

УДК 303.09
303.733.3

ІРИНА ПРИБИТКОВА,

доктор економічних наук, професор, провідний науковий співробітник відділу соціальних структур Інституту соціології НАН України

Статистичний аналіз і вимірювання в соціологічних дослідженнях

Анотація

Розглянуто методи статистичного аналізу та вимірювання, застосовувані в соціологічних дослідженнях для здійснення первинної систематизації й узагальнення даних статистичного спостереження і повторного опрацювання їх з метою глибшого розуміння сутності й закономірностей розвитку масових суспільних процесів і структур за конкретних умов місця і часу. Обговорюється, зокрема, використання марковських моделей, що дають імовірнісний опис механізму перерозподілу індивідів за стратифікаційними групами при моделюванні процесів соціальної мобільності. Викладено результати застосування крокового регресійного аналізу для прогнозування територіальної мобільності населення. Продемонстровано можливості використання графічних моделей статистичних розподілів (кривої Лоренца, кумулятивної кривої) при вивченні просторової структури переміщень сільського населення з трудовими цілями.

Ключові слова: *кількісний аналіз як складова теорії, методи повторного опрацювання й аналізу статистичної інформації, моделювання і прогнозування процесів просторової соціальної мобільності населення, графічні моделі статистичних розподілів*

Кількісний аналіз як складова теорії

Систематизація нагромаджених у процесі дослідження фактів, узагальнення, аналіз й інтерпретація їх, установлення зв'язку між ними і вже описаними фактами є важливим моментом розроблення теорії. Проте не кожен

факт можна вважати науковим. Лише за умови включення в певну теоретичну систему він набуває якості науковості. Науково-теоретичне осмислення конкретних фактів, що відображають характер, тенденції й закономірності розвитку масових суспільних процесів і явищ, дає змогу розшифрувати інформацію про їхню сутність і взаємозв'язки.

На підставі кількісного аналізу емпіричних даних формуються теоретичні гіпотези, що дають можливість в узагальненому вигляді пояснити соціальні, економічні й демографічні явища та процеси. Таке пояснення необхідно для того, щоб спрямувати дослідження на пошук систематичних зв'язків між фактами. З цієї точки зору кількісний аналіз можна інтерпретувати як складову теорії. Однак, коли розглядають конкретні способи опрацювання й упорядкування емпіричних даних і перевірки гіпотез статистичними методами, кількісний аналіз слід віднести до царини власне статистики.

Е.Кейн зазначає, що головну роль у кількісному аналізі завжди відіграють формулювання гіпотези (специфікація) та її перевірка. Значення статистики здебільшого полягає в тому, що вона може по-новому висвітлити теоретичні композиції, за якими передбачається описати і спрогнозувати ймовірні напрями, динаміку й інтенсивність розвитку різноманітних структур і процесів [Кейн, 1977: с. 9–10].

На думку Е.Кейна, основним предметом професійного інтересу статистиків переважно є *гіпотези, що можуть підлягати спростуванню*, тобто такі, хибність яких можна в принципі довести. Він стверджує, що тільки тоді, коли гіпотеза ризикує бути спростованою, існує вибір між нею й альтернативним теоретичним припущенням. І доходить висновку, згідно з яким для досвідченого аналітика поняття *спростованість* і *перевірка* — це майже синоніми.

Розвиток теоретичних уявлень у будь-якій царині наукового знання, за Е.Кейном, відбувається в такій послідовності:

- розроблення теорій, які дедалі більше піддаються перевірці на підставі емпіричних даних;
- верифікація теоретичних припущень шляхом зіставлення їх з емпіричними даними до того моменту, коли ці теорії будуть спростовані;
- заміна спростованих теорій новими.

У результаті нечіткі теорії набувають відточенішої форми і беруть участь в подальших випробуваннях і перевірках за емпіричними даними. Теорії, що суперечать фактам, підлягають перегляду. Натомість інші, котрі успішно здолали черговий тур випробувань, зазнають нових — складніших і більш трудомістких [Кейн, 1977: с. 10].

Конче важливо пам'ятати про неприпустимість пристосовування (адаптації) своєї теорії до тих чи інших конкретних фактів. Не менш важливим завданням є ретельне розмежування перевірки вже наявної теорії й аналізу спостережень для формулювання нової теорії, котра іще очікує на перевірку. Е.Кейн стверджує, що різні теорії, створені на одній множині даних, неможливо перевірити на підставі тих самих фактів, оскільки така перевірка не забезпечує незалежного підтвердження теорії. Незалежним під-

твердженням можуть слугувати лише нові факти або відомості, отримані в результаті подальшого дослідження старих фактів [Кейн, 1977: с. 40].

Кількісні закономірності масових суспільних процесів виявляються у формі емпіричних статистичних закономірностей. Виявлення і дослідження їх лежить в основі узагальненої кількісної характеристики цих процесів.

Статистика як наука

Згадування про статистику як науку з'явилося вперше у другій половині XVII століття у працях політичних арифметиків. Засновником нової науки К.Маркс вважав В.Петті, а його метод характеризував як нетрадиційний. "...Замість набору цілої низки слів у вищому і найвищому ступені та спекулятивних аргументів, він (В. Петті) вирішив ... говорити за посередництва чисел, ваг і вимірників, користуватися винятково аргументами, взятими з чуттєвого досвіду, і розглядати лише такі причини... що мають взаємне підґрунтя у природі", — писав класик [Маркс, Енгельс, т. 13: с. 39]. В.Петті показав наукову силу та переконливість нового методу і необхідність застосування його на противагу доказам, ґрунтованим лише на умоглядних аргументах. Подальшого розвитку статистична теорія набула у працях представників різних наукових шкіл, а сама статистика і як наука, і як галузь практичної діяльності досягла значних успіхів у техніці збирання, опрацювання, аналізу й узагальнення статистичної інформації.

Як наука статистика вивчає кількісний аспект масових суспільних явищ у нерозривному зв'язку з їхнім якісним аспектом, досліджує кількісне вираження закономірностей суспільного розвитку за конкретних умов місця і часу [Ряузов, 1979: с. 8]. Кількісні характеристики суспільних явищ статистика подає в цифрах, що змінюються в часі та просторі. Близьке до цього визначення тлумачення змісту статистики як галузі знання надає також поширений американський довідник з класифікації робіт в адміністративних установах США: статистика є *наукою про збирання, класифікацію і кількісне оцінювання фактів як підґрунтя для висновків*.

Важливе значення у статистиці має поняття *сукупності*, що являє собою множину окремих одиниць, об'єднаних спільною якісною основою, але відмінних одне від одного за низкою ознак. Самі статистичні цифри бувають двоякого роду: одні з них становлять узагальнену характеристику обсягів сукупності (наприклад, загальної чисельності населення, розмірів трудових ресурсів чи житлового фонду в країні); інші дають узагальнені характеристики сукупностей за низкою ознак, що мають різне кількісне або якісне вираження. Такі ознаки називають варіаційними (змінюваними).

Узагальнені характеристики масових суспільних процесів і явищ статистика використовує для пошуку закономірностей, що виявляються у певному порядку розташування, співвідношення або зміни статистичних цифр. Це можуть бути *закономірності розвитку (динаміки) різноманітних процесів або явищ; закономірності зміни їхньої структури, розподілу одиниць усередині сукупності і, зрештою, зв'язної зміни різних ознак, що варіюють, всередині сукупностей* (наприклад, розподіл зайнятих певних професій і кваліфікації за величиною стажу практичної роботи).

Закономірності громадського життя завжди відображають дію складних причин, співвідношення яких і сукупний результат змінюються мірою розвитку суспільства і, отже, вочевидноються *лише за умов конкретного місця і часу*. Найважливіше завдання статистичних досліджень полягає у виявленні й вимірюванні закономірностей масових процесів і явищ методами, розробленими статистикою як наукою [Рязов, 1979: с. 14]. Специфічні прийоми, за допомогою яких статистика вивчає свій предмет, утворюють статистичну методологію. У кожному статистичному дослідженні можна виокремити три послідовні стадії: *статистичне спостереження* (збирання первинного статистичного матеріалу); *зведення результатів спостереження в певні сукупності*; *аналіз отриманих зведених матеріалів*. На всіх цих трьох стадіях як інструмент статистичної методології використовують спеціальні методи.

Методи статистичного аналізу вможливають первинну систематизацію й узагальнення цифрових даних, що описують досліджувані сукупності. Глибшого розуміння сутності досліджуваних процесів або явищ можна досягти лише в результаті *вторинного опрацювання даних статистичного спостереження*, сенс якого полягає в подальшому узагальненні вже отриманих результатів і науковій інтерпретації їх.

Методи вторинного опрацювання й аналізу даних зводяться зрештою до математико-статистичного моделювання, яке можна здійснювати різними засобами. Графіки, використовувані в математико-статистичному аналізі, являють собою одну з форм математичного моделювання, цілком рівноправну з аналітичною (алгебраїчною) моделлю і наділену низкою переваг перед нею. Зокрема, в моделях, що описують статистичні сукупності зі складною структурою розподілу значень варіативної ознаки, доречнішою є графічна форма запису. Основні функції графічних моделей полягають у відображенні обсягу та структури статистичних сукупностей, зміни їх у часі та розподілу в просторі, у виявленні закономірностей, властивих цим сукупностям.

Наприклад, аналіз структурної діаграми, відомої у практиці демографічної статистики як *вікова піраміда*, дає змогу скласти уявлення про статево-вікову структуру населення і виявити відхилення вікового складу населення від нормального. Крім того, вікова піраміда забезпечує можливість демографічного прогнозу на найближче майбутнє.

Використовуючи графічне зображення динамічних рядів, можна наочно уявити характер розвитку того чи іншого процесу, явища або структури. Для аналізу форм статистичного розподілу можна застосувати кілька моделей: гауссову криву нормального розподілу, моделі розподілу Пуассона, Пірсона, Фішера, Стьюдента, криву Лоренца та низку інших.

Завдання графічної моделі, відомої як крива Лоренца, полягають у визначенні й вимірюванні відмінностей у рівнях концентрації великих мас у певній кількості окремих великих одиниць, у встановленні питомої ваги чи відносного значення цих великих одиниць в усій генеральній сукупності. Як інструмент аналізу статистичних розподілів можна використати також графічну модель наростаючих або нагромаджених частот (кумулятивна крива), широко застосовувану в статистичній практиці.

Математичне моделювання входить до числа загальних наукових методів, застосовуваних в економіко-соціологічних і соціально-демографічних дослідженнях. На певному етапі побудови наукової теорії виникає потреба в її формалізації та забезпеченні строгого логічного підґрунтя для висловлених гіпотез. Математика забезпечує тут той логічний фундамент, на якому можна будувати теорію [Бунге, 1967: с. 58]. На думку Д. Гарвея, модель можна розглядати як формалізований виклад теорії засобами формальної логіки, теорії множин і математичної статистики, застосування яких дає змогу виявити й усунути внутрішню неузгодженість теорії [Харвей, 1976: с. 240].

Найпоширенішими моделями, застосовуваними в дослідженнях розвитку просторових систем і придатними для аналізу міграційних потоків, є гравітаційні, а також моделі потенціалів і просторової взаємодії. Поряд із гравітаційними моделями широко використовуються призначені для практичних розрахунків регресійні моделі, хоча вони нерідко дають невисокі значення коефіцієнтів множинної регресії. Тому регресійний аналіз доцільно доповнювати якісними дослідженнями й ретельною інтерпретацією отриманих результатів [Прибыткова, 2009: с. 87].

Обнадійливі результати забезпечує комплексне використання статистичних методів, зокрема поєднання варіаційного, дисперсного, кореляційного і регресійного аналізів.

Оскільки регресійний аналіз дає результати, придатні для інтерпретації, лише застосовно до вельми однорідних сукупностей, рекомендується розподілити вихідні дані на однорідні групи і вже потім застосовувати регресійний аналіз до кожної з них окремо [Айвазян, 1973: с. 968–983]. Однак при виборі одного із багатьох можливих методів регресійного аналізу слід урахувувати властивості отриманих у результаті класифікації груп. Пакети прикладних статистичних програм зазвичай включають низку методів класифікації та регресії.

Моделювання процесів соціальної мобільності

Об'єктом моделювання процесів соціальної мобільності є різноманітні види руху безлічі індивідів і груп у часі та просторі: соціальні переміщення, різні види руху трудових ресурсів (територіального, міжгалузевого, міжпрофесійного). А початкове завдання полягає в описі їхньої поведінки, що має низку характерних рис.

- Соціальна поведінка детермінована багатьма особистими мотивами, що нерідко мають суперечливий характер і визначаються у підсумку конкретними об'єктивними обставинами.
- Соціальна поведінка індивідів характеризується нестабільністю, позаяк чинники, що її визначають, перебувають у стані безупинної зміни.
- Соціальна поведінка є чітко вираженим імовірнісним процесом.
- Соціальна поведінка вирізняється суттєвою неоднорідністю сукупності осіб, залучених у процеси соціальної мобільності [Движение рабочих кадров, 1974: с. 40].

При моделюванні процесів соціальної мобільності сукупність залучених у них індивідів поділяють на k стратифікаційних груп, кожна з яких об'єднує індивідів з однорідними характеристиками. Передбачається, що індивіди можуть переходити з однієї стратифікаційної групи до іншої і k -вимірний вектор $x(t)$ задає чисельності індивідів у групах у момент часу t . Опис поведінки вектора $x(t)$ в часі — завдання першорядної важливості, розгляд якого за припущення про випадковий характер індивідуальних переходів неодмінно приводить до вивчення векторного випадкового процесу $[x(t), t \leq 0]$.

При моделюванні соціальної мобільності найчастіше використовують моделі марковського типу [Stewman, 1976: р. 201–245; Романов, Терехов, 1980: с. 272–291], які дають змогу враховувати можливість кожного індивіда обирати з деякою імовірністю один із кількох передбачуваних способів поведінки. Це дає можливість природним чином отримати ймовірнісний опис механізму перерозподілу індивідів за стратифікаційними групами. Найширше застосування дістають різноманітні модифікації найпростішої марковської моделі з дискретним часом (наприклад, за рік) і постійною в часі матрицею перехідних імовірностей. Серед них слід назвати модифікацію моделі, що враховує неоднорідність індивідуальної активності, а також модифікацію моделі, що враховує мінливість поведінки індивідів у часі [Орлов, 1983: с. 216].

Задовільних результатів від використання найпростішої марковської моделі очікують при моделюванні процесів, у яких показники мобільності стаціонарні, а також за умови, що індивіди з однієї стратифікаційної групи поведуться однаково, тобто група однорідна. Такими процесами є, до прикладу, процеси територіальної мобільності (переїзди на постійне місце проживання), соціальні переміщення індивідів, процеси міжгалузевої й міжпрофесійної мобільності.

Для побудови інтервального прогнозу розподілу індивідів за стратифікаційними групами потрібні дані про пересування індивідів у вигляді матриці. Отримання такої інформації — це самостійна проблема, і пов'язане воно, як правило, із проведенням трудомістких досліджень, як, наприклад, при побудові міжпрофесійного балансу трудових ресурсів.

Наостанок зазначу, що ймовірнісний підхід відкриває широкі можливості для пояснення просторових взаємодій (зв'язків, потоків) і вивчення різноманітних проявів просторової діяльності людини.

Використання прикладного регресійного аналізу в прогнозуванні процесів територіальної мобільності населення

Розвиток процесів територіальної мобільності населення відбувається під впливом безлічі різноманітних чинників. Деякі з них є результатом цілеспрямованої діяльності людини, проте інші не залежать від неї. Тому при прогнозуванні розмірів, напрямів, інтенсивності та структури потоків територіальної мобільності населення сукупність аналізованих чинників можна розподілити на дві основні групи: регульовані та нерегульовані. Між

ними існують складні взаємозв'язки, вони діють одночасно, і їхній вплив на розміри й інтенсивність міжпоселенських і міжкраїнових потоків індивідів має комплексний характер. Аналіз, оцінювання і прогнозування обсягів, напрямів і структури цих потоків можливі шляхом кількісного вивчення рівня впливу всіх чинників — як регульованих, так і нерегульованих — на досліджуваний показник.

Багатофакторний кореляційно-регресійний аналіз дає змогу оцінити міру впливу на досліджуваний результативний показник кожного із включених в аналіз чинників за фіксованого положення (на середньому рівні) всіх інших, а також у разі будь-яких можливих сполучень чинників знайти з певною мірою точності теоретичне значення цього показника. При цьому необхідною умовою є відсутність між ними функціонального зв'язку. На основі попереднього теоретичного аналізу формулюється гіпотеза про причинно-наслідкові зв'язки.

Лише після того, як такі припущення висунуто, багатофакторний кореляційно-регресійний аналіз може або підтвердити, або спростувати їх і дати кількісне оцінювання впливу різноманітних чинників.

У перебігу кореляційно-регресійного аналізу здійснюють побудову й аналіз статистичної моделі у вигляді рівняння регресії, що наближено виражає залежність досліджуваного показника від чинників, що впливають на нього. Математичне завдання формулюється так: із багатьох функцій потрібно вибрати аналітичне вираження, котре якнайкраще відобразить реальні зв'язки між досліджуваним показником і чинниками, тобто знайти функцію $Y = f(x_1, x_2, \dots, x_n)$, де в нашому випадку Y — територіальна мобільність населення, а x_1, x_2, \dots, x_n — чинники, що впливають на її формування.

Завдання полягає в тому, щоби з'ясувати характер і міру впливу чинників на показники територіальної мобільності населення. Якщо завдання успішно виконане, ми матимемо інструмент для прогнозування, а в деяких випадках — і для управління досліджуваним процесом. Вибір типу функції здійснюють, спираючись або на теоретичні знання про досліджуваний процес, або на попередній досвід аналогічних досліджень, або ж визначають емпірично, шляхом перебирання функцій різних типів. Оскільки в більшості випадків будь-яку функцію багатьох змінних шляхом їх заміни чи логарифмування можна звести до лінійного вигляду, рівняння множинної регресії будують у лінійній формі $Y = f(a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n)$.

Кожен коефіцієнт цього рівняння показує міру впливу того чи іншого фактора на аналізований показник за фіксованого стану решти факторів. Зі зміною кожного фактора на одиницю показник змінюється на величину відповідного коефіцієнта регресії. Вільний член рівняння не несе жодного соціально-економічного навантаження. Рівняння множинної регресії враховує середній вплив на аналізований показник не всіх, а лише найсуттєвіших факторів. Тому навіть у разі найбільш вдало обраної функції виникає розкид фактичних значень Y навколо обчислюваних. При побудові рівняння виникає проблема добору змінних, включених у регресійний аналіз.

Існує кілька статистичних методів добору змінних: метод усіх можливих регресій, метод виключення, метод включення, покроковий регресійний аналіз, дві варіації чотирьох попередніх методів, пощаблевий регресійний аналіз. Непогані результати забезпечує кроковий регресійний аналіз, що являє собою поліпшену модифікацію методу включення.

Метод включення дає змогу уникнути опрацювання більшої кількості змінних, ніж це необхідно. До аналізу включають змінні по черзі, доки регресійне рівняння стане задовільним. Порядок включення їх визначають на підставі часткового коефіцієнта кореляції, використовуюваного як вимірник важливості ще не включених у рівняння змінних. Проте цей метод не дає можливості досліджувати вплив чергової змінної на внесок іншої змінної, введеної в рівняння на попередній стадії. Цю ваду долає покроковий регресійний аналіз, що передбачає додаткове дослідження й оцінювання на кожній стадії міри важливості змінних, включених у модель на попередніх стадіях.

Для оцінювання значимості внеску кожної зі змінних на черговому кроці здійснюють дисперсійний аналіз, у рамках якого обчислюється частковий F -критерій (критерій Фішера). Дисперсійний аналіз дозволяє математично строго перевірити суттєвість аналізованих зв'язків. Якщо змінна, що робила суттєвий внесок у модель на ранніх стадіях її побудови, вводиться в регресійне рівняння після інших змінних, вона може виявитися зайвою через значну кореляцію з ними. Таким чином, F -критерій уможливує оцінювання відносного ефекту кожної змінної після інших і слугує для розв'язання питання щодо додавання або виключення членів моделі.

У процесі добору змінних для включення в регресійне рівняння варто з'ясувати, чи є та чи інша змінна тільки вимірюваною, або ж вона поряд із тим є ще й керованою. Якщо модель призначена для цілей управління, до неї мають бути включені керовані змінні, навіть якщо внесок їх у рівняння менш значимий.

Однак самі коефіцієнти регресії не дають підстав стверджувати, які з них справляють найбільший вплив на розміри потоків територіальної мобільності (оскільки вони виміряні різними одиницями), а також у розвитку яких чинників закладено найзначніші резерви регулювання цього процесу в заданому напрямі. Для цього обчислюють часткові коефіцієнти еластичності (E_i) і так звані β -коефіцієнти (β_i).

Завдяки частковим коефіцієнтам еластичності усуваються розбіжності в одиницях вимірювання чинників. Вони показують, на скільки відсотків у середньому змінюється аналізований показник зі зміною на 1% кожного фактора за фіксованого стану решти факторів. β -коефіцієнт показує, на яку частину середньоквадратичного відхилення змінюється досліджуваний показник (залежна змінна) зі зміною відповідного чинника на величину його середньоквадратичного відхилення. Таким чином, на підставі часткових коефіцієнтів еластичності та β -коефіцієнтів можна судити про внесок, що його роблять включені в рівняння фактори в динаміку показників територіальної мобільності населення.

Мною було зроблено спробу визначити найсуттєвіші фактори та силу їхнього впливу на формування потоків територіальної трудової мобільності

сільського населення в Україні. Як залежну змінну Y взято величину контингенту осіб, які щодня здійснюють трудові поїздки із сільської місцевості до місця докладання праці в міських поселеннях і назад. До аналізу включено всі суттєві фактори, що визначають умови праці, побуту та відпочинку, а також рівень добробуту та якості життя сільського населення країни: рівень трудових доходів, розміри грошових заощаджень і нагромадження предметів тривалого користування, можливості здобуття жителями села освіти та професії, доступність і якість медичної допомоги, розвиток дорожньої мережі, забезпеченість установами культурно-побутового обслуговування й торгівлі та ін. Для обчислення коефіцієнтів кореляції та регресії було використано динамічні ряди включених в аналіз показників тривалістю 15 років.

Вибір рівняння, яке б найліпшим чином відображало реальні зв'язки між розмірами потоків щоденної трудової міграції в міста із сільської місцевості і факторами, що впливають на цей процес, був здійснений у процесі покрокового регресійного аналізу. Рівняння набуло такого вигляду:

$$Y = 1779,818 + 1,346X_1 + 0,028X_2 - 255,158X_3 + 9,510X_4 \quad (1),$$

де X_1 — величина роздрібного товарообігу в розрахунку на душу населення в сільській місцевості;

X_2 — будівництво житлових будинків, тис. м² загальної/корисної площі;

X_3 — кількість клубних закладів у сільській місцевості в розрахунку на 10000 жителів;

X_4 — щільність мережі автомобільних шляхів із твердим покриттям, км на 1000 км² території.

Головна ідея покрокового регресійного аналізу полягає в пошуку регресії з кількома змінними у вигляді серій лінійних регресійних залежностей і в перебудові таблиці крок за кроком до кінця. Крім того, на кожному етапі обчислюємо сукупний коефіцієнт множинної кореляції R і сукупний коефіцієнт множинної детермінації R^2 .

Коефіцієнт множинної кореляції R є основним показником лінійного кореляційного зв'язку. Він вимірює одночасний вплив змінних-факторів на залежну змінну і є індикатором сили зв'язку між ними. Чим менше спостережувані значення досліджуваного показника відхиляються від лінії множинної регресії, тим кореляційний зв'язок є інтенсивнішим, а отже, величина R — ближчою до одиниці. Сукупний коефіцієнт множинної детермінації показує, яка частка варіації досліджуваного показника пояснюється впливом факторів, включених у рівняння регресії. Чим ближче R^2 до одиниці, тим більшою мірою варіація залежної змінної характеризується впливом відібраних чинників.

У нашому випадку R становить 0,9979, а R^2 дорівнює 0,9959, тобто зміна територіальної трудової мобільності сільського населення України на 99,6% зумовлювана 4 чинниками, включеними в регресійне рівняння.

Суттєве значення для розвитку процесу територіальної мобільності сільських жителів має їхня поінформованість про умови і характер праці в місцях їхнього передбачуваного працевлаштування. Переміщенню у просторі як факту завжди передує процес прийняття рішення про зміну місця і сфери докладання праці, характеру трудової діяльності й належності до певної соціальної групи населення. Включені до аналізу фактори передусім впливають на прийняття рішення. Саме ж переміщення індивідів має стосовно рішення вторинний характер, підпорядковане йому й слугує цілям його реалізації.

Отримане рівняння регресії (1) можна використати для цілей прогнозування. Для обчислення перспективних значень Y достатньо підставити в нього прогнозні значення X_1, X_2, X_3 і X_4 .

Графічні моделі як інструмент аналізу просторової нерівномірності розподілу трудових потоків сільських жителів у міста й інтенсивності їх в Україні

Територіальні переміщення сільського населення в міста із трудовою метою поширені на всіх теренах України. Однак рівні їхнього розвитку в окремих регіонах країни вирізняються значною нерівномірністю. Для виявлення міри рівномірності розподілу між сільськими адміністративними районами мас індивідів, що утворюють трудові потоки, я скористалася графічною моделлю, відомою як *крива Лоренца* (рис. 1). Графік концентрації (нерівномірності) розподілу за моделлю Лоренца має форму квадрата, вздовж сторін якого від нульової точки йдуть дві однакові відсоткові шкали від 0 до 100%. На горизонтальній шкалі подано збільшення відсотків числа сільських адміністративних районів, на вертикальній — збільшення відсотків контингентів мешканців цих районів, зайнятих у міських населених пунктах. Відсоткові шкали слугують координатами, за якими в полі графіка точками позначено розраховані у спеціальній таблиці показники за однойменними групами (табл. 1). Сполучені між собою точки утворюють плавні ввігнуті криві Лоренца.

Проведена в полі графіка з нульової точки діагональна лінія являє собою теоретичну межу рівномірності розподілу, практично недосяжну. Міра відхилення кривих Лоренца від діагоналі відображає концентрацію або ступінь нерівномірності розподілу трудових мігрантів із числа сільських жителів між сільськими районами України.

Криві Лоренца, суміщені в полі графіка, вказують, що впродовж 1968–1981 років нерівномірність розподілу сільських жителів між сільськими адміністративними районами України, які працюють щодня у місті, зменшилася. Так, у 1968-му майже половина всіх зайнятих у місті жителів села (48%) була зосереджена лише в 10% від загальної сукупності цих районів, а на 30% районів припадало 80% трудових мігрантів із числа сільських жителів. До 1974 року в 10% районів мешкало вже 41%, а на 30% районів — 75% загальної сукупності сільських жителів, які працюють у місті. До 1981-го у 10% районів концентрація трудових мігрантів, зайнятих у місті, зменшилася до 38%, а у 30% районів — до 72%. І якщо у 1968 році в по-

ловині сільських адміністративних районів мешкали всього 8% зайнятих у міських поселеннях трудових мігрантів, то на 1981 рік питома вага їх збільшилася до 13%. Таким чином, упродовж 1968–1981 років у межах України вочевиднилася тенденція до певної деконцентрації потоків трудової маятничкової міграції сільського населення, спрямованих у міста.

Таблиця 1

Трудова маятничкова міграція сільського населення в міста України на 01.01.1981 року (групування для впорядкування графіка концентрації трудових мігрантів за моделлю Лоренца)

Групи за розмірами контингентів маятничкових мігрантів, осіб	Кількість районів у групі	Середня величина інтервалу, осіб	Загальна кількість маятничкових мігрантів, осіб	Підсумки за зростанням			
				Абсолютні дані		Відсотки	
				Кількість районів	Загальна кількість маятничкових мігрантів, осіб	Кількість районів	Загальна кількість маятничкових мігрантів, осіб
До 100	24	50	1200	24	1200	5,3	0,1
101–200	21	150	3150	45	4350	10,0	0,3
201–400	33	300	9900	78	14250	17,2	0,9
401–600	33	500	16500	111	30750	24,6	2,0
601–800	29	700	20300	140	51050	31,0	3,2
801–1000	16	900	14400	156	65450	34,5	4,2
1001–2000	104	1500	156000	260	221450	57,5	14,1
2001–3000	48	2500	120000	308	341450	68,1	21,7
3001–4000	33	3500	115500	341	456950	75,4	29,0
4001–5000	22	4500	99000	363	555950	80,3	35,3
5001–6000	11	5500	60500	374	616450	82,7	39,2
6001–8000	23	7000	161000	397	777450	87,8	49,4
8001–10000	13	9000	117000	410	894450	90,7	56,8
10001–15000	22	12500	275000	432	1169450	95,6	74,3
15001–20000	14	17500	245000	446	1414450	98,7	89,8
20001–25000	2	22500	45000	448	1452450	99,1	92,7
Понад 25000	4	28750	115000	452	1574450	100,0	100,0

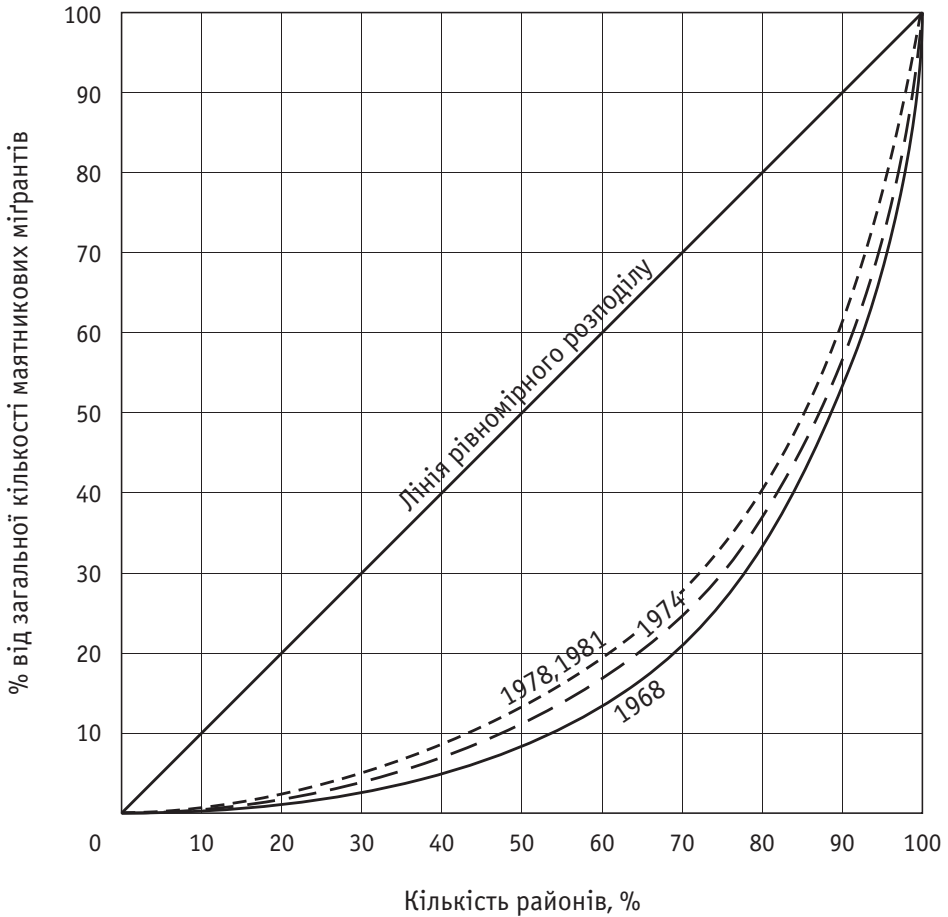


Рис. 1. Графік концентрації (нерівномірності) розподілу за моделлю Лоренца. Розподіл загальної кількості трудових маятникових мігрантів між сільськими адміністративними районами України у 1968, 1974, 1978 і 1981 роках

Аналіз розподілу сільських адміністративних районів України за рівнем інтенсивності трудової маятникової міграції сільського населення (рис. 2) був здійснений за графічною моделлю нагромаджених частот (кумулятивної кривої). Кумулятивний графік розподілу побудований у вигляді прямокутника, сторони якого являють собою градуйовані відсоткові шкали: на вертикальній шкалі розташовано питомі ваги трудових маятникових мігрантів із числа сільських жителів у загальній чисельності сільського працездатного населення адміністративного району; на горизонтальній — відсотки кількості районів за наростанням. Кумулятивні криві розраховано за принципом “не більш як”, розрахункову схему їх подано в таблиці 2.

Аналіз кумулятивних кривих, побудованих за даними 1968, 1974, 1978 і 1981 років, приводить нас до висновку про зменшення в цей період нерівномірності розподілу сільських адміністративних районів України за рівнем інтенсивності щоденних трудових потоків сільських жителів у міста.

У 1968 році переважна частина сільських адміністративних районів характеризувалися низькою інтенсивністю трудової маятникової міграції. У 80% районів питома вага зайнятих у містах сільських жителів не перевищувала 10% (у тому числі у 20% районів участь їх у цьому процесі позначалася показниками, меншими за 1%). У 90% загальної сукупності районів інтенсивність щоденних трудових поїздок сільських жителів у міста перебувала в межах 20%, і лишень 10% районів характеризувалися показниками понад 20% (від 20% до 69%).

У 1974 році кількість сільських районів, позначених показниками інтенсивності щоденних трудових потоків у міста понад 20%, досягла 16%; скоротилося число районів із мінімальним рівнем інтенсивності щоденних трудових поїздок сільських жителів у міста (менш як 1%): їх налічувалося лишень 10,3%. Для 80% районів показники інтенсивності цього процесу зросли до 17%, а для 9% районів сягнули рівня 27%.

У 1981 році інтенсивність трудових поїздок сільських жителів у міста в межах 1% була зафіксована лише у 7,7% районів. Частка районів із рівнем інтенсивності цього процесу понад 20% збільшилася до 22,4%; показники інтенсивності понад 30% спостерігалися вже в 11,2% районів.

Таким чином, уже до 1981 року відбулося скорочення числа районів із низьким рівнем інтенсивності територіальної мобільності сільського населення в напрямку міста із трудовими цілями за одночасного зростання частки районів із найвищими показниками розвитку цього процесу.

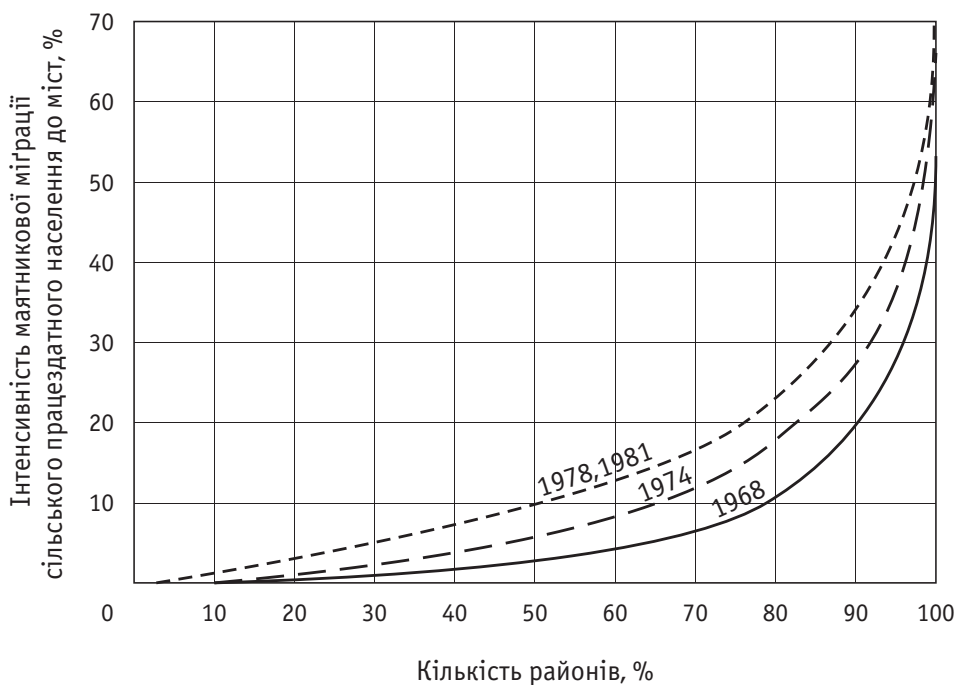


Рис. 2. Кумулятивний графік розподілу сільських адміністративних районів України за рівнем інтенсивності трудової маятникової міграції сільського працездатного населення у 1968, 1974, 1978 і 1981 роках

Таблиця 2

Розподіл сільських адміністративних районів України за рівнем інтенсивності трудової маятникової міграції населення на 01.01.1981 року (групування для впорядкування кумулятивного графіка розподілу)

Групи за рівнем інтенсивності трудової маятникової міграції сільського населення, у відсотках	Кількість районів у групі	Кумулята за кількістю районів ^a	
		В абсолютних значеннях	У відсотках
1,0	35	35	7,7
1,1–2,5	32	67	15,4
2,6–5,0	64	131	28,8
5,1–7,5	60	191	42,0
7,6–10,0	49	240	52,7
10,1–12,5	45	285	62,6
12,6–15,0	29	314	69,0
15,1–17,5	22	336	73,8
17,6–20,0	17	353	77,6
20,1–25,0	22	375	82,4
25,1–30,0	29	404	88,8
30,1–35,0	10	414	91,0
35,1–40,0	9	423	93,0
40,1–45,0	11	434	95,4
45,1–50,0	4	438	96,3
50,1–60,0	15	453	99,6
60,1–70,0	1	454	99,8
70,1–80,0	1	455	100,0

^a Кумуляту складено за принципом “не більш як”. Наприклад, у 35 (або 7,7%) сільських адміністративних районах інтенсивність трудової міграції сільського населення не більш як 1,0%; у 67 (15,4%) — не більш як 2,5% і т.д.

Висновки

Використовуючи статистичні моделі, соціологи отримують інструмент аналізу, завдяки якому можна зрозуміти багато речей, але ніколи не вдасться пізнати усе. Як правило, цього й не потрібно. Знати треба лише те, що насправді важливо. Утім, з іншого боку, ніколи не можна бути впевненим, що в моделі враховано всі чинники, конче важливі при формуванні теоретичних гіпотез або при розв’язанні прикладних задач. Статистика як наука намагається отримати відповіді на питання на підставі матеріалу, якого в принципі недостатньо. Тому вона визначає правила, що дають можливість

судити про рівень достатності або недостатності того чи іншого конкретного висновку. Центральною проблемою в опануванні статистики є теорія статистичного висновку, а поняття ймовірності можна назвати справжнім підґрунтям статистики.

Джерела

Айвазян С.А. Многомерный статистический анализ в социально-экономических исследованиях / Айвазян С.А. — М., 1973.

Бунге В. Теоретическая география / Бунге В. — М., 1967.

Движение рабочих кадров на промышленных предприятиях: теоретические и методические вопросы анализа текучести. — М., 1974.

Кейн Э. Экономическая статистика и эконометрия. Введение в количественный экономический анализ / Кейн Э. — М., 1977.

Маркс Ф. Сочинения / Ф. Маркс, Ф. Энгельс. — 2-е изд. — Т. 13.

Орлов В.И. О прогнозировании процессов социальной мобильности / В.И. Орлов // Прикладная статистика. — М., 1983.

Прибыткова И.М. Пространственная самоорганизация населения: теоретико-методологические предпосылки исследования / И.М. Прибыткова // Социология: теория, методы, маркетинг. — 2009. — № 4.

Романов А.К. Математические модели процессов мобильности : обзор / А.К. Романов, А.И. Терехов. — М., 1980.

Ряузов Н.А. Общая теория статистики / Ряузов Н.А. — М., 1979.

Харвей Д. Социальные процессы и пространственная форма. Анализ концептуальных проблем городского планирования / Д. Харвей // Новые идеи в географии ; вып. 2. — М., 1976.

Stewman S. Markov models of occupational mobility: theoretical development and empirical support / Stewman S. — Pt. I, II. 1976.