

ОТЗЫВ
официального оппонента на диссертационную работу
Савелова Максима Павловича
«Экстремальные характеристики критериев выбора
статистических гипотез»,
представленную на соискание учёной степени
кандидата физико-математических наук
по специальности 01.01.05 –
«Теория вероятностей и математическая статистика»

В диссертационной работе М.П. Савелова «Экстремальные характеристики критериев выбора статистических гипотез» рассматриваются задачи, связанные с построением статистических критериев различения $n \geq 3$ простых гипотез и доказательством их оптимальности в том или ином смысле (относительно некоторого функционала качества). Рассматриваемая задача различения произвольного конечного числа гипотез неоднократно изучалась в работах отечественных и зарубежных специалистов, среди которых особо следует отметить А.А. Ляпунова (1951 г.), К.Ш. Зигангирова (1977 г.), А.В. Иванова (1999 г.), К. Фукунага (1972 г.) и ряд других. В частности, в классическом учебнике А.А. Боровкова «Математическая статистика» (2010 г.) подробно рассмотрены байесовский и минимаксный подходы к построению соответствующих оптимальных критериев. В диссертации автором предложены несколько показателей оптимальности, в соответствии с которыми и отыскиваются оптимальные критерии.

Следует отметить, что решение задач подобного типа имеет большое теоретическое и практическое значение – например, для таких областей естествознания, как астрономия, теория связи (в частности, можно упомянуть монографию А.А. Харкевича «Борьба с помехами» (1963 г.)), статистическая радиотехника (см., например, монографию Б.Р. Левина «Теоретические основы статистической радиотехники» (1989 г.)), информационной безопасности и ряд других направлений. Задачи,

связанные с построением оптимальных критериев различения конечного числа гипотез, также интересны с теоретической точки зрения.

Вторая задача, решённая в диссертации, связана с изучением асимптотических распределений случайного процесса, определяемого статистикой последовательного r -кратного критерия хи-квадрат. Указанный критерий неоднократно изучался ранее в литературе (см., например, работы В.К. Захаров, О.В. Сарманов, Б.А. Севастьянов (1969 г.), А.П. Гермогенов, А.Ф. Ронжин (1984 г.), А.П. Гермогенов (1985 г.), Г.И. Ивченко, Б.А. Севастьянов (1986 г.), М. Мурвалиев (1988 г.), Б.И. Селиванов, В.П. Чистяков (1997, 1998 гг.)) и ряд других. Можно также упомянуть зарубежные работы по указанной и близкой к ней проблематике – например, G. Lindgren (1980 г.), G. Neuhaus, E. Cremer (1981 г.), J.H. Steiger, A. Shapiro, M.E. Browne (1985 г.) и ряд других. Достаточно объёмный пласт литературы по упомянутому направлению свидетельствует об определённом интересе к указанной области, что вызвано применением соответствующих результатов при решении ряда практических задач. В качестве наглядного примера можно упомянуть так называемую «задачу о разладке», в которой требуется по известным наблюдениям восстановить момент смены распределений (см., например, упомянутую работу А.Ф. Ронжина (1984 г.)).

Диссертация состоит из введения, трёх глав, разбитых на параграфы, заключения и списка литературы, который содержит 24 библиографические ссылки. Объём диссертации составляет 92 страницы. Проведём краткий анализ диссертации и полученных в ней результатов.

В первой главе рассмотрена задача различения $n \geq 3$ простых гипотез. Получен ряд оригинальных результатов об экстремальных значениях некоторых характеристик нерандомизированных статистических критериев для случая, когда известны попарные расстояния по вариации между распределениями, определяемыми соответствующими гипотезами. Основной результат первой главы состоит в том, что автору удалось свести исходную задачу поиска экстремальных значений кусочно-линейных непрерывных

функций от вероятностей ошибок к решению конечномерной задачи линейного программирования. Полученные результаты подробно иллюстрируются случаем трёх гипотез, для которого получены точные формулы. Для случая $n \geq 4$ гипотез получены оценки экстремальных значений конкретных функций от вероятностей ошибок. Одновременно доказано, что для любого $n \geq 2$ существуют непрерывные функционалы f качества критериев, для которых решение экстремальной задачи в классе безатомных мер качественно отличается от решения аналогичной задачи в условиях наличия атомов у соответствующих мер. Также показано, что рандомизация не позволяет улучшить характеристики оптимального нерандомизированного критерия в случае, когда качество критерия измеряется суммой вероятностей принятия правильных решений, или надёжностью критерия.

Во второй главе автором исследованы экстремальные значения интервальных нерандомизированных критериев различения $n \geq 3$ простых гипотез в условиях, когда известны попарные расстояния по вариации между распределениями, определяемыми соответствующими гипотезами. Автору удалось свести решение задачи о нахождении экстремальных значений соответствующих характеристик критериев к решению конечномерной задачи линейного программирования. Подробно исследовав случай $n = 3$ и получив содержательные результаты, автору удалось распространить их на случай произвольного $n \geq 4$.

В третьей главе рассмотрены вопросы, связанные с исследованием асимптотических распределений процесса, связанного с последовательным r -кратным критерием хи-квадрат. Найдены предельные распределения случайного процесса, определяемого r -мерным центральным хи-квадрат распределением. Автором доказан факт сходимости конечномерных распределений процесса, порождённого последовательными значениями статистики хи-квадрат, к конечномерным распределениям стационарного случайного процесса, который является нормированным квадратом процесса

Бесселя. Получены явные формулы плотности совместных распределений процесса Бесселя.

К изложению материала диссертации имеется ряд замечаний, основные из которых приведены ниже.

1) При введении обозначений необходимо пояснять их смысл. Так, на стр. 4 и далее по тексту автор использует обозначение $\prod_{i=1}^3 C_i$ для объединения непересекающихся множеств. Обычно операция объединения непересекающихся множеств обозначается в литературе символом «+» (в данном случае – $\sum_{i=1}^3 C_i$).

2) На стр. 4 и далее по тексту (например, стр. 13, 19) встречается фраза «будем считать, что качество критерия можно представить функцией $f(\bar{\mu}(\bar{C})) \dots$ ». Представляется, что данное утверждение должно быть строго математически обосновано.

3) На стр. 5 вместо обозначения $E_i \phi_i$ для сокращения записи можно использовать обозначение $E_i \phi_i$.

4) На стр. 10 утверждение теоремы 2 можно записать в виде $\bar{R}^{\Delta} = r^{(0)}, \bar{R}^+ = 2r^{(0)} + 1$. Аналогичное замечание относится и к Теореме 2 на стр. 21.

5) Вообще говоря, понятие «интервальный критерий» не является общеизвестным. По мнению рецензента, автору следовало бы, во-первых, сослаться на соответствующую литературу, в которой используются подобные конструкции, а во-вторых, более детально описать, как устроены такие критерии и как принимается решение о согласии с соответствующими гипотезами.

6) На стр. 13 и на стр. 63 фактически определены множества C_{ij} , по сути представляющие собой области, для элементов которых имеется согласие хотя бы с одной из гипотез из множества гипотез $\{H_i, H_j\}$. Однако ниже написано, что «при $\omega \in C_{ij}$ принимается решение о том, что

наблюдение не противоречит гипотезам H_i и H_j », что не вполне корректно, так как в этом случае подразумевается, что наблюдение одновременно не противоречит обоим гипотезам. Аналогичное замечание относится и к определению множества C_{123} .

7) Поскольку через Φ^* автор обозначает множество нерандомизированных критериев (см. стр. 12, 61), требует пояснений формула на стр. 13: $\bar{a}^* = \sup_{\phi \in \Phi^*} \frac{1}{n} \sum_{i=1}^n \mathbf{E}_{\mu_i} \phi_i$, так как в этом случае последняя сумма равна $\sum_{i=1}^n \mu_i(G_i)$, где через G_i обозначим область принятия гипотезы с номером i .

8) На стр. 16 и далее по тексту присутствует обозначение $[a]$. По всей видимости, автор обозначает так наибольшее целое число b , не превосходящее a . Однако далее для того же самого числа используется обозначение $[a]$ (см., например, стр. 86: $[nt_i]$). Необходимо использовать единообразные обозначения.

9) На стр. 23 и далее по тексту необходимо сослаться на упомянутую теорему Хана (например, см. Теорему Жордана-Хана о разложении на стр. 96 в книге М. Лозв, Теория вероятностей. – М.: Издательство иностранной литературы, 1962 г., 720 стр.).

10) В третьей главе (см. стр. 79, 80, 84, 86) автор использует обозначения cov и Cov . С одной стороны (см. стр. 79), через $\text{cov}(X, Y)$ обозначена обычная ковариация случайных величин X, Y . Однако на стр. 80 та же ковариация обозначена через Cov . В связи с тем, что на стр. 86 через $\text{Cov}(\bar{\Xi})$ автором обозначена уже матрица ковариаций соответствующего вектора, то Cov на стр. 80 также может быть воспринята как матрица ковариаций вектора. Необходимо привести обозначения к единообразию.

11) На стр. 80 в формуле для $Cov\left(\frac{1}{\sqrt{t}}W_j(t), \frac{1}{\sqrt{s}}W_k(t)\right)$ не определено, чему она равна в случае $s < t$, хотя понятно, что $Cov\left(\frac{1}{\sqrt{t}}W_j(t), \frac{1}{\sqrt{s}}W_k(t)\right) = Cov\left(\frac{1}{\sqrt{s}}W_k(t), \frac{1}{\sqrt{t}}W_j(t)\right)$.

12) При ссылке в тексте диссертации на конкретные теоремы, положения или результаты следует указывать не только источник, но и страницу, на которой они приведены. Так, например, на стр. 79 указано «(см., например, [2], [6], стр. 139)», откуда не совсем понятно, к какому из источников относится стр. 139. На той же странице ниже указаны ссылки [5], [9], [1], однако номера страниц не указаны.

13) В списке литературы следовало бы привести больше работ, связанных с построением критериев различения нескольких простых гипотез, а также с последовательным r -кратным критерием хи-квадрат.

14) В списке литературы желательно указывать номера страниц публикаций (см., например, 12), 13), 17), 22)).

Отметим, однако, что указанные недочёты в основном относятся к оформлению и не влияют на теоретическое содержание диссертации и суть полученных в ней результатов. Автореферат верно и со всей полнотой отражает содержание диссертации. Диссертация преимущественно носит теоретический характер. Основные утверждения и положения, выносимые на защиту, строго математически обоснованы и доказаны автором лично. Они имеют существенное значение для развития соответствующих областей теоретической математической статистики, а также могут найти применение в ряде практических приложений. Основные результаты диссертации опубликованы в четырёх печатных работах, три из которых присутствуют в перечне, рекомендованном ВАК Российской Федерации.

Считаю, что содержание и оформление диссертации соответствуют пп. 9-14 «Положения о присуждении учёных степеней», утверждённого Постановлением Правительства Российской Федерации № 842 от 24 сентября 2013 года.

На основании вышеизложенного считаю, что диссертация М.П. Савелова «Экстремальные характеристики критериев выбора статистических гипотез» удовлетворяет требованиям «Положения о присуждении учёных степеней», утверждённого Постановлением Правительства Российской Федерации № 842 от 24 сентября 2013 года, её содержание соответствует специальности 01.01.05 – «Теория вероятностей и математическая статистика», а её автор, Савелов Максим Павлович, заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 01.01.05 – «Теория вероятностей и математическая статистика».

Официальный оппонент:

кандидат физико-математических наук, 01.01.05 – «Теория вероятностей и математическая статистика», доцент по кафедре информационной безопасности, сотрудник войсковой части 43753



(Иванов Андрей Викторович)
«29» ноября 2016 года

Почтовый адрес: 121351, г. Москва, войсковая часть 43753

Телефон: 8-(499)-144-75-04

Адрес электронной почты: sesquialtera@bk.ru

Подпись Иванова Андрея Викторовича
заверяю



Командир подразделения
войсковой части 43753

 Д.В. Матюхин