Fundamental researches

STATISTICAL ANALYSIS OF MEDICAL TIME SERIES

Martynenko A., Raimondi G., Sotnikova-Meleshkina Zh., Danylenko H., Budreiko N.

Statistical analysis of data sets is a necessary component of any medical research. Modern methods of mathematical statistics and statistical application suites provide extensive capabilities for analysis of random values. However, when a data set is represented by a series of data ordered by time, or when structure and order of data are essential components of research, special approaches to statistical analysis become necessary.Presented in this article are special statistical methods developed by the authors for analysis of a time series: Time Series Mann-Whitney M-test is an analogue of the known nonparametric Mann-Whitney U-test for two Time Series with an equal number of elements; Nominal Time Series Measure is a statistical estimator of dynamics of a nominal series consisting of «0» (no) and «1» (yes); Time Series Entropy *EnRE* is a specially developed robust formula for a Time Series, intended for calculation of nonlinear stochastic measure of order or disorder, popular in various researches. Presented methods are accompanied by a detailed demonstration of capacity for statistical analysis of medical Time Series: Analysis of growth dynamics of boys and girls aged 6–7–8 years (data by World Health Organization); analysis of the number of seizures and choice of anti-epileptic drugs (data by The National Society for Epilepsy); Time series entropy *EnRE* for Detecting Congestive Heart Failure by standard 5-minutesHeart Rate Variability samples (data by Massachusetts Institute of Technology – Boston's Beth Israel Hospital RR database). It has been noted that, in every case, using the named special methods for statistical analysis of medical Time Series allows one to avoid errors in interpreting results received through statistical methods and substantially increases the accuracy of statistical analysis of medical Time Series.

KEY WORDS: medical time series, nonparametric test, nominal measure, entropy

INFORMATION ABOUT AUTHORS

Martynenko Alexander, D. Sc., Professor, Department of Hygiene and Social Medicine, V. N. Karazin Kharkiv National University, 6, Svobody sq., Kharkiv, Ukraine, 61022, e-mail: Alexander.v.martynenko@karazin.ua, ORCID ID: https://orcid.org/0000-0002-0609-2220

Gianfranco Raimondi, MD, PhD, Prof., Sapienza University of Rome (Italy), Piazzale Aldo Moro 5, 00185, Rome, Italy, e-mail: gianfrancoraimondi@uniroma1.it

Sotnikova-Meleshkina Zhanna, PhD in Medicine, Head of Department of hygiene and social medicine V. N. Karazin Kharkiv National University School of Medicine, 6, Svobody Sq., Kharkiv, Ukraine, 61022, e-mail: zhanna.v.sotnikova@karazin.ua, ORCID ID: https://orcid.org/0000-0001-5534-8264

Danylenko Heorhii., MD, PhD, Full Prof., State Enterprise «Institute for the Protection of Children and Adolescent Health of the National Academy of Medical Sciences of Ukraine», 52-a, Yubileiny Avenue, Kharkiv, Ukraine, 61153, e-mail: danylenkog@ukr.net, ORCID ID: https://orcid.org/0000-0001-7086-2720

Budreiko Nikita, Assistant, Department of hygiene and social medicine, V. N. Karazin Kharkiv National University School of Medicine, 6, Svobody sq., Kharkiv, Ukraine, 61022, e-mail: nbudreiko@protonmail.com

OBJECTIVE

Statistical analysis of data sets is a necessary component of any medical research. Modern methods of mathematical statistics and statistical application suites provide extensive capabilities for analysis of random values. However, when a data set is represented by a series of data ordered by time, or when structure and order of data are essential components of research, special approaches to statistical analysis become necessary. Presented in this article are the following special statistical methods developed by the authors for analysis of a Time Series:

1. Time Series Mann-Whitney M-test – an analogue of the known nonparametric Mann-Whitney U-test for two Time Series with an equal number of elements;

2. Nominal Time Series Measure – a statistical estimator of dynamics of a nominal series consisting of «0» (no) and «1» (yes). Such dichotomous Time Series are often used when there is a need for describing qualitative events in development or course of a disease;

3. Time Series Entropy _EnRE_ – specially developed robust formula for a Time Series, intended for calculation of nonlinear stochastic measure of order or disorder, popular in various researches. In medicine, such nonlinear methods have, to the greatest degree, proved their worth for analysis and prognosis of sudden changes in medical condition, such as atrial fibrillation, epileptic seizures, etc.

All presented methods are accompanied by a detailed demonstration of capacity for statistical analysis of medical Time Series.

MATERIALS AND METHODS

Used for statistical analysis of a time series have been methods, algorithms and programs developed by the authors, these being compared with known statistical methods presented in software suites «IBM SPSS Statistics» and «Statistica» by StatSoft.

In order to illustrate presented methods for statistical analysis of time series, developed by the authors, the following cases have been used:

Case 1: height-for-age data by World Health Organization (WHO) for school-aged children and adolescents [1]. The growth curves for ages 5 to 19 years were thus constructed using data from 18 months to 24 years. The final sample used for fitting the growth curves included 30 907 observations (15 537 boys, 15 370 girls) for the height-for-age curves;

Case 2: Epileptic patient case and example of anti-epileptic drugs (AEDs) therapy choice by The National Society for Epilepsy https://www.epilepsysociety.org.uk/;

Case 3: long-term HRV records by Massachusetts Institute of Technology – Boston'sBethIsraelHospital (MIT-BIH) from [28] (http://www.physionet.org) for Normal Sinus Rhythm (NSR) RR Interval Database includes beat annotation files for 54 long-term ECG recordings of subjects in normal sinus rhythm (30 men, aged 28.5 to 76, and 24 women, aged 58 to 73). Congestive Heart Failure (CHF) RR Interval Database includes beat annotation files for 29 long-term ECG recordings of subjects aged 34 to 79, with congestive heart failure (NYHA classes I, II, and III).

RESULTS AND DISCUSSION

1. Time Series Mann-Whitney M-test

The most popular statistical tests for comparison of two or more data sets are Student's t-test and Mann-Whitney U-test. The first one is used for data with normal distribution, and the second is a nonparametric alternative to the Student's t-test. Both tests identify differences of just the mean values in data sets and are insensitive to the order of data in a time series. If the named tests are used for a Time Series, this causes substantial errors. Specifically for comparison of Time Series with an equal number of elements we have developed a modification – Time Series Mann-Whitney M-test (MW M-test). Mann-Whitney U-test for data sets with number of elements N and sums of ranks R_1, R_2 is represented by a known formula for U [3]:

$$U = N^2 + \frac{N(N+1)}{2} - max(R_1, R_2)$$

Acquired value is compared with table Critical Values for chosen level of significance p (usually p = 0.05). There are numerous known modifications of the original method. In case of large samples, for one, U is approximately distributed normally, so standardized value use

$$z = \frac{U - m_U}{\sigma_U}; \; m_U = \frac{N^2}{2}; \sigma_U = N\sqrt{\frac{2N+1}{12}}$$

Additional correction of σ is performed in case of a need to take into consideration a small number of associated ranks. However, such approach is not applicable if all elements in data sets are interconnected as a single sequence of a time series.

In order to use the test for comparison of Time Series with an equal number of elements N, we have proposed a modification – Time Series Mann-Whitney M-test with a formula for M:

$$M = U - abs\left(\sqrt{D_1} - \sqrt{D_2}\right); \; D_1 = \sum_{i=0}^{N}\left(r_1^i - R_1^i\right)^2; D_2 = \sum_{i=0}^{N}\left(r_2^i - R_2^i\right)^2$$

In this case, r_1, r_2 are ranks of the elements in original time series, R_1, R_2 are ranks of time series in a general series after merging. In the proposed modification, we have taken into account the changed positions of elements of a time series before and after merging of data into a single sequence by calculating their total distances D_1, D_2. It should be noted that Critical Values of Time Series Mann-Whitney M-test are the same for Mann-Whitney U-test.

Case 1: Analysis of growth dynamics of boys and girls (6–7 years) and (6–8 years), according to data of the WHO [1].

In order to illustrate the use of developed Time Series MW M-test, we shall analyze the growth dynamics of boys and girls (6–7 years) and (6–8 years), according to data of the WHO [1]. The growth dynamics will be estimated by monthly increment of growth median in children. In table 1.a, it is shown that in boys and girls aged 6–7 years, growth dynamics differ at the mean level, which is confirmed, with significance level of $p < 0.05$, by all three conducted tests: Student's t-test, Mann-Whitney U-test and Time Series MW M-test. Should similar analysis be conducted in children aged 6–8 years, it becomes apparent that mean values ΔL have come closer, due to nonlinearity of curve to monthly increment of growth median in girls (Table 1.b). In this case, Student's t-test and Mann-Whitney U-test are not applicable, because while there is actually a greater change of growth dynamics than in children aged 6–7 years, these tests are insensitive to substantial changes in time series. Time Series MW M-test correctly takes into account veracious difference in time series at the level of $p < 0.05$, as well as decreasing value of correlation coefficient, from 0.986 (6–7 years) to 0.8 (6–8 years).
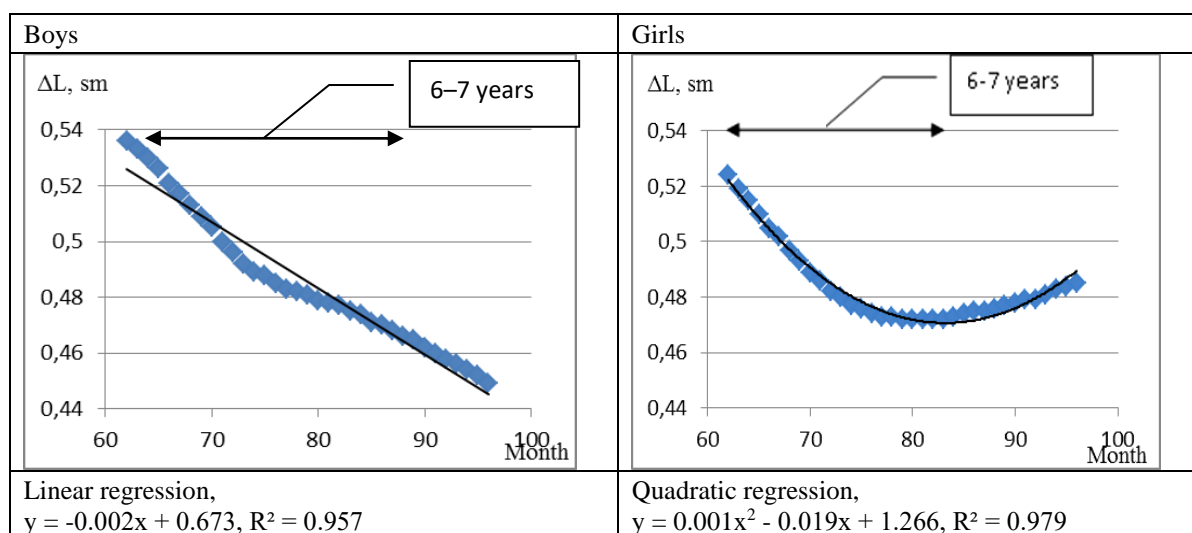
Table 1.a

Growth dynamics in children (6–7 years) and (6–8 years) based on monthly increment of growth median

	Age 61–84 m. (6–7 years)		Age 61–96 m. (6–8 years)	
	Boys	Girls	Boys	Girls
ΔL, sm (Mean ± Standard deviation)	0.499 ± 0.020	0.487 ± 0.017	0.485 ± 0.025	0.484 ± 0.015
Valid ΔL number, N	23	23	35	35
Student's t-test	$p < 0.05$		$p = 0.78$	
Mann-Whitney U-test	$p < 0.05$		$p = 0.84$	
Time Series MW M-test	$p < 0.05$		$p < 0.05$	
Pearson correlation	0.986		0.800	

Table 1.b

Chart of growth dynamics in children (6-7 years) and (6-8 years) based on monthly increment of growth median

Boys	Girls
Linear regression, $y = -0.002x + 0.673$, $R^2 = 0.957$	Quadratic regression, $y = 0.001x^2 - 0.019x + 1.266$, $R^2 = 0.979$

2. Nominal Time Series Measure

Using nominal data in medical practice is very common: everywhere the response options are represented by a dichotomy of «Yes» or «No», or either presence — «0» or absence – «1» of any negative trait is important, using special statistical methods is necessary. When comparing two or more data sets, the most prevalent is Cochran Q test, which enables one to estimate the

significance of their difference. In case sets of nominal data are represented by time series, a modification of aforesaid Time Series MW M-test is applicable. However, if there is only one nominal time series for which it is needed to assess the nature of changes occurring in time, using a measure developed specially for that purpose becomes necessary. We have proposed an appropriate Nominal Time Series Measure:

a. If positive events are designated as «1»

$$Mn = \frac{\sum_{i=1}^{N} X_i + \sum_{i=1}^{N} iX_i}{N + 0.5 * N * (N+1)}$$

b. If negative events are designated as «1»

$$Mn = 1 - \frac{\sum_{i=1}^{N} X_i + \sum_{i=1}^{N} iX_i}{N + 0.5 * N * (N+1)}$$

Let us mark peculiarities of interpretation of a Nominal Time Series Measure *Mn*:

- $Mn = 1$, only when all the $X_N = 1$;
- $Mn = 0$, only when all the $X_N = 0$;
- $Mn = 0.5$, threshold value;

- $Mn < 0.5$, from the position of an observer, in the end of a series, negative events or responses are prevalent;
- $Mn > 0.5$, from the position of an observer, in the end of a series, positive events or responses are prevalent.

Case 2: Analysis of number of seizures and choice of anti-epileptic drugs (according to data ofThe National Society for Epilepsy (NSE) https://www.epilepsysociety.org.uk/)

Number of epileptic seizures is a distinctive indicator for estimation of severity of an epilepsy and when choosing anti-epileptic drugs (AED) therapy. With the right AED, up to 70 % of people with epilepsy could have their seizures controlled or stopped. Shown in Table 2 are cases of epileptic seizure 2 and 4 weeks before treatment and during AED treatment. Facts of registered epileptic seizure are marked as «1», and the days without epileptic seizure are marked as «0». A simple estimator of the

Table 2

Seizures before and during AED therapy
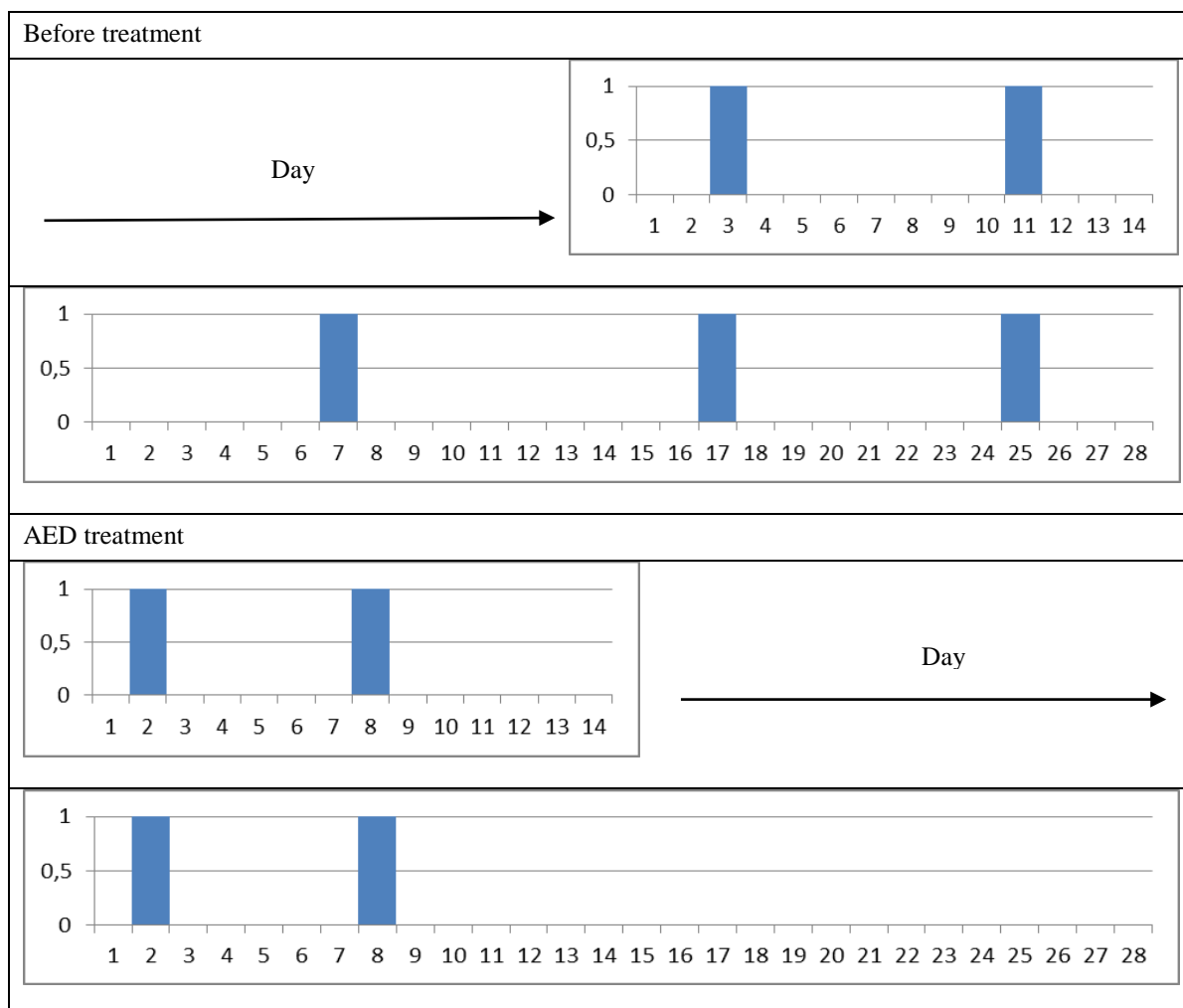
number of epileptic seizures 2 weeks before treatment and during AED treatment does not enable us to draw any conclusions as to the quality of applied AED therapy, because both before and during the therapy, equal number of epileptic seizures is observed. However, objective clinical picture has enabled a doctor to continue the existing AED therapy for another 2 weeks. The doctor's choice has proven correct, because in a 4-week timespan the number of epileptic seizures during AED treatment has decreased, when compared to the state of affairs 4 weeks before treatment. At the same time, if the doctor had an opportunity to quantitatively estimate the dynamics of change in a nominal series, the choice would be a statistically valid one:

- 2 weeks before treatment $Mn = 0.866$;
- 2 weeks during AED treatment $Mn = 0.899$;
- 4 weeks before treatment $Mn = 0.880$;
- 4 weeks during AED treatment $Mn = 0.972$.

Therefore, we can see that even during the first 2 weeks since the beginning of AED treatment, measure Mn has shown positive dynamics, and in a 4-week timespan since the beginning of AED treatment, objective situation of Mn estimates has greatly improved.

3. Time Series Entropy

Nonlinear statistical methods, such as entropy, have found widespread use and have shown great efficiency as part of analyzing various medical data [4, 5]. Various methods of implementing the entropy calculation were developed, some of which and use thereof in medicine even have special issues of magazines dedicated to them [6]. However, presented methods of entropy calculation share a common characteristic, which is their insensitivity to change in data structure, i.e. data could be either randomly shuffled or ordered, and entropy would not change. Also, known methods usually are very demanding to quantity of analyzed data – required number of those can exceed thousands, which is sometimes hardly achievable in medical research. We have developed and proposed a robust formula for calculating entropy of time series $EnRE$ [7]:

$$EnRE = ln\left(\frac{A}{N^{1/2}}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{(|x_i - x_j||x_j - MD|)^{1/k}}{(D_{ij})^{m/2}}\right)\right),$$

where MD is median of time series; D_{ij} – distance between observed data points X_i and X_j in time series; A, l, m, k – estimated coefficients. Search conditions for coefficients A, l, m, k are the following [7]:

1/ accurate approximation for known distributions of a random value;

2/ independence of $EnRE$ from N for initial time series and for series after sorting;

3/ independence of $EnRE$ from additive changes of mean.

In [7], it was established that, in case of time series represented by RR-intervals, the following coefficient values had been found:

$A = \sqrt{\frac{2\pi e}{5}}$, $l = 3$, $m = 1$, $k = 2$. Let us note important characteristics of the found generalized form of $EnRE$ and coefficients:

1/ $EnRE$ [7] and found coefficients l, m, k provide independence from additive change of mean series and from magnitude of selection N for basic series and for series after sorting;

2/ value $EnRE$ is sensitive to structural changes in series, such as, for example, sorting which increases the degree of order in series, decreasing the $EnRE$. This offers additional advantages in research, as shown below for a case of NSR and CHF groups classification;

3/ readjusting coefficient A alone may be required to find the best $EnRE$ value in another range of change in parameters of various random distributions, which can always be done using the method of least squares.

Case 3: Time series entropy $EnRE$ for Detecting Congestive Heart Failure by standard 5-minute HRV samples (MIT-BIH RR database) [2].

Let us demonstrate the usage of $EnRE$ for Detecting Congestive Heart Failure in short segments ($N = 500$) by MIT-BIH RR database. In [8], it has been shown that the minimal length of an RR-segment, for which it is possible to classify NSR and CHF groups by way of Multiscale Entropy Analysis, is $N = 1000$. The performance of such classification is: $Se = 0.70$; $Sp = 0.76$; $Acc = 0.74$. Given in Table 3 are Mean and Standard deviation of $EnRE$ for NSR and CHF for basic RR-intervals and series after sorting ($N = 500$). In both cases, the differences between groups are reliable to the degree of $p < 10^{-7}$.

Table 3

Mean and Standard deviation of *EnRE* for NSR and CHF (original and sorted data)

HRV Record ($N = 500$)	Original RR		Sorted Time Series	
	EnRE	*p*-value	*EnRE*(sort)	*p*-value
NSR	1.72 ± 0.47	10^{-7}	1.42 ± 0.51	10^{-7}
CHF	0.65 ± 0.76		0.55 ± 0.81	

Using discriminant analysis, let us assess the quality of division into NSR and CHF groups and build classification functions:

1. When using one parameter of classification *EnRE*, we receive the following indicators of quality of division into NSR and CHF groups

$Se = 0.66$; $Sp = 0.93$; $Acc = 0.83$

and classification function (1 variable *EnRE*; Wilks' Lambda: 0.56 approx. F $(1,81)$ = 62.43; $p < 10^{-7}$):

NSR = 4.96 * *EnRE* – 4.71

CHF = 1.87 * *EnRE* – 1.66

2. When using two parameters of classification, *EnRE* and *EnRE*(sort) for sorted series, we receive the following indicators of quality of division into NSR and CHF groups

$Se = 0.76$; $Sp = 0.98$; $Acc = 0.90$

and classification function (2 variables *EnRE* and *EnRE*(sort); Wilks' Lambda: 0.44; approx. F $(2,80)$ = 50.54; $p < 10^{-7}$):

NSR = 19.28 * *EnRE* – 13.85**EnRE*(sort) – 7.22

CHF = 6.91 * *EnRE* – 4.87**EnRE*(sort) – 1.97

Therefore, proposed generalized form for Robust Entropy Estimator *EnRE* allows, with high accuracy, to divide NSR and CHF groups in short records ($N = 500$), which had remained unachieved in [8] by the way of Multiscale Entropy Analysis, and presents additional advantages provided by *EnRE* in case of structural changes in series (such as sorting). Quality of classification achieved by using two variables *EnRE* and *EnRE*(sort) is superior to results received in [8] by way of Multiscale Entropy Analysis for RR segments with length of $N = 1000$, $N = 2000$ and $N = 5000$.

CONCLUSIONS

Developed by the authors and presented in this article are special methods for statistical analysis of Time Series:

1. Time Series Mann-Whitney M-test – an analogue of the known nonparametric Mann-Whitney U-test for two Time Series with an equal number of elements;

2. Nominal Time Series Measure – a statistical estimator of dynamics of a nominal series consisting of «0» (no) and «1» (yes);

3. Time Series Entropy *EnRE* — specially developed robust formula for a Time Series, intended for calculation of nonlinear stochastic measure of order or disorder, popular in various researches.

All presented methods are accompanied by a detailed demonstration of capacity for statistical analysis of medical Time Series:

1. Analysis of growth dynamics of boys and girls aged 6–7–8 years (WHO);

2. Analysis of number of seizures and choice of anti-epileptic drugs (NSE);

3. Time series entropy *EnRE* for Detecting Congestive Heart Failure by standard 5-minute HRV samples (MIT-BIH RR database).

It has been noted that, in every case, using the named special methods for statistical analysis of medical Time Series allows one to avoid errors in interpreting results received through statistical methods and substantially increases the accuracy of statistical analysis of medical Time Series.

REFERENCES

1. Mercedes de Onis et al. (2007). Development of a WHO growth reference for school-aged children and adolescents.Bulletin of the World Health Organization, 85, 660–667.

2. Goldberger, A.L. et al. (2000).Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation, 101, 215–220.

3. Mann, H.B., Whitney, D.R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Annals of Mathematical Statistics, 18(1),50–60. DOI:10.1214/aoms/1177730491.
4. Neal, A., Parra-Henao, G. (2007).Uses of entropy in medical research.Revista CES MEDICINA, 21(1), 65–75.
5. Boregowda, S., Handy, R., Sleeth, D., Merryweather, A. (2016). Measuring Entropy Change in a Human Physiological System.Journal of Thermodynamics, 1, 1–8. DOI:10.1155/2016/4932710.
6. Humeau-Heurtier, A. (ed.) (2016).Special Issue «Multiscale Entropy and Its Applications in Medicine and Biology».Entropy, 17–18, 88.
7. Martynenko, A., Raimondi, G., Budreiko, N. (2019). Robust Entropy Estimator for Heart Rate Variability.Journal of Clinical Informatics and Telemedicine, 14 (15), 67–73.
8. Chengyu, L., Rui, G. (2017).Multiscale Entropy Analysis of the Differential RR Interval Time Series Signal and Its Application in Detecting Congestive Heart Failure.Entropy, 19, 251. DOI:10.3390/e19060251.

СТАТИСТИЧНИЙ АНАЛІЗ МЕДИЧНИХ ЧАСОВИХ РЯДІВ

Мартиненко О., Раймонді Д., Сотнікова-Мелешкіна Ж., Даниленко Г., Будрейко М.

Статистичний аналіз даних є необхідним компонентом будь-якого медичного дослідження. Сучасні методи математичної статистики і пакети прикладних комп'ютерних програм надають широкі можливості для статистичного аналізу.Однак, коли набір даних представлений серією даних, упорядкованих за часом, або коли структура і порядок даних є важливими компонентами дослідження, стають необхідні спеціальні підходи до статистичного аналізу.У статті представлені спеціальні статистичні методи, що розроблені авторами для аналізу медичних часових рядів: Мана-Уітні М-тест часових рядів – аналог відомого непараметричного методу Мана-Уітні U-тесту для двох часових рядів з однаковим числом елементів; номінальна міра часових рядів – статистична оцінка динаміки номінального ряду,що складається з «0» (ні) і «1» (так); ентропія часових рядів *EnRE* – спеціально розроблена для часових рядів робастна формула обчислення популярного в наукових дослідженнях нелінійна стохастична міра порядку. Представлені методи супроводжуються докладною демонстрацією можливостей для статистичного аналізу медичних часових рядів: Аналіз динаміки ростухлопчиків и дівчат 6–7–8 років (за даними Всесвітньої Організації Охорони здоров'я); Аналіз числа епілептичних нападів і вибір протиепілептичної терапії (за даними Національної епілептичної спілки, Великобританія); ентропія часових рядів *EnRE* для діагностики серцевої недостатності на підгрунті стандартних 5-хвилинних записів варіабельності серцевого ритму (за даними Массачусетського інституту технологій – Бостонського госпіталю Бет-Ізраель RR база даних). Відзначено, що у всіх випадках застосуваннярозроблених спеціальних методів для статистичного аналізу медичнихчасових рядів дозволяєуникнутипомилок в інтерпретації отриманих результатіі суттєво підвищують точність статистичного аналізу медичних часових рядів.

КЛЮЧОВІ СЛОВА: едичні часові ряди, непараметричний тест, номінальна міра, ентропія

ИНФОРМАЦІЯ ПРО АВТОРІВ

Мартиненко Олександр, д.фіз-мат.н., професор кафедри гігієни та соціальної медицини Харківського національного університету імені В. Н. Каразіна, пл. Свободи, 6, Харків, Україна, 61022, e-mail: Alexander.v.martynenko@karazin.ua, ORCID ID: https://orcid.org/0000-0002-0609-2220

Ж. Раймонди, д.мед.н., проф., Римський університет Ла Сапієнца (Італія), Piazzale Aldo Moro 5, 00185, Рим, Італія, e-mail: gianfrancoraimondi@uniroma1.it

Сотнікова-Мелешкіна Жанна, к.мед.н., зав. кафедри гігієни та соціальної медицини Харківського національного університету імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022, e-mail: zhanna.v.sotnikova@karazin.ua, ORCID ID: https://orcid.org/0000-0001-5534-8264

Даниленко Георгій, д.мед.н., проф., ДУ «Інститут охорони здоров'я дітей та підлітків НАМН України», пр. Ювілейний, 52-а, Харків, Україна, 61153, e-mail: danylenkog@ukr.net, ORCID ID: https://orcid.org/0000-0001-7086-2720

Будрейко Микита, асистент кафедри гігієни та соціальної медицини Харківського національного університету імені В. Н. Каразіна**,** майдан Свободи, 6, Харків, Україна, 61022, e-mail: nbudreiko@protonmail.com

СТАТИСТИЧЕСКИЙ АНАЛИЗ МЕДИЦИНСКИХ ВРЕМЕННЫХ РЯДОВ

Мартыненко А., Раймонди Д., Сотникова-Мелешкина Ж., Даниленко Г., Будрейко Н.

Статистический анализ данных является необходимым компонентом любого медицинского исследования. Современные методы математической статистики и пакеты прикладных компьютерных программ предоставляют широкие возможности для статистического анализа. Однако, когда набор данных представлен серией данных, упорядоченных по времени, или когда структура и порядок данных являются важными компонентами исследования, становятся необходимы специальные подходы к статистическому анализу. В статье представлены специальные статистические методы, разработанные авторами для анализа медицинских временных рядов: Манна-Уитни М-тест временных рядов – аналог известного непараметрического метода Манна-Уитни U-теста для двух временных рядов с одинаковым числом элементов; номинальная мера временных рядов – статистическая оценка динамики номинального ряда, состоящего из «0» (нет) и «1» (да); энтропия временных рядов *EnRE* – специально разработанная для временных рядов робастная формула вычисления популярного в научных исследованиях нелинейная стохастическая мера порядка. Представленные методы сопровождаются подробной демонстрацией возможностью для статистического анализа медицинских временных рядов: Анализ динамики роста мальчиков и девочек 6–7–8лет (по данным Всемирной Организации Здравоохранения); Анализ числа эпилептических приступов и выбор антиэпилептической терапии (по данным Национального эпилептического общества, Великобритания); энтропия временных рядов *EnRE* для диагностики сердечной недостаточности на основании стандартных 5-минутных записей вариабельности сердечного ритма (Массачусетского института технологий – Бостонского госпиталя Бет-Израиль база данных RR).Отмечено, что во всех случаях использование указанных специальных методов для статистического анализа медицинских временных рядов позволяет избежать ошибок в интерпретации полученных результатов и существенно повышают точность статистического анализа медицинских временных рядов.

КЛЮЧЕВЫЕ СЛОВА: медицинские временные ряды, непараметрический тест, номинальная мера, энтропия

ИНФОРМАЦИЯ ОБ АВТОРАХ

Мартыненко Александр, д.физ-мат.н., профессор кафедры гигиены и социальной медицины Харьковского национального университета имени В. Н. Каразина, пл. Свободи, 6, Харьков, Украина, 61022, e-mail: Alexander.v.martynenko@karazin.ua, ORCID ID: https://orcid.org/0000-0002-0609-2220

Ж. Раймонди, д.мед.н., проф., Университет Рима «Сапиенза», Рим, Италия, Piazzale Aldo Moro 5, 00185, Рим, Италия, e-mail: gianfrancoraimondi@uniroma1.it

Сотникова-Мелешкина Жанна, к.мед.н., зав. кафедры гигиены и социальной медицины Харьковского национального университета имени В. Н. Каразина, пл. Свободи, 6, Харьков, Украина, 61022, e-mail: zhanna.v.sotnikova@karazin.ua, ORCID ID: https://orcid.org/0000-0001-5534-8264

Даниленко Георгий, д. мед. н., проф., ГУ «Институт охраны здоровья детей и подростков НАМН Украины», проспект Юбилейный, 52-а, Харьков, Украина, 61153, e-mail: danylenkog@ukr.net, ORCID ID: https://orcid.org/0000-0001-7086-2720

Будрейко Никита, асистент кафедры гигиены и социальной медицины Харьковского национального университета имени В. Н. Каразина**,** пл. Свободи, 6, Харьков, Украина, 61022, e-mail: nbudreiko@protonmail.com