Санкт-Петербургский государственный университет

Овечкин Григорий Владимирович

Выпускная квалификационная работа

Оценки времени релаксации последовательности независимых случайных величин при соответствии RSK

Образовательная программа бакалавриат «Математика»
Направление и код: 01.03.01 «Математика»
Шифр ОП: СВ.5000.2016

Научный руководитель:
главный научный сотрудник
Санкт-Петербургское отделение
Математического института
им. В. А. Стеклова РАН
доктор ф.-м. наук
Вершик Анатолий Моисеевич

Рецензент:
доцент
Академический университет
им. Ж.И. Алферова
кандидат ф.-м. наук
Горячко Евгений Евгеньевич

Санкт-Петербург
2020 год

Оглавление

1 Определения и постановка задачи

Рассматривается классическая схема Бернулли — последовательность независимых случайных величин одинаково распределенных (i.i.d) по мере Лебега m на отрезке $[0,1]$. Пространство реализаций этой схемы есть бесконечномерный куб с мерой Лебега $\mu = m^{\mathbb{N}}$: $\mathcal{X} = ([0,1]^{\mathbb{N}}, \mu)$. В работе доказывается (см. теорема 1.1), что при применении алгоритма RSK к $\mathcal{X}$ с любой вероятностью меньшей 1 количество вставок, которое потребуется координате для попадания в первый столбец, можно оценить сверху и снизу квадратичными функциями от номера координаты. Данная работа частично (параграфы 1, 2, 3, 4) содержится в заметке [1], написанной автором совместно с И. Азангуловым.

Конечная версия алгоритма RSK (см. [2], [3], [4]) была обобщена на случай бесконечных последовательностей независимых одинаково распределённых величин в работе [5] А.М. Вершика и С.В. Керова. А именно, доказано, что пространство реализаций схемы Бернулли $\mathcal{X}$ переводится с помощью последовательности нумерующих таблиц (Q-таблиц) в пространство бесконечных стандартных таблиц Юнга, при этом мера Лебега переходит в меру Планшереля. В недавней работе [6] было доказано, что этот гомоморфизм является изоморфизмом пространств с мерой. В этой работе мы изучаем статистику и асимптотику P-таблиц в процессе обобщенного RSK. По любой реализации $\{x_n\}_n \in \mathcal{X}$ строится последовательность конечных записывающих таблиц $\{P_n\}_n$ (P-таблиц), получающихся из начальных отрезков реализации: $x^n = \{x_1, \ldots, x_n\}$, с помощью вставок алгоритма RSK (см. обозначения в [4])

$$P_n = \varnothing \leftarrow x_1 \leftarrow \ldots \leftarrow x_n.$$

В этом процессе координаты реализаций меняют свое последовательное положение в P-таблице, при этом номер столбца, в котором содержится координата, убывает (нестрого). Заметим, что все координаты почти всех реализаций схемы Бернулли различны, и будем предпологать, что пространство $\mathcal{X}$ состоит только из таких реализаций. Тем самым, координата почти любой реализации однозначно определяется своим значением.

А.М. Вершик доказал (см. [7]), что любая координата x_n μ-почти всякой реализации $x \in \mathcal{X}$ попадает в некоторый момент $N(x,n)$ в первый столбец P-таблицы, и поставил задачу об оценке величины $N(x,n)$. Ответ на этот вопрос дает

Теорема 1.1. *Для любого $\varepsilon > 0$ существуют константы $k = k(\varepsilon)$, $K = K(\varepsilon)$ и число $n_0 \in \mathbb{N}$ такие, что оценка $kn^2 \leq N(x,n) \leq Kn^2$ верна для любого натурального числа $n > n_0$ и любой реализации x из некоторого подмножества реализаций $\mathcal{X}_{n,\varepsilon} \subset \mathcal{X}, \mu(\mathcal{X}_{n,\varepsilon}) \geq 1 - \varepsilon$. Более того, можно положить $k(\varepsilon) = c\varepsilon^5$, $K(\varepsilon) = C\varepsilon^{-5}$ для некоторых констант $c, C > 0$.*

2 Разбиение P-таблиц на уровни

Основная идея доказательства заключается в рассмотрении двух последовательностей "срезок" P-таблиц Юнга и оценки значения элемента, находящегося в последней клетке первого столбца. Обозначим через $\{X_n\}_n$ последовательность независимых одинаково распределенных случайных величин по мере Лебега на отрезке $[0, 1]$, а ее ралризации будем обозначать через $x = \{x_n\}_n$ или, если это существенно, $\{X_n(x)\}_n$. Пусть $\{P_n(x)\}_n$ — последовательность P-таблиц Юнга, полученная при помощи алгоритма RSK из реализации x. Для произвольного числа $c \in (0, 1]$ и реализации x рассмотрим последовательность "усеченных" таблиц $\{P_n^c(x)\}_n$, состоящих только из тех координат, которые не больше c:

$$P_n^c(x) = \varnothing \leftarrow x_{i_1} \leftarrow x_{i_2} \leftarrow \ldots \leftarrow x_{i_s},$$

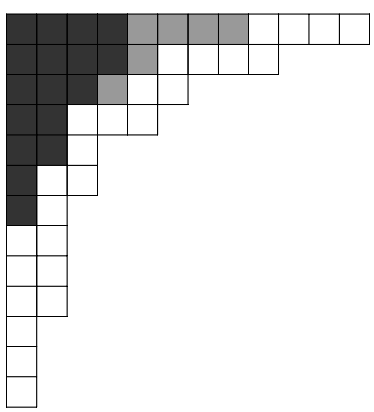
где $\{i_1, \ldots, i_s\} = \{i \leq n : x_i \leq c\}$.

Заметим, что вставки между различными усечениями таблиц согласованы в том смысле, что есть коммутативная диаграмма

$$\begin{tikzcd}
\mathbb{T} \arrow[rr, "i_x"] \arrow[d, "(-)^c"'] & & \mathbb{T} \arrow[d, "(-)^c"] \\
\mathbb{T} \arrow[r, "i_x"] & \mathbb{T} \arrow[r, "(-)^c"] & \mathbb{T},
\end{tikzcd}$$

где $\mathbb{T}$ — множество конечных таблиц Юнга со элементами из отрезка $[0, 1]$, i_x — операция всатвки числа x, а $(-)^c$ — операция усечения.

Из этого следует, что для любого $c \in (0, 1]$ таблица P_n^c есть подтаблица таблицы P_n, и если $c \leq c'$, то P_n^c есть подтаблица таблицы $P_n^{c'}$. На рис. 1 изображен пример срезок P-таблицы для параметров $0 < c < c' < 1$.



Черным цветом отмечены клетки, в которых стоят элементы x, $x \leq c$.
Серым цветом отмечены клетки, в которых стоят элементы x, $c < x \leq c'$.
В оставшихся клетках стоят элементы строго больше, чем c'.

Рис. 1. Разбиение P-таблицы на 3 уровня.

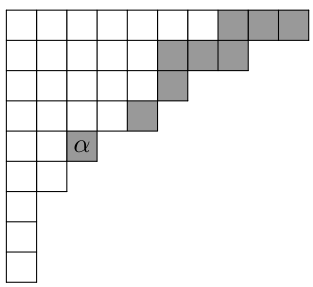
Для произвольного $c \in (0, 1]$ обозначим через $X_{c,n}$ случайную величину $\max\limits_{1 \leq i \leq n, X_i < c} X_i$. Рассмотрим произвольную реализацию x и положим $\alpha = x_n$, $\beta = X_{\alpha,n}(x)$. Тогда интересующие нас срезки — это срезки по параметрам α и β ($\beta < \alpha$). Из сказанного выше следует, что на динамику координаты α не влияет та часть таблицы, в которой стоят координаты большие α. В следующей лемме описываются последовательные положения координаты в зависимости от ее значения.

Лемма 2.1. *Зафиксируем реализацию x и обозначим $\alpha = x_n$, $\beta = X_{\alpha,n}(x)$. Тогда для любого $m \geq n$ в косой таблице $P_m^\alpha \backslash P_m^\beta$ самая нижняя строка содержит единственный элемент α (см. рис. 2).*

Доказательство. Докажем индукцией по m.

База: $m = n$. Косая таблица $P_n^\alpha(x) \backslash P_n^\beta(x)$ содержит единственный элемент α.

Переход: $m \mapsto m+1$. Пусть для m утверждение верно. Обозначим номер строки, содержащей элемент α, в таблице P_m^α через l. Рассмотрим $m+1$-ую вставку Шенстеда. Прежде всего заметим, что можно не обращать внимание на вставки координат больших α, потому как они не меняют таблицу P_m^α. Если процесс выбивания (см. терминологию в [4]) остановился до l-ой строки, то элементы в клетках, стоящих в строках с номерами не меньше l, не поменялись. Значит, утверждение леммы будет выполняться и для P_{m+1}^α. Пусть теперь процесс выбивания дошел до строки с номером l. Рассмотрим два случая: в l-ой строке выбили элемент α и в l-ой строке выбили элемент меньший, чем α. В первом случае элемент α перешел в $l+1$-ую строку, при этом в строчках от $l+1$-ой до последней остальные элементы не поменялись и были по предположению меньше, чем β. Во втором случае выбился элемент не большее, чем β, но его мог выбить только элемент меньший его, значит единственный новый элемент в строках от l-ой до последней меньше, чем β. $\square$



Серым цветом выделена косая таблица $P_m^\alpha(x) \backslash P_m^\beta(x)$.

Рис. 2. Позиция элемента α в процессе изменения P-таблицы.

В терминах доказанной леммы можно дать следующий критерий попадания координаты в первый столбец.

Следствие 2.2. Элемент α лежит в первом столбце таблицы P_N^α, а значит, и таблицы P_N, тогда и только тогда, когда в первом столбце таблицы P_N^α есть элемент больший β.

3 Оценка распределения значения последнего элемента в первом столбце

В данном параграфе рассматриваются только конечные начальные отрезки реализаций $x^n = \{x_i\}_1^n$ последовательности незвисимых случайных величин. Положим $P(x^n) = P_n(x)$. Задача состоит в оценке значения координаты, находящейся в последней клетке первого столбца таблицы $P(x^n)$. Для ее оценки сначала оценивается ее порядковый номер (т.е. место, на котором она бы стояла, если бы мы отсортировали последовательность по возрастанию) в части реализации x^n. Обозначим через Σ_n отображение, сопостовляющее реализации x перестановку из порядковых номеров ее первых n координат. Из предыдущего замечания о значениях координат ясно, что, $\sigma = \Sigma_n(X)$ есть однозначно определенная случайная перестановка, распределенная равномерно на S_n. Для простоты можно отождествить перестановки с элементами группы S_n, которые они задают.

Обозначим через $[n]$ множество $\{1, 2, \ldots, n\}$. Диаграмму, соответствующую таблице, обозначим через $\mathrm{sh}(.)$. А через $\dim(.)$ обозначим количество стандартных таблиц данной диаграммы. Пусть t — произвольная таблица с диаграммой λ ($t \vdash \lambda$). Тогда $\#\{\sigma \in S_n | P(\sigma) = t\} = \dim(\lambda)$. Значит, равномерная мера на S_n переходит в некоторую меру ν_n на P-таблицах, задаваемую по формуле $\nu_n(P) = \dim(\lambda)/(n!)$. Напомним, что это есть мера Планшереля на таблицах.

Для произвольной подстановки $\sigma \in S_{n-1}$ обозначим через $P(\sigma)$ и $f(\sigma)$ ее P-таблицу и последний элемент в первой строке этой таблице, соответственно. Для произвольной диаграммы $\lambda \vdash (n-1)$ обозначим через $\hat{\lambda}$ диаграмму Юнга из n клеток, отличающуюся от диаграммы λ ровно на одну клетку в первой строке.

Лемма 3.1. Математическое ожидание элемента $f(\sigma)$ вычисляется по формуле

$$E_{\nu_{n-1}}(f(\sigma)) = n - \frac{1}{(n-1)!} \sum_{\lambda \vdash (n-1)} \dim(\lambda) \cdot \dim(\hat{\lambda}). \tag{1}$$

Доказательство. По определению математического ожидания имеем

$$n! - (n-1)! E_{\nu_{n-1}}(f) = \sum_{\sigma} (n - f(\sigma)). \tag{2}$$

Остается показать, что

$$\sum_{\sigma} (n - f(\sigma)) = \sum_{\lambda \vdash (n-1)} \dim(\lambda) \cdot \dim(\hat{\lambda}). \tag{3}$$

Это легко доказать, используя биекцию между следующими множествами:

(i) Множество H_1 пар $(\sigma, k + 1/2)$, где $\sigma \in S_{n-1}$, а $k + 1/2 > f(\sigma)$ и $k \in [n-1]$,

(ii) Множество H_2 пар стандартных таблиц юнга (P, Q) одинаковой формы, где таблица P состоит в точности из элементов $1, 2, \ldots, n-1, k+1/2$, где $k \in [n-1]$, причем элемент $k + 1/2$ стоит на последнем месте первой строки, а таблица Q состоит из элементов $[n]$, причем элемент n стоит в первой строке.

Построим отображение $H_1 \to H_2$. Произвольный элемент σ задает пару таблиц $(P(\sigma), Q(\sigma))$, элемент $k + 1/2$ после вставки Шенстеда обязательно встанет в конец первой строки на свободное место. Значит, Q-таблица теперь будет содержать элемент n в первой строке. Перестановка при обратном отображении получается посредством композиции удалений элемента $k+1/2$ из P-таблицы и элемента n из Q-таблицы, а также обратного алгоритма RSK. Элемент $k + 1/2$ восстанавливается однозначно. Несложно видеть, что это взаимно обратные отображения.

Каждому $\sigma \in S_{n-1}$ соответствует ровно $n - x(\sigma)$ пар из первого множества. А значит мощность первого множества равна $\sum_{\sigma}(n - x(\sigma))$. Понятно, что стандартные P-таблицы и P-таблицы из второго множества состоят в естественной биекции, сохраняющей порядок между элементами. Поэтому мощность второго множества равна количеству пар стандартных таблиц Юнга из n клеток одинаковой формы, где Q-таблица содержит элемент n в первой строке. Количество таких пар равно $\sum_{\lambda \vdash (n-1)} \dim(\lambda) \cdot \dim(\hat{\lambda})$. Приравняв мощности, получим

$$\sum_{\lambda \vdash (n-1)} \dim(\lambda) \cdot \dim(\hat{\lambda}) = \sum_{\sigma} (n - f(\sigma)).$$

Вместе с равенством (2) это дает доказательство формулы (1). $\square$

Оценки асимптотики других моментов приведем ниже (см. параграф 5).

Из леммы несложно вывести оценки на $\mathbb{E}_{\nu_n}(f(\sigma))$. Для краткости положим $a_\lambda = \dim(\lambda), b_\lambda = \dim(\hat{\lambda})$. По неравенству Коши-Буняковского

$$(\sum_{\lambda\vdash(n-1)} \frac{a_\lambda b_\lambda}{(n-1)!})^2 \le (\sum_{\lambda\vdash(n-1)} \frac{a_\lambda^2}{(n-1)!})\cdot(\sum_{\lambda\vdash(n-1)} \frac{b_\lambda^2}{(n-1)!}) = n\cdot \sum_{\lambda\vdash(n-1)} \frac{b_\lambda^2}{n!} \le n.$$

Несложными преобразованиями получаем следующее

Следствие 3.2. Для любого $n>1$ верно, что $\frac{1}{\sqrt{n}}(n-\mathbb{E}_{\nu_{n-1}}f(\sigma)) \le 1$.

Вернемся к задаче, поставленную в начале параграфа. Обозначим через $W(x,n-1)$ значение, стоящее в последней клетки первого столбца $P(x^{n-1})$, и пусть $\sigma = \Sigma_{n-1}(x)$. Заметим, что порядковые номера элементов, стоящих в последней клетке первого столбца и последней клетке первой строки, распределены одинаково. Положим, $l \overset{d}{=} f(\sigma)$ есть случайная величина равная порядковому номеру W, где $W = W(x,n-1)$. Тогда

$$\mathbb{E}W = \mathbb{E}X_{(l)} = \sum_i \mathbb{P}\{l=i\}\mathbb{E}X_{(i)} = \sum_i \mathbb{P}\{l=i\}\cdot\frac{i}{n} = \frac{\mathbb{E}l}{n},$$

где $X_{(i)}$ обозначает i-ую порядковую статистику в конечной последовательности $\{X_j\}_1^{n-1}$. Получаем

Следствие 3.3. Имеется следующая оценка распределения случайной величины W ($W = W(x,n-1)$).

$$\mathbb{P}\{W > 1-\frac{t}{\sqrt{n}}\} \ge 1-\frac{1}{t}. \tag{4}$$

Доказательство. Из неравенства Маркова и следствия 3.2

$$\mathbb{P}\{W > 1-\frac{t}{\sqrt{n}}\} = \mathbb{P}\{(1-W)\sqrt{n} < t\} \ge 1-\frac{\mathbb{E}(1-W)\sqrt{n}}{t} \ge 1-\frac{1}{t}. \quad \square$$

4 Оценка сверху

В данном параграфе приводится доказательство оценки сверху на случайную величину $N(X,n)$.

Теорема 4.1. *Для любых $\varepsilon > 0$ и $n\in\mathbb{N}$ существуют константа $\tilde{K} = \tilde{K}(\varepsilon)$ и подмножество реализаций $\mathfrak{X}_{n,\varepsilon} \subset \mathfrak{X}, \mu(\mathfrak{X}_{n,\varepsilon}) \ge 1-\varepsilon$, такие, что для любой реализации $x\in\mathfrak{X}_{n,\varepsilon}$ имеет место оценка $N(x,n) \le \tilde{K}n^2$.*

Напомним, что $\{X_n\}_n$ есть последовательность независимых случайных величин одинаково распределенных по мере Лебега на отрезке $[0,1]$. Обозначим количество элементов не больших c в конечной последовательности $\{X_n\}_1^N$ через $q^c(N)$. Докажем две технические леммы о распределениях случайных величин $X_{c,n}$ и $q^c(N)$.

Лемма 4.2. *Если $0<\varepsilon\le\frac{1}{3}$, то для любого $\alpha\in(0,1]$ верно, что $\mathbb{P}\{X_{\alpha,n} \le \alpha(1-\frac{\varepsilon}{2n})\big|X_n=\alpha\} \ge 1-\varepsilon$.*

Доказательство. Заметим, что если $0<\varepsilon\le\frac{1}{3}$, то $1-\frac{\varepsilon}{2n} \ge \sqrt[n]{1-\varepsilon}$. Пользуясь независимостью в последовательности $\{X_i\}$, получаем

$$\begin{aligned}\mathbb{P}\{X_{\alpha,n} \le \alpha(1-\frac{\varepsilon}{2n})\big|X_n=\alpha\} &\ge \mathbb{P}\{X_{\alpha,n} \le \alpha\sqrt[n]{1-\varepsilon}\big|X_n=\alpha\} = \\ &= \mathbb{P}\{\forall 1\le i\le n\ X_i<\alpha \Rightarrow X_i \le \alpha\sqrt[n]{1-\varepsilon}\} = \prod_{X_i<\alpha}\mathbb{P}\{X_i\le\alpha\sqrt[n]{1-\varepsilon}\big|X_i<\alpha\} = \\ &= (1-\varepsilon)^{q^\alpha(n-1)/n} \ge (1-\varepsilon). \quad \square\end{aligned}$$

Лемма 4.3. *Пусть* $\alpha \in (0,1]$, $\varepsilon \in (0,1)$ *и* $n \le N$. *Если* $N \ge \frac{1}{4\varepsilon^3}$, *то* $\mathbb{P}\{|q^\alpha(N) - N\alpha| \le N\varepsilon \big| X_n = \alpha\} \ge 1 - \varepsilon$.

Доказательство. Требуемая оценка сразу вытекает из независимости $\{X_i\}_1^N$, неравенства Чебышёва и условия на N. $\square$

Далее выводится доказательство теоремы 4.1.

Доказательство. Пусть A и B есть случайные величины такие, что $A = X_n$, а $B = X_{A,n}$, где $X_{A,n}$ есть случайная величина равная $\max\limits_{1\le i\le n, X_i < A} X_i$.

Положим $\varepsilon' = \frac{\varepsilon}{5}$. По лемме 4.1 $\mathbb{P}\{\frac{B}{A} \le 1 - \frac{\varepsilon'}{2n}\} = \int \mathbb{P}\{\frac{B}{\alpha} \le 1 - \frac{\varepsilon'}{2n}\big|A = \alpha\}d\alpha \ge 1 - \varepsilon'$. Обозначим через случайную величину $W(x)$ элемент, находящийся в последней клетке первого столбца таблицы $P_N^A(x)$. Получаем, что

$$\mathbb{P}\{W > B\} = \mathbb{P}\{\frac{W}{A} > \frac{B}{A}\} \ge \mathbb{P}\{(\frac{W}{A} > 1 - \frac{\varepsilon'}{2n}) \,\&\, (1 - \frac{\varepsilon'}{2n} \ge \frac{B}{A})\} \ge$$
$$\ge \mathbb{P}\{\frac{W}{A} > 1 - \frac{\varepsilon'}{2n}\} + \mathbb{P}\{1 - \frac{\varepsilon'}{2n} \ge \frac{B}{A}\} - 1 \ge \mathbb{P}\{\frac{W}{A} > 1 - \frac{\varepsilon'}{2n}\} - \varepsilon'. \quad (5)$$

Для каждой реализации x^N оставим только координаты не большие $A(x)$, таких будет $q^A(N)(x)$ штук, и отнормируем их, поделив на $A(x)$. При фиксированном $q^A(N)$, переходя к условной мере, мы опять получим множество реализаций для конечной последовательности независимых случайных величин распределенных по мере Лебега на отрезке $[0,1]$. Значит, к ней применимы утверждения из параграфа 3. Тогда, используя неравенство (4), можно продолжить цепочку неравенств (5)

$$\mathbb{P}\{\frac{W}{A} > 1 - \frac{\varepsilon'}{2n}\} \ge \sum_i \mathbb{P}\{\frac{W}{A} > 1 - \frac{\varepsilon'}{2n}\big|q^A(N) = i\} \cdot \mathbb{P}\{q^A(N) = i\} \ge$$
$$\ge \sum_{i \ge 4n^2/\varepsilon'^4} (1 - \frac{2n}{\varepsilon'\sqrt{q^A(N)+1}}) \cdot \mathbb{P}\{q^A(N) = i\} \ge (1-\varepsilon') \cdot \mathbb{P}\{q^A(N) \ge \frac{4n^2}{\varepsilon'^4}\}. \quad (6)$$

Используя лемму 4.2, получаем, что

$$\mathbb{P}\{q^A(N) \ge \varepsilon' N\} \ge \int_{\alpha > 2\varepsilon'} \mathbb{P}\{q^\alpha(N) \ge \varepsilon' N \big| A = \alpha\}d\alpha \ge$$
$$\ge \int_{\alpha > 2\varepsilon'} \mathbb{P}\{q^\alpha(N) \ge (A - \varepsilon')N \big| A = \alpha\}d\alpha \ge (1-\varepsilon')(1-2\varepsilon') \ge 1 - 3\varepsilon'. \quad (7)$$

Тогда если $N \ge \frac{4n^2}{\varepsilon'^5}$, то по неравенствам (5), (6) и (7)

$$\mathbb{P}\{W > B\} \ge (1-\varepsilon') \cdot \mathbb{P}\{q^A(N) \ge \frac{4n^2}{\varepsilon'^4}\} - \varepsilon' \ge (1-\varepsilon')(1-3\varepsilon') - \varepsilon' \ge 1 - 5\varepsilon' = 1 - \varepsilon.$$

Остается применить следствие 2.2, и получить, что для $\tilde{K}(\varepsilon) = \frac{5^5 2^2}{\varepsilon^5}$ вероятность попадания n-ой координаты в первый столбец за $\tilde{K}(\varepsilon)n^2$ вставок не меньше, чем $1 - \varepsilon$. $\square$

5 Оценка снизу и вывод теоремы 1.1

В данном параграфе доказывается оценка снизу из теоремы 1.1.

Теорема 5.1. *Для любого $\varepsilon > 0$ существует константы $\tilde{k}$ и n_0 такие, что оценка $N(x,n) \geq \tilde{k}n^2$ верна для любого натурального числа $n > n_0$ и любой реализации x из некоторого подмножества реализаций $\mathfrak{X}_{n,\varepsilon} \subset \mathfrak{X}, \mu(\mathfrak{X}_{n,\varepsilon}) \geq 1-\varepsilon$.*

Напомним, что $f(\sigma)$ есть последний элемент в первой строке таблицы $P(\sigma)$. И. Азангулов изучал моменты такой величины и получил следующую асимптотику для флуктуаций последнего элемента (см. работу [8]), которая используется ниже. Если $\sigma \in S_N$ есть случайная перестановка с равномерным распределением, то для любого натурального k верно следующее равенство

$$\lim_{N\to\infty} \mathbb{E}\big(\frac{N-f(\sigma)}{\sqrt{N}}\big)^k = k!.$$

Напомним, что $\{X_n\}_n$ есть последовательность независимых случайных величин распределенных по мере Лебега на отрезке $[0,1]$, $W(x,N)$ (для краткости будем писать просто W) есть последний элемент в первом столбце P-таблицы, получившейся из начального отрезка $\{x_n\}_1^N$ реализации x, а $l=l(N)$ — номер W после упорядочивания начального отрезка $\{x_i\}_1^N$ по возрастанию. Заметим, что l и $f(\sigma)$ равны по распределению, следовательно, у случайных величин $(N-l)/\sqrt{N}$ и $(N-f(\sigma))/\sqrt{N}$ совпадают моменты. Имеется соотношение (равенство (8)) между этими моментами и моментами случайной величины $\sqrt{N}(1-W)$. Обозначим пушфорвард меры Лебега μ при отображении W через $\tilde{\mu}$, а пушфорвард условной меры $\mu(X|\Sigma_N(X)=\sigma)$ ($\Sigma_N(X)$ сопоставляет начальному отрезку последовательности X соответствующую перестановку) при том же отображении через $\tilde{\mu}_\sigma$. Также обозначим через $\beta_{a,b}$ меру соответствующую бета-распределению с плотностью $x^a(1-x)^b/B(a,b)$.

$$\begin{aligned}
\mathbb{E}(\sqrt{N}(1-W))^k &= N^{k/2}\int_0^1 (1-x)^k d\tilde{\mu} = N^{k/2}\sum_{\sigma\in S_N}\frac{1}{N!}\int_0^1(1-x)^k d\tilde{\mu}_\sigma \\
&= N^{k/2}\sum_{\sigma\in S_N}\frac{1}{N!}\int_0^1(1-x)^k d\beta_{f(\sigma),N+1-f(\sigma)} \\
&= N^{k/2}\sum_{s=1}^N \mathbb{P}\{f(\sigma)=s\}\int_0^1 x^k d\beta_{N+1-s,s} \\
&= N^{k/2}\sum_{s=1}^N \mathbb{P}\{f(\sigma)=s\}\frac{(N+1-s)^{\underline{k}}}{(N+1)^{\underline{k}}} \\
&= \frac{N^{k/2}}{(N+1)^{\underline{k}}}\sum_{s=1}^N\sum_{m=1}^k \mathbb{P}\{f(\sigma)=l\}c(k,m)(N+1-s)^m \\
&= \frac{N^{k/2}}{(N+1)^{\underline{k}}}\sum_{m=1}^k c(k,m)\mathbb{E}(N+1-f(\sigma))^m \\
&= \sum_{m=1}^k \frac{N^{k/2+m/2}}{(N+1)^{\underline{k}}}c(k,m)\mathbb{E}\big(\frac{N+1-f(\sigma)}{\sqrt{N}}\big)^m \\
&= \sum_{m=1}^k\sum_{i=0}^m c(k,m)\binom{m}{i}\frac{N^{k/2+m/2}}{(N+1)^{\underline{k}}N^{(m-i)/2}}\mathbb{E}\big(\frac{N-f(\sigma)}{\sqrt{N}}\big)^i, \qquad (8)
\end{aligned}$$

где $c(k,m)$ — беззнаковые числа Стирлинга первого рода.

Устремляя, N к бесконечности получаем, что все слагаемые, кроме одного, где $i=m=k$, устремятся к 0, значит,

$$\lim_{N\to\infty} \mathbb{E}(\sqrt{N}(1-W(x,N))^k = \mathbb{E}\big(\frac{N-f(\sigma)}{\sqrt{N}}\big)^k = k!.$$

Такие моменты имеет распределение $exp(1)$, оно единственно, так как эти моменты удовлетворяют условию Карлемана. Из сходимости моментов следует сходимость по распределению, следовательно, для любого $t > 0$ существует N_0 такое, что для любого $N > N_0$ выполняется неравенство

$$\mathbb{P}\{W < 1 - \frac{\varepsilon}{\sqrt{2N}}\} = \mathbb{P}\{\sqrt{N}(1-W) > \frac{\varepsilon}{2}\} \geq e^{-\varepsilon} > 1-\varepsilon \tag{9}$$

Далее приводится доказательство теоремы 5.1.

Доказательство. Положим $\varepsilon' = \varepsilon/6$. Пусть A и B — случайные величины, такие, что $A = X_n$, а $B = X_{A,n}$, где $X_{A,n} = \max_{1\leq i\leq n, X_i<A} X_i$. Рассмотрим теперь последовательность усеченных таблиц $\{P_m^\alpha\}_m$ и обозначим через W последний элемент в первом столбце усеченной таблицы P_N^α. Как известно из следствия 2.2, для того, чтобы координата x_n не лежала в первом столбце таблицы P_N достаточно, чтобы выполнялось неравенство $B \geq W(N)$. Для краткости будем писать q вместо $q^\alpha(n)$. Положим пока, что $A = \alpha$, и не будем писать это условие в вероятностях. Применяя неравенство Чебышева, получаем следующую цепочку неравенств

$$\begin{aligned}
\mathbb{P}\{\frac{B}{\alpha} > \frac{\alpha n}{\alpha n+2} - s\} &\geq \mathbb{P}\{\frac{B}{\alpha} > \frac{q}{q+1} - s \,\&\, q \geq \frac{\alpha n}{2}\} \\
&= \sum_i \mathbb{P}\{\frac{B}{\alpha} > \frac{q}{q+1} - s \,\&\, q \geq \frac{\alpha n}{2} | q = i\} \cdot \mathbb{P}\{q = i\} \\
&\geq \sum_{i\geq \alpha n/2} (1 - \frac{q}{(q+1)^2(q+2)s^2}) \cdot \mathbb{P}\{q = i\} \\
&\geq (1 - \frac{\alpha n/2}{(\alpha n/2+1)^2(\alpha n/2+2)s^2})(1 - \frac{4\alpha(1-\alpha)}{n})
\end{aligned}$$

При подстановке $s = 2/(n\alpha\sqrt{\varepsilon'})$ получается, что

$$\mathbb{P}\{\frac{B}{\alpha} > 1 - \frac{4}{\alpha\sqrt{\varepsilon'}} \cdot \frac{1}{n}\} \geq \mathbb{P}\{\frac{B}{\alpha} > \frac{\alpha n}{\alpha n+2} - \frac{2}{n\alpha\sqrt{\varepsilon'}}\} \geq (1 - \frac{\alpha^3 n^3 \varepsilon'}{(\alpha n+2)^2(\alpha n+4)})(1 - \frac{4\alpha(1-\alpha)}{n}) \geq 1 - \varepsilon' - \frac{1}{n}$$

Теперь оценим общую вероятность, т.е. без условия $A = \alpha$.

$$\mathbb{P}\{\frac{B}{A} > 1 - \frac{4}{\varepsilon'\sqrt{\varepsilon'}} \cdot \frac{1}{n}\} \geq \int_{\varepsilon}^{\prime 1} \mathbb{P}\{\frac{B}{\alpha} > \frac{4}{\varepsilon'\sqrt{\varepsilon'}} \cdot \frac{1}{n} | A = \alpha\} d\alpha \geq (1 - \varepsilon' - \frac{1}{n})(1-\varepsilon') \geq 1 - 2\varepsilon' - \frac{1}{n} \tag{10}$$

Для каждой реализации x^N оставим только координаты не большие $A(x)$, таких будет $q^A(N)(x)$ штук, и отнормируем их, поделив на $A(x)$. При фиксированном $q^A(N)$, переходя к условной мере, мы опять получим множество реализаций для конечной последовательности независимых случайных величин распределенных по мере Лебега на отрезке $[0,1]$. Тогда, используя неравенство (9), оценим значения W/A. Для краткости положим, что $q = q^\alpha(N)$.

$$\begin{aligned}
\mathbb{P}\{\frac{W}{A} < 1 - \frac{\varepsilon'}{2\sqrt{N}}\} &= \int_0^1 \sum_i \mathbb{P}\{\frac{W}{\alpha} < 1 - \frac{\varepsilon'}{2\sqrt{N}} | A = \alpha, q = i\} \cdot \mathbb{P}\{q = i\} d\alpha \\
&\geq \int_{\varepsilon'}^1 \sum_{i\geq \alpha N/2} \mathbb{P}\{\frac{W}{\alpha} < 1 - \frac{\varepsilon'}{2\sqrt{q}} | A = \alpha, q = i\} \cdot \mathbb{P}\{q = i\} d\alpha \\
&\geq \int_{\varepsilon'}^1 \sum_{i\geq \alpha N/2} (1-\varepsilon') \mathbb{P}\{q = i\} d\alpha \geq 1 - 2\varepsilon' - \frac{1}{N}
\end{aligned} \tag{11}$$

Заметим, что эта оценка верна, если $\alpha N/2 > N_0$, где N_0 из условия неравенства (9). Теперь, если мы положим $N = \lfloor \varepsilon'^5 n^2/64 \rfloor + 1$, то по неравенствам (10) и (11) для достаточно больших n выполнено

$$\mathbb{P}\{W \le B\} = \mathbb{P}\{\frac{W}{A} \le \frac{B}{A}\} \ge \mathbb{P}\{\frac{W}{A} \le 1 - \frac{\varepsilon'}{2\sqrt{N}} \,\&\, 1 - \frac{4}{\varepsilon'\sqrt{\varepsilon'}n} \le \frac{B}{A}\}$$
$$\ge \mathbb{P}\{\frac{W}{A} \le 1 - \frac{\varepsilon'}{2\sqrt{N}}\} + \mathbb{P}\{1 - \frac{4}{\varepsilon'\sqrt{\varepsilon'}n} \le \frac{B}{A}\} - 1 \ge 1 - 6\varepsilon' = 1 - \varepsilon \quad (12)$$

Под словами "для достаточно больших n" имеется ввиду для тех n, для которых $n > 1/\varepsilon' = 6/\varepsilon$ и $N = \lfloor \varepsilon^5 n^2/(2^9 3^3) \rfloor + 1 \ge \max(6/\varepsilon, 2N_0/\alpha)$. Из неравенства (12) следует, что для достаточно больших n доля тех последовательностей, в которых n-ая координата не попала в первый столбец за $N = \lfloor \varepsilon^5 n^2/(2^9 3^3) \rfloor + 1$ вставок, не меньше, чем $1 - \varepsilon$, следовательно, для всех меньших N это тоже верно, так как свойство быть в первом столбце монотонное, т.е. координата, однажды оказавшись в первом столбце, останется там навсегда. $\square$

Из всего вышесказанного несложно выводится теорема 1.1.

Доказательство. Из теорем 4.1 и 5.1 следует, что для любого числа $\varepsilon > 0$ существуют такие константы $\tilde{k}(\varepsilon/2), \tilde{K}(\varepsilon/2), n_0$, что для любого натурального числа $n > n_0$ существуют подмножества реализаций $\mathcal{X}^1_{n,\varepsilon}, \mathcal{X}^2_{n,\varepsilon} \subset \mathcal{X}$, $\mu(\mathcal{X}^1_{n,\varepsilon}) > 1-\varepsilon/2, \mu(\mathcal{X}^2_{n,\varepsilon}) > 1-\varepsilon/2$, такие, что для любой реализации $x \in \mathcal{X}^1_{n,\varepsilon}$ выполняется неравенство $N(x,n) \ge \tilde{k}(\varepsilon/2)n^2$, а для любой реализации $x \in \mathcal{X}^2_{n,\varepsilon}$ выполняется неравенство $N(x,n) \le \tilde{K}(\varepsilon/2)n^2$. Тогда, полагая $k(\varepsilon) = \tilde{k}(\varepsilon/2)$, $K(\varepsilon) = \tilde{K}(\varepsilon/2)$, $\mathcal{X}_{n,\varepsilon} = \mathcal{X}^1_{n,\varepsilon} \cap \mathcal{X}^2_{n,\varepsilon}$, получаем, что $\mu(\mathcal{X}_{n,\varepsilon}) > 1 - \varepsilon$, и для любой реализации $x \in \mathcal{X}_{n,\varepsilon}$ выполняются неравенства $k(\varepsilon)n^2 \le N(x,n) \le K(\varepsilon)n^2$. Утверждение о том, что $k(\varepsilon)$, $K(\varepsilon)$ можно положить равными $c\varepsilon^5$ и $C\varepsilon^{-5}$, соответственно, следует из замечаний сделанных в доказательствах теорем 4.1 и 5.1. $\square$

Литература

[1] И. Ф. Азангулов, Г. В. Овечкин, “Оценка времени попадания координаты схемы Бернулли в первый столбец таблицы Юнга”, *Функц. анализ и его прил.*, **54**:2 (2020), 78–84.

[2] Stanley R. (1999). Enumerative combinatorics. Vol 2. Cambridge university press. (Перевод на русский) Стенли Р. Перечислительная комбинаторика. Деревья, производящие функции и симметрические функции: Пер. с англ. — М.: Мир, 2009 — 767 с..

[3] Knuth D. (1970). Permutations, matrices and generalized Young tableaux. — *Pacific J. Math.*, **34**:3, 709–727.

[4] Fulton W. (1997). Young tableaux with application to representation theory and geometry. Cambridge university press.

[5] Kerov S., Vershik A. (1986). The characters of the infinite symmetric group and probability properties of the Robinson–Shensted–Knuth algorithm. — *SIAM J. Alg. Discr. Methods*, **7**:1, 116–123.

[6] Romik D., Sniady P. (2015). Jeu de taquin dynamics on infinite Young tableaux and second calss particles. — *The Ann. of Prob.*, **42**:7, 682–737.

[7] А.М. Вершик (2020). Комбинаторное кодирование схемы Бернулли и асимптотика таблиц Юнга. — *Функ. Анал. и его Прил.*, **54**:2, 3-24.

[8] И.Ф. Азангулов (2020). Распределение флуктуации P-таблиц схемы Бернулли при соответствии RSK. Выпускная квалификационная работа. СПбГУ.