Федеральное государственное автономное образовательное учреждение
высшего образования «Казанский (Приволжский) федеральный университет»

На правах рукописи

Заикин Артем Александрович

Асимптотическое разложение d-риска

01.01.05 – Теория вероятностей и математическая статистика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата физико-математических наук

Научный руководитель

к. ф.-м. н., доц.

Симушкин С.В.

Казань – 2017

Оглавление

Введение

Актуальность темы исследования. Существует ряд задач по применениям математической статистики, в которых имеется реальная последовательность статистических экспериментов с полученными результатами наблюдений случайных выборок и принятых решений относительно параметра вероятностной модели. Это задачи контроля качества, медицинской диагностики, анализа активности генов, исследования характеристик большого числа малых областей и т.д. Каждому статистическому эксперименту соответствует значение параметра, которое допускает трактовку как реализация случайного элемента. Качество решающей функции относительно значений этого параметра в байесовской статистике обычно определяется с помощью априорного риска. Априорный риск является интегральным показателем качества решающей функции, в то время как для многих задач, особенно связанных с проверкой гипотез, требуются более тонкие характеристики. Одной из таких характеристик является d-риск, описанный Л.Н.Большевым и представляющий собой условное среднее значение потерь относительно σ-алгебры, порожденной решающей функцией.

Актуальность математических задач, решаемых в диссертации, объясняется в первую очередь практической важностью использования процедур статистического вывода, гарантирующих заданные ограничения на функцию d-риска. Когда существует реальная последовательность статистических экспериментов, и когда естественно и, в большей степени, необходимо применять байесовские методы, выполнение ограничений на классическую функцию риска не решает проблему гарантийности статистического вывода по существу. В таких случаях надо гарантировать заданные ограничения не на среднюю величину потерь при фиксированном значении выводного параметра, а на среднюю величину потерь среди статистических экспериментов, завершившихся принятием решения одного и того же решения d. Построение оптимальной, в том или ином смысле, d-гарантийной процедуры зачастую связано с большими трудностями вычислительного характера. Естественно поставить вопрос о построении асимптотически оптимальных d-гарантийных процедур. Поскольку функция d-риска представляет собой условное математическое ожидание от апостериорного риска, асимптотический анализ поведения функции d-риска не возможен без соответствующего анализа апостериорного распределения.

Асимптотический анализ апостериорного распределения является основным инструментом байесовской статистики, с помощью которого решаются задачи построения асимптотически оптимальных статистических решений и асимптотического анализа их риска. Основным результатом, на который опирается большинство таких исследований, является теорема Бернштейна-фон Мизеса, которая состоит в асимптотической нормальности апостериорного распределения. Точности этого утверждения бывает недостаточно, и приходится прибегать к асимптотическим разложениям апостериорного распределения.

В современной литературе было предложено несколько асимптотических разложений апостериорного распределения, имеющих различные условия применимости, виды разложений и характеристики остатков.

Как теорема Бернштейна-фон Мизеса, так и асимптотические разложения допускают различные центрирования параметра. Полученные до сих пор разложения используют центрирование оценкой максимального правдоподобия, что затрудняет их применение для задач, связанных с вычислением асимптотики d-риска. Центрирование фиксированным (истинным) значением параметра представляется более удобным для вычисления d-риска в задаче проверки параметрических гипотез. Одной из интересных задач, решенных в диссертации, является задача выделения класса центрирующих функций, для которых справедливо асимптотическое разложение.

Важной принципиальной особенностью применения утверждений асимптотического характера при анализе поведения d-риска является требование равномерности остатков асимптотик и разложений апостериорного распределения. Поэтому актуальной задачей является также доказательство необходимой равномерности для полученных разложений.

Степень разработанности темы исследования.

Теория d-апостериорного подхода к проблеме статистического вывода была сформулирована в начале 80-ых годов (см. тезисы доклада [1]). Более подробно об этом подходе с введением понятия d-риска в общей проблеме статистического вывода описано в обзорной статье [2]. Естественно, что на начальном этапе развития нового подхода к определению риска рассматривались классические задачи математической статистики, подобные задачам проверки гипотез с минимальным d-риском и построения параметрических оценок с равномерно минимальным d-риском. Метод построения оценок с равномерно минимальным d-риском предложен в статьях [1], [3], [4]. Такие оценки существуют, как и в классическом подходе, только при наличии достаточных статистик. Асимптотические методы построения оценок с минимальным d-риском в схеме так называемых сжимающихся априорных распределений представлены в статье [5].

Наиболее проработанной частью d-апостериорного подхода является проблема различения двух гипотез о параметре распределения. В работе [6] приводится оптимальный критерий различения двух гипотез, фиксирующий d-риск первого рода и минимизирующий d-риск второго рода. В статье [7] доказывается замечательное свойство, связывающее d-риски d-оптимальной процедуры с вероятностями ошибок наиболее мощного критерия в задаче поиска оптимального момента остановки. Более близкими к диссертации по тематике являются работы, затронувшие асимптотику необходимого объема выборки (НОВ) при d-гарантийном различении двух гипотез: были получены формулы для НОВ в случае сжимающихся априори [8] и в случае фиксированного априорного распределения [9].

На сегодняшний день не опубликовано ни одного результата по асимптотическим разложениям d-риска.

Надо отметить, что методика работы с асимптотиками d-риска во всех случаях основывается на асимптотике апостериорного распределения и апостериорного риска.

Изучение асимптотического поведения апостериорного распределения восходит к известному результату Бернштейна-фон Мизеса об асимптотической нормальности апостериорного распределения (см., например, [10, с. 141]). Было предложено много вариаций этого утверждения, в том числе близкая по духу к результатам диссертации теорема в монографии Ле Кама [11, с. 619].

Первое разложение апостериорного распределение в достаточно широких условиях было предложено Джонсоном в его работе [12]. Им было найдено апостериорное разложение, старшие члены которого кроме оценки максимального правдоподобия содержали многочлены от производных статистики вклада и производных априорной плотности.

Однако тот простой факт, что даже в теореме Бернштейна-фон Мизеса распределение параметра может иметь любое среднее, отличающееся от оценки максимального правдоподобия не более чем на величину порядка $1/n$, показывает, что разложение Джонсона далеко не единственно даже в терминах главного члена разложения. Кроме того, так как итоговое разложение есть случайная величина, зависящая от выборки, то вполне естественно, что утверждение асимптотического характера должно содержать информацию о распределении коэффициентов разложения. Так, в теореме Бернштейна-фон Мизеса сходимость апостериорного распределения доказана в смысле сходимости по общей вариации (см., например, [10, с. 141]). Джонсон показал, что некотором n_0 коэффициенты и остаток разложения ограничены константой почти наверное для $n > n_0$.

Необходимо упомянуть другие разложения, связанные с апостериорным распределением, полученные ранее. Гусев в своей статье [13] находит разложение плотности апостериорного распределения в постановке, аналогичной Джонсону [12], при этом остаток, предложенного им разложения сходится равномерно по параметру. Ghosh в своей статье [14] показывает, что при дополнительных условиях разложение Джонсона также имеет равномерный остаток. В статье

Вэнга [15] делается аналогичное Джонсону разложение апостериорного распределения отклонения параметра от оценки максимального правдоподобия, однако другим путем — с применением равенства Штейна. Разложение при этом строится не на производных функций плотности, а на моментах апостериорного распределения. В более поздней статье [16] этот же автор сравнивает полученное им разложение с разложением Джонсона из статьи [12]. Оказалось, что члены разложения порядка $n^{-1/2}$ арифметически совпадают, в то время как уже начиная с n^{-1} члены разложения отличаются.

Цели и задачи диссертационной работы: Целью диссертационной работы является построение асимптотик и разложений d-риска для процедур оценивания и проверки гипотез и применение полученных асимптотик для решения смежных проблем, таких как нахождение приближения оценок с равномерно минимальным d-риском и нахождение асимптотики необходимого объема выборки для d-гарантийных процедур проверки гипотез.

Научная новизна. Асимптотическое разложение апостериорного распределения параметра, центрированного произвольной $\sqrt{n}$-состоятельной оценкой, является обобщением разложения Джонсона [12] в том смысле, что оценка максимального правдоподобия является $\sqrt{n}$-состоятельной оценкой.

Разложение d-рисков для проверки гипотез в схеме испытаний Бернулли является первым асимптотическим разложением d-риска.

Предложен новый подход к построению оценок с асимптотически равномерно минимальным d-риском, в частности, предложен новый метод к доказательству того, что оценка максимального правдоподобия имеет асимптотически равномерно минимальный d-риск.

Теоретическая и практическая значимость. Асимптотическое разложение апостериорного распределения является важным шагом для получения стохастических разложений оценок при больших объемах наблюдений. См. в связи с этим статью Бурнашева [17], а также статьи [13], [14], в которых применяются разложения апостериорного риска, легко получаемые из разложения апостериорного распределения.

Определение оценки с асимптотически равномерно минимальным d-риском является новым в своем роде определением оптимальности, относящейся к d-риску. Оно может быть использовано для формулировки новых определений оптимальности и получения связанных с ними результатов, в том числе и для неасимптотических случаев.

Для проверки гипотез в схеме испытаний Бернулли получено несколько результатов, касающихся поведения необходимого объема выборки, вероятностей ошибок и величины d-рисков в зависимости от значений параметров и вида проверяемых гипотез.

Основные результаты:

В диссертации проводится построение новое асимптотическое разложение апостериорного распределения параметра, отличительной чертой которого является центрирование истинным значением параметра. Приводится асимптотическое разложение апостериорного распределения парметра, центрированного произвольной $\sqrt{n}$-состоятельной оценкой. Для последнего разложения доказывается равномерная сходимость относительно истинных значений параметра распределения.

В диссертации выведена асимптотическая формула для необходимого объема выборки при d-гарантийном различении гипотез, аналогичная [9], при несколько иных условиях на вероятностную модель и доказательство которой опирается на результаты, полученные в диссертации.

Для схемы испытаний Бернулли выводится асимптотика дефекта размера оптимального критерия в классе нерандомизированных критериев. Для рандомизированного критерия приводится асимптотическое (по величине области безразличия) разложение необходимого объема выборки. Для этого же критерия строятся асимптотические разложения для d-рисков до порядка $n^{-3/2}$ включительно. Критерий, гарантирующий ограничения на вероятности ошибок первого и второго рода и критерий, гарантирующий ограничения на d-риски сравниваются на основе величины области безразличия и параметров априорного распределения.

Сформулировано определение оценки с асимптотически равномерно минимальным d-риском и доказано, что оценка максимального правдоподобия является таковой в классе $\sqrt{n}$-состоятельных оценок параметра.

Степень достоверности и апробация результатов. Основные результаты диссертации докладывались на следующих конференциях: Международная научная конференция студентов, аспирантов и молодых учёных «Ломоносов-2015»; Двенадцатая международная Казанская летняя научная школа-конференция «Теория функций, ее приложения и смежные вопросы»; Международная конференция по алгебре, анализу и геометрии и их приложениям. Был сделан также доклад на городском семинаре Санкт-Петербурга по теории вероятностей и математической статистике.

Публикации. Материалы диссертации опубликованы в статьях [18], [19], [20] и в тезисах конференций [21], [22].

Личный вклад автора. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Все представленные в диссертации результаты получены лично автором.

Структура и объем диссертации. Диссертация состоит из введения, 4 глав, заключения и библиографии. Общий объем диссертации 127 страниц. В диссертации содержится 3 рисунка и 1 таблица. Библиография включает 37 наименований.

Краткое содержание диссертации: Во всех главах вероятностная модель представлена выборкой $\mathbf{X} = (X_1, X_2, ..., X_n)$ с распределением наблюдения $\mathbf{P}_\theta$. Мы предполагаем всюду, что у $\mathbf{P}_\theta$ существует плотность $p(x|\theta)$ по некоторой мере ν. Значение параметра θ неизвестно и является реализацией случайной величины ϑ из абсолютно непрерывного распределения $\mathbf{G}$ с плотностью $g(\theta)$ по лебеговской мере. Апостериорное распределение определяется следующим образом:

$$\mathbf{P}\left(\vartheta \in A|\mathbf{X}\right) = \frac{\int_A \prod_{i=1}^n p\left(X_i|\theta\right) g(\theta) d\theta}{\int_\Theta \prod_{i=1}^n p\left(X_i|\theta\right) g(\theta) d\theta}.$$

В первой главе диссертации собраны результаты, относящиеся к асимптотическому поведению апостериорного распределения и апостериорного риска. Первая глава разделена на две части. Первая часть содержит результаты, относящиеся к разложению апостериорного распределения параметра, центрированного фиксированным значением параметра.

Условия $\mathbb{D}_{\theta_0}(M)$, упоминаемые в следующей теореме, перечислены на странице 22. Явный вид коэффициентов $H_m(z)$ приведен на странице 25.

Теорема 0.1. *Пусть семейство распределений* $\{\mathbf{P}_\theta, \theta \in \Theta\}$ *удовлетворяет условиям* $\mathbb{D}_{\theta_0}(M)$. *Тогда*

$$\mathbf{P}\left(\sqrt{n}(\vartheta - \theta_0) < z|\mathbf{X}\right) = \sum_{m=0}^{M} n^{-m/2} H_m(z) + n^{-(M+1)/2}\omega_n(z), \tag{1}$$

причем

$$\sup_z H_m(z) = O_{\mathbf{P}_{\theta_0}}(1), m = 0, \ldots, M,$$

$$\sup_z \omega_n(z) = O_{\mathbf{P}_{\theta_0}}(1),$$

где $O_{\mathbf{Q}}(1)$ *значит плотную (равномерно ограниченную) последовательность случайных величин относительно меры* $\mathbf{Q}$.

Это разложение уточняет теорему Бернштейна-фон Мизеса в том смысле, что главный член асимптотики (1) совпадает со значением в точке z предельной функции распределения упомянутой теоремы.

Метод построения указанного разложения основан на технике анализа асимптотического поведения логарифма правдоподобия в условиях локальной асимптотической нормальности, разработанной в трудах Ибрагимова и Хасьминского [23], [24] и модифицированной под наши цели.

Аналогичными методами находится разложение для математических ожиданий относительно апостериорного распределения для функций, имеющих полиномиальную мажоранту. Говорят, что функция $w : \mathbb{R} \mapsto \mathbb{R}$ имеет полиномиальную мажоранту, если существует такой полином $P(x)$, что $|w(x)| \leq P(x), \forall x \in \mathbb{R}$.

Коэффициенты разложения в следующей теореме приведены на странице 32.

Теорема 0.2. *Пусть семейство распределений $\{\mathbf{P}_\theta, \theta \in \Theta\}$ удовлетворяет условиям $\mathbb{D}_{\theta_0}(M)$. Тогда для любой функции w, имеющей полиномиальную мажоранту, найдется n_0, что для всех $n > n_0$ $\mathbf{P}_{\theta_0}$-почти наверное выполняется $\mathbf{E}\left\{w\left(\sqrt{n}(\vartheta - \theta_0)\right)\Big|\mathbf{X}\right\} < \infty$, и*

$$\mathbf{E}\left\{w\left(\sqrt{n}(\vartheta - \theta_0)\right)\Big|\mathbf{X}\right\} = \sum_{m=0}^{M} n^{-m/2} H_{m,w} + n^{-(M+1)/2}\omega_n, \tag{2}$$

причем

$$H_{m,w} = O_{\mathbf{P}_{\theta_0}}(1), m = 0, \ldots, M,$$

$$\omega_n = O_{\mathbf{P}_{\theta_0}}(1).$$

Замечательно, что теорема 0.2 не требует ввода дополнительных ограничений, и при этом содержит утверждение о существовании всех апостериорных моментов (случай $w(x) = x^r, r > 0$).

Во второй части первой главы получены результаты, относящиеся к асимптотике апостериорного распределения параметра, центрированного $\sqrt{n}$- состоятельной оценкой.

Определение 0.1. *Пусть K — некоторый компакт из Θ. Будем говорить, что оценка T_n принадлежит классу оценок $\mathbb{C}(K)$ параметра $\theta \in \Theta$, если T_n измерима и*

$$\forall \varepsilon > 0 \;\exists b > 0 \;\sup_n \mathbf{P}_{\theta,K}\left(\sqrt{n}|T_n - \theta| > b\right) < \varepsilon.$$

Тот факт, что оценка максимального правдоподобия, а также фиксированная точка являются $\sqrt{n}$-состоятельными оценками (оценка $T_n = \theta_0$ принадлежит $\mathbb{C}(\{\theta_0\})$), подразумевает, что полученные разложения обобщают более ранние результаты, в том числе полученные другими авторами. Однако для этого пришлось потребовать более строгие условия.

Следующая теорема устанавливает асимптотическую нормальность апостериорного распределения параметра, центрированного $\sqrt{n}$-состоятельной оценкой. Условия $\mathbb{D}(K)$ приведены на странице 37. Асимптотическое среднее обозначено $\mu_n = \Delta_1(T_n)(\sqrt{n}I(T_n))^{-1}$, асимптотическая дисперсия $\sigma_n^2 = (I(T_n))^{-1}$, где $\Delta_1(\theta)$ — функция вклада в точке θ, $I(\theta)$ — информация Фишера в точке θ.

Теорема 0.3. *Пусть K — компакт, лежащий в Θ вместе с некоторой своей окрестностью и оценка $T_n \in \mathbb{C}(K)$. Тогда при выполнении условий $\mathbb{D}(K)$*

$$\sup_{\theta\in K} \mathbf{P}_\theta \left(\left\| \mathbf{P}\left(\sqrt{n}(\vartheta - T_n) \in \cdot\ |\mathbf{X}\right) - \mathcal{N}\left(\ \cdot\ |\mu_n, \sigma_n^2\right) \right\|_{TV} > \varepsilon \right) \to 0,$$

где $\|\mathbf{P}(\cdot) - \mathbf{Q}(\cdot)\|_{TV}$ — расстояние по общей вариации между мерами $\mathbf{P}$ и $\mathbf{Q}$, а $\mathcal{N}(A|m, s^2)$ – распределение нормального закона со средним m и дисперсией s^2, вычисленное на борелевском множестве A.

Аналогичный результат этой теоремы был получен Ле Камом в его монографии [11]. Отличие заключается в том, что в теореме 0.3 сходимость имеет равномерный характер (по компакту K). Последнее обстоятельство играет существенную роль при выводе утверждений об асимптотике апостериорных средних, использующейся в главе 4.

Теорема 0.4. *Пусть выполнены условия $\mathbb{D}(K)$, оценка T_n такова, что величина $\sqrt{n}(T_n - \theta)$ имеет $r > 0$ моментов по распределению $\mathbf{P}_\theta, \theta \in K$, а функция w имеет полиномиальную мажоранту порядка не выше r. Тогда найдется n_0, что для всех $n > n_0$ и для всех $\varepsilon > 0$*

$$\sup_{\theta\in K} \mathbf{P}_\theta \left(\left| \mathbf{E}\left\{ w\left(\sqrt{n}(\vartheta - T_n)\right) \Big| \mathbf{X} \right\} - \int w(h)\varphi(h|\mu_n, \sigma_n^2)dh \right| > \varepsilon \right) \to 0,$$

где $\varphi(\cdot|\mu,\sigma)$ — функция плотности нормального распределения со средним μ и дисперсией σ^2.

Важным свойством является доказанная равномерность остатка по компакту K. В приложениях свойство асимптотической нормальности чаще всего целесообразно использовать в смысле сходимости по совместному распределением параметра и выборки $\mathbf{P}$.

Следствие 0.1. *Пусть для любого компакта $K \in \Theta$, входящего в Θ с некоторой окрестностью, оценка $T_n \in \mathbb{C}(K)$, и выполнены условия $\mathbb{D}(K)$. Тогда*

$$\left\|\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) \in B|\mathbf{X}\right) - \Phi\left(B|\mu_n, \sigma_n^2\right)\right\|_{TV} \xrightarrow{\mathbf{P}} 0.$$

Оставшаяся часть главы 1 посвящена асимптотическим разложениям. Для этого кроме условий $\mathbb{D}(K)$ необходимо потребовать существование дополнительных производных у плотности наблюдений и априорной плотности и соответствующих моментов. Зададимся некоторым неотрицательным целым числом M. Условия $\mathbb{D}(M,K)$, используемые для построения разложения, приведены на странице 52.

Теорема 0.5. *Пусть K — компакт, лежащий в Θ вместе с некоторой своей окрестностью и оценка $T_n \in \mathbb{C}(K)$. Пусть для некоторого целого $M \geq 0$ выполнены условия $\mathbb{D}(M,K)$. Тогда для любого $\varepsilon > 0$ найдется такое $C > 0$, что*

$$\sup_n \sup_{\theta\in K} \mathbf{P}_\theta\left[\left|\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) < z|\mathbf{X}\right) - \sum_{m=0}^{M} n^{-m/2} H_m(z)\right| > Cn^{-(M+1)/2}\right] < \varepsilon,$$

$$\sup_z H_m(z) = O_{\mathbf{P}_\theta}(1), \text{ равномерно по } \theta \in K,\ m = 0,\ldots,M.$$

Явный вид коэффициентов $H_m(z)$ приведен на странице 53.

Метод построения разложения теоремы 0.5 идентичен методу, примененному в теореме 0.1, однако доказательство равномерности по компакту K требует большего внимания работе с остатками разложений.

По аналогии с выводом теоремы 0.2 для случая центрирования фиксированной точкой, используя доказательство 0.5, выводится следующая

Теорема 0.6. *Пусть выполнены условия $\mathbb{D}(M, K)$, оценка T_n удовлетворяет $\sup_{\theta\in K}\mathbf{E}_\theta(\sqrt{n}|T_n-\theta|)^r<\infty, r>0$, а функция w имеет полиномиальную мажоранту, чей порядок не превышает r. Тогда существует n_0, для $n>n_0$ для любого $\varepsilon>0$ найдется такое $C>0$, что*

$$\sup_n\sup_{\theta\in K}\mathbf{P}_\theta\left[\left|\mathbf{E}\left\{w\left(\sqrt{n}(\vartheta-T_n)\right)\Big|\mathbf{X}\right\}-\sum_{m=0}^{M}n^{-m/2}H_{m,w}\right|>Cn^{-(M+1)/2}\right]<\varepsilon,$$

где коэффициенты $H_{m,w}$ определены на странице 61, причем

$$H_{m,w}=O_{\mathbf{P}_\theta}(1)\ \text{равномерно по}\ \theta\in K,\ m=0,\dots,M.$$

Для теорем 0.5 и 0.6 приведен аналог следствия 0.1, в котором утверждается справедливость асимптотических разложений по мере $\mathbf{P}$. Доказательство при этом остается идентичным.

Во второй главе ставится задача нахождения асимптотики (при стремящихся к нулю ограничениях на ошибки) необходимого объема выборки (НОВ) в задаче различения гипотез $H_0:\theta<\theta_0,\ \theta_0\in\Theta$ и $H_1:\theta\geq\theta_0$. Оптимальный критерий, гарантирующий ограничения на d-риски, известен [6], и основывается на статистике $R(\mathbf{X})=\mathbf{P}(\vartheta<\theta_0|\mathbf{X})$ — апостериорной вероятности того, что параметр принадлежит области нулевой гипотезы. Рассматривается ситуация, когда ограничения на d-риски $\beta_0,\beta_1\to 0$. В этом случае НОВ $n^*\to\infty$ и поэтому для анализа асимптотики d-рисков можно применять полученные ранее утверждения об асимптотическом распределении апостериорного распределения. В главе 2 используется лишь теорема Бернштейна-фон Мизеса, то есть асимптотическая нормальность апостериорного распределения.

Полученная в диссертации асимптотика совпадает с результатом статьи Володина и Новикова [9], однако использует несколько иной подход, и требует немного другие условия применения.

Упомянутые ниже в теореме условия $\mathbb{D}_{\theta_0}$ перечислены на странице 66.

Теорема 0.7. *Пусть выполнены условия* $\mathbb{D}_{\theta_0}$. *Пусть* $n^* = n^*(\beta_0, \beta_1)$ — *НОВ для различения гипотез* H_0 *и* H_1 *при ограничениях на d-риски* β_0 *и* β_1. *Пусть* $\tilde{c}$ *и* $\tilde{n}$ *определяются как* c *и* n, *которые являются решениями уравнений*

$$\frac{\varphi(q_c) + cq_c - q_c\beta_0}{\varphi(q_c) + cq_c - q_c(1-\beta_1)} = \frac{\beta_0 G_0}{\beta_1 G_1},$$

$$n = \frac{g(\theta_0)^2(2\varphi(q_c) + 2cq_c - q_c(1+\beta_0-\beta_1))^2}{I(\theta_0)(\beta_0 G_0 + \beta_1 G_1)^2}.$$

Если $\beta_0 \to 0$, $\beta_1 \to 0$ *так, что* $\beta_0/\beta_1 \to K > 0$, *то* $n^*/\tilde{n} \to 1$.

Третья глава посвящена приложению выработанных техник для случая, когда вероятностная модель сведена к схеме испытаний Бернулли с неизвестной вероятностью успеха θ. Так, рассматривается задача проверки гипотезы $H_0 : \theta < \theta_0$, и для гарантийной проверки этой гипотезы предлагаются два критерия: классический, для которого выполняются ограничения на вероятности ошибок, и d-апостериорный, который гарантирует ограничения на d-риски. Так как для классического гарантийного критерия необходимо задание альтернативы $H_1 : \theta > \theta_1$ с областью безразличия $\theta_1 > \theta_0$, а для d-гарантийного критерия необходимо задание априорного распределения $\mathbf{G}$ параметра θ, представляется естественным сравнение этих характеристик двух критериев (при прочих равных). Асимптотические результаты в этой главе получены с применением разложения Эджворта к распределению статистики $T = \sum_{i=1}^n X_i$.

Сначала в третьей главе оценивается дефект размера наилучшего критерия в классе нерандомизированных критериев, который основан на статистике T. Строится разложение вероятности ошибки до порядка n^{-1} включительно:

Теорема 0.8. *Если* $n \geq N$, *то при некотором* $R > 0$ *для дефекта критерия* $D(n) = \alpha - P(T \geq C_n(\alpha))$ *имеет место неравенство*

$$D(n) \leq \frac{V}{n} + R \cdot n^{-3/2},$$

где α — *уровень значимости,* $C_n(\alpha)$ — *критическая константа, соответствующая объему выборки* n *и уровню значимости* α, *а константа* V *известна и* *зависит только от* θ_0 *и* α.

Оценивается также асимптотика для минимального объема наблюдений $n(\alpha, \beta)$, при котором гарантируются ограничения на вероятности ошибок α и β для нерандомизированного критерия при малых размерах области безразличия:

Теорема 0.9. *Существует такая положительная константа* Δ^*, *что при всех* $\theta_1 - \theta_0 = \Delta < \Delta^*$ *значение* $n(\alpha, \beta)$ *не превосходит величины*

$$\frac{a^2}{\Delta^2} + \frac{2(b+2)}{\Delta} + \frac{2ca - (b+2)^2}{a^2} + 1,$$

где коэффициенты a, b, c *определены и зависят только от* $\theta_0, \theta_1, \alpha, \beta$.

Для рандомизированного критерия выводится асимптотическое разложение НОВ.

Теорема 0.10. *Пусть* n *– необходимый объем выборки рандомизированного критерия. Тогда существует такое число* Δ^*, *что для всех* $\Delta < \Delta^*$ *выполняется* $|n - n^*| \leq 1$, *где*

$$n^* = \frac{a^2}{\Delta^2} + \frac{2b}{\Delta} + \frac{2c^*a - b^2}{a^2}.$$

При d-гарантийной проверке гипотез необходимо введение априорного распределения. Оно полагается равным бета-распределению с известными параметрами a и b. Таким образом, проверяется гипотеза $H_0 : \theta < \theta_0$ против альтернативы $H_1 : \theta \geq \theta_0$. Для этого используется нормированная статистика

$$T_n = \frac{\sum_{i=1}^n X_i + U - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}},$$

где $U \sim \mathbb{U}([0;1])$ и не зависит от наблюдений X_i.

Для оптимального (основанного на T_n) критерия выводится асимптотическое разложение d-рисков

$$\mathcal{R}_0 = \mathbf{P}\left(\vartheta < \theta_0 | T_n \geq C\right), \quad \mathcal{R}_1 = \mathbf{P}\left(\vartheta \geq \theta_0 | T_n < C\right)$$

до порядка $n^{-3/2}$ включительно:

Теорема 0.11. *Пусть фиксировано C. Тогда для d-рисков $\mathcal{R}_0$ и $\mathcal{R}_1$ справедливы следующие асимптотические представления:*

$$\mathcal{R}_0 = \frac{Q_1}{\sqrt{n}} + \frac{Q_2}{n} + \frac{Q_3}{n^{3/2}} + O\left(n^{-2}\right),$$

$$\mathcal{R}_1 = \frac{W_1}{\sqrt{n}} + \frac{W_2}{n} + \frac{W_3}{n^{3/2}} + O\left(n^{-2}\right),$$

где явный вид коэффициентов $Q_1, Q_2, Q_3, W_1, W_2, W_3$ приведен на странице 99. Эти коэффициенты зависят только от C, θ_0 и параметров априорного распределения a и b.

В конце третьей главы проводится сравнение двух подходов на основе величины зоны безразличия и параметров априорного распределения при равных НОВ и ограничениях на вероятности ошибок и d-риски. Приводятся способ сравнения и графики, демонстрирующие различие двух критериев.

Последняя глава посвящена оценкам, оптимизирующим d-риск. Так как работа с оценками с равномерно минимальным d-риском сопряжена с не преодоленными пока трудностями, вводится новое понятие оценок с асимптотически равномерно минимальным d-риском.

Определение 0.2. *Измеримая оценка δ_n^* называется оценкой с ϕ_n-асимптотически ($\phi_n \to 0$) равномерно минимальным d-риском в классе оценок $\mathbb{T}$, если*

$$\mathbf{P}\left(\mathcal{R}_{\delta_n^*}(\delta_n^*) \geq \mathcal{R}_{\delta_n}(\delta_n^*) + \phi_n\right) \to 0$$

для любой другой оценки $\delta_n \in \mathbb{T}$.

Идея, преследуемая в главе — показать, что оценка максимального правдоподобия (ОМП) является оценкой с асимптотически равномерно минимальным d-риском. Само понятие введено потому как неясно, как доказать близость ОМП и оценки с равномерно минимальным d-риском. Более того, условия существования оценок с равномерно минимальным d-риском неизвестны, поэтому приходится обходиться только близостью d-риска ОМП к оптимальному уровню.

Используя результаты главы 1, касающиеся поведения апостериорного риска, удалось доказать два следующих результата. Условия, выдвигаемые к вероятностной модели, совпадают с условиями $\mathbb{D}(\Theta)$ (более точно, должны быть выполнены условия $\mathbb{D}(K)$ для любого внутреннего компакта $K \subset \Theta$) асимптотической нормальности апостериорного распределения параметра, центрированного $\sqrt{n}$-состоятельной оценкой.

Теорема 0.12. *Пусть выполнены условия $\mathbb{D}(\Theta)$, и задана функция потерь $L(\theta_1, \theta_2) = |\theta_1 - \theta_2|^r, r > 1$. Тогда найдется последовательность $\phi_n = o(n^{-r/2})$, такая, что оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с ϕ_n-асимптотически равномерно минимальным d-риском в классе оценок T_n, удовлетворяющих $\sup_{\theta \in K} \mathbf{E}_\theta(\sqrt{n}|T_n - \theta|)^r < \infty$.*

Теорема 0.13. *Пусть выполнены условия $\mathbb{D}(\Theta)$ и пусть Θ ограничено. Тогда найдется последовательность $\phi_n = o(1)$, такая, что оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с ϕ_n-асимптотически равномерно минимальным d-риском в классе $\sqrt{n}$-состоятельных оценок.*

Глава 1

Разложение апостериорного распределения

Результат об асимптотической нормальности апостериорного распределения известен давно как теорема Бернштейна-фон Мизеса. В данной главе этот результат уточняется, подобно тому, как это делается в работах Джонсона [12], Гусева [13], Гоша [14] и Вэнга [15]. Однако, в диссертации разложение делается центрированием не посредством оценки максимального правдоподобия, а центрированием выводным параметром и $\sqrt{n}$-состоятельной оценкой, что позволяет, в отличие от известных разложений, использовать полученные разложения для асимптотического анализа d-риска.

В этой главе выводятся асимптотические разложения для апостериорных распределений и апостериорных рисков. Глава разделена на две части. Первая часть посвящена выводу разложения апостериорного распределения и апостериорных моментов параметра, центрированного истинным значением параметра. Вторая часть посвящена выводу разложения апостериорного распределения и апостериорных моментов параметра, центрированного произвольной $\sqrt{n}$-состоятельной оценкой. Начнем с описания статистической задачи, в рамках которой проводятся все дальнейшие построения в диссертации.

Пусть $(\mathfrak{X}, \mathfrak{A})$ — измеримое пространство с заданным на нем семейством вероятностных мер $\{\mathbf{P}_\theta,\ \theta \in \Theta \subset \mathbb{R}\}$. Предполагается, что у распределения $\mathbf{P}_\theta$ существует плотность $p(x|\theta) = d\mathbf{P}_\theta/d\nu$ по некоторой σ-конечной мере ν, на $\mathfrak{A}$. В статистическом эксперименте наблюдается выборка $\mathbf{X} = (X_1, X_2, ..., X_n)$, элементы которой являются отображениями в пространство $(\mathfrak{X}, \mathfrak{A}, \mathbf{P}_\theta)$. Значение параметра θ неизвестно и является реализацией случайной величины ϑ, имеющей абсолютно непрерывную функцию распределения $\mathbf{G}$ с плотностью $g(\theta)$ по лебеговской мере. Пусть Θ — некоторое измеримое (по мере Лебега) подмножество числовой прямой. Всюду далее мы будем предполагать, что априорная плотность $g(\theta)$ задана так, что $g(\theta) > 0$ для всех $\theta \in \Theta$. Предполагается, что функция $p(x|\theta)$ задана всюду на $\mathfrak{X} \times \Theta$ и измерима относительно прямого произведения σ-алгебр $\mathfrak{A} \otimes \mathfrak{B}$, где $\mathfrak{B}$ — борелевская сигма-алгебра на Θ. Основным вероятностным пространством, использующимся в работе, будет $(\mathfrak{X}^\infty \times \Theta, \mathfrak{A}^\infty \otimes \mathfrak{B}, \mathbf{P})$, где $\mathbf{P}$ согласовано с условными распределениями $\mathbf{P}_\theta$ и распределением $\mathbf{G}$. Для упрощения записи будем записывать совместное распределение выборки $\mathbf{X}$ и параметра ϑ тем же символом $\mathbf{P}$, а условное распределение выборки при фиксированном параметре θ как $\mathbf{P}_\theta$. Соответствующие вероятностным мерам $\mathbf{P}_\theta$ и $\mathbf{P}$ математические ожидания обозначаются символами $\mathbf{E}_\theta$ и $\mathbf{E}$ соответственно. Плотность выборки $\mathbf{X}$ из $\mathbf{P}_\theta$ записывается как $p_n(x^{(n)})|\theta) = \prod_{i=1}^n p(x_i|\theta)$, $x^{(n)} = (x_1, \dots, x_n)$.

В данной главе изучаются асимптотические свойства функции апостериорного распределения, определенного для всех $A \in \mathfrak{B}$ по формуле

$$\mathbf{P}\left(\vartheta \in A|\mathbf{X}\right) = \frac{\int_A p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}{\int_\Theta p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}. \tag{1.1}$$

Приведем следующие обозначения, которые используются на протяжении всей главы. Пусть θ_0 — некоторая точка множества Θ, и пусть

$$l_m(x|\theta) = \left(\frac{\partial}{\partial\theta}\right)^m \ln p(x|\theta), \quad l_m(x) = l_m(x|\theta_0),$$

$$\Delta_m(\theta) = \sum_{i=1}^n l_m(x_i|\theta), \quad \Delta_m = \Delta_m(\theta_0),$$

$$\pi_m(\theta) = \left(\frac{d}{d\theta}\right)^m \ln g(\theta), \quad \pi_m = \pi_m(\theta_0),$$

$$Z_n(\theta, h) = \frac{p_n\left(x^{(n)}|\theta + h/\sqrt{n}\right)}{p_n\left(x^{(n)}|\theta\right)}, \quad \theta + h/\sqrt{n} \in \Theta.$$

$$\widetilde{Z}_n(\theta, h) = \frac{p_n\left(x^{(n)}|\theta + h/\sqrt{n}\right)}{p_n\left(x^{(n)}|\theta\right)} \frac{g\left(\theta + h/\sqrt{n}\right)}{g(\theta)}, \quad \theta + h/\sqrt{n} \in \Theta.$$

Как обычно, равенство $U_n = O_{\mathbf{P}}(1)$, означает, что последовательность случайных величин $U_n, n \geq 1$, равномерно плотна (ограничена по вероятности), т.е.

$$\forall \varepsilon > 0\ \exists C > 0\colon\ \sup_n \mathbf{P}(|U_n| > C) < \varepsilon.$$

1.1. Разложение апостериорного распределения с фиксированным центрированием параметра

В этом параграфе рассматривается задача нахождения асимптотического разложения апостериорного распределения по степеням $n^{-1/2}$. В качестве центра сосредоточения выбирается фиксированная (истинная) точка θ_0. Другими словами, для заданного целого $M \geq 0$ указать такие функции $H_m\left(z, x^{(n)}\right)$ $\mathbb{R} \times \mathfrak{X}^n \mapsto \mathbb{R}$, что для любого $\varepsilon > 0$, существует $C > 0$ такое, что для всех z

$$\sup_n \mathbf{P}_{\theta_0} \left\{ \left| \mathbf{P}\left(\sqrt{n}(\vartheta - \theta_0) < z | \mathbf{X}\right) - \sum_{m=0}^{M} n^{-m/2} H_m\left(z, \mathbf{X}\right) \right| > C n^{-(M+1)/2} \right\} < \varepsilon,$$

$$\sup_n \mathbf{P}_{\theta_0} \left\{ \left| H_m\left(z, \mathbf{X}\right) \right| > C n^{-m/2} \right\} < \varepsilon,\ 0 \leq m \leq M.$$

Отметим, что в приведенных здесь соотношениях параметр вероятностной меры $\mathbf{P}_{\theta_0}$ фиксирован, поэтому можно говорить, что центрирование ведется истинным значением параметра θ_0.

1.1.1. Условия регулярности и предварительные леммы

Перечислим условия, на плотность наблюдений и априорную плотность, при которых будет решаться задача построения указанного разложения. Легко заметить, что эти условия в большей степени повторяют условия статьи Гусева [13]. В отличие от [13] g есть функция плотности, а не некая весовая функция. Будем называть совокупность всех ниже перечисленных условий как $\mathbb{D}_{\theta_0}(M)$, где θ_0 — некоторая внутренняя точка Θ, а $M \geq 0$ — некоторое целое число.

1. Для $\theta \neq \theta'$
$$\int_{\mathfrak{X}} \left| p(x|\theta) - p(x|\theta') \right| \nu(dx) > 0.$$

2. Найдется $\delta_1 > 0$, такое что
$$\sup_{\theta \in \Theta} |\theta - \theta_0|^{\delta_1} \int_{\mathfrak{X}} \sqrt{p(x|\theta) p(x|\theta_0)} \nu(dx) < \infty.$$

3. Функция $p(x|\theta)$ определена для каждого $x \in \mathfrak{X}$ и имеет $M + 3$ ($M \geq 0$) непрерывных частных производных по θ для всех θ из внутренности Θ.

4. Пусть $I(\theta) = \mathbf{E}_\theta \left(l_1(X|\theta)\right)^2$ — информация Фишера.

   а. $0 < I(\theta) < \infty \quad \forall \theta \in \Theta$.

   б. $\mathbf{E}_{\theta_0} \left(l_2(X) - \mathbf{E} l_2(X)\right)^2 < \infty$.

   в. $\mathbf{E}_{\theta_0} |l_i(X)| < \infty, \ i = 3, ..., M+2$.

   г. Для некоторого $\delta_2 > 0$

$$\sup_{\lambda \in [\theta_0 - \delta_2; \theta_0 + \delta_2]} \mathbf{E}_{\theta_0} |l_{M+3}(X|\lambda)| < \infty.$$

5. Для некоторого $\delta_3 \geq 0$

$$\sup_{\theta \in \Theta} (1 + |\theta|^{\delta_3})^{-1} I(\theta) < \infty.$$

6. Априорная плотность $g(\theta)$ имеет $M + 1$ непрерывную производную во внутренности Θ.

Условие 4б до сих пор не вводилось в предыдущих работах по разложениям апостериорного распределения. Здесь оно введено для удобства записи асимптотической дисперсии (в варианте, представленном в теореме 1.1 этой главы, в качестве асимптотической дисперсии используется обратная информация Фишера $I^{-1}(\theta_0)$, в то время как возможно использование соответствующего выборочного момента $-n^{-1} \sum_{i=1}^n l_2(x_i|\theta_0)$, как, например, это было сделано Джонсоном [12]).

Теперь сформулируем несколько лемм, которые будут использованы при построении асимптотических разложений.

Лемма 1.1. *Пусть выполнены условия 1, 2, 4а, 5. Тогда для любых m, r найдется натуральное число n_0, константа $C_{m,r}$, зависящая только от m и r, такие, что для всех $n > n_0$ и $A > 0$ выполняется*

$$\mathbf{P}_{\theta_0}\left(\int\limits_{|h|>A}|h|^r Z_n(\theta_0,h)g\left(\theta_0+\frac{h}{\sqrt{n}}\right)dh > A^{-m}\right) \le C_{m,r}A^{-m-r}. \tag{1.2}$$

Доказательство. Разобьем область интегрирования $\{|h|>A\}$ на интервалы $|h|\in[A;A+1],[A+1;A+2],\ldots$:

$$\mathbf{P}_{\theta_0}\left(\int\limits_{|h|>A}|h|^r Z_n(\theta_0,h)g\left(\theta_0+\frac{h}{\sqrt{n}}\right)dh \ge \frac{1}{A^m}\right) \le$$
$$\le \sum_{k=A}^{\infty}\mathbf{P}_{\theta_0}\left(\sup_{k\le|h|\le k+1} Z_n(\theta_0,h) \ge \frac{1}{(k+1)^r k^m}\right). \tag{1.3}$$

В последнем выражении для сокращения записи положено $\sum_{k=A}^{\infty}b(k)=b(A)+b(A+1)+\ldots$ (A не обязательно целое число). Согласно теореме 2.3 статьи [23]

$$\mathbf{P}_{\theta_0}\left(\sup_{k\le|h|\le k+1} Z_n(\theta_0,h) \ge \frac{1}{k^m}\right) \le C_m k^{-m},$$

где константа C_m не зависит от k. Применяя это неравенство к правой части выражения (1.3), получим

$$\sum_{k=A}^{\infty}\mathbf{P}_{\theta_0}\left(\sup_{k\le|h|\le k+1} Z_n(\theta_0,h) \ge \frac{1}{(k+1)^r k^m}\right) \le \sum_{k=A}^{\infty}\frac{C_{m,r}}{k^{r+m}} \le \frac{C'_{m,r}}{A^{r+m}},$$

где константы $C_{m,r}$ и $C'_{m,r}$ не зависят от A, что доказывает утверждение леммы. ∎

В связи с леммой 1.1 см. также лемму 4 работы [25]), в которой также доказывается экспоненциальный характер убывания хвостов $Z_n(\theta_0,h)$ при больших h.

Следующая лемма общего характера применяется при анализе плотности семейства стохастических последовательностей.

Лемма 1.2. *Пусть дан конечный набор плотных по мере* $\mathbf{Q}$ *последовательностей* $Y_1 = (Y_{1,i})_{i=1}^{\infty}, ..., Y_K = (Y_{K,i})_{i=1}^{\infty}$. *Тогда*

$$Z_i = \prod_{j=1}^{K} Y_{j,i} = O_{\mathbf{Q}}(1), \quad S_i = \sum_{j=1}^{K} Y_{j,i} = O_{\mathbf{Q}}(1).$$

Доказательство.

Выпишем простое неравенство для вероятностей произведения случайных величин:

$$\mathbf{Q}(|Z_i| > C) \leq \sum_{j=1}^{K} \mathbf{Q}\left(|Y_{j,i}| > C^{1/K}\right).$$

Из конечности суммы следует результат о плотности последовательности Z_i.

Аналогично для суммы имеем

$$\mathbf{Q}(|S_i| > C) \leq \sum_{j=1}^{K} \mathbf{Q}\left(|Y_{j,i}| > C/K\right).$$

■

Из леммы 1.2 следует, что линейная комбинация элементов плотных последовательностей тоже плотна.

1.1.2. Асимптотическое разложение апостериорного распределения

Пусть $\Phi(x)$ и $\varphi(x)$ — соответственно функия распределения и функция плотности стандартного нормального закона, $I_1(m, v)$ – множество индексов $(i_1, ..., i_m)$, которые удовлетворяют тождеству

$$\sum_{j=1}^{m} i_j = v, \ i_j > 0, \ j = 1, \ldots, m.$$

Пусть также $\Phi(x|\mu, \sigma^2)$ и $\varphi(x|\mu, \sigma^2), x \in \mathbb{R}$ — соответственно функции распределения и плотности нормального закона со средним μ и дисперсией σ^2,

$$I = I(\theta_0), \ F_m(z|\mu, \sigma^2) = \int_{-\infty}^{z} h^m \varphi(h|\mu, \sigma^2) dh, \ z \in \mathbb{R},$$

$$\mu_n = \frac{\Delta_1}{\sqrt{n}I}, \sigma_n = I^{-1/2},$$

$$H_0(z) = f_0(z) = \Phi\left(z|\mu_n, \sigma_n^2\right),$$

$$H_m(z) = f_m(z) + \sum_{j=1}^{m} V_j f_{m-j}(z), \quad m = 1, \dots, M,$$

$$V_m = \sum_{j=1}^{m} (-1)^j \sum_{I_1(j,m)} \prod_{l=1}^{j} f_{i_l}(\infty), \quad m = 1, \dots, M,$$

$$f_m(z) = \sum_{j=1}^{m} \frac{1}{m!} \sum_{I_1(j,m)} \prod_{l=1}^{j} q_{i_l}(z), \quad m = 1, \dots, M,$$

$$q_1(z) = \sqrt{n} F_2(z|\mu_n, \sigma_n^2) \frac{n^{-1}\Delta_2 + I}{2} + F_3(z|\mu_n, \sigma_n^2) \frac{\Delta_3}{6n} + F_1(z|\mu_n, \sigma_n^2)\pi_1,$$

$$q_m(z) = F_{m+2}(z|\mu_n, \sigma_n^2) \frac{\Delta_{m+2}}{n(m+2)!} + F_m(z|\mu_n, \sigma_n^2) \frac{\pi_m}{m!}, \quad m = 2, \dots, M.$$

Величины $q_m(\infty)$ по определению полагаются равными $\lim_{z\to\infty} q_m(z)$. Заметим, что так как $\Delta_m, m \geq 1$ являются случайными величинами, то $H_m(z), m = 0, \dots, M$ также являются случайными величинами. Зависимость от $\mathbf{X}$ в записи функций H_m опущена с целью избежать перегруженности формул.

Теорема 1.1. *Пусть семейство распределений $\{\mathbf{P}_\theta, \theta \in \Theta\}$ удовлетворяет условиям $\mathbb{D}_{\theta_0}(M)$. Тогда*

$$\mathbf{P}\left(\sqrt{n}(\vartheta - \theta_0) < z|\mathbf{X}\right) = \sum_{m=0}^{M} n^{-m/2} H_m(z) + n^{-(M+1)/2}\omega_n(z), \tag{1.4}$$

причем

$$\sup_z H_m(z) = O_{\mathbf{P}_{\theta_0}}(1), m = 0, \dots, M,$$

$$\sup_z \omega_n(z) = O_{\mathbf{P}_{\theta_0}}(1).$$

Доказательство. Схема доказательства следующая: подынтегральное выражение в числителе и знаменателе (1.1) приводится к виду $\exp(\ln \widetilde{Z}_n(\theta_0, h))$, после чего раскладывается в ряд Тейлора $\ln \widetilde{Z}_n(\theta_0, h)$ относительно $h = 0$. Снова раскладываем в ряд Тейлора уже экспоненту относительно первых двух членов разложения логарифма (члены с h и h^2). Домножим оставшуюся экспоненту под знаком интеграла на множитель так, чтобы получилась плотность некоего нормального закона. Чтобы получить результат теоремы, доказываем, что остатки после интегрирования необходимо малы, а коэффициенты плотны относительно $\mathbf{P}_{\theta_0}$. После проведения интегрирования в числителе и знаменателе получаем рациональную функцию, которую приводим к искомому многочлену опять же путем разложения в ряд Тейлора.

Так как в определении теоремы присутствует масштабирование по параметру на $\sqrt{n}$, обозначим область интегрирования $\Theta_n = \sqrt{n}(\Theta - \theta_0)$. Ясно, что $\Theta_n \uparrow \mathbb{R}$, так что для упрощения записи везде, где происходит интегрирование по масштабированному параметру, считаем, что интегрирование происходит по указанной области, пересеченной с Θ_n, если не указано обратного. С учетом этого перепишем (1.1), сразу делая замену $\theta = \theta_0 + h/\sqrt{n}$. В числителе множитель $n^{-1/2}$ перед интегралом сократится с соответствующим множителем в знаменателе.

$$\mathbf{P}\left(\sqrt{n}(\vartheta - \theta_0) < z|\mathbf{X}\right) = \frac{\int_{-\infty}^{z} p_n\left(\mathbf{X}|\theta_0 + h/\sqrt{n}\right) g(\theta_0 + h/\sqrt{n})dh}{\int_{\mathbb{R}} p_n\left(\mathbf{X}|\theta_0 + h/\sqrt{n}\right) g(\theta_0 + h/\sqrt{n})dh}. \tag{1.5}$$

Разделим числитель и знаменатель (1.5) на величину $p_n(\mathbf{X}|\theta_0)g(\theta_0)$. В силу $\mathbf{P}_{\theta_0}\left(p_n(\mathbf{X}|\theta_0) = 0\right) = 0$ и условия $g(\theta) > 0$ для всех $\theta \in \Theta$ делением на ноль можно пренебречь. Числитель (1.5) после преобразования примет вид

$$\int_{-\infty}^{z} \widetilde{Z}_n(\theta_0, h)dh.$$

Согласно лемме 1.1 для $0 < \alpha < 1/2$, любой константы b и при некотором n_0 для $n > n_0$

$$\int_{-\infty}^{\infty} \widetilde{Z}_n(\theta_0, h)dh = \int_{-bn^{\alpha}}^{bn^{\alpha}} \widetilde{Z}_n(\theta_0, h)dh + n^{-(M+1)/2} O_{\mathbf{P}_{\theta_0}}(1). \tag{1.6}$$

Выпишем многочлен Тейлора логарифма $\widetilde{Z}_n(\theta_0, h)$ в точке $h = 0$ с остатком в форме Лагранжа:

$$\ln \widetilde{Z}_n(\theta_0, h) = \sum_{j=1}^{M+2} \left(\frac{h}{\sqrt{n}}\right)^j \frac{\Delta_j}{j!} + \sum_{j=1}^{M} \left(\frac{h}{\sqrt{n}}\right)^j \frac{\pi_j}{j!} + r_n(h),$$

$$r_n(h) = n^{-(M+1)/2}\left[\frac{h^{M+3}}{n(M+3)!}\Delta_{M+3}\left(\theta_0 + \frac{h'}{\sqrt{n}}\right) + \frac{h^{M+1}}{(M+1)!}\pi_{M+1}\left(\theta_0 + \frac{h''}{\sqrt{n}}\right)\right],$$

где $h', h'' \in [-bn^\alpha; bn^\alpha]$. В силу условий 3 и 6 разложение существует для каждого h.

Полученное разложение логарифма $\widetilde{Z}_n(\theta_0, h)$ можно представить в виде

$$\ln \widetilde{Z}_n(\theta_0, h) = \psi(h) + L_n(h),$$

где

$$\psi(h) = \frac{h\Delta_1}{\sqrt{n}} - \frac{h^2 I}{2},$$

$$L_n(h) = \frac{h^2}{2}\left(\frac{\Delta_2}{n} + I\right) + \sum_{m=1}^{M} n^{-m/2}\left(\frac{h^{m+2}}{(m+2)!}\frac{\Delta_{m+2}}{n} + \frac{h^m}{m!}\pi_m\right) + r_n(h).$$

Разложим экспоненту $\exp(\ln Z_n(\theta_0, h))$ в ряд Тейлора относительно точки $\psi(h)$ с остатком в форме Лагранжа:

$$\exp\{\ln \widetilde{Z}_n(\theta_0, h)\} = \exp\{\psi(h)\}\left[1 + L_n(h) + \frac{1}{2}\left(L_n(h)\right)^2 + \ldots + \frac{1}{M!}\left(L_n(h)\right)^M\right] + \\ + \exp\left\{\psi(h) + \widehat{L}_n(h)\right\}\frac{1}{(M+1)!}\left(L_n(h)\right)^{M+1}, \quad (1.7)$$

$\widehat{L}_n(h)$ — некая точка между 0 и $L_n(h)$.

Полученное коэффициенты разложения (1.7) согласно (1.6) теперь необходимо проинтегрировать по $h \in (-bn^\alpha, -bn^\alpha)$, откуда уже получается разложение знаменателя (1.5). Чтобы получить утверждение теоремы, необходимо доказать плотность проинтегрированных коэффициентов.

Сперва докажем, что

$$\sup_n \mathbf{P}_{\theta_0}\left(\int_{-bn^\alpha}^{bn^\alpha} \exp\{\psi(h)\}|r_n(h)|dh > Cn^{-(M+1)/2}\right) \to 0, \ C \to \infty.$$

Так как $h''/\sqrt{n} \in [-bn^{\alpha-0.5}; bn^{\alpha-0.5}] \subset [-b; b]$ $(0 < \alpha < 0.5)$ и функция $\pi_{M+1}(\theta)$ непрерывна в точке θ_0 вследствие условия 6, то $|\pi_{M+1}(\theta_0 + h''/\sqrt{n})|$ ограничено некоторой константой $\widetilde{\pi}_{M+1}$, не зависящей от n. Таким образом,

$$|r_n(h)| \le n^{-(M+1)/2} \left| \frac{h^{M+3}}{n(M+3)!} \Delta_{M+3} \left(\theta_0 + \frac{h'}{\sqrt{n}} \right) + \frac{h^{M+1}}{(M+1)!} \widetilde{\pi}_{M+1} \right|.$$

Далее,

$$\mathbf{P}_{\theta_0} \left(\int\limits_{-bn^\alpha}^{bn^\alpha} \exp\{\psi(h)\} |r_n(h)| dh > Cn^{-(M+1)/2} \right) \le$$

$$\le \mathbf{P}_{\theta_0} \left(\left\{ \int\limits_{-bn^\alpha}^{bn^\alpha} \exp\{\psi(h)\} |r_n(h)| dh > Cn^{-(M+1)/2} \right\} \cap \left\{ \frac{|\Delta_1|}{\sqrt{n}} \le C \right\} \right) +$$

$$+ \mathbf{P}_{\theta_0} \left(\frac{|\Delta_1|}{\sqrt{n}} > C \right).$$

Второе слагаемое в правой части последнего неравенства по неравенству Чебышева и условию 4a ограничено величиной IC^{-2}. Первое слагаемое можно также ограничить с помощью неравенства Маркова следующей величиной:

$$C^{-1} \mathbf{E}_{\theta_0} \left\{ \int\limits_{-bn^\alpha}^{bn^\alpha} \exp\{\psi(h)\} \left| \frac{h^{M+3}}{n(M+3)!} \Delta_{M+3} \left(\theta_0 + \frac{h'}{\sqrt{n}} \right) + \right. \right.$$

$$\left. \left. + \frac{h^{M+1}}{(M+1)!} \widetilde{\pi}_{M+1} \right| dh; \quad n^{-1/2} |\Delta_1| \le C \right\}. \quad (1.8)$$

Математическое ожидание здесь вычисляется по множеству $\{n^{-1/2}|\Delta_1| \le C\}$, где $\widetilde{\psi}(h) \stackrel{def}{=} C|h| - Ih^2/2 \ge \psi(h)$. Величина $\widetilde{\psi}(h)$ не зависит от $x^{(n)}$ (как переменной интегрирования $\mathbf{E}_{\theta_0}$), и, следовательно, можно поменять порядок интегрирования. Оценим величину математического ожидания оставшегося после указанных манипуляций случайного слагаемого $\mathbf{E}_{\theta_0}|n^{-1}\Delta_{M+3}(\theta_0 + h'/\sqrt{n})|$. В силу $h'/\sqrt{n} \in [-bn^{\alpha-0.5}; bn^{\alpha-0.5}] \subset [-b; b]$ имеем

$$\mathbf{E}_{\theta_0} \left| n^{-1} \Delta_{M+3} \left(\theta_0 + \frac{h'}{\sqrt{n}} \right) \right| \le \mathbf{E}_{\theta_0} \frac{1}{n} \sum_{i=1}^{n} \left| l_{M+3} \left(X_i \middle| \theta_0 + \frac{h'}{\sqrt{n}} \right) \right| =$$

$$= \mathbf{E}_{\theta_0} \left| l_{M+3} \left(X \middle| \theta_0 + \frac{h'}{\sqrt{n}} \right) \right| \le \sup_{t \in [-b;b]} \mathbf{E}_{\theta_0} \left| l_{M+3} (\theta_0 + t) \right|.$$

Выбирая здесь $b = \delta_2$ из условия 4г, получаем, что последнее выражение ограничено константой, не зависящей от n и h.

Далее, так как $\widetilde{\psi}(h)$ — квадратичная функция от h, и $I > 0$ в силу условия 4а, интеграл по h сводится к интегралу $\int_{\mathbb{R}} h^m e^{-h^2} dh$ и поэтому существует. Таким образом, (1.8) есть величина $O(C^{-1})$, и требуемая плотность остатков установлена.

Используя неравенство Маркова и условие 4в, таким же образом доказывается, что коэффициенты в $L_n(h)$ при $n^{-m/2}$ плотны по мере $\mathbf{P}_{\theta_0}$:

$$\mathbf{P}_{\theta_0}\left(n^{-m/2}\left|\int\limits_{-bn^\alpha}^{bn^\alpha} \exp\{\psi(h)\}\left[h^{m+2}\frac{\Delta_{m+2}}{n(m+2)!} + \frac{h^m}{m!}\pi_m\right]dh\right| > Cn^{-m/2}\right) \leq$$
$$\leq \mathbf{P}_{\theta_0}\left(\left\{\left|n^{-1}\Delta_{m+2}\right| > (C - \pi_m E_m)E_{m+2}^{-1}\right\} \cap \left\{n^{-1/2}|\Delta_1| \leq C\right\}\right) +$$
$$+\mathbf{P}_{\theta_0}\left(n^{-1/2}|\Delta_1| > C\right) = O\left(C^{-1}\right) + O\left(C^{-2}\right) = O\left(C^{-1}\right),$$

где

$$E_m = \frac{1}{m!}\int\limits_{-\infty}^{\infty} |h|^m \exp\left\{\widetilde{\psi}(h)\right\} dh.$$

Первое слагаемое в представлении $L_n(h)$ также является элементом плотной последовательности. Действительно, по неравенству Чебышева в силу условия 4б (заметим, что $\mathbf{E}_{\theta_0} l_2(X) = -I$ вследствие условия 4б):

$$\mathbf{P}_{\theta_0}\left(\frac{E_2}{2}\left|\frac{\Delta_2}{n} + I\right| > Cn^{-1/2}\right) = \mathbf{P}_{\theta_0}\left(\frac{n}{4}\left(\frac{\Delta_2}{n} + I\right)^2 > E_2^{-2}C^2\right) \leq$$
$$\leq C^{-2}E_2^2\mathbf{E}_{\theta_0}^n n\left(\frac{\Delta_2}{n} + I\right)^2 = C^{-2}E_2^2\mathbf{D}_{\theta_0} l_2(X) = O(C^{-2}).$$

Таким образом, доказана плотность последовательности коэффициентов при $n^{-m/2}$ в $\int L_n(h)\exp(\psi(h))dh$. Для оставшихся членов (1.7) вида $\int (L_n(h))^m \exp(\psi(h))dh$ применим лемму 1.2, что позволит нам утверждать, что все коэффициенты при $n^{-m/2}$ кроме остатка останутся также плотными как линейная комбинация элементов конечного числа плотных последовательностей.

Докажем теперь плотность интеграла от остатка разложения (1.7). Заметим сначала, что согласно предыдущим рассуждениям $L_n(h)$ есть произведение $n^{-1/2}$ на элемент плотной последовательности случайных величин. Поэтому согласно лемме 1.2 величина $(L_n(h))^{M+1} = n^{-(M+1)/2} O_{\mathbf{P}_{\theta_0}}(1)$. Имея ввиду, что $0 \le |\widehat{L}_n(h)| \le |L_n(h)|$, $hn^{-1/2} < b$ и то, что $h^{-2}L_n(h)$ зависит от h только через $hn^{-1/2}$, согласно неравенству Маркова мы имеем

$$\mathbf{P}_{\theta_0}\left(h^{-2}\left|\widehat{L}_n(h)\right| \ge C\right) \le \mathbf{P}_{\theta_0}\left(\left|\frac{L_n(bn^\alpha)}{b^2 n^{2\alpha}}\right| \ge C\right) = O\left(n^{-1/2}\right).$$

Таким образом,

$$\mathbf{P}_{\theta_0}\left(\int\limits_{-bn^\alpha}^{bn^\alpha} \exp\left\{\psi(h) + \frac{h^2}{h^2}\widehat{L}_n(h)\right\}(L_n(h))^{M+1}\,dh > Cn^{-(M+1)/2}\right) \le$$
$$\le \mathbf{P}_{\theta_0}\left(\left\{\int\limits_{-bn^\alpha}^{bn^\alpha} \exp\left\{\psi(h) + \frac{h^2}{h^2}\widehat{L}_n(h)\right\} h^m dh > C\right\} \cap \left\{\left|\frac{\widehat{L}_n(h)}{h^2}\right| < \frac{I}{4}\right\}\right) +$$
$$+\mathbf{P}_{\theta_0}\left(\left|\frac{\widehat{L}_n(h)}{h^2}\right| \ge \frac{I}{4}\right) \le \mathbf{P}_{\theta_0}\left(\int\limits_{-bn^\alpha}^{bn^\alpha} \exp\left\{h\frac{\Delta_1}{\sqrt{n}} - h^2\frac{I}{4}\right\} h^m dh > C\right) + O\left(\frac{1}{\sqrt{n}}\right).$$

Интегральное выражение, стоящее здесь под знаком вероятности, как было установлено выше, равномерно по n конечно для каждого m. Таким образом, первое слагаемое есть $O(C^{-1})$, так что плотность последовательностей коэффициентов разложения (1.7) доказана.

С целью получения удобного вида разложения числителя (1.5), домножим $\exp(\psi(h))$ на соответствующий коэффициент, чтобы прийти в результате к плотности нормального распределения:

$$\exp\left(\frac{h\Delta_1}{\sqrt{n}} - \frac{h^2 I}{2}\right) = \varphi(h|\mu_n, \sigma_n^2)\exp\left(\frac{\Delta_1^2}{2In}\right)\sqrt{\frac{2\pi}{I}}.$$

В силу условия 4a $I > 0$, и поэтому такое представление допустимо. Множитель при $\varphi(h|\mu_n, \sigma_n^2)$ сократится с подобным в знаменателе (1.5). Из правила

Лопиталя следует, что

$$F_m(z) - \int\limits_{-bn^\alpha}^{z} h^m \varphi(h|\mu_n, \sigma_n^2)dh = o\left(n^{m\alpha} z^m e^{-n^{2\alpha} z^2}\right), \tag{1.9}$$

поэтому мы можем заменить область интегрирования $[-bn^\alpha, bn^\alpha]$ на $(-\infty, \infty)$. После интегрирования и приведения подобных членов числитель примет вид

$$\Phi\left(z|\mu_n, \sigma_n^2\right) + \sum_{m=1}^{M} n^{-m/2} f_m(z) + n^{-(M+1)/2} R_1(z),$$

где $\sup_z R_1(z) = O_{\mathbf{P}_{\theta_0}}(1)$ согласно установленному выше.

Знаменатель (1.5) примет тот же вид лишь с заменой z на ∞ (заметим, что главный член здесь равен единице).Воспользуемся известным представлением дроби вида $(1+a)^{-1}$ в области $|a| < 1$:

$$\frac{1}{1+a} = \sum_{m=0}^{\infty} (-a)^m.$$

Так как $n^{-m/2} f_m(\infty) \xrightarrow{\mathbf{P}_{\theta_0}} 0$, можно подставить вместо a разложение знаменателя $\sum_{m=1}^{M} n^{-m/2} f_m(\infty) + R_1(\infty)$. Домножая на полученное ранее разложение числителя, приходим к утверждению теоремы. Плотность коэффициентов при $n^{-m/2}$ и остатка по $\mathbf{P}_{\theta_0}$ получим так же с помощью леммы 1.2. ■

1.1.3. Разложение апостериорных средних

Практически не меняя доказательства предыдущей теоремы, можно получить разложение средних значений относительно апостериорного распределения от функций w от параметра вероятностной модели. Пусть $w : \mathbb{R} \mapsto \mathbb{R}$ — измеримая функция. Рассмотрим условное математическое ожидание

$$\mathbf{E}\left\{w\left(\sqrt{n}(\vartheta - \theta_0)\right) \middle| \mathbf{X}\right\} = \frac{\int_\Theta w\left(\sqrt{n}(\theta - \theta_0)\right) p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}{\int_\Theta p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}. \tag{1.10}$$

Обозначим

$$F_{m,w} = \int\limits_{-\infty}^{\infty} h^m w(h) \varphi(h|\mu_n, \sigma_n^2) dh,$$

$$H_{0,w} = f_{0,w} = F_{0,w},$$
$$H_{m,w} = f_{m,w} + \sum_{j=1}^{m} V_j f_{m-j,w}, \quad m = \overline{1, M},$$
$$f_{m,w} = \sum_{j=1}^{m} \frac{1}{m!} \sum_{I_1(j,m)} \prod_{l=1}^{j} q_{i_l,w}, \quad m = \overline{1, M},$$
$$q_{1,w} = F_{1,w}\pi_1 + \sqrt{n} F_{2,w} \frac{n^{-1}\Delta_2 + I}{2} + F_{3,w} \frac{\Delta_3}{6n},$$
$$q_{m,w} = F_{m+2,w} \frac{\Delta_{m+2}}{n(m+2)!} + F_{m,w} \frac{\pi_m}{m!}, \quad m = \overline{2, M}.$$

Коэффициенты V_j приведены на странице 25 перед теоремой 1.1).

Говорят, что функция $w : \mathbb{R} \mapsto \mathbb{R}$ имеет полиномиальную мажоранту, если существует такой полином $P(x)$, что $|w(x)| \leq P(x), \forall x \in \mathbb{R}$.

Теорема 1.2. *Пусть семейство распределений* $\{\mathbf{P}_\theta, \theta \in \Theta\}$ *удовлетворяет условиям* $\mathbb{D}_{\theta_0}(M)$. *Тогда для любой функции* w, *имеющей полиномиальную мажоранту, найдется* n_0, *что для всех* $n > n_0$ $\mathbf{P}_{\theta_0}$*-почти наверное выполняется* $\mathbf{E}\left\{ w\left(\sqrt{n}(\vartheta - \theta_0)\right) \middle| \mathbf{X} \right\} < \infty$, *и*

$$\mathbf{E}\left\{ w\left(\sqrt{n}(\vartheta - \theta_0)\right) \middle| \mathbf{X} \right\} = \sum_{m=0}^{M} n^{-m/2} H_{m,w} + n^{-(M+1)/2}\omega_n, \tag{1.11}$$

причем

$$H_{m,w} = O_{\mathbf{P}_{\theta_0}}(1), m = 0, \ldots, M,$$
$$\omega_n = O_{\mathbf{P}_{\theta_0}}(1).$$

Доказательство. Достаточно показать, что теорема справедлива для случая $w(h) = h^r, r \geq 0$.

Получим результат теоремы в части существования апостериорного момента:

$$\mathbf{P}_{\theta_0}\left(\mathbf{E}\left\{ \left(\sqrt{n}(\vartheta - \theta_0)\right)^r \middle| \mathbf{X} \right\} < \infty \right) =$$
$$= 1 - \lim_{M \to \infty} \mathbf{P}_{\theta_0}\left(\mathbf{E}\left\{ \left(\sqrt{n}(\vartheta - \theta_0)\right)^r \middle| \mathbf{X} \right\} \geq M \right).$$

После замены $h = \theta_0 + hn^{-1/2}$ в интегралах (1.10) получим

$$\mathbf{P}_{\theta_0}\left(\mathbf{E}\left\{\left(\sqrt{n}(\vartheta-\theta_0)\right)^r \middle| \mathbf{X}\right\} \geq M\right) \leq$$
$$\leq \mathbf{P}_{\theta_0}\left(\int\limits_{\mathbb{R}} \widetilde{Z}_n(\theta_0,h)dh < 1/\sqrt{M}\right) + \mathbf{P}_{\theta_0}\left(\int\limits_{\mathbb{R}} h^r \widetilde{Z}_n(\theta_0,h)dh \geq \sqrt{M}\right).$$

Согласно утверждению леммы 3.1 статьи [24] первое слагаемое в правой части последнего неравенства ограничено величиной c/M, где c — некоторая константа.

Для второго слагаемого справедливо

$$\mathbf{P}_{\theta_0}\left(\int\limits_{\mathbb{R}} h^r \widetilde{Z}_n(\theta_0,h)dh \geq \sqrt{M}\right) \leq$$
$$\leq \mathbf{P}_{\theta_0}\left(\int\limits_{-M}^{M} h^r \widetilde{Z}_n(\theta_0,h)dh \geq \frac{\sqrt{M}}{2}\right) + \mathbf{P}_{\theta_0}\left(\int\limits_{|h|>M} h^r \widetilde{Z}_n(\theta_0,h)dh \geq \frac{\sqrt{M}}{2}\right).$$

Второе слагаемое правой части последнего неравенства стремится к нулю при $M \to \infty$ по лемме 1.1. Чтобы показать, что первое слагаемое стремится к нулю, достаточно показать, что коэффициенты разложения (1.11) есть плотная последовательность относительно $\mathbf{P}_{\theta_0}$, в чем, собственно, и заключается оставшаяся часть утверждения данной теоремы.

Доказательство остается таким же, как и в теореме 1.1, если учесть, что после замены $\theta = \theta_0 + h/\sqrt{n}$ под знаком интеграла остается h^r, и кроме того, в оценивающих неравенствах h присутствует только в интегральных выражениях типа $\int_{\mathbb{R}} h^m e^{-h^2} dh$, которые конечны для любого m. Наконец, простейшими преобразованиями можно убедиться, что формулы, приведенные выше, справедливы. ■

Замечание после теоремы 1.1 справедливо и для разложений моментов. Кроме того, стоит отметить, что хоть разложение (1.11) существует для всех функций $w(h) = |h|^r, r \geq 0$, коэффициенты перед $n^{-m/2}$ могут существенно возрастать по абсолютной величине вместе с r, поскольку они зависят от моментов нормального закона порядка $m + r$, в то время как нормирующий множитель есть $(m!)^{-1}$.

1.2. Разложение апостериорного распределения параметра, центрированного $\sqrt{n}$-состоятельной оценкой

Рассматривается асимптотика апостериорного распределения случайной величины $\sqrt{n}(\vartheta - T_n)$, где T_n — $\sqrt{n}$-состоятельная оценка параметра θ. Такой способ центрирования, по сравнению с центрированием фиксированной точкой θ_0, может позволить добиться того, что (1.4) будет выполняться равномерно по всем θ, входящих в область, относительно которой эта оценка состоятельна.

Здесь будет доказано, что апостериорное распределение $\sqrt{n}(\vartheta - T_n)$ асимптотически нормально со средним $\mu_n = n^{-1/2}\Delta_1(T_n)I^{-1}(T_n)$ и дисперсией $\sigma_n^2 = (I(T_n))^{-1}$, причем вероятность отклонения апостериорного распределения от асимптотики стремится к нулю по совместному распределению $\mathbf{P}$ выборки $\mathbf{X}$ и параметра ϑ. Построение разложения апостериорного распределения, аналогичное теореме 1.1, имеет место как относительно меры $\mathbf{P}$, так и для $\mathbf{P}_{\theta_0}$. Доказывается также аналог теоремы 1.2 разложения апостериорных средних.

1.2.1. Условия регулярности и предварительные леммы

Введем понятие $\sqrt{n}$-состоятельных оценок.

Определение 1.1. *Пусть K — некоторый компакт из Θ. Будем говорить, что оценка T_n принадлежит классу оценок $\mathbb{C}(K)$ параметра $\theta \in \Theta$, если T_n измерима и*

$$\forall \varepsilon > 0 \ \exists b > 0 \ \sup_n \sup_{\theta \in K} \mathbf{P}_\theta \left(\sqrt{n}|T_n - \theta| > b \right) < \varepsilon.$$

Например, оценки максимального правдоподобия и байесовские оценки принадлежат классу $\mathbb{C}(K)$ при довольно широких условиях (см., например [26]).

В работе [27] показывается, что для $T_n \in \mathbb{C}(K)$ величина

$$\tau_n = T_n + \frac{\Delta_1(T_n)}{nI(T_n)}$$

является асимптотически достаточной статистикой, что означает возможность асимптотического приближения плотности выборки $p_n(x^{(n)}|\theta)$ для $\theta \in K$ плотностью, зависящей только от $\tau_n(x^{(n)})$. Мы покажем далее, что добавок $\mu_n = \Delta_1(T_n)(\sqrt{n}I(T_n))^{-1}$ и будет апостериорным средним величины $\sqrt{n}(\vartheta - T_n)$.

Перечислим условия, при которых здесь будет осуществляться доказательство заявленной асимптотической нормальности.

Пусть K — некоторый компакт из Θ. Будем называть следующие условия условиями $\mathbb{D}(K)$. Несмотря на то, что часть условий не зависит от K, мы их объединили для удобства ссылок.

1. Параметрическое пространство Θ есть интервал (ограниченный или неограниченный) на $\mathbb{R}$.

2. Для $\theta \neq \theta'$
$$\int\limits_{\mathfrak{X}} |p(x|\theta) - p(x|\theta')| \, \nu(dx) > 0.$$

3. Найдется $\delta_1 > 0$, такое что
$$\sup_{\theta,\theta' \in \Theta} |\theta - \theta'|^{\delta_1} \int\limits_{\mathfrak{X}} \sqrt{p(x|\theta)p(x|\theta')}\nu(dx) < \infty.$$

4. Носитель распределения $\mathbf{P}_\theta$ не зависит от θ.

5. Для любого $x \in \mathfrak{X}$ функция $p(x|\theta)$ имеет две непрерывных частных производных по θ из внутренности Θ для любого x.

6. Пусть $I(\theta) = \mathbf{E}_\theta \left(l_1(X|\theta)\right)^2$.

   а. $0 < I(\theta) < \infty, \ \theta \in \Theta$.

   б. Найдется такое $\delta_2 > 0$, что
$$\mathbf{E}_\theta \left|l_1(X|\theta)\right|^{2+\delta_2} < \infty, \ \mathbf{E}_\theta \left|l_2(X|\theta)\right|^{1+\delta_2} < \infty, \theta \in K.$$

в. Пусть $U_\beta(\theta)$ — β-окрестность точки θ. Существуют такие $\beta_0, \beta_1 > 0$, что

$$|l_2(x|\theta') - l_2(x|\theta'')| \le |\theta' - \theta''| D(x|\theta),\ \forall x,\ \theta', \theta'' \in U_{\beta_0}(\theta) \cap \Theta,\ \theta \in K,$$

$$\mathbf{E}_\theta \sup_{t \in U_{\beta_1} \cap \Theta} D(X|t) < \infty,\ \theta \in K.$$

г. Для некоторого $\delta_3 \ge 0$

$$\sup_{\theta \in \Theta} (1 + |\theta|^{\delta_3})^{-1} I(\theta) < \infty.$$

7. Априорная плотность g непрерывна на внутренности Θ.

Сделаем несколько замечаний. Из представленных условий следует, что информация по Фишеру может быть вычислена как $I(\theta) = -\mathbf{E}_\theta l_2(X|\theta)$. В дальнейшем компакт K будет совпадать с компактом условия на класс $\sqrt{n}$-состоятельных оценок $\mathbb{C}(K)$, но, формально говоря, условия $\mathbb{D}(K)$ не зависят от оценки $T_n \in \mathbb{C}(K)$. Условие 1 введено для удобства выкладок и может быть опущено, учитывая необходимую плотность параметрического пространства Θ по мере Лебега. Это обстоятельство еще будет обсуждено позже. Условие 4 означает, что множество $\{x : p(x|\theta) > 0\}$ не зависит от θ, поэтому можно считать, что $\mathfrak{X} = \{x : p(x|\theta) > 0\}$. Условия 2, 3, 6г не будут использованы напрямую, а только лишь являются условиями используемых здесь утверждений из статей [23], [24]. Условия 6б и особенно 6в кажутся довольно сложными, однако они аналогичны условиям на асимптотическую нормальность оценки максимального правдоподобия (см. [26], теорема 8.1, с. 118).

Лемма 1.3. *Пусть выполнены условия $\mathbb{D}(K)$. Тогда для любых $r > 0$ и $m > 0$ найдутся натуральное число n_0 и константа $C_{m,r,K}$, зависящая только от m, r и компакта K такие, что для всех $n > n_0$ и всех $A > 0$*

$$\sup_{\theta \in K} \mathbf{P}_\theta \left(\sup_{|h| > A} |h|^r Z_n(\theta, h) > A^{-m} \right) \le C_{m,r,K} A^{-m-r}. \tag{1.12}$$

Доказательство. Для доказательства этого утверждения необходимо взять супремум обоих частей неравенства (1.2) из условия леммы 1.1. Условия леммы 1.1 выполняются в каждой точке θ за счет условий 6, и поэтому на компакте неравенство выполняется и для супремума.

См. также лемму 1 статьи [13], в которой приводится аналогичное утверждение. ∎

Лемма 1.4. *В условиях предыдущей леммы*

$$\sup_{\theta\in K} \mathbf{E}_\theta \left(\int\limits_{\sqrt{n}(\Theta-\theta)} Z_n(\theta,h) g\left(\theta + \frac{h}{\sqrt{n}}\right) dh \right)^{-1} < \infty.$$

Доказательство. Это утверждение является копией леммы 3.2 статьи [24] за исключением, как и в предыдущей лемме, утверждения о равномерности на компакте K. Равномерность на компакте K достигается за счет условий 6, так же, как и в лемме 1.3. ∎

Лемма 1.5. *Пусть выполнены условия $\mathbb{D}(K)$ и $T_n \in \mathbb{C}(K)$. Тогда существуют такие $A > 0, 0 < \alpha < 0.5, 0 < \delta < 1$, что для любого $\varepsilon > 0$*

$$\sup_{\theta\in K} \mathbf{P}_\theta \left(\sup_{t\in U_{A_n}(T_n)} n^{-\frac{1}{1+\delta}} \left| \sum_{i=1}^n (l_2(X_i|t) + I(T_n)) \right| > \varepsilon \right) \to 0,$$

где $A_n = An^{\alpha-0.5}$, $U_\delta(\theta)$ — δ-окрестность точки θ.

Доказательство. Так как $\mathbf{E}_\theta l_2(X|\theta) = -I(\theta)$, то

$$\sup_{\theta\in K} \mathbf{P}_\theta \left(\sup_{t\in U_{A_n}(T_n)} n^{-\frac{1}{1+\delta_2}} \left| \sum_{i=1}^n (l_2(X_i|t) + I(T_n)) \right| > \varepsilon \right) \le$$
$$\le \sup_{\theta\in K} \mathbf{P}_\theta \left(\sup_{t\in U_{A_n}(T_n)} n^{-\frac{1}{1+\delta}} \left| \sum_{i=1}^n (l_2(X_i|t) - l_2(X_i|T_n)) \right| > \frac{\varepsilon}{2} \right) +$$
$$+ \sup_{\theta\in K} \mathbf{P}_\theta \left(n^{-\frac{1}{1+\delta}} \left| \sum_{i=1}^n (l_2(X_i|T_n) + I(X_i|T_n)) \right| > \frac{\varepsilon}{2} \right).$$

Если положить A равным β_0 из условия 6в, то

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\sup_{t\in U_{A_n}(T_n)} n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (l_2(X_i|t)-l_2(X_i|T_n))\right|>\frac{\varepsilon}{2}\right)\le$$
$$\le\sup_{\theta\in K}\mathbf{P}_\theta\left(\beta_0 n^{\alpha-0.5}n^{-\frac{1}{1+\delta}}\sum_{i=1}^n D(X_i|T_n)>\frac{\varepsilon}{2}\right)\le$$
$$\le\sup_{\theta\in K}\mathbf{P}_\theta\left(\beta_0 n^{\alpha-0.5}n^{-\frac{1}{1+\delta}}\sum_{i=1}^n \sup_{t\in U_{\beta_1}(\theta)} D(X_i|t)>\frac{\varepsilon}{2}\right)+\sup_{\theta\in K}\mathbf{P}_\theta\left(|T_n-\theta|>\beta_1\right)$$

Второе слагаемое в правой части последнего выражения стремится к нулю вследствие $T_n\in\mathbb{C}(K)$. К первому слагаемому в правой части последнего выражения можно применить закон больших чисел вследствие условия 6в. Таким образом, если α и δ таковы, что $\alpha-0,5-(1+\delta)^{-1}<-1$, то первое слагаемое стремится к нулю. Ниже будет показано, что можно выбрать любое $0<\delta<\min(\delta_2,1)$, с константой δ_2 из условия 6б. Таким образом, предложенный выбор α возможен, так что

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (l_2(X_i|T_n)+I(X_i|T_n))\right|>\frac{\varepsilon}{2}\right)\le$$
$$\le\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (l_2(X_i|T_n)-l_2(X_i|\theta))\right|>\frac{\varepsilon}{6}\right)+$$
$$+\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (l_2(X_i|\theta)+I(\theta))\right|>\frac{\varepsilon}{6}\right)+$$
$$+\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (I(T_n)-I(\theta))\right|>\frac{\varepsilon}{6}\right). \quad (1.13)$$

Первое слагаемое в правой части (1.13) оценивается, как и раньше, с помощью условия 6в. Так как для любой последовательности $c_n\to\infty$ вероятность $\sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}|T_n-\theta|>c_n\right)\to 0$, то, выбирая $c_n=n^\beta$, получаем

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^n (l_2(X_i|T_n)-l_2(X_i|\theta))\right|>\frac{\varepsilon}{6}\right)\le$$
$$\le\sup_{\theta\in K}\mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}c_n n^{-1/2}\sum_{i=1}^n D(X_i|\theta)>\frac{\varepsilon}{6}\right)+\sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}|T_n-\theta|>c_n\right)\to 0$$

где бета таково, что $\beta - (1+\delta)^{-1} - 0,5 < -1$. Такое β, очевидно, существует.

Второе слагаемое в (1.13) стремится к нулю по усиленной версии закона больших чисел (см., например, [28], с. 257) в силу условия 6б (нужно лишь взять $\delta < \min(\delta_2, 1)$).

В силу условий $\mathbb{D}(K)$ выполняется следующее неравенство (см. [26], с.94): $|I(\theta+h) - I(\theta)| \le 4\left(I^{1/2}(\theta+h) + I^{1/2}(\theta)\right) Lh$ для некоторого $L > 0$. Таким образом, на компакте K функция $I(\theta)$ удовлетворяет условию Липшица, и поэтому существует такое $\gamma > 0$, что

$$\sup_{\theta\in K} \mathbf{P}_\theta\left(n^{-\frac{1}{1+\delta}}\left|\sum_{i=1}^{n}(I(T_n) - I(\theta))\right| > \frac{\varepsilon}{6}\right) \le \sup_{\theta\in K} \mathbf{P}_\theta\left(n^{1-\frac{1}{1+\delta}}|T_n - \theta| > \gamma\right) \to 0,$$

ибо $T_n \in \mathbb{C}(K)$ и $0 < 1 - (1+\delta)^{-1} < 1/2$ из выбора δ. ∎

Лемма 1.6. *При выполнении условий $\mathbb{D}(K)$ и при $T_n \in \mathbb{C}(K)$ последовательность величин $\mu_n = \Delta_1(T_n)(\sqrt{n}I(T_n))^{-1}$ является равномерно плотной по $\theta \in K$:*

$$\forall\varepsilon > 0 \;\; \exists M : \sup_n \sup_{\theta\in K} \mathbf{P}_\theta(|\mu_n| > M) < \varepsilon.$$

Доказательство. Имеем:

$$\frac{\Delta_1(T_n)}{\sqrt{n}} = \frac{\Delta_1(\theta)}{\sqrt{n}} + \sqrt{n}(T_n - \theta)\frac{\Delta_2(\tau_n)}{n},$$

где τ_n — точка между T_n и θ. Последовательность $n^{-1/2}\Delta_1(\theta)$ плотна по неравенству Чебышева:

$$\sup_{\theta\in K} \mathbf{P}_\theta\left(\left|n^{-1/2}\Delta_1(\theta)\right| > M\right) \le \frac{\sup_{\theta\in K} I(\theta)}{M^2} \to 0, \;\; M \to \infty,$$

так как информация $I(\theta)$ непрерывна, а значит, и ограничена на компакте K. Последовательность $\sqrt{n}(T_n - \theta)$ равномерно плотна на K, так как $T_n \in \mathbb{C}(K)$. Последовательность величин $n^{-1}\Delta_2(\tau_n)$ плотна по лемме 1.5. Отсюда в силу леммы 1.2 следует плотность $\Delta_1(T_n)n^{-1/2}$. Так как $I(\theta)$ непрерывна, а T_n состоятельна, то можно найти такое $\sigma > 0$, что

$$\sup_{\theta\in K} \mathbf{P}_\theta(I(T_n) > \sigma) \to 1.$$

Следовательно, последовательность μ_n является равномерно плотной по $\theta \in K$.

■

Обозначим $\mathcal{N}(A|m, s^2)$ нормальное распределение со средним m и дисперсией s^2, вычисленное на борелевском множестве $A \subset \mathbb{R}$, а $\Phi(x|m, s^2)$ и $\varphi(x|m, s^2)$ — соответствующие функцию распределения и плотность, вычисленные в точке x. Символами $\Phi(x)$ и $\varphi(x)$ без параметров будем обозначать функции распределения и плотности стандартного нормального закона. Положим $E_m(\mu, \sigma) = \int_{\mathbb{R}} |h|^m \varphi(h|\mu, \sigma^2) dh$.

Обозначим $\sigma_n^2 = (I(T_n))^{-1}$. Эта величина является асимптотической дисперсией апостериорного распределения $\sqrt{n}(\vartheta - T_n)$.

Лемма 1.7. *При выполнении условий* $\mathbb{D}(K)$ *и при* $T_n \in \mathbb{C}(K)$, *для любых* $m > 0, A_n = An^\alpha, M > 0, \alpha > 0$ *найдутся такие числа* $\gamma_1 > 0$ *и* $\gamma_2 > 0$, *что*

$$\sup_{\theta_0 \in K} \mathbf{P}_{\theta_0} \left\{ \int_{-A_n}^{A_n} |h|^m \varphi\left(h|\mu_n, \sigma_n^2\right) dh \geq E_m\left(\mu_n, \sigma_n^2\right) - \gamma_1 \exp(-n^{\gamma_2}) \right\} \to 1.$$

Доказательство. Применяя правило Лопиталя, легко устанавливается справедливость следующего асимптотического тождества ($x \to \infty$), справедливого при любых μ и $\sigma > 0$:

$$\int_x^\infty |t|^m \varphi\left(\frac{t-\mu}{\sigma}\right) dt = x^{m-1} \varphi\left(\frac{x-\mu}{\sigma}\right) (\sigma^2 + o(1)).$$

Следовательно,

$$\int_{-A_n}^{A_n} |h|^m \varphi\left(h|\mu_n, \sigma_n^2\right) dh \geq E_m\left(\mu_n, \sigma_n\right) - \frac{2A_n^{m-1}}{I(T_n)} \varphi\left(A_n| \left|\mu_n\right|, \sigma_n^2\right) (1 + \omega_n), \quad (1.14)$$

где $\omega_n \to 0$ при $n \to \infty$ по вероятности $\mathbf{P}_\theta, \theta \in K$. Если $\mathbf{X}$ принадлежит множеству $U_n = \{|\mu_n| \leq A_n/2\} \cap \{I(T_n) \geq \sigma\}$, то правая часть в (1.14) не меньше, чем

$$E_m\left(\mu_n, \sigma_n^2\right) - \frac{\sqrt{2} A_n^{m-1}}{\sqrt{\pi}\sigma} \exp\left(-\frac{\sigma^2 A_n^2}{8}\right) (1 + \omega_n).$$

Вследствие непрерывности $I(\theta)$ найдется такое $\sigma > 0$, что $\sup_{\theta\in K}\mathbf{P}_\theta(I(T_n) \geq \sigma) \to 1$. Вследствие леммы 1.6 $\sup_{\theta\in K}\mathbf{P}_\theta(|\mu_n| \leq A_n/2) \to 1$. Таким образом, вероятность события U_n стремится к единице.

При $\mathbf{X} \in U_n$ вследствие того, что $A_n = An^\alpha$, найдутся такие числа γ_1 и γ_2, что

$$\frac{\sqrt{2}(An^\alpha)^{m-1}}{\sqrt{\pi I(T_n)}}\exp\left(-\frac{I(T_n)A^2n^{2\alpha}}{8}\right)(1+\omega_n) \leq \gamma_1\exp(-n^{\gamma_2})(1+\omega_n).$$

Отсюда и $\sup_{\theta\in K}\mathbf{P}_\theta(U_n) \to 1$ следует результат леммы. ∎

Определим $B_n = [T_n - A_n/\sqrt{n}; T_n + A_n/\sqrt{n}]$, где $A_n = An^\alpha$ с константами $A > 0$ и $\alpha > 0$. Пусть q_n — апостериорная плотность, и $\widetilde{q}_n$ — апостериорная плотность, суженная на множество B_n:

$$q_n(\theta|\mathbf{X}) = \frac{p_n(\mathbf{X}|\theta)g(\theta)}{\int_\Theta p_n(\mathbf{X}|\theta)g(\theta)d\theta}, \quad \widetilde{q}_n(\theta|\mathbf{X}) = \mathbb{1}_{B_n}(\theta)\frac{p_n(\mathbf{X}|\theta)g(\theta)}{\int_{B_n} p_n(\mathbf{X}|\theta)g(\theta)d\theta},$$

где $\mathbb{1}_A$ — индикатор множества A.

Лемма 1.8. *При выполнении условий $\mathbb{D}(K)$ и при $T_n \in \mathbb{C}(K)$. Тогда для любого $m > 0$ вероятность*

$$\sup_{\theta_0\in K}\mathbf{P}_{\theta_0}\left(\int\limits_\Theta |q_n(\theta|\mathbf{X}) - \widetilde{q}_n(\theta|\mathbf{X})|\,d\theta > n^{-m}\right) \to 0. \tag{1.15}$$

Доказательство. Выпишем простое неравенство, которое нам пригодится несколько раз:

$$\left|\frac{S}{B} - \frac{s}{b}\right| = \left|\frac{S}{B} - \frac{S}{b} + \frac{S}{b} - \frac{s}{b}\right| \leq |b|^{-1}\left(\left|\frac{S}{B}\right||B-b| + |S-s|\right). \tag{1.16}$$

Пронормируем числители и знаменатели выражений q_n и $\widetilde{q}_n$ на $p(\mathbf{X}|\theta_0)$. Так как вероятность $\mathbf{P}_{\theta_0}(p(\mathbf{X}|\theta_0) = 0) = 0$ эта операция оправдана. Сделаем замену $\theta = \theta_0 + h/\sqrt{n}$ в числителе и знаменателе обеих плотностей в интеграле (1.15). Обозначим $\Theta_n = \sqrt{n}(\Theta - \theta_0), L_n = [-A_n + \sqrt{n}(T_n - \theta_0); A_n + \sqrt{n}(T_n -$

$\theta_0)], g_n(h) = g(\theta_0 + h/\sqrt{n})$. Из неравенства (1.16), (в качестве S и s берутся числители q_n и $\widetilde{q}_n$, а в качестве B и b — знаменатели q_n и $\widetilde{q}_n$) вытекает, что

$$\int_{\Theta} |\widetilde{q}_n(\theta|\mathbf{X}) - q_n(\theta|\mathbf{X})| d\theta \le \left(\int_{\Theta_n} Z_n(\theta_0, h) g_n(h) dh \right)^{-1} \times$$
$$\times \left(\int_{L_n^c \cap \Theta_n} Z_n(\theta_0, h) g_n(h) dh + \int_{\Theta} \widetilde{q}_n(\theta|\mathbf{X}) d\theta \int_{L_n^c \cap \Theta_n} Z_n(\theta_0, h) g_n(h) dh \right) =$$
$$= \left(\int_{\Theta_n} Z_n(\theta_0, h) g_n(h) dh \right)^{-1} 2 \int_{L_n^c \cap \Theta_n} Z_n(\theta_0, h) g_n(h) dh.$$

Обозначим правую часть последнего выражения $W(Z_n)$. Для случайной величины $W(Z_n)$ имеем оценку вероятности больших уклонений:

$$\sup_{\theta_0 \in K} \mathbf{P}_{\theta_0} \left(W(Z_n) > n^{-m} \right) \le$$
$$\le \sup_{\theta_0 \in K} \mathbf{P}_{\theta_0} \left(n^{-2m} \left(\int_{\Theta_n} Z_n(\theta_0, h) g_n(h) dh \right)^{-1} > n^{-3m/2} \right) +$$
$$+ \sup_{\theta_0 \in K} \mathbf{P}_{\theta_0} \left(\left\{ \int_{L_n^c \cap \Theta_n} Z_n(\theta_0, h) g_n(h) dh > n^{-3m/2} \right\} \cap \left\{ \sqrt{n} |T_n - \theta_0| < \frac{A_n}{2} \right\} \right) +$$
$$+ \sup_{\theta_0 \in K} \mathbf{P}_{\theta_0} \left(\sqrt{n} |T_n - \theta_0| > \frac{A_n}{2} \right).$$

Вследствие леммы 1.4 и неравенства Маркова первое слагаемое в правой части последнего неравенства стремится к нулю. Так как $L_n^c \cap \Theta_n \subset L_n^c$, то можно расширить область итегрирования во втором слагаемом до L_n^c. Далее, вследствие $A_n = An^{\alpha}$, ко второму слагаемому можно применить лемму 1.3, и получить, что второе слагаемое также стремится к нулю. Третье слагаемое стремится к нулю, так как $T_n \in \mathbb{C}(K)$. ∎

В следующей лемме устанавливается почти очевидный факт, что $\sqrt{n}$-состоятельная оценка с вероятностью, стремящейся к единице, принадлежит $1/\sqrt{n}$-окрестности компакта, определяющего эту оценку.

Лемма 1.9. *Пусть K — компакт, лежащий в интервале $(\theta_1;\theta_2)$, и оценка $T_n \in \mathbb{C}(K)$. Тогда для любой последовательности $A_n \to \infty$ такой, что $A_n/\sqrt{n} \to 0$, вероятности*

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}(T_n-\theta_1)<A_n\right)\to 0,\quad \sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}(\theta_2-T_n)<A_n\right)\to 0.$$

Доказательство.

Имеем

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}(T_n-\theta_1)<A_n\right)=\sup_{\theta\in K}\mathbf{P}_\theta\left(\sqrt{n}(T_n-\theta)<\sqrt{n}(A_n/\sqrt{n}+\theta_1-\theta)\right).$$

Так как $\theta_1-\inf_{\theta\in K}\theta<0$, то величина $\sqrt{n}(A_n/\sqrt{n}+\theta_1-\theta)\to-\infty$ при $n\to\infty$, поэтому вследствие $T_n\in\mathbb{C}(K)$ получим первое утверждение теоремы.

Второе утверждение получается аналогично.

■

1.2.2. Асимптотическая нормальность апостериорного распределения

Как известно, расстояние по вариации между мерами P и Q, имеющие плотности p и q по мере μ, может быть представлено в виде

$$\|P-Q\|_{TV}=\sup_{B\in\mathfrak{B}}|P(B)-Q(B)|=\frac{1}{2}\int|p-q|d\mu. \tag{1.17}$$

Теорема 1.3. *Пусть K — компакт, лежащий в Θ вместе с некоторой своей окрестностью и оценка $T_n\in\mathbb{C}(K)$. Тогда при выполнении условий $\mathbb{D}(K)$ для любого $\varepsilon>0$*

$$\sup_{\theta\in K}\mathbf{P}_\theta\left\{\left\|\mathbf{P}\left(\sqrt{n}(\vartheta-T_n)\in\cdot\,|\mathbf{X}\right)-\mathcal{N}\left(\,\cdot\,|\mu_n,\sigma_n^2\right)\right\|_{TV}>\varepsilon\right\}\to 0.$$

Доказательство. Утверждение теоремы будет доказано, если показать (см. (1.17)), что $\forall\varepsilon>0$

$$\sup_{\theta_0\in K}\mathbf{P}_{\theta_0}\left(\int\limits_{\Theta}\left|q_n\left(\theta|\mathbf{X}\right)-\varphi\left(\theta\middle|T_n+\frac{\Delta_1(T_n)}{nI(T_n)},\frac{1}{nI(T_n)}\right)\right|d\theta>\varepsilon\right)\to 0.$$

Выберем числа $A > 0$ и $0 < \alpha < 1/2$ из утверждения леммы 1.5. Определим последовательность чисел A_n, последовательность множеств B_n, и плотность $\widetilde{q}_n$ таким же образом, как и в условии леммы 1.8. Легко видеть, что справедливо неравенство:

$$\int\limits_{\Theta} \left| q_n\left(\theta|\mathbf{X}\right) - \varphi\left(\theta\Big|T_n + \frac{\mu_n}{\sqrt{n}}, \frac{\sigma_n^2}{n}\right)\right| d\theta \leq \int\limits_{\Theta} \left| q_n\left(\theta|\mathbf{X}\right) - \widetilde{q}_n\left(\theta|\mathbf{X}\right)\right| d\theta + \\ + \int\limits_{\Theta} \left| \widetilde{q}_n\left(\theta|\mathbf{X}\right) - \varphi\left(\theta\Big|T_n + \frac{\mu_n}{\sqrt{n}}, \frac{\sigma_n^2}{n}\right)\right| d\theta. \quad (1.18)$$

Первое слагаемое стремится к нулю по вероятности $\mathbf{P}_\theta$ равномерно по $\theta \in K$ в силу леммы 1.8. Остается показать, что

$$\sup_{\theta_0 \in K} \mathbf{P}_{\theta_0}\left(\int\limits_{\Theta} \left| \widetilde{q}_n\left(\theta|\mathbf{X}\right) - \varphi\left(\theta\Big|T_n + \frac{\mu_n}{\sqrt{n}}, \frac{\sigma_n^2}{n}\right)\right| d\theta > \varepsilon\right) \to 0. \quad (1.19)$$

В интеграле под знаком вероятности и в интеграле, определяющем $\widetilde{q}_n$, сделаем замену $\theta = T_n + h/\sqrt{n}$, после чего разделим числитель и знаменатель в определении $\widetilde{q}_n$ на величину $p(\mathbf{X}|T_n)g(T_n)$, которая с вероятностью единица отлична от нуля по условию 4. Тогда интеграл в (1.19) принимает вид

$$\int\limits_{-A_n}^{A_n} \left| \frac{\widetilde{Z}_n(T_n, h)}{\int_{-A_n}^{A_n} \widetilde{Z}_n(T_n, h) dh} - \varphi\left(h|\mu_n, \sigma_n^2\right)\right| dh, \quad (1.20)$$

при условии, что $T_n + h/\sqrt{n}$ не выходит за границы параметрического пространства Θ когда $h \in [-A_n; A_n]$. Очевидно, в случае неограниченного Θ такой ситуации не возникнет. Если же Θ — ограниченное множество, то без ограничения общности положим $\Theta = (\theta_1; \theta_2)$. Выход за границы параметрического пространства означает $\sqrt{n}(T_n - \theta_1) < A_n$ или $\sqrt{n}(\theta_2 - T_n) < A_n$. В силу леммы 1.9 вероятность этих событий стремится к нулю. Поэтому все дальнейшие рассуждения можно вести при условии, что $\mathbf{X}$ принадлежит множеству $\{[-A_n; A_n] \subset \sqrt{n}(\Theta - T_n)\}$.

Итак, остается показать, что интеграл (1.20) стремится к нулю по вероятности $\mathbf{P}_\theta$ равномерно по $\theta \in K$. Умножим числитель и знаменатель дроби под интегралом в (1.20) на

$$V(T_n) = \exp\left(-\frac{\Delta_1^2(T_n)}{2nI(T_n)}\right)\sqrt{\frac{I(T_n)}{2\pi}}.$$

Этот множитель пригодится в дальнейшем, чтобы привести $\widetilde{Z}_n(T_n,h)$ к функции плотности нормального закона. Воспользуемся неравенством (1.16), в котором $S = V(T_n)\widetilde{Z}_n(T_n,h)$, $B = V(T_n)\int_{-A_n}^{A_n}\widetilde{Z}_n(T_n,h)dh$, $s = \varphi(h|\mu_n,\sigma_n^2)$ и $b = 1$. Остается показать, что

$$\int\limits_{-A_n}^{A_n}\left|V(T_n)\widetilde{Z}_n(T_n,h) - \varphi\left(h\Big|\frac{\Delta_1(T_n)}{\sqrt{n}I(T_n)}, \frac{1}{I(T_n)}\right)\right| dh$$

по вероятности стремится к нулю, так как стремление $\int_{-A_n}^{A_n}\widetilde{Z}_n(T_n,h)dh$ к единице очевидным образом следует из этого факта и леммы 1.7, согласно которой

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\left|\int\limits_{-A_n}^{A_n}\varphi\left(h|\mu_n,\sigma_n^2\right)dh - 1\right| > \varepsilon\right) \to 1.$$

Разложим $\ln Z_n(T_n,h)$ по формуле Тейлора по степеням h до второго порядка относительно $h=0$:

$$\ln\widetilde{Z}_n(T_n,h) = \Delta_1(T_n)\frac{h}{\sqrt{n}} + \Delta_2\left(\widehat{T}_n\right)\frac{h^2}{2n} + \ln g\left(T_n + \frac{h}{\sqrt{n}}\right) - \ln g(T_n),$$

где $\widehat{T}_n$ — промежуточная точка интервала $[T_n; T_n + h/\sqrt{n}]$. Так как $A_n/\sqrt{n} \to 0$, то $h/\sqrt{n} < 1$ при достаточно большом n. Теперь разложим экспоненту от $\ln\widetilde{Z}_n(T_n,h)$ по формуле Тейлора относительно точки $\psi(T_n,h) = \Delta_1(T_n)hn^{-1/2} - I(T_n)h^2/2$:

$$\begin{aligned}\exp\ln\widetilde{Z}_n(T_n,h) = \exp\{\psi(T_n,h)\} + &\\ + \exp\{\psi(T_n,h)+\Lambda_n(h)\}\Bigg[&\left(I(T_n) + \frac{\Delta_2\left(\widehat{T}_n\right)}{n}\right)\frac{h^2}{2} + \\ &+ \ln g\left(T_n + \frac{h}{\sqrt{n}}\right) - \ln g(T_n)\Bigg], \quad (1.21)\end{aligned}$$

где $\Lambda_n(h) = \lambda\left(I(T_n) + n^{-1}\Delta_2\left(\widehat{T}_n\right)\right)h^2/2$, $0 < \lambda < 1$. Заметим, что

$$V(T_n)\exp(\psi(T_n, h)) = \varphi\left(h|\mu_n, \sigma_n^2\right).$$

Таким образом, нам нужно показать, что интеграл по h от модуля остатка в (1.21) (второе слагаемое), помноженного на V_n стремится к нулю по вероятности $\mathbf{P}_\theta$ равномерно по $\theta \in K$.

В силу условия 7 функция g непрерывна всюду, и для любых $\delta > 0$, $\gamma > 0$ вероятность

$$\sup_{\theta_0\in K}\mathbf{P}_{\theta_0}\left(\left\{\sup_{h\in[-A_n;A_n]}\left|g\left(T_n + \frac{h}{\sqrt{n}}\right) - \ln g(T_n)\right| > \delta\right\}\cap\{|T_n - \theta_0| < \gamma\}\right) \to 0. \tag{1.22}$$

В силу $T_n \in \mathbb{C}(K)$ вероятность $\sup_{\theta_0\in K}\mathbf{P}_{\theta_0}(|T_n - \theta_0| \geq \gamma) \to 0$.

Выберем числа A и α, удовлетворяющие условию леммы 1.5. Тогда последовательность $\left|I(T_n) + n^{-1}\Delta_2\left(\widehat{T}_n\right)\right| \to 0$ по вероятности $\mathbf{P}_\theta$ равномерно по $\theta \in K$ при любом $h \in [-A_n; A_n]$, поэтому и супремум по h будет стремиться к нулю. Вместе с результатом (1.22) остается показать, что для некоторого $c > 0$

$$\sup_{\theta_0\in K}\mathbf{P}_{\theta_0}\left\{\int\limits_{-A_n}^{A_n} h^2 V(T_n)\exp(\psi(T_n, h) + \Lambda_n(h))dh < c\right\} \to 1. \tag{1.23}$$

Определим случайную величину, связанную с подынтегральным выражением в (1.23):

$$\xi(T_n, h) = -\frac{\Delta_1^2(T_n)}{2nI(T_n)} + \Delta_1(T_n)\frac{h}{\sqrt{n}} - I(T_n)\frac{h^2}{2} + \Lambda_n(h).$$

Из определения Λ_n следует, что $2|\Lambda_n(h)| < \left|I(T_n) + n^{-1}\Delta_2\left(\widehat{T}_n\right)\right| h^2$.

Теперь введем события

$$U_n = \left\{n^{1-\zeta}\left|\sum_{i=1}^n l_2\left(X_i|\widehat{T}_n\right)\right| > 1\right\}, Y_n = \left\{\left|\frac{\Delta_1(T_n)}{\sqrt{n}}\right| > n^{\zeta/4}\right\}, J_n = \{I(T_n) > \sigma\},$$

где $0 < \zeta < 1/2$, $\sigma > 0$. В силу леммы 1.5 $\sup_{\theta\in K}\mathbf{P}_\theta(U_n) \to 0$, если положить ζ равным величине δ, которая присутствует в утверждении леммы 1.5. Вероятность $\sup_{\theta\in K}\mathbf{P}_\theta(Y_n) \to 0$ по лемме 1.6, и в силу непрерывности $I(\theta)$ найдется такое σ, что $\sup_{\theta\in K}\mathbf{P}_\theta(J_n)\to 0$. Таким образом, при $\mathbf{X}\in U_n^c\cap Y_n^c\cap J_n^c$ случайная величина

$$\xi(T_n,h)\leq -\frac{1}{2}\left(\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-n^{-\zeta})}-h\right)^2 I(T_n)+ \\ +\frac{\Delta_1^2(T_n)}{n}\left(\frac{1}{I(T_n)-n^{-\zeta}}-\frac{1}{I(T_n)}\right)\leq \\ \leq -\frac{1}{2}\left(\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-n^{-\zeta})}-h\right)^2 I(T_n)+n^{\zeta/2}\frac{n^{-\zeta}}{\sigma(\sigma-n^{-\zeta})}.$$

Итак, для $\mathbf{X}\in U_n^c\cap Y_n^c\cap J_n^c$ искомый интеграл

$$\int\limits_{-A_n}^{A_n} V(T_n)\exp(\psi(T_n,h)+\Lambda_n(h))dh\leq\int\limits_{-\infty}^{\infty} V(T_n)\exp(\psi(T_n,h)+\Lambda_n(h))dh\leq \\ \leq\exp\left(\frac{n^{-\zeta/2}}{\sigma(\sigma-n^{-\zeta})}\right)\int\limits_{-\infty}^{\infty}\varphi\left(h\Big|\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-n^{-\zeta})},\frac{1}{I(T_n)}\right)dh= \\ =\exp\left(\frac{n^{-\zeta/2}}{\sigma(\sigma-n^{-\zeta})}\right).$$

Последняя величина, очевидно, ограничена для $n\geq n_0$ при некотором n_0. Как было отмечено ранее, вероятность события $U_n^c\cap Y_n^c\cap J_n^c$ стремится к единице. Таким образом, показано, что интеграл (1.23) с вероятностью, стремящейся к единице, ограничен, и это доказывает теорему. ■

Условие 1, а также условие теоремы о том, что компакт K входит в Θ вместе со своей окрестностью, используется лишь один раз в доказательстве теоремы 1.3, а именно при получении (1.20), где мы ссылались на лемму 1.9. Очевидно, Θ может быть не только интервалом, но и, скажем, конечным объединением интервалов, поскольку требуется лишь конечность интервалов. Если учитывать, что K — компакт, то вместе с требованием о лебеговской измеримости Θ, можно утверждать, что теорема 1.3 выполняется для любого измеримого множества Θ.

Однако существование у K окрестности в Θ, по всей видимости, необходимо для утверждения теоремы, потому что на границах параметрического множества может и не быть асимптотической нормальности (см. пример после следствия). С другой стороны, с байесовской точки зрения не столь интересно равномерное $\theta \in K$ стремление к нулю по $\mathbf{P}_\theta$, сколь стремление к нулю по совместному распределению $\mathbf{P}$. В связи с этим докажем следующее утверждение, в котором условие 1 используется по существу.

Следствие 1.1. *Пусть для любого компакта $K \in \Theta$, входящего в Θ с некоторой окрестностью, оценка $T_n \in \mathbb{C}(K)$ и выполнены условия $\mathbb{D}(K)$. Тогда*

$$\left\|\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) \in \cdot|\mathbf{X}\right) - \mathcal{N}\left(\cdot|\mu_n, \sigma_n^2\right)\right\|_{TV} \xrightarrow{\mathbf{P}} 0.$$

Доказательство. Для произвольных ε и $B \in \mathfrak{B}$ ($\mathfrak{B}$ — борелевская σ-алгебра на $\mathbb{R}$) определим множество

$$U_n(B, \varepsilon) = \left\{\mathbf{X} : \left\|\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) \in B|\mathbf{X}\right) - \Phi\left(B|\mu_n, \sigma_n^2\right)\right\|_{TV} > \varepsilon\right\}.$$

Пусть $\Theta = (\theta_1; \theta_2)$. Положим для произвольного $\gamma > 0$ $K_m = [\theta_1 + \gamma/\sqrt{m}; \theta_2 - \gamma/\sqrt{m}]$. Для каждого отрезка K_m выполняются условия $\mathbb{D}(K_m)$, и по теореме 1.3 $\sup_B \mathbf{P}_{\theta_0, K_m}(U_n(B, \varepsilon)) \to 0$. Поэтому найдется такая последовательность K_{m_n}, что $\sup_B \mathbf{P}_{\theta_0, K_{m_n}}(U_n(B, \varepsilon)) \to 0$. Очевидно, $K_{m_n} \uparrow \Theta$, и поэтому

$$\sup_B \mathbf{P}(U_n(B, \varepsilon)) = \int_\Theta \sup_B \mathbf{P}_\theta(U_n(B, \varepsilon))\mathbf{G}(d\theta) \le$$

$$\le \sup_B \mathbf{P}_{\theta, K_{m_n}}(U_n(B, \varepsilon))\mathbf{G}\left(K_{m_n}\right) + \mathbf{G}\left(K_{m_n}^c\right) \to 0.$$

Тот же самый аргумент можно применить и для $\Theta = \mathbb{R}$. В таком случае $K_m = [-\gamma\sqrt{m}, \gamma\sqrt{m}]$. Все остальные случаи являются комбинациями этих двух, если принять во внимание условие 1. ∎

Условие следствия 1.1 на каждый внутренний компакт очень полезно для моделей, у которых параметрическое пространство есть открытый интервал. Например, в случае, когда выборка берется из распределения Бернулли, $\Theta =$

$(0;1)$, и условия $\mathbb{D}(\Theta)$ не выполняются, условия следствия 1, тем не менее, выполнены.

В связи с теоремой 1.3 и следствием 1.1 рассмотрим следующий пример. Пусть наблюдается выборка из показательного распределения с плотностью $p(x|\theta) = \theta \exp(-\theta x)$. В качестве априорного распределения возьмем треугольное распределение на $\Theta = [1;2]$. Несмотря на то, что Θ— компакт, и выполнены условия $\mathbb{D}(\Theta)$ (в том числе и непрерывность априорной плотности), численное моделирование показывает, что нормальная аппроксимация теоремы 1.3 не применима в граничных точках $\theta_0 = 1$ или $\theta_0 = 2$. Тем не менее, асимптотическая нормальность апостериорного распределения выполняется, если использовать совместное распределение $\mathbf{P}$. Этот пример интересен тем, что для естественного параметрического пространства $\Theta = (0;\infty)$ при любом априорном распределении, таком что $g(\theta) > 0, \theta \in \Theta$ точки $\theta_0 = 1$ и $\theta_0 = 2$ перестают быть особенными.

Идея асимптотической нормальности апостериорного распределения, центрированного $\sqrt{n}$-состоятельной оценкой, не нова. Так, в [11] на с. 619 приводится соответствующая теорема, которая, однако, утверждает лишь сходимость по мере $\mathbf{P}_{\theta_0}$ с фиксированным значением θ_0. Таким образом, текущий пункт настоящей статьи можно рассматривать как доказательство равномерной сходимости к нормальному пределу.

1.2.3. Асимптотическое разложение апостериорного распределения

Метод доказательства теоремы 1.3 можно применить для получения асимптотического разложения апостериорного распределения. Нужно лишь взять у $\ln Z_n(T_n, h)$ не только первые два члена разложения, а столько, сколько потребуется для достижения заданной точности. Соответственно, следует потребовать существования этих производных и их моментов. Формально, для любого компакта K и любого целого числа $M \geq 0$ следует найти такие функции $H_m(z, \mathbf{X})$:

$\mathbb{R} \times \mathfrak{X}^n \mapsto \mathbb{R}$, что для любого $\varepsilon > 0$ найдется такое $C > 0$, что при всех $z \in \mathbb{R}$

$$\sup_n \sup_{\theta \in K} \mathbf{P}_\theta \left(\left| \mathbf{P}\left(\sqrt{n}(\vartheta - T_n) < z|\mathbf{X}\right) - \sum_{m=0}^{M} n^{-m/2} H_m\left(z, \mathbf{X}\right) \right| > Cn^{-(M+1)/2} \right) < \varepsilon,$$

$$\sup_n \sup_{\theta \in K} \mathbf{P}_\theta \left(|H_m\left(z, \mathbf{X}\right)| > C \right) < \varepsilon, \ 0 \le m \le M.$$

Последнее условие означает плотность коэффициентов разложения. Плотность была выбрана в качестве альтернативы ограниченности почти наверное, как это, например, было сделано в [12].

Задавшись целым числом $M \ge 0$ и компактом K приведем условия на асимптотическое разложение. Как было отмечено ранее, это по большому счету те же условия $\mathbb{D}(K)$, только расширенные требованиями на существование производных высшего порядка у функций $p(X|\theta)$ и $g(\theta)$. Назовем эти условия $\mathbb{D}(M, K)$. Все последующие ссылки на условия в этой главе будут идти на список $\mathbb{D}(M, K)$, если не указано иначе.

1. Θ есть интервал (ограниченный или неограниченный) на $\mathbb{R}$.

2. Для $\theta \neq \theta'$
$$\int_{\mathfrak{X}} |p(x|\theta) - p(x|\theta')| \, \nu(dx) > 0.$$

3. Найдется $\delta_1 > 0$, такое что
$$\sup_{\theta, \theta' \in \Theta} |\theta - \theta'|^{\delta_1} \int_{\mathfrak{X}} \sqrt{p(x|\theta)p(x|\theta')} \nu(dx) < \infty.$$

4. Носитель распределения $\mathbf{P}_\theta$ не зависит от θ.

5. Функция $p(x|\theta)$ определена для каждого $x \in \mathfrak{X}$ и имеет $M + 3$ непрерывных частных производных по θ для всех θ из внутренности Θ.

6. Информация по Фишеру $I(\theta)$ строго положительна при $\theta \in \Theta$.

7. Существует такое $\delta_2 > 0$, что

а.

$$\mathbf{E}_\theta \sup_{\lambda\in[-\delta_2;\delta_2]} (l_2(X|\theta+\lambda))^2 < \infty, \theta\in\Theta.$$

б.

$$\mathbf{E}_\theta \sup_{\lambda\in[-\delta_2;\delta_2]} |l_m(X|\theta+\lambda)| < \infty, \theta\in K, m=0,\dots,M+3.$$

в.

$$\sup_{\theta\in\Theta}(1+|\theta|^{\delta_2})^{-1} I(\theta) < \infty.$$

8. Априорная плотность g имеет $M+1$ непрерывную производную на Θ.

Напомним, что $I_1(m,v)$ — множество индексов $(i_1,...,i_m)$, удовлетворяющее равенству

$$\sum_{j=1}^m i_j = v,\ i_j>0,\ j=\overline{1,m},$$

и $\Phi(x|\mu,\sigma^2), x\in\mathbb{R}$ — функция нормального распределения со средним μ, дисперсией σ^2. При фиксированном значении выборки $\mathbf{X}$ и любом целом $m\geq 0$ положим

$$F_m(z) = \int\limits_{-\infty}^{z} h^m \varphi\left(h\middle|\mu_n,\sigma_n^2\right) dh,$$

где по-прежнему $\mu_n = n^{-1/2}\Delta_1(T_n)I^{-1}(T_n)$, $\sigma_n^2 = (I(T_n))^{-1}$. Пусть, далее,

$$H_0(z) = f_0(z) = F_0(z),$$

$$H_m(z) = f_m(z) + \sum_{j=1}^m S_j f_{m-j}(z), \quad m=\overline{1,M},$$

$$S_m = \sum_{j=1}^m (-1)^j \sum_{I_1(j,m)} \prod_{l=1}^j f_{i_l}(\infty), \quad m=1,\dots,M,$$

$$f_m(z) = \sum_{j=1}^m \frac{1}{m!} \sum_{I_1(j,m)} \prod_{l=1}^j q_{i_l}(z), \quad m=1,\dots,M,$$

$$q_1(z) = F_1(z)\pi_1(T_n) + \sqrt{n}F_2(z)\frac{n^{-1}\Delta_2(T_n)+I(T_n)}{2} + F_3(z)\frac{\Delta_3(T_n)}{6n},$$

$$q_m(z) = F_{m+2}(z)\frac{\Delta_{m+2}(T_n)}{n(m+2)!} + F_m(z)\frac{\pi_m(T_n)}{m!}, \quad m=2,\dots,M.$$

Теорема 1.4. *Пусть K — компакт, лежащий в Θ вместе с некоторой своей окрестностью и оценка $T_n \in \mathbb{C}(K)$. Пусть для некоторого целого $M \geq 0$ выполнены условия $\mathbb{D}(M, K)$. Тогда*

$$\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) < z|\mathbf{X}\right) = \sum_{m=0}^{M} n^{-m/2} H_m(z) + n^{-(M+1)/2}\omega_n(z),$$

причем

$$\sup_{z\in\mathbb{R}} H_m(z) = O_{\mathbf{P}_\theta}(1) \text{ равномерно по } \theta \in K, m = 0, \ldots, M,$$

$$\sup_{z\in\mathbb{R}} \omega_n(z) = O_{\mathbf{P}_\theta}(1) \text{ равномерно по } \theta \in K.$$

Замечание. То, что в теореме коэффициенты разложения и остаток зависят от $\mathbf{X}$, мы опустили для сокращения записи.

Доказательство. Пусть $A = \delta_2$ из условия 7 и пусть $\alpha \in (0; 0,5)$ — произвольное число, $A_n = An^\alpha$, $B_n = [T_n - A_n/\sqrt{n}; T_n + A_n/\sqrt{n}]$. В силу леммы 1.8 (нужно применить ее тем же образом, что и при доказательстве теоремы 1.3) достаточно доказать, что для любого $\varepsilon > 0$ найдется такое число $c > 0$, что

$$\sup_{\theta\in K} \mathbf{P}_\theta \left(\left| \int_{-\infty}^{z} \frac{1}{\sqrt{n}} \widetilde{q}_n \left(T_n + \frac{h}{\sqrt{n}} \Big| \mathbf{X}\right) dh - \sum_{m=0}^{M} n^{-m/2} H_m(z) \right| > cn^{-(M+1)/2} \right) < \varepsilon,$$

где $\widetilde{q}$ определяется в лемме 1.8.

Как и в доказательстве теоремы 1.3, разделим числитель и знаменатель $\widetilde{q}$ на $p_n(\mathbf{X}|T_n)g(T_n)$, которое больше нуля по условию 4 с вероятностью единица. Сделаем замену $\theta = T_n + h/\sqrt{n}$ в знаменателе $\widetilde{q}$. Тогда достаточно показать, что величина

$$n^{(M+1)/2} \int_{-A_n}^{z} \frac{\widetilde{Z}_n(T_n, h)}{\int_{-A_n}^{A_n} \widetilde{Z}_n(T_n, h) dh} dh - \sum_{m=0}^{M} n^{-m/2} H_m(z) \tag{1.24}$$

плотна относительно $\mathbf{P}_\theta$ равномерно по $\theta \in K$. В силу леммы 1.9 вероятность того, что $T_n + h/\sqrt{n}$ не будет выходить за границы параметрического пространства Θ при $h \in [-A_n; A_n]$ стремится к нулю. Поэтому далее все события пересечены с событием $\{T_n + h/\sqrt{n} \in \Theta\}$ для всех h. Наконец, как и при доказательстве теоремы 1.3, умножим числитель и знаменатель дроби на

$$V_n(T_n) = \exp\left(-\frac{\Delta_1^2(T_n)}{2nI(T_n)}\right)\sqrt{\frac{I(T_n)}{2\pi}}.$$

Дальнейший ход доказательства следующий: разложим в многочлен Тейлора экспоненту от логарифма $\widetilde{Z}_n(T_n, h)$, и от полученного разложения вычислим интегралы в (1.24), после чего доказываем стремление у нулю интегралов от остаточных членов разложения, а также плотность коэффициентов разложения. Затем, получившуюся дробь вида $1/(1+b)$ снова раскладываем в ряд Тейлора, устанавливаем стремление к нулю остатков, что дает необходимый результат.

Разложим $\ln \widetilde{Z}_n(T_n, h)$ по формуле Тейлора относительно точки $h = 0$ с остатком в форме Лагранжа:

$$\ln \widetilde{Z}_n(T_n, h) = \sum_{m=1}^{M+2} \frac{\Delta_m(T_n)}{m!}\left(\frac{h}{\sqrt{n}}\right)^m + \sum_{m=1}^{M} \frac{\pi_m(T_n)}{m!}\left(\frac{h}{\sqrt{n}}\right)^m + r_1(h),$$

$$r_1(h) = \frac{\Delta_{M+3}(T_n + \xi_1(h))}{(M+3)!}\left(\frac{h}{\sqrt{n}}\right)^{M+3} + \frac{\pi_{M+1}(T_n + \xi_2(h))}{(M+1)!}\left(\frac{h}{\sqrt{n}}\right)^{M+1},$$

где $\xi_1(h)$ и $\xi_2(h)$ — некие точки между 0 и h. В силу условий 5 и 8 многочлен Тейлора существует для каждого h.

Положим

$$\psi(h) = \frac{h\Delta_1(T_n)}{\sqrt{n}} - \frac{h^2 I(T_n)}{2},$$

$$L_n(h) = \frac{h^2}{2}\left(\frac{\Delta_2(T_n)}{n} + I(T_n)\right) + \\ + \sum_{m=1}^{M} n^{-m/2}\left(\frac{h^{m+2}}{(m+2)!}\frac{\Delta_{m+2}(T_n)}{n} + \frac{h^m}{m!}\pi_m(T_n)\right) + r_1(h),$$

Тогда

$$\ln \widetilde{Z}_n(T_n, h) = \psi(h) + L_n(h).$$

Разложим экспоненту от $\ln \widetilde{Z}_n(T_n, h)$ по формуле Тейлора относительно точки $\psi(h)$:

$$\exp\{\ln \widetilde{Z}_n(h)\} = \exp\{\psi(h)\}\left[1 + L_n(h) + \frac{1}{2}\left(L_n(h)\right)^2 + \ldots + \frac{1}{M!}\left(L_n(h)\right)^M\right] + \\ + \exp\left\{\psi(h) + \widetilde{L}_n(h)\right\}\frac{1}{(M+1)!}\left(L_n(h)\right)^{M+1}, \quad (1.25)$$

$\widetilde{L}_n(h)$ — некая точка между 0 и $L_n(h)$.

Докажем плотность коэффициентов при $n^{-m/2}$ в получившемся разложении. Сначала установим, что

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\int_{-A_n}^{A_n} V_n(T_n)\exp\{\psi(h)\}|r_1(h)|dh > Cn^{-(M+1)/2}\right) \to 0, C \to \infty$$

равномерно для всех n.

Так как $\xi_2(h) \in [-A_n; A_n]$ и функция $\pi_{M+1}(\theta)$ вследствие условия 8 непрерывна, то величина $|\pi_{M+1}\left(T_n + \xi_2(h)/\sqrt{n}\right)|$ ограничена некоторой постоянной $\widetilde{\pi}_{M+1}$, не зависящей от n. Таким образом,

$$|r_1(h)| \le n^{-(M+1)/2}\left(\frac{|h|^{M+3}}{n(M+3)!}\sup_{|\xi|<A_n}\left|\Delta_{M+3}\left(T_n + \frac{\xi}{\sqrt{n}}\right)\right| + \frac{|h|^{M+1}}{(M+1)!}\widetilde{\pi}_{M+1}\right). \quad (1.26)$$

Для краткости положим

$$\sup_{|\xi|<A_n}\left|\Delta_{M+3}\left(\theta_0 + \frac{\xi}{\sqrt{n}}\right)\right| = \widetilde{\Delta}_{M+3}.$$

Поскольку

$$V_n(T_n)\exp\{\psi(h)\} = \varphi\left(h\middle|\mu_n, \sigma_n^2\right)$$

и $|a+b|^m \le 2^m(|a|^m + |b|^m)$, то

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\int\limits_{-A_n}^{A_n}\varphi\left(h\middle|\mu_n,\sigma_n^2\right)|r_1(h)|dh > Cn^{-(M+1)/2}\right)\leq$$

$$\leq\sup_{\theta\in K}\mathbf{P}_\theta\left(\int\limits_{-\infty}^{\infty}|h+\mu_n|^{M+3}\left(\frac{|\widetilde{\Delta}_{M+3}|}{n}+\widetilde{\pi}_{M+1}\right)\varphi\left(h\middle|0,\sigma_n^2\right)dh > C\right)\leq$$

$$\leq\sup_{\theta\in K}\mathbf{P}_\theta\left(2^{M+3}\left(E_{M+3}\sigma_n^{2(M+3)}+|\mu_n|^{M+3}\right)\left(\frac{|\widetilde{\Delta}_{M+3}|}{n}+\widetilde{\pi}_{M+1}\right) > C\right),$$

где $E_m=\int_{\mathbb{R}}|h|^m\varphi(h)dh$. Итак, в силу леммы 1.2, достаточно показать, что каждый из множителей под знаком вероятности есть плотная последовательность. Это обеспечивается леммой 1.6, непрерывностью $I(\theta)$ и неравенством Маркова ($\sup_{\theta\in K}\mathbf{E}_\theta|\widetilde{\Delta}_{M+3}|<\infty$ по свойству 7б).

Воспользовавшись неравенством Маркова и условием 7б и 8, можно таким же образом показать, что коэффициенты в сумме в $L_n(h)$ при $n^{-m/2}$ плотны:

$$\sup_{\theta\in K}\mathbf{P}_\theta\Bigg(n^{-m/2}\Bigg|\int\limits_{-A_n}^{A_n}\varphi\left(h\middle|\mu_n,\sigma_n^2\right)\Bigg[h^{m+2}\frac{\Delta_{m+2}(T_n)}{n(m+2)!}+$$

$$+\frac{h^m}{m!}\pi_m(T_n)\Bigg]dh\Bigg| > Cn^{-m/2}\Bigg)\to 0, C\to\infty.$$

Первое слагаемое в представлении $L_n(h)$ также является элементом плотной последовательности, ибо по неравенству Чебышева и условию 7а

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(\frac{E_2}{2}\left|\frac{\Delta_2(T_n)}{n}+I(T_n)\right| > Cn^{-1/2}\right)\leq$$

$$\leq\frac{E_2^2}{4C^2}\sup_{\theta\in K}\mathbf{E}_\theta\left(l_2(T_n)+I(T_n)\right)^2=O(C^{-2}).$$

Таким образом, установлено, что для интеграла от $L_n(h)$ коэффициенты являются плотными последовательностями случайных величин. Для оставшихся членов в (1.25) применим лемму 1.2, что позволит утверждать плотность всех коэффициентов при $n^{-m/2}$ кроме остатка.

Теперь обратимся к доказательству плотности остатка разложения экспоненты (1.25). Мы хотим показать, что последовательность

$$n^{(M+1)/2}\int\limits_{-A_n}^{A_n} V(T_n)\exp(\psi(T_n,h)+\widetilde{L}_n(h))\,(L_n(h))^{M+1}\,dh \tag{1.27}$$

плотна. Доказательство будет проводиться по аналогии с доказательством в теореме 1.3. Из определения $\widetilde{L}_n$ видно, что $0 \le |\widetilde{L}_n(h)| \le |L_n(h)|$. Имеем неравенство (модуль суммы не превосходит суммы модулей):

$$h^{-2}|L_n(h)| \le \frac{1}{2}\left|\frac{\Delta_2(T_n)}{n}+I(T_n)\right| + \\ + \sum_{m=1}^{M}\left[\left(\frac{h}{\sqrt{n}}\right)^m \frac{1}{(m+2)!}\frac{|\Delta_{m+2}(T_n)|}{n} + \left(\frac{h}{\sqrt{n}}\right)^{m-2}\frac{|\pi_m(T_n)|}{nm!}\right] + |r_1(h)|.$$

Величину $|r_1(h)|$ можно ограничить тем же образом, как и в последнем выражении, используя неравенство (1.26). Таким образом, $h^{-2}|L_n(h)|$ зависит от h только через $hn^{-1/2}$, и при этом $|h|n^{-1/2} \le An^{\alpha-1/2} \to 0$. Отсюда и из доказанной плотности коэффициентов при степенях $hn^{-1/2}$ для любого числа $0<\beta<1/2-\alpha$ и некоторого $C_1>0$ вероятность

$$\sup_{\theta\in K}\mathbf{P}_\theta\left(h^{-2}n^{\beta}|L_n(h)|>C_1\right)\to 0, n\to\infty.$$

Так же, как и при доказательстве теоремы 1.3, рассмотрим множества

$$U_n=\left\{h^{-2}n^{\beta}|L_n(h)|>C_1\right\}, Y_n=\left\{|\mu_n|>n^{\beta/4}\right\},$$

Только что было показано, что $\sup_{\theta\in K}\mathbf{P}_\theta\left(U_n\right)\to 0$, и $\sup_{\theta\in K}\mathbf{P}_\theta\left(Y_n\right)\to 0$ (см.

лемму 1.6). Для $\mathbf{X} \in U_n^c \cap Y_n^c$ одно из подынтегральных выражений в (1.27)

$$
\begin{aligned}
&V(T_n)\exp(\psi(T_n,h)+\widetilde{L}_n(h)) \le \\
&\qquad\le \varphi\left(h\middle|\mu_n,\sigma_n^2\right)\exp(|L_n(h)|) \le \varphi\left(h\middle|\mu_n,\sigma_n^2\right)\exp(h^2n^{-\beta}C_1) = \\
&\qquad= \varphi\left(h\middle|\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-2C_1n^{-\beta})},\frac{1}{I(T_n)-2C_1n^{-\beta}}\right)\times \\
&\qquad\times\exp\left[\frac{\Delta_1^2(T_n)}{2nI^2(T_n)}\left(\frac{2C_1n^{-\beta}I(T_n)}{I(T_n)-2C_1n^{-\beta}}\right)\right] \le \\
&\qquad\le \varphi\left(h\middle|\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-2C_1n^{-\beta})},\frac{1}{I(T_n)-2C_1n^{-\beta}}\right)\exp\left[\frac{C_1n^{-\beta/2}I(T_n)}{I(T_n)-2C_1n^{-\beta}}\right].
\end{aligned}
$$

Величина под экспонентой в правой части в последней цепочке неравенств стремится к нулю по вероятности $\mathbf{P}_\theta$ равномерно по $\theta \in K$. Таким образом, интеграл (1.27) не превосходит интеграла

$$
n^{(M+1)/2}\int_{-A_n}^{A_n}\varphi\left(h\middle|\frac{\Delta_1(T_n)}{\sqrt{n}(I(T_n)-2C_1n^{-\beta})},\frac{1}{I(T_n)-2C_1n^{-\beta}}\right)(L_n(h))^{M+1}\,dh. \tag{1.28}
$$

Как было показано ранее, интеграл вида $\int\sqrt{n}|L_n(h)|\varphi(h|m_n,s_n^2)dh$ является плотной последовательностью в случае, если последовательность средних m_n является плотной, а дисперсия s^2 не стремится к нулю. Эти условия выполнены для (1.28), поэтому это выражение является плотной последовательностью.

Обратимся снова к разложению (1.25). Как было показано, после интегрирования по h в пределах от $-A_n$ до A_n коэффициенты и остаток разложения (1.25) удовлетворяют утверждению теоремы. Для удобства заменим пределы интегрирования на бесконечности, что возможно сделать в силу леммы 1.7. Таким образом,

$$
\int_{-A_n}^{z}\widetilde{Z}_n(T_n,h)dh = \sum_{m=0}^{M}n^{-m/2}f_m(z)+n^{-(M+1)/2}R_1(z), \tag{1.29}
$$

где последовательность $\sup_{z\in\mathbb{R}}R_1(z)$ плотна относительно $\mathbf{P}_\theta$ равномерно по $\theta \in K$ согласно сказанному выше.

Знаменатель интеграла в (1.24) примет вид (1.29) лишь с заменой z на ∞ (заметим, что главный член окажется равным единице). Воспользуемся хорошо известным представлением:

$$\frac{1}{1+b} = \sum_{m=0}^{\infty} (-b)^m,$$

справедливым при $|b| < 1$. Так как $n^{-m/2} f_m(\infty) \xrightarrow{\mathbf{P}_\theta} 0$ равномерно по $\theta \in K$, то можно подставить в это разложение $b = \sum_{m=1}^{M} n^{-m/2} f_m(\infty) + n^{-(M+1)/2} R_1(z)$. Домножая далее на полученное ранее разложение числителя, получаем утверждение теоремы. Плотность относительно $\mathbf{P}_\theta$ равномерно по $\theta \in K$ коэффициентов при $n^{-m/2}$ и остатка доказывается так же с помощью леммы 1.2. ■

По аналогии со следствием 1.1, можно сформулировать следующее утверждение, доказательство которого полностью совпадает с доказательством следствия 1.1.

Следствие 1.2. *Пусть для любого компакта $K \in \Theta$, входящего в Θ с некоторой окрестностью, оценка $T_n \in \mathbb{C}(K)$ и выполнены условия $\mathbb{D}(M, K)$. Тогда*

$$\mathbf{P}\left(\sqrt{n}(\vartheta - T_n) < z|\mathbf{X}\right) = \sum_{m=0}^{M} n^{-m/2} H_m(z) + n^{-(M+1)/2}\omega_n, \quad \omega_n = O_{\mathbf{P}}(1).$$

1.2.4. Разложение апостериорных средних

Практически не меняя доказательства предыдущей теоремы, можно получить разложение моментов апостериорного распределения, или, более общо, средних значений от некоторых измеримых функций w относительно апостериорного распределения:

$$\mathbf{E}\left\{w(\sqrt{n}(\vartheta - T_n))\big|\mathbf{X}\right\} = \frac{\int_\Theta w(\sqrt{n}(\theta - T_n)) p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}{\int_\Theta p_n\left(\mathbf{X}|\theta\right) g(\theta) d\theta}.$$

Ширина класса функций w определяется узостью класса оценок T_n.

Определение 1.2. *Пусть K — некоторый компакт из Θ, а некоторое число $r \geq 0$. Будем говорить, что оценка T_n принадлежит классу оценок $\mathbb{C}(K, r)$ параметра $\theta \in \Theta$, если T_n измерима, $T_n \in \mathbb{C}(K)$ и*

$$\sup_{\theta \in K} \mathbf{E}_\theta |\sqrt{n}(T_n - \theta)|^r < \infty.$$

Обозначим

$$F_{m,w} = \int_{-\infty}^{\infty} h^m w(h) \varphi\left(h \middle| \mu_n, \sigma_n^2\right) dh,$$

$$H_{0,w} = f_{0,w} = F_{0,w},$$

$$H_{m,w} = f_{m,w} + \sum_{j=1}^{m} S_j f_{m-j,w}, \quad m = 1, \ldots, M,$$

$$f_{m,w} = \sum_{j=1}^{m} \frac{1}{m!} \sum_{I_1(j,m)} \prod_{l=1}^{j} q_{i_l,w}, \quad m = 1, \ldots, M,$$

$$q_{1,w} = F_{1,w}\pi_1(T_n) + \sqrt{n} F_{2,w} \frac{n^{-1}\Delta_2(T_n) + I(T_n)}{2} + F_{3,w} \frac{\Delta_3(T_n)}{6n},$$

$$q_{m,w} = F_{m+2,w} \frac{\Delta_{m+2}(T_n)}{n(m+2)!} + F_{m,w} \frac{\pi_m(T_n)}{m!}, \quad m = 2, \ldots, M,$$

а коэффициенты S_j введены перед утверждением теоремы 1.4) на странице 1.2.3.

Теорема 1.5. *Пусть выполнены условия $\mathbb{D}(M, K)$, оценка $T_n \in \mathbb{C}(K, r)$, $r \geq 0$, а измеримая функция w имеет полиномиальную мажоранту порядка не выше r. Тогда найдется n_0, что для всех $n > n_0$*

$$\mathbf{E}\left\{ w\left(\sqrt{n}(\vartheta - T_n)\right) \middle| \mathbf{X} \right\} = \sum_{m=0}^{M} n^{-m/2} H_{m,w} + n^{-(M+1)/2} \omega_n, \tag{1.30}$$

причем

$$H_{m,w} = O_{\mathbf{P}_\theta}(1) \text{ равномерно по } \theta \in K, m = 0, \ldots, M,$$

$$\omega_n = O_{\mathbf{P}_\theta}(1) \text{ равномерно по } \theta \in K.$$

Доказательство. Поскольку доказательство аналогично тому, что проводилось в теореме 1.4, то ограничимся лишь некоторыми моментами, в которых будут использоваться условия на w.

Первый этап доказательства состоит в замене области интегрирования с $\sqrt{n}(\Theta - T_n)$ на $[-A_n; A_n]$. Так как этот этап основан на применении леммы 1.8, то формально необходимо передоказать лемму для случая, когда в числителе стоит не $p_n(\mathbf{X}|\theta)g(\theta)$, а $w(\sqrt{n}(\theta - T_n))p_n(\mathbf{X}|\theta)g(\theta)$. В случае, когда w является полиномом, необходима версия леммы 1.3, в которой изучается поведение $|h + \sqrt{n}(T_n - \theta_0)|^r Z_n(\theta_0, h)$. Вследствие простого неравенства $|h + \sqrt{n}(T_n - \theta_0)|^r \leq 2^r|h|^r + 2^r|\sqrt{n}(T_n - \theta_0)|^r$ нужны утверждения по каждому из слагаемых. Утверждение о $|h|^r$ уже содержится в лемме 1.3, а интеграл

$$\sup_{\theta_0 \in K} \mathbf{E}_{\theta_0} \int\limits_{\mathbb{R}} \left|\sqrt{n}(T_n - \theta_0)\right|^r Z_n(\theta_0, h)dh$$

конечен вследствие условия $T_n \in \mathbb{C}(K, r)$ и той же леммы 1.3. Таким образом, сужение области интегрирования возможно.

Легко видеть, что формально применяя то же разложение Тейлора к величине $\widetilde{Z}_n(T_n, h)$, что и в теореме 1.4, после чего интегрируя и используя разложение $1/(1+b)$, получим коэффициенты, указанные в утверждении текущей теоремы. Остается показать плотность коэффициентов и остатка. Ранее это опиралось на существование интеграла вида $\int_{\mathbb{R}} w(h)h^m\varphi(h)dh$ и экспоненциальную (по $x \to \infty$) малость $\int_x^\infty w(h)h^m\varphi(h)dh$. Оба свойства выполняются для функций, ограниченных полиномиальной мажорантой, первое — по свойствам интеграла Эйлера, второе — из леммы 1.7. ■

Теорема 1.5 является прямым обобщением теоремы 1.4, в которой $w(h) = I_{\{h<z\}}(h)$, а условие $T_n \in \mathbb{C}(K, 0)$, очевидно, выполнено для любой оценки T_n.

Теорема 1.5 допускает обобщения для функций w, удовлетворяющих следующим условиям: величина $w(h)Z_n(\theta_0, h)$ экспоненциально убывает с ростом h, интеграл $\int_{\mathbb{R}} w(h)h^m\varphi(h)dh$ существует для всех m, и остаток $\int_x^\infty w(h)h^m\varphi(h)dh$ убывает экспоненциально по $x \to \infty$ для всех m.

До сих пор было приведено лишь утверждение о разложении апостериорных средних как аналога теоремы 1.4, однако можно также привести аналог теоремы 1.3 относительно моментных характеристик, то есть привести асимптотическое значение для средних. Понятно, что в качестве предела будет выступать среднее $\mathbf{E}w(Y)$, в котором $Y \sim N(\mu_n, \sigma_n^2)$, и которое совпадает с первым членом разложения (1.30).

Теорема 1.6. *Пусть выполнены условия* $\mathbb{D}(K)$, *оценка* $T_n \in \mathbb{C}(K, r)$, *а функция* w *есть измеримое отображение* $\mathbb{R} \mapsto \mathbb{R}$ *и имеет полиномиальную мажоранту, чей порядок не превышает* r. *Тогда найдется* n_0, *что для всех* $n > n_0$ *и любого* $\varepsilon > 0$

$$\sup_{\theta \in K} \mathbf{P}_\theta \left(\left| \mathbf{E} \left\{ w\left(\sqrt{n}(\vartheta - T_n)\right) \middle| \mathbf{X} \right\} - \mathbf{E}w(Y) \right| > \varepsilon \right) \to 0,$$

где распределение случайной величины Y *определено выше.*

Доказательство этой теоремы опущено в связи с тем, что рассуждения в нем идентичны рассуждениям при доказательстве теоремы 1.5.

Для асимптотики апостериорных средних и разложения апостериорных средних можно высказать аналог следствий 1.1 и 1.2 относительно равномерности остатка по мере $\mathbf{P}$.

Глава 2

Асимптотика необходимого объема выборки при d-гарантийном различении двух гипотез

Определение минимального объема наблюдений, при котором существует d-гарантийный критерий, является одной из важнейших задач приложений d-апостериорного подхода к задачам контроля качества. В статье Симушкина [6] показано, что критерий, минимизирующий объем наблюдений, должен быть основан на апостериорных вероятностях соответствующих гипотез, при этом d-риски первого и второго рода этого критерия должны в точности совпадать с заявленными ограничениями. Один из первых результатов в области асимптотического анализа необходимого объема выборки принадлежит работе Володина и Новикова [9], которые рассмотрели ситуацию, когда ограничения на d-риски стремятся к нулю. В этой главе полученная ими асимптотическая формула устанавливается при более слабых условиях, при этом доказательство опирается в основном на теорему Бернштейна-фон Мизеса, что значительно упрощает необходимые выкладки.

2.1. Постановка задачи на отыскание необходимого объема выборки

В этой главе рассматривается та же вероятностная модель с независимыми одинаково распределенными случайными величинами и $\Theta \subset \mathbb{R}$, что и в главе 1.

Рассматривается задача проверки гипотезы $H_0 : \theta < \theta_0,\ \theta_0 \in \Theta$ против альтернативы $H_1 : \theta \geq \theta_0$. Пусть d_0 — решение о принятии гипотезы H_0, d_1 — решение о принятии альтернативы. Критерий ϕ_n, проверяющий эти гипотезы, должен удовлетворять ограничениям β_0 и β_1 на d-риски

$$\mathcal{R}(d_0|\phi_n) = \mathbf{P}(\vartheta \in \Theta_1|\varphi_n(\mathbf{X}) = 0) \le \beta_0, \tag{2.1}$$

$$\mathcal{R}(d_1|\phi_n) = \mathbf{P}(\vartheta \in \Theta_0|\varphi_n(\mathbf{X}) = 1) \le \beta_1, \tag{2.2}$$

Требуется построить критерий ϕ_n, для которого указанные ограничения выполняются при минимально возможном объеме выборки n. Предлагается критерий (см. [6])

$$\phi_n^*(\mathbf{X}) = \begin{cases} 1, & R(\mathbf{X}) \ge C, \\ 0, & R(\mathbf{X}) < C, \end{cases}$$

где $R(\mathbf{X}) = \mathbf{P}(\vartheta \in \Theta_1|\mathbf{X})$, а критическая константа C выбирается так, чтобы выполнялись неравенства (2.1) и (2.2).

В этой главе решается задача отыскания асимптотической (при $\beta_0, \beta_1 \to 0$) формулы для минимального объема выборки n, при котором критерий ϕ_n^* удовлетворяет ограничениям (2.1) и (2.2).

Обозначим $H_n(c|\theta) = \mathbf{P}_\theta\left(R(\mathbf{X}) < c\right)$. Таким образом, выражения (2.1), (2.2) для рисков оптимальной процедуры ϕ_n^* примут вид

$$\mathcal{R}(d_0|\phi_n^*) = \frac{\int_{\Theta_1} H_n(C|\theta)g(\theta)d\theta}{\int_{\Theta} H_n(C|\theta)g(\theta)d\theta} \le \beta_0,$$

$$\mathcal{R}(d_1|\phi_n^*) = \frac{\int_{\Theta_0}(1 - H_n(C|\theta))g(\theta)d\theta}{\int_{\Theta}(1 - H_n(C|\theta))g(\theta)d\theta} \le \beta_1.$$

В дальнейшем будут использоваться следующие обозначения:

$$\Lambda_n(\theta) = \frac{\Delta_1(\theta)}{\sqrt{nI(\theta)}},$$

$$Q_n(\theta) = \sqrt{nI(\theta)}(\theta - \theta_0),$$

$$q_c = \Phi^{-1}(c),$$

где Φ^{-1} — квантиль нормального распределения.

2.2. Асимптотика необходимого объема выборки

2.2.1. Условия регулярности и предварительные утверждения

Будем предполагать выполнение следующих условий, которые мы назовем $\mathbb{D}_{\theta_0}$.

1. Точка θ_0, разграничивающая гипотезы, является внутренней точкой множества Θ.

2. Для $\theta \neq \theta'$
$$\int\limits_{\mathfrak{X}} |p(x|\theta) - p(x|\theta')|\, \nu(dx) > 0.$$

3. Найдется $\delta_1 > 0$, такое что
$$\sup_{\theta \in \Theta} |\theta - \theta_0|^{\delta_1} \int\limits_{\mathfrak{X}} \sqrt{p(x|\theta) p(x|\theta_0)} \nu(dx) < \infty.$$

4. Функция $p(x|\theta)$ для всех $x \in \mathfrak{X}$ дважды дифференцируема по θ в окрестности точки θ_0.

5. Плотность $p(x|\theta)$ выбрана так, что она непрерывна по x.

6. Информация по Фишеру $I(\theta) > 0$ для $\theta \in \Theta$.

7. Априорная плотность $g(\theta)$ непрерывна для всех внутренних точек $\theta \in \Theta$ и обладает непрерывной производной в точке θ_0.

8. Для всех $\theta \in \Theta$ и всех $\gamma > 0$
$$\inf_{|\theta' - \theta| > \gamma} r_2(\theta; \theta') = \inf_{|\theta' - \theta| > \gamma} \int \left(p^{1/2}(x|\theta) - p^{1/2}(x|\theta') \right)^2 d\nu = k_\theta(\gamma) > 0.$$

9. Для всех $\theta \in \Theta$
$$\int \sup_{|h| \le \delta} \left(p^{1/2}(x|\theta) - p^{1/2}(x|\theta + h) \right)^2 d\nu = \omega_\theta(\delta) \xrightarrow{\delta \to 0} 0.$$

10. Для всех θ из некоторой окрестности θ_0

$$\mathbf{E}_\theta (l_2(X|\theta))^2 < \infty.$$

11. Информация Фишера $I(\theta)$ ограничена для всех $\theta \in \Theta$.

Пусть, как и раньше, $\mathcal{N}(A|m, s^2)$ есть нормальное распределение со средним m и дисперсией s^2, вычисленное на борелевском множестве $A \subset \mathbb{R}$, а $\Phi(x|m, s^2)$ и $\varphi(x|m, s^2)$ — соответствующие функция распределения и плотность, вычисленные в точке x. Символами $\Phi(x)$ и $\varphi(x)$ без параметров будем обозначать функции распределения и плотности стандартного нормального закона.

Нам понадобится следующее утверждение, доказательство которого можно найти в [10], теорема 10.1. Определение дифференцируемости вероятностной модели в среднем квадратичном можно найти там же, с. 93.

Лемма 2.1 (Теорема Бернштейна-фон Мизеса) *Пусть в точке θ вероятностная модель дифференцируема в среднем квадратичном, $I(\theta) > 0$, $g(\theta) > 0$ и непрерывна в этой точке. Пусть, далее, существует критерий ϕ_n такой, что для всех $\delta > 0$*

$$\mathbf{E}_\theta \phi_n(\mathbf{X}) \xrightarrow{n\to\infty} 0, \qquad \sup_{|u-\theta|>\delta} \mathbf{E}_u (1 - \phi_n(\mathbf{X})) \xrightarrow{n\to\infty} 0.$$

Тогда

$$\left\| \mathbf{P}\left(\sqrt{n}(\vartheta - \theta) \in \cdot |\mathbf{X}\right) - \mathcal{N}\left(\sqrt{I(\theta)}(A - \Lambda_n(\theta))\right) \right\| \xrightarrow{\mathbf{P}_\theta} 0.$$

Сразу заметим, что условия предыдущей леммы выполнены для всех точек $\theta \in \Theta$ для вероятностной модели, удовлетворяющей условиям 1–9. Это следует из существования равномерно состоятельной оценки максимального правдоподобия (см. [26], теорема 4.3, замечание 4.1), что обеспечивает существование критерия, фигурирующего в условиях леммы. Дифференцируемость в среднем квадратичном следует из условий 4 и 6.

Таким образом, в условиях $\mathbb{D}_{\theta_0}$ справедливо представление

$$R(\mathbf{X}) = \Phi(\Lambda_n(\theta) + Q_n(\theta)) + \delta_{n,\theta}, \tag{2.3}$$

где $\mathbf{P}_\theta(|\delta_{n,\theta}| > \varepsilon) \to 0$ при всех $\varepsilon > 0$ для θ в окрестности θ_0. Однако, следует сказать, что этого утверждения недостаточно для аппроксимации распределения $R(\mathbf{X})$ по всему классу мер $\mathbf{P}_\theta, \theta \in \Theta$. Именно, для $|\theta - \theta_0| > M$ при некотором $M > 0$ сходимости $\delta_{n,\theta}$ к нулю по вероятности недостаточно, и требуется оценка вероятностей больших уклонений для апостериорного распределения. Следующая лемма содержит необходимое нам утверждение.

Лемма 2.2. *При выполнении условий $\mathbb{D}_{\theta_0}$ для любых $A > 0, a > 0$ существует константа C_a, которая зависит только от a, такая что*

$$\mathbf{P}_\theta\left(R(\mathbf{X}) > A^{-2a}\right) \le C_a A^{-a}$$

при всех $\sqrt{n}(\theta_0 - \theta) > A$, и

$$\mathbf{P}_\theta\left(R(\mathbf{X}) < 1 - A^{-2a}\right) \le C_a A^{-a}$$

при всех $\theta - \theta_0 > n^{-1/2}A$.

Доказательство. Докажем первую формулу в утверждении леммы. Вторая получится формальной заменой $R(\mathbf{X})$ на $1 - R(\mathbf{X})$ из первой.

Обозначим $A_n(\theta) = \sqrt{n}(\theta_0 - \theta)$. Имеем $A_n(\theta) < A$, и

$$\mathbf{P}_\theta\left(R(\mathbf{X}) > A^{-2a}\right) = \mathbf{P}_\theta\left(\frac{\int_{h > A_n(\theta)} Z_n(\theta,h) g(\theta + n^{-1/2}h)dh}{\int_{\mathbb{R}} Z_n(\theta,h) g(\theta + n^{-1/2}h)dh} > A^{-2a}\right) \le$$

$$\le \mathbf{P}_\theta\left(\int\limits_{h > A_n(\theta)} Z_n(\theta,h) g\left(\theta + \frac{h}{\sqrt{n}}\right) dh > A^{-a}\right) +$$

$$+ \mathbf{P}_\theta\left(\int\limits_{\mathbb{R}} Z_n(\theta,h) g\left(\theta + \frac{h}{\sqrt{n}}\right) dh < A^{-a}\right).$$

Для первого слагаемого согласно лемме 1.1 справедливо:

$$\mathbf{P}_\theta\left(\int\limits_{h>A_n(\theta)} Z_n(\theta,h)g\left(\theta+\frac{h}{\sqrt{n}}\right)dh > A^{-a}\right) \le$$

$$\le \mathbf{P}_\theta\left(\int\limits_{h>A_n(\theta)} Z_n(\theta,h)g\left(\theta+\frac{h}{\sqrt{n}}\right)dh > A_n(\theta)^{-a}\right) \le K_{a,\theta}A_n(\theta)^{-a}, \quad (2.4)$$

где $K_{a,\theta}$ — некоторая константа, зависящая только от a и θ. Нам необходимо показать, что $K_{a,\theta}A_n(\theta)^{-a} \le K_a A^{-a}$ для некоторой константы K_a. Так как для любого $M > 0$ супремум

$$\sup_{A/\sqrt{n}<\theta_0-\theta<M} K_{a,\theta} < \infty,$$

то остается доказать неравенство для $\theta_0 - \theta > M$. Для доказательства ограниченности супремума в этой области обратимся к доказательству теоремы 2.3 статьи [23]. Из леммы 2.8 работы [23] следует, что для всех $N > 0$ можно найти числа n_0 и $M_0 > 0$, что для $n > n_0$ при $|h| > \sqrt{n}M_0$

$$\mathbf{P}_\theta\left(Z_n(\theta,h) > |h|^{-N}\right) \le |h|^{-N}.$$

Из условия 3 следует, что числа M_0 и n_0 не зависят от θ. Таким образом, в области $\theta_0 - \theta > M_0$ константа $K_{a,\theta}$ не зависит от θ, что следует из леммы 2.9 статьи [23] (здесь уже потребуется условие 11). Поэтому можно принять $K_a = \sup_{A/\sqrt{n}<\theta_0-\theta} K_{a,\theta}$, и $K_a < \infty$.

Для того, чтобы ограничить второе слагаемое в правой части (2.4), заметим, что

$$\int Z_n(\theta,h)g\left(\theta+\frac{h}{\sqrt{n}}\right)dh \ge \int\limits_{-A^{-a}}^{A^{-a}} Z_n(\theta,h)g\left(\theta+\frac{h}{\sqrt{n}}\right)dh,$$

так что

$$\mathbf{P}_\theta\left(\int Z_n(\theta,h)g\left(\theta+\frac{h}{\sqrt{n}}\right)dh < A^{-a}\right) \le B_\theta A^{-2a},$$

что следует из леммы 3.1 статьи [24]. При этом величина

$$B_\theta = C \sup_{0<h<A^{-a}} I(\theta + h/\sqrt{n})$$

при некоторой константе $C > 0$. По условию 11 величина $B_\theta \leq B$ при некотором B, и поэтому не зависит от θ.

Итак, если выбрать $C_a = \max(K_a, B)$, то из приведенных выкладок следует утверждение леммы. ■

Лемма 2.3. *([26], замечание 4.1, стр. 55). Если выполнены условия 1, 8, 9, то*

$$\sup_{\theta\in\Theta} \omega_\theta(\delta) \xrightarrow{\delta\to 0} 0.$$

Лемма 2.4. *При выполнении условий 1 и 9 функция $H_n(c|\theta)$ равномерно непрерывна по $\theta \in \Theta$ равномерно по всем c.*

При выполнении условия 5 функция $H_n(c|\theta)$ непрерывна по $c \in [0;1]$ для всех $\theta \in \Theta$.

Доказательство. Зададим два числа θ и θ'. Имеем:

$$|H_n(c|\theta) - H_n(c|\theta')| = \left| \int\limits_{R(X)<c} (p_n(X|\theta) - p_n(X|\theta'))\, d\nu \right| \leq$$
$$\leq \int |p_n(X|\theta) - p_n(X|\theta')|\, d\nu =$$
$$= \int \left| \left(p_n^{1/2}(X|\theta) - p_n^{1/2}(X|\theta')\right) \left(p_n^{1/2}(X|\theta) + p_n^{1/2}(X|\theta')\right) \right| d\nu.$$

Квадрат последнего выражения можно оценить с помощью неравенства Коши-Буняковского:

$$\left(\int \left| \left(p_n^{1/2}(X|\theta) - p_n^{1/2}(X|\theta')\right) \left(p_n^{1/2}(X|\theta) + p_n^{1/2}(X|\theta')\right) \right| d\nu \right)^2 \leq$$
$$\int \left(p_n^{1/2}(X|\theta) - p_n^{1/2}(X|\theta')\right)^2 d\nu \int \left(p_n^{1/2}(X|\theta) + p_n^{1/2}(X|\theta')\right)^2 d\nu \leq 4r_2^2(\theta, \theta').$$

Вследствие леммы 2.3 расстояние по Хеллингеру непрерывно по θ всюду в Θ. Учитывая, что согласно условию 1 Θ — компакт, а также то, что последнее выражение не зависит от c, мы получим искомую равномерность.

Что касается непрерывности по c, это следует из непрерывности функции $R(\mathbf{X})$. В самом деле, из непрерывности $p(x|\theta)$ по x следует непрерывность по $x_1, \ldots, x_n$ интеграла вида $F(A, x_1, \ldots, x_n) = \int_A \prod_{i=1}^n p(x_i|\theta) g(\theta) d\theta$. Так как $R(\mathbf{X}) = F(\Theta_1, x_1, \ldots, x_n)/F(\Theta, x_1, \ldots, x_n)$, то $R(x_1, \ldots, x_n)$ также непрерывна по $x_1, \ldots, x_n$. Отсюда следует, что $H_n(c|\theta)$ непрерывна по c, как функция распределения непрерывного преобразования случайных величин, имеющих непрерывное распределение. ■

Следующая лемма описывает поведение распределения $R(\mathbf{X})$ вблизи θ_0, то есть при $\theta \in [\theta_0 - M_n n^{-1/2}; \theta_0 + M_n n^{-1/2}]$, $M_n \to \infty$. Заметим, что необходимость учета возможности стремления $M_n \to \infty$ с ростом n, доставляла основную сложность при доказательстве асимптотической нормальности апостериорного распределения и при построении асимптотических разложений в предыдущей главе (см. в связи с этим также доказательство утверждения леммы 2.1, [10], глава 10). В то же время классические теоремы о локальной асимптотической нормальности (см., например, [29], глава 2) описывают лишь случай ограниченных M_n.

Лемма 2.5. *Пусть выполнены условия $\mathbb{D}_{\theta_0}$. Если для некоторой числовой последовательности $\varepsilon_n \to 0$, такой, что $\varepsilon_n \sqrt{n} \to \infty$ выполняется*

$$\lim_{n\to\infty} \mathbf{P}_{\theta_0} \left(|\delta_{n,\theta_0}| > \varepsilon_n\right) = 0,$$

то для $|h| \le M_n$, $M_n/\sqrt{n} \to 0, \sqrt{n}\varepsilon_n/M_n \to \infty$

$$\lim_{n\to\infty} \mathbf{P}_{\theta_0 + hn^{-1/2}} \left(|\delta_{n,\theta_0+hn^{-1/2}}| > \varepsilon_n\right) = 0.$$

Доказательство. При доказательстве этой леммы будем использовать следующие обозначения:

$$\mathbf{P}_h = \mathbf{P}_{\theta_0 + hn^{-1/2}}, \;\; I_h = I(\theta_0 + hn^{-1/2}),$$

$$\Lambda_n(h) = \Lambda_n\left(\theta_0 + hn^{-1/2}\right), \ \delta_{n,h} = \delta_{n,\theta_0+hn^{-1/2}};$$

в дальнейшем эти обозначения встречаться не будут.

Имеем

$$\mathbf{P}_h\left(|\delta_{n,h}| > \varepsilon_n\right) \le \mathbf{P}_h\left(|\delta_{n,0}| + |\delta_{n,h} - \delta_{n,0}| > \varepsilon_n\right) \le$$
$$\le \mathbf{P}_h\left(|\delta_{n,0}| > \varepsilon_n\right) + \mathbf{P}_h\left(|\delta_{n,h} - \delta_{n,0}| > \varepsilon_n\right).$$

Вследствие того, что $|\mathbf{P}_h(A) - \mathbf{P}_0(A)| \le 2r_2(\theta_0, \theta_0 + h/\sqrt{n}) \to 0$ (что было показано в предыдущей лемме) и условия $\mathbf{P}_0(|\delta_{n,0}| > \varepsilon_n) \to 0$, получим $\mathbf{P}_h(|\delta_{n,0}| > \varepsilon_n) \to 0$.

Для второго слагаемого распишем ошибку нормального приближения $\delta_{n,h} = R(\mathbf{X}) - \Phi\left(\Lambda_n(h) + h\sqrt{I_h}\right)$:

$$\mathbf{P}_h\left(|\delta_{n,h} - \delta_{n,0}| > \varepsilon_n\right) = \mathbf{P}_h\left(\left|\Phi\left(\Lambda_n(h) + h\sqrt{I_h}\right) - \Phi(\Lambda_n(0))\right| > \varepsilon_n\right).$$

По формуле Тейлора ($z_n(h)$ — остаток в форме Пеано)

$$\Lambda_n(0) = \sqrt{\frac{I_h}{I_0}}\Lambda_n(h) - \frac{h}{n\sqrt{I_0}}\Delta_2\left(\theta_0 + \frac{h}{\sqrt{n}}\right) + z_n(h).$$

Вторая производная в окестности θ_0 существует в силу условия 4.

Опять же по формуле Тейлора для Φ имеем

$$\Phi\left(\Lambda_n(h) + h\sqrt{I_h}\right) = \Phi(\Lambda_n(0)) + \left[\Lambda_n(h)\left(1 - \sqrt{\frac{I_h}{I_0}}\right) + \right.$$
$$\left. + h\sqrt{I_h} + \frac{h}{n\sqrt{I_0}}\Delta_2\left(\theta_0 + \frac{h}{\sqrt{n}}\right) - z_n(h)\right]\varphi(\Lambda_n(0) + \nu),$$
$$0 < |\nu| < |\Lambda_n(0) - \Lambda_n(h)|.$$

Объединив два предыдущих тождества, получим

$$\mathbf{P}_h\left(\left|\Phi\left(\Lambda_n(h) + h\sqrt{I_h}\right) - \Phi(\Lambda_n(0))\right| > \varepsilon_n\right) \le$$
$$\le \mathbf{P}_h\left(\left|\Lambda_n(h)\left(1 - \sqrt{\frac{I_h}{I_0}}\right)\right| > \varepsilon_n\right) + \mathbf{P}_h\left(|z_n(h)| > \varepsilon_n\right) +$$
$$+ \mathbf{P}_h\left(\left|\sqrt{I_h} + \frac{1}{n\sqrt{I_0}}\Delta_2\left(\theta_0 + \frac{h}{\sqrt{n}}\right)\right|\varphi(0) > \frac{\varepsilon_n}{|h|}\right).$$

Вследствие выполнения условия Липшица для информации (см. [26], с.94) и условия 6

$$\left|1-\sqrt{\frac{I_h}{I_0}}\right| = \frac{1}{\sqrt{I_0}}\left|\sqrt{I_0}-\sqrt{I_h}\right| = \frac{1}{2\sqrt{I_0}\left(I_{\tilde{h}}\right)^{3/2}}|I_0-I_h| \le L\frac{h}{\sqrt{n}}$$

с некоторой константой L. Таким образом,

$$\mathbf{P}_h\left(\left|\Lambda_n(h)\left(1-\sqrt{\frac{I_h}{I_0}}\right)\right| > \varepsilon_n\right) \to 0$$

вследствие слабой сходимости $\Lambda_n(h)$ к $N(0,1)$ относительно $\mathbf{P}_h$.

Ввиду условия $\sqrt{n}\varepsilon_n/M_n \to \infty$ имеем

$$\mathbf{P}_h\left(|z_n(h)| > \varepsilon_n\right) \le \mathbf{P}_h\left(\left|\frac{z_n(h)}{h/\sqrt{n}}\right| > \frac{\varepsilon_n}{h/\sqrt{n}}\right) \le$$
$$\le \mathbf{P}_h\left(\sup_{|h|\le M_n}\left|\frac{z_n(h)}{h/\sqrt{n}}\right| > \frac{\varepsilon_n\sqrt{n}}{M_n}\right) \to 0.$$

Вследствие условия 10 величина

$$\frac{1}{\sqrt{n}}\left(\Delta_2\left(\theta_0+\frac{h}{\sqrt{n}}\right)+nI_h\right)$$

асимптотически нормальна. Поэтому

$$\mathbf{P}_h\left(\left|\sqrt{I_h}+\frac{1}{n\sqrt{I_0}}\Delta_2\left(\theta_0+\frac{h}{\sqrt{n}}\right)\right|\varphi(0) > \frac{\varepsilon_n}{|h|}\right) \le$$
$$\le \mathbf{P}_h\left(\sqrt{I_h}-\frac{I_h}{\sqrt{I_0}} > \frac{\varepsilon_n}{M_n\varphi(0)}\right)+$$
$$+\mathbf{P}_h\left(\sqrt{n}\left|\frac{1}{n}\Delta_2\left(\theta_0+\frac{h}{\sqrt{n}}\right)+I_h\right| > \frac{\sqrt{n}\varepsilon_n}{M_n\varphi(0)}\right),$$

что и доказывает требуемое утверждение. ■

Очевидно, последовательность ε_n из условия леммы 2.5 существует, так как $\mathbf{P}_{\theta_0}(|\delta_{n,\theta_0}| > \varepsilon) \to 0$ для всех $\varepsilon > 0$ вследствие (2.3).

Лемма 2.6. *Пусть выполнены условия 1–10. Тогда при $\beta_0 \to 0, \beta_1 \to 0$ НОВ $n^*(\beta_0,\beta_1) \to 0$, а критическая константа $C(\beta_0,\beta_1)$ критерия ϕ_n^*, соответствующего НОВ, отграничена от нуля и единицы при достаточно малых β_0,β_1: $0 < c_0 < C(\beta_0,\beta_1) < c_1 < 1$ с некоторыми c_0 и c_1.*

Д о к а з а т е л ь с т в о. Предположим противное, то есть существует такое n_0, что $n^* < n_0 < \infty$. Тогда легко понять, что

$$\mathcal{R}(d_0|\phi^*_{n_0}) = 0, \mathcal{R}(d_1|\phi^*_{n_0}) = 0. \tag{2.5}$$

Выпишем d-риски критерия $\phi_n(\mathbf{X}) = \mathbb{1}_{\{R(\mathbf{X})\geq C\}}$:

$$\mathcal{R}_0(n, C) = \left(\mathbf{P}\left(R(\mathbf{X}) \leq C\right)\right)^{-1} \int\limits_{\Theta_1} \mathbf{P}_\theta\left(R(\mathbf{X}) \leq C\right) g(\theta) d\theta,$$

$$\mathcal{R}_1(n, C) = \left(\mathbf{P}\left(R(\mathbf{X}) > C\right)\right)^{-1} \int\limits_{\Theta_0} \mathbf{P}_\theta\left(R(\mathbf{X}) > C\right) g(\theta) d\theta.$$

Отсюда и условия 7 следует, что

$$\begin{aligned} \mathbf{P}_\theta\left(R(\mathbf{X}) \leq C\right) &= H_{n_0}(C|\theta) = 0, \text{ для почти всех } \theta \in \Theta_1, \\ \mathbf{P}_\theta\left(R(\mathbf{X}) > C\right) &= 1 - H_{n_0}(C|\theta) = 0, \text{ для почти всех } \theta \in \Theta_0. \end{aligned} \tag{2.6}$$

По лемме 2.4 функция $H_n(C|\theta)$ непрерывна в области $\theta \in [0;1]$ при всех $C \in [0;1]$. Отсюда следует, что равенства (2.6) не могут выполняться одновременно, что приводит к противоречию.

Теперь покажем, что последовательность $C(\beta_0, \beta_1)$ не стремится к нулю или единице. Предположим противное, например, $C \to 1$. Тогда $\mathcal{R}_1(n, C) \to \mathcal{R}_1(n, 1) = G_0$, что противоречит тому, что $\beta_1 \to 0$.

По лемме 2.4 функция $H_n(C|\theta)$ непрерывна в области $C \in [0;1]$. Из этого следует непрерывность функции $\mathcal{R}_0(n_0, C)$ в области $C \in (0;1]$ и функции $\mathcal{R}_1(n_0, C)$ в области $C \in [0;1)$. Из непрерывности этих функций следует, что равенства (2.6) не могут выполняться для любого $C \in (0;1)$. Однако $\mathcal{R}_0(n, 1) = G_1$, $\mathcal{R}_1(n, 0) = G_0$, что противоречит гарантийности критерия. Из последнего же факта выходит, что при $C_n \to 0$ при $n \to \infty$ $\mathcal{R}_1(n, C_n) \to G_0$, а при $C_n \to 1$ $\mathcal{R}_0(n, C_n) \to G_1$. Из этого следует, что у критерия ϕ^*_n критическая константа не может стремиться к нулю или единице, что доказывает утверждение леммы.

■

2.2.2. Асимптотическая формула для необходимого объема выборки

Напомним, что $\delta_{n,\theta} = R(\mathbf{X}) - \Phi(\Lambda_n(\theta) + Q_n(\theta))$.

Теорема 2.1. *Пусть для некоторой числовой последовательности* $\varepsilon_n \to 0$, *такой, что* $\varepsilon_n\sqrt{n} \to \infty$, *выполняется равенство*

$$\lim_{n\to\infty} \mathbf{P}_{\theta_0}(|\delta_{n,\theta_0}| > \varepsilon_n) = 0.$$

При выполнении условий $\mathbb{D}_{\theta_0}$ *для любого* $c \in (\varepsilon_n; 1-\varepsilon_n)$ *справедливы следующие асимптотические (при* $n \to \infty$*) равенства:*

$$\int\limits_{\Theta} H_n(c|\theta)g(\theta)d\theta = G_0 + \frac{g(\theta_0)q_c}{\sqrt{nI(\theta_0)}} + o\left(\frac{1}{\sqrt{n}}\right), \tag{2.7}$$

$$\int\limits_{\Theta_1} H_n(c|\theta)g(\theta)d\theta = \frac{g(\theta_0)}{\sqrt{nI(\theta_0)}}\left(\varphi(q_c) + cq_c + o(1)\right), \tag{2.8}$$

$$\int\limits_{\Theta_0} (1 - H_n(c|\theta))g(\theta)d\theta = \frac{g(\theta_0)}{\sqrt{nI(\theta_0)}}\left(\varphi(q_c) - (1-c)q_c + o(1)\right). \tag{2.9}$$

Доказательство. Докажем первую формулу, остальные можно получить по аналогии. Выберем некоторую последовательность M_n так, чтобы $\Phi(M_n) = 1 - \exp(-n^\alpha)(1 + o(1))$ для некоторого $0 < \alpha < 1/2$. Согласно известной формуле аппроксимации $1 - \Phi(x) = x^{-1}\varphi(x)(1 + o(1))$ при $x \to \infty$ мы можем положить $M_n = Mn^{\alpha/2}$ при некотором достаточно большом $M > 0$. Имея ввиду последовательность ε_n, полученную в конце предыдущего пункта, возьмем последовательность M_n такой, чтобы она удовлетворяла условиям леммы 2.5. При условиях этой теоремы это всегда можно сделать. Разделим область интегрирования Θ на три части: $A_1 = \{\theta : \theta < \theta_0 - n^{-1/2}M_n\}$, $A_2 = \{\theta : \theta_0 - n^{-1/2}M_n < \theta < \theta_0 + n^{-1/2}M_n\}$, $A_3 = \{\theta : \theta > \theta_0 + n^{-1/2}M_n\}$.

К интегралам по множествам A_1 и A_3 можно применить лемму 2.2. На множестве A_1 имеем $\sqrt{n}(\theta_0 - \theta) \geq M_n$. Согласно лемме 2.2

$$\int\limits_{A_1} H_n(c|\theta)g(\theta)d\theta \geq \left(1 - C_aM_n^{-a}\right)\int\limits_{A_1} g(\theta)d\theta = \left(1 - C_aM_n^{-a}\right)G(\theta_0 - n^{-1/2}M_n)$$

для любого $a > 0$, при условии, что $M_n^{-a} < c$. Необходимое a всегда можно подобрать исходя из скорости сходимости ε_n. Так как $H(c|\theta) \le 1$, то можно получить неравенство с другой стороны:

$$\int\limits_{A_1} H_n(c|\theta)g(\theta)d\theta \le G(\theta_0 - n^{-1/2}M_n).$$

Таким образом,

$$\int\limits_{A_1} H_n(c|\theta)g(\theta)d\theta = G(\theta_0 - n^{-1/2}M_n)(1 + O\left(n^{-s}\right)),$$

причем число s можно сделать сколь угодно большим в силу произвольности a. Аналогично оценивается интеграл по множеству A_3, только в этом случае

$$\int\limits_{A_3} H_n(c|\theta)g(\theta)d\theta = O\left(n^{-s}\right).$$

Обратимся к интегралу по множеству A_2. Согласно (2.3) мы имеем

$$\begin{aligned} H_n\left(c|\theta\right) = \mathbf{P}_\theta\left(\Phi(\Lambda_n(\theta) + Q_n(\theta)) < c - \delta_{n,\theta}\right) = \\ = \mathbf{P}_\theta\left(\{\Phi(\Lambda_n(\theta) + Q_n(\theta)) < c - \delta_{n,\theta}\} \cap \{|\delta_{n,\theta}| \le \varepsilon_n\}\right) + \\ + \mathbf{P}_\theta\left(\{\Phi(\Lambda_n(\theta) + Q_n(\theta)) < c - \delta_{n,\theta}\} \cap \{|\delta_{n,\theta}| > \varepsilon_n\}\right). \end{aligned} \tag{2.10}$$

Второе слагаемое в правой части последнего неравенства ограничено сверху величиной $\mathbf{P}_\theta\left(B^c_{n,\theta}\right)$, где $B_{n,\theta} = \{|\delta_{n,\theta}| \le \varepsilon_n\}$. При $\theta \in A_2$ согласно лемме 2.5 эта величина стремится к 0 при $n \to \infty$. Обозначим $z_{n,\theta} = \mathbf{P}_\theta\left(B^c_{n,\theta}\right)$. Далее, вследствие условия $\varepsilon_n < c < 1 - \varepsilon_n$ величина $c - \delta_{n,\theta}$ отграничена от 0 и 1 на множестве $|\delta_{n,\theta}| \le \varepsilon_n$, и мы можем применить функцию Φ^{-1} к обеим частям неравенства под знаком вероятности в первом слагаемом правой части (2.10). Применяя формулу Тейлора к Φ^{-1} получим

$$\begin{aligned} \mathbf{P}_\theta\left(\{\Phi(\Lambda_n(\theta) + Q_n(\theta)) < c - \delta_{n,\theta}\} \cap B_{n,\theta}\right) = \\ \mathbf{P}_\theta\left(\left\{\Lambda_n(\theta) + Q_n(\theta) < \Phi^{-1}(c) + \left(\varphi\left(\Phi^{-1}\left(c + \tilde{\delta}\right)\right)\right)^{-1}\delta_{n,\theta}\right\} \cap B_{n,\theta}\right), \\ \left|\tilde{\delta}\right| \le |\delta_{n,\theta}|. \end{aligned} \tag{2.11}$$

Последнее выражение может быть оценено сверху и снизу, если заменить $\delta_{n,\theta}$ на $-\varepsilon_n$ и ε_n соответственно, а $\tilde{\delta}$ на ε_n (для простоты будем считать, что $c > 0.5$). Для того, чтобы получить дальнейшую оценку сверху, нужно просто избавиться от пересечения с множеством $B_{n,\theta}$. Для оценки снизу нужно избавиться от пересечения с $B_{n,\theta}$, и при этом вычесть $z_{n,\theta}$.

В итоге под знаком вероятности единственной случайной величиной останется $\Lambda_n(\theta)$. Так как распределение $\Lambda_n(\theta)$ по мере $\mathbf{P}_\theta$ асимптотически нормально со средним 0 и дисперсией 1мы можем записать верхнюю оценку для (2.11) как

$$\Phi\left(\Phi^{-1}(c) - \left(\varphi\left(\Phi^{-1}\left(c+\varepsilon_n\right)\right)^{-1}\varepsilon_n - Q_n(\theta)\right)\right) + z_{n,\theta} + v_{n,\theta} =$$
$$= \Phi\left(\Phi^{-1}(c) - Q_n(\theta)\right) - \frac{\varepsilon_n\varphi\left(\Phi^{-1}(c) - Q_n(\theta) - \nu\right)}{\varphi\left(\Phi^{-1}\left(c+\varepsilon_n\right)\right)} + z_{n,\theta} + v_{n,\theta},$$
$$0 \le \nu \le \left(\varphi\left(\Phi^{-1}\left(c+\varepsilon_n\right)\right)\right)^{-1}\varepsilon_n, \quad (2.12)$$

где $v_{n,\theta}$ — разница между вероятностью (2.11) и членом с Φ в левой части (2.12), причем $v_{n,\theta} \to 0, n \to \infty$ при любом фиксированном $\theta \in A_2$. Коэффициент при ε_n в последнем выражении ограничен при фиксированном c для всех $\theta \in A_2$.

Теперь в интеграле сделаем замену $\theta = \theta_0 + h/\sqrt{n}$.

Так как $\varepsilon_n \to 0$, а коэффициент при нем в правой части (2.12) ограничен, соответствующий член стремится к нулю с ростом n независимо от h. Величина $z_{n,\theta_0+h/\sqrt{n}}$ стремится к нулю при каждом фиксированном h на множестве $\theta_0 + h/\sqrt{n} \in A_2$ согласно лемме 2.5. Согласно лемме 2.4 $H_n(c|\theta_0+h/\sqrt{n})$ непрерывна по h, также как и $\Phi(\Phi^{-1}(c) - Q_n(\theta))$. Следовательно, $v_{n,\theta_0+h/\sqrt{n}}$ непрерывно по h, а значит и

$$\sup_{-M_n \le h \le M_n} v_{n,\theta_0+h/\sqrt{n}} \to 0,$$

и интеграл по h от $v_{n,\theta_0+h/\sqrt{n}}$ при $\theta_0+h/\sqrt{n} \in A_2$ также стремится к нулю. Теми же рассуждениями можно показать, что интеграл от $z_{n,\theta_0+h/\sqrt{n}}$ стремится к 0 с ростом n.

Таким образом, после интегрирования (2.12) по $h \in \sqrt{n}(A_2 - \theta_0)$ мы получим

$$\int\limits_{A_2} H_n(c|\theta)g(\theta)d\theta \le$$

$$= \frac{1}{\sqrt{n}}\left[\int\limits_{-M_n}^{M_n} \Phi\left(q_c - h\sqrt{I\left(\theta_0 + \frac{h}{\sqrt{n}}\right)}\right) g\left(\theta_0 + \frac{h}{\sqrt{n}}\right) dh + o(1)\right]. \quad (2.13)$$

Тот же самый результат можно получить после интегрирования нижней границы для (2.11), только с заменой неравенства на обратное (изменится только величина $o(1)$). Поэтому в (2.13) можно заменить знак неравенства на равенство.

Вследствие непрерывности информации $I(\theta)$ и априорной плотности $g(\theta)$ мы можем воспользоваться теоремой о среднем в интеграле в выражении (2.13) и вынести $I(\theta_0)$ и $g(\theta_0)$ за знак интеграла. Обозначим $V_n = M_n\sqrt{I(\theta_0)}$. Имеем

$$\int\limits_{-M_n}^{M_n} \Phi\left(q_c - h\sqrt{I(\theta_0)}\right) g(\theta_0)dh = \frac{g(\theta_0)}{\sqrt{nI(\theta_0)}}\Big[(V_n - q_c)\Phi(q_c - V_n) +$$

$$+ (V_n + q_c)\Phi(q_c + V_n) + \varphi(V_n + q_c) - \varphi(V_n - q_c)\Big].$$

Учитывая, что M_n выбиралось таким образом, чтобы $\Phi(M_n)$ было близко к 1 на порядок экспоненты, мы можем в итоге записать последнее выражение следующим образом ($\varphi(M_n)$, очевидно, также экспонента от некоторой степени n):

$$\int\limits_{-M_n}^{M_n} \Phi\left(q_c - h\sqrt{I(\theta_0)}\right) g(\theta_0)dh = \frac{g(\theta_0)}{\sqrt{nI(\theta_0)}}\left(\sqrt{I(\theta_0)}M_n + q_c\right) + O\left(\exp(-n^{\alpha})\right). \quad (2.14)$$

Вспомним, что интеграл по множеству A_1 равен $G(\theta_0 - n^{-1/2}M_n) +$ $+ O\left(\exp(-n^{\alpha})\right)$. Применяя к нему формулу Тейлора, получим

$$G(\theta_0 - n^{-1/2}M_n) = G(\theta_0) - g(\theta_0)\frac{M_n}{\sqrt{n}} + \frac{1}{2}g'(\theta)\left(\frac{M_n}{\sqrt{n}}\right)^2.$$

Второй член сократится с подобным в (2.14). Имея ввиду то, что остаток должен быть по порядку меньше, чем $n^{-1/2}$, мы должны потребовать, чтобы $\alpha \leq 1/2$, а это всегда возможно. Суммируя полученные результаты, мы получим первую формулу теоремы.

Остальные формулы получаются по аналогии.

■

Теорема 2.2. *Пусть выполнены условия $\mathbb{D}_{\theta_0}$. Пусть $n^* = n^*(\beta_0, \beta_1)$ — НОВ для различения гипотез H_0 и H_1 при ограничениях на d-риски β_0 и β_1. Пусть $\tilde{c}$ и $\tilde{n}$ определяются как c и n, которые являются решениями уравнений*

$$\frac{\varphi(q_c) + cq_c - q_c\beta_0}{\varphi(q_c) + cq_c - q_c(1-\beta_1)} = \frac{\beta_0 G_0}{\beta_1 G_1}, \tag{2.15}$$

$$n = \frac{g(\theta_0)^2(2\varphi(q_c) + 2cq_c - q_c(1+\beta_0-\beta_1))^2}{I(\theta_0)(\beta_0 G_0 + \beta_1 G_1)^2}. \tag{2.16}$$

Если $\beta_0 \to 0$, $\beta_1 \to 0$ так, что $\beta_0/\beta_1 \to K > 0$, то $n^*/\tilde{n} \to 1$.

Доказательство. По лемме 2.6 критическая константа отграничена от нуля и единицы, поэтому условия теоремы 2.1 выполнены.

Выражения (2.15) можно получить, если заменить числители и знаменатели (2.1) и (2.2) на их приближения, полученные в теореме 2.1, приравнять их соответственно к β_0 и β_1, и, наконец, разделить первое полученное уравнение на второе. После этих шагов получим следующее:

$$\frac{\varphi(q_c) + cq_c - q_c\beta_0 + z_{1,n^*}}{\varphi(q_c) + cq_c - q_c(1-\beta_1) + z_{2,n^*}} = \frac{\beta_0 G_0}{\beta_1 G_1}, \tag{2.17}$$

где $z_{i,n^*} \to 0$ при $n^* \to \infty$ (вообще говоря, величины z_{i,n^*} зависят также и от c, причем непрерывным образом согласно непрерывности $H_n(c|\theta)$). Аналогично можно получить выражение (2.16), если вместо деления просуммировать выражения для рисков и привести к

$$n^* = \frac{g(\theta_0)^2(2\varphi(q_c) + 2cq_c - q_c(1+\beta_0-\beta_1) + z_{1,n^*} + z_{2,n^*})^2}{I(\theta_0)(\beta_0 G_0 + \beta_1 G_1)^2}.$$

Пусть $C^* = C^*(\beta_0, \beta_1)$ истинное значение критической константы (оно определяется как решение (2.17)). Обозначим $F_0(c)$ и $F_1(c)$ соответственно числитель и знаменатель левой части (2.15). По лемме 2.6 $n^* \to \infty$, поэтому $z_{i,n^*} \to 0$ при $\beta_0 \to 0, \beta_1 \to 0$. Из этого факта, непрерывности по c функций $F_0(c)$ и $F_1(c)$, и того, что $\beta_0/\beta_1 \to K$, следует, что $|F_i(\tilde{c}) - F_i(C^*)| \to 0, i = 0, 1$, и, более того, $|\tilde{c} - C^*| \to 0$. Исходя из этого факта, легко получить результат теоремы:

$$\frac{n^*}{\tilde{n}} = \frac{F_0(C^*) + F_1(C^*) + z_{0,n^*} + z_{1,n^*}}{F_0(\tilde{c}) + F_1(\tilde{c})} \to 1.$$

■

К сожалению, нельзя показать, что $|n^* - \tilde{n}| \to 0$, поскольку эта величина, по всей видимости, стремится к бесконечности при $\beta_0, \beta_1 \to 0$. Для получения приближения $\tilde{n}$, стремящегося к n^* по абсолютной величине, необходимо уточнить (2.3), добавив хотя бы два члена разложения распределения $R(\mathbf{X})$. Если быть более точным, то при добавке одного члена разложения, по всей видимости, можно показать, что величина $|n^* - \tilde{n}|$ ограничена некоторой константой, а при добавлении двух членов разложения она будет стремиться к нулю.

Глава 3

Приложение к d-гарантийному контролю качества по альтернативному признаку

3.1. Введение

Статистический контроль качества предназначен для выборочного обследования выпускаемой продукции с целью снижения доли некондиционного продукта, получаемого потребителем. Однако существующие рекомендуемые статистические процедуры контроля качества [30], [31] гарантируют не вероятность получения потребителем получения некондиционного продукта, а так называемый риск потребителя — вероятность пропуска на контроле продукции, при условии, что она некондиционная. Таким образом, контроль качества с гарантией характеристики, интересующей потребителя, возможен только в рамках d-апостериорного подхода к данной проблеме.

Неблагополучие в определении гарантируемых характеристик в контроле качества упоминались достаточно давно, отметим хотя бы работу Бернштейна [32]. Наиболее корректная постановка задачи об изменении входного уровня качества таким образом, чтобы гарантировать выходной уровень, содержится в монографии Беляева [33]. По существу, эта задача о построении процедур контроля, гарантирующих d-риск первого рода — вероятность некондиционности продукции, при условии, что она принята на контроле. В рамках общей проблемы проверки гипотез о значении контролируемого «случайного» параметра θ кондиционность продукции обычно формулируется как гипотеза $H_0 : \theta \in \Theta_0$, в то время как недопустимое качество продукции означает справедливость альтернативной гипотезы $H_1 : \theta \in \Theta_1$. Таким образом, при d-апостериорном подходе к проблеме контроля критерий проверки H_0 должен гарантировать заданные ограничения β_0 на условную вероятность события $\theta \in \Theta_1$ при условии что критерий принял гипотезу H_0. Общая теория таких d-гарантийных критериев, обладающих свойством оптимальности (минимизируется d-риск второго рода), разработаны в работе Симушкина [7]. Естественно, эти критерии соответствуют необходимому объему выборки — минимальному числу наблюдений, при котором существует критерий с ограничениями на оба d-риска.

В этой главе рассматривается задача гарантийной проверки гипотез о параметре вероятности успеха в схеме испытаний Бернулли (контроль качества по альтернативному признаку). Сначала уточняются формулы для необходимого объема выборки при ограничениях на вероятности ошибок первого и второго рода. Далее строятся асимптотические разложения d-рисков оптимального критерия. Затем полученные в главе 2 результаты для d-гарантийных критериев применяются для построения процедур контроля качества по альтернативному признаку с ограничениями на d-риски. Необходимые объемы выборки в классическом и апостериорном подходах сравниваются в терминах области безразличия и априорного распределения при одинаковых значениях вероятности ошибок и «соответствующих» d-рисков.

В схеме испытаний Бернулли предполагается, что каждый тестируемый объект имеет некую вероятность p быть некондиционным. Контроля качества выпускаемой продукции сводится к проверке гипотезы $H_0 : p \leq p_0$ при альтернативе $H_1 : p > p_0$. Классический критерий с уровнем значимости α_1 основывается на статистике $T = \sum_{k=1}^{n} X_k$, имеющей биномиальное распределение. Обычно при гарантийной проверке продукции ставятся ограничения α_1 и α_2 на вероятности ошибок первого и второго рода соответственно, и находится такой объем выборки n, чтобы этот критерий выполнял заданные ограничения. Для этого необходимо задать зону безразличия гипотез, таким образом сдвигая альтернативу: $H_1 : p > p_1,\ p_1 > p_0$.

В случаях, когда проводится регулярное тестирование объектов, накапливается большая история испытаний, которую можно использовать для нахождения априорного распределения вероятности брака p. Используя знание априорной вероятности, можно использовать ограничения β_1 и β_2 для d-рисков, и, таким образом, избегать проблемы спецификации зоны безразличия. Поскольку T является достаточной статистикой, она также может быть использована в оптимальном критерии, удовлетворяющем ограничения на d-риски.

3.2. Уточнение асимптотических формул для необходимого объема выборки при ограничениях на вероятности ошибок первого и второго рода

Обозначим $\mathbf{P}_k$ вероятность, вычисленную при истинном значении параметра равным $p_k, k = 0, 1$. Далее,

$$\mu_k = p_k, \quad \sigma_k^2 = p_k(1-p_k), \quad \gamma_{1,k} = \frac{1-2p_k}{\sqrt{p_k\left(1-p_k\right)}}, \quad \gamma_{2,k} = \frac{1-6p_k\left(1-p_k\right)}{p_k\left(1-p_k\right)}.$$

Если индекс у этих величин не указан, значит, вычисление ведется при некотором неизвестном p, если не указано иначе.

При заданных ограничениях α_1 и α_2 на вероятности ошибок первого и второго рода минимальный объем наблюдений определяется посредством задания целочисленной критической константы C_n и объема выборки $n = n(\alpha_1, \alpha_2)$, удовлетворяющих неравенствам

$$\mathbf{P}_0\left(T \geq C_n\right) \leq \alpha_1, \qquad \mathbf{P}_1\left(T < C_n\right) \leq \alpha_2. \tag{3.1}$$

Определенный таким способом минимальный объем наблюдений, гарантирующий заданные ограничения $(\alpha,\ \beta)$ на вероятности ошибочных решений при использовании нерандомизированного критерия $T \geq C$, в дальнейшем будет называться *требуемым объемом наблюдений* (сокращенно – ТОН).

Называть ТОН необходимым объемом выборки некорректно. Дело в том, что в математической статистике *необходимым объемом выборки* (сокращенно – НОВ) называется минимальный объем наблюдений $n^* = n^*(\alpha, \beta)$, при котором *существует* критерий заданной силы (α, β). В данном случае такой критерий является рандомизированным и он основан на статистике $T+U$, где U – случайная величина с равномерным на интервале $(0;1)$ распределением. Распределение статистики $T+U$ принадлежит непрерывному типу и НОВ находится как наименьшее целое n, удовлетворяющее соотношениям

$$\mathbf{P}_0\left(T+U \geq C_n^*\right)=\alpha, \qquad \mathbf{P}_1\left(T+U < C_n^*\right) \leq \beta, \tag{3.2}$$

где C_n^* — критическая константа, соответствующая рандомизированному критерию.

Использование рандомизированного критерия приводит к сокращению объема наблюдений, гарантирующего заданные ограничения α и β на вероятности ошибочных решений. Асимптотическая оценка разности $n(\alpha,\beta)-n^*(\alpha,\beta)$ в терминах сближающихся гипотез $(\Delta=p_1-p_0)\to 0)$– одна из задач, решаемых в данной главе.

3.2.1. Дефект размера нерандомизированного критерия

Рассмотрим задачу нахождения асимптотики вероятности ошибки первого рода нерандомизированного критерия (размера критерия) при $n\to\infty$. Пусть T имеет биномиальное распределение с параметрами (n,p) и $F_n(x),\ x\in\mathbb{R},$ — функция распределения нормированной случайной величины $(T-n\mu)/\sigma\sqrt{n}$, $\mu=p,\ \sigma^2=p(1-p)$. Для аппроксимации F_n воспользуемся разложением Эджворта (смысл символа $\approx$ будет объяснен ниже):

$$F_n(x)\approx \widetilde{F}_n(x)=\Phi(x)+\varphi(x)\Bigg(\frac{\gamma_1}{6\sqrt{n}}(x^2-1)+ \\ +\frac{\gamma_2}{24n}(x^3-3x)+\frac{\gamma_1^2}{72n}(x^5-10x^3+15x)\Bigg), \tag{3.3}$$

где

$$\gamma_1=\frac{1-2p}{\sqrt{p\,(1-p)}}, \qquad \gamma_2=\frac{1-6p\,(1-p)}{p\,(1-p)}\ \ -$$

коэффициенты асимметрии и эксцесса биномиального $(1,p)$ распределения.

Однако применение разложения Эджворта к решетчатым распределениям обладает рядом существенных особенностей, и это в первую очередь касается точности аппроксимации. Так, в статье Сенатова [34] (теорема 5 статьи и утверждения на страницах 21 и 23 статьи (933 и 935 сборника)) устанавливается, что асимптотическая формула

$$F_n(x) = \widetilde{F}_n\left(\frac{x - 0.5 - n\mu}{\sqrt{n}\sigma}\right) + R_n,\ x \in \{0, 1, ..., n\}.$$

имеет остаточный член R_n порядка $O(n^{-3/2})$ только в точках решетки $D_n = \{0, 1, ..., n\}$. Вне этой решетки о точности аппроксимации сказать что-либо не представляется возможным. Это является ключевым фактом последующего исследования нерандомизированного критерия.

Разложение Корниша-Фишера есть корень уравнения $\widetilde{F}_n(x) = 1-\alpha$, то есть разложение Корниша-Фишера имеет отношение только к разложению Эджворта (а не собственно к распределению $F_n(x)$), и разложение Корниша-Фишера приближает корень этого уравнения с остатком r_n порядка $O(n^{-3/2})$:

$$\widetilde{F}_n^{-1}(1-\alpha) = \lambda_\alpha + \frac{\gamma_1}{6\sqrt{n}}\,(\lambda_\alpha^2 - 1) + \frac{\gamma_2}{24\,n}\,(\lambda_\alpha^3 - 3\lambda_\alpha) + \frac{\gamma_1^2}{72\,n}\,(\lambda_\alpha^5 - 4\lambda_\alpha^3 + 10\lambda_\alpha) + r_n, \quad (3.4)$$

где $\lambda_\alpha = \Phi^{-1}(1-\alpha)$.

Итак, критическая константа $C_n(\alpha)$ определяется как наименьшее целое C, удовлетворяющее неравенству $\mathbf{P}_0(T \geq C) \leq \alpha$. Пусть $\widetilde{C}_n(\alpha) = \sqrt{n}\sigma\widetilde{F}_n^{-1}(1 - \alpha) + n\mu$, где $\widetilde{F}_n^{-1}(1-\alpha)$ – разложение Корниша-Фишера (3.4) без остаточного члена, вычисленное при $p = p_0$. Положим $\mathbb{C}_n(\alpha) = \lfloor \widetilde{C}_n(\alpha) + 0.5 \rfloor + 1$, где $\lfloor x \rfloor$ – целая часть x.

Следующая лемма устанавливает асимптотическое разложение для размера критерия $T \geq \mathbb{C}_n(\alpha)$. Все вычисления ведутся при истинном значении параметра распределения $p = p_0$.

Лемма 3.1. *Для размера критерия* $T \geq \mathbb{C}_n(\alpha)$ *справедливо следующее* *асимптотическое разложение:*

$$F_n\left(\frac{\mathbb{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right)=1-\alpha+\varphi(\lambda_\alpha)\Bigg[\frac{1-\{\widetilde{C}_n(\alpha)+0.5\}}{\sqrt{n}}+$$
$$+\frac{1}{n}\left(\frac{\gamma_1^2}{72}(\lambda_\alpha^5-4\lambda_\alpha)-\frac{\gamma_1}{3}\lambda_\alpha-\frac{1}{2}\left(\frac{1-\{\widetilde{C}_n(\alpha)+0.5\}}{\sigma}\right)^2\lambda_\alpha\right)\Bigg]+O(n^{-3/2}), \quad (3.5)$$

где $\{x\}$ *– дробная часть* x*.*

Доказательство. В точке $\mathbb{C}_n(\alpha)$ значение функции распределения статистики T совпадает со значением функции распределения в точке $\lfloor\widetilde{C}_n(\alpha)+0.5\rfloor+0.5$. Кроме того, согласно вышесказанному, в точке $\lfloor\widetilde{C}_n(\alpha)+0.5\rfloor+0.5$ возможно разложение Эджворта с указанной точностью:

$$F_n\left(\frac{\mathbb{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right)=\widetilde{F}_n\left(\frac{\lfloor\widetilde{C}_n(\alpha)+0.5\rfloor+0.5-n\mu}{\sqrt{n}\sigma}\right)+O(n^{-3/2})\,.$$

Для того, чтобы получить утверждение леммы, достаточно разложить

$$\widetilde{F}_n\left(\frac{\lfloor\widetilde{C}_n(\alpha)+0.5\rfloor+0.5-n\mu}{\sqrt{n}\sigma}\right)=\widetilde{F}_n\left(\frac{\widetilde{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}+\frac{1-\{\widetilde{C}_n(\alpha)+0.5\}}{\sqrt{n}\sigma}\right)$$

в ряд Тейлора в окрестности точки $(\widetilde{C}_n(\alpha)-n\mu)/\sqrt{n}\sigma=F_n^{-1}(1-\alpha)$:

$$\widetilde{F}_n\left(\frac{\lfloor\widetilde{C}_n(\alpha)+0.5\rfloor+0.5-n\mu}{\sqrt{n}\sigma}\right)=\widetilde{F}_n(\widetilde{F}_n^{-1}(1-\alpha))+$$
$$\widetilde{F}_n'(\widetilde{F}_n^{-1}(1-\alpha))\left(\frac{1-\{\widetilde{C}_n(\alpha)+0.5\}}{\sqrt{n}\sigma}\right)+$$
$$\frac{1}{2}\widetilde{F}_n''(\widetilde{F}_n^{-1}(1-\alpha))\left(\frac{1-\{\widetilde{C}_n(\alpha)+0.5\}}{\sqrt{n}\sigma}\right)^2. \quad (3.6)$$

Остается разложить $\widetilde{F}_n$ и ее производных в точке $\widetilde{F}_n^{-1}(1-\alpha)$ по степеням $n^{-1/2}$. Для этого в (3.3) положим $x=\widetilde{F}_n^{-1}(1-\alpha)$ и разложим каждый член в окрестности точки $\lambda_\alpha=\Phi^{-1}(1-\alpha)$. В результате получим

$$\widetilde{F}_n\left(\widetilde{F}_n^{-1}(1-\alpha)\right)=1-\alpha+\varphi(\lambda_\alpha)\frac{\gamma_1^2}{72n}(\lambda_\alpha^5-4\lambda_\alpha)+O(n^{-3/2}), \qquad (3.7)$$

$$\widetilde{F}_n'\left(\widetilde{F}_n^{-1}(1-\alpha)\right) = 1 - \frac{\gamma_1}{3\sqrt{n}}\lambda_\alpha + O\left(\frac{1}{n}\right),$$

$$\widetilde{F}_n''\left(\widetilde{F}_n^{-1}(1-\alpha)\right) = \lambda_\alpha + O\left(\frac{1}{\sqrt{n}}\right).$$

После подстановки этих выражений в (3.6) получаем утверждение леммы.

■

Таким образом, если $\mathbb{C}_n(\alpha)$ совпадает с критической константой критерия $C_n(\alpha)$, то (3.5) представляет асимптотическое разложение размера нерандомизированного критерия. Связь между $C_n(\alpha)$ и $\mathbb{C}_n(\alpha)$ устанавливает

Теорема 3.1. *Существует такое $N > 0$, что величина $\mathbb{C}_n(\alpha) = \lfloor \widetilde{C}_n(\alpha) + 0.5 \rfloor + 1$ совпадает с критической константой $C_n(\alpha)$ рассматриваемого нерандомизированного критерия при всех $n \geq N$. Исключение составляет случай*

$$-\frac{\lambda_\alpha^5\,\gamma_{1,0}^2\,\sigma_0}{72\sqrt{n}} - \frac{\lambda_\alpha\,\gamma_{1,0}^2\,\sigma_0\theta^2}{18\sqrt{n}} + \frac{\lambda_\alpha^6\gamma_{1,0}^3\theta^2\,\sigma_0}{216n} + \frac{\lambda_\alpha^{11}\gamma_{1,0}^4\theta^2\,\sigma_0}{288n^{3/2}} < \\ < \lfloor \widetilde{C}_n(\alpha) \rfloor + 0.5 - \widetilde{C}_n(\alpha) < 0, \quad (3.8)$$

при некотором $\theta \in [0;1]$. В этом случае $\mathbb{C}_n(\alpha) = C_n(\alpha) + 1$.

Д о к а з а т е л ь с т в о.

Рассмотрим те случаи, когда $\mathbb{C}_n(\alpha) \neq C_n(\alpha)$. Так как погрешность R_n разложения Корниша-Фишера есть величина порядка $O(n^{-3/2})$,то существует такое N_1, что $|R_{N_1}| < 1$. При $n \geq N_1$ необходимо рассмотреть два случая.

$$1^0.\ \ F_n\left(\frac{\widetilde{C}_n(\alpha) - n\mu}{\sqrt{n}\sigma}\right) < 1-\alpha, \quad \widetilde{C}_n(\alpha) < \lfloor \widetilde{C}_n(\alpha) \rfloor + 0.5, \quad \mathbb{C}_n(\alpha) = C_n(\alpha) - 1.$$

$$2^0.\ \ F_n\left(\frac{\widetilde{C}_n(\alpha) - n\mu}{\sqrt{n}\sigma}\right) > 1-\alpha, \quad \widetilde{C}_n(\alpha) > \lfloor \widetilde{C}_n(\alpha) \rfloor + 0.5, \quad \mathbb{C}_n(\alpha) = C_n(\alpha) + 1. \quad (3.9)$$

В случае 1^0, используя равенство (3.7), получаем

$$F_n\left(\frac{\widetilde{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right)=\widetilde{F}_n\left(\frac{\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-n\mu}{\sqrt{n}\sigma}\right)+O(n^{-3/2})<$$
$$<1-\alpha=\widetilde{F}_n\left(\frac{\widetilde{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right)-\varphi(\lambda_\alpha)\frac{\gamma_1^2}{72n}(\lambda^5-4\lambda)+O(n^{-3/2}), \quad (3.10)$$

Так как $\widetilde{C}_n(\alpha)<\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5$, то данное неравенство не может выполняться (за счет слагаемого $\varphi(\lambda_\alpha)\gamma_1^2(\lambda^5-4\lambda)/72n$) при всех n больше некоторого N_2. Положим $N=\max\{N_1,N_2\}$.

В случае 2^0 имеем неравенство, противоположное (3.8). Для того, чтобы выяснить, при каких условиях на значения $\widetilde{C}_n(\alpha)$ выполняется это неравенство, разложим

$$\widetilde{F}_n\left(\frac{\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-n\mu}{\sqrt{n}\sigma}\right)$$

в ряд Тейлора в точке $\widetilde{F}_n^{-1}(1-\alpha)$. В результате получим неравенство

$$\frac{1}{\sqrt{n}\sigma}\left(\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-\widetilde{C}_n(\alpha)\right)\left(1-2\lambda_\alpha\frac{\gamma_1}{6\sqrt{n}}\right)-$$
$$-\frac{1}{2n\sigma^2}\left(\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-\widetilde{C}_n(\alpha)\right)^2\lambda_\alpha+O(n^{-3/2})>$$
$$>-\frac{\gamma_1^2}{72n}(\lambda_\alpha^5-4\lambda_\alpha)+O(n^{-3/2}).$$

Найдем те $\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-\widetilde{C}_n(\alpha)$, при которых выполняется это неравенство. Для этого выпишем соответствующее квадратное уравнение

$$\frac{1}{\sqrt{n}\sigma}x\left(1-2\lambda_\alpha\frac{\gamma_1}{6\sqrt{n}}\right)-\frac{\lambda_\alpha}{2n\sigma^2}x^2=-\frac{\gamma_1^2}{72n}(\lambda_\alpha^5-4\lambda_\alpha)$$

относительно переменной $x=\lfloor\widetilde{C}_n(\alpha)\rfloor+0.5-\widetilde{C}_n(\alpha)$. Вычисляя дискриминант $\mathcal{D}$ и применяя формулу Тейлора, находим, что

$$\sqrt{\mathcal{D}}=\frac{1}{\sqrt{n}\sigma_0}\sqrt{1-\frac{4\lambda_\alpha\gamma_1}{6\sqrt{n}}+\frac{\lambda_\alpha^6\gamma_1^2}{36n}}=$$
$$=\frac{1}{\sqrt{n}\sigma}\left(1-\frac{2\lambda_\alpha\gamma_1}{6\sqrt{n}}+\frac{\lambda_\alpha^6\gamma_1^2}{72n}-\frac{\lambda_\alpha^2\gamma_1^2\theta^2}{18n}+\frac{\lambda_\alpha^7\gamma_1^3\theta^2}{216n^{3/2}}+\frac{\lambda_\alpha^{12}\gamma_1^4\theta^2}{288n^2}\right)$$

при некотором $\theta \in [0;1]$ Подставляя это значение в формулу для корней уравнения, получим

$$\lfloor\widetilde{C}_n(\alpha)\rfloor + 0.5 - \widetilde{C}_n(\alpha) = -\frac{\lambda_\alpha^5\,\gamma_1^2\,\sigma}{72\sqrt{n}} - \frac{\lambda_\alpha\,\gamma_1^2\,\theta^2\sigma}{18\sqrt{n}} + \frac{\lambda_\alpha^6\gamma_1^3\theta^2\,\sigma}{216n} + \frac{\lambda_\alpha^{11}\gamma_1^4\theta^2\,\sigma}{288n^{3/2}},$$

$$\lfloor\widetilde{C}_n(\alpha)\rfloor + 0.5 - \widetilde{C}_n(\alpha) = 2\sqrt{n}\,\sigma + \frac{2\gamma_1\,\sigma}{3} - \frac{\lambda_\alpha^5\,\gamma_1^2\,\sigma}{72\sqrt{n}} - \frac{\lambda_\alpha\,\gamma_1^2\,\theta^2\sigma}{18\sqrt{n}} + \\ + \frac{\lambda_\alpha^6\gamma_1^3\theta^2\,\sigma}{216n} + \frac{\lambda_\alpha^{11}\gamma_1^4\theta^2\,\sigma}{288n^{3/2}}.$$

Учитывая, что $\lfloor\widetilde{C}_n(\alpha)\rfloor + 0.5 - \widetilde{C}_n(\alpha) < 0$, а второй корень уравнения больше нуля за счет слагаемого $2\sqrt{n}\sigma$, получим неравенство:

$$-\frac{\lambda_\alpha^5\,\gamma_1,^2\,\sigma}{72\sqrt{n}} - \frac{\lambda_\alpha\,\gamma_1^2\,\sigma\theta^2}{18\sqrt{n}} + \frac{\lambda_\alpha^6\gamma_1^3\theta^2\,\sigma}{216n} + \frac{\lambda_\alpha^{11}\gamma_1^4\theta^2\,\sigma}{288n^{3/2}} < \lfloor\widetilde{C}_n(\alpha)\rfloor + 0.5 - \widetilde{C}_n(\alpha) < 0.$$

Последние неравенства есть не что иное, как утверждение теоремы. Значение критической константы $C_n(\alpha)$ следует из условий 2^0 (3.9). Кроме того, видно, что случай 2^0 выполняется при близости $\{\widetilde{C}_n(\alpha)\}$ к 0.5 порядка $n^{-1/2}$. ■

Асимптотическое разложение (3.5) в утверждении леммы 3.1 можно использовать для построения верхней границы для разности между заданным уровнем значимости и размером нерандомизированного критерия при каждом значении n.

Следствие 3.1. *Если $n \geq N$, то при некотором $R > 0$ для дефекта критерия $D(n) = \alpha - P\left(T \geq C_n(\alpha)\right)$ имеет место неравенство*

$$D(n) \leq \varphi(\lambda_\alpha)\left[\frac{\gamma_1^2}{72n}(\lambda_\alpha^5 - 4\lambda_\alpha) + \left(\frac{1}{\sqrt{n}\sigma}\right)\left(1 - \frac{\gamma_1}{3\sqrt{n}}\lambda_\alpha\right) - \frac{1}{2}\left(\frac{1}{\sqrt{n}\sigma}\right)^2\lambda_\alpha\right] + \\ + R\cdot n^{-3/2}. \quad (3.11)$$

Доказательство. Утверждение следствия в случае $\mathbb{C}_n(\alpha) = C_n(\alpha)$ получается из простейшего неравенства $1 - \{\widetilde{C}_n(\alpha) + 0.5\} \leq 1$ в уравнении (3.5). Когда $\mathbb{C}_n(\alpha) = C_n(\alpha) + 1$ (случай 2^0 в доказательстве теоремы), то утверждение следствия вытекает из неравенства

$$1-\alpha \le F_n\left(\frac{C_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right) = \widetilde{F}_n\left(\frac{\lfloor \widetilde{C}_n(\alpha)\rfloor + 0.5 - n\mu}{\sqrt{n}\sigma}\right) + O(n^{-3/2}) <$$

$$< F_n\left(\frac{\widetilde{C}_n(\alpha)-n\mu}{\sqrt{n}\sigma}\right) = 1-\alpha+\varphi(\lambda_\alpha)\frac{\gamma_1^2}{72n}(\lambda_\alpha^5 - 4\lambda_\alpha) + O(n^{-3/2}).$$

■

Полученный результат иллюстрируется графически на Рис. 3.1 при $p_0 = 0.3$ и $\alpha = 0.01$. На оси абсцисс откладывается значение n (которое варьируется от 1 до 500), точками отмечены дефекты размера критерия $D(n)$. Линия представляет собой правую часть (3.11) при $R = 0.25$ (число R выбрано произвольным образом).

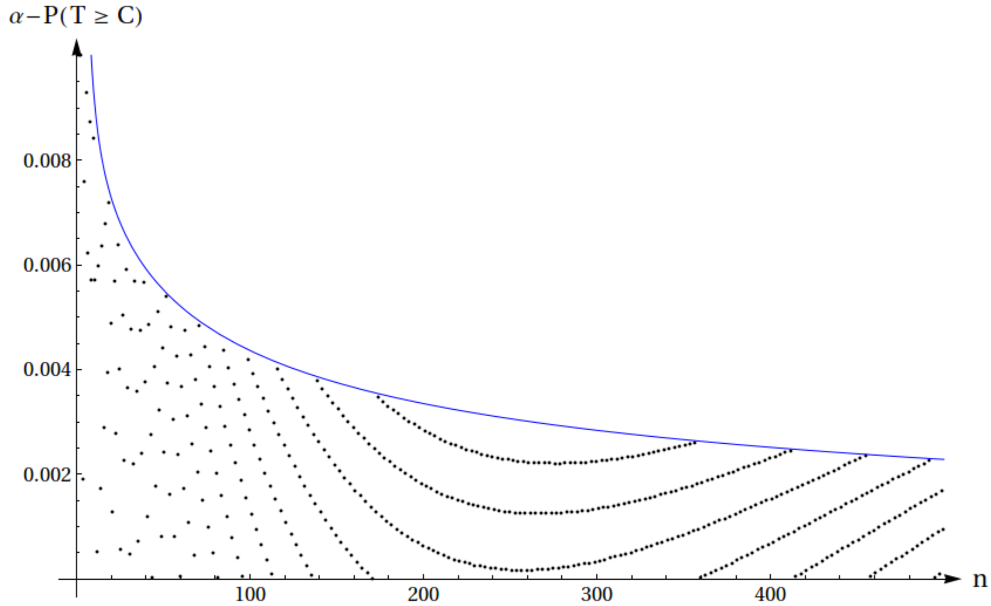

Рис. 3.1. Дефект размера критерия

В данном примере $N^* = 1$, $R < 0.25$. Случай (3.8) ошибки в определении критической константы встречается примерно для 4% значений n.

3.2.2. Асимптотика ТОН с использованием $\mathbb{C}_n(\alpha)$

Результаты предыдущего пункта позволяют построить асимптотику ТОН нерандомизированного критерия, если в качестве критической константы использовать $\mathbb{C}_n(\alpha)$, в которой $\widetilde{C}_n(\alpha) = \sqrt{n}\sigma_0 \widetilde{F}_{0,n}^{-1}(1-\alpha) + n\mu_0$, а $\widetilde{F}_{0,n}^{-1}(1-\alpha)$ – главный член в разложении Корниша-Фишера (3.4). При этом пренебрегается возможная ошибка (3.8) в определении критической константы.

Вероятность ошибки второго рода при критической константе $\mathbb{C}_n(\alpha)$ равна

$$F_{1,n}\left(\frac{\mathbb{C}_n(\alpha) - n\mu_1}{\sqrt{n}\sigma_1}\right) = \widetilde{F}_{1,n}\left(\frac{\lfloor \widetilde{C}_n(\alpha) + 0.5 \rfloor + 0.5 - n\mu_1}{\sqrt{n}\sigma_1}\right) + O(n^{-3/2}).$$

Мы должны найти такое n, чтобы удовлетворялось $F_{1,n}\left(\mathbb{C}_n(\alpha) - n\mu_1/\sqrt{n}\sigma_1\right) \le \beta$. По аналогии с постоянной $\mathbb{C}_n(\alpha)$ определим константу $\mathbb{C}_n(1-\beta) = \lfloor \widetilde{C}_n(1-\beta) - 0.5 \rfloor + 1$, где $\widetilde{C}_n(1-\beta) = \sqrt{n}\sigma_1 \widetilde{F}_{1,n}^{-1}(\beta) + n\mu_1$. Объем наблюдений $n(\alpha,\beta)$ определяется как наименьшее n, удовлетворяющее неравенству (см. [35]) $\mathbb{C}_n(\beta) \le \mathbb{C}_n(\alpha)$. Переходя к непрерывной переменной n, получаем уравнение для асимптотической оценки ТОН $n(\alpha,\beta)$:

$$\widetilde{C}_n(\alpha) - \widetilde{C}_n(1-\beta) + 1 - \{\widetilde{C}_n(\alpha) + 0.5\} + \{\widetilde{C}_n(1-\beta) - 0.5\} = 0, \tag{3.12}$$

где $\{a\}$ означает дробную часть числа a.

Положим $\zeta_n = 1 - \{\widetilde{C}_n(\alpha) + 0.5\} + \{\widetilde{C}_n(1-\beta) - 0.5\}$. При известном значении ζ_n, можно было бы выписать решение (5) из статьи [35], измененное с учетом ζ_n:

$$\widetilde{n} = \frac{a^2}{\Delta^2} + \frac{2\widetilde{b}}{\Delta} + \frac{2ca - \widetilde{b}^2}{a^2} + O(\Delta), \tag{3.13}$$

$$a = \sigma_0\,\lambda_\alpha + \sigma_1\,\lambda_\beta, \qquad b = \frac{1}{6}\left[(\lambda_\alpha^2 - 1)\sigma_0 \cdot \gamma_{1,0} - (\lambda_\beta^2 - 1)\sigma_1 \cdot \gamma_{1,1}\right], \quad \widetilde{b} = b + \zeta_n$$

$$c = \frac{1}{24}\left[(\lambda_\alpha^3 - 3\lambda_\alpha)\sigma_0 \cdot \gamma_{2,0} + (\lambda_\beta^3 - 3\lambda_\beta)\sigma_1 \cdot \gamma_{2,1}\right] -$$

$$- \frac{1}{72}\left[(\lambda_\alpha^5 - 4\lambda_\alpha^3 + 10\lambda_\alpha)\sigma_0 \cdot \gamma_{1,0}^2 + (\lambda_\alpha^5 - 4\lambda_\beta^3 + 10\lambda_\beta)\sigma_1 \cdot \gamma_{1,1}^2\right].$$

Здесь $a, \widetilde{b}, c$ – коэффициенты в уравнении (3.12) перед $\sqrt{n}$, 1, $1/\sqrt{n}$ соответственно.

Основная теорема статьи [35] утверждает, что для критериев, основанных на статистике, имеющей строго монотонную функцию распределения, можно с помощью формулы (3.13) аппроксимировать значение ТОН с точностью до единицы. В нашем случае предположение о монотонности функции распределения не выполняется. В то же время это предположение вводилось для определения функции квантили$F_n^{-1}(\varepsilon)$. В нашем случае $\mathbb{C}_n(\varepsilon)$ однозначно определяет квантиль для любого $0 < \varepsilon < 1$. Однако это не исчерпывает всех проблем дискретного распределения. Во-первых, функцию $\mathbb{C}_n(\varepsilon)$ нельзя считать всюду непрерывной функцией аргумента n (в связи с присутствием n в ζ_n), что затрудняет использование асимптотических методов обращения уравнения относительно n. Во-вторых, уравнение (3.12) может иметь не одно решение, в том смысле что для некоторых n, больших ТОН, условия (3.1) могут не выполняться. Однако имеет место следующее утверждение.

Теорема 3.2. *Существует такая положительная константа Δ^*, что при всех $\Delta < \Delta^*$ значение ТОН $n(\alpha, \beta)$ нерандомизированного критерия не превосходит величины*

$$\frac{a^2}{\Delta^2} + \frac{2(b+2)}{\Delta} + \frac{2ca - (b+2)^2}{a^2} + 1.$$

Доказательство. Пусть $\zeta_n = \zeta$ – фиксированное. Тогда в силу основной теоремы статьи [35] уравнение (3.12) имеет единственный корень $n^*(\zeta)$, полученный с помощью (3.13) с $\zeta_n = \zeta$. При этом $|n^*(\zeta) - n(\zeta)| \leq 1$, когда $\Delta < \Delta^*(\zeta)$, где $n(\zeta)$ – «истинное» значение ТОН при $\zeta_n = \zeta$, а $\Delta^*(\zeta)$ – некоторое число, которое, вообще говоря, зависит от $\zeta \in [0; 2]$. Однако, $[0; 2]$ – компакт, и судя по решению (3.13) $n^*(\zeta)$ является непрерывной функцией ζ. Следовательно, можно найти положительную константу $\Delta^* \leq \Delta^*(\zeta)$ при любом $\zeta \in [0; 2]$.

Далее, $n^*(\zeta)$, будучи многочленом второй степени, достигает своего максимума при

$$\zeta^* = \frac{1}{\Delta a^2} - \frac{b}{a^4}\,.$$

Если учесть, что при уменьшении Δ переменные a, b ведут себя как константы, то $\zeta^* > 2$ при достаточно малых Δ. Так что заменой $n^*(\zeta)$ на меньшее $n^*(2)$ получаем утверждение теоремы. ∎

3.2.3. Аппроксимация НОВ для рандомизированного критерия с использованием разложения Корниша-Фишера

Применим метод статьи Володина [35] для нахождения НОВ рандомизированного критерия. Представим функцию распределения G_n статистики $\mathfrak{T} = (T + U - n\mu)/(\sqrt{n}\sigma)$ в терминах F_n – нормированной функции биномиального распределения:

$$G_n(x) = \int\limits_0^1 \mathbf{P}\left(\frac{T - n\mu}{\sqrt{n}\sigma} < x - \frac{u}{\sqrt{n}\sigma}\,\Big|\, U = u\right) du = \int\limits_0^1 F_n\left(x - \frac{u}{\sqrt{n}\sigma}\right) du.$$

Заменим под интегралом F_n на ее приближение $\widetilde{F}_n$ из (3.3) и воспользуемся разложением Тейлора

$$\widetilde{F}_n\left(x - \frac{u}{\sqrt{n}\sigma}\right) \approx \widetilde{F}_n(x) - \frac{u}{\sqrt{n}\sigma}\widetilde{F}'_n(x) + \frac{1}{2}\left(\frac{u}{\sqrt{n}\sigma}\right)^2 \widetilde{F}''_n(x).$$

Отсюда, после вычисления интеграла получаем, что

$$G_n(x) \approx \widetilde{F}_n(x) - \frac{1}{2\sqrt{n}\sigma}\widetilde{F}'_n(x) + \frac{1}{6n\sigma^2}\widetilde{F}''_n(x).$$

После вычисления производных получаем, что

$$G_n(x) = \Phi(x) - \varphi(x)\bigg(\frac{1}{\sqrt{n}}\Big(\frac{1}{2\sigma} + \frac{\gamma_1}{6}(x^2 - 1)\Big) + \frac{1}{n}\Big(\frac{x}{6\sigma^2} + \frac{\gamma_1}{12\sigma}(x^3 - 3x) + \\ + \frac{\gamma_2}{24}(x^3 - 3x) + \frac{\gamma_1^2}{72}(x^5 - 10x^3 + 15x)\Big)\bigg) + O(n^{-3/2}). \quad (3.14)$$

Для вычисления обратной функции G_n^{-1} теперь уже можно применить разложение Корниша-Фишера. Вывод разложения опирается на разложение Тейлора функции $\Phi(x + (\xi - x))$ относительно точки x, где ξ – нормальная $(0, 1)$ случайная величина, а x – статистика, относительно которой строится разложение (в нашем случае это $\mathfrak{T}$). Т.е. для справедливости разложения требуется близость распределений x и ξ, а используемая в данном случае статистика $\mathfrak{T}$ не является асимптотически нормальной в условиях центральной предельной теоремы. Однако слагаемое $U/(\sqrt{n}\sigma)$ мало при достаточно больших n, а значит, ряд Тейлора для $\Phi(x + (\xi - x))$ будет сходиться при достаточно малых $\Delta = p_1 - p_0$.

Опуская длинные выкладки (более подробно о разложении см., например, [36, п. 6.25]), получим представление

$$G_n^{-1}(\varepsilon) = \lambda_\varepsilon + \frac{1}{\sqrt{n}}\Big(\frac{1}{2\sigma} + \frac{\gamma_1}{6}\,(\lambda_\varepsilon^2 - 1)\Big) + \\ + \frac{1}{n}\Big(\frac{\lambda_\varepsilon}{24\sigma^2} + \frac{\gamma_2}{24}\,(\lambda_\varepsilon^3 - 3\lambda_\varepsilon) + \frac{\gamma_1^2}{72}\,(\lambda_\varepsilon^5 - 4\lambda_\varepsilon^3 + 10\lambda_\varepsilon)\Big) + O(n^{-3/2}). \quad (3.15)$$

От разложения (3.4) последнее выражение отличается лишь на величину

$$\frac{1}{2\sqrt{n}\sigma} + \frac{\lambda_\varepsilon}{24n\sigma^2}.$$

Выпишем теперь аналог уравнения (3.12) для рандомизированного критерия. Здесь удобнее будет применить уравнение, использованное в [35] (уравнение немного изменено в целях использования в терминах статистики $\mathfrak{T}$):

$$n = \left(\frac{\sigma_0\, G_{0,n}^{-1}(1-\alpha) - \sigma_1\, G_{1,n}^{-1}(\beta)}{\mu_1 - \mu_0}\right)^2. \quad (3.16)$$

Подставляя полученные разложения (3.15) для $G_{i,n}^{-1}$ в (3.16), получаем уравнение для определения асимптотики НОВ, аналогичное (4) в [35]. Наконец, используя представление (3.13), получаем асимптотическое разложение для НОВ:

$$n^* = \frac{a^2}{\Delta^2} + \frac{2b}{\Delta} + \frac{2c^*a - b^2}{a^2} + O(\Delta), \quad (3.17)$$

где

$$c^* = c + r, \quad r = \left(\frac{\lambda_\alpha}{24\sigma_0} + \frac{\lambda_\beta}{24\sigma_1}\right).$$

Аналогично работе [37] можно сформулировать следующее утверждение:

Теорема 3.3. *Пусть n – необходимый объем выборки рандомизированного критерия, n^* получено из формулы (3.17). Тогда существует такое число Δ^*, что для всех $\Delta = p_1 - p_0 < \Delta^*$ выполняется $|n - n^*| \le 1$*

Судя по численным результатам, границы, полученные здесь (с использованием теоремы 3.2 и теоремы 3.17), довольно грубо оценивают разницу

$n(\alpha,\beta) - n^*(\alpha,\beta)$. По всей видимости, ζ_n, соответствующее ТОН n, убывает вместе с Δ, но медленнее чем $O(\Delta)$. В численных расчетах для построения адекватной границы $n(\alpha,\beta) - n^*(\alpha,\beta)$ было достаточно применить $\zeta \leq 0.5$.

Приведем таблицы, демонстрирующие эффект рандомизации в редукции ТОН и точность полученной в теореме 3.3 формулы. Слева вверху в ячейке НОВ (рандомизированный критерий), справа вверху ТОН (нерандомизированный критерий), значение внизу получено по формуле (3.17).

Таблица 3.1. *Сравнение рандомизированного и нерандомизированного критериев*

$\alpha = 0.05; \beta = 0.05$					
$p_0 \backslash p_1$		0.2	0.3	0.4	0.5
0,1		134 135 134	41 41 41	20 24 21	13 13 13
0.2	200 202 199		201 204 201	56 60 55	26 28 26
0.3	61 62 61	295 295 295		245 248 245	64 67 64
0.4	31 32 31	82 85 81	359 360 358		267 268 267
0.5	19 19 19	39 40 38	94 96 93	389 391 389	
$p_1 \backslash p_0$	0.1	0.2	0.3	0.4	
$\alpha = 0.05; \beta = 0.01$					

$\alpha = 0.01; \beta = 0.05$					
$p_0 \backslash p_1$		0.2	0.3	0.4	0.5
0.1		192 198 195	57 57 60	29 30 31	18 18 20
0.2	269 272 271		291 291 291	79 79 80	37 37 38
0.3	82 83 83	402 402 402		356 360 356	93 93 93
0.4	41 42 42	111 111 111	489 495 490		389 392 388
0.5	25 25 26	52 52 52	127 130 127	534 535 533	
$p_1 \backslash p_0$	0.1	0.2	0.3	0.4	
$\alpha = 0.01; \beta = 0.01$					

Стоит отметить, что применение разложение Корниша-Фишера остается эффективным даже при различении достаточно разнесенных гипотез (большое значение $\Delta = p_1 - p_0$).

3.3. Асимптотическое разложение d-рисков при контроле по альтернативному признаку

В этой части главы получена формула асимптотического разложения d-риска для схемы испытания Бернулли. Метод получения разложения заключается в применении разложения Эджворта для распределения статистики $\sum_{i=1}^{n} X_i + U$ и в последующем интегрировании по аналогии с получением разложения апостериорного распределения, проведенного в главе 1.

В схеме испытаний Бернулли с неизвестной вероятностью успеха θ рассматривается проблема проверки гипотезы $H_0 : \theta \le \theta_0$ против альтернативы $H_1 : \theta > \theta_1$. Для d-апостериорного подхода к решению этой задачи необходимо задание априорной плотности $g(\theta)$, которая для схемы испытаний Бернулли, как правило, полагается равной плотности бета-распределения с некоторыми неизвестными параметрами a и b: $g(\theta) = (B(a,b))^{-1}\theta^{a-1}(1-\theta)^{b-1}$. Учитывая, что распределение Бернулли удовлетворяет всем условиям теоремы 2.2, то при достаточно малых ограничениях на d-риски β_0 и β_1 можно воспользоваться формулами (2.7) и (2.8) для вывода асимптотик d-рисков. В таком случае d-риски первого и второго рода для критерия $\phi_n(\mathbf{X}) = \mathbb{1}_{\{R(\mathbf{X})>c\}}$ соответственно имеют вид

$$\mathcal{R}(d_0|\phi_n) = \frac{g(\theta_0)}{G_1\sqrt{nI(\theta_0)}} \left(\varphi(q_c) + cq_c + o(1)\right),$$

$$\mathcal{R}(d_1|\phi_n) = \frac{g(\theta_0)}{G_0\sqrt{nI(\theta_0)}} \left(\varphi(q_c) - (1-c)q_c + o(1)\right),$$

Для распределения Бернулли информация Фишера $I(\theta) = (\theta(1-\theta))^{-1}$.

Вывод предыдущих асимптотик заключался в нахождении асимптотического распределения асимптотики апостериорного распределения, и последующего интегрирования полученного результата по θ. Уточнить этот результат, вообще говоря, очень сложно, поскольку разложение (1.4) теоремы 1.1 апостериорного распределения $R(\mathbf{X})$ слишком сложно для вывода разложения $\mathbf{P}_\theta(R(\mathbf{X}) < c)$, которое затем использовалось бы в выражении для d-рисков

(2.1) и (2.2).

В случае схемы испытаний Бернулли статистику $R(\mathbf{X})$ можно [7] заменить статистикой $T = \sum_{i=1}^{n} X_i + U$, где случайная величина $U \sim \mathbb{U}([0;1])$, отвечающая за рандомизацию, не зависит от $\sum_{i=1}^{n} X_i$, так что оптимальный критерий отвергает H_0, если $T \geq C$ для некоторой критической константы C. Таким образом, ограничения для d-рисков примут вид

$$\mathcal{R}(d_0) = \mathbf{P}(\vartheta < \theta_0 | T \geq C) = \frac{\int_0^{\theta_0} \mathbf{P}_\theta(T \geq C) g(\theta) d\theta}{\int_0^1 \mathbf{P}_\theta(T \geq C) g(\theta) d\theta} \leq \beta_0, \tag{3.18}$$

$$\mathcal{R}(d_1) = \mathbf{P}(\vartheta \geq \theta_0 | T < C) = \frac{\int_{\theta_0}^{1} \mathbf{P}_\theta(T < C) g(\theta) d\theta}{\int_0^1 \mathbf{P}_\theta(T < C) g(\theta) d\theta} \leq \beta_1. \tag{3.19}$$

Получить разложение $\mathbf{P}_\theta(T < C)$ можно напрямую, используя разложение Эджворта, примененное ранее в этой главе. Сложность добавляет присутствие параметра θ в разложении (ранее использовались разложения Эджворта только для двух значений параметра p_0 и p_1). Далее получено асимптотическое разложение до порядка $n^{-3/2}$ включительно.

Метод, применяемый далее, заключается в повторении рассуждений главы 3 при выводе асимптотики d-рисков. Так, в интегралах в выражениях (3.18), (3.19) проводится замена $\theta = \theta_0 + I^{-1/2}(\theta_0)h/\sqrt{n}$, после чего интегрирование ведется по области $[-M_n; M_n]$, $M_n \to \infty$. Для того, чтобы интегралы по оставшейся области были малы, доказывается аналог утверждения о больших уклонениях (используется неравенство Хефдинга). В качестве аналога Леммы 2.5 о поведении искомого распределения вблизи θ_0 будет использоваться обычное разложение Тейлора полученного ранее разложения Эджворта в окрестности θ_0.

Однако, как и ранее в этой главе, при выводе асимптотики вместо статистики T удобнее использовать эквивалентную ей нормированную статистику $T_n = (T - n\theta_0)/\sqrt{n\theta_0(1-\theta_0)}$. Причина заключается в том, что критическая константа C в этом случае будет принимать значения в «нормальных» пределах, что позволит использовать ограничения на максимальное абсолютное значение константы C (также, как это было сделано в главе 3 с условием отграничения критической константы c от 0 и 1).

Обозначим далее $\sigma_\theta = \sqrt{\theta(1-\theta)}$ — стандартное отклонение распределения Бернулли с параметром θ. Для сокращения записи также положим $\sigma_0 = \sigma_{\theta_0}$. Аналогично

$$\gamma_{1,\theta} = \frac{1-2\theta}{\sqrt{\theta(1-\theta)}}, \quad \gamma_{2,\theta} = \frac{1-6\theta(1-\theta)}{\theta(1-\theta)}, \quad \gamma_{1,0} = \gamma_{1,\theta_0}, \quad \gamma_{2,0} = \gamma_{2,\theta_0}.$$

Итак, ближайшей целью исследования будет получение асимптотического разложения d-рисков

$$\mathcal{R}_0 = \mathbf{P}\left(\vartheta < \theta_0 | T_n \geq C\right), \quad \mathcal{R}_1 = \mathbf{P}\left(\vartheta \geq \theta_0 | T_n < C\right).$$

Теорема 3.4. *Пусть фиксировано C. Тогда для d-рисков $\mathcal{R}_0$ и $\mathcal{R}_1$ справедливы следующие асимптотические представления:*

$$\mathcal{R}_0 = \frac{1}{\sqrt{n}}\frac{B_1}{G_1} + \frac{1}{n}\frac{B_2G_1 + A_1B_1}{G_1^2} + \\ + \frac{1}{n^{3/2}}\frac{B_3G_1^2 + A_1B_2G_1 + A_2B_1G_1 + A_1^2B_1}{G_1^3} + O\left(n^{-2}\right),$$

$$\mathcal{R}_1 = \frac{1}{\sqrt{n}}\frac{D_1}{G_0} + \frac{1}{n}\frac{D_2G_0 - A_1D_1}{G_0^2} + \\ + \frac{1}{n^{3/2}}\frac{D_3G_0^2 - A_1D_2G_0 - A_2D_1G_0 - A_1^2D_1}{G_0^3} + O\left(n^{-2}\right),$$

где

$$G_0 = \mathbf{G}(\vartheta < \theta_0), \quad G_1 = 1 - G_0, \quad A_1 = g(\theta_0)\sigma_0 C,$$
$$A_2 = g(\theta_0)\sigma_0\left(\frac{C^2+1}{2}\zeta_1\sigma_0 + \frac{\gamma_{1,0}}{2} - \frac{1}{2\sigma_0}\right),$$

$$A_3 = g(\theta_0)\sigma_0\Bigg(\frac{C^3+3C}{6}\zeta_2\sigma_0^2 - 6C\frac{1}{4\sigma_0^2} + 4C + \frac{\gamma_{1,0}}{4\sigma_0} + 3C\frac{\gamma_{1,0}}{\sigma_0^2}m + \\ + (11C+2)\frac{\gamma_{1,0}^2}{12} - C\frac{\zeta_1}{2} + C\zeta_1\gamma_{1,0}\sigma_0\Bigg),$$

$$B_1 = g(\theta_0)\sigma_0 \left(-C(1-\Phi(C)) + \varphi(C)\right),$$

$$B_2 = g(\theta_0)\sigma_0\Bigg(\Big((C^2+1)(1-\Phi(C)) - C\varphi(C)\Big)\frac{\zeta_1\sigma_0}{2} + \\ + \left(C\varphi(C) - 3(1-\Phi(C))\right)\frac{\gamma_{1,0}}{6} + (1-\Phi(C))\frac{1}{2\sigma_0}\Bigg),$$

$$B_3 = g(\theta_0)\sigma_0\Bigg[\left[(C^3+3C)(1-\Phi(C)) - (2+C^2)\varphi(C)\right]\frac{\zeta_2\sigma_0^2}{6} + \\ + \Big[(C^2+1)\varphi(C) - 3(1-\Phi(C))\Big]\frac{\gamma_{1,0}}{12\sigma_0} + \left[\varphi(C) - C(1-\Phi(C))\right]\frac{\gamma_{1,0}}{4\sigma_0^2} + \\ + \left[36C(1-\Phi(C)) - 31\varphi(C)\right]\frac{1}{24\sigma_0^2} - 4(1-\Phi(C)) + 4\varphi(C) - \\ - \Big[(66C+12)(1-\Phi(C)) - (C^4 - 6C^2 - 6C + 99)\varphi(C)\Big]\frac{\gamma_{1,0}^2}{72} + \\ + (C^2+1)\varphi(C)\frac{\gamma_{2,0}}{24} - \left[6C(1-\Phi(C)) - 5\varphi(C)\right]\frac{\zeta_1\sigma_0\gamma_{1,0}}{6} + \\ + \left[C(1-\Phi(C)) - \varphi(C)\right]\frac{\zeta_1}{2}\Bigg],$$

$$D_1 = g(\theta_0)\sigma_0 \left(C\Phi(C) + \varphi(C)\right),$$

$$D_2 = g(\theta_0)\sigma_0\Bigg(\left[(C^2+1)\Phi(C) + \varphi(C)\right]\frac{\zeta_1\sigma_0}{2} + \\ + \left[3\Phi(C) + C\varphi(C)\right]\frac{\gamma_{1,0}}{6} - \Phi(C)\frac{1}{2\sigma_0}\Bigg),$$

$$D_3 = g(\theta_0)\sigma_0\Bigg[\Big((C^3+3C)\Phi(C) - (2+C^2)\varphi(C)\Big)\frac{\zeta_2\sigma_0^2}{6} + \\ - \left[36C\Phi(C) - 31\varphi(C)\right]\frac{1}{24\sigma_0^2} + 4C\Phi(C) + 4\varphi(C) + \\ + \Big[3C\Phi(C) + (C^2+1)\varphi(C)\Big]\frac{\gamma_{1,0}}{12\sigma_0} + \Big[C\Phi(C) + \varphi(C)\Big]\frac{\gamma_{1,0}}{4\sigma_0^2} + \\ + \Big((66C+12)\Phi(C) + (C^4 - 6C^2 - 6C + 99)\varphi(C)\Big)\frac{\gamma_{1,0}^2}{72} + \\ + (C^2+1)\varphi(C)\frac{\gamma_{2,0}}{24} + \left[6C\Phi(C) + 5\varphi(C)\right]\frac{\zeta_1\sigma_0\gamma_{1,0}}{6} - \left[C\Phi(C) + \varphi(C)\right]\frac{\zeta_1}{2}\Bigg],$$

$$\zeta_1 = \pi_1(\theta_0), \quad \zeta_2 = \pi_2(\theta_0) + \pi_1^2(\theta_0), \quad \zeta_3 = \pi_3(\theta_0) + 3\pi_2(\theta_0)\pi_1(\theta_0) + \pi_1^3(\theta_0),$$

$$\pi_k(\theta) = \frac{d^k}{d\theta^k} \ln g(\theta), \quad \pi_1(\theta) = \frac{a-1}{\theta} - \frac{b-1}{1-\theta},$$
$$\pi_2(\theta) = -\frac{a-1}{\theta^2} - \frac{b-1}{(1-\theta)^2}, \quad \pi_3(\theta) = \frac{2(a-1)}{\theta^3} - \frac{2(b-1)}{(1-\theta)^3}.$$

Доказательство.

Начнем с разложения d-риска первого рода. Так,

$$\mathcal{R}_0 = \frac{\int_0^{\theta_0} \mathbf{P}_\theta(T_n \geq C) g(\theta) d\theta}{\int_0^1 \mathbf{P}_\theta(T_n \geq C) g(\theta) d\theta}. \tag{3.20}$$

Возьмем некоторые числа $M > 0$ и $0 < \alpha < 1/2$ и положим $M_n = Mn^\alpha$.

Следуя доказательству теоремы 2.1, разделим отрезок $[0;1]$ на множества $\mathcal{A}_1 = [0; \theta_0 - M_n\sigma_0/\sqrt{n}]$, $\mathcal{A}_2 = [\theta_0 - M_n\sigma_0/\sqrt{n}; \theta_0 + M_n\sigma_0/\sqrt{n}]$, $\mathcal{A}_3 = [\theta_0 + M_n\sigma_0/\sqrt{n}; 1]$. Интеграл в знаменателе (3.20), таким образом, можно представить в виде суммы трех интегралов соответственно по областям $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$.

Интеграл по множеству $\mathcal{A}_1$ пренебрежимо мал. Это доказывается с помощью неравенства Хефдинга для схемы испытаний Бернулли:

$$\mathbf{P}_\theta\left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C\right) = \mathbf{P}_\theta\left(\frac{\sum X_i - n\theta}{\sqrt{n}\sigma_\theta} \geq \frac{\sqrt{n}\sigma_0 C + n(\theta_0 - \theta) - 1}{\sqrt{n}\sigma_\theta}\right) \leq$$
$$\leq \exp\left(-2C^2\frac{\sigma_0^2}{\sigma_\theta^2} - 4C\frac{\sigma_0}{\sigma_\theta^2}\sqrt{n}(\theta_0 - \theta) - \frac{2}{\sigma_\theta^2}n(\theta_0-\theta)^2 + O\left(\frac{1}{\sqrt{n}}\right)\right). \tag{3.21}$$

Дисперсия σ_θ ограничена сверху величиной $1/2$; критическая константа C фиксирована по условию теоремы, а отклонение параметра $\sqrt{n}(\theta_0 - \theta)$ ограничено снизу числом M_n на множестве $\mathcal{A}_1$. Поэтому последнее выражение асимптотически эквивалентно $\exp(-n^{2\alpha})$, так что при некотором $\gamma > 0$

$$\int\limits_{\mathcal{A}_1} \mathbf{P}_\theta(T_n \geq C) g(\theta) d\theta \leq \exp(-n^\gamma).$$

Интеграл по множеству $\mathcal{A}_2$ изучается с помощью получения разложений подинтегральных выражений и последующим интегрированием.

Выпишем разложение Эджворта для распределения T_n с членами до порядка $n^{-3/2}$ включительно (это разложение было получено раннее, см. (3.14)):

$$\mathbf{P}_\theta\left(\frac{T-n\theta_0}{\sqrt{n}\sigma_0}<C\right)=\mathbf{P}_\theta\left(\frac{T-n\theta}{\sqrt{n}\sigma_\theta}<\frac{\sqrt{n}\sigma_0 C+n(\theta_0-\theta)}{\sqrt{n}\sigma_\theta}\right)=$$
$$=\Phi(C_{n,\theta})-\varphi(C_{n,\theta})\Bigg(\frac{\gamma_{1,\theta}}{6\sqrt{n}}\left(C_{n,\theta}^2-1\right)+\frac{1}{2\sigma_\theta\sqrt{n}}+$$
$$+\frac{C_{n,\theta}}{6\sigma_\theta^2}+\frac{\gamma_{1,\theta}}{12\sigma_\theta}\left(C_{n,\theta}^3-3C_{n,\theta}\right)+\frac{\gamma_{2,\theta}}{24n}\left(C_{n,\theta}^3-3C_{n,\theta}\right)+$$
$$+\frac{\gamma_{1,\theta}^2}{72n}\left(C_{n,\theta}^5-10C_{n,\theta}^3+15C_{n,\theta}\right)\Bigg)+n^{-3/2}H_n(C,\theta)+\omega_{n,\theta}, \quad (3.22)$$

где $H_n(C,\theta)=O(1)$ — конкретный коэффициент перед $n^{-3/2}$ разложении Эджворта, $\omega_{n,\theta}=O(n^{-2})$ при $\sqrt{n}\sigma_0 C+n\mu_0\in\{-0.5,0.5,\ldots,n-0.5\}$ — остаток, а

$$C_{n,\theta}=\frac{\sqrt{n}\sigma_0 C+n(\theta_0-\theta)}{\sqrt{n}\sigma_\theta}.$$

Вследствие компактности множества $\mathcal{A}_2$ величина $\sup_{\theta\in\mathcal{A}_2}|\omega_{n,\theta}|=O(n^{-2})$. Поэтому $\int_{\mathcal{A}_2}|\omega_{n,\theta}|g(\theta)d\theta=O(n^{-2})$. Таким образом, далее будем работать только с основной частью разложения (3.22).

Сделаем замену $h=\sqrt{n}(\theta-\theta_0)\sigma_0^{-1}$ в интеграле по $\mathcal{A}_2$. Множитель, который появится перед интегралом, равен $\sigma_0/\sqrt{n}$.

Смысл во введении слагаемого $H_n(C,\theta)$ заключается в том, что неясно, как будет себя вести остаток разложения Эджворта $\omega_{n,\theta}$ после локальной замены параметра θ. Однако известен конкретный вид коэффициента $H_n(C,\theta)$, который после замены переменной не изменит своего поведения. Мы не будем здесь выписывать $H_n(C,\theta)$, и сошлемся на монографию [36, с. 230], где выписан конкретный вид очередного члена разложения Эджворта. Формально говоря, можно обойтись без введения $H_n(C,\theta)$, если контролировать порядок $\omega_{n,\theta}$ после перехода к h, аккуратно производя выкладки разложения Эджворта, однако этот путь слишком трудоемкий. В то же время для биномиального распределения существуют члены разложения Эджворта до любого порядка, и здесь мы этим воспользовались.

Для приведения к удобному для интегрирования виду необходимо преобразовать разложение (3.22) так, чтобы зависимость от параметра h была полиномиальная или экспоненциальная. Этого можно добиться, применив разложение Тейлора в окрестности точки θ_0 для всех величин, зависящих от θ:

$$C_{n,\theta} = (C-h)\left(1 - \frac{1-2\theta_0}{2\sigma_0}\frac{h}{\sqrt{n}} + \frac{3-8\sigma_0^2}{4\sigma_0^2}\frac{h^2}{n}\right) + n^{-3/2}\omega_{n,\theta,0}.$$

$$C_{n,\theta}^2 = (C-h)^2\left(1 - \frac{1-2\theta_0}{\sigma_0}\frac{h}{\sqrt{n}}\right) + O\left(n^{-1}\right).$$

$$\begin{aligned}\Phi(C_{n,\theta}) = \Phi(C-h) + \varphi(C-h)\Bigg(C_{n,\theta} - (C-h) -\\ - \frac{1}{2}(C-h)(C_{n,\theta}-(C-h))^2\Bigg) + n^{-3/2}\omega_{n,\theta,1} = \Phi(C-h) +\\ + \varphi(C-h)\left[\left(-\frac{1-2\theta_0}{2\sigma_0}\frac{h}{\sqrt{n}} + \frac{3-8\sigma_0^2}{4\sigma_0^2}\frac{h^2}{n}\right)(C-h) -\right.\\ \left. - \frac{h^2}{2n}(C-h)^3\left(\frac{1-2\theta_0}{2\sigma_0}\right)^2\right] + n^{-3/2}\omega_{n,\theta,1}.\end{aligned}$$

$$\varphi(C_{n,\theta}) = \varphi(C-h)\left(1 + (C-h)\frac{1-2\theta_0}{2\sigma_0}\frac{h}{\sqrt{n}}\right) + n^{-1}\omega_{n,\theta,2}.$$

$$\gamma_{1,\theta} = \gamma_{1,0} - \frac{h}{4\sqrt{n}\sigma_{\theta_0}^2} + n^{-1}\omega_{n,\theta,3}.$$

$$\frac{1}{\sigma_\theta} = \frac{1}{\sigma_0} - \frac{\gamma_{1,0}}{2\sigma_0^2}\frac{h}{\sqrt{n}} + n^{-1}\omega_{n,\theta,4}.$$

$$\gamma_{2,\theta} = \gamma_{2,\theta_0} + n^{-1/2}\omega_{n,\theta,5}.$$

Остатки $\omega_{n,\theta,k}, k = 0,\ldots,5$ имеют форму остатков Лагранжа. В таком случае коэффициент при остатке $\omega_{n,\theta,0} = O(h^4)$, коэффициенты $\omega_{n,\theta,1}$ и $\omega_{n,\theta,2}$ будут иметь вид полиномов от h, умноженных на $\varphi(C-h+w(h))$ или $\varphi(C-h)$, где $w(h) = \lambda(C_{n,\theta}-C+h), 0 < \lambda < 1$; аналогично коэффициенты $\omega_{n,\theta,3} = O(h^2)$, $\omega_{n,\theta,5} = O(h^2)$, $\omega_{n,\theta,5} = O(h)$. Остаток разложения $C_{n,\theta}^2$ выписан как $O(n^{-1})$, так как он является квадратичным полиномом от w_{n,θ_0}, и поэтому не представляет самостоятельного интереса.

Подставим полученные выше выражения в (3.22). В результате получим

$$\mathbf{P}_\theta\left(\frac{T-n\theta_0}{\sqrt{n}\sigma_{\theta_0}}<C\right)=\Phi(C-h)+\varphi(C-h)\left(\frac{1}{\sqrt{n}}V_1(h)+\frac{1}{n}V_2(h)\right)+ \\ +n^{-3/2}\left[\omega_n(h)+H_n(C,\theta)\right], \quad (3.23)$$

$$V_1(h)=-\frac{\gamma_{1,0}}{2}h(C-h)-\frac{\gamma_{1,0}}{6}\left((C-h)^2-1\right)-\frac{1}{2\sigma_0},$$

$$V_2(h)=\frac{(3-8\sigma_0^2)}{4\sigma_0^2}(C-h)h^2+\frac{1}{24\sigma_0^2}\left[h((C-h)^2-1)-4(C-h)\right]- \\ -\frac{\gamma_{2,0}}{24}\left[(C-h)^3-3(C-h)\right]-\frac{\gamma_{1,0}}{12\sigma_0}\left[(C-h)^3-(C-h)+3(C-h)h\right]+ \\ +\frac{\gamma_{1,0}}{4\sigma_0^2}-\frac{\gamma_{1,0}^2}{72}\big[(C-h)^5-10(C-h)^3+15(C-h)- \\ -6(C-h)h-12(C-h)^2h+6(C-h)^3h+9(C-h)^3h^2\big].$$

Коэффициент $\omega_n(h)$ является комбинацией коэффициентов $\omega_{n,\theta,k}, k=0,\dots,4$. Точнее, это сумма $\varphi(C-h+w(h))$ и $\varphi(C-h)$, умноженных на некие полиномы от h.

Ясно, что

$$\mathbf{P}_\theta\left(\frac{T-n\theta_0}{\sqrt{n}\sigma_{\theta_0}}\geq C\right)=1-\Phi(C-h)-\varphi(C-h)\left(\frac{1}{\sqrt{n}}V_1(h)+\frac{1}{n}V_2(h)\right)- \\ -n^{-3/2}\left[\omega_n(h)+H_n(C,\theta)\right], \quad (3.24)$$

Нам необходимо показать, что величина $\int_{\mathcal{A}_2}\omega_n(h)g(\theta_0+n^{-1/2}h\sigma_0)dh$ ограничена некоторой константой для всех n. Чтобы избавиться от априорной плотности g в интеграле, можно произвести разложение Тейлора нужного порядка относительно θ_0, таким образом получив полином от h. Далее, часть $\omega_n(h)$, соответствующая полиному от h, умноженному на $\varphi(C-h)$, ограничена вследствие существования интеграла $\int_{-\infty}^{\infty}h^k\varphi(C-h)dh$ для всех $k\geq 0$. Что касается слагаемого с $\varphi(C-h+w(h))$, то можно напрямую представить $w(h)=\lambda(C-h)\sigma_0\sigma_\theta^{-1}$. На компакте $\mathcal{A}_2$ величина σ_θ ограничена, поэтому $w(h)=O(h)$, и интеграл от полинома, умноженного на $\varphi(C-h+w(h))$ сводится к интегралу от нормальной плотности. Таким образом,

$$\sup_n \int\limits_{\mathcal{A}_2} \omega_n(h) g\left(\theta_0 + \frac{h\sigma_0}{\sqrt{n}}\right) dh < \infty.$$

Коэффициент $H_n(C,\theta)$ зависит от θ через первые пять полуинвариантов распределения Бернулли. Они же ограничены сверху и снизу на множестве $\mathcal{A}_2$ для любого n. Отсюда

$$\sup_n \int\limits_{-M_n}^{M_n} H_n\left(C, \theta_0 + \frac{\sigma_0 h}{\sqrt{n}}\right) g\left(\theta_0 + \frac{\sigma_0 h}{\sqrt{n}}\right) dh < \infty.$$

Далее необходимо проинтегрировать по h главную часть выражение (3.24), умноженное на $g(\theta_0 + \sigma_0 h/\sqrt{n})$. Как один из вариантов, можно сделать разложение Тейлора в окрестности точки θ_0 для функции g. Недостатком является то, что для экспоненциальных семейств (в частности, для используемого здесь бета-распределения) производные функций плотностей могут быть неудобны для вычисления. Однако производные логарифма легко выводимы, и они могут быть использованы при получении разложения $\exp(\ln g(\theta))$. С точки зрения изучения асимптотик оба подхода эквивалентны, однако мы пойдем путем, которым шли в Главе 1 и найдем разложение с применением производных логарифма $g(\theta)$.

Выпишем разложение априорной плотности g (при получении порядка остатка используется ограниченность производных логарифма при $h \in [-M_n; M_n]$):

$$\ln g\left(\theta_0 + \frac{\sigma_0 h}{\sqrt{n}}\right) = \pi_0(\theta_0) + \pi_1(\theta_0)\frac{\sigma_0 h}{\sqrt{n}} + \pi_2(\theta_0)\frac{\sigma_0^2 h^2}{2n} + \frac{\sigma_0^3 h^3}{6n^{3/2}}\pi_3\left(\theta_0\right) + \\ + \frac{\sigma_0^4 h^4}{24 n^2}\pi_4\left(\theta_0 + \lambda_1 \frac{\sigma_0 h}{\sqrt{n}}\right), \pi_k(\theta) = \frac{d^k}{d\theta^k}\ln g(\theta), 0 < \lambda_1 < 1,$$

$$\exp\left(\ln g\left(\theta_0+\frac{\sigma_0 h}{\sqrt{n}}\right)\right)=g(\theta_0)\Bigg(1+\pi_1(\theta_0)\frac{\sigma_0 h}{\sqrt{n}}+\left(\pi_2(\theta_0)+\pi_1^2(\theta_0)\right)\frac{\sigma_0^2h^2}{2n}+$$
$$+\left(\pi_3(\theta_0)+3\pi_2(\theta_0)\pi_1(\theta_0)+\pi_1^3(\theta_0)\right)\frac{\sigma_0^3h^3}{6n^{3/2}}\Bigg)+O\left(h^4n^{-2}\right). \quad (3.25)$$

Видно, что коэффициенты перед h^k в последнем разложении есть величины ζ_k из условия теоремы.

Константу α в степени $M_n = Mn^\alpha$ можно взять так, чтобы $n^{-2}M_n^4$ убывало не медленнее чем $n^{-3/2}$. Имеем

$$\int\limits_{-M_n}^{M_n}(1-\Phi(C-h))g\left(\theta_0+\frac{\sigma_0 h}{\sqrt{n}}\right)dh=$$
$$=\int\limits_{-M_n}^{M_n}(1-\Phi(C-h))g(\theta_0)\left[1+\zeta_1\frac{\sigma_0 h}{\sqrt{n}}+\zeta_2\frac{\sigma_0^2h^2}{2n}+\zeta_3\frac{\sigma_0^3h^3}{6n^{3/2}}+O\left(h^4n^{-2}\right)\right]dh=$$
$$=g(\theta_0)\Bigg[M_n-C+\frac{M_n^2-C^2-1}{2n^{1/2}}\sigma_0\zeta_1+\frac{M_n^3-C^3-3C}{6n}\sigma_0^2\zeta_2+$$
$$+\frac{M_n^4}{24n^{3/2}}\sigma_0^3\zeta_3\Bigg]+O\left(n^{-3/2}\right)+O\left(e^{-M_n^2}\right),$$

где величина $O(n^{-3/2})$ получена после интегрирования члена $O(h^4n^{-2})$, а экспоненциальный остаток получен из выражений $\Phi(-M_n) = O(M_n^{-1}e^{-M_n^2})$ и $\varphi(M_n) = O(e^{-M_n^2})$.

Далее необходимо найти интеграл от основной части разложения (3.24), умноженного на разложение (3.25) по множеству $h \in [-M_n; M_n]$. Легко видеть, что можно заменить область интегрирования на $h \in \mathbb{R}$, так как общее значение интеграла изменится на экспоненциально (по n) малую величину. Прямое вычисление в таком случае дает (для сокращения записи $\theta(h) = \sqrt{n}\sigma_0(\theta - \theta_0)$):

$$\int\limits_{-M_n}^{M_n} \mathbf{P}_{\theta(h)}\left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C\right) g(\theta(h))dh =$$

$$= g(\theta_0)\Bigg[M_n - C + \frac{1}{\sqrt{n}}\left(\frac{M_n^2 - C^2 - 1}{2}\zeta_1\sigma_0 - \frac{\gamma_{1,0}}{2} + \frac{1}{2\sigma_0}\right) +$$

$$+ \frac{1}{n}\Bigg(\frac{M_n^3 - C^3 - 3C}{6}\zeta_2\sigma_0^2 + 3C\frac{1}{4\sigma_0^2} - 4C - \frac{\gamma_{1,0}}{4\sigma_0} - 3C\frac{\gamma_{1,0}}{\sigma_0^2}m -$$

$$- (11C + 2)\frac{\gamma_{1,0}^2}{12} + C\frac{\zeta_1}{2} - C\zeta_1\gamma_{1,0}\sigma_0\Bigg) + \frac{M_n^4\sigma_0^3}{24n^{3/2}}\zeta_3\Bigg] + O\left(n^{-3/2}\right). \quad (3.26)$$

Перейдем к асимптотике интеграла по множеству $\mathcal{A}_3$. В этом случае из (3.21) следует, что

$$\mathbf{P}_\theta\left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C\right) \geq 1 - \exp\left(-M_n\right).$$

Поэтому

$$\int\limits_{\mathcal{A}_3} \mathbf{P}_\theta\left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C\right) g(\theta)d\theta = 1 - \mathbf{G}\left(\theta_0 + \frac{M_n\sigma_0}{\sqrt{n}}\right) + O\left(e^{-M_n}\right).$$

В последнем выражении разложим функцию $\mathbf{G}$ в ряд Тейлора относительно точки θ_0. Получим

$$\mathbf{G}\left(\theta_0 + \frac{M_n\sigma_0}{\sqrt{n}}\right) = G_0 + \frac{\sigma_0}{n^{1/2}}\int\limits_0^{M_n} g\left(\theta_0 + \frac{h\sigma_0}{\sqrt{n}}\right)dh =$$

$$= G_0 + g(\theta_0)\frac{\sigma_0}{n^{1/2}}\left[M_n + \zeta_1\frac{M_n^2\sigma_0}{2n^{1/2}} + \zeta_2\frac{M_n^3\sigma_0^2}{6n} + \zeta_3\frac{M_n^4\sigma_0^3}{24n^{3/2}}\right] + O(n^{-2}) \quad (3.27)$$

Остаток (3.27) является величиной порядка не менее n^{-2} вследствие интегрирования величины $O(h^4 n^{-2})$ и умножения на $n^{-1/2}$.

Остается объединить (3.26) и (3.27). Для этого необходимо умножить выражение (3.26) на $\sigma_0/\sqrt{n}$ в соответствии с произведенной заменой $\theta = \theta(h)$. Легко видеть, что члены со степенями M_n сократятся между собой, таким образом оставляя лишь члены порядка $n^{-1/2}$. Окончательный вид разложение интеграла в знаменателе (3.20) следующий:

$$\int\limits_{0}^{1} \mathbf{P}_\theta \left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C \right) g(\theta) d\theta = G_1 - \frac{A_1}{n^{1/2}} - \frac{A_2}{n} - \frac{A_3}{n^{3/2}} + O\left(n^{-2}\right).$$

Числитель (3.20) отличается от знаменателя тем, что интегрирование ведется по множеству $h \in [-\infty; 0]$. в таком случае не возникает слагаемых с M_n и не появляется слагаемого с априорным распределением $\mathbf{G}$. Очевидно, рассуждения о порядке остатков остаются теми же. Проделав все необходимые вычисления, можно получить

$$\int\limits_{0}^{\theta_0} \mathbf{P}_\theta \left(\frac{T - n\theta_0}{\sqrt{n}\sigma_0} \geq C \right) g(\theta) d\theta = \frac{B_1}{n^{1/2}} + \frac{B_2}{n} + \frac{B_3}{n^{3/2}} + O\left(n^{-2}\right).$$

Чтобы получить разложение d-риска $\mathcal{R}_0$, представленного формулой (3.20), необходимо разложить в ряд Тейлора дробь следующего вида:

$$\frac{b}{G_1 + a} = \frac{b}{G_1} \sum_{k=0}^{\infty} \left(-\frac{a}{G_1} \right)^k. \tag{3.28}$$

Полагая $a = -n^{-1/2}A_1 - n^{-1}A_2 - n^{-3/2}A_3$, $b = n^{-1/2}B_1 + n^{-1}B_2 + n^{-3/2}B_3$, получим асимптотическое разложение для d-риска первого рода $\mathcal{R}_0$ по степеням $n^{-1/2}$:

$$\mathcal{R}_0 = \frac{1}{\sqrt{n}} \frac{B_1}{G_1} + \frac{1}{n} \frac{B_2 G_1 + A_1 B_1}{G_1^2} + \\ + \frac{1}{n^{3/2}} \frac{B_3 G_1^2 + A_1 B_2 G_1 + A_2 B_1 G_1 + A_1^2 B_1}{G_1^3} + O\left(n^{-2}\right).$$

Аналогично получается выражение для d-риска второго рода

$$\mathcal{R}_1 = \frac{\int_{\theta_0}^{1} \mathbf{P}_\theta (T_n < C) g(\theta) d\theta}{\int_{0}^{1} \mathbf{P}_\theta (T_n < C) g(\theta) d\theta}. \tag{3.29}$$

Разложенный предложенным методом числитель (3.29) имеет вид

$$\int\limits_{\theta_0}^{1} \mathbf{P}_\theta \left(T_n < C\right) g(\theta) d\theta = \frac{D_1}{n^{1/2}} + \frac{D_2}{n} + \frac{D_3}{n^{3/2}} + O\left(n^{-2}\right),$$

а знаменатель вследствие $\mathbf{P}_\theta(T_n < C) = 1 - \mathbf{P}_\theta(T_n \geq C)$ примет вид

$$\int\limits_0^1 \mathbf{P}_\theta\left(T_n < C\right) g(\theta) d\theta = G_0 + \frac{A_1}{n^{1/2}} + \frac{A_2}{n} + \frac{A_3}{n^{3/2}} + O\left(n^{-2}\right).$$

Применяя тот же самый метод разложения дроби (3.28), получим результат теоремы для $\mathcal{R}_1$.

■

Теоретически результат предыдущей теоремы можно использовать для получения асимптотики необходимого объема выборки при стремящихся к нулю ограничениях на d-риски. Однако зависимость коэффициентов перед степенями $n^{-1/2}$ от C весьма сложная, что усложняет задачу.

3.4. Сравнение классического и апостериорного подходов к проблеме контроля качества по альтернативному признаку

В этой части главы будет применена формула (2.16) для НОВ при ограничениях на d-риски, полученная в главе 2. Она будет сравнена с асимптотической формулой (3.17), полученной в текущей главе в смысле зависимости величины области безразличия $p_1 - p_0$ от априорного распределения при равных ограничениях на вероятности ошибок и d-риски и равных необходимых объемах выборки. Приводятся таблицы и графики, демонстрирующие эту зависимость.

Наиболее характерным случаем при контроле качества по альтернативному признаку является довольно хорошее качество продукции при поступлении на контроль. В терминах вероятности появления дефекта θ это выражается в малых ожидаемых значениях θ. Эту информацию можно, например, задать с помощью априорного бета-распределения с параметрами $(1/b, b), b > 1$. Разумеется, существует множество подобных способов, однако нам в дальнейшем потребуется именно однопараметрическое априорное распределение.

Задача контроля качества выражается в проверке гипотезы $H_0 : \theta < \theta_0$ для некоторого близкого к нулю θ_0. Если необходимо контролировать ограничения на вероятности ошибок первого и второго рода α_0 и α_1 соответственно, то надо ввести зону безразличия, таким образом проверяя гипотезу H_0 против альтернативы $H_1 : \theta > \theta_1 = \theta_0 + \Delta$, где Δ — величина зоны безразличия. Введение Δ весьма неудобно, и в практических ситуациях возникает проблема задания этой величины. Выбор осложняется тем, что необходимый объем выборки существенно зависит от Δ, т.к. асимптотическая формула для необходимого объема выборки n^* (см. [35]) имеет вид

$$n^* \approx \frac{\left(\sigma_{\theta_0}\Phi^{-1}(1-\alpha_0) + \sigma_{\theta_1}\Phi^{-1}(1-\alpha_1)\right)^2}{\Delta^2}. \tag{3.30}$$

Гипотезу H_0 можно проверять без применения зоны безразличия (при альтернативе $H_1' : \theta > \theta_0$), при этом гарантируя ограничения β_0 и β_1 на d-риски первого и второго рода. Это требует ввода априорного распределения, однако, как было сказано ранее, у нас уже имеется некая априорная информация, которую можно эффективно использовать. Более того, обычно имеется большой архив экспериментов по контролю качества, на основании которого можно оценить априорное распределение.

Можно, таким образом, сказать, что априорное распределение для критерия, гарантирующего d-риски, выполняет функцию области безразличия критерия, гарантирующего вероятности ошибок. Следовательно, представляет интерес изучение связи между размером области безразличия и параметрами априорного распределения.

При достаточно малых ограничениях на вероятности ошибок асимптотическая формула дана в теореме 2.2 формулами (2.15) и (2.16). Если заменить статистику $R(\mathbf{X})$, примененную в теореме 2.2 на эквивалентную ей статистику $\Phi^{-1}(R(\mathbf{X}))$. В таком случае формулы (2.15) и (2.16) примут вид

$$\frac{\varphi(c) + \Phi(c)c - c\beta_0}{\varphi(c) + \Phi(c)c - c(1-\beta_1)} = \frac{\beta_0 G_0}{\beta_1 G_1}, \tag{3.31}$$

$$n \approx \frac{g(\theta_0)^2(2\varphi(c)+2\Phi(c)c-c(1+\beta_0-\beta_1))^2}{I(\theta_0)(\beta_0 G_0+\beta_1 G_1)^2}. \tag{3.32}$$

Переход к новой тестовой статистике упрощает изучение поведения НОВ при различных значениях параметров.

Интерес в данной части главы представляет сравнение асимптотических формул (3.30), (3.32) для НОВ критерия ϕ_d, гарантирующего ограничения на d-риски и НОВ критерия ϕ_p, удовлеторяющего ограничения на вероятности ошибок соответственно с целью выявления связи между размером области безразличия и параметрами априорного распределения. Для этого все прочие параметры фиксируются. Так, ограничения на вероятности ошибок полагаются равными ограничениям на d-риски: $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$. Необходимый объем выборки для двух критериев также полагается равным, так что исследуется уравнение

$$\frac{\left(\sigma_{\theta_0}\Phi^{-1}(1-\beta_0)+\sigma_{\theta_1}\Phi^{-1}(1-\beta_1)\right)^2}{\Delta^2} = \frac{g(\theta_0)^2(2\varphi(c)+2\Phi(c)c-c(1+\beta_0-\beta_1))^2}{I(\theta_0)(\beta_0 G_0+\beta_1 G_1)^2}, \tag{3.33}$$

где c определяется из уравнения (3.31). При фиксированных прочих параметрах, левая часть зависит только от Δ, а правая часть зависит только от b.

С целью вывода искомой зависимости можно провести численный эксперимент, в котором изучается поведение уравнения (3.33) при различных значениях параметра b. Результаты проведенного численного эксперимента показаны на рисунках 3.2 и 3.3. Значение параметров следующие: $\theta_0 = 0.1$, $\beta_0 = \beta_1 = 0.05$. Значение $b \approx 2.14108$, на котором возникают экстремумы на обоих графиках, соответствует равным априорным вероятностям гипотез $G_0 = \mathbf{G}(\vartheta < \theta_0) = \mathbf{G}(\vartheta \geq \theta_0) = 0.5$.

Характер зависимости Δ от b можно также выявить более аналитичным образом. Так, обратимся к формуле НОВ критерия ϕ_p (3.30). Разложение σ_θ в окрестности точки θ_0 дает

$$\sigma_\theta = \sigma_0 + O(\Delta)$$

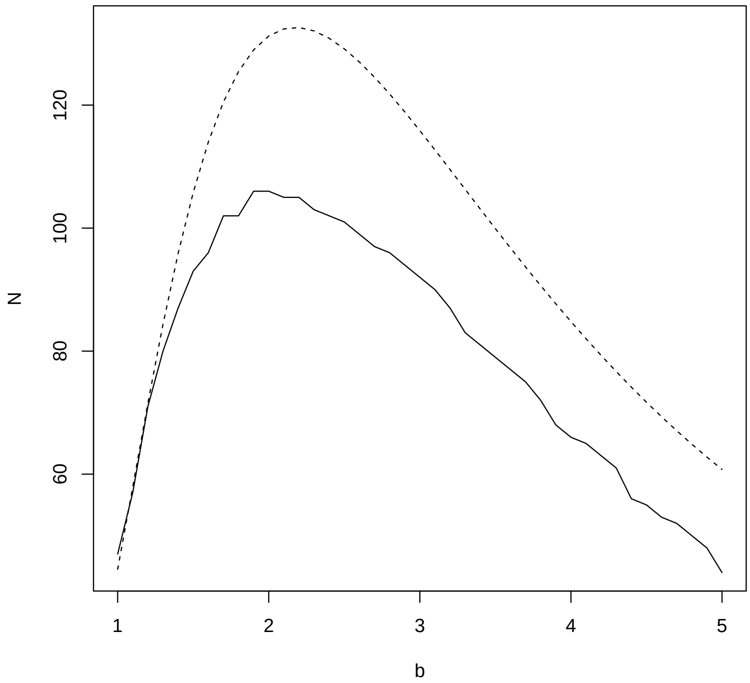

Рис. 3.2. Зависимость НОВ критерия ϕ_d от параметра b. Сплошная линия представляет собой истинный НОВ. Пунктирная линия получена по формуле (3.32).

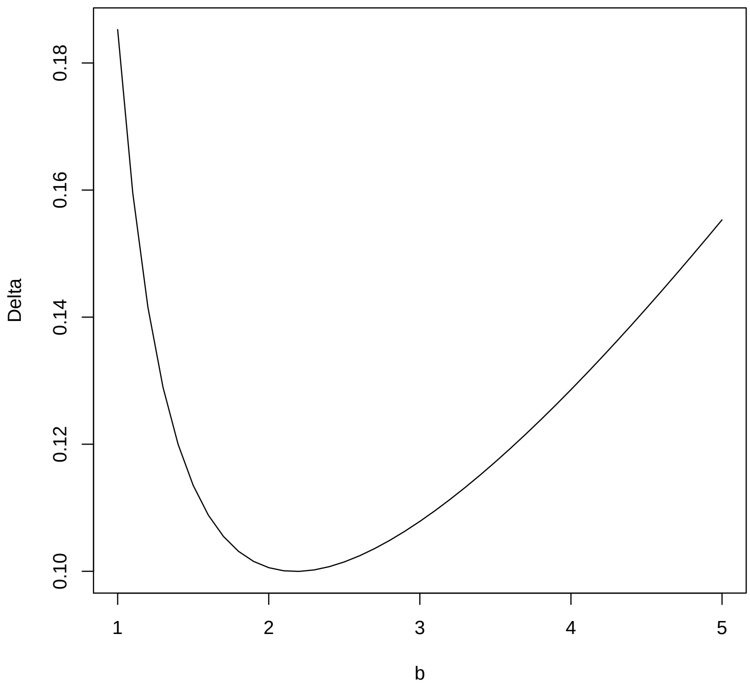

Рис. 3.3. Зависимость Δ от параметра b.

где $\sigma_0 = \sigma_{\theta_0}$. Подставляя это выражение в (3.30), получим

$$n \approx \frac{\sigma_0^2 \left(\Phi^{-1}(1-\beta_0) + \Phi^{-1}(1-\beta_1)\right)^2}{\Delta^2} + O\left(\Delta^{-1}\right).$$

Пусть $(\Phi^{-1}(1-\beta_0)+\Phi^{-1}(1-\beta_1))^2 = Q(\beta_0, \beta_1)$. Тогда после отбрасывания остатка порядка Δ^{-1} получим

$$\Delta \approx \sqrt{\frac{\sigma_0^2 Q(\beta_0, \beta_1)}{n}}.$$

Теперь вместо n можно подставить значение из формулы (3.32) или же истинное значение n для критерия ϕ_d.

Зависимость НОВ n критерия ϕ_d от b значительно сложнее, в первую очередь из-за присутствия нормировки бета-функцией $B(b^{-1}, b)$ в функции плотности бета-распределения $g(\theta) = (B(1/b, b))^{-1}\theta^{1/b-1}(1-\theta)^b, 0 < \theta < 1$.

Глава 4

Асимптотическая d-оптимальность оценки максимального правдоподобия

Исследование асимптотической оптимальности оценок максимального правдоподобия в рамках d-апостериорного подхода к определению риска процедуры статистического вывода проводилась в работе Володина и Новикова [5] в асимптотической схеме сжимающихся (к точке) априорных распределений. В этой главе будет доказана асимптотическая d-оптимальность для произвольных априорных распределений в рамках более общей схемы асимптотического анализа. Вводится новое определение асимптотической d-оптимальности любой процедуры статистического вывода. Это определение полезно для исследования асимптотик d-риска любой процедуры статистического вывода (не только для статистических правил оценки). Будет показано, что оценка максимального правдоподобия является оценкой с асимптотически равномерно минимальным d-риском в классе $\sqrt{n}$-состоятельных оценок.

4.1. Постановка задачи

Пусть дана выборка $\mathbf{X} = (X_1, ..., X_n)$ из распределения $\mathbf{P}_\theta, \theta \in \Theta \subset \mathbb{R}$, где параметр θ есть реализация случайной величины ϑ из распределения $\mathbf{G}$. Рассматривается задача оценивания, в которой определена некоторая функция потерь $L(\theta_1, \theta_2)$. Для измеримой оценки δ_n ее d-риск определяется как величина $\mathcal{R}_{\delta_n}(d) = \mathbf{E}\{L(\vartheta, \delta_n)|\delta_n = d\}$. Классически ([1], [4]) оценка с равномерно минимальным d-риском определяется как такая статистика δ_n^*, что ее d-риск меньше, чем d-риск любой другой оценки: $\mathcal{R}_{\delta_n^*}(d) \leq \mathcal{R}_{\delta_n}(d), d \in \Theta$. При этом предполагается существование регулярной условной вероятности так, чтобы последнее неравенство выполнялось для всех d из пересечения множеств значений функций δ_n и δ_n^*. Оценки с равномерно минимальным d-риском существуют для некоторых моделей (см., опять же, [1], [4]).

Одной из проблем последнего определения является то, что до сих пор не получены широкие условия существования оценок с равномерно минимальным d-риском. Не были также получены их асимптотические свойства, хотя во всех конкретных моделях оценки δ_n^*, оказываются близкими к оценке максимального правдоподобия $\widehat{\theta}_n$.

Идея заключается в том, чтобы показать, что d-риск оценки максимального правдоподобия асимптотически равномерно близок к оптимальному уровню. С этой целью вводится понятие решающей функции с асимптотически равномерно минимальным d-риском. Замечательным свойством этого определения является то, что правило с равномерно минимальным d-риском может и не существовать, так как условие ставится только на близость d-риска решающей функции к наименьшему возможному значению d-риска.

Показано, что при определенных условиях регулярности для степенной функции потерь оценка максимального правдоподобия является оценкой с асимптотически равномерно минимальным d-риском в определенном классе оценок.

К сожалению, не получилось показать, что сама оценка δ_n^* асимптотически близка к $\widehat{\theta}_n$.

4.2. Условия регулярности и некоторые леммы

В текущей главе используются обозначения главы 1, если явно не указано иначе.

Пусть дана некоторая последовательность $\epsilon_n \to 0, n \to \infty$. Сформулируем определение оценки с ϵ_n-асимптотически равномерно минимальным d-риском.

Определение 4.1. *Измеримая оценка δ_n^* называется оценкой с ϵ_n-асимптотически равномерно минимальным d-риском в классе оценок $\mathbb{T}$, ес-*

ли

$$\mathbf{P}\left(\mathcal{R}_{\delta_n^*}(\delta_n^*) \geq \mathcal{R}_{\delta_n}(\delta_n^*) + \epsilon_n\right) \to 0 \qquad (4.1)$$

для любой другой оценки $\delta_n \in \mathbb{T}$.

Для «хорошей» оценки последовательность ϵ_n должна быть как можно меньше. Скажем, для оценки с равномерно минимальным d-риском $\epsilon_n = 0$, а вероятность (4.1) равна нулю.

Основная идея асимптотической оптимальности заключается в том, что апостериорное распределение при определенных условиях асимптотически нормально, а для нормальной выборки с нормальным априори оценка с равномерно минимальным d-риском существует. Технически последовательность ϵ_n является точностью этой нормальной аппроксимации, хотя ее можно рассматривать также как допустимое превышение оценкой δ_n^* оптимального уровня. Следовательно, от вероятностной модели нужно потребовать условий, при которых апостериорное распределение асимптотически нормально.

Напомним, что

$$l_m(x|\theta) = \left(\frac{\partial}{\partial\theta}\right)^m \ln p(x|\theta), \quad \Delta_m(\theta) = \sum_{i=1}^{n} l_m(x_i|\theta).$$

1. Параметрическое пространство Θ есть интервал (ограниченный или неограниченный) на $\mathbb{R}$.

2. Для $\theta \neq \theta'$
$$\int |p(x|\theta) - p(x|\theta')|\, \nu(dx) > 0.$$

3. Найдется $\delta_1 > 0$, такое что
$$\sup_{\theta,\theta'\in\Theta} |\theta - \theta'|^{\delta_1} \int \sqrt{p(x|\theta)p(x|\theta')}\nu(dx) < \infty.$$

4. Носитель распределения $\mathbf{P}_\theta$ не зависит от θ.

5. Функция $p(x|\theta)$ имеет 2 непрерывных производных по $\theta \in \Theta$ для любого x.

6. Пусть $I(\theta) = \mathbf{E}_\theta \left(l_1(X|\theta)\right)^2$.

   а. $0 < I(\theta) < \infty,\ \theta \in \Theta$.

   б. Найдется такое $\delta_2 > 0$, что

$$\mathbf{E}_\theta \left|l_1(X|\theta)\right|^{2+\delta_2} < \infty,\ \mathbf{E}_\theta \left|l_2(X|\theta)\right|^{1+\delta_2} < \infty, \theta \in \Theta.$$

   в. Пусть $U_\beta(\theta)$ — шар вокруг точки θ с радиусом β. Существуют такие $\beta_0, \beta_1 > 0$, что

$$|l_2(x|\theta') - l_2(x|\theta'')| \leq |\theta' - \theta''| D(x|\theta), \quad \theta', \theta'' \in U_{\beta_0}(\theta) \quad \forall x, \theta \in \Theta,$$

$$\mathbf{E}_\theta \sup_{t \in U_{\beta_1}} D(X|t) < \infty, \theta \in \Theta.$$

   г. Для некоторого $\delta_3 \geq 0$

$$\sup_{\theta \in \Theta} (1 + |\theta|^{\delta_3})^{-1} I(\theta) < \infty.$$

7. Априорная плотность g непрерывна на Θ.

8. Функция потерь $L(\theta_1, \theta_2) = l(\theta_1 - \theta_2) = |\theta_1 - \theta_2|^r$ для некоторого $r > 0$.

Всюду далее, если не сказано обратного, мы будем работать в условиях 1–8.

Удалось доказать асимптотическую d-оптимальность оценки максимального правдоподобия только в классе (класс $\mathbb{T}$ из определения 4.1) $\sqrt{n}$-состоятельных оценок параметра θ.

Определение 4.2. *Будем говорить, что оценка δ_n принадлежит классу оценок $\mathbb{C}$ параметра $\theta \in \Theta$, если δ_n измерима и для каждой внутренней точки $\theta \in \Theta$*

$$\forall \varepsilon > 0\ \exists b > 0\ \sup_n \mathbf{P}_\theta \left(\sqrt{n}|T_n - \theta| > b\right) < \varepsilon.$$

По сути здесь повторяется определение 1.1, расширенное на все Θ, в соответствии с условиями следствия 1.1. Следующее определение есть расширенное на все Θ определение 1.2.

Определение 4.3. *Будем говорить, что оценка δ_n принадлежит классу оценок $\mathbb{C}(r), r > 0$ параметра $\theta \in \Theta$, если $\delta_n \in \mathbb{C}$ и*

$$\mathbf{E}_\theta \left|\sqrt{n}(T_n - \theta)\right|^r < \infty,\ \theta \in \Theta.$$

Обозначим

$$\mu_n(\delta_n) = \frac{\Delta_1(\delta_n)}{\sqrt{n} I(\delta_n)},\ \sigma_n^2(\delta_n) = \frac{1}{I(\delta_n)}.$$

Пусть $\mathcal{N}\left(B|\mu, \sigma^2\right)$ — нормальное распределение со средним μ и дисперсией σ^2. Следующая лемма есть копия следствия 1.1 главы 1.

Лемма 4.1. *При выполнении условий 1–7 для всех $\delta_n \in \mathbb{C}$ апостериорная вероятность асимптотически нормальна:*

$$\left\|\mathbf{P}\left(\sqrt{n}(\vartheta - \delta_n) \in \cdot|\mathbf{X}\right) - \mathcal{N}\left(\cdot\middle|\mu_n(\delta_n), \sigma_n^2(\delta_n)\right)\right\|_{TV} \xrightarrow{\mathbf{P}} 0,$$

где $\|\mathbf{P}(\cdot) - \mathbf{Q}(\cdot)\|_{TV}$ — расстояние по общей вариации между мерами $\mathbf{P}$ и $\mathbf{Q}$.

Для измеримой функции $w : \mathbb{R} \mapsto \mathbb{R}$ положим: $F(w|\mu, \sigma^2) = \int_{\mathbb{R}} w(x)\varphi(x|\mu, \sigma^2)dx$. Пусть $R_{\delta_n}(\mathbf{X}) = \mathbf{E}\{l(\vartheta - \delta_n)|\mathbf{X}\}$ — апостериорный риск оценки δ_n. Тогда справедливо следующее утверждение (о нем упомянуто в конце главы 1).

Лемма 4.2. *При выполнении условий 1–8 для всех $\delta_n \in \mathbb{C}(r)$ апостериорный риск*

$$\left|\mathbf{E}\left\{l\left(\sqrt{n}(\vartheta - \delta_n)\right)|\mathbf{X}\right\} - F\left(l\middle|\mu_n(\delta_n), \sigma_n^2(\delta_n)\right)\right| \xrightarrow{\mathbf{P}} 0.$$

Заметим, что $\mathcal{R}_{\delta_n}(\delta_n) = \mathbf{E}\{R_{\delta_n}(\mathbf{X})|\delta_n\}$.

4.3. Асимптотическая d-оптимальность оценки максимального правдоподобия

Теорема 4.1. *Пусть выполнены условия 1–8. Тогда найдется последовательность $\epsilon_n = o(n^{-r/2})$, такая, что оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с ϵ_n-асимптотически равномерно минимальным d-риском в классе $\mathbb{C}(r)$.*

Доказательство.

Обозначим $R^*(d) = n^{-r/2}F(l|0, \sigma_n^2(d))$. Вследствие леммы 4.2 и того, что $\Delta_1(\widehat{\theta}_n) = 0$, выполняется

$$n^{r/2}\left(R^*\left(\widehat{\theta}_n\right) - R_{\widehat{\theta}_n}(\mathbf{X})\right) \xrightarrow{\mathbf{P}} 0. \tag{4.2}$$

Так как в условиях теоремы

$$\mathbf{E}_\theta\left|\sqrt{n}\left(\widehat{\theta}_n - \theta\right)\right|^r < \infty$$

(см., например, [26], следствие 5.2, с. 64), поэтому

$$\mathbf{E}\left|\sqrt{n}\left(\widehat{\theta}_n - \vartheta\right)\right|^r < \infty. \tag{4.3}$$

Далее,

$$\mathbf{E}\left\{R^*\left(\widehat{\theta}_n\right)|\widehat{\theta}_n\right\} = R^*\left(\widehat{\theta}_n\right),$$

и вследствие существования математического ожидания (4.3) условное математическое ожидание от (4.2) относительно $\widehat{\theta}_n$ также сходится к нулю по вероятности:

$$n^{r/2}\left|\mathbf{E}\left\{R^*\left(\widehat{\theta}_n\right) - R_{\widehat{\theta}_n}(\mathbf{X})|\widehat{\theta}_n\right\}\right| = n^{r/2}\left|R^*\left(\widehat{\theta}_n\right) - \mathcal{R}_{\widehat{\theta}_n}\left(\widehat{\theta}_n\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{4.4}$$

Для любой другой оценки $\delta_n \in \mathbb{C}(r)$ вследствие леммы 2

$$\left|n^{r/2}R_{\delta_n}(\mathbf{X}) - F\left(l\,|\,\mu_n(\delta_n), \sigma_n^2(\delta_n)\right)\right| \xrightarrow{\mathbf{P}} 0. \tag{4.5}$$

Используя условие 8, можно применить лемму Андерсона (см. [26], лемма 10.2, с. 217), в силу которой

$$F\left(l\,|\,0,\sigma_n^2(d)\right) \leq F\left(l\,|\,\mu_n(d),\sigma_n^2(d)\right) \tag{4.6}$$

для всех d.

Из (4.5) и (4.6) следует, что

$$\mathbf{P}\left(R_{\delta_n}(\mathbf{X}) \geq R^*(\delta_n) - n^{-r/2}\right) \to 1. \tag{4.7}$$

Условное распределение $\mathbf{P}(\cdot|\delta_n)$ регулярно для измеримых оценок δ_n параметрического пространства $\Theta \subset \mathbb{R}$ (см., например, [28], с.379). Поэтому из (4.7) следует, что

$$\mathbf{P}\left(\mathcal{R}_{\delta_n}(\delta_n) \geq R^*(\delta_n) - n^{-r/2}\right) \to 1. \tag{4.8}$$

Подставим оценку δ_n^* в функцию $\mathcal{R}_{\delta_n}(\cdot)$, и, комбинируя (4.4) и (4.8), получим, что

$$\mathbf{P}\left(\mathcal{R}_{\delta_n}\left(\widehat{\theta}_n\right) \geq \mathcal{R}_{\widehat{\theta}_n}\left(\widehat{\theta}_n\right) - n^{-r/2}\right) \to 1.$$

Более того, найдется такая последовательность $\varepsilon_n \to 0$, что

$$\mathbf{P}\left(\mathcal{R}_{\delta_n}\left(\widehat{\theta}_n\right) \geq \mathcal{R}_{\widehat{\theta}_n}\left(\widehat{\theta}_n\right) - \varepsilon_n n^{-r/2}\right) \to 1.$$

Существование такой последовательности обосновывается тем, что для сходимости по вероятности к нулю некоторых величин Y_n можно найти такую последовательность $\varepsilon_n \to 0$, что $\mathbf{P}(|Y_n| > \varepsilon_n) \to 0$. Такая последовательность ε_n находится из выражений (4.4) и (4.5), из которых, собственно, и следует результат.

Поэтому, полагая $\epsilon_n = \varepsilon_n n^{-r/2}$, мы получим требуемый результат.

■

Ограничение на класс $\mathbb{C}(r)$ неприятно, однако само свойство $\sqrt{n}$-состоятельности само по себе может восприниматься как оптимальное свойство. Например, оценки максимального правдоподобия и байесовская принадлежат классу $\mathbb{C}(\infty)$ при довольно широких условиях (в частности, при условиях 1–8). Эта проблема, в сущности, заключается в том, что не доказано, что оценка с равномерно минимальным d-риском принадлежит классу $\mathbb{C}$ при сколько-нибудь широких условиях, что, однако, весьма вероятно. Тогда можно было бы сделать утверждение о том, что оценка $\widehat{\theta}_n$ является оценкой с асимптотически равномерно минимальным d-риском среди всех измеримых оценок параметра θ. Более того, можно было бы утверждать, что сама оценка максимального правдоподобия сходится к оценке с равномерно минимальным d-риском со скоростью $n^{-1/2}$.

В случае, когда Θ ограничено, можно сузить класс оценок, в котором $\widehat{\theta}_n$ оптимальна.

Теорема 4.2. *Пусть выполнены условия 1–8 и пусть Θ ограничено. Тогда найдется последовательность $\epsilon_n = o(1)$, такая, что оценка максимального правдоподобия $\widehat{\theta}_n$ является оценкой с ϵ_n-асимптотически равномерно минимальным d-риском в классе оценок $\mathbb{C}$.*

Доказательство.

По сути свойство $\delta_n \in \mathbb{C}(r)$ использовалось только при получении (4.5) с помощью леммы 4.2. Обозначим $p_n(\theta|\mathbf{X})$ функцию апостериорной плотности относительно меры Лебега. Для ограниченного Θ и любой оценки δ_n для некоторого M имеем $|\delta_n - \theta|^r \le M$, и

$$\left|R_{\delta_n}(\mathbf{X}) - n^{-r/2}F\left(l|\mu_n(\delta_n), \sigma_n^2(\delta_n)\right)\right| =$$
$$= \left|\int\limits_{\Theta} |\delta_n - \theta|^r p(\theta|\mathbf{X})d\theta - \int\limits_{\Theta} |\delta_n - \theta|^r d\Phi\left(\theta\middle|\delta_n + \frac{\mu_n(\delta_n)}{\sqrt{n}}, \frac{\sigma_n^2(\delta_n)}{n}\right)\right| \le$$
$$\le M\int\limits_{\Theta}\left|p(\theta|\mathbf{X}) - \varphi\left(\theta\middle|\delta_n + \frac{\mu_n(\delta_n)}{\sqrt{n}}, \frac{\sigma_n^2(\delta_n)}{n}\right)\right|d\theta,$$

где φ — нормальная функция плотности. Последнее выражение согласно лемме 4.1 стремится к нулю по вероятности (стремление к нулю по вариации равносильно стремлению к нулю интеграла от модуля разности плотностей). Поэтому найдется такая последовательность ε_n, что

$$\mathbf{P}\left(\left|R_{\delta_n}(\mathbf{X}) - n^{-r/2}F\left(l \mid \mu_n(\delta_n), \sigma_n^2(\delta_n)\right)\right| > \varepsilon_n\right) \to 0.$$

Таким образом, получен аналог выражения (4.5), в котором скорость стремления к нулю равна не $o(n^{-r/2})$, а $o(1)$. Повторяя рассуждения теоремы 4.1, можно получить

$$\mathbf{P}\left(\mathcal{R}_{\delta_n}\left(\widehat{\theta}_n\right) \geq \mathcal{R}_{\widehat{\theta}_n}\left(\widehat{\theta}_n\right) - \varepsilon_n\right) \to 1,$$

что и является утверждением теоремы.

■

По большому счету, разница в утверждениях теорем 4.1 и 4.2 заключается в доказуемой скорости сходимости апостериорного риска $R_{\delta_n}(\mathbf{X})$ к нулю. Это не слишком правильно, поскольку в конечном счете необходимо сравнивать d-риски оценок, которые могут существенно отличаться от апостериорных. Отсюда и получается утверждение теоремы 4.2, которое, несмотря на то, что к рассмотрению принимаются оценки с «худшими» асимптотическими свойствами, значительно слабее утверждения теоремы 4.1.

Здесь также уместно сравнение с оценками с равномерно минимальным риском (минимизирующим $\mathbf{R}_{\delta_n}(\theta) = \mathbf{E}_\theta L(\theta, \delta_n)$), которые существуют только в классе несмещенных оценок (или, обобщенно, оценок с одинаковым смещением). Однако в случае оценок с равномерно минимальным риском совершенно очевиден пример вырожденной оценки $\bar{\delta}_n = \theta_0$, для которой $\mathbf{R}_{\bar{\delta}_n}(\theta_0) = 0$, так что ограничение на класс оценок весьма естественно. Для случая d-риска такой пример пока неизвестен. Для двух различных вырожденных оценок вообще непонятно, как сравнивать их d-риски.

Заключение

В диссертации получены асимптотические разложения апостериорного распределения и апостериорных средних в различных постановках. Часть этих результатов была применена при выводе асимптотики необходимого объема выборки для d-гарантийного критерия во второй главе, и при выводе асимптотики d-риска оценки максимального правдоподобия в четвертой главе.

Техника разложений апостериорных распределений была применена в третьей главе, при выводе асимптотического разложения d-риска оптимального критерия в схеме испытаний Бернулли. Потребовалась техника разложения Эджворта для биномиального распределения, которая также была использована для получения асимптотических разложений для вероятностей ошибок и необходимого объема выборки критерия, гарантирующего ограничения на вероятности ошибок.

В четвертой главе было также введено новое понятие оценки с асимптотически равномерно минимальным d-риском. Это определение может быть использовано для прочих процедур, в том числе и процедур проверки гипотез.

Таким образом, поставленные цели и задачи диссертационной работы выполнены.

Список литературы

1. *Володин И. Н., Симушкин С. В.* О d-апостериорном подходе к проблеме статистического вывода. // Lect. Notes Math. — 1983. — Т. 1021. — С. 629—636.

2. *Volodin I. N.* Guaranteed statistical inference procedures (determination of the optimal sample size) // J.Soviet Math. — 1989. — Т. 44, № 5. — С. 568—600.

3. *Volodin I. N., Simushkin S. V.* Statistical Inference With the Minimal d-Risk // IV USSR-Japan Siposium on Prob. Th. and Mathem.Stat. — 1982. — Т. 2. — С. 249—250.

4. *Володин И. Н., Симушкин С. В.* Несмещенность и байесовость // Изв. вузов. Матем. — 1987. — Т. 1. — С. 3—7.

5. *Володин И. Н., Новиков А. А.* Статистические оценки с асимптотически минимальным d-риском // Теория вероятн. и ее примен. — 1993. — Т. 38, № 1. — С. 20—32.

6. *Симушкин С. В.* Оптимальные d-гарантийные процедуры различения двух гипотез // Рукопись деп. в ВИНИТИ. — 1981. — Т. 5547-81.

7. *Симушкин С. В.* Оптимальный обхем выборки при d-гарантийном различении двух гипотез // Изв. вузов. Матем. — 1982. — Т. 240, № 5. — С. 47—52.

8. *Володин И. Н., Новиков А. А.* Асимптотика необходимого объема выборки при d-гарантийном различении двух близких гипотез // Изв. вузов. Матем. — 1983. — Т. 11. — С. 59—66.

9. *Володин И. Н., Новиков А. А.* Асимптотика необходимого объема выборки при гарантийном различении параметрических гипотез // Исследования по прикладной математике. — 1999. — Т. 21. — С. 3—41.

10. *Vaart V.* Asymptotic statistics. — Cambridge : Cambridge University Press, 1998.

11. *Le Cam L. M.* Asymptotic methods in statistical theory. (Springer series in statistics). — New York : Springer-Verlag New York Inc., 1986.

12. *Johnson R. A.* Asymptotic expansions associated with posterior distributions // The Annals of Mathematical Statistics. — 1970. — Т. 41, № 3. — С. 851—864.

13. *Гусев С. И.* Асимптотические разложения, связанные с некоторыми статистическими оценками в гладком случае. I. Разложения случайных величин // Теория вероятностей и ее применения. — М, 1975. — Т. 20, № 3. — С. 488—514.

14. *Ghosh J. K.*, *Sinha B. K.*, *Joshi S. N.* Expansions for posterior probability and integrated Bayes risk // Statistical Decision Theory and Related Topics III. — 1982. — Т. 1. — С. 403—456.

15. *Weng R. C.* A Bayesian Edgeworth expansion by Stein's Identity // Bayesian Analysis. — 2010. — Т. 5, № 4. — С. 741—764.

16. *Weng R. C.*, *Hsu C.-H.* A Study of Expansions of Posterior Distributions // Communications in Statistics – Theory and Methods. — 2013. — Т. 42. — С. 343—364.

17. *Бурнашев М. В.* Исследование свойств второго порядка статистических оценок в схеме независимых наблюдений // Изв. АН СССР. Сер. матем. — 1981. — Т. 45, № 3. — С. 509—539.

18. *Заикин А. А.* Дефект размера нерандомизированного критерия и влияние рандомизации на сокращение необходимого объема выборки при тестировании вероятности успеха в схеме испытаний Бернулли // Теория вероятн. и ее примен. — 2014. — Т. 59, № 3. — С. 417—435.

19. *Zaikin A. A.* On asymptotic expansion of posterior distribution // Lobachevskii Journal of Mathematics. — 2016. — Т. 37, № 4. — С. 515—525.

20. *Заикин А. А.* Асимптотическое разложение апостериорного распределения параметра, центрированного $\sqrt{n}$-состоятельной оценкой // Записки научных семинаров ПОМИ. — 2016. — Т. 454. — С. 121—150.

21. *Заикин А. А.* Асимптотическое разложение апостериорного распределения параметра вероятностной модели // Труды математического центра имени И. Н. Лобачевского: Теория функций, ее приложения и смежные вопросы. Материалы двенадцатой международной Казанской летней научной школы-конференциии. Т. 51. — Казань, 2015. — С. 190—192.

22. *Заикин А. А.* Оценки с асимптотически равномерно минимальным d-риском // Материалы международной конференции по алгебре, анализу и геометрии. — Казань : Казанский университет, издательство академии наук РТ, 2016. — С. 171—172.

23. *Ибрагимов И. А., Хасьминский Р. З.* Асимптотическое поведение некоторых статистических оценок в гладком случае. I. Исследование отношения правдоподобия // Теория вероятностей и ее применения. — М, 1972. — Т. 17, № 3. — С. 469—486.

24. *Ибрагимов И. А., Хасьминский Р. З.* Асимптотическое поведение некоторых статистических оценок в гладком случае. II. Предельные теоремы для апостериорной плотности и байесовских оценок // Теория вероятностей и ее применения. — М, 1973. — Т. 18, № 1. — С. 78—93.

25. *Ибрагимов И. А., Хасьминский Р. З.* О моментах обобщенных байесовских оценок и оценок максимального правдоподобия // Теория вероятностей и ее применения. — М, 1973. — Т. XVIII, № 3. — С. 535—546.

26. *Ибрагимов И. А., Хасьминский Р. З.* Асимптотическая теория оценивания. — Москва : Наука, 1979. — С. 528.

27. *Le Cam L. M.* On the asymptotic theory of estimation and testing // Proc. 3rd Berkeley Symp. Math. Statist. Prob. — 1956. — Т. 1. — С. 129—156.

28. *Лоэв М.* Теория вероятностей. — Москва : Издательство иностраной литературы, 1962. — С. 720.

29. *Русас Д.* Контигуальность вероятностных мер. — Москва : Мир, 1975.

30. ГОСТ Р 50779.30-95. Статистические методы. Приемочный контроль качества. Общие требования. — Москва : Изд-во стандартов, 2008. — С. 24.

31. ГОСТ Р 50779.52-95. Статистические методы. Приемочный контроль качества по альтернативному признаку. — Москва : Изд-во стандартов, 2004. — С. 230.

32. *Бернштейн С. Н.* О "доверительных"вероятностях Фишера // Изв. АН СССР. Сер. матем. — 1941. — Т. 5, № 2. — С. 85—94.

33. *Беляев Ю. К.* Вероятностные методы выборочного контроля. — Москва : Физматлит, 1975. — С. 408.

34. *Сенатов В. В.* О реальной точности аппроксимаций в центральной предельной теореме // Сибирский математический журнал. — 2011. — Т. 52, № 4. — С. 727—746.

35. *Володин И. Н.* О числе наблюдений, необходимых для различения двух близких гипотез // Теория вероят. и ее применен. — 1967. — Т. 12, № 3. — С. 575—581.

36. *Кендалл М.*, *Стюарт А.* Теория распределений. — Москва : Наука, 1966. — С. 587.

37. *Большев Л. Н.* Асимптотически пирсоновские преобразования // Теория вероят. и ее применен. — 1963. — Т. 8, № 2. — С. 129—155.