МАТЕМАТИКА

Теория вероятностей и математическая статистика (01.01.05)

УДК 517.519
DOI: 10.24160/1993-6982-2019-3-130-133

Оценки ошибок в тексте по результатам нескольких независимых экспертиз

А.Н. Архангельский, Г.М. Пиголкин

Построена система оценки качества рабочего результата (текста, изделия и т. п.) любой сложности и любого предназначения. Подход к решению подобного рода проблемы состоит в том, что один и тот же информационный источник, полностью описывающий оцениваемый объект, тиражируется тем или иным образом столько раз, сколько имеется субъектов (экспертов), каждый из которых (независимо от остальных) способен проделать нужную экспертизу. Результаты экспертиз обрабатываются и устанавливаются некие недостатки (ошибки), обнаруженные каждым из экспертов. При этом фактические признаки обнаруженных ошибок должны быть определены так, чтобы можно было указать как количество ошибок, обнаруженных каждым экспертом, так и количество ошибок, одновременно найденных каждым подмножеством из данного множества экспертов. Таким образом, результатом множественной экспертизы считаются все ошибки по всем экспертам.

Используя принцип максимального правдоподобия и оценки, связанные с неравенством П.Л. Чебышева, на основе математической модели, в работе сделаны (доведены до окончательных формул) некоторые основные оценки качеств выполненных экспертиз и текста (изделия). Кроме этого, кратко представлен спектр возможных исследований по данным множественной экспертизы.

Ключевые слова: независимые испытания, испытания Бернулли, неравенство Чебышева, конституенты, функции и оценки наибольшего правдоподобия, оценки на подмножествах, интервальные оценки.

Для цитирования: Архангельский А.Н., Пиголкин Г.М. Оценки ошибок в тексте по результатам нескольких независимых экспертиз // Вестник МЭИ. 2019. № 3. С. 130—133. DOI: 10.24160/1993-6982-2019-3-130-133.

Assessments of Errors in a Text from the Results of Several Independent Examinations

A.N. Arkhangelsky, G.M. Pigolkin

A system for estimating the quality of a working result (a text, a product, etc.) of any complexity and for any purpose is constructed. The approach to solving a problem of such kind consists in the following. The same information source, which completely describes the estimated object is replicated in some or other way in the number of times equal to the number of available subjects (experts), each of which is capable to perform the necessary examination independently of other experts. The examination results are processed, and deficiencies (errors) revealed by each expert are established. In this case, the actual signs of the revealed errors must be defined so that it would be possible to indicate both the number of errors revealed by each expert and the number of errors found within each subset from the given set of experts. Thus, all errors for all experts are regarded as the result of a multiple examination.

In this study, certain basic estimates of the qualities of the accomplished examinations and text (product) are made (brought to final formulas) based on applying the mathematical model and using the maximum likelihood principle, and assessments connected with Chebyshev inequality. In addition, the range of possible investigations based on multiple examination data is briefly outlined.

Key words: independent tests, Bernoulli tests, Chebyshev inequality, constituents, maximum likelihood functions and estimates, subset-based estimates, interval estimates.

For citation: Arkhangelsky A.N., Pigolkin G.M. Assessments of Errors in a Text from the Results of Several Independent Examinations. Bulletin of MPEI. 2019;3:130—133. (in Russian). DOI: 10.24160/1993-6982-2019-3-130-133.

Постановка задачи

Задачи, связанные с использованием экспертных оценок, часто встречаются в практических приложениях и задачах. Много примеров можно найти в работах [1, 2]. Математические методы в большинстве случаев используются в для оценивания количества и качества выбираемой экспертной группы [3, 4]. В настоящей публикации рассмотрены оценки, связанные с качеством проведенных экспертиз уже выбранной экспертной группы.

Пусть один и тот же текст (документ) исследуется каждым из нескольких экспертов независимо от остальных так, что каждый из экспертов обнаруживает недостатки или ошибки в тексте независимо друг от друга. Обнаруженные каждым экспертом ошибки окончательно признаются и образуют результаты в виде набора числовых массивов:

- n_k, $1 \le k \le E$ — количество ошибок, обнаруженных k-м экспертом (всего экспертов $E \ge 2$);
- $n_{k_1 \dots k_m}$, $1 \le k_1 < k_2 < \dots < k_m \le E$, $1 \le m \le E$, $1 \le k \le E$— количество ошибок, совместно обнаруженных m экспертами (номера экспертов, обнаруживших эти ошибки $k_1, k_2, \dots, k_m$).

Таким образом, из всего количества данных E независимых экспертиз есть $2^E - 1$ чисел, упорядоченных как очевидные подмножества полного множества ошибок. Определение полного числа ошибок в тексте является главной целью построения оценок по указанным результатам экспертиз. Одновременно строят оценки качества проведения самих экспертиз.

Математическая модель

Описанная задача может быть сформулирована в терминах независимых испытаний Бернуллиевского типа. На каждой из ошибок, существующих в тексте, каждый из экспертов проходит испытание, обнаруживая ошибку с вероятностью p_i (для эксперта с номером $i (1 \le i \le E)$), а $q_i = 1 - p_i$ — вероятность не выявить встреченную ошибку.

Структуру подмножеств обнаруженных ошибок и их пересечений, возникающую по результатам экспертиз, рассмотрим как систему так называемых конституент $\omega(i_1, i_2, \dots, i_E)$, $i_j \in \{0, 1\}$, т. е. удобных элементарных подмножеств в алгебре дискретных множеств [5, с. 47], означающих множество ошибок, обнаруженных теми и только теми экспертами (с номерами k), чьи индексы равны единице ($i_k = 1$). Конституента $\omega(0, 0, \dots, 0)$ содержит те и только те ошибки, которые все E экспертов пропустили.

Обозначим $n(i_1, i_2, \dots, i_E) = |\omega(i_1, i_2, \dots, i_E)|$ количество элементов в той или иной конституенте, т. е. число вышеописанных ошибок, обнаруженных лишь теми экспертами, чьи индексы равны единице. Таким образом, для E экспертов всего конституент (включая вышеуказанную нулевую конституенту) равно 2^E, причем $n_{k_1,\dots,k_m} = \sum\limits_{i_{k_1} = \dots = i_{k_m} = 1} n(i_1, \dots, i_E)$, то есть число ошибок k, совместно найденных выбранными $k_1, \dots, k_m$ экспертами, равно сумме количеств элементов во всех конституентах с единичными индексами этих экспертов.

Отметим следующие очевидные факты:

- все 2^E конституент попарно не пересекаются;
- полное количество ошибок в тексте:

$$\sum_{i_1 + \dots + i_E \ge 0} n(i_1, \dots, i_E) = N;$$

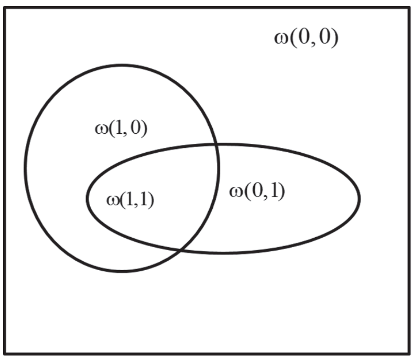
- полное количество ошибок, обнаруженных во всех E экспертизах:

$$\sum_{i_1 + \dots + i_E \ge 1} n(i_1, \dots, i_E) = S.$$

Структура конституент в случае двух экспертов продемонстрирована на рисунке.

Оценка случайного события экспертиз

Считаем, как подразумевается в испытаниях типа Бернулли (в случаях дискретных распределений — принципы построения схемы Бернулли) [6, с. 178], что все существующие ошибки при попадании в те или иные конституенты равноправны. Учитывая попарную несовместность конституент и очевидную кратность фиксированного по объемам набора 2^E конституент в $N!$ перестановок имеющегося (пока совершенно неиз-



$|\omega(1,0)| = n(1,0)$ – обнаружил только первый;

$|\omega(0,1)| = n(0,1)$ – обнаружил только второй;

$|\omega(1,1)| = n(1,1)$ – обнаружили оба;

$|\omega(0,0)| = (N - S)$ – оба пропустили;

$n_1 = n(1,0) + n(1,1)$ – все, обнаруженные первым;

$n_2 = n(0,1) + n(1,1)$ – все, обнаруженные вторым;

$n_{12} = n(1,1)$ – все, обнаруженные обоими.

Структура конституент

вестного) полного числа всех ошибок в тексте, получим вероятность случайного результата рассматриваемой системы E экспертиз:

$$P(n(i_1,...,i_E \mid i_j \in \{0,1\}) = \\ = \frac{N!}{(N-S)! \prod\limits_{i_1+...+i_E \geq 1} n(i_1,...,i_E)!} \prod_{k=1}^{E} p_k^{n_k} (1-p_k)^{N-n_k}. \quad (1)$$

Используя для распределения (1) свойства функции правдоподобия (абстрактное обоснование [7] или конкретизацию для системы независимых наблюдений [8]), примем принцип наибольшего правдоподобия. Согласно нему максимизируем вероятность (1) по параметрам, подлежащим оценке:

$$p_{i=} \frac{n_i}{N}, \ \ i = 1, 2, ..., E. \quad (2)$$

Заметим, что (2) максимизирует вероятность (1) при любом известном количестве ошибок $N \geq S$. Условия экстремума по N дает (вместе с (2)) оценку для N:

$$\ln \prod_{k=1}^{E} \left(1 - \frac{n_i}{N}\right) + \sum_{k=0}^{S-1} \frac{1}{N-k} = 0. \quad (3)$$

Не станем исследовать всю оценку (2), (3) максимума правдоподобия случайного события, отвечающего вероятности (1), по двум причинам. Во-первых, смысл оценки (3), мягко говоря, туманен. А, во-вторых, в приведенных оценках слишком слабо участвуют результаты ошибок, совместно обнаруженных несколькими экспертами. Данные $n_{k_1 ... k_m}$ лишь косвенно присутствуют в определении количества S всех обнаруженных ошибок, еще менее учитываются в оценках размеры конституент.

Попытаемся учесть все $n_{k_1 ... k_m}$, если не в полной мере, то хотя бы достаточно для реально действующих оценок.

Оценки на некоторых событиях, входящих в полный результат экспертиз

Займемся точечными оценками по методу максимального правдоподобия, добавив к оценкам (2) аналогичные оценки для фигурирующих в экспертизах испытаний на конституентах. Пусть у такой конституенты несколько индексов (не менее, чем один) единичны, остальные индексы — нули, и номера единичных индексов отличны от числа i_0. Тогда, по методу максимального правдоподобия, аналогично оценкам (2):

$$p_{i_0 =} \frac{n_{k_1 ... i_0 ... k_m}}{n_{k_1 ... k_m}}; \\ 1 \leq k_1 < k_2 < ... < k_m \leq E; \ \ i_0 \neq k_j; \ \ 1 \leq j \leq m. \quad (4)$$

Сопоставив оценки (4) с оценками (2), получим правдоподобные оценки для полного количества N ошибок в тексте:

$$N = \frac{n_i \cdot n_{k_1 ... k_m}}{n_{k_1 ... i ... k_m}}; \ \ 1 \leq k_1 < k_2 < ... < k_m \leq E; \\ 1 \leq i \leq E, i \neq k_j; \ \ 1 \leq j \leq m. \quad (5)$$

Нетрудно подсчитать количество оценок (5) для N. В случае E экспертов оно равно

$$\left(E 2^{E-1} - \frac{E(E+1)}{2}\right),$$

если $E = 2$, то оценка (5) всего одна:

$$N\big|_{E=2} = \frac{n_1 n_2}{n_{12}},$$

а при $E = 3$ их уже шесть:

$$N\big|_{E=3} \in \left\{ \frac{n_1 \cdot n_2}{n_{12}}, \frac{n_1 \cdot n_3}{n_{13}}, \frac{n_2 \cdot n_3}{n_{23}}, \frac{n_1 \cdot n_{23}}{n_{123}}, \frac{n_2 \cdot n_{13}}{n_{123}}, \frac{n_3 \cdot n_{12}}{n_{123}} \right\}.$$

Легко проверить, что при $E = 4$ оценок будет 22, при $E = 5$ — 65, при $E = 6$ — 171, при $E = 7$ – 420, при $E = 8$ — 988 и т. д.

Выбирая наибольшую из имеющихся оценок для фиксированного числа экспертов E, получим то, что называется пессимистической оценкой. Некоторую корректировку может дать использование неравенства Чебышева и интервальных оценок.

Применение неравенства Чебышева

Учитывая простые факты схемы Бернулли и применяя неравенство Чебышева [5, с. 21, 53] к описанным сериям экспертиз, найдем:

$$\begin{cases} P\left(\left|p_i - \dfrac{n_i}{N}\right| < \varepsilon\right) > 1 - \dfrac{1}{4N\varepsilon^2}; \\ P\left(\left|p_i - \dfrac{n_{k_1 ... i ... k_m}}{n_{k_1 ... k_m}}\right| < \varepsilon\right) > 1 - \dfrac{1}{4 n_{k_1 ... k_m} \varepsilon^2}. \end{cases}$$

Путем простых огрублений получим:

$$P\left(\left|\frac{n_i}{N} - \frac{n_{k_1 ... i ... k_m}}{n_{k_1 ... k_m}}\right| < 2\varepsilon\right) > 1 - \frac{1}{4 n_{k_1 ... k_m} \varepsilon^2}. \quad (6)$$

Таким образом, выбрав пессимистический вариант точечной оценки полного числа N ошибок в тексте, для соответствующей выбранной оценки индексации i; k_1, ..., k_m уточним принятый вариант в виде интервальной оценки (6).

Таким образом, в рассмотренной задаче оценок по нескольким экспертизам в реальных случаях хороших экспертов возможно более подробное и всестороннее оценивание по составляющим случайных испытаний.

Литература

1. **Семенов С.С., Воронов Е.М., Полтавский А.В., Крянев А.В.** Методы принятия решений в задачах оценки качества и технического уровня сложных технических систем. М.: Ленанд, 2016.
2. **Benini A. e. a.** The Use of Expert Judgment in Humanitarian Analysis — Theory, Methods, Applications. Geneva, Assessment Capacities Project, 2017.
3. **Крянев А.В., Тихомирова А.Н., Сидоренко Е.В.** Групповая экспертиза инновационных проектов с использованием байесовского подхода // Экономика и математические методы. 2013. Т. 49. № 2. С. 124—139.
4. **Крянев А.В., Семенов С.С.** К вопросу о качестве и надежности экспертных оценок при определении технического уровня сложных систем // Надежность. 2013. Вып. 4. С. 90—99.
5. **Смирнов Н.В., Дунин-Барковский И.В.** Краткий курс математической статистики для технических приложений. М.: Физматлит, 1959.
6. **Прохоров Ю.В., Пономаренко Л.С.** Лекции по теории вероятностей и математической статистике. М.: Изд-во МГУ им. М.В. Ломоносова, 2004.
7. **Барра Ж.-Р.** Основные понятия математической статистики. М.: Мир, 1974.
8. **Де Гроот М.** Оптимальные статистические решения. М.: Мир, 1974.

References

1. **Semenov S.S., Voronov E.M., Poltavskiy A.V., Kryanev A.V.** Metody Prinyatiya Resheniy v Zadachakh Otsenki Kachestva i Tekhnicheskogo Urovnya Slozhnykh Tekhnicheskikh Sistem. M.: Lenand, 2016. (in Russian).
2. **Benini A. e. a.** The Use of Expert Judgment in Humanitarian Analysis — Theory, Methods, Applications. Geneva, Assessment Capacities Project, 2017.
3. **Kryanev A.V., Tikhomirova A.N., Sidorenko E.V.** Gruppovaya Ekspertiza Innovatsionnykh Proektov s Ispol'zovaniem Bayesovskogo Podkhoda. Ekonomika i Matemticheskie Metody. 2013;49;2:124—139. (in Russian).
4. **Kryanev A.V., Semenov S.S.** K Voprosu o Kachestve i Nadezhnosti Ekspertnykh Otsenok pri Opredelenii Tekhnicheskogo Urovnya Slozhnykh Sistem. Nadezhnost'. 2013;4:90—99. (in Russian).
5. **Smirnov N.V., Dunin-Barkovskiy I.V.** Kratkiy Kurs Matematicheskoy Statistiki dlya Tekhnicheskikh Prilozheniy. M.: Fizmatlit, 1959. (in Russian).
6. **Prokhorov Yu.V., Ponomarenko L.S.** Lektsii po Teorii Veroyatnostey ш Matematicheskoy Statistike. M.: Izd-vo MGU im. M.V. Lomonosova, 2004. (in Russian).
7. **Barra Zp.-R.** Osnovnye Ponyatiya Matematicheskoy Statistiki. M.: Mir, 1974. (in Russian).
8. **De Groot M.** Optimal'nye Statisticheskie Resheniya. M.: Mir, 1974. (in Russian).

Сведения об авторах:

Архангельский Алексей Николаевич — кандидат физико-математических наук, доцент кафедры высшей математики НИУ «МЭИ», e-mail: ArkhangelskyAN@mpei.ru
Пиголкин Генрих Михайлович — кандидат физико-математических наук, доцент кафедры высшей математики НИУ «МЭИ»

Information about authors:

Arkhangelsky Aleksey N. — Ph.D. (Phys.-Math.), Assistant Professor of Higher Mathematics Dept., NRU MPEI, e-mail: ArkhangelskyAN@mpei.ru
Pigolkin Genrikh M. — Ph.D. (Phys.-Math.), Assistant Professor of Higher Mathematics Dept., NRU MPEI

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов
Conflict of interests: the authors declare no conflict of interest

Статья поступила в редакцию: *26.01.2018*
The article received to the editor: *26.01.2018*