Теорія ймовірностей, статистика та ймовірнісні процеси

Лекція 11.

Розора Ірина Василівна

Київ, 2022

Зміст

Дані. Моделі

Математична статистика

Математична статистика – це розділ математики, який базується на теорії ймовірностей та призначений для формулювання і доведення статистичних висновків про властивості ймовірнісного простору за результатами спостережень над відповідним стохастичним експериментом.

Висновок про ймовірнісний простір можна віднести до однієї з груп:

- висновок про кількісне значення деякої величини (параметра),
- якісний висновок про значення параметру чи іншу властивість імовірнісного простору.

Дані. Моделі

Математична статистика

Математична статистика – це розділ математики, який базується на теорії ймовірностей та призначений для формулювання і доведення статистичних висновків про властивості ймовірнісного простору за результатами спостережень над відповідним стохастичним експериментом.

Висновок про ймовірнісний простір можна віднести до однієї з груп:

- висновок про кількісне значення деякої величини (параметра),
- якісний висновок про значення параметру чи іншу властивість імовірнісного простору.

Задачі математичної статистики

Тому умовно задачі математичної статистики можна розбити на групи:

- задачі статистичного оцінювання, що спрямовані на побудову кількісних оцінок невідомих параметрів,
- задачі перевірки статистичних гіпотез, в яких встановлюються якісні властивості ймовірнісного простору.

Задачі математичної статистики

Тому умовно задачі математичної статистики можна розбити на групи:

- задачі статистичного оцінювання, що спрямовані на побудову кількісних оцінок невідомих параметрів,
- задачі перевірки статистичних гіпотез, в яких встановлюються якісні властивості ймовірнісного простору.

Приклади

A Генеральна сукупність (популяція) з N елементів, наприклад, відправлення виготовлених предметів. Невідоме число $N\theta$ цих елементів браковані.

Перевірити всі предмети – занадто дорого. Отже, щоб отримати інформацію про θ, відбирають на перевірку n предметів (без повторень).

Отримані дані - це число виявлених дефектів у вибірці.

Приклади

B Експериментатор робить n незалежних зчитувань фізичної сталої μ. Його вимірювання залежать від випадкових флуктуацій (похибок), тому дані можна розглядати як μ плюс деякі випадкові похибки.

Приклади

C Ми хочемо порівняти ефективність двох способів зробити щось за подібних умов, наприклад заварювання кави, зменшення забруднення, лікування захворювання, виробництво енергії, вивчення лабіринту тощо.

Це можна розглядати як задачу порівняння ефективності двох методів, застосованих до членів певної популяції.

Ми запускаємо $m+n$ незалежних експериментів так: $m+n$ членів генеральної сукупності (популяції) відбирають навмання, і m з них призначають першому методу, а решту n призначають другому. Так ми отримуємо одну або кілька кількісних чи якісних вимірювань ефективності експериментів.

Приклади

C Ми хочемо порівняти ефективність двох способів зробити щось за подібних умов, наприклад заварювання кави, зменшення забруднення, лікування захворювання, виробництво енергії, вивчення лабіринту тощо.

Це можна розглядати як задачу порівняння ефективності двох методів, застосованих до членів певної популяції.

Ми запускаємо $m + n$ незалежних експериментів так: $m + n$ членів генеральної сукупності (популяції) відбирають навмання, і m з них призначають першому методу, а решту n призначають другому. Так ми отримуємо одну або кілька кількісних чи якісних вимірювань ефективності експериментів.

Приклад 1. Вибіркове дослідження

Імовірнісна модель для X – числа бракованих елементів у вибірці.

- $\Omega = \{\omega\} = \{0, 1, \ldots, n\}$.
- Параметр θ: частка бракованих елементів у поставці

$$\Theta = \{\theta\} = \left\{0, \frac{1}{N}, \frac{2}{N} \ldots, \frac{N}{N}\right\}.$$

- Імовірнісний розподіл ξ – гіпергеометричний $\mathcal{H}(N\theta, N, n)$:

$$\mathbf{P}(\xi = k) = \frac{C_{N\theta}^{k} C_{N-N\theta}^{n-k}}{C_N^n}$$

Приклад 1. Вибіркове дослідження

Імовірнісна модель для X – числа бракованих елементів у вибірці.

- $\Omega = \{\omega\} = \{0, 1, \ldots, n\}$.
- Параметр θ: частка бракованих елементів у поставці

$$\Theta = \{\theta\} = \left\{0, \frac{1}{N}, \frac{2}{N} \ldots, \frac{N}{N}\right\}.$$

- Імовірнісний розподіл ξ – гіпергеометричний $\mathcal{H}(N\theta, N, n)$:

$$\mathbf{P}(\xi = k) = \frac{C_{N\theta}^{k} C_{N-N\theta}^{n-k}}{C_N^n}$$

Приклад 1. Вибіркове дослідження

Імовірнісна модель для X – числа бракованих елементів у вибірці.

- $\Omega = \{\omega\} = \{0, 1, \ldots, n\}$.
- Параметр θ: частка бракованих елементів у поставці

$$\Theta = \{\theta\} = \left\{0, \frac{1}{N}, \frac{2}{N} \ldots, \frac{N}{N}\right\}.$$

- Імовірнісний розподіл ξ – гіпергеометричний $\mathcal{H}(N\theta, N, n)$:

$$\mathbf{P}(\xi = k) = \frac{C_{N\theta}^{k} C_{N-N\theta}^{n-k}}{C_N^n}$$

Приклад 1 (продовження)

Основна відмінність нашої моделі від імовірнісної полягає в тому, що $N\theta$ – **невідоме**.

Тому, незважаючи на те, що вибірковий простір добре визначений, ми не можемо вказати імовірнісну структуру повністю, а лише запропонувати сімейство ймовірнісних розподілів $\{\mathcal{H}(N\theta, N, n)\}$ для X, будь-який з яких міг породити спостережені дані.

Приклад 2. Відбирання з популяції. Модель з однією вибіркою.

Приклад B можна розглядати як узагальнення ситуації A, бо вимірюють кількісну характеристику, а не просто записують "брак" чи ні.

Якщо вимірювання скалярні, ми спостерігаємо x_1, ..., x_n, які моделюємо як реалізації X_1, ..., X_n – н.о.р.в.в. зі спільною невідомою функцією розподілу F.

X_1, ..., X_n часто називають **випадковою вибіркою** із F. Модель повністю описується постульованим сімейством $\mathcal{F}$ розподілів.

Приклад 2 (продовження)

Можемо записати n зчитувань μ як

$$X_i = \mu + \varepsilon_i, \quad 1 \leq i \leq n,$$

де $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ – вектор випадкових похибок.

Приклад 2 (продовження)

Які треба накласти припущення на ε, щоб разом із μ повністю описати сумісний розподіл X_1, ..., X_n? Очевидно, це залежить від проведеного експерименту. Відповідно до В:

- Значення похибки на одному зчитуванні не впливає на значення похибки при інших зчитуваннях. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – незалежні.
- Розподіл похибки на одному зчитуванні такий самий, як і для інших. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – однаково розподілені.
- Розподіл ε не залежить від μ.

Приклад 2 (продовження)

Які треба накласти припущення на ε, щоб разом із μ повністю описати сумісний розподіл X_1, ..., X_n? Очевидно, це залежить від проведеного експерименту. Відповідно до В:

- Значення похибки на одному зчитуванні не впливає на значення похибки при інших зчитуваннях. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – незалежні.
- Розподіл похибки на одному зчитуванні такий самий, як і для інших. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – однаково розподілені.
- Розподіл ε не залежить від μ.

Приклад 2 (продовження)

Які треба накласти припущення на ε, щоб разом із μ повністю описати сумісний розподіл X_1, ..., X_n? Очевидно, це залежить від проведеного експерименту. Відповідно до В:

- Значення похибки на одному зчитуванні не впливає на значення похибки при інших зчитуваннях. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – незалежні.
- Розподіл похибки на одному зчитуванні такий самий, як і для інших. Тобто $\varepsilon_1, \ldots, \varepsilon_n$ – однаково розподілені.
- Розподіл ε не залежить від μ.

Приклад 2 (продовження)

Еквівалентно, X_1, ..., X_n – випадкова вибірка і, якщо G – розподіл ε_1, а F – розподіл X_1, то

$$F(x) = G(x - \mu)$$

і модель можна альтернативно визначити через $\mathcal{F}$ або через

$$\{(\mu, G)\colon \mu \in \mathbf{R}, G \in \mathcal{G}\},$$

де $\mathcal{G}$ – постульоване сімейство дозволених розподілів похибок. Зазвичай, $\mathcal{G}$ – всі розподілені із математичним сподіванням 0.
Класична параметрична модель така:

- Розподіл похибок $\mathcal{N}(0, \sigma^2)$, де σ^2 – невідомий параметр. Тобто X_i – вибірка з $\mathcal{N}(\mu, \sigma^2)$.

Приклад 2 (продовження)

Еквівалентно, X_1, ..., X_n – випадкова вибірка і, якщо G – розподіл ε_1, а F – розподіл X_1, то

$$F(x) = G(x - \mu)$$

і модель можна альтернативно визначити через $\mathcal{F}$ або через

$$\{(\mu, G) \colon \mu \in \mathbf{R}, G \in \mathcal{G}\},$$

де $\mathcal{G}$ – постульоване сімейство дозволених розподілів похибок. Зазвичай, $\mathcal{G}$ – всі розподілені із математичним сподіванням 0. Класична параметрична модель така:

- Розподіл похибок $\mathcal{N}(0, \sigma^2)$, де σ^2 – невідомий параметр. Тобто X_i – вибірка з $\mathcal{N}(\mu, \sigma^2)$.

Приклад 3. Модель із двома вибірками

Нехай $x_1, \dots, x_m$ – відгуки від m хворих осіб, яким дали ліки А, а $y_1, \dots, y_n$ – відгуки від n хворих такою ж хворобою осіб, які отримали лікування В.

Наприклад, А – плацебо і x_i – контрольні спостереження, а В – новий метод лікування і y_j описують спостережений ефект від лікування порівняно із плацебо.

Природні припущення:

- x і y – реалізації X_1, ..., X_n – випадкової вибірки з F та Y_1, ..., Y_n – – випадкової вибірки з G, тобто модель визначена сімейством можливих пар (F, G)

Приклад 3. Модель із двома вибірками

Нехай $x_1, \ldots, x_m$ – відгуки від m хворих осіб, яким дали ліки А, а $y_1, \ldots, y_n$ – відгуки від n хворих такою ж хворобою осіб, які отримали лікування В.

Наприклад, А – плацебо і x_i – контрольні спостереження, а В – новий метод лікування і y_j описують спостережений ефект від лікування порівняно із плацебо.

Природні припущення:

- x і y – реалізації X_1, ..., X_n – випадкової вибірки з F та Y_1, ..., Y_n – – випадкової вибірки з G, тобто модель визначена сімейством можливих пар (F, G)

Приклад 3 (продовження)

Щоб точніше визначати це сімейство, часто роблять припущення про сталий лікувальний ефект.

- Припустимо, що якби деякій особі призначили лікування А, отриманий відгук був би x. Тоді якщо натомість цій самій особі призначили б В, то отриманий відгук дорівнюватиме

$$y = x + \Delta,$$

де Δ не залежить від x. Звідси, якщо F – контрольний розподіл, то

$$G(x) = F(x - \Delta).$$

Приклад 3 (продовження)

Нарешті, остаточне спрощення:

- Контрольна група нормально розподілена. Тоді якщо F – це $N(\mu, \sigma^2)$, а G – це $N(\mu + \Delta, \sigma^2)$, отримуємо нормальну дво-вибіркову модель із рівними дисперсіями.

Як встановити набір припущень?

З досвіду і фізичних умов.

Як встановити набір припущень?

З досвіду і фізичних умов.

Зміст

Статистичний простір

Статистичним простором називають трійку

$$(\Omega, \mathcal{F}, (\mathbf{P}_\theta : \theta \in \Theta)),$$

що складається з таких елементів:

- Ω – простір елементарних подій деякого стохастичного експерименту,
- $\mathcal{F} \subset 2^\Omega$ – сигма-алгебра випадкових подій – підмножин Ω,
- $(\mathbf{P}_\theta : \theta \in \Theta)$ – деяка параметрична сім'я ймовірностей на $\mathcal{F}$,
- Θ – параметричний простір – множина довільної природи.

Статистичний простір

Статистичним простором називають трійку

$$(\Omega, \mathcal{F}, (\mathbf{P}_\theta : \theta \in \Theta)),$$

що складається з таких елементів:

- Ω – простір елементарних подій деякого стохастичного експерименту,
- $\mathcal{F} \subset 2^\Omega$ – сигма-алгебра випадкових подій – підмножин Ω,
- $(\mathbf{P}_\theta : \theta \in \Theta)$ – деяка параметрична сім'я ймовірностей на $\mathcal{F}$,
- Θ – параметричний простір – множина довільної природи.

Статистичний простір

Статистичним простором називають трійку

$$(\Omega, \mathcal{F}, (\mathbf{P}_\theta : \theta \in \Theta)),$$

що складається з таких елементів:

- Ω – простір елементарних подій деякого стохастичного експерименту,
- $\mathcal{F} \subset 2^\Omega$ – сигма-алгебра випадкових подій – підмножин Ω,
- $(\mathbf{P}_\theta : \theta \in \Theta)$ – деяка параметрична сім'я ймовірностей на $\mathcal{F}$,
- Θ – параметричний простір – множина довільної природи.

Статистичний простір

Статистичним простором називають трійку

$$(\Omega, \mathcal{F}, (\mathbf{P}_\theta : \theta \in \Theta)),$$

що складається з таких елементів:

- Ω – простір елементарних подій деякого стохастичного експерименту,
- $\mathcal{F} \subset 2^\Omega$ – сигма-алгебра випадкових подій – підмножин Ω,
- $(\mathbf{P}_\theta : \theta \in \Theta)$ – деяка параметрична сім'я ймовірностей на $\mathcal{F}$,
- Θ – параметричний простір – множина довільної природи.

Вважається, що вигляд залежності ймовірностей $\mathbf{P}_\theta(A)$ від подій $A \in \mathcal{F}$ при заданому значенні параметра θ повністю відомий, а сам параметр θ – невідомий статистику.

Найчастіше $\Theta \subset \mathbf{R}^d$, тобто ймовірнісний розподіл на $\mathcal{F}$ вважається відомим повністю за винятком d числових параметрів – координат $\theta \in \Theta$. У цьому випадку говорять про **параметричну статистику**.

Якщо ж множина Θ є підмножиною функціонального простору (наприклад, простору всіх функцій розподілу), то можна говорити про **непараметричну статистику**.

Математичне сподівання

Якщо ξ – в.в. на ймовірнісному просторі $(\Omega, \mathcal{F}, \mathbf{P}_\theta)$, математичне сподівання ξ (абстрактний інтеграл Лебега) за ймовірністю $\mathbf{P}_\theta$ будемо позначати через

$$\mathbf{M}_\theta \xi = \int_\Omega \xi(\omega) \mathbf{P}_\theta(\,\mathrm{d}\omega).$$

Аналогічний зміст має позначення $\mathbf{D}_\theta \xi$ для дисперсії.

Математичне сподівання

Якщо ξ – в.в. на ймовірнісному просторі $(\Omega, \mathcal{F}, \mathbf{P}_\theta)$, математичне сподівання ξ (абстрактний інтеграл Лебега) за ймовірністю $\mathbf{P}_\theta$ будемо позначати через

$$\mathbf{M}_\theta \xi = \int_\Omega \xi(\omega) \mathbf{P}_\theta(\,\mathrm{d}\omega).$$

Аналогічний зміст має позначення $\mathbf{D}_\theta \xi$ для дисперсії.

Зміст

Статистичні висновки про ймовірність $\mathbf{P}_\theta$ будемо робити на підставі спостережень, тобто значень певних випадкових величин, векторів та інших функцій від елементарних подій.

Статистична вибірка

Статистичною вибіркою називають довільну вимірну функцію $X : \Omega \to S$ зі значеннями у вимірному вибірковому просторі (S, Σ, λ), де:

- S – деяка множина (вибірковий простір),
- Σ – сигма-алгебра підмножин S,
- λ – деяка сигма-скінченна міра на Σ.

Вважається, що значення $X(\omega) = x$ є відомим (спостерігається) і може використовуватись для отримання статистичних висновків.

Статистичні висновки про ймовірність $\mathbf{P}_\theta$ будемо робити на підставі спостережень, тобто значень певних випадкових величин, векторів та інших функцій від елементарних подій.

Статистична вибірка

Статистичною вибіркою називають довільну вимірну функцію $X : \Omega \to S$ зі значеннями у вимірному вибірковому просторі (S, Σ, λ), де:

- S – деяка множина (вибірковий простір),
- Σ – сигма-алгебра підмножин S,
- λ – деяка сигма-скінченна міра на Σ.

Вважається, що значення $X(\omega) = x$ є відомим (спостерігається) і може використовуватись для отримання статистичних висновків.

Статистичні висновки про ймовірність $\mathbf{P}_\theta$ будемо робити на підставі спостережень, тобто значень певних випадкових величин, векторів та інших функцій від елементарних подій.

Статистична вибірка

Статистичною вибіркою називають довільну вимірну функцію $X : \Omega \to S$ зі значеннями у вимірному вибірковому просторі (S, Σ, λ), де:

- S – деяка множина (вибірковий простір),
- Σ – сигма-алгебра підмножин S,
- λ – деяка сигма-скінченна міра на Σ.

Вважається, що значення $X(\omega) = x$ є відомим (спостерігається) і може використовуватись для отримання статистичних висновків.

Статистичні висновки про ймовірність $\mathbf{P}_\theta$ будемо робити на підставі спостережень, тобто значень певних випадкових величин, векторів та інших функцій від елементарних подій.

Статистична вибірка

Статистичною вибіркою називають довільну вимірну функцію $X : \Omega \to S$ зі значеннями у вимірному вибірковому просторі (S, Σ, λ), де:

- S – деяка множина (вибірковий простір),
- Σ – сигма-алгебра підмножин S,
- λ – деяка сигма-скінченна міра на Σ.

Вважається, що значення $X(\omega) = x$ є відомим (спостерігається) і може використовуватись для отримання статистичних висновків.

Надалі вибірковим простором буде обиратися переважно евклідів простір $S = \mathbf{R}^n$ із борелєвою сигма-алгеброю $\Sigma = \mathcal{B}(\mathbf{R})$, тому під вибіркою слід розуміти звичайний випадковий вектор, що спостерігається в стохастичному експерименті.
У більшості випадків мірою λ слугує або точкова міра – відносно неї кожна одноточкова множина з певного класу має одиничне значення міри (у випадку дискретної вибірки X), або ж міра Лебега, якщо $S \subset R^n$ і вибірка X має сумісну щільність.

Часто у статистиці використовується схема багатократних спостережень

Вимірний простір $(S, \Sigma, \lambda) = (\mathbf{R}, \mathcal{B}, \nu)^n$ є n-кратним прямим добутком вимірного простору $(\mathbf{R}, \mathcal{B}, \nu)$, якщо

$$S = \mathbf{R}^n \equiv \{x = (x_1, ..., x_n), x_k \in \mathbf{R}\},$$

$$\Sigma = \mathcal{B} \otimes ... \otimes \mathcal{B} \equiv \sigma[B_1 \times ... \times B_n, B_k \in \mathcal{B}, k = \overline{1, n}],$$

а міра $\lambda \equiv \nu \times ... \times \nu$ визначається на прямокутниках як добуток

$$\lambda(B_1 \times ... \times B_n) \equiv \nu(B_1)...\nu(B_n).$$

Наприклад, $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n), L_n) = (\mathbf{R}, \mathcal{B}(\mathbf{R}), L_1)^n$, де L_n – n-вимірна міра Лебега (довжина, площа, об'єм...).

Нехай вибірковий простір (S, Σ, λ) є n-кратним прямим добутком $(\mathbf{R}, \mathcal{B}, \nu)^n$.

n-кратна вибірка

Випадковий вектор $X : \Omega \to S$ називається n-кратною вибіркою, якщо $X = (\xi_1, ..., \xi_n)$ складається з незалежних у сукупності однаково розподілених величин ξ_k, $k = 1, ..., n$, зі значеннями у просторі $(\mathbf{R}, \mathcal{B}, \nu)$, тобто

$$\mathbf{P}_\theta(X \in B_1 \times ... \times B_n) = \prod_{k=1}^{n} \mathbf{P}_\theta(\xi_1 \in B_1), \quad \forall B_k \in \mathcal{B}, \theta \in \Theta.$$

Об'єм вибірки

Число n називається об'ємом вибірки X.

Генеральна сукупність

Для кратної вибірки $X = (\xi_1, ..., \xi_n)$ множина всіх випадкових величин ξ, що мають однакову з ξ_1 функцію розподілу, називається генеральною сукупністю (популяцією).

У зв'язку з цим вибірку X можна уявляти як результат n-кратного незалежного послідовного вибору представників з генеральної сукупності.

Зміст

Статистика

Статистикою називається довільна вимірна функція (скалярна чи векторна) від вибірки: $T = T(X)$, яка не містить значень невідомого параметра θ

Множина значень статистики є довільним вимірним простором $(C, \mathcal{C})$. Найчастіше це евклідів простір:

$$(C, \mathcal{C}) = (\mathbf{R}^m, \mathcal{B}(\mathbf{R}^m))$$

Зауваження

Словом "статистика" будемо одночасно визначати як саму функціональну залежність

$$T(x) : S \to C$$

від вибірки, так і її значення

$$T(\omega) = T(X(\omega)) : \Omega \to C,$$

яке отримується після підстановки вибірки $x = X(\omega)$. Тлумачення випливатиме зі змісту відповідного аргумента.

Оцінка

Оцінкою невідомого параметра $\theta \in \Theta$ називається будь-яка статистика зі значеннями у параметричному просторі Θ.

Щоб підкреслити спеціальний характер оцінки, часто її зображають у вигляді

$$\hat{\theta}.$$

Очевидно, оцінка є засобом для прогнозування, передбачення, оцінювання значення невідомого параметра θ на підставі спостережень.

Зауваження

Якщо при кожному n спостерігається кратна вибірка X об'єму n, а спосіб, яким утворена оцінка, один і той самий (не залежить від об'єму вибірки n), то поняття "оцінка" використовують також у широкому розумінні як "послідовність оцінок", що утворені за одним правилом при різних значеннях об'єму вибірки.

Послідовності оцінок позначаються через

$$\hat{\theta}_n.$$

Приклад

Вибіркове середнє є класичною оцінкою положення і дорівнює середньому арифметичному спостережень, які утворюють вибірку. Однак фактично це є послідовність оцінок, оскільки при кожному значенні об'єму вибірки маємо окрему статистику

$$\theta = \frac{1}{n} \sum_{k=1}^{n} \xi_k.$$

Властивості оцінок

Безперечно, якість тієї чи іншої оцінки потребує порівняльного аналізу. Для порівняння оцінок чи їх послідовностей будемо використовувати такі поняття.

Незміщена оцінка

Оцінка $\hat{\theta}$ називається (незсунутою) незміщеною, якщо її математичне сподівання збігається з точним значенням θ:

$$\mathsf{M}_\theta \hat{\theta} = \theta, \quad \forall \theta \in \Theta.$$

Асимптотично незміщена оцінка

Оцінка $\hat{\theta}_n$ називається асимптотично незміщеною (незсунутою), якщо якщо має місце асимптотична збіжність середніх

$$\mathbf{M}_\theta \hat{\theta}_n \to \theta, \quad n \to \infty, \ \forall \theta \in \Theta.$$

Консистентна оцінка

Оцінка $\hat{\theta}_n$ називається консистентною (або слушною), якщо вона збігається за ймовірністю до істинного значення θ

$$\hat{\theta}_n \xrightarrow{\mathbf{P}_\theta} \theta, \quad n \to \infty, \ \forall \theta \in \Theta,$$

тобто

$$\mathbf{P}_\theta(|\hat{\theta}_n - \theta| \geq \varepsilon) \to 0, \quad n \to \infty, \ \forall \varepsilon > 0, \ \forall \theta \in \Theta.$$

Достатні умови консистентності

Теорема (Достатні умови консистентності)

Нехай оцінка $\hat{\theta}_n$ є незміщеною (асимптотично незміщеною) оцінкою параметра θ і $\mathbf{D}\hat{\theta}_n \to 0$, $n \to \infty$. Тоді оцінка $\hat{\theta}_n$ буде консистенною.

Доведення.

Доведемо це твердження для незміщених оцінок, а саме $\mathbf{M}\hat{\theta}_n = \theta$. За нерівністю Чебишева

$$\mathbf{P}_\theta(|\hat{\theta}_n - \theta| \geq \varepsilon) = \mathbf{P}_\theta(|\hat{\theta}_n - \mathbf{M}\hat{\theta}_n| \geq \varepsilon) \leq \frac{\mathbf{D}\hat{\theta}_n}{\varepsilon^2} \to 0, \quad n \to \infty$$

□

Строго консистентна оцінка

Оцінка $\hat{\theta}_n$ називається строго консистентною, якщо вона збігається з імовірністю 1 до істинного значення θ

$$\hat{\theta}_n \xrightarrow{\mathbf{P}_\theta 1} \theta, \quad n \to \infty, \ \forall \theta \in \Theta,$$

тобто

$$\mathbf{P}_\theta(\exists \lim_{n\to\infty} \hat{\theta}_n = \theta) = 1, \quad \forall \theta \in \Theta.$$

Зауваження

Наведені властивості стосуються оцінок $\hat{\theta}$ для значення невідомого параметру θ. У випадку, коли цей параметр – векторний, доцільно розглядати також оцінки $\hat{\tau}$ для значень деякої функції

$$\tau = \tau(\theta)$$

від параметра θ. Сформульовані вище означення поширюються також і на дану схему, якщо замінити θ на $\tau(\theta)$, а $\hat{\theta}$ на $\hat{\tau}$.

Асимптотично нормальна оцінка

Оцінка $\hat{\theta}_n$ параметра θ називається асимптотично нормальною, якщо знайдеться числова нормуюча послідовність $c_n = c_n(\theta)$ така, що має місце слабка збіжність

$$c_n(\hat{\theta}_n - \theta) \overset{W_\theta}{\to} \zeta \cong N(0, \sigma^2), \quad n \to \infty, \ \forall\theta \in \Theta.$$

Тобто

$$\mathbf{P}_\theta(c_n(\hat{\theta}_n - \theta) \leq x) \to \Phi(x), \quad n \to \infty, \ \forall x \in \mathbf{R}, \ \forall\theta \in \Theta,$$

де Φ – стандартна нормальна функція розподілу.

Задача

Нехай $\xi_1, \ldots, \xi_n$ – вибірка з розподілу зі щільністю

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & \text{якщо } x \in [a, b]; \\ 0, & \text{якщо } x \notin [a, b]. \end{cases}$$

Розглядають оцінки

$$\hat{\theta}_1 = \max\{\xi_1, \ldots, \xi_n\}, \quad \hat{\theta}_2 = \min\{\xi_1, \ldots, \xi_n\},$$

$$\hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n} \xi_n, \quad \hat{\theta}_4 = \frac{\xi_{n-1} + \xi_n}{2}.$$

Які з цих оцінок і для яких параметрів є незміщеними? Консистентними?

Розв'язування

Знайдемо розподіл $\hat{\theta}_1$. Оскільки $\xi_1, \ldots, \xi_n$ – н.о.р. в.в. з рівномірного розподілу,

$$\begin{aligned} F_{\hat{\theta}_1}(x) &= \mathbf{P}(\hat{\theta}_1 \le x) = \mathbf{P}(\max\{\xi_1, \ldots, \xi_n\} \le x) \\ &= \mathbf{P}(\xi_1 \le x, \ldots, \xi_n \le x) = \prod_{i=1}^{n} \mathbf{P}(\xi_i \le x) \\ &= \prod_{i=1}^{n} \int_{-\infty}^{x} f(y; a, b)\,\mathrm{d}y = \left(\int_{-\infty}^{x} f(y; a, b)\,\mathrm{d}y \right)^n . \end{aligned}$$

Розв'язування : $\hat{\theta}_1$

Отже, $\hat{\theta}_1$ – абсолютно неперервна в.в. зі щільністю

$$
\begin{aligned}
f_{\hat{\theta}_1}(x) &= \frac{\mathrm{d}}{\mathrm{d}x} F_{\hat{\theta}_1}(x) = nf(x; a, b) \left(\int_{-\infty}^{x} f(y; a, b)\, \mathrm{d}y \right)^{n-1} \\
&= \begin{cases} \frac{n}{(b-a)^n}(x-a)^{n-1}, & \text{якщо } x \in [a, b]; \\ 0, & \text{якщо } x \notin [a, b]. \end{cases}
\end{aligned}
$$

Розв'язування : $\hat{\theta}_1$

$$
\begin{aligned}
\mathbf{M}\hat{\theta}_1 &= \int_{-\infty}^{\infty} f_{\hat{\theta}_1}(x)\,\mathrm{d}x \\
&= \int_a^b x\frac{n}{(b-a)^n}(x-a)^{n-1}\,\mathrm{d}x = \frac{n}{n+1}b + \frac{a}{n+1}.
\end{aligned}
$$

Отже, оцінка $\hat{\theta}_1$ не є незсунутою оцінкою ні параметра a, ні b. Але при $n \to \infty$

$$
\mathbf{M}\hat{\theta}_1 = \frac{n}{n+1}b + \frac{a}{n+1} \to b.
$$

Тобто $\hat{\theta}_1$ є асимптотично незміщеною оцінкою параметра b.

Розв'язування : $\hat{\theta}_1$

Перевіримо, чи $\hat{\theta}_1$ є консистентною оцінкою параметра b. Для достатньо малих $\varepsilon > 0$

$$
\begin{aligned}
\mathbf{P}\{|\hat{\theta}_1 - b| > \varepsilon\} &= \mathbf{P}\{\hat{\theta}_1 \in (-\infty, b-\varepsilon) \cup (b+\varepsilon, +\infty)\} \\
&= \int_{-\infty}^{b-\varepsilon} f_{\hat{\theta}_1}(x)\,\mathrm{d}x + \int_{b+\varepsilon}^{+\infty} f_{\hat{\theta}_1}(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{b-\varepsilon} f_{\hat{\theta}_1}(x)\,\mathrm{d}x \\
&= \int_{a}^{b-\varepsilon} \frac{n}{(b-a)^n}(x-a)^{n-1}\,\mathrm{d}x \\
&= \left(1 - \frac{\varepsilon}{b-a}\right)^n \to 0, \quad n \to \infty.
\end{aligned}
$$

Розв'язування : $\hat{\theta}_1$

Отже, при $n \to \infty$

$$\mathbf{P}\{|\hat{\theta}_1 - b| > \varepsilon\} \to 0,$$

тобто $\hat{\theta}_1$ є консистентною оцінкою параметра b.

Оцінка $\hat{\theta}_2$ досліджується аналогічно. $\hat{\theta}_2$ є асимптотично незсунутою і консистентною оцінкою параметра a.

Розв'язування : $\hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}\xi_i$

Оскільки ξ_i рівномірно розподілені на $[a,b]$ в.в. і $m_1 = \mathbf{M}\xi_i = \frac{a+b}{2}$, то

$$\mathbf{M}\hat{\theta}_3 = \mathbf{M}\frac{1}{n}\sum_{i=1}^{n}\xi_i = \frac{a+b}{2}.$$

Отже, $\hat{\theta}_3$ є незсунутою оцінкою середини відрізка $[a,b]$.

Згідно ЗВЧ $\hat{\theta}_3$ збіжна за ймовірністю до $m_1 = \frac{a+b}{2}$, т.б. $\hat{\theta}_3$ – консистентна оцінка параметра $m_1 = \frac{a+b}{2}$.

Розв'язування : $\hat{\theta}_4 = (\xi_{n-1} + \xi_n)/2$

$$\mathsf{M}\hat{\theta}_4 = \mathsf{M}\frac{\xi_{n-1} + \xi_n}{2} = \frac{a+b}{2}.$$

Отже, $\hat{\theta}_4$ є незсунутою оцінкою параметра $m_1 = \frac{a+b}{2}$.

Оцінка $\hat{\theta}_4$ не збігається за ймовірністю до жодної сталої c. Дійсно, щільність

$$\begin{aligned} f_{\hat{\theta}_4}(x) &= \frac{\mathrm{d}}{\mathrm{d}x} F_{\hat{\theta}_1}(x) = nf(x;a,b)\left(\int_{-\infty}^{x} f(y;a,b)\,\mathrm{d}y\right)^{n-1} \\ &= \begin{cases} \frac{4}{(b-a)^2}(x-a), & \text{якщо } x \in [a, \frac{a+b}{2}]; \\ \frac{4}{(b-a)^2}(b-x), & \text{якщо } x \in [\frac{a+b}{2}, b]; \\ 0, & \text{якщо } x \notin [a,b]. \end{cases} \end{aligned}$$

Розв'язування : $\hat{\theta}_4 = (\xi_{n-1} + \xi_n)/2$

Тоді, для кожного достатньо малого ε

$$\mathsf{P}\{|\hat{\theta}_4 - c| > \varepsilon\} = \int_{\{x:\ |x-c|>\varepsilon\}} f_{\hat{\theta}_4}(x)\,\mathrm{d}x$$

є сталою, що не залежить від n, а значить, не збігається до нуля при $n \to \infty$.

ПИТАННЯ?