УДК 324.311
DOI https://doi.org/10.32782/cusu-pmtp-2024-1-11

ПРИКЛАД ПОБУДОВИ ТЕОРЕТИЧНОГО ЗАКОНУ РОЗПОДІЛУ ЗА ЕМПІРИЧНИМИ ДАНИМИ ДЛЯ ПРАКТИЧНИХ ЗАНЯТЬ В КУРСІ ТЕОРІЇ ІМОВІРНОСТЕЙ

Селезньова Надія Петрівна,
кандидат фізико-математичних наук,
доцент кафедри математичної фізики та диференціальних рівнянь
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID ID: 0000-0003-0849-3092

Кушлик-Дивульська Ольга Іванівна,
кандидат фізико-математичних наук,
доцент кафедри математичної фізики та диференціальних рівнянь
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID ID: 0000-0002-4999-6641

Рудик Тетяна Олександрівна,
кандидат фізико-математичних наук,
доцент кафедри математичної фізики та диференціальних рівнянь
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID ID: 0000-0003-1121-4963

Робота присвячена методиці викладання важливого розділу з курсу теорії імовірностей – побудові та перевірці статистичних гіпотез.

У статті розглянуто можливості застосування методів стохастики до задач, пов'язаних з електоральними дослідженнями, що є актуальним у процесі викладання курсу теорії імовірностей студентам соціологічного напряму. Опанування методів цієї дисципліни дає змогу будувати моделі стохастичних процесів і явищ та аналізувати їх. На основі кількісних опитувань та анкетування студенти мають навчитися визначати обсяг вибірки, тобто скількох людей потрібно опитати, щоб за їхніми відповідями можна було зробити коректні висновки, а також як саме обробляти й аналізувати зібрані статистичні дані.

У статті наведено приклад опитування, проведеного серед студентів факультету соціології і права НТУУ «КПІ ім. І. Сікорського», і обробки його результатів. Студентам було запропоновано самостійно сформувати емпіричні дані шляхом анкетування своїх колег. Було поставлено завдання зібрати рейтингові оцінки п'яти самих популярних партій на той час. Таким чином, досліджувалися політичні погляди студентів напередодні виборів, а саме було поставлено завдання оцінити в балах запропоновані п'ять партій, виділити ті, які їм подобаються найбільше і найменше.

Для аналізу результатів опитування вводиться випадкова величина – кількість респондентів, які поставили максимальний бал певній партії, і також розглянуто розподіл, де випадковою величиною є кількість респондентів, які поставили мінімальний бал певній партії, тобто кількість тих, хто поставив конкретній партії максимальні бали – п'ятірки та мінімальні бали – одиниці. Для цих розподілів за допомогою критерію Пірсона хі-квадрат перевіряються гіпотези про їх належність до біноміального та Пуассонівського розподілу. Ці розподіли тісно пов'язані з нормальним розподілом генеральної сукупності, який покладено в основу багатьох критеріїв. На цьому прикладі продемонстровано, що недостатньо покладатися тільки на графічну інтерпретацію розподілу навіть тоді, коли візуально графіки емпіричного й теоретичного розподілів практично збігаються.

Ключові слова: *критерій Пірсона хі-квадрат, біноміальний розподіл, Пуассонівський розподіл, електоральні дослідження.*

Seleznova Nadiia, Kushlyk-Dyvulska Olga, Rudyk Tetiana. An example of constructing a theoretical distribution law based on empirical data for practical classes in probability theory

The article is devoted to the methodology of teaching an important section of the probability theory course, namely, the construction and testing of statistical hypotheses.

The article considers the possibilities of applying stochastic methods to problems related to electoral research, which is relevant in the process of teaching a course on probability theory to students of sociology. Mastering the methods of this discipline allows you to build models of stochastic processes and phenomena and analyze them. On the basis of quantitative surveys and questionnaires, students should learn how to determine the sample size, i.e., how many people should be interviewed to draw correct conclusions from their answers, as well as how to process and analyze the collected statistical data.

The article presents an example of a survey conducted among students of the Faculty of Sociology and Law of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» and the processing of its results. Students were suggested to generate empirical data by questioning their colleagues. The task was to collect ratings of the five most popular parties at that time. Thus, the political views of students on the eve of the election were studied, namely, the task was to evaluate the proposed five parties in points, to highlight those they liked the most and least.

To analyze the results of the survey, a random variable is introduced – the number of respondents who gave the maximum score to a particular party, and a distribution is also considered where the random variable is the number of respondents who gave the minimum score to a particular party, i.e. the number of those who gave a particular party the maximum score – five, and the minimum score – one. For these distributions, the hypotheses that they belong to the binomial and Poisson distributions are tested using the Pearson chi-square test. These distributions are closely related to the normal distribution of the population, which is the basis for many criteria. This example demonstrates that it is not enough to rely only on the graphical interpretation of the distribution, even when the graphs of the empirical and theoretical distributions are virtually identical.

Key words: *Pearson's chi-square criterion, binomial distribution, Poisson distribution, electoral studies.*

Вступ. Компетентність випускника факультету будь-якого вищого навчального закладу потрібно формувати в процесі вивчення не тільки гуманітарних, але й інших дисциплін, серед яких чільне місце займає математика, особливо окремий її розділ – теорія ймовірностей та математична статистика. Адже саме вона є базою для розуміння змісту й методів математичного моделювання, що потрібні для задач соціології та менеджменту, прикладної лінгвістики, дає можливість адекватно інтерпретувати отримані результати. За останні десятиліття математика перетворилася з методу обчислень у метод досліджень, визначаючи методологію проведення експерименту та розуміння його результатів. Інакше кажучи, математика стала дедуктивною наукою про структури і має чітке розмежування аналізу математичної структури та її інтерпретацій.

Для практичної чи науково-дослідницької діяльності студентам потрібно засвоїти знання про світ випадкових явищ – подій, величин і процесів, про правила перевірки гіпотез та оцінки достовірності висновків і, нарешті, про аналіз даних спостережень чи експерименту. При цьому необхідним є сполучення у викладанні абстрактного і конкретного, індукції та дедукції. Особливо слід приділити увагу наданню простих прикладів у сполученні аналітичних формалізмів із геометричною наочністю. У запропонованій роботі представлено ілюстративні приклади дослідження передвиборчої ситуації стохастичними методами.

Аналіз досліджень і публікацій. Потреба у використанні в соціологічних дослідженнях математичних методів стала очевидною завдяки роботам Конта, Кетле, Парето. Але тільки на початку XX століття було застосовано найпростіші стохастичні засоби до задач соціології [1; 2]. У 1900 році Карл Пірсон запропонував простий та ефективний спосіб перевірки згоди між передбаченням моделі й емпіричними даними. Запропонований ним метод був універсальним, адже більшість завдань, пов'язаних з оцінкою невідомих параметрів моделі та перевірки за допомогою критерію згоди хі-квадрат моделі і емпіричних даних, можна вирішити за його допомогою. Саме застосуванню цього критерію згоди в електоральних дослідженнях і присвячена запропонована робота. Але не тільки в задачах соціологічних досліджень застосовується критерій згоди хі-квадрат. Приклади його застосування можна знайти і в медичних дослідженнях [4]. У роботі [3] розглянуто концептуальні підходи статистичної методології в дослідженні взаємозв'язків економічних явищ і процесів. Загальні проблеми визначення обсягу вибірки й ілюстрації формул для обчислення цього обсягу розглянуто в роботі [5].

Авторами цієї роботи розроблено й описано методики застосування статистичних методів до задач соціології, а саме дослідження електорального вибору напередодні парламентських і президентських виборів [6; 7]. Критерій Пірсона хі-квадрат описано в [4; 8; 9; 10].

Матеріали та методи. Матеріалом дослідження запропонованої роботи є зібрані емпіричні дані опитування певної групи респондентів з приводу їх ставлення до певних політичних партій напередодні парламентських виборів. У ході дослідження використовувалася низка методів, як-от аналіз наукових і методичних джерел, з яких брався матеріал з теорії ймовірностей і математичної статистики, потрібний для опрацювання зібраних емпіричних даних стосовно дослідження структури статистичних розподілів сукупностей випадкових подій і техніки експериментальної перевірки гіпотез.

Результати. Метою запропонованої статті є методика вивчення окремих розділів статистичного аналізу даних і розвиток практичних навичок їх використання в курсі викладання теорії імовірностей та математичної статистики. Студентам запропоновано самостійно зібрати статистичні дані, а саме провести опитування студентів свого факультету з приводу їх ставлення до п'яти найбільш популярних партій напередодні чергових виборів до Верховної Ради України. Слід зауважити, що в масштабах країни ця вибірка є непрезентативною. Досвід зі створення такої вибірки та її аналіз методами теорії імовірностей стане корисним студентам у професійній діяльності У подальшому студенти проводять первісну статистичну обробку даних – групують їх за певними ознаками, обчислюють мінімальну кількість респондентів, яка б надала репрезентативність вибірці, обчислюють точкові оцінки та надають їм певну інтерпретацію, встановлюють закон розподілу отриманих даних.

Дослідження будувалося за такою схемою: респондентам було запропоновано проранжувати п'ять найбільш популярних партій (позначимо їх A, B, C, D, E) таким чином: 5 балів отримує партія, яка є найкращою, на думку виборця (респондента), 4 – партія, яка стоїть на другому місці… 1 – найгірша партії, з погляду респондентів, 0 – партія, до якої респонденти є байдужими або не мають про неї інформації. У досліджені взяли участь n = 223 випадково обраних студентів факультету соціології і права НТУУ «КПІ ім. І. Сікорського». Далі до одержаних даних застосували методи статистичної обробки.

Випадковою величиною в нашому дослідженні є кількість партій, які обирає респондент серед п'яти партій. Обирає – це означає, що таким партіям ставлять п'ять балів, не обирає – означає, що необраним партіям ставлять по одному балу. Звісно, що є партії, які отримують не по одному чи п'ять балів, але вони не є об'єктами наших досліджень саме в цій роботі. Введену випадкову величину позначимо через X.

Процес вибору партій респондентом представляємо як п'ять дослідів, щодо яких зробимо такі припущення:

1) ці досліди незалежні, тобто ймовірність обрання будь-яким респондентом будь-якої партії не залежить від того, скільки буде обрано чи не обрано ним інших партій;

2) імовірність обрання респондентом будь-якої окремо взятої партії є тією самою та дорівнює p, а ймовірність необрання респондентом певної партії – $q = 1 - p$.

Звісно, що стосовно цих припущень можуть виникати деякі сумніви, але, можливо, вони не суперечитимуть результатам спостережень, що надалі ми і перевіримо за допомогою критерію χ^2 Пірсона.

Спочатку з'ясуємо, яку ж кількість респондентів слід опитати для отримання достатньо достовірних результатів. За генеральної сукупності N = 450 (усього студентів на факультеті) з надійністю $\gamma = 0{,}95$ і граничною похибкою $\delta = 0{,}05$, якщо $p = q = 0{,}5$, обсяг вибірки n обчислюємо за відомою формулою [5; 8], t для підстановки у формулу знаходимо з рівняння $\Phi(t)$, де $\Phi(t)$ – функція Лапласа:

$$n = \frac{N \cdot t^2 \cdot p \cdot q}{t^2 \cdot p \cdot q + N \cdot \Delta^2} = \frac{450 \cdot 3{,}84 \cdot 0{,}25}{3{,}84 \cdot 0{,}25 + 450 \cdot 0{,}0025} = 207.$$

Отже, кількість опитаних респондентів у цьому дослідженні має бути не меншою за 207. Тоді довірча ймовірність, з якою проведено дослідження, буде $\Delta \geq 0{,}05$.

Для прийнятих припущень маємо справу з повторними дослідами Бернуллі, тому число обраних партій серед п'яти буде мати біноміальний закон розподілу, тобто ймовірність того, що респондент обере x партій, обчислюється за відомою формулою Бернуллі [6; 10; 11].

$$P_5(X = x) = C_5^x \cdot p^x \cdot (1-p)^{5-x};\ x = 0, 1, 2, 3, 4, 5. \tag{1}$$

Знайдемо оцінку параметра p, що входить у модель. Зауважимо, що в умовах випробувань Бернуллі незміщеною оцінкою ймовірності є частість. У цьому випадку p – імовірність того, що респондент поставить певній партії п'ять балів, тому частість цієї події обчислюємо таким чином:

$$\hat{p} = \frac{\sum_{i=1}^{6} x_i n_i}{5 \cdot 223} = \frac{0 \cdot 105 + 1 \cdot 102 + 2 \cdot 13 + 3 \cdot 2 + 4 \cdot 0 + 5 \cdot 1}{1115} = \frac{139}{1115} = 0{,}12,$$

де n_i – кількість респондентів, які поставили бали $\tilde{o}_s$ ($\tilde{o}_s : 0,1,2,3,4,5$) різним партіям. Числові дані наведено в табл. 2.

Перейдемо до конкретики.

На основі зібраних статистичних даних обчислено такі точкові оцінки щодо кожної з партій (табл. 1).

Таблиця 1

Основні точкові оцінки розподілу голосів респондентів відносно партій

	A	**B**	**C**	**D**	**E**	**ABCDE**
Хсер	2,12	3,03	2,89	1,98	1,51	2,31
D	2,46	2,77	2,16	1,97	1,68	2,53
Sigma (σ)	1,57	1,66	1,47	1,40	1,30	1,59
Мо	1	5	4	2	1	1
Ме	2	3	3	2	1	2
Е	–0,71	–1,24	–0,92	–0,75	0,61	–1,10
А	0,62	–0,31	–0,34	0,28	1,05	0,27
V (%)	74	55	51	71	86	69

У табл. 1 обчислено точкові оцінки окремо для кожної з партій $A, B, C, D, E.$ $ABCDE$ – усі вибіркові дані щодо всіх партій у сукупності: Хсер – середня, Sigma(σ) – середнє квадратичне відхилення, Мо – мода, Ме – медіана, Е – ексцес, А – асиметрія, V – варіація.

Відомо, що нормальний закон розподілу характеризується тим, що $Mo = Me = X_{\tilde{n}\eth\eth}$, $A = E = 0$ [7; 8; 9]. Виходячи з даних табл. 1, бачимо, що найближчим до нормального розподілу є розподіл рейтингових оцінок у партії D та загальний розподіл оцінок $ABCDE$.

Також важливою характеристикою є коефіцієнт варіації. Якщо він високий, то це вказує на неодностайність уподобань виборців щодо певного кандидата (вибірка є неоднорідною).

Стійкість результатів голосування характеризує середнє квадратичне відхилення. Партії, у яких Sigma (σ) набуває найбільшого значення, мають найбільш стійку позицію, тобто мають більше шансів на виборах (за умови проведення їх в межах досліджуваної сукупності) отримати той же результат, що і в цьому дослідженні. Партії з найменшим Sigma мають найменш стійку позицію.

Партія з найбільшим середнім балом користується найбільшою довірою виборців.

Результати попередніх обчислень статистичних показників наших спостережень досить часто є достатніми для того, щоб сформулювати гіпотезу про модель закону розподілу нашої випадкової величини. Далі можна підставити у вибрану модель зібрані дані й оцінити її параметри.

Перейдемо безпосередньо до оцінки закону розподілу. Кожен з опитаних респондентів проранжував по 5 партій. Розглянемо подію $A - \{\text{поява } 5\}$. За результатами одержали такий розподіл випадкової величини: X – число появи події A в оцінках кожного респондента. Тоді загалом щодо всіх респондентів отримаємо такий розподіл:

Таблиця 2

Розподіл голосів усіх респондентів

x_i	0	1	2	4	4	5
n_i	105	102	13	2	0	1

Спробуємо встановити, якому закону розподілу випадкової величини відповідають дані табл. 2, шляхом перевірки гіпотези про розподіл генеральної сукупності X. Ми вже створили вибірку з генеральної сукупності, на основі якої перевіримо гіпотезу про біноміальний розподіл нашої сукупності за допомогою критерію χ^2 Пірсона. Цей критерій засновано на вимірюванні відхилень між емпіричними та теоретичними частотами, що відповідають різним інтервалам гістограми побудованої на емпіричних частотах n_i і теоретичних частотах np_i для кожного інтервалу i. Міра відхилення між теоретичними й емпіричними частотами визначається за формулою [4; 9; 10]:

$$\chi^2 = \sum_{i=1}^{N} \frac{(n_i - np_i)^2}{n_i}.$$

Коли $N \to \infty$, велична χ^2 асимптотично прямує до щільності ймовірності χ^2 – розподілу з $N-k-1$ степенями свободи, де k – кількість параметрів, оцінених на основі первісної інформації про теоретичний закон розподілу. Для біноміального закону розподілу $k=1$.

Нульова гіпотеза, яка стверджує, що отримана вибірка з теоретичного розподілу, приймається, якщо $\chi^2 < \chi^2_{N-k-1,1-\alpha}$, де $\chi^2_{N-k-1,1-\alpha}$ – значення χ^2 за $N-k-1$ степенях свободи, α – рівень значущості критерію.

Обчислення відповідно до критерію χ^2 Пірсона представлено такою таблицею:

Таблиця 3

Розрахункова таблиця для визначення статистики χ^2 розподілу голосів усіх респондентів

i	**Число 5** m_i	**Число виборців** m_i	**Частість** $\hat{p} = \frac{m_i}{n}$	$p_i^T = C_5^{x_i} \cdot p^{x_i} \cdot (1-p)^{5-x_i}$	$m_i^T = np_i^T$	$\left(m_i - m_i^T\right)^2$	$\frac{\left(m_i - m_i^T\right)^2}{m_i^T}$
1	2	3	4	5	6	7	8
1	0	105	0,47	0,51	115	100	0,8
2	1	102	0,46	0,37	82	400	4,8
3	2	13	0,06	0,1	26	100	3,8
4	3	2	0,01	0,01	3		
5	4	0	0	0	0		
6	5	1	0,004	0	0		

(літерою T в таблиці позначено те, що відповідні величини є теоретичними ймовірностями чи частотами, тобто це не степінь, а індекс)

Отже, якщо $n = 223$, виходячи з даних таблиці, маємо: $\chi_{ñïîñòåðåæ} = 9,4$.

Оскільки розбіжність між теоретичними й емпіричними частотами (стовпці 4, 5 табл. 3) не досить велика, то в першому наближенні можна прийняти біноміальну модель. Графічно це підтверджує рис. 1, на якому крива теоретичних імовірностей є досить близькою до кривої частостей.

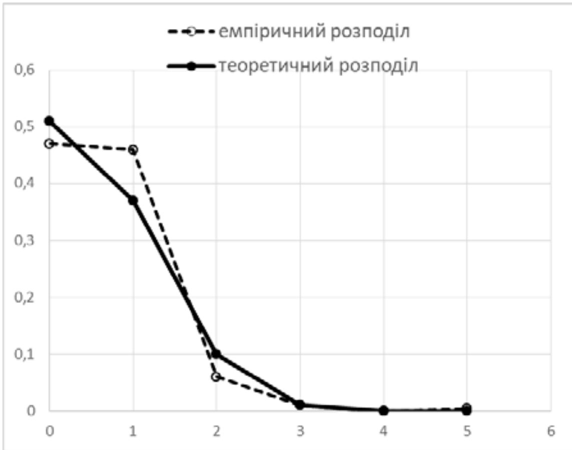

Рис. 1. Многокутники теоретичного (біноміального) та емпіричного розподілів голосів респондентів

Тепер застосуємо більш точний, глибший метод прийняття рішення про адекватність моделі закону розподілу, а саме критерій узгодженості. Порівняємо емпіричні і теоретичні частоти за допомогою критерію Пірсона. Розрахунки для цього представлено в стовпчиках 7 та 8 табл. 3, причому об'єднуємо рядки з частотами, меншими за п'ять, і відповідні їм теоретичні ймовірності так, щоб $m_i > 5$. За таблицею критичних точок розподілу χ^2 за рівнем значущості $\alpha = 0,01$ і числу степенів свободи $k = 3-1 = 2$ знаходимо критичну точку правосторонньої критичної області $\chi^2_{êð}(0,01;2) = 9,2$. Оскільки $\chi^2_{ñïîñòåðåæ} > \chi^2_{êð}$, то гіпотезу про біноміальну модель відхиляємо.

Тепер розглянемо подію {поява 1}. Число поставлених одиниць респондентами відповідним партіям є випадковою величиною. Позначимо її через *X* (табл. 4).

Таблиця 4

Розподіл числа одиниць поставлених респондентами партіям

X	**0**	**1**	**2**	**3**	**4**	**5**
n_i	59	90	44	17	9	4

Спробуємо встановити закон розподілу цієї величини. Перевіримо гіпотезу про те, що *X* – розподілена згідно із законом Пуассона. У такому випадку за оцінку параметра λ розподілу Пуассона візьмемо вибіркову середню:

$$\lambda = \frac{\sum_{i=1}^{6} x_i n_i}{n} = \frac{0\cdot59+1\cdot90+2\cdot44+3\cdot17+4\cdot9+5\cdot4}{223} = \frac{285}{223} = 1,28.$$

На закон розподілу Пуассона також вказує той факт, що дисперсія нашої випадкової величини *X* – (*D(X) = 1,3)*, що мало відрізняється від її математичного сподівання.

Наша дискретна випадкова величина *X*, якщо розподілена за законом Пуассона з параметром λ, то набуває значень 0, 1, 2, 3, 4, 5 з імовірностями $P_{223}(x_i) = \frac{(\lambda)^{x_i} \cdot e^{-\lambda}}{x_i\,!}$ [8; 9]. Аналогічно до вже розглянутої табл. 3 будуємо табл. 5.

Таблиця 5

Розрахункова таблиця для визначення статистики χ^2 розподілу одиниць, поставлених респондентами партіям

i	Число 1 x_i	Число виборців n_i	$\frac{n_i}{223}$	$P_{223}(x_i) = (\lambda)^{x_i} \times$ $\times e^{-\lambda} / x_i\,!$	$n_i^T = np_i^T$	$\left(n_i - n_i^T\right)^2$	$\frac{\left(n_i - n_i^T\right)^2}{n_i^T}$
1	2	3	4	5	6	7	8
1	0	59	0,26	0,28	62	9	0,15
2	1	90	0,40	0,36	79	121	1,53
3	2	44	0,20	0,23	51	49	0,93
4	3	17	0,08	0,1	22	25	1,14
5	4	9	0,04	0,031	7	16	1,78
6	5	4	0,02	0,01	2		

Графічну ілюстрацію закону розподілу зображено на рис. 2. На ньому спостерігаємо, що крива теоретичних імовірностей є досить близькою до кривої емпіричних імовірностей.

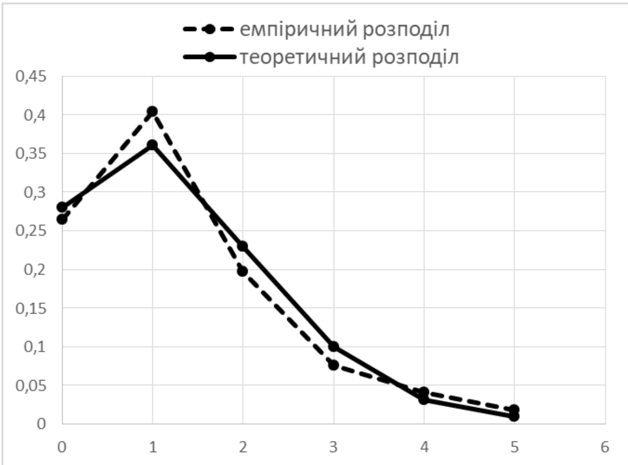

Рис. 2. Многокутники теоретичного (Пуасонівського) та емпіричного розподілів голосів респондентів

Порівняємо теоретичні й емпіричні частоти за допомогою критерію Пірсона. Для цього створили розрахункову табл. 5. З неї знаходимо емпіричне значення критерію Пірсона: $\chi^2_{\text{ăì ï³ðè÷}} = 5,5$.

За таблицею критичних точок розподілу χ^2 [6; 11] за рівнем значущості $\alpha = 0,05$ і числом степенів свободи $k = 5 - 2 = 3$ знаходимо критичну точку правосторонньої критичної області: $\chi^2_{\hat{e}\eth}(0,05;3) = 7,8$. Оскільки χ^2, то гіпотезу про розподіл випадкової величини *X* за законом Пуассона приймаємо.

Висновки. Проведено дослідження передвиборчої ситуації в Україні, а саме зібрано емпіричні дані про політичні вподобання певного прошарку суспільства – студентів НТУУ «КПІ ім. І. Сікорського», показано, як на основі масиву даних можна ввести певну випадкову величину, обчислити її стандартні характеристики та спробувати визначити закон розподілу, що достатньо добре моделює таку величину. На основі аналізу обчислених стандартних характеристик зроблено припущення про закон розподілу введеної до розгляду випадкової величини. За допомогою критерію згоди хі-квадрат Пірсона перевірено належність емпіричних даних до біноміального та Пуассонівського законів розподілу. Встановлено, що розподіл відповідає закону розподілу Пуассона. Окремо цікавим є те, що для перевірки того, наскільки ж добре підходить обраний закон розподілу, існує чіткий критерій, що дає однозначну відповідь «так» або «ні», однієї наочної подібності графіків недостатньо.

З наведених прикладів можна зробити висновок, що встановлювати закон розподілу випадкової величини, спираючись тільки на графічну ілюстрацію та співвідношення точкових оцінок, не можна. Обов'язково слід перевіряти гіпотезу про розподіл випадкової величини за допомогою одного з критеріїв згоди. У цій роботі було застосовано критерій згоди Пірсона. Теорія ймовірностей та її методи є важливою частиною дисциплін під час вивчення як математики, так і дисциплін економічного, психологічного, педагогічного, соціологічного й інших напрямів у ЗВО. Тому викладений матеріал є корисним для студентів і викладачів ЗВО, адже подібні приклади обов'язково слід демонструвати студентам, оскільки когнітивні упередження підштовхують до того, щоб покладатися на наочність. Це ще раз підкреслює те, що математика, і теорія ймовірностей зокрема, є контрінтуітивною наукою.

Література:

1. Kendall M., Stuart A. The Advanced theory of statistics. London: Charles Griffin & Company limited, Vol. 2: Inference and Relationship, 1966. 877 p.

2. Опря А. Т. Статистика. Бібліотека українських підручників. URL: https://westudents.com.ua/knigi/579-statistika-oprya-at.html#google_vignette (дата звернення: 09.03.2024).

3. Опря А. Т. Наукова концепція статистичної методології: методи, показники, критерії надійності. *Вісник Полтавської державної аграрної академії.* 2013. № 2. С. 109–119. URL: https://doi.org/10.31210/visnyk2013.02.31.

4. Ясинська Е. Ц. Застосування критерію Хі-квадрат для виявлення соціально-культурних чинників на виникнення порушень ритму і провідності серця. *Буковинський медичний вісник.* 2007. Т. 11, № 4. С. 153–155.

5. Bartlett J.E., Kotrlik J.W., Higgins C.C. Organizational research: Determining appropriate sample size for survey research. Information Technology, *Learning, and Performance Journal*. Vol. 19 (1). 2001. P. 43–50.

6. Селезньова Н. П., Сараєва Ю. О. Точкові оцінки числових характеристик дискретного розподілу в контексті виборів президента. *Математика в сучасному технічному університеті* : матеріали VIII Міжнар. наук.-практ. конф., м. Київ, 26–27 груд. 2019 р. / НТУУ «КПІ ім. І. Сікорського», 2020. С. 147–152.

7. Селезньова Н. П., Сараєва Ю. О. Математичне моделювання оцінок впливу політичних партій на прикладі виборів в Україні 2019 року. *Молодий вчений*. Херсон. 2020. № 2 (78). С. 207–213.

8. Карташов М. В. Імовірність, процеси, статистика. Київ : ВПЦ Київ, ун-т, 2007. 503 с.

9. Гіхман Й. І., Скороход А. В., Ядренко М. Й. Теорія ймовірностей і математична статистика. Київ : Вища школа, 1988. 440 с.

10. Meyer M.C. Probability and Mathematical Statistics: Theory, Applications, and Practice in R. SIAM – Society for Industrial and Applied Mathematics. 2019. 707 p.

References:

1. Kendall, M., & Stuart, A. (1966). The Advanced theory of statistics. London: Charles Griffin & Company limited, Vol. 2: Inference and Relationship. 877 p. [in English].

2. Opria, A.T. Statystyka. Biblioteka ukrainskykh pidruchnykiv [Statistics. Library of Ukrainian textbooks]. Retrieved from: https://westudents.com.ua/knigi/579-statistika-oprya-at.html#google_vignette (date of access: 09.03.2024) [in Ukrainian].

3. Opria, A.T. (2013). Naukova kontseptsiia statystychnoi metodolohii: metody, pokaznyky, kryterii nadiinosti [Scientific concept of statistical methodology: methods, indicators, reliability criteria]. *Visnyk Poltavskoi derzhavnoi ahrarnoi akademii*, 2, p. 109–119 [in Ukrainian].

4. Yasyns'ka, E.Ts. (2007). Zastosuvannia kryteriiu Khi-kvadrat dlia vyiavlennia sotsialno-kulturnykh chynnykiv na vynyknennia porushen rytmu i providnosti sertsia [The use of χ^2 – criterion to detect a combined effect of social– cultural factors on the onset of disturbances of the heart rate and cardial conduction]. Bukovinian State Medical University (Chernivtsi) *Buk. Med. Herald.* Vol. 11, № 4. P. 153–155 [in Ukrainian].

5. Bartlett, J.E., II; Kotrlik, J.W., & Higgins, C. (2001). Organizational research: Determining appropriate sample size for survey research. Information Technology, *Learning, and Performance Journal*. 19 (1). P. 43–50 [in English].

6. Seleznova, N.P., Saraieva, Yu.O. (2019). Tochkovi otsinky chyslovykh kharakterystyk dyskretnoho rozpodilu v konteksti vyboriv prezydenta. Matematyka v suchasnomu tekhnichnomu universyteti [Point estimates of numerical characteristics of a discrete distribution in the context of presidential elections. Mathematics in a modern technical university]. Materialy VIII Mizhnarodnoi naukovo-praktychnoi konferentsii 26–27 hrudnia. P. 147–152 [in Ukrainian].

7. Seleznova, N.P., & Saraieva, Yu.O. (2020). Matematychne modeliuvannia otsinok vplyvu politychnykh partii na prykladi vyboriv v Ukraini 2019 roku [Mathematical modeling of estimates of the influence of political parties on the example of the 2019 elections in Ukraine]. *"Molodyi vchenyi"*, № 2 (78), P. 207–213 [in Ukrainian].

8. Kartashov, M.V. (2007) *Imovirnist, protsesy, statystyka [Probability, processes, statistics]*. Kyiv: VPTs Kyivskyi universytet. 503 p. [in Ukrainian].

9. Hykhman, Y.I., Skorokhod, A.V., & Yadrenko, M.Y. (1988). *Teoriia iimovirnosteii i matematychna statistika [Probability Theory and Mathematical Statistics]*. Kyiv: Higher School. 440 p. [in Ukrainian].

10. Meyer, M.C. (2019). Probability and mathematical statistics: Theory, applications, and practice in R. SIAM, 24 Juny. 690 p. [in English].