

УДК 311.4

К.А. Метешкин<sup>1</sup>, О.И. Морозова<sup>2</sup>, А.Ю. Кащавцева<sup>1</sup><sup>1</sup> Харьковская национальная академия городского хозяйства, Харьков<sup>2</sup> Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков

## СТАТИСТИЧЕСКИЙ АНАЛИЗ ПРОФЕССИОНАЛЬНО-ОРИЕНТИРОВАННОЙ РАБОТЫ НА КАФЕДРЕ

В статье исследуются вопросы, связанные со статистическим анализом профессионально-ориентированной работы на кафедре. Для примера рассматривается кафедра «Геоинформационные системы и геодезии» Харьковской национальной академии городского хозяйства. Целью данного научного исследования является прогнозирование потока абитуриентов и востребованности специальности «Геоинформационные системы и технологии» с использованием web-технологий.

**Ключевые слова:** статистический анализ, профессионально-ориентированная работа, Google Analytics, web-технологии, системы поддержки образовательных процессов, корреляционный анализ.

### Введение

На сегодняшний день актуальной для ВУЗов является задача набора абитуриентов на новый учебный год. Три года назад высшие учебные заведения Украины столкнулись с проблемой снижения количества абитуриентов. Данная тенденция вызвана снижением рождаемости, которую повлекли ряд глобальных общественно-экономических факторов. На графике (рис. 1) представлена динамика рождаемости в Украине в период с 1989 по 2006 год по данным Государ-

ственного комитета статистики Украины [1]. Данная демографическая ситуация повлияла на то, что перед каждым вузом Украины возникла задача набора студентов. И данная проблема, как видно из графика, будет актуальна еще ближайшие семь лет. Рынок образования в Украине весьма велик, по данным Государственного комитета статистики, сегодня в Украине функционируют 854 высших учебных заведения, из них 345 – III-IV уровней аккредитации, в том числе 228 государственной формы собственности. Исследуем предпочтения абитуриентов.

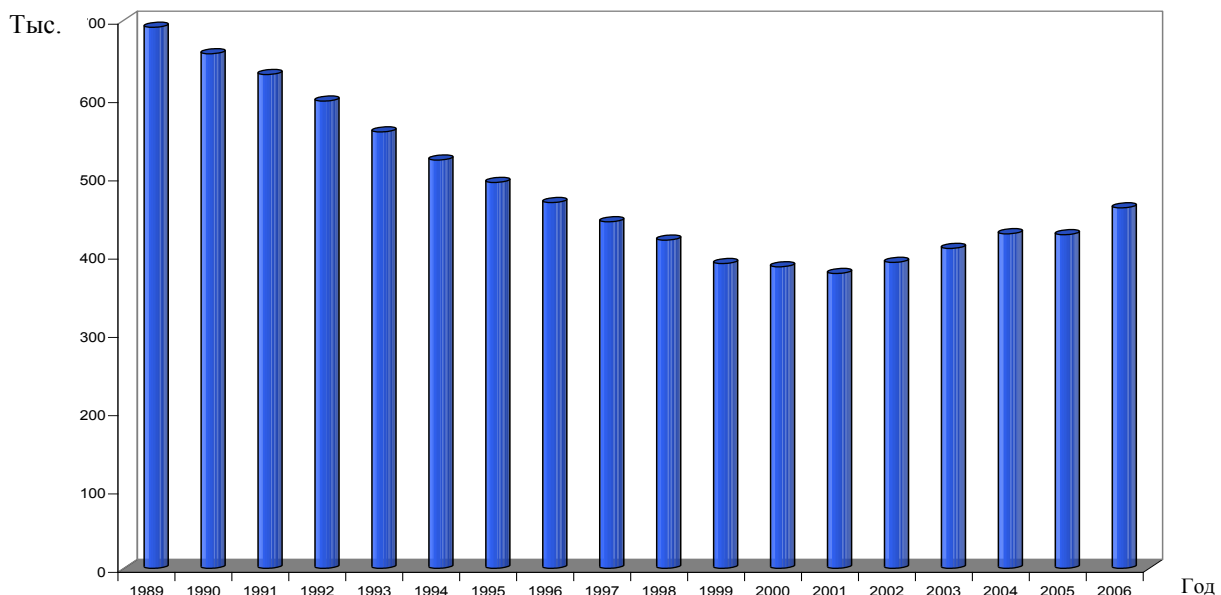


Рис. 1. Динамика рождаемости в Украине (тысяч человек)

При выборе специальности огромное влияние оказывают личные предпочтения абитуриента, его интересы и увлечения. При выборе высшего учебного заведения большую роль играет его престиж и стоимость обучения. Однако сориентироваться в многообразии профессий и знаний, которые можно получить, сейчас достаточно сложно. На помощь

абитуриенту при выборе профессии приходят друзья и родственники, и вместе они решают сложную многокритериальную задачу оценки и выбора ВУЗа и специальности. Одним из действенных источников информации для принятия правильного решения для абитуриента является информация, размещаемая на сайтах ВУЗов и соответствующих кафедр.

## 1. Системы поддержки образовательных процессов

Отдельные кафедры разрабатывают на базе web-технологий системы поддержки образовательных процессов, которые реализуются в виде динамических сайтов, т.е. управляемых сайтов, отражающих в динамике профессиональную деятельность преподавателей и кафедры в целом. Для абитуриента информация с таких сайтов является наиболее ценной, так как они, как будущие студенты, могут оценить учебную среду, в которой придется им учиться. Обычно такие сайты снабжаются специальным приложением Google Analytics. Это инструментальное средство веб-аналитики корпоративного уровня позволяет оценить трафик на веб-сайт и эффективность различных маркетинговых мероприятий. Мощные, гибкие и очень простые в использовании функции гарантируют удобство просмотра и анализа данных по трафику [2].

Данное приложение позволяет оценивать посещение сайта пользователями, большую часть которых составляют абитуриенты в период выбора вуза и специальности обучения. Приложение Google

Analytics позволяет оценить посещение каждой страницы сайта. Это дает возможность выделить основные, с точки зрения абитуриентов, страницы сайта и проанализировать частоту их посещения в летний период, когда они решают задачу выбора вуза и специальности.

На кафедре геоинформационных систем и геодезии Харьковской национальной академии городского хозяйства разработана система поддержки образовательных процессов [3], с помощью которой осуществлялось исследование процесса выбора абитуриентами ВУЗа и специальности обучения. Внешний вид его главной страницы с примером вызова информации о посещении сайта приведен на рис. 2.

Использование инструментальных средств сбора статистических данных о посещении пользователями страниц сайта актуализировало задачу прогнозирования процессов набора студентов для обучения по профилю кафедры.

Учитывая вышеизложенное, возникла задача – на основе данных, полученных с помощью Google Analytics, проанализировать эффективность использования сайта при профессионально-ориентированной работе с абитуриентами.



Рис. 2. Иллюстрация главной страницы сайта с информацией о ее посещении

## 2. Определение эффективности использования сайта при профессионально-ориентированной работе с абитуриентами

С помощью данного приложения были проанализированы посещения сайта пользователями в те-

чение некоторого времени – одного месяца, а именно 9 следующих web-страниц, которые мы представим в виде переменных: «Как поступить учиться» (переменная А); «Главная страница» (переменная В); «Направления подготовки» (переменная С); «Как нас найти» (переменная D); «Трудоустройство» (переменная Е); «Научно-педагогические работники» (пе-

ременная F); «Перспективы развития кафедры» (переменная G); «Открытые занятия» (переменная H); «Мы стоим на плечах гигантов» (переменная I). Данные о посещениях пользователями страниц сайта табулированы и представлены в виде таблиц. Рассмотрим содержание таблицы на примере одной из переменных, а именно переменной В – «Главная

страница» (см. табл. 1). Посещение страниц сайта было проанализировано с помощью методов математической статистики. Оценка на основе вычисления коэффициентов корреляции является одним из самых востребованных методов математической статистики в психологических и педагогических исследованиях [4 – 6].

Таблица 1

Статистика посещение главной страницы сайта

Дата	Кол-во посетителей	Дата	Кол-во посетителей	Дата	Кол-во посетителей
05.05.12	11	16.05.12	7	27.05.12	5
06.05.12	13	17.05.12	7	28.05.12	13
07.05.12	22	18.05.12	0	29.05.12	11
08.05.12	10	19.05.12	6	30.05.12	6
09.05.12	6	20.05.12	6	31.05.12	7
10.05.12	10	21.05.12	6	01.06.12	8
11.05.12	3	22.05.12	7	02.06.12	11
12.05.12	5	23.05.12	13	03.06.12	4
13.05.12	2	24.05.12	5	04.06.12	3
14.05.12	6	25.05.12	1	05.06.12	8
15.05.12	3	26.05.12	4	06.06.12	1

Коэффициент корреляции показывает степень статистической зависимости между двумя числовыми переменными. Он вычисляется таким образом:

$$r = \frac{\text{cov}(X, Y)}{\sqrt{D[X] \cdot D[Y]}}, \quad (1)$$

где  $\text{cov}(X, Y)$  – ковариация величин  $X$  и  $Y$ ,  $D[X]$ ,  $D[Y]$  – дисперсия величин  $X$  и  $Y$ .

Для вычисления коэффициента корреляции необходимо определить две величины: ковариацию и дисперсию.

Используем тот факт, что ковариация в теории вероятностей – это мера линейной зависимости случайных величин, которая вычисляется по формуле:

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)], \quad (2)$$

где  $E$  – математическое ожидание.

Как известно из теории вероятностей, математическое ожидание – это среднее значение случайной величины в теории вероятностей. И теперь осталось вычислить дисперсию. Дисперсия случайной величины – это мера разброса данной случайной величины, т.е. отклонение от математического ожидания. Дисперсия вычисляется по формуле:

$$D[X] = E[(X - EX)^2]. \quad (3)$$

Если коэффициент корреляции близок к 1, то между переменными наблюдается положительная корреляция. Иными словами, отмечается высокая степень связи входной и выходной переменных. В данном случае, если значения входной переменной  $x$  будут возрастать, то и выходная переменная также будет увеличиваться.

Если коэффициент корреляции близок к -1, это означает, что между переменными наблюдается отрицательная корреляция. Иными словами, поведение выходной переменной будет противоположным поведению входной. Если значение  $x$  будет возрастать, то  $y$  будет уменьшаться, и наоборот. Промежуточные значения, близкие к 0, будут указывать на слабую корреляцию между переменными и, соответственно, низкую зависимость. Иными словами, поведение входной переменной  $x$  не будет совсем (или почти совсем) влиять на поведение  $y$ .

Используя формулы (1) – (3), вычислили значение коэффициента корреляции для переменных  $A$  и  $B$ , расчет приведен ниже:

$$\text{cov}(A, B) = 9.91,$$

$$D[A] = 10.90,$$

$$D[B] = 19.72,$$

$$r_{AB} = 0.6762.$$

В соответствии с вышеизложенным, для каждой пары переменных для страниц сайта был рассчитан коэффициент корреляции.

Каждое полученное значение оценено по  $t$ -критерию Стьюдента.

Критерий Стьюдента направлен на оценку различий величин средних  $\bar{X}$  и  $\bar{Y}$  двух выборок  $X$  и  $Y$ , которые распределены по нормальному закону. Одним из главных достоинств критерия является широта его применения. Он может быть использован для сопоставления средних у связанных и несвязанных выборок, причем выборки могут быть не равны по величине. Значение  $t$ -критерия Стьюдента определяется по формуле:

$$t = \frac{r_{xy} \sqrt{n-2}}{\sqrt{1-r_{xy}^2}} \quad (4)$$

Критическое значение t-критерия Стьюдента принято 2.042, что соответствует уровню значимости 0.05. Результаты вычислений представлены в виде матрицы корреляции (см. табл. 2).

Жирным шрифтом в таблице выделены значения корреляции, t-критерий которых выше указанного критического значения, т.е. можно говорить, что данные страницы сайта взаимосвязаны.

Для подтверждения результатов данного исследования во время вступительной компании абитуриентам, подававшим документы на специальность «Геоинформационные системы и техноло-

гии», была предложена анкета, в которой необходимо было отметить – какие страницы сайта кафедры они посещали. В результате опроса было выяснено, что 81 % абитуриентов посещали страницу «Как поступить учиться», 74 % заинтересовались направлениями подготовки и 66 % заинтересовались страницей «Как нас найти», также следует отметить, что 50% абитуриентов зашедших на сайт изучала страницу о трудоустройстве.

Таким образом, если сопоставить результаты корреляционного анализа данных, полученных с помощью Google Analytics о трафике на веб-сайте кафедры, и данных анкетирования абитуриентов, можно сделать вывод, что они взаимосвязаны и подтверждают друг друга.

Таблица 2

Матрица корреляции посещения страниц сайта

	A	B	C	D	E	F	G	H	I
A	0	<b>0.6762</b>	0.1305	<b>0.5350</b>	<b>0.5341</b>	0.0928	0.1155	-0.0989	0.0383
B	<b>0.6762</b>	0	0.3342	<b>0.5584</b>	<b>0.6408</b>	0.1254	0.2981	0.1486	0.0702
C	0.1305	0.3342	0	<b>0.3777</b>	<b>0.4098</b>	<b>0.4976</b>	0.2530	<b>0.4091</b>	0.2763
D	<b>0.5350</b>	<b>0.5584</b>	<b>0.3777</b>	0	<b>0.7447</b>	0.2131	0.2871	0.2134	-0.0584
E	<b>0.5341</b>	<b>0.6408</b>	<b>0.4098</b>	<b>0.7447</b>	0	<b>0.3997</b>	0.2901	0.1812	0.2296
F	0.0928	0.1254	<b>0.4976</b>	0.2131	<b>0.3997</b>	0	<b>0.3790</b>	0.1228	0.2065
G	0.1155	0.2981	0.2530	0.2871	0.2901	<b>0.3790</b>	0	0.1155	0.1989
H	-0.0989	0.1486	<b>0.4091</b>	0.2134	0.1812	0.1228	0.1155	0	<b>0.3773</b>
I	0.0383	0.0702	0.2763	-0.0584	0.2296	0.2065	0.1989	<b>0.3773</b>	0

## Заключение

Проведя статистический анализ профессионально-ориентированной работы и проанализировав результаты корреляционного анализа данных, полученных с помощью Google Analytics, и данных анкетирования абитуриентов было получено, что еще один вид профессионально-ориентированной работы, а именно с использованием web-технологий, имеет право на жизнь и является весьма эффективным.

## Список литературы

1. Державна служба статистики України [Електронний ресурс]. – Режим доступа к ресурсу: <http://www.ukrstat.gov.ua/>. – 30.09.2012.
2. Веб-аналитика корпоративного уровня [Електронний ресурс]. – Режим доступа к ресурсу: <http://www.google.com/intl/ru/analytics/>. – 30.09.2012.

3. Сайт кафедры ГИС и геодезии Харьковской национальной академии городского хозяйства [Електронний ресурс]. – Режим доступа к сайту: <http://www.kaf-gis.kh.ua/>. – 30.09.2012.

4. Попов О.А. Коэффициент корреляции [Електронний ресурс] / О.А. Попов. – Режим доступа к ресурсу: <http://psystat.at.ua/publ/1-1-0-17>. – 30.09.2012.

5. Гайдышев И.П. Анализ и обработка данных [Текст] / И.П. Гайдышев. – СПб.: Питер, специальный справочник, 2001. – 750 с.

6. Гельман В.Я. Решение математических задач средствами Excel: практикум [Текст] / В.Я. Гельман. – СПб.: Питер, 2003. – 240 с.

Поступила в редколлегию 15.10.2012

Рецензент: д-р техн. наук, проф. О.Е. Федорович, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.

## СТАТИСТИЧНИЙ АНАЛІЗ ПРОФЕСІЙНО-ОРІЄНТОВАНОЇ РОБОТИ НА КАФЕДРІ

К.О. Метешкін, О.І. Морозова, А.Ю. Кашавцева

У статті досліджуються питання, пов'язані зі статистичним аналізом професійно-орієнтованої роботи на кафедрі. Для прикладу розглядається кафедра «Геоінформаційні системи і геодезії» Харківської національної академії міського господарства. Метою даного наукового дослідження є прогнозування потоку абитуриентів і затребуваності спеціальності «Геоінформаційні системи і технології» з використанням веб-технологій.

**Ключові слова:** статистичний аналіз, професійно-орієнтована робота, Google Analytics, веб-технології, системи підтримки освітніх процесів, кореляційний аналіз.

---

**STATISTICAL ANALYSIS OF PROFESSION-ORIENTED WORK IN THE DEPARTMENT**

K.O. Meteshkin, O.I. Morozova, A.Ur. Kaschavtseva

*This article investigates issues related to the statistical analysis of professional-oriented work in the department. The department of "Geodesy and Geographic Information Systems" of Kharkov national academy of municipal economy is considered for an example. The purpose of this research is to predict the flow of students and demand of the specialty "Geographic Information Systems and Technologies" with using of web-technologies.*

**Keywords:** *statistical analysis, professionally-oriented work, Google Analytics, web-technologies, systems support educational processes, correlation analysis.*