

ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

DOI: <https://doi.org/10.32839/2304-5809/2021-3-91-25>

УДК 519.22

Токовило Т.С.

Херсонська державна морська академія

ОСНОВНІ ПОНЯТТЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ.

ВИКОРИСТАННЯ ЕЛЕМЕНТІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ В СУЧАСНОМУ СВІТІ

Анотація. У статті розглянуто застосування математичної статистики в дослідженні виробничих процесів, їх ефективність, а саме пошук закономірностей величин різних технологічних процесів і прогнозування їх зміни. Показані можливі напрямки науково-дослідної та самостійної роботи студентів при вивченні математичної статистики. Математичні методи є основним механізмом аналізу виробничих процесів, спрямованих на розробку теоретичних моделей, що дасть можливість відображати існуючі зв'язки в житті, прогнозувати поведінку суб'єктів. Математичне моделювання стає мовою сучасного життя, однаково зрозумілим для вчених всіх країн світу. Використання методів математичної статистики в житті надзвичайно широко і різноманітно.

Ключові слова: математична статистика, абсолютна похибка, відносна похибка, вибіркова середня, дисперсія, вибірка.

Tokovilo Tetyana

Kherson State Maritime Academy

BASIC CONCEPTS OF MATHEMATICAL STATISTICS.

THE USE OF ELEMENTS OF MATHEMATICAL STATISTICS IN THE MODERN WORLD

Summary. The statistic has a clear view of the mathematics statistics in the prelude to viral processes, their efficiency, and the very noise of the regularities of the magnitudes of new technological processes and the forecast of these changes. Demonstrations of possible direct scientific-pre-modern and independent robotics of students with the introduction of mathematical statistics. Mathematical methods e the main mechanism for the analysis of viral processes, which are directed to the development of theoretical models, to give the ability to visualize real connections in life, to predict the behavior of subject. Mathematical modeling of the life of a bitter life, however, is sensible for all countries of the world. Victory of methods of mathematical statistics in life is superbly wide and versatile. Mathematical statistics is the science of mathematical methods for analyzing data obtained during mass observations (measurements, experiments). Depending on the mathematical nature of specific observation results, mathematical statistics are divided into statistics of numbers, multivariate statistical analysis, analysis of functions (processes) and time series, statistics of objects of non-numeric nature. An essential part of mathematical statistics is based on probabilistic models. Highlight the general tasks of data description, estimation and hypothesis testing. They also consider more specific tasks associated with conducting sample surveys, restoring dependencies, building and using classifications (typologies), etc. The widespread introduction of statistical research methods was facilitated by the appearance in the second half of the 20th century of electronic computers and, in particular, personal computers. Statistical software packages made these methods more accessible and visual, since the time-consuming work on calculating statistics, parameters, characteristics, constructing tables and graphs was mainly performed by a computer, and the researcher was mainly left with creative work: setting the problem, choosing the solution methods and interpreting the results.

Keywords: mathematical statistics, absolute misbehavior, obvious misbehavior, vibrating average, variance, vibrating.

Постановка проблеми. Актуальність вивчення статистики викликана тим, що статистичні міркування є невід'ємною частиною інтелектуального багажу сучасної людини. Вони потрібні і для повсякденного життя в сучасному цивілізованому суспільстві, і для продовження освіти практично у всіх сферах людської діяльності, наприклад, таких, як соціологія, економіка, право, медицина, демографія. Завдання математичної статистики полягає в створенні методів збору і обробки статистичних даних для отримання наукових і практичних висновків.

Аналіз останніх досліджень та публікацій. На сьогоднішній день існує велика кількість публікацій, яка б охоплювала дану тему. Переглянуті публікації описують використання елементів математичної статистики для розрахунків різних процесів. Тема даної статті розглядалися в працях Габдрахманова К.Ф., Гуторов Ю.А., Ларін П.А., Абдулхамітова Д.Х., Долгополова А.Ф., Волошенко А.Б., Джалладова І.А. та інших.

Виділення не вирішених раніше частин загальної проблеми. Незважаючи на наявність істотних досліджень в сфері математичної статистики, на даний час залишається актуальна та цікава

тема щодо використання елементів математичної статистики для дослідження сучасного світу.

Мета статті. Згадати та виділити основні поняття математичної статистики. У своїй статті наочно продемонструю можливість математичної статистики на прикладі задач в різних сферах життя.

Виклад основного матеріалу. Математична статистика – розділ математики, присвячений математичним методам систематизації, обробки та використання статистичних даних для наукових і практичних висновків. При цьому статистичними даними називаються відомості про кількість об'єктів в будь-якої більш-менш великої сукупності, що володіють тими чи іншими ознаками [2].

Найбільш відомі статистики – відносна частота, вибіркові середні, дисперсія. Коли зростає обсяг вибірки n , багато вибіркові статистики сходяться по ймовірності до відповідних параметрів теоретичного розподілу величини X . Тому кожному вибірку розглядають як вибірку з теоретично нескінченної генеральної сукупності, розподіл ознаки в якій збігається з теоретичним розподілом ймовірності випадкової величини. У багатьох випадках теоретична генеральна сукупність є ідеалізація дійсної сукупності, з якої отримана вибірка [1].

Різні значення спостережуваних ознак, що зустрічається в сукупності, називаються варіантами. Частоти варіантів висловлюють частки елементів сукупності з однаковими значеннями ознаки. Варіаційним рядом називається ранжируваних в порядку зростання або зменшення ряд варіантів з відповідним їм частотами.

Значення, що знаходяться в середині варіаційного ряду, прийнято ділити на власне середні і структурні середні. Власне середнє – це арифметичне середнє. Структурні середні – мода і медіана. Крім того, щоб охарактеризувати структуру варіаційного ряду, використовують квартили, квінти, децили і процентилю [3].

Приклад 1. Знайти середню температуру повітря. У таблиці 1 дані значення середньої температури повітря в місті Херсон в 2020 році.

Таблиця 1

Місяць	°C
Січень	-2,7
Лютий	-4,0
Березень	44,2
Квітень	110,0
Травень	115,3
Червень	220,0
Липень	221,3
Серпень	222,0
Вересень	116,1
Жовтень	99,1
Листопад	33,9
Грудень	-0,8

Розв'язок. Знайдемо середню температуру повітря як середнє значення для негрупованої вибірки:

$$\bar{x} = \frac{1}{12}((-2,7) + (-0,4) + 4,2 + 10,0 + 15,3 + 20,0 + 21,3 + 22,0 + 16,1 + 9,1 + 3,9 + (-0,8)) = 9,83$$

Приклад 2. Частка бракованої продукції в 1 партії виробів склала 1%, у 2 партії – 1,5%, а в третій – 2%. Перша партія становить 35% всієї продукції, друга – 40%. Визначити середній відсоток бракованої продукції.

Розв'язок. За даними задачі складемо таблицю 2:

Таблиця 2

№ партії	Частка бракованої продукції, % (x)	Питома вага кожної партії в загальному обсязі продукції, (d)
1	1	0,35
2	1,5	0,40
3	2	0,25
Всього	–	1

Середній відсоток бракованої продукції визначимо за формулою середньої арифметичної зваженої:

$$\bar{x} = \sum xd = 1 \cdot 0,35 + 1,5 \cdot 0,40 + 2 \cdot 0,25 = 1,45\%$$

Приклад 3. У таблиці 3 – дані про угруповання сільських господарств за врожайністю зернових. Знайти середню врожайність зернових.

Таблиця 3

Урожайність зернових в центнерах з га	Число сільських господарств – абсолютна	Питома вага сільських господарств – у відсотках
до 5,0	4244	6,2
5,1-10,0	10446	15,2
10,1-15,0	18956	27,5
15,1-20,0	20207	29,3
20,1-25,0	8159	11,9
25,1-30,0	4145	6,0
30,1-35,0	1316	1,9
35,1-40,0	792	1,2
40,1-45,0	183	0,3
45,1-50,0	182	0,3
50,1-55,0	161	0,2
Всього	68791	100,0

Розв'язок. Маємо тільки груповані дані і невідома середня врожайність кожної групи, як наближені значення до середньої кожної групи приймемо центри інтервалів (таблиця 4):

Таблиця 4

Центри інтервалів x_i	n_i	$x_i \cdot n_i$
2,5	4222	10610,0
7,5	10446	78345,0
12,5	18956	236950,0
17,5	20207	363622,5
22,5	8159	183577,5
27,5	4145	113987,5
32,5	1316	42770,0
37,5	792	29700,0
42,5	183	7777,5
47,5	182	8645,0
52,5	161	8452,5
Всього	68791	1074437,5

Знайдемо необхідну в умові завдання середню врожайність зернових:

$$\bar{x} = \frac{1}{68791} (2,5 \cdot 4244 + 7,5 \cdot 10446 + 47,5 \cdot 182 + 52,5 \cdot 161) = 15,6$$

Отже, середня врожайність по вибірці становить 15,6 центнерів з га.

Приклад 4. Середній вік жителів одного з регіонів 30 років. При цьому середній вік сільських жителів, які становлять 60% всіх жителів, 32 роки при 7 років, а міських жителів 27 років при 8 років. Визначте загальну дисперсію віку жителів регіону.

Розв'язок.

За умовою,

$\bar{x}_0 = 30$ років – середній вік жителів одного з регіонів;

$\bar{x}_c = 32$ роки – середній вік сільських жителів;

$d_c = 0,6$ – питома вага сільських жителів в загальній кількості жителів регіону;

$\sigma_c = 7$ років – середньоквадратичне відхилення (для сільських жителів);

$\bar{x}_m = 27$ років – середній вік міських жителів;

$\sigma_m = 8$ років – середньоквадратичне відхилення (для міських жителів).

Знайдемо питому вагу міських жителів у загальній кількості жителів регіону:

$$d_m = 1 - d_c = 1 - 0,6 = 0,4$$

Розрахуємо загальну дисперсію $\sigma_{\text{заг}}^2$ за правилами додавання дисперсій:

$$\sigma_{\text{заг}}^2 = \overline{\sigma_{\text{заг}}^2} + \sigma_{\text{гр.сп}}^2$$

де $\overline{\sigma_{\text{заг}}^2}$ – середня із внутрішньо групових дисперсій;

$$\overline{\sigma_{\text{заг}}^2} = \sum \sigma_i^2 d_i = 7^2 \cdot 0,6 + 8^2 \cdot 0,4 = 55;$$

$\sigma_{\text{гр.сп}}^2$ – дисперсія групових середніх;

$$\sigma_{\text{гр.сп}}^2 = \sum (\bar{x}_i - \bar{x}_0)^2 d_i = (32 - 30)^2 \cdot 0,6 + (27 - 30)^2 \cdot 0,4 = 6.$$

Звідси, загальна дисперсія буде:

$$\sigma_{\text{заг}}^2 = \overline{\sigma_{\text{заг}}^2} + \sigma_{\text{гр.сп}}^2 = 55 + 6 = 61.$$

Приклад 5. Скільки з 1000 жителів району необхідно обстежити в порядку випадкової вибірки (без повторювань) для визначення середнього віку, щоб з ймовірністю 0,954 ($t = 2$), можна було гарантувати помилку не більше 5 років. Передбачуване середньоквадратичне відхилення 20 років.

Розв'язок.

Для без повторювань відбору чисельність вибірки (при визначенні середнього віку) знайдемо за формулою:

$$n = \frac{t^2 \cdot \sigma^2 \cdot N}{\Delta^2 \cdot N + t^2 \cdot \sigma^2}.$$

За умовою,

$N = 1000$ (чол.) – загальне число жителів регіону (чисельність генеральної сукупності);

$\Delta = 5$ (років) – гранична помилка вибірки;

$t = 2$ – коефіцієнт кратності помилки (так як ймовірність $P = 0,954$);

$\sigma = 20$ (років) – вибіркове середнє квадратичне відхилення.

$$n = \frac{2^2 \cdot 20^2 \cdot 1000}{5^2 \cdot 1000 + 2^2 \cdot 20^2} \approx 60 \text{ (чол.)}$$

Приклад 6. На економічному факультеті вибірково методом (відбір повторний) був визначений середній вік студента. Виявилось, що він дорівнює 21,5 року при середньому квадратичному відхиленні 4 роки. Скільки треба обстежити студентів, щоб помилка при визначенні середнього віку не перевищила 1 рік з ймовірністю 0,997 ($t = 3$).

Розв'язок.

$\bar{x}_e = 21,5$ (років) – середній вік студента в вибірковій сукупності;

$\sigma_x = 4$ (рока) – вибіркове середнє квадратичне відхилення;

$t = 3$ – коефіцієнт кратності помилки (так як ймовірність $P = 0,997$);

$\Delta_x = 1$ (рік) – гранична помилка при визначенні середнього віку.

Для повторного відбору чисельність вибірки (при визначенні середнього розміру помилки ознаки) знаходиться за формулою:

$$n = \frac{t^2 \cdot \sigma^2}{\Delta^2};$$

$$n = \frac{3^2 \cdot 4^2}{1^2} = 9 \cdot 16 = 144 \text{ (чол.)}$$

Приклад 7. Планувалося знизити трудомісткість виготовлення продукції «А» на 3,6%, фактично вона була знижена на 5%. Визначте виконання плану по зниженню трудомісткості.

Розв'язок.

Відносна величина виконання плану являє собою відношення фактично досягнутого в даному періоді рівня до запланованого:

$$i_{\text{вик.пл.}} = \frac{105,0}{103,6} \cdot 100\% = 101,4\%,$$

тобто план по зниженню трудомісткості перевищено на 1,4% (101,4 – 100).

Висновки і пропозиції. У сучасній науці вважається, що будь-яка область досліджень не може бути справжньою наукою до тих пір, поки в неї не проникне математика. У цьому сенсі математична статистика є повноважним представником математики в будь-який інший науці і забезпечує науковий підхід до досліджень. Можна сказати, що науковий підхід починається там, де в дослідженні з'являється математична статистика. Статистика з необхідністю з'являється там, де відбувається перехід від одичного спостереження до множинного. Якщо у вас є безліч спостережень, вимірів і даних – то без математичної статистики вам не обійтися [1].

Розглянувши використання методів математичної статистики при завдань в різних сферах життя можна сказати наступне, що в сучасному світі математична статистика застосовна з тієї причини, що отримані дані часто є статистичні відомості, тобто відомості про однорідні сукупності об'єктів і явищ. Такими однорідними сукупностями можуть бути випущені промисловістю вироби, персонал промисловості, дані про прибутки підприємств, видача кредитів у банку, температурні зміни повітря та багато іншого.

Список літератури:

1. Использование методов математической статистики при решении банковских задач. Москва, 2014. URL: <http://www.top-technologies.ru/ru/article/view?id=34025> (дата звернення: 20.03.2021).
2. Габдрахманова К.Ф., Гуторов Ю.А., Ларин П.А. Теория вероятностей и математическая статистика в примерах и задачах по разработке нефтяных месторождений. Уфа, 2013. 134 с.
3. Волошенко А.Б., Джалладова І.А. Теорія ймовірностей та математична статистика : навч. метод. посібник для самостійного вивчення дисципліни. Київ, 2003. 256 с.

References:

1. Abdulhamitova D., Dolgoplova A. (2014) *Ispolzovanie metodov matematicheskoy statistiki pri reshenii bankovskih zadach* [Using the methods of mathematical statistics in solving banking problems (electronic journal)]. Available at: <http://www.top-technologies.ru/ru/article/view?id=34025> (accessed 26 January 2020).
2. Gabdrahmanova K., Gutorov Yu., Larin P. (2013) *Teoriya veroyatnostey i matematicheskaya statistika v primerah i zadachah po razrabotke neftyanyih mestorozhdeni* [Probability theory and mathematical statistics in examples and tasks for the development of oil fields]. Ufa: USPTU. (in. Russian)
3. Voloshenko A., Dzhalladova I. (2003) *Teoriya jmovirnostej ta matematichna statistika: navch. metod. posibnik dlya samostijnogo vivchennya disziplini* [Probability Theory and Mathematical Statistics: teaching. method. a manual for independent study of discipline]. Kiev: KNEU. (in Ukrainian)