УДК 519.23 + 37.062]:330.4

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ ЧИСЛЕННОСТИ ИНОСТРАННЫХ СТУДЕНТОВ, ОБУЧАЮЩИХСЯ В ВОЛГОГРАДСКОМ ТЕХНИЧЕСКОМ УНИВЕРСИТЕТЕ

Годенко А.Е., Симонова И.Э., Симонов А.Б., Тарасова И.А.

ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, e-mail: simonova-vstu@mail.ru, forstud@vstu.ru

Статья посвящена исследованию динамики изменения числа иностранных студентов, приезжающих на довузовский этап обучения в ВолгГТУ. Изучались данные о числе приглашений за 11 лет. Проведено статистическое исследование соответствующего ряда динамики с использованием ППП Statgraphics. Выявлены основные статистические закономерности. Методами однофакторного дисперсионного анализа построены расслоения исходных данных по годам, а затем по месяцам. Это дало возможность провести дальнейший анализ годовых показателей методами регрессионного анализа. Проанализированы особенности рядов динамики, выявлены вызывающие их причины, намечены пути улучшения показателей. Исследован характер сезонных колебаний, вычислены сезонные (месячные) индексы. Исследована также динамика числа студентов, поступивших на первый курс ВолгГТУ, и взаимосвязь двух показателей – числа иностранных студентов на довузовском и вузовском этапах обучения.

Ключевые слова: временные ряды, статистические модели, иностранные студенты

STARISTICAL MODELING OF DYNAMICS OF THE NUMBER OF FOREING STUDENTS, STUDING IN VOLGOGRAD TECHNICAL UNIVERSITY

Godenko A.E., Simonova I.E., Simonov A.B., Tarasova I.A.

Volgograd State Technical University, Volgograd, e-mail: simonova-vstu@mail.ru, forstud@vstu.ru

The article is devoted to investigation of dynamics of changes in the number of foreign students coming on preliminary grade level in VSTU. The data of number of invitations for 11 years were studied. Statistical research of corresponding time series using statistical program package Statgraphics is carried on. The basic statistical regularities are revealed. Stratifications of the initial data on years, and then on months were constructed by One-Factor ANOVA methods. It has enabled to carry out the further analysis of annual parameters by the methods of regression analysis. Special features of the time series are analysed, the reasons causing them are revealed, ways of improvement of parameters are planned. The character of seasonal fluctuations is investigated, seasonal (monthly) indexes are calculated. Dynamics of number of the students who have arrived on the first rate of VSTU and interrelation between two parameters – numbers of foreign students on preliminary and high school grade levels is investigated too.

Keywords: time series, statistical models, foreing students

В работе проведено статистическое исследование динамики изменения числа иностранных студентов, обучающихся в Волгоградском государственногм техническом университете (ВолгГТУ). Подробно изучены данные об иностранных студентах, приезжавших на довузовский этап обучения на факультете подготовки иностранных специалистов (ФПИС) ВолгГТУ за период 2005–2015 гг. В качестве показателя взято месячное число приглашений на обучение X_t, оформлявшихся на ФПИС. Проанализирована также динамика числа иностранных студентов, поступивших на 1 курс бакалавриата ВолгГТУ, и изучена взаимосвязь между этими двумя показателями. Актуальность и важность изучения количества иностранных студентов, обучающихся в вузах, обусловлена тем, что доля иностранных студентов в общем числе студентов – один из основных показателей, по которым при аттестации оценивается деятельность вуза. Работа по привлечению иностранных учащихся продвигает российское образование на мировой уровень, поднимает его престиж. Кроме того, средства, поступающие за обучение иностранных граждан, являются важной составляющей бюджета вуза и региона.

Исследование процесса оформления приглашений для иностранных студентов, оформлявшихся на ФПИС ВолгГТУ

Для выявления характерных особенностей процесса X_t построен график числа приглашений (рис. 1).

Визуальный анализ графика приводит к выводу об очень больших колебаниях числа приглашений, существенной неоднородности этого процесса (число приглашений колеблется от 0 до 35 за месяц). Показателем изменчивости экспериментальных данных является коэффициент

вариации K, который вычисляется как отношение среднеквадратического отклонения к среднему числу приглашений. Для числа приглашений X_t значение коэффициента К = 0,694, т.е. колебания сравнимы со средним числом приглашений, которое оказалось равным 13,06 приглашений в месяц. В исследуемом процессе доля неопределенности и случайности велика. Однако методы математической статистики позволяют выявлять закономерности и в таких неоднородных процессах получать обоснованные прогнозы. В свою очередь, такой анализ становится основой для эффективного принятия управленческих решений.

Таблица 1

Среднемесячное число приглашений в 2005–2015 гг.

год	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
среднее число	8,67	12,0	12,08	14,17	8,92	9,92	14,83	14,08	14,17	17,08	17,75

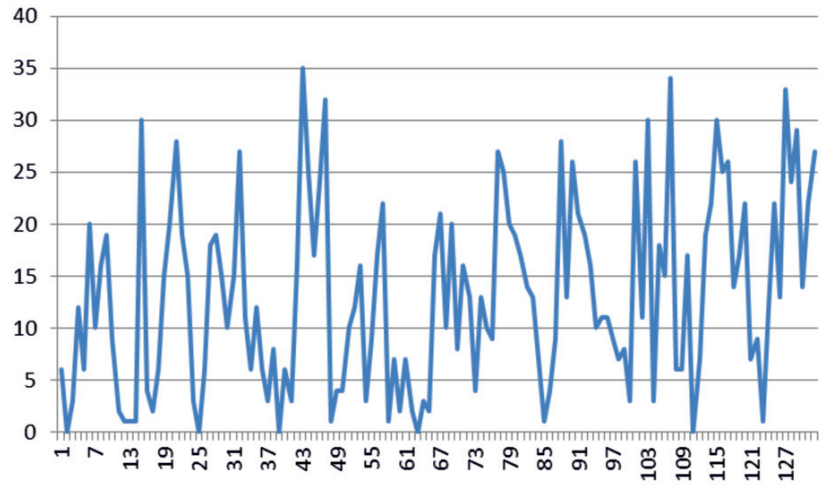

Рис. 1. Число приглашений по месяцам

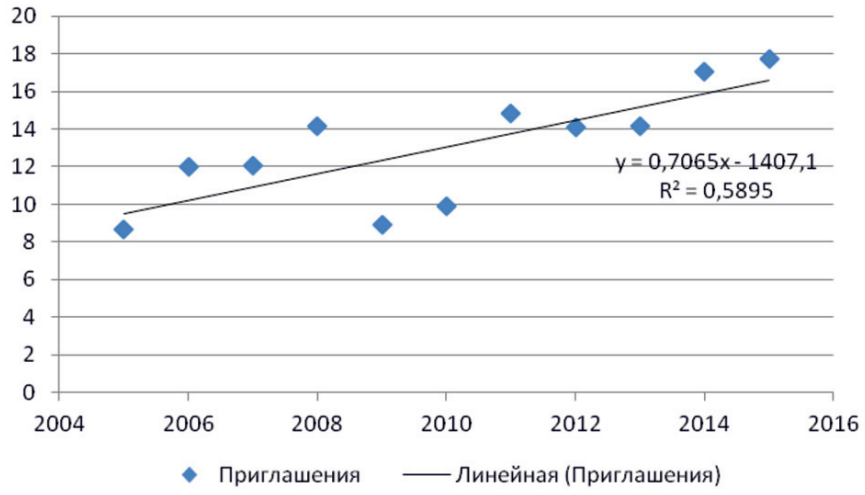

Рис. 2. Модель линейной регрессии числа приглашений

Современные статистические ППП (EViews, SPSS, Statgraphics, Statistica) позволяют осуществлять профессиональную обработку статистической информации, выявлять скрытые закономерности в различных областях – экономике, маркетинге, технике, социологии и др. (см., например, [1, 2]). Для проведения исследований авторами использовался профессиональный статистический пакет Statgraphics. В ВолгГТУ этот пакет (триальная версия) активно используется при проведении прикладных исследований. Так, например, в работе [3] анализируется опыт использования информационных технологий в образовательном процессе. А в [4, 5] приведены результаты статистического исследования деятельности региональных предприятий.

На первом этапе исследований было рассчитано среднемесячное число приглашений Y_t для каждого года. Сами значения представлены в табл. 1. Здесь и ниже значения округлены до сотых.

Исследование этих данных позволяет выделить периоды роста числа приглашений в 2005–2007 годах, существенное снижение числа прглашений в 2009 и 2010 годах и стабильное увеличение среднего числа приглашений в 2011–2015 гг. Так, например, в 2014 г. среднее число приглашений выросло до 17,08 (на 120% за год). Заметим, что среднее число приглашений за весь период равно 13,06, минимальное среднее 12,0 было в 2005 г., а максимальное значение 17,75 достигнуто в 2015 г. Налицо существенный рост среднегодового числа приглашений, несмотря на очень большие колебания по месяцам в течение каждого года.

Для дальнейшего анализа была построена модель линейной регрессии (см. рис. 2). Она позволила выявить тренд (основную тенденцию развития процесса), а также визуально и аналитически выявить колебания около линии тренда.

Действительно, видны периоды превышения над линией тренда и периоды сравнительно низких значений показателя. Относительно высокие средние показатели в 2006–2008 гг. (см. табл. 1) объясняются приездом больших групп китайских студентов в этот период. В связи с изменением ситуации на мировом образовательном рынке китайских студентов в 2009, 2010 гг. приехало гораздо меньше.

В 2011–2014 гг. положение существенно улучшилось благодаря тому, что в ВолгГТУ была проведена большая работа по привлечению студентов из других стран ближнего и дальнего зарубежья. Кроме того, налажено сотрудничество с новыми эффективными партнерами (рекрутинговыми агентствами).

Исследование сезонных колебаний

Второй этап исследований был посвящен изучению колебаний числа приглашений по месяцам и выявлению причин, вызывающих эти колебания. Такие колебания называются сезонными. Для изучения сезонных колебаний был проведен дисперсионный анализ (One-Way ANOVA) процесса X_t (invitations). В качестве фактора взята переменная «месяц» (month).

В первой строке табл. 2 приведены среднемесячные показатели (за 2005–2015 гг.). Во второй строке приведены значения сезонных индексов, вычисленные для каждого месяца как умноженное на 100% отношение среднего за данный месяц к среднемесячному числу приглашений за весь период, равному 13,06 приглашений в месяц (значения сезонных индексов округлены до целых).

Проанализируем причины, вызывающие сезонные (месячные) колебания. В период с июля по сентябрь будущие студенты получают результаты национальных выпускных экзаменов в своих странах. В соответсвтии с этим они принимают решение о выборе страны и вуза обучения. Поэтому в июле, августе и сентябре оформляется наибольшее (в среднем) число приглашений для студентов, поступающих на полный курс. Он составляет 10 месяцев. Обучение начинается в сентябре, октябре или ноябре (по мере комплектования групп). Это наиболее многочисленные группы, в которых занимаются учащиеся, которые готовятся к поступлению в ВолгГТУ и другие Российские вузы. В ноябре оформляют приглашения те, кто по тем или иным причинам не поступил в вуз в текущем году. Для них формируется группа со сроком обучения 1,5 года. Кроме того, на ФПИС организуется летняя языковая школа по изучению русского языка, которая работает в мае – июне. Приглашения для обучающихся в этой школе оформляются в декабре – январе.

Таблица 2

Среднемесячные показатели и сезонные индексы

январь	февраль	март	апрель	май	июнь	июль	август	сент.	октябрь	ноябрь	декабрь
4,36	6,45	8,09	10,45	13,73	15,36	21,73	18,64	20,27	12,18	16,45	9,00
33%	49%	62%	80%	105%	118%	166%	143%	156%	93%	126%	69%

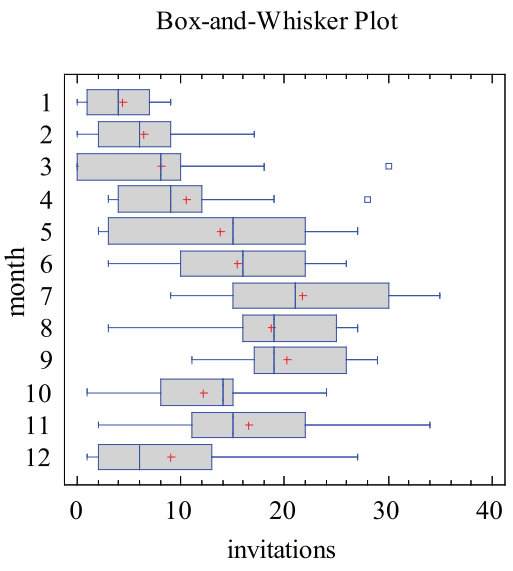

Рис. 3. График «ящик с усами»

Для выявления влияния фактора «месяц» (month) на число приглашений (invitations) был проведен однофакторный дисперсионный анализ данных (One-Way ANOVA).

Расчетное значение статистики Фишера (равное отношению межгрупповой дисперсии к внутригрупповой) равно *F-ratio = 5,91*. Так как соответсвующее ему значение вероятности *P-value < 0,01*, то существует статистически значимое различие между приглашениями в разные месяцы.

Более подробно исследовать колебания по месяцам позволяет график «ящик с усами». Он дает возможность визуализировать как закономерную, так и случайную компоненту динамического процесса X_t. Знак «+» отмечает среднемесячное значение, приведенное в табл.2 . Границы ящика для каждого месяца – это значения нижней и верхней квартилей, т.е. ящик содержит половину «средних» наблюдений для соответсствующего месяца. Вертикальная линия отмечает медиану. «Усы» простираются до наибольшего и наименьшего значений, если они не являются «резко выделяющимися». «Резко выделяющиеся» наблюдения, соответствующие большим значениям наблюдений, изображены точками. Они зафиксированы в марте и апреле.

Наибольшая изменчивость числа приглашений наблюдалась в марте. При среднем числе приглашений, равном 8, 27, три года (2008, 2010, 2014 гг.) в марте приглашения вообще не оформлялись. А в 2006 году было оформлено 30 приглашений. Особенно большое число приглашений, равное 35 и 34 соответственно, было зафиксировано в июле 2008 г. и ноябре 2013 г.

При исследовании временных рядов часто применяется метод декомпозиции временного ряда. При использовании этого метода на первом этапе выявляются тренд и сезонная компонента исходного ряда динамики, а затем строится аддитивная или мультипликативная модель ряда. Данный метод успешно применен одним из авторов статьи в [6]. В настоящей работе в связи с очень большими колебаниями исходного процесса X_t попытка выявить его тренд приводит к неудовлетворительным результатам (построенная регрессионная модель оказывается незначимой). Поэтому для процесса X_t применен метод укрупнения интервалов (переход к годовым показателям). Это позволило выявить тренд процесса $y = 0{,}7\,x - 1407$ (см. рис. 2) и проанализировать социально-экономические причины, вызвавшие колебания около линии тренда.

Исследование числа иностранных студентов Z_t, поступивших на 1 курс ВолгГТУ в 2005–2015 гг.

График процесса Z_t приведен на рис. 4 (в эту модель не включены украинские студенты). Особенностью этого ряда динамики является очень большой разброс наблюдений – от наименьшего значения 45 в 2010 г. до наибольшего значения 88 в 2013. Это привело к тому, что построенная по Z_t модель линейной регрессии $Z_t = 55{,}36 + 1{,}4 \cdot t$ оказалась не значимой (значение статистики критерия Фишера F = 1,73, P-value = 0,22), а ее коэффициент детерминации (доля объясненной дисперсии) равен 16 %.

Более подходящая модель $Z_t = 69{,}7 + \varepsilon_t$. Здесь 69,7 – среднее число студентов, поступивших на 1 курс, а ε_t – случайные колебания. Далее исследовалась взаимосвязь между количеством обучающихся на ФПИС (Y_t) и числом иностранных студентов, зачисленных на 1 курс (Z_t).

Apriori естественно предположить, что существует взаимосвязь между этими процессами, так как многие выпускники ФПИС поступают на 1 курс ВолгГТУ. Однако коэффициент кросскорреляции между этими процессами оказался равен $r = 0{,}41$, что свидетельствует об отсутствии такой взаимосвязи (точнее, ее незначительности). Это вызвано тем, что часть студентов ФПИС поступает в другие вузы или ограничивается языковой подготовкой. С другой стороны, часть иностранных студентов, поступивших на 1 курс, не училась на ФПИС. Это, в основном, студенты из ближнего зарубежья, уже имеющие языковую и общую подготовку, достаточную для поступления в вуз.

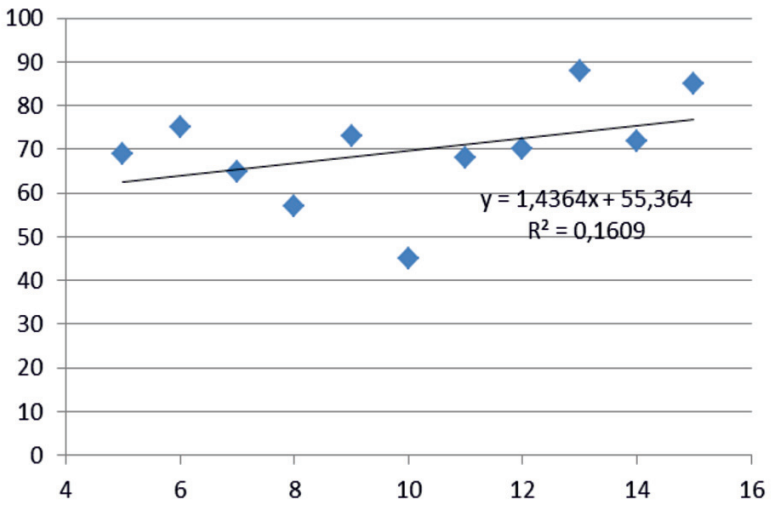

Рис. 4. Число студентов, поступивших на 1 курс

Выводы

Авторами собраны и проанализированы данные о числе приглашений для иностранных студентов, приезжающих на довузовский этап обучения в ВолгГТУ, за 11 лет (132 мес.), а также данные о количестве иностранных студентов, поступивших на первый курс. Проведено статистическое исследование этих данных с использованием ППП Statgraphics. Методами дисперсионного анализа построены расслоения исходных данных по годам, а затем по месяцам. Это дало возможность провести дальнейший анализ годовых данных методами регрессионного анализа. Для выявления сезонных колебаний использовались также методы анализа временных рядов (time series).

Удалось проанализировать характерные особенности соответствующих рядов динамики, выявить вызывающие их причины, наметить пути улучшения показателей. Построенные статистические модели отражают тенденции, определяющие рост числа иностранных студентов, приезжающих на довузовский и вузовский этапы обучения в Волгоградский государственный технический университет.

Список литературы

1. Бююль А., Цефель П. SPSS: искусство обработки информации: пер. с нем. – Москва-СПб.: ДиаСофт. 2002. – 602 с.

2. Наследов А. SPSS 19 профессиональный статистический анализ данных. – Москва-СПб.: Питер. 2011. – 400 с.

3. Лунева И.Г., Симонова И.Э. Основные направления компьютеризации математического образования будущего инженера: компетентностный подход. Приборы и системы. Управление, контроль, диагностика. – 2012. – № 1. – С. 55–58.

4. Савченко О. А., Симонов А.Б., Симонова И.Э. Применение методов анализа рядов динамики в стратегическом управлении предприятием // Инновации, качество и сервис в технике и технологиях. Сб. научн. статей IV Международной научно-практической конференции (Россия, г. Курск, 04–05 июня 2014 г.) / Курск, Юго -Запад. ГУ, 2014. – С. 273–277.

5. Симонова И.Э., Симонов А.Б., Тарасова И.А. Применение моделей временных рядов в стратегическом управлении предприятиями региона // Управление стратегическим потенциалом регионов России: методология, теория, практика. Материалы Всерос. науч.-практ. конф. (Россия, г. Волгоград, 16–18 апр. 2014 г.). – Волгоград, ВолгГТУ, 2014. – С. 209–212.

6. Щукина Н.А., Тарасова И.А. Моделирование динамики продаж новых легковых автомобилей в России // Фундаментальные исследования. – 2016. – № 10–3. – С. 672–677.

Referenses

1. Bjujul A., Cefel P. SPSS: iskusstvo obrabotki informacii: per. s nem. Moskva-SPb.: DiaSoft. 2002. 602 p.

2. Nasledov A. SPSS 19 professionalnyj statisticheskij analiz dannyh. Moskva-SPb.: Piter. 2011. 400 p.

3. Luneva I.G., Simonova I.Je. Osnovnye napravlenija kompjuterizacii matematicheskogo obrazovanija budushhego inzhenera: kompetentnostnyj podhod. Pribory i sistemy. Upravlenie, kontrol, diagnostika. 2012. no. 1. pp. 55–58.

4. Savchenko O. A., Simonov A.B., Simonova I.Je. Primenenie metodov analiza rjadov dinamiki v strategicheskom upravlenii predprijatiem // Innovacii, kachestvo i servis v tehnike i tehnologijah. Sb. nauchn. statej IV Mezhdunarodnoj nauchno-prakticheskoj konferencii (Rossija, g. Kursk, 04–05 ijunja 2014 g.) / Kursk, Jugo -Zapad. GU, 2014. pp. 273–277.

5. Simonova I.Je., Simonov A.B., Tarasova I.A. Primenenie modelej vremennyh rjadov v strategicheskom upravlenii predprijatijami regiona // Upravlenie strategicheskim potencialom regionov Rossii: metodologija, teorija, praktika. Materialy Vseros. nauch.-prakt. konf. (Rossija, g. Volgograd, 16–18 apr. 2014 g.). Volgograd, VolgGTU, 2014. pp. 209–212.

6. Shhukina N.A., Tarasova I.A. Modelirovanie dinamiki prodazh novyh legkovyh avtomobilej v Rossii // Fundamentalnye issledovanija. 2016. no. 10–3. pp. 672–677.