Метод последовательного анализа

Дана выборка, состоящая из независимых одинаково распределенных наблюдений $x = (x_1, \ldots, x_n)$.

Существует прекрасный метод (последовательный критерий Вальда или метод последовательного отношения правлоподобия) проверки простой гипотезы $H_0: \vartheta = \vartheta_0$ против простой альтернативы $H_1: \vartheta = \vartheta_1$.

Примечание Так же как и в случае леммы Неймана–Пирсона продолжаемый на сложные односторонние гипотезы $H_0: \vartheta \le \vartheta_0$ против $H_1: \vartheta \ge \vartheta_1$, $\vartheta_1 > \vartheta_0$ в случае монотонного отношения правдоподобия, то есть когда

$$L(x) = \frac{p_{\vartheta_1}(x_1, \ldots, x_n)}{p_{\vartheta_0}(x_1, \ldots, x_n)} = f_{\vartheta_0 \vartheta_1}(T(x)),$$

где f при всяких $\vartheta_1 > \vartheta_0$ монотонно неубывает, $T(x)$ — некоторая одномерная статистика.

Логарифм отношения правдоподобия в этом случае является суммой независимых случайных величин,

$$S_n = \ln L(x) = \sum_{i=1}^{n} \ln \frac{p_{\vartheta_1}(x_i)}{p_{\vartheta_0}(x_i)} = \sum_{i=1}^{n} \eta_i.$$

При нулевой гипотезе математическое ожидание слагаемых отрицательно, при альтернативной — положительно. то есть блуждание "уходит" вниз или вверх в зависимости от того, какая гипотеза верна.

Существуют такие границы A — верхняя и -B — нижняя, что вероятности ошибок 1-го α и второго рода β (выхода не в ту сторону, меры рассматриваются на бесконечных последовательностях) равны заданным величинам. При этом математические ожидания времен выхода при обеих гипотезах будут минимальны в классе всех так называемых последовательных критериев. то есть таких, где потенциально бесконечная выборка наблюдается до тех пор (марковский момент!!!). пока не накопится информация, достаточная для принятия решения в пользу одной из гипотез, что происходит с вкроятностью единица! Обычно время принятия решения в последовательном анализе много меньше объема выборки, необходимой для проверки "обычного" критерия с теми же ошибкам.

Можно показать, что

$$A_1 = \epsilon^A \le \frac{1-\beta}{\alpha}, \; A_0 = \epsilon^{-B} \ge \frac{\beta}{1-\alpha}.$$

Причем эти величины получаются как значения границ при приближенном методе их нахождения. Настоящие же ошибки α' и β' обладают следующими свойствами: $\alpha' \le \alpha/(1-\beta)$ $\beta' \le \beta/(1-\alpha)$ и наконец $\alpha' + \beta' \le \alpha + \beta$. В том случае, если настоящие ошибки можно считать равными, они очевидно меньше заданных. Времена выхода τ — принятия решения подчиняются неравенствам: $M_0\tau \ge w(\alpha, \beta)/(-M_0\eta_1$, $M_1\tau \ge w(\beta, \alpha)/(-M_1\eta_1$, где $w(x, y) = x\ln(x/(1-y)) + (1-x)\ln((1-x)/y)$ — функция Вальда.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Задание 1.

В модели$N(\vartheta, \sigma_0^2)$ проверить гипотезу $H_0: \ \vartheta = 0$ против альтернативы $H_1: \ \vartheta = 1$ при единичной дисперсии и $\alpha = \beta = 0.05$. воспользовавшись приближенными значениями границ. Построить графики блужданий (пучки) при обеих гипотезах. Можно на одном чертеже, выделив принадлежность траектории гипотезе цветом, графики доводить только до момента выхода за границу, то есть принятия гипотезы. Оценить вероятности выхода "не в ту сторону", то есть истинные ошибки. Используя симметрию задачи (A=B), подобрать (а лучше придумать итеративный алгоритм!) границы соответствующие заданным вероятностям ошибок (в форме $\ln A_i, \ i = 0, 1$). Найти средние времена выхода, сравнить среднее время принятия решения для каждой гипотезы с теоретической нижней оценкой.