ФГБОУ ВО
Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

На правах рукописи
УДК 519.213.22+519.233.32

Савелов Максим Павлович

Экстремальные характеристики критериев выбора статистических гипотез

01.01.05 – теория вероятностей и математическая статистика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва 2016

Работа выполнена на кафедре математической статистики и случайных процессов механико-математического факультета Московского государственного университета имени М.В. Ломоносова.

Научный руководитель: доктор физико-математических наук
Зубков Андрей Михайлович

Официальные оппоненты:
доктор физико-математических наук, Боровков Константин Александрович, School of Mathematics and Statistics, The University of Melbourne (Школа математики и статистики Мельбурнского университета, Австралия, профессор)

кандидат физико-математических наук, доцент Иванов Андрей Викторович, доцент кафедры «Компьютерная безопасность» Департамента прикладной математики МИЭМ НИУ ВШЭ

Ведущая организация:
Федеральное государственное бюджетное учреждение науки Центральный экономико-математический институт РАН

Защита диссертации состоится «23» декабря 2016 года в 16 часов 45 минут на заседании диссертационного совета Д 501.001.85 при МГУ имени М.В. Ломоносова по адресу: 119234, ГСП-1, Москва, Ленинские Горы, МГУ, механико-математический факультет, аудитория 16-24.

С диссертацией можно ознакомиться в Фундаментальной библиотеке ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова» по адресу Москва, Ломоносовский проспект, 27, сектор А и на сайтах механико-математического факультета:
http://mech.math.msu.su/~snark/index.cgi,
http://istina.msu.ru/dissertations/37196249/

Автореферат разослан «   » ноября 2016 года.

Ученый секретарь диссертационного
совета Д 501.001.85 на базе МГУ,
доктор физико-математических наук,
профессор Власов В.В.

Общая характеристика работы

Диссертация посвящена изучению экстремальных характеристик критериев выбора статистических гипотез.

Актуальность темы.

Одной из важных статистических задач является задача о различении нескольких (скажем, n) простых гипотез. Если $n = 2$, то ее решение (для широкого класса случаев) может быть получено с помощью леммы Неймана-Пирсона. Рассмотрим ситуацию, когда $n = 3$. Для любой фиксированной тройки вероятностных мер $\vec{\mu} := (\mu_1, \mu_2, \mu_3)$ можно найти оптимальный в том или ином смысле критерий и определить его «качество» (например, понимая под качеством вероятность ошибки в худшем случае) как функцию от тройки мер. В связи с этим естественным образом возникает вопрос: в каких пределах меняется качество критерия, если тройки вероятностных мер принадлежат некоторому множеству. Формализуем вышесказанное.

Пусть M - множество вероятностных мер на измеримом пространстве (Ω, Σ). Для любых трех вероятностных мер μ_1, μ_2, μ_3 на (Ω, Σ) можно рассмотреть задачу различения гипотез H_i: наблюдение ξ имеет распределение μ_i $(i = 1, 2, 3)$. Каждому нерандомизированному критерию в этой задаче соответствуют измеримое разбиение $\vec{C} = \{C_1, C_2, C_3\}$ пространства Ω ($\bigsqcup_{i=1}^{3} C_i = \Omega$; если наблюдение $\xi \in C_i$, то принимается гипотеза H_i, $i = 1, 2, 3$) и вероятности $\mu_j(C_k)$ принятия гипотезы H_k в случае, когда верна гипотеза H_j, $j, k \in \{1, 2, 3\}$. Совокупность всех разбиений Ω на три не пересекающихся измеримых подмножества обозначим $\mathcal{C}(\Omega)$. Далее, через $\vec{\mu}$ обозначим тройку (μ_1, μ_2, μ_3), через $\vec{\mu}(\vec{C})$ обозначим упорядоченный набор $(\mu_1(C_1), \mu_1(C_2), ..., \mu_3(C_3))$. Будем считать, что качество критерия можно представить функцией $f(\vec{\mu}(\vec{C}))$ со значениями в $\mathbb{R}$; иногда будет использоваться также обозначение $f(\{\mu_j(C_k)\}_{j,k=1}^{3})$. Фиксируем такие числа r_{ij}, $1 \leq i < j \leq 3$, что $0 \leq r_{ij} \leq 1$, $1 \leq i < j \leq 3$, $r_{12} + r_{13} \geq r_{23}$, $r_{12} + r_{23} \geq r_{13}$ и $r_{13} + r_{23} \geq r_{12}$. Рассмотрим семейство всевозможных троек вероятностных мер на Σ, между которыми заданы попарные расстояния по вариации следующим образом: $\rho(\mu_i, \mu_j) = r_{ij}$, $1 \leq i < j \leq 3$, где

$$\rho(\mu_i, \mu_j) := \sup_{A \in \Sigma} |\mu_i(A) - \mu_j(A)|.$$

Будем называть такие тройки мер $\vec{r}$-согласованными и обозначать это следующим образом: $\vec{\mu} \in D(\vec{r})$. Естественным образом возникает вопрос о получении удобных выражений для чисел

$$\bar{R} = \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f(\vec{\mu}(\vec{C})), \quad \underline{R} = \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f(\vec{\mu}(\vec{C})). \tag{1}$$

Отметим, что величины

$$\inf_{\vec{\mu}\in D(\vec{r})} \inf_{\vec{C}\in\mathcal{C}(\Omega)} f(\vec{\mu}(\vec{C})) \quad \text{и} \quad \sup_{\vec{\mu}\in D(\vec{r})} \inf_{\vec{C}\in\mathcal{C}(\Omega)} f(\vec{\mu}(\vec{C}))$$

сразу получаются из величин $\bar{R}$ и $\underline{R}$ за счет замены f на $-f$.

Величины $\bar{R}$ и $\underline{R}$ характеризуют качество оптимального нерандомизированного критерия в условиях, когда известны только попарные расстояния по вариации между мерами μ_1, μ_2, μ_3.

Одним из классических результатов в обсуждаемой тематике является лемма Фано[1], позволяющая оценить вероятности ошибок в терминах расстояния Кульбака-Лейблера. Усиление этого результата (также в терминах расстояния Кульбака-Лейблера) есть в статье L.Birge[2]. Оценка вероятностей ошибок через более общие f-дивергентные расстояния приведена в статье А.А. Гущина[3], результаты которой позволяют усилить результаты статьи L.Birge, а также получить следующие оценки вероятностей ошибок в терминах расстояний по вариации.

Пусть $\mathbb{E}=(\Omega, \mathbb{F}, (P_\theta)_{\theta\in\Theta})$— статистическая модель, причем $\Theta=\{1, 2, ..., n\}$ и пространство D решений совпадает с множеством Θ, функция потерь $L(\theta, d) = \mathbb{I}_{\theta\neq d}$, функция $\phi = (\phi_1, ..., \phi_n)$ задает рандомизированный критерий ($\phi_i : \Omega \to [0, 1]$ измеримы для каждого $1 \leq i \leq n$ и $\sum_{i=1}^n \phi_i = 1$) следующим образом: $\phi_i(\omega)$ - это вероятность принять решение под номером i, если происходит ω. Множество таких критериев обозначим через Φ. Далее, $a_M := \sup_{\phi\in\Phi} \min_{1\leq i\leq n} \mathbf{E}_{P_i}\phi_i$ (число $1-a_M$ равно вероятности ошибки «в худшем случае»). Положим $\bar{a} := \sup_{\phi\in\Phi} \frac{1}{n}\sum_{i=1}^n \mathbf{E}_{P_i}\phi_i$. С помощью теоремы 1 из статьи А.А. Гущина[3] легко получить следующую оценку:

$$a_M \leq \bar{a} \leq \frac{1}{n} \cdot (1 + \min_j \sum_{i=1}^n \rho(P_i, P_j)). \tag{2}$$

В случае, когда $n = 2$ и критерии рандомизированные, величина $\bar{a}$ выража-

[1]Fano R.M., «Class notes for transmission of information», MIT, Cambridge, MA, Course 6.574, (1952)

[2]Birge L., A new look at an old result: Fano's lemma, Prepublication n^0 632 du Labotatoire de Probabilites & Modeles Aleatores, Univesites de Paris 6 & 7, January 2001.

[3]Gushchin A.A. , On Fano's lemma and similar inequalities for the minimax risk, Theor.Probability and Math.Statist. No. 67, (2003).

ется через расстояние по вариации (см. [4], стр. 461-462): $\bar{a} = \frac{1+\rho(P_1,P_2)}{2}$.

Ряд результатов, связывающих вероятность ошибки и характеристики соответствующих распределений (например, энтропию), представлен в [5 – 7]. Схожие вопросы рассмотрены также в статье [8], основная теорема которой состоит в следующем. Пусть $\mu_1, ..., \mu_n$ - безатомные вероятностные меры на измеримом пространстве $(S, \mathcal{B})$, Π_S - множество измеримых разбиений S,

$$M = M(\mu_1, ..., \mu_n) := \sup\left\{\sum_{i=1}^{n} \mu_i(A_i) : \{A_i\}_{i=1}^{n} \in \Pi_S\right\},$$

$$v = v(\mu_1, ..., \mu_n) := \sup\left\{\min_{1\le i\le n} \mu_i(A_i) : \{A_i\}_{i=1}^{n} \in \Pi_S\right\},$$

тогда $(n - M + 1)^{-1} \le v \le Mn^{-1}$.

Задача, решаемая в первой главе диссертации, отличается от вышеперечисленных тем, что, ограничившись только множеством Φ^* нерандомизированных критериев, мы получим удобные формулы для величин $\underline{R}$ и $\bar{R}$, которые дают неулучшаемые оценки для величины $\sup_{\vec{C}\in\mathcal{C}(\Omega)} f(\vec{\mu}(\vec{C}))$ в условиях, когда заданы только расстояния по вариации r_{ij}, $1 \le i < j \le 3$. Будет показано, что нахождение экстремальных значений кусочно-непрерывных функций от вероятностей ошибок в широком классе случаев сводится к решению конечномерной задачи линейного программирования.

Некоторые результаты диссертации относятся и к случаю с рандомизированными критериями. В частности, мы докажем, что в условиях, когда $n = 3$ и известны только расстояния по вариации между мерами, оценка (2) является неулучшаемой, причем доказательство позволяет одновременно получить нижнюю оценку для $\bar{a}$ (в том числе и для произвольного n).

Отметим, что рассматриваемая тематика отчасти связана с проблемой деления торта (cake-cutting problem), см., например, [8]. В самом деле, задача о делении торта (Ω) на куски (измеримые множества C_i) между несколькими людьми (каждому человеку соответствует вероятностная мера μ_i на Ω, $1 \le i \le n$) может быть формализована, в частности, одним из следующих

[4] Ширяев А.Н., Вероятность, т.1. М.: МНЦМО, 2004.

[5] Ben-Bassat M., f-Entropies, Probability of Error, and Feature Selection, Information and Control, 39, (1978), 227-242.

[6] Golic J., On the Relationship Between the Information Measures and the Bayes Probability of Error, Information Theory, IEEE Transactions, Vol. 33, Issue: 5, Sep 1987, pp. 681-693.

[7] Golic J., On the Relationship Between the Separability Measures and the Bayes Probability of Error, Information Theory, IEEE Transactions, Vol. 33, Issue: 5, Sep 1987, pp. 694-701.

[8] Elton J., Hill T.P. and Kertz R., Optimal-partitioning inequalities for nonatomic probability measures, Transactions of the American mathematical Society, **296**:2, August 1986.

способов ([9], стр.62). Можно назвать деление «честным», если, например, выполняется одно из следующих условий:
1. $\mu_i(C_i) \ge \frac{1}{n}$ для каждого i, $1 \le i \le n$ (обычное честное деление),
2. $\mu_i(C_i) \ge \mu_i(C_j)$ для всех пар $1 \le i, j \le n$ («отсутствие зависти»).

Кроме того, естественно считать, что деление тем лучше, чем больше величина $\sum_{1\le i\le n} \mu_i(C_i)$ («суммарная удовлетворенность»). В рамках этих трех подходов можно полагать, что деление тем лучше, чем больше значения следующих функций:
1. $f_+(\vec{\mu}(\vec{C})) = f_1(\vec{\mu}(\vec{C})) = \sum_i \mu_i(C_i)$,
2. $f_\Delta(\vec{\mu}(\vec{C})) = f_2(\vec{\mu}(\vec{C})) = \min_{i\ne j}(\mu_i(C_i) - \mu_i(C_j))$,
3. $f_{\min}(\vec{\mu}(\vec{C})) = f_3(\vec{\mu}(\vec{C})) = \min_i \mu_i(C_i)$.

Далее, естественным образом возникает задача о нахождении числовых характеристик оптимальных разбиений, т.е. задача о нахождении

$$\inf_{\vec{\mu}\in A} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_l(\vec{\mu}(\vec{C})) \quad \text{и} \quad \sup_{\vec{\mu}\in A} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_l(\vec{\mu}(\vec{C})), \ \ 1 \le l \le n,$$

где A — некое множество наборов из n мер на Ω. Мы будем рассматривать случай, когда $A = D(\vec{r})$, и использовать следующие обозначения:

$$\underline{R}^{+} := \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_+(\vec{\mu}(\vec{C})), \ \ \bar{R}^{+} := \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_+(\vec{\mu}(\vec{C})), \tag{3}$$

$$\underline{R}^{\Delta} := \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_\Delta(\vec{\mu}(\vec{C})), \ \ \bar{R}^{\Delta} := \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_\Delta(\vec{\mu}(\vec{C})), \tag{4}$$

$$\underline{R}^{\min} := \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_{\min}(\vec{\mu}(\vec{C})), \ \ \bar{R}^{\min} := \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{C}(\Omega)} f_{\min}(\vec{\mu}(\vec{C})). \tag{5}$$

Задача о нахождении этих величин также освещена в диссертации.

Другой круг исследованных в диссертации задач связан с r-кратным критерием χ^2 (см., например, [10 – 14]).

Рассмотрим схему независимых испытаний с N исходами. Будем считать,

[9] Robertson J., Webb W., Cake-Cutting Algorithms. Be Fair, if You Can, Natick, Massachusetts: A K Peters, Ltd, (1998).

[10] Захаров В.К., Сарманов О.В., Севастьянов Б.А.,. Последовательный критерий χ^2, Матем. сб., **79(121)**:3(7) (1969), 444-460.

[11] Селиванов Б.И., Чистяков В.П., Многомерное распределение хи-квадрат для неоднородной полиномиальной схемы, Дискрет. матем., **10**:2 (1998), 52-61.

[12] Ронжин А.Ф., Предельное распределение процесса хи-квадрат при наличии разладки, Последовательный критерий хи-квадрат, Теория вероятн. и ее примен., **29**:3 (1984), 590-594.

[13] Гермогенов А.П., Ронжин А.Ф., Последовательный критерий хи-квадрат, Теория вероятн. и ее примен., **29**:2 (1984), 387-392.

[14] Туманян С.Х., Асимптотическое распределение критерия χ^2 при одновременном возрастании объема наблюдений и числа групп, Теория вероятн. и ее примен., **1**:1 (1956), 131-145.

что вероятность j-го исхода равна $p_j > 0, j = 1, ..., N$. Рассмотрим величины $I_j(t)$, которые равны 1, если при t-м испытании осуществился j-й исход, и $I_j(t) = 0$ в остальных случаях. Статистика Пирсона вычисляется по формуле $X(n) := \sum_{j=1}^{N} \frac{1}{np_j}(\sum_{t=1}^{n} I_j(t) - np_j)^2$. На основании набора статистик

$$X = (X(n_1), X(n_2), \ldots, X(n_r)), \quad n_1 < n_2 < \ldots < n_r,$$

строится последовательный r-кратный критерий согласия χ^2. Рассмотрим случайный процесс с дискретным временем $X(n)$, $n \geq 1$. В диссертации доказано, что при надлежащей нормировке по времени конечномерные распределения процесса, образованного последовательными значениями статистики Пирсона для выборок увеличивающихся со временем объемов, сходятся к конечномерным распределениям стационарного случайного процесса — нормированного квадрата процесса Бесселя. Полученные в [10] результаты о предельных совместных распределениях статистики Пирсона используются для вывода явных формул для плотности совместных распределений процесса Бесселя.

Цели работы. Нахождение экстремальных и асимптотических характеристик критериев выбора и проверки некоторых статистических гипотез.

Научная новизна. Результаты диссертации являются новыми и состоят в следующем:

- В задаче различения $n \geq 3$ гипотез с заданными расстояниями по вариации между распределениями наблюдений:
  А) показано, что нахождение экстремальных значений произвольных кусочно-линейных непрерывных функций от вероятностей ошибок сводится к решению конечномерной задачи линейного программирования;
  Б) для экстремальных значений нескольких конкретных функций от вероятностей ошибок найдены точные формулы (при $n = 3$) и оценки (при $n \geq 3$).

- Доказано, что конечномерные распределения последовательности значений статистики Пирсона, построенных по выборкам растущих объемов при надлежащей нормировке по времени сходятся к конечномерным распределениям стационарного случайного процесса (нормированного квадрата процесса Бесселя). Опубликованные в 1969 г. формулы для предельных совместных распределений статистики Пирсона использованы при выводе явных формул для плотности совместных распределений процесса Бесселя.

Основные методы исследования. В работе используются аналитические методы теории вероятностей, теории случайных процессов и методы оптимизации.

Теоретическая и практическая ценность работы. Диссертация имеет теоретический характер. Ее результаты могут быть полезны специалистам в области математической статистики.

Апробация работы. Результаты настоящей диссертации докладывались автором на следующих конференциях:

- Международная конференция «Statistics meets Stochastics», Снегири, Россия, 26.11.2014-29.11.2014.
- Международная конференция «Ломоносов», Москва, Россия, 11.04.2016-15.04.2016.
- Научная конференция «Ломоносовские чтения», Москва, Россия, 18.04.2016-27.04.2016.
- Международная конференция «Computer Data Analysis & Modeling», Минск, Беларусь, 06.09.2016-10.09.2016.
- Международная конференция «Анализ, вероятность и геометрия», Москва, Россия, 26.09.2016-01.10.2016.

Публикации. Основные результаты диссертации опубликованы в 4 работах (из них 3 в журналах из перечня ВАК), список которых приведён в конце автореферата.

Структура и объём работы. Диссертация состоит из введения, трех глав, заключения и списка литературы из 24 наименований. Общий объем диссертации составляет 92 страницы.

Краткое содержание диссертации

Приведём основные результаты диссертации. Нумерация утверждений совпадает с их нумерацией в соответствующих главах диссертации.

Во **введении** приведён обзор литературы по тематике работы, изложены цели исследования, а также перечислены основные полученные результаты.

Первая глава диссертации посвящена изучению чисел $\bar{R}$ и $\underline{R}$, которые определены в (1) и являются экстремальными значениями характеристик нерандомизированных статистических критериев выбора одной из нескольких гипотез в случае, когда известны только попарные расстояния по вариации между соответствующими им распределениями.

Во-первых, мы рассмотрим конкретную функцию качества критерия, которая имеет вид $f_+(\vec{\mu}(\vec{C})) = \sum_{i=1}^{3} \mu_i(C_i)$. Выбор ее связан с тем, что (как показано в лемме 4) оптимальным с точки зрения этой функции f_+ будет критерий, который тесно связан с оценкой максимального правдоподобия.

Кроме введенных выше величин $\bar{R}^+$, $\underline{R}^+$, $\underline{R}^\Delta$, $\bar{R}^\Delta$, $\bar{R}^{\min}$ и $\underline{R}^{\min}$ (см. (3)–(5)) нам потребуется новое обозначение. Перенумеруем три числа r_{12}, r_{13}, r_{23} в порядке неубывания, получим

$$r^{(1)} \le r^{(2)} \le r^{(3)}, \quad \left\{r^{(1)}, r^{(2)}, r^{(3)}\right\} = \{r_{12}, r_{13}, r_{23}\}.$$

В главе 1 установлена следующая теорема.

Теорема 1. *Пусть $n = 3$ и в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 6 непустых попарно не пересекающихся множеств. Тогда выполняются следующие равенства:*

$$\underline{R}^+ = \max(r^{(1)} + r^{(2)} + r^{(3)}, 1 + r^{(3)}), \quad \bar{R}^+ = 1 + r^{(1)} + r^{(2)}.$$

Более того, показано, что экстремальные значения функции f_+ не изменятся и в случае, если экстремум берется по множеству рандомизированных критериев.

В главе 1 также установлены следующие теоремы.

Теорема 2. *Пусть $n = 3$ и в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 6 непустых попарно не пересекающихся множеств. Пусть $r^{(1)} = r^{(2)}$ и $r^{(3)} \le \frac{2r^{(1)}+1}{3}$. Тогда $\bar{R}^\Delta = \frac{\bar{R}^+-1}{2} = \frac{r^{(1)}+r^{(2)}}{2}$.*

Теорема 3. *Пусть $n = 3$ и в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 6 непустых попарно не пересекающихся множеств. Пусть $r^{(2)} \le \frac{r^{(1)}+1}{2}$. Тогда $\bar{R}^{\min} = \frac{\bar{R}^+}{3} = \frac{1+r^{(1)}+r^{(2)}}{3}$.*

Отметим, что условия теоремы 2 сильнее, чем условия теоремы 3, которые, в свою очередь, очевидно сильнее, чем условия теоремы 1.

Пример 1. *Пусть $n = 3$ и все $r_{ij} = r$, тогда $\bar{R}^\Delta = r$, $\bar{R}^{\min} = \frac{1+2r}{3}$, $\bar{R}^+ = 1 + 2r$, $\underline{R}^+ = \max(3r, 1 + r)$.*

В первой главе диссертации доказано, что в случае $n = 3$ можно провести «понижение размерности», т.е. свести исходную задачу о нахождении

чисел $\bar{R}$ и $\underline{R}$ к поиску экстремума функции на выпуклых замкнутых многогранниках (полное описание которых дано в диссертации) $N_{18} \subset [0,1]^{54}$ и $N_6 \subset [0,1]^{18}$ соответственно, а именно, справедлива следующая теорема.
Теорема 4. *Предположим, что в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 18 непустых попарно не пересекающихся множеств. Тогда для любой функции $f : [0,1]^9 \to \mathbb{R}$ выполнены следующие равенства:*

$$\bar{R} = \sup_{\vec{x} \in N_{18}} f\left(\left\{\sum_{\sigma \in S_3} x_\sigma^{(j),k}\right\}_{j,k=1}^{3}\right),$$

$$\underline{R} = \inf_{\vec{x} \in N_6} \max_{\vec{C} \in \mathcal{C}(S_3)} f\left(\left\{\sum_{\sigma \in C_k} x_\sigma^{(j)}\right\}_{j,k=1}^{3}\right),$$

где S_3 - множество перестановок третьего порядка.

Эта теорема обобщается на случай произвольного конечного числа мер (теорема 9).

Таким образом, задача о нахождении $\bar{R}$ сводится к задаче о нахождении супремума по подмножеству $\mathbb{R}^{54}$, а задачу о нахождении $\underline{R}$ удается свести к задаче о нахождении минимакса в $\mathbb{R}^{18}$. Для непрерывных кусочно-линейных функций f обе полученные формулы сводят задачу о нахождении $\bar{R}$ и $\underline{R}$ к задачам линейного программирования.

Кроме того, в первой главе диссертации доказана следующая теорема.
Теорема 5. *Пусть Q_6, $G_6 \subset \mathbb{R}^6$ - множества, которые строятся по N_6 и N_{18} и полное определение которых дано в диссертации. Пусть в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 18 непустых попарно не пересекающихся множеств. Для любого $\vec{z} = (z_1, z_2, z_3, z_4, z_5, z_6) \in Q_6 \bigcup G_6$ положим*

$$f^*(z_1, z_2, z_3, z_4, z_5, z_6) := f(z_1, z_2, 1 - z_1 - z_2, z_3, z_4, 1 - z_3 - z_4, z_5, z_6, 1 - z_5 - z_6).$$

Тогда

$$\bar{R} = \sup_{\vec{z} \in Q_6} f^*(z_1, z_2, z_3, z_4, z_5, z_6),$$

$$\underline{R} = \inf_{\vec{z} \in G_6} f^*(z_1, z_2, z_3, z_4, z_5, z_6).$$

Отметим, что в теореме 4 множества, по которым берется экстремум (т.е.

N_6 и N_{18}), устроены в некотором смысле проще, чем множества G_6 и Q_6 из теоремы 5, однако последние, в отличие от первых, лежат в пространстве меньшей размерности.

Кроме того, доказано, что если в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере 18 непустых попарно не пересекающихся множеств, то экстремум в задаче о нахождении чисел $\underline{R}$ и $\bar{R}$ достигается на множестве дискретных мер.

Для случая произвольного $n \geq 3$ в первой главе получены следующие результаты.

Теорема 7. *Пусть в σ-алгебре измеримого пространства (Ω, Σ) содержится по крайней мере $n!$ непустых попарно не пересекающихся множеств. Тогда*

$$1 + \frac{2 \cdot \sum_{i<j} r_{ij}}{n(n-1)} \leq \underline{R}^{+} \leq \bar{R}^{+} \leq 1 + \frac{2 \cdot \sum_{i<j} r_{ij}}{n}.$$

Отметим, что верхняя оценка для $\bar{R}^{+}$ совпадает с верхней оценкой для экстремума по множеству рандомизированных критериев, которая установлена в теореме 3 статьи [3], но слабее оценки $\bar{R}^{+} \leq 1 + \min_i \sum_j r_{ij}$, которую легко получить из теоремы 1 (статьи [3]).

Установлена следующая теорема.

Теорема 8. *Пусть $\delta := \left(\frac{n}{2} - 1\right)(1 + \max_{i<j} r_{ij}) + 1 + \frac{(\max_{i<j} r_{ij} - 1)}{2} \cdot (n \mod 2)$. Для чисел $\underline{R}^{+}$, $\underline{R}^{\Delta}$, $\underline{R}^{\min}$, $\bar{R}^{+}$, $\bar{R}^{\Delta}$ и $\bar{R}^{\min}$ выполняются следующие соотношения:*

$$\underline{R}^{+} - \delta \leq \underline{R}^{\min} \leq \frac{\underline{R}^{+}}{n}, \quad \bar{R}^{+} - \delta \leq \bar{R}^{\min} \leq \frac{\bar{R}^{+}}{n},$$
$$2\underline{R}^{+} - 1 - 2\delta \leq \underline{R}^{\Delta} \leq \frac{\underline{R}^{+} - 1}{n-1}, \quad 2\bar{R}^{+} - 1 - 2\delta \leq \bar{R}^{\Delta} \leq \frac{\bar{R}^{+} - 1}{n-1},$$
$$2\underline{R}^{\min} - 1 \leq \underline{R}^{\Delta} \leq \frac{n\underline{R}^{\min} - 1}{n-1}, \quad 2\bar{R}^{\min} - 1 \leq \bar{R}^{\Delta} \leq \frac{n\bar{R}^{\min} - 1}{n-1}.$$

Далее, пусть $\mathbb{E} = (\Omega, \mathbb{F}, (P_\theta)_{\theta \in \Theta})$ — статистическая модель, причем $\Theta = \{1, 2, ..., n\}$ и пространство D решений совпадает с множеством Θ, функция потерь $L(\theta, d) = \mathbb{I}_{\theta \neq d}$, функция $\phi = (\phi_1, \phi_2, ..., \phi_n)$ задает рандомизированный критерий ($\phi_i : \Omega \to [0, 1]$ измеримы для каждого $1 \leq i \leq n$ и $\sum_{i=1}^{n} \phi_i = 1$) следующим образом: $\phi_i(\omega)$ - это вероятность принять решение под номером i, если происходит ω. Множество таких критериев обозначим через Φ, а множество нерандомизированных критериев обозначим через Φ^*. В «рандомизированном» случае байесовский риск p, соответствующий равномерному распределению на Θ, равен $\inf_{\phi\in\Phi} \frac{1}{n}\sum_{i=1}^n \mathbf{E}_{P_i}(1-\phi_i)$. Удобнее рассматривать величину $\bar{a} := 1-p = \sup_{\phi\in\Phi} \frac{1}{n}\sum_{i=1}^n \mathbf{E}_{P_i}\phi_i$. Аналогом величины $\bar{a}$ в нерандомизированном случае является величина $\bar{a}^* = \sup_{\phi\in\Phi^*} \frac{1}{n}\sum_{i=1}^n \mathbf{E}_{P_i}\phi_i$. Неулучшаемые оценки величин $\bar{a}$ и $\bar{a}^*$ позволяет получить следующая теорема из главы 1.

Теорема 10. *Выполнены следующие равенства:*

$$\sup_{\vec{\mu}\in D(\vec{r})} \sup_{\phi\in\Phi} \sum_{i=1}^n \mathbf{E}_{P_i}\phi_i = \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\phi\in\Phi^*} \sum_{i=1}^n \mathbf{E}_{P_i}\phi_i = \bar{R}^+,$$
$$\inf_{\vec{\mu}\in D(\vec{r})} \sup_{\phi\in\Phi} \sum_{i=1}^n \mathbf{E}_{P_i}\phi_i = \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\phi\in\Phi^*} \sum_{i=1}^n \mathbf{E}_{P_i}\phi_i = \underline{R}^+.$$

Другими словами, установлено, что рандомизация не позволяет улучшить оптимальный нерандомизированный критерий, если качество критерия понимать как сумму вероятностей принятия правильных решений.

Во **второй главе** рассмотрено обобщение предыдущей задачи на случай «интервальных критериев».

Пусть M - множество вероятностных мер на измеримом пространстве (Ω, Σ). Для любых трех вероятностных мер μ_1, μ_2, μ_3 на (Ω, Σ) можно, как в главе 1, рассмотреть задачу различения трех гипотез H_i: наблюдение ξ имеет распределение μ_i, $i = 1, 2, 3$. Теперь мы рассмотрим «интервальные критерии». Каждому нерандомизированному интервальному критерию в этой задаче соответствуют измеримое разбиение $\vec{C} = \{C_1, C_2, ..., C_7\}$ пространства Ω ($C_i : \bigsqcup_{i=1}^7 C_i = \Omega$), где через C_1, C_2, C_3, $C_4 = C_{12}$, $C_5 = C_{13}$, $C_6 = C_{23}$, $C_7 = C_{123}$ обозначены множества, определяющие критерий: если $\omega \in C_i$, $1 \le i \le 3$, то принимается решение о том, что наблюдение не противоречит гипотезе H_i и противоречит остальным гипотезам (сокращенно запишем: $C_i = \{\omega : H_i\}$, $1 \le i \le 3$), аналогично $C_{ij} = \{\omega : H_i \text{ или } H_j\}$, $1 \le i < j \le 3$, $C_{123} = \{\omega : H_1 \text{ или } H_2 \text{ или } H_3\}$, т.е. при $\omega \in C_{ij}$ принимается решение о том, что наблюдение не противоречит гипотезам H_i и H_j и противоречит третьей гипотезе; при $\omega \in C_{123}$ принимается решение о том, что наблюдение не противоречит ни одной из трех гипотез.

Далее, через $\vec{\mu}$ обозначим тройку мер (μ_1, μ_2, μ_3), через $\vec{\mu}(\vec{C})$ обозначим упорядоченный набор $(\mu_1(C_1), \mu_1(C_2), ..., \mu_3(C_7))$. Будем считать, что качество критерия можно представить функцией $f(\vec{\mu}(\vec{C}))$ со значениями в $\mathbb{R}$; иногда будет использоваться также обозначение $f(\{\mu_j(C_k)\}_{1\le j\le 3}^{1\le k\le 7})$. Фиксируем такие числа r_{ij}, $1 \le i < j \le 3$, что $0 \le r_{ij} \le 1$, $1 \le i < j \le 3$, $r_{12}+r_{13} \ge r_{23}$, $r_{12}+r_{23} \ge r_{13}$ и $r_{13}+r_{23} \ge r_{12}$. Рассмотрим семейство всевозможных троек вероятностных мер на Σ, между которыми заданы попарные расстояния по вариации следующим образом: $\rho(\mu_i, \mu_j) = r_{ij}$, $1 \le i < j \le 3$, где

$$\rho(\mu_i, \mu_j) := \sup_{A\in\Sigma} |\mu_i(A) - \mu_j(A)|.$$

Будем называть такие тройки мер $\vec{r}$-согласованными и обозначать это следующим образом: $\vec{\mu} \in D(\vec{r})$. Далее, для фиксированной тройки мер, числа α и функций $f, g : [0,1]^{21} \to \mathbb{R}$ обозначим через $\mathcal{E}_g^\alpha(\vec{\mu})$ класс таких разбиений $\vec{C} = \{C_1, C_2, ..., C_7\}$ пространства Ω, которые удовлетворяют условию $g(\vec{\mu}(\vec{C})) \ge \alpha$. Через $\underline{Ri}$ и $\overline{Ri}$ обозначим следующие величины:

$$\underline{Ri} = \inf_{\vec{\mu}\in D(\vec{r})} \inf_{\vec{C}\in\mathcal{E}_g^\alpha(\vec{\mu})} f(\vec{\mu}(\vec{C})), \quad \overline{Ri} = \sup_{\vec{\mu}\in D(\vec{r})} \inf_{\vec{C}\in\mathcal{E}_g^\alpha(\vec{\mu})} f(\vec{\mu}(\vec{C})).$$

При этом по определению полагаем, что инфимум по пустому множеству равен $+\infty$.

Теперь приведем примеры функций f и g, рассмотрение которых естественно. Для этого нам потребуются следующие определения:

$$C^{(1)} := C_1 \bigsqcup C_{12} \bigsqcup C_{13} \bigsqcup C_{123}, \quad C^{(2)} := C_2 \bigsqcup C_{12} \bigsqcup C_{23} \bigsqcup C_{123},$$
$$C^{(3)} := C_3 \bigsqcup C_{13} \bigsqcup C_{23} \bigsqcup C_{123}.$$

Другими словами, $C^{(i)}$ - это такое множество, на котором не отвергается гипотеза H_i (и, быть может, еще некоторые гипотезы). Далее, рассмотрим случайную величину $\#C$ (число неотвергаемых гипотез), которая равна 1 на C_i, $1 \le i \le 3$, равна 2 на C_{ij}, $1 \le i < j \le 3$ и равна 3 на C_{123}. Пусть $g(\vec{\mu}(\vec{C})) = \min_{1\le i\le 3}(\mu_i(C^{(i)}))$ и $f = \max_i(\mathbf{E}_i(\#C))$ (где символ $\mathbf{E}_i$ означает математическое ожидание по i-й мере). Тогда условие $g = \min_{1\le i\le 3}(\mu_i(C^{(i)})) \ge \alpha$ означает, что правильная гипотеза гарантированно принимается с вероятностью, не меньшей α (заданный уровень значимости), а минимизация f означает, что минимизируется максимум средних чисел принятых гипотез. Наряду с указанной парой функций f и g естественно рассмотреть пару функций $g = \sum_i \mu_i(C^{(i)})$, $f = \sum_i \mathbf{E}_i(\#C)$.

Итак, задача, решаемая во второй главе диссертации, состоит в сле-

дующем: для заданных чисел r_{ij}, $1 \le i < j \le 3$, числа α и функций $f, g : [0,1]^{21} \to \mathbb{R}$ требуется получить удобные формулы для величин $\underline{Ri}$ и $\overline{Ri}$.
Отметим, что величины

$$\underline{Ri} = \inf_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{E}_g^\alpha(\vec{\mu})} f(\vec{\mu}(\vec{C})) \quad \text{и} \quad \overline{Ri} = \sup_{\vec{\mu}\in D(\vec{r})} \sup_{\vec{C}\in\mathcal{E}_g^\alpha(\vec{\mu})} f(\vec{\mu}(\vec{C}))$$

сразу получаются из величин $\underline{Ri}$ и $\overline{Ri}$ за счет замены f на $-f$.

Во второй главе диссертации установлено, что в случае, когда функции f и g являются кусочно-линейными, данная задача сводится к стандартной задаче линейного программирования. Получено обобщение данного результата на случай произвольного конечного количества мер.

В **третьей главе** изучается случайный процесс, связанный с последовательным критерием χ^2 (см. [10]).

Рассмотрим схему независимых испытаний с $N \geqslant 2$ исходами. Будем считать, что вероятность j-го исхода равна $p_j > 0, j = 1, \dots, N$. Для $t = 1, 2, \dots$ и $j \in \{1, \dots, N\}$ введем индикаторы $I_j(t)$: $I_j(t) = 1$, если при t-м испытании осуществился j-й исход, и $I_j(t) = 0$ в противном случае. Тогда $\sum_{t=1}^n I_j(t)$ — частота исхода j в первых n испытаниях. Выражение

$$X(n) := \sum_{j=1}^N \frac{1}{np_j} \left(\sum\nolimits_{t=1}^n I_j(t) - np_j\right)^2$$

называется статистикой Пирсона. Рассмотрим случайный процесс с дискретным временем $X(n)$, $n \ge 1$. Установлена следующая теорема.
Теорема 15. *Если* $1 \leqslant s \leqslant t$, *то*

$$\mathbf{E}X(s) = N-1,\ \operatorname{cov}(X(s), X(t)) = 2(N-1)\frac{s}{t} + \frac{1}{t}\left(2 - 2N - N^2 + \sum_{k=1}^N \frac{1}{p_k}\right).$$

В частности, если $s, t \to \infty$ *так, что* $\frac{t}{s} \to C \geqslant 1$, *то*

$$\operatorname{cov}(X(s), X(t)) = \frac{2}{C}(N-1) + O(t^{-1}).$$

В силу теоремы 15 дисперсия $\mathbf{D}X(t) = 2(N-1) + \frac{1}{t}\cdot(2 - 2N - N^2 + \sum_{k=1}^N \frac{1}{p_k})$; это выражение для дисперсии хорошо известно, см., например, [14–16].

[15] Крамер Г., Математические методы статистики. М.: МИР, 1975, 648 с.
[16] Боровков А.А., Математическая статистика. СПб.: Лань, 2010, 704 стр.

Следствие 2. *Пусть $s \geqslant 1$ фиксировано. Тогда $\operatorname{cov}(X(s), X(t))$ строго монотонно убывает по t при $t \geqslant s$, за исключением случая $s = 1$, $p_1 = \ldots = p_N = \frac{1}{N}$: тогда $\operatorname{cov}(X(1), X(t)) = 0$ при всех $t \geqslant 1$.*

Пусть $W_1(t), W_2(t), \ldots$ — независимые стандартные винеровские процессы, т. е. независимые процессы с независимыми приращениями, $W_k(0) = 0$ и приращение $W_k(t+s) - W_k(t)$ имеет нормальное распределение со средним 0 и дисперсией s при любых $k = 1, 2, \ldots$ и $t, s \geqslant 0$. При любом целом $M > 0$ случайный процесс $W^{(M)}(t) = (W_1(t), \ldots, W_M(t))$ является стандартным винеровским M-мерным процессом. Случайный процесс

$$\mathrm{Bes}_M(t) = \|W^{(M)}(t)\|_2 = \left(\sum\nolimits_{k=1}^{M} W_k^2(t)\right)^{1/2}$$

называется M-мерным процессом Бесселя; очевидно, $\mathrm{Bes}_M^2(t) = \sum_{k=1}^{M} W_k^2(t)$ и поэтому $\frac{1}{t}\,\mathrm{Bes}_M^2(t)$ имеет распределение хи-квадрат с M степенями свободы. В третьей главе диссертации установлена следующая теорема.

Теорема 16. *Последовательность случайных процессов $X(\lfloor nt \rfloor)$, $t > 0$, при $n \to \infty$ сходится в смысле конечномерных распределений к процессу $\frac{1}{t}\,\mathrm{Bes}_{N-1}^2(t)$.*

Теорема 16, очевидно, является естественным обобщением общеизвестной теоремы о предельном распределении статистики Пирсона.

Кроме того, установлено, что теорему 16 можно переформулировать следующим образом.

Теорема 16а. *Конечномерные распределения случайных процессов $X(\lfloor ne^t \rfloor)$, $t \in \mathbb{R}$, при $n \to \infty$ сходятся к конечномерным распределениям стационарного в узком смысле процесса $e^{-t}\mathrm{Bes}_{N-1}^2(e^t)$.*

Тем самым доказано, что при надлежащей нормировке по времени конечномерные распределения процесса, образованного последовательными значениями статистики Пирсона для выборок увеличивающихся со временем объемов, сходятся к конечномерным распределениям стационарного случайного процесса — нормированного квадрата процесса Бесселя.

Далее, обозначим через $I_\delta(x)$ функцию Инфельда, т. е. модифицированную функцию Бесселя первого рода[17] (см. также [18], с. 139); она является

[17]Ватсон Г.Н., Теория бесселевых функций. М.: Изд. иностр. лит., 1949, 800 с.
[18]Лебедев Н.Н., Специальные функции и их приложения. М.: ФИЗМАТЛИТ, 1963, 359 с.

каноническим решением дифференциального уравнения

$$x^2\frac{d^2\omega}{dx^2} + x\frac{d\omega}{dx} - (x^2+\delta^2)\omega = 0,$$

разлагается в степенной ряд: $I_\delta(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+\delta}}{k!\Gamma(k+\delta+1)}$ и имеет различные интегральные представления. В третьей главе диссертации получены явные формулы для плотностей конечномерных распределений процесса Бесселя. А именно, верна следующая теорема.

Теорема 17. *При любом $k \geq 2$ и любых $0 < t_1 < ... < t_k$ совместное распределение* $(\mathrm{Bes}_N(t_1), \ldots, \mathrm{Bes}_N(t_k))$ *имеет плотность распределения*

$$b_{t_1,\ldots,t_k}(x_1,\ldots,x_k) =$$
$$= \frac{x_1}{t_1}\left(\frac{x_1^2 x_k^2}{4t_1^2}\right)^{\frac{N-2}{4}} \frac{e^{\frac{x_2^2}{2t_2}+\ldots+\frac{x_{k-1}^2}{2t_{k-1}}}}{\Gamma(\frac{N}{2})} \prod_{i=1}^{k-1} \frac{x_{i+1}}{t_{i+1}-t_i} \exp\left\{-\frac{\frac{x_i^2}{2t_i}+\frac{x_{i+1}^2}{2t_{i+1}}}{1-\frac{t_i}{t_{i+1}}}\right\} I_{\frac{N}{2}-1}\left(\frac{x_i x_{i+1}}{t_{i+1}-t_i}\right)$$

при $x_1,\ldots,x_k \geqslant 0$; $b_{t_1,\ldots,t_k}(x_1,\ldots,x_k) = 0$, если $\min\{x_1,\ldots,x_k\} < 0$.

Доказательство теоремы 17 опирается на теорему 16 и результаты статьи [10].

Заключение. Основные результаты диссертации состоят в следующем:

- В задаче различения $n \geq 3$ гипотез с заданными расстояниями по вариации между распределениями наблюдений:
  А) показано, что нахождение экстремальных значений произвольных кусочно-линейных непрерывных функций от вероятностей ошибок сводится к решению конечномерной задачи линейного программирования;
  Б) для экстремальных значений нескольких конкретных функций от вероятностей ошибок найдены точные формулы (при $n = 3$) и оценки (при $n \geq 3$).

- Доказано, что конечномерные распределения последовательности значений статистики Пирсона, построенных по выборкам растущих объемов при надлежащей нормировке по времени сходятся к конечномерным распределениям стационарного случайного процесса (нормированного квадрата процесса Бесселя). Опубликованные в 1969 г. формулы для предельных совместных распределений статистики Пирсона использованы при выводе явных формул для плотности совместных распределений процесса Бесселя.

Благодарности. Автор глубоко благодарен своему научному руководителю д.ф.-м.н. Андрею Михайловичу Зубкову за постановку интересных задач, плодотворные обсуждения, ценные замечания и постоянное внимание к работе.

Публикации автора по теме диссертации

[1] Зубков А.М., Савелов М.П., Сходимость последовательности значений статистики Пирсона к квадрату нормированного процесса Бесселя, Дискрет. матем., **28**:3 (2016), 49-58. [Постановка задач, а также редактирование окончательного варианта текста принадлежит А.М. Зубкову. Основные результаты получены М.П. Савеловым.]

[2] Савелов М.П., Экстремальные характеристики критериев выбора гипотез с заданными попарными расстояниями по вариации, Теория вероятн. и ее примен., **61**:3 (2016), 440-464.

[3] Savelov M.P., Extremal Problems for Hypotheses Testing with Set-Valued Decisions, Mathematical Methods of Statistics, **25**:1 (2016), 67-77.

[4] Savelov M.P., On the Sequential Chi-Square Test, Computer Data Analysis and Modeling: Theoretical and applied stochastics: Proc. of the Eleventh Intern. Conf., Minsk, Sept. 6-10, 2016, Publishing center of BSU, Minsk, 105-106.