УДК 519.612:632.4

ЭВРИСТИЧЕСКИЙ ВЫВОД ФОРМУЛЫ ДЛЯ АСИМПТОТИЧЕСКОГО ДЕФЕКТА В ОДНОЙ ЗАДАЧЕ РЕГРЕССИОННОГО ТИПА

Алексеев Ю.Г.[1], **Бенинг В.Е.**[1,2]
[1]Московский государственный университет им. М.В. Ломоносова, г. Москва
[2]Институт проблем информатики Российской академии наук, г. Москва

Поступила в редакцию 06.04.2014, после переработки 20.04.2014.

В статье развивается подход и идеи из работ [1], [5]. На эвристическом уровне получена формула для асимптотического дефекта в задаче регрессионного типа для случая, когда асимптотически эффективный критерий основан на первом члене стохастического разложения логарифма отношения правдоподобия. При этом показано, что разность между мощностью этого критерия и наилучшего критерия может стремиться к нулю с достаточно произвольной скоростью.

Ключевые слова: асимптотически эффективный критерий, задача регрессионного типа, асимптотический дефект, локальная асимптотическая нормальность, отношение правдоподобия, стохастическое разложение.

Вестник ТвГУ. Серия: Прикладная математика. 2014. № 3. С. 37–54.

1. Введение

Предположим, что имеются независимые одинаково распределенные наблюдения $\mathbf{X}_n = (X_1, \ldots, X_n)$, каждое из которых принимает значения в произвольном измеримом пространстве $(\mathcal{X},\ \mathcal{A})$ и имеет неизвестную с точностью до параметра θ плотность $p(x,\theta)$ относительно некоторой σ-конечной меры $\nu(\cdot)$ на $\mathcal{A}$. Рассмотрим задачу проверки простой гипотезы в случае однопараметрического семейства распределений. Пусть неизвестный параметр θ принадлежит открытому множеству $\Theta \subset \mathbb{R}^1$, содержащему ноль. Обозначим через $\mathsf{P}_{n,\theta},\ \mathsf{E}_{n,\theta}$ соответственно распределение и математическое ожидание $\mathbf{X}_n$, а через $\mathsf{P}_\theta,\ \mathsf{E}_\theta$ соответственно распределение и математическое ожидание случайного элемента X_1.

Предположим, что мы хотим проверить простую гипотезу

$$\mathsf{H}_0:\ \theta\ =\ 0 \tag{1.1}$$

против сложной альтернативы $\theta \neq 0$. Хорошо известно (см., например, [7], глава 3), что в общем случае наилучшего, то есть равномерно наиболее мощного критерия не существует, и поэтому рассматривается асимптотический подход, при котором $n \to \infty$. Рассмотрим сначала простую альтернативу

$$\mathsf{H}_1 : \ \theta \ = \ \theta_1 \ \neq \ 0. \tag{1.2}$$

Здесь θ_1 предполагается известным. Заметим, что точка $\theta = 0$ в гипотезе H_0 может быть заменена на любую фиксированную точку $\theta_0 \in \Theta$. Этот случай сводится к предыдущему с помощью рассмотрения семейства плотностей вида $p(x, \theta_0 + \xi)$, где ξ принадлежит некоторой окрестности нуля.

Применяя фундаментальную лемму Неймана – Пирсона (см., [7], глава 3), наилучший (наиболее мощный) критерий для проверки гипотезы H_0 против простой альтернативы H_1 основан на логарифме отношения правдоподобия

$$\Lambda_n(\theta) \ = \ \sum_{i=1}^{n} \left(l(X_i, \theta) \ - \ l(X_i, 0)\right), \tag{1.3}$$

где $l(x, \theta) \ = \ \log p(x, \theta)$, и этот критерий отвергает гипотезу H_0 при «больших» значениях логарифма отношения правдоподобия, то есть в случае, если

$$\Lambda_n(\theta_1) \ > \ c_n,$$

где критическое значение c_n выбирается из условия

$$\mathsf{P}_{n,0}(\Lambda_n(\theta_1) \ > \ c_n) \ = \ \alpha, \tag{1.4}$$

и $\alpha \ \in \ (0, 1)$ – фиксированный уровень значимости. Мы для простоты изложения предполагаем непрерывность распределения логарифма отношения правдоподобия $\Lambda_n(\theta_1)$ при гипотезе H_0, то есть считаем, что

$$\mathsf{P}_{n,0}(\Lambda_n(\theta_1) \ = \ c_n) \ = \ 0.$$

Далее предположим также, что существуют все необходимые моменты случайной величины $l(X_1, \theta)$ и все необходимые производные по θ функции $l(x, \theta)$.

Обозначим через

$$\mu(\theta) \ = \ \mathsf{E}_\theta(l(X_1, \theta_1) \ - \ l(X_1, 0)), \tag{1.5}$$

$$\sigma^2(\theta) \ = \ \mathsf{D}_\theta(l(X_1, \theta_1) \ - \ l(X_1, 0)) \tag{1.6}$$

математическое ожидание и дисперсию случайной величины $l(X_1, \theta_1) \ - \ l(X_1, 0)$ при распределении P_θ.

Поскольку логарифм отношения правдоподобия $\Lambda_n(\theta)$ есть сумма независимых одинаково распределенных случайных величин, то согласно центральной предельной теореме имеем при $n \to \infty$

$$\mathcal{L}\left(\frac{\Lambda_n(\theta_1) \ - \ n\mu(0)}{\sigma(0)\sqrt{n}} \ \Big| \ \mathsf{H}_0\right) \ \to \ \mathcal{N}(0, 1), \tag{1.7}$$

где $\mathcal{L}(Z \mid \mathsf{H}_i)$ означает распределение Z при гипотезе H_i, $i \ = \ 0, \ 1$, $\mathcal{N}(\mu, \sigma^2)$ – нормальный закон с параметрами μ и σ^2, а символ $\to$ здесь и далее означает слабую сходимость распределений.

Из соотношения (1.4) теперь следует (поскольку сходимость функций распределения к нормальной функции распределения равномерна), что

$$c_n \;=\; \sqrt{n}\sigma(0)u_\alpha \;+\; n\mu(0) \;+\; o(1), \tag{1.8}$$

где $u_\alpha \;=\; \Phi^{-1}(1 \,-\, \alpha)$ и $\Phi(x)$ – функция распределения стандартного нормального закона.

Обозначим через $\beta_n^*(\theta_1)$ мощность наиболее мощного критерия для проверки гипотезы H_0 против простой альтернативы H_1 (см. (1.1), (1.2)), основанного на логарифме отношения правдоподобия $\Lambda_n(\theta_1)$, то есть

$$\beta_n^*(\theta_1) \;=\; \mathsf{P}_{n,\theta_1}(\Lambda_n(\theta_1) \;>\; c_n). \tag{1.9}$$

Хорошо известно, что этот критерий состоятелен (см., например, [8], Лемма 1.1), то есть

$$\beta_n^*(\theta_1) \;\to\; 1, \;\; n \;\to\; \infty.$$

Факт стремления мощности $\beta_n^*(\theta_1)$ к единице, а точнее скорость сходимости к единице, может быть использован для сравнения различных состоятельных критериев (см., например, [9] и обзор, приведенный там). Однако мы здесь рассмотрим несколько иной, асимптотический подход к сравнению различных критериев. Это так называемый подход Питмэна (см. [10]), согласно которому для получения нетривиального предела мощности $\beta_n^*(\theta_1)$, заключенного между α и 1, рассматривают последовательность альтернатив $\theta_1 \;=\; \theta_n$, стремящуюся к нулю. Из соотношения (1.7) и центральной предельной теоремы для схемы серий (см., например, [11], теорема 8.4.5) следует, что в регулярном случае для выполнения этого должно быть

$$\mu(\theta_n) \;-\; \mu(0) \;=\; \mathcal{O}\big(n^{-1/2}\big), \;\; \theta_n \;=\; \mathcal{O}\big(n^{-1/2}\big). \tag{1.10}$$

Поэтому мы будем рассматривать задачу проверки простой гипотезы H_0 (см. (1.1)) против последовательности сложных близких альтернатив вида

$$\mathsf{H}_{n,1}: \; \theta \;=\; \frac{t}{\sqrt{n}}, \;\; 0 \;<\; t \;\leqslant\; C, \;\; C \;>\; 0, \tag{1.11}$$

где параметр t неизвестен.

Для любого фиксированного $t \;\in\; (0, C]$ наилучший критерий для проверки гипотезы H_0 против простой альтернативы

$$\mathsf{H}_{n,t}: \; \theta \;=\; \frac{t}{\sqrt{n}} \tag{1.12}$$

основан на логарифме отношения правдоподобия

$$\Lambda_n(t) \;\equiv\; \sum_{i=1}^{n} \big(l(X_i, tn^{-1/2}) \;-\; l(X_i, 0)\big). \tag{1.13}$$

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 $\mathsf{H}_{n,1}$. Однако $\beta_n^*(t)$, это так называемая огибающая функция мощности, которая дает верхнюю границу для мощности любого критерия при проверке гипотезы H_0 против фиксированной альтернативы $\mathsf{H}_{n,t}$, $t > 0$, и может служить стандартом при сравнении различных критериев.

При естественных условиях регулярности хорошо известно (см., например, [8], (1.26)) предельное значение для мощности $\beta_n^*(t)$:

$$\beta_n^*(t) \to \beta^*(t) = \Phi\big(t\sqrt{I} - u_\alpha\big), \tag{1.14}$$

где I – фишеровская информация (см. Теорему 2.2 с $I = I(0)$).

Заметим, что для проверки гипотезы H_0 против альтернативы $\mathsf{H}_{n,1}$ существуют критерии, основанные на статистиках, отличных от $\Lambda_n(t)$, и имеющие ту же предельную мощность $\beta^*(t)$. Такие критерии называются асимптотически наиболее мощными (АНМ) (точнее, локально АНМ, поскольку альтернатива $\mathsf{H}_{n,1}$ имеет локальный характер). Таковы, например, критерии основанные на статистиках T_n (см. раздел 2), $\Lambda_n(t_0)$, где $t_0 > 0$ фиксировано, оценках максимального правдоподобия и т.п. Заметим, что все эти статистики не зависят от неизвестного параметра t, и поэтому могут быть использованы при проверке гипотезы H_0 против альтернативы $\mathsf{H}_{n,1}$.

Соотношение (1.14) создает естественную основу для асимптотического сравнения различных АНМ критериев, однако, для различения критериев такого рода, то есть удовлетворяющих соотношению

$$\beta_n(t) \to \beta^*(t), \tag{1.15}$$

где $\beta_n(t)$ – мощность конкретного рассматриваемого критерия, нужны следующие члены асимптотического разложения $\beta_n(t)$, то есть представление типа

$$\beta_n(t) = \beta^*(t) + \frac{1}{\sqrt{n}}h_1(t) + \frac{1}{n}h_2(t) + \cdots . \tag{1.16}$$

Асимптотическим разложениям в статистике посвящены работы [12] и [13]. Получая формулы типа (1.16) для различных критериев, было замечено, что 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 [5], [1], [13]. При этом величина

$$r(t) = \lim_{n\to\infty} n(\beta_n^*(t) - \beta_n(t)) \tag{1.17}$$

допускает статистическую интерпретацию в терминах необходимого числа наблюдений и позволяет находить асимптотический дефект (см. [14], [5], [1]).

Соотношение (1.17) может быть эвристически объяснено следующим образом. Предположим, что статистику $T_n = T_n(\mathbf{X}_n)$ АНМ критерия можно монотонным преобразованием (не меняющим мощности критерия) преобразовать в статистику $S_n(t)$ такую, что величина

$$\Delta_n(t) \equiv S_n(t) - \Lambda_n(t) \to 0, \quad n \to \infty \tag{1.18}$$

по вероятности относительно распределений $\mathsf{P}_{n,0}$ и $\mathsf{P}_{n,t/\sqrt{n}}$. Тогда критерий, основанный на статистике $S_n(t)$, имеет те же предельные распределения при гипотезах

H_0 и $\mathsf{H}_{n,t}$, что и критерий, основанный на $\Lambda_n(t)$, и, следовательно, ту же предельную мощность $\beta^*(t)$ (см. (1.14)).

В том типичном случае, когда $\Delta_n(t)$ имеет порядок $n^{-1/2}$, то есть разность между $S_n(t)$ и $\Lambda_n(t)$ имеет тот же порядок, можно ожидать, что мощность $\beta_n(t)$ критерия, основанного на $S_n(t)$ (или на T_n), отличается от $\beta_n^*(t)$ на величину порядка $n^{-1/2}$. Однако было обнаружено, что для широкого класса АНМ критериев это отличие имеет порядок n^{-1} (см. [4], [5], [1]). Первоначально выражения для $r(t)$ (см. (1.17)) строились с помощью асимптотических разложений для $\beta_n^*(t)$ и $\beta_n(t)$ (см. [13], [4]). Этот подход технически очень трудоемкий и громоздкий. Однако в работах [4], [5], [1] была получена общая формула для величины $r(t)$ без построения асимптотических разложений. Эта формула имеет наглядный вид в терминах условных дисперсий. Для ее демонстрации обозначим через $\Lambda(t)$ нормальную случайную величину вида $\mathcal{N}(-\frac{1}{2}t^2 I, t^2 I)$, тогда в силу формулы (1.21) из работы [8]

$$\mathcal{L}(\Lambda_n(t) \mid \mathsf{H}_0) \;\to\; \mathcal{L}(\Lambda(t)), \tag{1.19}$$

и предположим, что при гипотезе H_0 случайный вектор

$$(\sqrt{n}\Delta_n(t),\; \Lambda_n(t)) \tag{1.20}$$

имеет предельное распределение (типичным образом двумерное нормальное), совпадающее с распределением вектора

$$(\Delta(t),\; \Lambda(t)), \tag{1.21}$$

тогда в работах [5] и [1] показано, что

$$r(t) \;=\; \frac{1}{2t\sqrt{I}}\,\varphi(u_\alpha \;-\; t\sqrt{I})\;\mathsf{D}(\Delta(t) \mid \Lambda(t) \;=\; c_t), \tag{1.22}$$

где

$$c_t \;=\; t\sqrt{I}u_\alpha \;-\; \frac{1}{2}t^2 I$$

и $\varphi(x) \;=\; \Phi'(x)$. Заметим, что величина $r(t)$ тесно связана с асимптотическим дефектом ([14], [1]). Точнее, при заданных ошибках первого и второго рода $\alpha \;>\; 0$ и $\gamma \;>\; 0$, $\alpha \;+\; \gamma \;<\; 1$ асимптотический дефект (см. [14], [5]) равен

$$d \;=\; \frac{2\,r(t_{\alpha,\gamma})}{(u_\alpha \;+\; u_\gamma)^2\;\varphi(u_\gamma)}, \quad t_{\alpha,\gamma} \;=\; I^{-1/2}\,(u_\alpha \;+\; u_\gamma).$$

Или в несколько другой постановке (см. формулу (1.5.39) из книги [1])

$$d \;=\; \frac{2\,r(t)}{t\,\sqrt{I}\;\varphi(t\sqrt{I} \;-\; u_\alpha)}.$$

В работе [1] рассмотрен общий случай в терминах общего статистического эксперимента и приведена общая теорема (см. [1], Теорема 3.2.1), дающая достаточные условия для существования предела

$$r \;=\; \lim_{n\to\infty} \tau_n^{-2}(\beta_n^* \;-\; \beta_n) \;=\; \frac{1}{2}e^b p(b)\;\mathsf{D}(\Delta \mid \Lambda \;=\; b), \tag{1.23}$$

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

$$\Phi_1(x) = \Phi\Big(\frac{x + \frac{1}{2}t^2 I}{t\sqrt{I}}\Big), \quad p(x) = \Phi_1'(x)$$

и $\tau_n \to 0$ – 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 критерия T_n.

Цель настоящей статьи – привести пример, состоящий в рассмотрении задачи регрессионного типа с регулярными коэффициентами, в котором в формуле (1.23) может возникнуть случай $\tau_n \neq n^{-1/2}$ и достаточно произвольной скорости стремления к нулю τ_n. Отметим, что в работе [8] рассмотрен случай распределения Лапласа и для него доказано, что $\tau_n = n^{-1/4}$.

2. Альтернативы регрессионного типа

Пусть, как и выше, мы имеем вектор наблюдений $\mathbf{X}_n = (X_1, \ldots, X_n)$, состоящий из независимых случайных элементов, но теперь эти случайные элементы распределены, вообще говоря, неодинаково. Рассмотрим так называемую задачу регрессионного типа (см., например, [15] стр.125). Предположим, что каждое наблюдение X_i принимает значения в произвольном измеримом пространстве $(\mathcal{X}, \mathcal{A})$ и имеет неизвестную с точностью до параметра θ_i плотность $p(x, \theta_i)$ относительно некоторой σ-конечной меры $\nu(\cdot)$ на $\mathcal{A}$. Пусть неизвестные параметры θ_i принадлежат множеству $\Theta \subset \mathbb{R}^1$, содержащему отрезок вида $[-\delta, \delta]$, $\delta > 0$. Таким образом каждое наблюдение X_i имеет плотность $p(x, \theta_i)$ с некоторым $\theta_i \in \Theta$, $i = 1, \ldots, n$.

Предположим, что при каждом n задан набор известных констант $c_{n,1}, \ldots, c_{n,n}$ и рассмотрим задачу проверки простой гипотезы

$$\mathsf{H}_0: \ \theta_i = 0, \ \ i = 1, \ldots, n \tag{2.1}$$

против последовательности сложных альтернатив вида

$$\mathsf{H}_{n,1}: \ \theta_i = c_{n,i}\, t, \ \ i = 1, \ldots, n, \ 0 < t \leqslant C, \ \ C > 0. \tag{2.2}$$

Заметим, что при

$$c_{n,1} = \ldots = c_{n,n} = \frac{1}{\sqrt{n}}$$

мы имеем альтернативу $\mathsf{H}_{n,1}$ (см. формулу (1.13)), рассмотренную выше.

Для любого фиксированного $t \in (0, C]$ наилучший критерий для проверки гипотезы H_0 против простой альтернативы

$$\mathsf{H}_{n,t}: \ \theta_i = c_{n,i}\, t, \ \ i = 1, \ldots, n \tag{2.3}$$

основан на логарифме отношения правдоподобия

$$\Lambda_n(t) \equiv \sum_{i=1}^{n} \big(l(X_i, {}_{n,i}\, t) - l(X_i, 0)\big), \tag{2.4}$$

где $l(x,\theta) = \log p(x,\theta)$, $\theta \in \Theta$.

Обозначим через $\mathsf{P}_{n,t}$, $\mathsf{E}_{n,t}$, соответственно, распределение и математическое ожидание $\mathbf{X}_n$ при распределении, определяемом гипотезой $\mathsf{H}_{n,t}, t > 0$, а через P_0, E_0, соответственно, распределение и математическое ожидание случайного элемента X_1 при распределении, определяемом гипотезой H_0.

Будем предполагать, что коэффициенты $c_{n,i}$, $i = 1,\ldots,n$ удовлетворяют следующему условию

Условие 1.

1. Условие нормировки

$$\sum_{i=1}^{n} c_{n,i}^2 = 1.$$

2. Условие равномерной малости

$$C_n \equiv \max_{1\leqslant i\leqslant n} |c_{n,i}| \to 0, \quad n \to \infty.$$

Заметим, что Условие 1 выполняется, например, если

$$c_{n,1} = \ldots = c_{n,n} = \frac{1}{\sqrt{n}},$$

или, если для некоторого натурального $m_n \in \{1,\ldots,n\}$, $m_n \to \infty$, $n \to \infty$

$$c_{n,i} = \pm\frac{1}{\sqrt{m_n}}, \quad i \in \{i_1,\ldots,i_{m_n}\}; \quad c_{n,i} = 0, \quad i \notin \{i_1,\ldots,i_{m_n}\}.$$

Лемма 2.1.

1. *При выполнении Условия 1 для любых $r > 2$ и $p \geqslant 3$ справедливы неравенства*

$$\sum_{i=1}^{n} |c_{n,i}|^r \leqslant C_n^{r-2} \to 0,$$

$$\sum_{i=1}^{n} |c_{n,i}|^p \geqslant n^{-1/2} \sum_{i=1}^{n} |c_{n,i}|^{p-1} \geqslant n^{-p/2+1}.$$

2. *Для любых констант $c_{n,1},\ldots,c_{n,n}$ и любых чисел $r > 0$ и $p \geqslant 1$ справедливы неравенства*

$$\sum_{i=1}^{n} |c_{n,i}|^r \leqslant C_n \sum_{i=1}^{n} |c_{n,i}|^{r-1}, \quad \sum_{i=1}^{n} |c_{n,i}|^r \geqslant C_n^r, \ C_n = \max_{1\leqslant i\leqslant n} |c_{n,i}|,$$

$$\sum_{i=1}^{n} |c_{n,i}|^{rp} \geqslant \frac{1}{n^{p-1}} \Big(\sum_{i=1}^{n} |c_{n,i}|^r\Big)^p.$$

Доказательство следует из соотношения

$$\sum_{i=1}^{n} |c_{n,i}|^r = \sum_{i=1}^{n} |c_{n,i}|^{r-2} c_{n,i}^2 \leqslant C_n^{r-2} \sum_{i=1}^{n} c_{n,i}^2 = C_n^{r-2} \to 0$$

и неравенства Гельдера.

В работе [15] (см. также [6], глава 7) доказана следующая теорема, позволяющая получить стохастическое разложение (это свойство локальной асимптотической нормальности) для логарифма отношения правдоподобия $\Lambda_n(t)$ (см. (2.4)).

Теорема 2.2. *Пусть* $\mathcal{X} = \mathbb{R}^1$, $\mathcal{A}$ *– борелевская* σ*-алгебра на прямой,* $\nu(\cdot)$ *– мера Лебега на прямой и плотность* $p(x,\theta)$ *удовлетворяет следующим условиям:*

1. При каждом $x \in \mathbb{R}^1$ *плотность* $p(x,\theta)$ *абсолютно непрерывна по* θ *в некоторой окрестности точки* $\theta = 0$.

2. При каждом θ *из этой окрестности производная*

$$p^{(1)}(x,\theta) \equiv \frac{\partial}{\partial\theta}\, p(x,\theta)$$

существует при почти всех (по мере Лебега) $x \in \mathbb{R}^1$.

3. Функция

$$I(\theta) \equiv \mathsf{E}_\theta \left(\frac{p^{(1)}(X_1,\theta)}{p(X_1,\theta)}\right)^2 < \infty$$

положительна и непрерывна в этой окрестности.

Пусть также коэффициенты $c_{n,i}$, $i = 1,\ldots,n$ *удовлетворяют Условию 1.*

Тогда для любой ограниченной последовательности $t_n \in \mathbb{R}^1$ *справедливо представление*

$$\Lambda_n(t_n) = t_n\, T_n - \frac{t_n^2\, I}{2} + \rho_n(t_n),$$

где

$$T_n = \sum_{i=1}^n c_{n,i}\, l^{(1)}(X_i), \qquad I = I(0), \qquad l^{(1)}(x) = \frac{\partial}{\partial\theta}\, l(x,\theta)\Big|_{\theta=0}$$

и для любого $\epsilon > 0$

$$\mathsf{P}_{n,0}(|\rho_n(t_n)| > \epsilon) \to 0, \quad n \to \infty.$$

Причем, если $t_n \to t < \infty$, *то*

$$\mathcal{L}(T_n|\mathsf{H}_0) \to \mathcal{N}(0,I), \qquad \mathcal{L}(T_n|\mathsf{H}_{n,t_n}) \to \mathcal{N}(tI,I),$$

$$\mathcal{L}(\Lambda_n(t_n)|\mathsf{H}_0) \to \mathcal{N}\Big(-\frac{t^2 I}{2}, t^2 I\Big), \qquad \mathcal{L}(\Lambda_n(t_n)|\mathsf{H}_{n,t_n}) \to \mathcal{N}\Big(\frac{t^2 I}{2}, t^2 I\Big), \quad n \to \infty.$$

Из этой Теоремы следует, что наилучший критерий для проверки гипотезы H_0 против последовательности простых альтернатив $\mathsf{H}_{n,t}$, основанный на логарифме отношения правдоподобия $\Lambda_n(t)$, имеет мощность

$$\beta_n^*(t) \to \beta^*(t) = \Phi\big(t\sqrt{I} - u_\alpha\big), \tag{2.5}$$

где $I = I(0)$ – фишеровская информация.

Из этой Теоремы также следует, что логарифм отношения правдоподобия $\Lambda_n(t)$ является линейной функцией статистики T_n с точностью до остаточного члена, сходящегося к нулю по вероятности. Поэтому следует ожидать, что критерий, основанный на статистике T_n, будет иметь туже предельную мощность $\beta^*(t)$, что и критерий, основанный на логарифме отношения правдоподобия $\Lambda_n(t)$. Действительно, обозначая мощность критерия, основанного на статистике T_n через $\beta_n(t)$ из этой Теоремы непосредственно следует, что

$$\beta_n(t) \;\to\; \beta^*(t) \;=\; \Phi\big(t\sqrt{I} \;-\; u_\alpha\big). \tag{2.6}$$

Таким образом, критерий, основанный на статистике T_n, является локально асимптотически наиболее мощным (АНМ) критерием и в следующем разделе будет исследовано асимптотическое поведение разности мощностей $\beta^*_n(t) \;-\; \beta_n(t)$ при $n \;\to\; \infty$.

3. Асимптотическое поведение мощности АНМ критерия, основанного на статистике T_n

Применяя общие результаты работы [1] к частному случаю, рассмотренному в разделе 2, можно точно сформулировать достаточные условия регулярности для выполнения равенства

$$r(t) \;=\; \lim_{n\to\infty} \tau_n^{-2}\,(\beta^*_n(t) \;-\; \beta_n(t)) \;=\; \frac{1}{2t\sqrt{I}}\,\varphi(u_\alpha \;-\; t\sqrt{I})\,\mathsf{D}(\Delta(t) \mid \Lambda(t) \;=\; c_t), \tag{3.1}$$

где

$$c_t \;=\; t\sqrt{I}u_\alpha \;-\; \frac{1}{2}t^2 I,$$

$\tau_n \;\to\; 0$ – малый параметр и $\varphi(x) \;=\; \Phi'(x)$.

Эвристически эти условия сводятся к существованию слабого предела

$$\mathcal{L}((\tau_n^{-1}\;\Delta_n(t), \Lambda_n(t))) \;\to\; \mathcal{L}((\Delta(t), \Lambda(t))), \quad n \;\to\; \infty,$$

где (см. формулу (3.2))

$$\Delta_n(t) \;=\; S_n(t) \;-\; \Lambda_n(t).$$

Из Теоремы 2.2 следует, что при выполнении ее условий, случайная величина $\Lambda(t)$ является нормальной случайной величиной вида $\mathcal{N}\Big(-\frac{t^2 I}{2},\; t^2 I\Big)$.

Цель этого раздела – найти малый параметр τ_n, описать двумерное распределение случайного вектора $(\Delta(t),\;\Lambda(t))$ и, как следствие, найти условную дисперсию в формуле (3.1).

Из соотношения (2.4) и формулы Тейлора следует приближенное равенство

$$\Lambda_n(t) \;\equiv\; \sum_{i=1}^{n} \big(l(X_i, {}_{n,i}\; t) \;-\; l(X_i, 0)\big) \;\approx\; \bar{\Lambda}_n(t) \;\equiv$$

$$\equiv\; t\;T_n \;+\; \frac{t^2}{2}\,\sum_{i=1}^{n}\,c_{n,i}^2\;l^{(2)}(X_i) \;+\; \frac{t^3}{6}\,\sum_{i=1}^{n}\,c_{n,i}^3\;l^{(3)}(X_i),$$

где

$$T_n \;=\; \sum_{i=1}^{n}\,c_{n,i}\;l^{(1)}(X_i).$$

Поэтому для линейного монотонного преобразования статистики T_n

$$S_n(t) \;=\; t\,T_n \;-\; \frac{t^2 I}{2} \;+\; \frac{t^3\;\mathsf{E}_0 l^{(3)}(X_1)}{6}\;\sum_{i=1}^{n}\;c_{n,i}^3 \tag{3.2}$$

справедливо равенство

$$\Delta_n(t) \;=\; S_n(t) \;-\; \Lambda_n(t) \;\approx\; \bar{\Delta}_n(t) \;\equiv$$

$$\equiv\; -\frac{t^2}{2}\;\sum_{i=1}^{n}\;c_{n,i}^2\;\big(l^{(2)}(X_i) \;+\; I\big) \;-\frac{t^3}{6}\;\sum_{i=1}^{n}\;c_{n,i}^3\;\big(l^{(3)}(X_i) \;-\; \mathsf{E}_0\;l^{(3)}(X_1)\big). \tag{3.3}$$

Найдем предельное распределение нормированной случайной величины $\bar{\Delta}_n(t)$. Пусть

$$\tau_n \;=\; \sqrt{\sum_{i=1}^{n}\;c_{n,i}^4}\;,$$

тогда при выполнении Условия 1, согласно Лемме 2.1 $\tau_n\;\to\;0,\;\;n\;\to\;\infty$. Сформулируем еще одно условие на асимптотическое поведение констант $c_{n,i},\;\;i\;=\;1,\ldots,n$.

Условие 2.

Предположим, что

$$\frac{C_n^2}{\sqrt{\sum_{i=1}^n\;c_{n,i}^4}} \;=\; C_n^2\;\tau_n^{-1}\;\to\;0,\quad n\;\to\;\infty,\quad C_n\;=\;\max_{1\leqslant i\leqslant n}\;|c_{n,i}|.$$

Заметим, что Условие 2 выполняется, например, если

$$c_{n,1} \;=\; \ldots\; = c_{n,n} \;=\; \frac{1}{\sqrt{n}},$$

или, если для некоторого натурального $m_n\;\in\;\{1,\ldots,n\}$, $m_n\;\to\;\infty,\;n\;\to\;\infty$

$$c_{n,i} \;=\; \pm\;\frac{1}{\sqrt{m_n}},\quad i\;\in\;\{i_1,\ldots,i_{m_n}\};\quad c_{n,i}\;=\;0,\quad i\;\notin\;\{i_1,\ldots,i_{m_n}\}.$$

Заметим, что в этом случае

$$\tau_n \;=\; \frac{1}{\sqrt{m_n}}\;\to\;0,\quad n\;\to\;\infty$$

и τ_n может стремиться к нулю достаточно медленно, например, как $\frac{1}{\log\log(n\;+\;1)}$.

Лемма 3.1.

1. Пусть выполнено Условие 1, тогда

$$\mathcal{L}(\bar{\Lambda}_n(t)|\mathsf{H}_0)\;\to\;\mathcal{N}\Big(-\frac{t^2\;I}{2},\;t^2 I\Big),\quad n\;\to\;\infty.$$

2. Пусть выполнены Условия 1, 2, тогда

$$\mathcal{L}(\tau_n^{-1}\;\bar{\Delta}_n(t)|\mathsf{H}_0)\;\to\;\mathcal{N}\Big(0,\;\frac{t^4}{4}\;\mathsf{D}_0\;l^{(2)}(X_1)\Big),\quad n\;\to\;\infty.$$

Доказательство. 1. Докажем, что третий член в определении $\bar{\Lambda}_n(t)$ стремится к нулю по вероятности. Это следует из неравенства Чебышева

$$\mathsf{P}_0\Big(\Big|\sum_{i=1}^n c_{n,i}^3\ l^{(3)}(X_i)\Big|\ \geqslant\ \epsilon\Big)\ \leqslant\ \frac{\mathsf{E}_0\ |l^{(3)}(X_1)|\ \sum_{i=1}^n\ |c_{n,i}|^3}{\epsilon}\ \leqslant$$

$$\leqslant\ \frac{\mathsf{E}_0\ |l^{(3)}(X_1)|\ C_n\ \sum_{i=1}^n\ c_{n,i}^2}{\epsilon}\ =\ \frac{\mathsf{E}_0\ |l^{(3)}(X_1)|\ C_n}{\epsilon},\quad \epsilon\ >\ 0.$$

В силу Условия 1 $C_n\ \to\ 0,\ \ n\ \to\ \infty$. Таким образом, величина $\bar{\Lambda}_n(t)$ асимптотически ведет себя так же, как

$$t\ \sum_{i=1}^n\ c_{n,i}\ l^{(1)}(X_i)\ +\ \frac{t^2}{2}\ \sum_{i=1}^n\ c_{n,i}^2\ l^{(2)}(X_i). \tag{3.4}$$

К первому слагаемому этой суммы применима Теоремы 4.1.2 из книги [15] (стр. 198), согласно которой

$$\mathcal{L}\Big(t\ \sum_{i=1}^n\ c_{n,i}\ l^{(1)}(X_i)\Big|\mathsf{H}_0\Big)\ \to\ \mathcal{N}\Big(0,\ t^2 I\Big),\quad n\ \to\ \infty.$$

Ко второму слагаемому суммы (3.4) применим закон больших чисел специального вида из работы [6] (стр. 170, Теорема 8.3.5), поскольку выполнено Условие 1, имеем

$$\frac{t^2}{2}\ \sum_{i=1}^n\ c_{n,i}^2\ l^{(2)}(X_i)\ \xrightarrow{\mathsf{P}_{n,0}}\ \frac{t^2}{2}\ \mathsf{E}_0\ l^{(2)}(X_1)\ =\ -\frac{t^2}{2}\ I,\quad n\ \to\ \infty.$$

2. Докажем, что второй член в определении $\tau_n^{-1}\ \bar{\Delta}_n(t)$ стремится к нулю по вероятности. Это следует из неравенства Чебышева вида

$$\mathsf{P}_0\Big(\Big|\tau_n^{-1}\ \sum_{i=1}^n\ c_{n,i}^3\ \big(l^{(3)}(X_i)\ -\ \mathsf{E}_0\ l^{(3)}(X_1)\big)\Big|\ \geqslant\ \epsilon\Big)\ \leqslant\ \frac{\mathsf{D}_0\ l^{(3)}(X_1)\ \tau_n^{-2}\ \sum_{i=1}^n\ c_{n,i}^6}{\epsilon^2}\ \leqslant$$

$$\leqslant\ \frac{\mathsf{D}_0\ l^{(3)}(X_1)\ \tau_n^{-2}\ C_n^4\ \sum_{i=1}^n\ c_{n,i}^2}{\epsilon^2}\ =\ \frac{\mathsf{D}_0\ l^{(3)}(X_1)\ \tau_n^{-2}\ C_n^4}{\epsilon^2},\quad \epsilon\ >\ 0.$$

В силу Условия 2

$$\tau_n^{-2}\ C_n^4\ \to\ 0,\quad n\ \to\ \infty.$$

Таким образом, величина $\tau_n^{-1}\ \bar{\Delta}_n(t)$ асимптотически ведет себя так же, как

$$-\tau_n^{-1}\ \frac{t^2}{2}\ \sum_{i=1}^n\ c_{n,i}^2\ \big(l^{(2)}(X_i)\ +\ I\big).$$

Теперь утверждение Леммы следует из Условия 2 и Теоремы 4.1.2 из книги [15] (стр. 198). Лемма доказана. □

Сформулируем еще одно условие на асимптотическое поведение констант $c_{n,i},\ i\ =\ 1,\dots,n$.

Условие 3.
Предположим, что существует конечный предел

$$\tau_n^{-1} \sum_{i=1}^{n} c_{n,i}^3 \to \rho < \infty, \; n \to \infty.$$

Заметим, что Условие 3 выполняется, например, если

$$c_{n,1} = \ldots = c_{n,n} = \frac{1}{\sqrt{n}},$$

то $\rho = 1$, или, если для некоторого натурального $m_n \in \{1, \ldots, n\}$, $m_n \to \infty, n \to \infty$

$$c_{n,i} = \pm \frac{1}{\sqrt{m_n}}, \quad i \in \{i_1, \ldots, i_{m_n}\}; \quad c_{n,i} = 0, \quad i \notin \{i_1, \ldots, i_{m_n}\},$$

то $\rho = \pm 1$. Если

$$c_{n,i} = \frac{1}{\sqrt{n}}, \quad i \in \{1, \ldots, [n/2]\}, \quad c_{n,i} = -\frac{1}{\sqrt{n}}, \quad i \in \{[n/2]+1, \ldots, n\},$$

то $\rho = 0$.

Лемма 3.2. *Пусть выполнены Условия 1, 2 и 3, тогда*

$$\mathcal{L}\big((\tau_n^{-1}\, \bar\Delta_n(t),\; \bar\Lambda_n(t))|\mathsf{H}_0\big) \to$$

$$\to \mathcal{N}_2\Big(0,\; -\frac{t^2\, I}{2},\; \frac{t^4}{4}\, \mathsf{D}_0\, l^{(2)}(X_1),\; t^2 I,\; -\frac{t^3}{2}\, \mathsf{E}_0\, l^{(1)}(X_1) l^{(2)}(X_1)\, \rho\Big), \quad n \to \infty.$$

Доказательство. Для доказательства Леммы применим прием Крамера-Уолда, согласно которому достаточно доказать, что для любого вектора $(s_1, s_2) \neq (0,0)$ справедливо соотношение

$$\mathcal{L}\big(s_1 \tau_n^{-1}\, \bar\Delta_n(t) + s_2 \bar\Lambda_n(t)|\mathsf{H}_0\big) \to \mathcal{L}\big(s_1 Z_1 + s_2 Z_2\big), \quad n \to \infty, \tag{3.5}$$

где $(Z_1,\; Z_2)$ двумерный нормальный вектор вида

$$(Z_1,\; Z_2) \sim \mathcal{N}_2\Big(0,\; -\frac{t^2\, I}{2},\; \frac{t^4}{4}\, \mathsf{D}_0\, l^{(2)}(X_1),\; t^2 I,\; -\frac{t^3}{2}\, \mathsf{E}_0\, l^{(1)}(X_1) l^{(2)}(X_1)\, \rho\Big).$$

Из доказательства Леммы 3.1 следует, что для доказательства соотношения (3.5) достаточно доказать, что

$$\mathcal{L}\Big(-s_1 \tau_n^{-1}\, \frac{t^2}{2} \sum_{i=1}^{n} c_{n,i}^2\, \big(l^{(2)}(X_i) + I\big) + s_2\Big(t \sum_{i=1}^{n} c_{n,i}\, l^{(1)}(X_i) - \frac{t^2 I}{2}\Big)\Big) \to$$

$$\to \mathcal{N}\Big(-\frac{t^2 I s_2}{2},\; \frac{t^4 s_1^2 \mathsf{D}_0\, l^{(2)}(X_1)}{4} + s_2 t^2 I - t^3\, s_1\, s_2\, \mathsf{E}_0\, l^{(1)}(X_1) l^{(2)}(X_1)\, \rho\Big),\; n \to \infty. \tag{3.6}$$

Для доказательства соотношения (3.6) достаточно доказать, что

$$\mathcal{L}\Big(-s_1\tau_n^{-1}\ \frac{t^2}{2}\ \sum_{i=1}^{n}\ c_{n,i}^2\ \big(l^{(2)}(X_i)\ +\ I\big)\ +\ s_2\ t\ \sum_{i=1}^{n}\ c_{n,i}\ l^{(1)}(X_i)\Big)\ \to$$

$$\to\ \mathcal{N}\Big(0,\ \frac{t^4 s_1^2 \mathsf{D}_0\ l^{(2)}(X_1)}{4}\ + s_2 t^2 I\ -\ t^3\ s_1\ s_2\ \mathsf{E}_0\ l^{(1)}(X_1) l^{(2)}(X_1)\ \rho\Big),\ \ n\ \to\ \infty. \quad (3.7)$$

Для доказательства соотношения (3.7) применим хорошо известную Теорему Линдеберга - Феллера (см., например, [16]), согласно которой для того, чтобы сумма независимых случайных величин вида

$$\sum_{i=1}^{n}\ Y_{n,i}$$

с

$$\mathsf{E}Y_{n,i}\ =\ 0,\ \ i\ =\ 1,\ldots,n,\ \ \sum^{n}\ \mathsf{E}\ Y_{n,i}^2\ =\ 1$$

слабо сходилась к стандартному нормальному закону, необходимо и достаточно, чтобы для любого $\epsilon\ >\ 0$ выполнялось условие

$$\sum_{i=1}^{n}\ \mathsf{E}\ Y_{n,i}^2\ \mathbf{I}\{|Y_{n,i}| > \epsilon\}\ \to\ 0,\quad n\ \to\ \infty, \quad (3.8)$$

где $\mathbf{I}\{A\}$ – индикаторная функция события A.

Применим эту Теорему к случайным величинам вида

$$Y_{n,i}\ \equiv\ \frac{-s_1\tau_n^{-1}\ \frac{t^2}{2}\ c_{n,i}^2\ \big(l^{(2)}(X_i)\ +\ I\big)\ +\ s_2\ t\ \ c_{n,i}\ l^{(1)}(X_i)}{\sqrt{\mathsf{E}_0\ \Big(-s_1\tau_n^{-1}\ \frac{t^2}{2}\ \sum_{i=1}^{n}\ c_{n,i}^2\ \big(l^{(2)}(X_i)\ +\ I\big)\ +\ s_2\ t\ \sum_{i=1}^{n}\ c_{n,i}\ l^{(1)}(X_i)\Big)^2}}, \quad (3.9)$$

при этом

$$\mathsf{E}_0\ \Big(-s_1\tau_n^{-1}\ \frac{t^2}{2}\ \sum_{i=1}^{n}\ c_{n,i}^2\ \big(l^{(2)}(X_i)\ +\ I\big)\ +\ s_2\ t\ \sum_{i=1}^{n}\ c_{n,i}\ l^{(1)}(X_i)\Big)^2\ \equiv$$

$$\equiv\ \sigma_n^2(t)\ =\ \frac{t^4}{4}\ s_1^2\ \mathsf{D}_0\ l^{(2)}(X_1)\ +\ s_2^2\ t^2\ I\ -\ t^3\ s_1\ s_2\ \mathsf{E}_0\ l^{(1)}(X_1) l^{(2)}(X_1)\ \tau_n^{-1}\ \sum_{i=1}^{n}\ c_{n,i}^3. \quad (3.10)$$

При выполнении Условия 3

$$\sigma_n^2(t)\ \to\ \sigma^2(t)\ \equiv$$

$$\equiv\ \frac{t^4}{4}\ s_1^2\ \mathsf{D}_0\ l^{(2)}(X_1)\ +\ s_2^2\ t^2\ I\ -\ t^3\ s_1\ s_2\ \mathsf{E}_0\ l^{(1)}(X_1) l^{(2)}(X_1)\ \rho,\quad n\ \to\ \infty. \quad (3.11)$$

Теперь для доказательства Леммы достаточно проверить выполнение условия (3.8) применительно к случайным величинам (3.9).

Обозначим

$$Z_{n,i} \ \equiv \ -s_1 \tau_n^{-1} \ \frac{t^2}{2} \ c_{n,i}^2 \ \left(l^{(2)}(X_i) \ + \ I\right) \ + \ s_2 \ t \ \ c_{n,i} \ l^{(1)}(X_i),$$

тогда

$$Y_{n,i} \ = \ \frac{Z_{n,i}}{\sigma_n(t)},$$

и в силу соотношения (3.11) при всех достаточно больших n

$$\sigma_n^2(t) \ \geqslant \ \frac{\sigma^2(t)}{4},$$

поэтому справедливо неравенство

$$\sum_{i=1}^{n} \ \mathsf{E}_0 \ Y_{n,i}^2 \ \mathbf{I}\{|Y_{n,i}| > \epsilon\} \ \leqslant \ \frac{2}{\sigma(t)} \ \sum_{i=1}^{n} \ \mathsf{E}_0 \ Z_{n,i}^2 \ \mathbf{I}\{|Z_{n,i}| > \sigma(t)\epsilon/2\}. \tag{3.12}$$

Обозначим через

$$\bar{C}_n \ = \ \max\,\{C_n, \ \tau_n^{-1} \ C_n^2\},$$

тогда в силу Условий 1 и 2 $\bar{C}_n \to 0, n \to \infty$. Теперь в силу неравенства

$$|Z_{n,i}| \ \leqslant \ \bar{C}_n \ \Big(|s_1| \ \frac{t^2}{2} \ \left|l^{(2)}(X_i) \ + \ I\right| + |s_2| \ t \ |l^{(1)}(X_i)|\Big),$$

с учетом неравенства (3.12) имеем

$$\sum_{i=1}^{n} \ \mathsf{E}_0 \ Y_{n,i}^2 \ \mathbf{I}\{|Y_{n,i}| > \epsilon\} \ \leqslant$$

$$\leqslant \ \frac{2}{\sigma(t)} \ \sum_{i=1}^{n} \ \mathsf{E}_0 \ Z_{n,i}^2 \ \mathbf{I}\{|s_1| \ t^2/2 \ |l^{(2)}(X_i) \ + \ I| \ + \ |s_2| \ t \ |l^{(1)}(X_i)| \ > \ \sigma(t)\epsilon/(2\bar{C}_n)\}. \tag{3.13}$$

Используя теперь неравенство $(a+b)^2 \ \leqslant \ 2(a^2+b^2)$, оценим случайные величины $Z_{n,i}$ следующим образом

$$Z_{n,i}^2 \ \leqslant \ 2 \ \left(s_1^2 \ t^2/4 \ (l^{(2)}(X_i) \ + \ I)^2 \ \tau_n^{-2} \ c_{n,i}^4 \ + \ s_2^2 \ t^2 \ (l^{(1)}(X_i))^2 \ c_{n,i}^2\right).$$

Теперь с учетом неравенства (3.13) получим

$$\sum_{i=1}^{n} \ \mathsf{E}_0 \ Y_{n,i}^2 \ \mathbf{I}\{|Y_{n,i}| > \epsilon\} \ \leqslant$$

$$\leqslant \ \frac{4}{\sigma(t)} \ \Big(\tau_n^{-2} \ \sum^{n} \ c_{n,i}^4 \ s_1^2 \ \frac{t^2}{4} \ \mathsf{E}_0 \ (l^{(2)}(X_1) \ + \ I)^2 \ \times$$

$$\times \mathbf{I}\{|s_1| \ t^2/2 \ |l^{(2)}(X_i) \ + \ I| \ + \ |s_2| \ t \ |l^{(1)}(X_i)| \ > \ \sigma(t)\epsilon/(2\bar{C}_n)\} \ +$$

$$+ \ s_2^2 \ t^2 \ \mathsf{E}_0 \ (l^{(1)}(X_1))^2 \ \sum_{i=1}^{n} \ c_{n,i}^2 \times$$

$$\mathbf{I}\{|s_1|\ t^2/2\ |l^{(2)}(X_i)\ +\ I|\ +\ |s_2|\ t\ |l^{(1)}(X_i)|\ >\ \sigma(t)\epsilon/(2\bar{C}_n)\}\Big). \tag{3.14}$$

В силу определения τ_n и Условия 1 правая часть неравенства (3.14) приобретает вид

$$\sum_{i=1}^{n} \mathsf{E}_0\ Y_{n,i}^2\ \mathbf{I}\{|Y_{n,i}| > \epsilon\}\ \leqslant$$

$$\leqslant\ \frac{4}{\sigma(t)}\ \Big(s_1^2\ \frac{t^2}{4}\ \mathsf{E}_0\ (l^{(2)}(X_1)\ +\ I)^2\ \times$$

$$\times\mathbf{I}\{|s_1|\ t^2/2\ |l^{(2)}(X_i)\ +\ I|\ +\ |s_2|\ t\ |l^{(1)}(X_i)|\ >\ \sigma(t)\epsilon/(2\bar{C}_n)\}\ +$$

$$+\ s_2^2\ t^2\ \mathsf{E}_0\ (l^{(1)}(X_1))^2\ \times$$

$$\mathbf{I}\{|s_1|\ t^2/2\ |l^{(2)}(X_i)\ +\ I|\ +\ |s_2|\ t\ |l^{(1)}(X_i)|\ >\ \sigma(t)\epsilon/(2\bar{C}_n)\}\Big). \tag{3.15}$$

Оба слагаемых в правой части неравенства (3.15) стремятся к нулю, поскольку $\bar{C}_n\ \to\ 0,\ \ n\ \to\ \infty$ и интегрирование в математических ожиданиях ведется по стягивающейся к пустому множеству области. Лемма доказана. □

Лемма 3.3. *Пусть выполнены Условия 1, 2 и 3, тогда*

$$r(t)\ =\ \lim_{n\to\infty}\ \tau_n^{-2}\ (\beta_n^*(t)\ -\ \beta_n(t))\ =\ \frac{1}{2t\sqrt{I}}\ \varphi(u_\alpha\ -\ t\sqrt{I})\ \mathsf{D}(\bar{\Delta}(t)\ |\ \bar{\Lambda}(t)\ =\ c_t)\ =$$

$$=\ \frac{t^3}{8\sqrt{I}}\ \varphi(u_\alpha\ -\ t\sqrt{I})\ \Big(\mathsf{D}_0\ l^{(2)}(X_1)\ -\ \frac{\rho^2}{I}\ \big(\mathsf{E}_0\ l^{(1)}(X_1)l^{(2)}(X_1)\big)^2\Big). \tag{3.16}$$

Доказательство Леммы состоит в применении Леммы 3.2 и формулы для условной дисперсии двумерного нормального закона. Заметим, что утверждение Леммы 3.3 согласуется с формулой (1.4.10) из книги [1] в случае $\rho\ =\ 1$.

Заключение

Таким образом, на эвристическом уровне (предварительные результаты доказаны в этой работе, однако строгое доказательство основной Леммы 3.3 будет приведено в другой статье) получена формула для асимптотического дефекта (см. (3.16)) в задаче регрессионного типа для случая, когда асимптотически эффективный критерий основан на первом члене стохастического разложения логарифма отношения правдоподобия. Статистика критерия в этом случае представляет собой взвешенную сумму независимых одинаково распределенных случайных величин с коэффициентами, зависящими от n. Таким образом естественно возникает схема серий. Найдены предельные законы распределений статистики критерия и логарифма отношения правдоподобия как при гипотезе, так и при альтернативе. Доказаны необходимые предельные теоремы. При этом показано, что разность между мощностью рассматриваемого критерия и наилучшего критерия, основанного на логарифме отношения правдоподобия, может стремиться к нулю с достаточно произвольной скоростью, зависящей от поведения регрессионных констант. Приведены необходимые примеры, иллюстрирующие возможные здесь ситуации.

Список литературы

[1] Bening V.E. Asymptotic theory of testing statistical hypotheses. Utrecht: VSP, 2000. 277 p.

[2] Bening V.E. A forumula for dificiency: one sample L- and R-tests. I, II // Mathematical Methods of Statistics. 1995. № 4. Pp. 167–188, 274–293.

[3] Bickel P.J., Chibisov D.M., Van Zwet W.R. On efficiency of first and second order // International Statistical Review. 1981. Vol. 49, № 2. Pp. 169–175.

[4] Chibisov D.M. Asymptotic expansions and deficiencies of tests // Proceedings of the International Congress of Mathematicians. Warsaw: PWN, 1984. Pp. 1063–1079.

[5] Чибисов Д.М. Вычисление дефекта асимптотически эффективных критериев // Теория вероятностей и ее применения. 1985. Т. 30, № 2. С. 269–288.

[6] Чибисов Д.М. Лекции по асимптотической теории ранговых критериев. М.: Лекционные курсы НОЦ, 2009.

[7] Леман Э. Проверка статистических гипотез. М.: Наука, 1964. 499 с.

[8] Королев Р.А, Тестова А.В., Бенинг В.Е. О мощности асимптотически оптимального критерия в случае распределения Лапласа // Вестник ТвГУ. Серия: Прикладная математика. 2008. № 8. С. 7–27.

[9] Никитин Я.Ю. Асимптотическая эффективность непараметрических критериев. М.: Наука, 1995. 250 с.

[10] Pitman E.J.G. Lecture notes on nonparametric statistical inference, Lectures given for the University of North Carolina. Institute of Statistics, 1948.

[11] Боровков А.А. Теория вероятностей. М.: УРСС, 2003. 470 с.

[12] Bickel P.J. Edgeworth expansions in nonparametric statistics // Annals of Statistics. 1974. Vol. 2, № 1. Pp. 1–20.

[13] Pfanzagl J. Asymptotic expansions in parametric statistical theory // In Developments in Statistics, Ed. by P.R. Krishnaiah. New York, London: Academic Press, 1989. Vol. 3. Pp. 1–97.

[14] Hodges J.L., Lehmann E.L. Deficiency // Annals of Statistics. 1970. Vol. 41, № 5. Pp. 783–801.

[15] Гаек Я., Шидак З. Теория ранговых критериев. М.: Наука, 1971. 375 с.

[16] Лоэв М. Теория вероятностей. М.: Иностранная Литература, 1962. 719 с.

Библиографическая ссылка

Алексеев Ю.Г., Бенинг В.Е. Эвристический вывод формулы для асимптотического дефекта в одной задаче регрессионного типа // Вестник ТвГУ. Серия: Прикладная математика. 2014. № 3. С. 37–54.

Сведения об авторах

1. **Алексеев Юрий Георгиевич**
аспирант факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова.

Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова.
E-mail: yury_ alekseev@yahoo.com

2. **Бенинг Владимир Евгеньевич**
профессор факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова, старший научный сотрудник ИПИ РАН.

Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова.
E-mail: bening@yandex.ru.

HEURISTIC PROOF OF THE FORMULA FOR THE ASYMPTOTIC DEFICIENCY IN THE CASE OF REGRESSION MODEL

Alekseev Yuriy Georgievich
PhD student of Computational Mathematics and Cybernetics faculty,
Lomonosov Moscow State University
Russia, 119992, Moscow, GSP-1, Vorobyovi gory, Lomonosov MSU.
E-mail: yury_alekseev@yahoo.com

Bening Vladimir Evgenyevich
Professor of Mathematical Statistics department,
Lomonosov Moscow State University
Russia, 119992, Moscow, GSP-1, Vorobyovi gory, Lomonosov MSU.
E-mail: bening@yandex.ru

Received 06.04.2014, revised 20.04.2014.

In the present paper, following [1], [2], we heuristically obtain a formula (see (3.16)) for the limit of the difference between the power of the asymptotically optimal test and the power of the asymptotically most powerful test based on the stochastic expansion of the logarithm of the likelihood ratio in the regression model. It is demonstrated that this difference tends to zero with arbitrary rate due to behavior of regression coefficients in contrast to regular cases for which this order equals n^{-1}.

Keywords: asymptotically efficient test, regression model, asymptotic deficiency, local asymptotic normality, likelihood ratio, stochastic expansion.

Bibliographic citation

Alekseev Yu.G., Bening V.Ye. Heuristic proof of the formula for the asymptotic deficiency in the case of regression model. *Vestnik TvGU. Seriya: Prikladnaya matematika* [Herald of Tver State University. Series: Applied Mathematics], 2014, no. 3, pp. 37–54. (in Russian)