УТОЧНЕНИЕ ГРАНИЦ ДЛЯ ВЕРОЯТНОСТИ ОШИБКИ ПЕРВОГО РОДА ПРОЦЕДУРЫ БОНФЕРРОНИ В СЛУЧАЕ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ СТАТИСТИК КРИТЕРИЕВ

И. С. Милованова

ВВЕДЕНИЕ

Задача множественной проверки гипотез возникает во многих прикладных областях: медицине, генетике, криптографии и т.д.

Например, при применении статистических методов в медицине для изучения эффективности методики лечения на двух группах больных, одна из которых использовала изучаемую методику, сравнение проводят по m различным показателям, и для каждого показателя используется свой статистический критерий. В данном случае возникает проблема принятия итогового решения об эффективности методики лечения [4]. В криптографии подобная проблема возникает при статистической проверке качества бинарных последовательностей — выходных последовательностей криптографических алгоритмов или генератора псевдослучайных чисел — с использованием множества критериев для обнаружения различных отклонений от равномерного распределения [3].

Для принятия итогового решения по результатам отдельных критериев разработаны процедуры множественной проверки гипотез. Наиболее известной процедурой является процедура Бонферрони [5]. Известно [5], что если статистики критериев зависимы, то данная процедура является консервативной, т.е. вероятность ошибки первого рода значительно меньше заданного уровня значимости. В данной статье предложено уточнение границ для вероятности ошибки первого рода процедуры Бонферрони для семейства критериев с нормальным распределением статистик. Уточнение границ получено с учетом попарных зависимостей между статистиками.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Пусть для проверки гипотез $\mathrm{H}_{0,1},\ldots,\mathrm{H}_{0,\mathrm{m}}$ против альтернатив $\mathrm{H}_{1,1},\ldots,\mathrm{H}_{1,\mathrm{m}}$ соответственно имеется m критериев со статистиками $s_1,\ldots,s_m$.

Предполагается, что при $\mathrm{H}_{0,\mathrm{i}}$ статистика s_i имеет стандартное нормальное распределение и при $\mathrm{H}_0 = \bigcap_{i=1}^m \mathrm{H}_{0,\mathrm{i}}$ совместное распределение статистик $\{s_i\}$ также является нормальным с известной ковариационной матрицей:

$$\mathrm{L}\{s=(s_1,...,s_m)'\}=N(0,\Sigma),\ \mathrm{Cov}\{s_i,s_j\}=\rho_{ij},\ i,j=\overline{1,m}.$$

Будем рассматривать двухсторонние критерии следующего вида:

$$\textit{принимается}\begin{cases}\mathrm{H}_{0,\mathrm{i}},\ \textit{если}\ |s_i|<\Delta_i=\Phi^{-1}(1-\alpha_i/2),\\ \mathrm{H}_{1,\mathrm{i}},\ \textit{если}\ |s_i|\geq\Delta_i,\end{cases}\quad i=\overline{1,m}, \tag{1}$$

где статистики $s_1,...,s_m$ вычисляются по наблюдаемой выборке X, $\{\alpha_i\}$ — уровни значимости индивидуальных критериев. Переходя от статистик критериев $s_1,\ldots,s_m$ к Р-значениям $p_1,\ldots,p_m$, $p_i=p_i(s_i)=2\Phi(-|s_i|)$, получаем эквивалентный вид критериев:

$$\textit{принимается}\begin{cases}\mathrm{H}_{0,\mathrm{i}},\ \textit{если}\ p_{\mathrm{i}}>\alpha_i,\\ \mathrm{H}_{1,\mathrm{i}},\ \textit{если}\ p_{\mathrm{i}}\leq\alpha_i,\end{cases}\quad i=\overline{1,m}.$$

Необходимо проверить «объединенную» гипотезу $\mathrm{H}_0=\bigcap_{i=1}^m\mathrm{H}_{0,\mathrm{i}}$ против альтернативы $\mathrm{H}_1=\bigcup_{i=1}^m\mathrm{H}_{1,\mathrm{i}}$.

Введем в рассмотрение обобщенную вероятность ошибки первого рода [4]:

$$\varepsilon=\mathbf{P}\{\textit{отвергнуть хотя бы одну}\ \mathrm{H}_{0,\mathrm{i}}\,|\,\mathrm{H}_0\}. \tag{2}$$

Процедура Бонферрони, использующая для принятия решения Р-значения, полученные при проверке индивидуальных гипотез, имеет вид [5]:

$$\textit{принимается}\quad\begin{cases}\mathrm{H}_0,\quad \textit{если}\ \bigcap_{\mathrm{i}=1}^{\mathrm{m}}\{p_i>\alpha_i\},\\ \mathrm{H}_1,\quad \textit{иначе.}\end{cases}$$

Для значения обобщенной вероятности ошибки первого рода ε данной процедуры известны следующие границы [4]:

$$\max_{1\leq i\leq m}\alpha_i\leq\varepsilon\leq\sum_{i=1}^m\alpha_i\,.$$

Тогда, выбирая $\alpha_i=\alpha_c=\alpha/m,\quad i=\overline{1,m}$, для процедуры Бонферрони будет выполняться $\varepsilon\leq\alpha$. Однако в случае сильной зависимости между статистиками критериев такой выбор уровня значимости индивидуальных критериев может привести к завышенной оценке сверху [5]: $\varepsilon/\alpha<<1$.

Поэтому для улучшения процедуры Бонферрони необходим учет зависимостей между статистиками критериев и уточнение границ для значения обобщенной вероятности ошибки первого рода.

Теорема. Для обобщенной вероятности ошибки первого рода (2) процедуры Бонферрони для семейства критериев (1) выполняется:

$$\alpha_-(\{\alpha_i\}) \le \varepsilon \le \alpha_+(\{\alpha_i\}),$$

$$\alpha_-(\{\alpha_i\}) = \sum_{i=1}^{m}\alpha_i - \sum_{j=1}^{m}\sum_{i\ne j}^{m}\left(\int_{-\Delta_i}^{\Delta_i}\int_{-\Delta_j}^{\Delta_j} n(u,v\,|\,0,\rho_{ij})\,du\,dv + \alpha_i + \alpha_j - 1\right),$$

$$\alpha_+(\{\alpha_i\}) = \sum_{i=1}^{m}\alpha_i - \max_{1\le j\le m}\sum_{i\ne j}^{m}\left(\int_{-\Delta_i}^{\Delta_i}\int_{-\Delta_j}^{\Delta_j} n(u,v\,|\,0,\rho_{ij})\,du\,dv + \alpha_i + \alpha_j - 1\right),$$

где $n(u,v\,|\,0,\rho_{ij})$ — маргинальная плотность распределения вектора (s_i,s_j), $\rho_{ij} = \mathrm{Cov}\{s_i,s_j\}$.

Доказательство основано на принципе включения-исключения и на неравенстве Бонферрони второго порядка, приведенного в [2].

*Следствие***.** В случае одинаковых уровней значимости для всех критериев $\alpha_i = \alpha_c, i = \overline{1,m}$, верхняя оценка для вероятности (2) процедуры Бонферрони равна

$$\alpha_+ = 1 + (m-2)(1-\alpha_c) - \max_{1\le j\le m}\sum_{i\ne j}^{m}\int_{-\Delta}^{\Delta}\int_{-\Delta}^{\Delta} n(u,v\,|\,0,\rho_{ij})\,du\,dv$$

и справедливо следующее разложение в ряд

$$\alpha_+ = m\alpha_c + (m-1)\alpha_c^2 - \max_{1\le j\le m}\sum_{i\ne j}^{m}\left[\sum_{k=1}^{\infty}\frac{4\left(n_1^{(2k-1)}(\Delta\,|\,0,1)\right)^2}{(2k)!}\rho_{ij}^{2k}\right], \tag{3}$$

где $n_1^{(k)}(\Delta\,|\,0,1)$ – значение $k-$й производной стандартной нормальной плотности распределения в точке $\Delta = \Phi^{-1}(1-\alpha_c/2)$.

В разложении (3) первое слагаемое — это оценка сверху для обобщенной вероятности ошибки первого рода процедуры Бонферрони без учета попарной зависимости между статистиками критериев. Следующие два слагаемых дают уточнение границы.

ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

В качестве математической модели выборки использовалась модель независимых симметричных испытаний Бернулли и семейство критериев поиска шаблонов 1,11,111,1111. Коэффициенты корреляции статистик близки к единице [1]. Теоретический уровень значимости полагался равным 0.05. На рисунке приведены оценки уровня значимости процедуры Бонферрони (обозначение «○») с 95 % доверительным интервалом. Из рисунка можно видеть, что уточненная граница (пунктир), равная $\alpha_+ = 0.03198369$, попадает в доверительный интервал, а классическая граница (непрерывная линия) находится выше. Как видно, полученная оценка на 36% меньше первоначальной, что иллюстрирует консервативность процедуры Бонферрони в случае сильно коррелированных статистик.

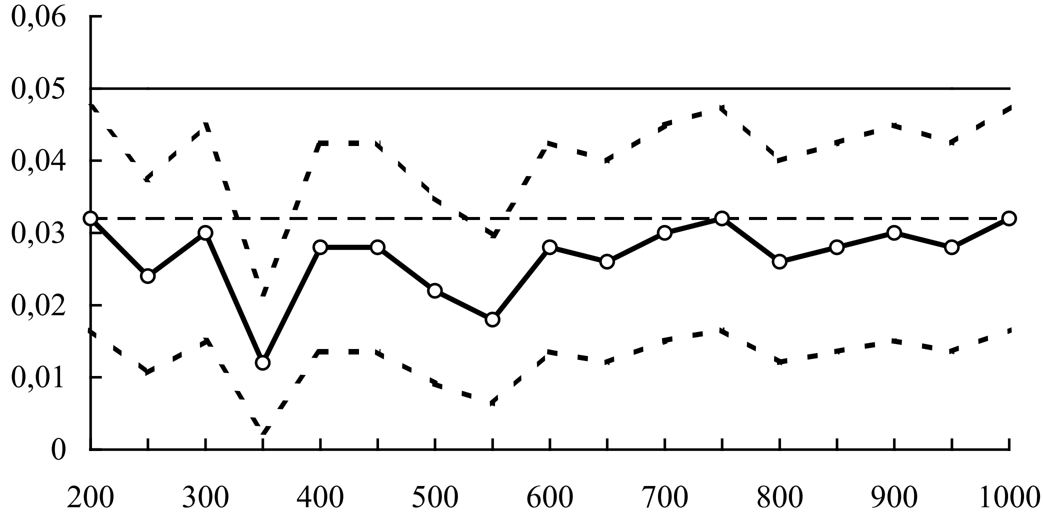

Рис. 1. Оценка вероятности ошибки первого рода для критерия шаблонов

Литература

1. *Милованова И. С.* О критерии поиска набора «шаблонов» для тестирования бинарных последовательностей // III республиканская научная конференция молодых ученых и студентов. Брест, 2003.
2. *Kounias E. G.* Bounds for the probability of a union of events with applications // Ann. Math. Statist.,1968, v.39.
3. *Rukhin A. et al.* Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST Special Publication 800-22, 2000.
4. *Sarkar S. K.* Recent Advances in Multiple Testing. – 2002. http://www.sbm.temple.edu/~sanat/multtest.pdf
5. *Simes R. J.* An Improved Bonferroni Procedure for Multiple Tests of Significance // Biometrika, 1986, v. 73, p. 751–754.