

# Методологія наукових досліджень

## Scientific research methodology

Шановні колеги! У рубриці „Методологія наукових досліджень” редакція продовжує публікацію матеріалів, що пов’язані з найважливішими аспектами наукової і навчальної діяльності: організаційно-методичним забезпеченням наукових видань, загальними принципами статистичного, біометричного і математичного супроводження досліджень, а також оригінальними методичними підходами вітчизняних і зарубіжних морфологів.

К.І.Дяговець  
Д.Г.Марченко  
С.Б.Морозова  
Н.С.Петрук  
Л.А.Романенко  
Н.В.Станішевська  
Н.М.Султанова  
Л.А.Філімонова  
І.С.Хріпков  
К.Л.Шамелашвілі  
К.М.Шевченко  
І.В.Твердохліб

ДЗ «Дніпропетровська медична академія МОЗ України»

**Ключові слова:** біометрія, багатовимірні об’єкти, інтегральний параметр, політетична класифікація.

Надійшла: 20.10.2016

Прийнята: 28.11.2016

DOI: <https://doi.org/10.26641/1997-9665.2016.4.84-87>

УДК 578.087.1:61(083.3)

### ВИКОРИСТАННЯ ІНТЕГРАЛЬНИХ ПАРАМЕТРІВ БАГАТОВИМІРНИХ ОБ’ЄКТІВ ДЛЯ РОЗВ’ЯЗАННЯ МЕДИКО-БІОЛОГІЧНИХ ЗАВДАНЬ

**Реферат.** У роботі вказано на деякі помилки та обмеження екстенсивного підходу до вибору математичного апарату для біометричної обробки результатів медико-біологічних досліджень. Виходячи з необхідності оперувати узагальненими характеристиками процесів і явищ, запропоновано використання політетичного кластерного аналізу для визначення інтегральних параметрів об’єктів з урахуванням всіх досліджуваних властивостей і математичним розрахунком ступеня значущості кожного з них. Виділено два роди завдань, які можуть бути вирішені за допомогою запропонованої математичної процедури: політетична класифікація та зіставлення інтегральних параметрів об’єктів.

**Morphologia.** – 2016. – Т. 10, № 4. – С. 84-87.

© К.І.Дяговець, Д.Г.Марченко, С.Б.Морозова, Н.С.Петрук, Л.А.Романенко, Н.В.Станішевська, Н.М.Султанова, Л.А.Філімонова, І.С.Хріпков, К.Л.Шамелашвілі, К.М.Шевченко, І.В.Твердохліб, 2016  
✉ [ivt-08@mail.ru](mailto:ivt-08@mail.ru)

Dyagovets K.I., Marchenko D.G., Morozova S.B., Petruk N.S., Romanenko L.A., Stanishevskaya N.V., Sultanova N.M., Filimonova L.A., Khripkov I.S., Shamelashvili K.L., Shevchenko K.M., Tverdokhlib I.V. Applying of integral parameters of multidimensional objects to solve the medical and biological problems.

**SUMMARY.** The article points out at certain mistakes and limitations of extensive approach towards the choice of mathematical instrument for biometrics processing of medical and biologic research results. Out of the necessity to operate generalizing properties of processes and events, it is suggested to use polythetic cluster analysis for determining integral parameters of the objects considering all properties of interest as well as mathematical determination of the degree of significance of each of them. Two types of tasks are marked out that could be easily solved by means of the suggested procedure: polythetic classification and object integral parameters comparison.

**Key words:** biometrics, multi-dimensional objects, integral parameter, polythetic classification.

#### Citation:

Dyagovets KI, Marchenko DG, Morozova SB, Petruk NS, Romanenko LA, Stanishevskaya NV, Sultanova NM, Filimonova LA, Khripkov IS, Shamelashvili KL, Shevchenko KM, Tverdokhlib IV. [Applying of integral parameters of multidimensional objects to solve the medical and biological problems]. *Morphologia*. 2016;10(4):84-7. Ukrainian.

Напевно можна стверджувати, що практично будь-який розділ математики (навіть такий абстрактний, як топологія) використовується сьогодні для рішення тих або інших медико-біологічних задач. У останні роки поступово розповсюдилася і зміцнилася думка, що якість дослідження, нарівні з іншими обставинами, визначається кількістю використаних методик, поєднанням методів суміжних наукових галузей і, нарешті, збільшенням кількості об'єктів і їх діапазону. При цьому з'явилися тенденції до максимального залучення багатьох різновидів математичного аналізу в обробку результатів медико-біологічного дослідження. Потрібно визнати, що ми розуміємо важливість і перспективність цих тенденцій, але вимушені, проте, проаналізувати типові помилки і обмеження екстенсивного підходу у виборі математичного апарату. Вельми поширеними є помилки "технічного" характеру: наприклад, проведення кореляційного аналізу між параметрами, які перебувають у математичному взаємозв'язку (тобто один показник виступає в ролі арифметичної похідної від іншого); необгрунтоване припущення про лінійний характер взаємозалежності між параметрами при проведенні регресійного аналізу, що приводить до втрати інформації і хибного уявлення про реальні взаємовідносини елементів системи.

Існує, однак, і більш глибока причина, що часто і закономірно приводить до спотворення результатів медико-біологічного дослідження. Причина ця пов'язана з природою об'єктів, що вивчаються. Справа в тому, що в медицині існує лише невелика кількість різновидів об'єктів, про які можна з упевненістю судити як про однорідні. І цілком природно, що практично в кожному керівництві з біометрії оцінка отриманих первинних результатів обов'язково починається з аналізу статистичних розподілень, які визначають вибір математичного апарату для аналізу результатів [1-3]. На жаль, у повсякденній дослідницькій практиці початок математичної обробки часто співпадає з розрахунком середнього значення параметра і інших варіаційних його характеристик, тобто *a priori* припускається однорідність об'єктів дослідження за цим параметром. Аналогічні припущення відбуваються і відносно інших кількісних показників, внаслідок чого формуються невірні уявлення як про об'єкт загалом, так і про окремі його характеристики.

Лише в тому випадку, якщо емпіричний варіаційний ряд відповідає нормальному (Гауссовому) розподіленню, можна приступати до розрахунку і аналізу основних математичних характеристик варіюючих об'єктів. Однак вірогідні ситуації, коли отримані в ході дослідження первинні дані не зможуть формувати варіаційних кривих, що наближаються до нормального або одного з стандартних розподілень. На цьому етапі необхідно застерегти від необгрунтованого

формального визначення кількості класів об'єктів, а також середніх значень параметрів для кожного класу. Рішення першої з означених задач пов'язане з графічною оцінкою кривих розподілення для кожного з параметрів окремо. Потрібно враховувати, що об'єкти можуть розрізнятися за одними параметрами (тобто бути за ними різнорідними), але за іншими показниками складати єдину сукупність зі статистичним розподіленням, близьким до нормального. Отже, кількість класів об'єктів відповідає максимальній кількості піків у кривій одного або декількох параметрів. Складніше стає справа з середніми величинами показників, оскільки максимуми частот не обов'язково повинні співпадати з ними через невідому величину трансгресії в змішаній вибірковій сукупності об'єктів [4]. У зв'язку з цим виникає необхідність вже на етапі набору даних знати, до якого саме класу належить той або інший об'єкт, що в даний момент часу досліджується.

Таким чином, ми наблизились до проблеми, яка пов'язана не стільки з помилками (технічними або методологічними), скільки з обмеженнями тих методів математичного аналізу, які так часто необгрунтовано або передчасно використовуються при обробці результатів досліджень. Одним з центральних ускладнень у цьому відношенні є необхідність кількісного зіставлення тих або інших явищ або процесів не по окремих параметрах, а по їх комплексу, тобто по об'єкту загалом [5]. У зв'язку з цим інтерес дослідників був спрямований на пошук і можливість оперувати інтегральними характеристиками, що включають всі вивчені показники [6-8]. Найчастіше для цієї мети використовується стратегія політетичної класифікації, яка заснована на кластерному аналізі [9]. Різниця  $d$  між сукупностями властивостей двох об'єктів  $M$  і  $N$  визначається формулою:

$$d = \sqrt{\left\{ \sum_{i=1}^{i=n} \left( \frac{x_{iM} - x_{iN}}{\sigma_i} \right)^2 a_i^2 \right\}}$$

де  $n$  - число властивостей, що складають сукупність;  $x_i$  - позначення  $i$ -ї властивості;  $a_i$  - коефіцієнт вагомості (значущості)  $i$ -ї властивості серед інших властивостей;  $\sigma_i$  - середнє квадратичне відхилення  $i$ -ї властивості по всій вибірці об'єктів, в яку входять  $M$  та  $N$ .

При використанні формули виходять з того, що сукупність всіх властивостей об'єкта означається точкою в  $n$ -мірному евклідовому просторі, а  $d$  є вектор, що характеризує різниці положень двох точок, що послідовно складаються по кожній з координат. Узагальнюючі властивості об'єктів можуть бути самими різними і, отже, самими різними можуть бути і їх розмірності. Внаслідок цього числа, що характеризують окремі властивості, можуть відрізнятися на порядки, що робить їх роль в формулі абсолютно не-

рівноцінною. Крім того, якщо не застосувати який-небудь тип коректування, доводиться порівнювати величини з різною розмірністю, що неприпустимо. Введення в формулу величин  $\sigma_i$  дозволяє вирішити ці проблеми [7]. При цьому стає очевидним, що значущість вивчених параметрів для інтегральної характеристики неоднакова: основна трудність в розрахунку інтегрального параметра полягає саме в точному (математичному), а не довільному (гаданому) визначенні коефіцієнтів значущості.

У цій ситуації ми пропонуємо логічне, на наш погляд, рішення задачі: повернемося до того етапу в дослідженні, коли виявилася необхідність знати, до якого класу належить об'єкт, що вивчається в даний момент. Отже, вивчення знає об'єкт, що належить до одного з класів. За деякими параметрами цей об'єкт ближче до середнього значення першого класу, за іншими - до значення, характерного для другого. При цьому участь кожного параметра в кінцевій діагностичній оцінці (1-й або 2-й клас) буде визначатися мірою його значущості в розрахунку інтегральної характеристики. Не інтуїтивне, а точне її визначення, як сказано, - основне ускладнення в рішенні задачі. Рішення це пов'язане з аналізом розподілень значень параметрів. При цьому, зрозуміло, чим вище піки в кривій розподілення і чим глибше між ними межа, тим більше даний параметр визначає приналежність об'єкта до одного з класів. З мірою стирання цієї межі параметр стає все менш і менш важливим аргументом, що відносить об'єкт до того або іншого класу. Якщо ж, нарешті, межа не визначається, то це свідчить про однорідність об'єктів (але тільки по даному, нормально розподіленому параметру) - отже, значущість його у визначенні класу об'єкта наближається до нуля. Таким чином, виникає необхідність кількісно (!) оцінити глибину розділу (виразність межі) між піками в розподіленнях всіх досліджуваних параметрів. Серед істотних характеристик статистичних розподілень найбільший інтерес в цьому відношенні являє величина ексцесу ( $E_x$ ), яка кількісно виражає міру вертикального відхилення емпіричного розподілення від нормального. При поглибленні меж між максимумами частот в розподіленнях параметрів зростає абсолютне значення  $E_x$ , що виражає значущість кожної з ознак в розрахунку інтегрального параметра. Показник  $E_x$  визначається формулою [2]:

$$E_x = \left[ \frac{\sum_{i=1}^n f_i (x_i - x)^4}{n} \right] / S_x^4 - 3$$

де  $n$  - об'єм вибірки;  $x_i$  - числове значення  $i$ -ї варіанти;  $x$  - середня арифметична;  $f_i$  - емпірична частота  $i$ -ї варіанти;  $S_x$  - помилка вибіркової середньої.

Отже, кожний об'єкт характеризується своєю величиною інтегрального параметра, визначеного на основі всіх використаних показників з урахуванням їх значущості. Знаючи величини інтегральних параметрів кожного класу, ми можемо безпомилково виявити "класову приналежність" кожного з об'єктів вже на етапі квантифікації, - з усіма наслідками, що звідси витікають: визначенням середнього значення кожного параметра для типу об'єкта, виявленням кількісного співвідношення між об'єктами різних класів в обмеженому просторі і т.п. Після відповідного угруповання первинних даних стає можливим розрахувати статистичні характеристики об'єктів диференційно по кожному з виявлених класів.

Отримавши принципову можливість точного визначення коефіцієнтів значущості кожного з параметрів багатовимірних об'єктів і, далі, точного розрахунку інтегрального параметра кожного об'єкта, ми маємо реальну можливість виявити відсутність або наявність гетерогенності системи. Іншими словами, стає можливим визначити наявність класів і, якщо вони існують, прослідити динаміку їх появи, розвитку, взаємної трансформації на певному хронологічному відрізку. Найбільш зручним підходом в цьому відношенні є графічний аналіз вказаної динаміки. Цікаво, що отримана вибірка інтегральних параметрів, що мають своє власне статистичне розподілення, також може бути охарактеризована величиною ексцесу (по стандартній процедурі), але в нових умовах значення ексцесу є кількісною оцінкою такого, здавалося б, складного і абстрактного поняття, як гетерогенність системи.

Особливий інтерес становить використання політетичного аналізу в клінічних дослідженнях, де однією з найважливіших задач є оцінка ефективності методів лікування, що застосовуються; часто ця оцінка побудована на суб'єктивному порівнянні стану пацієнта в певні терміни реконвалесцентного періоду і має в своїй основі велими спрощену шкалу. У цій ситуації, зрозуміло, стає проблематичним кількісно охарактеризувати клінічний ефект терапії, порівняти ефективність двох або декількох методів, визначити залежність терапевтичного ефекту від цікавлячих дослідника чинників (початкового рівня пошкодження, типу і характеру патології, віку і т.п.), а також вирішити ряд інших конкретних клінічних задач. Крім того, численні параметри, що досліджуються, зазнають неоднаправлених і несинхронних зрушень на етапах терапії, що не дозволяє однозначно оцінювати її ефективність. З іншого боку, перед клініцистами стоїть задача охарактеризувати не окремі показники (що не має певного клінічного значення), а їх комплекс, тобто стан хворого загалом. У цьому відношенні використання інтегральних характеристик, розрахованих по означеній процедурі, представляється цілком доречним для подолання вказаного ускладнен-

ня.

У конкретному випадку розрахунок інтегрального параметра, наприклад, термографічного комплексу у хворого N виглядає таким чином. У даного хворого через 1 місяць після проведення реконструктивно-відновного втручання з приводу травми кисті величина термоасиметрії становила 1,63°C; термографічний індекс - 4,08; середня температура пошкодженої кисті - 32,4°C; максимальна  $t^{\circ}$  - 32,9°C; мінімальна  $t^{\circ}$  - 31,4°C. Коефіцієнти значущості наведених 5 параметрів, що розраховані при визначенні ексцесу статистичних розподілень, склали відповідно 1,07; 1,32; 0,42; 0,30; 0,28. Заздалегідь розрахувавши середнє квадратичне відхилення і підставивши в формулу ці значення, ми отримуємо абстрактне число 4,21; виконавши таку ж процедуру зі значеннями тих же параметрів, зафіксованими в кінці 3-го місяця лікування, отримано значення 4,97. Таким чином, за минулий 2-місячний період інтегральна характеристика термографічного комплексу у хворого N поліпшилася на 18,1%. Аналогічним чином розраховуються інтегральні параметри для всіх комплексів первинних даних і блоків більш високих рівнів. Важливо підкреслити, що величина інтегрального параметра

не має ні розмірності, ні конкретного клінічного змісту, хоча розраховується вона на основі комплексу цілком певних клініко-лабораторних і соціально-психологічних критеріїв. Основне значення використання абстрактних інтегральних характеристик полягає саме в узагальненій кількісній оцінці виділених блоків і стану хворого загалом, що дозволяє оцінювати динаміку течії реконвалесцентного періоду у кожного пацієнта і об'єктивно порівнювати характеристики груп або категорій хворих між собою.

#### Підсумок

У принциповому вигляді використання інтегральних параметрів багатовимірних об'єктів дозволяє вирішити медико-біологічні задачі двох напрямків. По-перше, стає можливим проведення політетичної (поліпараметричної) класифікації об'єктів дослідження в тому випадку, коли неможливо заздалегідь визначити, чи відокремлені класи певно існують, чи справа просто в широкому варіюванні характеристик об'єктів. По-друге, запропонований спосіб оцінки об'єктів з багатьма властивостями дає змогу описати той чи інший процес загалом (не тільки окремі його прояви) з математично розрахованими ступенями значущості кожної вивченої характеристики.

#### Список літератури

#### References

1. Borovkov AA. [Mathematical statistics]. Moscow: Nauka; 1984. 286 p. Russian.
2. Lakin GF. [Biometrics]. 4th ed rev and enl. Moscow: Vysshaya shkola; 1990. 352 p. Russian.
3. Urbach VY [Statistical analysis of biological and statistical research]. Moscow: Nauka; 1975. 346 p. Russian.
4. Tverdokhle IV, Shponka IS, Mashtalir MA. [Applied biometrics for morfologist].- Dnepropetrovsk: Porogi; 1996. 226 p. Russian.
5. Gubler EV, Genkin AA. [The use of nonparametric statistics in biomedical research]. Lenin-grad: Meditsina; 1973. 284 p. Russian.
6. Gorelik AL, Skripkin VA. [Methods of recognition]. Moscow: Statistika; 1977. 382 p. Russian.
7. Tverdokhle IV. [Mathematical aspects of geteromorfy analysis // Bulletin of the problems of modern medicine]. 1994;5:37-45. Russian.
8. Peifer U. The evaluation of large test fields for morphometric studies in electron microscopy. Path Res Pract. 1980;166:188-202.
9. Bailey K. Systems as clusters. Behav Sci. 1985;30(2):98-107.

**Дяговец Е.И., Марченко Д.Г., Морозова С.Б., Петрук Н.С., Романенко Л.А., Станишевская Н.В., Султанова Н.Н., Филимонова Л.А., Хрипков И.С., Шамелашвили К.Л., Шевченко Е.Н., Твердохлеб И.В. Использование интегральных параметров многомерных объектов для решения медико-биологических задач.**

**Реферат.** В работе указаны некоторые ошибки и ограничения экстенсивного подхода к выбору математического аппарата для биометрической обработки результатов медико-биологических исследований. Исходя из необходимости оперировать обобщающими характеристиками процессов и явлений, предложено использование политетического кластерного анализа для определения интегральных параметров объектов с учетом всех изучаемых свойств и математическим расчетом степени значимости каждого из них. Выделено два рода задач, которые могут быть решены с помощью предложенной математической процедуры: политетическая классификация и сопоставление интегральных параметров объектов.

**Ключевые слова:** биометрия, многомерные объекты, интегральный параметр, политетическая классификация.