

УДК53-7:519.25

ПІДВИЩЕННЯ ДОСЯЖНОСТІ У МАРКЕТИНГОВИХ І СОЦІОЛОГІЧНИХ ДОСЛІДЖЕННЯХ

МИКОЛА ЧУРИЛОВ,
доктор соціологічних наук, генеральний директор «TNS UKRAINE»

У багатьох випадках, коли говорять про досяжність, дослідники не усвідомлюють, а для чого ж вона вимірюється і як її використовувати. Проблема аналізу досяжності ніколи не була предметом спеціального вивчення, і автори, які описують процес збирання інформації в маркетингових і соціологічних дослідженнях, як правило, не звертають уваги на цей бік питання. У пропонованій статті ми спробуємо описати проблему та узагальнити досвід як соціологічних, так і маркетингових компаній, які працюють в Україні.

Досяжність, слід зазначити, відіграє досить помітну роль лише у разі проектування випадкових, а не спрямованих вибірок. І це відбувається на етапі розрахунку обсягу проектованої вибіркової сукупності. Проаналізовані в безлічі підручниках зі статистики формули для розрахунку обсягу випадкової вибірки справедливі не для всіх її типів, а тільки для одноступінчатих імовірнісних вибірок.

Отже, обчислимо теоретичний обсяг вибіркової сукупності n_{eff} , який, за умов його повної реалізації, забезпечить плановану точність – випадкову помилку вибірки. На практиці все проходить не зовсім так, як намічено за планом. Реально отримувана у ході дослідження вибірка m зазвичай менша за обсягом, тому що частина респондентів, які мають брати участь в опитуванні, з різних причин відмовляються відповідати або брати участь в опитуванні, або відсутні, а отже, були замінені іншими респондентами.

Оскільки переважна більшість проектованих дослідниками вибірок не одноступінчаті, то на кожному наступному рівні також з'являються помилки у вибірці (враховуємо лише випадкові помилки вибірки), що спричинює зменшення точності кінцевих оцінок вибірки. Таке зменшення називають дизайн-ефект вибірки (або $DEFF$) і вимірюють за формулою

$$DEFF = \frac{\text{реальний обсяг вибіркової сукупності}}{\text{розрахований обсяг імовірнісної вибірки}} = \frac{m}{n_{eff}} \quad (1)$$

Таким чином,

$$n_{eff} = \frac{m}{DEFF} \quad (2)$$

Дослідник повинен мати можливість прогнозувати значення $DEFF$ для пропонованого плану вибірки з метою встановлення, скільки інтерв'ю необхідно провести, щоб досягти певного заданого значення n_{eff} . Наприклад, два компоненти $DEFF$ потрібно враховувати на етапі проектування дизайн-ефекту вибірки, що виникає на всіх подальших рівнях відбору. $DEFF_p$ – дизайн ефект, який виникає на проміжних ступіннях вибірки і дизайн-ефект від класифікації – кількість одиниць спостереження, що реалізуються у кожній точці опитування на останньому ступені відбору $DEFF_c$. Отже, матимемо такий сумарний дизайн-ефект:

$$DEFF = DEFF_p \cdot DEFF_c \quad (3)$$

Далі дослідник зобов'язаний на основі наявних даних передбачити рівень досяжності (частка недоступних респондентів у теоретичній вибірці) з метою визначити розмір базової вибірки n , необхідної для того, щоб взяти приблизно m інтерв'ю.

Наприклад, ми оцінили дизайн-ефект на кожному проміжному рівні вибірки, а також визначили дизайн-ефект на останньому рівні після кластеризації. Таким чином, сумарне значення дизайн-ефекту після всіх наших обчислень дорівнюватиме $DEFF = DEFF_{p1} \times DEFF_{p2} \times DEFF_c = 1,68$ (невисоке значення дизайн-ефекту). Це означатиме, що якщо ми проектуємо вибірку розміром 1500 одиниць спостереження, то реально вона повинна становити 2520 респондентів.

Наприклад, ми очікуємо рівень досяжності на рівні 70%, тоді

$$m = \frac{2520}{0,7} = 3600 \quad (4)$$

Звідси витікає, щоб реалізувати теоретично розраховану на 1500 респондентів вибірку, яку відшукуємо за допомогою багатоступеневої вибірки, ми повинні звернутися не до 1500 респондентів, а не менше як до 3600. Це за умови, що ми не враховували дизайн-ефект, що утворюється на останньому, четвертому, ступені, при відборі респондентів всередині визначених для проведення опитування домогосподарств.

Таким чином, некоректно розраховувати випадкову помилку вибірки виходячи з обсягу вибірки 3600 одиниць спостереження. Дослідник має перш за все орієнтуватися на теоретично розраховану вибірку в 1500 респондентів. На жаль, на практиці трапляється інакше.

Так, наприклад, дослідник реалізував вибірку на 1500 респондентів, і з цієї вибіркою обчислює випадкову помилку вибірки. Однак теоретичний обсяг вибірки буде при цьому значно меншим. Припустимо, що була реалізована багатоступенева вибірка, і досяжність при цьому була достатньо великою – на рівні 70%. Припустимо, що сумарний дизайн-ефект на проміжних і останньому рівні відбору так само дорівнював 1,68. Звідси матимемо

$$n_{eff} = \frac{m}{DEFF} = \frac{1500 \times 0,7}{1,68} = 625 \quad (5)$$

Отже, теоретичний обсяг вибіркової сукупності становитиме не 1500 (за якою, як правило, дослідники розраховують випадкову вибірку), а лише 625 одиниць спостереження. У цьому прикладі випадкова помилка вибірки становитиме приблизно 4%, а не 2,6, як для вибірки у 1500 респондентів.

Цілком очевидно, що дослідники повинні прагнути до того, аби зменшити дизайн-ефект вибірки, тому що опитування в понад вдвічі більшої вибірки потребує майже триразового збільшення початкового бюджету і строків проведення дослідження. Постає слушне запитання, за рахунок чого можна зменшити дизайн-ефект вибірки? По-перше, скоротити кількість ступенів відбору і знизити розмір кластера – навантаження на точку опитування, а по-друге, збільшити досяжність респондентів у процесі відбору.

Як для випадкових, так і спрямованих вибірок існує значна різниця між теоретично розрахованою вибіркою (вибіркова сукупність, яку розробляє дослідник до початку польового етапу дослідження) і реальною вибіркою (вибіркова сукупність, яка реалізована інтерв'юерами в процесі збирання інформації). Це пояснюється тим, що частина респондентів виявляються недосяжними для інтерв'юерів з огляду

на низку причин суб'єктивного й об'єктивного характеру. До суб'єктивних належать причини, що повністю залежать від учасників процесу спілкування – респондента та інтерв'юера:

- ◆ невмотивована відмова респондента від опитування;
- ◆ невміння інтерв'юера зацікавити респондента брати участь в опитуванні;
- ◆ стан респондента, коли починати опитування не має жодного сенсу;
- ◆ низька виконавча дисципліна інтерв'юера, що виявляє себе у відсутності бажання повторно або втретє відвідувати респондента, який з різних причин не зміг у перший прихід інтерв'юера взяти участь в опитуванні.

І цей перелік можна продовжити.

Однак існують й об'єктивні причини недосяжності респондентів, до яких належать такі:

- 1) відсутність респондента (з поважних причин) під час проведення опитування (відрадження, хвороба, відпустка, посилення на зайнятість тощо);
- 2) помилки в зазначеній адресі респондента (відсутність записаної в завданні інтерв'юеру вулиці, будинку або квартири, неправильно написана адреса, наявність даного респондента за вказаною адресою та ін.);
- 3) відсутність будь-кого в помешканні в момент відвідування інтерв'юера;
- 4) немає можливості потрапити до помешкання через консьєржку в під'їзді або дверей з вмонтованими домофонами і кодовими замками;
- 5) недосяжність респондента через те, що:
 - а) респондент проживає в так званому елітному будинку або закритих селищах, які охороняються, і допуск куди обмежений для сторонніх осіб;
 - б) дослідники не знають фактичної адреси проживання тих чи тих категорій необхідних для опитування респондентів;
 - в) дані категорії респондентів надто малі за кількістю, що вільний пошук їх інтерв'юерами не потрібний.

Саме остання причина характерна для багатьох важкодосяжним груп респондентів. Перш за все до цих груп належать так звані елітні респонденти: політична, наукова і бізнес-еліта та ін. Як правило, ці групи респондентів надзвичайно зайняті, погано йдуть на контакт з інтерв'юерами, важкодосяжні (живуть в елітних будинках і замських селищах, їхні адреси і номери телефонів відсутні у загальнодоступних довідниках). Все це ускладнює контакти з експертами вищого рівня та елітою.

Однак не менше труднощів, а може, і більше, становить пошук респондентів, що належать до так зва-

них груп ризику: ВІЛ-інфіковані, нарко- і алкозалежні, повії, бомжі та ін. Також вкрай складно проводити маркетингові дослідження тих груп респондентів, які користуються рідкісними або елітними товарами чи послугами.

У ході опитування представників бізнес-, політичної і наукової еліти, топ-менеджменту компаній та ін. дослідницькі компанії використовують в основному офіційні дані про місцезнаходження підприємств, організацій, фірм і товариств. Однак у цьому разі дослідники стикаються з низкою факторів, які суттєво ускладнюють успішне виконання роботи інтерв'юєрів. До таких чинників належать помилки в адресній базі. У багатьох випадках і адреси, і номери телефонів, що містяться в цій базі, не відповідають дійсності, тому місце юридичної реєстрації підприємства та організацій й їхнє фактичне і справжнє місце знаходження переважно не збігається.

Топ-менеджмент підприємств, організацій і установ не схильний йти на контакт з інтерв'юєрами та інформувати їх про справи своїх організацій, тому що не завжди чітко розуміють мету досліджень і, що найголовніше, які змістові висновки з результатів дослідження будуть зроблені. Це стосується тих випадків, коли в запитальнику містяться питання, які пов'язані з показниками економічної та фінансової діяльності цих підприємств, організацій або установ.

Не можна не зважати і на високу зайнятість топ-менеджменту цих організацій, що також ускладнює можливість їх контактів з інтерв'юєрами.

Ці та дія низки інших чинників призводять до того, що в дослідженнях, які проводяться на підприємствах, установах і в організаціях серед топ-менеджерів досяжність рідко перевищує 5-10%.

У цьому випадку дослідницьким компаніям доводиться різко змінювати стратегію пошуку та опитування так званих елітних респондентів. Крім традиційного пошуку респондентів методом «сніжної грудки» ряд компаній використовують власні нововведення. Так, наприклад, опитування елітних респондентів відбувається на елітній мийці машин, де клієнт приблизно протягом півгодини гає свій час в кімнаті очікування. За попередньою домовленістю з власником автомийок (на взаємовигідних для дослідників і господарів мийок умовах) інтерв'юєри працюють з клієнтами, пропонуючи їм відповіді на запитання, із пропозицією, що миття їхнього авто буде безкоштовним. Зазвичай така схема спрацьовує у 60-70-ти випадках зі ста.

Коли проводиться опитування важкодосяжних респондентів з груп ризику, то використовують дещо інші підходи у відборі та опитуванні. У деяких євро-

пейських країнах разом з традиційними методами пошуку респондентів (спосіб «сніжної грудки» або опитування респондентів у місцях їх традиційного перебування) застосовується й інша методика. До початку масових досліджень серед груп ризику проводиться ґрунтовна підготовча робота з пошуку і підготовки інтерв'юєрів із середовища, яке планується вивчати. І як показує практика, у такий спосіб вдається опитати від 30 до 40 процентів осіб із загальної кількості респондентів, опитаних у даних дослідженнях.

Хотілося б відповісти на одне досить просте запитання: для чого потрібно розраховувати досяжність у дослідженні? Коли йдеться про випадкові вибірки, відповідь уже дана вище – для того, аби оцінити обсяг теоретично розрахованої вибірки (якщо до початку дослідження не оцінювався дизайн-ефект і рівень досяжності у процесі опитування), за якою і визначається випадкова помилка вибірки. Для спрямованих (невипадкових) вибірок досяжність свідчить про ефективність обраної стратегії пошуку респондентів для опитування. Якщо досяжність низька, отже, стратегія пошуку респондентів була вибрана неправильно і наступного разу її варто переглянути. Іншого її застосування я не бачу.

На мій погляд, слід зробити ще одне зауваження щодо обсягу вибіркової сукупності для спрямованих вибірок. Вони, як правило, не репрезентативні, тому що в переважній більшості маркетингових досліджень (коли йдеться про вивчення поглядів, ставлення, цінностей і настанов споживачів стосовно товарів і послуг) нічого не відомо про генеральну сукупність (точніше, немає статистичного опису об'єкта дослідження). Дослідник вибудовує стратегію вибірки за принципом «методу основного масиву», виходячи із припущень на зразок «мені так здається», «це найбільш доцільно», «мій досвід підказує» та ін. Встановлюючи обсяг вибіркової сукупності, дослідник виходить з того, що аналізовані їм масиви і підмасиви були статистично наповненими. І ще раз наголошую, помилки вибірки (ні випадкові, ні спрямовані) в даному випадку не розраховуються.

На закінчення хотілося б зупинитись на такому. У ряді опитувань непрофесійні дослідники намагаються поєднати непоєднане – схеми випадкового відбору респондентів чи підприємств зі схемами квотного добору цих об'єктів. Наш досвід показує, що досяжність у цьому випадку становитиме вже не відсотки, а лише частки відсотків. Так, наприклад, в одному з експериментів, проведених в нашій компанії, щоб знайти десять респондентів із заданими параметрами (стать, вік і освіта) інтерв'юєр повинен був відвідати 891 випадково відібрану адресу. У цьому разі досяжність становила лише 0,11%.