

На рис. 3 Приведены графики зависимости средней длительности пребывания на виртуальном устройстве  $I$  сообщений различных потоков.

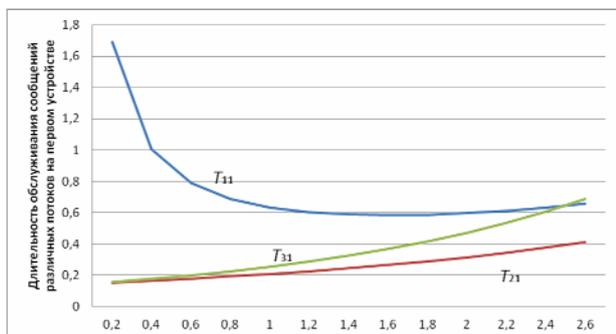


Рис. 3. Зависимость средней длительности пребывания сообщений различных потоков на виртуальном устройстве  $I$

### Выводы

Приведенные в статье результаты дают возможность вычислять характеристики системы с тиражи-

рованием обрабатываемых сообщений в зависимости от параметров тиражирования, что позволяет также решать задачи выбора оптимальных значений параметров тиражирования при проектировании подобных систем.

Возможно развитие результатов для случая, когда длительность обработки на виртуальных устройствах зависит от количества виртуальных устройств.

### Библиографические ссылки

1. Клейнрок Л. Вычислительные системы с очередями : пер. с англ. / под ред. Б. С. Цыбакова. – М. : Мир, 1979. – 600 с.
2. Никитин Е. В., Саксонов Е. А. Управление потоками данных в многосерверных системах обработки информации // Информатика и системы управления. – 2010. – № 25. – С. 3–9.
3. Далингер Я. М. Модель системы обработки с тиражированием поступающих сообщений // Вестник ИжГТУ. – 2014. – № 2(62). – С. 155–159.

Ya. M. Dalinger, PhD in Engineering, Associate Professor, Saint-Petersburg State University of Civil Aviation

### Analysis of System with Message Copying While Processing

*A mathematical model for calculation of the features of the system with copying the maintained messages is presented.*

**Keywords:** flow of messages, copying of messages, information processing, mathematical model.

Получено 11.08.2014

УДК 519.23

**М. В. Радионова**, кандидат физико-математических наук, Национальный исследовательский университет «Высшая школа экономики», Пермь

**В. В. Чичагов**, кандидат физико-математических наук, Пермский государственный национальный исследовательский университет

## ОБ ОДНОМ КЛАССЕ КРИТЕРИЕВ СОГЛАСИЯ ТИПА ХИ-КВАДРАТ\*

*Предложен новый класс асимптотических критериев для проверки гипотезы о виде распределения, принадлежащего однопараметрическому экспоненциальному семейству. Каждый критерий основывается на заданном множестве параметрических функций, допускающих несмещенную оценку. На примере проверки гипотезы о распределении Пуассона против гипотезы о геометрическом распределении проведено сравнение предложенного критерия с критерием, основывающимся на обобщенном методе моментов.*

**Ключевые слова:** экспоненциальное семейство, несмещенная оценка, критерий согласия, мощность.

**Р**ешение задачи проверки гипотезы о виде распределения случайной величины  $\xi$  по выборке  $X_1, \dots, X_n$ , элементами которой являются независимые случайные величины, имеющие то же распределение, что и  $\xi$ , является одной из фундаментальных проблем математической и прикладной статистики. Проверка этой гипотезы осуществляется обычно с помощью критерия согласия, который может основываться на близости эмпирической и теоретической функции

распределения, на близости плотности распределения и гистограммы, на характеристике распределения свойствами определенных статистик выборки. В эконометрических приложениях, связанных с идентификацией распределения  $\xi$ , нашли широкое применение статистические критерии моментных условий (см., например, [1]). Основу каждого такого  $J$ -критерия, так мы будем его называть, составляет статистика вальдовского типа, представляемая обычно следующим образом [2]:

$$J = n \left( \frac{1}{n} \sum_{i=1}^n \mathbf{m}^\top(X_i, \hat{\theta}) \right) \tilde{Q}_{\mathbf{m}, \mathbf{m}}^{-1} \left( \frac{1}{n} \sum_{i=1}^n \mathbf{m}(X_i, \hat{\theta}) \right), \quad (1)$$

где  $\mathbf{m}(\xi, \theta)$  – заданная векторная функция моментов размерности  $L \times 1$  такая, что  $\mathbf{M}[\mathbf{m}(\xi, \theta)] = \mathbf{0}$  и  $L \geq k$ ;  $\tilde{Q}_{\mathbf{m}, \mathbf{m}}$  – некоторая состоятельная оценка ковариационной матрицы  $Q_{\mathbf{m}, \mathbf{m}} = \mathbf{M}[\mathbf{m}(\xi, \theta), \mathbf{m}^\top(\xi, \theta)]$  размерности  $L \times L$ , а  $\hat{\theta}$  – оценка неизвестного параметра  $\theta$  размерности  $k \times 1$  по обобщенному методу моментов (ОММ-оценка):

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \left( \frac{1}{n} \sum_{i=1}^n \mathbf{m}^\top(X_i, \theta) \right) \tilde{Q}_{\mathbf{m}, \mathbf{m}}^{-1} \left( \frac{1}{n} \sum_{i=1}^n \mathbf{m}(X_i, \theta) \right). \quad (2)$$

Известно, что  $J$ -статистика (1) в условиях справедливости гипотезы  $H_0 : \mathbf{M}[\mathbf{m}(\xi, \theta)] = \mathbf{0}$  и невырожденности ковариационной матрицы  $Q_{\mathbf{m}, \mathbf{m}}$  имеет при  $n \rightarrow \infty$  асимптотическое распределение  $\chi^2$  с  $L - k$  степенями свободы.

В соответствии с  $J$ -критерием асимптотического уровня значимости  $\alpha$  гипотеза  $H_0$  принимается, если  $J < x_{1-\alpha}[\chi_{L-k}^2]$ , где  $x_p[\chi_\nu^2]$  – квантиль уровня  $p$  распределения  $\chi^2$  с  $\nu$  степенями свободы.

Статистические критерии моментных условий сводят проверку гипотезы о виде распределения к проверке гипотезы о том, что наблюдаемая случайная величина удовлетворяет заданным моментным ограничениям. Отметим также, что применение  $J$ -критерия сопряжено с определенными трудностями вычислительного характера, поскольку требует решения экстремальной задачи (2).

В данной работе предлагается иной подход к построению статистических критериев, учитывающий моментные ограничения при следующих предположениях.

**A1.** Распределение вероятностей случайной величины  $\xi$  принадлежит естественному однопараметрическому экспоненциальному семейству (см., например, [3]), определяемому выражением

$$f(x; a) = \exp\{\Phi(a)T(x) - \kappa(\Phi(a)) + d(x)\}, \quad x \in X_0 \subset \mathbf{R}. \quad (3)$$

Здесь  $f(x; a)$  – плотность распределения случайной величины  $\xi$  относительно меры  $\mu(x)$ , являющейся либо мерой Лебега, если  $\xi$  имеет абсолютно непрерывное распределение, либо считающей мерой, когда  $\xi$  имеет решетчатое распределение;  $d(x), T(x)$  – известные борелевские функции;  $\nu = \Phi(a)$  – заданная непрерывно дифференцируемая функция неизвестного параметра  $a = \mathbf{MT}(\xi) = \kappa'(\Phi(a))$ ,  $a \in \mathbf{A} \subset \mathbf{R}$ ;  $\kappa[\nu]$  – кумулянтное преобразование рас-

пределения  $\xi$ ;  $X_0$  – носитель распределения  $\xi$ , не зависящий от параметра  $a$ ; при этом  $b^2 = \mathbf{DT}(\xi) > 0$ .

**A2.** Если  $\mu(x)$  – мера Лебега, то для каждого  $a \in \mathbf{A}$  существует  $n_0 \in \mathbf{N}$  такое, что нормированная сумма  $Z_n = \frac{S_n - na}{b\sqrt{n}}$ , где  $S_n = \sum_{i=1}^n T(X_i)$ , имеет непрерывную ограниченную плотность. Если  $\mu(x)$  – считающая мера, то носитель  $X_0$  не содержится ни в какой подрешетке решетки  $\mathbf{Z}$ .

Введем дополнительно следующие обозначения:

$h_1(T), h_2(T), \dots, h_L(T)$  – заданный набор пробных функций,  $T = T(\xi)$ ;

$G_j \equiv G_j(a) \equiv \mathbf{M}h_j(T(\xi))$  – математическое ожидание пробной функции  $h_j(T)$ ;

$$G_j^* = \frac{1}{n} \sum_{i=1}^n h_j(T(X_i));$$

$\hat{G}_j$  – несмещенная оценка с равномерно минимальной дисперсией (НОРМД) функции  $G_j(a)$ ;

$$\mathbf{G} = (G_1, G_2, \dots, G_L)^\top, \quad \mathbf{G}^* = (G_1^*, G_2^*, \dots, G_L^*)^\top, \quad \hat{\mathbf{G}} = (\hat{G}_1, \hat{G}_2, \dots, \hat{G}_L)^\top;$$

$\mathbf{I}(A)$  – индикатор события  $A$ ;

$\mathbf{cov}(\mathbf{V})$  – ковариационная матрица случайного вектора  $\mathbf{V}$ .

Далее в работе изучается поведение статистики

$$\chi_G = n(\mathbf{G}^* - \hat{\mathbf{G}})^\top \tilde{\Sigma}^{-1}(\mathbf{G}^* - \hat{\mathbf{G}}), \quad (4)$$

которая и предлагается в качестве альтернативы статистике (1). Здесь  $\tilde{\Sigma}$  – некоторая состоятельная оценка асимптотической ковариационной матрицы вектора  $\sqrt{n}(\mathbf{G}^* - \hat{\mathbf{G}})$ .

Отметим, что при выборе в качестве пробной функции  $h_j(T)$  индикаторной функции  $h_j(T) = \mathbf{I}(T(\xi) \in J(j))$ , где  $J(j), j = \overline{1, L+1}$ , – атомы разбиения числовой прямой  $\mathbf{R}$ , статистика (4) представляет собой обобщение статистики хи-квадрат критерия Пирсона. Этот случай рассмотрен в работе [4].

**Критерий согласия  $\chi_G$**

Приведем теоретическое обоснование критерия, основанного на статистике  $\chi_G$  для пробной функции вида  $h_j(T) = T^{j+1}$ . В этом случае  $G_j = \mathbf{MT}^{j+1}(\xi)$ ,

$$G_j^* = \frac{1}{n} \sum_{i=1}^n T^{j+1}(X_i), \quad \hat{G}_j = \mathbf{M}[T^{j+1}(X_1) | S_n].$$

Чтобы сформулировать основное утверждение, воспользуемся определением, заимствованным из [5, п. 3.5].

**Определение.** Случайные величины  $Y_1, \dots, Y_k$  существенно линейно зависимы, если существует такой набор констант  $c_1, \dots, c_k$ , каждая из которых отлична от нуля, что  $c_1 Y_1 + \dots + c_k Y_k$  является вырожденной случайной величиной.

**Теорема 1.** Пусть выполнены предположения **A1-A2**,  $h_j(T) = T^{j+1}$ ,  $j = \overline{1, L}$ , случайные величины  $T(\xi), T^2(\xi), \dots, T^{L+1}(\xi)$  не являются существенно линейно зависимыми. Тогда при  $n \rightarrow \infty$

1) последовательность случайных величин  $\mathbf{Y}_n = \sqrt{n}(\mathbf{G}^* - \hat{\mathbf{G}})$  асимптотически нормальна с нулевым вектором средних и невырожденной ковариационной матрицей  $\Sigma = \Sigma_0 - [\Phi'(a)]^{-1} \mathbf{G}'(\mathbf{G}')^T$ , где

$$\Sigma_0 = \left\| \mathbf{K}(T^{i+1}, T^{j+1}) \right\|_{i,j=1}^L, \quad \mathbf{G}' = (G'_1, \dots, G'_L)^T;$$

2) асимптотическим распределением последовательности случайных величин (4) является распределение  $\chi^2$  с  $L$  степенями свободы.

**Доказательство.** Согласно [3], если распределение случайной величины  $\xi$  принадлежит экспоненциальному семейству вида (3), то случайная величина  $T(\xi)$  имеет моменты любого порядка. Поэтому при каждом  $j = 0, 1, 2, \dots$  функцию

$$G_j = \int_{X_0} T^{j+1}(x) \times \exp\{\Phi(a)T(x) - \kappa(\Phi(a)) + d(x)\} d\mu(x) \quad (5)$$

можно дифференцировать по  $a$  любое число раз. Дифференцируя (5), получим

$$G'_j \equiv \frac{dG_j}{da} = \Phi'(a)(G_{j+1} - G_j \cdot a) = \Phi'(a)K(T, T^{j+1}). \quad (6)$$

Используя следствие 2 работы [6] и (6), нетрудно при  $n \rightarrow \infty$  получить следующие стохастические разложения:

$$\begin{aligned} \hat{G}_j - G_j &= \frac{G'_j Z_n}{\sqrt{n\Phi'(a)}} + \mathbf{O}_P\left(\frac{1}{n}\right) = \\ &= \sqrt{\frac{\Phi'(a)}{n}} K(T, T^{j+1}) Z_n + \mathbf{O}_P\left(\frac{1}{n}\right); \end{aligned} \quad (7)$$

$$(\hat{G}_i - G_i)(\hat{G}_j - G_j) = \frac{1}{n\Phi'(a)} G'_i G'_j Z_n^2 + \mathbf{O}_P\left(\frac{1}{n}\right). \quad (8)$$

Разложение (7) позволяет представить вектор  $\mathbf{Y}_n$  в следующем виде:

$$\mathbf{Y}_n = \sqrt{n}(\mathbf{G}^* - \mathbf{G}) + (\Phi'(a))^{-1/2} \mathbf{G}' Z_n + \mathbf{O}_P(n^{-1/2}). \quad (9)$$

Имея в виду, что  $\mathbf{M}[G'_i \hat{G}_j | S_n] = \hat{G}_i \hat{G}_j$ , установим связь между ковариационными матрицами векторов  $\mathbf{Y}_n$ ,  $\mathbf{G}^*$  и  $\hat{\mathbf{G}}$ :

$$\begin{aligned} \Sigma_n = \text{cov}(\mathbf{Y}_n) &= n\mathbf{M}[(\mathbf{G}^* - \hat{\mathbf{G}})(\mathbf{G}^* - \hat{\mathbf{G}})^T] = \\ &= n\mathbf{M}\{\mathbf{M}[(\mathbf{G}^* - \hat{\mathbf{G}})(\mathbf{G}^* - \hat{\mathbf{G}})^T | S_n]\} = \\ &= n\mathbf{M}[(\mathbf{G}^* - \mathbf{G})(\mathbf{G}^* - \mathbf{G})^T] - n\mathbf{M}[(\hat{\mathbf{G}} - \mathbf{G})(\hat{\mathbf{G}} - \mathbf{G})^T] = \\ &= n\text{cov}(\mathbf{G}^*) - n\text{cov}(\hat{\mathbf{G}}). \end{aligned}$$

При этом элементы матрицы  $\Sigma_0 = n\text{cov}(\mathbf{G}^*)$  определяются выражениями

$$\begin{aligned} n\mathbf{K}(G_i^*, G_j^*) &= \frac{1}{n} \mathbf{K}\left(\sum_{\ell=1}^n T^{i+1}(X_\ell), \sum_{m=1}^n T^{j+1}(X_m)\right) = \\ &= \frac{1}{n} \sum_{\ell=1}^n \mathbf{K}(T^{i+1}(X_\ell), T^{j+1}(X_\ell)) = \mathbf{K}(T^{i+1}, T^{j+1}). \end{aligned}$$

Поскольку  $\mathbf{M}\mathbf{G}^* = \mathbf{M}\hat{\mathbf{G}} = \mathbf{G}$ , то  $\mathbf{M}\mathbf{Y}_n = \mathbf{0}$ . Поэтому по теореме 20 из [7] при любом  $k \geq 2$  верна оценка  $\mathbf{M}[\sqrt{n}(\mathbf{G}^* - \mathbf{G}_j)]^k \leq C_{jk}$ , где  $C_{jk}$  – положительная константа, зависящая только от  $k$  и  $j$ . Отсюда при любом  $k \geq 2$  по теореме Рао – Блэквелла – Колмогорова

$$\mathbf{M}[\sqrt{n}(\hat{G}_j - G_j)]^k \leq \mathbf{M}[\sqrt{n}(G_j^* - G_j)]^k \leq C_{jk},$$

а в силу неравенства Коши – Буняковского для любого  $n$

$$|n\mathbf{K}(\hat{G}_i, \hat{G}_j)| \leq \sqrt{n\mathbf{D}[\hat{G}_i]} \cdot \sqrt{n\mathbf{D}[\hat{G}_j]} \leq \sqrt{C_{i2}C_{j2}}. \quad (10)$$

Поскольку при  $n \rightarrow \infty$  последовательность случайных величин  $\{Z_n\}$  сходится к случайной величине  $\eta$ , имеющей стандартное нормальное распределение, то в силу (7), (8) и равномерной ограниченности ковариаций (10) при  $n \rightarrow \infty$  по теореме непрерывности для моментов [8] имеем

$$[\Phi'(a)]^{-1} n\mathbf{K}(\hat{G}_i, \hat{G}_j) \rightarrow \mathbf{M}[G'_i G'_j \eta^2] = G'_i G'_j.$$

$$\text{Отсюда} \quad \Sigma = \lim_{n \rightarrow \infty} \Sigma_n = \Sigma_0 - [\Phi'(a)]^{-1} \mathbf{G}'(\mathbf{G}')^T,$$

и из (9) в силу многомерной центральной предельной теоремы для независимых одинаково распределенных векторов следует справедливость первого утверждения теоремы. Для доказательства второго утверждения теоремы осталось показать, что  $\Sigma$  – невырожденная матрица.

По формуле обращения матрицы при малоранговой модификации (см., например, [9]) имеем соотношение

$$\Sigma^{-1} = \Sigma_0^{-1} + \frac{\Sigma_0^{-1} \mathbf{G}'(\mathbf{G}')^T \Sigma_0^{-1}}{\Phi'(a) - (\mathbf{G}')^T \Sigma_0^{-1} \mathbf{G}'}$$

По утверждению 3.5.2 работы [5] из того факта, что случайные величины  $T(\xi), T^2(\xi), \dots, T^{L+1}(\xi)$  не являются собственно линейно зависимыми, следует положительная определенность их ковариационной матрицы, а с ней и положительная определенность матрицы  $\Sigma_0$ . В силу (6) выражение  $[\Phi'(a)]^{-1}(\mathbf{G}')^T \Sigma_0^{-1} \mathbf{G}'$  представляет собой множественный коэффициент корреляции между случайной величиной  $T(\xi)$  и случайными величинами  $T^2(\xi), \dots, T^{L+1}(\xi)$ , который в условиях отсутствия их линейной зависимости не может быть равен 1. Отсюда следует невырожденность матрицы  $\Sigma$ , что завершает доказательство теоремы.

*Примечание.* Так как оценки максимального правдоподобия и НОРМД в условиях **A1-A2** асимптотически эквивалентны [10], то утверждение теоремы сохранит силу, если вместо НОРМД  $\hat{G}_j$  функции  $G_j, j = \overline{1, L}$  использовать ее оценку максимального правдоподобия.

Опираясь на утверждение теоремы, определим  $\chi_G$ -критерий для проверки нулевой гипотезы

$$H_0 : \mathbf{M}T^{j+1}(\xi_0) = \mathbf{M}T^{j+1}(\xi), j = \overline{1, L}. \quad (11)$$

Здесь  $\xi_0$  – гипотетическая случайная величина, имеющая распределение из семейства (3). Нулевую гипотезу (11) будем принимать, если верно неравенство

$$\chi_G = n(\mathbf{G}^* - \hat{\mathbf{G}})^T \hat{\Sigma}^{-1}(\mathbf{G}^* - \hat{\mathbf{G}}) < x_{1-\alpha}[\chi_L^2]. \quad (12)$$

Следствие.  $\chi_G$ -критерий, определяемый неравенством (12), имеет при  $n \rightarrow \infty$  асимптотический уровень значимости  $\alpha$ .

**Сравнение по мощности  $\chi_G$ -критерия с  $J$ -критерием методом статистического моделирования**

Сопоставим по мощности  $\chi_G$ -критерий с  $J$ -критерием на примере проверки гипотезы о том, что  $\xi$  имеет распределение Пуассона

$$\mathbf{P}(\xi = x) \equiv f_1(x; a) = \frac{a^x}{x!} e^{-a} = \exp\{x \ln a - a - \ln x!\}, x = 0, 1, \dots$$

с параметром  $a > 0$  против альтернативы, что  $\xi$  имеет геометрическое распределение с неизвестным параметром  $p = (1+a)^{-1}$ :

$$\mathbf{P}(\xi = x) \equiv f_2(x; p) = p \cdot (1-p)^x, x = 0, 1, \dots,$$

используя пробные функции  $h_1(T) = T^2$  и  $h_2(T) = T^3$ .

В данном примере  $L = 2$ , число неизвестных параметров распределения  $k = 1$ . Значение параметра  $p = (1+a)^{-1}$  определяет геометрическое распределение, имеющее наименьшее расстояние Кульбака – Лейблера до распределения Пуассона.

Исследование мощности критериев осуществлялось при значениях параметра  $a$  от 0,1 до 1 с шагом 0,05 с использованием метода статистического моделирования по следующей схеме:

Моделируем  $N = 50000$  выборок  $X_1, \dots, X_n$  объема  $n$ , имеющих геометрическое распределение с параметром  $p = (1+a)^{-1}$ .

Подсчитываем  $\mu[\chi_G]$  – количество имитаций, в которых оказалось неверным неравенство (12), и  $\mu[J]$  – количество имитаций, в которых оказалось неверным неравенство

$$\frac{1}{n} \sum_{i=1}^n \mathbf{m}^T(X_i, \hat{\theta}) \tilde{Q}_{\mathbf{m}, \mathbf{m}}^{-1} \sum_{i=1}^n \mathbf{m}(X_i, \hat{\theta}) < x_{1-\alpha}[\chi_{L-k}^2].$$

Вычисляем эмпирическую мощность критериев как относительные частоты  $\mu[\chi_G]/N$  и  $\mu[J]/N$ .

При проведении расчетов с помощью  $\chi_G$ -критерия использовались следующие формулы (выражения несмещенных оценок заимствованы из [11]):

$$\Phi(a) = \ln a, T(x) = x, S_n = \sum_{i=1}^n X_i;$$

$$G_1(a) = \mathbf{M}[\xi^2] = a^2 + a;$$

$$G_2(a) = \mathbf{M}[\xi^3] = a^3 - 3a^2 + a;$$

$$G_1^* = \frac{1}{n} \sum_{i=1}^n X_i^2, G_2^* = \frac{1}{n} \sum_{i=1}^n X_i^3;$$

$$\hat{G}_2 = \frac{S_n(S_n - 1)}{n^2} + \frac{S_n}{n};$$

$$\hat{G}_3 = \frac{S_n(S_n - 1)(S_n - 2)}{n^3} + \frac{3S_n(S_n - 1)}{n^2} + \frac{S_n}{n};$$

$$\Sigma = \begin{pmatrix} 2a^2 & 6a^2 + 6a^3 \\ 6a^2 + 6a^3 & 18a^2 + 42a^3 + 18a^4 \end{pmatrix}.$$

Оценка матрицы  $\hat{\Sigma}$  получена из  $\Sigma$  заменой параметра  $a$  на его оценку  $a_* = \frac{S_n}{n}$ .

При проведении расчетов с помощью  $J$ -критерия использовались следующие соотношения:

$$\mathbf{m}(x, a) = \begin{pmatrix} x^2 - a^2 - a \\ x^3 - a^3 + 3a^2 - a \end{pmatrix}; \overline{X^j} = \frac{1}{n} \sum_{i=1}^n X_i^j;$$

$$\hat{Q}_{m,m} = \begin{pmatrix} \overline{X^4} - 2\overline{X^2}(a_*^2 + a_*) + (a_*^2 + a_*)^2 & \overline{X^5} - 2\overline{X^3}(a_*^2 + a_*) - \overline{X^2}(a_*^3 - 3a_*^2 + a_*) \\ \overline{X^5} - 2\overline{X^3}(a_*^2 + a_*) - \overline{X^2}(a_*^3 - 3a_*^2 + a_*) & \overline{X^6} - 2\overline{X^3}(a_*^3 - 3a_*^2 + a_*) + (a_*^3 - 3a_*^2 + a_*)^2 \end{pmatrix}.$$

В табл.1 представлены результаты расчета мощности критериев при разных объемах выборки  $n$ , параметра  $a$  и заданном уровне значимости  $\alpha = 0,05$ . Расчеты производились с использованием пакета *Wolfram Mathematica 9*.

Анализ табл. 1 показывает, что предложенный критерий оказался более мощным при различении альтернатив «распределение Пуассона – геометриче-

ское распределение» при всех значениях входных параметров. Максимальная разность мощности критериев варьируется от 0,033 при  $n = 500$  до 0,079 при  $n = 20$ , минимальная – от 0,000 при объемах  $n = 300$  и  $n = 500$  до 0,019 при  $n = 100$ . Значения мощности, которым соответствуют минимальные и максимальные их разности, при каждом объеме выборки выделены в таблице жирным шрифтом.

Таблица 1. Мощность  $\chi_G$ - и  $J$ - критерия

$a$	$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 300$		$n = 500$	
	$\chi_G$	$J$										
0,10	<b>0,043</b>	<b>0,037</b>	0,123	0,074	0,142	0,117	0,216	0,184	<b>0,275</b>	<b>0,236</b>	<b>0,375</b>	<b>0,342</b>
0,15	0,067	0,055	0,140	0,110	0,217	0,191	0,321	0,303	0,384	0,380	0,515	0,502
0,20	0,104	0,079	0,172	0,154	<b>0,274</b>	<b>0,254</b>	0,401	0,390	0,524	0,507	0,697	0,676
0,25	0,125	0,096	0,208	0,190	0,345	0,320	0,527	0,493	0,673	0,635	0,848	0,816
0,30	0,148	0,116	<b>0,248</b>	<b>0,231</b>	0,414	0,383	<b>0,651</b>	<b>0,603</b>	0,786	0,747	0,926	0,901
0,35	0,172	0,137	0,301	0,279	0,498	0,455	0,742	0,697	0,869	0,833	0,972	0,956
0,40	0,207	0,166	0,351	0,318	0,575	0,529	0,822	0,776	0,923	0,889	0,990	0,981
0,45	0,219	0,179	0,408	0,374	0,647	0,591	0,876	0,833	0,963	0,937	0,997	0,991
0,50	0,254	0,209	0,466	0,420	0,705	0,653	0,917	0,884	0,981	0,961	0,999	0,997
0,55	0,281	0,225	0,514	0,462	0,756	0,700	0,951	0,916	0,991	0,980	1,000	0,999
0,60	0,308	0,252	0,560	0,505	0,813	0,758	0,971	0,946	0,997	0,990	1,000	1,000
0,65	0,338	0,280	0,612	0,552	<b>0,850</b>	<b>0,792</b>	0,984	0,964	0,998	0,994	1,000	1,000
0,70	0,369	0,309	0,652	0,588	0,885	0,831	0,990	0,978	0,999	0,997	1,000	1,000
0,75	0,395	0,328	0,692	0,625	0,914	0,864	0,995	0,985	1,000	0,998	1,000	1,000
0,80	0,418	0,356	0,727	0,659	0,933	0,889	0,997	0,990	1,000	0,999	1,000	1,000
0,85	0,436	0,364	<b>0,769</b>	<b>0,697</b>	0,952	0,912	0,999	0,994	1,000	1,000	1,000	1,000
0,90	<b>0,469</b>	<b>0,390</b>	0,791	0,723	0,959	0,928	1,000	0,998	1,000	1,000	1,000	1,000
0,95	0,503	0,427	0,822	0,758	0,974	0,949	1,000	0,998	1,000	1,000	1,000	1,000
1,00	0,525	0,449	0,846	0,781	0,982	0,955	<b>1,000</b>	<b>0,999</b>	1,000	1,000	1,000	1,000

Одной из возможных причин того, что  $\chi_G$ -критерий оказался более мощным, является то обстоятельство, что реальный уровень значимости оказался у него выше, чем у  $J$ -критерия (см. табл. 2), и ближе к заданному значению 0,05, что опять же говорит в пользу предложенного критерия.

Таблица 2. Реальный уровень значимости критериев

Объем выборки	$n = 20$	$n = 50$	$n = 100$	$n = 200$	$n = 300$	$n = 500$
$\chi_G$ -критерий	0,033	0,040	0,044	0,046	0,046	0,048
$J$ -критерий	0,023	0,039	0,041	0,044	0,044	0,047

**Библиографические ссылки**

- Hansen L. P. Large Sample Properties of Generalized Method of Moments Estimators // *Econometrica*. – 1982. – Vol. 50. – No. 4. – Pp. 1029–1054.
- Anatolyev S. GMM, GEL, serial correlation and asymptotic bias // *Econometrica*. – 2005. – Vol. 73. – No. 3. – Pp. 983–1002.
- Brown L. D. Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory //

Lecture Notes-Monograph Series. – Vol. 9. – Institute of Mathematical Statistics, 1986.

- Чичагов В. В. Поведение статистики хи-квадрат с использованием несмещенных оценок в случае однопараметрического распределения из экспоненциального семейства // *Статистические методы оценивания и проверки гипотез* : межвуз. сб. науч. тр. – Пермь, 2006. – С. 78–89.

- Уилкс С. Математическая статистика. – М. : Наука, 1967. – 632 с.

- Чичагов В. В. Стохастические разложения несмещенных оценок в случае однопараметрического экспоненциального распределения // *Информатика и ее приложения*. – 2008. – Т. 2. – Вып. 2. – С. 62–70.

- Петров В. В. Предельные теоремы для сумм независимых случайных величин. – М. : Наука, 1987. – 320 с.

- Боровков А. А. Математическая статистика: оценка параметров и проверка гипотез. – М. : Наука, 1984. – 472 с.

- Хорн Р., Джонсон Ч. Матричный анализ. – М. : Мир, 1989. – 666 с.

- Portnoy S. Asymptotic efficiency of minimum variance unbiased estimators // *Annals of Statistics*. – 1977. – Vol. 5. – No. 3. – Pp. 522–529.

- Воинов В. Г., Никулин М. С. Несмещенные оценки и их применение. – М. : Наука, 1989. – 440 с.

## On a New Class of Goodness-of-Fit Tests as a Type of Chi-Square

A new class of asymptotic tests is suggested for testing the hypothesis of distribution belonging to the one-parameter exponential family. Each criterion is based on a given set of parametric functions that allow an unbiased estimate. We compared the proposed criterion with the criterion based on the generalized method of moments on the example of testing the hypothesis of the Poisson distribution against the hypothesis of the geometric distribution.

**Keywords:** exponential family, unbiased estimator, goodness-of-fit test, power.

Получено 03.09.2014

УДК 51-72:531

С. А. Груздь, Ижевский государственный технический университет имени М. Т. Калашникова

## ВЛИЯНИЕ СПОСОБОВ РОСТА КЛАСТЕРОВ НА СКОРОСТЬ КОНДЕНСАЦИИ

Предложена математическая модель, рассматривающая процесс образования кластера за счет столкновения групп молекул различных размеров. Варианты размеров, получаемых при образовании зародышей, представлены в виде рядов. Устанавливается зависимость термодинамических параметров системы от способов образования кластеров в насыщенных парах.

**Ключевые слова:** моделирование, гомогенная конденсация, числовые ряды, образование зародыша.

### Теория быстрой коагуляции для групповых столкновений

Согласно теории Смолуховского [1], если любое столкновение частиц приводит к их слипанию, говорят о быстрой коагуляции, что наблюдается при стремительном охлаждении. Скорость коагуляции является функцией счетной концентрации частиц и интенсивности броуновского движения, характеризуемого коэффициентом диффузии. Таким образом, при конденсации важно знать, какой набор агрегатов частиц будет присутствовать в системе.

Выражая скорость коагуляции критического кластера через уменьшение числа частиц в единице объема за единицу времени, записываем:

$$I_{кр} = kn_{i-2}n_{i-1}, \quad (1)$$

где  $I_{кр}$  – скорость образования критического кластера;  $n$  – частичная концентрация кластеров, из которых произошло образование критического кластера;  $k$  – постоянная скорости коагуляции, характеризующая вероятность сближения.

Термодинамические параметры системы, учитывающие скорость коагуляции (1) для потоков, движущихся по соплу, получены из численного эксперимента и представлены в [2].

### Процесс образования кластера

Классическая теория гомогенной конденсации, представленная работами Френкеля, Зельдовича, Майера и др. [3], предполагает, что рост кластера происходит за счет присоединения к нему мономолекулы:

$$g_i = g_{i-1} + 1, \quad (2)$$

где  $g_i$  – размер  $i$ -го кластера.

Построенная на этой основе модель расчетов параметров системы, находящейся в состоянии термодинамического равновесия, дает существенную погрешность относительно экспериментальных данных [4, 5].

Предположим, что групповые столкновения в процессе конденсации являются наиболее устойчивыми и химически уравновешенными. В связи с этим рост кластеров будет происходить за счет присоединения одних кластеров к другим, а не путем присоединения одиночных молекул. Описать данный процесс можно следующими рядами.

1. Ряд Фибоначчи, где  $g_i = g_{i-1} + g_{i-2}$ :

$$1, 2, 3, 5, 8, 13, 21, 34, 55, \dots \quad (3)$$

2. Геометрическая прогрессия с шагом 2, где  $g_i = 2 \cdot g_{i-1}$ :

$$1, 2, 4, 8, 16, 32, 64, 128, \dots \quad (4)$$

Таким образом, в процессе конденсации в системе присутствуют кластеры, значение числа молекул в которых представлено в выбранном ряду.

### Анализ полученных данных

Значения параметров системы, движущейся в потоке и имеющей набор кластеров согласно ряду Фибоначчи, представлены в [5]. Результаты показывают, что температура и давление системы совпадают с экспериментальными данными [6]. Скорости образования кластеров паров воды по новому ряду, а также для классической теории представлены в таблице.

По классической теории момент «обрушения» системы в жидкую фазу происходит очень быстро, о чем нам говорят сравнительно большие значения по скоростям конденсации. Если исходить из экспериментальных данных [7] для паров воды на линии