

А.І. Пилипенко

ПЕРСПЕКТИВИ ЗАСТОСУВАННЯ МЕТОДІВ АНАЛІТИЧНОЇ СТАТИСТИКИ В УПРАВЛІННІ ПРОЕКТАМИ ТА ПРОГРАМАМИ

Виявлено перспективи застосування непараметричної статистики, статистики нечислових даних і статистики інтервальних даних як оптимальних інструментів побудови формальних моделей в управлінні проектами. Табл.1, дж 15.

Ключові слова: управління проектами, регресійний аналіз, непараметрична статистика, статистика нечислових даних, статистика інтервальних даних.

JEL O22

ВСТУП

Постановка проблеми у загальному вигляді. Немає необхідності говорити про важливість формальної постановки задачі дослідження, оптимального планування цього дослідження, грамотної статистичної обробки результатів у будь-якій предметній області. Поєднання розуміння задачі дослідження та приведення її у відповідний математико-статистичний запис – це функція аналітичного статистика. Автор оригінальних методів статистичного оцінювання та перевірки гіпотез Цейтлін Н.А. пише: «Аналітичний статистик повинен мати по суті дві освіти – математика-статистика і предметне» [1].

Розглянемо особливості застосування аналітичної статистики у такій предметній області як управління проектами і програмами. На сучасному етапі розвитку методології управління проектами відбувається розширення спектру математичних методів і підходів, які використовуються на всіх етапах життєвого циклу проекту [2]. Застосування статистичних методів найбільш поширене в управлінні якістю та управлінні ризиками проекту. Зокрема, можна навести такі статистичні методи [3]: *аналіз очікуваного грошового значення* (EMV), зазвичай цей метод використовується в рамках аналізу дерева рішень для оцінки наслідків ланцюга можливих варіантів в умовах невизначеності; *планування експериментів* (DOE) для визначення значущих чинників, які впливають на певні параметри продукту чи процесу, що знаходиться на стадії розробки або виробництва; *регресійний аналіз* (Regression Analysis), метою якого є створення математичної або статистичної залежності.

Статистичні методи постійно розвиваються, сучасний математичний апарат глибокий і майже завжди складний, розроблено потужні комп'ютерні програми. Тому важливо розуміти як вони можуть бути застосовані і яку користь привнесуть в управління проектами та програмами.

Аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми і виділення невирішених раніше частин. Значну увагу розвитку і становленню аналітичної статистики приділили у своїх роботах проф. Бараник З.П. [4] і проф. Орлов О.І. [5]. У вітчизняних фахових виданнях («Статистика України», «Прикладна статистика: проблеми теорії та практики», «Моделювання та інформаційні системи в економіці», «Теорія ймовірностей і математична статистика», «Український математичний журнал» та інші) приділяється увага розробці програмно-методологічних та організаційних засад статистичного спостереження; формуванню інтегрованих інформаційних баз даних за результатами статистичної звітності; статистичному вимірюванню взаємозв'язків економічних, соціальних та демографічних явищ і процесів;

статистичному оцінюванню стану і розвитку ринкової економіки; статистичній оцінці ризику економічної діяльності, тощо. Але, на жаль, публікації не містять рекомендацій зі застосування статистичних методів в управлінні проектами. Так само не приділяється достатньо уваги сучасним методам аналітичної статистики і у фахових виданнях з управління проектами.

Мета статті є виявлення перспективних методів аналітичної статистики, які можуть стати кількісною основою в управлінні проектами та програмами.

ОСНОВНІ РЕЗУЛЬТАТИ

Методи та методику дослідження. Методологічною основою роботи є сукупність наукових знань, методів і технологій, що використовуються в процесі дослідження, такі як: порівняння, аналіз, синтез, узагальнення, конкретизація.

Основна частина дослідження. Коло проблем аналітичної статистики обмежується, переважно, рішенням п'яти задач [1]:

формалізація задачі експериментального дослідження;

оптимальне планування експерименту (при дослідженні керованих об'єктів) або розробка раціональної тактики збору статистичної інформації (при «небезпечному» спостереженні некерованих об'єктів);

коригування плану експериментів в процесі проведення експериментальних робіт або тактика обстеження;

статистична обробка результатів спостережень із залученням (при необхідності в складних випадках) «чистих» математиків і програмістів;

інтерпретація результатів розрахунків в термінах вихідної предметної області у складанні звіту про проведену роботу.

На практиці можуть зустрічатися прості і складні статистичні задачі. Простими вважаються задачі точечного й інтервального оцінювання параметрів однієї випадкової величини, перевірки гіпотез щодо цих параметрів, а також задачі кореляційного і регресійного аналізів зв'язків між двома змінними. Такі задачі в управлінні проектами зустрічаються досить часто. Наприклад, побудова інтервальної оцінки ризиків в банківських проектах [6], застосування кореляційного аналізу в освітніх проектах для підтвердження зв'язку між вищою і професійною освітою та інноваціями [7], застосування регресійних моделей у проектах розвитку морських портів для прогнозування вантажообігу [8] та інші.

До складних задач аналітичної статистики відносяться задачі багатофакторного регресійного аналізу. Цей метод так само знайшов широке застосування в управлінні проектами. Так, в роботі [9] за допомогою багатофакторного регресійного аналізу проведено моделювання процесів бюджетування в проектах будівництва складних енергетичних об'єктів.

Сучасні методи моделювання стають настільки досконалими, а рівень знань настільки високим, що складна математична модель робиться неосяжно великою і громіздкою. Статистичний аналіз складної математичної моделі може показати, що прості математичні моделі у складі складної моделі практично не працюють. Наявність непрацюючих моделей створює непотрібну ілюзію «повного знання», хоча насправді ні в складній моделі, ні в житті непрацюючі прості моделі значущої ролі не грають. Задача редукції (скорочення) складної математичної моделі вирішується за допомогою елімінування [1]. Цим елімінування принципово відрізняється від регресійного аналізу. В управлінні проектами елімінування може бути застосовано в управлінні ризиками. Прикладом елімінування факторів ризику, які негативно впливають на вартість не фінансових компаній, є робота проф. Швеця С.К. [9].

Перспективними, такими що постійно розвиваються, є непараметричні методи аналітичної статистики. Непараметрична статистика використовується у випадках, коли закони розподілу, як правило, або невідомі, або для їх

встановлення потрібна велика обчислювальна робота. Популярність непараметричних методів пояснюється широкою областю їх застосування, стійкістю висновків, простотою математичних засобів. Непараметричні методи в даний час складають сформовану систему обробки даних, яку за своїми можливостями можна порівняти з гаусівською [10]. Ці методи можуть бути застосовані в управлінні термінами проекту, наприклад для оцінки параметрів часу затримок виконання розкладу проекту.

За оцінками багатьох учених центральною областю прикладної статистики у XXI ст., стане статистика нечислових даних, або її ще іменують нечисловою статистикою, оскільки містить найзагальніші підходи і результати [4]. Основою її математичного апарату є використання відстаней між об'єктами нечислової природи і рішень оптимізаційних задач, а не операцій підсумовування даних, як в інших областях статистики [4]. У загальному випадку під нечисловими даними розуміють елементи просторів, які не є лінійними (векторними), в яких немає операцій додавання елементів і їх множення на дійсне число. Крім результатів вимірювань за якісними ознаками, прикладами є послідовності з 0 і 1, бінарні відносини (ранжирування, розбиття, толерантності); безлічі (в тому числі плоскі зображення та об'ємні тіла); нечіткі (розмиті, розпливчасті, fuzzy) числа й множини, їх окремих випадок — інтервали; результати парних порівнянь та інші об'єкти, що виникають в прикладних дослідженнях.

У роботі [11] Орлов О.І. розглядає приклад порівняння восьми проектів, запропонованих для включення в план стратегічного розвитку фірми. Оскільки відповіді експертів у багатьох процедурах експертного опитування не числа, а такі об'єкти нечислової природи, як градації якісних ознак, ранжирування, розбиття, результати парних порівнянь, нечіткі переваги і т.д., то для їх аналізу виявляються корисними методи статистики нечислових даних. Відповіді експертів часто носять нечисловий характер, тому що люди не мислять числами. У мисленні людини використовуються образи, слова, але не числа. Відповіді експерта зазвичай виміряні шкалою порядку, або є ранжировками, результатами парних порівнянь та іншими об'єктами нечислової природи. Професор Орлов О.І. вважає поширеною помилкою те, що відповіді експертів намагаються розглядати як числа, займаються «оцифровуванням» їх думок, приписуючи цим думкам чисельні значення — бали, які потім обробляють за допомогою різних методів прикладної статистики як результати звичайних фізико-технічних вимірювань. У разі довільної «оцифровки» висновки можуть не мати відношення до реальності. Подамо базову інформацію нової парадигми аналізу статистичних та експертних даних, яка запозичена з роботи [5], в табличній формі (табл. 1).

Таблиця 1

Порівняння основних характеристик традиційного та нечислового статистичного аналізу

Характеристика	Стара парадигма	Нова парадигма
Типові вихідні дані	Числа, вектора, функції	Об'єкти нечислової природи
Основний підхід до моделювання даних	Розподіл з параметричних сімейств	Довільні функції розподілу

Основний математичний апарат	Суми і функції від сум	Відстані і алгоритми оптимізації
Джерела постановок нових задач	Традиції, що сформувалися до середини ХХ століття	Сучасні прикладні потреби аналізу даних (XXI століття)
Ставлення до питань стійкості висновків	Практично відсутній інтерес до стійкості висновків	Розвинута теорія стійкості (робастності) висновків
Оцінювані величини	Параметри розподілів	Характеристики, функції і щільності розподілів, залежності, правила діагностики та ін.
Можливість застосування	Наявність повторюваного комплексу умов	Наявність обґрунтованої ймовірносно-статистичної моделі
Центральна частина теорії	Статистика числових випадкових величин	Нечислова статистика
Роль інформаційних технологій	Тільки для розрахунку таблиць (інформатика знаходиться поза статистики)	Інструменти отримання висновків (датчики псевдовипадкових чисел, розмноження вибірок, та ін.)
Точність даних	Дані повністю відомі	Облік невизначеності даних, зокрема, інтервальності і нечіткості
Типові результати	Граничні теореми (при зростанні обсягів вибірок)	Рекомендації для конкретних обсягів вибірок
Вид постановок задач	Окремі задачі оцінювання параметрів і перевірки гіпотез	Високі статистичні технології (технологічні процеси аналізу даних)
Стиковка алгоритмів	Не розглядається	Вельми важлива при розробці процесів аналізу даних
Роль моделювання	Мала (окремі системи аксіом)	Системи моделей - основа аналізу даних
Аналіз експертних оцінок	Окремі алгоритми	Прикладне «дзеркало» загальної теорії
Роль методології	Практично відсутня	Основоположна

Окремим видом об'єктів нечислової природи є нечіткі множини. Тому при обробці вибірки, елементами якої є нечіткі множини, можуть бути використані різні методи аналізу статистичних даних довільної природи – розрахунок середніх, непараметричних оцінок щільності, побудова діагностичних правил і т.д. Нечіткі множини можуть бути використані для опису якісних характеристик при порівнянні варіантів реалізації інноваційного проекту і вибору кращого з них [12].

Останні десятиліття швидко розвивається нова область статистичних досліджень — математична статистика інтервальних даних. Відбувається розвиток методів прикладної математичної статистики в ситуації, коли статистичні дані — не числа, а інтервали, зокрема, породжені накладенням помилок вимірювання на значення випадкових величин. У цьому розділі прикладної статистики розроблено принципово нові (порівняно з класичною математичною статистикою) підходи, засновані на поняттях нотни та раціонального обсягу вибірки. Наведемо ці визначення [13].

Вводяться позначення $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$, $\varepsilon = (\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n)$. Статистичні висновки ґрунтуються на статистиці $f: R_n \rightarrow R_1$, яку використовують для оцінювання параметрів і характеристик розподілу, перевірки гіпотез та вирішення інших статистичних задач. Принципово важлива для статистики інтервальних даних така ідея: статистик знає тільки $f(y)$, але не $f(x)$.

Величину максимально можливого (по абсолютній величині) відхилення, викликаного похибками спостережень ε , відомого статистику значення $f(y)$ від істинного значення $f(x)$, тобто

$$N_{f(x)} = \sup |f(y) - f(x)|,$$

де супремум береться по множині можливих значень вектора похибок ε і називається *нотною*.

Доведено, що на відміну від класичної математичної статистики недоцільно безмежно збільшувати обсяг вибірки, оскільки середній квадрат помилки залишається завжди більшим квадрата нотни. Тому ввели поняття *раціонального обсягу вибірки* N_{rat} , при досягненні якого продовжувати спостереження недоцільно.

Методи статистики інтервальних даних можуть бути корисними не тільки в традиційних технічних та економетричних задачах, але і в управлінні проектами. Наприклад, при ухваленні рішення про вигідність або невигідність інвестиційного проекту. Припускають, що більш правильно було б вважати банківський відсоток q і дисконт C нечисловими величинами, а саме, інтервалами $[q_1; q_2]$ й $[C_1; C_2]$. Отже, економічні висновки повинні бути досліджені на стійкість по відношенню до можливих відхилень.

Як було сказано вище, застосування статистичних методів активно застосовується в управлінні ризиками проекту. Розглянемо це більш детально. Одним з найбільш широко застосовуваних методів кількісного аналізу ризиків є імітація. Імітації, як правило, проводяться за допомогою методу Монте-Карло. При імітації модель проекту розраховується безліч разів (ітеративно), при цьому для кожної ітерації вхідні значення вибираються довільно з розподілів ймовірностей цих змінних. При аналізі ризиків вартості методом імітації використовуються оцінки вартості. При аналізі ризиків розкладу використовуються діаграма мережі розкладу та оцінки тривалості [3]. Детальний приклад застосування методу Монте-Карло розглянуто в статті [14].

Воцинін О.П. запропонував новий підхід до планування обчислювального експерименту для оцінки впливу змінних на результат рішення задачі. В роботі [15] наведена методологія оптимального обчислювального експерименту і

показано його перевагу перед методом Монте-Карло як за кількістю дослідів, так і за точністю аналітичної моделі.

Розроблена методологія може бути використана для оцінки ефективності інвестиційних проектів. В рамках інтервальної парадигми під ризиком для обраного критерію ефективності розуміється можливість отримання негативного результату, оцінювана числом r , $0 \leq r \leq 1$. Поняття можливості аналогічно поняттю ймовірності, але не спирається на гіпотезу про випадковість і не припускає встановлення щільності ймовірності на інтервалі невизначеності. Наприклад, при заданому інтервалі невизначеності критерію ефективності $[y_{\min}, y_{\max}]$ ризик, відповідно до методу, описаного в статті [15], оцінюється за формулою

$$r_{y < C} = \begin{cases} 0, & y_{\min} > C, \\ (C - y_{\min}) / (y_{\max} - y_{\min}), & y_{\min} < C < y_{\max}, \\ 1, & y_{\max} < C; \end{cases} \quad (1)$$

де C – гранично допустиме значення ризику.

Якщо на інтервалі $[y_{\min}, y_{\max}]$ задати рівномірний розподіл випадкової величини і розрахувати ймовірність того, що вона буде менше константи C , то результат співпадає з формулою (1).

При наявності аналітичної моделі оцінка ризиків можлива також у припущенні нормального розподілу змінних всередині заданих діапазонів. При цьому випадкові незалежні величини x_i описуються математичним очікуванням μ_i і дисперсією σ_i^2 . Критерій ефективності y , як лінійна комбінація x_i , також має нормальний розподіл $\varphi(y)$ з математичним очікуванням $\mu_y = b_0 + \sum \mu_i$ і дисперсією $\sigma_y^2 = \sum b_i^2 \sigma_i^2$ ($i = 1, \dots, m$).

Ймовірність (тобто ризик r) того, що критерій y буде менше константи C , знаходиться за відомою формулою теорії ймовірностей:

$$r = P(y < C) = \int_{-\infty}^C \varphi(y) dy.$$

Висновки і перспективи подальших досліджень у даному напрямку.

Область застосування методів аналітичної статистики неозора. У статті показана їх роль як інструменту побудови формальних моделей в управлінні проектами. Відзначено широке застосування регресійного аналізу в управлінні якістю та ризиками проекту, а так само обмежене застосування сучасних методів аналітичної статистики.

Наведено основні напрямки розвитку методів аналітичної статистики, які можуть стати кількісною основою моделей управління проектами та програмами:

елімінування для редукції (скорочення) складної математичної моделі;

непараметричні методи аналітичної статистики для моделей з невідомими законами розподілу;

статистика нечислових даних для експертних моделей, якщо змінними є такі об'єкти нечислової природи, як градації якісних ознак, ранжирування, розбиття, результати парних порівнянь, нечіткі переваги тощо;

математична статистика інтервальних даних в ситуації, коли статистичні дані — не числа, а інтервали, зокрема, породжені накладенням помилок вимірювання на значення випадкових величин.

Показано, що методи статистики інтервальних даних доцільно використовувати для оцінки значень ризиків проекту.

ЛІТЕРАТУРА

1. Цейтлин Н. А. Из опыта аналитического статистика [Текст] : монография / Н. А. Цейтлин. - М.: Солар, 2007. – 912 с.
2. Россошанская О.В. Качественная основа количественного аспекта компетентностной методологии в управлении проектами / О.В. Россошанская // Управління проектами та розвиток виробництва: Зб.наук.пр. – Луганськ: вид-во СНУ ім. В.Даля, 2009. – № 1(29). – С. 75–81.
3. A Guide to the Project Management Body of Knowledge (PMBOK® Guide) - Fifth Edition. – [5-th edition]. – Project Management Institute, Inc., 2013, 589 p.
4. Бараник З. П. Прикладна статистика: історичне становлення та перспективи розвитку / З. П. Бараник, С. С. Ващасв, В. В. Майба // Моделювання та інформ. системи в економіці: зб. наук. праць / М-во освіти і науки, молоді та спорту України, ДВНЗ "Київ. нац. екон. ун-т ім. Вадима Гетьмана"; відп. ред. В. К. Галіцин. – 2011. – Вип. 84. – С. 65–86.
5. Орлов А.И. Новая парадигма прикладной статистики // Заводская лаборатория. Диагностика материалов. 2012. Том 78. №1, часть I. С.87-93.
6. Кузьменко О. В. Інтервальні оцінки ризиків в інноваційних банківських проектах / О. В. Кузьменко // Вісник Національного банку України. – 2008. – № 1. – С. 40–42.
7. Полотай О.І. Модель знань – основа ефективного управління інноваційними освітніми проектами / О.І. Полотай // Восточно-Европейский журнал передовых технологий. - № 1/10 (61). – 2013. . – С. 55–58.
8. Гусєва Ю. Ю. Управління вартістю програми регіонального розвитку [Електронний ресурс] / Ю. Ю. Гусєва, В. К. Доля, Н. О. Манакова // Системи обробки інформації . - 2013. - Вип. 8. - С. 249-252. - Режим доступу: http://nbuv.gov.ua/j-pdf/soi_2013_8_51.pdf
9. Швец С.К. Элиминирование рисков нефинансовых компаний / С.К. Швец. Актуальные вопросы современной экономической науки [Текст]: Сборник докладов VII-й Международной научной конференции (Липецк, 26 ноября 2011 г.). / Отв. ред. А.В. Горбенко. – Липецк: Издательский центр «Гравис», 2012. – С. 139-145.
10. Рунион Р. Справочник по непараметрической статистике: современный подход / Р. Рунион ; пер. с англ. Е. З. Демиденко. - М. : Финансы и статистика, 1982. - 198 с.
11. Орлов А.И. Организационно-экономическое моделирование. Ч.1. Нечисловая статистика. М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. 541 с.
12. Загонова Н.С. Эконометрическая поддержка контроллинга инноваций. Нечеткий выбор / Н.С. Загонова А.И.Орлов // Российское предпринимательство. 2004. №4. С.54-57.
13. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с.
14. Годлевский М. Д. Управление рисками в высокотехнологичных проектах: состояние и подходы управления / М. Д. Годлевский, А. А. Поляков // Пробл. програмув. - 2006. - 2-3 [спец. вип.]. - С. 311-319.
15. Вошинин А. П. Построение аналитических моделей по данным вычислительного эксперимента в задачах анализа чувствительности и оценки экономических рисков / А. П. Вошинин, П. В. Бронз // Заводская лаборатория. – 2007. – Т. 72, № 1. – С. 101–109.