Министерство образования и науки Российской Федерации
Казанский (Приволжский) федеральный университет

ВВЕДЕНИЕ В ПОСЛЕДОВАТЕЛЬНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

Ан.А. Новиков, П.А. Новиков, Ан.Ан. Новиков

Казань 2015

Новиков Андрей Алексеевич
к.ф.-м.н.
Департамент математики
Universidad Autónoma Metropolitana – Iztapalapa
(Мексика)

Новиков Петр Андреевич
к.ф.-м.н.
Кафедра программной инженерии
Высшая школа информационных технологий и информационных систем
Казанский (Приволжский) федеральный университет
(г. Казань)

Новиков Андрей Андреевич
Кафедра математического анализа
Институт математики и механики им. Н.И. Лобачевского
Казанский (Приволжский) федеральный университет
(г. Казань)

Рецензент:

1 Введение

Последовательный анализ является полноправным разделом математической статистики уже более шестидесяти лет. Наиболее известным результатом в этой области является оптимальность статистического критерия, известного как «последовательный критерией отношения вероятностей» (*sequential probability ratio test, SPRT*), первое доказательство которого было опубликовано в 1948 г. (см. A. Wald, J. Wolfowitz [11]).

Классическим результатам последовательного анализа ныне находится место и в монографиях, таких, например, как классические монографии Э. Лемена [4] и Ш. Закса [3], и даже в учебниках, например, в книге А.А. Боровкова [1]. Впрочем, несмотря на то, что в упомянутых изданиях (конечно, равно как и в [11]) и приводятся некоторые неформальные аргументы в пользу оптимальности SPRT, ни одна из подобных аргументаций не может быть признана математически строгой.

В таких изданиях, как [2] или [8], дается более или менее строгое обоснование, но при этом опускается множество существенных деталей, что делает их слишком сложными для студентов, даже для самых подготовленных и прилежных.

Также существуют весьма специализированные монографии, в которых излагаемый материал соответствует всем критериям математической строгости (см. [7], [6]) но в таких монографиях основной упор делается на теорию *оптимальной остановки*, которая сама по себе весьма обширна и технически сложна. В итоге получается, что почти все внимание в этих монографиях уделяется этой теории и в гораздо меньшей степени – статистическим приложениям. Так, например, во всей книге И. Чао и др. [7] не нашлось места для доказательства оптимальности SPRT, при том, что вся теоретическая база для этого в книге есть. Значительно больше внимания уделяется статистическим приложениям в монографии А.Н. Ширяева [6], также в основном посвященной развитию теории оптимальной остановки (для марковских процессов), но и в ней имеется заметный пробел в обосновании оптимальности SPRT, заполняемый ссылкой на уже цитированную монографию Лемана [4].

В заключение обзора следует отметить те книги, в которых доказательство оптимальности SPRT присутствует, но не выдерживает критики (см. [9]), и те, в которых оно вообще опускается ввиду его «технической сложности» (например, [10]).

Цель настоящего курса лекций – предложить студенту, интересующемуся математическими основами последовательного анализа, ясное, краткое и самодостаточное доказательство оптимальности SPRT.

Ограничим рассмотрение случаем наблюдений из дискретных распределений (к таковым относятся распределения Бернулли, равномерное дискретное, гипергеометрическое, Пуассона и т.д.). Читатель, владеющий теорией меры и интегрирования сможет повторить те же рассуждения для непрерывных случайных величин, или даже для общего случая случайных элементов измеримого пространства.

Мы рассматриваем дискретный случай, для понимания которого не требуется знать больше, чем основные понятия мат. анализа (сходимость числовых последовательностей и числовых рядов) и основные понятия теории вероятности (распределение и мат. ожидание) и статистики (проверка гипотез). Таким образом, курс подходит студентам-бакалаврам, знакомым с курсом мат. статистики.

У читателей, вообще не знакомых с мат. статистикой, эти лекции могут вызвать некоторые сложности, которые, впрочем, не окажутся непреодолимыми, поскольку курс в известной мере самодостаточен. Материал данного курса лекций даже может быть использован в качестве вводного в теорию статистических критериев и принятия решений.

2 Математическая модель и основные понятия

В этой части даются строгие определения последовательного статистического эксперимента, статистических гипотез, статистических критериев и их числовых характеристик.

2.1 Последовательный статистический эксперимент

В этом разделе дается строгое определение последовательного статистического эксперимента.

В рамках любого статистического эксперимента производится сбор данных. Наиболее распространенной моделью является *выборка из распределения*, в которой предполагается, что некоторое число n случайных величин, происходящих из одного и того же вероятностного распределения, наблюдается независимым образом. Результат такого наблюдения называется *случайной выборкой объема* n из данного распределения.

Идея же последовательного анализа заключается в том, чтобы не производить заданное число наблюдений единовременно, а формировать выборку из наблюдений *последовательно*, прекращая наблюдение в наиболее подходящий или наиболее удобный момент времени.

Придадим понятию последовательного эксперимента необходимую строгость, ограничиваясь случаем дискретного распределения с конечным числом значений.

Пусть X – дискретная случайная величина. Такая случайная величина характеризуется не более чем счетным множеством значений x и набором вероятностей $f(x) = P(X = x)$, соответствующих этим значениям. Для того, чтобы $f(x)$ было вероятностным распределением, необходимо выполнение равенства

$$\sum f(x) = 1, \tag{1}$$

в котором суммирование ведется по всем значениям x (сумма можеть быть бесконечной). Здесь и далее мы опускаем пределы суммирования, если речь идет о сумме по *всем* возможным значениям переменной (-ных) под знаком суммирования. Так, единственная переменная в формуле (1) – это x, и суммирование ведется по всем ее значениям.

Последовательно получаемые экспериментальные данные $X_1, X_2, \ldots, X_n, \ldots$ – это *независимые* копии случайной величины X. Иными словами,

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} f(x_i). \tag{2}$$

для всех $n = 1, 2, 3 \ldots$.

Момент остановки эксперимента будем определять по некоторому *правилу*, зависящему от последовательно получаемых наблюдений. На n-м шаге эксперимента, то есть при наблюдениях $X_1, X_2, \ldots, X_n$, правило должно определить, следует ли завершить наблюдения на данном шаге, или наоборот, требуется провести еще больше наблюдений.

Пусть $\psi_n = \psi_n(X_1, X_2, \ldots, X_n)$ – функция, принимающая значения 0 или 1 (*индикаторная функция*). В качестве правила остановки на n-шаге, в принципе, может служить любая функция такого вида, если трактовать ее значения следующим образом:

- остановить эксперимент, если $\psi_n = 1$,
- продолжить наблюдение, если $\psi_n = 0$.

Назовем последовательность ψ функций такого вида

$$\psi = (\psi_1, \psi_2, \ldots, \psi_n, \ldots)$$

правилом остановки.

Правило остановки ψ порождает следующий алгоритм последовательного эксперимента:

1. Произвести первое наблюдение X_1 и положить n (номер шага) $= 1$.
2. Если $\psi_n(X_1, X_2, \ldots, X_n) = 1$, то завершить эксперимент,
3. в противном случае произвести наблюдение X_{n+1}, увеличить n на 1 и перейти к шагу 2.

Иначе говоря, эксперимент прекращается в первый момент n при котором выполнится $\psi_n(X_1, X_2, \ldots, X_n) = 1$. Общее число произведенных наблюдений будет равно

$$\nu = \begin{cases} \min\{n \geq 1 : \psi_n(X_1, X_2, \ldots, X_n) = 1\} & \text{если это число конечно} \\ \infty & \text{в противном случае.} \end{cases} \tag{3}$$

Естественно, могут существовать такие правила остановки, для которых ν будет бесконечным (например, $\nu \equiv \infty$, если $\psi_n \equiv 0$, для любого $n \geq 1$). Из соображений практики следует избегать применения правил, допускающих подобное. Ниже мы придадим этому требованию более строгий смысл.

Назовем (случайную) величину ν, определенную в (3), *моментом остановки.* Распределение случайной величины ν задается с помощью формулы

$$P(\nu = n) = P(\psi_1 = 0, \psi_2 = 0, \ldots, \psi_{n-1} = 0, \psi_n = 1). \tag{4}$$

Выразим эту вероятность в терминах правила остановки и закона распределения случайной величины X.

Пусть

$$\chi_n = \chi_n(X_1, X_2, \ldots, X_n) = (1 - \psi_1)(1 - \psi_2) \ldots (1 - \psi_{n-1})\psi_n \tag{5}$$

для $n = 1, 2, \ldots$. Легко видеть, что χ_n – индикаторная функция, а $\chi_n = 1$ тогда и только тогда, когда $\nu = n$. Следовательно,

$$\begin{aligned} P(\nu = n) &= E\chi_n(X_1, X_2, \ldots, X_n) & (6) \\ &= \sum \chi_n(x_1, x_2, \ldots, x_n) \prod_{i=1}^{n} f(x_i), & (7) \end{aligned}$$

вследствие (2).

Суммы вида (7) еще будут встречаться в нашем курсе. Выше мы указывали, что $\psi_n = \psi_n(X_1, X_2, \ldots, X_n)$ и $\chi_n = \chi_n(X_1, X_2, \ldots, X_n)$. Но в формулах, подобных (7), под тем же самым обозначением, скажем, χ_n будем также понимать и $\chi_n(x_1, x_2, \ldots, x_n)$. Это не должно вызывать путаницы, если принимать во внимание контекст: если функция F_n стоит под знаком математического ожидания или вероятности (как в случае с (6)), то под ней будем понимать $F_n = F_n(X_1, X_2, \ldots, X_n)$; если же она стоит под знаком суммирования (как в (7)), то $F_n = F_n(x_1, x_2, \ldots, x_n)$.

С учетом этих обозначений формулы (6)–(7) запишутся в виде

$$P(\nu = n) = E\chi_n = \sum \chi_n \prod_{i=1}^{n} f(x_i). \tag{8}$$

2.2 Статистическая постановка. Проверка гипотез.

Статистический контекст присутствует во всех случаях, когда речь идет о неопределенности распределения, из которого берутся данные.

Так, если в статистической модели (2) вероятность $f(x)$ зависит от параметра θ

$$f(x) = f_\theta(x),$$

то любой вопрос, касающийся истинного значения параметра θ, когда мы располагаем лишь выборочными данными $X_1, X_2, \ldots, X_n$, входит в круг вопросов статистики.

Например, для выборки из распределения Бернулли

$$f(x) = P(X = x) = \begin{cases} p & \text{, если } x = 1 \\ 1 - p & \text{, если } x = 0 \end{cases}$$

а параметр θ – это вероятность «успеха» $p \in [0, 1]$.

Одна из важнейших задач статистики – это проверка гипотез. В нашем курсе мы рассмотрим простейший случай проверки гипотез: случай *двух простых гипотез*. Гипотеза называется *простой*, если ей отвечает *единственное* распределене.

Итак, формулировка задачи, рассматриваемой в нашем курсе, следующая: на основании полученных данных определить, которая из двух гипотез, $H_0 : \theta = \theta_0$ или $H_1 : \theta = \theta_1$, истинна.

Так, в примере с распределением Бернулли вопрос может быть поставлен следующим образом: правда ли, что «успехи» и «неудачи» в наблюдаемой последовательности происходят с одинаковой частотой? Другими словами, нас интересует простая гипотеза $H_0 : p = 0.5$. Если в качестве альтернативы этому имеется другое предположение (H_1) о значении параметра p, например, $p = 0.6$, то следует задаться вопросом, как, используя выборочные данные $X_1, X_2, \ldots, X_n$, принять решение в пользу первой или второй из двух гипотез.

Естественно, тот же самый вопрос может быть поставлен и в контексте последовательного эксперимента, описанного в предыдущем разделе.

И первое, что нам необходимо сделать, – это дать строгое определение процедуры *проверки гипотезы* в последовательной постановке.

Как мы видели в предыдущем разделе, эксперимент может быть завершен после любого числа шагов $n = 1, 2, 3, \ldots$, так что следует задаться вопросом решения в пользу H_0 или H_1 для любого конечного числа наблюдений n. Решение состоит либо в принятии H_0, либо в его отклонении в пользу H_1, и в качестве *решающего правила* нам подойдет, как и раньше, индикаторная функция, то есть функция, принимающая значения только 0 или 1.

Ввиду вышесказанного определим решающее правило ϕ как последовательность индикаторных функций $\phi_n = \phi_n(X_1, X_2, \dots X_n)$, $n = 1, 2, , \dots,$

$$\phi = (\phi_1, \phi_2, \dots, \phi_n, \dots), \tag{9}$$

которые будем трактовать следующим образом: при завершении последовательного эксперимента на шаге $\nu = n$ следует

- *принять* гипотезу H_0, если $\phi_n(X_1, X_2, \dots, X_n) = 0$ и
- *отклонить* H_0 в пользу H_1, если $\phi_n(X_1, X_2, \dots, X_n) = 1$

Последовательный критерий – это пара (ψ, ϕ) с правилом остановки ψ и решающим правилом ϕ.

Классические критерии, основанные на выборке фиксированного объема n, являются частным случаем последовательного критерия, а именно, случаем, когда $\psi_1 \equiv \psi_2 \equiv \dots \equiv \psi_{n-1} \equiv 0$ и $\psi_n \equiv 1$. Следовательно, классическому критерию может быть поставлено в соответствие просто решающее правило ϕ_n, которое для завершенного эксперимента с набором данных $X_1, X_2, \dots, X_n$ предписывает принятие гипотезы H_i с $i = \phi_n(X_1, X_2, \dots, X_n)$.

Видно, что класс последовательных критериев куда более обширен, чем класс критериев с фиксированным объемом выборки, поскольку каждый критерий, помимо решающего правила, содержит и правило остановки, определяющее окончательное число наблюдений в эксперименте. Это касается даже *усеченных* критериев, для которых возможное число наблюдений ограничено сверху, так как такие критерии позволяют остановиться на любом шаге вплоть до n.

2.3 Характеристики критериев

Последовательный критерий задает алгоритм хода получения выборочных данных и принятия решения по завершении эксперимента.

Ход алгоритма, основанного на случайных данных, может в итоге привести к ошибке.

Может оказаться, что при истинной H_0 окончательным решение будет «отклонить». И наоборот, при истинной H_1 окончательным решением будет «принять H_0». Первый случай называется «ошибка I рода» (также «False Negative», FN) а второй – «ошибка II рода» (также «False Positive», FP).

Если нам важно качество статистического вывода, нам следует убедиться, что ошибки происходят нечасто. Мера частоты – вероятность, поэтому при рассмотрени критериев следует принимать во внимание вероятности ошибок. Отсюда вытекают определения *вероятности ошибки I рода*

$$\alpha(\psi, \phi) = P_{\theta_0}(\text{отклонить } H_0) \tag{10}$$

и *вероятности ошибки II рода*

$$\beta(\psi, \phi) = P_{\theta_1}(\text{принять } H_0). \tag{11}$$

Другая важная характеристика – это общее число наблюдений ν; на практике мы заинтересованы в скорейшем прекращении эксперимента, потратив как можно меньше времени, денег, ресурсов и т .д.

Так как ν – случайная величина, естественным будет рассмотреть ее мат. ожидание. Но мат. ожидание может быть вычислено как при H_0, так и при H_1, что дает две характеристики:

$$N(\theta_0;\psi) = E_{\theta_0}\nu \quad \text{и} \quad N(\theta_1;\psi) = E_{\theta_1}\nu. \tag{12}$$

Прежде всего, выразим характеристики (10-12) в терминах элементов статистической модели и критерия.

В общем случае назовем вероятность *отклонить* H_0 при истинном значении параметра θ

$$m(\theta;\psi,\phi) = P_\theta(\text{отклонить } H_0)$$

функцией мощности критерия (ψ,ϕ).

Из (10) следует

$$\alpha(\psi,\phi) = m(\theta_0;\psi,\phi), \tag{13}$$

а из (11) –

$$\beta(\psi,\phi) = 1 - m(\theta_1;\psi,\phi), \tag{14}$$

так что перейдем к вычислению $m(\theta;\psi,\phi)$.

Так как

$$m(\theta;\psi,\phi) = P_\theta(\text{отклонить } H_0) = \sum_{n=1}^{\infty} P_\theta(\text{отклонить } H_0 \text{ при } \nu = n) \tag{15}$$

$$= \sum_{n=1}^{\infty} P_\theta(\phi_n = 1,\ \nu = n),$$

с учетом того, что $\nu = n$ тогда и только тогда, когда $\chi_n = 1$ (см. (5) и следующий за ним абзац), имеем

$$m(\theta;\psi,\phi) = \sum_{n=1}^{\infty} E_\theta \chi_n \phi_n = \sum_{n=1}^{\infty} \sum \chi_n \phi_n \prod_{i=1}^{n} f_\theta(x_i). \tag{16}$$

Для того, чтобы выражение (15) имело смысл, необходимо выполнение условия

$$P_\theta(\nu < \infty) = 1, \quad \text{para} \quad \theta = \theta_0 \quad \text{y} \quad \theta = \theta_1, \tag{17}$$

поэтому будем предполагать, что это условие выполнено. Выполнение этого условия можно обосновать тем, что из соображений практики бесконечное значение ν лишено смысла.

Аналогично (15)-(16) получаем

$$N(\theta;\psi) = \sum_{n=1}^{\infty} n P_\theta(\nu = n) = \sum_{n=1}^{\infty} n E_\theta \chi_n = \sum_{n=1}^{\infty} n \sum \chi_n \prod_{i=1}^{n} f_\theta(x_i), \tag{18}$$

чем расчет характеристик последовательных критериев завершается.

3 Оптимальность последовательных критериев

Это основная часть курса, содержащая постановку задачи оптимальности последовательного критерия и ее исчерпывающее решение.

3.1 Постановка задачи

Классическая задача последовательного анализа – нахождение формы критерия, минимизирующего средний объем наблюдений в классе всех критериев, вероятности ошибок которых не превосходят заданных ограничений.

Другими словами, рассмотрим класс $\Delta(\alpha, \beta)$ критериев (ψ, ϕ) удовлетворящих неравенствам

$$\alpha(\psi, \phi) \leq \alpha \quad \text{и} \quad \beta(\psi, \phi) \leq \beta, \tag{19}$$

где $0 < \alpha < 1$, $0 < \beta < 1$ – некоторые константы. Цель – найти критерий в классе $\Delta(\alpha, \beta)$ со значением $N(\theta_0; \psi)$ (и/или $N(\theta_1; \psi)$), минимальным в данном классе.

Прежде всего отметим, что вопрос о существовании критерия, минимизирующего *оба* мат. ожидания $N(\theta_0; \psi)$ и $N(\theta_1; \psi)$ *одновременно*, остается открытым. В связи с этим поставим менее амбициозную задачу: найти критерий с минимальным значением $N(\theta_0; \psi)$ в классе $\Delta(\alpha, \beta)$.

В действительности искомый критерий будет также минимизировать и $N(\theta_1; \psi)$. Однако нам этот факт, по крайней мере, с теоретической точки зрения, не представляется очень важным по следующей причине. Так как θ_0 и θ_1 взаимозаменяемы, то, найдя решение задачи минимизации $N(\theta_0; \psi)$, мы тут же получим критерий, минимизирующий $N(\theta_1; \psi)$ в $\Delta(\alpha, \beta)$. То, что получается тот же самый критерий, – результат любопытный, но получен он как чистое совпадение, а не как следствие нашей воли, желания или усилий. По этой причине оставим рассмотрение этого результата за рамками курса.

3.2 Сведение к задаче оптимизации без ограничений. Вариационный метод Лагранжа.

Задачу из предыдущего раздела удобно представлять в терминах стоимости эксперимента.

Пусть стоимость проведения одного наблюдения равна $c > 0$, средняя общая стомость последовательного эксперимента, проводимого в соответствии с правилом остановки ψ, равна $C(\psi) = E_{\theta_0}(c\nu) = cN(\theta_0; \psi)$. Таким образом, задача из предыдущего раздела эквивалентна задаче минимизации средней стоимости проведения эксперимента $C(\psi)$ при условии, что вероятности ошибок критерия удовлетворяют условиям (19).

Поставленная задача является задачей оптимизации с ограничениями с целевой функцией $C(\psi)$ и ограничениями типа неравенств (19).

Задачу оптимизации с ограничениями можно решить методом Лагранжа. С учетом того, что целевая функция $C(\psi)$ – функция от *функции*, метод Лагранжа будет называться *вариационным.*

Суть метода заключается в сведении задачи оптимизации с ограничениями к задаче оптимизации без ограничений путем включения ограничений в новую целевую функцию, называемую функцией Лагранжа.

В нашем случае функция Лагранжа запишется как

$$L(\psi, \phi) = C(\psi) + \lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi), \tag{20}$$

где $\lambda_0 \geq 0$ и $\lambda_1 \geq 0$ – константы, называемые *множителями Лагранжа.*

Следующая теорема сводит задачу оптимизации с ограничениями к задаче оптимизации без ограничений, в чем, собственно, и состоит суть метода Лагранжа.

Теорема 1. *Пусть* (ψ^*, ϕ^*) *– критерий, такой, что для любого другого последовательного критерия* (ψ, ϕ) *выполняется*

$$L(\psi, \phi) \geq L(\psi^*, \phi^*), \tag{21}$$

и

$$\alpha(\psi^*, \phi^*) = \alpha, \quad \text{у} \quad \beta(\psi^*, \phi^*) = \beta. \tag{22}$$

Тогда для любого критерия (ψ, ϕ)*, удовлетворяющего условиям*

$$\alpha(\psi, \phi) \leq \alpha, \quad \text{и} \quad \beta(\psi, \phi) \leq \beta \tag{23}$$

выполнено неравенство

$$C(\psi) \geq C(\psi^*). \tag{24}$$

Неравенство в (24) строгое, если по крайней мере одно из неравенств в (23) строгое.

Замечание 1. Очевидно, что условие (22) гарантирует принадлежность $(\psi^*, \phi^*) \in \Delta(\alpha, \beta)$, так что условие (24) представляет собой условие оптимальности (ψ^*, ϕ^*) в классе $\Delta(\alpha, \beta)$ с точки зрения средней стоимости эксперимента.

Доказательство. Пусть для (ψ^*, ϕ^*) выполнены условия (21) и (22). Тогда для любого другого критерия (ψ, ϕ), удовлетворяющего условию (23), имеем

$$C(\psi) + \lambda_0\alpha + \lambda_1\beta \geq C(\psi) + \lambda_0\alpha(\psi, \phi) + \lambda_1\beta(\psi, \phi) \tag{25}$$

$$= L(\psi, \phi) \geq L(\psi^*, \phi^*) = C(\psi^*) + \lambda_0\alpha(\psi^*, \phi^*) + \lambda_1\beta(\psi^*, \phi^*) \tag{26}$$

$$= C(\psi^*) + \lambda_0\alpha + \lambda_1\beta. \tag{27}$$

Сравнивая левую часть цепочки неравенств (25)-(27) с правой частью (27), получаем

$$C(\psi) \geq C(\psi^*),$$

что доказывает первое утверждение теоремы. При этом если теперь положить $C(\psi) = C(\psi^*)$ для некоторого критерия (ψ, ϕ), то *все* неравенства в (25)–(27) в действительности превратятся в равенства, из чего следует $\alpha(\psi, \phi) = \alpha$ *и* $\beta(\psi, \phi) = \beta$, что завершает доказательство. ■

Теорема 1 позволяет свести нашу задачу к задаче минимизации $L(\psi, \phi)$ (см. условие (21)). В критерии (ψ^*, ϕ^*) участвуют две константы (λ_0, λ_1), и это дает надежду на то, что можно добиться выполнения условий (22) подбором этих констант.

В некоторой степени условия (22) выполняются автоматически, так как раз существует критерий (ψ^*, ϕ^*), удовлетворяющий (21), условие (22) выполняется, если взять $\alpha = \alpha(\psi^*, \phi^*)$ и $\beta = \beta(\psi^*, \phi^*)$. В действительности, это именно тот смысл, в котором понимается оптимальность SPRT (см. [1],[2],[8],[4],[6],[11],[3]).

3.3 Сведение к задаче оптимальной остановки

Теорема 1 говорит нам о том, что для нахождения формы оптимального критерия необходимо уметь минимизировать функцию Лагранжа $L(\psi, \phi)$. В этом разделе рассматривается частное решение этой задачи: можно найти решающее правило ϕ^*, оптимальное в том смысле, что для любого другого решающего правила справедливо

$$L(\psi, \phi) \geq L(\psi, \phi^*).$$

Для определения ϕ^* потребуются следующие обозначения.
Пусть

$$f_i^n = f_i^n(X_1, X_2, \ldots, X_n) = \prod_{j=1}^{n} f_{\theta_i}(X_j), \quad i = 0, 1.$$

Будем обозначать I_A индикаторную функцию события A, т. е.,

$$I_A = \begin{cases} 1, & \text{если } A \text{ произошло,} \\ 0, & \text{в противном случае.} \end{cases}$$

В частности, пусть

$$\phi_n^* = I_{\{\lambda_0 f_0^n \leq \lambda_1 f_1^n\}} = \begin{cases} 1, & \text{если } \lambda_0 f_0^n \leq \lambda_1 f_1^n, \\ 0, & \text{в противном случае.} \end{cases} \tag{28}$$

Наконец, пусть $\phi^* = (\phi_1^*, \phi_2^*, \ldots, \phi_n^*, \ldots)$.

Теорема 2. *Для любого правила остановки ψ и любого правила принятия решения ϕ справедливо неравенство*

$$L(\psi, \phi) \geq L(\psi, \phi^*). \tag{29}$$

Значение $L(\psi, \phi^*)$ может быть записано в виде

$$L(\psi, \phi^*) = C(\psi) + \sum_{n=1}^{\infty} \sum \chi_n \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\} \tag{30}$$

где χ_n задается формулой (5).

Доказательство. Из определения функции Лагранжа (20) следует, что достаточно показать выполнение условий

$$\lambda_0 \alpha(\psi, \phi) + \lambda_1 \beta(\psi, \phi) \geq \lambda_0 \alpha(\psi, \phi^*) + \lambda_1 \beta(\psi, \phi^*), \tag{31}$$

и

$$\lambda_0 \alpha(\psi, \phi^*) + \lambda_1 \beta(\psi, \phi^*) = \sum_{n=1}^{\infty} \sum \chi_n \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\} \tag{32}$$

Ключевой для этого доказательства (равно как и для ряда нижеследующих) будет следующая

Лемма 1. *Пусть x – переменная, принимающая конечное число значений и пусть $F(x)$ – любая функция от x. Тогда для любых функций $\phi(x), \chi(x)$, $0 \leq \phi(x), \chi(x) \leq 1$, справедливо неравенство*

$$\sum \chi(x)\phi(x)F(x) \geq \sum \chi(x)F(x)I_{\{F(x) \leq 0\}} \tag{33}$$

Доказательство. Покажем, что

$$\sum \chi(x)\phi(x)F(x) - \sum \chi(x)F(x)I_{\{F(x)\le 0\}} \ge 0, \tag{34}$$

или, что то же самое,

$$\sum \chi(x)F(x)(\phi(x) - I_{\{F(x)\le 0\}}) \ge 0, \tag{35}$$

Для этого заметим, что $F(x)(\phi(x) - I_{\{F(x)\le 0\}}) \ge 0$ для любого x, так как

- если $F(x) \le 0$, то $\phi(x) - I_{\{F(x)\le 0\}} = \phi(x) - 1 \le 0$, и их произведение неотрицательно,
- если $F(x) > 0$, то $\phi(x) - I_{\{F(x)\le 0\}} = \phi(x) \ge 0$, и их произведение также неотрицательно.

Мы показали, что все слагаемые в (35) неотрицательны, из чего следует выполнение (35), (34) и далее (33).■

Вернемся к доказательству (31). Используя формулы (13),(14) и (16), представим левую часть (31) в виде

$$\lambda_0\alpha(\psi,\phi) + \lambda_1\beta(\psi,\phi) = \lambda_1 + \sum_{n=1}^{\infty}\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)\phi_n. \tag{36}$$

Применив лемму 1 к выражению $\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)\phi_n$ в правой части (36), получаем

$$\lambda_0\alpha(\psi,\phi) + \lambda_1\beta(\psi,\phi) \ge \lambda_1 + \sum_{n=1}^{\infty}\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)I_{\{\lambda_0 f_0^n - \lambda_1 f_1^n \le 0\}} \tag{37}$$

$$= \lambda_1 + \sum_{n=1}^{\infty}\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)\phi_n^* = \lambda_0\alpha(\psi,\phi^*) + \lambda_1\beta(\psi,\phi^*). \tag{38}$$

Неравенство (31) доказано. Также мы приблизились к доказательству (32), так как согласно (37)-(38),

$$\lambda_0\alpha(\psi,\phi^*) + \lambda_1\beta(\psi,\phi^*) = \lambda_1 + \sum_{n=1}^{\infty}\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)I_{\{\lambda_0 f_0^n - \lambda_1 f_1^n \le 0\}}. \tag{39}$$

Учитывая условие (17), имеем

$$P_{\theta_1}(\nu < \infty) = \sum_{n=1}^{\infty}\sum \chi_n f_1^n = 1,$$

и правую часть (39) можно записать в виде

$$\lambda_1 \sum_{n=1}^{\infty}\sum \chi_n f_1^n + \sum_{n=1}^{\infty}\sum \chi_n(\lambda_0 f_0^n - \lambda_1 f_1^n)I_{\{\lambda_0 f_0^n - \lambda_1 f_1^n \le 0\}} \tag{40}$$

$$= \sum_{n=1}^{\infty}\sum \chi_n\left[(\lambda_0 f_0^n - \lambda_1 f_1^n)I_{\{\lambda_0 f_0^n - \lambda_1 f_1^n \le 0\}} + \lambda_1 f_1^n\right] \tag{41}$$

$$= \sum_{n=1}^{\infty} \sum \chi_n \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\} \tag{42}$$

из чего следует (32) и далее (30). ■

Введем обозначение

$$L(\psi) = C(\psi) + \sum_{n=1}^{\infty} \sum \chi_n \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\}. \tag{43}$$

По теореме 2 $L(\psi)$ является минимумом $L(\psi, \phi)$ по *всем решающим правилам* ϕ. Таким образом, задача минимизации функции $L(\psi, \phi)$ сведена к задаче оптимальной остановки, так как $L(\psi)$ зависит только от правила остановки.

Теперь наша цель – найти правило остановки ψ^*, такое что

$$L(\psi) \geq L(\psi^*) \tag{44}$$

для любого правила остановки ψ.

После этого задача нахождения оптимального критерия будет решена, так как для любого критерия (ψ, ϕ) по теореме 2

$$L(\psi, \phi) \geq L(\psi, \phi^*) = L(\psi) \geq L(\psi^*) = L(\psi^*, \phi^*),$$

так что оптимальным критерием будет (ψ^*, ϕ^*).

3.4 Задача оптимальной остановки. Усеченный случай

В этом разделе задача оптимальной остановки из предыдущего раздела будет решена для класса усеченных моментов остановки.

Пусть последовательный эксперимент состоит не более чем из N шагов. Иначе говоря, допустимы лишь такие правила остановки ψ, для которых $\psi_N \equiv 1$, а остальные ψ_n произвольны. Тогда единственной варьируемой частью в правилах остановки будут $\psi_1, \psi_2, \ldots, \psi_{N-1}$, что даст возможность исчерпывающего решения задачи.

Для начала запишем функцию $L(\psi)$ в виде более удобном

$$L(\psi) = C(\psi) + \sum_{n=1}^{N} \sum \chi_n \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\}$$

$$= \sum_{n=1}^{N} \sum \chi_n (cnf_0^n + \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\})$$

и более явном

$$L(\psi) = \sum_{n=1}^{N} \sum (1-\psi_1) \ldots (1-\psi_{n-1}) \psi_n (cnf_0^n + \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\}).$$

С учетом того, что $\psi_N \equiv 1$, введя обозначение $l_n = \min\{\lambda_0 f_0^n, \lambda_1 f_1^n\}$, получим

$$L(\psi) = \sum_{n=1}^{N-1} \sum (1-\psi_1) \ldots (1-\psi_{n-1}) \psi_n (cnf_0^n + l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_{N-1})\left(cNf_0^N+l_N\right). \qquad (45)$$

Следующая лемма несет бо́льшую часть технической нагрузки в построении оптимального усеченного правила остановки.

Лемма 2. *Пусть $r\geq 2$, а $v_r=v_r(x_1,x_2,\dots,x_r)$ – некоторая функция. Тогда*

$$\begin{aligned}
&\sum_{n=1}^{r-1}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)\\
&\quad+\sum(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+v_r\right)\\
&\geq\sum_{n=1}^{r-2}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n) \qquad (46)\\
&\quad+\sum(1-\psi_1)\dots(1-\psi_{r-2})\left(c(r-1)f_0^{r-1}+v_{r-1}\right),
\end{aligned}$$

где

$$v_{r-1}=\min\{l_{r-1},cf_0^{r-1}+\sum_{x_r}v_r\}, \qquad (47)$$

а в неравенстве в (46) достигается равенство при

$$\psi_{r-1}=I_{\{l_{r-1}\leq cf_0^{r-1}+\sum_{x_r}v_r\}}. \qquad (48)$$

Доказательство. Для доказательства неравенства (46) достаточно показать

$$\begin{aligned}
&\sum(1-\psi_1)\dots(1-\psi_{r-2})\psi_{r-1}(c(r-1)f_0^{r-1}+l_{r-1})\\
&\quad+\sum(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+v_r\right)\\
&\geq\sum(1-\psi_1)\dots(1-\psi_{r-2})\left(c(r-1)f_0^{r-1}+v_{r-1}\right). \qquad (49)
\end{aligned}$$

Левая часть (49) равна

$$\begin{aligned}
&\sum_{x_1,\dots,x_{r-1}}(1-\psi_1)\dots(1-\psi_{r-2})\psi_{r-1}(c(r-1)f_0^{r-1}+l_{r-1})\\
&\quad+\sum_{x_1,\dots,x_r}(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+v_r\right)\\
&=\sum_{x_1,\dots,x_{r-1}}(1-\psi_1)\dots(1-\psi_{r-2})\Big[\psi_{r-1}(c(r-1)f_0^{r-1}+l_{r-1})\\
&\quad+(1-\psi_{r-1})\sum_{x_r}\left(crf_0^r+v_r\right)\Big] \qquad (50)
\end{aligned}$$

По определению f_0^r

$$\sum_{x_r}crf_o^r=\sum_{x_r}cr\prod_{i=1}^{r}f_{\theta_0}(x_i)=cr\prod_{i=1}^{r-1}f_{\theta_0}(x_i)\sum_{x_r}f_{\theta_0}(x_r)=crf_0^{r-1},$$

по (1), и правая часть (50) принимает вид

$$\sum_{x_1,\ldots,x_{r-1}} (1-\psi_1)\ldots(1-\psi_{r-2})[c(r-1)f_0^{r-1}$$

$$+\psi_{r-1}l_{r-1} + (1-\psi_{r-1})(cf_0^{r-1} + \sum_{x_r} v_r)]. \tag{51}$$

Применяя лемму 1, где

$$\begin{aligned} \chi &= (1-\psi_1)\ldots(1-\psi_{r-2}), \\ \phi &= \psi_{r-1} \\ F &= l_{r-1} - (cf_0^{r-1} + \sum_{x_r} v_r), \end{aligned}$$

получаем, что (51) больше или равно

$$\begin{aligned} \sum_{x_1,\ldots,x_{r-1}} & (1-\psi_1)\ldots(1-\psi_{r-2})[c(r-1)f_0^{r-1} + (cf_0^{r-1} + \sum_{x_r} v_r) \\ & +(l_{r-1} - (cf_0^{r-1} + \sum_{x_r} v_r))I_{\{l_{r-1} \le cf_0^{r-1} + \sum_{x_r} v_r\}}] \\ = & \sum_{x_1,\ldots,x_{r-1}} (1-\psi_1)\ldots(1-\psi_{r-2})[c(r-1)f_0^{r-1} \\ & + \min\{l_{r-1}, cf_0^{r-1} + \sum_{x_r} v_r\}], \end{aligned} \tag{52}$$

что с учетом определения (47) равно правой части (49).

При этом по той же лемме 1 (51) равно (52) при

$$\psi_{r-1} = I_{\{l_{r-1} \le cf_0^{r-1} + \sum_{x_r} v_r\}}.$$

■

Лемму 2 можно непосредственно применить к функции $L(\psi)$ (см. (45) с $v_N \equiv l_N$). Обозначим $V_N^N = v_N$. По лемме 2

$$L(\psi) \ge \sum_{n=1}^{N-2} \sum (1-\psi_1)\ldots(1-\psi_{n-1})\psi_n(cnf_0^n + l_n)$$

$$+\sum (1-\psi_1)\ldots(1-\psi_{N-2})\left(c(N-1)f_0^{N-1} + v_{N-1}\right), \tag{53}$$

где $v_{N-1} = \min\{l_{N-1}, cf^{N-1} + \sum_{x_N} v_N\}$. Пусть также $V_{N-1}^N = v_{N-1}$. По лемме 2 в равенстве в (53) достигается равенство при

$$\psi_{N-1} = I_{\{l_{N-1} \le cf_0^{N-1} + \sum_{x_N} V_N^N\}}. \tag{54}$$

Вновь применяя лемму 2 в правой части (53), получаем

$$L(\psi) \ge \sum_{n=1}^{N-3} \sum (1-\psi_1)\ldots(1-\psi_{n-1})\psi_n(cnf_0^n + l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_{N-3})\left(c(N-2)f_0^{N-2}+v_{N-2}\right), \qquad (55)$$

где $v_{N-2}=\min\{l_{N-2}, cf^{N-2}+\sum_{x_{N-1}}v_{N-1}\}$. Положим $V_{N-2}^N=v_{N-2}$. Равенство в (55) достигается, если выполняется (54) и

$$\psi_{N-2}=I_{\{l_{N-2}\le cf_0^{N-2}+\sum_{x_{N-1}}V_{N-1}^N\}}. \qquad (56)$$

Вновь применяя лемму 2 в правой части (55), получим новую, более нижнюю границу для $L(\psi)$ и новое ψ_{N-3} при котором она достигается. Будем продолжать процесс, пока не получим границу

$$L(\psi)\ge\sum\left(cf_0^1+v_1\right), \qquad (57)$$

и условия на ψ, начиная с (54), (56), и т. д., достаточные для достижения этой границы. Получится, что эти условия задают оптимальное правило остановки.

Таким образом, доказана следующая

Теорема 3. *Пусть ψ – усеченное правило остановки ($\psi_N\equiv 1$). Тогда для любого $1\le r\le N-1$ выполняются следующие неравенства:*

$$\begin{aligned} L(\psi) &\ge \sum_{n=1}^{r}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)\\ &+\sum(1-\psi_1)\dots(1-\psi_r)\left(c(r+1)f_0^{r+1}+V_{r+1}^N\right) \qquad (58)\\ &\ge\sum_{n=1}^{r-1}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)\\ &+\sum(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+V_r^N\right), \qquad (59)\end{aligned}$$

где $V_N^N=l_N$ и для любого $k<N$

$$V_k^N=\min\{l_k, cf_0^k+\sum_{x_{k+1}}V_{k+1}^N\}. \qquad (60)$$

Нижняя граница в (59) достигается при

$$\begin{aligned}\psi_{N-1} &= I_{\{l_{N-1}\le cf_0^{N-1}+\sum\limits_{x_N}V_N^N\}}\\ \psi_{N-2} &= I_{\{l_{N-2}\le cf_0^{N-2}+\sum\limits_{x_{N-1}}V_{N-1}^N\}} \qquad (61)\\ &\dots\\ \psi_r &= I_{\{l_r\le cf_0^r+\sum\limits_{x_{r+1}}V_{r+1}^N\}}.\end{aligned}$$

В частности, для $r=1$ формулы (61) исчерпывающим образом определяют оптимальное усеченное правило остановки.

3.5 Задача оптимальной остановки. Неусеченный случай

Этот раздел посвящен нахождению вида оптимального правила остановки в классе всех последовательных критериев. Изложение будет основано на результатах предыдущего раздела.

Прежде всего, для любого правила остановки ψ определим

$$L_N(\psi) = \sum_{n=1}^{N-1} \sum (1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n + l_n)$$

$$+ \sum (1-\psi_1)\dots(1-\psi_{N-1})\left(cNf_0^N + l_N\right), \tag{62}$$

функцию Лагранжа, соответствующую правилу остановки ψ, усеченному до шага N (ср. с (45)), т. е. правилу остановки с компонентами $(\psi_1, \psi_2, \dots, \psi_{N-1}, 1, \dots)$. Это усеченное правило остановки; для него справедливы результаты предыдущего раздела, в частности, неравенства из теоремы 3. Идея нижеследующего заключается в переходе к пределу в этих неравенствах при $N \to \infty$ для получения нижних границ $L(\psi)$.

Возникает первый вопрос: что происходит с $L_N(\psi)$, когда $N \to \infty$?

Лемма 3. *Для любого правила остановки* ψ

$$\lim_{N\to\infty} L_N(\psi) = L(\psi)$$

Доказательство. Пусть $L(\psi) < \infty$; случай $L(\psi) = \infty$ рассмотрим в конце доказательства. Вычислим разницу между $L(\psi)$ и $L_N(\psi)$, чтобы доказать ее стремление к нулю при $N \to \infty$. По (62)

$$L(\psi) - L_N(\psi) = \sum_{n=1}^{\infty} \sum \chi_n(cnf_0^n + l_n)$$

$$- \sum_{n=1}^{N-1} \sum \chi_n(cnf_0^n + l_n) - \sum (1-\psi_1)\dots(1-\psi_{N-1})\left(cNf_0^N + l_N\right)$$

$$= \sum_{n=N}^{\infty} \sum \chi_n(cnf_0^n + l_n) - \sum (1-\psi_1)\dots(1-\psi_{N-1})\left(cNf_0^N + l_N\right). \tag{63}$$

Первое слагаемое стремится к нулю при $N \to \infty$ как «хвост» сходящегося ряда ($L(\psi) < \infty$).

Также

$$\sum (1-\psi_1)\dots(1-\psi_{N-1})l_N \le \lambda_0 \sum (1-\psi_1)\dots(1-\psi_{N-1})f_0^N$$

$$= P_{\theta_0}(\nu > N-1) = \sum_{n=N}^{\infty} P_{\theta_0}(\nu = n) \to 0$$

при $N \to \infty$, так как $P_{\theta_0}(\nu < \infty) = \sum_{n=1}^{\infty} P_{\theta_0}(\nu = n) = 1$ сходится.

Остается показать, что

$$\sum (1-\psi_1)\dots(1-\psi_{N-1})Nf_0^N = NP_{\theta_0}(\nu \ge N) \to 0 \text{ cuando } N \to \infty. \tag{64}$$

Это также следует из условия леммы 3, гарантирующей, что $L(\psi) < \infty$, вследствие чего $E_{\theta_0}\nu = \sum_{n=1}^{\infty} nP_{\theta_0}(\nu = n) < \infty$.

Из сходмости ряда $\sum_{n=N}^{\infty} nP_{\theta_0}(\nu = n) \to 0$ и неравенства Чебышева следует

$$NP_{\theta_0}(\nu \geq N) \leq E_{\theta_0}\nu I_{\{\nu\geq N\}} = \sum_{n=N}^{\infty} nP_{\theta_0}(\nu = n) \to 0$$

при $N \to \infty$, что доказывает (64).

Пусть теперь $L(\psi) = \infty$. Это означает, что

$$\sum_{n=1}^{\infty} \sum \chi_n(cnf_0^n + l_n) = \infty,$$

и поэтому

$$L_N(\psi) \geq \sum_{n=1}^{N-1} \sum \chi_n(cnf_0^n + l_n) \to \infty.$$

■

Следующий вопрос касается поведения функций V_r^N из неравенств теоремы 3 при $N \to \infty$.

Лемма 4. Для любого $r \geq 1$, для любого $N \geq r$

$$V_r^N \geq V_r^{N+1}. \tag{65}$$

Доказательство по индукции по $r = N, N-1, \dots, 1$.

Пусть $r = N$. Тогда по (60)

$$V_N^{N+1} = \min\{l_N, cf_0^N + \sum_{x_{N+1}} V_{N+1}^{N+1}\} \leq l_N = V_N^N.$$

В предположении, что (65) справедливо для любого r, $N \geq r > 1$,

$$V_{r-1}^N = \min\{l_{r-1}, cf_0^{r-1} + \sum_{x_r} V_r^N\}$$

$$\geq \min\{l_{r-1}, cf_0^{r-1} + \sum_{x_r} V_r^{N+1}\} = V_{r-1}^{N+1}$$

т. к. (65) выполняется для $r-1$, что завершает индукцию. ■

Из леммы 4 следует, что для любого $r \geq 1$ последовательность V_r^N убывает. Следовательно, существует предел такой последовательности

$$V_r = \lim_{N\to\infty} V_r^N. \tag{66}$$

Переходя к пределу при $N \to \infty$ во всех равенствах и неравенствах теоремы 3, получаем

Теорема 4. *Пусть ψ – произвольное правило остановки. Тогда для любого $r \geq 1$ выполняются следующие неравенства:*

$$L(\psi) \geq \sum_{n=1}^{r} \sum (1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n + l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_r)\left(c(r+1)f_0^{r+1}+V_{r+1}\right) \tag{67}$$

$$\geq \sum_{n=1}^{r-1}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+V_r\right). \tag{68}$$

где

$$V_k=\min\{l_k, cf_0^k+\sum_{x_{k+1}}V_{k+1}\} \tag{69}$$

для любого $k\geq 1$.

В частности, при $r=1$ *получается следующая нижняя граница:*

$$L(\psi)\geq c+\sum V_1. \tag{70}$$

В формулировке теоремы 4 отсутствует один важный элемент, а именно, вид правила остановки, при котором достигается нижняя граница, то есть оптимального правила остановки. Неформально переходя к пределу при $N\to\infty$ в уравнениях (61) (начиная с последнего) получаем вид оптимального правила остановки:

$$\psi_r^*=I_{\{l_r\leq cf_0^r+\sum\limits_{x_{r+1}}V_{r+1}\}},\ r=1,2,3,\dots. \tag{71}$$

Основная цель данного раздела – показать, что правило (71) действительно оптимальное.

Теорема 5. *Для любого последовательного критерия* ψ

$$L(\psi)\geq L(\psi^*)=c+\sum V_1. \tag{72}$$

Доказательство. Пусть ψ – правило остановки. По теореме 4 для любого фиксированного $r\geq 1$ выполняются неравенства:

$$L(\psi)\geq \sum_{n=1}^{r}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_r)\left(c(r+1)f_0^{r+1}+V_{r+1}\right) \tag{73}$$

$$\geq \sum_{n=1}^{r-1}\sum(1-\psi_1)\dots(1-\psi_{n-1})\psi_n(cnf_0^n+l_n)$$

$$+\sum(1-\psi_1)\dots(1-\psi_{r-1})\left(crf_0^r+V_r\right). \tag{74}$$

$$\geq\dots$$

$$\geq\sum\psi_1(cf_0^1+l_1)+\sum(1-\psi_1)\left(c2f_0^2+V_2\right) \tag{75}$$

$$\geq\sum\left(cf_0^1+V_1\right). \tag{76}$$

Применяя лемму 2 и учитывая (69), легко видеть, что во всех неравенствах, за исключением первого, будут равенства при

$$(\psi_1^*, \psi_2^*, \dots, \psi_r^*, 1, \dots).$$

Из этого, в частности, следует, существование предела

$$\lim_{r\to\infty}[\sum_{n=1}^{r}\sum(1-\psi_1^*)\dots(1-\psi_{n-1}^*)\psi_n^*(cnf_0^n+l_n)$$

$$+\sum(1-\psi_1^*)\dots(1-\psi_r^*)\left(c(r+1)f_0^{r+1}+V_{r+1}\right)]=\sum\left(cf_0^1+V_1\right). \qquad (77)$$

Из этого непосредственно следует существование предела

$$\lim_{r\to\infty}\sum_{n=1}^{r}\sum[(1-\psi_1^*)\dots(1-\psi_{n-1}^*)\psi_n^*(cnf_0^n+l_n)\le\sum\left(cf_0^1+V_1\right). \qquad (78)$$

В левой части (78) стоит $L(\psi^*)$, поэтому

$$L(\psi^*)\le\sum\left(cf_0^1+V_1\right). \qquad (79)$$

С другой стороны, с учетом (73)-(76)

$$L(\psi^*)\ge\sum\left(cf_0^1+V_1\right),$$

что, вместе с (79), дает

$$L(\psi^*)=\sum\left(cf_0^1+V_1\right),$$

и из (73)-(76) следует оптимально правила остановки ψ^*. ■

3.6 Последовательный критерий отношения вероятностей (SPRT)

В этом разделе будет показано, что оптимальный критерий из предыдущего раздела эквивалентен последовательному критерию отношения вероятностей (*Sequential Probability Ratio Test, SPRT*).

Чтобы избежать несущественных технических сложностей, предположим для всех x

$$f_{\theta_0}(x)>0, \quad \text{у} \quad f_{\theta_1}(x)>0.$$

Определим величину

$$Z_n=\prod_{i=1}^{n}\frac{f_{\theta_1}(x_i)}{f_{\theta_0}(x_i)}=\frac{f_1^n}{f_0^n} \qquad (80)$$

называемую *отношением вероятностей*, $n=1,2,\dots$

Чтобы выразить элементы оптимального правила остановки в терминах Z_n, на основе функций V_r^N введем новые функции U_r^N:

$$U_r^N=V_r^N/f_0^r, \qquad (81)$$

Несложно видеть, что все функции U_r^N зависят от наблюдений $x_1, x_2, \ldots, x_r$ только через Z_r. Действительно,

$$U_N^N = U_N^N(Z_N) = \min\{\lambda_0, \lambda_1 Z_N\} = g(Z_N) \tag{82}$$

(пусть по определению $g(z) = \min\{\lambda_0, \lambda_1 z\}$), и для $r < N$

$$U_r^N = U_r^N(Z_r) = \min\left\{g(Z_r), c + \sum_{x_{r+1}} f_{\theta_0}(x_{r+1}) U_{r+1}^N\left(Z_r \frac{f_{\theta_1}(x_{r+1})}{f_{\theta_0}(x_{r+1})}\right)\right\} \tag{83}$$

В частности, из (82)-(83) видно, что все функции U_r^N могут быть получены из следующей последовательности функций $\rho_n = \rho_n(z)$: положим

$$\rho_0(z) = g(z), \tag{84}$$

и далее рекуррентно по $n = 1, 2, \ldots$

$$\rho_n(z) = \min\left\{g(z),\ c + \sum_x f_{\theta_0}(x)\rho_{n-1}\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)\right\}. \tag{85}$$

В такой записи

$$U_N^N = \rho_0(Z_N),\ U_{N-1}^N = \rho_1(Z_{N-1}), \ldots, U_r^N = \rho_{N-r}(Z_r). \tag{86}$$

Результаты предыдущего раздела дают нам существование

$$\lim_{N\to\infty} V_r^N = V_r,$$

из чего, вследствие (81), следует существование

$$\lim_{N\to\infty} U_r^N = U_r = V_r / f_0^r.$$

По (86),

$$U_r = \lim_{N\to\infty} \rho_{N-r}(Z_r) = \rho(Z_r), \tag{87}$$

где по определению

$$\rho(z) = \lim_{n\to\infty} \rho_n(z).$$

Переходя к пределу при $n \to \infty$ в (85) получаем следующее неравенство для функции ρ:

$$\rho(z) = \min\left\{g(z),\ c + \sum_x f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)\right\}. \tag{88}$$

Назовем это неравенство *уравнением динамического программирования.*

Теперь можно выразить оптимальное правило остановки из теоремы 5 через отношение вероятностей Z_n. Выражая V_{r+1} и прочие элементы формулы (71) через Z_r, имеем

$$\psi_r^* = I_{\left\{g(Z_r) \le c + \sum_x f_{\theta_0}(x)\rho\left(Z_r \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right)\right\}}. \tag{89}$$

Формула (89) уже более явная, чем (71). Но ее можно записать еще в более конкретом виде, показав, что неравенство, задающее момент остановки (89):

$$g(z) \leq c + \sum_x f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right) \tag{90}$$

равносильно простому условию

$$z \notin (A, B) \tag{91}$$

с некоторыми константами A и B, такими, что $0 < A < B < \infty$, из чего (89) приобретает вид

$$\psi_r^* = I_{\{Z_r \notin (A,B)\}}.$$

Оставшаяся часть раздела будет посвящена техническим деталям доказательства эквивалентности между (90) и (91).

Изучим свойства функций $\rho_n(z)$ и их предела $\rho(z)$.

Лемма 5. *Функции $\rho_n(z)$, задаваемые как (84)-(84), и функция $\rho(z)$ обладают следующими свойствами:*

i) вогнутость и непрерывность на $[0, \infty)$,

ii) неубывание на $[0, \infty)$,

iii) $\rho_n(0) = \rho(0) = 0$, $\rho_n(\infty) = \rho(\infty) = \lambda_0$,

Доказательство.

i) Докажем по индукции, что ρ_n вогнуты.

Для n=0 функция $\rho_0(z) = g(z)$ вогнута как минимум двух вогнутых функций (в нашем случае, линейных).

Предположим вогнутость $\rho_n(z)$. Тогда по (85) функция ρ_{n+1} – минимум между вогнутой функцией и суммой вогнутых функций. Следовательно, ρ_{n+1} также вогнута.

Поточечный придел последовательности вогнутых функций сам является вогнутой функцией, так что функция $\rho(z) = \lim_{n\to\infty} \rho_n(z)$ также вогнута.

Из того, что $\rho(z) \geq 0$ (как предел неотрицательных функций), а $\rho(z)$ вогнута, следует, что функция ρ непрерывна на $(0, \infty)$ (см., к примеру, раздел 3.18 [5]). При этом она непрерывна в точке 0, так как $\rho(z) \leq g(z) \to 0 = \rho(0)$, $z \to 0$, и следовательно, $\rho(z)$ непрерывна на $[0, \infty)$.

ii) Также докажем по индукции. Для $n = 0$ – очевидно.

Если $r_n(z)$ не убывает, то по (85) функция ρ_{n+1} также не убывает.

Последовательный предел неубывающих функций сам является неубывающей функцией, так что функция $\rho(z)$ также не убывает.

iii) Свойство $\rho_n(0) = 0$ легко доказывается по индукции. Поэтому $\rho(0) = \lim_{n\to\infty} \rho_n(0) = 0$.

Также по индукции доказывается $\rho_n(\infty) = \lambda_0$.

По *ii)* предел $\lim_{z\to\infty} \rho(z) = \lambda$ существует. Переходя к пределу при $z \to \infty$ в (88), получаем $\lambda = \min\{\lambda_0, c + \lambda\}$, что дает $\lambda = \lambda_0$

■

Пусть
$$h(z)=\sum_x f_{\theta_0}(x)\rho\left(z\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}\right).$$

Определим
$$A=\sup\{z:0\le z\le\lambda_0/\lambda_1,\ c+h(z)\ge g(z)\}$$
$$B=\inf\{z:\lambda_0/\lambda_1\le z,\ c+h(z)\ge g(z)\}$$

Из того, что $c+h(0)=c>g(0)=0$, $c+h(\infty)=c+\lambda_0>g(\infty)=\lambda_0$ следует то, что $A>0$ и $B<\infty$.

С другой стороны, если $A=B=\lambda_0/\lambda_1$, то $c+h(z)\ge g(z)$ для любого $z\ge 0$, следовательно, оптимальным моментом остановки будет $\nu=1$.

В противном случае, $A<B$. Будем предполагать именно это.

Лемма 6. *$A<z<B$ если и только если $g(z)>c+h(z)$.*

Доказательство. Из непрерывности всех функций $g(A)=c+h(A)$ и $g(B)=c+h(B)$. Кроме того, по определению $c+h(z)<g(z)$ для $z\in(A,\lambda_0/\lambda_1]$. Также по определению $c+h(z)<g(z)$ для $z\in[\lambda_0/\lambda_1,B)$. Теперь, если $z\in[0,A)$, то существует $1\ge a>0$, такое, что $z=(1-a)A$. Так как $h(0)=0$, из вогнутости $h(z)$ получаем
$$h((1-a)A)\ge(1-a)h(A)=(1-a)(g(A)-c).$$

Таким образом,
$$c+h(z)\ge c+(1-a)(g(A)-c)=ac+(a-1)g(A)>(1-a)g(A)=g((1-a)A)=g(z).$$

то есть, $c+h(z)>g(z)$.

Аналогичным образом доказывается, что если $z\in(B,\infty)$, то $c+h(z)>g(z)$.■

Оптимальное решающее правило всегда имеет вид
$$\phi_n^*=I_{\{\lambda_0 f_0^n\le\lambda_1 f_1^n\}},$$

что эквивалентно
$$\phi_n^*=I_{\{\lambda_0/\lambda_1\le Z_n\}}.$$

Так как $B\ge\lambda_0/\lambda_1\ge A$, момент остановки оптимального критерия равен
$$\nu=\inf\{n\ge1:Z_n\notin(A,B)\},$$

а H_0 отклоняется при $Z_\nu\ge B$ и принимается при $Z_\nu\le A$.

Все вышесказанное можно сформулировать в виде

Теорема 6. *Критерий (ψ^*,ϕ^*) с ψ^* из теоремы 5 является последовательным критерием отношения вероятностей (SPRT) с моментом остановки*
$$\nu^*=\inf\{n\ge1:Z_n\notin(A,B)\};\tag{92}$$
H_0 отклоняется, если $Z_\nu\ge B$, H_0 принимается, если $Z_\nu\le A$.

Очевидно, что SPRT-критерий из теоремы 6 оптимален, если A и B определяются из уравнений
$$g(A)=c+\sum_x f_{\theta_0}\rho(A\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}),\quad A\le\lambda_0/\lambda_1,\tag{93}$$

$$g(B) = c + \sum_x f_{\theta_0} \rho(B \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)}) \quad B \geq \lambda_0/\lambda_1, \tag{94}$$

(см. определение A и B и доказательство леммы 6).

Однако только случай $1 \in (A, B)$ имеет практический смысл, так как иначе получается, что можно получить результат даже лучший вообще без проведения наблюдений.

Действительно, из $1 \notin (A, B)$ по лемме 6 следует

$$g(1) \leq c + \sum_x f_{\theta_0} \rho(\frac{f_{\theta_1}}{f_{\theta_0}}), \tag{95}$$

где правая часть (95) – абсолютный минимум функции $L(\psi)$ по всем моментам остановки, проводящим *как минимум одно* наблюдение (теорема 5). Но $g(1) = \min\{\lambda_0, \lambda_1\}$ – это, так сказать, «функция $L(\psi_0)$» критерия, не проводящего ни одного наблюдения. Естественно, если наблюдений не проводится, то стоимость эксперимента $C(\psi_0)$ равна нулю. При применении решающего правила

$$\phi_0 = I_{\{\lambda_0 \leq \lambda_1\}},$$

получаем

$$\alpha(\psi_0, \phi_0) = \begin{cases} 1, & \text{если } \lambda_0 \leq \lambda_1 \\ 0 & \text{если } \lambda_0 > \lambda_1 \end{cases}$$

$$\beta(\psi_0, \phi_0) = \begin{cases} 0, & \text{если } \lambda_0 \leq \lambda_1 \\ 1 & \text{если } \lambda_0 > \lambda_1 \end{cases},$$

из чего

$$L(\psi_0, \phi_0) = C(\psi_0) + \lambda_0 \alpha(\psi_0, \phi_0) + \lambda_1 \beta(\psi_0, \phi_0) = \min\{\lambda_0, \lambda_1\} = g(1).$$

Приходим к выводу, что если выполнено (95), то существует тривиальный критерий (ψ_0, ϕ_0), результат которого даже лучше, чем у критерия с как минимум одним наблюдением. И ничего, что мы боролись за оптимальную процедуру, производящую как минимум одно наблюдение, при том, что лучше получается вообще без наблюдений! Значит, последовательный критерий из теоремы 6 имеет практический смысл лишь в случае $A < 1 < B$.

«Предельный» случай $1 = A < B$ или $A < B = 1$ из рассмотрения исключать не нужно, по крайней мере, в теории, так как в этом случае тривиальный критерий совпадает с SPRT.

Вопрос определения констант A и B для SPRT из теоремы 6 сводится к решению уравнений (93) и (94), так что в явном виде их нет.

Однако представляется весьма правдоподобным, что *для любых* констант A и B, таких, что $0 < A \leq 1 \leq B < \infty$, $A < B$, существуют константы c, λ_0 и λ_1, при которых $L(\psi)$ минимизируется SPRT-критерием (92), поскольку требуется выполнение двух условий: (93) и (94) при трех переменных: c, λ_0 и λ_1.

Строгое доказательство этого факта будет чрезмерно технически насыщенным для такого курса, как этот, а ничего нового вероятностного или статистического оно не привнесет, так что оставим его для более подходящего случая. Как бы то ни было, приведем теорему, формализующую вышесказанное.

Теорема 7. *Пусть A и B – две константы, такие, что $0 < A \leq 1 \leq B < \infty$, $A < B$, и пусть (ψ^*, ϕ^*):*

$$\psi_n^* = I_{\{Z_n \notin (A,B)\}},$$

$$\phi_n^* = I_{\{Z_n \geq B\}},$$

критерий SPRT.

Тогда существуют $c > 0$, $\lambda_0 > 0$ и $\lambda_1 > 0$, такие, что для любого последовательного критерия (ψ, ϕ) с $\nu = \inf\{n \geq 0 : \psi_n = 1\}$

$$cE_{\theta_0}\nu + \lambda_0\alpha(\psi,\phi) + \lambda_1\beta(\psi,\phi) \geq cE_{\theta_0}\nu^* + \lambda_0\alpha(\psi^*,\phi^*) + \lambda_1\beta(\psi^*,\phi^*), \quad (96)$$

то есть ψ^* оптимально в смысле теоремы 5.

В сочетании с теоремой 1 немедленно получим

Теорема 8. *Пусть A и B – две константы, такие, что $0 < A \leq 1 \leq B < \infty$ и $A < B$, и пусть (ψ^*, ϕ^*):*

$$\psi_n^* = I_{\{Z_n \notin (A,B)\}},$$

$$\phi_n^* = I_{\{Z_n \geq B\}},$$

критерий SPRT с этими константами.

Тогда для любого критерия (ψ, ϕ) (с моментом остановки $\nu = \inf\{n \geq 1 : \psi_n = 1\}$), такого, что

$$\alpha(\psi,\phi) \leq \alpha(\psi^*,\phi^*) \quad \text{и} \quad \beta(\psi,\phi) \leq \beta(\psi^*,\phi^*) \quad (97)$$

справедливо

$$E_{\theta_0}\nu \geq E_{\theta_0}\nu^*. \quad (98)$$

Неравенство в (98) строгое, если строгое по меньшей мере одно из неравенств (97) строгое.

Если известно, что A и B удовлетворяют (93) и (94), оптимальность SPRT, установленная теоремой 8, полностью доказана.

4 Примеры построения SPRT-критерия

4.1 Распределение Бернулли

Распределение Бернулли задается функцией «плотности»

$$f_p(x) = P(X = x) = \begin{cases} p & \text{, если } x = 1 \\ 1 - p & \text{, если } x = 0, \end{cases}$$

или, что то же самое,

$$f_p(x) = p^x (1-p)^{1-x},$$

где $x \in \{0, 1\}$, $p \in [0, 1]$.

Рассмотрим задачу проверки простой гипотезы $H_0 : p = p_0$ при простой альтернативе $H_1 : p = p_1$, где $p_1 > p_0$.

Вычислим статистику Z_n:

$$Z_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{f_1(x_i)}{f_0(x_i)} = \frac{p_1^{\sum x_i}}{p_0^{\sum x_i}} \frac{(1-p_1)^{n-\sum x_i}}{(1-p_0)^{n-\sum x_i}}$$

$$Z_n(x_1, \ldots, x_n) = \exp\left\{ \left(\ln \frac{p_1}{p_0} - \ln \frac{1-p_1}{1-p_0} \right) \sum x_i + n \ln \frac{1-p_1}{1-p_0} \right\}$$

Последовательный критерий отношения вероятностей основан на статистике Z_n и предписывает

- продолжать эксперимент при $A < Z_n < B$,
- отклонить H_0 при $Z_n > B$.

Для удобства вычислений выражение Z_n можно заменить его логарифмом.

4.2 Распределение Пуассона

Дискретное распределение Пуассона задается функцией «плотности»

$$f_\lambda(x) = \frac{\lambda^x}{x!} e^{-\lambda},$$

где $x \in \{1, 2, \ldots\}$, $\lambda > 0$. Рассмотрим задачу проверки простой гипотезы $H_0 : \lambda = \lambda_0$ при простой альтернативе $H_1 : \lambda = \lambda_1$, $\lambda_1 > \lambda_0$.

Вычислим статистику Z_n:

$$Z_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{f_1(x_i)}{f_0(x_i)} = \exp\left((\ln \lambda_1 - \ln \lambda_0) \sum_{i=1}^{n} x_i - n(\lambda_1 - \lambda_0) \right)$$

Как известно, критерий с фиксированным объемом выборки для проверки гипотезы $H_0 : \lambda = \lambda_0$ при альтернативе $H_1 : \lambda = \lambda_1$ отвергает нулевую гипотезу при

$$Z_n > C.$$

Последовательный критерий отношения вероятностей также основан на статистике Z_n и предписывает

- продолжать эксперимент при $A < Z_n < B$,
- отклонить H_0 при $Z_n > B$.

Таким образом, форма критерия с фиксированным объемом выборки –

$$\sum_{i=1}^{n} X_i > C,$$

а форма последовательного критерия –

$$A < (\ln \lambda_1 - \ln \lambda_0) \sum_{i=1}^{n} X_i - n(\lambda_1 - \lambda_0) < B.$$

Для критерия с фиксированным объемом выборки функция мощности равна вероятности отклонить гипотезу при данном значении параметра λ. Для SPRT функция мощности вычислится как вероятность продолжать наблюдения до некоторого шага, а затем отклонить гипотезу, при данном значении параметра λ.

Напишем функции, вычисляющие значения функций мощности критериев с фиксированным объемом выборки и SPRT-критерия, на языке R. Для критерия с фиксированным объемом выборки также несложно написать функцию, численно вычисляющую критическую константу C по α и n.

```
# Функция мощности критерия с фиксированным объемом выборки
# n -- объем выборки
# lambda -- истинный параметр распределения Пуассона
# C -- критическая константа
fixed.pois.power <- function(n, lambda, C, maxiters=10000)
{
  rejcount <- 0.
  for(ri in 1:maxiters)      # Повторяем несколько тысяч раз
  {
    x <- rpois(n, lambda)    # Генерируем случайную выборку из распределения Пуассона
    if(sum(x) > C)           # Нулевая гипотеза отвергается
      rejcount <- rejcount+1 # Подсчитываем, сколько раз H0 отклонена
  }
  return(rejcount/maxiters)  # Возвращаем вероятность отклонить гипотезу
}

# Функция мощности SPRT-критерия
# lambda -- истинный параметр распределения Пуассона
# lambda0 -- параметр распределения Пуассона при нулевой гипотезе
# lambda1 -- параметр распределения Пуассона при альтернативе
# A, B -- границы интервала продолжения эксперимента
sprt.pois.power <- function(lambda, lambda0, lambda1, A, B, maxiters=10000)
{
  rejcount <- 0.
  sumn <- 0.
  for(ri in 1:maxiters)         # Повторяем несколько тысяч раз
  {
    x <- c()
```

```
    n <- 0.
    repeat                      # Последовательный эксперимент
    {
      n <- n+1.
      x[n] <- rpois(1, lambda)  # Производим одно новое наблюдение
      Z <- n*(lambda0-lambda1) + (log(lambda1)-log(lambda0))*sum(x)
      if(Z > B)                 # Отклонить H0
        rejcount <- rejcount+1  # Подсчитываем, сколько раз H0 отклонена
      if(Z < A || Z > B)        # Прекратить эксперимент
        break
    }
    sumn <- sumn + n            # Общее количество проведенных экспериментов
  }
  return(c(pow=rejcount/maxiters, n=sumn/maxiters))
  # Возвращает значение функции мощности и средний объем выборки
}

# Критическая константа критерия с фиксированным объемом выборки
# alpha -- вероятноть ошибки первого рода
# n -- объем выборки
# lambda0 -- значение параметра распределения Пуассона при нулевой гипотезе
fixed.pois.C <- function(alpha, n, lambda0, interval = c(-200, 200))
{
  f <- function(C) fixed.pois.power(n, lambda0, C) - alpha
  return(uniroot(f, interval))  # Численно находим корень выражения f
}
```

Рассмотрим задачу проверки гипотезы $H_0 : \lambda = 1$ при альтернативе $H_1 : \lambda = 1.5$. Зададим вероятности ошибки первого и второго рода 0.05. Построим критерий SPRT и критерий с фиксированным объемом выборки, удовлетворяющие указанным ограничениям на вероятности ошибок.

Для SPRT обеспечим выполнение ограничений на вероятности ошибок первого и второго рода подбором значений констант A и B, Получаем $A = -2.7$, $B = 2.55$: Удостоверимся в том, что условия выполняются:

```
# Вероятность ошибки первого рода = 0.05
sprt.pois.power(lambda=1.0, lambda0=1.0, lambda1=1.5, A=-2.7, B=2.55)
# 1 - Вероятность ошибки второго рода = 0.95
sprt.pois.power(lambda=1.5, lambda0=1.0, lambda1=1.5, A=-2.7, B=2.55)
```

Результаты выполнения:

```
   pow       n
 0.0516 27.4306
   pow       n
 0.9423 23.8242
```

Из результатов выполнения функции `sprt.pois.power` следует, что средний объем выборки для SPRT не превосходит $n = 28$.

Если мы построим критерий с фиксированным объемом выборки с тем же $n = 28$, то увидим, что условие на вероятность ошибки второго рода не выполняется:

```
C <- fixed.pois.C(alpha=0.05, n=28, lambda0=1.0)$root
# Вероятность ошибки первого рода = 0.05
fixed.pois.power(n=28, lambda=1.0, C)
# 1 - Вероятность ошибки второго рода = 0.80
fixed.pois.power(n=28, lambda=1.5, C)
```

Для выполнения ограничения на вероятности ошибок обоих родов необходимо взять $n = 54$ или больше. Тогда действительно,

```
C <- fixed.pois.C(alpha=0.05, n=54, lambda0=1.0)$root
# Вероятность ошибки первого рода = 0.05
fixed.pois.power(n=54, lambda=1.0, C)
# 1 - Вероятность ошибки второго рода = 0.95
fixed.pois.power(n=54, lambda=1.5, C)
```

Из численного сравнения критерия с фиксированным объемом выборки и последовательного критерия видно, что при равных вероятностях ошибок первого и второго родов число испытаний в последовательном эксперименте меньше (в рассмотренном случае – в 1,9 раза). Таким образом, стоимость проведения экспериментов при последовательной постановке будет существенно меньше при равных ограничениях на вероятности ошибок. В этом состоит выгода от использования последовательного анализа.

Список литературы

[1] Боровков А.А. *Математическая статистика.* М.: Наука, 1986.

[2] Де-Гроот, М. *Оптимальные статистические решения.* М.: Мир, 1974. [DeGroot, M.H. *Optimal Statistical Decisions.* McGraw Hill, New York, 1970.]

[3] Закс Ш. *Теория статистических выводов.* М.: Мир, 1975. [Zacks S. The Theory of Statistical Inference. John Wiley and Sons, New York-London-Sydney-Toronto, 1971.]

[4] Леман Э. *Проверка статистических гипотез.* М.: Наука, 1979. [Lehmann E. *Testing Statistical Hypotheses.* John Wiley & Sons, New York; Chapman & Hall, Ltd., London 1959.]

[5] Харди Г.Г., Литтльвуд Дж.Е., Полиа Г. *Неравенства.* М.: Гос. изд. иностр. лит-ры, 1948. [Hardy G.H., Littlewood J.E., Polya G. *Inequalities.* Cambridge University Press, 1934.]

[6] Ширяев А.Н. *Статистический последовательный анализ. Оптимальные правила остановки.* М.: Наука, 1976.

[7] Chow Y.S., Robbins H., Siegmund D. *Great Expectations: The Theory of Optimal Stopping.* Boston. Houghton Mifflin Company, 1971.

[8] Ferguson, T. S. *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York, 1967.

[9] Ghosh B. K., *Sequential tests of statistical hypotheses.* Addison-Wesley , Reading, Mass.-London-Don Mills, Ont., 1970.

[10] Govindarajulu Z. *Sequential Statistics.* World Scientific Publishing, Singapore, 2004.

[11] Wald A., Wolfowitz J. Optimum character of the sequential probability ratio test. *Ann. Math. Statistics* **19** (1948), pp. 326–339.