Совместное асимптотическое распределение статистик хи-квадрат тестов качества криптографических генераторов

В.А. Волошко

Для двоичных последоваельностей $\mathbf{x}_t \in \mathbf{V} ::= \{0,1\}$, $t \in \mathbb{Z}$, рассматривается многообразие $\mathbb{M}^\pi$ вероятностных мер периодических цепей Маркова

$$\mathsf{P}\{\mathbf{x}_t = 0 | \mathbf{x}_{t-i} = q_i, i \in \mathbb{N}\} = \frac{1+\boldsymbol{\omega}_t(q)}{2}, \quad \|\boldsymbol{\omega}\|_\infty = \max_{t,q} |\boldsymbol{\omega}_t(q)| < 1, \quad q = (q_i)_{i\in\mathbb{N}} \in \mathbf{V}^{\mathbb{N}},$$

где функция $\boldsymbol{\omega}_t(q)$ Δ-периодична по $t \in \mathbb{Z}$ и зависит лишь от конечного числа компонент $(q_i)_{i=1}^s$ для некоторых $s =:: \mathtt{ord}(\boldsymbol{\omega})$, $\Delta =:: \mathtt{per}(\boldsymbol{\omega})$, а распределение последовательности $\mathbf{x}$ инвариантно к сдвигу на период $(x_t) \mapsto (x_{t+\Delta})$. Равномерному распределению $\mathtt{u} \in \mathbb{M}^\pi$ (нулевая гипотеза) отвечает нулевая функция $\boldsymbol{\omega}_t(q) \equiv 0$, а освобожденные от условия $\|\boldsymbol{\omega}\|_\infty < 1$ функции $\boldsymbol{\omega}$ образуют бесконечномерное касательное пространство $\mathbb{T}^\pi$ (линеаризованную малую окрестность) многообразия $\mathbb{M}^\pi$ в точке $\mathtt{u}$ ($\|\boldsymbol{\omega}\|_\infty$ конечны). На касательном пространстве $\mathbb{T}^\pi$ определено скалярное произведение Фишера-Римана $\langle \boldsymbol{\omega}, \boldsymbol{\omega}' \rangle = \Delta^{-1} 2^{-s} \sum_{t=1}^{\Delta} \sum_{q\in\mathbf{V}^s} \boldsymbol{\omega}_t(q) \boldsymbol{\omega}'_t(q)$, $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathbb{T}^\pi$, где $\Delta = \mathtt{HOK}(\mathtt{per}(\boldsymbol{\omega}), \mathtt{per}(\boldsymbol{\omega}'))$ – общий период $\boldsymbol{\omega}$ и $\boldsymbol{\omega}'$, $s = \max\{\mathtt{ord}(\boldsymbol{\omega}), \mathtt{ord}(\boldsymbol{\omega}')\}$.

С каждым конечномерным касательным подпространством $\mathtt{T} \subset \mathbb{T}^\pi$ ассоциирован класс асимптотически эквивалентных статистик $\mathtt{S}_\mathtt{T} = \mathtt{S}_\mathtt{T}(\mathbf{x}_1^n)$, имеющих при $n \to +\infty$ и истинной нулевой гипотезе асимптотическое распределение $\chi^2_{\dim(\mathtt{T})}$ (нецентральное при контигуальной альтернативе). В семействе $\{\mathtt{S}_\mathtt{T}\}_{\mathtt{T}\subset\mathbb{T}^\pi}$ лежат статистики (либо их эквивалентные аналоги) всех известных «χ^2-тестов», основанных на частотах конечных битовых цепочек: тестов многомерной равномерности по пересекающимся и непересекающимся цепочкам, тестов независимости и однородности, энтропийных тестов, «монобит» и «знакоперемен». Не лежит в этом семействе, например, статистика спектрального теста (NIST Discrete Fourier Transform), использующая частотные характеристики битовых цепочкек произвольной длины. Для этого теста, однако, построенная теория также позволяет получить новые результаты.

Доклад посвящён совместному асимптотическому распределению статистик из семейства $\{\mathtt{S}_\mathtt{T}\}_{\mathtt{T}\subset\mathbb{T}^\pi}$. Для каждой пары пространств $\mathtt{T}^{(1)}, \mathtt{T}^{(2)} \subset \mathbb{T}^\pi$ асимптотическое распределение пары $(\mathtt{S}_{\mathtt{T}^{(1)}}, \mathtt{S}_{\mathtt{T}^{(2)}})$ задаётся количественными геометрическими характеристиками $\Lambda(\mathtt{T}^{(1)}, \mathtt{T}^{(2)})$ совместного расположения пространств. Во многих случаях удается явно найти характеристики Λ и выразить из них асимптотические функционалы зависимости статистик $(\mathtt{S}_{\mathtt{T}^{(1)}}, \mathtt{S}_{\mathtt{T}^{(2)}})$, такие как коэффициент корреляции или, при необходимости, смешанные моменты высоких порядков.