

УДК : 303.4 (477)

Т.Д. МАТВІЙЧУК

**ЗАСТОСУВАННЯ ПРОПОРЦІЙНОЇ ВІДСОТКОВОЇ ВИБІРКИ ДЛЯ ПРОВЕДЕННЯ
ВСЕУКРАЇНСЬКОГО СОЦІОЛОГІЧНОГО ДОСЛІДЖЕННЯ****Матвійчук Тарас Дмитрович** – Львівський інститут економіки і туризму, старший викладач кафедри соціально-гуманітарних дисциплін, кандидат соціологічних наук

У статті запропоновано спосіб розрахунку пропорційної відсоткової вибірки з використанням даних щодо кількості населення в областях та районах України. Вибірка сформована на основі даних щодо чисельності населення на 1 вересня 2011 року та середньої чисельності за січень-серпень 2011 року, отриманих на сайті Державної служби статистики України (<http://www.ukrstat.gov.ua>).

Ключові слова: вибірка, спосіб розрахунку вибірки, генеральна сукупність, вибіркова сукупність.

This article proposes a way to calculate the proportionate interest of the sample using data on population in regions and in districts of Ukraine. The sample is formed on the basis of data on population by 1 September 2011 and the average number for January-August 2011, received at the site of the State Statistics Service of Ukraine (<http://www.ukrstat.gov.ua>).

Key words: sample, a method of sample calculating, a population, survey sampling.

© Т.Д. Матвійчук, 2012

Мета статті – запропонувати колегам-соціологам один із доступних і простих варіантів розрахунку вибіркової сукупності.

Постановка проблеми.

Якісний розрахунок вибіркової сукупності, яка б максимально репрезентативно відображала особливості та характеристики генеральної сукупності, є одним із найважливіших елементів більшості соціологічних досліджень. Запропонований у статті спосіб розрахунку вибірки є достатньо простим для застосування й одночасно повністю репрезентативним для більшості видів соціологічних опитувань, генеральною сукупністю яких є все населення України.

Аналіз останніх досліджень і публікацій, на які спирається автор. Уперше даний спосіб розрахунку вибіркової сукупності був застосований та апробований автором під час написання кандидатської дисертації на тему: «Соціальні чинники розбудови громадянського суспільства в Україні», практична частина якої включала опитування студентської молоді всіх вищих навчальних закладів Львівської області III-IV рівнів акредитації (з повним урахуванням кількісного складу кожного факультету, всіх навчальних закладів відповідного типу). Тому дана вибірка може бути корисною для більшості досліджень, генеральна сукупність яких може бути певним чином структурована.

Формулювання цілей статті.

Центральна ідея розрахунку такої вибірки полягає у відсоткових пропорціях, а саме – у застосуванні чіткої числової відповідності між генеральною та вибірковою сукупностями.

Такий спосіб розрахунку вибіркової сукупності є одним з найпростіших, якщо мова йде про струк-

туроване та пропорційне охоплення генеральної сукупності.

Виклад основного матеріалу.

Приклад розрахунку вибіркової сукупності представлено як систему послідовних кроків. Серед доступних статистичних даних можна обрати таблицю чисельності населення станом на якусь дату (у нашому випадку, 1 вересня 2011 року), або таблицю середньої чисельності за певний період.

Використання таблиці з даними про кількість населення на певну дату доцільно використовувати, якщо певне соціологічне дослідження стартує одразу після надходження даних та після оперативного розрахунку вибірки (для прикладу ми використовуємо саме такі дані). Таблиця, що містить дані про середню кількість населення за певний період може вважатися більш відповідною для розрахунку вибіркової сукупності, оскільки певною мірою відображає плинність населення (яка відбувається через різні причини), проте складно врахувати всі особливості та відповідності змін кількісного складу населення областей та районів. У будь-якому випадку «схопити» точну вибіркову сукупність відповідно до точної генеральної сукупності (яка є постійно частково змінною, особливо у масштабах країни) непросто.

У Таблиці 1 наведено кількісні дані чисельності населення в Україні відповідно до статистичних звітів (Таблиця 1, колонки 2, 3). Для спрощеного візуального сприйняття числа розділені крапками на мільйони, тисячі та сотні. У колонці 4 Таблиці 1 з допомогою операційної програми Microsoft Excel розраховано відсоткову частку Автономної Республіки Крим, кожної області, міста Києва та міста Севастополя по відношенню до кількості всього на-

селення України. У колонці 5 Таблиці 1 відсоткові значення населення вказаних міст та областей заокруглено (оскільки у статті представлено зразок розрахунків).

Саме тому сума заокруглених відсоткових значень становить 101 (через накопичення цифр, які знаходяться після коми). Для більш точного та репрезентативного представлення (що залежить від кожного окремого дослідження) можна враховувати і цифри після коми (тобто відсоткові значення не заокруглювати).

Таблиця 1

Розрахунок вибірки для всеукраїнського соціологічного дослідження відповідно до даних про чисельність населення на 1 вересня 2011 року¹

	Територія генеральної сукупності	Все населення	%	%	N
1	2	3	4	5	6
A	Україна	45.668.028			
A1	Автономна Республіка Крим	1.962.552	4,297431	4	16
	Області				
1	Вінницька	1.637.480	3,585616	4	16
2	Волинська	1.039.011	2,275139	2	8
3	Дніпропетровська	3.324.236	7,279132	7	28
4	Донецька	4.412.617	9,662377	10	40
5	Житомирська	1.276.471	2,795109	3	12
6	Закарпатська	1.249.369	2,735763	3	12
7	Запорізька	1.795.027	3,930599	4	16
8	Івано-Франківська	1.380.367	3,022611	3	12
9	Київська	1.719.159	3,76447	4	16
10	Кіровоградська	1.005.977	2,202804	2	8
11	Луганська	2.279.442	4,99133	5	20
12	Львівська	2.541.430	5,565009	6	24
13	Миколаївська	1.180.323	2,584572	3	12
14	Одеська	2.385.719	5,224046	5	20
15	Полтавська	1.481.744	3,244598	3	12
16	Рівненська	1.153.913	2,526741	3	12
17	Сумська	1.156.020	2,531355	3	12
18	Тернопільська	1.082.097	2,369485	2	8
19	Харківська	2.740.738	6,001437	6	24
20	Херсонська	1.085.615	2,377188	2	8
21	Хмельницька	1.324.687	2,900688	3	12
22	Черкаська	1.281.711	2,806583	3	12
23	Чернівецька	904.167	1,979869	2	8
24	Чернігівська	1.093.086	2,393548	2	8
B	м.Київ	2.794.352	6,118837	6	24
B1	Севастополь (міськрада)	380.718	0,833664	1	4

Враховуючи дані Таблиці 2, можемо обрати для цього гіпотетичного дослідження вибірку сукупність обсягом 400 осіб (що також може коригуватися відповідно до об'єкта і предмета окремих досліджень).

Таблиця 2

Залежність обсягу вибірки (n) від обсягу генеральної сукупності (N) за припустимої похибки 5% (довірча ймовірність – 0,954)²

N	n	N	n
100	80	2000	333
200	133	2500	353
300	171	4000	364
400	200	5000	370
500	222	10 000	385
600	240		
700	255	100 000	398
800	267		
900	277	Нескінчена	400
1000	286		
1500	316		

Розраховуємо відсоткову частку у вибірковій сукупності відповідно до відсоткової частки кожної області у генеральній сукупності і отримуємо кількість респондентів, яких необхідно опитати у кожній з областей (Курсив автора. Т. М.). Та прописуємо результати у колонці 6 Таблиці 1.

Сума респондентів, яких необхідно опитати, становить 404 особи, тобто, внаслідок розрахунку відсотків (які не завжди є цілими числами) відбувся перебір вибірки на чотири респонденти. Чотирима особами представлено, наприклад, місто Севастополь, тому, щоб у результатах узагальненого аналізу не було зсуву у відповідях, варто відняти чотири анкети від тих областей чи міст, у яких визначено опитати найбільші кількості респондентів. Такий, на перший погляд, значний перебір вибіркової сукупності внаслідок розрахунків відбувся внаслідок того, що генеральна сукупність обсягом 45.668.028 осіб (все населення України) представлена вибірковою сукупністю обсягом лише 400 респондентів. При наближенні обсягів генеральної та вибіркової сукупностей (зменшення значення «N» при збільшенні значення «n») перебір вибірки зменшується і може перетворитися на недобір. (Курсив автора. Т. М.). Такі варіації трапляються через заокруглення, якого не вдається уникнути хоча б тому, що кількість респондентів повинна бути прописана цілими числами. Аналогічний спосіб розрахунку вибірки можемо застосувати до районів кожної з областей.

Отже продовжимо на прикладі районів Львівської області. За 100% братиметься вже не кількість населення в країні, а кількість населення в області. З метою спрощення подальших робіт з масивом табличної інформації можемо одночасно кодувати назви районів кожної з областей (наприклад, Львівська область у Таблиці 1 під номером

«12», відповідно нумерація районів цієї області також може починатися з цифри «12»).

Таблиця 3

Приклад розрахунку вибірки для районів Львівської області³

№	Назви районів	Кількість населення (тис.)	%	N
1	2	3	4	5
12/1	Бродівський	63,5	4,297509	1
12/2	Буський	50,8	3,438008	1
12/3	Городоцький	74,2	5,021657	1
12/4	Дрогобицький	76,3	5,163779	1
12/5	Жидачівський	80,5	5,448024	1
12/6	Жовківський	108,9	7,37006	2
12/7	Золочівський	74,5	5,04196	1
12/8	Кам'янка-Бузький	61,9	4,189226	1
12/9	Миколаївський	66,3	4,487006	1
12/10	Мостиський	61,9	4,189226	1
12/11	Перемишлянський	47,7	3,228208	1
12/12	Пустомитівський	111,9	7,573091	2
12/13	Радехівський	52,5	3,553059	1
12/14	Самбірський	74,6	5,048728	1
12/15	Сколівський	50,2	3,397401	1
12/16	Сокальський	98,1	6,639145	2
12/17	Старосамбірський	82,2	5,563075	1
12/18	Стрийський	63,3	4,283974	1
12/19	Турківський	54,8	3,708717	1
12/20	Яворівський	123,5	8,358148	2

Оскільки на Інтернет-сторінці Державної служби статистики не вдалося знайти дані щодо кількості населення у районах областей, такі цифри отримано з окремого джерела, чисельність населення в якому подано на 5.12.2001 року⁴.

Якщо додати населення районів Львівської області у Таблиці 3, отримаємо цифру 1477,6 тис. осіб, додамо населення міста Львова (732,8 тис. осіб) і отримаємо цифру 2,2104 тис. осіб (хоча відповідно до даних з Таблиці 1 населення Львівської області становить 2,541430 тис. осіб).

Така невідповідність спричинена різними джерелами даних та особливостями врахування кількостей населення великих міст (Курсив автора Т.М.).

Проте з метою презентації способу розрахунку вибірки можемо Таблицю 3 розглядати окремо.

Отже, відповідно до Таблиці 1, у Львівській області необхідно опитати 24 респонденти. Тож знаходимо відсоткову частку в числі «24» відповідно до відсоткової частки населення районів області порівняно із усім населенням області.

В результаті обчислень отримаємо кількість респондентів, яких необхідно опитати у кожному районі області.

Висновки і перспективи подальших розвідок у даному напрямку.

Запропонований спосіб розрахунку вибіркової сукупності можемо використовувати практично у будь-якому соціологічному дослідженні, генеральну сукупність якого можна певним чином структурувати.

Цілком можливо, що такий спосіб розрахунків вибіркової сукупності потребує вдосконалення. Тому буду вдячний за продуктивну дискусію та розвиток теми розрахунку доступних та цікавих вибірок.

ПРИМІТКИ:

¹ За даними Інтернет-сторінки Державної служби статистики <http://www.ukrstat.gov.ua> (7 листопада 2011 року 14:00).

² Паніотто В. І., Максименко В. С., Харченко Н. М. Статистичний аналіз соціологічних даних // Володимир Паніотто, Валентин Максименко, Наталя Харченко. – К.: Вид. дім «КМ Академія», 2004. – 270 с., С. 181.

³ Джерело кількісних даних щодо населення районів в областях України: Україна. Атлас: Бібліотека нового українця. – Київ: МАПА. – 2007. – 56 с.

⁴ Джерело кількісних даних щодо населення районів в областях України: Україна. Атлас: Бібліотека нового українця. – Київ: МАПА. – 2007. – 56 с.

ЛІТЕРАТУРА:

1. Паніотто В. І., Максименко В. С., Харченко Н. М. Статистичний аналіз соціологічних даних // Володимир Паніотто, Валентин Максименко, Наталя Харченко. – К.: Вид. дім «КМ Академія», 2004. – 270 с., С. 181.

2. Інтернет-сторінка Державної служби статистики <http://www.ukrstat.gov.ua>

3. Україна. Атлас: Бібліотека нового українця. – Київ: МАПА. – 2007. – 56 с.