Math-Net.Ru

Общероссийский математический портал

Г. И. Ивченко, Ю. И. Медведев, Случайные подстановки:
общая параметрическая модель, *Дискрет. матем.*, 2006,
том 18, выпуск 4, 105–112

DOI: <https://doi.org/10.4213/dm75>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.197.217.227

26 апреля 2023 г., 13:11:20

Случайные подстановки: общая параметрическая модель

© 2006 г. Г. И. Ивченко, Ю. И. Медведев

Предлагается новый подход к изучению структурных свойств случайных подстановок, в основе которого лежит введение на множестве S_n всех подстановок степени n некоторой общей параметрической меры, приписывающей подстановке с цикловой структурой $a = (a_1, \dots, a_n)$ вероятность, пропорциональную величине $\prod_{i=1}^n \theta_i^{a_i}$, где $\theta = (\theta_1, \dots, \theta_n)$ — параметр меры.

1. Описание модели

Пусть S_n есть множество всех подстановок степени n , то есть взаимно однозначных отображений конечного множества $X_n = \{1, 2, \dots, n\}$ в себя. Обозначим $a = (a_1, \dots, a_n)$ цикловую структуру подстановки $A \in S_n$, где a_i — число циклов подстановки A длины i , $i = 1, \dots, n$, $\sum_{i=1}^n i a_i = n$, и введем на множестве S_n вероятностную меру, зависящую от параметра $\theta = (\theta_1, \dots, \theta_n)$, $\theta_1, \dots, \theta_n \geq 0$, в соответствии с которой подстановка A со структурой $a = (a_1, \dots, a_n)$ наблюдается с вероятностью

$$P_\theta(A) = I\left(\sum_{i=1}^n i a_i = n\right) \prod_{i=1}^n \theta_i^{a_i} / H_n(\theta), \quad (1)$$

где $I(\cdot)$ — индикатор и $H_n(\theta)$ — нормирующий множитель.

Поскольку число подстановок циклового класса $a = (a_1, \dots, a_n)$ равно ([1])

$$K_n(a) = \frac{n!}{\prod_{i=1}^n a_i! i^{a_i}},$$

из условия нормировки

$$\sum_{A \in S_n} P_\theta(A) = \sum_{a: a_1 + 2a_2 + \dots + na_n = n} K_n(a) \prod_{i=1}^n \theta_i^{a_i} / H_n(\theta) = 1$$

находим, что

$$\begin{aligned}
 H_n(\theta) &= \sum_{a: a_1 + 2a_2 + \dots + na_n = n} K_n(a) \prod_{i=1}^n \theta_i^{a_i} \\
 &= n! \sum_{a: a_1 + 2a_2 + \dots + na_n = n} \prod_{i=1}^n \left(\frac{\theta_i}{i}\right)^{a_i} \frac{1}{a_i!} \\
 &= n! [z^n] \exp \left\{ \sum_{i=1}^n \frac{z^i}{i} \theta_i \right\}, \tag{2}
 \end{aligned}$$

где $[z^n]$ — коэффициент при z^n в далее стоящей функции.

Таким образом, $H_n(\theta)$ — производящая функция числа подстановок степени n заданного циклового класса в точке $\theta = (\theta_1, \dots, \theta_n)$.

Соотношениями (1)–(2) и определяется общая параметрическая модель случайной подстановки степени n . Достоинством этой модели является наличие в ней большого числа степеней свободы: зависимость от n -мерного параметра $\theta = (\theta_1, \dots, \theta_n)$ придает ей большую гибкость, так как выбирая соответствующим образом значения этого параметра, мы можем задавать на множестве S_n различные распределения, позволяющие изучать различные особенности подстановок. Такой подход открывает, на наш взгляд, новые возможности в исследовании этого популярного объекта вероятностной комбинаторики, позволяет решать не только стандартные перечислительные задачи, но и проводить более глубокий вероятностный и статистический анализ случайных подстановок и тем самым более полно отвечать на запросы практики.

Отметим некоторые частные случаи введенной модели, которые рассматривались ранее в литературе.

1. Если $\theta_i = 1$ для всех $i = 1, \dots, n$, то получаем равномерную меру на множестве S_n , приписывающую каждой подстановке $A \in S_n$ одну и ту же вероятность $P(A) = (n!)^{-1}$, поскольку

$$\begin{aligned}
 [z^n] \exp \left\{ \sum_{i=1}^n \frac{z^i}{i} \right\} &= [z^n] \exp \left\{ \sum_{i=1}^{\infty} \frac{z^i}{i} \right\} \\
 &= [z^n] \exp \{-\ln(1-z)\} = [z^n](1-z)^{-1} = 1.
 \end{aligned}$$

Большинство известных результатов для случайных подстановок, начиная с классической работы В. Л. Гончарова [1], получено в рамках именно этой равновероятной модели (см. также [2–6] и библиографию в них).

2. Пусть для заданного подмножества $R \subseteq X_n$ символ $S_n(R)$ обозначает подмножество всех тех подстановок $A \in S_n$, длины циклов которых принадлежат R (так называемые R -подстановки). Полагая в (1) $\theta_i = 0$ для $i \notin R$, получим меру, сосредоточенную на $S_n(R)$. Тем самым мы имеем возможность изучать различные структурные свойства случайных подстановок с учетом произвольного ограничения на длины циклов, задаваемого множеством R . Типичными примерами задания R являются $R = \{i: i \leq s\}$ при некотором $s < n$ (длины циклов не превосходят s) либо $R = \{i: i \equiv m \pmod{d}\}$ при некоторых m и d . Литература по равновероятным R -подстановкам (когда $\theta_i = 1$ для $i \in R$) также весьма значительна (библиографию см. в [3, 6]).

Неоднократно отмечалось, что при попытках изучения характеристик цикловой структуры случайных подстановок с распределениями, отличными от равномерного на всем множестве S_n подстановок степени n , или на некотором их подмножестве, возникают существенные трудности. Общая параметрическая модель позволяет в определенной мере преодолеть эти трудности.

3. Если в (1) положить $\theta_1 = \theta_2 = \dots = \theta_n = \theta > 0$, то получим однопараметрическую меру, приписывающую подстановкам $A \in S_n$ вероятности, пропорциональные $\theta^{|A|}$, где $|A| = a_1 + \dots + a_n$ — общее число циклов подстановки A . В этом случае из (2) получаем, что

$$\begin{aligned} H_n(\theta) &= n! [z^n] \exp\{-\theta \ln(1-z)\} \\ &= n! [z^n] (1-z)^{-\theta} = n! (-1)^n \binom{-\theta}{n} \\ &= \theta(\theta+1) \dots (\theta+n-1) = \theta_{(n)}, \end{aligned} \quad (3)$$

и формула (1) принимает вид

$$P_\theta(A) = I \left(\sum_{i=1}^n i a_i = n \right) \frac{\theta^{|A|}}{\theta_{(n)}}. \quad (4)$$

Впервые такая неравновероятная мера рассматривалась Эвансом [8], и эта модель достаточно детально исследована в литературе (см. [5, 7] и библиографию в них).

4. Введем другую однопараметрическую модель случайных подстановок, полагая в (1)

$$\theta_i = \frac{\theta}{(i-1)!}, \quad \theta > 0, \quad i = 1, \dots, n. \quad (5)$$

В этом случае получаем меру, не только приписывающую подстановкам $A \in S_n$ вероятности, пропорциональные $\theta^{|A|}$, как в схеме Эванса, но и учитывающую длины самих циклов. В этом случае из (2) получаем, что

$$\begin{aligned} H_n(\theta) &= n! [z^n] \exp \left\{ \sum_{i=1}^n \frac{z^i}{i!} \theta \right\} = n! [z^n] \exp\{\theta(e^z - 1)\} \\ &= \sum_{k=0}^n \sigma(n, k) \theta^k = e^{-\theta} \sum_{k=1}^{\infty} \frac{k^n \theta^k}{k!}, \end{aligned} \quad (6)$$

где $\sigma(n, k)$ — числа Стирлинга второго рода. Таким образом, в этом случае $H_n(\theta)$ — производящая функция чисел $\sigma(n, k)$ по второму аргументу в точке θ . Вероятностная мера в этой модели имеет вид

$$P_\theta(A) = I \left(\sum_{i=1}^n i a_i = n \right) \frac{\theta^{|A|}}{\prod_{i=1}^n ((i-1)!)^{a_i} H_n(\theta)}. \quad (7)$$

2. Цикловая структура случайной подстановки

Будем обозначать далее $c(n) = (c_1(n), \dots, c_n(n))$ цикловую структуру случайной подстановки в модели (1). Тогда для ее распределения справедливо представление

$$\begin{aligned} P_{\theta}(c(n) = a) &= I \left(\sum_{i=1}^n i a_i = n \right) K_n(a) \prod_{i=1}^n \theta_i^{a_i} / H_n(\theta) \\ &= I \left(\sum_{i=1}^n i a_i = n \right) n! \prod_{i=1}^n \left(\frac{\theta_i}{i} \right)^{a_i} \frac{1}{a_i!} / H_n(\theta) \end{aligned} \quad (8)$$

Из (8) с учетом (2) легко получить и производящую функцию случайного вектора $c(n)$:

$$\begin{aligned} F_{n,\theta}(t) &= E_{\theta} \prod_{i=1}^n t_i^{c_i(n)} = n! \sum_{a: a_1 + 2a_2 + \dots + na_n = n} \prod_{i=1}^n \left(\frac{t_i \theta_i}{i} \right)^{a_i} \frac{1}{a_i!} / H_n(\theta) \\ &= H_n(t\theta) / H_n(\theta), \end{aligned} \quad (9)$$

где $t = (t_1, \dots, t_n)$ и $t\theta = (t_1 \theta_1, \dots, t_n \theta_n)$.

Соотношение (9) может быть основой для изучения различных структурных свойств подстановок в рамках общей модели (1).

В качестве иллюстрации использования соотношения (9) получим общую формулу для смешанных факториальных моментов случайной структуры $c(n)$:

$$\begin{aligned} E_{\theta} \prod_j (c_j(n))_{r_j} &= \frac{\partial^{r_1 + \dots + r_n} F_{n,\theta}(t)}{\prod_j \partial t_j^{r_j}} \Big|_{t=(1,\dots,1)} \\ &= n! [z^n] \prod_j \left(\frac{\theta_j}{j} \right)^{r_j} z^{jr_j} \exp \left\{ \sum_{i=1}^n \frac{z^i}{i} \theta_i \right\} / H_n(\theta) \\ &= \frac{n! H_{n-m}(\theta)}{(n-m)! H_n(\theta)} \prod_j \left(\frac{\theta_j}{j} \right)^{r_j}, \end{aligned} \quad (10)$$

где

$$(a)_r = a(a-1) \dots (a-r-1), \quad m = \sum_j jr_j.$$

В частности, формулы для средних значений имеют вид

$$\begin{aligned} E_{\theta} c_j(n) &= \frac{n! \theta_j H_{n-j}(\theta)}{(n-j)! j H_n(\theta)} \\ &= \frac{\theta_j}{j} [z^{n-j}] \exp \left\{ \sum_{i=1}^n \frac{z^i}{i} \theta_i \right\} / [z^n] \exp \left\{ \sum_{i=1}^n \frac{z^i}{i} \theta_i \right\}. \end{aligned} \quad (11)$$

3. О методах анализа модели

Анализ предложенной модели может быть основан на классических методах производящих и характеристических функций в сочетании с методом перевала, а также методов условных распределений и рандомизации параметра n . Опишем кратко два последних подхода.

Теорема 1. Распределение случайной структуры $c(n) = (c_1(n), \dots, c_n(n))$ в модели (1)–(2) может быть представлено в виде условного распределения

$$\mathcal{L}_\theta(c(n)) = \mathcal{L}_\theta(Z_1, \dots, Z_n \mid T_n = n), \quad (12)$$

где случайные величины $Z_1, \dots, Z_n$ независимы, Z_i имеет распределение Пуассона $\Pi(\lambda_i)$ с параметром $\lambda_i = \theta_i/i$, $i = 1, \dots, n$, и

$$T_n = Z_1 + 2Z_2 + \dots + nZ_n. \quad (13)$$

Доказательство утверждения теоремы следует из следующей цепочки равенств:

$$\begin{aligned} & \mathbf{P}_\theta(X_i = a_i, i = 1, \dots, n \mid T_n = n) \\ &= \frac{\mathbf{P}_\theta(Z_i = a_i, i = 1, \dots, n, T_n = n)}{\mathbf{P}_\theta(T_n = n)} \\ &= I\left(\sum_{i=1}^n i a_i = n\right) \prod_{i=1}^n \mathbf{P}_\theta(Z_i = a_i) / \sum_{b: b_1 + 2b_2 + \dots + nb_n = n} \prod_{i=1}^n \mathbf{P}_\theta(Z_i = b_i) \\ &= \exp\left\{-\sum_{i=1}^n \lambda_i\right\} I\left(\sum_{i=1}^n i a_i\right) \prod_{i=1}^n \frac{\lambda_i^{a_i}}{a_i!} / \exp\left\{-\sum_{i=1}^n \lambda_i\right\} \sum_{b: b_1 + 2b_2 + \dots + nb_n = n} \prod_{i=1}^n \frac{\lambda_i^{b_i}}{b_i!} \\ &= \frac{H_n(\theta)}{n!} \mathbf{P}_\theta(c(n) = a) / \sum_{b: b_1 + 2b_2 + \dots + nb_n = n} \frac{H_n(\theta)}{n!} \mathbf{P}_\theta(c(n) = b) \\ &= \mathbf{P}_\theta(c(n) = a), \end{aligned}$$

поскольку

$$\sum_{b: b_1 + 2b_2 + \dots + nb_n = n} \mathbf{P}_\theta(c(n) = b) = 1.$$

Замечание 1. Из предыдущих соотношений следует также, что

$$\begin{aligned} \mathbf{P}_\theta(T_n = n) &= \exp\left\{-\sum_{i=1}^n \lambda_i\right\} \frac{H_n(\theta)}{n!} \\ &= \exp\left\{-\sum_{i=1}^n \frac{\theta_i}{i}\right\} [z^n] \exp\left\{\sum_{i=1}^n \frac{z^i}{i} \theta_i\right\} \end{aligned} \quad (14)$$

Теорема 2. Если степень подстановки является случайной величиной η , имеющей распределение типа степенного ряда

$$\mathbf{P}_\theta(\eta = n) = \frac{x^n [x^n] f_\theta(x)}{f_\theta(x)}, \quad n = 0, 1, \dots, \quad x > 0, \quad (15)$$

где функция $f_\theta(x)$ имеет вид

$$f_\theta(x) = \exp\left\{\sum_{i=1}^{\infty} \lambda_i x^i\right\}, \quad \lambda_i = \frac{\theta_i}{i}, \quad i = 1, 2, \dots, \quad (16)$$

то элементы структуры $c_i(\eta)$, $i = 1, 2, \dots$, являются независимыми и

$$\mathcal{L}_\theta(c_i(\eta)) = \Pi(\lambda_i x^i), \quad i = 1, 2, \dots$$

Доказательство утверждения теоремы следует из цепочки равенств (см. (2) и (9)):

$$\begin{aligned}
 \mathbf{E}_\theta \prod_{i=1}^{\infty} t_i^{c_i(n)} &= \sum_{n=0}^{\infty} \mathbf{P}_\theta(\eta = n) \mathbf{E}_\theta \prod_{i=1}^n t_i^{c_i(n)} \\
 &= \sum_{n=0}^{\infty} \frac{H_n(\theta)}{n! f_\theta(x)} x^n \sum_{a: a_1 + 2a_2 + \dots + na_n = n} \frac{n!}{H_n(\theta)} \prod_{i=1}^n \frac{(t_i \lambda_i)^{a_i}}{a_i!} \\
 &= f_\theta^{-1}(x) \sum_{n=0}^{\infty} \sum_{a: a_1 + 2a_2 + \dots + na_n = n} \prod_{i=1}^n \frac{(t_i \lambda_i x^i)^{a_i}}{a_i!} \\
 &= f_\theta^{-1}(x) \prod_{i=1}^{\infty} \sum_{a_i=0}^{\infty} \frac{(t_i \lambda_i x^i)^{a_i}}{a_i!} \\
 &= \prod_{i=1}^{\infty} \exp\{\lambda_i x^i (t_i - 1)\}.
 \end{aligned}$$

Замечание 2. В теореме 2 параметр x является свободным параметром рандомизации, значение которого можно подбирать специальным образом в зависимости от цели исследования.

4. R -подстановки

Пусть $R \subseteq X_n = \{1, \dots, n\}$ — произвольное подмножество. Тогда производящая функция цикловой структуры случайной R -подстановки получается из (9) при $\theta_i = 0$ для $i \notin R$ и равна

$$F_{n,\theta,R}(t) = \mathbf{E}_\theta \prod_{i \in R} t_i^{c_i(n)} = [z^n] \exp \left\{ \sum_{i \in R} \frac{z^i}{i} \theta_i t_i \right\} / [z^n] \exp \left\{ \sum_{i \in R} \frac{z^i}{i} \theta_i \right\}. \quad (17)$$

Формула (10) для факториальных моментов принимает вид

$$\mathbf{E}_\theta \prod_{j \in R} (c_j(n))_{r_j} = \prod_{j \in R} \left(\frac{\theta_j}{j} \right)^{r_j} [z^{n-m}] \exp \left\{ \sum_{i \in R} \frac{z^i}{i} \theta_i \right\} / [z^n] \exp \left\{ \sum_{i \in R} \frac{z^i}{i} \theta_i \right\}.$$

Приведем еще один асимптотический результат.

Пусть $n \rightarrow \infty$ и существует единственный ограниченный корень $z_n = z_n(\theta)$ уравнения

$$\sum_{i \in R} z^i \theta_i = n. \quad (18)$$

Тогда распределение числа $c_i(n)$ циклов фиксированной длины $i \in R$ является асимптотически Пуассоновским с параметром $z_n^i \theta_i / i$.

В частности, если $R = X_n$ и $\theta_i = b\theta^i$, $i = 1, \dots, n$, то $z_n \sim \theta^{-1}$ и параметр предельного пуассоновского распределения равен b/i , то есть не зависит от θ .

Если же $R = \{1, \dots, m\}$ и $\ln m \asymp \ln n$ (по-прежнему $\theta_i = b\theta^i$, $i \in R$), то

$$z_n = \frac{1}{\theta} + \frac{1}{m\theta} \left(\ln \frac{n}{m} + \ln \ln \frac{n}{m} - \ln b + o(1) \right)$$

и общее число циклов случайной R -подстановки будет асимптотически нормальным со средним и дисперсией, равными $\sum_{i \in R} z_n^i / i$.

В модели (1) с $\theta_i = \theta / (i - 1)!$, $i = 1, \dots, n$, уравнение (18) принимает вид

$$\sum_{i=1}^n \frac{z^i}{(i-1)!} = \frac{n}{\theta},$$

или при $n \rightarrow \infty$

$$ze^z = \frac{n}{\theta} + O\left(\frac{z^{(n+1)}}{n!}\right).$$

Решая это уравнение методом итераций, получаем, что

$$z_n = \ln \frac{n}{\theta} - \ln \ln \frac{n}{\theta} + \frac{\ln \ln n'}{\ln n'} + \frac{1}{2} \left(\frac{\ln \ln n'}{\ln n'} \right)^2 + O\left(\frac{\ln \ln n'}{\ln^2 n'}\right),$$

где $n' = n/\theta$. Таким образом, в этом случае $z \sim \ln(n/\theta)$, и этот корень ограничен при $\theta \sim \alpha n$, $0 < \alpha < 1$. В этом случае число циклов фиксированной длины i также имеет в пределе распределение Пуассона, а в случае, когда θ имеет порядок, меньший n , число циклов $c_i(n)$ асимптотически нормально.

5. Двупараметрическая модель

Рассмотрим двупараметрическую модель на подмножестве подстановок. Пусть θ и m — два заданных параметра, причем $\theta > 0$ и m — целое число, $1 \leq m < n$. Положим в (1)

$$\begin{aligned} \theta_i &= \binom{m}{i} \theta, & i &= 1, \dots, m, \\ \theta_i &= 0, & i &= m+1, \dots, n. \end{aligned} \quad (19)$$

Уравнение (18) принимает вид

$$\theta \sum_{i=1}^m \binom{m}{i} z^i = n$$

или

$$(z+1)^m - 1 = n/\theta,$$

откуда находим точное решение

$$z_n = (n/\theta + 1)^{1/m} - 1.$$

Предельное распределение числа $c_i(n)$ циклов фиксированной длины i определяется поведением величин z_n , $\theta = \theta(n)$ и $m = m(n)$ при $n \rightarrow \infty$.

Более детальное исследование случайных подстановок в общей модели (1) ждет своего времени.

Замечание 3. Изложенный в данной статье метод введения неравновероятной рандомизации на множестве подстановок степени n можно применять и в исследованиях других комбинаторных объектов, например, однозначных отображений конечного множества в себя, графов, многочленов над конечными полями, разбиений конечных множеств и целых чисел (см. [9]).

Список литературы

1. Гончаров В. Л., Из области комбинаторики. *Изв. АН СССР, сер. матем.* (1944) **8**, №1, 3–48.
2. Колчин В. Ф., Севастьянов Б. А., Чистяков В. П., *Случайные размещения*. Наука, Москва, 1976.
3. Колчин В. Ф., *Случайные отображения*. Наука, Москва, 1984.
4. Колчин В. Ф., *Случайные графы*. Физматлит, Москва, 2000.
5. Ивченко Г. И., Медведев Ю. И., О случайных подстановках. *Труды по дискретной математике* (2002) **5**, 73–92.
6. Сачков В. Н., *Введение в комбинаторные методы дискретной математики*. МЦНМО, Москва, 2004.
7. Ивченко Г. И., Медведев Ю. И., Метод В. Л. Гончарова и его развитие в анализе различных моделей случайных подстановок. *Теория вероятностей и ее применения* (2002) **47**, №3, 558–566.
8. Ewens W. J., The sampling theory of selectively neutral alleles. *Theoret. Popul. Biol.* (1972) **3**, 87–112.
9. Ивченко Г. И., Медведев Ю. И., Случайные комбинаторные объекты. *Докл. РАН* (2004) **396**, №2, 151–154.

Статья поступила 20.02.2005.