Записки научных
семинаров ПОМИ
Том 240, 1997 г.

Н. В. Цилевич

РАСПРЕДЕЛЕНИЕ СРЕДНЕГО ЗНАЧЕНИЯ ДЛЯ НЕКОТОРЫХ СЛУЧАЙНЫХ МЕР

1. Постановка задачи и формулировка результата. Пусть τ – произвольное вероятностное распределение на отрезке $[0,1]$. Рассмотрим случайную дискретную вероятностную меру

$$F=\sum_{i=1}^{\infty} P_i\delta_{X_i}, \tag{1.1}$$

где $X=(X_1,X_2,\dots)$ – последовательность независимых случайных величин с общим распределением τ, а $P=(P_1,P_2,\dots)\in\Sigma=\{Y=(Y_1,Y_2,\dots):\,Y_i\geqslant 0,\sum Y_i\leqslant 1\}$ – случайная последовательность нагрузок – независима от X и имеет распределение $\mathrm{GEM}(1)$, задаваемое простейшим *процессом ломания палки*: выбираем число P_1, равномерно распределенное на $[0,1]$; затем число P_2, равномерно распределенное на "оставшейся" части отрезка – т.е. на $[0,1-P_1]$ и так далее: P_n равномерно распределено на $[0,1-P_1-\dots-P_{n-1}]$. Таким образом,

$$P_n=W_n(1-W_1)\dots(1-W_{n-1}), \tag{1.2}$$

где $W_1,W_2,\dots$ – последовательность независимых случайных величин, равномерно распределенных на отрезке $[0,1]$.

Построенная мера называется *случайной мерой Дирихле* с параметрической мерой τ и обозначается $\mathcal{D}(\tau)$. Случайную меру (1.1) удобно также записывать в виде

$$F=\sum_{i=1}^{\infty}\tilde P_i\delta_{X_i}, \tag{1.3}$$

где $\tilde P_1\geqslant\tilde P_2\geqslant\dots$ – перестановка элементов последовательности (P_i) в невозрастающем порядке.

Рассмотрим функционал среднего значения $\int x dF(x)$ для $F\in\mathcal{D}(\tau)$ и обозначим через μ распределение этого функционала. В

Работа выполнена при поддержке гранта РФФИ 96-01-00676.

работе [2] было доказано, что исходная мера τ и полученная мера μ связаны замечательным соотношением

$$\int \frac{d\mu(u)}{z-u} = \exp \int \ln \frac{1}{z-u} d\tau(u), \quad z \in \mathbb{C} \setminus \mathbb{R}. \tag{1.4}$$

Пусть теперь α, θ – два вещественных параметра. Рассмотрим обобщенную модель, в которой положения нагрузок $X_1, X_2, \ldots$ выбираются так же, как и раньше, а величины нагрузок строятся с помощью *нестационарного* процесса ломания палки (*модели остаточного распределения*): случайная величина W_j в представлении (1.2) имеет бета-распределение $B(1-\alpha, \theta + j\alpha)$ с плотностью

$$\frac{\Gamma(1+\theta+(j-1)\alpha)}{\Gamma(1-\alpha)\Gamma(\theta+j\alpha)} t^{-\alpha}(1-t)^{\theta-1+j\alpha}, \quad t \in [0,1]. \tag{1.5}$$

Полученную случайную меру будем обозначать через $\mathcal{D}(\tau, \alpha, \theta)$ и называть *обобщенной мерой Дирихле.* Область допустимых параметров есть объединение множеств

$$\{(\alpha, \theta) : 0 \leqslant \alpha < 1, \theta > -\alpha\} \quad \text{и} \quad \{(\alpha, -m\alpha),\ \alpha < 0, m \in \mathbb{N}\}. \tag{1.6}$$

Пусть μ – распределение среднего $\int x dF(x)$ для $F \in \mathcal{D}(\tau, \alpha, \theta)$. В этой статье мы доказываем формулу, обобщающую (1.4). При $\alpha, \theta \neq 0$ она имеет вид

$$\left(\int (z-u)^{-\theta} d\mu(u)\right)^{-\frac{1}{\theta}} = \left(\int (z-u)^{\alpha} d\tau(u)\right)^{\frac{1}{\alpha}}, \quad z \in \mathbb{C} \setminus \mathbb{R}. \tag{1.7}$$

Автор благодарит С. В. Керова за постановку задачи, ссылки на литературу и многочисленные полезные обсуждения.

2. Классический процесс Дирихле. Классический процесс Дирихле (соответствующий параметру $\alpha = 0$) был введен Фергюсоном ([3]) с помощью следующего аксиоматического определения. Пусть сначала β – положительная дискретная мера, имеющая нагрузки $\beta_1, \ldots, \beta_n$ в точках $z_1, \ldots, z_n$. Множество дискретных вероятностных мер с нагрузками в $z_1, \ldots, z_n$ естественным образом отождествляется с симплексом $\Delta_n = \{x = (x_1, \ldots, x_n) : x_i \geqslant 0, \sum x_i = 1\}$. Рассмотрим на Δ_n т.н. распределение Дирихле $D(\beta_1, \ldots, \beta_n)$, имеющее плотность

$$\frac{\Gamma(\beta_1 + \ldots + \beta_n)}{\Gamma(\beta_1) \ldots \Gamma(\beta_n)} x_1^{\beta_1 - 1} \ldots x_n^{\beta_n - 1}$$

по мере Лебега $dx_1 \dots dx_{n-1}$. Случайная мера $F = \sum_{i=1}^n P_i \delta_{z_i}$, где $P = (P_1, \dots, P_n) \in \Delta_n$ имеет распределение $D(\beta_1, \dots, \beta_n)$, называется процессом Дирихле с параметром β. Для произвольной конечной положительной меры β на отрезке $[0,1]$ случайная мера F имеет распределение $\mathcal{D}(\beta)$, если для любого конечного разбиения $[0,1] = A_1 \cup \dots \cup A_n$ вектор $(F(A_1), \dots, F(A_n))$ имеет распределение Дирихле на Δ_n с параметрами $(\beta(A_1), \dots, \beta(A_n))$. (См. также [8], гл. 9.) Фергюсон показал, что случайная мера Дирихле представляется в виде (1.3) для описанной конструкции с параметрами $(\nu, 0, \theta)$, где θ есть полная масса меры β, а $\nu = \beta/\theta$. В этом случае процесс ломания палки стационарен – все W_i имеют одинаковое распределение с плотностью $\theta(1-x)^{\theta-1}$. Задаваемая этим процессом мера на симплексе Σ называется GEM-распределением с параметром θ, а ее монотонизация (т.е. распределение упорядоченной перестановки $(\tilde{P}_i)$ на симплексе монотонных последовательностей) есть знаменитое *распределение Пуассона–Дирихле* $PD(\theta)$.

Со случайными мерами удобно работать в терминах выборок. Выборкой из случайного распределения называется последовательность случайных величин, которая получается с помощью следующей двуступенчатой процедуры: сначала мы выбираем реализацию случайной меры, а затем – последовательность независимых величин с выбранным на первом шаге распределением. Такие последовательности обладают свойством перестановочной инвариантности (exchangeability), которое состоит в том, что совместные распределения последовательности инвариантны относительно любых конечных перестановок ее элементов. Согласно классической теореме де Финетти, любая инвариантная относительно перестановок последовательность есть выборка из некоторой случайной меры, однозначно определяемой этой последовательностью.

В статье [1] предложена конструкция, позволяющая строить выборку из распределения $\mathcal{D}(\beta)$ – так называемая *урновая схема Блэкуэлла–Маккуина*:

(1) X_1 имеет распределение $\frac{\beta}{\|\beta\|}$;

(2) условное распределение X_{n+1} при условии, что заданы величины $X_1, \dots, X_n$, равно $\frac{\beta_n}{\|\beta_n\|}$, где $\beta_n = \beta + \sum_{i=1}^n \delta_{X_i}$.

Здесь $\|\beta\|$ обозначает полную массу меры β.

Вариантом урновой схемы является *конструкция китайского ресторана* (см. [2]). Для выборки из процесса Дирихле с параметром $\beta = \theta\nu$, $\theta = \|\beta\|$ она описывается следующим образом. Пусть имеется бесконечно много круглых столов, за каждым из которых может разместиться бесконечно много людей. Снабдим эти столы метками $x_1, x_2, \ldots$, выбирая их независимо распределенными по мере ν. Первый человек садится за первый стол. Пусть уже размещены n человек. Очередной посетитель садится с вероятностью $\frac{\theta}{n+\theta}$ за первый незанятый стол, а с вероятностью $\frac{1}{n+\theta}$ занимает место справа от i-го уже сидящего человека $(i = 1, \ldots, n)$. Положим X_i равным метке стола, за которым сидит i-й посетитель. Полученная последовательность случайных величин $X_1, X_2, \ldots$ есть выборка из распределения $\mathcal{D}(\beta)$.

Заметим, что можно трактовать номера людей, сидящих за одним столом на n-м шаге, как цикл перестановки $\pi \in S_n$, так что попутно мы задали меры M_θ^n на симметрических группах S_n, причем

$$M_\theta^n(\pi) = \frac{\theta^{c(\pi)}}{\theta(\theta+1)\ldots(\theta+n-1)}, \tag{2.1}$$

где $c(\pi)$ – число циклов перестановки $\pi \in S_n$. Эти меры называются мерами Ювенса с параметром $\theta > 0$, они играют важную роль в комбинаторике и приложениях. Заметим, что M_1^n есть мера Хаара на S_n.

3. Обобщенный процесс Дирихле. С. В. Керов ([5]) и Дж. Питман ([11]) независимо предложили обобщение конструкций, приводящих к мерам Дирихле. В терминах представлений типа (1.4) они означают рассмотрение случайных мер вида

$$F = \sum \tilde{P}_i \delta_{X_i} + (1 - \sum \tilde{P}_i)\tau, \tag{3.1}$$

где $\{X_i\}$ – последовательность независимых случайных величин с распределением τ, а вектор $\tilde{P} = (\tilde{P}_1, \tilde{P}_2, \ldots)$ имеет *произвольное* распределение на симплексе

$$\tilde{\Sigma} = \{Y = (Y_1, Y_2, \ldots) : Y_1 \geqslant Y_2 \geqslant \ldots \geqslant 0, \quad \sum Y_i \leqslant 1\}.$$

Дж. Питман описал соответствующие урновые схемы. Мы приведем обобщенную модель китайского ресторана, принадлежащую С. В. Керову ([5]). Параметром этой модели, кроме вероятностной меры τ, служит семейство мер M^n на симметрических группах S_n, удовлетворяющих условию инвариантности относительно внутренних автоморфизмов (т.е. нагрузка перестановки зависит только от длин ее циклов) и условию согласованности в следующем смысле: для любой перестановки $\pi \in S_n$

$$M^n(\pi) = \sum_{\sigma'=\pi} M^{n+1}(\sigma), \tag{3.2}$$

где сумма берется по всем перестановкам $\sigma \in S_{n+1}$, таким что π получается из σ вычеркиванием элемента $n+1$ из его цикла. Как и прежде, снабдим столы случайными метками x_i, выбирая их независимо распределенными по мере τ. Обобщенное правило для размещения посетителей ресторана имеет следующий вид. Первый человек садится, как и ранее, за первый стол. Пусть уже сидящие n человек образуют перестановку $\pi_n \in S_n$. После того, как займет свое место $(n+1)$-й посетитель, получится перестановка $\pi_{n+1} \in S_{n+1}$, такая что $\pi'_{n+1} = \pi_n$. Положим вероятность получения такой перестановки равной $M^{n+1}(\pi_{n+1})/M^n(\pi_n)$. Как и ранее, возьмем X_i равным метке стола, за которым сидит i-й посетитель. Условия, наложенные на семейство мер M^n, обеспечивают инвариантность полученной последовательности $\{X_i\}$ относительно перестановок. Поэтому она будет задавать некоторую случайную меру F. Обозначим ее $\mathcal{D}(\tau, \{M^n\})$.

Вопрос о связях урновых схем и моделей китайского ресторана с представлениями (3.1) решается в рамках кингмановской теории случайных разбиений (см. [6], [7]).

Пусть μ – распределение среднего значения $\int x dF(x)$ для $F \in \mathcal{D}(\tau, \{M^n\})$. В работе [5] приведена формула, связывающая моменты мер μ и τ.

Теорема ([5], теорема 4.2.2). *Пусть $h_n = \int u^n d\mu(u)$ и $p_n = \int u^n d\tau(u)$ – моменты мер μ и τ. Имеет место следующее соотношение:*

$$h_n = \sum_{\pi \in S_n} M^n(\pi) \prod_{j \geqslant 1} p_j^{r_j(\pi)}, \quad n = 1, 2, \ldots \tag{3.3}$$

где $r_j(\pi)$ – количество циклов перестановки π, имеющих длину j.

В работах [10],[12] было показано, что меры Пуассона–Дирихле $PD(\theta)$ естественным образом включаются в двупараметрическое семейство $PD(\alpha, \theta)$, получаемое с помощью модели остаточного распределения, описанной в пункте 1: пусть $\{W_i\}$

– независимые случайные величины на отрезке $[0,1]$, причем W_i имеет бета-распределение $B(1-\alpha, \theta+i\alpha)$; рассмотрим последовательность $P_i = W_i \prod_{j=1}^{i-1}(1-W_j)$. Мера $PD(\alpha,\theta)$ есть распределение вектора $\{\tilde{P}_i\}$, где $\tilde{P}_1 \geqslant \tilde{P}_2 \geqslant \dots$ – перестановка $\{P_i\}$ в невозрастающем порядке. Область допустимых параметров есть множество (1.6).

Пусть τ – произвольная вероятностная мера на отрезке $[0,1]$, α,θ – допустимые параметры. Случайную меру $F = \sum_{i=1}^{\infty} \tilde{P}_j \delta_{X_j}$, где X – последовательность независимых величин с общим распределением τ, а вектор $\tilde{P}$ независим от X и $\tilde{P} \in PD(\alpha,\theta)$, будем называть обобщенной мерой Дирихле $\mathcal{D}(\tau,\alpha,\theta)$.

Соответствующее семейство мер на симметрических группах $M^n_{\alpha,\theta}$ задается следующей формулой:

$$M^n_{\alpha,\theta}(\pi) = \frac{(\theta+\alpha)(\theta+2\alpha)\dots(\theta+(k-1)\alpha)}{[\theta+1]_{n-1}} \prod_{j\geqslant 1} \left(\frac{[1-\alpha]_{j-1}}{(j-1)!}\right)^{r_j}, \tag{3.4}$$

где $k = c(\pi)$ – число циклов π, r_j – количество циклов π, имеющих длину j, $[x]_m = x(x+1)\dots(x+m-1)$ – символ Похгаммера.

Таким образом, выборка из $\mathcal{D}(\tau,\alpha,\theta)$ строится по следующему правилу. Выбираем $\{x_i\}$ – последовательность независимых величин с распределением τ. Полагаем $X_1 = x_1$. Пусть уже построены n элементов $X_1,\dots,X_n$, и из них n_j имеют значение x_j, $j = 1,\dots,k$. Тогда условное распределение X_{n+1} при условии, что заданы величины $X_1,\dots,X_n$, равно

$$\sum_{j=1}^{k} \frac{n_j-\alpha}{\theta+n}\delta_{x_j} + \frac{\theta+k\alpha}{\theta+n}\delta_{x_{k+1}}. \tag{3.5}$$

Предельным значениям параметров отвечают следующие схемы.

1. Если $\alpha = 1$, $\theta > -1$, то $M^n_{1,\theta}$ есть дельта-мера в тождественной перестановке, поэтому $X_i = x_i$ – независимые величины с распределением τ, т.е. $\mathcal{D}(\tau,1,\theta)$ с вероятностью 1 совпадает с τ (в представлении (3.1) все $\tilde{P}_i = 0$ п.н.).

2. Если $\theta = -\alpha$, $0 < \alpha < 1$, то $M^n_{\alpha,-\alpha}$ сосредоточена на перестановках, имеющих ровно один цикл. Отсюда $X_i = x_1$ при всех i, т.е. $\mathcal{D}(\tau,\alpha,-\alpha)$ есть δ_x, где x имеет распределение τ ($\tilde{P}_1 = 1$ п.н.). То же верно для $\theta \to \infty$, $\alpha \geqslant 0$.

3. Если $m = -\frac{\theta}{\alpha} \in \mathbb{N}$ фиксировано, $\alpha \to -\infty$, $\theta \to \infty$, то

$$M^n_{-\infty, m\infty} = \frac{1}{\prod_{i \geqslant 1}((i-1)!)^{r_i}} \cdot \frac{m(m-1)\dots(m-k+1)}{m^n}, \tag{3.6}$$

и в формуле (3.5) условное распределение X_{n+1} при заданных значениях $X_1, \dots, X_n$ принимает вид

$$\sum_{j=1}^{k} \frac{1}{m} \delta_{x_j} + \frac{m-k}{m} \delta_{x_{k+1}}. \tag{3.7}$$

В этом случае в представлении (3.1) $\tilde{P}_i = 1/m$, $i \leqslant m$, т.е.

$$F = \sum_{i=1}^{m} \frac{1}{m} \delta_{X_i} = \frac{1}{m} \sum_{i=1}^{m} \delta_{X_i}. \tag{3.8}$$

4. Формула распределения среднего значения по $\mathcal{D}(\tau, \alpha, \theta)$.

Теорема. *Пусть τ – вероятностная мера на отрезке $[0,1]$, (α, θ) – допустимые параметры, $F \in \mathcal{D}(\tau, \alpha, \theta)$. Пусть мера μ есть распределение функционала среднего значения $\int x dF(x)$. Тогда μ и τ связаны соотношением*

(1) *при* $\alpha, \theta \neq 0$

$$\Big(\int (z-u)^{-\theta} d\mu(u)\Big)^{-\frac{1}{\theta}} = \Big(\int (z-u)^{\alpha} d\tau(u)\Big)^{\frac{1}{\alpha}}, \quad z \in \mathbb{C} \setminus \mathbb{R}; \tag{4.1}$$

(2) *при* $\theta = 0$

$$\exp \int \ln(z-u)^{\alpha} d\mu(u) = \int (z-u)^{\alpha} d\tau(u), \quad z \in \mathbb{C} \setminus \mathbb{R}; \tag{4.2}$$

(3) *при* $\alpha = 0$

$$\int (z-u)^{-\theta} d\mu(u) = \exp \int \ln(z-u)^{-\theta} d\tau(u), \quad z \in \mathbb{C} \setminus \mathbb{R}. \tag{4.3}$$

В доказательстве теоремы мы используем следующее утверждение.

Лемма ([9], пример I.2.11). *Пусть*

$$f(t)=\sum_{n=1}^{\infty}\frac{f_n t^n}{n!},\quad g(t)=\sum_{n=1}^{\infty}\frac{g_n t^n}{n!} \tag{4.4}$$

– формальные степенные ряды. Рассмотрим композицию $H(t)=f(g(t))$ и ее разложение в степенной ряд $H(t)=\sum_{n=1}^{\infty}H_n t^n/n!$. Тогда коэффициенты H_n имеют вид

$$H_n=\sum_{k=1}^{n}f_k B_{n,k}(g), \tag{4.5}$$

где $B_{n,k}$ являются многочленами от коэффициентов ряда g. При этом имеет место следующая формула:

$$B_{n,k}=\sum_{\lambda}\frac{n!}{\prod_{i\geqslant 1}r_i!(i!)^{r_i}}\prod_{i\geqslant 1}g_i^{r_i}, \tag{4.6}$$

где суммирование ведется по всем разбиениям λ числа n, имеющим ровно k слагаемых, а r_j есть количество слагаемых λ, равных j.

Нам будет удобнее использовать вариант формулы (4.6), в котором суммирование ведется по элементам симметрической группы S_n. Поскольку для произвольного разбиения λ числа n количество перестановок, имеющих цикловую структуру λ, есть

$$z_\lambda=\frac{n!}{\prod_{i\geqslant 1}r_i!i^{r_i}}, \tag{4.7}$$

то (4.5) можно записать в виде

$$H_n=\sum_{\pi\in S_n}f_k\prod_{i\geqslant 1}\left(\frac{g_i}{(i-1)!}\right)^{r_i}, \tag{4.8}$$

где k – число циклов перестановки π, а r_j – количество циклов π, равных j.

Доказательство теоремы. Пусть $h_n=\int u^n d\mu(u)$, $p_n=\int u^n d\tau(u)$ – моменты мер μ и τ соответственно. По формуле (3.3)

$$\begin{aligned}h_n&=\sum_{\pi\in S_n}M^n_{\alpha,\theta}(\pi)\prod_{j\geqslant 1}p_j^{r_j}=\\&=\sum_{\pi\in S_n}\frac{(\theta+\alpha)\ldots(\theta+(k-1)\alpha)}{[\theta+1]_{n-1}}\cdot\prod\left(\frac{[1-\alpha]_{j-1}}{(j-1)!}\right)^{r_j}\prod p_j^{r_j}.\end{aligned} \tag{4.9}$$

Пусть $\alpha, \theta \neq 0$. Тогда (4.9) можно записать в виде

$$h_n = \sum_{\pi \in S_n} \frac{[\frac{\theta}{\alpha}]_k}{[\theta]_n} \prod_{j \geqslant 1} \left(\frac{-[-\alpha]_j p_j}{(j-1)!} \right)^{r_j}. \tag{4.10}$$

Обозначим $f_k = [\frac{\theta}{\alpha}]_k$, $g_j = -[-\alpha]_j p_j$, $H_n = [\theta]_n h_n$ $(k, j, n \geqslant 1)$. Тогда (4.10) примет вид

$$H_n = \sum_{\pi \in S_n} f_k \prod \left(\frac{g_j}{(j-1)!} \right)^{r_j}. \tag{4.11}$$

Но эта формула в точности совпадает с (4.8), поэтому для функций $f(z) = \sum_{k=1}^{\infty} \frac{f_k}{k!} z^k$, $g(z) = \sum_{k=1}^{\infty} \frac{g_k}{k!} z^k$ и $H(z) = \sum_{k=1}^{\infty} \frac{H_k}{k!} z^k$ выполнено соотношение

$$H(z) = f(g(z)). \tag{4.12}$$

Осталось вычислить функции f, g, и H. Нетрудно убедиться, что

$$\begin{aligned} f(z) &= (1-z)^{-\frac{\theta}{\alpha}} - 1; \\ g(z) &= -\sum_{k=1}^{\infty} \frac{[-\alpha]_k}{k!} p_k z^k = -\sum_{k=1}^{\infty} \frac{[-\alpha]_k}{k!} z^k \int u^k d\tau(u) = \\ &= 1 - \int (1-zu)^{\alpha} d\tau(u); \\ H(z) &= \int (1-zu)^{-\theta} d\mu(u) - 1. \end{aligned} \tag{4.13}$$

Подставляя эти выражения в формулу (4.12) и заменяя z на $\frac{1}{z}$, получаем искомое соотношение.

В случае $\theta = 0$ имеем

$$h_n = \sum_{\pi \in S_n} \frac{(k-1)!}{\alpha(n-1)!} \prod \left(\frac{-[-\alpha]_j p_j}{(j-1)!} \right)^{r_j}, \tag{4.14}$$

так что (4.11) выполнено для $H_n = \alpha(n-1)! h_n$, $f_k = (k-1)!$, $g_j = -[-\alpha]_j p_j$. Соответствующие функции $H(z) = -\int \ln(1-uz)^{\alpha} d\mu(u)$, $f(z) = \ln(1-z)$, $g(z) = 1 - \int (1-zu)^{\alpha} d\tau(u)$. Подставляя эти выражения в (4.12), получаем нужную формулу.

Соотношение для $\alpha = 0$ приведено в [5]. Для полноты укажем соответствующие формулы:

$$h_n = \sum_{\pi \in S_n} \frac{1}{[\theta]_n} \prod (\theta p_j)^{r_j}, \tag{4.15}$$

так что $H_n = [\theta]_n h_n$, $f_k = 1$, $g_j = \theta p_j (j-1)!$, т.е. $H(z) = \int \frac{d\mu(u)}{(1-zu)^\theta}$, $f(z) = e^z$, $g(z) = \int \ln \frac{1}{(z-u)^\theta} d\tau(u)$, откуда следует искомое соотношение.

Замечания. 1. На первый взгляд, не ясно, что правая часть формулы (4.1) стремится к правой части (4.3) при $\alpha \to 0$. Но это нетрудно увидеть, т.к.

$$\frac{1}{\alpha} \ln \int \left(1 - \frac{u}{z}\right)^\alpha d\tau(u) = \frac{1}{\alpha} \ln \int \left(1 + \alpha \ln\left(1 - \frac{u}{z}\right) + o(\alpha^2)\right) d\tau(u) =$$
$$= \frac{1}{\alpha} \ln \left(1 + \alpha \int \ln\left(1 - \frac{u}{z}\right) d\tau(u) + o(\alpha^2)\right) \to \int \ln\left(1 - \frac{u}{z}\right) d\tau(u).$$

2. При предельных значениях параметров рассматриваемое соотношение принимает следующую форму.

(1) При $\alpha = 1$, $\theta > -1$ формула (4.1) дает

$$\left(\int (z-u)^{-\theta} d\mu(u)\right)^{-\frac{1}{\theta}} = \int (z-u) d\tau(u) = z - p_1, \tag{4.16}$$

откуда μ есть дельта-мера в p_1. Это соответствует тому, что $\mathcal{D}(\tau, 1, \theta)$ равна τ почти наверное.

(2) При $\theta = -\alpha$, $0 < \alpha < 1$ (4.1) обращается в тождество, откуда получаем, что $\mu = \tau$. Действительно, в этом случае $\mathcal{D}(\tau, \alpha, -\alpha)$ есть δ_x, где x распределено по мере τ, так что распределение среднего случайной меры совпадает с τ.

(3) При $m \equiv -\frac{\theta}{\alpha} \in \mathbb{N}$, $\alpha \to -\infty$, $\theta \to \infty$, используя (3.6) и рассуждения аналогичные доказательству теоремы, получаем соотношение

$$\int e^{umz} d\mu(u) = \left(\int e^{uz} d\tau(u)\right)^m, \tag{4.17}$$

т.е. в терминах характеристических функций $f_\mu(t) = f_\tau(t/m)^m$. Это согласуется с формулой (3.8), которая показывает, что в данном случае μ есть распределение нормированной суммы $\frac{1}{m}\sum_{i=1}^m x_i$ независимых случайных величин с распределением τ.

3. Симметрия соотношения (4.1) относительно замены $(\tau, \alpha, \theta) \leftrightarrow (\mu, -\theta, -\alpha)$ показывает, что формула (4.9) останется справедливой после подстановки $h_i \leftrightarrow p_i$, $\alpha \to -\theta$, $\theta \to -\alpha$. Таким обра-

зом, имеет место следующая “формула восстановления” моментов меры τ по моментам меры μ:

$$p_n = \sum_{\pi\in S_n}(-1)^{k-1}\frac{(\alpha+\theta)\dots(\alpha+(k-1)\theta)}{[1-\alpha]_{n-1}}\cdot\prod\left(\frac{[1+\theta]_{j-1}}{(j-1)!}\right)^{r_j}\cdot\prod h_j^{r_j}.$$

4. Наше доказательство основано на том, что меры Ювенса–Питмана $M^n_{\alpha,\theta}$ представляются в виде

$$M^n(\pi) = \frac{a_k}{b_n}\cdot\prod_{i\geqslant 1}c_i^{r_i}. \tag{4.18}$$

Распределения такого вида рассматривались в работе [4], где они получаются с помощью т.н. модели Колчина, которая описывается следующим образом. Пусть $\varkappa = \{\varkappa_i\}$ и $\sigma = \{\sigma_i\}$ – два вероятностных распределения на множестве натуральных чисел. Выберем случайное число K, распределенное по мере $\varkappa$. Затем рассмотрим случайный K-вектор $(S_1,\dots,S_K)$, состоящий из независимых целочисленных случайных величин с общим распределением σ. Этот вектор задает случайное разбиение $\Lambda = (1^{R_1}2^{R_2}\dots)$ числа $N = S_1+\dots S_K$ $(R_j = \#\{i : S_i = j\})$. Для фиксированного $n\in\mathbb{N}$ рассмотрим условное распределение Λ при условии $N = n$; обозначим его P_n. Полученная мера на множестве разбиений числа n естественным образом соответствует инвариантной относительно внутренних автоморфизмов мере на симметрической группе S_n, а именно, $M^n(g) = P_n(\lambda)/z_\lambda$, где разбиение λ описывает цикловую структуру перестановки g, а z_λ есть число перестановок с цикловой структурой λ, задаваемое формулой (4.7). Нетрудно подсчитать, что

$$M^n(g) = \frac{k!\varkappa_k}{n!C_n}\cdot\prod_{i\geqslant 1}(i\sigma_i)^{r_i}. \tag{4.19}$$

Теперь из доказательства теоремы видно, что ее утверждение может быть записано в терминах модели Колчина следующим образом:

$$\int\varphi_N(uz)d\mu(u) = \varphi_K\left(\int\varphi_S(uz)d\tau(u)\right), \tag{4.20}$$

где φ_K и φ_S – производящие функции распределений $\varkappa$ и σ соответственно, а $\varphi_N(u) = \varphi_K(\varphi_S(u))$ – производящая функция распределения суммы $N = S_1+\dots+S_K$. Следует заметить, что при фиксированных α и θ мера Ювенса–Питмана $M^n_{\alpha,\theta}$ может быть получена в модели Колчина с различными параметрами $\varkappa$ и σ (см. [4]), но при подстановке в (4.20) соответствующих производящих функций получаются формулы, приводящиеся к (4.1)–(4.3) тривиальными преобразованиями. Отметим также, что, как показано в [4], среди всех распределений, получаемых в модели Колчина, только меры Ювенса–Питмана и их предельные варианты удовлетворяют условию согласованности (3.2).

Литература

1. D. Blackwell, J. B. MacQueen, *Ferguson distributions via Pólya urn schemes.* — Ann. Statist. **1** (1973), 353–355.
2. P. Diaconis, J. Kemperman, *Some New Tools for Dirichlet Priors.* — Preprint (1994), 1–12.
3. T. Ferguson, *A Bayesian analysis of some nonparametric problems.* — Ann. Statist. **1** (1973), 209–230.
4. S. Kerov, *Coherent Random Allocations, and the Ewens–Pitman Formula.* — Preprint (1995), 1–15.
5. S. Kerov, *Interlacing measures.* — Preprint (1996), 1–53.
6. J. F. C. Kingman, *Random partitions in population genetics.* — Proc. R. Soc. Lond. A **361** (1978), 1–20.
7. J. F. C. Kingman, *The representation of partition structures.* — J. London Math. Soc.(2) **18** (1978), 374–380.
8. J. F. C. Kingman, *Poisson Processes.* Oxford, Clarendon Press, 1993.
9. И. Макдональд, *Симметрические функции и многочлены Холла.* Москва, Мир, 1985.
10. J. Pitman, *The Two-parameter Generalization of Ewens' Random Partition Structure.* — Tech. Report **345** (1992), 1–23.
11. J. Pitman, *Some Developments of the Blackwell-MacQueen Urn Scheme.* — Preprint (1996).
12. J. Pitman, M. Yor, *The two-parameter Poisson–Dirichlet distribution.* — Tech. Report **427** (1995), 1–46.

Tsilevich N. V. Distributions of mean values for some random measures.

Let τ be a probability measure on $[0,1]$. We consider a generalization of the classic Dirichlet process – the random probability measure $F = \sum P_i \delta_{X_i}$, where $X = \{X_i\}$ is a sequence of independent random variables with the common distribution τ and $P = \{P_i\}$ is independent of X and has the two-parameter Poisson–Dirichlet distribution $PD(\alpha,\theta)$ on the unit simplex. The main result is the formula connecting the distribution μ of the random mean value $\int x dF(x)$ with the parameter measure τ.

С.-Петербургское отделение
Математического института
им. В. А. Стеклова РАН

E-mail: natalia@pdmi.ras.ru

Поступило 15 декабря 1996 г.