Math-Net.Ru
Общероссийский математический портал

Г. И. Ивченко, Ю. И. Медведев, Статистические выводы для случайных подстановок по неполным данным,
Тр. по дискр. матем., 2006, том 9, 66–76

https://www.mathnet.ru/tdm141

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

Параметры загрузки:
IP: 18.97.14.80
11 декабря 2025 г., 05:29:09

Труды по дискретной математике
Том 9, с. 66–76
М.: Гелиос АРВ, 2006

СТАТИСТИЧЕСКИЕ ВЫВОДЫ ДЛЯ СЛУЧАЙНЫХ ПОДСТАНОВОК ПО НЕПОЛНЫМ ДАННЫМ

Г. И. ИВЧЕНКО, Ю. И. МЕДВЕДЕВ

В модели Эванса случайных подстановок с неизвестным параметром решаются задачи оценивания параметра модели и проверки гипотезы о равновероятности подстановок по неполным данным, когда наблюдается лишь некоторое число начальных членов цикловой структуры подстановок; также рассматривается случай больших выборок и предлагается критерий однородности данных для него.

§ 1. ВВЕДЕНИЕ

Рассматривается параметрическая модель Эванса случайных подстановок [1], когда произвольная подстановка A из симметрической группы S_n наблюдается с вероятностью

$$\mathbf{P}_\theta(A) = \frac{\theta^k}{\theta_{(n)}}, \tag{1}$$

где k — число циклов подстановки A, $\theta > 0$ — некоторый параметр, $\theta_{(n)} = \theta(\theta+1)\cdots(\theta+n-1)$. Как известно, $\sum_{k=0}^{n} c_{nk}\theta^k = \theta_{(n)}$, где c_{nk} — число подстановок из S_n с числом циклов, равным k. В случае, когда параметр θ неизвестен, возникает проблема получения тех или иных статистических выводов о нем (точечное и интервальное оценивание, проверка статистических гипотез). Решение подобных задач существенно зависит от имеющейся в нашем распоряжении статистической информации, т. е. от того, что мы наблюдаем. Мы предполагаем, что имеется информация

© Г. И. Ивченко, Ю. И. Медведев, 2006

о цикловой последовательности $c(n) = (c_1(n), c_2(n), \ldots, c_n(n))$ наблюдаемой в модели (1) случайной подстановки (или нескольких наблюдаемых подстановок), здесь $c_i(n)$ — число циклов длины i подстановки, $i = 1, 2, ..., n$.

Когда известна вся цикловая последовательность $c(n)$, то говорят, что имеется полная информация о наблюдаемой подстановке. Этот случай рассматривался в предыдущей работе авторов [3], где получено полное решение перечисленных выше статистических задач. Для приложений наибольший интерес представляют ситуации, когда порядок подстановок n достаточно велик, и тогда все соответствующие выводы имеют асимптотический (при $n \to \infty$) характер. Но в этом случае не всегда является реалистичным предположение о том, что мы можем наблюдать всю цикловую последовательность $c(n)$. Зачастую нам бывает доступно для подсчета лишь какое-то ограниченное число k ее первых членов $c_1(n), ..., c_k(n)$; в этом случае мы будем говорить о неполных данных.

Настоящая работа посвящена рассмотрению статистических задач для модели (1) именно по таким неполным данным. При этом мы будем проводить анализ в асимптотической постановке, предполагая, что порядок подстановок $n \to \infty$, а параметр неполноты k фиксирован.

Основой для дальнейших выводов будет следующее утверждение об асимптотическом распределении наблюдаемых статистик $c_1(n), ..., c_k(n)$ [2] (ниже символом $\mathcal{L}$ обозначается распределение соответствующих случайных величин).

Теорема. *Для случайной подстановки A в модели* (1) *при $n \to \infty$ и фиксированном $k \geqslant 1$*

$$\mathcal{L}(c_1(n), ..., c_k(n)) \to \mathcal{L}(z_1, ..., z_k),$$

где $z_1, ..., z_k$ — независимые случайные величины, имеющие пуассоновское распределение со средними $\mathbf{E}z_i = \theta/i$, $i = 1, ..., k$.

Следствие. *Статистика $\xi_k(n) = c_1(n) + ... + c_k(n)$ имеет при $n \to \infty$ и фиксированном $k \geqslant 1$ пуассоновское распределение с параметром $\theta\lambda_k$, где $\lambda_k = 1 + 1/2 + ... + 1/k$, т. е.*

$$\mathcal{L}(\xi_k(n)) \to \prod(\theta\lambda_k). \tag{2}$$

Соотношение (2) сводит исходную проблему к соответствующим статистическим задачам для пуассоновски распределенных

случайных величин с неизвестным параметром, решение которых достаточно хорошо известно. Используя известные результаты для пуассоновской модели, мы в § 2 даем решение задач точечного и доверительного оценивания параметра θ модели (1), а в § 3 приводим решение задач проверки статистических гипотез о параметре θ. В § 4 анализируется многовыборочный случай в асимптотической постановке, когда наблюдается большое число независимых подстановок в модели (1), а в § 5 решается задача проверки гипотезы однородности для этой ситуации.

§ 2. ОЦЕНИВАНИЕ ПАРАМЕТРА

Уточним постановку задачи. В общем случае мы будем предполагать, что имеется $N \geqslant 1$ независимых подстановок $A_1, ..., A_N$ с распределением (1) при одном и том же (но неизвестном) значении параметра θ и для каждой из них наблюдаются лишь циклы с длинами, не превосходящими заданного числа $k \geqslant 1$. Пусть $\xi_k^{(j)}(n)$ есть наблюдаемое число циклов для подстановки A_j, $j = 1, ..., N$. Набор статистик (выборка) $(\xi_k^{(j)}(n),\ j = 1, ..., N)$ — это наша исходная статистическая информация о модели (1), и по этим данным требуется оценить неизвестный параметр θ или те или иные функции от него $\tau(\theta)$; при этом мы считаем, что порядок подстановок n велик ($n \to \infty$), а параметр k фиксирован.

В этих условиях статистики $\xi_k^{(j)}(n)$, $j = 1, ..., N$, независимы и одинаково распределены в соответствии с (2), поэтому для суммарного числа наблюдаемых циклов

$$T(N, n) = \sum_{j=1}^{N} \xi_k^{(j)}(n) \tag{3}$$

будет выполняться предельное соотношение

$$\mathcal{L}(T(N, n)) \to \prod (N\theta\lambda_k). \tag{4}$$

Соотношение (4) означает, что в нашем случае речь идет об оценивании неизвестного параметра пуассоновской модели по наблюдению над статистикой $T(N, n)$.

Известно (см., например, [4, § 2.3]), что оптимальной (т. е. несмещенной с минимальной дисперсией) оценкой сходящегося при всех $\theta > 0$ степенного ряда $\tau(\theta) = \sum_{j \geqslant 0} a_j \theta^j$ по одному

наблюдению над случайной величиной X с пуассоновским распределением $\Pi(\theta)$ является статистика $\tau^* = \sum_{j\geqslant 0} a_j (X)_j$, где $(X)_j = X(X-1)\cdots(X-j+1)$.

На основании этого результата и соотношения (4) мы можем сформулировать следующее общее утверждение.

Утверждение 1. *При сформулированных условиях асимптотически оптимальной оценкой сходящегося при всех $\theta > 0$ степенного ряда $\tau(\theta) = \sum_{j\geqslant 0} a_j \theta^j$ в модели* (1) *является статистика*

$$\tau^* = \sum_{j\geqslant 0} a_j \frac{(T(N,n))_j}{(N\lambda_k)^j}. \tag{5}$$

Следствие. *Статистика $T(N,n)/(N\lambda_k)$ является асимптотически оптимальной оценкой параметра θ модели* (1) *по неполным данным* (*см.* (3)).

Утверждение 1 решает для модели (1) задачу оценивания по неполным данным параметрических функций, представимых степенными рядами.

Наряду с точечным оцениванием неизвестного параметра модели в математической статистике используется оценивание с помощью доверительных интервалов. При интервальном оценивании скалярного параметра θ ищут две такие статистики (функции от наблюдаемых данных) T_1 и T_2, что $T_1 < T_2$, для которых при заданном доверительном уровне $\gamma \in (0,1)$ для всех θ выполняется условие

$$\mathbf{P}_\theta\{T_1 < \theta < T_2\} \geqslant \gamma.$$

Такой случайный интервал (T_1, T_2) называют γ-доверительным интервалом для параметра θ. Длина этого интервала $T_2 - T_1$ характеризует точность локализации неизвестного параметра, а доверительный уровень γ — его «надежность»: вероятность невыполнения условия $\theta \in (T_1, T_2)$ не превышает $1-\gamma$. При этом, если имеется некоторая точечная оценка T параметра θ и ее функция распределения $F(t;\theta)$ непрерывна и монотонна по θ, то, определив из уравнений (относительно θ)

$$F(T;\theta) = (1-\gamma)/2, \quad 1 - F(T-0;\theta) = (1-\gamma)/2$$

два числа T_i, $i = 1,2$, $T_1 < T_2$, получаем так называемый центральный γ-доверительный интервал (T_1, T_2) для θ.

Известно (см., например, [4, пример 2.35]), что для модели Пуассона центральный θ-доверительный интервал по одному наблюдению X имеет вид

$$(T_1, T_2) = \left(\frac{1}{2}\chi^2_{(1-\gamma)/2,2X}, \frac{1}{2}\chi^2_{(1+\gamma)/2,2(X+1)}\right), \tag{6}$$

где $\chi^2_{p,s}$ есть p-квантиль распределения χ-квадрат с s степенями свободы.

Соотношения (4) и (6) позволяют сформулировать следующее правило доверительного оценивания параметра θ в рассматриваемых условиях.

Утверждение 2. *Центральный γ-доверительный интервал для параметра θ асимптотически имеет вид*

$$\left(\frac{1}{2N\lambda_k}\chi^2_{(1-\gamma)/2,2T(N,n)}, \frac{1}{2N\lambda_k}\chi^2_{(1+\gamma)/2,2(T(N,n)+1)}\right). \tag{7}$$

§ 3. ПРОВЕРКА ГИПОТЕЗ

В этом параграфе мы рассматриваем проблему проверки статистических гипотез о параметре θ модели (1) по наблюдению над статистикой $T(N,n)$, определенной в (3). Как уже отмечалось выше, соотношение (4) сводит нашу задачу к проверке соответствующих гипотез для пуассоновской модели с неизвестным параметром. Поэтому мы вначале приведем краткую сводку основных фактов теории пуассоновской модели, соответствующих нашей ситуации (см., например, [4, §§ 4.2, 4.4]).

Итак, пусть по одному наблюдению над случайной величиной X, имеющей распределение Пуассона $\Pi(\theta)$ с неизвестным параметром $\theta > 0$, требуется проверить простую гипотезу H_0: $\theta = \theta_0$ против альтернативы H_1: $\theta = \theta_1$. Пусть для определенности $\theta_1 > \theta_0$. Тогда (см., например, [5, задача 3.41 и ее решение]) для построения критерия Неймана–Пирсона при заданной вероятности ошибки первого рода α определяется целое число t_α условием

$$\alpha'' \equiv \sum_{m=t_\alpha+1}^{\infty} e^{-\theta_0}\frac{\theta_0^m}{m!} < \alpha \leqslant \sum_{m=t_\alpha}^{\infty} e^{-\theta_0}\frac{\theta_0^m}{m!} \equiv \alpha'. \tag{8}$$

Если здесь $\alpha = \alpha'$, то искомый критерий является нерандомизированным и имеет вид (задается критической областью)

$\{X \geqslant t_\alpha\} \equiv \mathcal{X}_\alpha^+$. Если же в (8) имеет место строгое неравенство $\alpha < \alpha'$, то критерий Неймана–Пирсона является рандомизированным и задается критической функцией

$$\varphi_\alpha(X) = \begin{cases} 1 & \text{при } X \geqslant t_\alpha + 1, \\ \dfrac{\alpha - \alpha''}{\alpha' - \alpha''} & \text{при } X = t_\alpha, \\ 0 & \text{при } X \leqslant t_\alpha - 1. \end{cases} \tag{9}$$

В любом случае мощность критерия вычисляется по формуле

$$W(\theta_1) = \sum_{m=t_\alpha+1}^{\infty} e^{-\theta_1} \frac{\theta_1^m}{m!} + (\alpha - \alpha'') \left(\frac{\theta_1}{\theta_0}\right)^{t_\alpha} e^{\theta_0 - \theta_1}. \tag{10}$$

В случае $\alpha < \alpha'$ вместо рандомизированного критерия можно также использовать нерандомизированные наиболее мощные критерии $\mathcal{X}_{\alpha'}^+ = \{X \geqslant t_\alpha\}$ и $\mathcal{X}_{\alpha''}^+ = \{X \geqslant t_\alpha + 1\}$ с уровнями значимости α' и α'' соответственно.

Построенный критерий Неймана–Пирсона является одновременно равномерно наиболее мощным критерием проверки гипотезы H_0: $\theta = \theta_0$ против сложной односторонней альтернативы H_1^+: $\theta > \theta_0$.

Вернемся к нашей исходной задаче. Для приложений наиболее важной является гипотеза о равновероятности подстановок, т. е. гипотеза H_0: $\theta = 1$. Альтернативами в данном случае могут быть H_1^+: $\theta > 1$ и H_1^-: $\theta < 1$ и двусторонняя сложная гипотеза $H_1 = H_1^- \cup H_1^+$: $\theta \neq 1$. Обсудим более детально задачу (H_0, H_1^+) проверки гипотезы H_0 против правосторонней альтернативы H_1^+.

Из вышеизложенного (в частности, из соотношения (4)) следует, что в задаче (H_0, H_1^+) существует асимптотический равномерно наиболее мощный критерий, конструируемый на основании соотношений (8)–(10), в которых следует положить $X = T(N, n)$, $\theta_0 = N\lambda_k$, $\theta_1 = N\lambda_k\theta$.

Таким образом, если при $\theta_0 = N\lambda_k$ в (8) имеет место равенство $\alpha = \alpha'$, то критерий имеет вид

$$\mathcal{X}_\alpha^+(n) = \{T(N, n) \geqslant t_\alpha\}, \tag{11}$$

если же $\alpha < \alpha'$, то либо можно использовать критерий

$$\mathcal{X}_{\alpha'}^+(n) = \{T(N, n) \geqslant t_\alpha\} \tag{12}$$

с несколько большим уровнем значимости α', либо критерий

$$\mathcal{X}^{+}_{\alpha''}(n) = \{T(N,n) \geqslant t_\alpha + 1\} \tag{13}$$

с несколько меньшим уровнем значимости α''.

Аналогично анализируется задача (H_0, H_1^-) с левосторонней альтернативой. Наконец, при двусторонней альтернативе критерий задается объединением двух односторонних критических областей, т. е. имеет вид (см. [4, § 4.4])

$$\mathcal{X}_\alpha(n) = \mathcal{X}^{+}_{\alpha_1}(n) \cup \mathcal{X}^{-}_{\alpha_2}(n), \quad \alpha_1 + \alpha_2 = \alpha. \tag{14}$$

§ 4. СХЕМА БОЛЬШИХ ВЫБОРОК

В этом и следующем параграфах рассмотрим ситуацию, когда число наблюдаемых подстановок N велико, т. е. схему больших выборок. Для этого случая из соотношений (3) и (4) следует, что для статистики

$$\tilde{T}(N,n) = T(N,n)/(N\lambda_k) \tag{15}$$

справедлива нормальная аппроксимация

$$\mathcal{L}(\tilde{T}(N,n)) \sim \mathcal{N}(\theta, \theta/(N\lambda_k)), \tag{16}$$

и тем самым статистические выводы о неизвестном параметре θ можно получать, используя соответствующие результаты для нормальной модели. С аналогичной ситуацией мы уже встречались в предыдущей работе (см. соотношение (10) в [3]). Можно сформулировать следующие утверждения.

1. Статистика $\tilde{T}(N,n)$ (см. (15)) является асимптотически несмещенной и асимптотически эффективной оценкой для параметра θ; такими же свойствами обладает статистика $\tau(\tilde{T}(N,n))$ как оценка функции $\tau(\theta)$ для любой дифференцируемой функции τ.

2. Асимптотическим γ-доверительным интервалом для параметра θ является интервал $\left(\tilde{T}(N,n) \pm z_\gamma \sqrt{\tilde{T}(N,n)/(N\lambda_k)}\right)$, где z_γ определяется уравнением $\Phi(z_\gamma) = (1+\gamma)/2$, $\Phi(z)$ — функция распределения стандартного нормального закона.

Аналогичный интервал для дифференцируемой параметрической функции $\tau(\theta)$ с непрерывной производной $\tau'(\theta)$ есть

$$\left(\tau(\tilde{T}(N,n)) \pm z_\gamma \tau'(\tilde{T}(N,n)) \sqrt{\tilde{T}(N,n)/(N\lambda_k)}\right). \tag{17}$$

3. Критерий уровня значимости α для проверки гипотезы H_0: $\theta = 1$ при левосторонней альтернативе H_1^-: $\theta < 1$ асимптотически при $N \to \infty$ задается критической областью

$$\mathcal{X}_\alpha^-(N,n) = \{\tilde{T}(N,n) < 1 - u_\alpha/\sqrt{N\lambda_k}\}, \quad u_\alpha = \Phi^{-1}(1-\alpha); \quad (18)$$

«пороговыми» для такого критерия являются альтернативы вида

$$H_{1N}^-\colon \theta = \theta_N = 1 - t/\sqrt{N\lambda_k}, \quad t > 0,$$

и мощность $W_N^-(\theta) = \mathbf{P}_\theta\{\mathcal{X}_\alpha^-(N,n)\}$ критерия при таких «близких» альтернативах стремится к $\Phi(t - u_\alpha)$.

Аналогично, в задаче $(H_0, H_1^+\colon \theta > 1)$ критерий с уровнем значимости α асимптотически имеет вид

$$\mathcal{X}_\alpha^+(N,n) = \{\tilde{T}(N,n) > 1 + u_\alpha/\sqrt{N\lambda_k}\}, \quad (19)$$

и его мощность при «близких» альтернативах

$$H_{1N}^+\colon \theta = \theta_N = 1 + t/\sqrt{N\lambda_k}, \quad t > 0,$$

стремится к $\Phi(t - u_\alpha)$.

Наконец, в задаче с двусторонней альтернативой $(H_0, H_1 = H_1^- \cup H_1^+)$ критерий асимптотически имеет вид

$$\mathcal{X}_\alpha(N,n) = \{|\tilde{T}(N,n) - 1| > u_{\alpha/2}/\sqrt{N\lambda_k}\}, \quad (20)$$

и его мощность при «близких» альтернативах

$$H_{1N}\colon \theta = \theta_N = 1 + t/\sqrt{N\lambda_k}, \quad t \neq 0,$$

стремится к $\Phi(t - u_{\alpha/2}) + \Phi(-t - u_{\alpha/2})$.

§ 5. ГИПОТЕЗА ОДНОРОДНОСТИ

Выше предполагалось, что в процессе наблюдения над подстановками $A_1,\ldots, A_N$ параметр θ модели (1) остается неизменным, что на самом деле является лишь гипотезой, которая сама должна быть подвергнута статистической проверке; такая гипотеза и называется гипотезой однородности. Здесь предлагается асимптотический (для больших выборок) вариант возможного критерия однородности.

Итак, пусть наблюдается большое число независимых подстановок $A_1,\ldots, A_N$ в модели (1), но каждая из них получается,

вообще говоря, при своем (но неизвестном) значении параметра θ. Требуется по информации

$$(\xi_k^{(j)}(n), j = 1, ..., N) \equiv (X_1, X_2, ..., X_N) \tag{21}$$

проверить гипотезу однородности H_0: $\theta_1 = \theta_2 = ... = \theta_N$ (здесь θ_j — значение параметра θ для подстановки A_j, $j = 1, ..., N$).

В силу предположения о независимости подстановок и соотношения (2) можно считать, что при гипотезе H_0 данные (21) представляют собой независимую выборку из пуассоновского распределения $\Pi(\theta\lambda_k)$ при некотором $\theta > 0$ (далее для краткости полагаем $\lambda = \theta\lambda_k$).

Хорошо известно (см., например, [4, глава 1]), что для больших выборок (в нашем случае — при больших N) совместное распределение любого конечного числа выборочных моментов асимптотически нормально, в частности, совместное распределение выборочного среднего

$$\bar{X} = \frac{1}{N}\sum_{j=1}^{N} X_j$$

и выборочной дисперсии

$$S^2 = \frac{1}{N}\sum_{j=1}^{N}(X_j - \bar{X})^2$$

асимптотически нормально. Отсюда и распределение любой их линейной комбинации асимптотически нормально. Чтобы построить удобную тестовую статистику в нашей задаче, основываясь на $\bar{X}$ и S^2, рассмотрим вначале линейную комбинацию (см. [5, задача 1.60])

$$\zeta_N = \sqrt{\frac{N-1}{2}}\frac{1}{\lambda}\left(\bar{X} - \frac{N}{N-1}S^2\right) \tag{22}$$

и установим, что при $N \to \infty$

$$\mathcal{L}(\zeta_N) \to \mathcal{N}(0, 1). \tag{23}$$

Для этого достаточно вычислить среднее и дисперсию разности $\bar{X} - \dfrac{N}{N-1}S^2$.

Используя известные формулы для центральных моментов μ_r пуассоновского распределения $\Pi(\lambda)$:

$$\mu_2 = \mu_3 = \lambda, \quad \mu_4 = \lambda + 3\lambda^2$$

и следующие формулы (см. [5, задача 1.27]):

$$\mathbf{D}\bar{X} = \frac{\mu_2}{N}, \quad \mathbf{E}S^2 = \frac{N-1}{N}\mu_2, \quad \operatorname{cov}(\bar{X}, S^2) = \frac{N-1}{N^2}\mu_3,$$
$$\mathbf{D}S^2 = \frac{(N-1)^2}{N^3}\left(\mu_4 - \frac{N-3}{N-1}\mu_2^2\right),$$

непосредственными вычислениями получаем, что в нашем случае

$$\mathbf{E}\left(\bar{X} - \frac{N-1}{N}S^2\right) = 0,$$

$$\mathbf{D}\left(\bar{X} - \frac{N}{N-1}S^2\right) = \mathbf{D}\bar{X} + \left(\frac{N}{N-1}\right)^2 \mathbf{D}S^2 - \frac{2N}{N-1}\operatorname{cov}(\bar{X}, S^2) =$$
$$= \frac{\mu_2}{N} + \frac{1}{N}\left(\mu_4 - \frac{N-3}{N-1}\mu_2^2\right) - \frac{2\mu_3}{N} = \frac{2\lambda^2}{N-1}.$$

Отсюда следует (23).

Величина ζ_N еще не может служить тестовой статистикой для проверки гипотезы H_0, так как она зависит от неизвестного параметра $\lambda = \theta\lambda_k$. Но если воспользоваться тем фактом, что $\bar{X}/\lambda \xrightarrow{\mathrm{P}} 1$ при $N \to \infty$, и заменить в представлении (22) параметр λ на $\bar{X}$, то мы придем к статистике

$$T_N = \zeta_N \frac{\lambda}{\bar{X}} = \sqrt{\frac{N-1}{2}}\frac{1}{\bar{X}}\left(\bar{X} - \frac{N}{N-1}S^2\right), \tag{24}$$

также асимптотически нормальной $\mathcal{N}(0,1)$, если справедлива гипотеза H_0. Тем самым, основываясь на статистике T_N, можно предложить следующий критерий согласия для гипотезы H_0: при заданном уровне значимости α гипотеза однородности H_0 отвергается тогда и только тогда, когда $|T_N| \geqslant u_{\alpha/2}$, где $\Phi(-u_{\alpha/2}) = \alpha/2$.

Итак, если мы находимся в ситуации, описанной в § 4, то прежде чем применять изложенные в нем алгоритмы, следует предварительно провести тестирование данных (21) на однородность, используя статистику (24) и описанный критерий согла-

сия, и лишь если гипотеза однородности H_0 принимается (если $|T_N| < u_{\alpha/2}$), то эти алгоритмы можно использовать.

СПИСОК ЛИТЕРАТУРЫ

1. *Ewens W.J.* The sampling theory of selectively neutral alleles. — Theor. Popul. Biology, 1972, v. 3, p. 87–112.
2. *Arratia R.A., Barbour A.D., Tavare S.* Poisson process approximations for the Ewens Sampling Formula. — Ann. Appl. Probab., 1922, v. 2, p. 519–535.
3. *Ивченко Г.И., Медведев Ю.И.* Статистика параметрической модели случайных подстановок. — В сб.: Труды по дискретной математике. Т. 8. — М.: ФИЗМАТЛИТ, 2004, с. 116–127.
4. *Ивченко Г.И., Медведев Ю.И.* Математическая статистика. — М.: Высшая школа, 1992.
5. *Ивченко Г.И., Медведев Ю.И., Чистяков А.В.* Сборник задач по математической статистике. — М.: Высшая школа, 1989.