УДК 378.14:371.32

МЕТОДЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ СТУДЕНТОВ В ОБЛАСТИ ТОЧНОСТИ ВЕРОЯТНОСТНЫХ МОДЕЛЕЙ

Краснощеков В.В.[1], Семенова Н.В.[1], Алдармини С. [1,2]

[1]ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого», Санкт-Петербург, e-mail: krasno_vv@spbstu.ru;
[2]Высший институт прикладных наук и технологий, Дамаск, Сирия, e-mail: sokratdarmini@gmail.com

Авторы рассматривают проблемы преподавания вероятностных и статистических дисциплин вузовского курса математики. Несмотря на востребованность компетенций выпускников в области теории вероятностей и математической статистики, эти разделы часто оказываются в роли «падчерицы» для вузовских преподавателей математики. При традиционных способах подачи материала не только вопросы точности вероятностных моделей, но и сама тематика вероятностного моделирования уходит на второй план, уступая место логическим или вычислительным задачам. В то же время именно проблемы оценки точности моделирования наиболее востребованы практикой. Более того, они являются одной из главных составляющих вероятностного подхода к познанию, которому авторы уделили немало внимания в своих предыдущих работах. Для демонстрации возможностей наглядного представления проблематики точности авторы используют решения достаточно простых задач. В статье авторы ограничиваются примерами на применение схемы Бернулли, формулы Пуассона и интегральной теоремы Лапласа. На этих примерах авторы демонстрируют градации точности вероятностных моделей. Все расчеты снабжены графическими иллюстрациями, которые играют большую роль в методах восприятия информации современными студентами – представителями цифровых поколений. Предложенная методика, несомненно, способствует формированию компетенций студентов в области точности вероятностных и статистических моделей.

Ключевые слова: преподавание математики, вероятностные модели, точность моделирования, вероятностной подход к научному познанию, графическая интерпретация.

METHODS FOR FORMING STUDENTS' COMPETENCES IN THE FIELD OF PROBABILITY MODELS ACCURACY

Krasnoshchekov V.V.[1], Semenova N.V.[1], Aldarmini S.[1,2]

[1]Peter the Great St. Petersburg Polytechnic University, St. Petersburg, e-mail: krasno_vv@spbstu.ru
[2]Higher Institute for Applied Science and Technology, Damascus, Syria, e-mail: sokratdarmini@gmail.com

The authors consider the problems of teaching probabilistic and statistical disciplines in university mathematical courses. Despite the demand for the competencies of graduates in the field of probability theory and mathematical statistics, these sections often find themselves in the role of "stepdaughter" for university mathematics teachers. With traditional methods of presenting material, not only the issues of the accuracy of probabilistic models, but the very topic of probabilistic modeling goes into the background, giving way to logical or computational problems. At the same time, the problems of assessing the accuracy of modeling are demanding in practice. Moreover, they are one of the main components of the probabilistic approach to cognition, to which the authors have paid a lot of attention in their previous works. To demonstrate the possibilities of visual representation of accuracy problems, the authors use solutions to simple problems. In the article, the authors limit themselves to examples of the application of the Bernoulli scheme, Poisson's formula and Laplace's integral theorem. Using these examples, the authors demonstrate the gradations of the accuracy of probabilistic models. Authors provide all calculations with graphic illustrations, which play an important role in the methods of information perception by modern students as representatives of digital generations. The proposed methodology undoubtedly contributes to the formation of students' competencies in the field of accuracy of probabilistic and statistical models.

Keywords: teaching mathematics, probabilistic models, modeling accuracy, probabilistic approach to scientific cognition, graphic interpretation.

Компетенции, получаемые студентами вузов при изучении теории вероятностей и математической статистики, находят широкое применение в научных теоретических и, преимущественно, практико-ориентированных исследованиях, относящихся к различным отраслям человеческой деятельности [1, 2, 3]. В то же время преподавание теории вероятностей на 1–2-м курсах вуза не позволяет обучающимся в полной мере оценить возможности и ограничения вероятностных моделей процессов и явлений. Во многом это объясняется ориентацией значительной части преподавателей на развитие комбинаторных умений студентов и концентрацией их внимания на вычислительных процедурах [4]. Тем не менее авторы выявили несколько вузовских научно-методических групп, которые разрабатывают активные методы преподавания вероятностных дисциплин, такие как решение ситуационных задач [5], метод проектов, математические бои, деловые игры [6], а также проектируют программы дисциплины, ориентируясь на направление подготовки [7].

Именно моделирование является одним из основных компонентов научного и инженерного творчества. Вопросы точности стохастических моделей обычно всегда остаются вне поля зрения педагогов-математиков. Более того, эти вопросы практически не затрагиваются в популярных учебниках и руководствах по теории вероятностей. В настоящем исследовании авторы делают попытку найти достаточно простые методические приемы для ликвидации этого пробела.

Целью настоящего исследования является разработка таких методик преподавания некоторых разделов теории вероятности, которые позволяют наглядно продемонстрировать студентам подходы к оценке точности вероятностных моделей. Авторы предлагают для внедрения в практику преподавания несколько конкретных примеров, снабженных визуализацией, с целью облегчения восприятия молодыми людьми новой для них информации. Приведенные примеры могут быть легко трансформированы с целью соответствия предметной области обучающихся. Также авторы дают рекомендации по локализации разработанных методик для различных образовательных программ.

Материал и методы исследования

Во-первых, авторы изучили доступные источники по современным тенденциям в преподавании теории вероятностей и выбрали те из них, которые близки авторской концепции формирования у студентов вероятностного подхода к познанию [8].

Таким образом, авторы дают дополнительную аргументацию по необходимости переакцентировки преподавания курса теории вероятностей и математической статистики, который должен соответствовать не только историческому пути развития этой науки, но и современным задачам практической деятельности.

Во-вторых, авторы проводят численное исследование решений некоторых задач теории вероятностей, сравнивая точность моделей, основанных на схеме Бернулли, формуле Пуассона и интегральной теореме Лапласа. При этом на основании визуального графического анализа авторам удалось сформировать некоторые рекомендации по наглядному представлению студентам положений о точности вероятностных моделей. Кроме того, было проведено статистическое исследование погрешностей анализируемых моделей.

Результаты исследования и их обсуждение

Вопросы развития у студентов представлений о точности вероятностных моделей заинтересовали авторов в связи с развиваемой ими концепцией формирования вероятностного подхода к познанию [8]. В этой области существуют, видимо, независимые исследования ряда вузовских ученых. Прежде всего это публикации, группирующиеся вокруг Г.Д. Гефана [9], в том числе развивающие понятие «вероятностный стиль мышления» [10], близкое по смыслу к понятию «вероятностный подход», используемому авторами. Далее следует отметить работы М.Э. Григорян и ее соавторов [11], также связанные с формированием научного мировоззрения студентов через вероятностную картину мира [12]. Наконец, в вышедшей недавно работе Т.В. Васильева приводит примеры практических заданий, способствующих повышению стохастической культуры обучающихся [13].

Вероятностный подход к познанию, вероятностный стиль мышления, вероятностная картина мира и стохастическая культура с разных сторон и в разных измерениях подтверждают необходимость формирования у студентов компетенций в области теории вероятности, опираясь на которые, они смогут в профессиональной деятельности строить адекватные модели анализируемых процессов и явлений, принимая во внимание не только детерминистские, но и стохастические факторы.

Проблема точности вероятностных моделей является одним из компонентов вероятностного подхода, которому в преподавании, в том числе и авторами, уделяется недостаточное внимание. Разумеется, инженеры и экономисты забудут о методах оценки точности, даже если бы они изучали их на 2-м курсе, но само представление о необходимости такой оценки должно входить в компетентностную модель выпускника.

Авторы предлагают ряд примеров, которые можно демонстрировать студентам для иллюстрации положений о точности вероятностных моделей.

Задание 1. 25% проектов фирмы – крупные. Случайная величина (СВ) X – число крупных проектов. Построить полигон – график ряда распределения СВ X, если фирма выполняет n проектов для следующих случаев: 1) n = 5: 2) n = 10; 3) n = 15; 4) n = 20.

Решение: значения параметров:

$$p = 0.25, q = 1 - p = 0.75\ .$$

Используя формулу Бернулли $P(X = i) = C_n^i p^i q^{n-i}$, получаем графики (рис. 1).

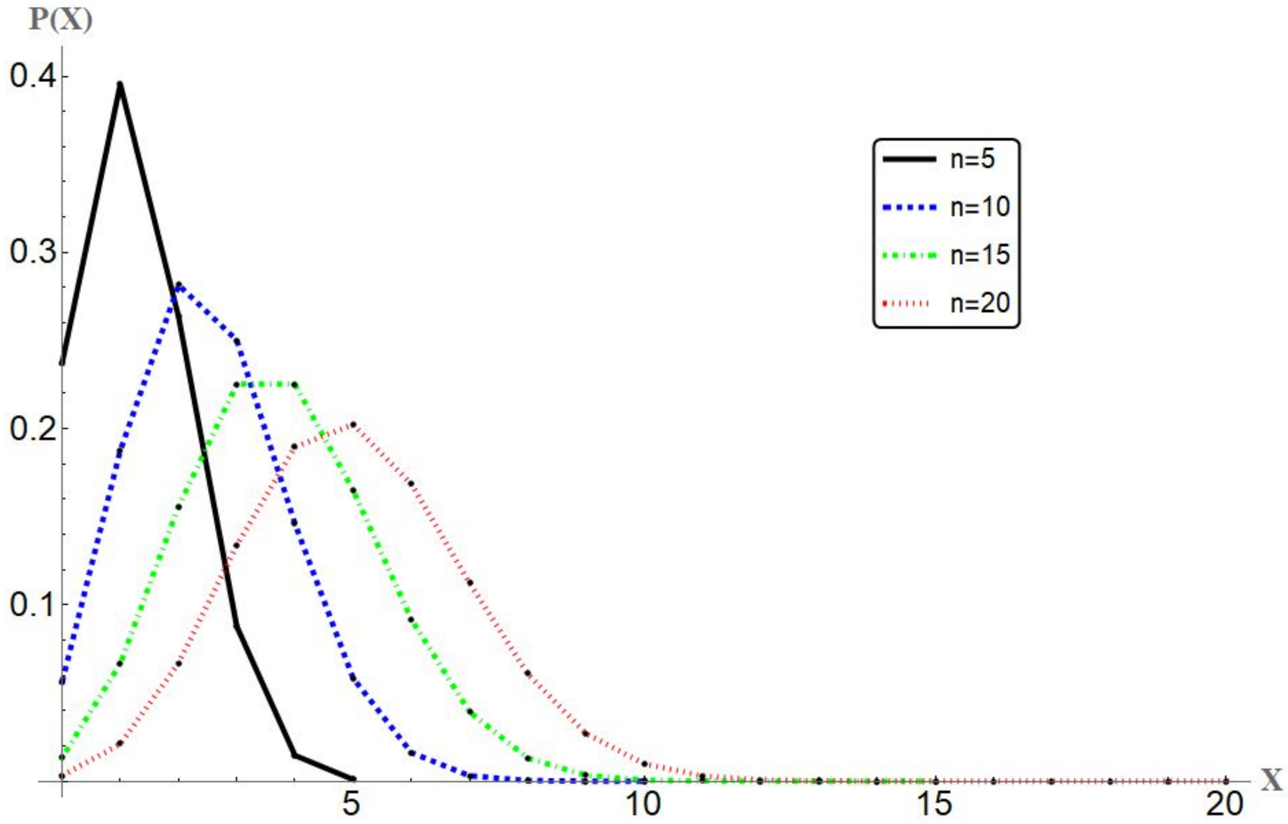

Рис. 1. Полигоны распределения по схеме Бернулли при p = 0,25

Величина p = 0,25 была выбрана для демонстрации несимметричности распределения Бернулли. Очень часто для иллюстрации схемы Бернулли выбирают p = 0,5, и у студентов может сложиться ошибочное мнение о симметричном характере распределения.

Дидактический вывод по Заданию 1: с ростом числа испытаний полигон распределения Бернулли приближается к дискретному аналогу нормального распределения. Это свидетельствует в пользу замены расчета по точной формуле Бернулли приближенным расчетом по теореме Лапласа.

Задание 2. 25% проектов фирмы – крупные. Случайная величина (СВ) X – число крупных проектов. Найти по формуле Бернулли и теореме Лапласа 1) P (4 ≤ X ≤ 6 при n = 20; 2) P (46 ≤ X ≤ 54) при n = 200; 3) P (486 ≤ X ≤ 514) при n = 2000.

Решение:

1) n = 20; P (4 ≤ X ≤ 6)

- по формуле Бернулли:

$$P(4 \le X \le 6) = \sum_{i=4}^{6} C_{20}^{i} p^{i} q^{20-i} = 0.560626$$

- по теореме Лапласа:

$$x_4 = \frac{m - n.p}{\sqrt{npq}} = \frac{4 - 20 \times 0.25}{\sqrt{20 \times 0.25 \times 0.75}} = -0.516398$$

$$x_6 = \frac{m - n.p}{\sqrt{npq}} = \frac{6 - 20 \times 0.25}{\sqrt{20 \times 0.25 \times 0.75}} = 0.516398$$

$$P(4 \le X \le 6) \approx \phi(x_6) - \phi(x_4) \approx 0.197212 - (-0.197212)$$
$$\approx 0.394423$$

Здесь далее *Ф* – значение нормированной функции Лапласа.

2) n = 200; P (46 ≤ X ≤ 54)

- по формуле Бернулли:

$$P(46 \le X \le 54) = \sum_{i=46}^{54} C_{200}^{i} p^{i} q^{200-i} = 0.5374979788968025$$

- по теореме Лапласа:

$$x_{46} = \frac{m - n.p}{\sqrt{npq}} = \frac{46 - 200 \times 0.25}{\sqrt{200 \times 0.25 \times 0.75}} = -0.6531972647421809$$

$$x_{54} = \frac{m - n.p}{\sqrt{npq}} = \frac{54 - 200 \times 0.25}{\sqrt{200 \times 0.25 \times 0.75}} = 0.6531972647421809$$

$$P(46 \le X \le 54) \approx \phi(x_{54}) - \phi(x_{46}) \approx 0.24318544330343794 - (-0.24318544330343794)$$
$$\approx 0.4863708866068759$$

3) n = 2000; P (486 ≤ X ≤ 514)

- по формуле Бернулли:

$$P(486 \le X \le 514) = \sum_{i=486}^{514} C_{2000}^{i} p^{i} q^{2000-i} \approx 0.5460024936070365$$

- по теореме Лапласа:

$$x_{486} = \frac{m - n.p}{\sqrt{npq}} = \frac{486 - 2000 \times 0.25}{\sqrt{2000 \times 0.25 \times 0.75}} = -0.7229568912920512$$

$$x_{514} = \frac{m - n.p}{\sqrt{npq}} = \frac{514 - 2000 \times 0.25}{\sqrt{2000 \times 0.25 \times 0.75}} = 0.7229568912920512$$

$$P(486 \le X \le 514) \approx \phi(x_{514}) - \phi(x_{486}) \approx 0.26514681531647766 - (-0.26514681531647766)$$
$$\approx 0.5302936306329553$$

Границы интервалов были выбраны на основании данных предварительного анализа таким образом, чтобы величины интервальной вероятности, вычисленной по формуле Бернулли P_B, были бы максимально близкими. Для того чтобы оценить погрешность результатов, полученных по приближенной формуле Лапласа P_L, по сравнению с точной формулой Бернулли P_B, были вычислены абсолютная ошибка $\Delta = P_B - P_L$ и относительная ошибка $\varepsilon = \Delta / P_B$ (табл.).

Абсолютная и относительная ошибки значений вероятности

по формуле Бернулли и по теореме Лапласа

n	Δ	ε	ε (%)
20	0.1662025162722412	0.29645886194510335	29,6%
200	0.05112709228992662	0.09512052937364221	9,5%
2000	0.01570886297408114	0.0287706799108265	2,8%

Графики относительной ошибки и ее возможных аппроксимаций приведены на рисунке 2, где по горизонтальной оси введена логарифмическая шкала.

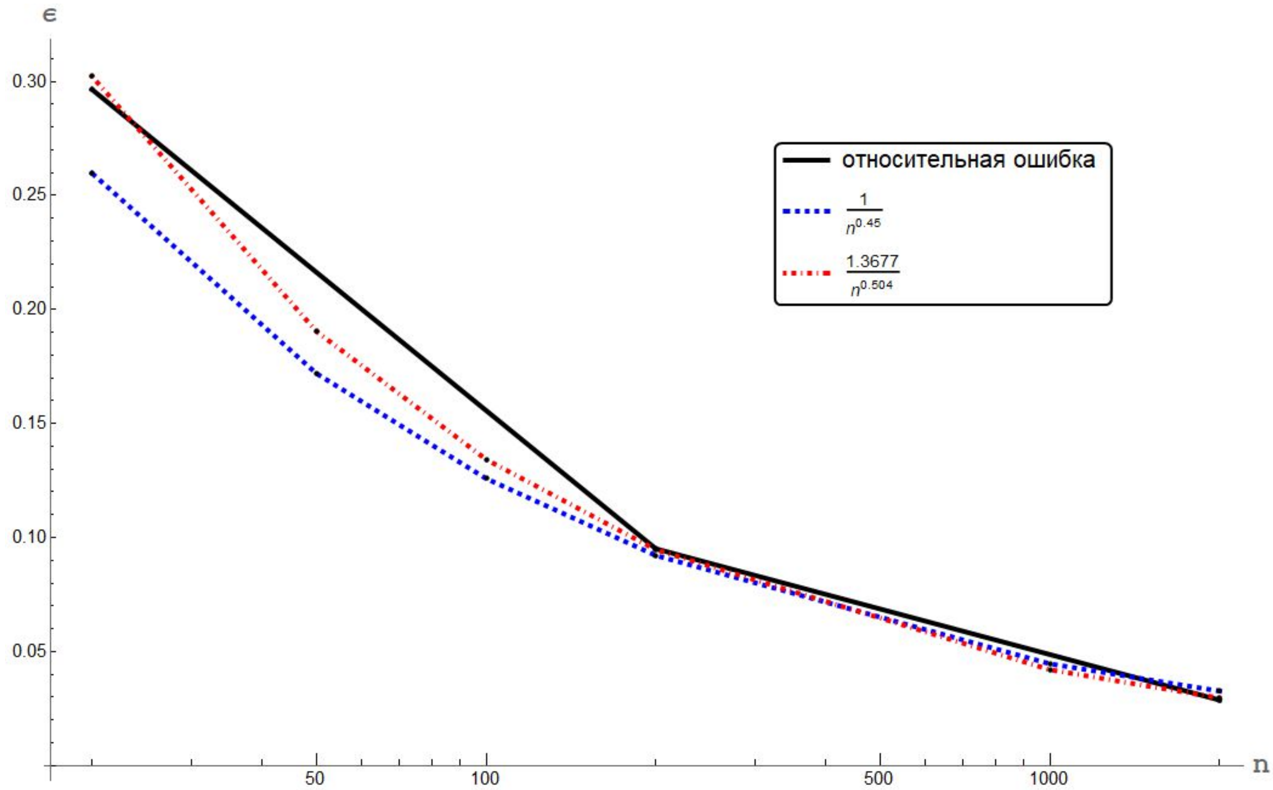

Рис. 2. Относительная ошибка вычисления вероятности по формуле Бернулли и по теореме Лапласа и ее возможные аппроксимации

Видно, что относительная ошибка уменьшается с ростом числа опытов. Характер убывания – обратный корневому, что согласуется с положениями теории погрешностей С.Г. Михлина.

Дидактический вывод по Заданию 2: с ростом числа испытаний растет точность вычислений значений интервальной вероятности по приближениям, опирающимся на интегральную теорему Лапласа. Это подтверждает визуальный вывод, который был сделан на основании выполнения Задания 1. В то же время даже при большом числе опытов значение относительной ошибки величиной около 3% нельзя назвать удовлетворительным.

Это значит, что пользоваться теоремой Лапласа вместо формулы Бернулли предпочтительнее для предварительного или качественного оценивания.

Задание 3. 1% проектов фирмы – крупные. Случайная величина (СВ) X – число крупных проектов. Построить полигон – график ряда распределения СВ X, если фирма выполняет проекты по формуле Бернулли и по формуле Пуассона для 1) n = 20; 2) n = 200.

Решение: значения параметров:

$$p = 0.01, q = 1 - p = 0.99 .$$

Очевидно, что событие «крупный проект» можно считать редким. Используя точную формулу Бернулли $P(X = i) = C_n^i p^i q^{n-i}$ и приближенную формулу Пуассона $P(X = i) = \frac{a^i}{i!} e^{-a}$, где $a = n.p$, получаем графики полигонов распределения (рис. 3).

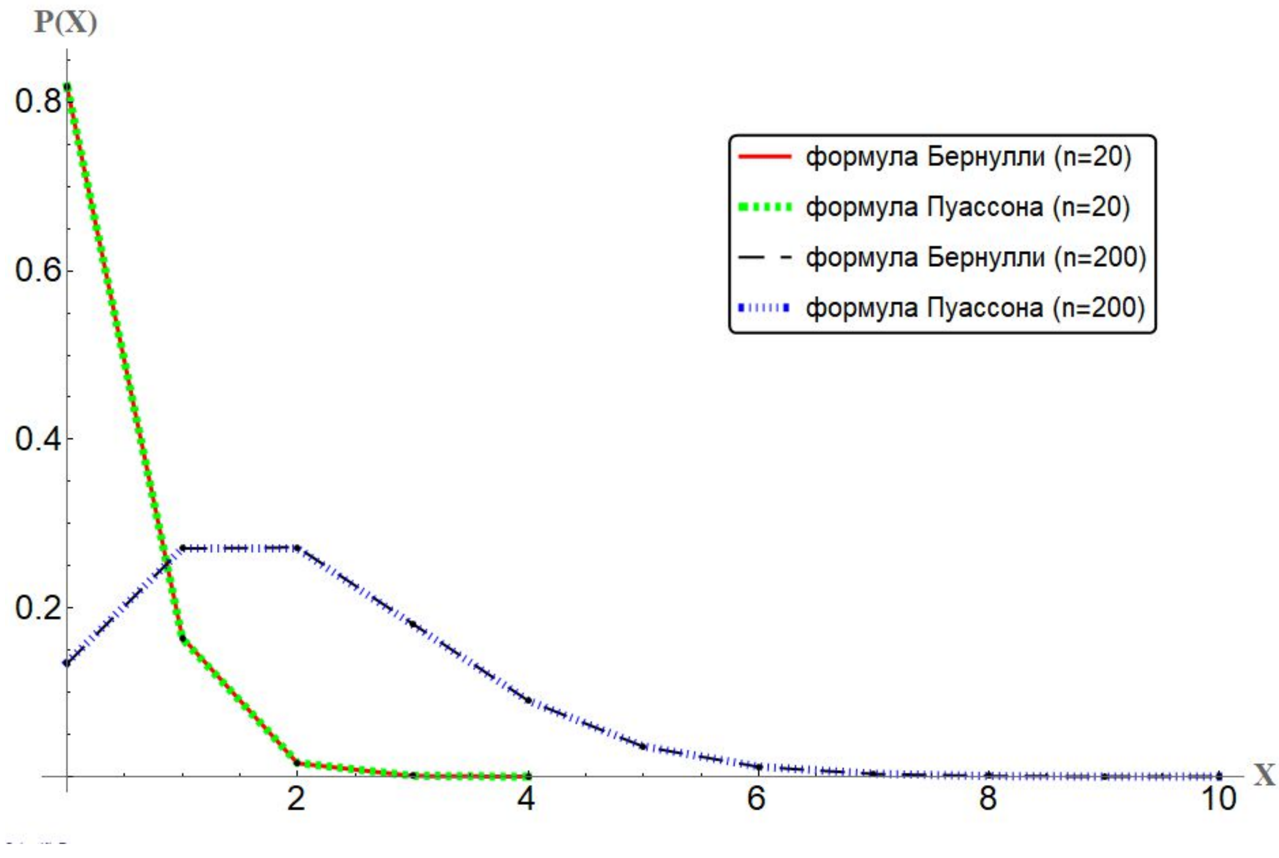

Рис. 3. Полигоны распределения числа редких событий

Дидактический вывод по Заданию 3: на основании числовых и визуальных данных можно констатировать, что вероятности редких событий (p = 0,1), вычисленные по точной формуле Бернулли и по приближенной формуле Пуассона, практически не различаются. Значит, для значений вероятности редкого события p ≤ 0,1 можно рекомендовать использование приближенной формулы Пуассона.

Задание 4. 1% проектов фирмы – крупные. Случайная величина (СВ) X – число крупных проектов. Найти по формуле Бернулли, по формуле Пуассона и по теореме Лапласа 1) P ($9 \le X \le 11$) при n=100; 2) P ($90 \le X \le 10$) при n = 1000.

Решение:

1) n = 100; P ($9 \le X \le 11$)

- по формуле Бернулли

$$P(9 \le X \le 11) = \sum_{i=9}^{11} C_{100}^{i} p^{i} q^{100-i} = 8.38046 \times 10^{-7}$$

- по формуле Пуассона

$$a = n.p = 1$$

$$P(9 \le X \le 11) = \sum_{i=9}^{11} \frac{a^{i}}{i!} e^{-a} = 1.12437 \times 10^{-6}$$

- по теореме Лапласа

$$x_9 = \frac{m - n.p}{\sqrt{npq}} = \frac{9 - 100 \times 0.01}{\sqrt{100 \times 0.01 \times 0.99}} = 8.0403$$

$$x_{11} = \frac{m - n.p}{\sqrt{npq}} = \frac{11 - 100 \times 0.01}{\sqrt{100 \times 0.01 \times 0.99}} = 10.0504$$

Обычно при $x > 5$ принимают значение функции Лапласа равным 0,5. Вычислив с большей точностью $\phi(x_{11}), \phi(x_9)$, получаем $P(9 \le X \le 11) \approx \phi(x_{11}) - \phi(x_9) \approx 4.42915 \times 10^{-16}$

2) n = 1000; P ($90 \le X \le 110$)

- по формуле Бернулли:

$$P(90 \le X \le 110) = \sum_{i=90}^{110} C_{1000}^{i} p^{i} q^{1000-i} = 1.28403 \times 10^{-54}$$

- по формуле Пуассона:

$$a = n.p = 10$$

$$P(90 \le X \le 110) = \sum_{i=90}^{110} \frac{a^{i}}{i!} e^{-a} = 3.43245 \times 10^{-53}$$

- по теореме Лапласа:

$$x_{90} = \frac{m - n.p}{\sqrt{npq}} = \frac{90 - 1000 \times 0.01}{\sqrt{1000 \times 0.01 \times 0.99}} = 25.4257$$

$$x_{110} = \frac{m - n.p}{\sqrt{npq}} = \frac{110 - 1000 \times 0.01}{\sqrt{1000 \times 0.01 \times 0.99}} = 31.7821$$

Обычно при $x > 5$ принимают значение функции Лапласа равным 0,5. Вычислив с большей точностью $\phi(x_{90}), \phi(x_{110})$, получаем $P(90 \le X \le 110) \approx \phi(x_{110}) - \phi(x_{90}) \approx 0$.

Дидактический вывод по Заданию 4: оценивание интервальной вероятности для редкого события (p = 0,1) по точной формуле Бернулли и по приближенной формуле Пуассона или по приближенной интегральной теореме Лапласа дает сходные результаты, отличающиеся в 6-м знаке после запятой и дальше. Значит, для интервальных значений вероятности редкого события $p \leq 0{,}1$ можно рекомендовать использование приближенной теоремы Лапласа, требующей минимума вычислительной работы.

Выводы

Авторами предложены варианты заданий, выполнение которых способствует формированию у обучающихся вероятностного подхода к познанию в части оценки точности вероятностных моделей. Несмотря на то что в отечественной средней школе изучаются основы теории вероятности, студенты 1–2-х курсов оказываются мало подготовленными к восприятию информации, связанной с построением вероятностных моделей явлений окружающего мира. С большим интересом и пониманием рассмотрение этих вопросов было встречено в аудитории иностранных обучающихся, готовящихся к поступлению в магистратуру [14, 15]. Это приводит к мысли о том, что организация программ предмагистерской подготовки была бы полезной для формирования качественного контингента магистратуры инженерных и экономических направлений подготовки.

Список литературы

1. Bonawitz E., Denison S., Griffiths T.L., Gopnik A. Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. Trends in Cognitive Sciences. 2014. V. 18 (10). P. 497-500. DOI: 10.1016/j.tics.2014.06.006.
2. Cusumano-Towner M.F., Mansinghka V.K. AIDE: An algorithm for measuring the accuracy of probabilistic inference algorithms. Advances in Neural Information Processing Systems. 30 (NIPS). MIT Press, 2017. P. 3004-3014.
3. Pita R., Mendonça E., Reis S., Barreto M., Denaxas S. (2017) A Machine Learning Trainable Model to Assess the Accuracy of Probabilistic Record Linkage. In: Bellatreche L., Chakravarthy S. (eds) Big Data Analytics and Knowledge Discovery. DaWaK 2017. Lecture Notes in Computer Science, V.10440. P. 214-227. Springer, Cham. DOI: 10.1007/978-3-319-64283-3_16.
4. Краснощеков В.В., Семенова Н.В. Инновационная методика преподавания теории вероятностей в больших потоках // Современные наукоемкие технологии. 2018. № 8. С. 199-203.
5. Болдыревский П.Б., Григорян М.Э., Зимина С.В. Ситуационные задачи в процессе обучения теории вероятностей как одно из средств реализации проектно-ориентированного

метода в высшей школе // Вестник Нижегородского университета им. Н.И. Лобачевского. 2014. № 3-4. С. 28-31.

6. Гефан Г.Д., Кузьмин О.В. Сравнительный анализ эффективности образовательных методик на примере обучения теории вероятностей и математической статистике // Вестник Томского государственного педагогического университета. 2017. № 4 (181). С. 49-56.

7. Григорян М.Э., Болдыревский П.Б. Компетентностные задачи по теории вероятностей как средство формирования профессиональной компетентности будущего экономиста // Успехи современного естествознания. 2015. № 3. С. 169-173.

8. Краснощеков В.В., Семенова Н.В. Формирование вероятностного подхода как методологии научного познания студентов вузов // Современные наукоемкие технологии. 2016. № 9-3. С. 515-519.

9. Гефан Г.Д. Теория вероятностей и математическая статистика: историко-философские аспекты в преподавании дисциплины // Омский научный вестник. 2015. № 1 (135). С. 123-126.

10. Гефан Г.Д., Кузьмин О.В. О вероятностном стиле мышления // Проблемы учебного процесса в инновационных школах / Под ред. О.В. Кузьмина. Иркутск, ИГУ, 2018. С. 32-37.

11. Григорян М.Э., Залесский М.Л. Методические аспекты изучения понятия вероятности случайного события // Образовательные технологии. 2018. № 1. С. 48-55.

12. Григорян М.Э., Болдыревский П.Б. Роль истории развития теории вероятностей в формировании у студентов научной картины мира как основы мировоззрения // Успехи современного естествознания. 2014. № 12-2. С. 133-137.

13. Васильева Т.В. Повышение стохастической культуры бакалавров в области информационной безопасности // Высшее образование сегодня. 2020. № 4. С. 14-16. DOI: 10.25586/RNU.HET.20.04.P.14

14. Краснощеков В.В., Рудь В.Ю., Давыдов В.В. Модели обучения иностранных предмагистрантов инженерных профилей подготовки // Современные наукоемкие технологии. 2018. № 12-1. С. 214-219. DOI: 10.17513/snt.37290.

15. Krasnoshchekov V., Arseniev D., Rud’ V., Switala F., Chetiy V. Improving the quality of pre-master training of foreign students in the field of environment. ECOBALTICA 2019, IOP Conf. Series: Earth and Environmental Science, V. 390, No 1, 012017. IOP Publishing LTd, 2019. URL:

16. https://ioscience.iop.org/artcle/10.1088/1755-1315/390/1/012017/pdf (access data 10.09.2020).DOI: 10.1088/1755-1315/390/1/012017.