

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО КОНТРОЛЯ ПО АЛЬТЕРНАТИВНОМУ ПРИЗНАКУ

© 2014

**В.Н. Клячкин**, доктор технических наук, профессор кафедры «Прикладная математика и информатика»

**Л.А. Козлова**, программист отдела Информационно-телекоммуникационного обеспечения

Института дистанционного и дополнительного образования

Ульяновский государственный технический университет, Ульяновск (Россия)

**Ключевые слова:** контроль по альтернативному признаку; доля несоответствующих единиц продукции; контрольная карта; средняя длина серий; статистические испытания.

**Аннотация:** Исследована эффективность двух подходов к проведению многомерного статистического контроля технологического процесса по альтернативному признаку. Сравнительный анализ проводится по критерию средней длины серий путем проведения статистических испытаний.

Стандартами статистического управления процессами предусмотрены четыре метода контроля по альтернативному признаку [1]. Это использование карт Шухарта для числа и доли несоответствующих единиц продукции ( $np$  и  $p$ -карты соответственно), а также числа и доли несоответствий ( $c$  и  $u$ -карты). При контроле многопараметрического технологического процесса можно использовать одномерные карты Шухарта по каждому параметру. Это адекватная процедура лишь в том случае, если параметры являются независимыми. На практике многие производственные показатели коррелированы [2], поэтому необходимо использовать специальные многомерные методы контроля.

Один из первых подходов к процедуре многомерного статистического контроля по альтернативному признаку, основанный на применении многомерного биномиального распределения или многомерного распределения Пуассона, предложен Пателем [3]. Позднее другой метод, основанный на расширении  $np$ -карты, предложен Лу с соавторами [4].

Представляет значительный интерес сравнение эффективности контроля процесса этими методами [5]. Основной характеристикой эффективности контроля является средняя длина серий ARL (Average Run Length) [6] – количество наблюдений от момента нарушения процесса до момента обнаружения этого нарушения при использовании того или иного метода. В некоторых случаях эту характеристику можно определить аналитически, чаще она оценивается по результатам статистических испытаний.

Предположим, что по альтернативному признаку контролируется  $k$  показателей. На практике по каждому набору показателей через определенные промежутки времени (как правило, равные) берется  $m$  выборок объемом по  $n$  наблюдений. По результатам контроля получаем набор данных из нулей и единиц (единица означает наличие несоответствующего объекта), состоящий из  $k$ -столбцов и  $mn$  строк. Обозначим через  $x_{ijt}$  результаты контроля в  $i$ -ом наблюдении ( $i=1, 2, \dots, n$ ) в  $t$ -ой выборке ( $t=1, 2, \dots, m$ ) по  $j$ -му показателю ( $j=1, 2, \dots, k$ ). Для каждой  $t$ -ой выборки можно определить вектор средних, компоненты которого – среднее число несоответствующих единиц продукции в  $t$ -й выборке по  $j$ -му показателю

$$\bar{X}_t = (\bar{x}_{t1} \dots \bar{x}_{tk})^T,$$

где

$$\bar{x}_{ij} = \frac{1}{n} \sum_{i=1}^n x_{ijt}.$$

Введем также вектор целевых средних, который в данном случае является вектором среднего числа несоответствующих единиц продукции по всей совокупности взятых выборок:

$$np_0 = (np_1 \dots np_k)^T,$$

где

$$np_j = \frac{1}{m} \sum_{t=1}^m \bar{x}_{ij} \quad (1)$$

Рассеяние значений показателей качества и корреляция между ними определяются компонентами ковариационной матрицы  $\Sigma$ . Матрица имеет размерность  $k \times k$ , её компоненты оцениваются по формуле:

$$s_{jr} = \frac{1}{m(n-1)} \sum_{t=1}^m \sum_{i=1}^n (x_{ijt} - p_j)(x_{irt} - p_r), \quad (2)$$

$$j, r = 1, \dots, k.$$

Пател предложил  $\chi^2$ -карту (типа Хотеллинга) для мониторинга наблюдений, имеющих многомерное биномиальное распределение или многомерное распределение Пуассона. При достаточно больших значениях  $m$  и  $n$  статистика

$$G_t = (\bar{X}_t - np_0)^T S^{-1} (\bar{X}_t - np_0), \quad (3)$$

определяемая для каждой  $t$ -ой выборки имеет приблизительно распределение хи-квадрат с  $k$  степенями свободы. В формуле (3) матрица  $S$ , компоненты которой определяются по формуле (2), – выборочная оценка ковариационной матрицы  $\Sigma$ . Предполагается, что ковариационная матрица  $\Sigma$  остается неизменной на всем протяжении процесса.

В этой ситуации можно использовать  $\chi^2$ -карту, чтобы обнаружить выход процесса из-под контроля. На карте откладываются значения статистики по формуле (3). На заданном уровне значимости  $\alpha$  верхний

контрольный предел UCL (Upper Control Limit) карты определяется как квантиль  $\chi^2_{1-\alpha}(k)$  распределения хи-квадрат с  $k$  степенями свободы порядка  $1-\alpha$ . Выход статистики (3) за контрольную границу означает нарушение стабильности процесса.

Другой подход предложен Лу с соавторами [4]: они рассмотрели многомерную контрольную карту по альтернативному признаку, которая является обобщением одномерной  $np$ -карты (авторы называют ее  $MNP$ -картой).

Введем статистику  $L_t$ , которая является взвешенной суммой несоответствующих единиц всех показателей качества процесса в выборке:

$$L_t = \sum_{j=1}^k \frac{\bar{x}_{tj}}{\sqrt{p_j}}$$

(доля несоответствующих единиц продукции  $p_j$  определяется из формулы (1)).

Статистика  $L$  имеет математическое ожидание

$$E(L) = n \sum_{j=1}^k \sqrt{p_j}$$

и дисперсию

$$D(L) = n \sum_{j=1}^k (1 - p_j) + 2n \sum_{r < j} \rho_{jr} \sqrt{(1 - p_j)(1 - p_r)},$$

где  $\rho_{jr}$  – это коэффициент корреляции между показателями  $X_j$  и  $X_r$ , который либо уже известен, либо может быть оценен из предварительных выборок.

Используя общие принципы построения контрольных карт Шухарта, в основе которых – правило «трех сигма», и предполагая, что вектор доли несоответствий и корреляционная матрица многомерного процесса известны или оценены, находим положение центральной линии

$$CL = n \sum_{j=1}^k \sqrt{p_j}$$

и контрольные пределы  $MNP$ -карты типа Шухарта

$$UCL = n \sum_{j=1}^k \sqrt{p_j} + 3 \sqrt{n \left\{ \sum_{j=1}^k (1 - p_j) + 2 \sum_{r < j} \rho_{jr} \sqrt{(1 - p_j)(1 - p_r)} \right\}},$$

$$LCL = n \sum_{j=1}^k \sqrt{p_j} - 3 \sqrt{n \left\{ \sum_{j=1}^k (1 - p_j) + 2 \sum_{r < j} \rho_{jr} \sqrt{(1 - p_j)(1 - p_r)} \right\}}.$$

Задачей исследования является сравнение чувствительности к нарушению процесса многомерных контрольных карт по альтернативному признаку, построенных по методам Пателя и Лу. Для того, чтобы найти среднюю длину серий, необходимо вначале исследовать процесс, который находится под контролем, чтобы оценить математические ожидания показателей и ковариационную матрицу. Последующие наборы выборок можно смоделировать из обучающего набора с помощью бутстреп-метода [7].

Смещение среднего уровня процесса моделировалось по одному или нескольким контролируемым показателям в долях от оценки соответствующего стандартного отклонения  $\delta_j s_j$  ( $s_j = \sqrt{s_{jj}}$ ;  $s_{jj}$  определяется из формулы (2)); значение  $\delta_j$  выбиралось в диапазоне от 0,25 до 3 с шагом 0,25. По каждому виду нарушений определялось среднее по смоделированным выборкам число наблюдений, взятое от момента нарушения до момента выхода контролируемой статистики за контрольную границу по методам Пателя и Лу.

Моделировались выборки, идентичные данным реального технологического процесса, в котором контролировались по альтернативному признаку три показателя. Был получен набор данных из 60 тысяч выборок по трем переменным, имеющим многомерное биномиальное распределение с вероятностями успеха  $p_1=0,06$ ,  $p_2=0,09$ ,  $p_3=0,07$  и коэффициентами корреляции  $\rho_{12}=0,41$ ,  $\rho_{23}=0,27$ ,  $\rho_{13}=0,33$ .

На рисунке 1 показаны результаты испытаний: кривые зависимости средней длины серий от смещения среднего уровня процесса, характеризующегося значением  $\delta$ , для контрольных карт по методу Пателя и соответствующие точки для карт по методу Лу. На рисунке 1, а показаны результаты испытаний, при которых моделировалось смещение процесса по двум показателям, на рисунке 1, б – по трем.

Результаты испытаний показывают, что метод Пателя дает во всех случаях среднюю длину серий меньшую, чем метод Лу, то есть соответствующие контрольные карты эффективнее с точки зрения чувствительности к обнаружению нарушений.

В то же время видно, что смещения процесса в диапазоне  $\delta=0,25 \div 1,75$  не обнаружены ни одной из использованных контрольных карт. Отсюда вытекает задача модификации методов для повышения их эффективности, при этом за основу должен быть взят метод Пателя, как показавший более высокую чувствительность. Основные направления модификации, по-видимому аналогичны методам повышения эффективности многомерных карт по количественному признаку: использование алгоритма кумулятивных сумм, экспоненциально взвешенных скользящих средних, или применение предупреждающей границы [2].

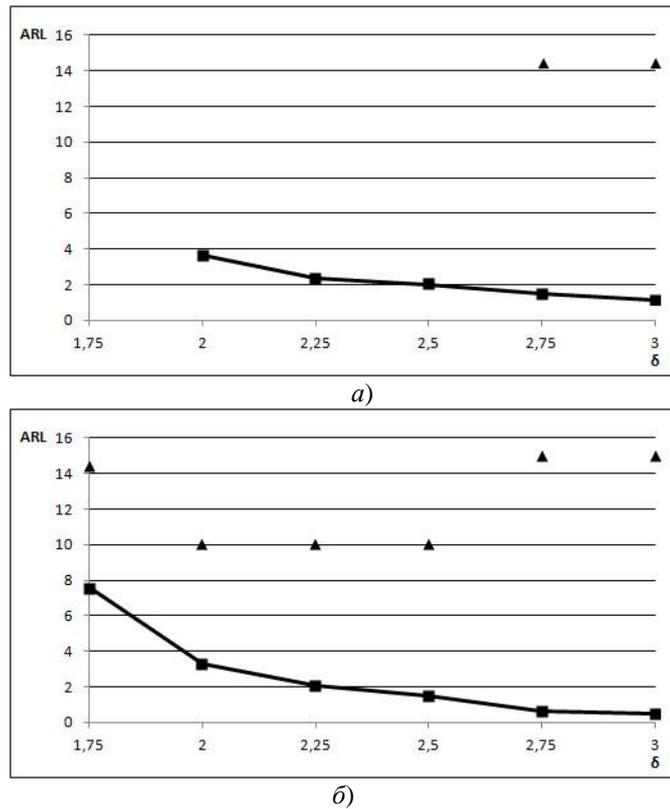


Рис. 1. Зависимость средней длины серий (ARL) от смещения среднего уровня процесса  $\delta$ : а) по двум показателям, б) по трем показателям (сплошная линия – метод Пателя, треугольные точки – метод Лу)

#### СПИСОК ЛИТЕРАТУРЫ

- ГОСТ Р 50779.42–99 (ИСО 8258-91) Статистические методы. Контрольные карты Шухарта.
- Клячкин, В.Н. Многомерный статистический контроль технологического процесса / В.Н. Клячкин. – М.: Финансы и статистика, 2003. – С. 192.
- Patel H. Quality control methods for multivariate binomial and Poisson distributions // *Technometrics*, 1973. № 15, p. 103–112.
- Lu X.S., Xie M., Goh T.N., Lai C.D. Control chart for multivariate attribute processes // *International Journal of Production Research*, 1998. № 12, p. 3477–3489.
- Козлова, Л.А. Сравнение методов многомерного статистического контроля по альтернативному признаку. / Л.А. Козлова // Современные проблемы создания и эксплуатации радиотехнических систем: Сборник научных трудов восьмой Всероссийской научно-практической конференции (с участием стран СНГ). – Ульяновск : УлГТУ, 2013. – С. 192–194.
- Клячкин, В.Н. Статистические методы в управлении качеством: компьютерные технологии / В.Н. Клячкин. – М.: Финансы и статистика, ИНФРА-М, 2009. – С. 304.
- Эфрон Б. Нетрадиционные методы многомерного статистического анализа. / Пер. с англ.; Под ред. Ю.П. Адлера. М.: Финансы и статистика, 1988. 263 с.

#### COMPARATIVE ANALYSIS OF THE MULTIVARIATE STATISTICAL CONTROL METHODS FOR ATTRIBUTES

© 2014

*V.N. Klyachkin*, doctor of technical sciences, professor of the department «Applied mathematics and informatics»

*L.A. Kozlova*, programmer of Information and telecommunication maintenance

Institute of distance and additional education

*Ulyanovsk State Technical University, Ulyanovsk (Russia)*

**Keywords:** control for attributes; fraction nonconforming product units; control chart; average run length; statistical tests.

**Annotation:** The efficiency of the two approaches to conducting multivariate statistical control for attributes of the technological process is investigated. Comparative analysis is performed by the average run length through statistical tests.