УДК 658.012.011

ГАРАНТИРОВАННАЯ ОЦЕНКА КВАЛИФИКАЦИИ ЭКСПЕРТОВ В ЗАДАЧАХ ПРИНЯТИЯ РЕШЕНИЙ

Л.С. Файнзильберг, О.А. Жуковская
НТТУ КПИ, МНУЦ ИТИС
e-mail: fainzilberg@svitonline.com

Термин «квалифицированный эксперт» довольно часто встречается в публикациях, в том числе, и в научной литературе. Интуитивное определение этого термина понятно: квалифицированным экспертом считается признанный специалист в определенной предметной области. Однако развитие компьютерных систем поддержки принятия решений, основанных на знаниях экспертов, настоятельно требует формализации представлений о квалификации экспертов.

В докладе приводятся результаты исследований, цель которых получить формальные условия, гарантирующих квалификацию эксперта при диагностике случайных состояний объекта, в частности, квалификацию врача при диагностике заболевания с известной распространенностью.

Постановка задачи. Пусть некоторый объект находится в одном из двух возможных состояний V_1, V_2, переходя случайным образом из одного состояния в другое с априорными вероятностями $P(V_1)$, $P(V_2) = 1 - P(V_1)$. Эксперт, опираясь на свои знания и доступную информацию, принимает решения δ о текущем состоянии объекта в виде индикаторной функции

$$\delta = \begin{cases} 1, \text{если эксперт решает в пользу } V_1, \\ 2, \text{если эксперт решает в пользу } V_2. \end{cases}$$

Предполагается, что, помимо правильных решений, эксперт может ошибаться, т.е. допускать ошибки пропуска цели и ложной тревоги. Будем, как это принято в теории статистических решений, характеризовать возможные потери платежной матрицей

$$L = \begin{Vmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{Vmatrix},$$

в которой L_{11} и L_{22} – потери, связанные с правильными решениями, а L_{12} и L_{21} – потери, связанные с ошибками первого и второго рода. Тогда средний риск R решений, принимаемых экспертом, определяется математическим ожиданием указанных потерь

$$R = \sum_{k=1}^{2} \sum_{j=1}^{2} L_{jk} P(V_k, \delta = j) .$$

где величина $P(V_k, \delta = j)$ обозначает вероятность совместного выполнения двух случайных событий: объект находится в состоянии V_k $(k = 1, 2)$, а эксперт принял решение $\delta = j$ в пользу j-го состояния V_j $(j = 1, 2)$.

Допустим, что по репрезентативной выборке наблюдений с известными состояниями объекта оценены частоты ошибок эксперта

$$P_{jk}^* = \frac{E_j}{n_k}, \quad k, j = 1, 2,$$

где n_k, E_j – соответственно число случаев, когда объект заведомо находился в состоянии V_k, а эксперт принял неверное решение ($j \neq k$).

Ставится задача получить достаточные условия, позволяющие на основе имеющейся априорной информации с заданной доверительной вероятностью β оценить квалификацию эксперта и проводить сравнение квалификаций двух экспертов с точки зрения следующих определений.

Определение 1. Эксперт квалифицирован, если средний риск R его решений меньше априорного риска R_0, т.е. выполняется строгое неравенство $R < R_0$.

Доказана следующая теорема, которая являются интервальным обобщением одного из результатов, полученных в работе [1].

Теорема 1. С доверительной вероятностью β эксперт квалифицирован, если выполняется одно из условий

$$P_{12}^c + r_{12} < \theta(1 - P_{21}^c - r_{21}) \text{ при } \theta \le 1, \tag{1}$$

$$P_{12}^c + r_{12} < 1 - \theta(P_{21}^c + r_{21}) \text{ при } \theta > 1, \tag{2}$$

и неквалифицирован, если выполняются одно из условий

$$P_{12}^c - r_{12} > \theta(1 - P_{21}^c + r_{21}) \text{ при } \theta \le 1, \tag{3}$$

$$P_{12}^c - r_{12} > 1 - \theta(P_{21}^c - r_{21}) \text{ при } \theta > 1, \tag{4}$$

где

$$\theta = \frac{(L_{21} - L_{11})P(V_1)}{(L_{12} - L_{22})[1 - P(V_1)]}$$

а P_{12}^c, P_{21}^c и r_{12}, r_{21} – соответственно центры и радиусы доверительных интервалов вероятностей ошибок пропуска цели и ложной тревоги, которые допустил эксперт на ограниченной экзаменационной выборке наблюдений с известными состояниями объекта.

Очевидно, что квалификация эксперта остается неопределенной, если не выполняется ни одно из условий (1)– (4).

В качестве иллюстрации на рис. 1 представлены области значений частот ошибок эксперта P_{12}^* и P_{21}^*, которые с вероятностью $\beta = 0{,}99$ подтверждают (область **A**) и опровергают (область **B**) квалификацию эксперта. Указанные области построены для следующих фиксированных величин: $P(V_1) = 0{,}15$, $L_{12} = 5$, $L_{21} = 1$, $L_{11} = L_{22} = 0$, т.е. $\theta = 0{,}882$. Как и следовало ожидать, область неопределенной квалификации эксперта (серая область) уменьшается по мере увеличения объема n наблюдений.

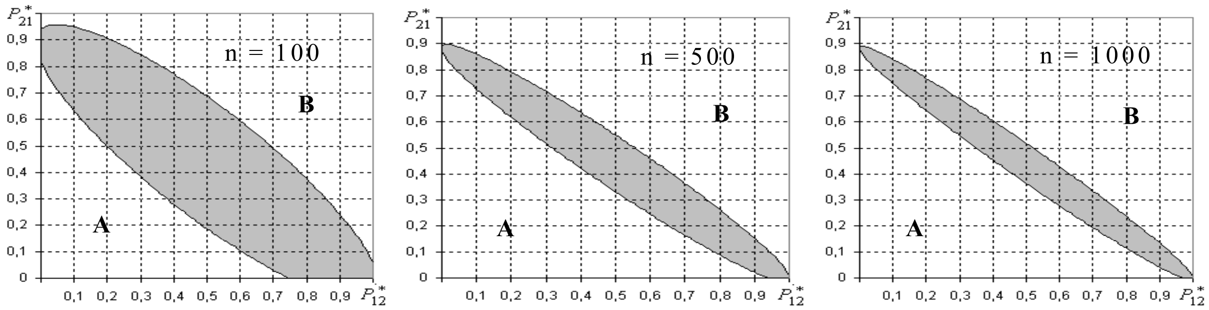

Рис.1

Важным для практики результатом исследований является следующая теорема.

Теорема 2. Для любой доверительной вероятности β и фиксированном значении θ существует такое число $n_0 > 0$, что после оценки частот P_{jk}^*, $j,k = 1, 2,\ j \ne k$ ошибок эксперта по репрезентативной выборке объемом $n > n_0$ можно с вероятностью β подтвердить или опровергнуть квалификацию эксперта, причем значение n_0 определяется решением уравнений

$$P_{12}^{c}(n) \pm r_{12}(n) = \theta[1 - P_{21}^{c}(n) \mp r_{21}(n)], \quad \text{если } \theta \le 1,$$

$$P_{12}^{c}(n) \mp r_{12}(n) = 1 - \theta[P_{21}^{c}(n) \pm r_{21}(n)], \quad \text{если } \theta > 1.$$

где

$$P_{jk}^{c} = \frac{P_{jk}^{*} + t_{\beta}^{2} / 2n}{1 + t_{\beta}^{2} / n},$$

$$r_{jk} = \frac{t_{\beta}\sqrt{\dfrac{P_{jk}^{*}(1 - P_{jk}^{*})}{n} + \dfrac{t_{\beta}^{2}}{4n^{2}}}}{1 + t_{\beta}^{2} / n},$$

$$t_{\beta} = \arg \Phi^{*}\left(\frac{1+\beta}{2}\right),$$

а $\Phi^{*}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{\tau^2}{2}} d\tau$ – нормальная функция распределения.

Для сравнения квалификаций двух экспертов введем определение.

Определение 2. Эксперт A_1 более квалифицирован, чем эксперт A_2, если средний риск R_1, основанный на решениях A_1, меньше среднего риска R_2, основанного на решениях A_2.

Теорема 3. С доверительной вероятностью β эксперт A_1 более квалифицирован, чем эксперт A_2, если

$$\theta(P_{21}^{c\,(1)} + r_{21}^{(1)} - P_{21}^{c\,(2)} + r_{21}^{(2)}) < P_{12}^{c\,(2)} - r_{12}^{(2)} - P_{12}^{c\,(1)} - r_{12}^{(1)}, \tag{5}$$

и менее квалифицирован, чем эксперт A_2, если

$$\theta(P_{21}^{c\,(1)} - r_{21}^{(1)} - P_{21}^{c\,(2)} - r_{21}^{(2)}) > P_{12}^{c\,(2)} + r_{12}^{(2)} - P_{12}^{c\,(1)} + r_{12}^{(1)}, \tag{6}$$

где $P_{12}^{c\,(i)}$, $P_{21}^{c\,(i)}$ и $r_{12}^{(i)}$, $r_{21}^{(i)}$ – соответственно центры и радиусы доверительных интервалов условных вероятностей ошибок i-го эксперта ($i = 1, 2$), которые зависят исключительно от доверительной вероятности β и частот ошибок, допущенных соответствующим экспертом на экспериментальных выборках известного объема.

На рисунке 2 приведены графики функций, представляющие собой зависимости от числа наблюдений n левой (кривая 1) и правой (кривая 2) частей неравенства (5), а также левой (кривая 3) и правой (кривая 4) частей неравенства (6). Указанные зависимости построены для значения доверительной вероятности $\beta = 0{,}99$ при частотах ошибок экспертов $P_{12}^{*\,(1)} = 0{,}01$ $P_{12}^{*\,(2)} = 0{,}07$, $P_{21}^{*\,(1)} = 0{,}04$, $P_{21}^{*\,(2)} = 0{,}02$, априорной вероятности $P(V_1) = 0{,}1$ и значениях элементов матрицы потерь $L_{11} = L_{22} = 0$, $L_{21} = 2$, $L_{12} = 1$, т.е. для случая $\theta \approx 0{,}22$.

Легко видеть, что в рассматриваемом случае сравнение квалификаций экспертов возможно лишь в том случае, когда частоты их ошибок оценены более чем по 430 наблюдениям.

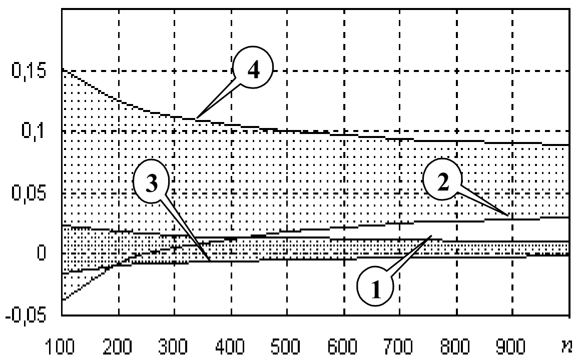

Рис. 2.

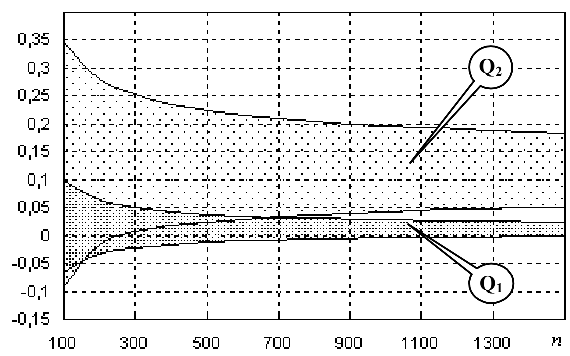

Рис. 3

Величина θ, фигурирующая в условиях (5), (6), является константой лишь в том случае, когда известны точные значения элементов платежной матрицы и априорные вероятности состояний объекта. Однако в общем случае разумно перейти от точечных значений L_{ij} и $P(V_1)$ к их интервальным аналогам $\left\langle L_{ij}^c, r_{ij} \right\rangle$ и $\left\langle P_{V_1}^c, r_{V_1} \right\rangle$ в форме центр-радиус. Легко показать, что в общем случае сравнение квалификаций экспертов возможно, если интервал

$$\mathbf{Q}_1 = \left\langle P_{21}^{c\,(1)} - P_{21}^{c\,(2)}, r_{P_{21}^{(1)}} + r_{P_{21}^{(2)}} \right\rangle \left\langle P_{V_1}^c, r_{V_1} \right\rangle \left\langle L_{21}^c - L_{11}^c, r_{L_{21}} + r_{L_{11}} \right\rangle$$

не пересекается с интервалом

$$\mathbf{Q}_2 = \left\langle P_{12}^{c\,(2)} - P_{12}^{c\,(1)}, r_{P_{12}^{(2)}} + r_{P_{12}^{(1)}} \right\rangle \left\langle 1 - P_{V_1}^c, r_{V_1} \right\rangle \left\langle L_{12}^c - L_{22}^c, r_{L_{12}} + r_{L_{22}} \right\rangle.$$

причем эксперт A_1 более квалифицирован, чем эксперт A_2, если выполняется условие $Q_1^c + r_{Q_1} < Q_2^c - r_{Q_2}$ и менее квалифицирован, чем эксперт A_2, если выполняется условие $Q_1^c - r_{Q_1} > Q_2^c + r_{Q_2}$, где Q_j^c, r_{Q_j} – центры и радиусы интервалов $\mathbf{Q_j}$, $j = 1,2$.

Используя выражения для арифметических операций над интервалами [2], определены значения Q_j^c, r_{Q_j} и показано, что и в рассматриваемом общем случае для любой доверительной вероятности β всегда можно за счет увеличения объема выборки обеспечить выполнение одного из условий, позволяющих сравнивать квалификации экспертов.

Приведем модельный пример (рис.3), который иллюстрирует этот факт для следующих условий: $\beta = 0{,}99$, $P_{12}^{*\,(1)} = 0{,}01$ $P_{12}^{*\,(2)} = 0{,}07$, $P_{21}^{*\,(1)} = 0{,}04$, $P_{21}^{*\,(2)} = 0{,}02$, $P_{V_1}^{*} = 0{,}1$, $L_{11} \in \langle 1;\, 0{,}1\rangle$, $L_{22} \in \langle 1;\, 0{,}1\rangle$, $L_{21} \in \langle 5;\, 0{,}5\rangle$, $L_{12} \in \langle 3;\, 0{,}3\rangle$. В отличие от предыдущего примера (рис.2), в данном случае сравнение квалификаций экспертов с доверительной вероятностью $\beta = 0{,}99$ возможно уже при большем значении числа наблюдений в экспериментальной выборке ($n > 710$).

1. *Файнзильберг Л.С.* Условия полезности диагностических тестов с позиции теории статистических решений // Проблемы управления и информатики. – 2003. – № 2. – С. 100-111.
2. *Жуковська О.А., Новицький В.В.* Прямий метод обчислення добутку інтервалів у формі центр-радіус // Наук. вісті НТУУ “КПІ”. – 2003. – № 1. – С. 138 – 144.