

МІНІСТЕРСТВО ОСВІТУ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ПОЛІТЕХНІЧНИЙ УНІВЕРСИТЕТ

ІНСТИТУТ КОМП'ЮТЕРНИХ СИСТЕМ

МАТЕРІАЛИ ДЕВ'ЯТОЇ
МІЖНАРОДНОЇ НАУКОВОЇ КОНФЕРЕНЦІЇ
СТУДЕНТІВ ТА МОЛОДИХ ВЧЕНІХ



ПРИСВЯЧЕНА 55-РІЧЧЮ
ІНСТИТУТУ КОМП'ЮТЕРНИХ СИСТЕМ

“ Сучасні інформаційні технології 2019 ”

“ Modern Information Technology 2019 ”



NetCracker®



23-24 травня

Одеса
«Екологія»
2019

УДК 658.512.2.011.56:612.846

**ДОСЛІДЖЕННЯ УМОВ ЗАСТОСУВАННЯ СТАТИСТИЧНИХ МЕТОДІВ ПРИ
АВТОМАТИЗАЦІЇ РІШЕНЬ ЗАДАЧ ПРОГНОЗУВАННЯ**

Комлева Н. О., Дубовий Р. В.

к.т.н., доцент каф. СПЗ Комлева Н.О.

Одеський Національний Політехнічний Університет, УКРАЇНА

АНОТАЦІЯ. Проведено аналіз та надано рекомендації щодо одержування автоматизованого рішення для найбільш відомих методів математичної статистики, що дозволяють отримувати прогностичне рішення з урахуванням припущення про можливість поширення минулих і теперішніх тенденцій і закономірностей на майбутній розвиток обраного об'єкта прогнозування.

Вступ. Прогнозування є важливою метою і завданням великого числа фахівців, що займаються аналізом даних. Сучасні методи статистичного прогнозування дозволяють з високою точністю прогнозувати практично всі можливі показники.

При аналізі часових рядів можна виділити дві основні мети: визначення природи часового ряду; прогнозування (передбачення майбутніх значень часового ряду по теперішнім і минулим значенням). Водночас треба пам'ятати, що не існує універсальних методів прогнозування на всі випадки життя. Вибір методу прогнозування та його ефективність залежать від багатьох умов, і зокрема від необхідної довжини або часу прогнозування [1, 2].

Мета роботи. Метою роботи є надання практичних рекомендацій щодо автоматизації певних методів математичної статистики для отримання прогностичного рішення з заданою точністю та з урахуванням особливостей закономірностей подання досліджуваного показника.

Основна частина роботи. Суть будь-яких методів прогнозування полягає в отриманні прогностичного значення досліджуваного параметра з певною точністю рішення з урахуванням припущення про можливість поширення минулих і теперішніх тенденцій і закономірностей на майбутній розвиток обраного параметра. Найбільш поширеними статистичними методами прогнозування є: метод рухомого середнього, найменших квадратів і експоненціального згладжування. Всі вони є методами екстраполяції – це головні способи прогнозу, що аналізують значення показників, які мали місце в минулі роки. Для того, щоб отримати прогностичне значення досліджуваного показника, потрібно мати його реальні значення за попередні часові відліки. Нажаль, не завжди протягом попереднього періоду зберігаються умови вимірювання реального значення параметра, для якого потрібно отримати прогноз.

Розглянемо завдання прогнозування результатів зовнішнього незалежного оцінювання (ЗНО) на наступний 2019 рік, ґрунтуючись на результатах минулих років, починаючи з 2010 року. Дані знаходяться на офіційному сайту українського центру оцінювання якості освіти [3]. На даному веб-ресурсі зберігаються дані про результати проведення тестування по кожному предмету по окремим регіонам, містам регіону, а також по окремим навчальним закладам.

Дані за 2010-2015 роки представлені у вигляді програмного забезпечення, що зберігає ці дані і дозволяє імпортувати їх в формат Microsoft Excel (рисунок 1). Дані за 2016-2018 роки представлені у вигляді архівів, що містять в собі інформацію в форматах CSV і Microsoft Excel. При імпорті даних з веб-сайту є можливість отримати дані в розрізі:

- 1) Дані про результати ЗНО за шкалою 100 – 200 балів.
- 2) Дані про результати ЗНО за шкалою 100 – 200 балів (в розрізі профілів навчання).
- 3) Дані про результати ДПА (Державної підсумкової атестації), проведеної в формі ЗНО (за типами навчальних закладів).
- 4) Дані про результати ДПА, проведеної в формі ЗНО (в розрізі профілів навчання).

Складність представляє той факт, що даний відкритий ресурс не містить усереднених значень оцінювання студентів, а лише процентний внесок входжень їх оцінок в діапазони. Крім того, в різні роки межі цих діапазонів також різні. Для того, щоб можна було використовувати

дані за всі попередні роки, необхідно визначити закон розподілу оцінок абітурієнтів на деякій репрезентативній вибірці, а потім поширити ці відомості на всю генеральну сукупність.

Україна	Назва	Район/Місто	Тип	Всього шкіл	% абітурієнтів, що отримали відповідний результат за шкалою 100-200 балів								
					від 100,0 до 123,5	від 124,0 до 135,5	від 136,0 до 150	від 150,5 до 161,5	від 162,0 до 172,5	від 173,0 до 183	від 183,5 до 193	від 193,5 до 195	від 195,5 до 199,5
Львівська область	Львівський район	Львів	середня загальноосвітня школа	47	8,51	12,77	27,66	27,66	21,28	2,13	0,00	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	213	7,04	12,68	26,76	27,84	20,29	9,39	4,23	1,41	0,47
Львівська область	Львівський район	Львів	середня загальноосвітня школа	144	14,58	39,44	34,02	37,36	6,94	5,56	2,08	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	105	3,81	9,52	30,48	25,71	35,29	14,29	0,00	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	182	30,44	11,54	40,11	11,54	14,84	8,79	3,25	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	299	9,36	35,03	29,43	22,07	35,72	5,35	0,67	0,33	0,00
Львівська область	Львівський район	Львів	спеціальна загальноосвітня школа	28	0,00	3,57	23,00	32,14	35,71	0,00	0,00	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	377	30,34	13,26	29,97	22,02	13,79	8,75	1,06	0,80	0,00
Львівська область	Львівський район	Львів	спеціальна загальноосвітня школа	3	0,00	30,00	60,00	30,00	0,00	0,00	0,00	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	145	4,14	7,59	30,34	39,97	11,03	12,41	2,76	0,69	2,07
Львівська область	Львівський район	Львів	середня загальноосвітня школа	110	30,00	7,27	40,00	30,00	15,45	6,36	0,00	0,91	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	266	6,39	9,77	30,45	25,94	38,80	6,01	0,75	1,50	0,38
Львівська область	Львівський район	Львів	середня загальноосвітня школа	192	6,25	11,90	34,48	27,08	35,67	9,90	2,60	1,04	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	79	30,13	11,39	46,84	35,46	12,66	6,53	0,00	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	64	6,25	7,81	29,69	32,81	14,06	6,25	1,56	1,56	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	47	6,38	17,02	27,66	21,28	17,02	4,26	2,26	0,00	2,13
Львівська область	Львівський район	Львів	середня загальноосвітня школа	102	8,82	9,80	33,33	26,47	11,76	6,86	2,94	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	168	7,14	11,31	36,90	33,21	11,31	8,33	1,79	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	196	2,04	4,08	22,45	25,00	39,90	14,80	7,14	3,06	1,02
Львівська область	Львівський район	Львів	середня загальноосвітня школа	29	0,00	0,00	30,34	20,69	34,48	20,69	6,90	6,90	0,00
Львівська область	Львівський район	Львів	спеціальна загальноосвітня школа	18	0,00	0,00	38,89	30,00	11,11	0,00	0,00	0,00	0,00
Львівська область	Львівський район	Львів	спеціальна загальноосвітня школа	42	2,38	2,38	14,29	23,81	11,90	19,05	11,90	7,14	4,76
Львівська область	Львівський район	Львів	середня загальноосвітня школа	104	2,58	4,64	23,77	19,07	22,08	17,53	6,19	3,51	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	30	0,00	0,00	6,67	0,00	23,33	23,33	20,00	36,67	6,67
Львівська область	Львівський район	Львів	середня загальноосвітня школа	160	2,50	3,13	30,13	21,88	23,75	35,25	6,25	3,13	2,75
Львівська область	Львівський район	Львів	середня загальноосвітня школа	78	8,97	11,54	33,33	19,23	17,95	7,69	0,00	1,38	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	30	6,67	6,67	23,33	23,33	35,67	30,00	6,67	2,13	3,13
Львівська область	Львівський район	Львів	середня загальноосвітня школа	25	0,00	3,45	3,45	20,69	31,03	27,59	13,79	0,00	0,00
Львівська область	Львівський район	Львів	середня загальноосвітня школа	229	2,13	7,42	33,78	19,63	23,59	35,36	6,55	2,06	2,13
Львівська область	Львівський район	Львів	спеціальна загальноосвітня школа	5	0,00	0,00	0,00	20,00	20,00	0,00	20,00	20,00	0,00

Рис. 1 – Видгляд подання результатів зовнішнього незалежного оцінювання

Якщо закон розподілу сукупності заздалегідь невідомий, але є підстави припустити, що він має певний вид, то перевіряють основну гіпотезу: сукупність розподілена за певним законом. Перевірка гіпотези про передбачуваний закон невідомого закону розподілу проводиться за допомогою спеціально підібраної випадкової величини – критерію згоди. Критерієм згоди називають критерій перевірки гіпотези про передбачуваний закон невідомого розподілу. У математичній статистиці розглядаються різні критерії згоди (Пірсона, Колмогорова–Смирнова, та ін.). Але найбільшого поширення отримав критерій згоди Пірсона, так як він застосовується для різних типів розподілу, в чому і полягає його основна перевага.

Алгоритм застосування критерію згоди Пірсона χ при перевірці гіпотези про нормальний розподіл випадкової величини наступний: 1) знаходять нульову гіпотезу про нормальний закон розподілу випадкової величини – усередненого значення балів; 2) визначають теоретичні частоти, що відповідають емпіричним частотам; 3) обчислюють спостережуване значення критерію Пірсона; 4) по заданому рівню значущості і числу ступенів свободи знаходять критичне значення критерію Пірсона; 5) якщо спостережуване значення критерію більше ніж критичне, нульова гіпотеза відкидається, інакше приймається. Отримані знання щодо вигляду закону розподілу застосовують на генеральну сукупність по усім досліджуваним параметрам.

Висновки. Було вирішено питання щодо визначення вигляду закону розподілу досліджуваного параметру з використанням критерію Пірсона. Автоматизація методів прогнозування з подальшою розробкою програмних засобів прогнозування передбачає попередню обробку даних з урахуванням вигляду закону розподілу. Після цього можливо застосування методів екстраполяції з отриманням прогностичного значення балів на наступний рік та обчислення точності цього рішення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Komleva N.O. Some aspects of analyzing and diagnosing complex systems // International Multidisciplinary Conference «Science and Technology of the Present Time: Priority Development Directions of Ukraine and Poland» Wolomin, Republic of Poland. V. 3. Wolomin: Izdevnieciba «Baltija Publishing», 2018. P. 106 – 109.
2. Komleva N.O., Cherniha K.S., Tymchenko B.I., Komlevoy O.M. Intellectual approach application for pulmonary diagnosis. *Proceedings of the 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP*, 2016, Article № 7583505, pp. 48-52.
3. Український центр оцінювання якості освіти. URL: <http://testportal.gov.ua/> (дата звернення 26.04.19).