

Моделювання стохастичних вибірок на основі методу Монте-Карло

Назарчук Микола Андрійович
 Науковий керівник – Кузовик В. Д., д.т.н.
 ПДС, НАУ
 Київ, Україна
 Skor pion2810@ukr.net

Анотація — Представлено алгоритм ітераційного моделювання для малої вибірки даних на основі методу Монте-Карло, що використовується з метою розрахунку нормованих інтервальних оцінок біологічних параметрів людини. В якості прикладу роботи алгоритму розраховані параметри крові та електроенцефалографії людини. Представлений підхід автоматизовано за допомогою статистичних методів програмного середовища *MatLab*.

Ключові слова — метод Монте-Карло, електроенцефалограф, біологічні параметри, інтегральна оцінка

I. ВСТУП

В сучасному світі є потреба якісного і швидкого оцінювання психофізіологічного стану (ПФС) здоров'я операторів екстремальних видів діяльності (ОЕВД), наприклад, льотчиків, полярників, спортсменів, водіїв та інші. Країнами світу витрачаються значні зусилля та кошти для якісного професійного відбору ОЕВД [1], яке ґрунтується на процесі оцінювання психофізіологічного стану організму операторів.

В попередніх дослідженнях [2] було виявлено, що лімбічна система синтезує інформаційно-енергетичне поле (ІЕП) організму людини, параметри якого в повній мірі характеризують психофізіологічний стан гомеостазу організму людини. Оцінювання та прогноз стану роботи ІЕП можна реалізувати на основі комплексу кефалоенцефалографу, із застосуванням попередньої класифікації операторів за типом темпераменту на основі побудованого системного програмного забезпечення. Враховуючи зв'язок лімбічної системи із гуморальною системою організму, додатковим методом ідентифікації стану ІЕП організму ОЕВД обрано аналіз параметрів крові.

Процес реєстрації біологічних даних потребує значних затрат часу та зусиль, через що робота із великою кількістю статистичної вибірки даних не завжди є можливою, що впливає на подальший розрахунок критеріїв професійного відбору ОЕВД. Тому, для підвищення ефективності розрахунку нормованих значень біологічних параметрів, які використовуються в процесі професійного відбору ОЕВД, застосовано ітераційне моделювання за методом Монте-Карло. Серед переваг запропонованого підходу є можливість моделювати стохастичні вибірки даних та контролювати умови, що накладаються на об'єм ітераційної вибірки, які використовуються для розрахунку нормованих оцінок щодо стану роботи ІЕП організму ОЕВД.

Представлений метод являє собою частину комп'ютеризованої експертної системи, що надає можливість медику-спеціалісту зменшити затрати часу та зусиль для

реалізації експериментального дослідження щодо професійного відбору ОЕВД.

II. ОБ'ЄКТ ДОСЛІДЖЕННЯ

Об'єктом дослідження є процес відбору операторів екстремальних видів діяльності та прогнозування їхнього психофізіологічного стану організму на основі визначення нормованих значень біологічних параметрів організму при наявності генеральної вибірки з малою кількістю експериментальних значень.

III. МЕТА ТА ЗАДАЧІ

У зв'язку з наведеним, метою дослідження є розробка алгоритму ітераційного моделювання, при умові малої вибірки даних, а також його комп'ютеризація, для отримання нормованих кількісних оцінок параметрів електроенцефалографії та аналізу крові, за допомогою якого медик-спеціаліст має можливість оцінити стан інформаційно-енергетичного поля організму.

Для досягнення поставленої мети вирішувались наступні задачі:

- 1) проаналізувати існуючі методи ітераційного моделювання стохастичних процесів;
- 2) розробити алгоритм визначення нормованих значень параметрів біооб'єктів;

IV. АНАЛІЗ ЛІТЕРАТУРНИХ ДАНИХ

Сучасні інформаційні технології, що активно розвиваються в медичній науці і практиці охорони здоров'я, орієнтовані на аналіз медико-біологічних даних, розроблення методів вилучення з них інформації, формування інтегральних оцінок стану біосистем [3].

Побудова інтегральних критеріїв оцінювання психофізіологічного стану (ПФС) організму ОЕВД є ефективним при професійному відборі ОЕВД [4-5], але їх реалізація можлива лише за умови отримання статистично достатньої кількості експериментальних даних, що не завжди вдається досягнути за умови проведення окремих експериментів з ОЕВД.

Статистичне моделювання даних стану біологічних систем є ефективним методом збільшення розміру статистичної вибірки, що дозволяє зменшити варіабельність інтегральної оцінки стану досліджуваної системи в умовах малих об'ємів вхідних експериментальних даних [3].

Останні результати зарубіжних досліджень показують ефективність застосування ітераційного моделювання біологічних параметрів для збільшення розміру вибірки [11, 12].

Проте, в представлених дослідженнях розрахунки проводяться з допущенням про нормальність закону розподілу експериментальних даних, що не завжди є коректним при роботі з біологічними об'єктами.

V. МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Для підвищення якості статистичної вибірки реалізовано аналіз випадкових відхилень значень вибірки методом Граббса:

$$G_n = (x_{(n)} - \bar{x}) / S, \quad (1)$$

де $x_{(n)}$ – значення вибірки, яка перевіряється на викид; \bar{x} – середнє арифметичне вибірки; S – середнє квадратичне відхилення вибірки.

Для реалізації ітераційного моделювання за методом Монте-Карло визначається закон розподілу вибірки. Враховуючи, що критерій Пірсона (χ^2) та складений критерій (d) для перевірки на нормальний закон розподілу (гн) не задовольняє або умові довжини вибірки ($n\chi^2 \geq 100$), або універсальності (d-критерій підходить лише для ідентифікації нормального закону розподілу), на початковому етапі реалізується перевірка вибірки на нормальний закон розподілу за критерієм Колмогорова-Смірнова (λ), використовуючи значення найбільшої абсолютної різниці:

$$\lambda = \sqrt{n} * \max(w_i - w_i^T), \quad (2)$$

де $w_i = \sum_{j=1}^i \frac{n_j}{n}$ – спостережувані значення сумарної частоти

включень в i -й інтервал гістограми закону розподілу; n_j – кількість включень випадкової величини в j -й інтервал гістограми закону розподілу; n – кількість спостережуваних випадкових чисел; $w_i^T = F(x_i)$ – теоретичні значення сумарної частоти включень в i -й інтервал гістограми закону розподілу, розраховані за законом розподілу $F(x)$ випадкової величини ζ ; x_i – середина i -ого інтервалу гістограми закону розподілу. Критерій Колмогорова-Смірнова також можна використовувати для перевірки на наступні закони розподілу вибірки:

- нормальний закон розподілу (гн);
- експоненціальний закон розподілу (ге);
- логіарифмічно-нормальний закон розподілу (гл);
- закон розподілу χ -квадрат (гх);
- закон розподілу Релея (гр);
- закон розподілу Вейбула (гв).

Таким чином, провівши аналіз, було виявлено, що критерій Колмогорова-Смірнова є універсальним для визначення виду закону розподілу, а також його можна використовувати для вибірок з малою кількістю даних.

Для збільшення ефективності ітераційного моделювання при умові виявлення нормальності закону розподілу вибірки в алгоритмі застосовується робастний метод обробки даних, який дозволяє імітувати наближення статистичних оцінок малої вибірки даних до нормального закону великої кількості даних. Розрахунок за робастним методом комп'ютеризовано за наступним алгоритмом [4]:

наявні вибіркові експериментальні дані ранжуються по наростанню;
 розраховується медіана ранжованого ряду: $M = \text{med}\{x_i\}$;
 розраховується абсолютне відхилення експериментальних даних від медіани: $|x_i - M|$;

ранжуються абсолютні відхилення по наростанню: $|x_i - M|$; $\min \dots |x_i - M|$;

розраховується медіана для ранжованого абсолютного відхилення: $MAD_n = \text{med}\{|x_i - M_n|\}$;

розраховується робастна оцінка СКВ: $S_{\text{роб}}(x_i) = K(n) \cdot MAD_n$.

Для ітераційного моделювання біологічних параметрів використано метод Монте-Карло, який комп'ютеризовано за наступним алгоритмом:

задавши законом розподілу (r), за експериментальною вибіркою визначаються статистичні оцінки цієї вибірки: середнє квадратичне відхилення S (СКВ); середнє арифметичне μ ;

завдяки лічильнику випадкових чисел, робота якого задана на основі СКВ розподілу експериментальних даних (S) та відхиленню (Sigma), задається шум вибірки (Noise);

додається сформований вектор шуму (NoiseArray) до середнього арифметичного реальної вибірки для розрахунку вихідного вектору імітованих значень (Y).

Використовуючи змодельовані значення параметрів електроенцефалографії та параметрів аналізу крові певної підгрупи операторів, розраховується інтервальна оцінка (V) при відомому середньому квадратичному відхиленні (СКВ) та при обраному рівні значущості $\alpha = 0,05$:

$$P \left(\bar{x} - t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n-1}} < v < \bar{x} + t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n-1}} \right) = 1 - \alpha, \quad (3)$$

де n – кількість вибірки; $t_{\alpha, n-1}$ – коефіцієнт Стюдента; \bar{x} – середнє арифметичне вибірки; S – середнє квадратичне відхилення.

VI. ВИСНОВКИ

Проведено аналіз існуючих методів ітераційного моделювання стохастичних процесів, серед яких обрано метод Монте-Карло як оптимальний для моделювання біологічних параметрів організму операторів екстремальних видів діяльності.

Розроблено та комп'ютеризовано алгоритм визначення нормованих інтервальних оцінок параметрів біооб'єктів, завдяки якому враховуються випадкові відхилення вибірки, закон розподілу вибірки, застосовується робастний метод, який збільшує достовірність моделювання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Бодров, В. А. Психология профессиональной пригодности [Текст]: учебное пособие / В. А. Бодров. – М.: ПЕР СЭ, 2006. – 511 с.
- [2] Кузовик, В. Д. Статистична обробка параметрів перехідних процесів біоритмів кори головного мозку [Текст] / В. Д. Кузовик, А. Д. Гордєєв // Технологічний аудит та резерви виробництва. – 2016. – № 4/2 (30). – С. 59–64. doi:10.15587/2312-8372.2016.74649
- [3] Фокин, В. А. Статистическое моделирование данных при оценке состояния биологических систем [Текст] / В. А. Фокин // Известия Томского политехнического университета. – 2007. – Т. 311, № 5. – С. 132–135.
- [4] Володарський, С. Т. Статистичне оцінювання професійної придатності операторів екстремальних видів діяльності [Текст] / С. Т. Володарський, О. В. Булигіна // Інформаційні технології та комп'ютерна інженерія. – 2012. – № 3. – С. 71–78.
- [5] Кузовик, В. Д. Апаратно-програмний комплекс для оцінювання психофізіологічного стану оператора [Текст] / В. Д. Кузовик, А. Д. Гордєєв // Технологічний аудит та резерви виробництва. – 2014. – № 1/5 (15). – С. 44–46. doi:10.15587/2312-8372.2014.21740