

8. Тупахина О. В. Религиозная биография как форма нарративизации жизненного мира верующих (на примере православных христиан) [Электронный ресурс] / О. В. Тупахина. – Режим доступа : http://jourssa.ru/sites/all/files/volumes/2012_1/Tupakhina_2012_1.pdf
9. Улановский А. М. Изучение религиозных переживаний и дискурса (на материале исследования кришнаитов) [Электронный ресурс] / А. М. Улановский, А. Б. Матюхина. – Режим доступа : <http://www.hse.ru/data/2010/05/14/1217166239/Ulanovsky%20Matyukhina%20Religious%20experience.pdf>
10. Шюц А. О множественности реальностей / А. Шюц // Социологическое обозрение. – Т. 3. – № 2. – 2003 – С. 18.
11. Ярская-Смирнова Е. Р. Нарративный анализ в социологии [Электронный ресурс] / Е. Р. Ярская-Смирнова // Социолог. журн. – 1997. – № 3. – Режим доступа : <http://www.socjournal.ru/article/343>

Миронович Дмитро. Наративізація релігійного досвіду в (авто)біографічних інтерв'ю. У межах статті у феноменологічній перспективі розглянуто можливості соціологічного аналізу релігійного досвіду. Запропоновано концептуалізацію понять «наратив» і «наративна ідентичність», які використовуються в аналізі релігійних (авто) біографій. Відзначено, що характер і склад релігійного досвіду позбавлені універсального характеру й тісно пов'язані з особливостями категоризації та інтерпретації подій у кожній із релігійних традицій. Припускається, що в межах певної релігійної традиції або чітко окресленої релігійної спільноти існують загальні колективні наративні зразки, що описують та обумовлюють певні моделі й правила включення в релігійні спільноти. Отже, наративна ідентичність, що формується в певному релігійному середовищі, постає обмеженою набором «готових форм» та інтерпретаційних моделей, що існують у цій наративній спільноті. Саме тому завдання дослідника – пошук відзначених «офіційних» зразків інтерпретацій і схем аргументації в релігійному наративі серед різних релігійних традицій.

Ключові слова: релігійний наратив, наративна ідентичність, наративні спільноти, релігійне навернення.

Myronovych Dmytro. Religious Narratives in Life-History Research. In this article we attempt to clarify the methodological basis of sociological analysis of the phenomena of religious experience through the prism of phenomenological perspective. The author describes conceptual frameworks for understanding personal religious experience in terms of narrative, narrative identity, and life history. The author suggests that the nature of religious experience devoid of universal character and closely related with specific features of categorization and interpretation of reality in each of the religious traditions. The paper argues that a particular religious tradition or a clearly defined religious community use the common narrative patterns. These common narrative patterns describe and prescribe certain models and rules of community inclusion. That is why the researcher's task is the search of such «official» narrative models and patterns of argumentation in religious narrative.

Key words: religious narrative, narrative identity, narrative community, religious conversion.

УДК 316.303.09, 316.303.8

*Микола Сидоров,
Євген Білоус*

Підходи до аналізу даних із невідповідями

Суперечність між необхідністю аналізу невідповідей та відсутністю узагальненої інформації щодо розуміння природи виникнення різного виду відсутніх значень актуалізує розгляд основних методів аналізу невідповідей в емпіричних соціологічних дослідженнях. У межах цієї статті наведено класифікацію методів, порівняно недоліки та переваги кожного з урахуванням того, що методи аналізу невідповідей різняться механізмами їх реалізації й обумовлюються подальшим аналізом та інтерпретацією даних: або 1) видалити невідповіді й реалізовувати процедури аналізу даних, ігноруючи їх; або заповнити пропуски, використовуючи певний спосіб; 2) простих чи складних методів імпутовання даних; 3) зважування; 4) математичного моделювання. У межах цього дослідження наведено не тільки класифікацію методів, але порівняно недоліки та переваги кожного з них.

Ключові слова: невідповідь, відсутні дані, методи аналізу невідповідей, модель.

Постановка наукової проблеми та її значення. Соціологія як наука об'єктом свого пізнання має соціальну реальність. Соціологічне дослідження цієї специфічної реальності складається з кількох етапів (підготовка дослідження, збір первинної інформації, обробка зібраних даних, аналіз).

Але часто трапляється так, що в момент збору первинної інформації частина респондентів відмовляється від участі в опитуванні у тій чи іншій формі взагалі або не відповідає на одне чи кілька запитань. І навіть після того, як первинна інформація зібрана, виникають помилки (під час перевірки анкет, кодування, уведення), які призводять до відсутності відповіді. У цих випадках дослідник зіштовхується з проблемою «невідповідей».

Можна виділити два загальних аспекти невідповідей: 1) запобігання цій проблемі до її виникнення; 2) методи оцінки, які спрямовані на врегулювання зміщення через невідповіді після того, як дані зібрано [9, 640].

Згадана проблема «невідповідей» виникла разом із першими опитуваннями. Проте осмислення відсутності відповіді починається лише з 1930-х років, а активне обговорення цієї проблеми в рамках статистики та соціології на шпальтах газет, журналів – із 1940-х років. При цьому статисти були більш зацікавлені способами пристосування до невідповідей, тоді як соціологи й методологи схильні зосереджуватися на тому, щоб зрозуміти та зменшити невідповіді [10, 28–30].

Аналіз досліджень цієї проблеми. Над вивченням методів аналізу невідповідей працювали Р. Грувс [6], Р. Літл та Д. Рубін [2], Д. Едіф, Т. Крезке, П. Лін [8] й інші. Що ж до вітчизняних науковців, то українські вчені мало уваги приділяють проблемі невідповідей у соціологічному дослідженні. У статті Ю. Сороки та Т. Зуб автори торкаються проблеми невідповідей у контексті розгляду відмов від відповіді як змістовної інформації [4, 222–236], публікація В. Снитюка присвячена одному з методів боротьби з відсутніми даними [3, 262–271].

Мета й завдання статті. Проблемою є суперечність між необхідністю використання методів аналізу невідповідей та недостатністю узагальненої інформації щодо таких методів. Саме тому ця робота присвячена систематизації методів аналізу невідповіді в соціологічних дослідженнях.

Виклад основного матеріалу та обґрунтування результатів дослідження. Під час вибору методу аналізу та інтерпретації результатів слід опиратися на знання механізму, що призводить до відсутності значень. Знаючи механізм породження пропусків, дослідник далі обирає спосіб аналізу даних із невідповідями. Тут можливі кілька варіантів:

- 1) відкидання невідповідей і проведення аналізу з подальшим їх ігноруванням;
- 2) заповнення пропусків, викликаних невідповідями, через застосування простих або складних методів імпутування даних;
- 3) заповнення відсутніх значень із застосування спеціально розроблених математичних методів аналізу неповних даних (зокрема зважування);
- 4) заміна відсутніх значень такими, які можна розрахувати методами, заснованими на математичному моделюванні [2, 14–15].

До першого варіанта відносять метод виключення некомплектних спостережень. Він полягає в тому, що за відсутності в деяких спостереженнях значень будь-яких змінних, здійснюється видалення таких некомплектних спостережень з аналізу та обробка даних без пропусків (без невідповідей). Отже, перевага методу виключення некомплектних спостережень полягає в тому, що він легко реалізується. Проте іноді цей метод призводить до серйозних зсувів і зазвичай не є дуже ефективним. Головний недолік такого підходу обумовлений утратою інформації¹. Тому цей метод рекомендують застосовувати у випадку, коли наявні велика вибірка й мала кількість невідповідей. Окрім цього, метод виключення некомплектних спостережень слід застосовувати, коли дані відсутні випадковим чином.

До другого варіанта належить метод заповнення пропусків, або імпутування [2].

Розглянемо спочатку методи із заповненням, які відносять до простих алгоритмів. Їх сутність полягає в тому, що пропущені значення вихідної вибірки заповнюються й отримані «повні» дані обробляють звичайними методами аналізу. До таких методів належать:

– заповнення з підбором – пропуски заповнюються значеннями, отриманими для іншого схожого об'єкта вибірки. Тут можуть використовуватися підстановка з підбором усередині груп (відбувається формування групи, пропуски в кожній групі заповнюються присутніми значеннями із цієї ж групи) або підбір ближнього сусіда (обраховують відстань між об'єктами в просторі інших

¹ У разі попарного видалення втрачається менше інформації, ніж при використанні видалення listwise.

змінних, далі вибирають підстановку об'єктом із присутнім значенням, ближчим до об'єкта з пропуском) [1, 52];

- заміна – метод обробки пропусків на етапі збору даних при дослідженні, що полягає в заміні об'єкта з відсутньою відповіддю на інший об'єкт, який не включений до вибірки. Наприклад, якщо домовласника опитати неможливо, то можна опитати його сусіда, не включеного до вибіркового списку опитуваних. Таку вибірку не можна вважати повною, оскільки той, хто дає відповіді, може систематично відрізнятися від того, кого не вдається опитати. Тому під час аналізу слід розглядати цю заміну як заповнення певного виду;

- заповнення без підбору. Пропуск заповнюють постійним значенням із зовнішнього джерела, наприклад значенням попереднього спостереження з цього ж дослідження. Як і при заміні, отримані дані прийнято розглядати як повну вибірку, тобто наслідки заповнення ігнорують;

- заповнення середніми. У цьому випадку підставляються середні із присутніх значень (вибір середнього повинен бути узгодженим із типом шкал, що використовується). Щодо цих двох методів Г. Толстова зазначає, що дослідник повинен розуміти ризик надання даним більш рівного характеру, ніж це є насправді [5, 138];

- заповнення за допомогою регресії. Пропуски замінюються значеннями, обчисленими за допомогою регресії, побудованої на наявних даних;

- стохастичне заповнення за регресією засноване на заміні пропуску значенням, що підставляється при заповненні за допомогою регресії, у сумі із залишком, що відображає невизначеність передбачуваного значення. При нормальній лінійній регресійній моделі звичайними є нормальні залишки з нульовим середнім і дисперсією, яка дорівнює залишковій дисперсії регресії. При бінарній змінній, як у логістичній регресії, передбачуване значення – ймовірність спостереження 1 або 0, а самі значення (1 або 0) вибираються з цієї ймовірністю [2, 61–62]. Обидва методи – заповнення за регресією й заповнення за стохастичною регресією – відносять до методів, заснованих на моделюванні.

Складні алгоритми розділяють на глобальні та локальні. У локальних алгоритмах в оцінюванні (прогнозі) кожного пропущеного значення беруть участь повні спостереження, що перебувають у деякій близькості прогнозованого об'єкта. До цієї групи належать алгоритми Zet і Zet Braid та методи багаторазового заповнення.

Суть алгоритму Zet полягає в підборі для кожного пропуску значення не з усієї сукупності повних спостережень, а з деякої її частини, яку називають компонентною матрицею, що складається з компонентних рядків і стовпців. Компонентність рядка або об'єкта є величиною, обернено пропорційною декартовій відстані до цільового рядка (неповного спостереження з пропуском) у просторі, осі якого задані змінними – розглянутими характеристиками об'єктів. За даними компонентної матриці будують функціональну залежність прогнозованого значення від відповідного значення в компонентній матриці, на основі якої прогнозують значення пропуску [3, 5–6].

Основна відмінність і перевага алгоритму ZetBraid над Zet полягає в тому, що в ньому закладено апарат для об'єктивного визначення розмірності компонентної матриці. У процесі роботи алгоритму відбувається послідовний почерговий відбір компонентних рядків і компонентних стовпців. При кожному новому відборі рядка або стовпця формується нова компонентна матриця. За заданим критерієм визначають її ефективність у прогнозі пропусків.

Методи багаторазового заповнення передбачають, що пропуск заповнюють декількома значеннями. Істотний недолік методів одноразового заповнення полягає в тому, що звичайні формули приводять для заповнених даних до систематично занижених оцінок дисперсії, навіть якщо правильна модель, що застосовується для обчислення значень. При багаторазовому заповненні отримують правильні оцінки дисперсії, які можна отримати звичайними методами аналізу [2, 74].

У глобальних алгоритмах в оцінюванні (прогнозі) кожного пропущеного значення беруть участь усі об'єкти цієї сукупності: метод Бартлетта, Resampling.

Метод Бартлета складається з двох етапів: підстановка, замість пропусків початкових значень на першому етапі й проведення на другому етапі коваріаційного аналізу цільової змінної та дихотомічного індикатора повноти спостереження за цільовою змінною. Індикатор повноти спостереження завжди дорівнює 0, за винятком одного-єдиного випадку, коли i -те значення цільової змінної пропущено, тільки в цьому випадку він приймає значення 1. Цей метод має низку переваг: 1) він не є ітеративним; 2) якщо структура пропусків має виродженість, цей метод «попереджає» дослідника, тоді як ітеративні методи приводять до недопустимих відповідей; 3) метод дає правильні оцінки й залишкові суми квадратів, точні стандартні помилки, суми квадратів і F-критерії. Незважаючи на ці переваги, його не можна часто застосовувати безпосередньо, адже спеціалізовані програми дисперсійного аналізу можуть не мати можливості вести обробку при багатьох супутніх змінних [1, 55–56].

Resampling метод застосовують для виконання завдання заповнення пропусків у неповних даних, коли значення для заповнення пропущених елементів вибирають випадковим чином із вихідної множини даних, де значення для заміни пропуску можна вибрати двома способами: із поверненням (раніше вибране значення може брати участь у заміні ще раз) і без повернення. Застосовують цей метод двома способами:

а) будують матрицю повних спостережень, для кожного пропуску вибирають із вибірки випадковим чином спостереження, яким заміщають пропущене значення. За даними, отриманими після такого заповнення, будують регресійну модель і знаходять оцінки коефіцієнтів $\hat{\beta}_i$ і вільного члена $\hat{\beta}_0$;

б) за наявними даними, без пропусків будують регресійну модель, заповнюють пропуски за нею, додаючи похибку моделі, обрану випадковим чином. За заповненою матрицею будують регресійну модель і знаходять оцінки коефіцієнтів $\hat{\beta}_i$ та вільного члена $\hat{\beta}_0$. Цю процедуру повторюють кілька разів і знаходять середні значення коефіцієнтів регресійної моделі. Ці значення і є результатом застосування resampling-методу.

Перевага на користь resampling-методу – повторне використання вихідних даних, адже збільшення числа підвибірок дає можливість найбільш повно й інформативно використовувати вихідну інформацію. З іншого боку, кількість нової інформації зменшується для кожної нової підвибірки, оскільки збільшується вірогідність того, що ці елементи вибірки були вже вибрані раніше – це основний недолік методу [1, 56–57].

Методи заповнення пропусків мають свої недоліки. Передусім, параметри для алгоритму заповнення пропусків обчислюють за наявними даними, що вносить залежність між спостереженнями. Звичайно, такої штучної залежності не виникає, якщо проводиться заповнення константою або випадковими значеннями, незалежними від присутніх спостережень у вибірці, або методом «підстановки без добору». Залежності можна також уникнути, розділяючи вихідну вибірку на дві підвибірки й обчислюючи підстановки (наприклад середнє вибіркове значення) для аналізованої підвибірки за значеннями спостережень у другій підвибірці. При такому підході жертвують частиною інформації, щоб заповнити пропущені значення.

По-друге, розподіл даних після заповнення відрізнятиметься від істинного, навіть якщо знехтувати залежністю, яка зазначена вище. Цей факт особливо наочний для простих методів заповнення – за середнім вибіркочним, за регресією тощо. Так, заповнення за вибіркочним середнім за наявними значеннями сформує розподіл у вигляді суміші, один компонент якої є справжнім розподілом за наявними значеннями, а іншими компонентами – розподіли, які відповідають некомплектним спостереженням із різними структурами пропусків.

До третього варіанта аналізу даних із невідповідями відносять такі методи зважування: 1) компенсація неправильної ймовірності вибору; 2) компенсація невідповідей; 3) створення зваженого розподілу вибірки для ключових змінних, що становлять інтерес.

Зважування здійснюють за такими кроками: 1) задають початкові ваги; 2) розділяють вибірку на підгрупи й обчислюють зважені рівні відповіді для кожної підгрупи; 3) використовують аналог рівня відповідей підгруп для регуляторів невідповідей; 4) розраховують ваги невідповіді [12, 8]. Зважування пов'язане із заповненням пропусків середніми. Наприклад, якщо заплановані ваги є

постійними в підгрупах вибірки, то заповнення пропусків у кожній із них середніми за підгрупами й зважування наявних значень за допомогою їх частки в кожній підгрупі ведуть до однакових оцінок середньої сукупності загалом, хоча оцінки вибіркової дисперсії різні, якщо тільки не використовують поправки на заповнення середніми.

Статистичну процедуру регулювання непокриття й невідповіді називають постстратифікацією. Вона зважає вибірку за певними змінними, щоб відповідати відомому розподілу населення. Норвезькі вчені [11] зазначають, що техніка постстратифікації особливо важлива й актуальна для аналізу результатів обстежень витрат домашніх господарств. Річ у тім, що в цих дослідженнях відповідей у цілому менше, ніж зазвичай. Крім того, у цих дослідженнях рівень відповідей варіюється в межах підкласів, які є однорідними стосовно витрат.

На практиці важливо правильно вибрати змінні для постстратифікації. Крім того, важливо не робити великої кількості постстрат, тому що користь від використання цього методу зменшується зі збільшенням їх кількості; занадто багато постстрат може призвести до того, що в деяких із них не буде респондентів.

Окремо слід виділити широкий клас методів, які ґрунтуються на побудові моделі породження пропусків. Висновки отримують за допомогою функції правдоподібності, побудованої за умови справедливості цієї моделі з оцінюванням параметрів методами типу максимальної правдоподібності.

У методах, що використовують функцію правдоподібності, реалізована ідея обробки неповних даних: заповнення пропусків оцінками пропущених значень; оцінювання параметрів; повторне оцінювання пропущених значень (оцінки параметрів вважаються точними); повторне оцінювання параметрів і так далі до збіжності процесу. Переваги такого підходу полягають у тому, що він гнучкий, дає змогу відмовитися від методів, розроблених для окремих випадків, та оцінювати в наближенні великої вибірки дисперсії оцінок за допомогою матриці других похідних функцій правдоподібності для неповних даних. Недолік алгоритму полягає в тому, що швидкість збіжності може бути дуже низькою, якщо пропущено багато даних [1, 54].

Із використанням максимальної правдоподібності виконується й пряма оцінка, для здійснення якої потрібне спеціалізоване програмне забезпечення. Неповні випадки включаються в аналіз, і якщо в якості методу оцінки використовується максимальна правдоподібність, мають на увазі, що дані відсутні випадково [7].

Метод максимальної правдоподібності дає можливість моделювати відсутні дані, ґрунтуючись на доступних даних і розглядаючи їх як репрезентативну вибірку деякого розподілу. Здебільшого цей метод є спробою створити модель, яка оптимізує імовірність знаходження відношень даних, що спостерігаються.

ЕМ-оцінювання (метод максимізації очікувань) дає можливість не тільки відновлювати пропущені значення з використанням двоетапного ітеративного алгоритму, а й оцінювати середні значення, коваріаційні та кореляційні матриці для кількісних змінних. ЕМ-алгоритм у найзагальнішому сенсі являє собою ітераційну процедуру, призначену для розв'язання завдань оптимізації деякого функціоналу через аналітичний пошук екстремуму цільової функції. Цей алгоритм реалізують у два етапи: на першому – Е (expectation), за сукупністю наявних абсолютно повних або частково (за цільовою змінною) повних спостережень розраховують умовні очікувані значення цільової змінної для кожного неповного спостереження. Потім, після отримання масиву повних спостережень, оцінюють основні статистичні параметри. У разі наявності неповних даних на Е-етапі визначають функцію умовного математичного сподівання логарифма повної функції правдоподібності при відомому значенні цільової змінної. Завдання алгоритму на другому етапі М (maximization) – максимізувати ступінь взаємної відповідності очікуваних і реальних даних, які підставлятимуться, а також відповідності структури імпутованих даних структурі даних повних спостережень [1].

Висновки та перспективи подальших досліджень. Отже, для успішної профілактики та «лікування» відсутніх даних потрібно визначити й зрозуміти причини виникнення невідповідей, особливо слід звернути увагу на випадковість чи не випадковість характеру відсутніх даних. Також варто врахувати, що методи аналізу невідповідей різняться механізмами їх реалізації та обумовлюються подальшим аналізом й інтерпретацією даних. Знаючи механізм породження пропусків, дослідник,

передусім, обирає спосіб аналізу даних із невідповідями, основними серед яких є такі: 1) відкидання невідповідей і здійснення аналізу з подальшим їх ігноруванням; 2) заповнення пропусків через застосування простих або складних методів імпутування даних; 3) заповнення відсутніх значень із застосування спеціально розроблених математичних методів аналізу неповних даних (зокрема зважування); 4) заміна відсутніх значень такими, які можна розрахувати методами, заснованими на математичному моделюванні.

Подальших напрацювань потребує кожна виокремлена група в напрямі виявлення не лише недоліків і переваг відповідного аналізу даних із невідповідями, а й деталізації кола завдань, до яких він застосовується.

Джерела та література

1. Злоба Е. Статистические методы восстановления пропущенных данных / Е. Злоба, И. Яцкив // Computer Modeling & New Technologies. – 2002. – Vol. 6. – № 1. – С. 51–61.
2. Литтл Р. Дж. А. Статистический анализ данных с пропусками. / Р. Дж. А. Литтл, Д. Б. Рубин // Статистический анализ данных с пропусками. – М. : Финансы и статистика, 1990. – 336 с.
3. Снитюк В. Е. Эволюционный метод восстановления пропусков в данных / В. Е. Снитюк // Интеллектуальный анализ информации : материалы междунар. конф. – Киев, 2006. – С. 262–271.
4. Сорока Ю. Г. «Отказ от богатства» и изменения в социальном восприятии неравенства / Ю. Г. Сорока, Т. С. Зуб // Соціологія: теорія, методи, маркетинг. – 2008. – № 1. – С. 222–236.
5. Толстова Ю. Н. Анализ социологических данных: методология, дескриптивная статистика, изучение связей между номинальными признаками / Ю. Н. Толстова. – М. : Науч. мир, 2000. – 352 с.
6. Groves R. Experiments in Producing Nonresponse Error / R. Groves, M. Couper, S. Presser, E. Singer, R. Tourangeau, G. Piani, L. Nelson // Public Opinion Quarterly. – 2006. – Vol. 70. – № 5. – P. 720–736.
7. Leeuw E. Prevention and Treatment of Item Nonresponse / E. Leeuw, J. Hox, M. Huisman // Journal of Official Statistics. – 2003. – Vol. 19. – № 2. – P. 153–176.
8. Lynn P. Extended field efforts to reduce the risk of non-response bias: do they pay off? / P. Lynn, J. Hall, V. Brown, G. Nicolaas // ISER working Paper Series, 2011. – № 24. – 31 p.
9. Marsden Peter V. Handbook of survey research. 2nd edition / P. V. Marsden, J. D. Wright. – Emerald Group Publishing, 2010. – 886 p.
10. Singer E. Nonresponse bias in household surveys / E. Singer // Public Opinion Quarterly. – 2006. – Vol. 70. – № 5. – P. 637–645.
11. Thomsen I. On the causes and effects of non-response: Norwegian Experiences / I. Thomsen, E. Siring. – Central Bureau of Statistics, Norway : Oslo. – 1980. – 52 p.
12. Yansaneh I. Construction and use of sample weights / I. Yansaneh // Draft Handbook on Designing of Household Sample Surveys. – 2003. – P. 1–14.

Сидоров Николай, Билоус Евгений. Подходы к анализу данных с неответами. Несоответствие между необходимостью анализа неответов и отсутствием обобщенной информации относительно природы возникновения разного рода отсутствующих данных актуализирует рассмотрение основных методов анализа неответов в эмпирических социологических исследованиях. Учитывая то, что методы анализа неответов разнятся механизмами их реализации и обуславливаются дальнейшим анализом и интерпретацией данных, исследователю предлагается определить способ анализа данных с неответами: либо 1) удалить неответы и реализовывать процедуры анализа данных, игнорируя их; либо заполнить пропуски, используя определённый способ; 2) простых или сложных методов импутирования данных; 3) взвешивания; 4) математического моделирования. В рамках этого исследования приведена не только классификация методов, но и сравнены недостатки и преимущества каждого.

Ключевые слова: неответ, пропущенные данные, методы анализа неответов, модель.

Sydorov Mykola, Bilous Yevhen. Approaches to the Analysis of Data with Nonresponse. The discrepancy between the need for analysis of non-response and a generalized information about the nature of a different kind of missing data are updated by the consideration of the basic methods of analysis of non-response in empirical social science research. Given that the non-response analysis methods have different mechanisms for their implementation, and are caused by the further analysis and interpretation of data, the researcher proposed to define a method for analyzing data from non-responses: either 1) remove the non-response and to implement procedures for data analysis, ignoring them, or fill in the blanks using a certain method; 2) simple or complex methods of data imputation;

3) weighing; 4) mathematical modeling. The study shows not only the classification methods, but compared the advantages and disadvantages of each.

Key words: nonresponse, missing data, methods of analysis of data with nonresponse, model.

УДК 316.334(477)

Марія Хомяк

Бідність в Україні у показниках соціальної статистики

Розглянуто основні концепції об'єктивного оцінювання бідності. Наведено кількість бідних в Україні відповідно до різних методик вимірювання. Зокрема, згідно з абсолютною концепцією бідності відповідно до міжнародних критеріїв, запропонованих ООН, 2,9 % українців вважаються бідними; за критерієм прожиткового мінімуму – 9,9 %; а за відносним критерієм бідних в Україні 26 %. Проте практичне використання об'єктивних оцінок бідності має певні недоліки, головний із яких – складність отримання своєчасної, повної та достовірної інформації щодо рівня доходів, витрат і наявних ресурсів домогосподарств.

Ключові слова: рівень бідності, абсолютна бідність, відносна бідність, споживчий кошик, межа бідності.

Постановка наукової проблеми та її значення. Подолання бідності – одна з найважливіших глобальних проблем сучасності. Вона і в глобальному, і в локальному вимірах є предметом наукового інтересу та жвавих дискусій і зарубіжних, і вітчизняних науковців протягом багатьох років, що свідчить про її надзвичайну актуальність. Однак, незважаючи на підвищену увагу вчених до розв'язання проблеми бідності в Україні, досить мало уваги приділяється обґрунтуванню підходів щодо об'єктивного (кількісного) оцінювання бідності.

Аналіз досліджень цієї проблеми. Проблематика бідності широко представлена у вітчизняній та зарубіжній науковій літературі. Варто також зазначити, що за часів існування СРСР проблема бідності не досліджувалася як така, адже ні бідності, ні суттєвої соціальної нерівності не існувало. Лише на початку 90-х років ХХ ст. аспекти цієї проблеми почали досліджувати вітчизняні економісти, соціологи й політологи. Серед них, зокрема, слід назвати С. Башкірова, В. Броницьку, О. Васильєва, О. Крикун, Е. Лібанову, М. Міщенко, С. Полякову, А. Реут, Н. Холод, С. Тютюнникову, Л. Шангіну, Л. Шевченко, В. Шишкіна, В. Юрчишина та ін. Можна відзначити також достатню увагу до проблеми бідності з боку Міжнародних організацій, а саме Організації Об'єднаних Націй, Світового банку, Всесвітньої організації охорони здоров'я, Міжнародної організації праці та ін.

Мета статті – обґрунтування підходів щодо кількісного оцінювання бідності, а також аналіз даних статистичних досліджень проблеми бідності в сучасній Україні.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Сучасна наука визначає бідність, як неможливість через брак коштів підтримувати спосіб життя, притаманний конкретному суспільству в конкретний період часу [1, 9]. Ідеться про те, що внаслідок матеріальних нестатків бідні верстви населення не можуть харчуватися відповідно до місцевих стандартів, не можуть оплачувати житло та комунальні послуги, виходячи зі своїх потреб; не можуть лікуватися й відпочивати, забезпечуючи відновлення втраченого через хворобу або перенавантаження здоров'я, зрештою, не можуть забезпечити оплату освіти собі та своїм дітям. Отже, маючи економічне походження, бідність – набагато ширше явище, що стосується різноманітних аспектів життя населення.

У дослідженні бідності важливу роль відіграє методика її кількісного оцінювання. Для комплексної оцінки бідності в соціальній статистиці використовують такі показники, як:

– межа крайньої форми бідності – вартісний поріг доходу (витрат), нижче якого задоволення основних потреб людини неможливе¹;

© Хомяк М., 2013

¹ Стандартом, що виступає базою для розрахунків, тобто межею бідності, можуть виступати медіанні витрати, середні, модальні, квартильні або децильні рівні доходу, а також фізіологічний та прожитковий мінімуми [1, 52–53; 10, 59–60].