

УДК 389.14

doi:10.20998/2413-4295.2017.53.07

## ИНФОРМАЦИОННЫЙ АНАЛИЗ МОДЕЛЕЙ ПЛАНИРОВАНИЯ МЕДИКО-БИОЛОГИЧЕСКОГО ЭКСПЕРИМЕНТА ПРИ КОНТРОЛЕ ДИСКРЕТНЫХ ИЗМЕНЕНИЙ УРОВНЕЙ ВЛИЯЮЩЕГО ФАКТОРА

**Р. С. ТОМАШЕВСКИЙ**

кафедра промышленной и биомедицинской электроники, НТУ «ХПИ», Харьков, УКРАИНА  
\*email: romiocat.khpi@gmail.com

**АННОТАЦИЯ** В статье рассмотрены две базовые статистические модели для обозначения эффектов уровня влияющего фактора, а именно параметрическая модель и модель компонент дисперсий. Проведен сравнительный анализ эффективности двух моделей однофакторного медицинского эксперимента с односторонней классификацией результатов по получаемой информации, учитывающей априорную неопределенность модели. Получены аналитические выражения для определения информационного содержания рассматриваемых моделей исходя из параметров плана медицинского эксперимента. Проведен анализ рассмотренных статистических моделей в результате, которого определены эффекты влияния каждого из параметров на эффективность использования модели и приведены соответствующие рекомендации по их использованию.

**Ключевые слова:** дисперсионный анализ; биомедицинский эксперимент; параметрическая модель; модель компонент дисперсий; параметр нецентральности; медицинские измерения; критериальная статистика.

## INFORMATION ANALYSIS OF MODELS OF PLANNING OF THE MEDICAL-BIOLOGICAL EXPERIMENT AT CONTROL OF DISCRETE CHANGES OF LEVELS OF THE IMPACTANT FACTOR

**R. S. TOMASHEVSKY**

Department of Industrial and Biomedical Electronics, NTU "KhPI", Kharkov, UKRAINE

**ABSTRACT** The paper considers the main approaches of modern evidence-based medicine and the role of the medical-biological experiment in it. It is established that one of the most used method in medical practice is the method of single-factor variance analysis (ANOVA). Two basic statistical models for designating effects of the level of the influencing factor (parametric model and the model of the variance components) are considered.

**Objective:** A comparative analysis of the effectiveness of two models of a single-factor medical experiment with a one-stage classification of the results from the received information, taking into account the a priori uncertainty of the model.

**Solution methods:** The positions of the information theory of measurements and the model of representation of the experimental results, taking into account the influence of the factor, were used for information analysis. In the analysis, two variants of the results of the biomedical experiment were considered: the factor influence takes place or does not take place in the groups under consideration.

**Expressions for variance** are obtained in case of confirmation of the main or alternative hypothesis. On the basis of the analysis, expressions were obtained to determine the information content of the models under consideration based on the parameters of the medical experiment plan. The analysis of expressions for estimating the amount of information contained allowed us to compare two basic statistical models for describing effects of the level of the influencing factor in the analysis of variance: parametric and random.

**Results:** For all other things being equal, the random model has a higher information content, since a priori it takes into account the presence of probabilistic properties in the levels of the influencing factor. Positive dependence of the information content of models on the number of groups of measurements is obtained. The model of the dispersion components is more sensitive to a change in this parameter of the experimental plan than the parametric one.

**Conclusions:** The obtained analytical expressions allow estimating the effect of the parameters of the experimental plan, namely the number of groups, the volume of samples, the level of significance, on the informational content of models, and use this parameter as the criteria of the choice of that one or another model.

**Keywords:** variance analysis; biomedical experiment; parametric model; model of variance components; noncentral parameter; medical measurements; criterial statistics.

### Введение

Современная медицинская наука, как и биология, использует научный эксперимент как основной показатель эффективности того или иного технологического инструмента, будь то терапевтический метод (агент) или диагностический тест [1-3]. Экспериментальная проверка и

подтверждение медицинского решения лежит в основе сравнительно нового направления в медицинской науке – доказательной медицины. При массовом использовании и внедрении в практику данного подхода чрезвычайно важными становятся показатели эффективности медико-биологического эксперимента, то есть отношение количества полученной информации к количеству затраченных ресурсов

(лабораторные образцы, реактивы, рабочее время и т.д.) [4-6]. При этом одной из важнейших задач будет количественная оценка информационной составляющей эксперимента в условиях высокой степени сложности и разнообразия биологических образцов и ограниченности размера выборки результатов измерений. Также немаловажным становится вопрос планирования медико-биологического эксперимента и выбор статистической модели факторной нагрузки.

Сфера использования научного эксперимента в медико-биологических науках неимоверно широка [1-3, 7-9] и может быть оценена по количеству задач, решаемых с помощью этого инструмента познания, а это:

- 1) Получение новых данных, для исследования новых явлений и свойств живых организмов.
- 2) Определение факта и степени влияния физического (химического, биологического) фактора на конкретный показатель состояния или функционирования живого организма, или его частей;
- 3) Определение диагностической ценности и эффективности информационного показателя биологического объекта, разработка диагностических протоколов;
- 4) Оценка степени негативного влияния на биологический объект определенного фактора (агента), прогнозирование последствий такого влияния.

В большинстве случаев при лабораторной постановке таких экспериментов исследователь оперирует ограниченным нормированным набором факторов воздействия. В простейшем случае это двухуровневая шкала со значениями «Наличие фактора/отсутствие фактора»; в более сложном набор уровней фактора, значение которых изменяется от минимума до максимума по равномерной или неравномерной, но детерминированной шкале; и наиболее сложном, где фактор изменяется по определенному закону и его значения в определенные моменты времени, могут быть сгруппированы по уровням.

Выбор определенной модели существенно зависит от начальных предположений и желаемых результатов и выводов исследования. Однако в ряде случаев исследователь становится перед сложным выбором статистической модели, от правильности которого будет зависеть эффективность диагностики и оценка динамики того или иного заболевания [9-12].

### Цель работы

Целью данной работы является сравнительный анализ эффективности двух моделей однофакторного медицинского эксперимента с односторонней классификацией результатов по получаемой информации, учитывающей априорную неопределенность модели.

### Анализ информационных моделей

Стандартные статистические методы однофакторного дисперсионного анализа (ANOVA) результатов медико-биологических экспериментов выделяют две базовых статистические модели для обозначения эффектов уровня влияющего фактора:

- параметрическая модель,
- случайная (модель компонент дисперсий).

Обе модели используют общее условие дискретности (скачкообразного изменения) уровней влияющего фактора (например,  $k$  способов лечения).

Однако, если в параметрической модели эти  $k$  уровней фиксированы в форме параметров и могут однозначно воспроизводиться ( $k$  конкретных способов лечения), то в случайной модели эти уровни характеризуются неопределенностью ( $k$  выборок, взятых сгруппированных по подмножествам, способов лечения).

Пусть  $X$  – измеряемая в ходе эксперимента физическая (биологическая) величина, представляемая, по окончании эксперимента, набором сгруппированных (по номерам  $j = \overline{1, k}$  уровней фактора)

$$X = \{x_{1n_1}, \dots, x_{jn_j}, \dots, x_{kn_k}\} \quad (1)$$

где  $n_1, \dots, n_j, \dots, n_k$  количество измерений в каждой из групп. Обозначим индексом  $i$  – номер отдельного измерения в  $j$  –й группе ( $i = \overline{1, n_j}$ ). Введем верхний индекс (1) или (2) для результата измерений, соответственно, для параметрической или случайной моделей ( $x_{ji}^{(1)}$  или  $x_{ji}^{(2)}$ ).

Тогда модели результатов измерений для планов параметрического представления уровней влияющего фактора будут иметь вид:

$$x_{ji}^{(1)} = \bar{X} + \gamma_j + z_{ji}, \quad (2)$$

$$x_{ji}^{(2)} = \bar{X} + u_j + z_{ji}, \quad (3)$$

где  $\gamma_j$  и  $u_j$  отклонения значений  $x_{ji}^{(1)}$  или  $x_{ji}^{(2)}$  от общего среднего  $\bar{X}$ ;  $z_{ji}$  – случайное остаточное отклонение.

Вероятностные свойства моделей (2) и (3), включая свойства составляющих  $\gamma_j$ ,  $u_j$ ,  $z_{ji}$  детально описаны в доступной литературе [13].

Поскольку, для проверки основной гипотезы ( $H_0$ : «факторное влияние отсутствует») используют критериальную  $F$  – статистику в форме отношения среднего квадрата отклонений  $\bar{Q}_1$  между группами к среднему квадрату внутри групп  $\bar{Q}_2$ ,

$$F = \frac{\bar{Q}_1}{Q_2}, \quad (4)$$

то имеет смысл оценить количество измерительной (и логической, т.е. диагностической или контрольной) информации, содержащейся в  $F$  – статистике (4).

Причем, такую оценку можно получить отдельно для параметрической и случайной моделей медико-биологического эксперимента (при справедливости или несправедливости основной гипотезы  $H_0$ ).

Хотя формально, обе модели планирования эксперимента реализуют тест на значимость [14], т.е. оперируют только вероятностью ошибки первого рода (уровнем значимости) и проверяют только единственную основную гипотезу  $H_0$ , введем дополнительно и альтернативную гипотезу ( $H_1$ : «имеет место факторное влияние»).

Вероятностные модели информационной теории измерений [15-17] позволяют получить уравнение для оценки количества  $I$  ожидаемой измерительной информации для любых вариантов законов распределения случайных результатов измерения

$$I = H_1 - H_2, \quad (5)$$

где  $H_1$  и  $H_2$  – исходная (до измерения) и остаточная (после измерения) энтропии случайной величины, неслучайный параметр которого содержит ожидаемую информацию.

Если, в общем случае,  $X$  – случайная измеримая величина,  $\Delta x$  – погрешность ее различения,  $f(x)$  – закон распределения величины  $X$ , а  $x_N$  – реализация результата измерения, то получим [16].

$$\begin{cases} H_1 = - \int_{-\infty}^{\infty} f(x) \log(x) dx - \log \Delta x \\ H_2 = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, x_N) \log \frac{f(x, x_N)}{f(x_N)} dx dx_N - \log \Delta x \end{cases}, \quad (6)$$

где  $f(x, x_N)$  – двумерная плотность распределения величин  $X$  и  $X_N$ .

Если  $f(x)$  и  $f(x, x_N)$  – одномерный и двумерный нормальные законы распределения, то

$$I = \frac{1}{2} \log \left[ 1 + \left( \frac{\sigma_x}{\sigma_{\Delta x}} \right)^2 \right], \quad (7)$$

где  $\sigma_x^2$  и  $\sigma_{\Delta x}^2$  – дисперсии величины  $X$  до и после измерения. Выражение (7) удобно для получения оценок количества информации для любых, даже несимметричных законов распределения, особенно для оценки энтропии  $H_1$  [18, 19]. Более того, [18, 16],

что при одинаковой дисперсии  $\sigma_{\Delta x}^2$ , погрешность, распределенная по нормальному закону, вносит большее дезинформационное влияние, чем погрешность, распределенная по закону, который отличается от нормального.

Случайная  $F$  – статистика (4) имеет плотность либо центрального (справедлива гипотеза  $H_0$ ), либо нецентрального (справедлива гипотеза  $H_1$ )  $F$  – распределения Фишера [14], причем повышение числа степеней свободы и параметра нецентральности только приближают это распределение к нормальному.

Используем информационную модель (7) для оценки количества получаемой, в ходе эксперимента, информации по измеряемым величинам  $F^{(1)}$  и  $F^{(2)}$  вычисляемым по отношению (4) для параметрического ( $x_{ji}^{(1)}$ ) и случайного ( $x_{ji}^{(2)}$ ) вариантов использования измеряемой величины  $X(1)$ .

Введем обозначения для дисперсий  $F$  – статистики, с учетом модели плана эксперимента и гипотез  $H_0$  и  $H_1$ , табл. 1

Таблица 1 – Обозначения для дисперсии критерияльной – статистики

| Модель эксперимента | Справедливость гипотезы |             |
|---------------------|-------------------------|-------------|
|                     | $H_0$                   | $H_1$       |
| Параметрическая     | $D_0^{(1)}$             | $D_1^{(1)}$ |
| Случайная           | $D_0^{(2)}$             | $D_1^{(2)}$ |

Обозначим символами  $q_0$  и  $q_1$  априорные вероятности гипотез  $H_0$  и  $H_1$ . Зададим план эксперимента в виде набора условий:

$$(N, k, n, \alpha), \quad (8)$$

где  $N$  – общее число всех результатов измерения величины  $X(1)$ ,  $k$  – число уровней влияющего фактора (число групп),  $n$  – число измерений в группе ( $n_1 = n_2 = \dots = n_j = \dots = n_k = n$ ),  $\alpha$  – уровень значимости.

Учтем, что  $F$  – статистика, при справедливости основной гипотезы  $H_0$ , имеет плотность центрального  $F$  – распределения ( $F_{v_1, v_2}$ ) с  $v_1 = k - 1$  и  $v_2 = N - k$  степенями свободы [20].

$$F \sim F_{v_1, v_2} \quad (9)$$

Тогда дисперсии  $D_0^{(1)}$  и  $D_0^{(2)}$  одинаковы и определяются выражениями [17, 19].

$$D_0 = D_0^{(1)} = D_0^{(2)} = \frac{2v_2^2}{(v_2 - 2) \cdot (v_2 - 4)} \cdot \left[ 1 + \frac{2}{v_1} \right]. \quad (10)$$

Введем обозначения  $\sigma_\gamma^2$ ,  $\sigma_u^2$  и  $\sigma^2$  для отклонений  $\gamma_j$ ,  $u_j$  и  $z_{ji}$  моделей (2), (3).

Если справедлива альтернативная гипотеза  $H_1$ , то  $F$  – статистика будет иметь уже плотность нецентрального  $F$  – распределения с  $\nu_1$ ,  $\nu_2$  степенями свободы и соответствующим (для параметрической или случайной модели плана эксперимента) параметром нецентральности  $\lambda^{(1)}$  или  $\lambda^{(2)}$ .

$$\begin{cases} F \sim F_{\nu_1, \nu_2, \lambda^{(1)}}, \\ F \sim F_{\nu_1, \nu_2, \lambda^{(2)}}. \end{cases} \quad (11)$$

Можно показать [13, 14], что при одинаковом ( $n$ ) числе измерений в группах, параметры нецентральности имеют, с учетом выражений (5), (3), вид

$$\begin{cases} \lambda^{(1)} = (k-1) \cdot n \frac{\sigma_\gamma^2}{\sigma^2}, \\ \lambda^{(2)} = (k-1) \cdot n \frac{\sigma_u^2}{\sigma^2}. \end{cases} \quad (12)$$

Для случайной модели факторного влияния существует удобный вариант вероятностного представления нецентрального  $F$  – распределения, как линейной комбинации [15, 17]

$$F_{\nu_1, \nu_2} \cdot \left(1 + n \frac{\sigma_u^2}{\sigma^2}\right), \quad (13)$$

что позволяет определить  $D_1^{(2)}$  в форме выражения

$$D_1^{(2)} = D_0 \left(1 + \frac{\lambda^{(2)}}{\nu_1}\right)^2. \quad (14)$$

Общее выражение для дисперсии  $D_1^{(1)}$  и  $D_1^{(2)}$  будут представлены уравнениями [19]

$$D_1^{(1)} = \frac{2\nu_2^2}{(\nu_2 - 2) \cdot (\nu_2 - 4)} \cdot \left[ \left( \frac{1 + \lambda^{(1)}}{(\nu_2 - 2)} \right)^2 + \left( \frac{1 + 2\lambda^{(1)}}{\nu_2} \right) \right], \quad (15)$$

$$D_1^{(2)} = \frac{2\nu_2^2 \cdot (\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2 \cdot (\nu_2 - 4)} \cdot \left[ 1 + \frac{\lambda^{(2)}}{\nu_1} \right]^2. \quad (16)$$

### Оценка количества информации

При оценке количества информации результатов медико-биологического эксперимента (7) учтем, что:

а) для параметрической модели факторного влияния

$$\sigma_x^2 = q_0 D_0 + q_1 D_1^{(1)}, \quad (17)$$

б) для случайной модели

$$\sigma_z^2 = q_0 D_0 + q_1 D_1^{(2)}, \quad (18)$$

Дисперсия  $\sigma_{\Delta x}^2$  определяется дисперсией критериальной  $F$  – статистики, которая, для параметрической модели, должна учитывать риск (уровень значимости)  $\alpha$  и, задаваемую, по этому риску с степенями свободы  $\nu_1$ ,  $\nu_2$  критическую статистику

$$F_{кр} = F_{\nu_1, \nu_2, \alpha}. \quad (19)$$

Для случайной модели, дисперсия  $\sigma_{\Delta x}^2$  должна учитывать не только риск первого рода  $\alpha$ , но и риск второго рода  $\beta$ . Эти риски обуславливают априори заданные соотношения дисперсий  $\sigma_{u_j}^2$  и  $\sigma_{z_{ji}}^2$  для отклонений  $u_j$  и  $z_{ji}$  модели результата измерения (3) [13].

Дисперсия  $\sigma_{\Delta x}^2$  может быть определена, как дисперсия дискретной случайной величины, имеющей только два возможных значения:

а) среднее  $m_{F_0}$  статистики  $F_{\nu_1, \nu_2}$  и значение, задаваемое величиной выражение (19) – для параметрической модели;

б) среднее  $m_{F_0}$  статистики  $F_{\nu_1, \nu_2}$  и среднее  $m_{F_1}$  статистики  $F_{\nu_1, \nu_2, \lambda^{(2)}}$  – для случайной модели.

Выражения для определения  $m_{F_0}$ ,  $m_{F_1}$  и  $F_{кр}$  имеют вид [13]

$$m_{F_0} = \frac{\nu_2}{\nu_2 - 2}; \quad (20)$$

$$m_{F_1} = \frac{\nu_2}{\nu_2 - 2} \left(1 + \frac{\lambda}{\nu_1}\right), \quad (21)$$

где  $\lambda = \lambda^{(1)}$  или  $\lambda = \lambda^{(2)}$ ,  $F_{кр} - \alpha$  – процентная точка центрального  $F$  – распределения  $f(F)$  Фишера – Снедекора, имеющего  $\nu_1$  и  $\nu_2$  степеней свободы, обеспечивающих выполнение условия [19]

$$\int_{F_{кр}}^{\infty} f(F) dF = \alpha. \quad (22)$$

Общие выражения для дисперсии  $\sigma_{\Delta x}^2$  по вариантам а) и б) имеют следующий вид:

$$\text{а) } \sigma_{\Delta x}^2 = \sigma_{0\Delta x}^2 = m_{F_0}^2 (1 - \alpha) + F_{кр}^2 \cdot \alpha - [m_{F_0} \cdot (1 - \alpha) + F_{кр} \cdot \alpha]^2; \quad (23)$$

$$\sigma_{\Delta x}^2 = \sigma_{1\Delta x}^2 = q_0 \left\{ (m_{F_0} - \bar{F}_0)(1-\alpha) + (m_{F_1} - \bar{F}_0)^2 \cdot \beta \right\} + q_1 \left\{ (m_{F_1} - \bar{F}_1)^2 (1-\beta) + (m_{F_0} - \bar{F}_1)^2 \cdot \beta \right\} \quad (24)$$

$$\begin{cases} \bar{F}_0 = m_{F_0} \cdot (1-\alpha) + m_{F_1} \cdot \alpha; \\ \bar{F}_1 = m_{F_0} \cdot \beta + m_{F_1} \cdot (1-\beta). \end{cases} \quad (25)$$

С учетом выражений (20) – (24), дисперсии  $\sigma_{0\Delta x}^2$  (для параметрической модели) и  $\sigma_{1\Delta x}^2$  (для случайной модели факторного влияния) будут иметь вид:

$$\begin{cases} \sigma_{0\Delta x}^2 = \alpha \cdot (1-\alpha) \left[ F_{sp} - \frac{v_2}{v_2-2} \right]^2; \\ \sigma_{1\Delta x}^2 = [\alpha \cdot (1-\alpha) + q_1 \beta (1-\beta)] \left( \frac{v_2}{v_2-2} \right)^2 \left( \frac{\lambda}{v_1} \right)^2. \end{cases} \quad (26)$$

Используя выражения (9), (13), (14), (26) получим следующие выражения для определения информационного содержания моделей  $I^{(1)}$  (параметрическая модель) и  $I^{(2)}$  (случайная модель), учитывая, что  $v_1=k-1$ ,  $v_2=N-k$ , а  $q_0=q_1=0,5$ :

$$I^{(1)} = \frac{1}{2} \ln \left\{ 1 + \frac{1 + (1+\lambda)^2 + (1+\lambda) \left[ \frac{k}{k-1}(n-1) - \frac{2}{k-1} \right]}{(N-k-4) \cdot \alpha \cdot (1-\alpha) \left[ F_{sp} \left( \frac{N-k-2}{N-k} \right) - 1 \right]^2} \right\}; \quad (27)$$

$$I^{(2)} = \frac{1}{2} \ln \left\{ 1 + \frac{2(N-3)(N-k-2)^2(k-1) \left[ \frac{2}{\lambda^2} + \frac{2}{\lambda(k-1)} + \frac{1}{(k-2)^2} \right]}{(N-k-4) \cdot (N-k)^2 [\alpha(1-\alpha) + \beta(1-\beta)]} \right\} \quad (28)$$

### Анализ полученных результатов

Анализ выражений (27-28) для оценки количества содержащейся информации позволяет сравнить две базовых статистические модели для обозначения эффектов уровня влияющего фактора при дисперсионном анализе – параметрическую и случайную. Особенно интересны случаи типичные для медико-биологического эксперимента, а именно сравнительно небольшое количество общих и внутригрупповых измерений.

В работе был проведен расчет ожидаемого количества информации для наиболее типичного случая (количество групп  $k = 3$ : контрольная, под воздействием фактора, «плацебо»; вероятность ошибки  $\alpha = 0,05$ ). На рисунке 1 представлена зависимость количества информации  $I$  от количества измерений  $n$  в каждой группе при двух значениях параметра нецентральности  $\lambda=1$  и  $\lambda=5$ .

Из рисунка 1 следует, что увеличение числа измерений в группах уменьшает информационное

содержание модели, поскольку уменьшается их априорная энтропия.

Однако, при прочих равных условиях, случайная модель имеет более высокое информационное содержание, так как априори учитывает наличие вероятностных свойств в уровнях влияющего фактора.

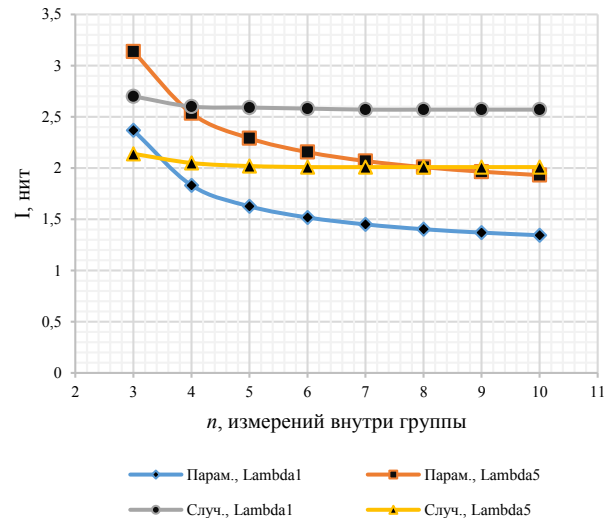


Рис. 1 – Результаты расчета количества ожидаемой информации для трех ( $k = 3$ ) групп измерений при использовании параметрической и случайной моделей

Увеличив количество групп измерений до пяти ( $k = 5$ ) можно наблюдать существенное сближение характеристик, описывающих информационное содержание моделей при значительных значениях параметра нецентральности (см. рис. 2).

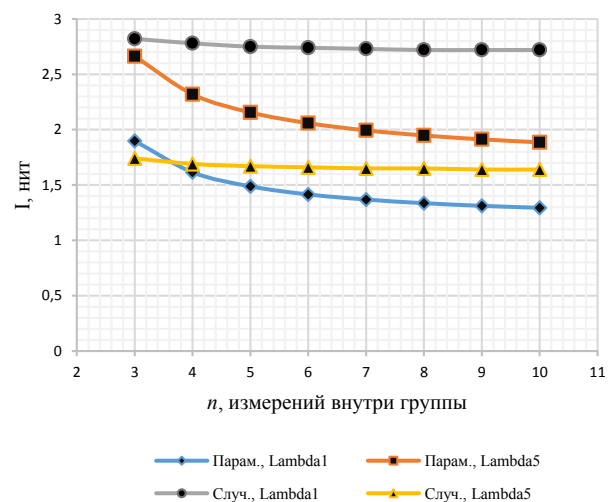


Рис. 2 – Результаты расчета количества ожидаемой информации для пяти ( $k = 5$ ) групп измерений при использовании параметрической и случайной моделей

Выражения (27-28) позволяют исследовать влияние количества групп измерения на информационное содержание моделей (см. рис. 3) при условии равенства общего количества измерений для каждого эксперимента  $N = n \cdot k$ . Исследования проводились для двух значений уровня значимости ( $\alpha=0,05$  и  $\alpha=0,01$ ).

Анализ результатов, представленных на рисунке 3 показывает положительную зависимость информационного содержания моделей от количества групп измерений, при этом модель компонент дисперсий более чувствительна к изменению этого параметра плана эксперимента, чем параметрическая.

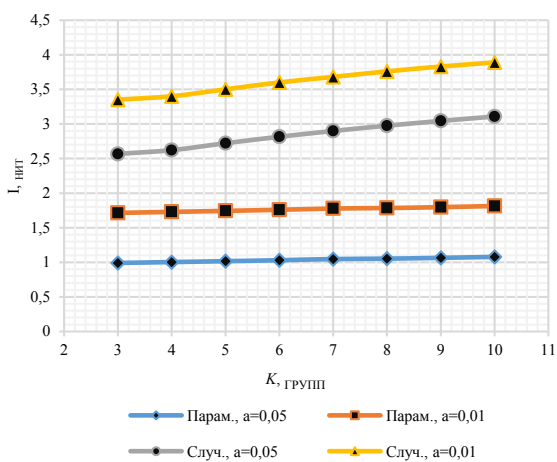


Рис. 3 – Результаты исследования влияния количества групп измерений k на информационное содержание моделей

Также в работе было проведено влияние параметра нецентральности  $\lambda$  на информационное содержание моделей ( $N = 50$ ;  $k = 5$ ;  $\alpha = 0.05$ ), результаты которого представлены на рис. 4.

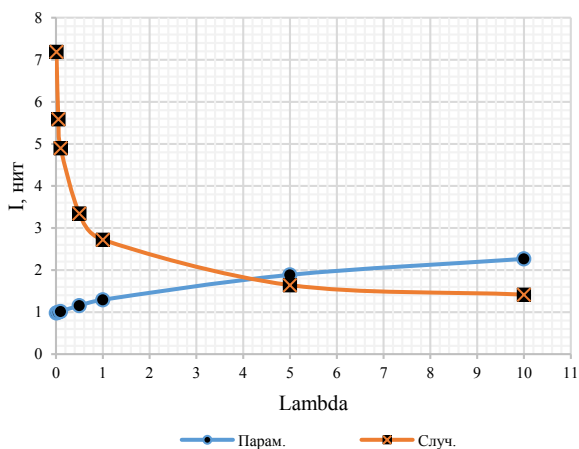


Рис. 4 – Результаты исследования влияния параметра нецентральности  $\lambda$  на информационное содержание моделей

Анализ полученных результатов показывает, что параметр нецентральности  $\lambda$  существенно влияет на информационное содержание статистических моделей для обозначения эффектов уровня влияющего фактора, причем для каждой комбинации параметров плана эксперимента ( $N, k, \alpha$ ) может выбрана та или иная модель, исходя из критерия информативности.

### Выводы

В результате работы проведен качественный и количественный анализ информационного содержания двух базовых (для результатов медико-биологических экспериментов) статистических моделей однофакторного дисперсионного анализа (ANOVA) – параметрической и случайной (модель компонент дисперсий).

Полученные аналитические выражения позволяют оценить влияние параметров плана эксперимента, а именно количества групп, объема выборок, уровня значимости, на информационное содержание моделей, а, следовательно, использовать этот параметр как критерий выбора той или иной модели.

### Список литературы

1. Tsanas, A. A methodology for the analysis of medical data / A. Tsanas, M. A. Little, P. E. McSharry. – Handbook of Systems and Complexity in Health, Springer, New York, 2013. – pp. 113-125.
2. Gandhi, P. Clinical Research Methodology / P. Gandhi B. S. Patel – Pharmacy College, Indian Journal of Pharmaceutical Research and Education, Mehsana, India, 2011. – pp. 199-209.
3. Health Research Methodology: A guide for training in research methods. / WHO Regional Office for the Western Pacific. – Second Edition. – Manila, 2001 – 246 p.
4. Royal, P. W. Dishonesty in Medical Research/ Peter Wilms Hurst Royal // Article in The Medico-legal journal. – 2007. – 75(Part 1). – pp. 3-12.
5. Amar-Sing, H. S. S. The medical research handbook. Planning a Research project. / H. S. S. Amar-Sing, A. Abu Bakar, S. Sararaks. – Kuala Lumpur, 2008.
6. Rockville, R. Methods Guide for Medical Test Reviews. / R. Rockville. – Agency for Healthcare Research and Quality. – 2010. – 168 p.
7. Taroni, F. Statistical hypothesis testing and common misinterpretations: Should we abandon p-value in forensic science applications? / F. Taroni, A. Biedermann, S. Bozza. // Forensic Science International. – 2016. – 259. – pp. 32-36.
8. Yusuff, H. Breast Cancer Analysis Using Logistic Regression / H. Yusuff, N. Mohamad, U. K. Ngah, A. S. Yahaya. – IJRRAS. – 2012. – 10 (1). – pp. 14-22.
9. Guy, B. Faguet. Regression Analysis in Medical Research / Guy B. Faguet, Harry C. Davis // Southern Medical Journal. – 1984. – pp. 722-729.
10. Мун, С. А. Регрессионный анализ в медико-биологических исследованиях: методические

- рекомендации / С. А. Мун, А. Н. Глушков, Т. А. Штернис, С. А. Ларин, С. А. Максимов. - ГБОУ ВПО КемГМА Минздрав-соцразвития России. - Кемерово: КемГМА, 2012. - 115 с.
11. **Перова, И. Г.** Нечеткая классификация данных медико-биологических исследований в условиях дефицита информации / **И. Г. Перова, Е. В. Бодянский** // *Харьковский национальный университет радиоэлектроники*. - Харьков, 2015. -11 (136). - с. 161-163.
  12. **Hopkins, W. G.** Progressive Statistics for Studies in Sports Medicine and Exercise Science / **W. G. Hopkins, S. W. Marshall, A. M. Batterham, J. Hanin** // *Med. Sci. Sports Exerc.* - 2009, 41(1). - pp. 3-13. - doi:10.1249/MSS.0b013e31818cb278.
  13. **Джонсон, Н.** Статистика и планирование эксперимента в технике и науке: Методы планирования эксперимента / **Н. Джонсон, Ф. Лион**; пер. с англ. под ред. Э.К. Лецкого. - М.: Мир, 1981. - 520 с.
  14. **Королюк, В. С.** Справочник по теории вероятностей и математической статистике / **В. С. Королюк, Н. И. Портенко**; под ред. **В. С. Королюка**. - К.: Наукова думка, 1978. - 584 с.
  15. **Орнатский, П. П.** Теоретические основы информационно-измерительной техники / **П. П. Орнатский** - К.: Вища школа, 1983. - 455 с.
  16. **Кисиль, И. С.** Метрологія, точність і надійність засобів вимірювань: [навч. посібник] / **И. С. Кисиль** - Івано-Франковськ: Видавництво „Факел”, 2000. - 400 с.
  17. **Щапов, П. Ф.** Повышение достоверности контроля и диагностики объектов в условиях неопределённости: монография / **П. Ф. Щапов, О. Г. Аврунин**. - Х.: НАДУ, 2011. - 191 с.
  18. **Маркова, Е. В.** Планирование эксперимента в условиях неоднородностей / **Е. В. Маркова, А. В. Лисенков**. - М.: Наука, 1973. - 220 с.
  19. **Patnaik, P. B.** The noncentral  $\chi^2$  and F-distributions and their approximations / **P. B. Patnaik** // *Biometrika*. - 1949. - vol. 36. - №2. - P. 202-232.
  20. **Вентцель, Е. С.** Теория вероятностей и её инженерные приложения: [учеб. пособие для вузов] / **Е. С. Вентцель, Л. А. Овчаров** - М.: Высш. шк., 2000. - 480 с.
  3. **Health Research Methodology: A guide for training in research methods.** WHO Regional Office for the Western Pacific. - Second Edition, Manila, 2001, 246 p.
  4. **Royal, P. W.** Dishonesty in Medical Research. *Article in The Medico-legal journal*, 2007, 75(Part 1), pp. 3-12.
  5. **Amar-Sing, H. S. S., Azman Abu, Bakar, Sondi, Sararaks** The medical research handbook. Planning a Research project. Kuala Lumpur, 2008.
  6. **Rockville, R.** Methods Guide for Medical Test Reviews. Agency for Healthcare Research and Quality, 2010, 168 p.
  7. **Taroni, F., A. Biedermann, S. Bozza.** Statistical hypothesis testing and common misinterpretations: Should we abandon p-value in forensic science applications? *Forensic Science International*, 2016, 259, 32-36.
  8. **Yusuff, H., Mohamad, N., Ngah, U. K., Yahaya, A. S.** Breast Cancer Analysis Using Logistic Regression. *IJRRAS* 10 (1), 2012, 14-22.
  9. **Guy, B. Faguet, Harry, C. Davis.** Regression Analysis in Medical Research. *Southern Medical Journal*, 1984, 722-729.
  10. **Mun, S. A., Glushkov, A. N., Shternis, T. A., Larin, S. A., Maksimov, S. A.** Regressionnyiy analiz v mediko-biologicheskikh issledovaniyah: metodicheskie rekomendatsii. GBOU VPO KemGMA Minzdravsotsrazvitiya Rossii. - Kemerovo: KemGMA, 2012. - 115 s.
  11. **Perova, I. G., Bodyanskiy, E. V.** Nечetkaya klassifikatsiya dannykh mediko-biologicheskikh issledovaniy v usloviyakh defitsita informatsii. *Harkovskiy natsionalniy universitet radioelektroniki*, 2015, 11 (136), 161-163.
  12. **Hopkins, W. G., Marshall, S. W., Batterham, A. M., Hanin, J.** Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Med. Sci. Sports Exerc.* 2009, 41(1), 3-13. - doi:10.1249/MSS.0b013e31818cb278.
  13. **Dzhonson, N., Lion, F.** Statistika i planirovanie eksperimenta v tehnikе i nauke: Metodyi planirovaniya eksperimenta. per. s angl. pod red. E.K. Letskogo. - M.: Mir, 1981. - 520 s.
  14. **Korolyuk, V. S., Portenko, N. I.; pod red. V. S. Korolyuka** Spravochnik po teorii veroyatnostey i matematicheskoy statistike. - K.: Naukova dumka, 1978. - 584 s.
  15. **Ornatskiy, P. P.** Teoreticheskie osnovy informatsionno-izmeritelnoy tehniky / K.: Vischa shkola, 1983. - 455 s.
  16. **Kisil, I. S.** Metrologiya, tochnist I nadlynst zasoblv vimiryuvan: [navch. poslbnyk] / Ivano-Frankovsk: Vidavnistvo „Fakel”, 2000. - 400 s.
  17. **Schapov, P. F., Avrunin, O. G.** Povyishenie dostovernosti kontrolya i diagnostiki ob'ektov v usloviyakh neopredelYonnosti: monografiya, H.: HNADU, 2011, 191 s.
  18. **Markova, E. V., Lisenkov, A. V.** Planirovanie eksperimenta v usloviyakh neodnorodnostey. M.: Nauka, 1973. - 220 s.
  19. **Patnaik, P. B.** The noncentral  $\chi^2$  and F-distributions and their approximations. *Biometrika*, 1949, 36, 2, 202-232.
  20. **Venttsel, E. S., Ovcharov, L. A.** Teoriya veroyatnostey i eYo inzhenernye prilozheniya: [ucheb. posobie dlya vuzov]. M.: Vyissh. shk., 2000. - 480 s.
- Bibliography (transliterated)**
1. **Tsanias, A., Little, M. A., McSharry, P. E.** A methodology for the analysis of medical data. - Handbook of Systems and Complexity in Health, Springer, New York, 2013, 113-125.
  2. **Gandhi, P., Patel, B. S.** Clinical Research Methodology. Pharmacy College, Indian Journal of Pharmaceutical Research and Education, Mehsana, India, 2011, 199-209.

**Сведения об авторах (About authors)**

**Томашевский Роман Сергеевич** – кандидат технических наук, доцент, Национальный технический университет «Харьковский политехнический институт», доцент кафедры промышленной и биомедицинской электроники; г. Харьков, Украина; e-mail: romiocat.khpi@gmail.com.

**Roman Tomashevsky** – candidate of technical sciences, associate professor, National Technical University "Kharkov Polytechnic Institute", associate professor of the Department of Industrial and Biomedical Electronics; Kharkov, Ukraine; e-mail: romiocat.khpi@gmail.com .

*Пожалуйста, ссылайтесь на эту статью следующим образом:*

**Томашевский, Р. С.** Информационный анализ моделей планирования медико-биологического эксперимента при контроле дискретных изменений уровней влияющего фактора / **Р. С. Томашевский** // *Вестник НТУ «ХПИ»*, Серия: *Новые решения в современных технологиях*. – Харьков: НТУ «ХПИ». – 2017. – № 53 (1274). – С. 45-52. – doi:10.20998/2413-4295.2017.53.07.

*Please cite this article as:*

**Tomashevsky, R. S.** Information analysis of planning models of medical-biological experiment at control of discrete changes in the levels of the influencing factor. *Bulletin of NTU "KhPI". Series: New solutions in modern technologies*. – Kharkiv: NTU "KhPI", 2017, **53** (1274), 45–52, doi:10.20998/2413-4295.2017.53.07.

*Будь ласка, посилайтесь на цю статтю наступним чином:*

**Томашевський, Р. С.** Інформаційний аналіз моделей планування медико-біологічного експерименту при контролі дискретних змін рівнів впливаючого фактора / **Р. С. Томашевський** // *Вісник НТУ «ХПІ»*, Серія: *Нові рішення в сучасних технологіях*. – Харків: НТУ «ХПІ». – 2017. – № 53 (1274). – С. 45-52. – doi:10.20998/2413-4295.2017.53.07.

**АНОТАЦІЯ** У статті розглянуті дві базові статистичні моделі для позначення ефектів рівня впливаючого фактора, а саме параметрична модель і модель компонент дисперсій. Проведено порівняльний аналіз ефективності двох моделей однофакторного медичного експерименту з односторонньою класифікацією результатів по одержуваній інформації, яка враховує апріорну невизначеність моделі. Отримано аналітичні вирази для визначення інформаційного змісту розглянутих моделей виходячи з параметрів плану медичного експерименту. Проведено аналіз розглянутих статистичних моделей в результаті якого визначено ефекти впливу кожного з параметрів на ефективність використання моделі та наведено відповідні рекомендації по їх використанню.

**Ключові слова:** дисперсійний аналіз; біомедичний експеримент; параметрична модель; модель компонент дисперсій; параметр нецентральності; медичні вимірювання; критеріальна статистика.

*Поступила (received) 31.11.2017*