Теорія ймовірностей, статистика та ймовірнісні процеси

Лекція 14.

Розора Ірина Василівна

Київ, 2020

Зміст I

Зміст

Нехай $X = (\xi_1, \cdots, \xi_n)$—кратна вибірка. $\Theta \subset \mathbf{R}^d$.

Інтервальна оцінка

Інтервальною оцінкою (або надійним інтервалом) невідомого параметра θ називається пара $\mathbf{R}^d$-значних статистик $(\hat{\theta}_1, \hat{\theta}_2)$ таких, що

$$\hat{\theta}_1 \leq \hat{\theta}_2 \quad \text{майже напевне.}$$

Оцінкою тут є паралелепіпед

$$[\hat{\theta}_1, \hat{\theta}_2) = \{\theta \in \mathbf{R}^d : \hat{\theta}_1 \leq \theta < \hat{\theta}_2\}.$$

Іноді щодо невідомого значення θ досить вказати обмеження лише з одного боку. У цьому разі на відміну від попередньої інтервальної оцінки, яка називається двобічною, розглядають також однобічні оцінки: лівобічну інтервальну оцінку $[\hat{\theta}, \infty)$ та правобічну інтервальну оцінку $(-\infty, \hat{\theta})$.

Інтервальна оцінка

Інтервальна оцінка $(\hat{\theta}_1, \hat{\theta}_2)$ невідомого параметра θ є незсунутою оцінкою надійності p, якщо для всіх $\theta \in \Theta$

$$\mathbf{P}_\theta(\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2) = p.$$

Зауваження

Значення p ще називають довірчою ймовірністю. Якщо $p = 1 - \alpha$, то α називають рівнем значущості.

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

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

За теоремою про еквівалентність слабкої збіжності та в основному асимптотична нормальність еквівалентна збіжності

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

де Φ – стандартна нормальна функція розподілу.

Побудова інтервальної оцінки

Якщо оцінка $\hat{\theta}_n$ є асимптотично нормальною, а $c_n \ \sqrt{n}/\sigma$, то

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{W}{\to} \sigma\zeta \equiv \eta \cong N(0, \sigma^2), \quad n \to \infty.$$

Асимптотична дисперсія

Величина σ^2 називається асимптотичною дисперсією оцінки $\hat{\theta}_n$.

Істинне її значення $\sigma^2 = \sigma^2(\theta)$ є відомою функцією параметра θ.

З наведеної слабкої збіжності випливає, що розподіл нормованої величини

$$\sqrt{n}(\hat{\theta}_n - \theta)/\sigma(\theta)$$

наближається до розподілу стандартної нормальної величини ζ.

Означення

Нехай γ деяка в.в. із функцією розподілу $F(x)$. Квантиллю порядку $p \in (0,1)$ називається число x_p таке, що

$$F(x_p) = P\{\gamma \le x_p\} = p.$$

Значення x_p для стандартного нормального розподілу шукаються з таблиць. Наприклад, $x_{0.997} \approx 3$ за правилом трьох сигма.

Надалі, через x_p будемо позначати квантилі для в.в.$\zeta \approx N(0,1)$.

Тоді при великих n наближено

$$\mathbf{P}_\theta\left(\sqrt{n}\frac{\hat{\theta}_n-\theta}{\sigma(\theta)}\le x_{1-\alpha/2}\right)\approx \mathbf{P}(|\zeta|\le x_{1-\alpha/2})=$$

із симетричності стандартої в.в . випливає, що

$$x_{\alpha/2}=-x_{1-\alpha/2},$$

тому

$$=\mathbf{P}(-x_{1-\alpha/2}<\zeta\le x_{1-\alpha/2})=\Phi(x_{1-\alpha/2})-\Phi(\alpha/2)=$$

$$=1-\alpha/2-\alpha/2=1-\alpha.$$

Припустимо, що функція $\sigma(\theta)$ неперервна, а оцінка $\hat{\theta}_n$ – консистентна. З цих припущень випливає збіжність за ймовірністю

$$\sigma(\hat{\theta}_n) \xrightarrow{\mathbf{P}} \sigma(\theta).$$

Тому можна наближено замінити $\sigma(\theta)$ на $\sigma(\hat{\theta}_n)$ під знаком імовірності і стверджувати, що при великих n подія

$$\left\{ \sqrt{n}\frac{\hat{\theta}_n - \theta}{\sigma(\hat{\theta}_n)} \leq x_{1-\alpha/2} \right\}$$

наближено теж має ймовірність $1-\alpha$.

Інтервальні асимптотично незсунуті оцінки з довірчою ймовірністю $1-\alpha$

Властивості асимптотичної нормальності та консистентності дають можливість наближеної побудови інтервальних асимптотично незсунутих оцінок з довірчою надійністю $1-\alpha$ для невідомого параметра, що мають вигляд

$$\mathbf{P}_\theta\left(\hat{\theta}_n - x_{1-\alpha/2}\frac{\sigma(\hat{\theta}_n)}{\sqrt{n}} \le \theta \le \hat{\theta}_n + x_{1-\alpha/2}\frac{\sigma(\hat{\theta}_n)}{\sqrt{n}}\right) \approx 1-\alpha, \quad \forall\theta\in\Theta.$$

https://edpuzzle.com/media/5f8e896a063dcb406bb5aaad

Зміст

Стандартний гауссовий розподіл

$\xi \sim N(0,1)$ тоді і тільки тоді, коли щільність цієї в.в.

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Властивості:

- Якщо $\xi \sim N(m, \sigma^2)$, то $z = \frac{\xi - m}{\sigma} \sim N(0,1)$.
- Якщо ξ_i— незалежні в.в. з $\xi_i \sim N(m_i, \sigma_i^2)$, $i = \overline{1,n}$, то

$$\sum_{i=1}^{n} \xi_i \sim N(\sum_{i=1}^{n} m_i, \sum_{i=1}^{n} \sigma_i^2).$$

Властивості

- Якщо $\xi \sim N(m, \sigma^2)$, то

$$P\{a < \xi < b\} = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right),$$

  де $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} dt$ — функція розподілу $N(0, 1)$.
- Якщо $z \sim N(0, 1)$, то для її функції розподілу справедлива властивість

$$\Phi(-a) = 1 - \Phi(a),$$

.

χ^2-розподіл з n ступенями свободи(вільності)

Розглянемо n незалежних в.в. xi_i, $i = \overline{1, n}$, що мають однаковий стандартний гауссовий розподіл $N(0, 1)$.

Означення

Кажуть, що в.в. $\chi^2(n)$ має χ^2-розподіл з n ступенями свободи, якщо її можна представити у вигляді

$$\chi^2(n) = \sum_{i=1}^{n} \xi^2.$$

χ^2-розподіл з n ступенями свободи, як і стандартний гауссовий, також табульований.

Статистики від нормальних вибірок

Теорема (про вибіркові моменти стандартної нормальної вибірки)

Нехай вибірка $\zeta = (\zeta_1, \dots, \zeta_n)$ є стандартним нормальним вектором. Тоді її перші два вибіркові моменти

$$\hat{\mu}_n = \frac{1}{n}\sum_{k=1}^{n}\zeta_k, \quad \hat{s}_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(\zeta_k - \hat{\mu}_n)^2$$

є незалежними випадковими величинами, причому

$$\sqrt{n}\hat{\mu}_n \cong N(0,1)$$

та

$$\frac{(n-1)\hat{s}_n^2}{\sigma^2} \cong \chi^2_{n-1},$$

Теорема (про вибіркові середнє та дисперсію нормальної вибірки)

Нехай $X = (\xi_1, \ldots, \xi_n)$ – кратна вибірка з нормальним розподілом спостережень: $\xi_k \cong N(\mu, \sigma^2)$. Тоді статистики $\hat{\mu}_n$ і $\hat{s}_n^2$ незалежні та

$$\sqrt{n}\frac{\hat{\mu}_n - \mu}{\sigma} \cong N(0,1), \qquad \frac{(n-1)\hat{s}_n^2}{\sigma^2} \cong \chi^2_{n-1},$$

$$n\frac{\hat{\sigma}_n^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{k=1}^{n}(\xi_k - \mu)^2 \cong \chi^2_n.$$

Припущення нормальності кратної вибірки X необхідне і достатнє для незалежності вибіркових середнього та дисперсії.

Статистики Стьюдента та Фішера

Розподіл Стьюдента

Випадкова величина τ_n має t-розподіл, або розподіл Стьюдента, з n ступенями свободи, якщо її можна подати у вигляді

$$\tau_n = \frac{\zeta}{\sqrt{\chi_n^2/n}},$$

де $\zeta \cong N(0,1)$ – стандартна нормальна величина, а χ_n^2 – незалежна від неї величина з хі-квадрат розподілом та n ступенями свободи.

Розподіл Фішера

Випадкова величина $\phi_{n,m}$ має розподіл Фішера, або ж розподіл Снедекора – Фішера, з n, m ступенями свободи, якщо її можна подати у вигляді

$$\phi_{n,m} = \frac{\chi_n^2/n}{\chi_m^2/m},$$

де χ_n^2, χ_m^2 – незалежні величини з хі-квадрат розподілом та n, m ступенями свободи відповідно.

Теорема (про статистику Стьюдента від нормальної вибірки)

Нехай $X = (\xi_1, ..., \xi_n)$ – кратна вибірка з нормальним розподілом $N(\mu, \sigma^2)$. Тоді статистика

$$\tau_{n-1} = \sqrt{n}\frac{\hat{\mu}_n - \mu}{\hat{s}_n}$$

має розподіл Стьюдента з $n-1$ ступенем свободи.

Історична довідка

William Sealy Gosset(1876 -1937) Canterbury, England.
Working in the Guinness brewery in Dublin he did important work on statistics. Gosset invented the t-test to handle small samples for quality control in brewing. He wrote under the name "Student".

Доведення

За теоремою про вибіркові середнє та дисперсію нормальної вибірки після ділення на σ чисельник

$$\sqrt{n}\frac{\hat{\mu}_n - \mu}{\sigma}$$

має стандартний нормальний розподіл, і не залежить від знаменника

$$\hat{s}_n/\sigma = \sqrt{(n-1)\hat{s}_n^2/\sigma^2(n-1)} \cong \sqrt{\chi^2_{n-1}/(n-1)}.$$

Тому весь дріб τ_{n-1} має розподіл Стьюдента.

Інтервальні оцінки параметрів нормальних спостережень

Побудова надійних інтервалів для нормальних спостережень ґрунтується на властивостях вибіркових моментів для нормальних виборок.

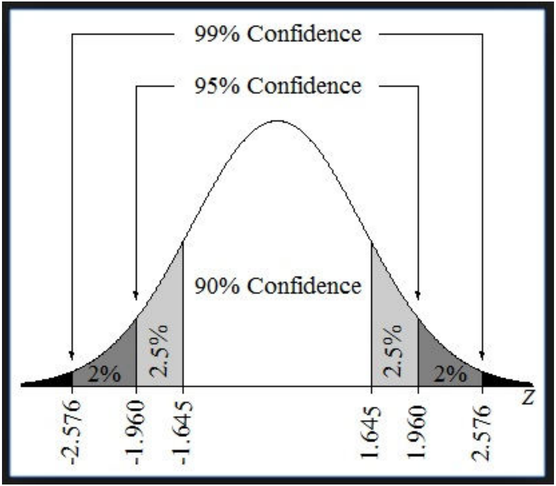
Позначимо через $x_{1-\alpha/2}$ квантиль порядку $1-\alpha/2$ для стандартного нормального розподілу: $\Phi(x_{1-\alpha/2}) = 1-\alpha/2$. Тоді

$$\mathbf{P}(|\zeta| \leq x_{1-\alpha/2}) = 1-\alpha.$$

Інтервальні оцінки параметрів нормальних спостережень

Зокрема, при побудові інтервальних оцінок можна скористатись тим, що при $1 - \alpha = 0.95$ надійний інтервал для середнього $N(0, 1)$ розподілу дорівнює:

- $(-1.96, 1.96)$ – у симетричному варіанті,
- $(-\infty, 1.645)$ – у правобічному варіанті,
- $(-1.645, \infty)$ – у лівобічному варіанті,
- $(-1.88, 2.05)$ – у правозкошеному варіанті,
- $(-2.05, 1.88)$ – у лівозкошеному варіанті.


99% Confidence
95% Confidence
90% Confidence
2%
2.5%
2.5%
2%
Z
-2.576
-1.960
-1.645
1.645
1.960
2.576

Інтервальні оцінки параметрів нормальних спостережень

Аналогічний зміст мають квантилі $y_{n,1-\alpha/2}$ для розподілу Стьюдента, та квантилі $z_{n,1-\alpha/2}$ для хі-квадрат розподілу:

$$\mathsf{P}(|\tau_n| \le y_{n,1-\alpha/2}) = 1 - \alpha, \quad \mathsf{P}(\chi^2_n \le z_{n,1-\alpha/2}) = 1 - \alpha/2.$$

Оцінка середнього при відомій дисперсії

Задана кратна вибірка $X = (\xi_1, \cdots, \xi_n)$ р розподілом $N(\mu, \sigma^2)$. Оскільки при відомій дисперсії σ^2 статистика

$$\zeta = \sqrt{n}(\hat{\mu}_n - \mu)/\sigma \cong N(0,1)$$

є стандартною нормальною, то

$$\mathbf{P}\left(\hat{\mu}_n - \frac{\sigma}{\sqrt{n}} x_{1-\alpha/2} \le \mu \le \hat{\mu}_n + \frac{\sigma}{\sqrt{n}} x_{1-\alpha/2}\right) =$$

$$= \mathbf{P}(|\zeta| \le x_{1-\alpha/2}) = 1 - \alpha.$$

Оцінка дисперсії при відомому середньому

У випадку відомого середнього μ для відповідно модифікованої вибіркової дисперсії

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mu)^2$$

зі співвідношення

$$n\hat{\sigma}_n^2/\sigma^2 \cong \chi_n^2$$

виводимо, що

$$\mathbf{P}\left(n\hat{\sigma}_n^2/z_{n,1-\alpha/2} \le \sigma^2 \le n\hat{\sigma}_n^2/z_{n,\alpha/2}\right) = \mathbf{P}\left(z_{n,\alpha/2} \le \chi_n^2 \le z_{n,1-\alpha/2}\right)$$

$$= 1 - \alpha/2 - \alpha/2 = 1 - \alpha.$$

Оцінка середнього при невідомій дисперсії

За теоремою про статистику Стьюдента від нормальної вибірки статистика

$$\tau_n = \sqrt{n}\frac{\hat{\mu}_n - \mu}{\hat{s}_n}$$

має розподіл Стьюдента з $n-1$ ступенем свободи. Тому

$$\mathbf{P}\left(\hat{\mu}_n - \frac{\hat{s}_n}{\sqrt{n}}y_{n-1,1-\alpha/2} \le \mu \le \hat{\mu}_n + \frac{\hat{s}_n}{\sqrt{n}}y_{n-1,1-\alpha/2}\right) =$$

$$= \mathbf{P}(|\tau_n| \le y_{n-1,1-\alpha/2}) = 1 - \alpha.$$

Оцінка дисперсії при невідомому середньому

Оскільки
$$\frac{(n-1)\hat{s}_n^2}{\sigma^2} \cong \chi^2_{n-1},$$
то
$$\mathbf{P}\left(\frac{(n-1)\hat{s}_n^2}{z_{n-1,1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)\hat{s}_n^2}{z_{n-1,\alpha/2}}\right)$$
$$= \mathbf{P}(z_{n-1,\alpha/2} \le \chi^2_{n-1} \le z_{n-1,1-\alpha/2}) = 1-\alpha/2-\alpha/2 = 1-\alpha.$$

Зміст

Статистична гіпотеза

Статистичною гіпотезою називається довільне припущення про розподіл вибірки, яка спостерігається у стохастичному експерименті.

Оскільки цей розподіл вважається повністю відомим при істинному значенні параметра θ, статистичні гіпотези часто формулюють у вигляді

$$H_0 : \theta \in \Theta_0,$$

де $\Theta_0 \subset \Theta$ – певна параметрична підмножина. Якщо параметричний простір Θ ототожнити з множиною всіх можливих розподілів вибірки, то будь-яка гіпотеза може бути зображена у наведеному вигляді.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Зокрема, розглядають гіпотези про те, що:

1. значення параметра дорівнює заданому,
2. значення параметра перевищує задане,
3. розподіл спостережень збігається з заданим,
4. розподіл спостережень належить заданому класу розподілів,
5. групи спостережень є однорідними (мають однакові розподіли),
6. групи спостережень є незалежними,
7. спостереження є випадковими.

Статистичні властивості (зокрема, якість) висновків щодо розподілу вибірки суттєво залежать не тільки від вигляду статистичної гіпотези, а й від множини значень параметра, що не задовольняють цю гіпотезу.

Вказана множина значень не завжди збігається з доповненням параметричної множини Θ_0.

Статистична альтернатива

Статистичною альтернативою називають таке припущення про розподіл вибірки, яке вважається виконаним у випадку, коли не справджується основна статистична гіпотеза.

Статистичні альтернативи мають вигляд

$$H_1 : \theta \in \Theta_1,$$

де $\Theta_1 \subset \Theta$.

Якщо

$$\Theta_1 = \Theta \backslash \Theta_0,$$

то альтернатива може не формулюватися явно. Якщо ж має місце строге включення

$$\Theta_0 \cup \Theta_1 \subset \Theta,$$

то основна гіпотеза H_0, на відміну від альтернативної, називається нульовою гіпотезою.

Проста альтернатива

Статистична гіпотеза

$$H_0 : \theta \in \Theta_0$$

називається простою гіпотезою, якщо множина

$$\Theta_0 = \{\theta_0\}$$

є одноелементною, причому розподіл

$$\mathbf{P}_{\theta_0}(X \in \cdot)$$

відомий повністю.

Приклади

1. Нормальна вибірка з відомою дисперсією. Невідомим параметром нормального розподілу є середнє: $\theta = \mu$, дисперсія σ^2 вважається відомою. При фіксації μ щільність спостережень відома, тому гіпотеза

$$H_0 : \theta = \mu_0$$

– проста.

2. Нормальна вибірка з невідомою дисперсією. Якщо параметр $\theta = (\mu, \sigma^2)$, то при виконанні гіпотези

$$H_0 : \mu = \mu_0$$

вибіркові щільності можуть мати різні дисперсії, тому ця гіпотеза не є простою.

Приклади

1. Нормальна вибірка з відомою дисперсією. Невідомим параметром нормального розподілу є середнє: $\theta = \mu$, дисперсія σ^2 вважається відомою. При фіксації μ щільність спостережень відома, тому гіпотеза

$$H_0 : \theta = \mu_0$$

   – проста.

2. Нормальна вибірка з невідомою дисперсією. Якщо параметр $\theta = (\mu, \sigma^2)$, то при виконанні гіпотези

$$H_0 : \mu = \mu_0$$

   вибіркові щільності можуть мати різні дисперсії, тому ця гіпотеза не є простою.

Статистика критерію, критична область

Абстрактно критерій для перевірки H_0 проти H_1 на підставі вибірки X можна визначити як функцію

$$\delta(X) : S \to \{H_0, H_1\},$$

що для кожного вектора X приймає висновок на користь однієї з гіпотез.

Однак на практиці статистичний висновок робиться на підставі розгляду значення певної функції від вибірки – статистики критерію. Ця функція

$$\hat{\kappa}(X) : S \to D$$

довільною статистикою (вимірною функцією від вибірки) зі значеннями в деякому вимірному просторі D.

На відміну від оцінки, статистика критерію може містити значення параметрів – у випадку, коли гіпотеза висувається саме щодо цих значень.

Для визначення результату статистичного висновку щодо якісної (дихотомічної) властивості розподілу спостережень на основі значення статистики $\hat{\kappa}$ досить розбити множину D на дві частини:

$$D = D_0 \cup D_1,$$

та

(0) у випадку включення $\hat{\kappa} \in D_0$ – приймати нульову гіпотезу,

(1) при $\hat{\kappa} \in D_1$ – приймати альтернативу і відхиляти нульову гіпотезу.

Критична область

Множина D_1, на якій відхиляється нульова гіпотеза, називається критичною областю статистики критерію.

Для визначення результату статистичного висновку щодо якісної (дихотомічної) властивості розподілу спостережень на основі значення статистики $\hat{\kappa}$ досить розбити множину D на дві частини:

$$D = D_0 \cup D_1,$$

та

(0) у випадку включення $\hat{\kappa} \in D_0$ – приймати нульову гіпотезу,

(1) при $\hat{\kappa} \in D_1$ – приймати альтернативу і відхиляти нульову гіпотезу.

Критична область

Множина D_1, на якій відхиляється нульова гіпотеза, називається критичною областю статистики критерію.

Очевидно, що довільний алгоритм дихотомічного вибору на підставі значення вибірки X можна подати у наведеному вигляді, обираючи, наприклад, $D = \{0, 1\}$, $D_0 = \{0\}$, $D_1 = \{1\}$ та відповідно конструюючи D-значну статистику критерію.

На практиці часто обирають

$$D = \mathbf{R}, \quad D_0 = (-\infty, x_0), \quad D_1 = [x_0, \infty).$$

У цьому випадку статистика критерію є числовою величиною, а x_0 визначає критичний рівень статистики критерію.

Нульова гіпотеза відкидається за умови перевищення статистикою критерію критичного рівня:

$$\hat{\kappa} \geq x_0.$$

У загальному випадку з критичною областю статистики D_1 можна пов'язати критичну область вибірки

$$W = \{x \in S : \hat{\kappa}(x) \in D_1\} = \{x \in S : \delta(x) = H_1\}.$$

При потраплянні вибіркового вектора X у критичну область нульова гіпотеза відкидається, у іншому випадку відкидається альтернатива.

Статистичний тест

Статистичним критерієм (статистичним тестом) називається пара

$$(\hat{\kappa}(X), D_1),$$

що утворена D-значною статистикою критерію

$$\hat{\kappa}(X)$$

та її критичною областю

$$D_1 \subset D.$$

Алгоритм перевірки статистичної гіпотези H_0 проти альтернативи H_1 за допомогою критерію $(\hat{\kappa}, D_1)$ виконується в два етапи:

(1) обчислюють значення

$$\hat{\kappa} = \hat{\kappa}(X),$$

(2) перевіряють включення

$$\hat{\kappa} \in D_1,$$

що еквівалентне $X \in W$,

(2.1) якщо воно справджується, то нульова гіпотеза H_0 на підставі спостережень X відкидається і приймається альтернатива H_1,

(2.0) якщо ж це включення не справджується, то нульова гіпотеза H_0 на підставі спостережень X не може бути відкинута, отже, приймається, а альтернатива H_1 відкидається.

Зауваження

У зв'язку зі статистичним характером перевірки гіпотез зауважимо таке.

Перевірка гіпотези ні в якому разі не є доведенням справедливості чи несправедливості припущення гіпотези.

Невдача у відхиленні H_0 означає лише, ще немає досить вагомих свідчень для відхилення H_0 – це і є зміст висновку про те, що ця гіпотеза приймається.

Про справедливість припущення можна казати лише при наявності багатосторонніх свідчень на його користь, як статистичних, так і інших.

Зміст

Рівень та потужність критерію

Наведені вище означення критерію перевірки гіпотези є суто технічними і безпосередньо не пов'язані з якістю статистичного висновку. Для визначення показників якості зауважимо, що алгоритм перевірки гіпотези спричиняє лише одне з двох можливих рішень: прийняття або відхилення нульової гіпотези.

Кожне з цих рішень може призвести до похибки.

Похибка першого роду

Похибкою першого роду статистичного критерію називається відхилення нульової гіпотези за умови, що вона справджується.

Похибкою другого роду

Похибкою другого роду називається прийняття нульової гіпотези за умови, коли справджується альтернатива.

Вказані помилкові рішення пов'язані з відповідними подіями. Їх імовірності називаються ймовірностями похибок першого та другого роду.


True positive
You're pregnant
True negative
You're not pregnant
False positive (Type I error)
You're pregnant
False negative (Type II error)
You're not pregnant

Вірогідний рівень (або критичний рівень) статистичного критерія

Нехай нульова гіпотеза має вигляд

$$H_0 : \theta \in \Theta_0,$$

а альтернатива –

$$H_1 : \theta \in \Theta_1.$$

Вірогідним рівнем (або критичним рівнем) статистичного критерія $(\hat{\kappa}, D_1)$ називається функція від θ, що задає ймовірності похибок першого роду:

$$\mathbf{P}_\theta(\hat{\kappa}(X) \in D_1) = \mathbf{P}_\theta(X \in W), \quad \theta \in \Theta_0.$$

Потужність критерію

Потужністю критерію називається ймовірність відсутності похибки другого роду (тобто ймовірність правильного – альтернативного – висновку при альтернативі), що задається функцією:

$$1 - \mathbf{P}_\theta(\hat{\kappa}(X) \in D_0) = \mathbf{P}_\theta(\hat{\kappa}(X) \in D_1)$$

$$= \mathbf{P}_\theta(X \in W), \theta \in \Theta_1.$$

P-значення

Якщо нульова гіпотеза є простою гіпотезою, то вірогідний рівень задається одним числом:

$$\mathbf{P}_{\theta_0}(X \in W),$$

оскільки при $\theta = \theta_0$ розподіл вибірки відомий. Це число називають P-значенням критерію.

Оперативна характеристикою критерію

Імовірність

$$\mathbf{P}_{\theta}(X \in W)$$

попадання вибірки у критичну область як функція від усіх значень параметра $\theta \in \Theta_0 \cup \Theta_1$ називається оперативною характеристикою критерію.

Маючи на меті одночасну мінімізацію ймовірностей похибок першого та другого роду, для створення якісного критерію треба відшукати таку статистику $\hat{\kappa}(X)$, яка була б "чутливою" до істинного значення параметра:

- при виконанні нульової гіпотези набувала переважно значень із множини D_0 (наприклад, "досить помірних" значень) – що зменшує ймовірність похибки першого роду,
- при виконанні альтернативи набувала переважно значень у доповненні $D_1 = D \backslash D_0$ (наприклад, "надмірно великих значень") – що зменшує ймовірність похибки другого роду.

Очевидно, що задача побудови оптимального статистичного критерію зводиться до одночасної мінімізації ймовірностей похибок першого та другого роду, або ж до мінімізації його вірогідного рівня при максимізації потужності. Ця задача є внутрішньо суперечливою, оскільки і рівень, і потужність (як значення однієї оперативної характеристики на різних множинах) одночасно монотонно (у напрямку зростання) залежать від критичної області вибірки W.

Зокрема, при $W = \emptyset$ імовірність похибки першого роду нульова, другого роду дорівнює одиниці, а потужність – нульова. При $W = S$ має місце протилежне: похибка першого роду і потужність одночасно одиничні. Тому задачу відшукання оптимального критерію зводять до задачі умовної оптимізації, тобто знаходженню найбільшої потужності при обмеженні похибки першого роду.

Альтернативним до викладеного є баєсівський підхід до побудови оптимального критерію. Він полягає у постулюванні апріорного розподілу між нульовою гіпотезою та альтернативою, причому оптимальність зводиться до мінімізації середньої похибки, що отримується відповідним усередненням імовірностей похибок першого та другого роду.

Ще одну альтернативу становить мінімаксний підхід, згідно з яким оптимальним є критерій, що мінімізує найбільшу з імовірностей похибок першого та другого роду.

Вірогідний рівень критерія

Критерій $(\hat{\kappa}, D_1)$ має вірогідний рівень α, якщо ймовірності похибок першого роду не перевищують α:

$$\mathbf{P}_\theta(\hat{\kappa}(X) \in D_1) = \mathbf{P}_\theta(X \in W) \leq \alpha, \quad \forall \theta \in \Theta_0,$$

і хоча б одна з цих нерівностей є рівністю (хоча б для одного θ).

Зауваження

Якщо нульова гіпотеза є простою гіпотезою, то в означенні рівня критерію обов'язково має місце рівність. Дійсно, в такому випадку маємо єдине значення параметру і одну нерівність, яка за озна- ченням перетворюється на рівність.

Вірогідний рівень критерія

Критерій $(\hat{\kappa}, D_1)$ має вірогідний рівень α, якщо ймовірності похибок першого роду не перевищують α:

$$\mathbf{P}_\theta(\hat{\kappa}(X) \in D_1) = \mathbf{P}_\theta(X \in W) \leq \alpha, \quad \forall \theta \in \Theta_0,$$

і хоча б одна з цих нерівностей є рівністю (хоча б для одного θ).

Зауваження

Якщо нульова гіпотеза є простою гіпотезою, то в означенні рівня критерію обов'язково має місце рівність. Дійсно, в такому випадку маємо єдине значення параметру і одну нерівність, яка за озна- ченням перетворюється на рівність.

Незсунутий критерій

Критерій $(\hat{\kappa}, D_1)$ є незсунутим критерієм, якщо його потужність не менша за його рівень α:

$$\mathbf{P}_\theta(\hat{\kappa}(X) \in D_1) = \mathbf{P}_\theta(X \in W) \geq \alpha, \quad \forall \theta \in \Theta_1.$$

Зауваження

Як і у випадку статистичних оцінок, послідовність критеріїв $(\hat{\kappa}_n, D_1)$, що побудовані для кожного n за кратною вибіркою об'єму n, за одним алгоритмом, також будемо називати критерієм.

ПИТАННЯ?