

Московский государственный университет им. М.В. Ломоносова
Факультет вычислительной математики и кибернетики

На правах рукописи

КОРОЛЕВ Роман Анатольевич

**О МОЩНОСТИ КРИТЕРИЯ ЗНАКОВ
В СЛУЧАЕ РАСПРЕДЕЛЕНИЯ ЛАПЛАСА**

01.01.05 — теория вероятностей и математическая статистика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2010

Работа выполнена на кафедре математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова.

Научный руководитель доктор физико-математических наук,
профессор
Бенинг Владимир Евгеньевич

Официальные оппоненты доктор физико-математических наук,
профессор
Зейфман Александр Израилевич

доктор физико-математических наук,
ведущий научный сотрудник
НИИ математики и механики
имени Н.Г. Чеботарева
Чупрунов Алексей Николаевич

Ведущая организация Институт проблем информатики
Российской академии наук (ИПИ РАН)

Защита состоится 24 декабря 2010 года в 11:00 ч. 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 ВМиК, аудитория 685.

С диссертацией можно ознакомиться в библиотеке факультета ВМиК МГУ. С текстом автореферата можно ознакомиться на официальном сайте факультета ВМиК МГУ <http://www.cmc.msu.ru> в разделе Наука — Работа диссертационных советов — Д 501.001.44.

Автореферат разослан 24 ноября 2010 года.

Ученый секретарь
диссертационного совета
профессор

Н.П. Трифонов

Общая характеристика работы

Актуальность темы. В 1774 году Пьер-Симон Лаплас в своей статье «Sur la probabilité des causes par les événements» (см. [1] и литературу там) предложил естественный вероятностный закон для ошибки измерений в такой формулировке: логарифм частоты ошибки есть линейная функция абсолютного значения ошибки. Назвав этот закон первым законом для ошибки измерений, который исторически является первым вероятностным распределением с неограниченным носителем, Лаплас уже через 4 года в своей фундаментальной работе «Théorie Analytique» (см. [1] и литературу там) рассматривает второй вероятностный закон, который гласит: логарифм частоты ошибки измерений есть квадратичная функция ошибки. Именно этот второй закон благодаря хорошим аналитическим свойствам будет детально исследоваться все последующее время, получит название «нормальное распределение» и займет главное место в теории вероятностей вследствие центральной предельной теоремы. Лишь спустя почти 150 лет известный экономист и математик Дж. Кейнс (см. [1] и литературу там) напомним о существовании первого закона для ошибки измерений и получит его вновь из предположения, что наиболее вероятное значение измеряемой величины есть ее медиана. Следом за ним известный математик Э. Уилсон (см. [1] и литературу там) с помощью непараметрических методов покажет на одном примере, что распределение отклонений от медианы измерений является скорее первым законом Лапласа, нежели нормальным законом. Спустя еще почти 50 лет в научной литературе (см. [1] и литературу там) все чаще стали появляться пожелания использовать первый закон Лапласа в качестве основного распределения для экономических, биометрических и демографических данных в противовес нормальному распределению.

В наши дни первый закон Лапласа называют распределением Лапласа. Это распределение задается характеристической функцией (см. обзор в

¹Kotz S., Kozubowski T. J., Podgórski K. The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance. — Birkhäuser Boston, USA, 2001.

работе [2] и литературу там)

$$f(s) = \frac{2}{2 + \sigma^2 s^2}, \quad s \in \mathbb{R}^1,$$

или плотностью

$$l(x) = \frac{1}{\sigma\sqrt{2}} \exp\left\{-\frac{\sqrt{2}|x|}{\sigma}\right\}, \quad \sigma > 0, \quad x \in \mathbb{R}^1.$$

Другое название — двойное экспоненциальное распределение — указывает на возможность получения его как разности двух независимых одинаково распределенных экспоненциальных величин, которые часто используются для описания продолжительности жизни наблюдаемых объектов. Недавно В.Е. Бенингом и В.Ю. Королёвым было показано (см. работу [2]), что распределение Лапласа естественно возникает как предельное распределение асимптотически нормальных статистик в случае выборок случайного объема.

В прикладных областях экономики и науки популярность распределения Лапласа как математической (вероятностной модели) обусловлена тем, что его хвосты тяжелее, чем у нормального распределения. Так в теории связи в задачах обнаружения постоянного сигнала в качестве вероятностной модели некоторых типов импульсных помех выбирают распределение Лапласа (см. [3], [4], [5]). В работе [6] распределение Лапласа рассматривается как модель для речевого сигнала в задачах кодирования и декодирования аналоговых сигналов. Использованию распределения Лапласа в задачах на разрушение

² Бенинг В. Е., Королёв В. Ю. Некоторые статистические задачи, связанные с распределением Лапласа. // Информатика и её применения, 2008, т. 2, вып. 2, с. 19–34.

³ Dadi M. I., Marks R. J. II. Detector relative efficiencies in the presence of Laplace noise. // IEEE Trans. Aerospace Electron. Systems, 1987, AES-23(4), p. 568–582.

⁴ Marks R. J., Wise G. L., Haldeman D. G., Whited J. L. Detection in Laplace noise. // IEEE Trans. Aerospace Electron. Systems, 1978, AES-14(6), p. 866–871.

⁵ Miller J. H., Thomas J. B. Detectors for discrete-time signals in non-Gaussian noise. // IEEE Trans. Inform. Theory, 1972, IT-18(2), p. 241–250.

⁶ Duttweiler D. L., Messerschmitt D. G. Nearly instantaneous companding for nonuniformly quantized PCM. // IEEE Trans. Comm., 1976, COM-24(8), p. 864–873.

устройств и излом материалов посвящена работа [7]. Работы [8] [9] обсуждают применение распределения Лапласа в аэродинамике, где градиент скорости ветра по отношению к периоду времени моделируется с помощью смесей распределения Лапласа и нормального распределения, распределение ошибки в навигации с использованием распределения Лапласа исследуется в [10].

Взросший интерес к распределению Лапласа делает актуальным его использование в математической статистике. Нерегулярность распределения Лапласа порождает известные трудности при использовании его в задачах проверки статистических гипотез. Однако развитые в последние десятилетия асимптотические методы проверки статистических гипотез (см. [11], [12], [13], [14] и литературу там) позволяют теперь решать многие из подобных задач. Одна из них — задача сравнения мощности асимптотически наиболее мощного критерия с мощностью наилучшего критерия — рассматривается в диссертации.

Пусть θ принадлежит открытому множеству $\Theta \subset \mathbb{R}^1$, содержащему ноль, $n \in \mathbb{N}$. Рассмотрим задачу проверки простой гипотезы

$$H_0 : \theta = 0 \quad (1)$$

против последовательности сложных близких альтернатив вида

$$H_{n,1} : \theta = \frac{t}{\sqrt{n}}, \quad 0 < t \leq C, \quad C > 0, \quad (2)$$

⁷ Epstein B. Application of the theory of extreme values in fracture problems. // J. Amer. Statist. Assoc., 1948, vol. 43, No. 243, p. 403–412.

⁸ Jones P. N. McLachlan G. J. Laplace-normal mixtures fitted to wind shear data. // J. Appl. Statistics, 1990, vol. 17, No. 2, p. 271–276.

⁹ Kanji G. K. A mixture model for wind shear data. // J. Appl. Statistics, 1985, vol. 12, No. 1, p. 49–58.

¹⁰ Hsu D. A. Long-tailed distributions for position errors in navigation. // Appl. Statist., 1979, vol. 28, No. 1, p. 62–72.

¹¹ Чибисов Д. М. Вычисление дефекта асимптотически эффективных критериев. // Теория вероятн. и ее примен., 1985, т. 30, вып. 2, с. 269–288.

¹² Bening V. E. Asymptotic Theory of Testing Statistical Hypotheses. — VSP, Utrecht, 2000.

¹³ Chibisov D. M. An asymptotic expansion for distributions of $C(\alpha)$ test statistics. // Lecture Notes in Statistics, 1980. vol. 2. p. 63–96.

¹⁴ Chibisov D. M., Van Zwet W. R. On the Edgeworth Expansion for the Logarithm of the Likelihood Ratio. I. // Теория вероятн. и ее примен., 1984, т. 29, вып. 3, с. 417–439.

с неизвестным параметром t , на основе выборки $\mathbf{X}_n = (X_1, \dots, X_n)$, состоящей из независимых и одинаково распределенных наблюдений, имеющих плотность

$$p(x, \theta) = \frac{1}{2} e^{-|x - \theta|}, \quad x, \theta \in \mathbb{R}^1. \quad (3)$$

Распределение случайной величины X_1 будем обозначать P_θ , а n -кратное произведение таких распределений при гипотезе H_0 и альтернативе $H_{n,1}$ обозначим, соответственно, через $P_{n,0}$ и $P_{n,1}$.

Для любого фиксированного $t \in (0, C]$ согласно лемме Неймана-Пирсона наилучший критерий для проверки гипотезы H_0 (см. (1)) против простой альтернативы

$$H_{n,t} : \theta = \frac{t}{\sqrt{n}} \quad (4)$$

в случае распределения Лапласа (3) основан на логарифме отношения правдоподобия

$$\Lambda_n(t) = \sum_{i=1}^n (|X_i| - |X_i - tn^{-1/2}|). \quad (5)$$

Обозначим через $\beta_n^*(t)$ мощность такого критерия уровня $\alpha \in (0, 1)$. Заметим, что поскольку t неизвестно, то мы не можем использовать статистику $\Lambda_n(t)$ для построения критерия проверки гипотезы H_0 против альтернативы $H_{n,1}$. Однако $\beta_n^*(t)$, это так называемая огибающая функция мощности, дает верхнюю границу для мощности любого критерия при проверке гипотезы H_0 против фиксированной альтернативы $H_{n,t}$ (см. (4)), $t > 0$, и может служить стандартом при сравнении различных критериев. Рассмотрим критерий уровня $\alpha \in (0, 1)$, основанный на знаковой статистике вида

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i), \quad (6)$$

обозначим через $\beta_n(t)$ его мощность. Очевидно, статистика (6) не зависит от неизвестного параметра t и может быть использована для проверки гипотезы H_0 против альтернативы $H_{n,1}$. В диссертации вычисляется предел вида

$$r(t) \equiv \lim_{n \rightarrow \infty} \sqrt{n}(\beta_n^*(t) - \beta_n(t)). \quad (7)$$

Число $r(t)$ имеет статистическую интерпретацию в терминах дополнительного числа наблюдений, необходимых критерию для достижения той же мощности, что и наилучшего критерия. В работе [12] получена общая теорема, дающая достаточные условия для справедливости представления

$$r(t) = \frac{1}{2} e^{b_t} p(b_t) D(\Delta(t) | \Lambda(t) = b_t), \quad (8)$$

где $b_t = \Phi_1^{-1}(1 - \alpha)$, $\Phi_1(x)$ — функция распределения, предельная для распределений логарифмов отношения правдоподобия $\Lambda_n(t)$ при гипотезе H_0 , $p(x) = \Phi_1'(x)$, $(\Delta(t), \Lambda(t))$ — случайный вектор, предельный для $(\sqrt[n]{n}\Delta_n(t), \Lambda_n(t))$, $\Delta_n(t) = S_n(t) - \Lambda_n(t)$, $S_n(t)$ — монотонное преобразование статистики T_n , не меняющие мощности критерия, вида

$$S_n(t) = \frac{t}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i) - \frac{t^2}{2}.$$

Однако достаточные условия из этой теоремы не выполняются в случае распределения Лапласа. Так теорема 3.2.1 работы [12] не может непосредственно быть применена к случаю распределения Лапласа, поскольку достаточное условие 3 (ii) (см. с. 79 работы [12]) — аналог условия Крамера (C) — не выполняется для характеристической функции решетчатой статистики $S_n(t)$. Также в связи с нерегулярностью распределения Лапласа статистика $\Delta_n(t)$ допускает нерегулярное стохастическое разложение порядка $n^{-1/4}$ в отличие от случая $n^{-1/2}$ в формулировке теоремы 2.1 работы [11].

В диссертации доказана общая теорема, обобщающая условие теоремы из работы [12], и распределение Лапласа является частным случаем применимости этой теоремы.

Цель работы. Целью данной диссертации является вычисление предела (7) для нормированной разности мощностей асимптотически оптимального критерия знаков и асимптотически наилучшего критерия в случае распределения Лапласа. При этом получена общая теорема, частным случаем которой является распределение Лапласа. Получены также асимптотические разложения для распределений логарифма отношения

правдоподобия как при гипотезе, так и при альтернативах. Проведена численная аппроксимация для мощности критерия и для дефекта асимптотически оптимального критерия знаков в случае распределения Лапласа.

Научная новизна. Все полученные в диссертации результаты являются новыми. Вычислен предел нормированной разности мощностей асимптотически оптимального критерия знаков и асимптотически наилучшего критерия в задаче проверки простой гипотезы против последовательности односторонних сложных локальных альтернатив в случае распределения Лапласа, а также получен дефект критерия знаков. Получены общие достаточные условия для справедливости формулы для предела нормированной разности мощностей критериев, в случае когда статистика асимптотически оптимального критерия имеет решетчатое распределение. Распределение Лапласа при этом является частным случаем.

Методы исследования. В работе используются методы математического и функционального анализа, а также методы теории вероятностей, в частности, методы сходимости условных моментов и условных мер, зависящих от параметра (см. [11], [12], [13]), неравенство сглаживания для расстояния Леви между распределениями (см. [13], [15], [16], [17]), метод характеристических функций.

Теоретическая и практическая значимость. Результаты диссертации имеют теоретический характер и одновременно допускают практическое применение для построения асимптотически оптимальных критериев, проверки гипотез на основе наблюдений в случае распределения Лапласа и вычисления дефекта критериев.

¹⁵ Золотарев В. М. Оценки различия распределений в метрике Леви. // Труды Матем. ин-та им. В. А. Стеклова, 1971, т. СХІІ, с. 224–231.

¹⁶ Золотарев В. М., Сенатов В. В. Двусторонние оценки метрики Леви. // Теория вероятн. и ее примен., 1975, т. 20, вып. 2, с. 239–250.

¹⁷ Абрамов В. А. Оценка расстояния Леви-Прохорова. // Теория вероятн. и ее примен., 1976, т. 21, вып. 2, с. 406–410.

Апробация работы. Результаты диссертации докладывались на научно-исследовательском семинаре “Теория риска и смежные вопросы” на факультете ВМиК МГУ (руководители В.Е. Бенинг и В.Ю. Королёв), на семинаре Школы математических наук Пекинского университета (май 2009 г.), на XXVIII международном научном семинаре по проблемам устойчивости стохастических моделей (31 мая - 5 июня 2009 г., Закопане, Польша), на семинаре 32-й Финской летней школы теории вероятностей (июнь 2010 г.) и нашли свое отражение в трудах упомянутых семинаров и конференций.

Публикации. Основные результаты диссертации опубликованы в 6 статьях (1, 2, 3, 4, 5, 6) в журналах, входящих в список ВАК «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук».

Структура и объем диссертации. Диссертация состоит из введения, трех глав, приложения, содержащего 3 таблицы и 7 рисунков, и списка литературы, содержащего 54 наименования. Объем работы 131 страница.

Благодарности. Работа выполнена под руководством доктора физико-математических наук, профессора Бенинга Владимира Евгеньевича, которому автор выражает искреннюю благодарность.

Содержание работы

Введение освещает актуальность темы, описывает объекты исследования и основные результаты диссертации.

В первой главе формулируется задача проверки простой гипотезы об одномерном параметре сдвига против последовательности односторонних сложных локальных альтернатив в случае распределения Лапласа на основе независимых одинаково распределенных наблюдений. Эвристически выводятся формулы для предела нормированной разности мощностей асимптотически наилучшего и асимптотически оптимального критериев и для дефекта асимптотически оптимального критерия.

При этом порядок разности мощностей критериев оказывается $n^{-1/2}$ в случае распределения Лапласа в отличие от регулярного случая, когда этот порядок есть n^{-1} . Получены предельные теоремы для последовательностей функций распределения статистики, на которой основан критерий.

Рассмотрим снова задачу проверки простой гипотезы H_0 (см. (1)) против последовательности односторонних сложных локальных альтернатив $H_{n,1}$ (см. (2)), где параметр t неизвестен. Согласно фундаментальной лемме Неймана-Пирсона для любого фиксированного $t \in (0, C]$ существует наиболее мощный критерий для проверки простой гипотезы H_0 (см. (1)) против простой альтернативы $H_{n,t}$ (см. (4)), который основан на логарифме отношения правдоподобия

$$\Lambda_n(t) = \sum_{i=1}^n (l(X_i, tn^{-1/2}) - l(X_i, 0)), \quad (9)$$

где $l(x, \theta) = \log p(x, \theta)$, и отвергает гипотезу H_0 в случае, если

$$\Lambda_n(t) > c_{n,t},$$

причем критическое значение $c_{n,t}$ выбирается из условия (предположим непрерывность распределения $\Lambda_n(t)$)

$$P_{n,0}(\Lambda_n(t) > c_{n,t}) = \alpha,$$

где $\alpha \in (0, 1)$ – фиксированный уровень значимости. Мощность наилучшего критерия $\beta_n^*(t)$ вычисляется по формуле

$$\beta_n^*(t) = P_{n,t/\sqrt{n}}(\Lambda_n(t) > c_{n,t}).$$

Функция $\beta_n^*(t)$ называется огибающей функцией мощности, поскольку $\beta_n^*(t)$ дает верхнюю границу для мощности любого критерия при $t > 0$. Критерий, основанный на $\Lambda_n(t)$, нельзя применить для сложных альтернатив ввиду неизвестности параметра $t \in (0, C]$, в регулярном случае существует алгоритм построения критериев, не зависящих от t и имеющих ту же предельную мощность, что и $\beta_n^*(t)$, при $n \rightarrow \infty$. Этот подход основан на стохастическом разложении логарифма отношения правдоподобия $\Lambda_n(t)$

$$\Lambda_n(t) = tL_n^{(1)} - \frac{1}{2}t^2I + \frac{1}{2}\frac{t^2}{\sqrt{n}}L_n^{(2)} + \dots, \quad (10)$$

где

$$L_n^{(j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n (l^{(j)}(X_i) - E_0 l^{(j)}(X_i)), \quad j = 1, 2, \dots,$$

$$l^{(j)}(x) = \frac{\partial^j}{\partial \theta^j} l(x, \theta) \Big|_{\theta=0}, \quad j = 1, 2, \dots$$

и $I = E_0(l^{(1)}(X_1))^2$ — фишеровская информация. При этом в выражении (10) опущены неслучайный член $\frac{1}{6} \frac{t^3}{\sqrt{n}} E_0 l^{(3)}(X_1)$, члены более высокого порядка малости, чем $n^{-1/2}$, и использован хорошо известный факт, состоящий в том, что

$$E_0 l^{(1)}(X_1) = 0, \quad E_0 l^{(2)}(X_1) = -I.$$

Рассмотрим вместо критерия, основанного на $\Lambda_n(t)$, критерий, основанный на $L_n^{(1)}$ из выражения (10). Хорошо известно, что

$$\begin{aligned} \mathcal{L}(\Lambda_n(t) | H_0) &\rightarrow \mathcal{N}\left(-\frac{1}{2}t^2 I, t^2 I\right), \quad c_{n,t} \rightarrow c_t = t\sqrt{I}u_\alpha - \frac{1}{2}t^2 I, \\ \mathcal{L}(\Lambda_n(t) | H_{n,t}) &\rightarrow \mathcal{N}\left(\frac{1}{2}t^2 I, t^2 I\right), \quad \beta_n^*(t) \rightarrow \beta^*(t) = \Phi(t\sqrt{I} - u_\alpha), \\ \beta_n(t) &\rightarrow \beta^*(t) = \Phi(t\sqrt{I} - u_\alpha), \end{aligned} \quad (11)$$

где $\beta_n(t)$ — мощность критерия, основанного на статистике $L_n^{(1)}$, $\mathcal{L}(\Lambda_n(t) | H_0)$, $\mathcal{L}(\Lambda_n(t) | H_{n,t})$ — распределения $\Lambda_n(t)$ при соответствующих гипотезах, $\mathcal{N}$ — нормальный закон с соответствующими параметрами и $u_\alpha = \Phi^{-1}(1 - \alpha)$.

Критерии, для которых выполняется (11), называются асимптотически наиболее мощными (АНМ). Таковы, например, критерии основанные на статистиках $L_n^{(1)}$, $\Lambda_n(t_0)$, где $t_0 > 0$ фиксировано, оценках максимального правдоподобия и т.п. Для сравнения таких критериев можно использовать следующие члены асимптотического разложения $\beta_n(t)$, то есть представления типа

$$\beta_n(t) = \beta^*(t) + \frac{1}{\sqrt{n}} h_1(t) + \frac{1}{n} h_2(t) + \dots \quad (12)$$

Асимптотическим разложениям в статистике посвящены работы [18] и [19].

¹⁸ Pfanzagl J. Asymptotic expansions in parametric statistical theory. in: Developments in Statistics, Ed. by P.R. Krishnaiah, New York-London, Academic Press, 1980, vol. 3, p. 1–97.

¹⁹ Bickel P. J. Edgeworth expansions in nonparametric statistics. // Ann. of Statist., 1974, vol. 2, No. 1, p. 1–20.

Было замечено, что в формуле (12) при выполнении естественных условий регулярности для АНМ критериев совпадают и члены $h_1(t)$, различия наступают в членах порядка n^{-1} . Этим вопросам посвящены работы [11], [12], [18]. При этом величина

$$r(t) = \lim_{n \rightarrow \infty} n(\beta_n^*(t) - \beta_n(t))$$

допускает статистическую интерпретацию в терминах необходимого числа наблюдений и позволяет находить асимптотический дефект (см. [11], [12], [20].)

В работе [12] рассмотрен общий случай в терминах общего статистического эксперимента и приведена общая теорема, дающая достаточные условия для существования предела $r(t)$.

Теорема 1. (Bening, 2000) *Пусть выполнены условия регулярности (см. теорему 3.2.1 из работы [12]) и $0 < \alpha < 1$. Тогда*

$$r \equiv \lim_{n \rightarrow \infty} \tau_n^{-2}(\beta_n^* - \beta_n) = \frac{1}{2} e^b p(b) D(\Delta | \Lambda = b),$$

где $b = \Phi_1^{-1}(1 - \alpha)$, $\Phi_1(x)$ — функция распределения, предельная для логарифма отношения правдоподобия Λ_n при гипотезе H_0 (выше было $\Phi_1(x) = \Phi\left(\frac{x + \frac{1}{2}t^2 I}{t\sqrt{I}}\right)$), $p(x) = \Phi'_1(x)$ и $\tau_n \rightarrow 0$ — малый параметр (выше было $\tau_n = n^{-1/2}$), (Δ, Λ) — случайный вектор, предельный для $(\tau_n^{-1}\Delta_n, \Lambda_n)$, $\Delta_n = S_n - \Lambda_n$, S_n — монотонное преобразование статистики АНМ критерия T_n , не меняющее мощности критерия.

Рассмотрим теперь случай распределения Лапласа (см. (3)). Для этого семейства не выполняются естественные условия регулярности, поскольку у $p(x, \theta)$ (см. (3)) не существует производной по θ в точке $\theta = x$. Справедлива следующая

Теорема 2. *В случае распределения Лапласа (см. (3)) справедливы следующие соотношения: фишеровская информация равна 1 и*

$$\mathcal{L}(\Lambda_n(t) | H_0) \rightarrow \mathcal{N}\left(-\frac{t^2}{2}, t^2\right),$$

$$\mathcal{L}(\Lambda_n(t) | H_{n,t}) \rightarrow \mathcal{N}\left(\frac{t^2}{2}, t^2\right),$$

²⁰ Hodges J. L., Lehmann E. L. Deficiency. // Ann. Math. Statist., 1970, vol. 41, No. 5, p. 783–801.

и поэтому

$$\beta_n^*(t) \rightarrow \beta^*(t) = \Phi(t - u_\alpha), \quad n \rightarrow \infty.$$

Эта лемма показывает, что отсутствие дифференцируемости по θ функции $p(x, \theta)$ (см. (3)) в точке $\theta = x$ качественно не влияет на порядок альтернатив θ_n (равный $n^{-1/2}$) и вид предельной мощности $\beta^*(t)$. Однако стохастическое разложение для $\Lambda_n(t)$

$$\Lambda_n(t) = \frac{t}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i) + 2 \sum_{i=1}^n (X_i - tn^{-1/2}) \mathbf{1}_{[0, tn^{-1/2}]}(X_i). \quad (13)$$

содержит решетчатую статистику и оказывается нерегулярным. Действительно, рассмотрим следующее монотонное преобразование ($t > 0$) знаковой статистики T_n (см. (6))

$$S_n(t) = tT_n - \frac{1}{2}t^2,$$

тогда

$$\Delta_n(t) = S_n(t) - \Lambda_n(t) = -\frac{1}{2}t^2 - 2 \sum_{i=1}^n (X_i - tn^{-1/2}) \mathbf{1}_{[0, tn^{-1/2}]}(X_i),$$

и справедлива следующая

Теорема 3. Для распределения Лапласа (см. (3)) выполнены следующие соотношения

$$\mathcal{L}(\sqrt[4]{n}\Delta_n(t) \mid H_0) \rightarrow \mathcal{N}\left(0, \frac{2t^3}{3}\right),$$

$$\mathcal{L}((\sqrt[4]{n}\Delta_n(t), \Lambda_n(t)) \mid H_0) \rightarrow \mathcal{N}_2\left(0, \frac{2t^3}{3}, 0, -\frac{t^2}{2}, t^2\right),$$

где $\mathcal{N}_2$ — двумерный нормальный закон с соответствующими параметрами.

Как видно, порядок величины $\Delta_n(t)$ равен $\tau_n = n^{-1/4}$ в отличие от регулярного случая $n^{-1/2}$. Из этой теоремы также следует, что случайные величины $\sqrt[4]{n}\Delta_n(t)$ и $\Lambda_n(t)$ асимптотически независимы, и поэтому, применяя формально указанную выше теорему 1, имеем

$$r(t) = \lim_{n \rightarrow \infty} \sqrt{n}(\beta_n^*(t) - \beta_n(t)) = \frac{1}{2t} \varphi(u_\alpha - t) D(\Delta(t) \mid \Lambda(t) = c_t) =$$

$$= \frac{1}{2t} \varphi(u_\alpha - t) D\Delta(t) = \frac{t^2}{3} \varphi(u_\alpha - t), \quad (14)$$

где $\Lambda(t)$, $\Delta(t)$ независимые нормальные случайные величины соответственно с параметрами $(-\frac{t^2}{2}, t^2)$ и $(0, \frac{2t^3}{3})$. Строгое доказательство этой формулы содержится в главе 3.

Величина $r(t)$ тесно связана с асимптотическим дефектом. Найдём асимптотическое представление для дефекта d_n (см. [12], [20]) критерия, основанного на статистике T_n (см. (6)), который определяется как разность $k_n - n$, где k_n число наблюдений, необходимых критерию, основанному на статистике T_n для достижения той же мощности, что и критерий, основанный на $\Lambda_n(t)$, при одинаковых альтернативах $tn^{-1/2}$. Предполагая, что d_n непрерывная переменная, получаем равенство для её определения

$$\beta_n^*(t) = \beta_{k_n}(t\sqrt{k_n n^{-1}}). \quad (15)$$

Из теорем 2, 3 следует, что в случае распределения Лапласа для мощностей $\beta_n^*(t)$ и $\beta_n(t)$ справедливы представления

$$\beta_n^*(t) = \Phi(t - u_\alpha) + \frac{1}{\sqrt{n}} \varphi(t - u_\alpha) h^*(t) + o(n^{-1/2}),$$

$$\beta_n(t) = \Phi(t - u_\alpha) + \frac{1}{\sqrt{n}} \varphi(t - u_\alpha) h(t) + o(n^{-1/2}),$$

где $h^*(t)$ и $h(t)$ некоторые полиномы по t и u_α . Из этих соотношений и равенства (15) следует, что $k_n \rightarrow \infty$, $\frac{k_n}{n} \rightarrow 1$ и

$$\begin{aligned} d_n &= \sqrt{n} \frac{2(h^*(t) - h(t))}{t} + o(\sqrt{n}) = \\ &= \sqrt{n} \frac{2r(t)}{t\varphi(u_\alpha - t)} + o(\sqrt{n}) = \sqrt{n} \frac{2t}{3} + o(\sqrt{n}). \end{aligned}$$

Таким образом, в отличие от регулярного случая, в котором $d_n \rightarrow d < \infty$ (см. [11], [12]), то есть существует конечный асимптотический дефект, в случае распределения Лапласа дефект d_n стремится к бесконечности со скоростью $\sqrt{n}$.

Для критерия, основанного на статистике T_n , в первой главе доказана следующая

Теорема 4. В случае распределения Лапласа (см. (3)) для любого $0 \leq \delta < \frac{1}{2}$ справедливо соотношение

$$n^\delta(\beta_n^*(t) - \beta_n(t)) \rightarrow 0, \quad n \rightarrow \infty, \quad 0 < t \leq C, \quad C > 0.$$

Во второй главе доказана формула (14) для предела нормированной разности мощностей асимптотически оптимального критерия знаков и асимптотически наилучшего критерия в случае распределения Лапласа с помощью асимптотических разложений для мощностей критериев. Проведен численный анализ аппроксимации для мощностей обоих критериев. Для критериев численно исследована точность аппроксимации функций мощности их асимптотическими разложениями до порядка n^{-1} . Численно исследовано поведение дефекта критерия знаков. Результаты расчетов в виде таблиц и графиков приведены в приложении.

В следующей теореме строится асимптотическое разложение для мощности критерия, основанного на знаковой статистике T_n .

Теорема 5. Для мощности $\beta_n(t)$ равномерно по $t \in [0, C]$, $C \geq 0$, справедливо следующее асимптотическое разложение

$$\begin{aligned} \beta_n(t) = & \Phi(t - u_\alpha) - \frac{t^2}{2\sqrt{n}} \varphi(t - u_\alpha) + \\ & + \frac{t}{2n} \varphi(t - u_\alpha) \left(\frac{t^2}{3} - \frac{t^3}{4}(t - u_\alpha) - \right. \\ & \left. - \frac{1}{6}(u_\alpha^2 + u_\alpha t - 3t^2 + 24\varepsilon_n(1 - \varepsilon_n) - 3) \right) + \mathcal{O}(n^{-3/2}), \end{aligned}$$

где

$$\varepsilon_n = \frac{n+1}{2} + \frac{\sqrt{n}u_\alpha}{2} - \left[\frac{n+1}{2} + \frac{\sqrt{n}u_\alpha}{2} \right] + \mathcal{O}(n^{-1/2}),$$

где $[y]$ — целая часть числа $y \in \mathbb{R}^1$.

Получено также асимптотическое разложение для критерия, основанного на $\Lambda_n(t)$.

Теорема 6. Для мощности $\beta_n^*(t)$ критерия уровня $\alpha \in (0, 1)$, основанного на логарифме отношения правдоподобия $\Lambda_n(t)$ справедливо

$$\beta_n^*(t) = \Phi(t - u_\alpha) - \frac{t^2}{6\sqrt{n}}\varphi(t - u_\alpha) - \\ - \frac{t}{12n}\left(\frac{t^4}{6} - \frac{t^2}{3} - 1 - u_\alpha\left(\frac{t^3}{6} - u_\alpha + t\right)\right)\varphi(t - u_\alpha) + \mathcal{O}(n^{-3/2}).$$

Полученные асимптотические разложения для мощностей критериев (теоремы 5, 6) позволяют непосредственно вычислить предел их нормированной разности. Кроме того, данные разложения могут быть использованы для численной аппроксимации соответствующих мощностей.

Теорема 6 дает упрощенную формулу для вычисления мощности наилучшего критерия, и, как замечено в проведенных в диссертации численных расчетах, эта аппроксимация дает вполне приемлемое приближение даже для малых n . Численная аппроксимация была проведена также для мощности критерия знаковой статистики и для дефекта знаковой статистики. Приложение диссертации содержит графики и таблицы результатов численных расчетов.

В третьей главе в терминах общего статистического эксперимента сформулирована и доказана общая теорема, позволяющая получать предел нормированной разности мощностей асимптотически оптимальных критериев и наилучшего критерия. В отличие от предшествующих работ (см. [11], [12]) данная формула справедлива в случаях, когда порядок отклонения мощностей может отличаться от n^{-1} , и когда условие Крамера (C) или его аналог может не выполняться для характеристической функции статистики, лежащей в основе асимптотически эффективного критерия. В качестве примера применения данной теоремы в последнем параграфе диссертации рассматривается случай распределения Лапласа.

Теорема 7. *В случае распределения Лапласа (см. (3)) предел нормированной разности мощностей критериев, основанных на статистиках Λ_n (см. (5)) и T_n (см. (6)), в задаче проверки гипотезы H_0 (см. (1)) против альтернативы $H_{n,1}$ (см. (2)) для заданного уровня значимости $\alpha \in (0, 1)$ вычисляется по формуле*

$$\lim_{n \rightarrow \infty} \sqrt{n}(\beta_n^*(t) - \beta_n(t)) = \frac{1}{2}e^{b_t}D[\Delta(t) \mid \Lambda(t) = b_t]p(b_t) = \frac{t^2}{3}\varphi(u_\alpha - t),$$

$$(\Delta(t), \Lambda(t)) \sim \mathcal{N}_2\left(0, \frac{2t^3}{3}, 0, -\frac{t^2}{2}, t^2\right),$$

$b_t = tu_\alpha - t^2/2$, $p(x) \equiv \varphi(\frac{x+t^2/2}{t})/t$ и $\varphi(x)$ и u_α — плотность и квантиль уровня $1 - \alpha$ стандартного нормального распределения.

Приложение содержит таблицы и графики численных расчетов, проведенных в главе 2.

Основные результаты работы

1. Для случая распределения Лапласа вычислен предел нормированной разности мощностей асимптотически оптимального критерия, основанного на знаковой статистике, и асимптотически наилучшего критерия, основанного на логарифме отношения правдоподобия. Получены асимптотические разложения для функции распределения логарифма отношения правдоподобия при гипотезе и альтернативе.
2. В терминах общего статистического эксперимента сформулирована и доказана общая теорема о формуле для предела нормированной разности мощностей критериев. Эта теорема применена к случаю распределения Лапласа.
3. Получены численные оценки точности аппроксимации функций мощности критерия знаков и асимптотически наилучшего критерия в случае распределения Лапласа их асимптотическими разложениями до порядка n^{-1} , а также численные оценки для асимптотического дефекта критерия знаков. Построены таблицы для значений мощностей критериев и дефекта.

Список публикаций по теме диссертации

1. Королев Р. А., Тестова А. В., Бенинг В. Е. О мощности асимптотически оптимального критерия в случае распределения Лапласа. // Вестник Тверского государственного университета, серия Прикладная математика, 2008, вып. 8, № 4(64), с. 5–23.
2. Королев Р. А., Бенинг В. Е. Асимптотические разложения для мощностей критериев в случае распределения Лапласа. // Вестник Тверского государственного университета, серия Прикладная математика, 2008, вып. 3(10), № 26(86), с. 97–107.
3. Королев Р. А. О численной аппроксимации мощностей критериев в случае распределения Лапласа. // Вестник Тверского государственного университета, серия Прикладная математика, 2009, вып. 1(12), № 8, с. 59–76.
4. Королев Р. А. Формула для предела нормированной разности мощностей критериев в случае распределения Лапласа. // Вестник Тверского государственного университета, серия Прикладная математика, 2010, вып. 1(16), № 9, с. 5–24.
5. Бенинг В. Е., Королев Р. А. О предельном поведении мощностей критериев в случае распределения Лапласа. // Информатика и её применения, 2010, т. 4, вып. 2, с. 63–74.
6. Korolev R. A., Bening V. E. On the power of an asymptotically optimal test for the case of Laplace distribution. // Banach Center Publications, 2010, V. 90, p. 27–38.