

7.2 Последовательный подход в статистической проверке гипотез

В разделе кратко обсуждаются особенности последовательного подхода в статистике, даются основные понятия последовательного анализа и приводятся классические результаты, касающиеся характеристик эффективности последовательных тестов, а также рассматриваются методы, позволяющие строить последовательные тесты для проверки сложных (составных) гипотез.

7.2.1 Особенности последовательного подхода

Большинство разработанных методов статистической проверки гипотез обладают общей особенностью: объем n наблюдаемой выборки $X_n = (x_1, \dots, x_n) \in \mathbb{R}^{nN}$, на основании которой принимается решение о принятии той или другой гипотезы, фиксируется априори. Такая ситуация возникает при обработке ранее полученных (архивных) экспериментальных данных, в так называемых пассивных экспериментах. Однако на практике часто встречаются ситуации, когда проверку гипотез необходимо осуществлять в ходе активного эксперимента и решать вопрос о его продолжении или прекращении.

В 1947 г. американский статистик А. Вальд опубликовал теорию, основанную на том, что априорно фиксировать объем выборки n необязательно. Целесообразно определять необходимый объем выборки в ходе самого эксперимента в зависимости от поступающих наблюдений. В таком подходе объем выборки становится случайной величиной, зависящей от экспериментальных данных: $n = n(X)$. Методы проверки гипотез, основанные на этой идее, получили название *последовательного анализа*. Рассмотрим простейшую модель последовательного анализа Вальда.

Пусть наблюдается случайная последовательность $x_1, x_2, \dots \in \mathbb{R}^N$ независимых в совокупности одинаково распределенных случайных векторов, имеющих п. р. в. $p(x; \theta)$, $x \in \mathbb{R}^N$, где $\theta \in \Theta \in \{\theta_0, \theta_1\}$ – неизвестное истинное значение параметра, принимающего одно из двух различных возможных значений. Определены две простые гипотезы:

$$\mathcal{H}_0 : \theta = \theta_0; \quad \mathcal{H}_1 : \theta = \theta_1.$$

Обозначим через $X_m = (x_1, \dots, x_m) \in \mathbb{R}^{mN}$ выборку, полученную при проведении m экспериментов ($m = 1, 2, \dots$); $d_m \in D = \{0, 1, 2\}$ – возможное решение после проведения m экспериментов.

Решение $d_m = 0$ (1) означает, что по результатам m экспериментов принимается гипотеза \mathcal{H}_0 (\mathcal{H}_1), а $d_m = 2$ – что для вынесения решения об истинности \mathcal{H}_0 или \mathcal{H}_1 с заданной точностью не хватает данных, содержащихся в X_m , и требуется осуществить $(m+1)$ -й эксперимент для принятия следующего решения по $m+1$ наблюдению.

Таким образом, последовательный анализ характеризуется схемой получения наблюдений и принятия решений, приведенной на рис. 7.2.

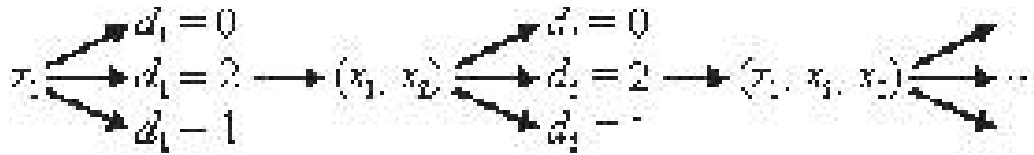


Рис. 7.2: Схема последовательного анализа Вальда

Процесс увеличения числа экспериментов m продолжается до тех пор, пока для некоторого случайного n не будет вынесено решение $d_n \in \{0, 1\}$ об окончании процесса наблюдения и принятии соответствующей гипотезы. Задать последовательное решающее правило (последовательный критерий, последовательный тест) – значит задать последовательность решающих правил:

$$d_m = d_m(X_m), \quad X_m \in \mathbb{R}^{mN}, \quad d_m \in \{0, 1, 2\}, \quad m = 1, 2, \dots$$

Эффективность последовательного критерия характеризуют вероятности ошибок I и II рода:

$$\alpha = P_{\theta_0}\{d_n = 1\}, \quad \beta = P_{\theta_1}\{d_n = 0\},$$

а также условные математические ожидания объемов выборок (средние длительности экспериментов):

$$t_i = E_{\theta_i}\{n\}, \quad i = 0, 1.$$

Принцип оптимальности Вальда состоит в том, что последовательное решающее правило $\{d_m = d_m(X_m)\}$ надлежит выбирать таким образом, чтобы вероятности ошибок I и II рода α и β не превосходили соответственно наперед заданных достаточно малых уровней α_0 и β_0 , и при этом условные средние объемы выборок t_0 , t_1 были бы минимальны среди всех тестов с указанными уровнями вероятностей ошибок.

Этому принципу удовлетворяет последовательный критерий отношения вероятностей Вальда:

$$d_m = d_m(X_m) = \begin{cases} 0, & L_m(X_m) \leq B; \\ 2, & B < L_m(X_m) < A; \\ 1, & L_m(X_m) \geq A, \end{cases} \quad (7.22)$$

где

$$L_m(X_m) = \frac{p(X_m; \theta_1)}{p(X_m; \theta_0)} - \quad (7.23)$$

статистика отношения правдоподобия;

$$p(X_m; \theta_i) = \prod_{j=1}^m p(x_j; \theta_i) - \quad (7.24)$$

плотность распределения случайной выборки X_m при $\theta = \theta_i$ ($i = 0, 1$), $A, B \in \mathbb{R}$ ($B < A$) – некоторые числовые параметры правила, называемые порогами и выбираемые из принципа оптимальности Вальда.

7.2.2 Последовательная статистическая проверка простых гипотез

Дадим строгое определение понятию «последовательный тест» и определим последовательный критерий отношения вероятностей как его специальный случай.

Пусть независимые одинаково распределенные (н. о. р.) случайные величины x_1, x_2, \dots заданы на вероятностном пространстве $(\Omega, \mathcal{F}, P_\theta)$, $x_t \in U \subseteq \mathbb{R}^N$, $t \in \mathbb{N}$, где $\theta \in \Theta = \{0, 1\}$ – ненаблюдаемое значение параметра распределения вероятностей. Имеются две простые гипотезы о значении параметра θ :

$$\mathcal{H}_0 : \theta = 0, \quad \mathcal{H}_1 : \theta = 1. \quad (7.25)$$

Пусть \mathcal{M} – множество марковских моментов τ относительно системы σ -алгебр $\{\mathcal{F}_n\}$: $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma\{\omega : x_1, \dots, x_n\}$, $n \in \mathbb{N}$; \mathcal{D}_τ – семейство \mathcal{F}_τ -измеримых функций, множество значений которых представляет Θ .

Пара функций $\delta = (\tau, d)$, где $\tau \in \mathcal{M}$, $d \in \mathcal{D}_\tau$, называется последовательным тестом (критерием, решающим правилом). Задание первой компоненты τ означает определение функции для вычисления случайного номера завершающего наблюдения, сразу же после которого принимается гипотеза \mathcal{H}_d в соответствии с функцией заключительного решения, заданной второй компонентой. Пусть Δ – множество всех последовательных тестов для проверки гипотез (7.25): $\delta \in \Delta$.

Без потери общности будем полагать, что для распределения вероятностей P_k , $k \in \Theta$, существует некоторая плотность распределения вероятностей (п. р. в.) $p_k(x)$ относительно некоторой меры $\mu(x)$.

Рассмотрим семейство последовательных тестов $\delta_\lambda = (\tau_\lambda, d_\lambda) \in \Delta$, основанных на функции $\lambda(x): U \rightarrow \mathbb{R}$, для которых

$$\tau_\lambda = \inf \{n : \Lambda_n \notin (C_-, C_+)\}, \quad d_\lambda = 1_{[C_+, +\infty)}(\Lambda_n), \quad (7.26)$$

где для $n \in \mathbb{N}$

$$\Lambda_n = \Lambda_n(x_1, \dots, x_n) = \sum_{t=1}^n \lambda(x_t) - \quad (7.27)$$

критериальная (тестовая) статистика, $C_-, C_+ \in \mathbb{R}$ – параметры (пороги) последовательного теста (7.26), (7.27), $C_- < C_+$; $1_A(\cdot)$ – индикаторная функция множества A .

Последовательный критерий отношения вероятностей (ПКОВ), предложенный А. Вальдом, основанный на статистике отношения правдоподобия, является элементом рассмотренного семейства последовательных тестов (7.26), (7.27) при

$$\lambda(u) = \lambda_W(u) = \log \frac{p_1(u)}{p_0(u)}, \quad u \in U. \quad (7.28)$$

Обычно значения порогов C_- , C_+ для ПКОВ вычисляют по формулам:

$$C_+ = \log \frac{1 - \beta_0}{\alpha_0}, \quad C_- = \log \frac{\beta_0}{1 - \alpha_0}, \quad (7.29)$$

где α_0 , β_0 – допустимые значения вероятностей ошибок I рода (принимается гипотеза \mathcal{H}_1 ($d_\lambda = 1$) при условии, что справедлива \mathcal{H}_0) и II рода (принимается \mathcal{H}_0 ($d_\lambda = 0$), когда справедлива \mathcal{H}_1).

Процесс изменения критериальной статистики в ПКОВ схематично представлен на рис. 7.3.

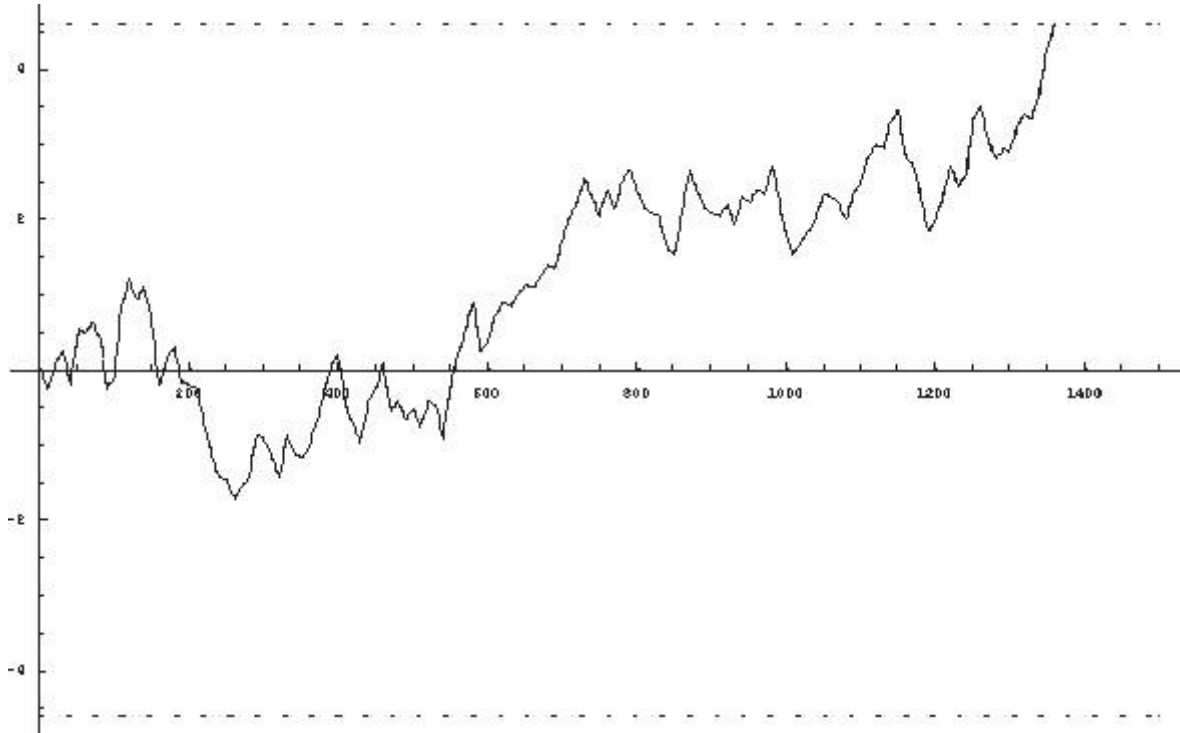


Рис. 7.3: Принятие ПКОВ гипотезы \mathcal{H}_1

7.2.3 Характеристики эффективности последовательных статистических тестов

Для теста, основанного на функции $\lambda(\cdot)$, обозначим

$$\begin{aligned} \alpha &= \alpha(\delta_\lambda) = E_0\{P_0\{d_\lambda = 1 \mid \tau_\lambda\}\}, \\ \beta &= \beta(\delta_\lambda) = E_1\{P_1\{d_\lambda = 0 \mid \tau_\lambda\}\} \end{aligned} \quad (7.30)$$

фактические значения вероятностей ошибок I и II рода соответственно, где $E_\theta\{\cdot\}$ означает математическое ожидание по распределению P_θ . Введем в рассмотрение также условное математическое ожидание объема выборки, когда верна гипотеза \mathcal{H}_k ($k \in \Theta$):

$$t_k = t_k(\delta_\lambda) = E_k\{\tau_\lambda\}. \quad (7.31)$$

Таким образом, множество характеристик эффективности последовательного теста $\delta_\lambda = (\tau_\lambda, d_\lambda) \in \Delta$ определяется четырьмя значениями: $\alpha(\delta_\lambda)$, $\beta(\delta_\lambda)$, $t_0(\delta_\lambda)$, $t_1(\delta_\lambda)$.

Доказано, что для ПКОВ выполняются неравенства

$$\alpha(\delta_{\lambda_w}) \leq \frac{\alpha_0}{1 - \beta_0}, \quad \beta(\delta_{\lambda_w}) \leq \frac{\beta_0}{1 - \alpha_0}, \quad \alpha(\delta_{\lambda_w}) + \beta(\delta_{\lambda_w}) \leq \alpha_0 + \beta_0.$$

На практике, однако, значения α_0 , β_0 могут существенно отличаться от фактических значений $\alpha(\delta_{\lambda_w})$, $\beta(\delta_{\lambda_w})$, и задача оценивания характеристик эффективности (7.30), (7.31) для последовательных тестов является важной.

Пусть $\Delta(\alpha, \beta)$ – совокупность всех решающих правил проверки гипотез (7.25), для которых вероятность ошибки I рода не превосходит α , а вероятность ошибки II рода не превосходит β .

Теорема 7.3 Пусть $\alpha, \beta > 0$ таковы, что $\alpha + \beta < 1$ и найдутся $C_-, C_+ \in \mathbb{R}$, $C_- < 0 < C_+$, для которых

$$\alpha(\delta_{\lambda_w}) = \alpha, \quad \beta(\delta_{\lambda_w}) = \beta.$$

Тогда ПКОВ (7.28) оптимален в классе $\Delta(\alpha, \beta)$ в том смысле, что одновременно выполняются неравенства

$$\forall \delta \in \Delta(\alpha, \beta) : t_0(\delta_{\lambda_w}) \leq t_0(\delta), \quad t_1(\delta_{\lambda_w}) \leq t_1(\delta).$$

Другими словами, теорема 7.3 утверждает, что ПКОВ удовлетворяет принципу оптимальности Вальда, изложенному выше.

Для условных математических ожиданий объема выборки известны нижние границы:

$$t_0(\delta_{\lambda_w}) \geq \frac{(1 - \alpha_0) \ln(\beta_0/(1 - \alpha_0)) + \alpha_0 \ln((1 - \beta_0)/\alpha_0)}{E_0 \{p_1(x_1)/p_0(x_1)\}},$$

$$t_1(\delta_{\lambda_w}) \geq \frac{\beta_0 \ln(\beta_0/(1 - \alpha_0)) + (1 - \beta_0) \ln(1 - \beta_0)/\alpha_0}{E_1 \{p_1(x_1)/p_0(x_1)\}}.$$

В общем случае получить точные выражения для характеристик эффективности последовательных тестов не удается.

7.2.4 Последовательная статистическая проверка сложных гипотез

Кратко опишем два основных подхода, применяемых в случае, когда требуется построение последовательного теста для проверки сложных параметрических гипотез.

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P_\theta)$ наблюдается последовательность независимых одинаково распределенных случайных величин с плотностью распределения вероятностей $f(x; \theta)$, зависящей от параметра $\theta = (\theta_1, \dots, \theta_k)$. Пусть множество значений Θ параметра θ разделено на три части:

$\Theta = \Theta_0 \cup \tilde{\Theta} \cup \Theta_1$. Имеются две сложные гипотезы о значении параметра θ : \mathcal{H}_0 и \mathcal{H}_1 . Требования к тесту проверки гипотез имеют следующий вид. Если $\theta \in \Theta_0$, то предпочтение должно отдаваться принятию гипотезы \mathcal{H}_0 ; если $\theta \in \Theta_1$, то предпочтение должно отдаваться гипотезе \mathcal{H}_1 ; множество $\tilde{\Theta}$ – так называемая «зона безразличия», при попадании значения параметра в которую нет строгих предпочтений в пользу какой-то из гипотез.

Таким образом, к тесту предъявляются следующие требования:

$$\sup P_{\theta}\{d = 1 \mid \theta \in \Theta_0\} \leq \alpha_0; \sup P_{\theta}\{d = 0 \mid \theta \in \Theta_1\} \leq \beta_0,$$

и тест должен быть конечен с вероятностью 1.

Пусть заданы две функции от θ , называемые весовыми, такие что

$$W_0(\theta) = 0, \theta \notin \Theta_0; W_1(\theta) = 0, \theta \notin \Theta_1,$$

и при этом

$$\int_{\Theta_0} W_0(\theta) d\theta = \int_{\Theta_1} W_1(\theta) d\theta = 1.$$

Задача проверки сложных (составных) гипотез $\mathcal{H}_0, \mathcal{H}_1$ заменяется задачей проверки простых гипотез $\mathcal{H}'_0(n), \mathcal{H}'_1(n)$; гипотеза $\mathcal{H}'_i(n)$ состоит в том, что при фиксированном n плотность распределения вероятностей выборки равна $\int_{\Theta_i} f(x_1; \theta) \cdots f(x_n; \theta) W_i(\theta) d\theta, i = 0, 1$.

При вычислении порогов последовательного теста следует положить

$$\alpha_0 = \int_{\Theta_0} W_0(\theta) \alpha(\theta) d\theta, \beta_0 = \int_{\Theta_1} W_1(\theta) \beta(\theta) d\theta,$$

где $\alpha(\theta), \beta(\theta)$ – условные вероятности ошибок I и II рода, когда значение параметра фиксировано и равно θ .

Изложенный метод построения последовательного теста носит название подхода, основанного на весовых функциях Вальда.

Второй возможный подход – преобразование наблюдений таким образом, чтобы распределение вероятностей преобразованных наблюдений имело бы один вид при справедливой сложной (составной) гипотезе \mathcal{H}_0 , и другой – при \mathcal{H}_1 . Этот подход основывается на теореме Кокса и приводит к построению так называемого последовательного t -теста.