

## КЛАСТЕРНЫЙ АНАЛИЗ СОВРЕМЕННОГО АРХИТЕКТУРНОГО ОБРАЗОВАНИЯ НА ОСНОВЕ ВСТУПИТЕЛЬНЫХ ИСПЫТАНИЙ

Михальчук А. А., Арефьев В. П.

*ФГБОУ ВПО «Национальный исследовательский томский политехнический университет», Томск, Россия (634050, Томск, пр. Ленина, 30), e-mail: [R577@tpu.ru](mailto:R577@tpu.ru)*

Проведен многомерный статистический анализ качества набора абитуриентов в российские архитектурно-строительные вузы на основе результатов вступительных испытаний 2012 г. Выявлена высоко значимая отрицательная корреляционная зависимость между долевым количеством абитуриентов, принятых по конкурсу баллов ЕГЭ и по целевому набору. Другой особенностью вступительных испытаний является практически нулевое доленое количество абитуриентов, принятых по олимпиадам. Построена статистически значимая семи кластерная качественная модель архитектурно-строительных вузов в трехмерном пространстве показателей вступительных испытаний 2012 г. (средний балл ЕГЭ 2012 г., его разность между 2012 и 2011 гг. и доленое количество абитуриентов, принятых по конкурсу баллов ЕГЭ). Такая модель позволяет выделять группы архитектурно-строительных вузов, однородных по совокупности показателей. Результаты подобной кластеризации вузов могут быть учтены в рамках проходящей реформы высшего образования, например, при финансировании вузов или сокращении их количества.

Ключевые слова: многомерный статистический (корреляционный, кластерный и дисперсионный) анализ, рейтинг, вступительные испытания.

## THE CLUSTER ANALYSIS OF MODERN ARCHITECTURAL EDUCATION ON THE BASIS OF INTRODUCTORY TRIALS

Mihalchuk A. A., Arefyev V. P.

*National research tomsk polytechnic university, Tomsk, Russia (634050, Tomsk, Lenin's avenue, 30), e-mail: [R577@tpu.ru](mailto:R577@tpu.ru)*

The multidimensional statistical analysis of quality of a gang of entrants is spent to the Russian is architectural-building high schools on the basis of outcomes of introductory trials of 2012 highly significant negative correlative association between a share amount of the entrants accepted on competition of points of Unified State Examination and on a target gang Is revealed. Other singularity of introductory trials is almost zero share amount of the entrants accepted on olympiad. The significant seven cluster quality model of is architectural-building high schools in three-dimensional space of indicators of introductory trials of 2012 (a mean score of Unified State Examination of 2012, its difference between 2012 both 2011 and a share amount of the entrants accepted on competition of points of Unified State Examination) is constructed statistically. Such model allows to select groups of the is architectural-building high schools homogeneous for a population of indicators. Outcomes similar кластеризации high schools can be considered within the limits of passing reform of higher education, for example, at financing of high schools or reduction of their amount.

Keywords: multidimensional statistic (correlative, cluster and dispersing) analysis, rating, introductory testing.

Современное российское архитектурное образование находится в стадии реформирования [7, 8]. В данной работе проведен статистический анализ российского архитектурно-строительного образования в начале XXI века на основе вступительных испытаний, в значительной степени опирающихся на результаты ЕГЭ как критерий оценки качества работы средней школы и качества набора абитуриентов в высшую школу [1, 2] с привлечением иногда для анализа методов математической статистики [6]. Рейтинг вузов РФ по среднему баллу ЕГЭ ( $m_{ЕГЭ}$ ) [4], показывающий уровень знаний абитуриентов 2012 года, поступивших на бюджетные места в государственные вузы страны, возглавляет МГИМО, у которого средний балл ЕГЭ – 92,3 по 100 балльной шкале, а среди государственных архитектурно-строительных вузов (АСВ) лидирует Уральская ГАХА (76,9) – 33-е место среди 491 вуза. В

табл. 1 из показателей вступительных испытаний (ПВИ) кроме  $m_{\text{ЕГЭ}}$  приведены также разность  $m_{\text{ЕГЭ}}$  между 2012 и 2011 гг. ( $Dm$ ), долевое количество абитуриентов (в % от общего количества бюджетных мест), принятых по конкурсу баллов ЕГЭ ( $N_{\text{ЕГЭ}}$ ), по олимпиадам ( $N_{\text{О}}$ ), по льготам ( $N_{\text{Л}}$ ) и по целевому набору ( $N_{\text{Ц}}$ ). Заметим, что подсистема долевых показателей является избыточной, так как  $N_{\text{ЕГЭ}}+N_{\text{Ц}}+N_{\text{О}}+N_{\text{Л}}=100\%$ .

**Таблица 1.** ПВИ 2012 г. государственных архитектурно-строительных университетов (ГАСУ) и архитектурно-художественных академий (ГАХА)

| Рейтинг АСВ ( $m_{\text{ЕГЭ}}$ ) | Рейтинг вузов ( $m_{\text{ЕГЭ}}$ ) | АСВ                      | $m_{\text{ЕГЭ}}$ | $Dm$ | $N_{\text{ЕГЭ}}$ | $N_{\text{О}}$ | $N_{\text{Л}}$ | $N_{\text{Ц}}$ |
|----------------------------------|------------------------------------|--------------------------|------------------|------|------------------|----------------|----------------|----------------|
| 1                                | 33                                 | Уральская ГАХА           | 76,9             | 3,1  | 89,14            | 0              | 3,17           | 7,69           |
| 2                                | 81                                 | Санкт-Петербургский ГАСУ | 71,5             | -0,2 | 87,84            | 0              | 4,64           | 7,52           |
| 3                                | 92                                 | Новосибирская ГАХА       | 70,8             | 4,2  | 90,65            | 0              | 2,16           | 7,19           |
| 4                                | 112                                | Московский АРХИ          | 69,1             | 1,1  | 92               | 0              | 2,29           | 5,71           |
| 5                                | 131                                | Казанский ГАСУ           | 67,7             | 1,1  | 85,1             | 0              | 3,05           | 11,8           |
| 6                                | 136                                | Московский ГСУ           | 67,5             | 1,3  | 90,63            | 0              | 3,79           | 5,58           |
| 7                                | 234                                | Тюменский ГАСУ           | 62,6             | -1,6 | 85,66            | 0              | 2,83           | 11,5           |
| 8                                | 238                                | Нижегородский ГАСУ       | 62,4             | 0,8  | 96,73            | 0,4            | 2,86           | 0              |
| 9                                | 239                                | Ростовский ГСУ           | 62,1             | -0,4 | 84,76            | 0              | 4,38           | 10,9           |
| 10                               | 247                                | Пензенский ГАСУ          | 61,9             | 0,6  | 92,54            | 1,1            | 2,63           | 3,73           |
| 11                               | 253                                | Самарский ГАСУ           | 61,6             | -2,2 | 90,38            | 1,5            | 2,83           | 5,28           |
| 12                               | 279                                | Воронежский ГАСУ         | 60,8             | -4,4 | 62,37            | 0              | 4,57           | 33,1           |
| 13                               |                                    | Ивановский ГАСУ (2011г.) | 58,5             | 4,2  | 95,86            | 0              | 1,48           | 2,66           |
| 14                               | 369                                | Волгоградский ГАСУ       | 57,3             | 0,3  | 97,83            | 0              | 1,81           | 0,36           |
| 15                               | 391                                | Московская ГАКХиС        | 56,2             | 2,7  | 98,51            | 0              | 1,49           | 0              |
| 16                               | 430                                | Томский ГАСУ             | 53,8             | 2    | 92,86            | 0              | 1,31           | 5,83           |
| 17                               | 464                                | Новосибирский ГАСУ       | 51,1             | -7,3 | 95,85            | 0              | 1,7            | 2,45           |

Согласно табл. 1,  $N_{\text{О}}$  практически равно нулю. Данное обстоятельство должно вызывать тревогу в рамках креативного подхода к архитектурному образованию [7].

В работах [3, 5] рассмотрено применение метода классификации вузов на мировом. В данной работе этот метод применен на федеральном уровне для классификации российских АСВ по совокупности ПВИ 2011–2012 гг. (табл. 1). Составляющими статистического метода исследования являются корреляционный, кластерный и дисперсионный анализы. Статистический анализ проводился в системе Statistica [9]. В силу разнородности ПВИ они были стандартизированы.

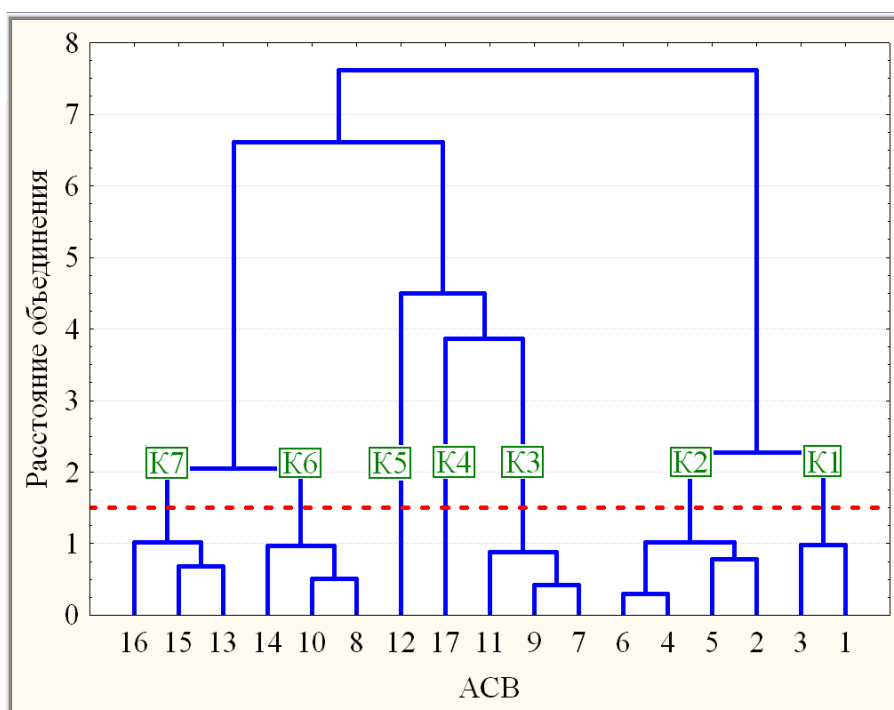
Результаты проверки ПВИ на корреляционную зависимость (матрица коэффициентов парных корреляций) приведены в табл. 2 (Пирсона  $r$  – в право-верхнем треугольнике над диагональю и Спирмена  $R$  – в лево-нижнем треугольнике под диагональю). Жирным шрифтом выделены статистически значимые корреляции: положительная между  $N_{\text{Л}}$  и  $N_{\text{Ц}}$ , отрица-

тельные – между  $N_{EGЭ}$  и  $N_{Ц}$ , а также  $N_{EGЭ}$  и  $N_{Л}$ . Наличие корреляционной связи долевых ПВИ позволяет уменьшить их количество до одного (основного  $N_{EGЭ}$ ) и провести дальнейший анализ в трехмерном пространстве ПВИ:  $\{m_{EGЭ}, Dm, N_{EGЭ}\}$ .

**Таблица 2.** Матрица коэффициентов парных корреляций Пирсона  $r$  и ранговых корреляций Спирмена  $R$  ПВИ

| ПВИ       | $m_{EGЭ}$ | $Dm$  | $N_{EGЭ}$    | $N_{Л}$      | $N_{Ц}$      |
|-----------|-----------|-------|--------------|--------------|--------------|
| $m_{EGЭ}$ | 1,00      | 0,42  | -0,22        | 0,51         | 0,17         |
| $Dm$      | 0,26      | 1,00  | 0,38         | -0,27        | -0,36        |
| $N_{EGЭ}$ | -0,51     | 0,36  | 1,00         | <b>-0,70</b> | <b>-0,99</b> |
| $N_{Л}$   | 0,61      | -0,37 | <b>-0,77</b> | 1,00         | <b>0,62</b>  |
| $N_{Ц}$   | 0,49      | -0,16 | <b>-0,93</b> | <b>0,59</b>  | 1,00         |

При проведении кластеризации АСВ в построенном трехмерном пространстве  $\{m_{EGЭ}, Dm, N_{EGЭ}\}$  в качестве меры близости выбрано евклидово расстояние, а в качестве правила объединения двух кластеров использован метод Уорда. Методом древовидной кластеризации построено иерархическое дерево (рис. 1).



**Рис. 1.** Дендрограмма АСВ в пространстве  $\{m_{EGЭ}, Dm, N_{EGЭ}\}$

В зависимости от выбора расстояния объединения можно получить соответствующее число кластеров. Так, например, расстоянию объединения, равного 1,5 (красная пунктирная горизонтальная прямая), соответствует 7 кластеров (К1-7), расстоянию объединения, равного 5, соответствует 3 кластера (К1+К2, К3+ К4+К5, К6+К7). Таким образом, выбор зна-

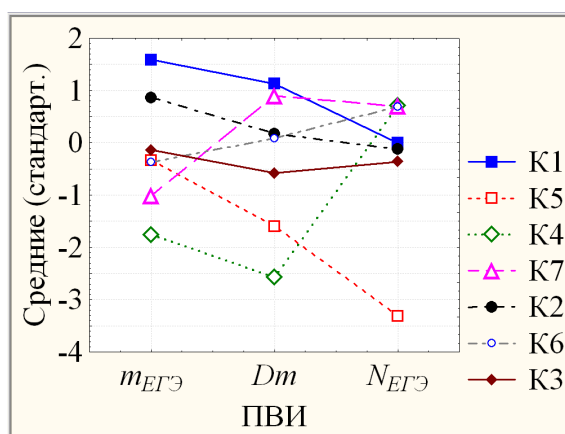
чения связующего расстояния позволяет проводить кластеризацию на любом уровне, то есть строить кластерную модель с любым наперед заданным числом кластеров.

Наряду с методом древовидной кластеризации применяется также метод  $K$ -средних, проводящий классификацию объектов (АСВ) по заданному количеству кластеров. Предлагается семикластерная высококачественная модель АСВ, согласно  $\lambda$ -критерию Уилкса высоко значимо (на уровне значимости  $p < 0,000005$ ) различающая 7 кластеров АСВ по совокупности  $m_{EG3}$ ,  $Dm$  и  $N_{EG3}$ . Качество кластеризации АСВ по каждому ПВИ оценено посредством параметрического дисперсионного анализа (табл. 3) как высоко значимое (на уровне значимости  $p < 0,00005$ ). Заметим, что согласно критериям непараметрического дисперсионного анализа (ранговому критерию Краскела-Уоллиса и медианному тесту) качество кластеризации АСВ по каждому ПВИ оценивается как статистически значимое (на уровне значимости  $0,01 < p < 0,05$ ).

**Таблица 3.** Результаты дисперсионного анализа кластеризации АСВ по каждому ПВИ

| Дисперсионный анализ (Кластер_АСВ) |             |    |              |    |       |                 |
|------------------------------------|-------------|----|--------------|----|-------|-----------------|
| перем                              | Между<br>SS | сс | Внутри<br>SS | сс | F     | значим.<br>p    |
| $m_{EG3}$                          | 14,79       | 6  | 1,21         | 10 | 20,31 | <b>0,000046</b> |
| $Dm$                               | 15,27       | 6  | 0,73         | 10 | 34,98 | <b>0,000004</b> |
| $N_{EG3}$                          | 14,86       | 6  | 1,14         | 10 | 21,64 | <b>0,000035</b> |

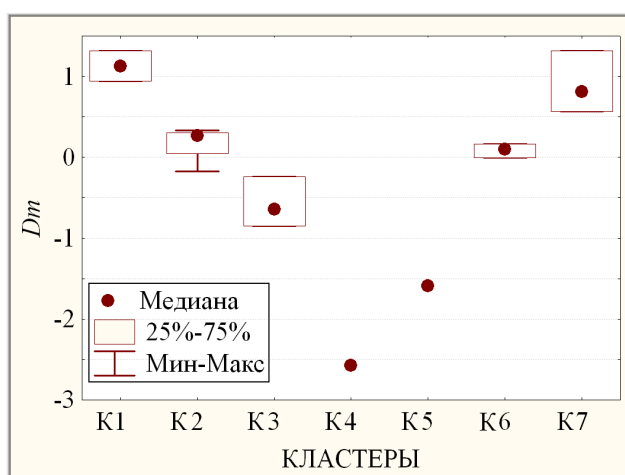
После получения результатов классификации рассчитываются средние значения показателей по каждому кластеру (рис. 2) для оценивания значимости их различий между собой.



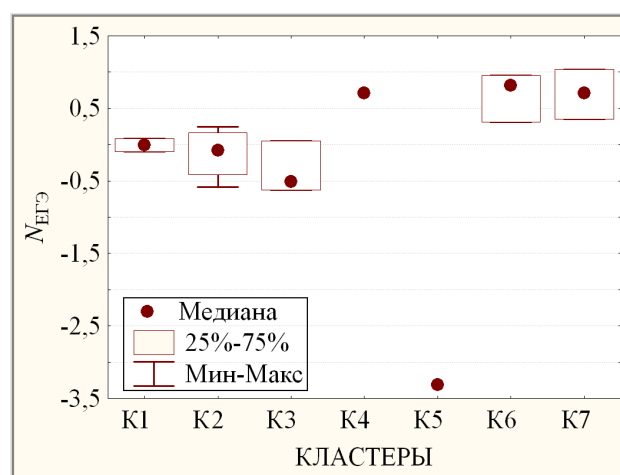
**Рис. 2.** Графики средних значений ПВИ по каждому кластеру

Для каждого ПВИ согласно апостериорному критерию наименьших значений разности (НЗР), ранговому критерию Краскела-Уоллиса и медианному тесту можно выделить однородные (различающиеся незначимо, то есть на уровне значимости  $p > 0,10$ ) группы кластеров (рис. 2), расположенные в порядке убывания средних:

- $m_{EGЭ}$ : {K1}, {K2}, {K3, K5, K6}, {K4}, {K7}. При этом, например, {K2} отличается от {K3, K5, K6} сильно значимо по НЗР (уровень значимости  $0,001 < p < 0,01$ ) и статистически значимо по критерию Краскела-Уоллиса и по медианному тесту (уровень значимости  $0,01 < p < 0,05$ ).
- $Dm$  (рис. 3): {K1, K7}, {K2, K6}, {K3}, {K5}, {K4}. При этом, например, {K1, K7} и {K2, K6} различаются сильно значимо по НЗР и по медианному тесту (уровень значимости  $0,001 < p < 0,01$ ), а по критерию Краскела-Уоллиса различаются статистически значимо ( $0,01 < p < 0,05$ ).
- $N_{EGЭ}$  (рис. 4): {K4, K6, K7}, {K1, K2, K3}, {K5}. При этом, например, по НЗР K1 отличается от K6 и K7 также, как K2 от K4 на уровне значимости  $p \approx 0,05$ , а {K4, K6, K7} и {K1, K2, K3} различаются статистически значимо по критерию Краскела-Уоллиса (уровень значимости  $0,01 < p \approx 0,038 < 0,05$ ) и по медианному тесту (уровень значимости  $0,01 < p \approx 0,023 < 0,05$ ).

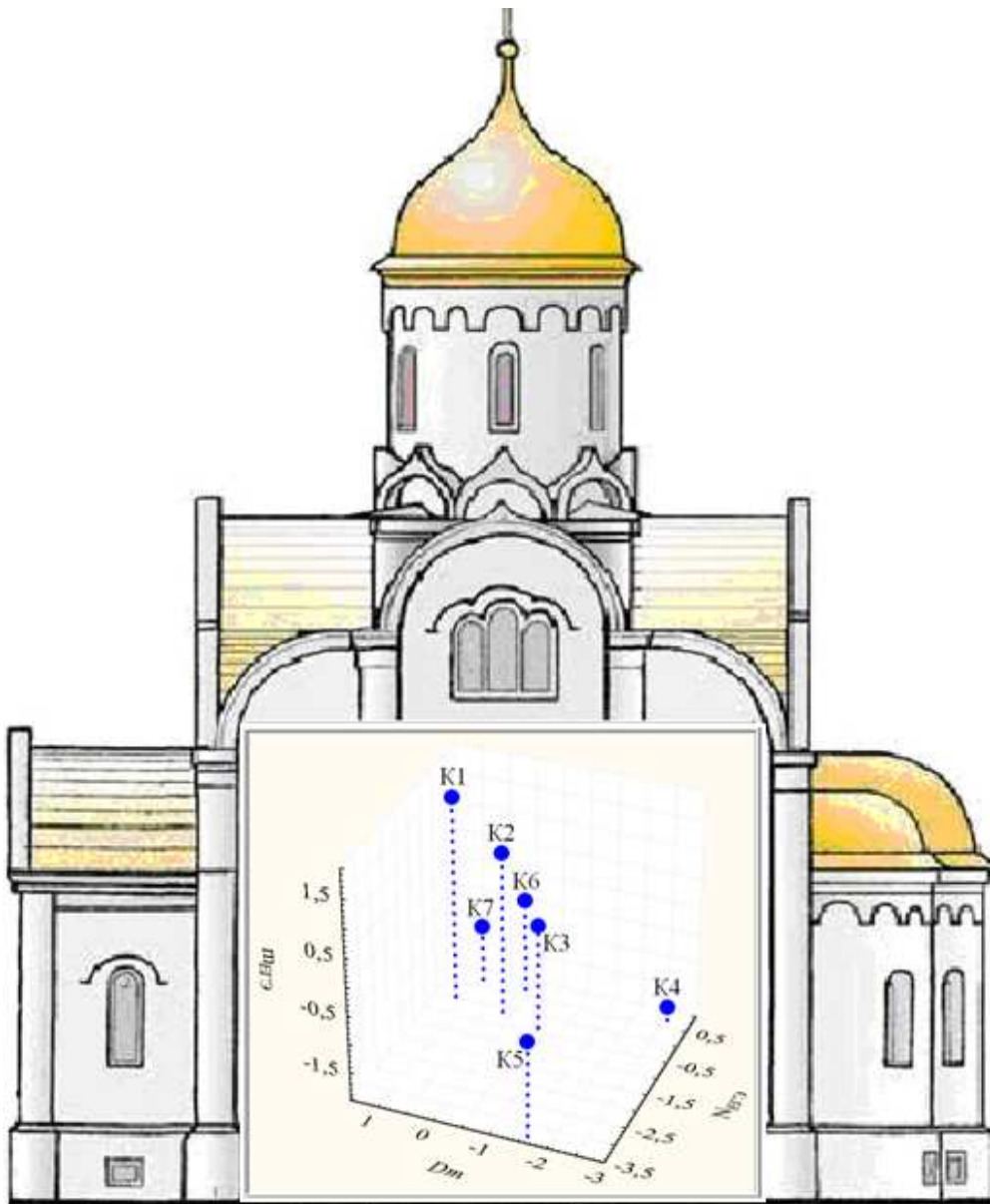


**Рис. 3.** Диаграмма размаха для каждого кластера по  $Dm$



**Рис. 4.** Диаграмма размаха для каждого кластера по  $N_{EGЭ}$

Графики средних кластеров в рамках семи кластерной модели АСВ представлены диаграммой рассеяния средних кластеров АСВ в трехмерном пространстве  $\{m_{EGЭ}, Dm, N_{EGЭ}\}$  в виде образной формы «храма потенциального высшего архитектурно-строительного образования» (рис. 5). Как было замечено выше (табл. 1), практически отсутствует «олимпиадная» составляющая фундамента «храма», необходимая в рамках креативного подхода к архитектурному образованию [7], что порождает ощущение неполноценности и неустойчивости всего сооружения.



**Рис. 5.** «Храм потенциального архитектурно-строительного образования»

Результаты кластерного анализа АСВ по совокупности показателей (рис. 2–5) позволяют провести качественную классификацию АСВ в номинальной шкале измерений (табл. 4), полагая в качестве уровня «Средний» – стандартизированный интервал (-0,5; +0,5) для ПВИ. Аномально высокие значения ( $> +2,0$ ) определяют уровень «**Лидер**», а аномально низкие значения ( $< -2,0$ ) определяют уровень «**Аутсайдер**». Промежуточные значения между средними и аномальными значениями определяют уровень «Выше среднего» и «Ниже среднего» соответственно.

**Таблица 4.** Качественная классификация АСВ

| Кластер | Уровень кластера |
|---------|------------------|
|---------|------------------|

|    | на фоне среднего по ПВИ |                  |                  |
|----|-------------------------|------------------|------------------|
|    | $m_{\text{ЕГЭ}}$        | $Dm$             | $N_{\text{ЕГЭ}}$ |
| К1 | Выше среднего           | Выше среднего    | Средний          |
| К2 | Выше среднего           | Средний          | Средний          |
| К3 | Средний                 | Ниже среднего    | Средний          |
| К5 | Средний                 | Ниже среднего    | <b>Аутсайдер</b> |
| К6 | Средний                 | Средний          | Выше среднего    |
| К7 | Ниже среднего           | Выше среднего    | Выше среднего    |
| К4 | Ниже среднего           | <b>Аутсайдер</b> | Выше среднего    |

### Выводы

1. Особенностью системы показателей вступительных испытаний 2012 г. в российские архитектурно-строительные вузы является практически нулевое доленое количество абитуриентов (в % от общего количества бюджетных мест), принятых по олимпиадам.
2. На основании корреляционного анализа показателей вступительных испытаний в архитектурно-строительные вузы выявлена высоко значимая наиболее сильная отрицательная корреляционная зависимость между долевым количеством абитуриентов, принятых по конкурсу баллов ЕГЭ и по целевому набору.
3. В рамках кластерного анализа в трех мерном пространстве ПВИ построена кластерная модель архитектурно-строительных вузов, позволяющая проводить кластеризацию на любом уровне, т. е. строить кластерную модель с любым наперед заданным числом кластеров, что дает возможность выделять группы архитектурно-строительных вузов, однородных по совокупности факторных показателей.
4. Подробно рассмотрена высоко значимая семикластерная модель. Для каждого показателя выделены группы однородных кластеров. На основе результатов кластеризации проведена качественная классификация АСВ.
5. Результаты подобной кластеризации вузов могут быть учтены в рамках проходящей реформы высшего образования, например, при финансировании вузов или сокращении их количества.

## Список литературы

1. Гоник И. Л., Москвичев С. М., Иванов Ю. В., Гурулев Д. Н. Различные формы сдачи вступительных испытаний как элемент формирования контингента абитуриентов // Известия Волгоградского государственного технического университета. – 2009. – Т. 10. – № 6. – С. 27–28.
2. Грязев М. В., Хадарцев А. А., Хрупачёв А. Г., Туляков С. П. Методика интегральной оценки знаний абитуриентов // Высшее образование в России. – 2010. – № 6. – С. 28–32.
3. Ильшев А. М., Шубат О. М. Многомерная классификация данных: особенности методики, анализ практики и перспектив применения // Вопросы статистики. – 2010. – № 10. – С. 34–40.
4. Качество приема в государственные вузы РФ по профилю – 2012 // РИА Новости. – URL: [http://www.ria.ru/ratings\\_rt2012/](http://www.ria.ru/ratings_rt2012/) (28.09.2012).
5. Кружалин В. И., Аршинова В. В., Попов Л. В., Чаплыгина А. А. Рейтинги мировых университетов как инструмент управления качеством образования // Alma mater (Вестник высшей школы). – 2010. – № 6. – С. 9–8.
6. Лапотникова И. Н. Методы математической статистики для оценки результатов ЕГЭ // Ярославский педагогический вестник. – 2008. – № 1. – С. 17–23.
7. Метленков Н. Ф. Креативное образование в архитектуре и проблемы его современной стандартизации // Alma mater (Вестник высшей школы). – 2010. – № 3. – С. 11–19.
8. Никитина Н. П. Архитектурное образование: от истоков до наших дней (современная потребность в инженерно-строительной подготовке архитектора) // Актуальные проблемы гуманитарных и естественных наук. – 2009. – № 12. С. 300–302.
9. Халафян А. А. Statistica 6. Статистический анализ данных. – М.: ООО «Бином-Пресс», 2008. – 512 с.

*Работа выполнена в рамках государственного задания «Наука» № 1.604.2011 и поддержана ФЦП «Научные и научно-педагогические кадры инновационной России» по контрактам П691.*

### Рецензенты:

Трифонов Андрей Юрьевич, д.ф.-м.н., профессор кафедры высшей математики и математической физики, ФГБОУ ВПО «Национальный исследовательский томский политехнический университет, г. Томск.

Арефьев Константин Петрович, д.ф.-м.н., профессор кафедры высшей математики, ФГБОУ ВПО «Национальный исследовательский томский политехнический университет, г. Томск.