ПО 10. Теорія ймовірностей
Домашнє завдання 1

Данило Тавров, Антон Цибульник

15 вересня 2025 р.

1. Організація посвяти студентів

У студентській раді, що складається з 8 дівчат та 6 хлопців, потрібно утворити тимчасову групу для організації посвяти першокурсників. Група має бути гендерно збалансованою, а саме складатися з 3 хлопців та 3 дівчат. Порахуйте загальну кількість груп, які можна утворити, за різних умов.

(а) У колективі двоє хлопців-базік постійно розмовляють між собою та заважають іншим, тож задля ефективности робочого процесу в робочу групу варто включати лише одного з них або взагалі не включати нікого з цієї пари.

(б) Дві дівчини в студраді мають невирішену фінансову суперечку, через яку спільна праця в групі призведе до надлишкового обговорення грошового питання. Тому доцільно обрати або одну з них, або жодну.

(в) Хлопець та дівчина, які раніше перебували в стосунках, не воліли б опинитися в одній групі, тому варто обрати або одного з цієї пари, або нікого.

(г) Двоє хлопців зі студради навчаються в одній академічній групі та мають бездоганний рівень співпраці, а отже якщо й обирати, то їх варто обрати в спільну групу лише разом.

2. Бутстреппінг

Нехай маємо вектор дійсних чисел $\mathbf{p} = (p_1, p_2, \ldots, p_n)^\top \in \mathbb{R}^n$, у якому всі координати різні. У статистиці *бутстреп-вибіркою* (bootstrap sample) називають вектор $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^n$, утворений процедурою випадкового вибору елементів p_i, $i = 1, \ldots, n$, *з повтореннями* і з *однаковими ймовірностями*. Наприклад, якщо $\mathbf{p} = (3, 1, 6)^\top$, то бутстреп-вибірка може містити такі вектори: $(1, 1, 3)^\top$, $(3, 6, 1)^\top$, $(6, 1, 1)^\top$ і т.п.

Бутстреппінг (bootstrapping) широко застосовують у статистиці для оцінювання характеристик розподілів чи тестування відповідних статистичних гіпотез, коли масове повторення експерименту коштовне або неможливе. Своє застосування процедура також має в машинному навчанні під назвою *бутстреп-агрегування* або *беґґінгу* (bootstrap aggregating, bagging), покликаному зменшити перенавчання моделей.

(а) Скільки всього можна згенерувати бутстреп-вибірок на основі вектора $\mathbf{p} = (p_1, p_2, \ldots, p_n)^\top \in \mathbb{R}^n$, якщо порядок елементів у вибірці має значення?

(б) Суть бутстреппінгу полягає в генеруванні значної кількости бутстреп-вибірок та обчислення для кожної з них деякого підсумкового значення (наприклад, середнього значення, медіани тощо). Масив одержаних підсумкових значень далі аналізують за допомогою статистичних методів. У цьому сенсі *порядок* слідування значень у вибірці *значення не має*, і бутстреп-вибірку можна розглядати не як вектор, а як *мультимножину* $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, тобто множину, елементи якої можуть повторюватися, але порядок слідування яких ролі не грає.

Скільки всього можна згенерувати бутстреп-вибірок на основі вектора $\mathbf{p} = (p_1, p_2, \ldots, p_n)^\top \in \mathbb{R}^n$, *якщо порядок слідування значень у вибірці не грає жодної ролі*?

(в) Незважаючи на те, що результати генерування бутстреп-вибірки *зі збереженням порядку* є рівноймовірні, вибірки *без збереження порядку* можуть мати різні ймовірності бути згенерованими. Наприклад, якщо $\mathbf{p} = (1,2)^\top$, то всі можливі вибірки зі збереженням порядку мають однакову ймовірність:

$$\mathbb{P}\left((1,1)^\top\right) = \mathbb{P}\left((1,2)^\top\right) = \mathbb{P}\left((2,1)^\top\right) = \mathbb{P}\left((2,2)^\top\right) = \frac{1}{4}\,,$$

але вибірки без збереження порядку мають різні ймовірності:

$$\mathbb{P}\left(\{1,1\}\right) = \mathbb{P}\left(\{2,2\}\right) = \frac{1}{4}\,, \qquad \mathbb{P}\left(\{1,2\}\right) = \mathbb{P}\left((1,2)^\top\right) + \mathbb{P}\left((2,1)^\top\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}\,.$$

Чому дорівнює ймовірність згенерувати деяку конкретну вибірку $\mathbf{x} = \{x_1, \ldots, x_n\}$ на основі вектора $\mathbf{p} = (p_1, p_2, \ldots, p_n)^\top \in \mathbb{R}^n$?

(г) Яка бутстреп-вибірка (чи вибірки) $\mathbf{b}_1$ має ***найбільшу ймовірність*** бути згенерованою? Яка бутстреп-вибірка (чи вибірки) $\mathbf{b}_2$ має ***найменшу ймовірність*** бути згенерованою? Чому дорівнює $\mathbb{P}(\mathbf{b}_1)/\mathbb{P}(\mathbf{b}_2)$?

Наприклад, якщо $\mathbf{p} = (1,2)^\top$, то вибірка $\mathbf{b}_1 = \{1,2\}$ має найбільшу ймовірність, а саме $\frac{1}{2}$, а вибірки $\mathbf{b}_2 = \{1,1\}$ чи $\mathbf{b}_2 = \{2,2\}$ — найменшу ймовірність, а саме $\frac{1}{4}$.

Відповідь на це питання тісно пов'язана з поняттям мультиномного розподілу, у якому всі ймовірності однакові.

(д) Оскільки бутстреп-вибірка на основі вектора $\mathbf{p} = (p_1, p_2, \ldots, p_n)^\top \in \mathbb{R}^n$ генерується процедурою вибору елементів *із повторенням*, виникає питання, яку частку елементів вектора $\mathbf{p}$ буде відібрано в деяку бутстреп-вибірку? (Тобто з якою ймовірністю буде відібрано деяке p_i.) Що відбуватиметься з цією ймовірністю, коли розмір вектора $\mathbf{p}$ стає «достатньо великим» (тобто прямує до нескінченности)?

3. Комбінаторна тотожність

Розгляньмо комбінаторну тотожність, яка полягає у виборі трьох студентів серед потоку другокурсників для призначення бонусного питання на іспиті з дисципліни «Теорія ймовірностей». Отже, на потоці маємо групи КМ-41, КМ-42 та КМ-43. Нехай кожна група складається з n студентів, а отже всього студентів на потоці — $3n$ осіб.

(а) Скільки всього існує можливих варіантів вибору 3 студентів із потоку?

(б) Поміркуйте, які існують принципово різні конфігурації вибору трьох студентів із погляду представництва різних груп? *Підказка*: з'ясуйте, з яких саме груп і в якій кількості може бути обрано студентів.

(в) Підрахуйте кількості варіантів у кожній конфігурації окремо та підсумуйте отримані результати.

(г) Використовуючи відповіді пунктів (а) та (в), запишіть комбінаторну тотожність.

4. Парі на влучність

Для вирішення суперечки друзі уклали парі, суть якого полягає в підкиданні п'ятигривневої монети з відстани метра на стіл, розкреслений невеликими квадратами. Перемагає в парі той, чия монета влучила у квадрат, тобто *повністю* лежить у його межах. Задано, що п'ятигривнева монета має діаметр 2 см, а розкреслені квадрати на столі мають сторону 3 см кожен. З'ясуйте, яка ймовірність виграшу в цьому парі. Вважайте, що промахнутися і взагалі не влучити в стіл неможливо.

5. Алгебра на основі заданої множини

Нехай маємо деякий простір елементарних подій Ω та алгебру на ньому $\mathcal{F}$. Розгляньмо деяку множину A таку, що $\emptyset \subset A \subset \Omega$. Покажіть, що клас множин $\mathcal{F}_A = \{B \subseteq \Omega : B = A \cap C\ ,\ C \in \mathcal{F}\}$ є алгеброю. Зверніть увагу, що роль «усього простору» для $\mathcal{F}_A$ грає множина A.