

Impact Factor:

ISRA (India) = 1.344	SIS (USA) = 0.912	ICV (Poland) = 6.630
ISI (Dubai, UAE) = 0.829	PIHII (Russia) = 0.207	PIF (India) = 1.940
GIF (Australia) = 0.564	ESJI (KZ) = 4.102	IBI (India) = 4.260
JIF = 1.500	SJIF (Morocco) = 2.031	

SOI: [1.1/TAS](#) DOI: [10.15863/TAS](#)

International Scientific Journal Theoretical & Applied Science

p-ISSN: 2308-4944 (print) e-ISSN: 2409-0085 (online)

Year: 2018 Issue: 05 Volume: 61

Published: 30.05.2018 <http://T-Science.org>

Andrey Leonidovich Gusev
Doctor of Technical Sciences
professor of Perm State
National Research University
(PSNRU)
guseval@mail.ru

Alexander Anatolevich Okunev
Post-graduate student of the Perm State
National Research University
(PSNRU)
alexander2510@mail.ru

SECTION 2. Applied Mathematics.
Math modeling.

DEFINITION OF THRESHOLD FOR INDICATORS

Abstract: The principal algorithm for determining thresholds for indicators is considered in the article. On the example of morbidity indicators, the "simplicity" and "complexity" of the algorithm for determining thresholds are shown. Some generalizations have been made for such algorithms. It is shown how the thresholds of indicators can be used in management tasks using the example of risk management for public health. The prospects of using thresholds in management are discussed.

Key words: management, health risk, indicator threshold, cause-effect relationship, choice marker.

Language: Russian

Citation: Gusev AL, Okunev AA (2018) DEFINITION OF THRESHOLD FOR INDICATORS. ISJ Theoretical & Applied Science, 05 (61): 67-71.

Soi: <http://s-o-i.org/1.1/TAS-05-61-14> **Doi:**  <https://dx.doi.org/10.15863/TAS.2018.05.61.14>

ОПРЕДЕЛЕНИЕ ПОРОГОВ ДЛЯ ПОКАЗАТЕЛЕЙ

Аннотация: В статье рассматривается принципиальный алгоритм определения порогов для показателей. На примере показателей заболеваемости показана «простота» и «сложность» алгоритма определения порогов. Сделаны некоторые обобщения для такого рода алгоритмов. Показано как пороги показателей могут быть использованы в задачах управления на примере управления рисками для здоровья населения. Обсуждаются перспективы использования порогов в управлении.

Ключевые слова: управление, риск для здоровья, порог показателя, причинно-следственная связь, маркер выбора.

Introduction

«Отец» кибернетики Н.Винер [1] положил начало постановке и решению задач управления. К типичным задачам управления относятся такие задачи как установление причинно-следственных связей между показателями (группами показателей), построение моделей управления, на основе установленных причинно-следственных связей, ситуационное моделирование на основе построенных моделей управления; оптимизационное моделирование на области определения управляющих факторов; прогнозирование управляемых факторов в предположении свершения управляющих действий. Авторы настоящей статьи с различных точек зрения рассматривали такие задачи [2-7]. Управление рисками здоровью населения, например, происходит на фоне совокупности неуправляемых и незарегистрированных факторов [8, 9]. Задача управления осложняется неполнотой статистических данных,

несопряженностью показателей (например, некоторые показатели замеряются непосредственно на территориях, другие же на территориальных округах), частотой изменения отчетных статистических форм.

В этих условиях важным моментом является определение порогов показателей, то есть тех конкретных значений показателей, начиная с которых, показатели существенно влияют на модель управления.

Materials and Methods

В настоящей статье авторами предложен алгоритм определения порогов на примере определения порогов для массовых неинфекционных заболеваемости. Эти пороги, как правило, используются для планирования надзорных мероприятий на территориях субъектов. Заболеваемость на территории по отдельным классам болезней и нозологическим формам выше этих порогов идентифицируется



Impact Factor:

ISRA (India) = 1.344	SIS (USA) = 0.912	ICV (Poland) = 6.630
ISI (Dubai, UAE) = 0.829	ПИИЦ (Russia) = 0.207	PIF (India) = 1.940
GIF (Australia) = 0.564	ESJI (KZ) = 4.102	IBI (India) = 4.260
JIF = 1.500	SJIF (Morocco) = 2.031	

как массовая. Это обстоятельство является определяющим при создании модели управления рисками здоровью населения.

На основе порогов массовой неинфекционной заболеваемости можно произвести выделение территорий для установления причин, выявления условий возникновения и распространения массовых неинфекционных заболеваний, которые осуществляется на основании анализа причинно-следственной связи показателей (факторов) среды обитания и массовой заболеваемости на этих территориях.

Далее возможно определение приоритетных объектов надзора при планировании надзорных мероприятий и управляющих действий на основании оценки опасности для здоровья источников загрязнения среды обитания на выбранных территориях, то есть для управления рисками здоровью населения.

Общий алгоритм управления и подготовка исходной информации

Расчет порогов индикативных показателей покажем на примере показателей (факторов), которые участвуют в моделях управления показателями здоровья населения. Сначала введем обозначения.

G - расчетный год.

K - количество лет перед расчетным годом, используемое для расчета порога массовой неинфекционной заболеваемости.

N - количество территорий субъекта.

$PЗ_i^j$ - показатель заболеваемости по исследуемому классу болезни или нозологической форме на 1 тысячу человек (промилле) по определенной возрастной группе (дети, подростки, взрослые, все или трудоспособные) по i -ой ($i = \overline{1, N}$) территории субъекта за j -ый год ($j = \overline{G-K, G}$).

P - порог заболеваемости.

Общий алгоритм управления, включающий определение порогов массовой неинфекционной заболеваемости и их использования в планировании надзорных мероприятий, осуществляются по следующей схеме:

1. Подготовка исходной информации.
2. Расчет порогов массовой неинфекционной заболеваемости для классов болезней и нозологических форм.
3. Идентификация территорий с массовой неинфекционной заболеваемостью.
4. Оценка опасности воздействия факторов среды обитания на

формирование классов болезней и нозологических форм с выделением территорий с наибольшей опасностью.

5. Выделение объектов надзора на территориях с массовой неинфекционной заболеваемостью для планирования приоритетных надзорных действий.

Подготовка исходной информации включает в себя следующую последовательность предварительных действий:

1. Выбор для исследования нозологической формы или класса болезней по данным социально-гигиенического мониторинга, которые формируются по формам статистической отчетности.
2. Определение возрастной группы: дети, подростки, взрослые или все в соответствии с поставленной задачей.
3. Определение расчетного года в соответствии с поставленной задачей.
4. Составление структурной таблицы. Каждая строка таблицы включает порядковый номер, номер территории субъекта и показатели заболеваемости за $G-K, G-K+1, \dots, G$ годы. Всего строк N по количеству территорий субъекта.

Алгоритм расчета порогов.

С использованием подготовленной структурной таблицы нужно рассчитать среднее арифметическое выборочное показателя заболеваемости для совокупности территорий составляющих субъект

$$\overline{PЗ} = \frac{\sum_{i=1}^N \sum_{j=G-K}^{G-1} PЗ_i^j}{K \cdot N}, \quad (1)$$

и среднее выборочное квадратичное отклонение

$$S = \sqrt{\frac{1}{K \cdot N - 1} \sum_{i=1}^N \sum_{j=G-K}^{G-1} (PЗ_i^j - \overline{PЗ})^2}. \quad (2)$$

Далее необходимо присвоить территориям маркеры выбора по правилу:

$$R_i^j = \begin{cases} 0, & \text{если } PЗ_i^j \leq \overline{PЗ} + mS \\ 1, & \text{если } PЗ_i^j > \overline{PЗ} + mS \end{cases}, \quad (3)$$

где m - некоторое действительное число. Из дальнейшего расчета порога массовой неинфекционной заболеваемости исключить значения показателя заболеваемости с маркером выбора «1». Подсчитать k_0 - количество

Impact Factor:

ISRA (India) = 1.344	SIS (USA) = 0.912	ICV (Poland) = 6.630
ISI (Dubai, UAE) = 0.829	ПИИЦ (Russia) = 0.207	PIF (India) = 1.940
GIF (Australia) = 0.564	ESJI (KZ) = 4.102	IBI (India) = 4.260
JIF = 1.500	SJIF (Morocco) = 2.031	

территорий за K лет перед расчетным годом с маркерами выбора равными «0».

Для расчета пороговых показателей необходимо вычислить $\overline{ПЗ}_0$ и S_0 аналогично формулам (1) и (2), за той разницей, что в вычисления берутся только те $ПЗ_i^j$, для которых маркер выбора с аналогичными индексами равен «0» ($R_i^j=0$). В этом случае, $\overline{ПЗ}_0$ - среднеарифметическое выборочное и S_0 - среднее выборочное квадратичное отклонение, соответственно, равны:

$$\overline{ПЗ}_0 = \frac{\sum_{i=1}^N \sum_{j=G-K}^{G-1} ПЗ_i^j}{k_0}, \quad (4)$$

$$S_0 = \sqrt{\frac{1}{k_0-1} \sum_{i=1}^N \sum_{j=G-K}^{G-1} (ПЗ_i^j - \overline{ПЗ}_0)^2}, \quad (5)$$

где k_0 – количество территорий за K лет перед расчетным годом с маркерами выбора равными «0». В формулах (4) и (5) участвуют $ПЗ_i^j$, для которых $R_i^j=0$.

В качестве порога массовой неинфекционной заболеваемости принимается значение

$$P = \overline{ПЗ}_0 + mmS_0, \quad (6)$$

где mm - действительное число.

Если для территории субъекта заболеваемость ниже порога массовой неинфекционной заболеваемости ($ПЗ_i^G \leq P$), то на территории в расчетном году не фиксируется массовая неинфекционная заболеваемость исследуемого вида и ей присваивается маркер выбора «0». В противном случае на этой территории в расчетном году массовая неинфекционная заболеваемость фиксируется, и ей присваивается маркер выбора «1».

Использование порогов показателей заболеваемости.

Для территорий с маркером выбора равным «1» необходимо установление связи заболеваемости данным классом болезней или нозологической формой с факторами риска среды обитания.

Для оценки возможной связи массовой неинфекционной заболеваемости с негативными факторами среды обитания, обусловленными объектами надзора на первом этапе необходимо осуществить сопоставление территорий, характеризующихся массовой неинфекционной заболеваемостью с показателями опасности развития этого класса болезней или нозологической формы. Для оценки возможной связи массовой неинфекционной заболеваемости с негативными факторами загрязнения

атмосферного воздуха используется индекс сравнительной неканцерогенной опасности, обозначаемый в предметной литературе как HRI.

Аналогично HRI могут быть использованы другие показатели. Например, доля проб воды в источниках централизованного водоснабжения, не отвечающая гигиеническим нормативам по санитарно-химическим показателям или по микробиологическим показателям в процентах. Также могут быть использованы: доля проб воды в водопроводах, не отвечающая гигиеническим нормативам по санитарно-химическим показателям или по микробиологическим показателям в процентах; доля нестандартных проб в процентах продовольствия по санитарно-химическим или по микробиологическим показателям; доля проб атмосферного воздуха, превышающих ПДК в процентах всего; доля проб почвы по санитарно-химическим показателям или по микробиологическим показателям, не отвечающих гигиеническим нормативам в процентах в селитебной зоне; доля рабочих мест, не отвечающих нормативам в процентах по шуму или по вибрации; и другие.

Выбранный показатель опасности статистически обрабатывается аналогично показателю здоровья. Далее территории в расчетном году сравниваются по индексу выбора. На тех территориях, где маркер выбора равен «1» для показателя здоровья и показателя опасности, причинно-следственная связь между показателем опасности и показателем здоровья считается установленной.

Следующим важным шагом является выявление объектов надзора, которые могли послужить источником массовой неинфекционной заболеваемости. В этом случае используются коэффициенты HRI для объектов надзора, с помощью которых ранжируются все объекты по степени риска для здоровья населения. Но это уже отдельная и узкоспециальная тема.

Интерпретация и обсуждение алгоритма.

Алгоритм выглядит крайне простым. При этом возникают вопросы о необходимости двойного расчета средних значений. Ведь для каждой нозологической формы группа экспертов всегда может довольно точно указать территории с массовым неинфекционным заболеванием. Следовательно, можно легко подобрать значение коэффициента m в формуле (3) такое, что без дополнительных расчетов по формуле (6) можно будет выделить все территории с массовым неинфекционным заболеванием. Однако такое можно проделать только для конкретной нозологической формы и только для

Impact Factor:

ISRA (India) = 1.344	SIS (USA) = 0.912	ICV (Poland) = 6.630
ISI (Dubai, UAE) = 0.829	ПИИЦ (Russia) = 0.207	PIF (India) = 1.940
GIF (Australia) = 0.564	ESJI (KZ) = 4.102	IBI (India) = 4.260
JIF = 1.500	SJIF (Morocco) = 2.031	

определенного субъекта. Понятно, что нозологические формы, принадлежащие разным классам заболеваний, будут серьёзно отличаться друг от друга по определению «аномальных» территорий даже в одном и том же субъекте. К тому же одна и та же нозологическая форма может себя по-разному проявлять в разных субъектах. Субъекты могут серьёзно отличаться друг от друга климатическими условиями и хозяйственной структурой (род промышленных предприятий на территории, доля аграрного сектора на территории и так далее), что, несомненно, влияет на интенсивность присутствия той или иной нозологической формы на той или иной территории. Все эти размышления наталкивают на мысль о создании более универсального алгоритма, коим, по сути, и является вышеописанный алгоритм.

Однако в алгоритме не определены следующие значения: K – количество лет, за которые нужно собрать статистическую информацию, значения коэффициентов m и mm в формулах (3) и (6) при множителях S и S_0 .

Здесь приходится в каждой конкретной предметной области математикам вместе со специалистами в этой предметной области, руководствуясь здравым смыслом, интуицией и знаниями математической статистики, решать конкретную задачу по нахождению значений K , m и mm . Нахождение конкретных значений коэффициентов продемонстрируем на задачах управления рисками здоровью населения.

Хорошо известно, что вся статистическая информация по показателям качества среды обитания и показателям здоровья фиксируется один раз в год по субъектам, которые состоят из 15-40 территорий. Первым вопросом, который нужно было решить, был вопрос о количестве лет, которые нужно исследовать перед расчетным годом. Путем всесторонних исследований было рекомендовано 4 года. Дело всё в том, что увеличение количества лет почти всегда гарантирует изменение параметров распределения показателей, как случайных величин. Это, прежде всего, связано с изменением структуры промышленных предприятий на территориях и изменением уровня медицинского обслуживания. Уменьшение количества лет существенно сокращает объем статистических данных, который и так не велик при 4 годах. В любой другой предметной области рассуждения про количество исследуемых временных периодов (год, месяц, неделя, день, час и так далее) должен иметь такую же смысловую подоплеку.

Экспериментальным путём по статистическим данным одного из субъектов,

состоящего из 53 территорий, были подобраны и установлены значения коэффициентов m и mm . По экспертным оценкам было установлено, что лучше всего для алгоритма использовать $m = mm = 2$. При этом было рассмотрено порядка 25-30 различных сочетаний m и mm . Каждое сочетание оценивалось экспертами по 25 показателям качества среды обитания и 96 показателям здоровья (все классы заболеваний и некоторые нозологические формы, которые эксперты определили как актуальные). Задача подбора коэффициентов состояла в том, чтобы в среднем алгоритм «ошибался» не более чем на 5% и, безусловно, «ошибался» как можно меньше. Настоящий алгоритм со значениями $m = mm = 2$ на вышеупомянутых данных имел ошибку эквивалентную 1%. Этот алгоритм был реализован в программных продуктах [10, 11] для решения двух разных задач управления рисками здоровью населения.

Затем алгоритм был проверен на двух других субъектах и на нескольких субъектах одновременно (то есть субъекты выступили в качестве территорий, составляющих единое целое). Алгоритм неизменно показывал ошибку не более одного процента.

Заключение

Отметим некоторые обобщения алгоритма определения порогов у произвольной случайной величины. В алгоритме определения порогов заболеваемости на первом этапе с помощью формулы (3) исключались аномально большие значения случайной величины, а на втором этапе с помощью формулы (6) устанавливался порог, с помощью которого определялись значения случайной величины, которые объявлялись недопустимыми (территории с массовым неинфекционным заболеванием). Таким образом, в конечном итоге выделяли значения случайной величины, которые являются предметом управления.

В других предметных областях при управлении случайные величины, очевидно, могут иметь принципиально другой смысл. В связи с этим, на первом этапе с помощью формулы (3) можно исключать аномально малые значения случайной величины. На втором этапе с помощью формулы (6) можно устанавливать порог, с помощью которого определять значения случайной величины, которые далее будут объявлены недопустимыми и будут являться предметом будущего управления. Принципиально важно в алгоритме, что он состоит из двух этапов. С помощью положительных и отрицательных значений



Impact Factor:

ISRA (India) = 1.344	SIS (USA) = 0.912	ICV (Poland) = 6.630
ISI (Dubai, UAE) = 0.829	ПИИЦ (Russia) = 0.207	PIF (India) = 1.940
GIF (Australia) = 0.564	ESJI (KZ) = 4.102	IBI (India) = 4.260
JIF = 1.500	SJIF (Morocco) = 2.031	

коэффициентов m и mm можно на первом этапе устанавливать аномально большие или аномально малые значения случайной величины, а на втором этапе устанавливать недопустимые значения случайной величины с точки зрения модели управления, как малые, так и большие.

Наверняка, существуют случайные величины, для которых на первом этапе алгоритма необходимо исключались аномально большие и аномально малые значения случайной величины одновременно. Тогда формула (3) примет следующий вид:

$$R_i^j = \begin{cases} 0, & \text{если } \bar{PZ} - m_1 S \leq PZ_i^j \leq \bar{PZ} + m_2 S, \\ 1, & \text{в противном случае} \end{cases}, (7)$$

где m_1 и m_2 некоторые действительные коэффициенты. И в этом случае, принципиально то, что алгоритм состоит из двух этапов.

Таким образом, при применении алгоритма нахождения порога случайной величины главным при его использовании является последовательность двух этапов. Сами этапы могут отличаться друг от друга наполнением в зависимости от поставленной задачи.

References:

1. Viner N. (1983) Kibernetika, ili upravlenie i svyaz v zhivotnom i mashine. Moscow: Nauka, 344.
2. Gusev A.L. (2011). Upravlenie i procedury gruppovykh proverok. Saarbrucken (Germany): LAP LAMBERT Academic Publishing GmbH & Co. KG, 75.
3. Gusev A.L. (2011). Metod otbora informacionnogo prostranstva dlya postoeniya nejronnoj seti, kak modeli upravleniya, v usloviyah zashumlennykh i nepolnykh dannykh. No 7. pp. 55-57.
4. Gusev A.L. (2012) Nepreryvnyj statisticheskij kontrol pri upravlenii. Los Angeles (USA): Createspace, 128.
6. Gusev A.L. (2012) Continuous Inspection with Memory. Statistics & Probability Letters. Vol. 82. pp. 303-307.
7. Gusev A.L. (2012) The optimal number of items in a group for group testing. Statistics & Probability Letters. Vol. 82. pp. 2083-2085.
8. Gusev A.L., Okunev A.A. (2017) Forecasting with incomplete set of factors determining the predicted factor. Neural network error extrapolation method. International Journal of Applied Mathematics and Statistics. Vol.56, №5, p.48-52.
9. Zaitseva N.V., Gusev A.L., Shur P.Z. (2010) Sovershenstvovanie metodicheskikh podhodov k planirovaniyu deyatelnosti organov i organizacij Rospotrebnadzora v ramkah regionalnykh vedomstvennykh celevykh program. Zdorovje naseleniya i sreda obitaniya. No 1, Vol. 214. pp. 4-7.
10. Zaitseva N.V., Gusev A.L., Shur P.Z., Babushkina E.V. (2010) Metodicheskie podhody k opredeleniyu vklada organov i organizacij Rospotrebnadzora v upravlenie riskom zdorovyu naseleniya. Zdorove naseleniya i sreda obitaniya. No. 11, Vol. 212. pp. 11-13.
11. Gusev A.L., Shur P.Z. (2013). Raschet porogov dlya pokazateley zdorovya i ih ispolzovaniye pro klassifikacii territorij s massovoj neinfekcionnoj zaboлеваemostyu. Programma dlya EVM №2013618581.
12. Gusev A.L., Shur P.Z. (2013). Raschet celevykh pokazateley neposredstvennogo i konechnogo rezultatov dlya upravlyayushey organizacii na terretoriyah s massovymi neinfekcionnymi zabolevaniyami. Programma dlya EVM № 2013618580.

