

Kompozitna multiskalarna entropija: poređenje ApEn-a i SampEn-a

Studentski rad

Nina Bertić

Student 3. godine prvog ciklusa studija
Fakultet tehničkih nauka
Novi Sad, Srbija
berticeva@gmail.com

Sažetak – Ovaj rad se bavi primenom multiskalarne entropije i upoređuje dva ključna metrička parametra – entropiju uzorka i aproksimativnu entropiju – u analizi eksperimentalnih podataka dobijenih farmakološkim tretmanom pacova. Farmakološki tretman je uključivao administraciju lekova za kardiovaskularni sistem, dok su mere entropije izračunate pre samog tretmana. Cilj analize je bio razlučivanje kompleksnosti i pravilnosti fizioloških odgovora na tretman pružajući uvid u efekte lekova na fiziološku dinamiku pacova. Takođe, primenjeni su i pseudoslučajno generisani signali koji simuliraju uniformnu i Gausovu raspodelu radi sticanja boljeg uvida u analitičke mogućnosti metode. Poredeći primenu istih mernih jedinica entropije na ove sintetičke signale, dodatni cilj rada je procena otpornosti i osetljivosti mernih jedinica na različite karakteristike signala. Upoređivanje rezultata između eksperimentalnih i pseudoslučajnih podataka pruža sveobuhvatan uvid u univerzalnost i ograničenja mernih jedinica entropije pri analizi razolikih skupova podataka. Rezultati ukazuju na značajne varijacije u merama entropije na različite distribucije signala, naglašavajući važnost prilagodavanja analize entropije prirodni podataka.

Ključne reči - entropija; aproksimativna entropija; entropija uzorka; multiskalarna entropija; kompozitna multiskalarna entropija; unakrsna entropija

I. UVOD

Svakodnevno smo okruženi signalima koji prenose informacije u različitim oblicima, bilo da je reč o zvuku, slici ili podacima. U pozadini ovih interakcija stoji ključna disciplina – obrada signala. Ova grana inženjeringa i nauke igra nezamenljivu ulogu u analiziranju, modifikaciji i interpretaciji signala, pružajući temelj za mnoge tehnološke inovacije koje definišu savremeni život.

Nekada je bila nezamisliva obrada signala u naučnim oblastima kao što je medicina, ali je teorema uzorkovanja [1] to promenila. U medicinskim istraživanjima su se brzo pokazale beneficije obrade signala [2] i uspešno primenjene aproksimativna entropija (*ApEn*) i entropija uzorka (*SampEn*)

kod kardiovaskularnih signala: u analizi varijabilnosti srčanog ritma kod pacijenata obolelih od dijabetesa (tip 2) [3], otkazivanja srca [4], kao i kod zdravih subjekata [5] tokom vežbi i u stanju odmora [6], tokom stresnih situacija [7,8], itd. *XEn* (unakrsna entropija) je predstavljala jednostavnu generalizaciju, izvedenu kao mera koja procenjuje međusobnu (ne)predvidivost dva istovremeno snimljena i međusobno povezana signala [7,9,10]. Ona može biti zasnovana na *ApEn*-u (*XApEn* – aproksimativna unakrsna entropija) i *SampEn*-u (*XSampEn* – unakrsna entropija uzorka).

Korišćeni signali u izradi ovog rada su veštački generisani pseudoslučajni nizovi uniformne i Gausove raspodele, dužine u okviru opsega 100-10000, kao i nizovi kardiovaskularnih signala laboratorijskih pacova pre davanja lekova dobijenim u farmakološkom eksperimentu [11]. Signali laboratorijskih životinja su snimljeni nad odraslim pacovima muškog pola u svesnom stanju, težine 300-350 grama. Pacovi su odvojeno smešteni u kontrolisanim laboratorijskim uslovima (temperatura: $-22 \pm 2^{\circ}\text{C}$, relativna vlažnost: 60-70%, osvetljenje: 12:24h - ciklus svetlo-tama). Merenje talasnih oblika krvog pritiska (*BP*) izvedeno je korišćenjem TL11M2-PA C50-PKS-DSI implantiranim u abdominalnu aortu. Učestalost uzorkovanja je bila 1000 Hz, sa 12-bitnim A/D konvertorom koji daje 4096 nivoa amplitude. Sistolni pritisak (*SBP*) i pulsni interval (*PI*) su izvedeni iz *BP* kao lokalni maksimumi i intervali između uzastopnih maksimalnih pozitivnih promena *BP*. Artefakti koji su detektovani poluautomatski, zajedno sa preostalim greškama, ručno su uklonjeni. Vremenski nizovi su detrendirani kako bi se uklonile veoma spore promene komponente signala pomoću filtra koji je predložen za biomedicinsku primenu u [12]. Za svaki segment signala je proverena stacionarnost u smislu srednje vrednosti i varijanse metodom iz [13].

Svi eksperimentalni postupci su u skladu sa Zakonom o dobrobiti 2009/6/RC, kao i Pravilnikom 2010/RC o zaštiti životinja, koji se koriste u naučne svrhe, direktive 2010/63/EU za nacionalne životinje. Protokol je odobren od strane Komisije za etičku reviziju Univerziteta u Beogradu (licenca broj 323-07-10519-2013-05/2).

Cilj ovog rada je upoređivanje metode multiskalarne i kompozitne multiskalarne entropije za slučaj kada se kao osnova metode koristi *SampEn*, i kada se kao osnova metode koristi *ApEn*. Metode koje su korišćene tokom eksperimenta su *ApEn*, *SampEn*, *MSE*, *CMSE*, *XEn*, *XSampEn*.

II. METODOLOGIJA

A. Aproksimativna entropija

Aproksimativna entropija (*ApEn*) je matematička metrika koja se koristi za kvantifikovanje sekvencijalne i temporalne iregularnosti ili kompleksnosti niza. Korisna je u analiziranju kompleksnih sistema gde suptilne promene mogu biti indikacija važnih osnovnih dinamika. [14] Niske vrednosti *ApEn*-a predstavljaju regularne, repetitivne i predvidljive nizove koji se šablonski ponavljaju kroz sekvencu. Visoke vrednosti *ApEn*-a predstavljaju suprotno. *ApEn* je pristrasna statistika; u pokušaju da se definiše nepristrasna statistika, entropija uzorka je predstavljena kao alternativna statistika regularnosti sa malo drugačijim osnovama. [15]

U ovom radu, algoritam *ApEn*-a, kao i entropije uzorka, je izveden iz naučnoistraživačkog rada [16].

Uzimajući u obzir sekvencu $u = \{u(1), u(2), \dots, u(N)\}$ dužine N , dužinu sekvenci koje se upoređuju $0 \leq m \leq N$ i realni broj $r > 0$ koji predstavlja toleransu za primanje istih parova. Parametri N , m i r moraju biti konstantni za svaku kalkulaciju. Vektori $x_m(j) = \{u(j), u(j+1), \dots, u(j+m-1)\}$ i $x_m(i) = \{u(i), u(i+1), \dots, u(i+m-1)\}$ su definisani. *Chebyshev*-a distanca između njih se računa preko jednačine (1):

$$d|x_m(i), x_m(j)| = \max(|u(i+k-1) - u(j+k-1)|) \quad (1)$$

Vrednost $C_i^m(r)$ je izvedena iz jednačine (2):

$$C_i^m(r) = \frac{\text{broj od } j \leq (N-m+1) \text{ kao što je } d|x_m(i), x_m(j)| \leq r}{(N-m+1)} \quad (2)$$

Numerator u jednačini (2) broji, s ograničenjem praga r , broj vektora uzastopnih vrednosti dužine m , što je slično datom vektoru i dozvoljava samoprebrojavanje. Pri izračunavanju $C_i^m(r)$, vektor $x_m(i)$ predstavlja šablon, dok vektor $x_m(j)$ u okviru praga r predstavlja podudarni šablon. Parametar $C_i^m(r)$ je verovatnoća da se bilo koji vektor $x_m(j)$ nalazi unutar praga r od $x_m(i)$.

U narednom koraku, magnituda $C_i^m(r)$ je izvedena iz jednačine (3):

$$C^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log C_i^m(r) \quad (3)$$

Na isti način, kad ugrađena dimenzija dostigne vrednost $m+1$, vrednost $C^{m+1}(r)$ se računa pomoću jednačine (3). Konačno, može se izracunati *ApEn* statistične vrednosti sekvence (4):

$$ApEn(m, r, N)(u) = C_m(r) - C_{m+1}(r) \quad (4)$$

gde je $m \geq 1$ i $ApEn(0, r, N)(u) = -C^1(r)$. Dobijeni $ApEn(m, r, N)$ iz jednačine (4) predstavlja statistički procenitelj parametra $ApEn(m, r)$:

$$ApEn(m, r) = \lim_{N \rightarrow \infty} |C^m(r) - C^{m+1}(r)|. \quad (5)$$

B. Entropija uzorka (*SampEn*)

Za datu sekvencu $u = \{u(1), u(2), \dots, u(N)\}$ dužine N , dužinu sekvenci koje se upoređuju $0 \leq m \leq N$ i realni broj $r > 0$ - distanca d jednačine (1) je ista kao i za aproksimativnu entropiju.

Broj vektora $x_m(j)$ unutar praga r od $x_m(i)$ bez dovoljavanja samoprebrojavanja je definisana jednačinom (6):

$$B_i^m(r) = \frac{1}{N-m+1} \sum_{j=1}^{N-m} (\text{broj puta gde je } d|x_m(i), x_m(j)| \leq r) \quad (6)$$

Nakon toga, konačan broj mogućih vektora $B^m(r)$ je izračunat na osnovu jednačine (7) i označava empirijsku verovatnoću da se dve sekvence poklapaju za m tačaka.

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \quad (7)$$

Slično tome, broj vektora $x_{m+1}(j)$ na distanci r od $x_{m+1}(i)$ bez dozvoljavanja samoprebrojavanja samopodudaranja je definisan jednačinom (8):

$$A_i^m(r) = \frac{1}{N-m-1} \sum_{j=1}^{N-m} (\text{broj puta gde je } d|x_m(i), x_m(j)| \leq r) \quad (8)$$

Broj sličnih $A^m(r)$ je izveden iz jednačine (9) i označava empirijsku mogućnost da su dve sekvence slične za $m+1$ tački (slični):

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \quad (9)$$

Pošto je broj sličnih ($A^m(r)$) uvek manji ili jednak broju sličnih vektora ($B^m(r)$), odnos $\frac{A^m(r)}{B^m(r)} < 1$ je uslovna verovatnoća [4]. U poslednjem koraku, *SampEn* statističke vrednosti sekvence u je procenjena kao:

$$SampEn(m, r, n)(u) = -\log\left(\frac{A^m(r)}{B^m(r)}\right) \quad (10)$$

SampEn(m, r, n) dat u jednačini (10) je statistički procenitelj parametra *SampEn(m, r)*:

$$SampEn(m, r) = \lim_{N \rightarrow \infty} [-\log(\frac{A^m(r)}{B^m(r)})] \quad (11)$$

Za regularne, ponavljajuće podatke, odnos $\frac{A^m(r)}{B^m(r)}$ jednačine (11) se približava jedinici, pa se stoga $SampEn$ približava nuli.

Prema *Delgado-Bonal*-u i *Marshak*-u [17], tri glavne razlike između aproksimativne entropije i entropije uzorka glase:

1) $ApEn$ dozvoljava samoprebrojavanje za razliku od $SampEn$ -a ($j \neq 1$) što predstavlja prednost entropije uzorka. Algoritam $SampEn$ -a ne dozvoljava poređenje vektora samim sa sobom pošto se ne dobijaju korisne informacije.

2) Mana $ApEn$ -a je to da ovaj algoritam uključuje faktor $\frac{1}{N-m}$, što $ApEn$ čini direktno zavisnim od veličine niza N .

3) Pošto je suma svih šablonskih vektora unutar logaritma u $SampEn$ -u, to podrazumeva da $SampEn$ posmatra čitav niz i ako šablon pronade svog para, onda je $SampEn$ već definisan. $ApEn$ -u je potrebno podudaranje za svaki šablon. Takođe, vrednosti $SampEn$ -a su veće od $ApEn$ vrednosti za isti niz podataka.

C. Multiskalarna entropija (MSE)

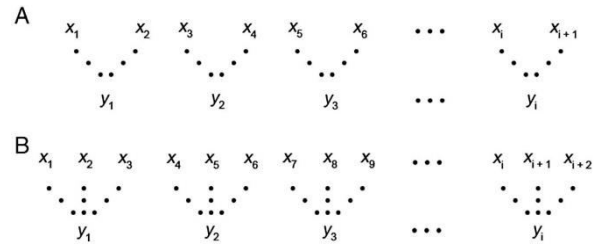
Algoritam za multiskalarnu entropiju je izveden iz naučnoistraživačkog rada [18].

Od jednodimenzionalnog diskretnog niza $\{x_1, \dots, x_i, \dots, x_N\}$ se mogu konstruisati uzastopni grubi kvantovani nizovi $y^{(T)}$ koji odgovaraju faktoru razmene T . Prvi korak je da se originalni niz podeli u delove čije se dužine, T , ne preklapaju. Nakon toga se usredsređuje na tačke podataka unutar svakog dela. Svaki element grubog kvantovanog niza je izračunat na osnovu jednačine (12):

$$y_j^{(T)} = \frac{1}{T} \sum_{i=(j-1)T+1}^{jT} x_i, \quad 1 \leq j \leq \frac{N}{T} \quad (12)$$

gde y_j predstavlja tačku podataka u novoizgrađenom nizu, x_i tačku podataka u originalnom nizu, dok N predstavlja dužinu originalnog niza.

Za prvu skalu, niz $y^{(1)}$ predstavlja originalni niz. Dužina svakog niza je jednaka dužini originalnog niza podeljenog sa faktorom skale T . Nakon toga, za svaki novonastali niz se računa $SampEn$ kao funkcija faktora T .



Slika 0. Procedura grubog kvantovanja [19] - (A) niz 1, (B) niz 2, gde x predstavlja originalni niz, dok je y novonastali niz (12)

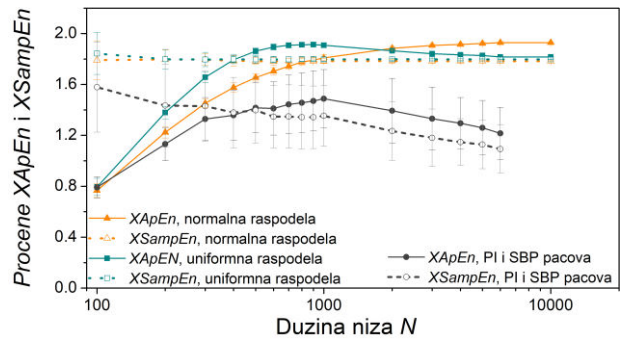
D. Kompozitna multiskalarna entropija (CMSE)

Nakon što su izračunate vrednosti $SampEn$ -a svih nizova, može se izračunati kompozitna multiskalarna entropija po formuli:

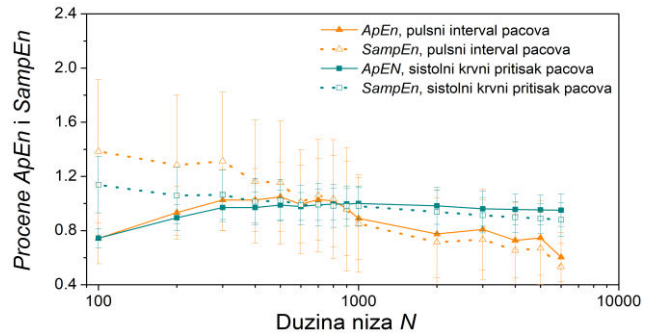
$$CMSE(x, T, m, r) = \frac{1}{T} \sum_{k=1}^T (y_k^{(T)}, m, r) \quad (13)$$

gde $x = \{x_1, x_2, x_3, \dots, x_N\}$ predstavlja jednodimenzionalni niz dužine N , m dimenziju svakog para sličnih vektora za računanje $SampEn$, dok je r prethodno definisan prag tolerancije. [20]

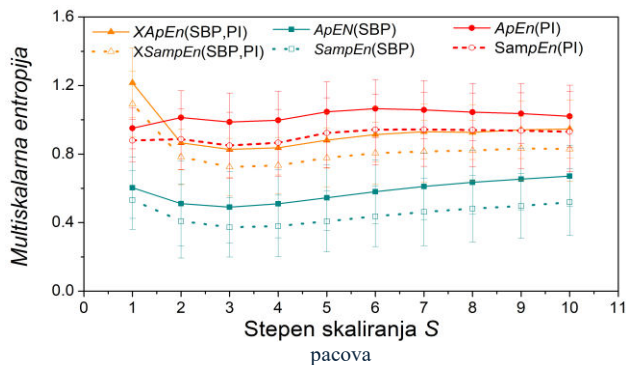
III. REZULTATI



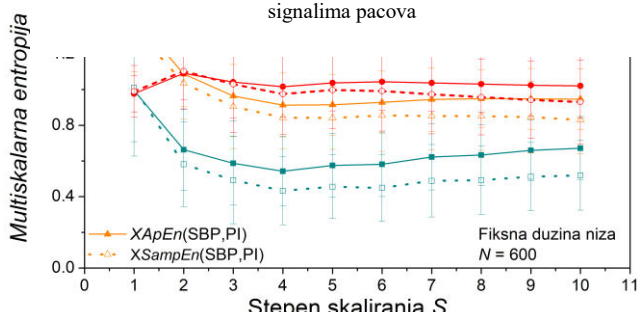
Slika 2. Procene $XApEn$ -a i $XSampEn$ -a primenjene nad pseudoslučajnim nizovima normalne i uniformne raspodele, kao i nad PI i SBP signalima



Slika 3. Procene $ApEn$ -a i $SampEn$ -a primenjene nad PI i SBP signalima pacova



Slika 4. *SampEn*, *ApEn*, *XSampEn* i *XapEn* primenjeni nad *PI* i *SBP* signalima pacova



Slika 5. *SampEn*, *ApEn*, *XSampEn* i *XapEn* primenjeni nad *PI* i *SBP* signalima pacova sa fiksnom dužinom niza $N=600$

IV. ANALIZA I ZAKLJUČAK

Na pseudoslučajnim nizovima, procena *XSampEn*-a je stabilna. S druge strane, vrednost *XApEn*-a zavisi od dužine niza, što je čini nestabilnijom i manje pogodnom za primenu (Slika 2). S obzirom na to da se entropije samih pseudoslučajnih nizova isto ponašaju, one nisu prikazane.

Razlika se primećuje tokom primene unakrsne entropije *SBP* i *PI* signala parova, kao i entropije isključivo *SBP*, odnosno *PI* signala (Slika 4.). *SampEn* se prikazao kao stabilniji od *ApEn*-a za male vrednosti dužine niza N , dok je za veće vrednosti N i entropiju niza *PI* prikazan blago opadajući trend za unakrsnu entropiju (Slika 3). Nizovi sistolnog pritiska su pokazali stabilno ponašanje, što ukazuje na to da je reč o mehaničkom procesu koji je inertniji od srčanog ritma.

Primenom multiskalarnu entropije dobijeno je nekoliko interesantnih zaključaka. Budući da procene entropije zavise od dužine analiziranih nizova, stabilniji rezultati su dobijeni kada je za analizu bila korišćena fiksna dužina niza za sve stepene skaliranja s . U ovom radu je za fiksnu dužinu izabrano $N = 600$ (Slika 2). Kompozitnim usrednjavanjem multiskalarnu entropije, rezultat se dodatno stabilizovao. S obzirom na to da aproksimativna entropija odviše zavisi od

dužine niza, nije se koristila u multiskalarnom okruženju. Takođe je primećeno da ne postoji bitnih razlika između multiskalarnu i kompozitne multiskalarnu entropije.

U svim analizama dobijeni su slični rezultati. Pokazalo se da je entropija, nezavisno od načina procene, veća za nizove većeg varijabiliteta (*PI*), dok je manja za mehanički filtrirane nizove sistolnog krvog pritiska. Pri tome, dominantan uticaj na unakrsnu entropiju je imao niz *PI*. Za unakrsnu entropiju i entropiju *SBP* nizova, blago je izražen porast entropije kao posledica skaliranja i centralne granične teoreme, što je očekivano. Očekivani porast nije očit na nizovima *PI*. Moguće je da je to posledica pojave uočene u istraživanju [21] u kojem su korišćeni isti signali kao i u ovom radu: naime, laboratorijske životinje, kao izvor signala, podvrgnute su eksperimentu u kom su svakog dana primali veliku dozu različitih lekova s uticajem na kardiovaskularni sistem. Analize u [21] su pokazale da su signali pacova, pre primljenih lekova, potpuno poremećeni. Isto je bilo odraženo i u rezultatima ovog rada.

V. ZAHVALNICA

Zahvaljujem se prof. dr Nini Japundžić-Žigon što je dozvolila da studenti Fakulteta tehničkih nauka koriste signale snimljene u njenoj laboratoriji za laboratorijske vežbe, diplomske i master radove, kao i za studentske radove na konferencijama.

LITERATURA

- [1] Shannon, C.E. A mathematical theory of communication. Bell Syst. Tech. J. 1948, 27, 623–656.
- [2] Jelinek, H.F.; Cornforth, D.J.; Khandoker, A.H. ECG Time Series Variability Analysis: Engineering and Medicine, 1st ed.; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, 2017; pp. 1–450. ISBN 978-1482243475
- [3] Li, X.; Yu, S.; Chen, H.; Lu, C.; Zhang, K.; Li, F. Cardiovascular autonomic function analysis using approximate entropy from 24-h heart rate variability and its frequency components in patients with type 2 diabetes. J. Diabetes Investig. 2015, 6, 227–235
- [4] Krstacic, G.; Gamberger, D.; Krstacic, A.; Smuc, T.; Milicic, D. The chaos theory and non-linear dynamics in heart rate variability in patients with heart failure. In Proceedings of the Computers in Cardiology, Bologna, Italy, 14–17 September 2008; pp. 957–959
- [5] Storella, R.J.; Wood, H.W.; Mills, K.M.; Kanters, J.K.; Højgaard, M.V.; Holstein-Rathlou, N.-H. Approximate entropy and point correlation dimension of heart rate variability in healthy subjects. Integr. Psychol. Behav. Sci. 1998, 33, 315–320
- [6] Tulppo, M.P.; Makikallio, T.H.; Takala, T.E.; Seppanen, T.; Huikuri, H.V. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. Am. J. Physiol. Heart Circ. Physiol. 1996, 271, H244–H252
- [7] Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol. 2000, 278, H2039–H2049
- [8] Boskovic, A.; Loncar-Turukalo, T.; Sarenac, O.; Japundžić-Žigon, N.; Bajic, D. Unbiased entropy estimates in stress: A parameter study. Comput. Biol. Med. 2012, 42, 667–679

- [9] Pincus, S.M.; Mulligan, T.; Iranmanesh, A.; Gheorghiu, S.; Godschalk, M.; Veldhuis, J.D. Older males secrete lutenizing hormone and testosterone more irregularly, and jointly more asynchronously than younger males. *Proc. Natl. Acad. Sci. USA* 1996, 93, 14100–14105
- [10] Skoric, T.; Sarenac, O.; Milovanovic, B.; Japundzic-Zigon, N.; Bajic, D. On consistency of cross-approximate entropy in cardiovascular and artificial environments. *Complexity* 2017, 1–15
- [11] Milutinovic, S.; Murphy, D.; Japundzic-Zigon, N. The role of central vasopressin receptors in the modulation of autonomic cardiovascular controls: A spectral analysis study. *Am. J. Physiol. Regul. Integr.* 2006, 291, r1579–r1591
- [12] Tarvainen, M.P.; Ranta-Aaho, P.O.; Karjalainen, P.A. An advanced detrending approach with application to HRV analysis. *IEEE Trans. Biomed. Eng.* 2002, 42, 172–174
- [13] Bendat, J.S. Piersol, A.G. *Random Data Analysis and Measurement Procedures*; Wiley Series in Probability and Statistics: New York, NY, USA, 1986.
- [14] Pincus, S., 1991. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* 88, 2297–2301. doi:10.1073/pnas.88.6.2297
- [15] A. Delgado-Bonal, A. Marshak, “Approximate Entropy and Sample Entropy: A Comprehensive Tutorial” , published online 2019. May 28, doi: 10.3390/e21060541
- [16] Pincus, S., 1991. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* 88, 2297–2301. doi:10.1073/pnas.88.6.2297
- [17] Delgado-Bonal, A., Marshak, A., 2019. Approximate Entropy and Sample Entropy: A comprehensive tutorial. *Entropy* 21, 541. doi:10.3390/e21060541
- [18] M. Costa, A. L. Goldberger, C. K. Peng, “Multiscale entropy analysis of biological signals”, published 18.2.2005.
- [19] M. A. Busa, R. E. A. van Emmerik, “Multiscale entropy: A tool for understanding the complexity for postural control”, *Journal of Sport and Health Science*, March 2016, pp. 44-51
- [20] R. Zhang, Y. He, K. Liu, L. Zhang, S. Zhang, E. Pickwell-MacPherson, Y. Yao, C. Zhang, “Composite multiscale entropy analysis of reflective terahertz signals for biological tissues”, *Optics Express*, vol.25, issue 20, 2017.
- [21] S. Jovanović, doktorska disertacija, FTN, Novi Sad, 2017.

ABSTRACT

This study explores the application of multiscale entropy and compares two key metric parameters – sample entropy and approximate entropy – in the analysis of experimental data obtained through pharmacological treatment of rats. The pharmacological treatment involved the administration of drugs targeting the cardiovascular system, and entropy measures were calculated prior to the treatment. The aim of the analysis was to discern the complexity and regularity of physiological responses to the treatment, providing insights into the effects of drugs on the physiological dynamics of rats. Additionally, pseudorandomly generated signals simulating uniform and Gaussian distributions were used to extend the analysis. By comparing the application of the same entropy metrics to these synthetic signals, the study aims to assess the resilience and sensitivity of entropy metrics to different signal characteristics. Comparing results between experimental and pseudorandom data provides a comprehensive insight into the universality and limitations of entropy metrics in analyzing diverse datasets. The results indicate significant variations in entropy measures for different signal distributions, emphasizing the importance of adapting entropy analysis to the nature of the data.

COMPOSITE MULTISCALAR ENTROPY: A COMPARISON OF APEN AND SAMPEN

Nina Bertić