УДК 004.421.6: 519.23

ОБ ИМИТАЦИИ СЛУЧАЙНЫХ ДАННЫХ ДЛЯ ОЦЕНКИ КАЧЕСТВА СТАТИСТИЧЕСКИХ ТЕСТОВ В КРИПТОГРАФИИ

В. Ю. Палуха[1)], Н. А. Прохорчик[2)], Ю. С. Харин[3)]

Учреждение Белорусского государственного университета «НИИ прикладных проблем математики и информатики»,
пр. Независимости, 4, 220030, г. Минск, Беларусь, [1)] palukha@bsu.by, [2)] prohorchikNA@bsu.by, [3)] kharin@bsu.by

Рассматривается задача моделирования вектора, задающего распределение вероятностей, равномерно распределённое на гиперсфере заданного радиуса.

Ключевые слова: математическое моделирование, распределение вероятностей, гиперсфера.

Введение

Пусть нам необходимо смоделировать двоичный временной ряд с K-мерным распределением вероятностей, которое выбирается случайным равновероятным образом среди расположенных на гиперсфере заданного радиуса. Такая задача может возникнуть, например, для генерации модельных данных, позволяющих проверить работоспособность статистического теста, проверяющего сложную нулевую гипотезу [1]. Для этого необходимо решить задачу моделирования такого вектора вероятностей. В данной статье предложен алгоритм генерации искомого вектора вероятностей.

1. Математическая модель

Введём обозначения: $p = (p_0,\dots, p_{K-1})'$ – вектор-столбец вероятностей размера K, $\tilde{p} = (p_0,\dots, p_{K-2})'$ – вектор-столбец, состоящий из первых K – 1 координат вектора p, $p_* = \left(\frac{1}{K},\dots,\frac{1}{K}\right)'$ – вектор-столбец равномерного распределения вероятностей размера K, I_K – единичная матрица размера $K\times K$, 1_K – вектор размера K, в котором все элементы равны 1, $1_{K\times K}$ – матрица размера $K\times K$, в которой все элементы равны 1.

Обозначим искомое множество векторов распределений вероятностей как

$$P_0^{\varepsilon} = \left\{ p = (p_k) : p_k \ge 0, \sum_{k=0}^{K-1} p_k = 1, \sum_{k=0}^{K-1}\left(p_k - \frac{1}{K}\right)^2 = \varepsilon^2 \right\}, \qquad (1)$$

представляющее собой пересечение единичного симплекса и гиперсферы радиуса $\varepsilon > 0$ с центром в точке p_*.

Обозначим через C_K квадратную матрицу размера $K \times K$ вида

$$C_K = \left(c_{ij}\right) = \begin{pmatrix} 2\,1\dots 1 \\ 1\,2\dots 1 \\ \dots\dots\dots \\ 1\,1\dots 2 \end{pmatrix} = I_K + 1_{K\times K}, \quad c_{ij} = \begin{cases} 2, & i = j; \\ 1, & i \ne j. \end{cases} \qquad (2)$$

Представим C_K в виде произведения

$$C_K = \left(C_K^{1/2}\right)' C_K^{1/2}, \tag{3}$$

где $C_K^{1/2}$ – верхнетреугольная матрица, т.е. в виде разложения Холецкого [2].

Теорема 1. Для координат вектора *p*, равномерно распределённого на множестве (1), справедливы выражения

$$p_k = \begin{cases} \dfrac{1}{K} + \varepsilon\left(C_{K-1}^{-1/2}\xi\right)_k, & k = 0, 1, \ldots, K-2, \\ p_{K-1} = 1 - \sum\limits_{i=0}^{K-2} p_i, \end{cases} \tag{4}$$

где ξ – вектор, равномерно распределённый на гиперсфере единичного радиуса в $\mathbb{R}^{K-1}$, и выполняются ограничения

$$p_k \geq 0, \quad \sum_{k=0}^{K-2} p_k \leq 1, \quad k = 0, 1, \ldots, K-2. \tag{5}$$

Для того, чтобы воспользоваться выражениями (4), необходимо вычислить матрицу $C_{K-1}^{-1/2}$. Воспользуемся формулами для обратной матрицы к треугольной из [3].

Лемма 1. Для элементов матрицы $C_K^{-1/2} = \left(\gamma_{ij}\right)$ справедливо представление

$$\gamma_{ij} = \begin{cases} -\dfrac{1}{\sqrt{j(j+1)}}, & i < j; \\ 0, & i > j; \\ \sqrt{\dfrac{i}{i+1}}, & i = j. \end{cases} \tag{6}$$

Следствие 1. Алгоритм моделирования вектора *K*-мерного распределения вероятностей, который выбирается случайным равновероятным образом среди расположенных на гиперсфере радиуса ε состоит из следующих шагов:

1) Вычислить матрицу $C_{K-1}^{-1/2}$ по формуле (6);

2) Сгенерировать вектор ξ, равномерно распределённый на гиперсфере единичного радиуса в $\mathbb{R}^{K-1}$ методом из §3.13 [4];

3) Методом исключения сгенерировать вектор *p* по формуле (4): в случае, если не выполняется условие (5), вернуться к шагу 2.

2. Результаты компьютерных экспериментов

Для демонстрации работы алгоритма произведено моделирование векторов при $K = 3$. Данное значение *K* позволяет графически отобразить сгенерированные точки. В соответствии с алгоритмом сгенерировано 1000 векторов при $K = 3$ и $\varepsilon = 0{,}05$. Как следует из (1), при $K = 3$ искомые вектора лежат на окружности радиуса ε с центром в точке (1/3, 1/3, 1/3) на плоскости $x + y + z = 1$. Расположение сгенерированных трёхмерных векторов в пространстве показано на рис. 1. Как видно из рисунка, точки действительно лежат на заданной выше окружности.

Для того, чтобы убедиться, что сгенерированные точки действительно расположены равномерно на допустимом множестве, перейдём к полярным координатам и построим гистограмму значений угла φ. Для этого нужно преобразовать декартовы координаты таким образом, чтобы точки лежали на окружности с центром (0, 0, 0) на плоскости *Oxy*. Сначала вычтем из каждой координаты 1/3, получим окружность радиуса ε с центром в точке (0, 0, 0) на плоскости $x + y + z = 0$. Данная плоскость пересекает плоскость $z = 0$ по прямой $y = -x$ под углом $\psi = \arccos\dfrac{1}{\sqrt{3}}$ [5]. Для того, чтобы преобразовать плоскость $x + y + z = 0$ к плоскости $z = 0$,

необходимо выполнить композицию поворотов: на угол $\frac{\pi}{4}$ по оси *Oz* и на угол $\psi = \arccos\frac{1}{\sqrt{3}}$ по оси *Ox*. Матрица поворота равна [6]

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \qquad (7)$$

После умножения нормированных векторов на матрицу (7) их третьи координаты будут равны 0. Теперь мы можем перейти к полярным координатам в соответствии с известными формулами [5]. Гистограмма значений угла φ полученных точек приведена на рис. 2. Как видно из этого рисунка, распределение точек действительно близко к равномерному.

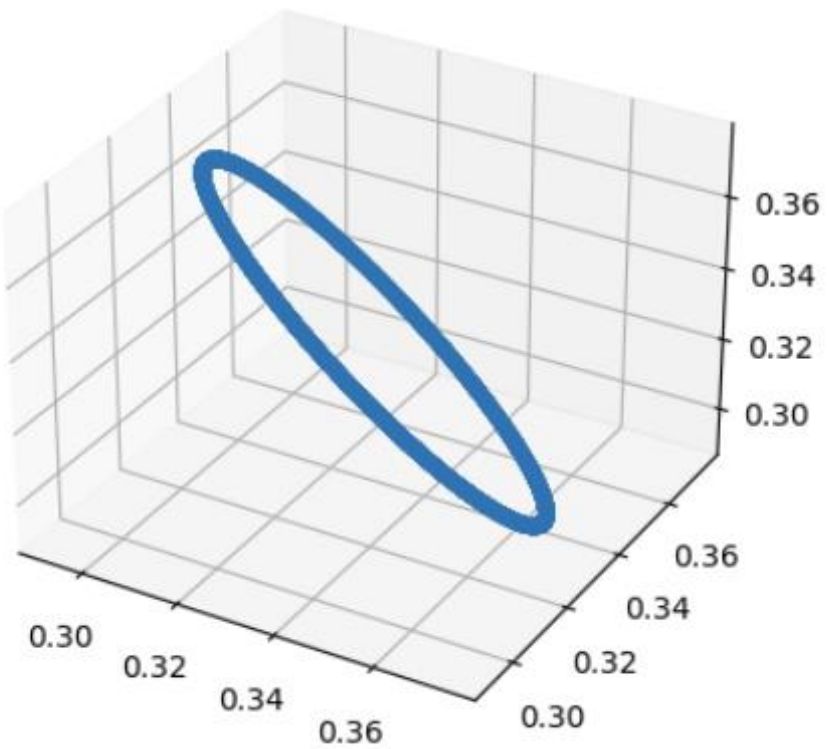

Рис. 1. Расположение в пространстве сгенерированных векторов

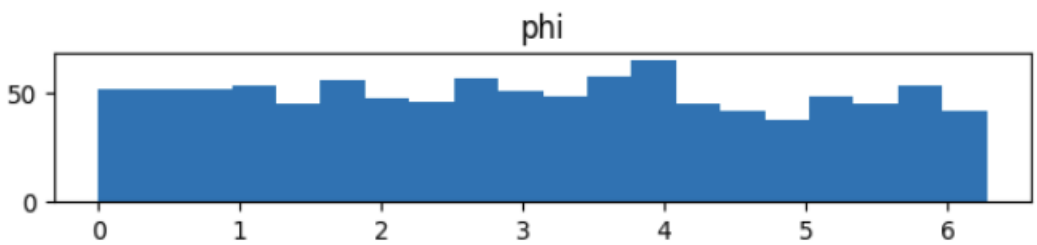

Рис. 2. Гистограмма угла φ

Библиографические ссылки

1. Палуха В. Ю., Харин Ю. С. Статистическое тестирование криптографических генераторов на основе сложной нулевой гипотезы // Теоретическая и прикладная криптография: материалы II Международной научной конференции, Минск, 19-20 октября 2023 г. / Белорусский государственный университет; редколлегия: Ю. С. Харин (гл. ред.) [и др.]. Минск: БГУ, 2023. С. 185-193.

2. Прудников А. П., Брычков Ю. А, Маричев О. И. Интегралы и ряды. Москва: Наука, 1981. 800 с.

3. Каплан И. А. Практические занятия по высшей математике. Часть V. Харьков: Издательство Харьковского университета, 1972. 413 с.

4. Харин Ю. С., Степанова М. Д. Практикум на ЭВМ по математической статистике. Минск: Университетское, 1987. 304 с.

5. Милованов М. В., Тышкевич Р. И., Феденко А. С. Алгебра и аналитическая геометрия: Часть 1. Минск: Вышэйшая школа, 1984. 302 с.

6. Лурье А. И. Аналитическая механика. Москва: Физматлит, 1961. 824 с.