Світлана Панова, *старший лаборант кафедри менеджменту та правового забезпечення господарської діяльності Бердянського державного педагогічного університету Запорізька обл.*

ДЕЯКІ ТЕОРЕТИЧНІ АСПЕКТИ ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ У ПЕРЕВІРЦІ НАУКОВИХ ГІПОТЕЗ В ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ

У статті представлене сучасне наукове бачення проблеми застосування методів математичної статистики у перевірці наукових гіпотез в економічних дослідженнях. Викладено деякі теоретичні аспекти досліджуваної проблеми.

Постановка проблеми. У дослідженнях різноманітних економічних процесів, явищ, стану ринків та інших економічних аспектів нашого життя велика кількість іноземних та вітчизняних фахівців відзначають, що використовуючи математичну статистику, багато результатів можна отримати і обґрунтувати теоретично, на базі існуючого емпіричного матеріалу. Застосування математичних методів в економіці дає змогу відокремити та формально описати найважливіші та найсуттєвіші зв'язки економічних змінних процесів і об'єктів, а також індуктивним шляхом отримати нові знання про об'єкт. Крім того, математичною мовою можна описати твердження економічної теорії, формулювати її поняття та висновки.

Методи математичної статистики використовують в економічних, технічних дослідженнях, у менеджменті, маркетингу, медицині та інших наукових сферах.

Математичну статистику можна визначити як галузь математичних знань, присвячених вивченню закономірностей, які мають місце в масових явищах і статистичних закономірностях. Вона розробляє раціональні прийоми (способи) систематизації, обробки і аналізу даних статистичних спостережень масових явищ з метою встановлення характерних для них статистичних закономірностей, використання для обґрунтування результатів наукових і практичних досліджень. Більшість методів обробки статистичних даних ґрунтується на імовірнісній природі цих даних. Галузь застосування таких статистичних методів обмежена з метою підпорядкованості явищ, які досліджуються, достатньо визначеним імовірнісним закономірностям. Математична статистика абстрагується від матеріального змісту масових явищ, які вона характеризує, озброює дослідника математичним апаратом.

Її основними розділами є опис даних, оцінка та перевірка гіпотез. Усі вони тісно пов'язані між собою. У практичній та науковій діяльності часто доводять справедливість будь-якого факту за допомогою висловлювання гіпотез, які можуть бути перевірені на основі даних вибіркового спостереження.

Аналіз останніх досліджень та публікацій. Проблема застосування методів математичної статистики висвітлена у наукових дослідженнях багатьох зарубіжних і вітчизняних фахівців. Так, І.А. Орлов, Д.А. Новіков, Б.В. Гнеденко, А.Л. Аліфанов та інші присвячували наукові роботи основним методам сучасної прикладної математичної статистики для фахівців різних галузей науки.

Метою даної статті є аналіз теоретичного змісту методів математичної статистики при перевірці наукових гіпотез в економічних дослідженнях.

Виклад основного матеріалу. Статистичною гіпотезою (H) називається припущення щодо параметрів або форми розподілу генеральної сукупності, яке перевіряється на основі даних вибіркового спостереження [6, 106].

Статистична гіпотеза – це деяке припущення про закон розподілу випадкової величини або про параметри цього закону, що формулюється на основі вибірки. Прикладами статистичних гіпотез є припущення: генеральна сукупність розподілена за експонентним законом; математичні сподівання двох експоненціально розподілених вибірок рівні один одному. У першій з них висловлене припущення про вид закону розподілу, а в другій – про параметри двох розподілів. Гіпотези, в основі яких немає ніяких припущень про конкретний вид закону розподілу, називають *непараметричними*, у противному випадку – *параметричними*.

Гіпотезу, яка затверджує, що розходження між порівнюваними характеристиками відсутнє, а відхилення, що спостерігаються, зумовлюються лише випадковими коливаннями у вибірках, на підставі яких робиться порівняння, *називають* нульовою (основною) гіпотезою і *позначають* H_0. Поряд з основною гіпотезою розглядають і *альтернативну* (конкуруючу, суперечну) їй гіпотезу H_1. І якщо нульова гіпотеза буде відкинута, то буде мати місце альтернативна гіпотеза.

Розрізняють прості і складні гіпотези. Гіпотезу називають *простою*, якщо вона однозначно характеризує параметр розподілу випадкової величини. Наприклад, якщо λ є параметром експонентного розподілу, то гіпотеза H_0 про рівність λ =10 – проста гіпотеза. *Складною* називають гіпотезу, що складається з кінцевої або нескінченної безлічі простих гіпотез. Складна гіпотеза H_0 про нерівність λ >10 складається з нескінченної безлічі простих гіпотез H_0 про рівність $\lambda = b_i$, де b_i – будь-яке число, більше 10. Гіпотеза H_0 про те, що математичне сподівання нормального розподілу дорівнює двом при невідомій дисперсії, теж є складною. Складною гіпотезою буде припущення про розподіл випадкової величини X за нормальним законом, якщо не фіксуються конкретні значення математичного сподівання і дисперсії [2, 3, 9].

Перевірка гіпотези ґрунтується на обчисленні деякої випадкової величини – критерію, точний або наближений розподіл якого відомий. Позначимо цю величину через z, її значення є функцією від елементів вибірки $z = z(x_1, x_2, \ldots, x_n)$... Процедура перевірки гіпотези пропонує кожному значенню критерію одне з двох рішень – прийняти або відкинути гіпотезу. Тим самим увесь вибірковий простір і відповідно нескінченна множина значень критерію поділяються на дві непересічні підмножини S_0 і S_1. Якщо значення критерію z потрапляє в область S_0, то гіпотеза приймається, а якщо в область S_1, – гіпотеза відхиляється. Нескінченна множина S_0 називається *областю прийняття гіпотези або областю припустимих значень*, а Нескінченна множина S_1 – *областю відхилення гіпотези або критичною областю*. Вибір однієї області однозначно визначає й іншу область.

Прийняття або відхилення гіпотези H_0 за випадковою вибіркою відповідає істині з деякою імовірністю і, відповідно, можливі два роди помилок. Помилка першого роду виникає з імовірністю, а тоді, коли відкидається правильна гіпотеза H_0 і приймається конкуруюча гіпотеза H_1. Помилка другого роду виникає з імовірністю b у тому випадку, коли приймається неправильна гіпотеза H_0, у той час як справедлива конкуруюча гіпотеза H_1. *Довірча імовірність* – це імовірність не зробити помилку першого роду і прийняти правильну гіпотезу H_0. Імовірність відкинути помилкову гіпотезу H_0 називається *потужністю критерію*. Отже, при перевірці гіпотези можливі чотири варіанти результатів, табл. 1.

Таблиця 1

Можливі результати при перевірці гіпотези

Гіпотеза H_0	Рішення	Імовірність	Примітка
Правильна	Приймається	$1-\alpha$	Довірча імовірність
	Відхиляється	α	Імовірність помилки першого роду
Неправильна	Приймається	β	Імовірність помилки другого роду
	Відхиляється	$1-\beta$	Потужність критерію

Залежно від сутності гіпотези, що перевіряється, і використовуваних мір розбіжності оцінки характеристики від її теоретичного значення, застосовують різні критерії. До числа найбільше часто застосовуваних критеріїв для перевірки гіпотез про закони розподілу відносять критерії χ^2 – Пірсона, Колмогорова, Мизеса, Вилкоксона, про значення параметрів – критерії Фішера, Стьюдента.

Розглянемо теоретичні аспекти застосування деяких критеріїв перевірки гіпотез.

За допомогою нормального розподілу визначаються три розподіли, що у даний час часто використовуються при статистичній обробці даних [7, 79-80].

Розподіл Пірсона c^2 (хі - квадрат) – розподіл випадкової величини

$$X = X_1^2 + X_2^2 + \ldots + X_n^2, \qquad (1)$$

де випадкові величини $X_1, X_2, \ldots, X_n$ незалежні і мають один і той же розподіл $N(0,1)$. При цьому числа що складаються, тобто n, називаються "числом ступенів свободи" розподілу хі-квадрат.

Розподіл хі-квадрат використовують при оцінюванні дисперсії (за допомогою довірчого інтервалу), при перевірці гіпотез згоди, однорідності, незалежності, насамперед для якісних перемінних, приймаючи кінцеве число значень, і в багатьох інших задачах статистичного аналізу даних [1, 4, 5, 10].

Розподіл t Стьюдента – це розподіл випадкової величини

$$T = \frac{U\sqrt{n}}{\sqrt{X}}, \tag{2}$$

де випадкові величини U і X незалежні, U має розподіл стандартний нормальному розподілу $N(0,1)$, а X – розподіл χ^2 з n ступенями свободи. При цьому n називається "числом ступенів свободи" розподілу Стьюдента.

Розподіл Стьюдента було введено в 1908 р. англійським статистиком В. Госсетом, що працював на фабриці, яка випускала пиво. Ймовірнісно-статистичні методи використовувалися для прийняття економічних і технічних рішень на цій фабриці, тому її керівництво забороняло В. Госсету публікувати наукові статті під своїм ім'ям. Таким способом охоронялася комерційна таємниця, "ноу-хау" у вигляді ймовірнісно-статистичних методів, розроблених В. Госсетом. Однак він мав можливість публікуватися під псевдонімом "Стьюдент". Історія Госсета – Стьюдента показує, що ще сто років тому менеджерам Великобританії була очевидна велика економічна ефективність ймовірносно-статистичних методів [7, 80].

У даний час розподіл Стьюдента – один з найбільш відомих розподілів серед тих, що використовуються при аналізі реальних даних. Його застосовують при оцінюванні математичного сподівань, прогнозного значення й інших характеристик за допомогою довірчих інтервалів, при перевірці гіпотез про значення математичних сподівань, коефіцієнтів регресійної залежності, гіпотез однорідності вибірок і тощо. [1, 4, 5, 10].

Розподіл Фішера – це розподіл випадкової величини

$$F = \frac{\frac{1}{k_1}X_1}{\frac{1}{k_2}X_{21}}, \tag{3}$$

де випадкові величини X_1 і X_2 незалежні і мають розподіл хі – квадрат з числом ступенів своюоди k_1 і k_2 відповідно. При цьому пари (k_1, k_2) – пари "чисел ступенів своюоди" розподілу Фішера, а саме, k_1 – число ступенів свободи чисельника, а k_2 – число ступенів свободи знаменника. Розподіл випадкової величини F названо на честь великого англійського статистика Р. Фішера (1890 – 1962), що активно використовував його у своїх роботах.

Розподіл Фішера застосовують при перевірці гіпотез про адекватність моделі в регресійному аналізі, про рівність дисперсій і в інших задачах прикладної статистики [1, 4, 5, 10,].

Вираження для функцій розподілу χ^2 – Пірсонса, Стьюдента і Фішера, їх характеристик, а також таблиці, необхідні для їхнього практичного використання, можна знайти в спеціальній літературі (наприклад [1]).

Висновки. Отже, у статті ми виклали теоретичні аспекти застосування методів математичної статистики у перевірці гіпотез, але обсяг даної статі не дозволяє розглянути усю кількість сучасних статистичних критеріїв, тому ми зосередили увагу лише на деяких з них.

1. *Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.). – 150 с.*

2. *Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1999. – 576 с.*

3. *Елисеева И.И., Юзбашев М.М. Общая теория статистики. – М.: Финансы и статистика, 1999. – 640 с.*

4. *Кендалл М.Дж., Стьюарт А. Статистические выводы и связи. – М.: Наука, 1973. – 900 с.*

5. *Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.*

6. *Мармоза А.Т. Практикум з математичної статистики: Навч. посіб. – К: Кондор, 2004. – 264 с.*

7. *Орлов А.И. Математика случая. Вероятность и статистика – основные факты. Учебное пособие. М.: МЗ-Пресс, 2004. – 110 с.*

8. *Орлов А.И. Прикладная статистика. Учебник. / А.И.Орлов. – М.: Издательство "Экзамен", 2004. – 656 с.*

9. *Рябинин И.А. Надежность и безопасность структурно-сложных систем. – СПб.: Политехника, 2000. – 785 с.*

10. *Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 512 с.*