

чем в СОК. Это позволяет использовать для корректирующих ИМ кодов предельно простые проверочные соотношения и алгоритмы декодирования.

Список литературы

1. Касами Т., Токура Н., Ивадари Е., Инагаки Я. Теория кодирования. М., 1978.
2. Журавлев Ю. П., Котелюк Л. А., Циклинский Н. И. Надежность и контроль ЭВМ. М., 1978.
3. Дадаев Ю. Г. Теория арифметических кодов. М., 1981.
4. Акушский И. Я., Юдицкий Д. И. Машинная арифметика в остаточных классах. М., 1968.
5. Торгашев В. А. Система остаточных классов и надежность ЦВМ. М., 1973.
6. Mandelbaum D. M. // IEEE Trans. Inform. Theory. 1976. V. 23. N. 1. P. 85.
7. Краснобаев В. А., Давыдов И. Б. Устройство для обнаружения ошибок в системе остаточных классов СОК: А. с. 1013957 СССР // БИ 1983. № 15.
8. Коляда А. А. Устройство для исправления ошибок в непозиционном коде: А. с. 1136165 СССР // БИ 1985. № 3.

Поступила в редакцию 10.03.86.

УДК 519.234.8

Т. Н. ДЫЛЬКО

ВЫБОР КОЭФФИЦИЕНТОВ СТАТИСТИКИ ДЛЯ ПРОВЕРКИ ГИПОТЕЗЫ ОДИНАКОВОЙ РАСПРЕДЕЛЕННОСТИ В ЗАДАЧАХ АНАЛИЗА ДИХОТОМИЧЕСКОЙ ИНФОРМАЦИИ

1. При формировании группы экспертов по принципу согласованности мнений [1] возникает задача проверки гипотезы о равенстве распределений.

Например, группа экспертов проводит классификацию изделий в N классов (изделия 1-го сорта, 2-го и т. д.) при m -кратной экспертизе, причем классификация проводится независимо между экспертами и между экспертизами. Ответы j -го эксперта можно описывать набором независимых случайных величин следующим образом: $\alpha_{ts}^j = 1$ с вероятностью P_{ts}^j , если при s -й экспертизе изделие отнесено в t -й класс и $\alpha_{ts}^j = 0$ с вероятностью $1 - P_{ts}^j$ в противном случае.

Случайные величины $\{\alpha_{ts}^j, t = 1, \dots, N; s = 1, \dots, m\}$ при сделанных предположениях независимы в совокупности и представляют собой конечную последовательность независимых испытаний Бернулли, т. е. образуют люсиан [2].

В соответствии с методом, предложенным в [1] и подробно изложенным в [2], строится критерий для проверки гипотезы одинаковой распределенности, основанный на расстоянии между люсианами и содержащий некоторые коэффициенты. Возникает вопрос: каким условиям должны удовлетворять эти коэффициенты, чтобы функция распределения статистики была близка к предельной функции распределения (в пределах заданной точности)?

2. **Математическая постановка задачи.** Рассматривается последовательность n люсианов $A_j, j = 1, \dots, n, A_j = \{\alpha_{11}^j, \dots, \alpha_{N1}^j, \alpha_{12}^j, \dots, \alpha_{N2}^j, \dots, \alpha_{1m}^j, \dots, \alpha_{Nm}^j\}, N \geq 2$, где α_{ts}^j определены в п. 1. Проверяется гипотеза

$$H_0: P_{ts}^j = P_{ts}, t = 1, \dots, N, s = 1, \dots, m \text{ для любого } j. \quad (1)$$

(В рассмотренном выше примере это соответствует согласованности мнений экспертов, т. е. тому, что вероятность классификации изделия в один класс одинакова для всех экспертов). Рассматривается асимпто-

тика $n = \text{const}$, $m \rightarrow \infty$, т. е. число параметров P_{ts} растет вместе с объемом выборки.

3. Построение критерия. При построении статистики для проверки гипотезы (1) используется расстояние между ляснанами: $d(A_k, A_q) = \sum_{t=1}^N \sum_{s=1}^m |\alpha_{ts}^k - \alpha_{ts}^q|$, удовлетворяющее естественной системе аксиом [1].

Введем некоторые обозначения:

$$\gamma_{ts} = \sum_{j=1}^n \alpha_{ts}^j \quad (2)$$

(в рассматриваемом примере это общее число голосов, отданных всеми экспертами за t -й класс при s -й экспертизе)

$$\beta_{ts} = \frac{2}{n-1} \left[\gamma_{ts} - \frac{\gamma_{ts}^2}{n} \right], \quad \xi_{ts}^{kq} = |\alpha_{ts}^k - \alpha_{ts}^q| - \beta_{ts}, \quad (3) - (4)$$

$$\eta_{ts} = \sum_{(k,q) \in G_1} c_{kq} \xi_{ts}^{kq}, \quad (5)$$

G_1 — множество, состоящее из $n_1 = \left[\frac{n}{2} \right]$ пар (k, q) , не имеющих общих элементов.

$$D\eta_{ts} = \sum_{i=1}^{n_1} c_{k_i q_i}^2 D\xi_{ts}^{k_i q_i} + \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} c_{k_i q_i} c_{k_j q_j} E \{ \xi_{ts}^{k_i q_i} \xi_{ts}^{k_j q_j} \}, \quad (6)$$

$$D\xi_{ts}^{k_i q_i} = \left(2 - \frac{4}{n} \right) P_{ts} (1 - P_{ts}) - \frac{4(n-2)(n-3)}{n(n-1)} P_{ts}^2 (1 - P_{ts})^2, \quad (7)$$

$$E \{ \xi_{ts}^{k_i q_i} \xi_{ts}^{k_j q_j} \} = \frac{-4}{n} P_{ts} (1 - P_{ts}) + \frac{8(2n-3)}{n(n-1)} P_{ts}^2 (1 - P_{ts})^2. \quad (8)$$

При практическом использовании величины (6) — (8) заменяются их несмещенными оценками, которые вычисляются аналогично [2]. Обозначим S_{ts} несмещенную оценку $D\eta_{ts}$. Тогда в соответствии с [2, 3] имеем следующую статистику для проверки (1):

$$\xi = \frac{\sum_{s=1}^m \sum_{t=1}^N \eta_{ts}}{\left(\sum_{s=1}^m \sum_{t=1}^N S_{ts} \right)^{1/2}}. \quad (9)$$

Как вытекает из работ [1, 2], распределение ξ при росте m сходится к стандартному нормальному распределению $\Phi(x)$. Обозначим $F_m(x) = P(\xi \leq x)$.

4. Ограничения на коэффициенты c_{kq} . Получим ограничения на коэффициенты c_{kq} , используя обобщенное неравенство Эссеена. Мы предлагаем выбирать коэффициенты c_{kq} таким образом, чтобы

$$\sup_x |F_m(x) - \Phi(x)| \leq \alpha^v, \quad (10)$$

где $0 < \alpha < 1$ — уровень значимости, задаваемый статистиком при проверке гипотезы (1) по распределению статистики ξ , $v \geq 1$. В противном случае гипотеза H_0 может быть необоснованно отвергнута или принята.

Для введенных в (5) случайных величин η_{ts} обобщенное неравенство Эссеена имеет вид:

Теорема [4. С. 139]. Пусть η_{ts} , $t=1, \dots, N$, $s=1, \dots, m$ — независимые случайные величины.

$$E\eta_{ts} = 0, E(|\eta_{ts}|^3) < \infty \text{ для } t=1, \dots, N, s=1, \dots, m. \quad (11)$$

$$\sigma_{ts}^2 = E\eta_{ts}^2 = D\eta_{ts}, \quad B_m = \sum_{s=1}^m \sum_{t=1}^N \sigma_{ts}^2 = \sum_{s=1}^m \sum_{t=1}^N D\eta_{ts}, \quad (12)$$

$$F_m(x) = P \left\{ B_m^{-1/2} \sum_{s=1}^m \sum_{t=1}^N \eta_{ts} < x \right\}.$$

Тогда

$$\sup_x |F_m(x) - \Phi(x)| \leq \frac{A \cdot \sum_{s=1}^m \sum_{t=1}^N E(|\eta_{ts}|^3)}{(B_m)^{3/2}}, \quad (13)$$

где $A = 0,9051$.

Ясно, что (11) выполняется. Оценим сверху $E(|\eta_{ts}|^3)$:

$$\begin{aligned} E(|\eta_{ts}|^3) &\leq \sum_{i=1}^{n_1} |c_{k_i q_i}|^3 E\{|\xi_{ts}^{k_i q_i}|^3\} + \sum_{\substack{i=1 \\ i \neq j}}^{n_1} \sum_{j=1}^{n_1} |c_{k_i q_i}| \times \\ &\times |c_{k_j q_j}|^2 E\{|\xi_{ts}^{k_i q_i}| |\xi_{ts}^{k_j q_j}|^2\} + \sum_{\substack{i=1 \\ i \neq j \neq s_1}}^{n_1} \sum_{j=1}^{n_1} \sum_{s_1=1}^{n_1} |c_{k_i q_i}| \times \\ &\times |c_{k_j q_j}| |c_{k_{s_1} q_{s_1}}| \times E\{|\xi_{ts}^{k_i q_i}| |\xi_{ts}^{k_j q_j}| |\xi_{ts}^{k_{s_1} q_{s_1}}|\}, \end{aligned} \quad (14)$$

причем последняя сумма равна нулю, если $n < 6$.

Теорема 1. При $n \geq 6$ справедливы равенства:

$$E\{|\xi_{ts}^{kq}|^3\} = \sum_{i=0}^2 \sum_{l=0}^{n-2} |y_l^{(i)}|^3 C_{n-2}^l P_{ts}^l (1 - P_{ts})^{n-2-l} P(A_i^{kq}), \quad (15)$$

$$\begin{aligned} E\{|\xi_{ts}^{kq}| |\xi_{ts}^{k_1 q_1}|^2\} &= \sum_{i=0}^2 \sum_{j=0}^2 \sum_{l=0}^{n-4} |y_l^{(i,j)}| \times (l_j - l_i + y_l^{(i,j)})^2 \times \\ &\times C_{n-4}^l P_{ts}^l (1 - P_{ts})^{n-4-l} \times P(A_i^{kq}) P(A_j^{k_1 q_1}), \end{aligned} \quad (16)$$

$$\begin{aligned} E\{|\xi_{ts}^{kq}| |\xi_{ts}^{k_1 q_1}| |\xi_{ts}^{k_2 q_2}|\} &= \sum_{i=0}^2 \sum_{j=0}^2 \sum_{s_1=0}^2 \sum_{l=0}^{n-6} |y_l^{(i,j,s_1)}| \times \\ &\times |l_j - l_i + y_l^{(i,j,s_1)}| \times |l_{s_1} - l_i + y_l^{(i,j,s_1)}| \times C_{n-6}^l P_{ts}^l \times \\ &\times (1 - P_{ts})^{n-6-l} P(A_i^{kq}) P(A_j^{k_1 q_1}) P(A_{s_1}^{k_2 q_2}), \end{aligned} \quad (17)$$

где

$$A_i^{kq} : \alpha_{ts}^k + \alpha_{ts}^q = i, \quad i = 0, 1, 2, \quad l_i = \begin{cases} i, & i = 0, 1, \\ 0, & i = 2, \end{cases} \quad (18)-(19)$$

$$y_l^{(i)} = l_i + \tilde{y}_l^{(i)}, \quad y_l^{(i,j)} = l_i + \tilde{y}_l^{(i,j)}, \quad (20)$$

$$y_l^{(i,j,s_1)} = l_i + \tilde{y}_l^{(i,j,s_1)}, \quad \tilde{y}_l^{(i)} = -\frac{2}{n-1} \left[l + i - \frac{1}{n} (l+i)^2 \right],$$

$$P(A_0^{kq}) = (1 - P_{ts})^2, \quad P(A_1^{kq}) = 2P_{ts}(1 - P_{ts}), \quad P(A_2^{kq}) = P_{ts}^2. \quad (21)$$

Доказательство. Проведем для формулы (16). Формулы (15) и (17) получаются аналогично.

Введем обозначение:

$$z = \sum_{j=1}^n \alpha_{ts}^j, \quad j \neq k, k_1, q, q_1. \quad (22)$$

Тогда, учитывая (22), (2) — (4), имеем:

$$\xi_{ts}^{kq} = |\alpha_{ts}^k - \alpha_{ts}^q| - \frac{2}{n-1} \left[z + \alpha_{ts}^k + \alpha_{ts}^q + \alpha_{ts}^{k_1} + \alpha_{ts}^{q_1} - \right.$$

$$-\frac{1}{n} (z + \alpha_{ts}^k + \alpha_{ts}^q + \alpha_{ts}^{k_1} + \alpha_{ts}^{q_1})^2 \Big].$$

Ясно, что целочисленная случайная величина $z \in \{0, 1, \dots, n-4\}$. Поэтому случайная величина ξ_{ts}^{kq} при выполнении условий A_i^{kq} и $A_j^{k_1q_1}$ может принимать значения:

$$\xi_{ts}^{kq} = l_i - \frac{2}{n-1} \left[l + i + j - \frac{1}{n} (l + i + j)^2 \right] = y_l^{(i,j)},$$

$$l = 0, 1, \dots, n-4. \quad (23)$$

В силу определения множества G_1 , $P(A_i^{kq}, A_j^{k_1q_1}) = P(A_i^{kq}) P(A_j^{k_1q_1})$. Выразим β_{ts} из (23), подставим в выражение для вычисления $\xi_{ts}^{k_1q_1}$, вычислим $E\{|\xi_{ts}^{kq}| |\xi_{ts}^{k_1q_1}|^2\}$ по формуле полного математического ожидания и получим (16). (21) получаем, учитывая (18) и определение случайных величин α_{ts}^j .

Теорема 2. Если параметры c_{kq} статистики (9) удовлетворяют условиям:

$$c_{kq} \leq a_m, (k, q) \in G_1, \sum_{G_1} c_{kq} \leq b, c_{kq} \geq 0, \sum_{G_1} c_{kq}^2 \geq 1, \quad (24) - (25) - (26)$$

где b удовлетворяет условию:

$$1 < b < \sqrt{\frac{n}{4} \left(2 - \frac{4(n-2)(n-3)}{n(n-1)} \max_{P_{ts}} P_{ts} (1 - P_{ts}) \right)}, \quad (27)$$

$$a_m = \frac{\alpha^v M_4}{A \cdot M_1} - \frac{b(M_2 + (n_1 - 1)M_3)}{M_1}, \quad (28)$$

$$M_1 = \sum_{s=1}^m \sum_{t=1}^N E\{|\xi_{ts}^{kq}|^3\}, M_2 = \sum_{s=1}^m \sum_{t=1}^N E\{|\xi_{ts}^{kq}| |\xi_{ts}^{k_1q_1}|^2\}, \quad (29)$$

$$M_3 = \sum_{s=1}^m \sum_{t=1}^N E\{|\xi_{ts}^{kq}| |\xi_{ts}^{k_1q_1}| |\xi_{ts}^{k_2q_2}|\}, \quad (30)$$

$$M_4 = \left(\sum_{s=1}^m \sum_{t=1}^N \left[2 - \frac{4b^2}{n} - \frac{4(n-2)(n-3)}{n(n-1)} P_{ts} (1 - P_{ts}) \right] P_{ts} (1 - P_{ts}) \right)^{3/2}, \quad (31)$$

$A=0,9051$, $0 < \alpha < 1$ — уровень значимости проверки гипотезы (1) по статистике (9), $v \geq 1$, то выполняется (10).

Если объем выборки m удовлетворяет условию

$$a_m > \frac{1}{\sqrt{n_1}}, \quad (32)$$

то множество решений (24) — (26) не пусто.

Доказательство. Оценим числитель (13) сверху, используя (25) и неравенство $\left(\sum_{i=1}^{n_1} c_{k_iq_i} \right) \leq n_1 \sum_{i=1}^{n_1} c_{k_iq_i}^2$. Знаменатель оценим снизу, учитывая (25), (26). Используя обозначения (29) — (31) и требуя выполнения (10), получаем:

$$\sum_{i=1}^{n_1} \left\{ \left[\frac{\alpha^v M_4}{A \cdot M_1} - \frac{b}{M_1} (M_2 + (n_1 - 1)M_3) \right] c_{k_iq_i}^2 - c_{k_iq_i}^3 \right\} \geq 0. \quad (33)$$

Введем обозначения:

$$c_{k_iq_i} = \Theta_i, i = 1, \dots, n_1, \quad (34)$$

$$\Theta = \begin{bmatrix} \Theta_1 \\ \vdots \\ \Theta_{n_1} \end{bmatrix} - n_1\text{-вектор, } \Xi = \begin{bmatrix} \Theta_1 & 0 \\ & \ddots \\ 0 & \Theta_{n_1} \end{bmatrix}, M = \begin{bmatrix} a_m & 0 \\ & \ddots \\ 0 & a_m \end{bmatrix} \quad (35)$$

— $n_1 \times n_1$ -матрицы.

Тогда, учитывая (34) — (35), (33) запишем в виде:

$$\Theta' \{ \Xi - M \} \Theta \geq 0. \quad (36)$$

Согласно [5], квадратичная форма (36) неотрицательно определена тогда и только тогда, когда матрица этой квадратичной формы имеет неотрицательные характеристические числа, т. е. получаем (24).

Неравенство (27) получаем, учитывая (26) и требуя, чтобы подкоренное выражение в M_4 было больше нуля; (32) следует из неравенств (26) и (24). Подставляя в (32) выражение из (28), имеем неравенство относительно m .

5. Пример. При $n=4$, $N=3$, $\alpha=0,1$, $\nu=1$, $P_{1s}=0,5$, $P_{2s}=P_{3s}=0,25$, $s=1, \dots, m$, $b=1,01$, $a_m=1,0000129$, $m=5637$, коэффициенты c_{12} , c_{34} можно выбрать следующим образом: $c_{12}=1$, $c_{34}=0,001$.

Теоретическая оценка объема выборки m по (10) является завышенной. Так, если подставить полученные значения m , c_{hv} в (13), будем иметь: $\sup_x |F_m(x) - \Phi(x)| \leq 0,0345268$, т. е. на порядок лучше, чем предполагали.

Список литературы

1. Орлов А. И. Устойчивость в социально-экономических моделях. М., 1979.
2. Орлов А. И. // Алгоритмическое и программное обеспечение прикладного статистического анализа. М., 1980.
3. Орлов А. И. // Экспертные оценки в задачах уравнения. М., 1982.
4. Петров В. В. Суммы независимых случайных величин. М., 1972.
5. Мальцев А. И. Основы линейной алгебры. М., 1970.

Поступила в редакцию 16.04.86.

УДК 517.968.7

У. У. ДИАЛЛО

К ТЕОРИИ ЛИНЕЙНЫХ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В ПРОСТРАНСТВЕ НЕПРЕРЫВНЫХ ФУНКЦИЙ

В настоящей работе изучается линейное дифференциальное уравнение

$$\frac{dx}{dt} = A(t)x \quad (1)$$

с действующим в пространстве C непрерывных на отрезке $[a, b]$ вещественных функций оператором

$$A(t)x(s) = c(t, s)x(s) + \int_a^b k(t, s, \sigma)x(\sigma)d\sigma. \quad (2)$$

Здесь $c(t, s)$ и $k(t, s, \sigma)$ — заданные на $\mathbf{R} \times [a, b]$ и $\mathbf{R} \times [a, b] \times [a, b]$ вещественные измеримые функции.

Уравнения вида (1) наряду с бесконечными системами линейных уравнений являются естественным бесконечномерным аналогом конечных линейных систем. Однако если бесконечным системам посвящена значительная литература (см., например, [1]), интегро-дифференциальные уравнения (1), несмотря на то, что встречаются во многих задачах анализа (в частности, в теории вероятностей), почти не исследовались (отдельные результаты см., например, [2—5]).