales el segmento de información perdida es inferior a 20s (en registros de 5 minutos, 7%) [22]. En nuestro análisis ApEn ha resultado estable en todos los registros al menos hasta un 40% de pérdidas ( $C_{xy} \geq 0.8$ ).

En SampEn pueden tolerarse pérdidas de hasta 60 s con  $r = 0.15$  y  $m = 2$  (20%). Se considera que SampEn(2,0.15) es la medida más estable para estudiar la entropía de HRV [22]. Estos parámetros son muy similares a los utilizados en este trabajo, reportando estabilidad en SampEn para todos los registros considerados hasta un 70% de muestras pérdidas (registros EGM).

MSE, como era esperado, refleja el mismo comportamiento que SampEn para todos los registros analizados. La complejidad se mantiene estable en todo el rango de pérdidas. En señales EEG, la complejidad se mantiene más o menos estable entre 1.8 y 2 para señales CT mientras que aumenta con la escala hasta valores similares para PT. La influencia de la escala se traduce en una reducción en la separabilidad entre las clases, por ello no es recomendable considerar escalas superiores a 8. Los registros RR también mantienen su complejidad constante tanto en diferencia entre clases, como en valores en función de la escala y diezmado. En diezmados altos aparece una tendencia creciente. Esto puede asociarse a la eliminación de un mayor número de latidos ectópicos de forma que se aprecia la regularidad intrínseca de los registros RR sin éstos. Similar comportamiento se observa en EGM donde la complejidad en cada grupo aumenta en función de la escala y contrario a señales EEG la distancia entre clases aumenta ligeramente.

En DFA se observa también un comportamiento estable donde el coeficiente se mantiene más o menos constante para todo el rango de pérdidas tanto para señales RR como para EGM. En señales EEG observamos que la estabilidad se mantiene

hasta el 20% de muestras perdidas. En esta tesis se ha realizado una caracterización con un único coeficiente de escalado de acuerdo al algoritmo original definido por Peng en [8], donde se han considerado todas las escalas posibles desde  $0.477 \leq \log(n) \leq 2.574$ . En [120] se sugiere buscar el régimen de escalado óptimo como el régimen de escalado que presenta el mejor ajuste. En [116] se sugiere un algoritmo de DFA tiempo variante. En [121] se muestra cómo en sujetos de control y narcolépticos se necesitan dos coeficientes para caracterizar completamente el comportamiento fractal de un EEG. En [22] se señala que DFA puede considerarse robusta siempre y cuando el segmento RR perdido sea inferior a 10 s (3.3%). Govindan [19] establece que DFA es robusta hasta cuando se pierden 50 minutos consecutivos en registros de ritmo cardíaco fetal de 8h muestreados a 312.5 Hz (10.4%). En el presente estudio se ha obtenido que DFA es consistente hasta aproximadamente un porcentaje de pérdidas del 80% en EEG, 70% en RR y 60% en EGM, de acuerdo a su coeficiente de correlación cruzada ( $C_{xy} \geq 0.8$ ) [105].

## 8.2 Señales de corta duración

En esta tesis se han considerado señales de corta duración como aquellas de menos de 500 muestras. Se han considerado registros de glucemia de 24h en pacientes críticos y ambulatorios con 288 muestras y registros de dinámicas de movilidad con 312 muestras. Se ha considerado un segundo grupo como señales de muy corta duración por tratarse de registros con menos de 100 muestras, como los registros de glucemia post-operatorios o los registros de hipertensión arterial.

Se ha realizado una caracterización paramétrica de ApEn, SampEn y MSE y un estudio del efecto de las ventanas deslizantes, el orden de integración y los múltiples coeficientes en DFA para las señales de más de 100 muestras. Para los registros de hipertensión arterial y glucemia post-operatoria se han considerado medidas más adecuadas a su longitud y de reciente consideración como son la entropía cuadrada de Rényi o entropía muestral de orden dos (QSE) y el coeficiente de entropía muestral (CosEn), propuestas por Lake en [17] y [23], respectivamente. Estos algoritmos basados en la estimación del número de patrones óptimos para la segmentación presentan alto coste computacional y por ello no se considera su aplicación en señales de mayor longitud.

Sobre las señales de glucemia en sujetos críticos, ApEn no es capaz de establecer diferencias significativas si no es con un valor de  $r$  alto, aún considerando dimensiones  $m$  bajas y solo entre sujetos SV y NSV. Pincus en [14] establece que se pueden obtener probabilidades razonablemente estables con  $10^m - 30^m$  muestras. Se ha observado cómo en este caso no sucede así. Por el contrario, en sujetos ambulatorios, ApEn(1,0.1) consigue segmentar sujetos DM de NDM y MS de DM. Esto sugiere que el estado crítico de los pacientes influye en las series de glucemia algo que ya se indica en el artículo original [54] donde las series han sido analizadas

mediante DFA. Esta influencia enmascara las variaciones en las que ApEn se basa para establecer diferencias significativas. En [55] se establece que el aumento en la regularidad indica un agravamiento en la patología diabética. Este mismo comportamiento se observa en ApEn siempre y cuando la selección de parámetros sean adecuados. En las señales de dinámicas de movilidad se observa que ApEn sólo es capaz de discernir entre LSA y CT y en algunos casos entre HD y CT. De nuevo se observa que únicamente con combinaciones de parámetros pequeñas (1,0.1), (3,0.4) se consigue establecer diferencias significativas entre CT y algunos de los registros patológicos, respetando los resultados establecidos en la literatura que asocian valores inferiores de complejidad a situaciones patológicas [10, 14, 17, 38].

SampEn no consigue establecer diferencias significativas entre ninguno de los grupos de los registros de glucemia en pacientes críticos. Sobre pacientes ambulatorios las combinaciones paramétricas que establecen diferencias significativas entre grupos es más limitada que para ApEn. Lake en [17] dice que con  $m$  suficientemente pequeño y  $r$  suficientemente grande para garantizar un número adecuado de patrones coincidentes, SampEn puede considerarse de acuerdo a una distribución normal. En este caso observamos cómo las distribuciones son normales y solo se ajustan a la afirmación de Lake en SampEn (1,0.2), donde solo se ofrece segmentación entre sujetos DM y NDM. Se observa que en función del agravamiento de la patología de los sujetos, la regularidad disminuye. Resultado en consonancia con los presentados en [55] y con la literatura existente [10, 14, 17, 38]. Con las señales de dinámicas de movilidad se observa que los resultados incluso empeoran al únicamente establecer diferencias entre HD y CT. Según [10, 17] SampEn es más estable, presenta mayor consistencia que ApEn, por lo que debería ofrecer unos resultados más robustos. Bajo esta premisa se puede intuir que ninguna de las dos son adecuadas para analizar estas señales.

En MSE se observa el mismo comportamiento que para ApEn y SampEn. No se encuentran diferencias significativas de forma homogénea ni acordes con la literatura en todos los casos. Tampoco se han podido considerar escalas superiores a 3 debido a la longitud limitada de los registros. Los valores de  $r$  para los cuales se consigue establecer segmentación entre algunas de las variables en la base de datos de glucemia ambulatoria y en dinámicas de movilidad son pequeños ( $r=0.1$ ) lo que no garantiza que el número de patrones coincidentes sean suficientes para obtener una estimación robusta de las probabilidades marginales y en consecuencia de las entropías. Esto en cierta forma era un resultado esperado ya que de acuerdo a [15], MSE es un algoritmo diseñado para registros de larga duración. De hecho, en [122] se realiza un estudio de dinámicas de movilidad con MSE pero sobre señales de más de 1500 muestras.

La mayoría de los estudios realizados sobre registros de glucemia se realizan mediante el uso de DFA [54, 55, 56, 58, 123, 124]. Incluso de otros registros de corta duración como EEG [8, 45] o registros HRV [19]. Generalmente, en estos algoritmos se aplica bien polinomio de eliminación de tendencia de orden 1 [54, 55] u

orden 3 [56, 123]. No hay una clara definición de la motivación para uso de uno u otro.

En esta tesis se ha presentado una caracterización de DFA en función del orden de integración sobre dos bases de datos de glucemia y sobre registros de dinámicas de movilidad. En los resultados se ha observado cómo  $\alpha$  aumenta con el orden de integración y la función de fluctuación reduce su ordenada. De acuerdo a [88], el orden de DFA está relacionado con  $\alpha$ , de forma que se debe seleccionar un orden  $l$  tal que  $l \geq \alpha$ . Para los registros de glucemia con un orden  $l = 2$  era suficiente y para registros de dinámicas de movilidad bastaba con  $l = 1$ . Se ha observado como utilizar  $l = 3$  en registros de glucemia incluso empeoraba los resultados de separabilidad entre clases, descartando este valor. En segundo lugar el algoritmo de DFA definido originalmente utiliza ventanas disjuntas [8] para el análisis de la señal. En [56, 58, 123] se utilizan ventanas deslizantes sin una argumentación pausable. El uso de ventana deslizante fue descrito inicialmente en [50] como alternativa para obtener mejores estadísticos. En esta tesis se ha realizado una comparativa entre los resultados obtenidos con ventanas disjuntas frente a ventanas deslizantes. En la figura 7.25 se observa que la única variación existente es que la función de fluctuación aparece más suavizada. En particular para escalas temporales altas  $\log(n) \leq 1.5$ , donde el número de ventanas es limitado. Los resultados no mejoran de forma notoria como para justificar el coste computacional asociado.

En señales de glucemia, aunque el ajuste ( $R^2$ ) es bueno, se han observado que existen dos regiones de escalado y un punto de *cross-over*. De acuerdo a [116] sólo se puede aceptar el método tracional con un único  $\alpha$  si  $R^2 \geq 0.95$  y no se observa un *cross-over*. En la mayoría de trabajos que tratan series con más de un coeficiente de escalado, el punto de *cross-over* se ha estimado mediante inspección visual [48, 51, 89]. En otros mediante la estimación del cruce entre las dos regresiones lineales [90] o mediante Akaike [56, 92]. En esta tesis se han unificado las dos últimas para hacer una estimación más robusta mediante el uso de Akaike y la minimización del error de aproximación entre  $\log(F(n))$  y su regresión lineal. Se ha observado como se obtienen resultados similares a los mostrados por Ogata [56, 123], donde el *cross-over* aparece en escalas temporales similares para los grupos y en DM aparece retrasado con respecto a NDM. En [51] se establece que los exponentes de escalado están influenciados por los ciclos circadianos y por la condición patológica del sujeto y en [52] se establece que cambios fisiológicos o patológicos en el sistema pueden cambiar la estructura de los  $\alpha$  en la señal. Los resultados obtenidos en esta tesis, muestran que los puntos de *cross-over* aparecen en escalas similares para los mismos grupos de sujetos independientemente de la situación crítica de los mismos,  $\log(n) = 1.20$  en NDM críticos y 1.38 en NDM ambulatorios, 1.6 para DM críticos y 1.9 para DM ambulatorios, Ogata establece el *cross-over* en 1.38 para sujetos sanos y 1.53 para sujetos DM [56]. Se observa cómo la tendencia es similar. A corto plazo  $\alpha_{sr}$ , la señal es más regular que a largo plazo,  $\alpha_{lr}$ , pero también es donde se encuentran las diferencias entre los distintos

grupos tanto para glucemia ambulatoria como glucemia en pacientes críticos donde [54] no era capaz de encontrar diferencias entre DM y NDM. Estos resultados sugieren que el punto de *cross-over* puede determinar el grado de la patología de los sujetos con síndrome metabólico de forma que se puede evitar el desarrollo de la patología en los mismos. En señales de dinámicas de movilidad, al igual que sucedía en situaciones anteriores, no se encuentran diferencias significativas entre grupos salvo CT y HD. Esto sugiere que ninguna de las medidas consideradas consigue dilucidar las diferencias existentes entre los patrones de movimiento.

Finalmente se han caracterizado métricas usando señales de hipertensión arterial (HTA) de apenas 55 muestras de longitud y señales de glucosa con longitudes inferiores a 108 muestras. Considerando que ApEn, SampEn y MSE no han sido capaces de establecer resultados estadísticamente robustos sobre señales de 300 muestras, difícilmente van a poder analizar estas señales. DFA tampoco es factible puesto que las escalas temporales en las que analizar se limitarían a 4 o 5 ventanas distintas, lo que no podría ofrecer un resultados suficientemente robusto. Se investigó entonces la existencia de otras medidas que fueran más indicadas para registros de estas longitudes como la entropía cuadrada de Rényi o entropía muestral de orden 2 (QSE) [17] o el coeficiente de entropía muestral (CosEn) [23, 82], consiguiendo identificar AF con apenas 12 latidos. En esta tesis se aplicó QSE y CosEn a las señales de HTA, obteniendo diferencias significativas con ambas medidas de regularidad, siendo mejores en CosEn que QSE. Esto es un hecho muy relevante, ya que nunca antes se habían analizado registros de presión arterial en humanos de tan poca longitud. Generalmente se han analizado sobre animales [69, 70] o derivadas de alguna otra señal biológica [51, 125, 126, 127], con una longitud mínima de al menos 1000 muestras que permite utilizar técnicas de análisis no lineal tradicionales.

Las señales de glucemia post-operatoria se han analizado únicamente con CosEn ya que en HTA ofrecía mejor capacidad de segmentación entre grupos que QSE. Como se ha mostrado en los resultados, esta medida de regularidad consiste en estimar el número de patrones coincidentes ( $M$ ) que maximiza la distancia entre clases y calcular el umbral  $r$  en función del mismo a través de una relación directa. El algoritmo desarrollado para optimizar  $M$  es costoso ya que integra métodos de *bootstrapping* para reducir el sesgo asociado a un limitado número de señales. Se ha mostrado que este método depende en gran medida de la técnica a utilizar para la estimación del máximo de las curvas AUC, en particular el filtro utilizado para reducir el ruido de las mismas. Se ha mostrado cómo la entropía se ve claramente afectada y que se debe buscar el mejor filtro posible. En este trabajo se ha mostrado filtrado de media móvil de distintas longitudes y filtrado mediante técnicas Wavelet. Los resultados indican que se debe aplicar un filtro adaptado a la señal (AUC) que reduzca de forma eficiente las variaciones rápidas y detecte de forma unánime el máximo de la AUC.

## Capítulo 9

# Conclusión

En esta tesis se ha abordado una caracterización de medidas de entropía frente a la pérdida de muestras o la ausencia de éstas en registros de corta duración. Se ha caracterizado el efecto de la pérdida de muestras tanto de forma aleatoria como uniforme sobre tres bases de datos de señales biológicas: EEG, RR y EGM. Se han caracterizado señales de corta duración donde  $N < 300$ : glucemia y dinámicas de movilidad; y de muy corta duración como hipertensión arterial ( $N = 55$ ) o glucemia post-operatoria ( $N \leq 108$ ). Las medidas de complejidad caracterizadas han sido seleccionadas por ser de las más comunes en este contexto biomédico como ApEn, SampEn, MSE y DFA o bien por ser de reciente desarrollo y adecuadas a la corta longitud de los registros como QSE o CosEn.

La pérdida de muestras desorganizada (de forma aleatoria) provoca un incremento en el valor de la entropía asociado a la aleatorización introducida en la señal. Este incremento es tanto mayor cuanto más regular es en origen la señal analizada. La pérdida de un segmento continuo de muestras (de forma uniforme) se traduce en una conservación de la entropía independientemente del número de muestras que ha perdido la señal.

ApEn presenta menor consistencia que SampEn. En igualdad de condiciones (mismos porcentajes de pérdida de muestras y registros analizados), presenta menores coeficientes de correlación cruzada. ApEn es también muy sensible a la longitud de los registros y no debería aplicarse siempre que el número de muestras perdidas en una señal sea superior al 50 % cuando la pérdida de muestras se produce de forma aleatoria o un 40 % cuando es uniforme. Por debajo de esos porcentajes de muestras se asegura que la entropía mantenga validez clínica ( $C_{xy} \geq 0.8$ ).

SampEn es muy insensible a la pérdida de muestras siempre y cuando la señal a analizar no sea determinista. Se puede aplicar en condiciones inestables donde

el número de muestras descartadas o pérdidas es elevado, hasta un 70 %, ya que mantiene una alta correlación cruzada con la entropía de las series originales tanto para pérdidas tanto uniformes como aleatorias.

En cuanto a MSE, no se puede recomendar su uso en escalas mayores a 2 aún cuando aumenta la separabilidad entre clases en las tres bases de datos cuando las muestras se han perdido de forma aleatoria. En escalas mayores, a medida que aumenta el diezmado, muestra un comportamiento en la complejidad contrario a lo establecido en la literatura donde refiere mayor complejidad a señales patológicas. Para pérdida de muestras de forma uniforme el comportamiento es más estable, puede aplicarse sin llegar a usar escalas elevadas, superiores a 6 para evitar variaciones en el nivel de la complejidad o perder la capacidad de segmentación.

DFA refleja el mismo comportamiento ofreciendo una segmentación estable en todo el rango en señales EEG, no siendo el comportamiento del coeficiente de escalado similar en los tres tipos de registros cuando la pérdida de muestras se produce de forma aleatoria. Con pérdidas de forma uniforme EEG presenta un comportamiento distinto a EGM o RR, manteniendo la capacidad de segmentación en los tres tipos de registros. La variación en los comportamientos ante las mismas condiciones de pérdida de muestras puede deberse a que las señales dejen de ser monofractales y exista la necesidad de utilizar más de un coeficiente para su completa caracterización. De acuerdo a lo anterior y al coste computacional que tiene asociado el cálculo de DFA, no se recomienda esta técnica de análisis.

De acuerdo a los resultados expuestos, se puede concluir que para realizar análisis de entropía, la medida más estable es SampEn ya que se ve mínimamente alterada ante pérdida de muestras y reporta un comportamiento similar en las tres bases de datos analizadas. Aunque no consigue establecer una segmentación entre registros RR, esto puede deberse a que se está analizando la complejidad de los latidos ectópicos y no de la variabilidad propia de los registros RR. Para señales con *spikes* o estimaciones anómalas como son las derivadas de los latidos ectópicos en registros RR, se recomienda el uso de MSE en escalas 2 ó 3 que al aplicar un suavizado y decorrelación entre los subvectores ofrece mejores resultados en términos de segmentación y estabilidad de la medida.

Para señales de longitud inferior a 1000 muestras y de dinámicas de variación lentas como son glucemia y dinámicas de movilidad, no es aconsejable el uso de ApEn, SampEn ni MSE ya que no ofrecen resultados robustos al no encontrar un número suficiente de patrones que garanticen el cálculo de las probabilidades marginales de forma fiable. Tampoco consiguen establecer separación entre grupos de forma significativa. Por ello no se recomienda su uso para el análisis de estas señales. DFA consigue establecer diferencias allá donde otros métodos no lo podían hacer en estas señales. Utilizando un algoritmo de más de un coeficiente de escalado se encuentran diferencias significativas entre los grupos DM y NDM en registros de glucemia en pacientes críticos. El *cross-over* puede indicar el grado de patología de un sujeto al

---

aparecer más retrasado en MS y DM respectivamente o el tiempo necesario para la activación completa de sistema glucorregulador, ya que la regularidad de la señal tras este punto es similar en los tres grupos de los registros de glucosa ambulatoria y el método es incapaz de establecer diferencias significativas.

Finalmente, para señales de muy corta duración ( $N \leq 100$ ), se sugiere el uso de CosEn donde se debe desarrollar un filtro genérico que se adapte a las AUC para que elimine de forma eficiente las transiciones rápidas y permita determinar el  $M$  en el que se maximiza la AUC y sí establecer una separación entre clases óptima. Si el número de señales es pequeño ( $N \leq 30$ ), se recomienda el uso de *Leave-One-Out* para evitar el sesgo asociado a *K-fold*. En caso de seleccionar el segundo se sugiere  $K=10$ .



## Capítulo 10

# Líneas futuras

En este trabajo se ha presentado una caracterización de ApEn, SampEn, MSE y DFA sobre registros de EEG, RR y EGM que han sufrido pérdida de muestras tanto de forma aleatoria como uniforme. De las 4 medidas caracterizadas, SampEn, y en consecuencia MSE en algunas escalas eran las medidas más robustas frente a la pérdida de muestras. Existe poca literatura en la cual se valoran estas situaciones y casi toda se centra en el análisis de registros conciernes a la actividad cardíaca [19, 22]. Se debería profundizar en un estudio poco abordado, y seguir caracterizando distintos tipos de registros de uso común en el diagnóstico de patologías ante estas situaciones. Por ejemplo, en [22] se propone el uso de la representación *Poincaré* ya que es más robusta para realizar el análisis de HRV cuando se han perdido intervalos RR de los cuales se ha derivado. También se debería ampliar el espectro de medidas caracterizadas a otras de uso común como podrían ser la complejidad de Lempel–Ziv [29], entropía Wavelet [128, 129], dimensión de correlación [130], exponentes de Lyapunov [131], dimensión fractal de Higuchi [132], información mútua [133], entropía de Rényi [11] o a otras más recientes como la entropía borrosa [134] o la entropía permutada [135], entre otras.

En señales de menos de 1000 muestras las medidas caracterizadas no ofrecen resultados robustos, siendo la que mejor comportamiento refleja DFA. Se propone el estudio de la entropía modificada en escalas múltiples, reajustada para poder procesar señales a partir de 500 muestras [136].

Todas las señales no presentan un comportamiento monofractal puro y algunas de ellas necesitan más de un exponente de escalado ya que su comportamiento fractal cambia a lo largo de las distintas escalas temporales. Por ello se recomienda el estudio de la entropía multifractal [28, 137, 138] donde se ofrece un espectro completo de la dimensión fractal de las señales. Se ha visto cómo el punto de *cross-over* puede dar información relativa al grado de patología diabética en los

sujetos, lo que resulta de especial interés a la hora de analizar la probabilidad de un sujeto para desarrollar o no la enfermedad. Conviene implementar un método robusto que determine este punto de forma exacta y profundizar un poco más en su estudio en colaboración con clínicos expertos.

En cuanto a señales de muy corta duración ( $N < 100$ ), no hay muchos estudios en la literatura, excepto los presentados en este trabajo sobre RSE y CosEn. Se han propuesto en la literatura la entropía borrosa aproximada [139] donde la distancia para calcular los patrones coincidentes pasa de ser una función escalón a una gaussiana, y puede extrapolarse a otras medidas de regularidad. La entropía muestral generalizada (qSampEn) [140] donde el logaritmo ya no es natural sino en base  $q$ . La entropía muestral modificada (mSampEn) [141] aplicada sobre procesos MIX de apenas 50 muestras con resultados robustos.

Para finalizar, se ha visto que todas las medidas dependen de algún tipo de parámetro que condiciona su comportamiento. En el caso de ApEn y SampEn, tenemos  $m$  y  $r$ , para MSE tenemos también la escala  $s$ , con DFA tenemos el orden de integración  $l$  y la longitud máxima de la secuencia de inventariado, qSampEn introduce la base del logaritmo  $q$  como un nuevo parámetro. En este trabajo que han realizado parametrizaciones para las señales de corta duración o aplicado parámetros estándar en los que se asume que el comportamiento de las medidas es adecuado, por ello se propone analizar la influencia de dichos parámetros y valorar su optimización en función de las señales de forma que sean establecidos de forma autónoma por el propio algoritmo. Actualmente se han propuesto algunas metodologías para optimizar  $r$  [23, 42, 142, 143] o seleccionar  $m$  [17, 23] pero dichas técnicas incrementan considerablemente el coste computacional del algoritmo de cómputo de la medida sin introducir una mejora sustancial en el resultado.

# Capítulo 11

## Aportaciones de la tesis

### 11.1 Contribuciones a congresos

1. E.M.C. Roldán, A. Molina-Picó, D. Cuesta-Frau, P.M. Martínez, and S.O. Crespo. Characterization of entropy measures against data loss: Application to eeg records. In *Eng. Med. Biol. Soc. (EMBC) 2011 Ann. Int. Conf. IEEE*, pages 6110 – 6113, 2011
2. E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, and P. Miró-Martínez. Comparative study between sample entropy and detrended fluctuation analysis performance on eeg records under data loss. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 4233 – 4236, 2012
3. E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, P. Miró-Martínez, L. Vigil-Medina, and M. Varela-Entrecanales. Characterization of detrended fluctuation analysis in the context of glycemic time series. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 4225 – 4228, 2012
4. E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, P. Miró-Martínez, L. Vigil-Medina, and M. Varela-Entrecanales. Customization of entropy estimation measures for human arterial hypertension records segmentation. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 33 – 36, 2012

## 11.2 Artículos en revistas

1. E.M. Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, S. Oltra-Crespo, L. Vigil-Medina, and M. Varela-Entrecanales. A new algorithm for quadratic sample entropy optimization for very short biomedical signals: Application to blood pressure records. *Comp. Meth. Prog. Biomed.*, 114(3):231 – 239, 2014
2. E.M. Cirugeda-Roldán, A. Molina-Picó, D. Novak, D. Cuesta-Frau, and V. Křemen. Atrial electrogram complex fractionated entropy study. *Exp. Clin. Cardiol.*, 2014
3. L. Vigil, E. Condés, M. Varela, C. Rodriguez, A. Colas, B. Vargas, M. Lopez, and E. Cirugeda. Glucose series complexity in hypertensive patients. *J. Am. Soc. Hyperten. (in press)*, 2014
4. E.M. Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, S. Oltra-Crespo, L. Vigil-Medina, and M. Varela-Entrecanales and C. Rodriguez-de Castro. Unsupervised crossover estimation in detrended fluctuation analysis: Application to blood glucose data. *submitted to Med. Eng. Phys.*, 2014
5. E.M. Cirugeda-Roldán, D. Novak, V. Křemen, D. Cuesta-Frau, M.W Keller, C. Schilling, O. Doessel, C. Schmitt, and A. Luik. A new algorithm for quadratic sample entropy optimization for very short biomedical signals: Application to blood pressure records. *submitted to PlosOne*, 2014
6. E.M. Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, and S. Oltra-Crespo. Comparative study of entropy sensitivity to missing biosignal data. *submitted Entropy*, 2014
7. J.L. Rodríguez-Sotelo, D. Cuesta-Frau A. Osorio-Forero, A. Jiménez-Rodríguez, E. Cirugeda-Roldán, and D. Peluffo. Automatic sleep stages classification using eeg entropy features and unsupervised pattern analysis techniques. *submitted Entropy*, 2014

# Bibliografía

- [1] A. Babloyantz and A. Destexhe. Is the normal heart a periodic oscillator? *Biol. Cyber.*, 58(3):203 – 211, 1988.
- [2] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiol. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circ.*, 93(5):1043 – 1065, 1996.
- [3] J.M. Lee, D.J. Kim, I.Y. Kim, K.S. Park, and S.I. Kim. Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data. *Comp. in Biol. and Med.*, 32(1):37 – 47, 2002.
- [4] S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C.M. Hill, and P.R. White. Signal processing techniques applied to human sleep EEG signals – a review. *Biomed. Sig. Process. Control.*, 10(0):21 – 33, 2014.
- [5] F. Censi, V. Barbaro, P. Bartolini, G. Calcagnini, A. Michelucci, G.F. Gensini, and S. Cerutti. Recurrent patterns of atrial depolarization during atrial fibrillation assessed by recurrence plot quantification. *Ann. Biomed. Eng.*, 28(1):61 – 70, 2000.
- [6] D.T. Kaplan and R.J. Cohen. Is fibrillation chaos? *Circ. Res.*, 67(4):886 – 892, 1990.
- [7] A. Casaleggio, A. Corana, and S. Ridella. Correlation dimension estimation from electrocardiograms. *Chaos Solitons Fractals*, 5(3 – 4):713 – 726, 1995.
- [8] D.K. Peng, S. Havlin, H.E. Stanley, and Goldberger A.L. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5(1):82 – 87, 1995.
- [9] S.M. Pincus. Approximate entropy as a measure of system complexity. *Phys. Rev. E*, 88(6):2297 – 2301, 1991.

- [10] J.S. Richman and J.R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart. Circ. Physiol.*, 278(6):H2039 – 2049, 2000.
- [11] A. Rényi. On measures of entropy and information. In *Proc. Fourth Berkely Symp. Math. Statist. Prob. 1960*, volume 1, pages 547 – 561. University of California Press, Berkeley, 1961.
- [12] A. Wolf, J.B Swift, H.L. Swinney, and J.A. Vastano. Determining lyapunov exponents from a time series. *Phys. D: Nonlin. Phenom.*, 16(3):285 – 317, 1985.
- [13] P. Grassberger and I. Procaccia. Estimation of the kolmogorov entropy from a chaotic signal. *Phys. Rev. A*, 28(4), 1983.
- [14] S. Pincus, I. Gladstone, and R. Ehrenkranz. A regularity statistic for medical data analysis. *J. Clin. Monit. Comput.*, 7:335 – 345, 1991.
- [15] M. Costa, A.L. Goldberger, and C.K. Peng. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.*, 89(6):068102–1 – 068102–4, 2002.
- [16] J. Bruhn, J. Röpcke, B. Rehberg, T. Bouillon, and A. Hoeft. Electroencephalogram approximate entropy correctly classifies the occurrence of burst suppression pattern as increasing anesthetic drug effect. *Anesthesiol.*, 93(4):981 – 985, 2000.
- [17] D.E. Lake, J.S. Richman, M.P. Griffin, and J.R. Moorman. Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 283(3):R789 – 797, 2002.
- [18] M.A. García-González, M. Fernández-Chimeno, and J. Ramos-Castro. Errors in the estimation of approximate entropy and other recurrence-plot-derived indexes due to the finite resolution of RR time-series. *IEEE Trans. Biomed. Eng.*, 56(2):345 – 351, 2009.
- [19] R.B. Govindan, J.D. Wilson, H. Preissl, H. Eswaran, J.Q. Campbell, and C.L. Lowery. Detrended fluctuation analysis of short datasets: An application to fetal cardiac data. *Phys. D: Nonlin. Phenom.*, 226(1):23 – 31, 2007.
- [20] D.E. Lake. Rényi entropy measures of heart rate gaussianity. *IEEE Trans. Biomed. Eng.*, 53(1):21 – 27, 2006.
- [21] H. Jing and M. Takigawa. Low sampling rate induces high correlation dimension on electroencephalograms from healthy subjects. *Psych. Clin. Neurosci.*, 54(4):407 – 412, 2000.

- 
- [22] K.K. Kim, H.J. Baek, Y.G. Lim, and K.S. Park. Effect of missing RR-interval data on nonlinear heart rate variability analysis. *Comp. Meth. Prog. Biomed.*, 106(3):210 – 218, 2012.
- [23] D.E. Lake and J.R. Moorman. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Physiol. Heart. Circ. Physiol.*, 300(1):H319 – H325, 2011.
- [24] A.L. Goldberger, V. Bhargava, B.J. West, and A.J. Mandell. On a mechanism of cardiac electrical stability. the fractal hypothesis. *Biophys. J.*, 48(3):525 – 528, 1985.
- [25] A.L. Goldberger and J. Bruce. Fractals in physiology and medicine. *Yale J. Biol. Med.*, 60(5):421 – 435, 1987.
- [26] H.S. Greenside, A. Wolf, J. Swift, and T. Pignataro. Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors. *Phys. Rev. A*, 25:3453 – 3456, 1982.
- [27] M. Costa, A.L. Goldberger, and C.K. Peng. Multiscale entropy to distinguish physiologic and synthetic RR time series. *Comp. Cardiol.*, 29:137 – 140, 2002.
- [28] P.Ch. Ivanov, L.A.N. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z.R. Struzik, and H.E. Stanley. Multifractality in human heartbeat dynamics. *Nat.*, 399(6735):461 – 465, 1999.
- [29] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inform. Theory*, 22(1):75 – 81, 1976.
- [30] S. Zozor, P. Ravier, and O. Buttelli. On Lempel–Ziv complexity for multidimensional data analysis. *Phys. A: Stat. Mech. Appl.*, 345(1 – 2):285 – 302, 2005.
- [31] H.B. Xie, W.T. Chen, W.X. He, and H. Liu. Complexity analysis of the biomedical signal using fuzzy entropy measurement. *Appl. Soft Comput.*, 11(2):2871 – 2879, 2011.
- [32] G.E. Powell and I.C. Percival. A spectral entropy method for distinguishing regular and irregular motion of hamiltonian systems. *J. Phys. A: Mathem. and Gen.*, 12(11):2053, 1979.
- [33] M. Mäenpää, T. Laitio, T. Kuusela, J. Penttilä, K. Kaisti, S. Aalto, S. Hinkka-Yli-Salomäki, and H. Scheinin. Delta entropy of heart rate variability along with deepening anesthesia. *Anesth. Analg.*, 112(3):587 – 592, 2011.
- [34] C. Shannon. Bell sys. *Tech. J.*, 27:379, 1948.

- [35] D. Simpelaere. Correlation dimension. *J. Stat. Phys.*, 90(1 – 2):491 – 509, 1998.
- [36] O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schürmann, and E. Basar. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *J. Neurosci. Meth.*, 105(1):65 – 75, 2001.
- [37] S. Pincus. Approximate entropy ( $\{ApEn\}$ ) as a complexity measure. *Chaos*, 5(1):110 – 117, 1995.
- [38] S.M. Pincus and A.L. Goldberger. Physiological time-series analysis: what does regularity quantify? *Am. J. Physiol.*, 266(4 pt.2):H1643 – 1656, 1994.
- [39] M. Ferrario, M.G. Signorini, G. Magenes, and S. Cerutti. Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress. *IEEE Trans. Biomed. Eng.*, 53(1):119 – 125, 2006.
- [40] H Ocak. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Syst. Appl.*, 36(2, Part 1):2027 – 2036, 2009.
- [41] H.M. Al-Angari and A.V. Sahakian. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *IEEE Trans. Biomed. Eng.*, 54(10):1900 – 1904, 2007.
- [42] R. Alcaraz and J.J. Rieta. A novel application of sample entropy to the electrocardiogram of atrial fibrillation. *Nonlin. Anal.: Real World Appl.*, 11(2):1026 – 1035, 2010.
- [43] R. Alcaraz and J.J. Rieta. Non-invasive organization variation assessment in the onset and termination of paroxysmal atrial fibrillation. *Comp. Meth. Prog. Biomed.*, 93(2):148 – 154, 2009.
- [44] R. Alcaraz and J.J. Rieta. Surface  $\{ECG\}$  organization analysis to predict paroxysmal atrial fibrillation termination. *Comp. in Biol. and Med.*, 39(8):697 – 706, 2009.
- [45] R.C. Hwa and T.C. Ferree. Scaling properties of fluctuations in the human electroencephalogram. *Phys. Rev. E*, 66(2):021901, 2002.
- [46] D. Abasolo, R. Hornero, J. Escudero, and P. Espino. A study on the possible usefulness of detrended fluctuation analysis of the electroencephalogram background activity in alzheimer’s disease. *IEEE Trans. Biomed. Eng.*, 55(9):2171 – 2179, 2008.
- [47] R.L. Burr, C.J. Kirkness, and P.H. Mitchell. Detrended fluctuation analysis of intracranial pressure predicts outcome following traumatic brain injury. *IEEE Trans. Biomed. Eng.*, 55(11):2509 – 2518, 2008.

- 
- [48] J.M. Lee, D.J. Kim, I.Y. Kim, K.S. Park, and S.I. Kim. Nonlinear-analysis of human sleep EEG using detrended fluctuation analysis. *Med. Eng. Phys.*, 26(9):773 – 776, 2004.
- [49] R. Ferri, F. Rundo, O. Bruni, M.G. Terzano, and C.J. Stam. Dynamics of the {EEG} slow-wave synchronization during sleep. *Clin. Neurophysiol.*, 116(12):2783 – 2795, 2005.
- [50] S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Matsu, C.K. Peng, M. Simons, and H.E. Stanley. Long-range correlation properties of coding and noncoding DNA sequences: Genbank analysis. *Phys. Rev. E*, 51:5084 – 5091, 1995.
- [51] P. Castiglioni, G. Parati, A. Civijian, L. Quintin, and M. Di Rienzo. Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: Effects of posture, exercise, and aging. *IEEE Trans. Biomed. Eng.*, 56(3):675 – 684, 2009.
- [52] P. Castiglioni, G. Parati, M. Di-Rienzo, R. Carabalona, A. Cividjian, and L. Quintin. Scale exponents of blood pressure and heart rate during autonomic blockade as assessed by detrended fluctuation analysis. *J. Physiol.*, 589(2):355 – 369, 2011.
- [53] G.M. Viswanathan, C.K. Peng, H.E. Stanley, and A.L. Goldberger. Deviations from uniform power law scaling in nonstationary time series. *Phys. Rev. E*, 55:845 – 849, 1997.
- [54] D. Lundelin, L. Vigil, S. Bua, I. Gómez-Mestre, T. Honrubia, and M. Varela. Differences in complexity of glycemic profile in survivors and nonsurvivors in an intensive care unit: a pilot study. *Crit. Care Med.*, 38:849 – 854, 2010.
- [55] J. Churruca, L. Vigil, E. Luna, J. Ruiz-Galiana, and M. Varela. The route to diabetes: Loss of complexity in the glycemic profile from health through the metabolic syndrome to type 2 diabetes. *Diab., Metab. Syndr. and Obes.*, 1:3 – 11, 2008.
- [56] H. Ogata, K. Tokuyama, S. Nagasaka, T. Tsuchita, I. Kusaka, S. Ishibashi, H. Suzuki, N. Yamada, K. Hamano, K. Kiyono, Z.R. Struzik, and Y. Yamamoto. The lack of long-range negative correlations in glucose dynamics is associated with worse glucose control in patients with diabetes mellitus. *Metabolism*, 61(7):1041 – 1050, 2012.
- [57] H. Ogata, K. Tokuyama, S. Nagasaka, A. Ando, I. Kusaka, N. Sato, A. Go-to, S. Ishibashi, K. Kiyono, Z.R. Struzik, and Y. Yamamoto. Long-range correlated glucose fluctuations in diabetes. *Meth. Inf. Med.*, 46:222 – 226, 2007.

- [58] N.A. Khovanova, I.A. Khovanov, L. Sbrano, F. Griffiths, and T.A. Holt. Characterisation of linear predictability and non-stationarity of subcutaneous glucose profiles. *Comp. Meth. Prog. Biomed.*, 110(3):260 – 267, 2013.
- [59] A. Phinyomark, P. Phukpattaranont, and C. Limsakul. Fractal analysis features for weak and single-channel upper-limb EMG signals. *Expert Syst. Appl.*, 39(12):11156 – 11163, 2012.
- [60] R.A. Thuraisingham and G.A. Gottwald. On multiscale entropy analysis for physiological data. *Phys. A. Stat. Mech. Appl.*, 366:323 – 332, 2006.
- [61] M. Baumert, M. Javorka, A. Seeck, R. Faber, P. Sanders, and A. Voss. Multiscale entropy and detrended fluctuation analysis of qt interval and heart rate variability during normal pregnancy. *Comp. Biol. Med.*, 42(3):347 – 352, 2012.
- [62] A. Catarino, O. Churches, S. Baron-Cohen, A. Andrade, and H. Ring. Atypical EEG complexity in autism spectrum conditions: A multiscale entropy analysis. *Clin. Neurophysiol.*, 122(12):2375 – 2383, 2011.
- [63] T. Mizuno, T. Takahashi, R.Y. Cho, M. Kikuchi, T. Murata, K. Takahashi, and Y. Wada. Assessment of EEG dynamical complexity in alzheimer’s disease using multiscale entropy. *Clin. Neurophysiol.*, 121(9):1438 – 1446, 2010.
- [64] T. Takahashi, R.Y. Cho, T. Mizuno, M. Kikuchi, T. Murata, K. Takahashi, and Y. Wada. Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: A multiscale entropy analysis. *NeuroImage*, 51(1):173 – 182, 2010.
- [65] R. Carvajal, N. Wessel, M. Vallverdú, P. Caminal, and A. Voss. Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy. *Comp. Meth. Prog. Biomed.*, 78(2):133 – 140, 2005.
- [66] S. Janjarasjitt, M.S. Scher, and K.A. Loparo. Nonlinear dynamical analysis of the neonatal eeg time series: The relationship between neurodevelopment and complexity. *Clin. Neurophysiol.*, 119(4):822 – 836, 2008.
- [67] T. Kobayashi, K. Misaki, H. Nakagawa, S. Madokoro, T. Ota, H. Ihara, K. Tsuda, Y. Umezawa, J. Murayama, and K. Isaki. Correlation dimension of the human sleep electroencephalogram. *Psychiatry Clin. Neurosci.*, 54(1):11 – 16, 2000.
- [68] H. Strengé, G. Fritzer, R. Göder, U. Niederberger, W.D. Gerber, and J. Aldenhoff. Non-linear electroencephalogram dynamics in patients with spontaneous nocturnal migraine attacks. *Neurosci. Lett.*, 309(2):105 – 108, 2001.

- 
- [69] Y. Almog, O. Oz, and S. Akselrod. Correlation dimension estimation: can this nonlinear description contribute to the characterization of blood pressure control in rats? *IEEE Trans. Biomed. Eng.*, 46(5):535 – 547, 1999.
- [70] R. Hornero, M. Aboy, C. Gómez, D.S. Hagg, and C.R. Phillips. Complexity analysis of arterial pressure during periods of abrupt hemodynamic changes. *IEEE Trans. Biomed. Eng.*, 55(2):797 – 801, 2008.
- [71] Jing H., Jianbo G., and J.C. Principe. Analysis of biomedical signals by the Lempel–Ziv complexity: the effect of finite data size. *IEEE Trans. Biomed. Eng.*, 53(12):2606 – 2609, 2006.
- [72] A. Bershadskii, E. Dremencov, D. Fukayma, and G. Yadid. Multifractal properties of brain neuron signals. *Europhys. Lett.*, 58(2):306, 2002.
- [73] Y. Zheng, J. Gao, J.C. Sanchez, J.C. Principe, and M.S. Okun. Multiplicative multifractal modeling and discrimination of human neuronal activity. *Phys. Lett. A*, 344(2-4):253 – 264, 2005.
- [74] A.L. Goldberger, L.A.N. Amaral, J.M. Hausdorff, P.Ch. Ivanov, C.K. Peng, and H.E. Stanley. Fractal dynamics in physiology: Alterations with disease and aging. *Proc. Nat. Acad. Sci. USA*, 99(Suppl 1):2466 – 2472, 2002.
- [75] P. Laguna and L. Sörnmo. *Bioelectrical signal processing in cardiac and neurological applications*. Elsevier Academic Press, 2005.
- [76] E.S.G. Carotti, J.C. De Martin, R. Merletti, and D. Farina. Compression of multidimensional biomedical signals with spatial and temporal codebook–excited linear prediction. *IEEE Trans. Biomed. Eng.*, 56(11):2604 – 2610, 2009.
- [77] A.J. Casson and E. Rodríguez-Villegas. Toward online data reduction for portable electroencephalography systems in epilepsy. *IEEE Trans. Biomed. Eng.*, 56(12):2816 – 2825, 2009.
- [78] J.R. Gallego, A. Hernandez-Solana, M. Canales, J. Lafuente, A. Valdovinos and, and J. Fernandez-Navajas. Performance analysis of multiplexed medical data transmission for mobile emergency care over the umts channel. *IEEE Trans. Inform. Technol. Biomed.*, 9(1):13 – 22, 2005.
- [79] J.C. Chan and P.W. Tse. A novel, fast, reliable data transmission algorithm for wireless machine health monitoring. *IEEE Trans. Rel.*, 58(2):295 – 304, 2009.
- [80] S. Pavlopoulos, E. Kyriacou, A. Berler, S. Dembeyiotis, and E. Koutsouris. A novel emergency telemedicine system based on wireless communication technology–{AMBULANCE}. *IEEE Trans. Inform. Technol. Biomed.*, 2(4):261 – 267, 1998.

- [81] R. Watts. Some design considerations for narrow-band medical telemetry over the switched-message network. *IEEE Trans. Wireless Commun.*, 19(3):246 – 255, 1971.
- [82] P. Langley, M. Dewhurst, L.Y. Di-Marco, P. Adams, F. Dewhurst, J.C. Mwita, R. Walker, and A. Murray. Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings. *Med. Eng. Phys.*, 34(10):1441 – 1447, 2012.
- [83] J.E. Deffeyes, R.T. Harbourne, W.A. Stuberg, and N. Stergiou. Approximate entropy used to assess sitting postural sway of infants with developmental delay. *Infant Behav. and Develop.*, 34(1):81 – 99, 2011.
- [84] J.S. Richman. Sample entropy statistics and testing for order in complex physiological signals. *Com. in Stat. – Theory Meth.*, 36(5):1005 – 1019, 2007.
- [85] M.J. Lewis and A.L. Short. Sample entropy of electrocardiographic RR and QT time-series data during rest and exercise. *Physiol. Meas.*, 28(6):731 – 744, 2007.
- [86] K. Hu, P.C. Ivanov, Z. Chen, P. Carpena, and E.H. Stanley. Effect of trends on detrended fluctuation analysis. *Phys. Rev. E*, 64:011114, 2001.
- [87] M. Jospin, P. Caminal, E.W. Jensen, H. Litvan, M. Vallverdu, M.M.R.F. Struys, H.E.M. Vereecke, and D.T. Kaplan. Detrended fluctuation analysis of EEG as a measure of depth of anesthesia. *IEEE Trans. Biomed. Eng.*, 54(5):840 – 846, 2007.
- [88] Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galván, H.E. Stanley, and P.C. Ivanov. Effect of nonlinear filters on detrended fluctuation analysis. *Phys. Rev. E*, 71:011104, 2005.
- [89] R. Nagarajan and R.J. Kavasseri. Minimizing the effect of periodic and quasi-periodic trends in detrended fluctuation analysis. *Chaos, Solitons & Fractals*, 26(3):777 – 784, 2005.
- [90] J. Yue, X. Zhao, and P. Shang. Effect of trends on detrended fluctuation analysis of precipitation series. *Math. Prob. Eng.*, 2010:15, 2010.
- [91] N. Xu, P. Shang, and S. Kamae. Minimizing the effect of exponential trends in detrended fluctuation analysis. *Chaos, Solitons & Fractals*, 41(1):311 – 316, 2009.
- [92] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Automat Cont.*, 19(6):716 – 723, 1974.
- [93] C.E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, 1949.

- 
- [94] P. Henríquez, J.B. Alonso, M.A. Ferrer, C.M. Travieso, J.I. Godino-Llorente, and F. Díaz-de María. Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE Trans. Speech Audio Processing*, 17(6):1186 – 1195, 2009.
- [95] L. Golshani, E. Pasha, and G. Yari. Some properties of Rényi entropy and Rényi entropy rate. *Inform. Sci.*, 179(14):2426 – 2433, 2009.
- [96] D. Erdogmus, K.E. Hild-II, and J.C. Principe. Blind source separation using Rényi  $\alpha$ -marginal entropies. *Neurocomp.*, 49(1 – 4):25 – 38, 2002.
- [97] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, and O.J.J. Michel. Measuring time–frequency information content using the rényi entropies. *IEEE Trans. Inform. Theory*, 47(4):1391 – 1409, 2001.
- [98] J. Gao, J. Hu, T. Buckley, K. White, and C. Hass. Shannon and Rényi entropies to classify effects of mild traumatic brain injury on postural sway. *PLoS ONE*, 6(9):e24446, 2011.
- [99] Tom Fawcett. Roc graphs: Notes and practical considerations for researchers. Technical report, 2004.
- [100] K.H. Zou, A.J. O’Malley, and L. Mauri. Receiver–operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circ.*, 115(5):654 – 657, 2007.
- [101] T.A. Lasko, J.G. Bhagwat, K.H. Zou, and L. Ohno-Machado. The use of receiver operating characteristic curves in biomedical informatics. *J. Biomed. Inf.*, 38(5):404 – 415, 2005. Clinical Machine Learning.
- [102] T. Fushiki. Estimation of prediction error by using k–fold cross–validation. *Stat. Comp.*, 21(2):137 – 146, 2011.
- [103] D. Clonda, J.M. Lina, and B. Goulard. Complex daubechies wavelets: properties and statistical image modelling. *Signal Proc.*, 84(1):1 – 23, 2004.
- [104] M. Jarque, A.K. Bera, C.M. Jarque, and A.K. Bera. A test for normality of observations and regression residuals. *Int. Statist. Rev.*, pages 163 – 172, 1987.
- [105] K. Shimizu. Telemedicine by mobile communication. *IEEE Eng. Med. Biol. Mag.*, 18(4):32 – 44, 1999.
- [106] R.G. Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David, and C.E. Elger. Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *IEEE Trans. Knowledge Data Eng.*, 64(061907):1 – 8, 2001.

- [107] M.A. Brown, M.L. Buddle, M. Bennett, B. Smith, R. Morris, and J.A. Whitworth. Ambulatory blood pressure in pregnancy: Comparison of the Space-labs 90207 and Accutracker II monitors with intraarterial recordings. *Am. J. Obstet. Gynecol.*, 173(1):218 – 223, 1995.
- [108] W.A. Knaus, E.A. Draper, D.P. Wagner, and J.E. Zimmerman. APACHE II: A severity of disease classification system. *Crit. Care Med.*, 13(10):818 – 829, 1985.
- [109] F.J. Service, G.D. Molnar, J.W. Rosevear, E. Ackerman, L.C. Gatewood, and W.F. Taylor. Mean amplitude of glycemic excursions, a measure of diabetic instability. *Diabetes*, 19(9):644 – 655, 1970.
- [110] S.M. Grundy, H.B. Brewer, J.I. Cleeman, S.C. Smith, C. Lénfant, and for the Conference Participants. Definition of metabolic syndrome. *Circ.*, 109(3):433 – 438, 2004.
- [111] A.S. Jadidi, E. Duncan, S. Miyazaki, N. Lellouche, A.J. Shah, A. Forclaz, I. Nault, M. Wright, L. Rivard, X. Liu, D. Scherr, S.B. Wilton, F. Sacher, N. Derval, S. Knecht, S.J. Kim, M. Hocini, S. Narayan, M. Haïssaguerre, and P. Jaïs. Functional nature of electrogram fractionation demonstrated by left atrial high-density mapping. *Circ. Arrhythm. Electrophysiol.*, 5(1):32 – 42, 2012.
- [112] V. Jacquemet and C.S. Henríquez. Genesis of complex fractionated atrial electrograms in zones of slow conduction: A computer model of microfibrosis. *Heart Rhythm*, 6(6):803 – 810, 2009.
- [113] P.K. Stein, R.E. Kleiger, P.P. Domitrovich, K.B. Schechtman, and J.N. Rottman. Clinical and demographic determinants of heart rate variability in patients post myocardial infarction: Insights from the cardiac arrhythmia suppression trial {(CAST)}. *Clin. Cardiol.*, 23(3):187 – 194, 2000.
- [114] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circ.*, 101(23):e215 – e220, 2000.
- [115] S.M.S. Jalaliddine, C.G. Hutchens, R.D. Strattan, and W.A. Coberly. ECG data compression techniques – a unified approach. *IEEE Trans. Biomed. Eng.*, 37(4):329 – 343, 1990.
- [116] L. Berthouze and S.F. Farmer. Adaptive time-varying detrended fluctuation analysis. *J. Neurosci. Meth.*, 209(1):178 – 188, 2012.
- [117] E.M. Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, S. Oltra-Crespo, L. Vigil-Medina, and M. Varela-Entrecanales. A new algorithm for quadratic

- sample entropy optimization for very short biomedical signals: Application to blood pressure records. *Comp. Meth. Prog. Biomed.*, 114(3):231 – 239, 2014.
- [118] M.D. Costa, W.T. Schnettler, C. Amorim-Costa, J. Bernardes, A. Costa, A.L. Goldberger, and D. Ayres-de Campos. Complexity–loss in fetal heart rate dynamics during labor as a potential biomarker of acidemia. *Early Human Develop.*, 90(1):67 – 71, 2014.
- [119] T. Penzel, J.W. Kantelhardt, L. Grote, J.H. Peter, and A. Bunde. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans. Biomed. Eng.*, 50(10):1143 – 1151, 2003.
- [120] J.H. Lim, E.J. Khang, T.H. Lee, I.H. Kim, S.E. Maeng, and J.W. Lee. Detrended fluctuation analysis and Kolmogorov Sinai entropy of electroencephalogram signals. *Phys. Lett. A*, 377(38):2542 – 2545, 2013.
- [121] J.W. Kim, H. Shin, E.J. Kim, Y.J. Koo, B. Choi, K.S. Park, and D.U. Jeong. Characteristic time scales of electroencephalograms of narcoleptic patients and healthy controls. *Comput. Biol. Med.*, 40(10):831 – 838, 2010.
- [122] M. Costa, C.K. Peng, A.L. Goldberger, and J.M. Hausdorff. Multiscale entropy analysis of human gait dynamics. *Phys. A. Stat. Mech. Appl.*, 330(1):53 – 60, 2003.
- [123] H. Ogata, K. Tokuyama, S. Nagasaka, A. Ando, I. Kusaka, N. Sato, A. Goto, S. Ishibashi, K. Kiyono, Z.R. Struzik, and Y. Yamamoto. Long–range negative correlation of glucose dynamics in humans and its breakdown in diabetes mellitus. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 291(6):R1638 – R1643, 2006.
- [124] R. Brunner, G. Adelsmayr, H. Herkner, C. Madl, and U. Holzinger. Glycemic variability and glucose complexity in critically ill patients: a retrospective analysis of continuous glucose monitoring data. *Crit. Care*, 16(5):R175, 2012.
- [125] E. Kaniusas, H. Pfutzner, L. Mehnen, J. Kosel, C. Tellez-Blanco, G. Varoneckas, A. Alonderis, T. Meydan, M. Vazquez, M. Rohn, A.M. Merlo, and B. Marquardt. Method for continuous nondisturbing monitoring of blood pressure by magnetoelastic skin curvature sensor and ECG. *IEEE Sensors J.*, 6(3):819 – 828, 2006.
- [126] B. Subramaniam, K.R. Khabbaz, T. Heldt, A.B. Lerner, M.A. Mittleman, R.B. Davis, E.L. Goldberger, and M.D. Costa. Blood pressure variability: Can nonlinear dynamics enhance risk assessment during cardiovascular surgery? *J. Cardiothor. Vasc. Anesth.*, 28(2):392 – 397, 2014.

- [127] L. Angelini, R. Maestri, D. Marinazzo, L. Nitti, M. Pellicoro, G.D. Pinna, S. Stramaglia, and S.A. Tuppiti. Multiscale analysis of short term heart beat interval, arterial blood pressure, and instantaneous lung volume time series. *Artific. Intell. Med.*, 41(3):237 – 250, 2007.
- [128] B. Audit, E. Bacry, J.F. Muzy, and A. Arneodo. Wavelet–based estimators of scaling behavior. *IEEE Trans. Inform. Theory*, 48(11):2938 – 2954, 2002.
- [129] E. Bacry, J.F. Muzy, and A. Arneodo. Singularity spectrum of fractal signals from wavelet analysis: Exact results. *J. of Stat. Phys.*, 70(1,2):635 – 674, 1993.
- [130] R. Hornero, P. Espino, A. Alonso, and M. López. Estimating complexity from EEG background activity of epileptic patient: Calculating correlation dimensions of chaotic dynamic attractors to compare EEGs of normal and epileptic subjects. *IEEE Eng. Med. Biol. Mag.*, 18(6):73 – 79, 1999.
- [131] D.J. Kim, J. Jeong, J.H. Chae, S. Park, S.Y. Kim, H.J. Go, I.H. Paik, K.S. Kim, and B. Choi. An estimation of the first positive lyapunov exponent of the EEG in patients with schizophrenia. *Psych. Res. – Neuroim.*, 98(3):177 – 189, 2000.
- [132] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Phys. D: Nonlin. Phenom.*, 31(2):277 – 283, 1988.
- [133] J. Jeong, J.C. Gore, and B.S. Peterson. Mutual information analysis of the EEG in patients with Alzheimer’s disease. *Clin. Neurophys.*, 112(5):827 – 835, 2001.
- [134] C. Liu, K. Li, L. Zhao, F. Liu, D. Zheng, C. Liu, and S. Liu. Analysis of heart rate variability using fuzzy measure entropy. *Comp. Biol. Med.*, 43(2):100 – 108, 2013.
- [135] E. Ferlazzo, N. Mammone, V. Cianci, S. Gasparini, A. Gambardella, A. Labate, M.A. Latella, V. Sofia, M. Elia, F.C. Morabito, and U. Aguglia. Permutation entropy of scalp EEG: A tool to investigate epilepsies: Suggestions from absence epilepsies. *Clin. Neurophys.*, 125(1):13 – 20, 2014.
- [136] S.D. Wu, C.W. Wu, K.Y. Lee, and S.G. Lin. Modified multiscale entropy for short–term time series analysis. *Phys. A. Stat. Mech. Appl.*, 392(23):5865 – 5873, 2013.
- [137] J.F. Muzy, E. Bacry, and A. Arneodo. Multifractal formalism for fractal signals: The structure–function approach versus the wavelet-transform modulus–maxima method. *Phys. Rev. E*, 47(2):875 – 884, 1993.
- [138] J.F. Muzy, E. Bacry, and A. Arneodo. Wavelets and multifractal formalism for singular signals: Application to turbulence data. *Phys. Rev. Lett.*, 67(25):3515 – 3518, 1991.

- [139] H.B. Xie, J.Y. Guo, and Y.P. Zheng. Fuzzy approximate entropy analysis of chaotic and natural complex systems: Detecting muscle fatigue using electromyography signals. *Ann. Biomed. Eng.*, 38(4):1483 – 1496, 2010.
- [140] L.E. Virgilio-Silva and L. Otavio-Murta. Evaluation of physiologic complexity in time series using generalized sample entropy and surrogate data analysis. *Chaos: Interdis. J. Nonlin. Sci.*, 22(4):-, 2012.
- [141] H.B. Xie, W.X. He, and H. Liu. Measuring time series regularity using nonlinear similarity-based sample entropy. *Phys. Lett. A*, 372(48):7140 – 7146, 2008.
- [142] K.H. Chon, C.G. Scully, and S. Lu. Approximate entropy for all signals. *IEEE Eng. Med. Biol. Mag.*, 28(6):18 – 23, 2009.
- [143] S. Lu, S. Chen, J.K. Kanters, I.C. Solomon, and K.H. Chon. Automatic selection of the threshold value  $r$  for approximate entropy. *IEEE Trans. Biomed. Eng.*, 55(8):1966–1972, 2008.
- [144] E.M.C. Roldán, A. Molina-Picó, D. Cuesta-Frau, P.M. Martínez, and S.O. Crespo. Characterization of entropy measures against data loss: Application to eeg records. In *Eng. Med. Biol. Soc. (EMBC) 2011 Ann. Int. Conf. IEEE*, pages 6110 – 6113, 2011.
- [145] E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, and P. Miró-Martínez. Comparative study between sample entropy and detrended fluctuation analysis performance on eeg records under data loss. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 4233 – 4236, 2012.
- [146] E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, P. Miró-Martínez, L. Vigil-Medina, and M. Varela-Entrecanales. Characterization of detrended fluctuation analysis in the context of glycemic time series. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 4225 – 4228, 2012.
- [147] E.M. Cirugeda-Roldán, A. Molina-Picó, D. Cuesta-Frau, S. Oltra-Crespo, P. Miró-Martínez, L. Vigil-Medina, and M. Varela-Entrecanales. Customization of entropy estimation measures for human arterial hypertension records segmentation. In *Eng. Med. Biol. Soc. (EMBC) 2012 Ann. Int. Conf. IEEE*, pages 33 – 36, 2012.
- [148] E.M. Cirugeda-Roldán, A. Molina-Picó, D. Novak, D. Cuesta-Frau, and V. Křemen. Atrial electrogram complex fractionated entropy study. *Exp. Clin. Cardiol.*, 2014.

- [149] L. Vigil, E. Condés, M. Varela, C. Rodriguez, A. Colas, B. Vargas, M. Lopez, and E. Cirugeda. Glucose series complexity in hypertensive patients. *J. Am. Soc. Hyperten. (in press)*, 2014.
- [150] E.M. Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, S. Oltra-Crespo, L. Vigil-Medina, and M. Varela-Entrecanales nd C. Rodriguez-de Castro. Unsupervised crossover estimation in detrended fluctuation analysis: Application to blood glucose data. *submitted to Med. Eng. Phys.*, 2014.
- [151] E.M. Cirugeda-Roldán, D. Novak, V. Křemen, D. Cuesta-Frau, M.W Keller, C. Schilling, O. Doessel, C. Schmitt, and A. Luik. A new algorithm for quadratic sample entropy optimization for very short biomedical signals: Application to blood pressure records. *submitted to PlosOne*, 2014.
- [152] E.M Cirugeda-Roldán, D. Cuesta-Frau, P. Miró-Martínez, and S. Oltra-Crespo. Comparative study of entropy sensitivity to missing biosignal data. *submitted Entropy*, 2014.
- [153] J.L. Rodríguez-Sotelo, D. Cuesta-Frau A. Osorio-Forero, A. Jiménez-Rodríguez, E. Cirugeda-Roldán, and D. Peluffo. Automatic sleep stages classification using eeg entropy features and unsupervised pattern analysis techniques. *submitted Entropy*, 2014.

# Índice de abreviaturas

## Approximate Entropy

ApEn, 6, 7, 13–18, 24, 45, 46, 50–  
53, 55–61, 63, 64, 66, 68, 69, 76,  
83, 86, 91, 111–118, 120, 125

## Coefficient of Sample Entropy

CosEn, 6, 13, 26, 27, 44, 46, 108,  
109, 111, 117, 120, 126

## Detrended Fluctuation Analysis

DFA, 6, 13, 19, 20, 22, 45, 46, 92–  
97, 99, 102–104, 111, 112, 115–  
120, 125

## MultiScale Entropy

MSE, 6, 13, 18, 45, 46, 70, 73, 74,  
80, 83, 86, 87, 91, 111, 112, 114–  
118, 120, 125

## Rényi Entropy

QSE, 6, 13, 23–25, 27, 44, 46, 111,  
117, 118, 120

## Sample Entropy

SampEn, 13, 16–18, 24, 26, 45, 46,  
62–70, 76, 83, 86, 90, 91, 111,  
112, 114–118, 120