

Интернет-журнал «Мир науки» / World of Science. Pedagogy and psychology <https://mir-nauki.com>

2018, №4, Том 6 / 2018, No 4, Vol 6 <https://mir-nauki.com/issue-4-2018.html>

URL статьи: <https://mir-nauki.com/PDF/64PDMN418.pdf>

Статья поступила в редакцию 20.08.2018; опубликована 08.10.2018

Ссылка для цитирования этой статьи:

Никаноркина Н.В., Степанова А.С., Гераева Е.В. Статистическое исследование уровня знаний студентов в области информационных технологий // Интернет-журнал «Мир науки», 2018 №4, <https://mir-nauki.com/PDF/64PDMN418.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ.

For citation:

Nikanorkina N.V., Stepanova A.S., Geraeva E.V. (2018). Statistical survey of the level of knowledge of students in information technology. *World of Science. Pedagogy and psychology*, [online] 4(6). Available at: <https://mir-nauki.com/PDF/64PDMN418.pdf> (in Russian)

УДК 378.1

ГРНТИ 14.35.07

Никаноркина Наталия Владимировна

ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»
Филиал в г. Калуга, Россия
Доцент кафедры «Высшая математика и статистика»
Кандидат педагогических наук
E-mail: nika4061@yandex.ru
РИНЦ: http://elibrary.ru/author_profile.asp?id=691631

Степанова Анна Сергеевна

ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»
Филиал в г. Калуга, Россия
Студент
E-mail: a3stepanova@gmail.com
ORCID: <http://orcid.org/0000-0002-2697-8205>
РИНЦ: http://elibrary.ru/author_profile.asp?id=974504

Гераева Екатерина Васильевна

ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации»
Филиал в г. Калуга, Россия
Студент
E-mail: geraeva.katya@yandex.ru

**Статистическое исследование уровня знаний
студентов в области информационных технологий**

Аннотация. В статье представлены результаты тестирования знаний студентов экономического вуза различных направлений и профилей подготовки в области современных информационных компьютерных технологий, а также результаты обработки полученных данных и их анализ. Авторами статьи была создана опросная онлайн-форма, случайным образом отобраны студенты для тестирования, проведено тестирование и проанализированы его результаты с целью выявления закономерностей в распределении знаний студентов по информатике. Первичная обработка полученных данных производилась с помощью возможностей языка программирования Python для систематизации и визуализации эмпирического материала. Применение методов математической статистики позволило авторам представить данные тестирования в виде интервального статистического ряда. На основе выборочных данных были выдвинуты две статистические гипотезы: первая – о том, что

случайная величина количества правильных ответов студентов на тест подчиняется нормальному закону распределения, вторая – о влиянии курса обучения студента в университете на его общие представления и знания в сфере использования современных информационных компьютерных технологий. Для проверки этих статистических гипотез были выбраны наиболее мощные статистические критерии. Первая статистическая гипотеза, относящаяся к группе непараметрических, проверялась с помощью критерия согласия Пирсона. Для проверки второй, параметрической, статистической гипотезы был использован критерий Стьюдента (t-критерий). В ходе исследования авторы статьи пришли к выводу, что распределение общих знаний студентов об информационных технологиях подчиняется нормальному закону, а также отклонено статистическое предположение о существенной зависимости уровня знаний студентов от курса обучения.

Ключевые слова: обучение студентов; тестирование; обработка данных Python; анализ данных; информационные технологии; методы математической статистики; статистическая гипотеза; нормальное распределение; критерий согласия

Введение

Современный период развития цивилизованного общества характеризуется бурным процессом информатизации, который связан с внедрением информационных средств и технологий в науку, производство, образование, быт и т. д. Информатизация общества представляет собой глобальный социальный процесс, особенностью которого является то, что доминирующим видом деятельности в сфере общественного производства является умение работать с информацией с помощью различных средств информационного обмена [1]. Современному специалисту любого профиля важно уметь свободно ориентироваться в информационных потоках, уметь получать, обрабатывать и использовать информацию с помощью компьютеров и других технических средств.

Активными участниками процесса информатизации общества являются студенты. Выпускнику в будущем предстоит существовать в социуме, в котором умение использовать современные информационные технологии позволит продуктивно решать широкий круг профессиональных задач, и как следствие, будет во многом определять его жизненный и профессиональный успех. При этом научиться использовать информационные технологии можно, лишь активно применяя их в учебном процессе [2, 3].

В контексте сказанного особую актуальность приобретает проблема выявления уровня знаний студентов и степени владения ими современными компьютерными информационными технологиями.

Целью исследования является выявление закономерностей в распределении знаний студентов по информатике и определение степени влияния курса, на котором обучается студент, на уровень общих представлений и знаний в области информационных компьютерных технологий.

Материалы и методы

Оценка качества обучения с помощью различных технологий тестирования является одним из основных способов оценивания качества образования [4]. В связи с этим для достижения поставленной цели была выбрана технология онлайн-тестирования, как наиболее эффективный и быстрый способ получить необходимую информацию.

Измерение уровня знаний студентов в информационной сфере было проведено с помощью онлайн-тестирования средствами универсальной площадки по созданию онлайн-опросов – Google Формы. Данный инструмент позволяет создавать гибкие опросные формы, связанные с онлайн-документом. Созданная опросная форма включала 10 вопросов разного уровня сложности, но рассчитанных на общие знания в области информационных технологий. Опрос проводился в условиях полной конфиденциальности и анонимности опрашиваемых.

В проводимом исследовании роль генеральной совокупности играли все студенты бакалавриата Калужского филиала Финансового университета при Правительстве Российской Федерации (1-4 курсы).

Для проведения тестирования случайным образом из генеральной совокупности было отобрано 100 студентов 1-3 курсов (что составляет около 20 % от общего количества студентов 1-3 курсов очной формы обучения), обучающихся на разных направлениях и профилях бакалавриата. В используемый нами вариант теста было включено 10 заданий. Все они имеют закрытую форму и сопровождаются четырьмя альтернативными ответами, один из которых верный, а остальные правдоподобные, но ошибочные. В ходе выполнения заданий студенты указывали номера верных по их мнению ответов.

Вопросы теста не требовали узкопрофессиональных знаний студентов в области информационных компьютерных технологий. В тесте были представлены вопросы о функциях оперативной системы, об интерфейсе и работе текстового редактора Word, о горячих клавишах Windows, о поисковых системах сети Интернет, об антивирусных программах и т. п.

Результаты исследования и их обсуждение

В результате проведённого опроса были получены данные, которые в дальнейшем были систематизированы и обработаны средствами языка программирования Python и его библиотек [5, 6], а затем исследованы и проанализированы методами математической статистики [7-10].

Первоначально данные из Google Форм были выведены в таблицу Excel, которая была сохранена в формате CSV (это текстовый формат для представления табличных данных, от английского *Comma-Separated Values* – значения, разделённые запятыми). Затем информация из CSV-таблицы была загружена в двумерный массив table (листинг 1).

```
Infile= open('Book1.csv', 'r')
table = []
for row in csv.reader(infile):
    table.append(row)
infile.close()
```

Листинг 1. Загрузка данных из CSV-таблицы (составлено авторами)

После этого была проведена обработка в программе ответов на тест таким образом, что правильному ответу присвоено значение «1», а неправильному – «0» (листинг 2).

```
For r in range(0,len(table)):  
if table[r][7] = 'распределение ресурсов':  
table[r][7] = '1'  
else:  
table[r][7] = '0'
```

Листинг 2. Обработка ответов на вопросы теста (составлено авторами)

Следующим шагом обработки данных был подсчёт количества правильных ответов каждого студента (листинг 3) и представление этих результатов графически в виде гистограммы (рис. 1).

```
row = [0.0]*len(table)  
for c in range(0, len(row)):  
s = 0  
for r in range(7, 16):  
s += int(table[c][r])  
row[c] = s  
table.append (row)  
plt.ylabel ('количество правильных ответов')  
plt.plot(row)  
plt.show()
```

Листинг 3. Сумма правильных ответов для каждого студента (составлено авторами)

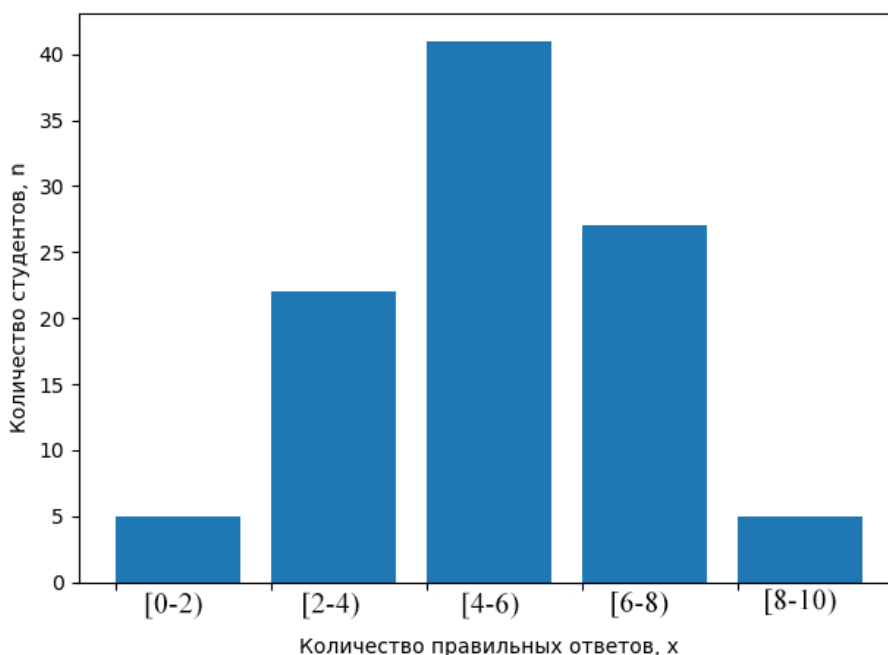


Рисунок 1. Гистограмма частот правильных ответов студентов (составлено авторами)

На основе имеющихся данных нами был построен интервальный статистический ряд количества правильных ответов студентов (случайной величины X), который представлен в таблице 1.

Таблица 1

Интервальный статистический ряд количества правильных ответов студентов

Количество правильных ответов, $[x_i - x_{i+1})$	[0-2)	[2-4)	[4-6)	[6-8)	[8-10)
Частота, n_i	5	22	41	27	5
Середина интервала, x'_i	1	3	5	7	9

Составлено авторами

По виду гистограммы распределения изучаемой случайной величины X , представленной на рисунке 1, можно выдвинуть предположение о нормальном законе распределения признака X .

Для дальнейшей обработки данных проверим, согласуется ли высказанное предположение о генеральной совокупности с эмпирическим распределением выборки, полученной из этой совокупности. Проверку проведем с помощью критерия согласия Пирсона (критерия «хи»-квадрат) [7].

При использовании критерия согласия в основной гипотезе H_0 утверждается, что эмпирическое и теоретическое распределения согласованы, т. е. в качестве основной гипотезы примем предположение о том, что случайная величина X подчиняется нормальному закону распределения с параметрами a и σ (символически это можно записать так $H_0: X \sim N(a, \sigma)$).

Альтернативная гипотеза H_1 по умолчанию предполагает противоположное – распределения не согласованы (символически $H_1: X \sim \bar{N}(a, \sigma)$).

Уровень значимости выбран стандартно равным 5 %, т. е. $\alpha = 0,05$.

Необходимым условием применения критерия согласия Пирсона является наличие в каждом из интервалов не менее 5 наблюдений, т. е. $n_i > 5$, [8]. В данном случае, как видно из таблицы 1, количество наблюдений в каждом интервале не меньше 5, поэтому можно воспользоваться критерием согласия Пирсона.

Применение критерия «хи»-квадрат предполагает необходимость вычисления вероятностей попадания случайной величины X в каждый из интервалов, для чего мы предварительно вычислили несмещенные точечные оценки параметров нормального распределения a и σ , а именно выборочную среднюю \bar{x}_s и исправленное выборочное среднее квадратическое отклонение s .

Выборочная средняя определяется по формуле:

$$a^* = \bar{x}_s = \frac{\sum x_i n_i}{n},$$

где x_i – значение признака, n_i – частота значения x_i в выборке.

Получим оценку генеральной средней или параметра a : $a^* = \frac{510}{100} = 5,1$.

Исправленное среднее квадратическое отклонение вычисляется так:

$\sigma^* = S = \sqrt{S^2} = \sqrt{\frac{n}{n-1} D_s}$, где S^2 – исправленная выборочная дисперсия, D_s – выборочная дисперсия.

Вычислим сначала выборочную дисперсию D_e , воспользовавшись формулой:

$$D_e = \bar{x}_e^2 - [\bar{x}_e]^2 = \sum \frac{x_i^2}{n} - [\bar{x}_e]^2 = 29,56 - 26,01 = 3,55$$

Тогда исправленная выборочная

дисперсия равна $S^2 = \frac{100}{99} 3,55 \approx 3,586$. Далее найдем значение исправленного среднего квадратического отклонения, которое и будет являться искомой оценкой параметра σ , т. е. $\sigma^* = S = \sqrt{3,586} \approx 1,89$.

Проведенные расчеты для параметров нормального распределения позволили нам переформулировать выдвигаемую основную гипотезу следующим образом: случайная величина X распределена нормально с параметрами $a = 5,1$ и $\sigma = 1,89$ (или символически $H_0: X \sim N(5,1, 1,89)$).

Далее мы определили теоретические вероятности p_i попадания случайной величины X в каждый из интервалов $[x_i; x_{i+1})$ по формуле:

$$p_i = P(\alpha < X < \beta) = \Phi_0\left(\frac{\beta - a}{\sigma}\right) - \Phi_0\left(\frac{\alpha - a}{\sigma}\right) = \Phi_0\left(\frac{\beta - \bar{x}_e}{s}\right) - \Phi_0\left(\frac{\alpha - \bar{x}_e}{s}\right),$$

где $\Phi_0(x)$ – нормированная функция Лапласа, значения которой определяются по соответствующей таблице.

Так как случайная величина X , распределённая по нормальному закону, определена на всей числовой прямой $(-\infty; +\infty)$, то крайние интервалы в ряде распределения следовало сделать открытыми: $[0; 2) \rightarrow (-\infty; 2)$ и $[8; 10) \rightarrow (8; +\infty)$.

Используя таблицу нормированной функции Лапласа [7], находим:

$$p_1 = P(-\infty < X < 2) = \Phi_0\left(\frac{2 - 5,1}{1,89}\right) - \Phi_0(-\infty) \approx 0,0505$$

$$p_2 = P(2 < X < 4) = \Phi_0\left(\frac{4 - 5,1}{1,89}\right) - \Phi_0\left(\frac{2 - 5,1}{1,89}\right) \approx 0,23046$$

$$p_3 = P(4 < X < 6) = \Phi_0\left(\frac{6 - 5,1}{1,89}\right) - \Phi_0\left(\frac{4 - 5,1}{1,89}\right) \approx 0,4068$$

$$p_4 = P(6 < X < 8) = \Phi_0\left(\frac{8 - 5,1}{1,89}\right) - \Phi_0\left(\frac{6 - 5,1}{1,89}\right) \approx 0,25619$$

$$p_5 = P(8 < X < +\infty) = \Phi_0(+\infty) - \Phi_0\left(\frac{8 - 5,1}{1,89}\right) \approx 0,06301$$

Условие $\sum_{i=1}^n p_i = 1$ выполняется.

Полученные результаты вычислений сведем в таблицу 2.

Таблица 2

Интервальный статистический ряд эмпирических и теоретических частот

$[x_i; x_{i+1})$	$(-\infty; 2)$	$[2; 4)$	$[4; 6)$	$[6; 8)$	$[8; \infty)$
n_i	5	22	41	27	5
$n_i' = np_i$	5,05	23,05	40,68	25,62	6,3

Составлено авторами

Затем мы вычислили наблюдаемое значение статистики критерия по формуле:

$$\chi_n^2 = \sum_{i=1}^m \frac{n_i^2}{n_i'} - n$$

где n_i – эмпирическая частота варианты x_i , $n_i' = p_i \cdot n$ – теоретическая частота, n – объем выборки. В итоге получили $\chi_n^2 \approx 0,2425$.

Далее найдем критическую точку, воспользовавшись таблицей распределения Пирсона [7], по уровню значимости $\alpha = 0,05$ и числу степеней свободы $k = m - r - 1$, где m – количество интервалов наблюдения, r – число параметров предполагаемого закона распределения. В данном случае $m = 5$, $r = 2$ (т. к. нормальный закон распределения задается двумя параметрами). Следовательно, число степеней свободы равно $k = m - r - 1 = 5 - 2 - 1 = 2$. В итоге получим $\chi_{кр.}^2 = (0,05; 2) = 5,991$.

Так как наблюдаемое значение не превосходит критического, т. е. $\chi_n^2 = 0,2425 < \chi_{кр.}^2 = 5,991$, то, согласно критерию Пирсона, основная гипотеза не противоречит опытным данным, и нет оснований ее отвергнуть. Гипотеза о выбранном теоретическом нормальном законе распределения $H_0: X \sim N(5,1,1,89)$ согласуется с опытными данными.

Таким образом, результаты проведенной проверки позволяют сделать следующий вывод: на уровне значимости 5 % можно утверждать, что полученное нами эмпирическое распределение согласуется с нормальным законом, т. е. количество правильных ответов студентов на тест описывается нормальным законом распределения.

Сказанное выше наглядно иллюстрирует одно из применений методов математической статистики в проверке предположений о виде неизвестного закона распределения случайной величины, т. е. в проверке непараметрических гипотез. Кроме того, полученные в результате проведенного тестирования данные можно использовать для проверки еще одного типа гипотез, которые называются параметрическими. В том случае, когда распределение случайной величины известно, по выборке наблюдений можно проверять предположения о значении параметров этого распределения [7].

Продemonстрируем сказанное на примере гипотезы о зависимости между курсом обучения студента и уровнем его общих знаний по информатике, т. е. сформулируем и проверим гипотезу о равенстве генеральных средних по двум независимым выборкам из нормальных совокупностей.

С этой целью необходимо сначала сформировать две генеральные совокупности X и Y . Мы разбили всех студентов филиала университета на студентов первого курса (это первая группа студентов) и студентов второго и третьего курсов (это вторая группа студентов). Генеральная совокупность (или случайная величина) X обозначает количество правильных ответов студентов первого курса. Количество правильных ответов студентов второго и третьего курсов есть вторая генеральная совокупность (или случайная величина) Y . Исследование будем

проводить в предположении, что обе генеральные совокупности имеют нормальный закон распределения.

Результаты опроса студентов первого курса (выборка из генеральной совокупности X) были обработаны и представлены в виде интервального статистического ряда (таблица 3). Общее количество студентов первого курса в выборке составило 52, т. е. объем выборки из генеральной совокупности X: $n_x = 52$.

Таблица 3

**Интервальный статистический ряд
количества правильных ответов студентов первого курса**

$[x_i-x_{i+1})$	[0-2)	[2-4)	[4-6)	[6-8)	[8-10)
Частота, n_i	2	11	21	15	3
Середина интервала, x_i'	1	3	5	7	9

Составлено авторами

Аналогичным образом была обработана выборка из генеральной совокупности Y объемом $n_y = 48$ и представлена в виде интервального статистического ряда (таблица 4).

Таблица 4

**Интервальный статистический ряд количества
правильных ответов студентов второго и третьего курсов**

$[y_i-y_{i+1})$	[0-2)	[2-4)	[4-6)	[6-8)	[8-10)
Частота, n_i	3	11	20	12	2
Середина интервала, y_i'	1	3	5	7	9

Составлено авторами

Далее мы рассчитали выборочные характеристики (выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию) для обеих совокупностей по формулам, приведенным выше. В результате получили:

$$\bar{x}_g = 5,23, D_x = 3,485, S_x^2 = 3,553$$

$$\bar{y}_g = 4,958, D_y = 3,585, S_y^2 = 3,659$$

Для выяснения вопроса о влиянии курса обучения на качество общих знаний студентов в области информационных технологий мы выдвинули следующие нулевую и конкурирующую гипотезы:

- H_0 – курс обучения существенно не влияет на уровень общих знаний студентов по информатике, т. е. генеральные средние двух совокупностей равны. Символически $H_0: \bar{x}_{ген} = \bar{y}_{ген}$.
- H_1 – курс обучения влияет на уровень общих знаний студентов по информатике, то есть средний балл студентов 2 и 3 курса выше среднего балла студентов 1 курса, $H_1: \bar{x}_{ген} < \bar{y}_{ген}$. Гипотеза H_1 – левосторонняя.

Так как в данном случае генеральные дисперсии совокупностей не известны, то приступать к проверке гипотезы о равенстве генеральных средних двух нормальных совокупностей с неизвестными генеральными дисперсиями можно только в том случае, если генеральные дисперсии равны [9]. Поэтому сначала необходимо проверить гипотезу о

равенстве генеральных дисперсий с помощью критерия Фишера-Снедекора (F-критерия), т. е. гипотезу $H_0': \sigma_x^2 = \sigma_y^2$.

В качестве альтернативной гипотезы в таких случаях, как правило, используют правостороннюю гипотезу. В нашем исследовании это предположение о том, что генеральная дисперсия знаний студентов 1 курса больше, чем генеральная дисперсия знаний студентов 2 и 3 курсов, т. е. $H_1': \sigma_x^2 > \sigma_y^2$.

Наблюдаемое значение статистики критерия вычисляется как отношение большей исправленной выборочной дисперсии к меньшей, т. е.

$$F_n = \frac{S_{\max}^2}{S_{\min}^2} = \frac{S_y^2}{S_x^2},$$

где S_x^2 и S_y^2 – это исправленные выборочные дисперсии количества правильных ответов студентов первой и второй групп соответственно, которые вычисляются по формуле,

представленной выше. В итоге получим $F_n = \frac{3,659}{3,553} \approx 1,029$.

Так как альтернативная гипотеза правосторонняя, то и критическая область – правосторонняя. По таблице распределения Фишера-Снедекора [7] найдем правую критическую точку $F_{кр.пр.}$ на заданном уровне значимости α ($\alpha = 5\%$) с k_1 и k_2 степенями свободы (здесь $k_1 = n_2 - 1$ и $k_2 = n_1 - 1$, т. к. выборочная дисперсия второй совокупности больше, чем первой).

$$F_{кр.пр.} = F(\alpha, k_1, k_2) = F(0,05; 47; 51) = 1,6.$$

Так как наблюдаемое значение критерия меньше критического, т. е. F_n не попадает в критическую область, то основная гипотеза H_0' принимается на уровне значимости 0,05. Следовательно, согласно F-критерию, с вероятностью 95 % можно считать равными генеральные дисперсии рассматриваемых совокупностей X и Y.

Теперь можно перейти к проверке основной гипотезы H_0 о равенстве генеральных средних двух совокупностей. Для этого воспользуемся критерием Стьюдента (или t-критерием) [10].

Наблюдаемое значение статистики критерия рассчитывается по формуле:

$$t_{набл} = \frac{\bar{x}_g - \bar{y}_g}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

где $S = \sqrt{S^2}$, а S^2 – это объединенная оценка дисперсии, которая в свою очередь вычисляется следующим образом:

$$S^2 = \frac{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}{n_1 + n_2 - 2} = \frac{51 \cdot 3,553 + 47 \cdot 3,659}{52 + 48 - 2} \approx 3,604$$

$$S = \sqrt{3,604} \approx 1,898$$

Получаем наблюдаемое значение:

$$t_{\text{набл}} = \frac{5,23 - 4,958}{1,898 \sqrt{\frac{1}{52} + \frac{1}{48}}} \approx 0,716$$

Далее находим критическую точку по таблице распределения Стьюдента для односторонней критической области [7].

$$t_{\text{кр.пр.}} = t_{\text{одностор.}}(\alpha; k = n_1 + n_2 - 2) = t_{\text{одностор.}}(0,05; k = 98) = 1,66$$

Поскольку в нашем случае выбрана левосторонняя альтернативная гипотеза H_1 , то и критическая область будет левосторонняя. А значит, критическая точка будет находится слева от нуля, т. е. $t_{\text{кр.лев.}} = -t_{\text{кр.пр.}} = -1,66$.

Так как наблюдаемое значение статистики критерия превосходит критическое, т. е. $t_{\text{набл}} > t_{\text{кр.лев.}}$, то $t_{\text{набл}}$ не попадает в критическую область и основная гипотеза H_0 принимается на уровне значимости 0,05.

Таким образом, согласно критерию Стьюдента, с вероятностью 95 % можно утверждать, что нет существенных различий в уровне знаний студентов первого и последующих курсов, то есть курс обучения не влияет на уровень общих знаний студентов в области информационных технологий. Еще раз подчеркнем, что речь идет именно об общих представлениях и знаниях студентов в сфере использования информационных компьютерных технологий.

Заключение

Результаты проведенного исследования позволили сделать следующие выводы:

1. Для проведения различных онлайн-опросов и тестирований респондентов по интересующим исследователя вопросам удобно использовать универсальный онлайн-сервис под названием Google Формы, который дает возможность составлять тесты с любым количеством заданий и вопросов, с открытой и закрытой формой ответов, не требует специального программного обеспечения и позволяет оперативно получить нужную информацию.

В нашем случае был сконструирован тест, позволяющий выявить общие представления и знания студентов в области применения информационных компьютерных технологий.

2. Язык программирования Python и библиотеки для него используются для эффективного решения широкого спектра задач, связанных с анализом данных.

В нашем исследовании средства Python позволили удобным образом обработать, систематизировать и визуализировать полученные данные опроса для дальнейшей работы с ними.

3. Методы математической статистики дают возможность выявлять различные закономерности в массовых случайных явлениях по выборочным данным, выдвигать статистические гипотезы и делать обоснованные выводы относительно генеральных совокупностей.

В проведенном нами исследовании было выявлено, что уровень общих представлений и знаний студентов бакалавриата экономического вуза в области современных информационных компьютерных технологий можно описать нормальным законом распределения и отклонена статистическая гипотеза о существенных различиях в уровне знаний студентов в рассматриваемой области, обучающихся на различных курсах.

ЛИТЕРАТУРА

1. Роберт И.В. Современные информационные технологии в образовании: дидактические проблемы; перспективы использования. М.: ИИО РАО, 2010. 140 с.
2. Головина О.В., Никаноркина Н.В. Об одном из аспектов использования информационных технологий при изучении курса теории вероятностей и математической статистики в экономическом вузе // Развитие управленческих и информационных технологий, их роль в региональной экономике: сборник научных статей по материалам II Международной открытой научно-практической конференции. Москва: Издательство: ООО "ТРИП", 2016. С. 295-301.
3. Головина О.В., Зайчикова И.В., Никаноркина Н.В. К вопросу об использовании информационных технологий в преподавании курсов математической статистики и эконометрики // Калужский экономический вестник. 2016. №2. С. 56-60.
4. Ozhereleva T.A. The Use of Information Technology for Education Quality Management // European Journal of Economic Studies. 2013. V.6. №. 4. p. 221-226.
5. Маккинни У. Python и анализ данных / пер. с англ. Слинкин А.А. М.: ДМК-Пресс, 2015. 482 с.
6. Демьянов Р.С. Использование языка Python как инструмента визуализации данных временных рядов // Nauka-rastudent.ru. – 2017, №03 (039) / [Электронный ресурс] – Режим доступа. – URL: <http://nauka-rastudent.ru/39/4072/> (дата обращения 15.08.2018).
7. Мхитарян В.С. Теория вероятностей и математическая статистика: учебник. М.: Издательский центр «Академия», 2012. 416 с.
8. Кобзарь А.И. Прикладная математическая статистика. М.: ФИЗМАТЛИТ, 2012. 816 с.
9. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н., Чимитова Е.В. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография. Новосибирск: Изд-во НГТУ, 2011. 888 с.
10. Граничина О.А. Математико-статистические методы психолого-педагогических исследований. Спб.: Издательство ВВМ, 2012. 115 с.

Nikanorkina Nataliia Vladimirovna

Financial university under the government of the Russian Federation
Kaluga branch, Russia
E-mail: nika4061@yandex.ru

Stepanova Anna Sergeevna

Financial university under the government of the Russian Federation
Kaluga branch, Russia
E-mail: a3stepanova@gmail.com

Geraeva Ekaterina Vasilyevna

Financial university under the government of the Russian Federation
Kaluga branch, Russia
E-mail: geraeva.katya@yandex.ru

Statistical survey of the level of knowledge of students in information technology

Abstract. In this paper presents the results of testing the knowledge of students of economic University of different directions and profiles of training in the field of modern information computer technology, as well as the results of processing the data and their analysis. The authors of the article created an online survey form, randomly selected students for testing, tested and analyzed its results in order to identify patterns in the distribution of students knowledge in computer science. The primary processing of the data was performed using the Python programming language for systematization and visualization of empirical material. The use of mathematical statistics methods allowed the authors to present the test data in the form of an interval statistical series. On the basis of sample data was proposed for two statistical hypotheses: the first is that the covariate of the number of correct student responses on the test subject to the normal distribution, the second is about the impact of the training of the student in the University in its overall presentation and knowledge in the use of modern information and computer technologies. The most powerful statistical criteria were chosen to test these hypotheses. The first statistical hypothesis relating to the nonparametric group was tested using Pearson's goodness-of-fit test. To test the second, parametric, statistical hypothesis, Student's criterion (t-test) was used. In a study, the authors of the article came to the conclusion that the distribution of General knowledge of students about information technology is subject to the normal distribution, as well as the statistical assumption about the significant dependence of the level of knowledge of students on the course of study is rejected.

Keywords: student learning; testing; data processing Python; data analysis; information technology; methods of mathematical statistics; statistical hypothesis; normal distribution; goodness of fit