

Ігор Ігорович Дзеверін,
доктор біологічних наук,
завідувач відділу еволюційної морфології Інституту зоології імені І. І.
Шмальгаузена НАН України,
професор кафедри теорії ймовірностей, статистики та актуарної математики
Київського національного університету імені Тараса Шевченка

У мене часто запитують, для чого потрібна статистика. Іноді з щирим нерозумінням, іноді з іронічними коментарями на кшталт “є три стадії брехні: просто брехня, нахабна брехня і статистика”. Часто згадують у цьому контексті випадки маніпулювання статистичними даними і спотворення даних, наприклад горезвісний “гребінець Чурова” (попри те, що цей “гребінець” було сфабриковано не статистиками, навпаки, саме праця статистиків-професіоналів уможливила виявлення та доведення фальсифікації). Навіть у науковців часто трапляється ставлення до статистичних методів як до чогось необов’язкового, такого, що нібито покликано підтвердити висновки, які і без цього є очевидними.

Частково це пов’язане з контрінтуїтивністю багатьох статистичних методів для тих, хто не вивчав спеціально статистику, або вивчав у надто обмеженому обсязі. Якщо, наприклад, мета і спосіб визначення середнього значення зазвичай проблем не викликає, то вже фізичний сенс і методику обчислення стандартного відхилення та стандартної похибки пояснити набагато складніше. Чому при обчисленні дисперсії треба ділити суму квадратів відхилень не на N , а на $(N-1)$? Чому в дослідженні, наприклад, прибутків населення (або, хоча б даних по зарплатні у власній установі) краще застосовувати медіану, ніж середнє? Чому наявність кореляції у загальному випадку не свідчить про причинно-наслідковий зв’язок, і, якщо це справді так, то навіщо ця кореляція взагалі потрібна? Чому некоректно застосовувати критерій Стюдента до випадкових величин із дискретним варіюванням? І так далі. Нерозуміння теорії, що лежить в основі методів, викликає сумніви у самих методах.

Якщо ж говорити не про обґрунтування та розуміння статистичних методів, а про мету їх застосування у науковому дослідженні, то вона є, як мінімум, потрібною. По-перше, є суто технічні підстави для застосування статистики. Наукове пізнання ґрунтується на індуктивному методі, який у загальному випадку не гарантує нам правильного результату. Відтак ми не можемо остаточно ні підтвердити, ні спростувати наші гіпотези. Якщо ми завжди спостерігали певну послідовність якихось явищ, то це не означає, що в майбутньому вона не зміниться, і ми надалі її спостерігатимемо. Те, що якесь явище постійно відбувається після якогось явища, не означає, що воно є його наслідком.

Втім, розуміння цих обмежень не заважає науковцям працювати. Я думаю, більшість науковців погодиться з тим, що, так, дійсно, індукція не гарантує правильного висновку, але, якщо якась ситуація часто повторюється, то імовірність того, що вона не є випадковим збігом обставин, збільшується. Якщо якісь події часто відбуваються разом, то це може свідчити про зв’язок між ними. Тут принциповим є питання, як розуміти цей прислівник “часто”? Наскільки часто події мають траплятися разом, щоб ми припустили, що це щось більше, ніж випадковий збіг? Якщо явища добре нам відомі, то інтуїтивно ми легко проведемо ту межу, після якої зможемо говорити про наявність зв’язку. А що робити, коли ми мало що знаємо про ці явища, і нам саме і треба їх дослідити?

Саме в такій ситуації статистика стає нам потрібною, бо статистичні тести, що ґрунтуються на теорії ймовірностей, дозволяють нам визначити з максимально можливою (для того обсягу інформації, котрим ми володіємо) точністю, де щонайшвидше був просто випадковий збіг обставин, а де можна припускати наявність певних чинників, зв'язків та взаємодій. Про це дуже влучно написав колись один з найвидатніших класиків теорії ймовірностей П. С. Лаплас: *“la théorie des probabilités n'est au fond, que le bon sens réduit au calcul : elle fait apprécier avec exactitude ce que les esprits justes sentent par une sorte d'instinct, sans qu'ils puissent souvent s'en rendre compte”* [“теорія ймовірностей – це по суті не що інше, як здоровий глузд, зведений до обчислення: вона змушує оцінювати з точністю те, що гострі розуми відчують неначе інстинктом, часто не усвідомлюючи цього”].

По-друге, теорія ймовірностей та статистика корисні там, де треба описувати дуже складні процеси та явища. Об'єкт наших досліджень може бути цілковито детерміністичним, усі можливі його стани можуть однозначно визначатися взаємодіями його елементів та зовнішніми впливами, але, якщо елементів дуже багато, взаємодії між ними дуже складні, і нам не все відомо як про елементи, так і про взаємодії, то нам краще застосовувати при описі такого об'єкта стохастичні підходи, розглядати спектр можливих станів як варіювання певної випадкової величини, що залежить від дії та взаємодії багатьох чинників. Можливо, пізніше ми розберемося в деталях і зрозуміємо детерміністичний механізм, але на початку (а, можливо, і не тільки на початку) нас влаштовує приблизна стохастична модель.

Нарешті, по-третє, є й принципова стохастичність природних процесів та явищ природи, унаслідок чого суто детерміністичний опис цих процесів не є вичерпним. Це стохастичність квантово-механічної природи. Наскільки відомо сучасній фізиці, в основі квантової механіки не лежить жодний детерміністичний процес. Стохастичність нашого світу на квантово-механічному рівні первинна, однак у комплексах з великої кількості квантово-механічних об'єктів (якими і є тіла макросвіту) стохастичні відхилення нівелюються, і зазвичай ми можемо вивчати макрооб'єкти як детерміністичні, без урахування квантово-механічної стохастичності.

Зазвичай, але не завжди. Іноді стохастичність квантово-механічного походження нагадує про себе і в нашому макросвіті. Прикладом цього є явища живої природи.

Досить складно визначити, що таке життя, і чим живі істоти відрізняються від об'єктів неживої природи. Шкільні підручники та посібники найчастіше дають переліки характерних рис живих об'єктів, деякі з цих визначень мають винятки, деякі властиві не тільки живим істотам. Серед таких рис згадують розмноження, обмін речовин, подразливість, рух, ріст, розвиток, саморегуляцію і багато іншого. Я міг би внести дуже багато пунктів до цього переліку, але не буду, бо не дерзаю змагатися з підручником.

Насправді принциповою особливістю живих істот є здатність до активного самопідтримування та самовідтворення, яке відбувається з використанням вільної енергії середовища. Забезпечується ця здатність наявністю програми для побудови цієї істоти. Будь-який організм – це, по суті, кібернетична система, яка окрім двигунів, енергосистем, систем захисту та орієнтації, бортового комп'ютера (коротше кажучи, фенотипу) містить іще програму для утворення усього переліченого (генотип). Цю програму записано за допомогою дискретних сигналів на певних носіях, отже програми, створені людиною, дуже схожі на неї – не складністю та досконалістю (до цього нашої техніці ще як до неба), але принципами запису та виконання. Далеко не кожна фізична чи хімічна система може бути носієм інформації. Головні умови тут – надійність збереження та ефективність функціонування. Системи з малої кількості молекул не годяться для цього: вони нестійкі через

теплові рухи; через це, на що колись звернув увагу О. А. Ляпунов, носій інформації може бути або макроскопічним, або, навпаки, мономолекулярним. У останньому випадку інформація кодується станами окремих молекул. Саме такий шлях збереження і реалізації генетичної інформації властивий живим істотам.

Але під час відтворення інформації завжди можливі випадкові помилки, і, якщо інформацію закодовано на мономолекулярному рівні, то помилки ці виникатимуть на рівні частин молекули, тобто на рівні, де вже не можна нехтувати квантово-механічними ефектами. Кожна мутація (тобто, кожна помилка у генетичній програмі) є через це принципово випадковою і непередбачуваною. Фенотип – це макроскопічна система, його розміри на багато порядків більші, ніж розміри генотипу. Відповідно, під час побудови фенотипу на основі генотипу, наслідки мутації також зростуть багаторазово, і наслідки одиначної події, що відбулася на квантовому рівні, тепер впливатимуть на світ на макрорівні.

Ця особливість живих істот (так званий принцип підсилювача, сформульований свого часу М. В. Тимофєєвим-Ресовським) зазвичай не потрапляє до списків універсальних рис життя в підручниках. А даремно, бо цей принцип є універсальнішим, ніж майже всі такі риси, більше того, є для них важливою передумовою. Б. М. Медников цілком справедливо відніс принцип підсилювача до числа аксіом біології – найзагальніших положень, що становлять фундамент науки про живу природу. Завдяки принципу підсилювача квантово-механічна стохастичність стає помітною на макрорівні, принаймні, якщо йдеться про живу природу.

А, оскільки людина також є частиною живої природи, то все сказане повною мірою стосується і нашого життя, і нашої культури. Ми живемо в принципово непередбачуваному, загадковому світі, світі випадкових подій та явищ. Тому розуміння законів випадковості, якого ми досягли, і яке систематизували як теорію ймовірностей та статистику, – це для нас ніби компас у нашому непевному світі.