

ДИСКРЕТНЫЕ ЗАДАЧИ В ТЕОРИИ ВЕРОЯТНОСТЕЙ

В. А. Иванов, Г. И. Ивченко, Ю. И. Медведев

ВВЕДЕНИЕ

Дискретные задачи занимают значительное место в исследованиях по теории вероятностей. Среди основных направлений в этой области можно выделить такие, как комбинаторные задачи теории вероятностей, задачи о размещении частиц по ячейкам, случайные отображения и графы, задачи выборочного обследования конечных совокупностей.

Большое внимание к этой проблематике объясняется, в первую очередь, многочисленными приложениями соответствующих результатов в самых различных областях науки и техники: математической статистике, теории автоматов, вычислительной технике, статистической физике, биологии, медицине, социологии, экономике. Во многом этому способствует все более широкое применение при проведении практических расчетов современных быстродействующих вычислительных машин.

В указанной проблематике одно из центральных мест занимают направления, связанные со случайными размещениями частиц по ячейкам. Это — наиболее интенсивно развивающаяся в настоящее время ветвь дискретной теории вероятностей, отличающаяся простотой и четкостью постановок задач и возможностью применения для их решения практически всех современных методов асимптотического анализа.

Состояние этой области исследований до 1975—1976 годов с исчерпывающей полнотой отражено в вышедшей в 1976 году монографии В. Ф. Колчина, Б. А. Севастьянова и В. П. Чистякова «Случайные размещения» [75], а также в обзоре В. Ф. Колчина и В. П. Чистякова [76], изданном ВИНТИ в серии «Итоги науки».

За время, прошедшее после выхода этих публикаций, исследования по дискретным задачам теории вероятностей продолжали интенсивно развиваться. Появились новые подходы и методы для решения соответствующих задач, накопилось большое число новых интересных результатов. Поэтому назрела

необходимость в написании нового обзора, который отразил бы достигнутый на сегодняшний день уровень развития данной проблематики.

Предлагаемый вниманию читателя обзор и преследует эту цель. Вместе с тем, ограниченный объем статьи не позволяет охватить все направления дискретной теории вероятностей. Некоторые из этих направлений заслуживают отдельного самостоятельного изложения. Так, в данный обзор не включены вопросы анализа случайных графов. Как нам известно, в настоящее время по этому направлению готовится специальная монография. Мы также не рассматриваем большой класс комбинаторных задач теории вероятностей, которые в значительной мере отражены в недавно вышедших книгах В. Н. Сачкова [98, 99, 100], и вопросы общей теории случайных отображений, изложенные в монографии В. Ф. Колчина [74].

§ 1. Разделимые статистики

При статистическом анализе дискретных (или сводящихся к ним при соответствующей группировке наблюдений) случайных последовательностей часто возникает задача проверки гипотезы H_0 о том, что данная последовательность получена по схеме независимых испытаний с N исходами и вероятностями исходов $\mathbf{q} = (q_1, \dots, q_N)$, $q_k > 0$, $k = 1, \dots, N$, $\sum_{k=1}^N q_k = 1$, при конкурирующей сложной гипотезе $H = \{H(\mathbf{p})\}$, сохраняющей предположения независимости и однородности испытаний, но оставляющей произвольными вероятности $\mathbf{p} = (p_1, \dots, p_N)$ исходов. При этом представляют интерес как случаи, когда с ростом числа испытаний n число возможных исходов N остается конечным или медленно растет: при $n \rightarrow \infty$ величина $a = \min_{1 \leq m \leq N} np_m \rightarrow \infty$,

так и случаи, когда последнее условие не выполняется. В этих ситуациях (при $a \leq c < \infty$) известные классические методы решения задач, основанные на асимптотической нормальности вектора чисел реализаций (частот) исходов $\mathbf{h} = (h_1, \dots, h_N)$, уже не дают должного эффекта и требуются другие подходы к их решению. Схемы, в которых величины np_m остаются ограниченными при $n \rightarrow \infty$, называют схемами с малыми выборками [80]. Рассмотрению таких схем уделяется в последние годы значительное внимание в литературе и соответствующее направление исследований связано, прежде всего, с понятием разделимой статистики (р. с.).

1.1. Характеристическая функция и моменты р. с. Пусть n — число испытаний с возможными исходами E_1, \dots, E_N , вероятности которых равны, соответственно, p_1, \dots, p_N , $p_k > 0$, $k = 1, \dots, N$, $p_1 + \dots + p_N = 1$ (гипотеза $H(\mathbf{p})$). Обозначим через

h_k число испытаний с исходом E_k , $k=1, \dots, N$, а через $\mathbf{h} = (h_1, \dots, h_N)$ — вектор частот исходов. Как известно, случайная величина (сл. в.) \mathbf{h} имеет полиномиальное распределение с параметрами n и $\mathbf{p} = (p_1, \dots, p_N)$.

Пусть $f_m(x)$, $m=1, \dots, N$, — функции, определенные во всех целых неотрицательных точках $x=0, 1, 2, \dots$. Функцию от вектора \mathbf{h} называют разделимой статистикой [80], если она представима в виде

$$L_N(\mathbf{h}) = \sum_{m=1}^N f_m(h_m). \quad (1.1)$$

В (1.1) функции $f_m(x)$ могут зависеть от N .

Этот класс сл. в. встречается во многих статистических задачах, где рассматриваются схемы с конечным числом исходов. Он включает в себя (при соответствующей нормировке и центрировке) статистики, лежащие в основе наиболее важных критериев для проверки статистических гипотез относительно вероятностей появления исходов. Другая особенность этого класса определяется тем, что он поддается асимптотическому анализу с единых позиций и едиными методами. Это позволяет исследовать мощность и другие характеристики соответствующих статистических критериев в зависимости от поведения параметров изучаемой схемы в различных условиях.

Приведем важнейшие примеры р. с.

1. Для проверки гипотезы $H_0 = H(\mathbf{q})$ часто используется статистика

$$\begin{aligned} X_N^2 &= \sum_{m=1}^N \frac{(h_m - nq_m)^2}{nq_m} = \frac{1}{n} \sum_{m=1}^N \frac{h_m^2}{q_m} - n = \\ &= \frac{N}{n} \sum_{m=1}^N f_m(h_m) - n \equiv \frac{N}{n} \chi_1^2 - n, \end{aligned} \quad (1.2)$$

где $f_m(x) = x^2/(Nq_m)$, $m=1, \dots, N$. Таким образом, статистика классического критерия χ^2 Пирсона является р. с.

2. Для проверки той же гипотезы иногда используют отношение правдоподобия

$$\lambda_N = \frac{Q(\mathbf{q})}{\max_{\mathbf{p}} Q(\mathbf{p})} = \prod_{m=1}^N \left(\frac{nq_m}{h_m} \right)^{h_m}, \quad Q(\mathbf{p}) = p_1^{h_1} \dots p_N^{h_N}.$$

Следовательно,

$$-\ln \lambda_N = \sum_{m=1}^N h_m \ln \left(\frac{h_m}{Nq_m} \right) - n \ln \left(\frac{n}{N} \right) = \sum_{m=1}^N f_m(h_m) - n \ln \left(\frac{n}{N} \right), \quad (1.3)$$

где $f_m(x) = x \ln \left(\frac{x}{Nq_m} \right)$, $m=1, \dots, N$.

3. Если в (1.1) все функции $f_m(x)$ представляют собой некоторые многочлены одной и той же степени r , то такие р. с. можно назвать полиномиальными степени r . При $r=1$ имеем подкласс линейных р. с., при $r=2$ — квадратичных и т. д. Иногда удобнее рассматривать факториальную р. с. степени $r \geq 2$:

$$L_N = \sum_{m=1}^N c_m h_m^{[r]} = \sum_{m=1}^N c_m h_m (h_m - 1) \dots (h_m - r + 1). \quad (1.4)$$

4. Пусть A — некоторое выделенное подмножество исходов $\{1, 2, \dots, N\}$. Положив в (1.1) $f_m(x) \equiv 0$ при $m \notin A$, получим р. с.

$$L_N = \sum_{m \in A} f_m(h_m), \quad (1.5)$$

учитывающую информацию, заключенную в наблюдаемых частотах выделенной группы исходов.

5. Положим в (1.1) $f_m(x) = d_m$ при $x \geq r$ и $f_m(x) = 0$ при $x < r$, где d_m — заданные константы и $r \geq 0$ — заданное целое число. Тогда

$$L_N = \sum_{m: h_m \geq r} d_m \quad (1.6)$$

есть сумма «весов» d_m тех исходов, каждый из которых реализовался не менее r раз.

6. Статистика вида

$$L_N = \sum_{m=1}^N a_m^{h_m} \quad (1.7)$$

называется показательной р. с.

7. Важный подкласс р. с. образуют симметрические статистики (с. р. с.), когда в (1.1) все функции $f_m(x)$ одинаковы:

$$L_N = \sum_{m=1}^N f(h_m). \quad (1.8)$$

Класс с. р. с. совпадает с классом линейных комбинаций $\sum_{r=0}^n c_r \mu_r$, где $\mu_r = \mu_r(n, N)$ — число исходов, реализовавшихся при n испытаниях ровно r раз каждый.

Точные формулы для распределений р. с. при больших значениях n и N довольно сложны даже для равновероятной схемы ($p_1 = \dots = p_N = \frac{1}{N}$) и не пригодны для практических расчетов.

Поэтому основное внимание при изучении р. с. уделяется исследованию их асимптотического поведения при $n, N \rightarrow \infty$. Для существования моментов у сл. в. $L_N(h)$ необходимо несколько

ограничить рост функций $f_m(x)$ при $x \rightarrow \infty$. В дальнейшем предполагается, что при некоторых постоянных $a, b > 0$ и всех N выполняются неравенства

$$|f_m(x)| \leq a e^{bx}, \quad m=1, \dots, N. \quad (1.9)$$

При этом ограничимся случаем регулярных схем, т. е. когда при всех N $\min_{1 \leq m \leq N} N p_m \geq C_1 > 0$, и малых выборок: $\min_{1 \leq m \leq N} n p_m \leq C_2 < \infty$.

Вывод предельных при $n, N \rightarrow \infty$ распределений р. с. $L_N(\mathbf{h})$ основан на том хорошо известном факте, что распределение полиномиального вектора $\mathbf{h} = (h_1, \dots, h_N)$ совпадает с условным распределением независимых сл. в. X_1, \dots, X_N при условии $\sum_{m=1}^N X_m = n$, где сл. в. X_m имеет распределение Пуассона $\Pi(np_m)$ ($X_m \sim \Pi(np_m)$), $m=1, \dots, N$. Использование интегрального представления для характеристической функции $E \exp\{itL_N(\mathbf{h})\}$ позволяет применять при изучении асимптотического поведения сл. в. $L_N(\mathbf{h})$ метод перевала. Именно, имеет место формула

$$E e^{itL_N(\mathbf{h})} = \frac{n!}{N^{n+2\pi i}} \oint e^{Nz} \prod_{m=1}^N \left(\sum_{k=0}^{\infty} \pi_k(N p_m z) e^{itf_m(k)} \right) \frac{dz}{z^{n+1}}, \quad (1.10)$$

где $\pi_k(\lambda) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k=0, 1, \dots$, и интегрирование ведется по любому замкнутому контуру, обходящему точку $z=0$.

Предельное поведение статистики $L_N(\mathbf{h})$ во многом определяется асимптотикой ее дисперсии. При сформулированных выше условиях для моментов $L_N(\mathbf{h})$ справедливы следующие представления:

$$\begin{aligned} \mu_N = EL_N(\mathbf{h}) &= \left[\sum_{m=1}^N E f_m(X_m) \right] (1 + o(1)), \\ \sigma_N^2 = DL_N(\mathbf{h}) &= \\ &= \left[\sum_{m=1}^N D f_m(X_m) - \frac{1}{n} \left(\sum_{m=1}^N \text{cov}(f_m(X_m), X_m) \right)^2 \right] (1 + o(1)). \end{aligned} \quad (1.11)$$

Выделим некоторые частные случаи этих общих соотношений. Если при некотором целом $r \geq 2$ $f_m(0) = f_m(1) = \dots = f_m(r-1) = 0$, $f_m(r) \neq 0$, $m=1, \dots, N$ (например, для сл. в. μ_r или полиномиальных статистик вида (1.4)), то при

$$n, N \rightarrow \infty, \quad n/N \rightarrow 0, \quad \max_m N p_m \leq C < \infty \quad (1.12)$$

$$\begin{aligned}\mu_N &= \frac{n^r}{r!} \left(\sum_{m=1}^N p_m^r f_m(r) \right) (1 + o(1)), \\ \sigma_N^2 &= \frac{n^r}{r!} \left(\sum_{m=1}^N p_m^r f_m^2(r) \right) (1 + o(1)).\end{aligned}\tag{1.13}$$

Пусть теперь при некотором m $f_m(1) \neq 0$ и выполнено (1.12) (без потери общности, можно предполагать, что $f_m(0) = 0$ при всех m).

Тогда

$$\begin{aligned}\mu_N &= nE\xi(1 + o(1)), \\ \sigma_N^2 &= nD\xi + \frac{n^2}{2} \left[\sum_{m=1}^N p_m^2 (f_m^2(2) - 4f_m^2(1)) - \right. \\ &\quad \left. - 4E\xi \sum_{m=1}^N p_m^2 \Delta^2 f_m(0) \right] + O\left(\frac{n^3}{N^2}\right),\end{aligned}$$

где через ξ обозначена сл. в., принимающая значения $f_m(1)$ с вероятностями p_m , $m = 1, \dots, N$, и Δ — оператор разности $\Delta f(x) = f(x+1) - f(x)$. Эти формулы, в частности, для с. р. с при $f(2) \neq 2f(1)$ принимают вид

$$\begin{aligned}\mu_N &= nf(1)(1 + o(1)), \\ \sigma_N^2 &= \frac{n^2}{2} (f(2) - 2f(1))^2 \sum_{m=1}^N p_m^2 (1 + o(1)).\end{aligned}$$

1.2. Предельные теоремы. Для разделимых статистик в качестве предельных законов (при соответствующих нормировках) могут быть нормальный и пуассоновский законы и их свертки.

Пусть при $n, N \rightarrow \infty$ выполняются условия

$$\max_{1 \leq m \leq N} np_m = o(1), \quad n^2 \sum_{m=1}^N p_m^2 \rightarrow 2\lambda, \quad \lambda < \infty,\tag{1.14}$$

и при этом

$$\begin{aligned}\sigma \equiv f(2) - 2f(1) &\neq 0, \quad f(x) = \sum_{m=1}^N p_m f_m(x), \\ nE\eta^2 &\rightarrow \gamma^2 < \infty, \quad \eta = \frac{1}{\sigma} (\xi - E\xi),\end{aligned}$$

$$nE|\eta|^3 \rightarrow 0, \quad n^2 \sum_{m=1}^N p_m^2 (|f_m(1) - f(1)| + |f_m(2) - f(2)|) \rightarrow 0.$$

Тогда предельным законом распределения сл. в. $\chi = \frac{1}{\sigma} [L_N(\mathbf{h}) - nE_{\chi}]$ при гипотезе $H(\mathbf{p})$ является свертка законов $\Pi(\lambda)$ и нормального $\mathcal{N}(0, \gamma^2)$; если при этом $\gamma = 0$ ($\lambda = 0$), то сл. в. χ распределена в пределе по закону $\Pi(\lambda)$ ($\mathcal{N}(0, \gamma^2)$).

Следует заметить, что в этих условиях $\sigma^{-2}DL_N(\mathbf{h}) \rightarrow \lambda + \gamma^2$, т. е. этот случай соответствует конечной дисперсии.

В рамках условий (1.14) может оказаться, что р. с. выбрана так, что $nD\xi \rightarrow \infty$. В этом случае $DL_N(\mathbf{h}) \rightarrow \infty$ и сл. в. $L_N(\mathbf{h})$ распределена в пределе по нормальному закону. Условия (1.14) означают, что $n = O(\sqrt{N})$. В других ситуациях, т. е. когда $n^2/N \rightarrow \infty$, рост дисперсии $DL_N(\mathbf{h})$, как правило, влечет за собой асимптотическую нормальность сл. в. $L_N(\mathbf{h})$. Если же с ростом n $DL_N(\mathbf{h})$ остается ограниченной, то $L_N(\mathbf{h})$ может быть в пределе распределена по закону Пуассона. Например, имеет место следующий результат: пусть r — минимальное целое значение k , при котором $f_m(k) \neq kf_m(1)$ хотя бы при одном $m = 1, \dots, N$ и $r \geq 2$. При этом

$$\max_{1 \leq m \leq N} \left| f_m(r) - \sum_{m=1}^N p_m f_m(r) \right| = o(1), \quad \frac{n^r}{r!} \sum_{m=1}^N p_m r^m \rightarrow \lambda, \quad 0 < \lambda < \infty.$$

Тогда сл. в. $L_N(\mathbf{h})$ при гипотезе $H(\mathbf{p})$ распределена в пределе по закону $\Pi(\lambda)$.

Перенесем некоторые общие результаты на важные частные случаи, например, на статистику X_1^2 . Пусть $n, N \rightarrow \infty$ и $n = O(\sqrt{N})$. Тогда (см. (1.2)) $D\chi_1^2$ остается ограниченной, если и только если

$$\gamma^2(N) = \frac{1}{N^2} \left(\sum_{m=1}^N \frac{p_m}{q_m^2} - \left(\sum_{m=1}^N \frac{p_m}{q_m} \right)^2 \right) = O\left(\frac{1}{n}\right). \quad (1.15)$$

Это условие означает, что полиномиальная схема при основной гипотезе $H(\mathbf{q})$ должна быть в определенном смысле близкой к равновероятной.

Пусть $q_m = \frac{1}{N} \left(1 + \frac{d_m}{\sqrt{n}} \right)$, $|d_m| \leq C < \infty$, $m = 1, \dots, N$, $\frac{n^2}{2} \sum_{m=1}^N p_m^2 \rightarrow \lambda$, $\lambda < \infty$, $\frac{1}{4} \left(\sum_{m=1}^N p_m d_m^2 - \left(\sum_{m=1}^N p_m d_m \right)^2 \right) \rightarrow \gamma^2$, $\gamma^2 < \infty$. Тогда при гипотезе $H(\mathbf{p})$ предельным законом распределения сл. в. $\frac{nx_N^2}{N\delta} + \frac{n^2}{2N} - \frac{n}{2}$, где $\delta = \frac{2}{N} \sum_{m=1}^N \frac{p_m}{q_m}$, является свертка $\Pi(\lambda) * \mathcal{N}(0, \gamma^2)$.

Если же при $N, n \rightarrow \infty$ выполняется одно из условий

- а) $n^2/N \leq C < \infty$, $n^{1/3}\gamma^2(N) \rightarrow \infty$;
- б) $n^2/N \rightarrow \infty$, $n/N \rightarrow 0$,

$$\frac{n^2}{ND\chi_1^2} = o\left(\frac{1}{n} + \frac{1}{\max_{m,k} |q_m^{-1} + q_k^{-1}|}\right); \quad (1.16)$$

в) $0 < C_1 \leq \frac{n}{N} \leq C_2 < \infty$, $DX_1^2 \sim CN$, $C > 0$, то сл. в. X_N^2 ,

соответствующим образом нормированная и центрированная, распределена в пределе по закону $\mathcal{N}(0,1)$. Во всех указанных случаях $DX_N^2 \rightarrow \infty$ при $n, N \rightarrow \infty$.

Приведенные результаты получены в работах Ю. И. Медведева [80, 81, 82], Холста [148] и Морриса [160]. Уточнение нормальной предельной теоремы для линейной separable статистики проведено Э. В. Хмаладзе [117].

Перенесение результатов на важные частные случаи р. с. $(X_N^2, -\ln \lambda_N, \text{с. р. с.})$ проведено в [82]. Предельные распределения сл. в. X_N^2 в схеме серий при $n, N \rightarrow \infty$ и $\min_{1 \leq m \leq N} np_m \rightarrow \infty$ исследовались в работах С. Х. Туманяна [109, 110]. Стеком в [172] доказана нормальная предельная теорема для сл. в. X_N^2 в случае «сближающихся» гипотез ($n \max_{1 \leq m \leq N} (p_m - q_m)^2 \rightarrow 0$) в условиях в) соотношений (1.16). При тех же условиях ($n/N \geq C > 0$) для равновероятной основной гипотезы ($q_m = \frac{1}{N}$, $m = 1, \dots, N$) асимптотическая нормальность статистики X_N^2 доказана Харрисом и Парком [143]. Локальная нормальная теорема для сумм квадратов частот при $N \rightarrow \infty$ получена в работе В. Ф. Колчина [72]. Эмпирическое исследование статистик X_N^2 и $-\ln \lambda_N$ при $n/N \rightarrow \alpha$, $0 < \alpha_1 \leq \alpha \leq \alpha_2 < \infty$ проведено Кёлером и Ларнтзом [158].

В случае, когда нулевая гипотеза соответствует равновероятной схеме, статистика отношения правдоподобия эквивалентна статистической оценке энтропии $\sum_{m=1}^N h_m \ln h_m$. Предельные распределения этой статистики при фиксированном N или в схеме серий при некоторых ограничениях на параметры n , p_m были получены в работах А. М. Зубкова [23], Т. А. Азларова и Р. Мухамедхановой [1] и Р. Мухамедхановой [90, 91].

Понятие separable статистики можно распространить и на обобщенные схемы размещения, т. е. когда вместо полиномиального вектора частот $\mathbf{h} = (h_1, \dots, h_N)$ рассматривается некоторый целочисленный сл. вектор $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)$, компоненты которого неотрицательны, $\sum_{i=1}^N \eta_i = n$ и их совместное распределение представимо в виде

$$P(\eta_i = k_i, i = 1, \dots, N) = \\ = P(\xi_i = k_i, i = 1, \dots, N | \xi_1 + \dots + \xi_N = n),$$

где ξ_1, \dots, ξ_N — независимые одинаково распределенные неотрицательные целочисленные сл. в. Такие схемы впервые рассматривались в работе В. Ф. Колчина [70]. Холстом [147] для такой схемы рассматривалась сл. в. вида

$$\sum_{m=1}^M f(\eta_m), \quad M \leq N,$$

и исследовались ее предельные распределения при $n, M \rightarrow \infty$.

Симметрическим разделимым статистикам посвящена обширная литература. Весьма полная библиография по этой проблематике содержится в монографии В. Ф. Колчина, Б. А. Севастьянова и В. П. Чистякова [75]. Результаты, полученные после выхода этой книги, кратко изложены в § 4 данного обзора.

1.3. Разделимые статистики для нескольких полиномиальных схем. Для решения ряда статистических задач, в частности, задачи о проверке однородности s полиномиальных выборок, необходимо расширить понятие р. с. Пусть заданы независимые в совокупности случайные векторы h_1, \dots, h_s , где $h_j = (h_{1j}, \dots, h_{Nj})$ имеет полиномиальное распределение $M(n_j, p_{1j}, \dots, p_{Nj}) \left(\sum_{m=1}^N h_{mj} = n_j, \sum_{m=1}^N p_{mj} = 1 \right)$, $j = 1, \dots, s$, и $k \geq 1$ наборов функций $\{f_{1j}(x_1, \dots, x_s), \dots, f_{Nj}(x_1, \dots, x_s)\}$, $j = 1, \dots, k$, определенных при неотрицательных целочисленных значениях аргументов x_1, \dots, x_s . Образум с помощью этих функций k -мерную статистику $L_N = (L_{N1}, \dots, L_{Nk})$, где

$$L_{Nj} = \sum_{m=1}^N f_{mj}(h_{m1}, \dots, h_{ms}), \quad j = 1, \dots, k. \quad (1.17)$$

По аналогии с одномерным случаем, статистики вида (1.17) также называют разделимыми. Р. с. для нескольких полиномиальных последовательностей в схеме малых выборок изучались Г. И. Ивченко и В. В. Левиным в работах [59, 78, 79]. Ими доказана предельная теорема о сходимости совместного распределения статистик L_{N1}, \dots, L_{Nk} , соответствующим образом центрированных и нормированных, к многомерному нормальному закону в условиях, когда $N, n_1, \dots, n_s \rightarrow \infty$ так, что $n_j/N \rightarrow \alpha_j$, $0 < \alpha_j < \infty$, $j = 1, \dots, s$.

В тех случаях, когда функции f_{mj} не зависят от m , говорят о симметрических р. с. (с. р. с.). Класс с. р. с. совпадает с классом линейных комбинаций сл. в. μ_{r_1, \dots, r_s} , где μ_{r_1, \dots, r_s} есть число тех m , для которых реализуется событие $\{h_{m1} = r_1, \dots, h_{ms} = r_s\}$.

Величины μ_{r_1, \dots, r_s} можно интерпретировать в терминах размещения частиц по ячейкам как число ячеек, содержащих r_1 частиц 1-го типа, \dots , r_s частиц s -го типа после размещения заданных чисел частиц s различных типов по N ячейкам. В схеме с частицами двух типов ($s=2$), трехмерная ($k=3$) с. р. с. $(\mu_{00}, \sum_{r \geq 0} \mu_{0r}, \sum_{r \geq 0} \mu_{r0})$ изучалась в работе Т. Ю. Поповой [92].

Асимптотическая нормальность с. р. с. для s типов частиц доказана в работе А. Асимова [4].

1. 4. Сходимость к случайным процессам. Рассмотрим р. с.

$\xi_N(n) = \sum_{m=1}^N f_m(h_m)$, $h_1 + \dots + h_N = n$, как случайную функцию параметра n (этот процесс введен и исследован в работе Г. И. Ивченко и В. В. Лёвина [59]). Если положить

$$\xi_N(n_1 + \dots + n_j) = L_{Nj} = \sum_{m=1}^N f_m(h_{m1} + \dots + h_{mj}), \quad j = 1, \dots, s,$$

то исследование конечномерных распределений процесса $\xi_N(n)$ сводится к изучению р. с. (L_{N1}, \dots, L_{Ns}) в случае, когда $p_{mj} = p_m$, $f_{mj}(x_1, \dots, x_s) = f_m(x_1 + \dots + x_{j1})$, $j = 1, \dots, s$, $m = 1, \dots, N$. Если $\xi_N(n)$ есть с. р. с. и при этом $f(x) = 1$ при $x = r \geq 0$ и $f(x) = 0$ в остальных точках, то приходим к случаю, изучавшемуся в работах Б. А. Севастьянова [102] и Ю. В. Болотникова [12].

Для схемы с несколькими полиномиальными выборками можно по аналогии ввести понятие случайного поля $\xi_N(n_1, \dots, n_s)$, как случайной функции переменных n_1, \dots, n_s . Г. И. Ивченко и В. В. Лёвиным изучены условия сходимости распределения этого случайного поля.

Другой тип случайного процесса в полиномиальной схеме возникает при рассмотрении накопленной суммы

$$\eta_N(l) = \sum_{m=1}^l f_m(h_m)$$

как случайной функции параметра l , $1 \leq l \leq N$. Исследование конечномерных распределений процесса $\eta_N(l)$ эквивалентно изучению р. с. (L_{N1}, \dots, L_{Nh}) , где $L_{Nj} = \eta_N(l_j)$, $1 \leq l_1 < l_2 < \dots < l_h \leq N$. В работе Г. И. Ивченко и В. В. Лёвина [59] приводятся условия сходимости этого процесса (в смысле сходимости конечномерных распределений) к гауссовскому процессу.

В работах Э. В. Хмаладзе [118, 119] развивается мартингалльный подход к изучению р. с. Запишем р. с. в виде

$$L_N = \sum_{m=1}^N f_N\left(h_m, \frac{m}{N}\right) \quad (1.18)$$

и рассмотрим процесс $X^{n,N}$, образованный частичными суммами

$$X^{n,N}(t) = \frac{1}{\sqrt{N}} \sum_{m \leq Nt} \left[f_N \left(h_m, \frac{m}{N} \right) - E f_N \left(h_m, \frac{m}{N} \right) \right]. \quad (1.19)$$

Этот процесс является семимартингалом относительно потока σ -алгебр $\{\mathcal{F}_m^n\}$, $m=1, \dots, N$, где \mathcal{F}_m^n порождена сл. в. h_1, \dots, h_{m-1} , ($h_h = h_h(n)$). Показывается, что к мартингальной составляющей этого семимартингала применимы предельные теоремы для семимартингалов. Получено утверждение о сходимости по распределению процесса $X^{n,N}(t)$ к процессу Ито $X(t)$ в случае, когда $n=N$. Приведено также конструктивное описание предельного процесса $X(t)$.

1.5. Рандомизированные р. с. Сопоставим каждому исходу полиномиальной схемы случайную функцию дискретного аргумента $\eta_m(k)$, $k=0, 1, 2, \dots$, $m=1, \dots, N$. Пусть для любых наборов (k_1, \dots, k_N) , $k_m=0, 1, 2, \dots$, $m=1, \dots, N$, сл. в. $\eta_1(k_1), \dots, \eta_N(k_N)$ независимы в совокупности.

Случайная функция вида

$$\mathcal{L}_N(\mathbf{h}) = \sum_{m=1}^N \eta_m(h_m) \quad (1.20)$$

называется рандомизированной разделимой статистикой (р. р. с.) [32]. В работах В. А. Иванова и С. А. Лапина [32, 36] доказывается асимптотическая нормальность р. р. с. при $n, N \rightarrow \infty$, $\alpha = n/N \geq \alpha_0 > 0$ и при некоторых ограничениях, наложенных на распределения сл. функций $\eta_m(k)$, $m=1, \dots, N$, и поведение вероятностей p_1, \dots, p_N . Показывается полезность введения таких статистик при рассмотрении задачи проверки простой гипотезы о виде вектора вероятностей переходов конечного вероятностного автомата без памяти против сложной альтернативы.

1.6. Обобщенные разделимые статистики. Пусть $f_m(x_1, \dots, x_{s+1})$, $m=1, \dots, N$, — функции от $s+1$ переменного. Обобщенными разделимыми статистиками назовем сл. в. вида

$$L_{Ns}(\mathbf{h}) = \sum_{m=1}^M f_m(h_m, \dots, h_{m+s}), \quad (1.21)$$

где M может быть равным $N-s$ или N . В последнем случае сложение в индексах берется по mod M . Статистики такого типа изучались в работах А. П. Баранова [5], Холста [150, 153], Г. И. Ивченко и С. В. Цуканова [69] и др. авторов не только для полиномиальной схемы, но и для других схем размещения, в которых совместное распределение сл. в. (h_1, \dots, h_N) представимо в виде

$$\begin{aligned} P(h_i = k_i, i = 1, \dots, N) = \\ = P(\xi_i = k_i, i = 1, \dots, N \mid \xi_1 + \dots + \xi_N = n), \end{aligned} \quad (1.22)$$

где ξ_1, \dots, ξ_N — независимые неотрицательные целочисленные сл. в. Получены предельные теоремы для таких обобщенных схем размещения, в частности, найдены условия асимптотической нормальности сл. в. $L_{Ns}(\mathbf{h})$ при $n, N \rightarrow \infty, n/N \rightarrow a$.

1.7. Р. с. в схеме выбора без возвращения. В теории выборочного обследования конечных совокупностей часто рассматриваются выборочные процедуры, соответствующие схеме простого случайного выбора без возвращения. В этих случаях основой для получения статистических выводов служит выборка из n различных элементов ($n < N, N$ — число элементов совокупности), причем предполагается, что все C_N^n возможных вариантов выбора равновероятны.

Пусть в соответствии с описанной схемой извлечено s независимых выборок объемами n_1, \dots, n_s . Поставим в соответствии l -й выборке двоичный вектор $\mathbf{h}_l = (h_{l1}, \dots, h_{lN})$, где $h_{lm} = 1$, если m -й элемент совокупности вошел в выборку и $h_{lm} = 0$ в противном случае. Разделимой статистикой в этой схеме будем называть (многомерную) статистику (S -статистику): $S_N(\mathbf{h}) = (S_{N1}(\mathbf{h}), \dots, S_{Nk}(\mathbf{h}))$, где

$$S_{Nj}(\bar{\mathbf{h}}) = \sum_{m=1}^N f_{jm}^{(N)}(h_{1m}, \dots, h_{sm}), \quad j = 1, \dots, k, \quad (1.23)$$

и $f_{jm}^{(N)}(x_1, \dots, x_s)$ — заданные функции, определенные на множестве s -мерных двоичных векторов (x_1, \dots, x_s) .

Работа Г. И. Ивченко и В. В. Лёвина [60] посвящена исследованию условий асимптотической нормальности статистик вида (1.23), когда объемы совокупности N и выборок n_1, \dots, n_s неограниченно возрастают.

Приведем несколько примеров. В случае $s = 1$ (см. [9] и библиографию к ней) часто рассматриваются статистики вида $\sum a_m(N)$, где суммирование ведется по номерам элементов, вошедших в выборку. Это частный случай S -статистики, когда $f_m^{(N)}(0) = 0$, а $f_m^{(N)}(1) = a_m(N)$.

Пусть все функции $f_{jm}^{(N)}(x_1, \dots, x_s)$, $m = 1, \dots, N$, одинаковы и равны $f_j^{(N)}(x_1 + \dots + x_s)$. Тогда

$$S_{Nj}(\bar{\mathbf{h}}) = \sum_{r=0}^s f_j^{(N)}(r) \mu_r, \quad (1.24)$$

где μ_r есть число элементов совокупности, каждый из которых вошел ровно в r некоторых выборок (о таком элементе говорят, что он «задет» ровно r раз). По аналогии со схемой с возвращением статистики вида (1.24) называются симметрическими. Сл. в. μ_r , также являющиеся, таким образом, симметрическими S -статистиками, изучаются обычно в задачах размещения час-

тиц по ячейкам, причем этот случай соответствует схеме размещения частиц комплектами (см. § 3 данного обзора).

Если рассматривать накопленную сумму

$$\eta_N(l) = \sum_{m=1}^l f_m^{(N)}(h_{1m}, \dots, h_{sm})$$

как сл. функцию параметра l , $1 \leq l \leq N$, то исследование конечномерных распределений этого процесса сводится к изучению многомерной S -статистики. При $s=1$ и $f_m^{(N)}(x) \equiv x$ этот процесс изучался в работе Ю. К. Беляева и Л. В. Рыковой [11], а в общем случае сходимость этого процесса к некоторому гауссовскому процессу получена в [60].

§ 2. Статистические выводы, основанные на делимых статистиках

2.1. Постановка задачи. Изложенные в § 1 общие предельные теоремы для р. с. позволяют рассчитывать параметры статистических критериев, основанных на применении р. с. и служащих для проверки различных гипотез о вероятностях p_1, \dots, p_N исходов в полиномиальной схеме в случае растущего числа исходов N .

Пусть требуется проверить простую гипотезу $H_0 = H(\mathbf{q})$, задаваемую вероятностями $\mathbf{q} = (q_1, \dots, q_N)$, внутри сложной гипотезы $H = \{H(\mathbf{p})\}$.

Обычно наибольший интерес представляет проблема различения сближающихся по той или иной мере при $N \rightarrow \infty$ гипотез, поэтому далее будем предполагать, что гипотеза H задается в следующем виде:

$$H = \{H(\mathbf{p}) : p_m = q_m(1 + \varepsilon_m), \quad m = 1, \dots, N, \quad \max_m |\varepsilon_m| = o(1)\}. \quad (2.1)$$

Для удобства дальнейшего изложения будем считать, что вероятности p_m задаются с помощью некоторой непрерывной на $[0, 1]$ функции плотности $p(y) > 0$ следующим образом:

$$p_m = \int_{(m-1)/N}^{m/N} p(y) dy, \quad m = 1, \dots, N,$$

а соответствующую функцию плотности при гипотезе H_0 обозначать $q(y)$. Наконец, будем считать, что используемая для построения критерия проверки гипотезы H_0 статистика имеет вид

$$L_N = L_N(f) = \sum_{m=1}^N f_N\left(h_m, \frac{m}{N}\right), \quad (2.2)$$

где при каждом $x=0, 1, 2, \dots$ и почти всюду по $y \in [0, 1]$ существует

$$\lim_{N \rightarrow \infty} f_N(x, y) = f(x, y). \quad (2.3)$$

В этом случае о критерии, порождаемом статистикой (2.2), будем говорить для краткости как о критерии f .

При любой гипотезе $H(p)$ из класса (2.1) тип предельного закона р. с. (2.2), как правило, один и тот же, а зависимость от гипотезы проявляется лишь через параметры нормировки в соответствующей предельной теореме для $L_N(\hat{f})$. Это позволяет свести задачу сравнения асимптотической мощности различных критериев к исследованию следующего функционала

$$A_N = [E(L_N | p) - E(L_N | q)] / \sqrt{D(L_N | q)}. \quad (2.4)$$

В этой задаче следует различать два принципиально разных случая: симметрический, т. е. когда, во-первых, гипотеза H_0 соответствует равновероятной схеме: $q(y) \equiv 1$ и, во-вторых, статистика L_N симметрическая: $f_N(x, y) \equiv f_N(x)$, и не симметрический, т. е. когда, по крайней мере, одно из этих двух условий не выполняется.

2.2. Симметрический случай. Рассмотрим сначала симметрический случай и будем обозначать класс всех симметрических критериев (т. е. критериев, порождаемых р. с.) через S . Предположим, что $n, N \rightarrow \infty$, $n/N \rightarrow \alpha$, $0 < \alpha < \infty$. Тогда любой критерий $f \in S$ позволяет отличать от H_0 лишь такие «близкие» альтернативы $H(p)$, для которых

$$p(y) = 1 + \frac{g(y)}{n^{1/4}} (1 + o(1)), \quad (2.5)$$

где $g(y)$ — произвольная функция из $L_2(0, 1)$, удовлетворяющая условиям

$$\int_0^1 g(y) dy = 0, \quad g = \int_0^1 g^2(y) dy > 0 \quad (2.6)$$

(о такой альтернативе будем говорить кратко как об альтернативе g). Справедливы следующие утверждения:

1) При произвольной альтернативе g (и фиксированном уровне значимости γ) асимптотическая мощность критерия f является монотонной функцией от

$$\rho_\alpha(f) = \frac{\alpha}{\sqrt{2}} |E\Delta^2 f(X)| \left(Df(X) - \frac{1}{\alpha} \text{cov}^2(f(X), X) \right)^{-1/2}, \quad (2.7)$$

$$X \sim \Pi(\alpha),$$

и равна $\Phi\left(g \sqrt{\frac{\alpha}{2}} \rho_\alpha(f) - u_\gamma\right)$, где $\Phi(t)$ — функция стандартного нормального распределения, $\Phi(-u_\gamma) = \gamma$ и Δ — оператор разности.

2) Значения функционала $\rho_\alpha(f)$ заключены между 0 и 1, и при этом

$$\rho_\alpha(f) \equiv 1 \Leftrightarrow f(x) = x^2$$

(с точностью до преобразования $af(x) + bx + c$, $a \neq 0$); тем самым критерий χ^2 , порождаемый статистикой

$$X_N^2 = \sum_{m=1}^N h_m^2, \quad (2.8)$$

является асимптотически равномерно наиболее мощным в классе S .

3) Справедливо соотношение

$$\rho_\alpha(f) = \lim_{N \rightarrow \infty} |\text{corr}_{H_0}(L_N(f), X_N^2)|.$$

Ввиду утверждений 1 и 2 функционал $\rho_\alpha(f)$ задает меру (асимптотической) эффективности критериев из класса S , а свойство 3 раскрывает наглядный смысл этой меры как предельного значения модуля коэффициента корреляции (при нулевой гипотезе) статистики критерия и оптимальной статистики X_N^2 .

Для ряда наиболее интересных и известных критериев f мера $\rho_\alpha(f)$ вычислена в явном виде и проведено ее аналитическое и численное исследование как функции параметра α .

В качестве примеров приведем следующие результаты.

1) Функция $f(x) = x \ln x$ задает критерий отношения правдоподобия. Для этого случая функция $\rho_\alpha(f)$ выпукла (по α) вниз, $\rho_\alpha(f)$ монотонно возрастает при $\alpha \downarrow 0$ и $\alpha \rightarrow \infty$, и $\min_{\alpha} \rho_\alpha(f) = \rho_{\alpha^*}(f) = 0,9335\dots$, где $\alpha^* = 2,375\dots$

2) Функция $f(x) = 1$ при $x = 0$ и $f(x) = 0$ при $x \neq 0$ порождает критерий пустых ящиков, и для него

$$\rho_\alpha(f) = l(\alpha) = \frac{\alpha}{\sqrt{2}} (e^\alpha - 1 - \alpha)^{-1/2}.$$

3) Для показательной статистики ($f(x) = (1-a)^x$, $a \neq 0$)

$$\rho_\alpha(f) = l(\alpha a^2).$$

4) Эффективность $\rho_\alpha(f)$ степенной статистики степени $k \geq 2$ ($f(x) = x^k$) убывает с ростом k при любом α . Наконец, приведем вид критической области для произвольного симметрического критерия f (при уровне значимости γ):

$$\left\{ L_N(f) > NEf(X) + u_\gamma \sqrt{N \left[Df(X) - \frac{1}{\alpha} \text{cov}^2(f(X), X) \right]} \right\} \quad (2.9)$$

и для критерия χ^2 :

$$\{X_N^2 > N\alpha(1+\alpha) + u_\gamma \alpha \sqrt{2N}\}. \quad (2.10)$$

Перечисленные результаты получены в работах Г. И. Ивченко и Ю. И. Медведева [62, 63], А. Ф. Глибоченко, Г. И. Ивченко и Ю. И. Медведева [18] и Холста [148].

В работе Г. И. Ивченко и С. В. Цуканова [69] исследованы критерии, основанные на обобщенных р. с. вида (1.21). Обозначим класс всех таких критериев S_s . Показано, что в классе S_s существуют критерии, асимптотическая мощность которых является монотонно возрастающей функцией параметра s . В частности, таким свойством обладает обобщенный критерий χ^2 , основанный на статистике

$$X_{N,s}^2 = \sum_{m=1}^N (h_m + h_{m+1} + \dots + h_{m+s})^2.$$

Таким образом, данная методика приводит к более эффективному критерию, чем обычный критерий χ^2 , основанный на статистике $X_N^2 = X_{N,0}^2$. Но оказывается, что обобщенный критерий χ^2 при $s \geq 1$ уже не является асимптотически оптимальным в классе S_s (в отличие от случая $s=0$), а таким свойством обладает критерий, основанный на статистике

$$Y_{N,s}^2 = X_{N,s}^2 - X_{N,s-1}^2 \quad (X_{N,-1}^2 = 0).$$

Критическая область этого критерия имеет вид

$$\{Y_{N,s}^2 > \alpha(1 + (2s+1)\alpha)N + u_{\gamma\alpha} \sqrt{2N(2s+1)}\},$$

а его мощность при альтернативе g равна в пределе

$$\Phi\left(g \sqrt{\frac{2s+1}{2}} \alpha - u_{\gamma}\right).$$

В. В. Лёвиным [78] предложен и рассчитан многомерный вариант симметричного критерия, основанный на векторной статистике $L_N = (L_{N1}, \dots, L_{Nk})$, где

$$L_{Nj} = \sum_{m=1}^N f(h_{1m} + \dots + h_{jm}), \quad j=1, \dots, k,$$

и h_{jm} — число реализаций m -го исхода в j -й серии из $n_j = \sum_{m=1}^N h_{jm}$ испытаний. Критерий задается критической областью вида $\{L_{N1} > c_1, \dots, L_{Nk} > c_k\}$. Таким образом, в данном случае проверка гипотезы H_0 проводится последовательно, в несколько (не более чем в k) этапов, причем принять гипотезу можно на любом этапе (H_0 принимается на j -ом этапе по первым $n_1 + \dots + n_j$ испытаниям, если осуществляется событие $\{L_{N1} > c_1, \dots, L_{N,j-1} > c_{j-1}, L_{Nj} \leq c_j\}$), а отвергнуть — лишь на последнем, k -ом этапе. Идея построения подобных критериев последовательного типа предложена в работе В. К. Захарова, О. В. Сарманова и Б. А. Севастьянова [22], где был построен и расчи-

тан последовательный критерий χ^2 (при фиксированном числе классов N). В [78] вычисляется предельное значение мощности критерия при произвольной альтернативе g в условиях: $N, n_1, \dots, n_k \rightarrow \infty, n_i/N \rightarrow \alpha_i, 0 < \alpha_i < \infty, i = 1, \dots, k$.

В работе В. В. Лёвина [79] рассматривается задача проверки гипотезы однородности (и равновероятности) для s полиномиальных выборок. Построен критерий, основанный на с. р. с.

$L_N = \sum_{m=1}^N f(h_{1m}, \dots, h_{sm})$, где h_{jm} — число реализаций m -го исхода в j -й выборке, и вычислено предельное значение его мощности на множестве «близких» альтернатив вида (2.5), когда $N \rightarrow \infty$, а объемы выборок сравнимы с N .

2.3. Несимметрический случай. В несимметрическом случае роль «близких» гипотез играют альтернативы вида

$$p(y) = q(y) \left(1 + \frac{b(y)}{\sqrt{n}} (1 + o(1)) \right). \quad (2.11)$$

Таким образом, в этом случае можно строить критерии, отличающиеся от H_0 гипотезы, приближающиеся к ней со скоростью $n^{-1/2}$. Обозначим через \mathcal{K} класс критериев, функции мощности которых при фиксированном уровне значимости и $n, N \rightarrow \infty, n/N \rightarrow \alpha, 0 < \alpha < \infty$, не вырождаются (не стремятся к уровню значимости) на некотором непустом подмножестве близких альтернатив вида (2.11). Об альтернативе (2.11) далее будем говорить кратко как об альтернативе b , а критерии из класса \mathcal{K} будем называть высокочувствительными.

Класс \mathcal{K} не пуст и, в частности, включает в себя все локально наиболее мощные критерии [62]. В то же время при $q(y) \equiv 1$ в этот класс не входят симметрические критерии, в частности, критерий χ^2 , отношения правдоподобия, пустых ящиков, и потому к их практическому использованию надо подходить более критично. В работе Г. И. Ивченко и Ю. И. Медведева [63] предлагается некоторая классификация высокочувствительных критериев, описываются их общие свойства и формулируется ряд положений в пользу критериев из подкласса $\mathcal{L} \subset \mathcal{K}$ линейных критериев (т. е. когда функция $f_N(x, y)$ в (2.2) линейна по x). Эти результаты кратко могут быть сформулированы следующим образом.

Введем на множестве альтернатив $\{b\}$ функционал

$$B_\alpha(b, f) = \frac{1}{\alpha} \int_0^1 b(y) \operatorname{cov}(f(X_y, y), X_y) dy, \quad X_y \sim \Pi(\alpha q(y)),$$

и пусть $H_\alpha^1(f) = \{b : B_\alpha(b, f) \neq 0\}$.

1) $f \in \mathcal{K} \Leftrightarrow H_\alpha^1(f) \neq \emptyset$. При этом альтернативы из $H_\alpha^0(f) = \overline{H_\alpha^1(f)}$ критерием f асимптотически от H_0 не отличимы.

Например, для критерия χ^2 функция $f(x, y) = x^2/q(y)$, $B_\alpha(b, f) = \int_0^1 b(y) dy$ и если $q(y) \equiv 1$, то $H_\alpha^1(f) = \emptyset$.

2) Класс \mathcal{K} можно представить в виде $\mathcal{K} = \mathcal{K}_0 \cup \mathcal{K}_1$, где $\mathcal{K}_0 = \{f: Df(1, \xi) = 0\}$, $\mathcal{K}_1 = \overline{\mathcal{K}_0}$ (здесь и далее сл. в. ξ имеет плотность $q(y)$). Если $f \in \mathcal{K}_0$, то $\lim_{\alpha \rightarrow 0} B_\alpha(b, f) = 0$ для любой альтернативы b , т. е. с уменьшением α чувствительность любого критерия из подкласса \mathcal{K}_0 ухудшается. В частности, все симметрические критерии (при $q(y) \equiv 1$) входят в \mathcal{K}_0 .

3) Пусть $f(x, y) = xl(y)$ (критерий f — линейный) и обозначим $\mathcal{L} = \{l: Dl(\cdot, \cdot) \neq 0\}$ подкласс «невырожденных» линейных критериев. Для произвольного критерия $l \in \mathcal{L}$ функционал $B_\alpha(b, l)$ не зависит от α и имеет представление

$$B(b, l) = Eb(\cdot, \cdot)l(\cdot, \cdot) = \text{cov}(b(\cdot, \cdot), l(\cdot, \cdot)).$$

От параметра α не зависят также множество $H_\alpha^1(f) \equiv H^1(l)$ и предельное значение функции мощности $W_n(b, l)$. В частности, для односторонних альтернатив: $Eb(\xi)l(\xi) > 0$

$$\lim_{n \rightarrow \infty} W_n(b, l) = \Phi(\text{korr}(b(\cdot, \cdot), l(\cdot, \cdot)) \sqrt{Db(\hat{g}) - u_\gamma}),$$

где $\Phi(-u_\gamma) = \gamma$ — заданный уровень значимости. Отметим также, что альтернатива $b = l \in H^1(l)$ и при ней критерий l при любом α имеет асимптотически наибольшую мощность, равную $\Phi(\sqrt{Db(\hat{g})} - u_\gamma)$, в классе всех критериев \mathcal{K} (это по существу критерий Неймана — Пирсона для альтернативы l).

4) Каждому критерию $f \in \mathcal{K}_1$ можно поставить в соответствие линейный критерий $l \in \mathcal{L}$ по правилу: $l(y) = f(1, y)$. Такой критерий l называется линейным аналогом для f и обозначается l_f . Тогда мощность критерия l_f равна предельной при $\alpha \rightarrow 0$ мощности критерия f . При этом в ряде случаев (например, для некоторого подкласса квадратичных критериев, включающего критерий χ^2) линейный аналог l_f оказывается равномерно мощнее критерия f при любом $\alpha > 0$.

5) На класс \mathcal{K} можно вести некоторый функционал $I_\alpha(f)$ как меру асимптотической эффективности критерия f :

$$I_\alpha^2(f) = \int_0^1 \text{korr}^2(g(X_y, y), X_y) Dg(X_y, y) dy \Bigg/ \int_0^1 Dg(X_y, y) dy,$$

где $g(x, y) = f(x, y) - \delta x$, $\delta = \frac{1}{\alpha} \int_0^1 \text{cov}(f(X_y, y), X_y) dy$. Значения $I_\alpha(f)$ заключены между 0 и 1 и

$$I_\alpha(f) \equiv 1 \Leftrightarrow f \in \mathcal{L}.$$

Детальному исследованию поведения мощности критерия χ^2 при $N \rightarrow \infty$ посвящены работы А. А. Боровкова [13], Л. Г. Гванцеладзе и Д. М. Чибисова [139] и С. Х. Гуманяна [111], а критерия пустых ящиков и его обобщениям — более ранние работы Б. А. Севастьянова, В. П. Чистякова, И. И. Викторовой [103, 17] и др.

§ 3. Схема размещения частиц комплектами

В последние годы появилось большое число работ, посвященных изучению асимптотического поведения различных сл. в. в схеме размещения частиц комплектами.

Пусть n комплектов частиц объема s каждый (под объемом комплекта понимается число содержащихся в нем частиц) независимо размещаются по N ячейкам так, что частицы каждого комплекта размещаются в ячейках по одной и все C_N^s возможных размещений считаются равновероятными. Общее число размещенных частиц обозначим $n' = ns$. Таковую схему будем называть равновероятной схемой размещения частиц комплектами.

Наибольшее количество результатов относится к исследованию асимптотического поведения (при $N \rightarrow \infty$) следующих сл. в.: $\mu_r = \mu_r(n, N, s)$ ($\xi_r = \xi_r(n, N, s)$) — число ячеек, в которых после размещения n комплектов оказалось по r частиц (не более чем по r частиц), $\nu_m = \nu_m(N, k, s)$ — число комплектов, после размещения которых впервые какие-либо k ячеек будут содержать не менее m частиц каждая.

Изучение предельных распределений в схеме размещения частиц комплектами имеет свою специфику. Традиционный для доказательства многих предельных теорем в классической схеме размещения частиц по ячейкам метод производящих функций в данной ситуации мало пригоден из-за отсутствия удобных для аналитических исследований представлений соответствующих производящих функций. Эффективными в данном случае оказались другие методы асимптотического анализа.

Первые существенные результаты, положившие начало интенсивному развитию этого направления, были получены в работах Б. А. Севастьянова [101] и Г. И. Ивченко и Ю. И. Медведева [61]. В работе [101] предложен метод реализации сл. в. $\mu_r(n, N, s)$ и $\mu_r(ns, N, 1)$ на одном и том же вероятностном пространстве. Это позволило в случае $N, n \rightarrow \infty$ и при некоторых ограничениях на рост параметра s аппроксимировать распределение сл. в. $\mu_r(n, N, s)$ распределением хорошо изученной сл. в. $\mu_r(ns, N, 1)$ в равновероятной схеме при $n, N \rightarrow \infty$ в следующих зонах изменения параметров [75]: 1) $\alpha = \frac{ns}{N} \rightarrow 0, s = o\left(\frac{1}{\alpha}\right)$; 2) $0 < \alpha_0 \leq \alpha \leq \alpha_1 < \infty, s = o(N^{1/4})$; 3) $\alpha \rightarrow \infty, s = \text{const}$. При

этих ограничениях предельные распределения сл. в. $\mu_r(n, N, 1)$ и $\mu_r(ns, N, 1)$ совпадают и все результаты, полученные для классической схемы, переносятся на равновероятную схему размещения комплектов частиц. В работе [61] получены предельные распределения и асимптотические формулы для математического ожидания и дисперсии сл. в. $\nu_1(N, k, s)$ при $N \rightarrow \infty$, $s = \text{const}$ для всего спектра изменения параметра k , $1 \leq k \leq N$.

Дальнейшие исследования схемы размещения частиц комплектами велись в двух направлениях: продолжалось изучение предельных распределений в равновероятной схеме для других областей изменения параметров и одновременно проводились исследования различных ее обобщений.

3.1. Равновероятная схема размещения частиц комплектами. Остановимся подробнее на исследовании равновероятной схемы размещения комплектов частиц.

В работе В. А. Иванова и С. Ю. Теребулина [42] получены предельные распределения сл. в. $\nu_m(N, k, s)$ при $N \rightarrow \infty$, $m = \text{const} \geq 2$, $s = \text{const} \geq 2$ для следующих областей изменения параметра k : 1. $1 \leq k \leq C$, 2. $k \rightarrow \infty$, $\lambda = \frac{k}{N} \rightarrow 0$, 3. $0 < \lambda_1 \leq \lambda \leq \lambda_2 < 1$, 4. $N - k \rightarrow \infty$, $\lambda = \frac{k}{N} \rightarrow 1$, 5. $N - k \leq C$, где λ_1, λ_2, C — некоторые константы. Исследование предельных распределений сл. в. $\mu_r(n, N, s)$ проведено в работах В. Г. Михайлова [84, 86, 87]. В [84] показано, что если параметры схемы n, N, s изменяются так, что $N, n \rightarrow \infty$ и при некотором целом $r \geq 0$

$$E\mu_r = C_n r N \left(\frac{s}{N}\right)^r \left(1 - \frac{s}{N}\right)^{n-r} \rightarrow \lambda > 0,$$

то распределение μ_r сходится к распределению Пуассона $\Pi(\lambda)$. Отмечается также, что в случае $n, N \rightarrow \infty$, $\alpha \rightarrow 0$, $\frac{1}{2} N\alpha^2 \rightarrow \lambda > 0$ распределения величин $\mu_0 - N + sn$ и $\frac{1}{2}(sn - \mu_1)$ сходятся к распределению $\Pi(\lambda)$.

В [87] методом моментов доказана асимптотическая нормальность величин $\mu_r(n, N, s)$ и их линейных комбинаций. Сл. в. μ_r асимптотически нормальна, если s, n, N меняются так, что $n, N \rightarrow \infty$ и 1) $\alpha = \frac{sn}{N} \rightarrow 0$, $r \geq 2$, $N\alpha^r \rightarrow \infty$ или 2) $0 < \alpha_0 \leq \alpha \leq \alpha_1 < \infty$. В последнем случае доказана асимптотическая нормальность вектора (μ_0, \dots, μ_T) при любом T . Величины μ_0 и μ_1 асимптотически нормальны при $n, N \rightarrow \infty$, $\alpha \rightarrow 0$, $N\alpha^2 \rightarrow \infty$.

В [86] методом моментов доказана асимптотическая нормальность μ_0 в случае, когда параметры s, n, N меняются так, что $n \geq 2$,

$$E\mu_0 = N(1-p)^n \rightarrow \infty, \quad \alpha^k N \rightarrow \infty \quad \forall k < \infty.$$

Харт, Мачек, Степан и Ворличкова [155] исследовали сл. в. $\mu_n = \mu_n(n, N, s)$ и $\tau = \min\{n : \mu_n = 0\}$. Получены точные распределения μ_n и τ . Для μ_n доказаны предельная пуассоновская (при $N \rightarrow \infty$, $\frac{s}{N} = p \in (0, 1)$ — фиксировано, $n \rightarrow \infty$ так, что $Np^n \rightarrow \lambda$, $0 < \lambda < \infty$) и нормальная (при $N \rightarrow \infty$, $p \in (0, 1)$) теоремы.

В работе Агравала [122] доказывается формула

$$E\left(\frac{1}{N - \mu_0}\right) = \sum_{r=0}^{l_n} \frac{1}{N-r} \left[\prod_{i=1}^n C_{N-r}^{s_i} (C_N^{s_i})^{-1} \right], \quad l_n = \min_{1 \leq i \leq n} (N - s_i),$$

в схеме размещения n комплектов объемов s_1, s_2, \dots, s_n , соответственно.

В работе Саперстейна [168] найдена вероятность того, что никакие m последовательных ячеек не содержат k или более частиц после размещения одного комплекта объема s .

В ряде работ рассматривалась схема размещения конечного числа ($n = \text{const}$) комплектов растущих объемов ($s \rightarrow \infty$). Г. И. Ивченко и В. В. Лёвин [60] показали, что при $s, N, N-s \rightarrow \infty$, $s^n/N^{n-1} \rightarrow \infty$, $(N-s)^n/N^{n-1} \rightarrow \infty$ вектор (μ_0, \dots, μ_n) асимптотически нормален. В [60] найдены также предельные распределения сл. в. $v_m(N, k, s)$ при $N, k, s \rightarrow \infty$, $0 < C_1 \leq p = \frac{s}{N} \leq C_2 < \infty$, $0 < C_3 \leq \lambda = \frac{k}{N} \leq C_4 < 1$, приведены условия, при которых предельное распределение является одноточечным или двухточечным.

В. А. Иванов и С. Ю. Теребулин [43], используя асимптотическую независимость частот заполнения ячеек $h_{j_1}(n), \dots, h_{j_l}(n)$ для любых $1 \leq j_1 < \dots < j_l \leq N$ при $N, s, N-s \rightarrow \infty$, $l = o(\sqrt{p(1-p)N})$, указали условия, при которых распределения сл. в. μ_r , $0 < r < n$, $\mu_0 - N + sn$, $\mu_n - N + (N-s)n$ сходятся к распределению Пуассона. Полученные результаты позволили с учетом соотношения

$$P(v_m(N, k, s) \leq n) = P(\sum_{m=1}^n \mu_m \leq N - k)$$

доказать ряд теорем об асимптотическом поведении сл. в. $v_m(N, k, s)$ при $N, s, N-s \rightarrow \infty$ в случаях $p \rightarrow 0$ и $p \rightarrow 1$. С. Ю. Теребулин [106] изучал асимптотическое поведение различных членов вариационного ряда для частот $h_1(n), \dots, h_N(n)$ при $n = \text{const}$, $N, s, N-s \rightarrow \infty$; предельное поведение максимального члена вариационного ряда $h_{(N)} = \max_{1 \leq j \leq N} h_j$ исследовал Е. Р. Хакимуллин. Результаты для случая $n, N, \alpha/\ln N \rightarrow \infty$, $p = \frac{s}{N} \rightarrow 0$ изложены в [115].

3.2. Обобщения равновероятной схемы. Предположим, что в ячейки, занумерованные натуральными числами $1, 2, \dots, N, \dots$, размещаются n комплектов частиц так, что j -ый комплект

представляет собой случайное подмножество множества натуральных чисел, выбираемое в соответствии с вероятностями $\{p_{i_1, i_2, \dots}^j\}$, $1 \leq i_1 < i_2 < \dots$, $\sum_{i_1 < i_2 < \dots} p_{i_1, i_2, \dots}^j = 1$, $j = 1, \dots, n$ (вероятность выбора подмножества i_1, i_2, \dots равна $p_{i_1, i_2, \dots}^j$). Многие авторы изучали различные частные случаи этой общей схемы. Г. И. Ивченко [51] исследовал схему равновероятного размещения комплектов частиц случайных объемов. Пусть ξ_1, ξ_2, \dots — последовательность независимых сл. в., принимающих значения $1, 2, \dots, N$ с вероятностями $p_s = P(\xi_j = s)$, $\sum_{s=1}^N p_s = 1$. Предполагается, что j -ый комплект имеет объем ξ_j и равновероятно размещается по N ячейкам. В [51] найдены предельные распределения сл. в. μ_0 и ν_1 при $N \rightarrow \infty$ в ситуации, когда занятыми оказались почти все ячейки. Для первой сл. в. это условие конечности $E\mu_0$, для второй — условие конечности параметра $N - k$.

Более общую схему размещения комплектов частиц по N ячейкам, в которой комплекты произвольных объемов k ($k = 1, \dots, N$) могут размещаться в ячейки с номерами i_1, \dots, i_k с вероятностями $p_{i_1, \dots, i_k}^j = p_{i_1, \dots, i_k}$, исследовал В. Г. Михайлов [84].

С помощью метода моментов в форме Б. А. Севастьянова [104] в [84] найдены условия, при которых распределение сл. в. μ_r сходится к распределению Пуассона. В работе [85] В. Г. Михайлов получил ряд неравенств для расстояния по вариации между распределениями Пуассона и сл. в. μ_r в схеме размещения частиц комплектами, являющимися случайными подмножествами множества натуральных чисел. Единственными характеристиками распределений комплектов, используемыми в формулировках теорем и при их доказательстве в [85], являются $a_m(j)$ — вероятность того, что частица j -го комплекта попадет в ячейку с номером m и $a_{lm}(j)$ — вероятность того, что частицы j -го комплекта попадут в ячейки с номерами l и m .

Изучению схемы неравновероятного размещения частиц комплектами заданного объема посвящена работа В. А. Иванова и С. Ю. Теребулина [42]. Предполагается, что распределение вероятностей на множестве всех комплектов объема s может быть задано в виде

$$p_{i_1, \dots, i_s} = \frac{a_{i_1} \dots a_{i_s}}{\sum_{1 < j_1 < \dots < j_s < N} a_{j_1} \dots a_{j_s}},$$

где a_1, \dots, a_N — произвольные положительные числа. Обозначим $p_j = a_j / \sum_{i=1}^N a_i$, $j = 1, \dots, N$. В [42] с помощью метода перевала

исследовано асимптотическое поведение сл. в. $\xi_m(n, N, 1)$ и $v_m(k, N, 1)$ в неравновероятной схеме независимого размещения частиц по ячейкам (вероятность попадания частицы в j -ую ячейку равна p_j , $0 < C_1 \leq N p_j \leq C_2 < \infty$, $j = 1, \dots, N$, $\sum_{j=1}^N p_j = 1$) при $N \rightarrow \infty$ в промежуточных областях изменения параметра k : $k \rightarrow \infty$, $\frac{k}{N} \rightarrow 0$ (в случае $m = 1$ при $\frac{k^2}{N} \rightarrow \infty$) и $N - k \rightarrow \infty$, $\frac{k}{N} \rightarrow 1$ и получены предельные распределения сл. в. $v_m(k, N, s)$ при $N \rightarrow \infty$, $m = \text{const}$, $s = \text{const}$ во всех зонах изменения параметра k , $1 \leq k \leq N$. При доказательстве теорем использовался метод одного вероятностного пространства, предложенный Б. А. Севастьяновым [101].

А. Е. Жуков и В. П. Чистяков [20] исследовали схему размещения частиц по N ячейкам комплектами, в которой объемы ξ_1, \dots, ξ_n комплектов являются сл. в., а при фиксированных значениях $\xi_j = s_j$, $j = 1, \dots, n$, множества ячеек, образующие комплекты, независимы, и распределение j -го комплекта совпадает с распределением множества непустых ячеек в схеме полиномиального размещения частиц по N ячейкам в момент времени, когда число непустых ячеек становится равным s_j . Авторы, используя аппроксимацию распределений, возникающих при размещении частиц комплектами, хорошо изученными распределениями, соответствующими размещению частиц по полиномиальной схеме, получили предельные теоремы для распределений сл. в. $\mu_r(n, N)$.

В работе А. Е. Жукова [19] рассмотрена схема независимого размещения n комплектов частиц объема $Ss \leq N$ каждый в SN ячеек, расположенных в виде матрицы размера $S \times N$. При этом предполагается, что на множестве всевозможных размещений Ss частиц в SN ячеек со следующими ограничениями: 1) в каждую строку матрицы попадает ровно s частиц; 2) в каждый столбец матрицы попадает не более одной частицы, задана равномерная мера.

Доказано, что если $r = \text{const}$, $N, S \rightarrow \infty$, $\frac{n}{N^2} \rightarrow 0$, $C_n r \frac{NS}{Nr} \times \times \left(1 - \frac{s}{N}\right)^n \rightarrow \lambda$, $0 < \lambda < \infty$, то $\mu_r(n, N)$ имеет в пределе распределение $\Pi(\lambda s^r)$.

В. А. Ватутин и В. Г. Михайлов [14] исследовали схему независимого и равновероятного размещения n комплектов частиц объемов s_1, \dots, s_n . Получены интегральная и локальная центральные предельные теоремы, а также все возможные «пуассоновские» предельные теоремы с оценками скорости сходимости для сл. в. $\mu_0 = \mu_0(N, s_1, \dots, s_n)$ — числа ячеек, оставшихся пустыми после размещения всех n комплектов. Метод

доказательства основан на том факте, что производящая функция сл. в. μ_0 имеет лишь отрицательные действительные корни, что позволяет представить распределение сл. в. μ_0 в виде композиции некоторых двухточечных распределений. В работе Парка [162] этот метод применен для доказательства асимптотической нормальности сл. в. $\mu_0(n, N, s)$ в схеме размещения частиц комплектами объема s при $N, n \rightarrow \infty, \frac{n}{N} \rightarrow \alpha > 0$. Показано, что $\mu_0 = \sum_{j=1}^{N-s} \eta_j$, где η_j — независимые бернуллиевские сл. в. с параметрами γ_j , где γ_j^{-1} — корни функции $E(1+z)^{\mu_0}$.

Применяемый в [14, 162] метод является весьма эффективным при решении целого ряда комбинаторных задач. Поэтому остановимся подробнее на некоторых результатах, связанных с использованием подобных методов.

3.3. Нули производящих функций и предельные теоремы. Пусть $\{\xi_n\}$ — последовательность целочисленных сл. в. с распределениями

$$P(\xi_n = k) = p_{nk}, \quad k = 0, 1, \dots, m_n, \quad \sum_{k=0}^{m_n} p_{nk} = 1, \quad p_{n0} p_{nm_n} > 0.$$

Обозначим z_{n1}, \dots, z_{nm_n} нули производящей функции (п. ф.) $P_n(z) = E z^{\xi_n}$ и пусть $a_{nk} = \frac{1}{1-z_{nk}}, k = 1, \dots, m_n$. Моменты сл. в. ξ_n легко выражаются через числа a_{nk} , в частности,

$$E \xi_n = \mu_n = \sum_{k=1}^{m_n} a_{nk}, \quad D \xi_n = \sigma_n^2 = \sum_{k=1}^{m_n} a_{nk} (1 - a_{nk}).$$

Если известны нули z_{nk} , то асимптотическое поведение сл. в. ξ_n при $n \rightarrow \infty$ легко описывается в терминах поведения z_{nk} . Случай сходимости распределения ξ_n к закону Пуассона характеризуется при этом условием конечности среднего μ_n , а случай асимптотической нормальности (а. н.) (с параметрами μ_n и σ_n^2) — условиями неограниченного возрастания дисперсии σ_n^2 и величины $\gamma_n = \sigma_n \min_{1 < k < m_n} |1 - z_{nk}|$.

Приведем точные формулировки соответствующих утверждений.

1) Условия $E \xi_n \rightarrow \lambda, 0 < \lambda < \infty, \sum_{k=1}^{m_n} |z_{nk}|^{-2} \rightarrow 0$ достаточны, а при $\sup_{n,k} \operatorname{Re} z_{nk} \leq 0$ и необходимы, чтобы при $n, m_n \rightarrow \infty$ распределение ξ_n сходилось к распределению $\Pi(\lambda)$.

2) Если при $n, m_n \rightarrow \infty$ выполняются условия: $\sigma_n \rightarrow \infty$, $\gamma_n \rightarrow \infty$, $\mu_n / \sigma_n^3 \rightarrow 0$, $\sum_{k=1}^{m_n} |a_{nk}|^3 / \sigma_n^3 \rightarrow 0$, то сл. в. ξ_n асимптотически нормальна; при этом для выполнения указанных условий достаточно, чтобы $n^{-1} \gamma_n^2 \min(\gamma_n, \sigma_n) \rightarrow \infty$.

Эти утверждения доказываются непосредственным анализом поведения п. ф. $P_n(z)$ при $n \rightarrow \infty$.

Отметим также некоторые интересные специальные случаи, при которых имеет место а. н. ξ_n .

а) Пусть нули z_{nh} при всех $n \geq n_0$ лежат в левой полуплоскости $\operatorname{Re} z \leq 0$. Тогда условие $\sigma_n \rightarrow \infty$ необходимо и достаточно для а. н. величины ξ_n .

Доказательство этого факта основано на том, что в этом случае распределение ξ_n является композицией некоторого числа двухточечных и трехточечных распределений с носителями в точках 0, 1 и 0, 1, 2, соответственно.

б) Пусть нули z_{nh} при всех $n \geq n_0$ лежат на единичной окружности $|z|=1$. Тогда для а. н. величины ξ_n достаточно выполнения условия $\gamma_n \rightarrow \infty$.

Случай, когда все нули z_{nh} действительны (и следовательно, лежат в левой полуплоскости) подробно рассматривался Харпером [141], Харрисом и Парком [143], Парком [161], В. А. Ватутиным и В. Г. Михайловым [14], результаты для общего случая получены Ю. И. Медведевым.

3.4. Размещение комплектов на окружности. Рассмотрим еще одну схему размещения комплектов. Расположим все ячейки на окружности так, что за ячейкой с номером j , $j=1, \dots, N-1$, будет следовать ячейка с номером $j+1$, а за ячейкой с номером N следует первая ячейка. В ячейки независимо бросаются n комплектов частиц объема s так, что каждый комплект может с равными вероятностями занять любые s подряд идущих ячеек. М. Н. Савчук [96] изучил предельное поведение сл. в. $\eta_r = \eta_r(n, N)$, равной числу комплектов, имеющих пересечения ровно с r другими комплектами и сл. в. $\xi_r = \sum_{j=0}^r \eta_j$. В [96] получены асимптотические выражения для фак-

ториальных моментов $E\eta_r^{(k)}$ при $n, s, N \rightarrow \infty$, $\alpha = \frac{ns}{N} \rightarrow 0$ и фиксированных $r \geq 1$, $k \geq 1$. Показано, что если выполнено дополнительное условие $E\eta_r \rightarrow \gamma$, $0 < \gamma < \infty$, то распределение η_r сходится к распределению линейной комбинации r независимых пуассоновских случайных величин с различными параметрами.

3.5. Процессы в схеме размещения частиц комплектами. М. Н. Савчуком [95] рассматривались следующие процессы: $\mu_N^i(t) = \frac{1}{\sqrt{N}} (X_N^i(t) - EX_N^i(t))$, $t \geq 0$, $i=0, \dots, n$,

$X_N^i(t)$ — число ячеек, содержащих ровно i частиц каждая после размещения $n = [t/p]$ комплектов объема s , $p = s/N$, $[x]$ — целая часть числа x ; $\xi_N^i(t) = \frac{1}{\sqrt{N}} (Y_N^i(t) - EY_N^i(t))$, $0 \leq t \leq 1$, $i = 0, \dots, \dots, n$, $Y_N^i(t)$ — число ячеек из первых $[Nt]$, содержащих по i частиц каждая; $\kappa_N^i(t) = \frac{1}{\sqrt{N}} (Z_N^i(t) - EZ_N^i(t))$, $0 \leq t \leq 1$, $i = 1, \dots, n$, $Z_N^i(t)$ — число частиц в i -ом комплекте, попавших в первые $[Nt]$ ячеек, $\kappa_N(t) = \frac{1}{\sqrt{N}} (Z_N(t) - EZ_N(t))$, $0 \leq t \leq 1$, $Z_N(t)$ — число частиц, попавших в первые $[Nt]$ ячеек после размещения n комплектов частиц объема s каждый.

Доказана сходимость этих процессов при соответствующих условиях к некоторым гауссовским диффузионным процессам и исследована структура предельных процессов.

§ 4. Другие задачи о размещении частиц

В книге В. Ф. Колчина, Б. А. Севастьянова и В. П. Чистякова [75] подробно рассмотрены различные схемы размещения частиц по ячейкам: равновероятные и полиномиальные размещения, размещения со случайным числом частиц, размещения частиц комплектами и т. д. Наиболее интенсивно развивающееся в последние годы направление, связанное с исследованием размещений частиц комплектами, выделено в настоящем обзоре в отдельный параграф. Хотя исследование других схем размещения велось не столь интенсивно, однако по ним также получен целый ряд новых результатов. Отметим наиболее интересные из них.

4.1. Полиномиальные размещения. Рассмотрим схему размещения частиц, в которой каждая из n частиц независимо от остальных попадает в j -ю ячейку с вероятностью p_j , $j = 1, \dots, N$, $\sum_{j=1}^N p_j = 1$. Пусть $h_j = h_j(n)$ — число частиц, оказавшихся в j -ой ячейке, и μ_r — число ячеек, содержащих ровно r частиц каждая после размещения n частиц, $h_{(1)} \leq \dots \leq h_{(N)}$ — вариационный ряд (в. р.) частот h_1, \dots, h_N .

Статистика $\zeta_{n,N} = \sum_{r=0}^n c_r \mu_r$, где c_r — некоторые константы,

является симметрической разделимой статистикой, из которой специальным подбором констант c_r можно получить статистики критериев пустых ящиков, χ^2 и отношения правдоподобия. Новые результаты о сходимости распределения сл. в $\zeta_{n,N}$ к законам Пуассона, нормальному и их смесям получены в работе Ю. И. Медведева [82]. Поведение $\zeta_{n,N}$ в равновероятной схеме

размещения при $n/N \rightarrow 0$ изучалось Квином [165]. Инглунд [134] получил оценку скорости сходимости распределения сл. в. $\mu = \mu(n, N) = N - \mu_0$, равной числу заполненных ячеек, к предельному нормальному закону. Им доказаны следующие неравенства (в случае равновероятной схемы):

$$\frac{0,087}{\max(3, D\mu)} \leq \sup_x \left| P(\mu < x) - \Phi\left(\frac{x - E\mu}{\sqrt{D\mu}}\right) \right| \leq \frac{10,4}{D\mu}.$$

Локальные предельные теоремы для μ , в случае, когда $n, N \rightarrow \infty$ и вектор (p_1, \dots, p_N) имеет симметричное распределение Дирихле с параметром β , доказаны в работе Чен Вэн Чена [128]. Там же рассмотрена аналогичная задача для схемы размещения со случайным числом частиц.

Предельные распределения максимальной частоты $h_{(N)}$ (и других характеристик в.р. $h_{(1)}, \dots, h_{(N)}$) при $n, N \rightarrow \infty$ были исследованы в статьях И. И. Викторовой, Б. А. Севастьянова [15, 16], В. Ф. Колчина [71] и Г. И. Ивченко [50]. Е. Р. Хакимуллин [116] исследовал скорость сходимости распределения $h_{(N)}$ в равновероятной полиномиальной схеме к предельному распределению. В работе Й. С. Юсаса [121] выведены неравенства для уровня значимости критерия, основанного на максимальной частоте, и проведено сравнение функций мощности этого критерия и критерия типа χ^2 в случае конечных n и N . Асимптотическое поведение максимальной частоты $h_{(N)}$ для цепи Маркова с растущим числом исходов N изучалось А. С. Амбросиовым [2].

Страусс [173] изучает распределение сл. в. τ , равной числу интервалов, состоящих сплошь из заполненных ячеек (все ячейки располагаются друг за другом на прямой или на окружности) при равновероятном и независимом размещении n частиц по N ячейкам. Он показал, что, например, в случае расположения ячеек на окружности при $N \rightarrow \infty$, $\frac{n}{N} = -\log \lambda = \text{const}$, $0 < \lambda < 1$,

справедливы асимптотические формулы $E\tau \sim N\lambda(1-\lambda)$, $D\tau \sim N(1-\lambda)(1-3\lambda+\lambda^2)$. Аналогичная задача для схемы размещения комплектов частиц рассматривается в [163]. В работе Вейса [175] при некоторых условиях на параметры схемы рассматриваются вопросы нормальной аппроксимации распределения частот исходов полиномиальной схемы с растущим числом исходов.

В работе [167] изучаются предельные распределения числа пустых ячеек и времени ожидания до наступления некоторого события, связанного с процессом заполнения ячеек, для полиномиальных размещений с параметрами $p_1 = \dots = p_{N-1} = \mathcal{P}/(N-1)$, $p_N = 1 - \mathcal{P}$, $0 < \mathcal{P} < 1$.

4.2. Зависимые размещения. Пусть в N ячеек последовательно по одной бросаются частицы. Обозначим $\mu(k)$ число

занятых ячеек после размещения k частиц. Предположим, что $(k+1)$ -я частица может быть размещена в любую фиксированную свободную ячейку с вероятностью $(N + \theta \mu(k))^{-1}$ и в любую занятую ячейку с вероятностью $(1 + \theta)(N + \theta \mu(k))^{-1}$, $\theta > -1$. В работе Б. А. Севастьянова [105] получены предельные распределения числа пустых ячеек после размещения n частиц при 1) $\alpha = \frac{n}{N} \rightarrow 0$, $\frac{n^2}{2N} \rightarrow \lambda$; 2) $\alpha \rightarrow 0$, $N\alpha^2 \rightarrow \infty$; 3) $0 < \alpha_0 \leq \alpha \leq \alpha_1 < \infty$.

В статье Лам Хин—Кама [159] рассматривается схема последовательного размещения n частиц по N ячейкам с фиксированными уровнями t_1, \dots, t_N (t_j — натуральное число, $j = 1, \dots, N$). Предполагается, что вероятность попадания k -ой частицы в j -ю ячейку равна $p_j(\sum' p_m)^{-1}$, где суммирование ведется по тем m , для которых $h_m(k-1) < t_m$, если $h_j(k-1) < t_j$, и равна 0, если при размещении первых $k-1$ частиц в j -ю ячейку уже попало t_j частиц. Найдено распределение сл. в. (h_1, \dots, h_N) и число возможных значений этого вектора.

Херрман [145] рассматривает модель марковского размещения частиц по ячейкам, в которой вероятность попадания частицы в данную ячейку зависит как от количества содержащихся в ней частиц, так и от их состава. Различные комбинаторные модели, связанные с заполнением ячеек, рассмотрены в работе В. Н. Сачкова [97]. Связь задачи о размещении частиц с исследованием случайных отображений рассмотрена в работе В. Ф. Колчина [73].

4.3. Случайное заполнение таблиц. Пусть имеется N ячеек, расположенных в виде таблицы, состоящей из s строк, j -ая трока которой содержит N_j ячеек, $j = 1, \dots, s$. Таблица случайным образом заполняется n частицами, причем каждая частица независимо от других с равной вероятностью попадает в любую из N ячеек. Пусть $\mu_r^{(j)}$ — число ячеек j -ой строки, содержащих r частиц каждая, $\xi_m^{(j)}$ — число ячеек j -ой строки, содержащих не более m частиц каждая после бросания n частиц.

В работе Г. И. Ивченко и Т. В. Ивановой [57] находятся многомерные предельные распределения векторов, составленных из сл. в. $\mu_r^{(j)}$ и $\xi_m^{(j)}$, а также предельные распределения времени ожидания некоторых событий, связанных с процессом заполнения таблиц (см. § 5). Близкие к описанной схема размещения частиц по ячейкам со случайными уровнями и схема с выделенной группой ячеек исследовались в работах Т. В. Ивановой [45, 46] (см. § 5).

4.4. Размещения со счетным числом ячеек. Пусть n частиц независимо размещаются по счетному множеству ячеек $\{1, 2, \dots\}$, причем для каждой частицы вероятность попадания

в k -ю ячейку равна p_k , $\sum_{k=1}^{\infty} p_k = 1$. Без ограничения общности

будем предполагать, что $p_1 \geq p_2 \geq p_3 \geq \dots$. Обозначим $\mu_r = \mu_r(n)$ число ячеек, содержащих ровно r частиц каждая после размещения n частиц. В работе В. Г. Михайлова [88] указываются достаточные условия сходимости (при таком изменении параметров n и $\{p_k\}$, что $E\mu_r \rightarrow \infty$) распределения сл. в. μ_r к нормальному закону и оценивается скорость этой сходимости. Вопрос об оценке точности пуассоновской аппроксимации для сл. в. μ_r в описанной схеме рассматривается в работе А. М. Зубкова и В. Г. Михайлова [26]. Ш. К. Форманов и А. Асимов в [114] исследовали предельное поведение при $n \rightarrow \infty$

сл. в. $\mu_r(n)$ и $\mu(n) = \sum_{r=1}^n \mu_r(n)$ в предположении, что функция

$\alpha(x) = \max \left\{ j : p_j \geq \frac{1}{x} \right\}$, $x \in (0, \infty)$, имеет вид $\alpha(x) = x^\gamma L(x)$, где $0 < \gamma < 1$, $L(x)$ — медленно меняющаяся на бесконечности функция. В работе Чаки и Фелдеша [130] найдено распределение наименьшего среди номеров пустых ячеек в предположении, что $p_k \sim ck^{-\alpha}$, $k \rightarrow \infty$. Статья [136] посвящена изучению предельных распределений оценок и построению доверительных интервалов для сл. в. $\tau = \sum' p_n$, где суммирование ведется по тем k , для которых $h_n(n) > 0$. В работе Руо [166] при некоторых условиях на распределение $\{p_k\}$ находятся вероятностные пределы (при $n \rightarrow \infty$) отношений $\mu_r(n)/\mu(n)$ и $h_{(r)}(n)/n$.

А. М. Зубков и Н. Н. Попов [28] вводят на множество вероятностных распределений $P = \{p = (p_1, p_2, \dots)\}$ отношение частичного порядка: если ζ_1, ζ_2, \dots — независимые сл. в., имеющие распределение p , и $\mu(n, p)$ — число различных элементов в $\{\zeta_1, \dots, \zeta_n\}$, то $p \leq p'$ тогда и только тогда, когда $Ef(\mu(n, p)) \leq Ef(\mu(n, p'))$, $n = 1, 2, \dots$, для любой невозрастающей функции $f(x) \geq 0$, $x \geq 0$.

Показывается, что $p \leq p'$ тогда и только тогда, когда

$$p_1 + \dots + p_k \leq p'_1 + \dots + p'_k, \quad k = 1, 2, \dots$$

В работе В. Г. Михайлова [89] рассматривается схема m -зависимого размещения частиц по счетному множеству ячеек. С использованием метода моментов получены достаточные условия асимптотической нормальности сл. в. $\mu_r(n)$, $\mu(n)$ и η_r (число r -кратных повторений).

4.5. Размещения частиц нескольких типов. В работах Ш. К. Форманова и А. Асимова [3, 113] изучаются схемы размещения частиц нескольких типов. Пусть в N ячеек бросается n_1 частиц первого типа и n_2 частиц второго типа, причем вероятность попадания частицы k -го типа в i -ю ячейку равна $p_i^{(k)}$ ($k = 1, 2$, $i = 1, \dots, N$). Обозначим $\mu_{r_k}^{(k)}$ число ячеек, содержащих

ровно r_k частиц k -го типа, μ_{r_1, r_2} — число ячеек, содержащих ровно r_1 частиц первого типа и r_2 частиц второго типа каждая. В [113] при некоторых условиях на параметры $n_1, n_2, N, \{p_i^k\}$ доказывается локальная предельная теорема для сл. вектора $(\mu_{r_1}^{(1)}, \mu_{r_2}^{(2)}, \mu_{r_1, r_2})$. Пусть теперь $\mu_i(n_i)$ — число ячеек, занятых частицами i -го типа и $\mu(n_1, n_2)$ — общее число занятых ячеек. В [3] найдена характеристическая функция сл. вектора $(\mu_1(n_1), \mu_2(n_2), \mu(n_1, n_2))$ и доказаны интегральная и локальная предельные теоремы для схемы размещения со счетным множеством ячеек.

4.6. Марковские размещения. Пусть задана последовательность независимых сл. в.

$$\xi_1, \xi_2, \dots, \xi_l, \quad (4.1)$$

принимающих значения $1, 2, \dots, N$ с вероятностями p_1, p_2, \dots, p_N ,

$\sum_{i=1}^N p_i = 1$. Назовем l -значной комбинацией упорядоченный набор $\gamma^l = (\gamma_1, \dots, \gamma_l)$ из l натуральных чисел $1, 2, \dots, N$. Общее число различных l -значных комбинаций равно N^l . Обозначим

$$\xi_s^l = (\zeta_s, \dots, \zeta_{s+l-1}), \quad s = 1, 2, \dots, n. \quad (4.2)$$

Тогда по последовательности (4.1) длины $t = n + l - 1$ можно построить марковскую сл. последовательность вида (4.2), состоящую из n l -значных комбинаций. Сл. последовательности (4.2) сопоставим схему размещения n частиц по N^l ячейкам, занумерованным всевозможными наборами $\gamma^l = (\gamma_1, \dots, \gamma_l)$, в которой k -ая частица размещается в ячейку с номером γ^l , если $\xi_k^l = \gamma^l$. Таким образом, частицы размещаются не независимо, как в классической схеме [75], а в соответствии с цепью Маркова (4.2).

По аналогии с классической схемой обозначим $\mu_r(n, l)$ число ячеек, в которых после размещения n частиц оказалось ровно по r частиц. Тогда $\mu_0(n, l)$ — число l -значных комбинаций, не реализовавшихся в последовательности (4.1) длины $t = n + l - 1$.

В случае $l = 1$, сл. в. $\mu_0(n, 1) = \mu_0(n)$ является хорошо изученной величиной, называемой числом пустых ячеек [75]. Общий случай $l = \text{const} \geq 2$, $N \rightarrow \infty$, исследовался в работе В. Ф. Колчина и В. П. Чистякова [77]. Доказана теорема о сходимости при $\alpha = \frac{n}{N} \rightarrow \infty$, $np^{l+1} \rightarrow 0$, $p = \max_{1 \leq j \leq N} p_j$, распределения $\mu_0(n, l)$ к композиции распределения Пуассона и последовательности двухточечных распределений (в частности, к распределению Пуассона или биномиальному распределению). В работе А. М. Зубкова [25] этот результат доказан при менее ограничительных условиях как следствие из общих предельных теорем

для числа состояний, появившихся в отрезке траектории цепи Маркова заданное число раз.

Сходимость распределения $\mu_0(n)$ к закону Пуассона при $N, n \rightarrow \infty, \alpha = \frac{n}{N} \rightarrow \infty$ для цепей Маркова (простых и сложных), приближающихся к последовательности независимых равновероятных испытаний, была доказана ранее П. Ф. Беляевым [6, 7].

В работе А. М. Зубкова и В. Г. Михайлова [27] изучались распределения сл. в. $\mu_r(n, l), r \geq 2$, для последовательностей (4.2) в случае, когда сл. в. (4.1) принимают счетное множество значений, при условиях, обеспечивающих выполнение соотношений

$$E(\mu_r(n, l) + \mu_{r+1}(n, l) + \dots) \rightarrow \infty, n \rightarrow \infty.$$

В [24] А. М. Зубков предложил способ, позволяющий перенести на цепи Маркова, приближающиеся к последовательности независимых испытаний, предельные теоремы для распределений различных функционалов, доказанные для последовательностей независимых испытаний.

Обозначим множество состояний цепи Маркова S и для каждого состояния $k \in S$ зададим подмножество $R(k) \subset \{0, 1, 2, \dots\}$.

Пусть $\mu_R(n) = \sum_n \chi(h_k(n) \in R(k))$, где $\chi(A)$ — индикатор события A , $h_k(n)$ — частота появлений исхода k в реализации цепи Маркова длины n . Заметим, что если $R(k) = \{r\}$, то $\mu_R(n) = \mu_r(n)$.

В работе П. Ф. Беляева [8] доказана предельная теорема о сходимости распределения $\mu_R(n)$ к пуассоновскому при $R(k) = \{m: m < n\pi_k - \lambda_{1,k} \sqrt{n\pi_k} \text{ или } m > n\pi_k + \lambda_{2,k} \sqrt{n\pi_k}\}$ (при соответствующем выборе $\lambda_{1,j}$) для последовательности цепей Маркова с $S = \{1, 2, \dots, N\}$ при условиях $n^{-1}N \ln^3 N \rightarrow 0, n \rightarrow \infty, \sup Np \leq c < \infty, \sup \min \{s: N \min p_{ij}^{(s)} > \varepsilon\} < \infty$ для некоторого $\varepsilon > 0$, где $p_{ij}^{(s)}$ — вероятность перехода из состояния i в состояние j за s шагов, $\{\pi_k\}$ — стационарное распределение цепи, $p = \sup_{i,j \in S} p_{ij}^{(1)}$. В работе А. М. Зубкова [25] этот результат доказан при более слабых ограничениях.

Асимптотические оценки необходимого объема выборки для различения двух сближающихся гипотез в схеме марковского размещения с двумя ячейками получены в работе В. А. Иванова [30].

Перейдем теперь к рассмотрению задач, в которых изучается заполнение некоторой выделенной группы ячеек. Комбинацию длины l вида $\gamma^l(j) = \underbrace{(j, \dots, j)}_{l \text{ раз}}$ назовем l -членной j -серией или просто (l, j) -серией, комбинацию вида $\gamma^{l+2}(i, j, k) = (i, j, \dots$

..., j, k), где $i \neq j, k \neq j$, назовем изолированной l -членной j -серией (изолированной (l, j) -серией). К изолированным сериям (в литературе их часто называют просто сериями) обычно причисляют также начальную $\gamma^{l+1}(j, k) = (j, \dots, j, k), j \neq k$, и финальную $\tilde{\gamma}^{l+1}(i, j) = (i, \underbrace{j, \dots, j}_{l \text{ раз}}), i \neq j, (l, j)$ -серии, которые реализуются в последовательности ξ_1, \dots, ξ_t длины $t = l + n - 1$, если $\xi_1^{l+1} = \gamma^{l+1}(j, k)$ и $\xi_{n-1}^{l+1} = \tilde{\gamma}^{l+1}(i, j)$, соответственно. Серии вида $\gamma^{l+1}(j, k)$ и $\tilde{\gamma}^{l+1}(i, j), i \neq j \neq k$ называются изолированными права и слева (l, j) -сериями, соответственно.

В работах В. А. Иванова и А. Е. Новикова [38, 39, 40] проведено систематическое исследование асимптотического поведения следующих сл. в.: $\mu_r(n, N, M, l)$ — число исходов среди первых M, l -членные серии которых встретились в реализации сл. последовательности длины n ровно по r раз; $\xi_r(n, N, M, l)$ — число исходов среди первых M, l -членные серии которых встретились в реализации длины n не более чем по r раз; $\theta(n, N, M, l)$ — число l -членных серий первых M исходов в реализации длины n ; $\nu_m(N, k, M, l)$ — минимальная длина реализации сл. последовательности, содержащая не менее чем по m l -членных серий некоторых k из первых M исходов; $\eta(N, k, M, l)$ — минимальная длина реализации сл. последовательности, содержащая k l -членных серий первых M исходов.

В приведенных определениях под термином «серия» следует понимать любую из определенных выше разновидностей серий.

В [38] изучено асимптотическое поведение сл. в. $\nu_1(N, k, N, l)$ и $\mu_0(n, N, N, l)$ для равновероятной полиномиальной схемы при $N \rightarrow \infty, l = \text{const}$ в различных областях изменения параметров n и $k, 1 \leq n < \infty, 1 \leq k \leq N$.

В [39] исследовано асимптотическое поведение сл. в. $\mu_0(n, N, M, l), \theta(n, N, M, l), \nu_1(N, k, M, l), \eta(N, k, M, l)$ для равновероятной полиномиальной схемы во всем диапазоне изменения параметра $M, 1 \leq M \leq N$ и таких N, l, n, k , что $Nln \rightarrow \infty$ и $Nlk \rightarrow \infty$.

Наконец, в [40] результаты работ [38, 39] обобщены на цепи Маркова с матрицами вероятностей переходов вида

$$P = \| p_{\alpha\beta} \|_1^N, \quad p_{\alpha\beta} = p_\beta + \delta(1 - \delta_{\alpha\beta}), \quad 0 < p_\alpha < 1, \quad \alpha = 1, \dots, N,$$

$$\delta = \left(1 - \sum_{\alpha=1}^N p_\alpha \right) / (N - 1),$$

$\delta_{\alpha\beta}$ — символ Кронекера. Стационарное распределение цепи имеет вид (π_1, \dots, π_N) , где $\pi_\beta = \frac{p_\beta + \delta}{1 + \delta}, \beta = 1, \dots, N$.

В случае $\delta=0$ рассматриваемая цепь Маркова превращается в полиномиальную схему с вероятностями p_1, \dots, p_N , $\sum_{\beta=1}^N p_\beta = 1$, в случае $p_\beta = p > 0$, $\beta = 1, \dots, N$, получаем факторизуемую цепь Маркова [21].

В [40] приведены точные формулы для производящих функций, математических ожиданий и дисперсий сл. в. $\mu_0(n, N, N, l)$, $\theta(n, N, N, l)$, $v_1(N, k, N, l)$, $\eta(N, k, N, l)$ для описанной цепи Маркова.

Для факторизуемой цепи Маркова получен полный спектр предельных распределений указанных выше сл. в. в областях изменения параметров N, n, k, l, p , удовлетворяющих соотношениям $kl[p(1-p)]^{-1} \rightarrow \infty$, $Nl[p(1-p)]^{-1} \rightarrow \infty$, $nl[p(1-p)]^{-1} \rightarrow \infty$. В основе доказательства теорем в [40] лежит метод производящих функций (п. ф.).

Пусть $h_{jl}(n)$ — число изолированных (l, j) -серий (включая начальную и финальную) в реализации сл. последовательности длины n .

Тогда имеет место равенство

$$\sum_{j=1}^N \sum_{l=1}^n l h_{jl}(n) = n,$$

а общее число изолированных серий будет равно $\sum_{j=1}^N \sum_{l=1}^n h_{jl}(n)$.

Обозначим $\mathbf{h}(n) = (h_{11}(n), \dots, h_{N1}(n), \dots, h_{1l}(n), \dots, h_{Nl}(n), \dots)$, $\mathbf{x} = (x_{11}, \dots, x_{N1}, \dots, x_{1l}, \dots, x_{Nl}, \dots)$, $\mathbf{m} = (m_{11}, \dots, m_{N1}, \dots, m_{1l}, \dots, m_{Nl}, \dots)$, где m_{jl} — неотрицательные целые числа,

$\mathbf{x}^{\mathbf{m}} = \prod_{j=1}^N \prod_{l=1}^{\infty} x_{jl}^{m_{jl}}$, $f_\beta(\mathbf{x}) = \sum_{l=1}^{\infty} (p_\beta + \delta) p_\beta^{l-1} x_{\beta l}$, $\beta = 1, \dots, N$, и $p_{\alpha\beta}(\mathbf{m}) = P_{\alpha\beta}(\mathbf{h}(n) = \mathbf{m})$ — вероятность того, что в реализации рассматриваемой цепи Маркова длины n содержится m_{jl} изолированных (l, j) -серий ($j = 1, \dots, N$, $l = 1, \dots, n$), начальной реализовалась α -серия и финальной — β -серия ($\alpha = 1, \dots, N$, $\beta = 1, \dots, N$).

Введем п. ф.

$$P_{\alpha\beta}(\mathbf{x}) = \sum_{n=1}^{\infty} \sum_{\mathbf{m}} p_{\alpha\beta}(\mathbf{m}) \mathbf{x}^{\mathbf{m}}, \quad \alpha, \beta = 1, \dots, N, \quad (4.3)$$

где суммирование ведется по всевозможным векторам \mathbf{m} с целочисленными неотрицательными координатами. Функции (4.3) удовлетворяют системе уравнений

$$P_{\alpha\beta}(x) = f_{\beta}(x) \left(\frac{\delta_{\alpha\beta}}{1+\delta} + \sum_{\substack{\gamma=1 \\ \gamma \neq \beta}}^N P_{\alpha\gamma}(x) \right), \quad \alpha, \beta = 1, \dots, N, \quad (7.3)$$

решение которой имеет вид

$$P_{\alpha\beta}(x) = \begin{cases} \frac{f_{\alpha}(x)}{(1+\delta)(1+f_{\alpha}(x))} \left(1 - \sum_{\substack{j=1 \\ j \neq \alpha}}^N \frac{f_j(x)}{1+f_j(x)} \right) \left(1 - \sum_{i=1}^N \frac{f_i(x)}{1+f_i(x)} \right)^{-1}, & \beta = \alpha, \\ \frac{f_{\alpha}(x)f_{\beta}(x)}{(1+\delta)(1+f_{\alpha}(x))(1+f_{\beta}(x))} \left(1 - \sum_{i=1}^N \frac{f_i(x)}{1+f_i(x)} \right)^{-1}, & \beta \neq \alpha. \end{cases} \quad (4.4)$$

С помощью соотношений (4.4) получены п. ф. сл. в. μ_r , ξ_r , θ ν_m и η .

§ 5. Время ожидания в схемах размещения частиц по ячейкам

Сколько необходимо разместить в N ячейках частиц, чтобы наступило то или иное событие, определяемое заполнением ячеек? Задачи такого типа составляют класс задач о времени ожидания в схемах размещения. Типичный и наиболее часто встречающийся в литературе пример такой задачи следующий: как распределено число частиц $\nu_m(N, k)$, после размещения которых по N ячейкам впервые появятся k ячеек ($1 \leq k \leq N$), содержащих не менее m частиц каждая ($m \geq 1$)?

5.1. Равновероятное размещение частиц. Впервые предельное при $N \rightarrow \infty$ распределение сл. в. $\nu_m(N, k)$ при $k=N$ и $m = \text{const}$ в схеме равновероятного размещения частиц получено Эрдешем и Реньи [135]; одно из многочисленных обобщений их результата состоит в следующем (Каплан [157]):

$$\max_{1 \leq k \leq N} |P(\nu_m(N, k)/N - \ln N - (m-1) \ln \ln N \leq x) - \Pi_{N-k} \left(\frac{e^{-x}}{(m-1)!} \right)| = o(1), \quad (5.1)$$

где $\Pi_m(\lambda) = \sum_{r=0}^m \pi_r(\lambda)$. Последующими исследованиями выявлена

полная картина асимптотического поведения сл. в. $\nu_m(N, k)$ в равновероятной схеме при всех возможных соотношениях между параметрами N , k и m . Обзор и библиография соответствующих результатов, охватывающие период до 1972 г. включительно, содержится в работе В. Ф. Колчина и В. П. Чистякова [76]. Большое внимание этим задачам уделено также в моно-

графин В. Ф. Колчина, Б. А. Севастьянова и В. П. Чистякова [75]. Это направление продолжало развиваться и в последующем. Так, в работе Г. И. Ивченко [54] изучено совместное распределение сл. в. $v_m(N, 1)$ и $v_m(N, N)$ при $N \rightarrow \infty$ и любом $m = o(N)$ и их разности $\tau_m(N) = v_m(N, N) - v_m(N, 1)$, т. е. числа размещаемых частиц после того, как была заполнена до уровня m одна ячейка, до заполнения, по крайней мере, до этого уровня всех N ячеек. Один из результатов этой работы гласит, что при $N, m, m/\ln N \rightarrow \infty, m = o(N)$ сл. в.

$$v_m^*(N, 1) = \sqrt{\frac{2 \ln N}{m} \cdot \frac{v_m(N, 1)}{N}} - \sqrt{2m \ln N} (1 + u_1(t)) - \ln \sqrt{4\pi}$$

и

$$v_m^*(N, N) = \sqrt{\frac{2 \ln N}{m} \cdot \frac{v_m(N, N)}{N}} - \sqrt{2m \ln N} (1 + u_1(t)) + \ln \sqrt{4\pi} \quad (5.2)$$

асимптотически независимы и имеют предельные функции распределения $1 - e^{-e^x}$ и $e^{-e^{-x}}$, соответственно, а функция распределения сл. в. $\tau_m^*(N) = v_m^*(N, N) - v_m^*(N, 1)$ сходится к $\int_0^\infty \exp\left\{-y - \frac{e^{-x}}{y}\right\} dy$. В выражениях (5.2) $t = (\ln N - \frac{1}{2} \ln \ln N) / m$ и функции $u_1(t) \leq 0, u_2(t) \geq 0$ являются решениями уравнения $u - \ln(1 + u) = t, t \geq 0$. В этой же работе проведено полное исследование предельных при $N \rightarrow \infty$ и $m = o(N)$ распределений максимального (минимального) заполнения ячеек в случайный момент $v_m(N, N)$ (соответственно, $v_m(N, 1)$).

В работе Флатто [138] изучалось совместное распределение сл. в. $v_i(N, N), i = 1, \dots, m$, и получен следующий результат: для любых x_1, \dots, x_m при $N \rightarrow \infty$

$$P(v_i(N, N)/N - \ln N - (i-1) \ln \ln N \leq x_i, i = 1, \dots, m) \rightarrow \prod_{i=1}^m \exp\left\{-\frac{e^{-x_i}}{(i-1)!}\right\}. \quad (5.3)$$

5.2. Полиномиальная схема размещения. Пусть частицы распределяются по ячейкам так, что каждая частица независимо от остальных попадает в ячейку с номером j с вероятностью $p_j = \frac{a_j}{N}, j = 1, \dots, N (a_1 + \dots + a_N = N)$. Если ограничиться классом регулярных схем, т. е. когда выполняется условие

$$0 < C_1 \leq \min_j a_j \leq \max_j a_j \leq C_2 < \infty, \quad (5.4)$$

то многие результаты об асимптотическом поведении времени ожидания $v_m(N, k)$, известные для равновероятной схемы раз-

мещения, переносятся и на полиномиальную схему размещения. Приведем формулировки соответствующих теорем:

1) Если $N, k \rightarrow \infty$ так, что $0 < \lambda_1 \leq \lambda = \frac{k}{N} \leq \lambda_2 < 1$, и выполнено условие (5.4), то сл. в. $v_m(N, k)$ асимптотически нормальна с параметрами $Nu_m(\lambda)$ и $Nv_m^2(\lambda)$, где для каждого $t \in [0, 1]$ функция $u_m(t)$ определяется как решение уравнения

$$1 - \frac{1}{N} \sum_{j=1}^N \Pi_{m-1}(a_j u) = t \quad (5.5)$$

и

$$v_m^2(\lambda) = \left(1 - \lambda - \frac{1}{N} \sum_{j=1}^N \Pi_{m-1}^2(u_m(\lambda) a_j) \right) (u_m'(\lambda))^2 - u_m(\lambda).$$

2) Пусть $N, k \rightarrow \infty$ так, что $s = N - k = \text{const}$, и при этом $a_j = 1 + \frac{g_j}{\ln N}$, $|g_j| \leq C < \infty$, $j = 1, \dots, N$, и $\frac{1}{N} \sum_{j=1}^N e^{-g_j} \rightarrow g < \infty$.

Тогда

$$P(v_m(N, k)/N - \ln N - (m-1) \ln \ln N \leq x) \rightarrow \Pi_s \left(g \frac{e^{-x}}{(m-1)!} \right). \quad (5.6)$$

3) Пусть при $N, k \rightarrow \infty$ выполняется условие

$$\frac{k^2}{2} \sum_{j=1}^N p_j^2 \rightarrow \lambda, \quad 0 < \lambda < \infty.$$

Тогда

$$P(v_1(N, k) - k = r) \rightarrow \pi_r(\lambda), \quad r = 0, 1, 2, \dots \quad (5.7)$$

4) Если $N \rightarrow \infty$, а $k = \text{const}$, то при $m \geq 2$

$$P \left(\left(\sum_{j=1}^N p_j^m \right)^{1/m} v_m(N, k) \leq x \right) \rightarrow 1 - \Pi_{k-1} \left(\frac{x^m}{m!} \right). \quad (5.8)$$

Отметим, что при $k = o(\sqrt{N})$ сл. в. $v_1(N, k)$ в пределе вырождается в точку k , в то время как при $m \geq 2$ сл. в. $v_m(N, k)$ по вероятности неограниченно возрастает при любом конечном k , имея при этом порядок $O \left(N^{\frac{m-1}{m}} \right)$.

Эти результаты содержатся в работах Г. И. Ивченко [53] и Г. И. Ивченко и Т. В. Ивановой [58], соотношение (5.7) получено ранее Холстом [146]. В работах [48, 93] получены точные формулы для производящей функции и моментов сл. в. $v_m(N, k)$ в полиномиальной схеме размещения.

Изучение сл. в. $v_m(N, k)$ с помощью соотношения

$$P(v_m(N, k) \leq n) = P(i_m(n, N) \leq N - k)$$

сводится к изучению разделимой статистики $\xi_m(n, N) = \sum_{j=1}^N f_m(h_j)$, где $h_j = h_j(n)$, $j = 1, \dots, N$, — заполнения ячеек

после размещения n частиц, а функция $f_m(x)$ принимает значение 1 в точках $0, 1, \dots, m-1$ и 0 — в остальных точках. Полный спектр предельных теорем для сл. в. $\xi_m(n, N)$ при условии (5.4) приведен в работе [58].

В работе Т. В. Ивановой и Г. И. Ивченко [49] рассматривалась сл. в. $v_m(N, k)$ как функция параметра k — процесс ожидания $\{v_m(N, k), k = 1, \dots, N\}$ и доказана сходимость этого процесса (в смысле сходимости конечномерных распределений) при $N \rightarrow \infty$ и соответствующей центрировке и нормировке к некоторому гауссовскому процессу.

Вопросы применения статистик типа времени ожидания для проверки статистических гипотез относительно вероятностей исходов в полиномиальной схеме в случае растущего числа исходов N рассматривались в работах [44, 47, 49, 53, 68]. Пусть гипотеза H_0 означает, что $a_1 = \dots = a_N = 1$ (равновероятная схема), а H_1 — произвольная «близкая» к H_0 альтернатива вида

$$H_1: a_j = 1 + b_j(N) N^{-\frac{1}{4}}, \quad j = 1, \dots, N, \quad |b_j(N)| \leq C < \infty, \quad (5.9)$$

и при $N \rightarrow \infty$

$$\frac{1}{N} \sum_{j=1}^N b_j^2(N) \rightarrow b^2 > 0.$$

В работе Г. И. Ивченко [53] построен критерий для различения этих гипотез, критическая область которого имеет вид $\{v_m(N, \lambda N) > C\}$ или $\{v_m(N, \lambda N) < C\}$, и исследовано поведение мощности такого критерия при $N \rightarrow \infty$ в зависимости от соотношения между параметрами λ и m . В частности, наибольшую асимптотическую мощность среди простейших критериев (случай $m = 1$) имеет критерий, основанный на статистике $v_1(N, \lambda^* N)$, где λ^* — корень уравнения $\frac{4\lambda}{4-3\lambda} - \ln \frac{1}{1-\lambda} = 0$. Соответствующая критическая область при уровне значимости γ имеет вид

$$\left\{ v_1(N, \lambda^* N) > N \ln \frac{1}{1-\lambda^*} + t_\gamma \sqrt{N \left(\frac{\lambda^*}{1-\lambda^*} - \ln \frac{1}{1-\lambda^*} \right)} \right\},$$

а предельное значение мощности при альтернативе (5.9) равно $\Phi \left(\frac{b^2}{2} \Phi(\lambda^*) - t_\gamma \right)$. Здесь $\Phi(x)$ — функция стандартного нормального распределения, $\Phi(-t_\gamma) = \gamma$ и $\Phi(\lambda) = 8\lambda(1-\lambda)^{1/2}(4-3\lambda)^{-3/2}$.

Табулирование различных характеристик рассматриваемых критериев проведено в работе [68]; в частности, $\lambda^* = 0,972 \dots$, $\Phi(\lambda^*) = 1,145 \dots$. Отметим также, что для $m \geq 2$ имеет место эффект падения мощности при $\lambda = \lambda_m = 1 - P_{m-1}(m-1)$: критерий

рий, основанный на статистике $v_m(N, \lambda_m N)$, асимптотически не отличается альтернативы вида (5.9) от нулевой гипотезы.

В работе Т. В. Ивановой и Г. И. Ивченко [49] предложен и рассчитан многомерный вариант рассматриваемого критерия, задаваемый критической областью вида

$$\{v_m(N, \lambda_1 N) > C_1, \dots, v_m(N, \lambda_r N) > C_r\}, \\ 1 - \prod_{m=1}^r (m-1) < \lambda_1 < \dots < \lambda_r < 1.$$

Т. В. Ивановой [44] изучены критерии, основанные на статистиках $v_m(N, k)$ при $k = \text{const}$ и $N - k = \text{const}$, а также рассмотрен случай проверки неравновероятной нулевой гипотезы [47]. В последнем случае при $k = \lambda N$, $0 < \lambda < 1$, соответствующий критерий позволяет различать гипотезы, сближающиеся со скоростью $N^{-1/2}$, однако он уже перестает быть критерием согласия: некоторый подкласс таких близких альтернатив он не отличает в пределе от нулевой гипотезы.

5.3. Схема со случайными уровнями. В полиномиальной схеме размещения свяжем с j -й ячейкой неотрицательную целочисленную сл. в. v_j ($P(v_j = 0) < 1$) и будем называть эту ячейку заполненной в «момент» n (после размещения n частиц), если $h_j(n) \geq v_j$ ($j = 1, \dots, N$). Предполагается, что «уровни» v_1, \dots, v_N независимы между собой и не зависят от $h_1(n), \dots, h_N(n)$, $n = 1, 2, \dots$. Таким образом, при каждом $n = 0, 1, 2, \dots$ определено число заполненных ячеек $\xi(n, N)$. Определим время ожидания до заполнения k ячеек как

$$v(N, k) = \min \{n : \xi(n, N) = k\}.$$

Отметим, что если $v_1 = \dots = v_N = m = \text{const}$, т. е. уровни не случайны и одинаковы, то это время ожидания совпадает с рассмотренной выше сл. в. $v_m(N, k)$. Асимптотическое при $N \rightarrow \infty$ поведение сл. в. $v(N, k)$ для схемы со сл. уровнями v_1, \dots, v_N изучалось в работе В. А. Иванова, Г. И. Ивченко и А. М. Протасова [35], а для случая одинаково распределенных уровней v_j и равновероятного размещения частиц — в работе В. А. Иванова [33].

Положим

$$R_{v_j}(z) = \sum_{n=0}^{\infty} P(v_j > n) \pi_n(z), \quad H_j(t) = 1 - R_{v_j}(a_j t), \quad j = 1, \dots, N,$$

и введем функции

$$F_{k,N}(t) = 1 - \sum_{l=0}^{k-1} \sum_{1 < j_1 < \dots < j_l < N} \prod_{s=1}^l H_{j_s}(t) \prod_{j \in I_{N,l}} (1 - H_j(t)),$$

где $I_{N,l} = \{1, \dots, N\} \setminus \{j_1, \dots, j_l\}$. Тогда для производящей функции сл. в. $v(N, k)$ справедливо представление

$$Ez^{\nu(N,k)} = \int_0^{\infty} \exp\left\{-\frac{1-z}{z} Nt\right\} dF_{k,N}(t), \quad 0 < z \leq 1. \quad (5.10)$$

Здесь $F_{k,N}(t)$ представляет собой функцию распределения k -й порядковой статистики некоторой совокупности из N независимых сл. в., и потому соотношение (5.10) сводит задачу исследования предельных при $N \rightarrow \infty$ распределений времени ожидания $\nu(N, k)$ для различных значений параметра k к изучению асимптотического поведения соответствующих порядковых статистик.

Если уровни ν_1, \dots, ν_N одинаково распределены, то для случая равновероятного размещения частиц можно получить исчерпывающую картину асимптотического поведения $\nu(N, k)$ для различных значений параметра k . Обозначим для этого случая $R_\nu(z)$ общую производящую функцию хвостов распределения уровней ν_j , $H(t) = 1 - R_\nu(t)$ и пусть $u(\lambda)$ — функция, обратная к $H(t)$: $H(u(\lambda)) \equiv \lambda$. Тогда имеют место следующие предельные теоремы (ограничимся случаем $P(\nu > 0) = 1$).

1) Пусть $N \rightarrow \infty$, а $k = \text{const}$. Тогда если $a_1 = P(\nu = 1) > 0$, то

$$P(\nu(N, k) = k + n) \rightarrow C_{k+n-1}^n a_1^k (1 - a_1)^n, \quad n = 0, 1, 2, \dots; \quad (5.11)$$

если $m = \min\{l: P(\nu = l) > 0\} > 1$ и $a_m = P(\nu = m)$, то

$$P\left(N^{\frac{1-m}{m}} \nu(N, k) \leq x\right) \rightarrow 1 - \Pi_{k-1}\left(\frac{a_m}{m!} x^m\right). \quad (5.12)$$

2) Пусть $N, k \rightarrow \infty$ так, что $\frac{k}{N} \rightarrow \lambda$, $0 < \lambda < 1$. Тогда сл. в. $\nu(N, k)$ асимптотически нормальна с параметрами $Nu(\lambda)$ и $N\sigma^2(\lambda)$, где $\sigma^2(\lambda) = \lambda(1-\lambda)(u'(\lambda))^2 - u(\lambda)$.

3) Пусть для любого $\tau > 0$ имеет место соотношение

$$R_\nu(t)/R_\nu(\tau t) \rightarrow \tau^\alpha, \quad t \rightarrow \infty. \quad (5.13)$$

Тогда при $N, k \rightarrow \infty$ и $s = N - k = \text{const}$

$$P\left(\frac{\nu(N, k)}{Nu(1-1/N)} \leq x\right) \rightarrow \Pi_s(x^{-\alpha}), \quad x > 0. \quad (5.14)$$

4) Определим две последовательности констант $b_N = u\left(1 - \frac{1}{N}\right)$ и $a_N = u\left(1 - \frac{1}{eN}\right) - u\left(1 - \frac{1}{N}\right)$ и пусть $b_N/a_N \rightarrow \infty$, $b_N/(Na_N^2) \rightarrow 0$ при $N \rightarrow \infty$. Тогда при $N, k \rightarrow \infty$, $s = N - k = \text{const}$

$$P\left(\frac{\nu(N, k) - Nb_N}{Na_N} \leq x\right) \rightarrow \Pi_s(e^{-x}), \quad |x| < \infty. \quad (5.15)$$

Из приведенных результатов видно, что в рассматриваемой схеме возникают новые типы предельных распределений, не встречающихся при изучении схемы с детерминированными уровнями. Это отрицательное биномиальное распределение

(5.11), когда $a_1 < 1$, и распределение (5.14). Так если при некотором целом $m \geq 1$

$$P(v > n) = \frac{m!}{(n+1)\dots(n+m)}, \quad n=0, 1, \dots,$$

то $R_v(t) = \frac{m!}{t^m} (1 - \Pi_{m-1}(t))$, и имеет место (5.13) с $\alpha = m$, а $u\left(1 - \frac{1}{N}\right) \sim (m!N)^{1/m}$ при $N \rightarrow \infty$. В этом случае время ожида-

ния $v(N, k)$ при k , близком к N , имеет порядок $O\left(N^{1+\frac{1}{m}}\right)$, а $m-1$ есть число (конечных) моментов у сл. уровня v . В частности, если v имеет бесконечное среднее, то $v(N, k)$ растет как $O(N^2)$, если v имеет конечное среднее, но бесконечный второй момент, то $v(N, k)$ растет как $O(N^{3/2})$ и т. д. Соотношение (5.15) имеет место, например, когда при $t \rightarrow \infty$

$$R_v(t) = at^\alpha e^{-\beta t} (1 + o(1)), \quad a, \beta > 0, \quad \alpha \geq 0.$$

Для этого случая

$$b_N = \frac{1}{\beta} \left(\ln N + \alpha \ln \ln N - \ln \frac{\beta^\alpha}{a} \right) + o(1), \quad a_N = \frac{1}{\beta} + o(1).$$

Влияние параметра $p = P(v=0)$ при $p > 0$ на предельное поведение времени ожидания $v(N, k)$ проанализировано в [33].

Для общего случая (неодинаково распределенные уровни и (или) неравновероятность размещения частиц) доказаны теоремы следующего типа: если при каждом $t > 0$ существует

$$\lim_{N \rightarrow \infty} \prod_{j=1}^N R_{v_j}(a_j, t) = R(t) > 0, \quad (5.16)$$

тогда при $N \rightarrow \infty$

$$P(v(N, 1)/N \leq t) \rightarrow 1 - R(t), \quad t > 0. \quad (5.17)$$

Условие (5.16) означает, что необходимо выполнение по крайней мере одного из следующих условий: 1) $\lim_{j \rightarrow \infty} P(v_j > n) = 1$, $n=1, 2, \dots$, 2) $\lim_{j \rightarrow \infty} a_j = 0$, т. е. с большей вероятностью первыми будут заполняться ячейки с малыми номерами. Так, в случае равновероятного размещения ($a_1 = \dots = a_N = 1$) справедливо следующее утверждение: если $S(N)$ означает номер первой заполненной ячейки, то при условии (5.16) существует

$$\lim_{N \rightarrow \infty} P(S(N) = j, v(N, 1)/N \leq t) = - \int_0^t \frac{R'_{v_j}(t)}{R_{v_j}(t)} R(t) dt, \quad (5.18)$$

$$j = 1, 2, \dots$$

Т. В. Ивановой [46] проведено исследование асимптотического поведения сл. в. $\xi(n, N)$ — числа заполненных ячеек —

при различных соотношениях между n и N в схеме равновероятного размещения частиц и при однородных уровнях.

5.4. Распределение других характеристик схемы, связанных с временем ожидания. Пусть $h_{(1)} \leq \dots \leq h_{(N)}$ — вариационный ряд (в. р.) из заполнений ячеек в случайный «момент» $\nu_m(N, k)$ при последовательном и равновероятном размещении частиц по N ячейкам. В работе Г. И. Ивченко [52] проведено исследование предельных при $N \rightarrow \infty$ распределений различных характеристик этого вариационного ряда (экстремальных значений $h_{(1)}$ и $h_{(N)}$; участков постоянства μ_r , т. е. числа членов в. р., принявших значение r ; числа членов, имеющих максимальное значение и др.) для различных значений параметра k и произвольном фиксированном m . В данном случае частоты h_j оказываются условно независимыми сл. в., распределения которых задаются урезанными пуассоновскими законами с одним и тем же случайным параметром. Это дает возможность записать распределение любой характеристики в. р. в виде смеси соответствующего параметрического семейства распределений; при этом распределение рандомизированного параметра является бета-распределением $B(N-k+1, k)$. В качестве примера укажем вид распределения максимального заполнения: если $s < m$, то $P(h_{(N)} \leq s) = 0$, а при $s \geq m$

$$P(h_{(N)} \leq s) = NC_{N-1}^{k-1} \int_0^1 \left[\frac{1}{1-\theta} \sum_{r=m}^s \pi_r(t_m(\theta)) \right]^{k-1} \theta^{N-k} \times \\ \times (1-\theta)^{k-1} d\theta, \quad (5.19)$$

где $t_m(x)$ — функция, обратная к функции $\Pi_{m-1}(t)$.

Асимптотическое поведение сл. в. $h_{(N)}$ для случая $k=N$ и сл. в. $h_{(1)}$ для случая $k=1$ в зависимости от соотношения между уровнем m и числом ячеек N исследовано в работе Г. И. Ивченко [54]. В частности, при $m = \text{const}$ максимальное заполнение $h_{(N)}$ имеет порядок $e \ln N$, не зависящий от m , а минимальное заполнение $h_{(1)}$ «отрывается от нуля» лишь при $m/\ln N \rightarrow e$.

Распределение сл. в. μ_r , в случае $k=N$, изучалось также Флатто [138], а Холстом [154] рассматривались сл. в. вида $z = f(h_1) + \dots + f(h_N)$.

5.5. Время ожидания в других моделях. Пусть N ячеек, в которых последовательно и равновероятно размещаются частицы, расположены в виде таблицы из s строк, состоящих из N_1, \dots, N_s ячеек, соответственно, ($N_1 + \dots + N_s = N$). Через $S(m, k)$, $m = (m_1, \dots, m_s)$, $k = (k_1, \dots, k_s)$ обозначим состояние таблицы, характеризующееся наличием в j -й строке не менее k_j ячеек, заполненных не менее m_j частицами каждая ($j=1, \dots, s$). Обозначим время ожидания до момента достижения состояния $S(m, k)$ через $\nu_m(N, k)$, $N = (N_1, \dots, N_s)$. В работе Г. И. Ивченко и Т. В. Ивановой [57] исследовано асимптотическое поведение этой сл. в. при $N \rightarrow \infty$ и различном характере

поведения величин N_1, \dots, N_s и k_1, \dots, k_s (параметры m_1, \dots, m_s фиксированы).

Распределение сл. в. $v_m(N, k)$, когда выбор строки, в которой размещается очередная частица, управляется эргодической цепью Маркова, рассмотрено в работе А. Асимова [4].

Время ожидания в схемах распределения частиц комплектами изучалось в работах [43, 51, 60] (см. § 3).

В работе Г. И. Ивченко и Т. В. Ивановой [48] рассматривалось время ожидания до заполнения, по крайней мере, до уровня m каждой из k выделенных ячеек в полиномиальной схеме размещения частиц по $N > k$ ячейкам и получена производящая функция и моменты соответствующей сл. в. Т. В. Ивановой [45] проведено исследование асимптотического поведения времени ожидания $v_m(N, k)$ в схеме с одной «особой» ячейкой, заполнение которой нас «не интересует». Выяснено влияние «веса» этой особой ячейки (т. е. вероятности попадания в нее частиц) на предельное поведение сл. в. $v_m(N, k)$.

Б.6. Схемы с фиксированным числом ячеек. Пусть ξ_1, ξ_2, \dots — последовательность сл. в., принимающих значения $1, 2, \dots, N$, и $h_j(n)$ — число появлений j -го исхода в n испытаниях. Вектор частот $h(n) = (h_1(n), \dots, h_N(n))$ порождает некоторое случайное блуждание в N -мерном евклидовом пространстве R^N .

В ряде приложений возникает необходимость изучения сл. в. типа

$$\mu'(A) = \min \{n: h(n) \notin A\}, \quad v'(B) = \min \{n: h(n) \in B\},$$

где A и B — заданные области в R^N . Величина $\mu'(A)$ определяет момент выхода случайного блуждания $h(n)$, $n=1, 2, \dots$, из области A , величина $v'(B)$ — момент достижения области B .

В. А. Иванов и Г. И. Ивченко [34] исследовали случай удаляющихся от начала координат областей с границами, параллельным осям:

$$A = A(l) = \{(x_1, \dots, x_N): x_j < l_j, j=1, \dots, N\},$$

$$B = B(m) = \{(x_1, \dots, x_N): x_j \geq m_j, j=1, \dots, N\},$$

где $l = (l_1, \dots, l_N)$ и $m = (m_1, \dots, m_N)$ — заданные векторы с целочисленными положительными координатами.

Обозначим $\mu(l) = \mu'(A(l))$ и $v(m) = v'(B(m))$, а в случае

$$l_j \equiv l, \quad m_j \equiv m, \quad \mu(l) \equiv \mu(l), \quad v(m) \equiv v(m).$$

Тогда $\mu(l)$ ($v(m)$) есть момент достижения уровня $l(m)$ случайным функционалом $\max_{1 < j < N} h_j(n)$ ($\min_{1 < j < N} h_j(n)$).

Соотношение

$$P(\mu(l) > n_1, v(m) \leq n_2) = P(h(n_1) < l, h(n_2) \geq m)$$

позволяет сводить изучение совместного распределения сл. в. $\mu(l)$ и $v(m)$ к изучению двумерных распределений процесса

$\{h(n)\}$. В [34] получены предельное совместное распределение сл. в. $\mu(l)$ и $\nu(m)$, и асимптотика их моментов при $l, m \rightarrow \infty$, $l/m \rightarrow \alpha \leq 1$, где $l = \min_j \frac{l_j}{p_j} \leq m = \max_j \frac{m_j}{p_j}$, в случае, когда последовательность $\xi_{n1}, \xi_{n2}, \dots$ является однородной эргодической цепью Маркова с N состояниями, заданной матрицей вероятностей переходов и стационарными вероятностями (p_1, \dots, p_N) .

Например, для последовательности независимых и равновероятных сл. в. предельная функция распределения величины $(\nu(m) - Nm)/(N\sqrt{Nm})$ при $m \rightarrow \infty$, $N = \text{const}$ имеет вид

$$H(x) = \int_{\Delta(x)} \varphi(z_1, \dots, z_{N-1}) dz_1 \dots dz_{N-1}, \quad x > 0,$$

где $\Delta(x) = \{z_1, \dots, z_{N-1} : z_i < x, \quad i = 1, \dots, N-1, \quad z_1 + \dots + z_{N-1} > -x\}$,

$$\varphi(z_1, \dots, z_{N-1}) = N^{N/2} (2\pi)^{-(N-1)/2} \exp \left\{ -\frac{N}{2} \left(\sum_{i=1}^N z_i^2 - \left(\sum_{i=1}^N z_i \right)^2 \right) \right\}.$$

Для моментов сл. в. $\nu(m)$ справедливы формулы

$$E\nu(m) = Nm + \delta_1 N \sqrt{Nm} (1 + o(1)), \quad D\nu(m) \sim (\delta_2 - \delta_1^2) N^3 m,$$

где $\delta_r = \int_0^\infty x^r dH(x)$ и для $r = 1, 2$ $\delta_1 = \sqrt{N} \int_0^1 x^{N-1} \Phi^{-1}(x) dx$,

$$\delta_2 = \int_0^1 x^{N-1} (\Phi^{-1}(x))^2 dx, \quad \Phi^{-1}(x) \text{ — функция, обратная к } \Phi(x).$$

Для получения более точных асимптотических разложений в некоторых частных случаях был использован метод непосредственного асимптотического анализа допредельных выражений.

В работе В. А. Иванова [29] приведены асимптотические разложения для первых двух моментов сл. в. $\mu(l)$ и $\nu(m)$ в случае бернуллиевской последовательности, а в работе В. А. Иванова и А. Е. Новикова [37] — для цепи Маркова с двумя состояниями. Так, для дважды стохастической цепи Маркова с двумя состояниями 0 и 1, с вероятностями переходов, задаваемыми значением δ , где $p_{01} = \frac{1}{2}(1-\delta)$, при $m_1 = m_2 = m \rightarrow \infty$ имеют место соотношения

$$E\nu(m) = 2m + 2\sqrt{\frac{(1+\delta)}{\pi(1-\delta)}} m + \frac{\delta}{1-\delta} - \frac{1-3\delta^2}{4(1-\delta^2)} \sqrt{\frac{1+\delta}{\pi(1-\delta)}} m + O\left(\frac{1}{m\sqrt{m}}\right),$$

$$D\nu(m) = 2\left(1 - \frac{2}{\pi}\right) \frac{1+\delta}{1-\delta} m + \frac{2}{\sqrt{\pi}} \left(\frac{1+\delta}{1-\delta}\right)^{3/2} \sqrt{m} +$$

$$+ \frac{\delta(1+\delta)}{(1-\delta)^2} + \frac{1-3\delta^2}{\pi(1-\delta^2)} + O\left(\frac{1}{\sqrt{m}}\right);$$

для равновероятной бернуллиевской последовательности

$$Dv(m) = 2\left(1 - \frac{2}{\pi}\right)m + \frac{2}{\sqrt{\pi}}\sqrt{m} + \frac{1}{\pi} - \frac{1}{4\sqrt{\pi m}} - \frac{1}{8\pi m} + \\ + O\left(\frac{1}{m\sqrt{m}}\right).$$

Случай двух блуждающих на плоскости точек рассмотрен в работе В. А. Иванова [31].

§ 6. Оценивание параметров для конечных совокупностей

В ряде прикладных задач, связанных с выборочным обследованием конечных совокупностей, мы имеем дело с ситуацией, когда число элементов (объем) N исследуемой совокупности \mathcal{U} является априори либо неизвестной величиной, либо о нем известно лишь, что его значение находится в некоторых заданных пределах $N_1 \leq N \leq N_2$. В этих случаях возникает задача получения тех или иных статистических выводов относительно N на основе имеющейся статистической информации об \mathcal{U} . Такая информация представляет собой обычно выборку элементов из \mathcal{U} , извлеченную по некоторому заданному стохастическому закону, и решение задачи, естественно, существенно зависит от типа этого закона. Чаще всего в приложениях имеют дело со схемами повторной и бесповторной выборки.

6.1. Точечные оценки. Простейший пример подобной задачи обсуждается в книге В. Феллера [112, стр. 56—58], где речь идет об оценке числа рыб в озере по результатам двух независимых уловов, когда каждый улов представляет собой бесповторную выборку. В этом случае статистические выводы о неизвестном числе рыб N формулируются на основании изучения статистики μ_2 — числа рыб, попавших в оба улова (величина μ_2 имеет, как известно, гипергеометрическое распределение). Если обозначить объемы выборок через m_1 и m_2 , соответственно, то в данном случае для оценки максимального правдоподобия (о. м. п.) \hat{N} параметра N справедливо соотношение:

$$\hat{N} \approx \frac{m_1 m_2}{\mu_2} = \frac{m_1 m_2}{m_1 + m_2 - \eta},$$

где $\eta = m_1 + m_2 - \mu_2$ — число разных рыб, пойманных за оба улова. Таким образом, в данном случае (две независимые бесповторные выборки) изучение статистики μ_2 фактически эквивалентно изучению статистики η . Детальный анализ такой двухвыборочной схемы проводится в работе Чэпмана [126]. Обобщение этой схемы на случай произвольного числа s выборок рассматривалось в работах Г. И. Ивченко и Е. Е. Тимошиной [64, 65, 66].

Пусть m_1, \dots, m_s ($s \geq 2$) — объемы независимых бесповторных выборок из совокупности \mathcal{U} , $m = \max_{1 \leq i \leq s} m_i$, $n = m_1 + \dots + m_s$ и μ_r обозначает число элементов \mathcal{U} , каждый из которых вошел ровно в r некоторых выборок, $r = 1, \dots, s$. Тогда вся совокупность данных представляется векторной статистикой $\mu = (\mu_1, \dots, \mu_s)$, и по этой информации требуется оценить неизвестный параметр N и, более общо, произвольную параметрическую функцию $\tau(N)$.

В работе [66] доказывается, что статистика $\eta = \mu_1 + \dots + \mu_s$ — общее число наблюдаемых элементов \mathcal{U} — является полной достаточной статистикой для параметра N и, основываясь на этом, описывается класс функций $\tau(N)$, для которых существуют несмещенные (а значит, и с минимальной дисперсией) оценки $\hat{\tau}(\eta)$, и указывается, как по τ определить ф. Оказывается, что если $N \leq n$, то оптимальная оценка существует для любой функции $\tau(N)$ и она имеет вид

$$\hat{\tau} = \Phi(\eta) = \frac{1}{\Delta^n \psi(0)} \sum_{j=m}^n (-1)^{n-j} C_n^j \tau(j) \psi(j), \quad (6.1)$$

где $\psi(j) = \prod_{i=1}^s C_j^{m_i}$ и Δ — оператор разности: $\Delta f(x) = f(x+1) - f(x)$. В частности, оптимальная оценка для параметра N имеет вид (при $N \leq n$)

$$\hat{N} = \Delta^n (0 \cdot \psi(0)) / \Delta^n \psi(0). \quad (6.2)$$

Если же $N (\geq m)$ может быть априори любым натуральным числом, то оптимальные оценки можно построить лишь для функций вида $\tau(N) = f(N) / \psi(N)$, где $f(N)$ — многочлен степени не выше n , удовлетворяющий условиям $f(x) = 0$ при $x = 0, 1, \dots, m-1$. Если $\tau(N)$ — такая функция, то оптимальная оценка для нее дается формулой

$$\hat{\tau} = \Delta^n (\tau(0) \psi(0)) / \Delta^n \psi(0). \quad (6.3)$$

В частности, несмещенная оценка с минимальной дисперсией для функции $\tau(N) = \frac{1}{N}$ всегда существует и имеет вид

$$\hat{\tau} = \Delta^n (0^{-1} \cdot \psi(0)) / \Delta^n \psi(0). \quad (6.4)$$

Если же ограничиться классом оценок, являющихся линейными функциями от μ_1, \dots, μ_s , то в этом классе единственной несмещенной оценкой для $\tau(N) = 1/N$ является статистика

$$\bar{\tau} = \sum_{r=2}^s r(r-1) \mu_r \left/ \sum_{i \neq j} m_i m_j \right. \quad (6.5)$$

Для параметра N построена также о. м. п. \hat{N} , которая для случая выборок одинакового объема ($m_1 = \dots = m_s = m$) находится из условия [65]

$$S_m(\hat{N} - 1, \eta) > s \geq S_m(\hat{N}, \eta), \quad (6.6)$$

если $\eta > m$, и $\hat{N} = m$, если $\eta = m$. Здесь

$$S_m(N, k) = \left(\ln \frac{N+1}{N-k+1} \right) / \ln \frac{N+1}{N-m+1} \text{ при } N \geq k$$

и

$$S_m(k-1, k) = +\infty.$$

Показано также [64], что с помощью линейных оценок вида

$\sum_{r=1}^s c_r \mu_r$ несмещенным образом можно оценивать лишь полиномы от $1/N$ степени не выше $s-1$.

Положив в предыдущих формулах $m_1 = \dots = m_s = 1$, можно получить соответствующие выводы для схемы простой повторной выборки, которая изучалась Харрисом [142] (его результат содержится в формуле (6.2)) и Дримлом и Ульрихом [133] (их результат содержится в (6.6)).

Основываясь на распределении статистики η , для схемы с s бесповторными выборками равных объемов m рассчитаны доверительные интервалы для N для значений $n = sm = 5$ (5) 25, всех возможных комбинаций s и m и доверительных уровней 0,9; 0,95; 0,99.

6.2. Асимптотическое оценивание. На практике часто объем совокупности представляет собой априори весьма большую величину и потому представляет интерес асимптотическая постановка задачи, предполагающая неограниченное возрастание как параметра N , так и объема выборки.

Пусть $m_1 = \dots = m_s = m \rightarrow \infty$, $N \rightarrow \infty$ и при этом $p = \frac{m}{N} \in [p_1, p_2]$, где $0 < p_1 < p_2 < 1$ — заданные границы, в которых находится неизвестное значение параметра p . В качестве оценки для p , согласно (6.5), можно рассматривать статистику

$$\bar{p}_m = \frac{1}{s(s-1)m} \sum_{r=2}^s r(r-1)\mu_r,$$

которая является единственной несмещенной оценкой, линейной относительно μ_1, \dots, μ_s . В рассматриваемых асимптотических условиях эта оценка является состоятельной и (равномерно по p) асимптотически нормальной (с. а. н.) $\mathcal{N}\left(p, \frac{2pq^2}{s(s-1)m}\right)$, $q = 1 - p$.

Выделим в классе линейных оценок $\mathcal{L} = \left\{ l_m = \frac{1}{m} \sum_{r=1}^s f(r)\mu_r \right\}$

подкласс $\mathcal{L}' \subset \mathcal{L}$ с монотонной (по p) функцией $a_f(p) = \frac{1}{p} E f(X)$, где X — биномиальная случайная величина с параметрами s и p ; при этом должно выполняться условие $a_f'(p) \neq 0$ при $p \in [p_1, p_2]$. Введем новый класс оценок $\mathfrak{M}_s = \{\zeta_m = a_f^{-1}(l_m) : l_m \in \mathcal{L}'\}$, где $a_f^{-1}(t)$ — функция, обратная к $a_f(p)$. При любом s оценка $\bar{p}_m \in \mathfrak{M}_s$, но при $s > 2$ класс \mathfrak{M}_s содержит и другие оценки, которые при рассматриваемых асимптотических условиях также являются с. а. н. оценками параметра p . В работе [65] доказано, что в классе \mathfrak{M}_s существует единственная асимптотически оптимальная (т. е. с асимптотически минимальной дисперсией) оценка параметра p , которая имеет вид $\zeta_m^* = a^{-1}\left(\frac{\eta}{m}\right)$, где $a(p) = \frac{1 - q^s}{p}$. Эта оценка представляет собой асимптотический аналог о. м. п. $\hat{p}_m = m/\hat{N}$, где \hat{N} определяется, согласно (6.6).

В [65] решаются также задачи асимптотического оценивания (точечного и доверительного) произвольных параметрических функций $\tau(p)$, а также проверки гипотез (простых и сложных) о параметре p . В частности, асимптотически наикратчайшим доверительным интервалом для p , отвечающим доверительному уровню $1 - \gamma$, является интервал $(\zeta_m^* \pm t_{\gamma/2} \sqrt{\Phi(\zeta_m^*)/m})$, где $\Phi(p) = \frac{p^3 q^s}{1 - q^s - s p q^{s-1}}$, $\Phi(-t_\delta) = \delta$ и $\Phi(x)$ — функция стандартного нормального распределения.

В работе [67] получены аналоги этих результатов для случая простой повторной выборки, когда ее объем n и объем совокупности N удовлетворяет условиям: $n, N \rightarrow \infty$, $\alpha = \frac{n}{N} \in [\alpha_1, \alpha_2]$, где $0 < \alpha_1 < \alpha_2 < \infty$ — заданные границы, в которых по условию находится неизвестное значение параметра α .

В [64] проведено сравнение схем повторной и бесповторной выборки с точки зрения эффективности оценивания параметра N , когда в качестве соответствующих оценок рассматриваются статистики $\bar{N}_1 = m/\bar{p}_m$ (схема с s бесповторными выборками равных объемов m) и $\bar{N}_{11} = n/\bar{\alpha}_n$ (схема повторной выборки объема $n = sm$, $\bar{\alpha}_n$ — несмещенная линейная оценка α). Показано, что при $n \leq N/2$ (основной случай в приложениях) оценка \bar{N}_{11} асимптотически более эффективная, чем \bar{N}_1 , но при $n > N/2$ всегда можно указать значение s , при котором более эффективной является оценка \bar{N}_1 .

6.3. Расслоенные совокупности. Пусть элементы конечной совокупности \mathcal{U} разбиты по некоторому признаку на k классов (слоев) $\mathcal{U}_1, \dots, \mathcal{U}_k$ ($\mathcal{U}_i \cap \mathcal{U}_j = \emptyset$, $i \neq j$, $\mathcal{U}_1 \cup \dots \cup \mathcal{U}_k = \mathcal{U}$) размерами, соответственно, N_1, \dots, N_k . Параметры N_1, \dots, N_k и $N = N_1 + \dots + N_k$ неизвестны и подлежат оцениванию. Для слу-

чая, когда имеется простая повторная выборка элементов из \mathcal{U} , эта задача решена в работе Г. И. Ивченко [55].

Пусть объем выборки есть n и η_j обозначает число элементов j -го слоя, вошедших в выборку, $j=1, \dots, k$. Тогда вектор $\eta = (\eta_1, \dots, \eta_k)$ является полной достаточной статистикой для параметров $N = (N_1, \dots, N_k)$. Пусть требуется оценить заданную параметрическую функцию $\tau(N)$. Тогда задача сводится к выяснению условий разрешимости и отысканию при этих условиях явного вида решения уравнения несмещенности

$$E_N \Phi(\eta) = \tau(N), \quad \forall N \neq 0. \quad (6.7)$$

Доказано, что если возможные значения параметров N ограничены условием $N \leq n$, то уравнение (6.7) разрешимо для любой функции $\tau(N)$ и при этом оптимальной оценкой для $\tau(N)$ является статистика

$$\hat{\tau} = \Phi(\eta) = (\Delta_1^{\eta_1} \dots \Delta_k^{\eta_k} - (-1)^n I) f(0) / \Delta^n 0^n, \quad (6.8)$$

где $\eta = \eta_1 + \dots + \eta_k$, $f(x) = (x_1 + \dots + x_k)^n \tau(x)$, $x = (x_1, \dots, x_k)$, Δ_j — оператор разности по j -й переменной и I — тождественный оператор.

Если же $N \neq 0$ априори может быть любым целочисленным вектором, то уравнение (6.7) разрешимо лишь для функций вида $\tau(N) = f(N) / N^n$, где $f(N)$ — многочлен от N_1, \dots, N_k степени не выше n , удовлетворяющий условию $f(0) = 0$. Для любой такой функции оптимальная оценка имеет вид (6.8).

В частности, несмещенным образом можно оценивать лишь полиномиальные функции степени не выше n от долей $\beta_j = \frac{N_j}{N}$, $j=1, \dots, k$. Задача оценивания долей β_1, \dots, β_k с использованием выборочных частот $\tilde{\beta}_i = \frac{h_i}{n}$, h_i — число наблюдений элементов i -го класса ($i=1, \dots, k$), обсуждается в [137].

Для параметров N, N_1, \dots, N_k построены также оценки максимального правдоподобия.

Если объем совокупности $N \rightarrow \infty$, а объем выборки n и размеры классов N_1, \dots, N_k связаны с N соотношениями

$$\begin{aligned} n &= \alpha N, \quad 0 < C_1 \leq \alpha \leq C_2 < \infty, \\ N_j &= \beta_j N, \quad 0 < C_3 \leq \beta_j \leq C_4 < 1, \quad j=1, \dots, k, \end{aligned} \quad (6.9)$$

то оценивать параметры α и $\beta = (\beta_1, \dots, \beta_k)$ можно на основании следующих предельных теорем:

1. Статистика $\hat{\alpha}_n = m^{-1}(\eta/n)$, где $m(\alpha) = (1 - e^{-\alpha})/\alpha$ и $m^{-1}(t)$ — функция, обратная к $m(\alpha)$, является с. а. н. $\mathcal{N}(\alpha, b^2(\alpha)/n)$ оценкой для α , где $b^2(\alpha) = \alpha^3 / (e^\alpha - 1 - \alpha)$, а асимптотическим доверительным интервалом для α с доверительным уровнем $1 - \gamma$ является интервал $(\hat{\alpha}_n \pm t_{\gamma/2} b(\hat{\alpha}_n) / \sqrt{n})$, $\Phi(-t_\delta) = \delta$.

2. Статистика $\hat{\beta}_n = \left(\frac{\eta_1}{\eta}, \dots, \frac{\eta_k}{\eta} \right)$ является с. а. н. $\mathcal{N}(\bar{\beta}, \frac{1}{n} G_n B_n G_n')$ оценкой для β , где G есть матрица с элементами $\partial g_i(x) / \partial x_j$ при $x = m(\alpha) \beta$, $g_i(x) = \frac{x_i}{x_1 + \dots + x_k}$, $i = 1, \dots, k$, B — матрица с элементами

$$b_{ij} = \begin{cases} \frac{1}{\alpha_i} e^{-\alpha} \left(1 - e^{-\alpha} - \frac{\alpha^2}{\alpha_i} e^{-\alpha} \right), & i = j, \\ -\frac{\alpha^2}{\alpha_i \alpha_j} e^{-2\alpha}, & i \neq j, \end{cases}$$

$\alpha_i = \alpha / \beta_i$, $i = 1, \dots, k$, а G_n , B_n — значения этих матриц при $\alpha = \hat{\alpha}_n$, $\alpha_i = \hat{\alpha}_{in} = \frac{\eta}{\eta_i} \hat{\alpha}_n$, $i = 1, \dots, k$.

6.4. Другие выборочные схемы. Имеется большое число работ как математических, так и сугубо прикладного характера, посвященных задаче оценивания неизвестного размера конечной совокупности, в которых рассматриваются специфические статистические эксперименты, «высвечивающие» некоторое случайное подмножество элементов совокупности. Одно перечисление подобных работ заняло бы слишком много места (их список только за последние несколько лет насчитывает несколько десятков названий), поэтому мы ограничимся упоминанием лишь отдельных последних публикаций [10, 107, 108, 120, 124, 125, 127, 129, 131, 152, 164, 169, 171] и монографии Ю. К. Беляева [9], в которой также затрагиваются близкие вопросы. В качестве одного соответствующего примера рассмотрим следующую модель эксперимента, которая может быть наглядно интерпретирована в терминах распространения слухов [56, 123]. Именно, пусть имеется коллектив из N лиц, среди которых распространяется некоторая новость. В начальный момент новость сообщается некоторому числу Y_1 лиц. В следующий момент каждое из этих Y_1 лиц независимо от остальных передает ее какому-нибудь другому лицу, выбираемому с равной вероятностью из всех остальных $N-1$ членов коллектива. В результате таких Y_1 контактов новость станет известной некоторому дополнительному числу лиц $Y_2 (\leq Y_1)$, которые по описанному алгоритму продолжают распространение слуха на третьем этапе ит. д. Таким образом, процесс распространения слуха может быть описан невозрастающей последовательностью $\{Y_1 \geq Y_2 \geq \dots\}$, где Y_t есть число лиц, узнающих новость в результате случайных контактов на t -м этапе, $t = 2, 3, \dots$. Поскольку $\sum_{t \geq 1} Y_t \leq N$, а размер N коллектива конечен, то последовательность $\{Y_t\}$ также конечна. Пусть $v = \min\{t: Y_t = 0\}$ есть момент прекращения распространения слуха. Тогда $S_v = Y_1 + \dots + Y_v$ — число лиц, которым в конеч-

ном итоге станет известна новость, — является полной достаточной статистикой для N , а оптимальная (т. е. несмещенная с минимальной дисперсией) оценка для N существует и имеет вид (при $Y_1=r$)

$$\widehat{N}_r = r + \varphi_r(S_v), \quad (6.10)$$

где функция

$$\varphi_r(s) = S(s-1, s-r-1; r) / S(s-1, s-r; r-1)$$

и

$$S(n, k; x) = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} C_k^j (x+j)^n = \frac{\Delta^k x^n}{k!}$$

есть так называемые обобщенные числа Стирлинга 2-го рода [83]. Если S_v велико, то вместо (6.10) можно использовать простую приближенную формулу $\widehat{N}_r \sim S_v^2/2r$. При $N \rightarrow \infty$ и фиксированном r статистика $\frac{2r}{N} \widehat{N}_r$ распределена асимптотически по закону $\chi^2(2r)$ и при этом $D\widehat{N}_r \sim N^2/r$, а асимптотическим γ -доверительным интервалом для N является интервал

$$\Delta(\gamma) = \left[S_v^2/t_{2r}\left(\frac{1+\gamma}{2}\right), S_v^2/t_{2r}\left(\frac{1-\gamma}{2}\right) \right],$$

где $t_{2r}(\delta)$ есть δ -квантиль закона $\chi^2(2r)$.

ЛИТЕРАТУРА

1. *Азларов Т. А., Мухамедханова Р.*, Класс предельных распределений статистической оценки энтропии в полиномиальной схеме. Докл. АН СССР, 1982, 264, № 4, 777—780 (РЖМат, 1982, 10В100)
2. *Амбросимов А. С.*, Об асимптотическом поведении максимума в однородной цепи Маркова с большим числом состояний. Теория вероятностей и ее применения, 1978, 23, № 2, 438—445 (РЖМат, 1978, 11В85)
3. *Асимов А.*, Предельные теоремы в задаче размещения двух типов частиц по бесконечным ячейкам. Редкол. ж. «Изв. АН УзССР. Сер. физ.-мат. н.» Ташкент, 1981, 14 с. Библиогр. 8 назв. (Рукопись деп. в ВИНТИ 22 июня 1981 г., № 3023—81Деп.) (РЖМат, 1981, 11В36ДЕП)
4. —, О многомерной предельной теореме в задачах размещения нескольких типов частиц. Изв. АН УзССР. Сер. физ.-мат. н., 1982, № 3, 3—6 (РЖМат, 1982, 12В35)
5. *Баранов А. П.*, Некоторые предельные теоремы для условных распределений U -статистик. Науч. тр. Кубан. ун-т, 1979, № 282/2, 27—38 (РЖМат, 1980, 5В164)
6. *Беллев П. Ф.*, О вероятности неоявления заданного числа исходов. Теория вероятностей и ее применения, 1964, 9, № 3, 541—547 (РЖМат, 1965, 2В61)
7. —, О вероятности неоявления заданного числа s -цепочек в сложных цепях Маркова. Теория вероятностей и ее применения, 1965, 10, № 3, 547—551 (РЖМат, 1966, 1В46)
8. —, О совместном распределении частот исходов в цепях Маркова с большим числом состояний. Теория вероятностей и ее применения, 1977, 22, № 3, 534—545 (РЖМат, 1978, 1В112)
9. *Беллев Ю. К.*, Вероятностные методы выборочного контроля. М.: Наука, 1975, 407 с. (РЖМат, 1976, 4В276К)

10. —, Выборочные методы. Теория вероятностей и ее применения, 1980, 25, № 3, 658—659
11. —, Рыкова Л. В., Непараметрический критерий Колмогорова для выборок из конечных совокупностей. Докл. АН СССР, 1973, 210, № 6, 1261—1264 (РЖМат, 1973, 10В150)
12. Болотников Ю. В., Предельные процессы в неравновероятной схеме размещения частиц по ячейкам. Теория вероятностей и ее применения, 1968, 13, № 3, 534—542 (РЖМат, 1969, 8В19)
13. Боровков А. А., О мощности критерия χ^2 при увеличении числа групп. Теория вероятностей и ее применения, 1977, 22, № 2, 375—379 (РЖМат, 1977, 9В151)
14. Ватутин В. А., Михайлов В. Г., Предельные теоремы для числа пустых ячеек в равновероятной схеме размещения частиц комплектами. Теория вероятностей и ее применения, 1982, 27, № 4, 684—692 (РЖМат, 1983, 4В32)
15. Викторова И. И., Об асимптотическом поведении максимума в равновероятной полиномиальной схеме. Мат. заметки, 1969, 5, № 3, 305—316 (РЖМат, 1969, 8В8)
16. —, Севастьянов Б. А., О предельном поведении максимума в полиномиальной схеме. Мат. заметки, 1967, 1, № 3, 331—338 (РЖМат, 1967, 10В23)
17. —, Чистяков В. П., Некоторые обобщения критерия пустых ящиков. Теория вероятностей и ее применения, 1966, 11, № 2, 306—313 (РЖМат, 1967, 2В104)
18. Глибоченко А. Ф., Ивченко Г. И., Медведев Ю. И., Об эффективности симметрических критериев в полиномиальной схеме и корреляционных свойствах разделимых статистик. Вероятн. процессы и их прилож. (междуз. сб.), М., МИЭМ, 1983, 3—17
19. Жуков А. Е., Некоторые теоремы для одной схемы размещения комплектами. Мат. заметки, 1979, 25, № 1, 149—155 (РЖМат, 1979, 7В30)
20. —, Чистяков В. П., О схемах размещения комплектов, аппроксимируемых полиномиальной схемой. Мат. заметки, 1979, 25, № 3, 455—463 (РЖМат, 1979, 8В10)
21. Захаров В. К., Сарманов О. В., О законе распределения числа серий в однородной цепи Маркова. Докл. АН СССР, 1968, 179, № 3, 526—528 (РЖМат, 1968, 12В67)
22. —, —, Севастьянов Б. А., Последовательный критерий χ^2 . Мат. сб., 1969, 79, № 3, 444—460 (РЖМат, 1969, 12В157)
23. Зубков А. М., Предельные распределения статистической оценки энтропии. Теория вероятностей и ее применения, 1973, 18, № 3, 643—650 (РЖМат, 1973, 12В163)
24. —, Цепи Маркова, близкие к последовательности независимых испытаний. Мат. заметки, 1979, 25, № 3, 465—478 (РЖМат, 1979, 7В37)
25. —, Неравенства для вероятностей переходов с запрещениями и их применения. Мат. сб., 1979, 109, № 4, 491—532 (РЖМат, 1979, 12В82)
26. —, Михайлов В. Г., Оценка точности пуассоновской аппроксимации в задаче о размещении частиц по ячейкам. Теория вероятностей и ее применения, 1978, 23, № 4, 815—823 (РЖМат, 1979, 4В11)
27. —, —, О повторениях s -цепочек в последовательности независимых величин. Теория вероятностей и ее применения, 1979, 24, № 2, 267—279 (РЖМат, 1979, 10В19)
28. —, Попов Н. Н., Отношение частичного порядка, порожденное распределениями числа занятых ячеек. Мат. заметки, 1982, 32, № 1, 97—102 (РЖМат, 1982, 11В6)
29. Иванов В. А., Некоторые задачи о простом случайном блуждании на прямой. Кибернетика, 1974, № 5, 122—126 (РЖМат, 1975, 4В76)
30. —, О необходимом числе наблюдений для различения сближающихся гипотез. В сб. «Методы исслед. операций и теории надежности в анализе системы». Киев, 1976, 3—23 (РЖМат, 1977, 5В128)

31. —, О блуждании двух точек на плоскости. Науч. тр. Кубан. ун-т, 1979, № 282/2, 3—7 (РЖМат, 1980, 5В54)
32. —, Рандомизированные разделимые статистики. В сб. «Теория и практика применения математических методов и средств вычислительной техники», Киев, ИК АН УССР, 1981, 67—75
33. —, Предельные теоремы в схеме размещения со случайными уровнями. Мат. заметки, 1982, 31, № 4, 619—631 (РЖМат, 1982, 8В16)
34. —, *Ивченко Г. И.*, О некоторых граничных функционалах для марковских блужданий. Мат. заметки, 1978, 23, № 2, 315—326 (РЖМат, 1978, 6В34)
35. —, —, *Протасов А. М.*, Порядковые статистики и их применение к задаче о размещении. Вероятн. процессы и их прилож. (межвуз. сб.), М., МИЭМ, 1984, 89—100
36. —, *Липин С. А.*, Об асимптотической нормальности рандомизированных разделимых статистик в полиномиальной схеме. Мат. заметки, 1983, 34, № 5, 745—756 (РЖМат, 1984, 2В47)
37. —, *Новиков А. Е.*, О цепях Маркова с двумя состояниями. Тр. Моск. ин-та электрон. машиностр., 1975, вып. 44, 160—167 (РЖМат, 1977, 9В42)
38. —, —, О распределении времени до появления заданного числа различных l -членных серий. Теория вероятностей и ее применения, 1977, 22, № 3, 546—555 (РЖМат, 1978, 2В11)
39. —, —, О распределении серий в полиномиальной схеме. Науч. тр. Кубан. ун-т, 1979, № 282/2, 12—26 (РЖМат, 1980, 7В107)
40. —, —, О распределении серий в факторизуемой цепи Маркова. Вероятн. процессы и их прилож. (межвуз. сб.). М., МИЭМ, 1983, 29—40
41. —, *Теребулин С. Ю.*, О времени ожидания в схеме размещения частиц комплектами. Тр. Моск. ин-та электрон. машиностр., 1976, № 57, 225—231 (РЖМат, 1978, 11В30)
42. —, —, Некоторые предельные теоремы в неравновероятной схеме размещения частиц комплектами. Теория вероятностей и ее применения, 1978, 23, № 3, 643—650 (РЖМат, 1979, 3В13)
43. —, —, Некоторые предельные теоремы в схеме размещения частиц растущими комплектами. Науч. тр. Кубан. ун-т, 1979, № 282/2, 71—79 (РЖМат, 1980, 4В218)
44. *Иванова Т. В.*, О некоторых статистических критериях в полиномиальной схеме. Тр. Моск. ин-та электрон. машиностр., 1976, № 57, 149—162 (РЖМат, 1978, 9В176)
45. —, Предельные теоремы в схеме последовательного заполнения ячеек при разбиении их на группы. Науч. тр. Кубан. ун-т, 1977, 247, 24—32 (РЖМат, 1978, 2В61)
46. —, К задаче о размещении частиц по ячейкам. Вероятн. процессы и их прилож. (межвуз. сб.), М., МИЭМ, 1984, 41—44
47. —, О статистических критериях, основанных на времени ожидания. Статистические методы оценивания и проверки гипотез (межвуз. сб.). Пермь, 1984, 23—28
48. —, *Ивченко Г. И.*, О времени ожидания в полиномиальной схеме. Тр. Моск. ин-та электрон. машиностр., 1975, вып. 44, 77—95 (РЖМат, 1977, 9В12)
49. —, —, Об одном критерии последовательного типа. Теория вероятностей и ее применения, 1976, 21, № 4, 854—857 (РЖМат, 1977, 4В227)
50. *Ивченко Г. И.*, О предельных распределениях порядковых статистик полиномиальной схемы. Теория вероятностей и ее применения, 1971, 16, № 1, 94—107 (РЖМат, 1971, 9В76)
51. —, Некоторые предельные теоремы в схеме размещения. Тр. Моск. ин-та электрон. машиностр., 1973, вып. 32, 111—119 (РЖМат, 1974, 11В36)
52. —, Время ожидания и вариационный ряд частот в полиномиальной схеме. Тр. Моск. ин-та электрон. машиностр., 1973, вып. 32, 39—64 (РЖМат, 1975, 1В42)
53. —, Время ожидания и проверка гипотез в полиномиальной схеме. Теория вероятностей и ее применения, 1974, 19, № 4, 839—844 (РЖМат, 1975, 4В178)

54. —, О некоторых распределениях, связанных с временем ожидания в полиномиальной схеме. Теория вероятностей и ее применения, 1975, 20, № 3, 557—570 (РЖМат, 1976, 3В33)
55. —, О некоторых задачах статистического вывода для расслоенных совокупностей. Теория вероятностей и ее применения, 1984, 29, № 1, 113—118 (РЖМат, 1984, 6В150)
56. —, *Воскресенский В. Н.*, Об одной модели распространения слухов и ее применение в статистике. 16 Всес. шк.-коллоквиум по теории вероятностей и мат. стат., Бакуриани, 26 февр.—5 марта, 1982. Материалы. Тбилиси, 1982, 21—31 (РЖМат, 1984, 1В474)
57. —, *Иванова Т. В.* О случайном заполнении таблиц. Тр. Моск. ин-та электрон. машиностр., 1975, вып. 44, 111—130 (РЖМат, 1977, 9В30)
58. —, —, Предельные теоремы для времени ожидания. Тр. Моск. ин-та электрон. машиностр., 1975, вып. 44, 96—110 (РМат, 1977, 7В107)
59. —, *Лёвшин В. И.*, Асимптотическая нормальность одного класса статистик в полиномиальной схеме. Теория вероятностей и ее применения, 1976, 21, № 1, 190—195 (РЖМат, 1976, 11В135)
60. —, —, Асимптотическая нормальность в схеме выбора без возвращения. Теория вероятностей и ее применения, 1978, 23, № 1, 98—108 (РЖМат, 1978, 12В12)
61. —, *Медведев Ю. И.*, Асимптотическое поведение числа комплектов частиц в классической задаче о размещении. Теория вероятностей и ее применения, 1966, 11, № 4, 701—708 (РЖМат, 1967, 4В22)
62. —, —, Разделимые статистики и проверка гипотез. Случай малых выборок. Теория вероятностей и ее применения, 1978, 23, № 4, 796—806 (РЖМат, 1979, 5В206)
63. —, —, Разделимые статистики и проверка гипотез для группированных данных. Теория вероятностей и ее применения, 1980, 25, № 3, 549—560 (РЖМат, 1981, 1В211)
64. —, *Тимошина Е. Е.*, Об оценивании при выборе из конечной совокупности. Мат. заметки, 1980, 28, № 4, 623—633 (РЖМат, 1981, 2В146)
65. —, —, О некоторых задачах оценивания для выборок из конечной совокупности. Изв. АН УзССР. Сер. физ.-мат. н., 1981, № 4, 27—33 (РЖМат, 1982, 2В173)
66. —, —, Об оптимальном оценивании для конечной совокупности, размер которой не известен. Мат. заметки, 1982, 31, № 4, 633—640 (РЖМат, 1982, 8В178)
67. —, —, Об оценивании размера конечной совокупности. Теория вероятностей и ее применения, 1982, 27, № 2, 380—384 (РЖМат, 1982, 10В102)
68. —, *Хонов С. А.*, Проверка статистических гипотез в полиномиальной схеме. Расчет критериев. Тр. Моск. ин-та электрон. машиностр., 1976, № 57, 101—135 (РЖМат, 1979, 2В195)
69. —, *Цуканов С. В.*, О новом способе обработки частот в методе группировки наблюдений и оптимальности критерия χ^2 . Докл. АН СССР, 1984, 277, № 2, 287—291
70. *Колчин В. Ф.*, Один класс предельных теорем для условных распределений. Лит. мат. сб., 1968, 8, № 1, 53—63 (РЖМат, 1969, 4В20)
71. —, О предельном поведении крайних членов вариационного ряда в полиномиальной схеме. Теория вероятностей и ее применения, 1969, 14, № 3, 476—487 (РЖМат, 1970, 4В26)
72. —, О распределении одной статистики в полиномиальной схеме. Тр. Моск. ин-та электрон. машиностр., 1973, вып. 32, 73—91 (РЖМат, 1974, 11В38)
73. —, Задача о размещении частиц по ячейкам и случайные отображения. Теория вероятностей и ее применения, 1976, 21, № 1, 48—62 (РЖМат, 1976, 8В39)
74. —, Случайные отображения, М.: Наука, 1984, 208 стр.
75. —, *Севастьянов Б. А.*, *Чистяков В. П.*, Случайные размещения. М.: Наука, 1976, 224 с. (РЖМат, 1976, 12В20К)
76. —, *Чистяков В. П.*, Комбинаторные задачи теории вероятностей. В сб.

- «Теория вероятностей. Мат. статистика. Теор. кибернетика. Т. 11 (Итоги науки и техн. ВИНТИ АН СССР)». М., 1974, 5—45 (РЖМат, 1974, 8В3)
77. —, Предельные распределения числа непооявившихся s -цепочек в полиномиальной схеме. Теория вероятностей и ее применения, 1974, 19, № 4, 855—864 (РЖМат, 1975, 4В26)
 78. *Лёвин В. В.*, Общая предельная нормальная теорема в задаче о размещении и ее приложение к статистике. Тр. Моск. ин-т электрон. машиностр., 1976, Вып. 57, 54—70 (РЖМат, 1978, 11В195)
 79. —, Критерии однородности для s выборок из полиномиального распределения. Науч. тр. Кубан. ун-т, 1977, 247, 38—45 (РЖМат, 1978, 2В171)
 80. *Медведев Ю. И.*, Некоторые теоремы об асимптотическом распределении статистики χ^2 . Докл. АН СССР, 1970, 192, № 5, 987—989 (РЖМат, 1970, 11В19)
 81. —, Разделимые статистики и полиномиальной схеме. I. Теория вероятностей и ее применения, 1977, 22, № 1, 3—17 (РЖМат, 1977, 7В106)
 82. —, Разделимые статистики в полиномиальной схеме. II. Теория вероятностей и ее применения, 1977, 22, № 3, 623—631 (РЖМат, 1978, 1В183)
 83. —, *Ивченко Г. И.*, Асимптотические представления конечных разностей от степенной функции в произвольной точке. Теория вероятностей и ее применения, 1965, 10, № 1, 151—156 (РЖМат, 1965, 9А133)
 84. *Михайлов В. Г.*, Предельная теорема Пуассона в схеме размещения частиц комплектами. Теория вероятностей и ее применения, 1977, 22, № 1, 155—160 (РЖМат, 1977, 8В37)
 85. —, Оценка скорости сходимости к распределению Пуассона при размещении частиц комплектами. Теория вероятностей и ее применения, 1977, 22, № 3, 566—574 (РЖМат, 1978, 1В28)
 86. —, Асимптотическая нормальность числа пустых ячеек при размещении частиц комплектами. Теория вероятностей и ее применения, 1980, 25, № 1, 83—91 (РЖМат, 1980, 8В8)
 87. —, Сходимость к многомерному нормальному закону в равновероятной схеме размещения частиц комплектами. Мат. сб., 1980, 111, № 2, 163—185 (РЖМат, 1980, 5В28)
 88. —, Центральная предельная теорема для схем независимого размещения частиц по ячейкам. Тр. Мат. ин-та АН СССР, 1981, 157, 138—152 (РЖМат, 1982, 2В20)
 89. —, Асимптотическая нормальность в схеме конечно-зависимого размещения частиц по ячейкам. Мат. сб., 1982, 119, № 4, 509—520 (РЖМат, 1983, 3В31)
 90. *Мухамедханова Р.*, Предельные распределения статистической оценки энтропии в случае биномиальной схемы. Изв. АН УзССР. Сер. физ.-мат. н., 1980, № 1, 10—15 (РЖМат, 1980, 8В97)
 91. —, Об асимптотическом поведении распределения статистической оценки энтропии. Докл. АН УзССР 1980, № 4, 3—4 (РЖМат, 1980, 9В139)
 92. *Попова Т. Ю.*, Предельные теоремы в одной модели размещения частиц двух типов. Теория вероятностей и ее применения, 1968, 13, № 3, 542—548 (РЖМат, 1969, 8В20)
 93. *Прогасов А. М.*, Производящая функция времени ожидания в полиномиальной схеме. Тр. Моск. ин-та электрон. машиностр., 1976, № 57, 239—240 (РЖМат, 1978, 11В26)
 94. *Савчук М. Н.*, Формулы для параметров распределения некоторых случайных величин в задаче о размещении частиц комплектами. В сб. «Математические методы исследования операций и теории надежности». Киев, Институт кибернетики АН УССР, 1978, 81—87
 95. —, Некоторые предельные теоремы в схеме равновероятного размещения частиц комплектами. Теория вероятностей и мат. стат. (Киев), 1983, № 28, 122—130 (РЖМат, 1983, 7В29)
 96. —, О предельных распределениях некоторых случайных величин, связанных с раскрытиями сообщений в дискретных системах передачи информации. Препринт, 82—56. Некоторые вопросы применения марковских

- процессов для моделирования систем. Киев, Институт кибернетики АН УССР, 1983, 29—36
97. *Сачков В. Н.*, Перечислительные задачи комбинаторного анализа. В сб. «Вопр. кибернетики», М., 1973, 146—164 (РЖМат, 1974, 1В301)
 98. —, Комбинаторные методы дискретной математики. М.: Наука, 1977. 319 с. (РЖМат, 1978, 4В321К)
 99. —, Вероятностные методы в комбинаторном анализе. М.: Наука, 1978. 288 с. (РЖМат, 1979, 2В7)
 100. —, Введение в комбинаторные методы дискретной математики. М.: Наука, 1982, 384 с.
 101. *Севастьянов Б. А.*, Предельные теоремы в одной схеме размещения частиц по ячейкам. Теория вероятностей и ее применения, 1966, 11, № 4, 696—700 (РЖМат, 1967, 4В21)
 102. —, Сходимость к гауссовскому и пуассоновскому процессам распределения числа пустых ящиков в классической задаче о дробинках. Теория вероятностей и ее применения, 1967, 12, № 1, 144—154 (РЖМат, 1967, 9В30)
 103. —, Критерий «пустых ящиков» и его обобщения. Тр. Ин-т прикл. мат. Тбилис. ун-та, 1969, № 2, 229—233 (РЖМат, 1970, 6В154)
 104. —, Предельный закон Пуассона в схеме сумм независимых случайных величин. Теория вероятностей и ее применения, 1972, 17, № 4, 733—738 (РЖМат, 1973, 4В53)
 105. —, Об одной схеме независимых размещений. Изв. АН УзССР. Сер. физ.-мат. н., 1981, № 2, 37—41 (РЖМат, 1981, 9В14)
 106. *Теребулин С. Ю.*, Вариационный ряд частот в схеме размещения конечного числа растущих комплектов. Стат. методы. Пермь, 1982, 153—159 (РЖМат, 1982, 10В334)
 107. *Тимошина Е. Е.*, Об оценивании вероятности появления нового элемента при выборе из конечной совокупности. Вероятн. процессы и их прилож. (межвуз. сб.). М., МИЭМ, 1983, 60—69
 108. —, Оценивание параметрических функций для конечной совокупности, состоящей из неизвестного числа классов одинакового объема М.: МИЭМ 1983, 12 с. (Рукопись деп. в ВИНТИ 13.12.83 № 6740—83 Деп)
 109. *Туманян С. Х.*, Об асимптотическом распределении критерия χ^2 . Докл. АН СССР, 1954, 94, № 6, 1011—1012 (РЖМат, 1954, 4862)
 110. —, Асимптотическое распределение критерия χ^2 при одновременном возрастании объема наблюдений и числа групп. Теория вероятностей и ее применения, 1956, 1, № 1, 131—145 (РЖМат, 1957 5013)
 111. —, О мощности критерия χ^2 относительно «близких» альтернатив. Математика (межвуз. сб.). I. Ереван, 1982, 6—18
 112. *Феллер В.*, Введение в теорию вероятностей и ее приложения. М.: Мир, 1, 1967, 500 с.
 113. *Форманов Ш. К., Асимов А.*, О локальной предельной теореме для одной схемы размещения частиц. Предел. теоремы для случайн. процессов и стат. выводы. Ташкент, 1981, 202—214 (РЖМат, 1982, 2В19)
 114. —, —, О предельных теоремах в задаче размещения частиц по бесконечным ячейкам. Докл. АН УзССР, 1982, № 3, 3—4 (РЖМат, 1982, 9В23)
 115. *Хакимуллин Е. Р.*, О предельном поведении максимального заполнения в равновероятной схеме размещения частиц комплектами. Мат. заметки, 1981, 30, № 2, 277—289 (РЖМат, 1981, 12В47)
 116. —, Скорость сходимости распределения максимального заполнения в равновероятной схеме размещения частиц. М., МИЭМ, 1981, 11 с. Библиогр. 3 назв. (Рукопись деп. в ВИНТИ 21 июля 1981, № 3651—81ДЕП) (РЖМат, 1981, 10В26 ДЕП)
 117. *Хмаладзе Э. В.*, Замечание о слабой сходимости линейных разделимых статистик. Теория вероятностей и ее применения, 1980, 25, № 3, 633—635 (РЖМат, 1981, 1В166)
 118. —, Некоторые применения теории мартигалов в статистике. Успехи мат. наук, 1982, 37, № 6, 193—212 (РЖМат, 1983, 3В224)

119. —, Мартингалные предельные теоремы для разделимых статистик. Теория вероятностей и ее применения, 1983, 28, № 3, 504—520 (ПЖМат, 1984, 1В61)
120. Шлоссер А., Об оценке объема словаря длинного текста на основе выборки. Изв. АН СССР. Техн. кибернет., 1981, № 1, 117—122 (ПЖМат, 1981, 6В511)
121. Юсас Я. С., О распределении максимальной частоты в полиномиальной схеме. Теория вероятностей и ее применения, 1972, 17, № 4, 755—761 (ПЖМат, 1973, 3В151)
122. Agrawal M. C., On averaging over distinct units in replicated samples. Math. Operationsforsch. and statist. Ser. Statist., 1982, 13, № 1, 91—97 (ПЖМат, 1982, 12В10)
123. Berg S., On showball sampling, random mappings and related problems. J. Appl. Probab., 1981, 18, 283—290 (ПЖМат, 1981, 9В387)
124. Castledine B. J., A Bayesian analysis of a multiple-recapture sampling for a closed population. Biometrika, 1981, 68, № 1, 197—210 (ПЖМат, 1981, 8В183)
125. Chao A., On estimating the probability of discovering a new species. Ann. Statist., 1981, 9, № 6, 1339—1342 (ПЖМат, 1982, 5В138)
126. Chapman D. G., Some properties of the hypergeometric distributions with applications to zoological sample censuses. University of California Publ. Statist., 1951, 1, 131—160
127. Charalambides Ch. A., On a restricted occupancy model and its applications. Biometr. J., 1981, 23, № 6, 601—610 (ПЖМат, 1982, 3В163)
128. Chen Wen Chen, Some local limit theorems in the symmetric Dirichlet-multinomial urn models. Ann. Inst. Statist. Math., 1981, 33, № 3, 405—415 (ПЖМат, 1982, 8В18)
129. Chromy J. R., Sequential sample selection methods. Proc. Sec. Surv. Res. Meth. Pap. Annu. Meet. Washington, D. C., 1979. Washington, D. C., s. a., 401—406 (ПЖМат, 1981, 11В204)
130. Csáki E., Földes A., On the first empty cell. Stud. sci. math. hung., 1976, 11, № 3—4, 373—382 (ПЖМат, 1980, 8В7)
131. Darroch J. N., Ratcliff D., A note on capture-recapture estimation. Biometrics, 1980, 36, № 1, 149—153 (ПЖМат, 1980, 11В419)
132. Defays D., Etude du comportement asymptotique de schémas d'urnes. Bull. Soc. roy. sci. Liège, 1974, 43, № 1—2, 26—34 (ПЖМат, 1974, 12В2)
133. Drimi M., Ullrich M., Maximum likelihood estimate of the number of types. Acta techn. CSAV, 1967, 12, № 3, 300—303 (ПЖМат, 1968, 8В94)
134. Englund G., A remainder term estimate for the normal approximation in classical occupancy. Ann. Probab., 1981, 9, № 4, 684—692 (ПЖМат, 1982, 6В23)
135. Erdős P., Rényi A., On a classical problem of probability theory. Magyar tud. akad. Mat. Kutató int. közl., 1961, 6, № 1—2, 215—220 (ПЖМат, 1962, 6В2)
136. Eshy W. W., Confidence intervals for the coverage of low coverage samples. Ann. Statist., 1982, 10, № 2, 190—196 (ПЖМат, 1982, 9В106)
137. Finch P. D., Population proportions and sample frequencies: a new look at an old problem. Math. Sci., 1981, 6, № 2, 103—109 (ПЖМат, 1981, 10В476)
138. Flatto L., Limit theorems for some random variables associated with urn models. Ann. Probab., 1982, 10, № 4, 927—934 (ПЖМат, 1983, 6В19)
139. Gvancladze L. G., Chibisov D. M., On tests of fit based on grouped data. Contributions to statistics. Acad. Prague, 1979, 79—89
140. Hajner R., Neuer Beweis eines klassischen Besetzungsproblems. Sitzungsber. Österr. Akad. Wiss. Math.-naturwiss. Kl., 1973, Abt. 2, 181, № 8—10, 269—289 (ПЖМат, 1975, 1В44)

141. *Harper L.*, Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 1961, 38, № 2, 410—414
142. *Harris B.*, Statistical inference in the classical occupancy problem unbiased estimation of the number of classes. *J. Amer. Statist. Assoc.*, 1968, 63, № 323, 837—847 (PЖMat, 1969, 6B87)
143. —, *Park C. J.*, A note on the asymptotic normality of the distribution of the number of empty cells in occupancy problem. *Math. research center summary report*, Univ. of Wisconsin., 1968, № 955
144. —, —, The distribution of linear combinations of the sample occupancy numbers. *Proc. Kon. ned. akad. wetensch.*, 1971, 474, № 2, 121—132 (PЖMat, 1971, 12B69)
145. *Herrmann H.*, Besetzungsmodelle. *Biometr. J.*, 1979, 21, № 7, 629—643 (PЖMat, 1980, 7B439)
146. *Holst L.*, Limit theorems for some occupancy and sequential occupancy problems. *Ann. Math. Stat.*, 1971, 42, № 5, 1671—1680 (PЖMat, 1972, 5B24)
147. —, Asymptotic normality in a generalized occupancy problem. *Z. Wahrscheinlichkeitstheor. and verw. Geb.*, 1972, 21, № 2, 109—120 (PЖMat, 1972, 4B16)
148. —, Asymptotic normality and efficiency for certain goodness-of-fit tests. *Biometrika*, 1972, 59, № 1, 137—145 (PЖMat, 1972, 8B141)
149. —, Some asymptotic results for occupancy problems. *Ann. Probab.*, 1977, 5, № 6, 1028—1035 (PЖMat, 1978, 8B28)
150. —, Two conditional limit theorems with applications. *Ann. Statist.*, 1979, 7, № 3, 551—557 (PЖMat, 1980, 1B202)
151. —, A unified approach to limit theorems for urn models. *J. Appl. Probab.*, 1979, 16, № 1, 154—162 (PЖMat, 1980, 3B35)
152. —, Some asymptotic results for incomplete multinomial and Poisson samples. *Scand. J. Statist. Theory and Appl.*, 1981, 8, № 4, 243—246 (PЖMat, 1982, 6B189)
153. —, Some conditional limit theorems in exponential families. *Ann. Probab.*, 1981, 9, № 5, 818—830 (PЖMat, 1982, 5B22)
154. —, On sequential occupancy problems. *J. Appl. Probab.*, 1981, 18, № 2, 435—442 (PЖMat, 1981, 12B46)
155. *Hurt J., Machek J., Stepán J., Vorličková D.*, The intersections of random finite sets. *Math. slov. (CSSR)*, 1982, 32, № 3, 229—237 (PЖMat, 1983, 1B35)
156. *Johnsson N. L., Kotz S.*, Urn models and their application. New York, Wiley, 1977, 405 p.
157. *Kaplan N.*, A generalization of a result of Erdős and Rényi. *J. Appl. Probab.*, 1977, 14, № 1, 212—216 (PЖMat, 1976, 2B60)
158. *Koehler K. J., Larntz K.*, An empirical investigation of goodness-of-fit statistics for sparse multinomials. *J. Amer. Statist. Assoc.*, 1980, 75, № 370 326—334 (PЖMat, 1981, 2B124)
159. *Lam Hing-Kam*, Bounded multinomial distribution. *Commun. Statist. Theory and Meth.*, 1982, 11, № 16, 1869—1880 (PЖMat, 1983, 1B14)
160. *Morris C.*, Central limit theorems for multinomial sums. *Ann. Statist.*, 1975, 3, № 1, 165—188 (PЖMat, 1975, 10B104)
161. *Park C. J.*, A note on the classical occupancy problem. *Ann. Math. Stat.*, 1972, 43, № 5, 1698—1701 (PЖMat, 1973, 4B54)
162. —, On the distribution of the number of unobserved elements when m -samples of size n are drawn from a finite population. *Commun. Statist.*, 1981, A10, № 4, 371—383 (PЖMat, 1981, 10B199)
163. *Pease Roger W., Jr.*, General solution to the occupancy problem with variably sized runs of adjacent cells occupied by single balls. *Math. Mag.*, 1975, 48, № 3, 131—134 (PЖMat, 1976, 4B422)
164. *Pollock K. H.*, Capture-recapture models allowing for age-dependent survival and capture rates. *Biometrics*, 1981, 37, № 3, 521—529 (PЖMat, 1982, 4B393)

165. *Quine M. P.*, Three limit theorems for scores based on occupancy numbers. *Ann. Probab.*, 1980, 8, № 1, 148—156 (PЖMat, 1980, 12B31)
 166. *Ronault A.*, Propriétés asymptotiques d'un n -échantillon d'une variable aléatoire dénombrable connues sous le nom de lois de Zipf. *C. r. Acad. sci.*, 1976, 283, № 6, A379—A380 (PЖMat, 1977, 4B137)
 167. *Samuel-Cahn E.*, Asymptotic distributions for occupancy and waiting time problems with positive probability of falling through the cells. *Ann. Probab.*, 1974, 2, № 3, 515—521 (PЖMat, 1975, 4B25)
 168. *Saperstein B.*, The generalized birthday problem. *J. Amer. Statist. Assoc.*, 1972, 67, № 338, 425—428 (PЖMat, 1973, 1B15)
 169. *Seber G. A. F., Felton R.*, Tag loss and the Petersen mark-recapture experiment. *Biometrika*, 1981, 68, № 1, 211—219 (PЖMat, 1981, 8B486)
 170. *Sobel M., Uppuluri V. R.*, Sparse and crowded cells and Dirichlet distributions. *Ann. Statist.*, 1974, 2, № 5, 977—987 (PЖMat, 1975, 4B150)
 171. *Starr N.*, Linear estimation of the probability of discovering a new species. *Ann. Statist.*, 1979, 7, № 3, 644—652 (PЖMat, 1980, 1B353)
 172. *Steck G. P.*, Limit theorems for conditional distributions. *Univ. Calif. Publ. Statist.*, 1957, 2, № 12, 237—284 (PЖMat, 1958, 5B15)
 173. *Strauss D.*, Runs of occupied cells. *Biometrika*, 1977, 64, № 1, 170—171 (PЖMat, 1977, 11B4)
 174. *Uksánović J. B.*, Random selection of mutually exclusive samples. *Mat. vech.*, 1981, 5, № 2, 233—234 (PЖMat, 1983, 2B196)
 175. *Weiss L.*, The normal approximation to the multinomial with an increasing number of classes. *Nav. Rev. Log. Quart.*, 1976, 23, № 1, 139—149 (PЖMat, 1977, 4B24)
-