

МАТЕМАТИКА И ИНФОРМАТИКА

УДК 378.048.2

doi:10.18720/SPBPU/2/id25-410

Краснощеков Виктор Владимирович¹,

доцент, канд. техн. наук, доцент;

Семенова Наталья Валентиновна²,

доцент, канд. физ.-мат. наук, доцент;

Заиб Шах³,

магистр, студент предаспирантуры;

Зарак Али Хан⁴,

магистр, студент предаспирантуры

О КУРСЕ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ ДЛЯ ИНОСТРАННЫХ ПРЕД-ПОСТДИПЛОМНИКОВ

Россия, Санкт-Петербург,

Санкт-Петербургский политехнический университет Петра Великого,

¹krasno_vv@spbstu.ru, ²nvsemenova@spbstu.ru,

³shahzaibbutt893@gmail.com, ⁴zarakalikhan37@gmail.com

Аннотация. Авторы представили дополнительные принципы проектирования и реализации курса математической статистики для иностранных пред-постдипломников в сравнении с таким же курсом для студентов бакалавриата. В качестве иллюстрации этих идей авторы привели подробные решения двух задач математической статистики. Это проверка гипотезы о равенстве вероятностей (долей) и нахождение степени корреляционной связи при дихотомическом оценивании.

Ключевые слова: иностранные студенты, поствузовская подготовка, гипотеза о равенстве долей, корреляции, дихотомическое оценивание.

Victor V. Krasnoshchekov¹,

Docent, PhD in Mechanical Engineering, Associate Professor;

Natalia V. Semenova²,

Docent, PhD in Physics and Mathematics, Associate Professor;

Zaib Shah³,

MSc, pre-postgraduate student;

Khan Zarak Ali⁴,

MSc, pre-postgraduate student

ABOUT THE COURSE IN MATHEMATICAL STATISTICS FOR INTERNATIONAL PRE-POSTGRADUATE STUDENTS

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia,

¹ krasno_vv@spbstu.ru, ² nvsemenova@spbstu.ru,

³ shahzaibbutt893@gmail.com, ⁴ zarakalikhan37@gmail.com

Abstract. The authors presented additional principles for designing and implementing a course in mathematical statistics for international pre-postgraduate students in comparison with the same course for undergraduate students. To illustrate these ideas, the authors provided detailed solutions to two problems in mathematical statistics. This is a test of the hypothesis of equality of probabilities (percentages) and finding the degree of correlation in dichotomous assessment.

Key words: international students, postgraduate training, equal percentages hypothesis, correlations, dichotomous assessment.

ВВЕДЕНИЕ

Под пред-постдипломной подготовкой иностранных граждан понимается подготовка иностранных граждан, имеющих высшее образование, к освоению профессиональных образовательных программ на русском языке. Это предмагистерская, предаспирантская и предординаторская подготовка, которые были впервые упомянуты в распорядительной документации в Приказе Министерства науки и высшего образования Российской Федерации от 18.10.2023 № 998 (<http://publication.pravo.gov.ru/document/0001202311220006>). Этот документ зафиксировал наличие уже фактически существующего уже в ряде ведущих университетов страны отдельного направления предвузовской подготовки иностранных граждан [1, 2].

Среди публикаций первых лет реализации программ предпостдипломной подготовки превалировали работы, связанные как с необходимостью их вычленения из массива программ подготовки в вуз, так и с решением организационно-методических вопросов [3]. В последующие годы авторы рассматривали как вопросы содержания дисциплин программ пред-постдипломной

подготовки, так и проблемы адаптации предпостдипломников [4, 5]. Собственно дидактические вопросы наиболее подробно были освещены преподавателями русского языка как иностранного [6]. Настоящая работа восполняет этот пробел в отношении курса математической статистики.

Материалы и методы

Дисциплина «Теория вероятностей и математическая статистика» введена в программу предпостдипломной подготовки инженерно-технического и технологического профиля в Санкт-Петербургском политехническом университете Петра Великого (СПбПУ) по двум основным причинам. Во-первых, в СПбПУ традиционно осуществляется усиленная подготовка бакалавров всех направлений по вероятностным дисциплинам. Это значит, что иностранные пред-постдипломники должны не только ориентироваться в понятиях дисциплины, но и уметь решать задачи на уровне, востребованном в бакалавриате СПбПУ. Во-вторых, существует значительная терминологическая разница в русскоязычной и англоязычной версиях вероятностных и, особенно, статистических дисциплин [7], которую иностранным пред-постдипломникам надлежит нивелировать в ходе освоения дисциплины.

Цель работы – сформулировать основные принципы построения курса математической статистики для пред-постдипломников и дать примеры реализации этих принципов.

В СПбПУ ежегодно обучается на программах предмагистерской и предаспирантской подготовки инженерно-технического и технологического направления от 60 до 85 иностранных студентов.

Курс теории вероятностей и математической статистики для пред-постдипломников построен на основе учебного пособия [8], теоретическая часть которого имеет русско-английский [9] и русско-китайский [10] аналоги. После каждого занятия студенты получают в качестве домашнего задания индивидуальную задачу (всего пособие содержит по 80 вариантов каждой из 18 задач). Пособие ориентировано на освоение дисциплины «Теория вероятностей и математическая статистика» студентами бакалавриата.

В настоящей работе приведены два примера (задачи № 14 и № 17) реализации принципов построения курса математической статистики для студентов пред-постдипломных программ. Решения заданий на двух уровнях бакалавров и пред-постдипломников разработаны студентами программы предасpirантской подготовки 2025 года выпуска из Республики Пакистан и проанализированы в настоящей работе.

Результаты и обсуждение

Целью пред-постдипломной программы является подготовка иностранных студентов к освоению профессиональных образовательных программ магистратуры или аспирантуры на русском языке. Поэтому и в номенклатуре дисциплин программы, и в их содержании одним из ведущих является научно-исследовательский компонент [11, 12], подготовка к участию в научных мероприятиях и к командной работе в научных группах [13]. В СПбПУ и других ведущих вузах, реализующих пред-постдипломную подготовку иностранных граждан, таких как Томский политехнический университет, Казанский (Приволжский) федеральный университет, Волгоградский государственный технический университет практикуется участие пред-постдипломников в научных публикациях, например [14 – 17]. В настоящей работе мы успешно использовали потенциал предасpirантов для совершенствования методического обеспечения курса математической статистики для пред-постдипломников.

При проектировании содержания и методики реализации курсов пред-постдипломной подготовки следует придерживаться ряда дополнительных к традиционным принципов, а именно, усиливать:

- вариативность подходов и решений;
- концептуальные аспекты;
- дискуссионные аспекты;
- обобщающие аспекты.

Реализация этих принципов продемонстрирована на примерах решения задач № 14 и № 17 учебного пособия [8] на двух уровнях – уровне бакалавров и уровне пред-постдипломников.

Задание № 14. Преподаватель A поставил на экзаменах в потоке из 200 студентов 42 двойки, а преподаватель B – 64 двойки в потоке из 270 студентов. Какой преподаватель строже и почему, если точность данных $+5\%$?

Рассмотрим способ решения задачи, рекомендованный для **уровня подготовки бакалавров**. Определим значения параметров задачи.

Дано:

Преподаватель A :

Двоек: 42

Общее число студентов: 200

$$\text{Доля двоек: } P_A = \frac{42}{200} = 0,21 (21\%)$$

Точность: $\varepsilon_A = \pm 5\% = \pm 0,05$

Преподаватель B :

Двоек: 64

Общее число студентов: 270

$$\text{Доля двоек: } P_B = \frac{64}{270} = 0,237 (23,7\%)$$

Точность: $\varepsilon_B = \pm 5\% = \pm 0,05$

Доля двоек – это оценка вероятности того, что преподаватель поставит двойку на экзамене.

Доверительный интервал для оценки вероятности события A имеет вид:

$$(P_A - \varepsilon_A; P_A + \varepsilon_A) = (0,21 - 0,05; 0,21 + 0,05) = (0,16; 0,26).$$

Выдвигаем основную и альтернативную гипотезы:

Гипотеза H_0 : вероятности равны $P_A = P_B$.

Гипотеза H_1 : вероятности не равны $P_A \neq P_B$.

Правило проверки статистической гипотезы о равенстве или неравенстве долей таково: если оценка вероятности события B принадлежит доверительному интервалу оценки вероятности события A :

$$P_B \in (P_A - \varepsilon_A; P_A + \varepsilon_A),$$

то вероятности событий A и B можно считать статистически неразличимыми, т. е. равными. Принимаем гипотезу H_0 . В противном случае вероятности событий A и B равными считать нельзя. Принимаем альтернативную гипотезу H_1 .

Таким образом, согласно правилу проверки гипотезы о равенстве вероятностей событий имеем:

$$P_B = 0,237 \in (0,16; 0,26),$$

т. е. оценка вероятности события B входит в пределы доверительного интервала оценки вероятности события A . Это значит, что вероятности событий A и B считаются статистически неразличимыми, то есть говорить о том, кто из преподавателей строже не представляется возможным. Различия в оценках долей двоек можно объяснить исключительно случайными (стохастическими) причинами.

Дополнительно можно завершить анализ проверкой двухстороннего критерия. А именно, проверить, входит ли вероятность события A в доверительный интервал события B . Построим доверительный интервал для вероятности события B , т. е. доли двоек, поставленных преподавателем B :

$$(P_B - \varepsilon_B; P_B + \varepsilon_B) = (0,237 - 0,05; 0,237 + 0,05) = (0,187; 0,287).$$

Согласно правилу проверки гипотезы о равенстве или неравенстве долей имеем:

$$P_A = 0,21 \in (0,187; 0,287),$$

т. е. оценка вероятности события A входит в пределы доверительного интервала оценки вероятности события B .

Вывод по результатам проверки гипотезы:

Вероятности событий A и B считаются статистически неразличимыми, что подтверждает выводы одностороннего критерия.

Теперь рассмотрим способ решения задачи, рекомендованный для **уровня подготовки пред-постдипломников**. Текст задания остался неизменным (см. выше). Определим значения параметров задачи.

Дано:

Преподаватель A :

Двоек: 42

Общее число студентов: 200

Доля двоек: $\hat{P}A = \frac{42}{200} = 0,21 (21\%)$

Точность: $\varepsilon_{PA} = \pm 5\%$

Преподаватель B :

Двоек: 64

Общее число студентов: 270

Доля двоек: $\hat{P}B = \frac{64}{270} = 0,237 (23,7\%)$

Точность: $\varepsilon_{PB} = \pm 5\%$

Доля двоек – это оценка вероятности того, что преподаватель поставит двойку на экзамене. Поэтому выдвигаем две гипотезы:

Основная гипотеза H_0 – преподаватели одинаково строгие, т. е. доли поставленных двоек неразличимы;

Альтернативная гипотеза H_1 – преподаватели отличаются по измерению строгости, т. е. доли поставленных ими двоек статистически различимы.

Проверку гипотезы о равенстве долей можно проводить методом доверительных интервалов.

Прежде всего, для доли двоек, поставленных каждым преподавателем, построим доверительный интервал с учетом точности $\pm 5\%$.

Доверительный интервал для доли двоек преподавателя A :

$$21\% \pm 5\% \Rightarrow [16\% \div 26\%].$$

Доверительный интервал для доли двоек преподавателя B :

$$23,7\% \pm 5\% \Rightarrow [18,7\% \div 28,7\%].$$

Сравним доверительные интервалы. Найдем пересечение интервалов:

Нижняя граница интервала для преподавателя A (16%) $<$ верхней границы интервала для преподавателя B ($28,7\%$).

Верхняя граница интервала для преподавателя A (26%) $>$ нижней границы интервала для преподавателя B ($18,7\%$).

Интервалы пересекаются в диапазоне $[18,7\%; 26\%]$. Ширина пересечения интервалов:

$$\Delta = 26\% - 18,7\% = 7,3\%.$$

Вывод по проверке гипотезы методом доверительных интервалов:

С учетом погрешности $\pm 5\%$ нельзя однозначно утверждать, что один преподаватель строже другого. Таким образом, нельзя отвергнуть основную о равенстве долей поставленных преподавателями A и B двоек.

Можно предложить другой вариант проверки гипотезы о равенстве долей. **Проверим, является ли разница в долях статистически значимой при заданной точности.**

Разница долей:

$$\delta = \hat{P}B - \hat{P}A = 23,7\% - 21\% = 2,7\%.$$

Суммарная погрешность:

$$\varepsilon = \varepsilon_{PA} - \varepsilon_{PB} = 5\% + 5\% = 10\%.$$

Сравнение:

$$\delta = 2,7\% < 10\% = \varepsilon.$$

Вывод по оценке статистической значимости равенства долей:

Разница долей δ меньше суммарной погрешности ε , поэтому различия в долях двоек, поставленных преподавателями A и B статистически не значимы.

Наконец, можно проверить гипотезу о статистической неразличимости долей поставленных двоек с помощью Z-теста.

Для этого необходимо по заданной точности данных ($\varepsilon_{PA} = \varepsilon_{PB} = \pm 5\%$) найти соответствующие доверительные вероятности, т.е. надёжности.

Расчет доверительной вероятности γ проводится по формуле Муавра-Лапласа. Формула для точности ε :

$$\varepsilon = Z_\gamma \sqrt{\frac{p \cdot q}{n}},$$

где $\varepsilon = 0,05$ (5%); p – доля двоек; $q = 1 - p$, n – объем выборки; Z_γ – квантиль стандартного нормального распределения, соответствующий доверительной вероятности γ .

Для каждого преподавателя найдем соответствующие доверительные вероятности γ :

Для преподавателя A :

$$0,05 = Z_{\gamma A} \sqrt{\frac{0,21 \cdot 0,79}{200}}$$

$$Z_{\gamma A} = \frac{0,05}{\sqrt{\frac{0,21 \cdot 0,79}{200}}} = \frac{0,05}{0,0288} = 1,736.$$

По таблице стандартного нормального распределения:

$$\gamma_A = \Phi(1,736) = 0,9588 = 95,88\%.$$

Уровень значимости:

$$\alpha_A = 1 - \gamma_A = 0,0412 = 4,12\%.$$

Соответствующее критическое значение $Z_{\text{крит}}$:

$$Z_{\text{крит}} = 1,736 \text{ (для } \gamma = 95,88\%)$$

Для преподавателя B :

$$0,05 = Z_{\gamma B} \sqrt{\frac{0,237 \cdot 0,763}{270}}$$
$$Z_{\gamma B} = \frac{0,05}{\sqrt{\frac{0,237 \cdot 0,763}{270}}} = \frac{0,05}{0,0258} = 1,938.$$

По таблице стандартного нормального распределения:

$$\gamma_B = \Phi(1,938) = 0,9736 = 97,36\%.$$

Уровень значимости:

$$\alpha_B = 1 - \gamma_B = 0,0264 = 2,64\%.$$

Выбираем наименьшую надёжность 95,88%, т. е. наибольший уровень значимости:

$$\alpha = \max(\alpha_A, \alpha_B) = 0,0412 \approx 0,04.$$

Соответствующее критическое значение $Z_{\text{крит}}$:

$$Z_{\text{крит}} = 1,938 \text{ (для } \gamma = 95,88\%)$$

Проведем Z -тест для двух пропорций. Найдем объединённую долю двоек, поставленных обоими преподавателями:

$$\hat{P} = \frac{42 + 64}{200 + 270} = \frac{106}{470} = 0,2255 = 22,55\%.$$

Найдем стандартную ошибку разницы:

$$SE = \sqrt{\hat{P}(1 - \hat{P}) \left(\frac{1}{200} + \frac{1}{270} \right)} = \sqrt{0,2255 \cdot 0,7745 \cdot 0,0087} = 0,039.$$

Вычислим значение Z -статистики:

$$Z = \frac{\hat{P}B - \hat{P}A}{SE} = \frac{0,237 - 0,21}{0,39} = 0,692.$$

Так как

$$|Z| = 0,692 < 1,938,$$

то разница статистически не значима.

Вывод по проверке гипотезы о равенстве долей:

На уровне значимости 0,04 доли двоек, поставленных преподавателями A и B не являются статистически различимыми, значит нельзя дать заключение о том, какой из преподавателей строже.

Задание № 17. Респондент отвечает на вопросы анкеты X и Y «да» (1 балл) или «нет» (0 баллов). Процентные доли различных вариантов ответов сведены в матрицу корреляции. Установить степень тесноты связи вопросов X и Y .

X		Y	0	1
X	0	P_{11}	P_{12}	
	1	P_{21}	P_{22}	

X		Y	0	1
X	0	0,1	0,1	
	1	0,75	0,05	

Рассмотрим способ решения задачи, рекомендованный для **уровня подготовки бакалавров**.

Построим ряды распределения случайных величин X и Y по отдельности.

$$P_0^X = P_{11} + P_{12} = 0,1 + 0,1 = 0,2;$$

$$P_1^X = P_{21} + P_{22} = 0,75 + 0,05 = 0,8;$$

$$P_0^Y = P_{11} + P_{21} = 0,1 + 0,75 = 0,85;$$

$$P_1^Y = P_{12} + P_{22} = 0,1 + 0,05 = 0,15.$$

Тогда отдельные ряды распределения имеют вид:

X	0	1
P	0,2	0,8

Y	0	1
P	0,85	0,15

Заметим, что условие нормировки в каждом ряде соблюдается – сумма вероятностей равна 1.

Найдем математические ожидания:

$$M_X = P_1^X = 0,8; M_Y = P_1^Y = 0,15.$$

Найдем дисперсии:

$$D_X = P_1^X - \left(P_1^X \right)^2 = 0,8 - 0,8^2 = 0,8 - 0,64 = 0,16;$$

$$D_Y = P_1^Y - \left(P_1^Y \right)^2 = 0,15 - 0,15^2 = 0,15 - 0,0225 = 0,1275.$$

Найдем среднеквадратические отклонения:

$$\sigma_X = \sqrt{D_X} = \sqrt{0,16} \approx 0,4; \sigma_Y = \sqrt{D_Y} = \sqrt{0,1275} \approx 0,3571.$$

Найдем ковариацию:

$$K_{XY} = P_{22} - P_1^X \cdot P_1^Y = 0,05 - 0,8 \cdot 0,15 = 0,05 - 0,12 = -0,07.$$

Найдем коэффициент корреляции:

$$\rho_{XY} = \frac{K_{XY}}{\sigma_X \cdot \sigma_Y} \approx \frac{-0,07}{0,4 \cdot 0,3571} \approx -0,49.$$

Вывод по оценке тесноты связи между вопросами анкеты:

Теснота связи между вопросами анкеты X и Y – отрицательная умеренная, что свидетельствует, скорее всего, о приемлемом качестве ее составления.

На уровне подготовки предмагистрантов текст задачи можно усложнить.

Задание № 17. Респондент отвечает на вопросы анкеты X и Y «да» (1 балл) или «нет» (0 баллов). Всего в опросе участвовали 200 человек. Число ответивших по-разному сведено в таблицу вариантов ответов сведены в матрицу корреляции. Установить степень тесноты связи вопросов X и Y .

X	Y	0	1
0	a	b	
1	c	d	

X	Y	0	1
0		20	20
1		150	10

Коэффициент корреляции вычисляется по формуле:

$$\begin{aligned}\rho_{XY} &= \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \\ &= \frac{20 \cdot 10 - 20 \cdot 150}{\sqrt{(20+20)(150+10)(20+150)(20+10)}} = \\ &= \frac{200 - 300}{\sqrt{40 \cdot 160 \cdot 170 \cdot 30}} = \frac{-2800}{\sqrt{32640000}} \approx \frac{-2800}{5713} \approx -0,49.\end{aligned}$$

Теперь составляем корреляционную матрицу, вычисляя процентные доли ответивших «да» и «нет» на вопросы анкеты:

$$P_{11} = \frac{a}{n}; P_{12} = \frac{b}{n}; P_{21} = \frac{c}{n}; P_{22} = \frac{d}{n}; n = 200.$$

Сразу расширяем корреляционную матрицу, складывая процентные доли по срокам и по столбцам:

$X \backslash Y$	0	1	$P(X)$
0	0,1	0,1	0,2
1	0,75	0,05	0,8
$P(Y)$	0,85	0,15	1

Значение «1» в правой нижней графе означает выполнение условия нормировки для корреляционной матрицы. Для бинарных переменных коэффициент корреляции Пирсона ρ_{XY} вычисляется по формуле:

$$\rho_{XY} = \frac{P(X=1, Y=1) \cdot P(X=0, Y=0) - P(X=1, Y=0) \cdot P(X=0, Y=1)}{\sqrt{P(X=1)P(X=0)P(Y=1)P(Y=0)}}$$

$$\rho_{XY} = \frac{(0,05 \cdot 0,01) - (0,75 \cdot 0,1)}{\sqrt{0,8 \cdot 0,2 \cdot 0,15 \cdot 0,85}} = \frac{0,005 - 0,075}{\sqrt{0,0204}} \approx \frac{-0,07}{0,1428} \approx -0,49.$$

Очевидно, что результат совпадает с результатом расчёта коэффициента корреляции для абсолютного числа ответивших на вопросы анкеты. Следует отметить, что в ряде источников формула для дихотомического коэффициента корреляции ρ_{XY} называется формулой коэффициента четырехклеточной сопряженности Пирсона (<https://studfile.net/preview/9783499/page:135/>).

Вывод о степени тесноты связи вопросов анкеты:

Значение $\rho_{XY} \approx -0,49$ указывает на умеренную отрицательную связь между вопросами. Это означает, что ответ «да» на один вопрос чаще сочетается с ответом «нет» на другой.

Предмагистрантам, разумеется, необходимо **проверить гипотезы о значимости корреляционной связи**. Для проверки гипотезы значимости корреляционной связи можно использовать критерий Стьюдента (t -тест):

$$t = \frac{\rho_{XY} \sqrt{n-2}}{\sqrt{1-\rho_{XY}^2}},$$

где $n = 200$ – общее количество ответов на два вопроса анкеты.

$$t = \frac{0,49 \sqrt{200-2}}{\sqrt{1-0,49^2}} \approx \frac{0,49 \cdot 14,071}{0,872} \approx 7,907.$$

Вычислим число степеней свободы:

$$k = n - 1 - 1 = 200 - 2 = 198,$$

поскольку оценивается только один параметр – коэффициент корреляции.

В таблице критических значений распределения Стьюдента приходится брать $k = \infty$. На уровне значимости $\alpha = 0,05$:

$$t_{\text{крит}}^{\infty, 0,05} = 1,96.$$

Так как

$$t = 7,907 > 1,96 = t_{\text{крит}}^{\infty, 0,05},$$

то корреляционная связь статистически значима с надёжностью 95%, т. е. на уровне значимости 0,05.

Поскольку число анкетируемых было велико, то можно проверить значимость корреляции с большей надёжностью, т.е. при меньшем значении уровня значимости. Возьмём $\alpha = 0,001$ (надёжность 99,9%):

$$t_{\text{крит}}^{\infty, 0,001} = 3,29.$$

Так как

$$t = 7,907 > 3,29 = t_{\text{крит}}^{\infty, 0,001},$$

то корреляционная связь статистически значима с надёжностью 99,9%, т. е. на уровне значимости 0,001.

В качестве альтернативных способов оценки знака связи для бинарных данных также можно использовать отношение шансов *OR*:

$$OR = \frac{P(X=1, Y=1) \cdot P(X=0, Y=0)}{P(X=1, Y=0) \cdot P(X=0, Y=1)} = \frac{0,05 \cdot 0,1}{0,75 \cdot 0,1} \approx 0,0667.$$

Значение *OR* < 1 подтверждает отрицательную корреляционную связь.

Вывод по проверке гипотезы о связи между вопросами анкеты:

Предмагистранты оценили тесноту связи между вопросами анкеты как для абсолютных значений числа ответивших, так и для процентных долей, вычислив коэффициент корреляции Пирсона, а также проверили гипотезу о значимости корреляционной связи по критерию Стьюдента, варьируя уровни значимости. Полученное значение тесноты связи между вопросами анкеты свидетельствует о приемлемом качестве её составления в отношении вопросов *X* и *Y*. Дополнительное подтверждение знака связи получили вычислением отношения шансов.

Выводы

На примерах решения двух задач математической статистики авторы продемонстрировали реализацию принципов построения курса «Теория вероятностей и математическая статистика» программы пред-постдипломной подготовки иностранных граждан в СПбПУ. Авторы провели сравнительный анализ решения задач на уровне предъявления материала в бакалавриате и в предмагистратуре/предасpirантуре. Основное внимание авторы уделили вариативности подходов к решению и всестороннему обсуждению результатов решения задач.

ЛИТЕРАТУРА

1. Гузарова Н.И., Кашкан Г.В., Шахова Н.Б. Предмагистерская подготовка иностранных граждан // Высшее образование в России, 2013. 2. 84 – 89. URL: <https://vovr.elpub.ru/jour/article/view/3475>.

2. Горбенко В.Д., Краснощеков В.В., Новикова О.А. Проблемы проектирования и реализации программ подготовки иностранных граждан в магистратуру // III Международный конгресс преподавателей и руководителей подготовительных факультетов (отделений) вузов РФ «Довузовский этап обучения в России и мире: язык, адаптация, социум, специальность. IV Всероссийская научно-практическая конференция «Актуальные вопросы реализации образовательных программ на подготовительных факультетах для иностранных граждан». Москва, 16-18 октября 2019 г. / Отв. ред. М.Н. Русецкая, А.В. Колтакова. М., ГИРЯ им. А.С. Пушкина, 2019. 663 – 667. EDN: YLGFAM. URL: <https://www.elibrary.ru/item.asp?id=42410449>

3. Поморцева Н.В., Куновски М.Н. Научно-практические исследования по предмагистерской, предаспирантской и предординаторской подготовке иностранных граждан в российских вузах // Вопросы методики преподавания в вузе, 2021. 10 (37). 30 – 56. DOI: 10.18720/HUM/ISSN 2227-8591.37.03.

4. Краснощеков В.В. Развитие программы предмагистерской подготовки иностранных граждан в 2021 и 2022 годах // Предмагистерская подготовка иностранных граждан в вузах России. Сборник статей IV международной научно-практической онлайн-конференции. Санкт-Петербург, 08 июля 2022 г. / Отв. ред. Игнатьева Д.А., ред. колл. Краснощеков В.В., Рудь В.Ю. СПб., ПОЛИТЕХ-ПРЕСС, 2023. 17 – 25. DOI: 10.18720/SPBPU/2/id23-434.

5. Краснощеков В.В. Развитие программы предмагистерской подготовки иностранных граждан в 2023 году // Неделя науки высшей школы международных образовательных программ. Материалы научной конференции. С международным участием. СПб., 11 – 15 декабря 2023 года / Ред. Д.А. Игнатьева, В.В. Краснощеков, В.Д. Горбенко. СПб., ПОЛИТЕХ-ПРЕСС, 2023. 78 – 84. DOI: 10.18720/SPBPU/2/id23-728.

6. Горбенко В.Д., Бусурина Е.В., Куралева И.Р. Формирование soft skills у иностранных слушателей программы предмагистерской подготовки // V Международный конгресс преподавателей и руководителей подготовительных факультетов (отделений) вузов РФ «Довузовский этап обучения в России и мире: язык, адаптация, социум, специальность». VI Всероссийской научно-практической конференции «Актуальные вопросы реализации образовательных программ на подготовительных факультетах для иностранных граждан». Москва, 17 – 20 ноября 2021 г. Сборник материалов. Отв. ред. Н.В. Татаринова. М.: Государственный институт русского языка им. А.С. Пушкина, 2022. 138 – 142. URL: https://elibrary.ru/download/elibrary_49407420_36046883.pdf

7. Krasnoshchekov V.V., Semenova N.V. Forming Of Probabilistic Approach To Cognition As Component Of Students Professional Culture. In: O.D. Shipunova & D.S. Bylieva (Eds) European Proceedings of Social and Behavioural Sciences EpSBS. 20th Professional Culture of the Specialist of the Future (PCSF 2020) & 12th Communicative Strategies of Information Society (CSIS 2020). St. Petersburg, 26 – 27 November, 2020 & 23 - 24 October, 2020. European Publisher, 2020. Vol. 98. 14. 139 – 149. DOI: 10.15405/epsbs.2020.12.03.14.
8. Краснощеков В.В., Семенова Н.В. Математика. Тестовые задания по теории вероятностей и математической статистике. Учебное пособие для студентов высших учебных заведений, обучающихся по направлению подготовки бакалавров «Технологические машины и оборудование». СПб., Изд-во Политехн. ун-та, 2013. 222 с. DOI: 10.18720/SPBPU/2/si20-1051.
9. Краснощеков В.В., Семенова Н.В. Теория вероятностей и математическая статистика. Практикум. СПб., Изд-во Политехн. ун-та, 2017. 176 с. [Текст. парал. на рус. и англ. яз.]. DOI: 10.18720/SPBPU/2/i17-419.
10. Краснощеков В.В., Семенова Н.В., Цуй Лун. Теория вероятностей и математическая статистика. СПб., ПОЛИТЕХ-ПРЕСС, 2023. 161 с. [Текст на рус. и кит. яз.]. DOI: 10.18720/SPBPU/2/i23-318.
11. Krasnoshchekov V., Rud' V., Davydov R., Semenova N., Diuldin M., Kharlamova N., Ionkina E., Shimkovich E. Formation of environmental research competencies of foreign pre-masters' students for sustainable region development. In Pukhkhal V., Uvarova S. (eds) International Scientific Conference Energy Management of Municipal Facilities and Environmental Technologies (EMMFT-2023). Voronezh, Russia, October 9 – 13, 2023. E3S Web of Conferences, Vol. 458. 06018. DOI: 10.1051/e3sconf/202345806018.
12. Краснощеков В.В., Рудь В.Ю., Давыдов В.В. Модели обучения иностранных предмагистрантов инженерных профилей подготовки // Современные научно-исследовательские технологии. 2018. 12 (1). 214 – 219. DOI: 10.17513/snt.37290.
13. Shakhova N.B., Sherina E.A., Kashkan G.V. Formation of new generation of young scientists by supporting the scientific initiative of foreign students // Предмагистерская подготовка иностранных граждан в вузах России. Сборник статей IV межвузовской научно-практической онлайн-конференции. Санкт-Петербург, 08 июля 2022 г. / Отв. ред. Д.А. Игнатьева, ред. колл. В.В. Краснощеков, В.Ю. Рудь. СПб., ПОЛИТЕХ-ПРЕСС, 2023. 46 – 55. DOI: 10.18720/SPBPU/2/id23-437.
14. Краснощеков В.В., Семенова Н.В., Алдармини С.С. Методы формирования компетенций студентов в области точности вероятностных

моделей // Современные проблемы науки и образования, 2020. 5. 34.
DOI: 10.17513/spno.30142.

15. Краснощеков В.В., Семенова Н.В., Алсалама А.М., Михолитис А.Г. О точных и приближенных моделях в вузовском курсе теории вероятности // Современные научноемкие технологии, 2021. 10. 149 – 154.
DOI: 10.17513/snt.38869.

16. Краснощеков В.В., Семенова Н.В., Мухамед Б.М.М., Баккар М.М.А. О выборе из конечного и бесконечного поля в курсе теории вероятности // Современные научноемкие технологии, 2022. 9. 138 – 143.
DOI: 10.17513/snt.39322.

17. Краснощеков В.В., Семенова Н.В., Аббас А., Шбиб Х. Формирование у студентов вуза представлений о точности и надежности оценки вероятности // Современные научноемкие технологии, 2024. 7. 163 – 170.
DOI: 10.17513/snt.40102.

УДК 004.056.5

doi:10.18720/SPBPU/2/id25-411

Пап Адам¹

бакалавр, студент предмагистратуры;

Рышавы Ондржей²

доцент, кандидат технических наук

АВТОМАТИЧЕСКОЕ ОБНАРУЖЕНИЕ АКТИВНОСТИ ВРЕДОНОСНЫХ ПРОГРАММ В ЛОКАЛЬНЫХ СЕТЯХ

¹ Россия, Санкт-Петербург,

Санкт-Петербургский политехнический университет

Петра Великого,

^{1,2} Чешская Республика, Брно, Технический университет в Брно,

¹adam.pap146@gmail.com, ²rysavy@fit.vut.cz

Аннотация. В рамках работы был создан набор данных, из которого были извлечены IoC для каждого семейства вредоносных программ. Полученные IoC были проверены через платформу AlienVault OTX для подтверждения их релевантности. На тестовых данных обе IoC-модели, созданные из наборов данных, достигли точности 99,337% и 94,732% для набора данных №1 и №2 соответственно. IoC-модели набора №1 в реальной эксплуатации ложно классифицировали 3,03% коммуникационных окон как вредоносные. IoC-модели набора № 2 классифицировали