

На правах рукописи

Кареев Искандер Амирович

**Нижние границы для  
среднего объёма наблюдений в  
процедурах отбора и упорядочивания**

Специальность 01.01.05 — теория вероятностей и  
математическая статистика

АВТОРЕФЕРАТ

диссертации на соискание учёной степени  
кандидата физико-математических наук

Москва — 2014

Работа выполнена на кафедре математической статистики Института вычислительной математики и информационных технологий Казанского (Приволжского) федерального университета.

Научный руководитель: доктор физико-математических наук,  
профессор  
Володин Игорь Николаевич

Официальные оппоненты: доктор физико-математических наук,  
зав. лаб. «Когнитивные методы анализа дан-  
ных и моделирования» ИСА РАН,  
профессор  
Бернштейн Александр Владимирович

кандидат физико-математических наук,  
старший научный сотрудник кафедры мате-  
матической статистики фак. ВМК МГУ,  
старший научный сотрудник  
Пагурова Вера Игнатьевна

Ведущая организация: Вологодский государственный университет

Защита диссертации состоится 10 июня 2014 года в 11 часов 00 минут на заседании диссертационного совета Д 501.001.44 при Московском государственном университете имени М.В. Ломоносова по адресу: *119991, ГСП-1, Москва, Ленинские горы, МГУ, 2-й учебный корпус, факультет ВМК, аудитория 685.*

Желающие присутствовать на заседании диссертационного совета должны сообщить об этом за 2 дня по тел. *(495) 939-30-10* (для оформления заявки на пропуск).

С диссертацией можно ознакомиться в Фундаментальной библиотеке МГУ.

С текстом автореферата и диссертации можно ознакомиться на официальном сайте факультета ВМК МГУ <http://cs.msu.ru/> в разделе «Диссертации».

Автореферат разослан \_\_\_\_\_

Ученый секретарь  
диссертационного совета,  
к. т. н, ведущий научный сотрудник

В. А. Костенко

# Общая характеристика работы

## **Актуальность темы исследования:**

В теории статистических выводов существует класс особых многовыборочных проблем, при решении которых требуется выполнение заданных ограничений на вероятность корректного решения, принятого после проведения наблюдений. Одна из таких проблем – отбор „наилучшей“ популяции или упорядочивание популяций в соответствии с определенным показателем их предпочтения. Естественно, для реализации этого требования необходимо предварительно, до постановки статистического эксперимента, планировать объем испытаний. В связи с этим возникает *актуальная* и важная в практических применениях процедур отбора и упорядочивания задача нахождения минимального (среднего) объема наблюдений, ниже которого процедур с заданной вероятностью корректного решения не существует. Решению этой задачи посвящена представляемая диссертация.

## **Степень разработанности:**

В математической статистике существует ряд неравенств, устанавливающих нижние границы для различных характеристик процедур статистического вывода. Обзор таких границ следует начать с неравенства Рао-Крамера и его обобщения на последовательные процедуры, данные Хёфдингом<sup>1</sup>. Неравенство Хёфдинга, разрешённое относительно среднего объёма наблюдений, позволяет построить нижние границы для среднего объёма выборки, необходимого для несмещённого оценивания с заданными ограничениями на дисперсию. Другое, не менее известное, неравенство Вальда<sup>2</sup> для среднего объёма выборки при проверке гипотез легко обобщается на случай различения сложных гипотез и позволяет строить нижние границы для среднего объёма выборки, необходимого для различения двух односторонних гипотез, разделённых областью безразличия, с заданными ограничениями на вероятность ошибок первого и второго рода.

В последующим Саймонс<sup>3</sup> обобщил границы Вальда на случай различения более, чем двух простых гипотез. Аналог границ Саймонса для случая многовыборочных проблем приводятся в монографии Бекхофера-Кифера-

---

<sup>1</sup> *Hoeffding, W.* A lower bound for the average sample number of a sequential test / W. Hoeffding // The Annals of Mathematical Statistics. 1953. Vol. 24, №1. pp. 127–130.

<sup>2</sup> *Wald, A.* Statistical decision functions / A. Wald. Oxford, England: Wiley. 1950. 179 p.

<sup>3</sup> *Simons, G.* Lower bounds for average sample number of sequential multihypothesis tests / G. Simons // The Annals of Mathematical Statistics. 1967. Vol. 38, №5. pp. 1343–1364.

Собеля<sup>4</sup> и приписываются к неопубликованным на тот момент результатам Хёфдинга. Использование этих границ для построения процедур различения многих простых гипотез с заданными ограничениями на все элементы матрицы ошибок дано в работах Володина<sup>5,6</sup>.

Наконец, Володиным были получены универсальные нижние границы для среднего объёма наблюдений в процедурах статистического вывода<sup>7,8</sup>, справедливые для любой статистической проблемы. Такие границы содержат как частный случай границы Хёфдинга, Вальда и Саймонса. Реализация этих границ для гарантийных процедур статистического вывода связана с решением достаточно сложных задач на экстремум. Решая такие задачи, Володин получил нижние границы для среднего объёма наблюдений в критериях согласия, однородности<sup>9</sup>, инвариантности<sup>10</sup> и независимости. В дальнейшем Малютовым<sup>11</sup> эти границы были распространены на случай управления наблюдениями, с семейством наблюдаемых случайных величин произвольной мощности. Это позволило улучшить нижние границы Володина в проблеме проверки однородности более чем двух распределений<sup>12</sup>, а также построил границы объёма наблюдений в задачах регрессии и планирования статистических экспериментов<sup>13</sup>. Указанные результаты Володина и Малютова содержатся в обзорной статье Володина<sup>14</sup>.

---

<sup>4</sup> *Bechhofer, R.E.* Sequential identification and ranking procedures / R.E. Bechhofer, J. Kiefer, M. Sobel. Chicago: University of Chicago Press. 1968. 420 p.

<sup>5</sup> *Володин, И.Н.* Оценки необходимого объёма наблюдений в задачах статистической классификации. I / И.Н. Володин // М.: ТВП. 1977. 22:2. С. 347–357.

<sup>6</sup> *Володин, И.Н.* Оценки необходимого объёма наблюдений в задачах статистической классификации. II / И.Н. Володин // М.: ТВП. 1977. 22:4. С. 749–765.

<sup>7</sup> *Володин, И.Н.* Нижние границы для среднего объёма выборки и эффективность процедур статистического вывода / И.Н. Володин // М.: ТВП. 1979. т. 24. вып. 1. С. 119–129.

<sup>8</sup> *Володин, И.Н.* Нижние границы для среднего объёма выборки в процедурах с управлением / И.Н. Володин // М.: ТВП. 1981. т. 26. вып. 3. С. 630–631.

<sup>9</sup> *Володин, И.Н.* Нижние границы для среднего объёма выборки в критериях согласия и однородности / И.Н. Володин // М.: ТВП. 1979. т. 24. вып. 3. С. 637–645.

<sup>10</sup> *Володин, И.Н.* Нижние границы для среднего объёма выборки в критериях инвариантности / И.Н. Володин // М.: ТВП. 1980. т. 25. вып. 2. С. 359–364.

<sup>11</sup> *Малютов, М.Б.* Нижние границы для средней длительности последовательно планируемых экспериментов / М.Б. Малютов // Изв. ВУЗов, сер. матем.. 1983. №11. С. 19–41.

<sup>12</sup> *Malyutov, M.B.* One bound for the mean duration on sequential testing homogeneity / L.I. Galtchouk, M.B. Malyutov // MODA4 — Advances in model-oriented data analysis contributions to statistics. 1995. Part 1. pp. 49–56.

<sup>13</sup> *Малютов, М.Б.* Нижние границы для средней длительности последовательно планируемых экспериментов / М.Б. Малютов // Изв. ВУЗов, сер. матем.. 1983. №11. С. 19–41.

<sup>14</sup> *Володин, И.Н.* Нижние границы для среднего объёма наблюдений в гарантийных процедурах статистического вывода / И.Н. Володин // Исследования по прикладной математике и информатике, Казань:

Построение нижних границ для среднего объёма наблюдений в гарантийных процедурах отбора и упорядочивания было очередной задачей в рамках этих исследований. В этом направлении имелаась единственная статья Новикова<sup>15</sup>, в которой получены границы в случае отбора наилучшей нормальной популяции и исследована эффективность процедуры Бекхофера. Цель данной диссертации — дать полное решение задачи нижней оценки среднего числа наблюдений в гарантийных процедурах отбора и упорядочивания.

Проведём небольшой обзор существующих методов отбора и упорядочивания статистических популяций.

Обзор следует начать с основополагающих работ Бекхофера<sup>16</sup> (1954) и Гупты<sup>17</sup> (1956). Суть подхода Бекхофера к построению гарантийных процедур отбора и упорядочивания заключается в введении так называемой зоны безразличия: от таких процедур требуется гарантировать заданный уровень вероятности корректного решения лишь при таких конфигурациях популяций, при которых их значения параметров находятся на достаточном расстоянии друг от друга. Это ограничение на степень близости значений параметров у разных популяций и называют зоной безразличия. Гарантия вероятности корректного отбора достигается посредством выбора требуемого числа наблюдений. Заметим, что эта работа Бекхофера представляет наиболее ранние результаты в области отбора и упорядочивания и впервые формулирует в современном виде проблему отбора и упорядочивания.

Другой подход в построении гарантийных процедур отбора, который не будет рассматриваться в диссертации, — это подход Гупты<sup>18</sup>. Здесь заданная вероятность корректного отбора гарантируется при любом числе наблюдений, и представляет собой вероятность отбора некоторого множества популяций, содержащего наилучшую. Согласно этому подходу, в результате применения процедуры отбора формируется некоторое „доверительное“ множество, которое с заданной вероятностью накрывает наилучшую популяцию; увеличение объёма наблюдений позволяет сузить это множество, увеличивая, таким об-

---

Изд-во Казан. ун-та. 2011. Вып. 27. С. 70–116.

<sup>15</sup> Новиков, А.А. Эффективность процедур отбора / А.А. Новиков // Казань: Исслед. по прикл. матем.. 1984. Т. 11, №2. С. 43–51.

<sup>16</sup> *Bechhofer, R.E.* A single-sample multiple decision procedure for ranking means of normal populations with known variances / R.E. Bechhofer // *The Annals of Mathematical Statistics*. 1954. Vol. 25, №1. pp. 16–39.

<sup>17</sup> *Gupta, S.S.* On a decision rule for a problem in ranking means / S.S. Gupta. University of North Carolina at Chapel Hill. 1956. 208 p.

<sup>18</sup> *Gupta, S.S.* On a decision rule for a problem in ranking means / S.S. Gupta. University of North Carolina at Chapel Hill. 1956. 208 p.

разом, точность отбора.

Задачи отбора и упорядочивания в основном рассматривались только для наиболее распространённых вероятностных моделей наблюдаемых характеристик популяций. Это нормальная модель, при которой популяции распределены согласно нормальному закону с общей известной дисперсией, а целевым неизвестным параметром является среднее. В экспоненциальной модели целевым является масштабный параметр показательного распределения. В биномиальной модели, естественно, отбирается популяция с наибольшей вероятностью успешного исхода, или популяции упорядочиваются по величине этой вероятности. Для пуассоновской модели, очевидно, целевым параметром является интенсивность пуассоновского потока. Наконец, особый случай представляет мультиномиальная модель, в которой каждая популяция соответствует некоторой компоненте мультиномиального случайного вектора, и задачей является выявление компоненты с наибольшей вероятностью успеха.

Все указанные задачи рассматриваются при различных способах введения зоны безразличия. Например, для параметра сдвига нормальной модели чаще всего рассматривается зона безразличия, основанная на разности параметров. С другой стороны, для параметров масштаба или, например, вероятности успеха в биномиальной модели зачастую используется зона безразличия, основанная на отношении значений параметров популяций.

В процедурах отбора и упорядочивания можно фиксировать объём наблюдения заранее или использовать последовательные схемы выбора с управлением обхода популяций. Обычно последовательные процедуры требуют в среднем меньшего суммарного объёма наблюдений, чем процедуры с фиксированным числом наблюдений, и в этом отношении являются более предпочтительными. Заметим, однако, что на практике процедуры с фиксированным числом наблюдений зачастую являются значительно более удобными с организационной точки зрения, что с избытком компенсирует их меньшую эффективность в отношении требуемого объёма наблюдений.

Опишем некоторые процедуры, выполняющие отбора и упорядочивания в рамках различных моделей. Первая процедура, решающая задачу отбора в случае нормальной модели, когда дисперсии популяций известны и равны, была построена Бекхофером<sup>19</sup>. Эта процедура требовала фиксированного числа наблюдений и по окончании эксперимента выбирала в качестве наилуч-

---

<sup>19</sup> *Bechhofer, R.E.* A single-sample multiple decision procedure for ranking means of normal populations with known variances / R.E. Bechhofer // *The Annals of Mathematical Statistics*. 1954. Vol. 25, №1. pp. 16–39.

шей популяцию с наибольшим значением выборочного среднего. Требуемый объём наблюдений в этой процедуре определяется на основании оценки вероятности корректного отбора при наименее благоприятном случае, то-есть когда параметры популяций находятся настолько близко друг от друга, насколько позволяет введённая зона безразличия.

Некоторым усовершенствованием этой процедуры с фиксированным числом наблюдением стала последовательная процедура Бекхофера-Кифера-Собея. В этой процедуре используется довольно простое управление — на каждом шаге эксперимента из каждой популяции берётся по одному наблюдению, после чего решается вопрос о продолжении эксперимента. Такого рода управление часто обозначается термином *vector-at-once* (вектор за раз). По окончании эксперимента в качестве наилучшей выбирается популяция с наибольшим значением выборочного среднего.

Важной положительной особенностью процедур Бекхофера с фиксированным числом наблюдений и последовательных процедур Бекхофера-Кифера-Собея является их универсальность. Этими процедурами могут решаться как задачи отбора, так и задачи упорядочивания для широкого класса моделей популяции. В частности, они применимы ко всем рассматриваемым в данной работе моделям: нормальной, экспоненциальной, биномиальной, пуассоновской, мультиномиальной.

Вероятно, наиболее экономичной процедурой отбора нормальной популяции, с точки зрения числа наблюдений, можно назвать последовательную процедуру Као-Лай<sup>20</sup>. Она во многом схожа с последовательной процедурой Бекхофера-Кифера-Собея, однако в ней представлен механизм раннего экранирования популяций с наименьшими значениями параметра. На первом шаге эксперимента процедура производит по одному наблюдению в каждой популяции. После этого на основании полученных данных выявляются популяции, которые с достаточно малой вероятностью являются наилучшими. Такие процедуры исключаются из дальнейшего рассмотрения. На следующем шаге процедура вновь берёт по наблюдению из каждой популяции, кроме уже исключённых. И так далее. Наконец, эксперимент заканчивается, когда в рассмотрении остаётся лишь одна, последняя популяция, которая и объявляется наилучшей. Заметим, что наибольший выигрыш такого экранирования достигается при сильно различающихся значениях параметров популяций. На-

---

<sup>20</sup> *Kao, S.C.* Sequential selection procedures based on confidence sequences for normal populations / *S.C. Kao, T.L. Lai* // *Communications in Statistics — Theory and Methods.*— 1980. Vol. 9, №16. pp. 1657–1676.

против, как показывают результаты статьи, при наименее благоприятном для отбора случае, когда популяции максимально похожи, асимптотический средний объём наблюдений процедуры Као-Lai оказывается на том же уровне, что и у более простой последовательной процедуры Бекхофера-Кифера-Собеля.

Ряд процедур упорядочивания по параметрам масштаба и сдвига для широкого класса распределений представлены в статьях Скафера и Рутемиллера<sup>21</sup>, Бишопа и Дудовича<sup>22</sup>. Бейрлант, Дудевич и ван дер Меулен<sup>23</sup> предложили двухступенчатую процедуру упорядочивания нормальных популяций по средним значениям при неизвестных дисперсиях и привели примеры её использования на реальных данных. Проблема использования различных функций потерь в задачах упорядочивания обсуждается в статье Собела<sup>24</sup>.

Отметим процедуру отбора биномиальной популяции, предложенную Бекхофером и Кулкарни<sup>25</sup>. Эта последовательная процедура основана на принципе, схожем с выбором по последнему успеху (*play-the-winner*). В их работе было показано, что эта процедура обеспечивает не меньшую вероятность корректного успеха, чем процедура отбора с фиксированным числом наблюдений, требуя при этом меньшее число наблюдений.

Мулекар и Матежик<sup>26</sup> предложили процедуру отбора популяции с наименьшим средним среди пуассоновских популяций с фиксированным числом наблюдений. Для построения такой процедуры им понадобилось рассмотреть зону безразличия, контролирующую как разность между параметрами популяции, так и отношение между ними. Мулекар и Матежик определили точное выражение для вероятности корректного решения при наименее благоприятном для отбора случае, на использовании которого и построен выбор

---

<sup>21</sup> *Schafer, R.E.* Some characteristics of a ranking procedure for population parameters based on chi-square statistics / R.E. Schafer, H.C. Rutenmiller // *Technometrics*. 1975. Vol. 17, №3. pp. 327–331.

<sup>22</sup> *Dudewicz, E.J.* Complete ranking of reliability-related distributions / E.J. Dudewicz, T.A. Bishop // *IEEE Transactions on Reliability*. 1977. Vol. R-26, №5. pp. 362–365. (подробное изложение см. в Stanford university, Technical report No. 114. 1976)

<sup>23</sup> *Dudewicz, E.J.* Complete Statistical ranking of populations, with tables and applications / E.J. Dudewicz, J. Beirlant, E.C. van der Meulen // *Journal of Computational and Applied Mathematics*. 1982. Vol. 8, №3. pp. 187–201.

<sup>24</sup> *Sobel, M.J.* Complete ranking procedures with appropriate loss functions / M.J. Sobel // *Communications in Statistics - Theory and Methods*. 1990. Vol. 19, №12. pp. 4525–4544.

<sup>25</sup> *Bechhofer, R.E.* On the performance characteristics of a closed adaptive sequential procedure for selecting the best bernoulli population / R.E. Bechhofer, R.V. Kulkarni // *Communications in Statistics. Part C: Sequential Analysis: Design Methods and Applications*. 1983. Vol. 1, №4. pp. 315–354.

<sup>26</sup> *Mulekar, M.S.* Determination of sample size for selecting the smallest of k possible population means / M.S. Mulekar, F.J. Matejck // *Communications in Statistics — Simulation and Computation*. 2000. Vol. 29, №1. pp. 37–48.

объёма наблюдения в процедуре<sup>27</sup>. Кроме того, Мулекар и Собэл построили аналогичную процедуру для отбора пуассоновской популяции с наибольшим средним<sup>28</sup>.

Для решения задачи отбора в мультиномиальной модели, Бекхофер, Элмаграби и Морсе<sup>29</sup> была предложена процедура с фиксированным числом наблюдений. Объём выборки в этой классической процедуре отбора определяется как наименьшее целое, при котором при наименее благоприятном для отбора случае ещё соблюдается ограничение на вероятность корректного отбора. По окончании наблюдений в качестве наилучшей выбирается компонента с наибольшим числом успехов.

В дальнейшем, для ещё большего повышения эффективности отбора в мультиномиальной модели Бекхофером и Голдсманом<sup>30</sup> была предложена последовательная процедура, являющаяся модификацией процедуры отбора Бекхофера-Кифера-Собея<sup>31</sup>. В процедуре Бекхофера-Голдсмана дополнительно к оригинальному правилу останова добавляется ограничение сверху  $n_0$  на объём наблюдений, по достижению которого процедура прекращает наблюдение и выносит вердикт. По определению,  $n_0$  выбирается как число наблюдений, гарантирующее заданную вероятность корректного отбора. Такой простой приём позволяет существенно сократить число наблюдений. В последующей статье<sup>32</sup> Бекхофер и Голдсман предложили ещё более эффективную последовательную процедуру отбора.

В книгах Гиббонса, Олкина, Собея<sup>33</sup> и Гупты, Панчапакесана<sup>34</sup> пред-

---

<sup>27</sup> *Mulekar, M.S.* On selecting a process with the smallest number of unfortunate events / M.S. Mulekar, F.J. Matejcek // The Journal of the Operational Research Society. 2006. Vol. 57, №4. P. 416–422.

<sup>28</sup> *Mulekar, M.S.* Fixed-sample-size selection problem for Poisson populations / M.S. Mulekar, M. Sobel // Statistics & Decisions. Supplemental Issue No. 4. 1999. pp. 69–85.

<sup>29</sup> *Bechhofer, R.E.* A single-sample multiple-decision procedure for selecting the multinomial event which has the highest probability / R.E. Bechhofer, S. Elmaghraby, N. Morse // The Annals of Mathematical Statistics. 1959. Vol. 30, №1— pp. 102–119.

<sup>30</sup> *Bechhofer, R.E.* Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the multinomial event which has the largest probability / R.E. Bechhofer, D.M. Goldsman // Communications in Statistics — Simulation and Computation. 1985. Vol. 14, №2. P. 283–315.

<sup>31</sup> *Bechhofer, R.E.* Sequential identification and ranking procedures / R.E. Bechhofer, J. Kiefer, M. Sobel. Chicago: University of Chicago Press. 1968. 420 p.

<sup>32</sup> *Bechhofer, R.E.* Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the multinomial event which has the largest probability (II): extended tables and an improved procedure / R.E. Bechhofer, D.M. Goldsman // Communications in Statistics — Simulation and Computation. 1986. Vol. 15, №3. pp. 829–851.

<sup>33</sup> *Gibbons, J. D.* Selecting and ordering populations / J. D. Gibbons, I. Olkin, M. Sobel. New York: Wiley. 1977. 569 p.

<sup>34</sup> *Gupta, S.S.* Multiple decision procedures: theory and methodology of selecting and ranking populations / S. S. Gupta, S. Panchapakesan. New York: Wiley. 1979. 573 p.

ставлен обширный и детальный обзор задач отбора и упорядочивания. В них рассматриваются различные подходы к постановке и решению этих задач, производится их детальное описание и исследование, представлены наиболее значимые процедуры отбора и упорядочивания.

### **Научная новизна:**

1. Впервые получены нижние границы для среднего объёма наблюдений в широком классе задач отбора и упорядочивания. Полученные результаты представляют решение нового класса статистических задач по планированию объёма испытаний.
2. Построенные границы получили применение к новому подходу в исследовании эффективности существующих процедур отбора и упорядочивания для различных моделей задач отбора и упорядочивания.

### **Цели и задачи:**

Целью данной работы является построение нижних границ для среднего объёма наблюдений последовательных гарантийных процедур отбора и упорядочивание, изучение их основных свойств, а также их применение для исследования эффективности некоторых из наиболее значимых процедур отбора и упорядочивания.

### **Методология и методы исследования:**

Решение поставленных задач в диссертации производилось с привлечением методов математического анализа, теории вероятностей. Главным инструментом, на котором базируется вывод основных результатов диссертации, явилась универсальная нижняя граница для среднего объёма наблюдений последовательных гарантийных процедур статистического вывода Володина-Малютова.

В дополнение к аналитическим результатам, приведены и различные численные исследования, выполненные с использованием языков программирования Python, C++ (с использованием компилятора GNU GCC), пакета статистических вычислений GNU R.

### **Теоретическая и практическая значимость:**

Практическая ценность построенных нижних границ состоит в их использовании как критерия недостаточности имеющегося у экспериментатора объёма наблюдений для существования гарантийных процедур. Кроме того, такие границы являются некоторым ориентиром в поиске оптимальных с точки зрения объёма наблюдений гарантийных процедур отбора и упорядочивания, что говорит о теоретической ценности работы.

### **Степень достоверности и апробация результатов:**

Степень достоверности полученных результатов определяется строгим математическим доказательством всех утверждений, присутствующих в диссертации. Результаты работы докладывались на всероссийской научной конференции "XII Всероссийский симпозиум по прикладной и промышленной математике" (Казань, 1–8 мая 2011 г.), на международной конференции "XXX Международный семинар по проблемам устойчивости стохастических моделей" (Светлогорск, 24–30 сентября 2012 г.), на всероссийской конференции "XIV Всероссийский симпозиум по прикладной и промышленной математике" (Йошкар-Ола, 12–18 мая 2013 г.), на научном семинаре кафедры Математической статистики факультета ВМК МГУ (2013 г.). Кроме того, работа многократно докладывалась на семинарах кафедры Математической статистики и итоговых научных конференциях института ВМиИТ КФУ (2011 – 2013 гг.).

### **Публикации:**

Материалы диссертации опубликованы в 6 печатных работах ([1]-[6]), из них 2 статьи опубликованы в журналах, включённых в перечень ВАК ([3], [6]).

### **Положения, выносимые на защиту**

На защиту выносятся следующие результаты:

1. Нижние границы для среднего объёма наблюдений в задачах отбора и упорядочивания общего вида, применимые к классу проблем со строго монотонной информацией по Кульбаку-Лейблеру.
2. Применение нижних границы общего вида к конкретным задачам отбора и упорядочивания. Построены границы для нормальной, экспоненциальной, биномиальной и пуассоновской моделей.
3. Результаты исследования эффективности некоторых из наиболее значимых процедур отбора и упорядочивания, основанного на использовании полученных в диссертации нижних границ.

### **Структура и объём диссертации:**

Диссертация состоит из введения, трёх глав и списка литературы, содержащего 33 наименований. В диссертации присутствуют 2 рисунка и 11 таблиц. Общий объём работы составляет 113 страниц.

## **Содержание работы**

**Первая глава** диссертации посвящена введению основных определений, а также вопросу построения нижних границ для среднего объёма наблюдений в последовательных процедурах отбора и упорядочивания, применимых для широкого класса распределений наблюдаемых характеристик популяций. В этой главе формулируются задачи отбора и упорядочивания в общем виде, вводится основной инструмент, на котором базируются результаты диссертации — универсальная нижняя граница Володина-Малютова для среднего объёма наблюдений в последовательных процедурах статистического вывода. Здесь же формулируются и доказываются основополагающие результаты диссертации, представляющие нижние границы для среднего объёма наблюдений в задачах отбора и упорядочивания, применяемые для широкого класса моделей со строго монотонной различающей информацией по Кульбаку-Лейблеру.

**В параграфе 1.1** формулируются задачи отбора и упорядочивания  $m \geq 2$  статистических популяций. Все  $m$  популяций имеют распределение общего вида и отличаются только значениями скалярного параметра  $\theta \in \Theta$ . Таким образом, значение  $\theta_i$  параметра  $\theta$  полностью определяет распределение наблюдаемого случайного элемента  $\xi_i$ , относящегося к  $i$ -ой популяции,  $1 \leq i \leq m$ . Задача отбора "наилучшей" популяции состоит в выборе популяции с наибольшим (наименьшим) значением параметра.

В ходе статистического эксперимента с управлением наблюдений последовательные процедуры отбора и упорядочивания доставляют в каждой популяции случайное число наблюдений  $\nu_i$ ,  $1 \leq i \leq m$  с суммарным объёмом выборки  $\nu = \nu_1 + \dots + \nu_m$ , представляющим момент остановки статистического эксперимента.

Значения параметра популяций составляют вектор значений  $\theta = (\theta_1, \dots, \theta_m) \in \Theta^m$ . Пусть также  $\theta_{[1]} \leq \dots \leq \theta_{[m]}$  — упорядоченные по возрастанию значения параметров. С целью сокращения обозначений, квадратные скобки в индексах  $\theta$  иногда будут убираться, так что в дальнейшем всегда будет подразумеваться, что  $\theta_1 \leq \dots \leq \theta_m$ .

Формулируется задача построения нижней границы для среднего объёма выборки  $E_{\theta}\nu$  в процедурах отбора или упорядочивания, гарантирующих заданное ограничение  $1 - \alpha$  на вероятность корректного решения — отбора наилучшей популяции, отделённой от остальных популяций некоторой "зоной безразличия" размера  $\Delta$ . Используются универсальные нижние границы Володина-Малютова. Основная техническая сторона проблемы состоит в ре-

шении достаточно сложной задачи на минимакс. Идея решения состоит в некотором аккуратном сужении параметрического пространства  $\Theta^m$  с убранной зоной безразличия. Сужение осуществляется таким образом, чтобы полученная нижняя граница была не слишком далека от границы, получаемой при точном решении задачи на минимакс, а иногда и совпадала с такой точной границей.

При формулировки последующих результатов диссертации будут использоваться следующие обозначения и понятия. Пусть  $I(\theta, \vartheta)$  — различающая информация по Кульбаку-Лейблеру между распределениями, индексированными параметрами  $\theta$  и  $\vartheta$ . Информация  $I(\theta, \vartheta)$  будет называться *строго монотонной*, если при каждом фиксированном  $\theta$  она строго убывает по  $\vartheta < \theta$  и строго возрастает, когда  $\vartheta > \theta$ .

**Параграф 1.2** состоит в построении нижней границы для среднего объёма наблюдений в процедурах отбора и исследовании случаев, при которых эта граница является точной — совпадает с исходной нижней границей.

Параметрическое пространство с зоной безразличия в задачах отбора вводится как

$$\Theta_{\Delta}^s = \{\boldsymbol{\theta} \in \Theta^m : \theta_{[m-1]} \leq r_{\Delta}(\theta_{[m]})\},$$

где  $r_{\Delta}(\theta)$  — определяющая вид зоны безразличия функция, сдвигающая аргумент  $\theta$  на величину  $\Delta$  влево.

Основной результат параграфа состоит в доказательстве следующего утверждения.

**ТЕОРЕМА 1.1.** Пусть семейство распределений в популяциях таково, что различающая информация строго монотонна. Тогда для любого  $\boldsymbol{\theta} \in \Theta_{\Delta}^s$  справедливо неравенство

$$\mathbf{E}_{\boldsymbol{\theta}} \nu \geq \sum_{i=1}^{m-1} \frac{\omega(\alpha, \alpha)}{I(\theta_i, r_{\Delta}^{-1}(r_t(\theta_m)))}$$

где  $t \geq 0$  должно удовлетворять условиям:

$$r_t(\theta_m) \in \Theta, \quad r_{\Delta}^{-1}(r_t(\theta_m)) \in \Theta,$$

$$\sum_{i=1}^{m-1} \frac{I(\theta_m, r_t(\theta_m))}{I(\theta_i, r_{\Delta}^{-1}(r_t(\theta_m)))} \leq 1,$$

$$r_t(\theta_m) \geq \theta_{m-1}.$$

Таким образом, всё сводится к выбору значения  $t$ , удовлетворяющего сформулированным условиям. Эти условия не однозначно определяют значение  $t$ . Заметим, что наибольшее значение  $t$  соответствует наибольшему значению нижней границе.

К сожалению, результат теоремы 1.1 существенным образом основан на некотором огрублении исходной универсальной нижней границы Володина-Малютова. С другой стороны, в некоторых ситуациях её точность всё же остаётся не меньшей, чем исходная граница.

Пусть выполняются условия Теоремы 1.1 и:

- (i)  $\theta_1 = \dots = \theta_{m-1}$  ;
- (ii) замкнутый интервал  $[r_{\Delta}(\theta_1); r_{\Delta}^{-1}(\theta_m)] \subseteq \Theta$  ;
- (iii) выполняется более строгий вариант условия на  $t$  Теоремы 1.1:

$$(m - 1)I(\theta_m, r_t(\theta_m)) = I(\theta_1, r_{\Delta}^{-1}(r_t(\theta_m)));$$

- (iv) существует значение  $w^* \in [0; 1]$  такое, что

$$\inf_{\vartheta \in \Theta} \left( \frac{1 - w^*}{m - 1} I(\theta_1, r_{\Delta}^{-1}(\vartheta)) + w^* I(\theta_m, \vartheta) \right) = I(\theta_m, r_t(\theta_m)).$$

Тогда граница, указанная в теореме 1.1, представляет точное решение задачи на минимакс в границе Володина-Малютова.

Вкратце, это предложение утверждает следующие: при некоторых условиях на распределение популяций (например, если различающая информация выпукла по второму аргументу) и если параметры всех популяций, кроме наилучшей, совпадают (то-есть  $\theta_1 = \dots = \theta_{m-1}$ ), то значение нижней границы, полученной в этом параграфе, совпадает со значением исходной универсальной нижней границы Володина-Малютова.

Наконец, в конце параграфа приводятся некоторые технические утверждения, касающиеся свойств условий, наложенных на  $t$  в теореме 1.1. Они применяются в главе 2 при рассмотрении некоторых задач отбора с конкретными распределениями.

**В параграфе 1.3** рассматривается вопрос построения нижних границ для среднего объёма наблюдений в процедурах упорядочивания, а также приводятся некоторые вспомогательные результаты, упрощающие применение этих границ в дальнейшем.

Параграф начинается с описание параметрического пространства с зоной безразличия, соответствующей задаче упорядочивания. Аналогично параграфу 1.2, для некоторого обобщения вида зоны безразличия это пространство определяется через функцию  $r_\Delta(\theta)$  и имеет вид

$$\Theta_\Delta^r = \{\theta \in \Theta^m : \theta_{[i]} \leq r_\Delta(\theta_{[i+1]}), 1 \leq i \leq m-1\}.$$

Далее определяются некоторые объекты, необходимые при формулировке основного результата параграфа. Это  $B_\Delta(\theta)$  — множество векторов  $\vartheta \in \Theta_\Delta^r$ , корректное решение для которых отличается от корректного решения для вектора  $\theta$ . Иными словами, в рамках задачи упорядочивания, это множество векторов, у которых порядок элементов по возрастанию отличается от порядка элементов по возрастанию у вектора  $\theta$ . Исходя из множества  $B_\Delta(\theta)$  определяется множество  $B_\Delta^i(\theta)$  (где  $1 \leq i \leq m-1$ ) значений  $\vartheta \in [\theta_i; r_\Delta^{-1}(\theta_{i+1})]$  таких, что выполняется условие

$$(\theta_1, \dots, \theta_{i-1}, \vartheta, r_\Delta(\vartheta), \theta_{i+2}, \dots, \theta_m) \in B_\Delta(\theta).$$

Наконец в виде теоремы формулируется основной результат параграфа — нижние границы для среднего объёма наблюдений в процедурах упорядочивания популяций.

**ТЕОРЕМА 1.2.** Пусть семейство распределений  $\mathcal{P}$  таково, что различающая информация  $I(u, v)$  строго монотонна. Тогда для любого  $\theta \in \Theta_\Delta^r$  справедлива оценка:

если  $m = 2$ , то

$$\mathbf{E}_\theta \nu \geq \frac{\omega(\alpha, \alpha)}{\sup_{w \in [0;1]} \inf_{\vartheta \in B_\Delta^1(\theta)} (wI(\theta_1, \vartheta) + (1-w)I(\theta_2, r_\Delta(\vartheta)))};$$

если  $m \geq 3$ , то

$$\mathbf{E}_\theta \nu \geq \sum_{i=1}^{m-1} \frac{\omega(\alpha, \alpha)}{V_i},$$

где

$$V_1 = \sup_{w \in [0;1]} \inf_{\vartheta \in B_\Delta^1(\theta)} (wI(\theta_1, \vartheta) + 2(1-w)I(\theta_2, r_\Delta(\vartheta)));$$

$$V_i = \sup_{w \in [0;1]} \inf_{\vartheta \in B_\Delta^i(\theta)} 2(wI(\theta_i, \vartheta) + (1-w)I(\theta_{i+1}, r_\Delta(\vartheta))), \quad 2 \leq i \leq m-2;$$

$$V_{m-1} = \sup_{w \in [0;1]} \inf_{\vartheta \in B_\Delta^{m-1}(\theta)} (2wI(\theta_{m-1}, \vartheta) + (1-w)I(\theta_m, r_\Delta(\vartheta))).$$

Полученная нижняя граница является существенно более простой в

вычислительном отношении, чем универсальная нижняя граница Володина-Малютова. Однако, как и в случае нижних границ для задачи отбора, её вывод существенным образом основан на огрублении исходной границы.

В окончании параграфа приводятся некоторые вспомогательные результаты, несколько облегчающие вычисление величин  $V_i$  в дальнейшем.

**Во второй главе** рассматривается вопрос построения нижних границ для среднего объёма наблюдений в конкретных задачах отбора и упорядочивания с заданными распределениями популяций.

**В параграфе 2.1** рассматриваются задачи отбора и упорядочивания при нормальном распределении популяций. При этом отбор и упорядочивание производятся относительно среднего с зоной безразличия, основанной на разности параметров. Дисперсии популяций полагаются совпадающими и известными, равными  $\sigma^2$ . Детализируется вид параметрического пространства с зоной безразличия:

$$\Theta_{\Delta}^s = \{\boldsymbol{\theta} \in \Theta^m : \theta_{[m]} - \theta_{[m-1]} \geq \Delta\}.$$

Нижняя граница для среднего объёма наблюдений в задаче отбора, основанная на теореме 1.1, имеет следующий вид:

$$E_{\boldsymbol{\theta}} \nu \geq \sum_{i=1}^{m-1} \frac{2\sigma^2 \omega(\alpha, \alpha)}{(\theta_m - \theta_i + \Delta - t)^2} \quad \forall \boldsymbol{\theta} \in \Theta_{\Delta}^s,$$

где

$$t = \frac{\theta_m - \theta_{m-1} + \Delta}{\sqrt{m-1} + 1}.$$

В случае наименее благоприятной для отбора конфигурации она является не менее точной, чем граница Володина-Малютова, и имеет вид

$$E_{\boldsymbol{\theta}} \nu \geq \frac{(\sqrt{m-1} + 1)^2 \omega(\alpha, \alpha)}{2\Delta^2} \sigma^2.$$

Рассмотрим теперь задачу упорядочивания популяций. Этой задаче соответствует параметрическое пространство

$$\Theta_{\Delta}^r = \{\boldsymbol{\theta} \in \Theta^m : \theta_{i+1} - \theta_i \geq \Delta, 1 \leq i \leq m-1\}.$$

Имеет место неравенство. Если  $m \geq 3$ , то

$$E_{\boldsymbol{\theta}} \nu \geq \sum_{i=1}^{m-1} \frac{4\sigma^2 \omega(\alpha, \alpha)}{(\theta_{i+1} - \theta_i + \Delta)^2}.$$

**Параграфы 2.2, 2.3, 2.4** посвящёны построению нижних границ для среднего объёма выборки в задачах отбора и упорядочивания для, соответственно, показательной, биномиальной и пуассоновской моделей. Результаты этих параграфов в целом аналогичны нормальной модели

**В параграфе 2.5** рассматривается задача отбора наиболее вероятного исхода в мультиномиальной схеме испытаний  $\mathcal{M}(\boldsymbol{\theta}, m, \nu)$ , где  $\boldsymbol{\theta}$  — вероятности соответствующих  $m$  исходов,  $\nu$  — момент останова в мультиномиальной схеме наблюдений.

Нижняя граница здесь строится при зоне безразличия, основанной на отношении значений параметра:

$$\Theta_{\Delta} = \{\boldsymbol{\theta} : \Delta\theta_{[m-1]} < \theta_{[m]}\}.$$

Рассматриваемая в этом параграфе задача отбора представляет собой особый случай, концептуально несколько отличающийся от ранее рассматриваемых задач. Здесь деление на популяции носит условный характер, так как по-сути присутствует лишь одна популяция, из которой наблюдается векторная случайная характеристика.

Устанавливается, что

$$\begin{aligned} \mathbb{E}_{\theta\nu} \geq \omega(\alpha, \alpha) & \left( (\theta_{m-1} + \theta_m) \ln \frac{1 + \Delta}{\theta_{m-1} + \theta_m} + \right. \\ & \left. \theta_{m-1} \ln \frac{\theta_{m-1}}{\Delta} + \theta_m \ln \theta_m \right)^{-1}. \end{aligned}$$

Заметим, что, в отличие от других построенных в этой главе нижних границ, нижняя граница для указанной мультиномиальной модели строится путём непосредственного решения максиминной задачи в универсальной нижней границе Володина-Малютова. Этот результат, таким образом, является точным в том смысле, что полученная нижняя граница принимает в точности те же значения, что и исходная универсальная граница.

**В третьей главе** полученные нижние границы применяются для исследования эффективности некоторых процедур отбора и упорядочивания для различных моделей. Исследование в основном касается понятия асимптотической эффективности при  $\alpha \rightarrow 0$ . Эта величина будет обозначаться как:

$$\mathcal{E}(\boldsymbol{\theta}) = \lim_{\alpha \rightarrow 0} \frac{\inf_{\psi \in H(\alpha)} \mathbb{E}_{\theta\nu\psi}}{\mathbb{E}_{\theta\nu\varphi}},$$

где  $\varphi$  — гарантийная процедура отбора или упорядочивания, эффективность которой рассматривается, а  $H(\alpha)$  — множество всех гарантийных процедур, вероятность корректного отбора или упорядочивания в которых не меньше  $1 - \alpha$  для любого  $\theta$ , принадлежащий соответствующему параметрическому пространству с зоной безразличия.

В параграфе 3.1 оценивается эффективность процедур отбора для нормальной модели. Заметим, что, обычно, наибольший интерес вызывает поведение процедур при наименее благоприятной для отбора конфигурации, которая в данном нормальном случае имеет вид:

$$\theta_1 = \dots = \theta_{m-1} = \theta_m - \Delta.$$

В первую очередь исследуется классическая процедура отбора Бекхофера с фиксированным числом наблюдений, объёма наблюдений в которой имеет следующее асимптотическое (при  $\alpha \rightarrow 0$ ) поведение:

$$n_\alpha = -\frac{4m\sigma^2}{\Delta^2} \ln \alpha (1 + O(1)).$$

Применяя соответствующую нижнюю границу, получаем оценку асимптотической эффективности вида:

$$\mathcal{E}(\theta) \geq \frac{1}{2m} \sum_{i=1}^{m-1} \frac{\Delta^2}{(\theta_m - \theta_i + \Delta - t)^2}, \quad t = \frac{\theta_m - \theta_{m-1} + \Delta}{\sqrt{m-1} + 1},$$

и, в частности, для наименее благоприятной конфигурации

$$\mathcal{E}(\theta) \geq \frac{(\sqrt{m-1} + 1)^2}{8m}.$$

Далее рассматривается эффективность последовательной процедуры отбора Бекхофера-Кифера-Собея, обладающей достаточно простым управлением: на каждом шаге эксперимента из каждой популяции берётся ровно по одному наблюдению, после чего решается вопрос о продолжении эксперимента (*vector-at-once*). Для среднего объёма наблюдений этой процедуры верна следующая асимптотическая (при  $\alpha \rightarrow 0$ ) формула:

$$E_{\theta\nu} = -\frac{m\sigma^2}{\Delta(\theta_m - \theta_{m-1})} \ln \alpha (1 + O(1)).$$

Таким образом, исходя из полученной нижней границ получаем оценку для асимптотической эффективности:

$$\mathcal{E}(\theta) \geq \frac{2}{m} \sum_{i=1}^{m-1} \frac{\Delta(\theta_m - \theta_{m-1})}{(\theta_m - \theta_i + \Delta - t)^2}, \quad t = \frac{\theta_m - \theta_{m-1} + \Delta}{\sqrt{m-1} + 1},$$

или , в случае наименее благоприятной конфигурации:

$$\sup_{\boldsymbol{\theta} \in \Theta_{\Delta}^s} \mathcal{E}(\boldsymbol{\theta}) \geq \frac{(\sqrt{m-1} + 1)^2}{2m}.$$

Ещё более совершенная последовательная процедура отбора была предложена Као и Lai. Её управление основано на экранировании: в процессе эксперимента процедура может исключать из дальнейшего рассмотрения популяции, маловероятно являющиеся наилучшими. Средний объём наблюдений этой процедуры имеет следующий асимптотический (при  $\alpha \rightarrow 0$ ) вид:

$$\begin{aligned} E_{\theta^{\nu}} = & - \left( \sum_{i=1}^{m-1} \frac{4\sigma^2}{(\Delta + \theta_m - \theta_i)^2} + \frac{4\sigma^2}{(\Delta + \theta_m - \theta_{m-1})^2} \right) \\ & \cdot \ln \alpha (1 + O(1)). \end{aligned}$$

Применение нижней границы позволяет построить следующую оценку асимптотической эффективности:

$$\begin{aligned} \mathcal{E}(\boldsymbol{\theta}) \geq & \frac{1}{2} \sum_{i=1}^{m-1} \frac{1}{(\theta_m - \theta_i + \Delta - t)^2} \\ & \cdot \left( \sum_{i=1}^{m-1} \frac{1}{(\Delta + \theta_m - \theta_i)^2} + \frac{1}{(\Delta + \theta_m - \theta_{m-1})^2} \right)^{-1}, \end{aligned}$$

где

$$t = \frac{\theta_m - \theta_{m-1} + \Delta}{\sqrt{m-1} + 1}.$$

В частном случае наименее благоприятной конфигурации это принимает вид:

$$\sup_{\boldsymbol{\theta} \in \Theta_{\Delta}^s} \mathcal{E}(\boldsymbol{\theta}) \geq \frac{(\sqrt{m-1} + 1)^2}{2m}.$$

Кроме того, была получена оценка в наглядной форме, показывающая поведение эффективности процедуры при большом количестве рассматриваемых популяций. Пусть дана последовательность векторов

$$\boldsymbol{\theta}^k = (\theta_1^k, \dots, \theta_{m_k}^k) \in \Theta_{\Delta}^k,$$

такая, что  $m_k \rightarrow \infty$  при  $k \rightarrow \infty$  и существует константа  $C > 0$  такая, что

$$|\theta_i^k - \theta_j^k| < C, \quad 1 \leq i, j \leq m_k.$$

Тогда

$$\lim_{m \rightarrow \infty} \mathcal{E}(\boldsymbol{\theta}^k) \geq \frac{1}{2}.$$

В рамках **параграфа 3.2** производится исследование асимптотической эффективности последовательной процедуры упорядочивания Бекхофера-Кифера-Собеля. Для этой процедуры верна асимптотическая (при  $\alpha \rightarrow 0$ ) формула:

$$E_{\theta\nu} = -\frac{m \ln \alpha}{\Delta \min_{1 \leq i \leq m-1} (\theta_{i+1} - \theta_i)} (1 + O(1)).$$

Применяя нижнюю границу для рассматриваемой задачи упорядочивания нормальных популяций по значениям их средних, построенную во второй главе, немедленно получаем оценку асимптотической эффективности процедуры при  $m \geq 3$ :

$$\mathcal{E}(\boldsymbol{\theta}) \geq \sum_{i=1}^{m-1} \frac{4\Delta}{m(\theta_{i+1} - \theta_i + \Delta)^2} \min_{1 \leq i \leq m-1} (\theta_{i+1} - \theta_i),$$

или для наименее благоприятной конфигурации:

$$\sup_{\boldsymbol{\theta} \in \Theta_{\Delta}^r} \mathcal{E}(\boldsymbol{\theta}) \geq \frac{m-1}{m}.$$

Аналогично, в **параграфах 3.3 и 3.4** исследуется эффективность процедур отбора и упорядочивания для биномиальной и мультиномиальной моделей. Эти исследования осуществляются в основном численными методами.

**Параграф 3.5** изучает эффективность процедур отбора для мультиномиальной модели при фиксированном ограничении на вероятность корректного отбора. Получены численные результаты об эффективности классической процедуры отбора Бекхофера-Элмаграби-Морсе, а также более эффективной последовательной процедуры отбора Бекхофера-Голдсмана.

## Заключение

В диссертации выполнены следующие основные задачи:

1. Получены нижние границы для среднего объёма наблюдений в задачах отбора и упорядочивания общего вида, применимые для широкого класса проблем, распределение популяции которых удовлетворяет условию строгой монотонности различающей информации по Кульбаку-Лейблеру.

2. С помощью общих границ получены нижние границы для нескольких конкретных, наиболее часто рассматриваемых, моделей: нормальной, экспоненциальной, биномиальной, пуассоновской, мультиномиальной. Установлены их основные свойства.
3. В качестве приложения сконструированных во второй главе нижних границ, рассмотрено исследование эффективности нескольких процедур отбора и упорядочивания для различных моделей.

Полученные нижние границы можно рекомендовать практикам в качестве, например, критерия недостаточности для гарантийного отбора или упорядочивания объёма наблюдений, которым располагает статистик.

В перспективе могут быть решены следующие задачи:

1. Получение нижних границ для среднего объёма наблюдений в гарантийных процедурах отбора, построенных согласно подходу Гупты.
2. Построение нижних границ для задач отбора и упорядочивания в непараметрической модели.
3. Решение аналогичных задач в байесовском подходе проблемы статистического вывода.

Работа выполнена под руководством доктора физико-математических наук, профессора Игоря Николаевича Володина, которому автор выражает искреннюю благодарность.

## ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. *Кареев, И.А.* Нижняя граница для среднего объёма выборки и эффективность процедур отбора / И.А. Кареев // Обзорное прикладное и промышленное математическое издание.— 2011.— Т. 18, №1.— С. 84–85.
2. *Кареев, И.А.* Нижняя граница для среднего объёма выборки и эффективность процедур отбора / И.А. Кареев // Итоговая научно-образовательная конференция студентов КФУ 2011.— С. 107–108.
3. *Кареев, И.А.* Нижние границы для среднего объёма выборки и эффективность последовательных процедур отбора / И.А. Кареев // Теория вероятностей и её применения.— 2012.— Т. 57, №2.— С. 278–295.

4. *Kareev, I.A.* Lower bounds for average observations number in selection and ranking of binomial and Poisson populations / I.A. Kareev // XXX International Seminar on Stability Problems for Stochastic Models and VI International Workshop „Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems“, Book of abstracts.— 2012.— pp. 29–31.
5. *Кареев, И.А.* Нижние границы для среднего объёма наблюдений / И.А. Кареев, И.Н. Володин // Обозрение прикладной и промышленной математики.— 2013.— Т. 20, №2.— С. 139.
6. *Кареев, И.А.* Нижние границы для среднего объёма выборки и эффективность последовательных процедур упорядочивания / И.А. Кареев // Теория вероятностей и её применения.— 2013.— Т. 58, №3.— С. 591–597.