(1)

Звичайно друге. За кожного з зазначених принципів ми могли б 7 разів промахнутись і 3 рази влучити, але ймовірність такого результату стрільця буде найбільшою коли $p=0{,}3$.

Ці міркування ~~лягли~~ лягли в основу методу максимальної вірогідності.

Припустимо спочатку, що розподіл F_θ дискретний, і позначимо $f(\theta, t) = P(\xi_1 = t)$. Має сенс розглядати лише ті значення t для яких ці ймовірності додатні.

Продовження лекції №2

Нехай далі $t = (t_1, \dots, t_n)$

$$f(\theta, t) = P(\xi_1 = t_1, \dots, \xi_n = t_n) = \prod_{i=1}^{n} f(\theta, t_i)$$

– ймовірність того, що вибірка набуде конкретне значення $(t_1, \dots, t_n)$.

Якщо в результаті експериментів реалізувалася вибірка ξ, то підставивши її у функцію f, отримаємо $f(\theta, \xi)$, що за умови фіксованого значення вибірки дорівнює ймовірності її появи.

Функція $f(\theta, \xi)$ називається функцією вірогідності.

непарні 25.02.2011 непарний тиж.

Ідея методу полягає в наступному: ми підбираємо таке значення θ, за якого ймовірність отримати нашу вибірку максимальна. Інакше, ми підбираємо найвірогіднішу з точки зору отриманого результату значення параметру.

Аналітично це означає, що ми маємо (8) дослідити на максимум функцію вірогідності і взяти за оцінку методу максимальної вірогідності (ММВ – оцінки) те значення θ^*, за якого

$$f(\theta^*, \xi) = \max_{\theta} f(\theta, \xi).$$

Тому, що $f(\theta, \xi) = \prod_{i=1}^{n} f(\theta, \xi_i)$, то часто досліджувати цю функцію на максимум зручніше, попередньо взявши від неї логарифм.

03.03.2020(2)

$$l(\theta, \xi) = \ln f(\theta, \xi) = \sum_{i=1}^{n} \ln f(\theta, \xi).$$

Функція $l(\theta, \xi)$ назив. логарифмічною функцією вірогідності. Точки максимуму $l(\theta, \xi)$ і $f(\theta, \xi)$ співпадають, а з сумами працювати зручніше ніж з добутками.

Якщо похідна за θ існує і неперервна, то точку екстремуму можливо знайти із рівняння

$$\frac{\partial l(\theta, \xi)}{\partial \theta} = 0.$$

Переконавшись, що в знайденій точці дійсно досягається максимум, а не мінімум, ми знаходимо методом максимальної вірогідності оцінку як розв'язок цього рівняння.

9

Нехай тепер функції розподілу F_θ абсолютно неперервна. Позначимо $f(\theta,t)$ відповідну їй щільність розподілу.

Нехай для $t=(t_1,\dots,t_n)$

$$f(\theta,t)=\prod_{i=1}^{n} f(\theta,t_i)$$

– щільність розподілу випадкового вектора ξ.

По аналогії з дискретним випадком функцією вірогідності будемо називати $f(\theta,\xi)$. ММВ-оцінкою називається те значення $\theta=\theta^*$, яке максимізує функцію вірогідності:

$$f(\theta^*,\xi)=\max_{\theta} f(\theta,\xi).$$

Можливо ввести логарифмічну функцію вірогідності $l(\theta,\xi)=\ln f(\theta,\xi)$ і працювати з нею.

Приклад. Нехай $\xi\in E_\alpha$ (експоненціальний закон). Через те, що всі спостереження за таких умов додатні, то маємо

$$f(\alpha,\xi)=\prod_{i=1}^{n}\alpha e^{-\alpha\xi_i}=\alpha^n e^{-\alpha(\xi_1+\dots+\xi_n)}$$

Диференціювання приведе нас до точки ~~екстремуму~~ максимуму.

$$l(\alpha,\xi)=n\ln\alpha-\alpha(\xi_1+\dots+\xi_n);$$

$$\frac{\partial l(\alpha,\xi)}{\partial\alpha}=\frac{n}{\alpha}-(\xi_1+\dots+\xi_n)=0 \quad \alpha^*=\frac{n}{\xi_1+\dots+\xi_n}=\frac{1}{\bar{\xi}}.$$

Зауваження

1. Якщо $\theta = (\theta_1, \dots, \theta_k)$, то функція максимальної вірогідності досліджується на максимум, як функція k-змінних. Наприклад, пошук максимуму за допомогою диференціювання приведе до системи рівнянь

$$\begin{cases} \dfrac{\partial l(\theta, \xi)}{\partial \theta_1} = 0, \\ \dots \\ \dfrac{\partial l(\theta, \xi)}{\partial \theta_k} = 0. \end{cases}$$

2. Якщо функція вірогідності досягає максимального значення в декількох точках, то всі вони, за визначенням, вважаються ММВ-оцінками.

3. ММВ-оцінки, як правило, є асимптотично <u>незміщеними</u> і <u>спроможними</u>.

4. В багатьох випадках ММ-оцінки співпадають з ММВ-оцінками, але це відбувається не завжди.

10.03.2020

Л3

Порівняння оцінок

Нехай $\xi \in F_\theta$ $\theta \in \mathbb{R}$ — невідомий параметр. Отримано дві оцінки θ_1^* і θ_2^*. Обидві хороші оцінки: незміщені (або асимптотично незміщені), спроможні. Яку з них вибрати?

Оцінка точніша, якщо відхилення оцінки від невідомого параметра менше.

Для досліджень застосовуватимемо середньоквадратичне відхилення.

<u>Визначення</u>. Будемо вважати, що оцінка θ_1^* краща за оцінку θ_2^*, якщо за усіх значень θ

$$M_\theta(\theta_1^*-\theta)^2 \le M_\theta(\theta_2^*-\theta)^2$$

і хоча б за одного значення θ нерівність строга.

Якщо оцінка θ^* незміщена, то $M_\theta(\theta^*-\theta)^2 = D_\theta\theta^*$.

Отже кращою оцінкою є та, у якої дисперсія менша.

<u>Визначення</u>. Якщо серед незміщених оцінок (а їх нескінченно багато) знайдеться та у якої дисперсія мінімальна, то така оцінка називається <u>ефективною</u>.

Ефективні оцінки є найточнішими. Необхідно здійснювати пошук саме таких оцінок. Це досить складно, але є розроблені методи.

В МС відома нерівність Рао-Крамера: якщо $f(\theta, t)$ як функція змінної θ характеризується властивостями регулярності, що для будь-якої незміщеної оцінки (2)

θ^*

$$D_\theta \theta^* \geq C(\theta, n).$$

Константа $C(\theta, n)$ легко обчислюється, після чого з нею порівнюють дисперсії оцінок. Якщо для деякої оцінки в нерівності Рао-Крамера виконується рівність, то ця оцінка має найменшу дисперсію.

Довірчі інтервали

Деякі розподіли, повязані з нормальним

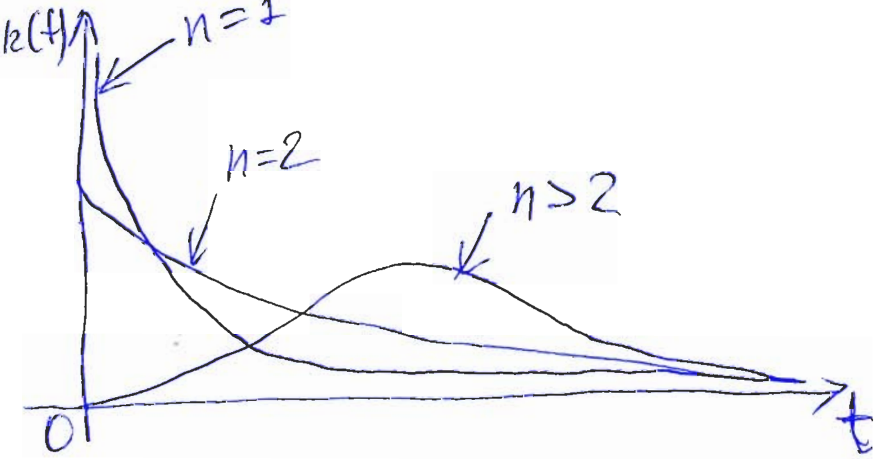
Розглянемо деякі розподіли, що відіграють важливу роль у МС.

I. Розподіл χ-квадрат.

Розподілом χ^2-квадрат с n ступенями свободи називається $\chi^2_n = \Gamma_{1/2, n/2}$. Цей розподіл має щільність

$$\gamma_{1/2, n/2}(t) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} t^{\frac{n}{2}-1} e^{-\frac{t}{2}}, & t > 0, \\ 0, & t \leq 0 \end{cases}$$

Графік $\gamma_{1/2, n/2}(t)$



Властивості розподілу χ-квадрат. (3)

1. Якщо випадкові величини ξ_1 і ξ_2 незалежні, $\xi_1 \in \chi^2_{n_1}$, $\xi_2 \in \chi^2_{n_2}$, то $\xi_1 + \xi_2 \in \chi^2_{n_1+n_2}$.

2. Нехай випадкові величини $\xi_1, \dots, \xi_n$ незалежні і $\xi_i \in \Phi_{0,1}$, коли $i = 1, \dots, n$. Тоді

$$\xi_1^2 + \dots + \xi_n^2 \in \chi^2_n.$$

Доведення. Достатньо перевірити твердження для $n=1$, а потім застосувати властивість 1°. Маємо $F_{\xi_1^2}(x) = P(\xi_1^2 < x) = 0$ для $x \le 0$.

Якщо $x > 0$, то

$$F_{\xi_1^2}(x) = P(\xi_1^2 < x) = P(-\sqrt{x} < \xi_1 < \sqrt{x}) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{x}}^{\sqrt{x}} \exp\left\{-\frac{u^2}{2}\right\} du = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{x}} \exp\left\{-\frac{u^2}{2}\right\} du.$$

Здійснемо заміну $t = u^2$ і отримаємо

$$F_{\xi_1^2}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x t^{-\frac{1}{2}} \exp\left\{-\frac{t}{2}\right\} dt = \Gamma_{1/2, n/2}(x).$$

3. Коли значення n велике розподіл $\chi^2_n(x)$ апроксимується нормальним розподілом

Дійсно застосовуючи позначення попереднього твердження, знаходимо

$$\chi^2_n(x) = P(\xi_1^2 + \dots + \xi_n^2 < x) =$$

$$P\left(\frac{\xi_1^2+\dots+\xi_n^2-nM\xi_1^2}{\sqrt{nD\xi_1^2}}<\frac{x-nM\xi_1^2}{\sqrt{nD\xi_1^2}}\right).$$

Якщо n велике, то можливо застосувати Центральну граничну теорему (ЦГТ).
Тут $M\xi_1^2=1$, $M\xi_1^4=3$, тому $D\xi_1^2=2$.
Маємо, що

$$\chi_n^2(x)\simeq\Phi_{0,1}\left(\frac{x-n}{\sqrt{2n}}\right).$$

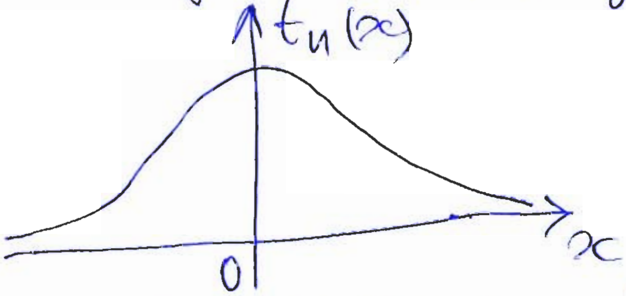
II Розподіл Ст'юдента.
Нехай випадкові величини ξ і η_n незалежні, $\xi\in\Phi_{0,1}$, $\eta_n\in\chi_n^2$. Тоді розподіл T_n випадкової величини

$$T_n=\frac{\xi}{\sqrt{\frac{1}{n}\eta_n}}$$

називається розподілом Ст'юдента з n ступенями свободи.
Замість випадкової величини η_n у цьому визначенні можна застосувати $\xi_1^2+\dots+\xi_n^2$, де $\xi, \xi_1, \dots, \xi_n$ незалежні і $\xi_i\in\Phi_{0,1}$, коли $i=1,\dots,n$.
Щільність розподілу Ст'юдента має такий вигляд

$$t_n(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}.$$

Графіком щільності є симетрична крива, (5) яка нагадує графік щільності розподілу Коші. Розподіл Коші співпадає з розподілом Стьюдента з одним ступенем свободи.



Розподіл введений англійським математиком В. С. Госсетом, який підписував роботи псевдонімом Student.

Доводиться, що $t_n(x) \to \varphi_{0,1}(x)$, коли $n \to \infty$.

Із закону великих чисел і властивостей збіжності за ймовірністю випливає, що

$$\frac{\xi}{\sqrt{\frac{\xi_1^2 + \dots + \xi_n^2}{n}}} \xrightarrow{P} \xi \in \Phi_{0,1}, \text{ через те, що}$$

$$\frac{\xi_1^2 + \dots + \xi_n^2}{n} \xrightarrow{P} M\xi_1^2 = 1.$$

Випадкові величини $\xi, \xi_1, \dots, \xi_n$ незалежні і мають стандартний нормальний розподіл.

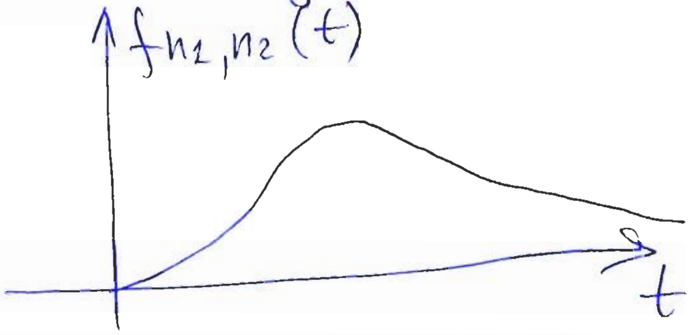
$\underline{\text{III}}$ Розподіл Фішера. Нехай η_1 і η_2 незалежні, $\eta_1 \in \chi^2_{n_1}$, $\eta_2 \in \chi^2_{n_2}$. Тоді розподіл випадкової величини

$$\frac{\eta_1/n_1}{\eta_2/n_2}$$

КУ21, КІ21, КБ21, КБ22
10.03.2020
Дав визначення розподілу Фішера

називається розподілом Фішера з числом ступенів свободи (n_1, n_2) і позначається F_{n_1,n_2}

Графік щільності розподілу Фішера приблизно має вигляд



Всі три розподіли широко застосовуються в статистичних розрахунках. Є таблиці.

Властивості вибірок з нормальним розподілом.

Наступна лема є основою багатьох статистичних висновків.

Лема Фішера. Нехай випадкові величини $\xi_1, \dots, \xi_n$ незалежні, $\xi_i \in \Phi_{0,1}$, $i = 1, \dots, n$ і

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = A \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix},$$

де A — ортогональна матриця. Тоді для будь-якого $r = 1, \dots, n-1$

$$\sum_{i=1}^{n} \xi_i^2 - \eta_1^2 - \dots - \eta_r^2 \in \chi^2_{n-r}$$

і ця випадкова величина не залежить від $\xi_1, \dots, \xi_r$.

Доведення. Раніше було встановлено 7, що в наших умовах випадкові величини $\xi_1, \dots, \xi_n$ незалежні і мають розподіл $\Phi_{0,1}$. Ортогональне перетворення не змінює довжини вектора: $\sum_{i=1}^{n} \xi_i^2 = \sum_{i=1}^{n} \eta_i^2$, тому

$$\sum_{i=1}^{n} \xi_i^2 - \eta_1^2 - \dots - \eta_r^2 = \eta_{r+1}^2 + \dots + \eta_n^2 \in \chi^2_{n-r}$$

і ця випадкова величина не залежить від $\eta_1, \dots, \eta_r$. □

Теорема про властивості вибірок із нормального розподілу.

Нехай $\xi = (\xi_1, \dots, \xi_n) \in \Phi_{a,\sigma^2}$.

Тоді

1) $\dfrac{\bar{\xi} - a}{\sigma}\sqrt{n} \in \Phi_{0,1}$;

2) $\dfrac{nS^2}{\sigma^2} \in \chi^2_{n-1}$;

3) $\bar{\xi}$ і S^2 незалежні.

Доведення. 1. Відомо (доведено), що $\bar{\xi} \in \Phi_{a, \frac{\sigma^2}{n}}$. Застосуємо операцію стандартизації:

$$\frac{\bar{\xi} - M\bar{\xi}}{\sqrt{D\bar{\xi}}} = \frac{\bar{\xi} - a}{\sigma}\sqrt{n} \in \Phi_{0,1}.$$

8

2. Позначимо $\eta_i = \frac{\xi_i - a}{\sigma}$, $i = 1, \dots n$.
Тоді $\eta_i \in \Phi_{0,1}$, $\frac{\overline{\xi} - a}{\sigma} = \overline{\eta}$ і

$$\frac{S^2}{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\xi_i - \overline{\xi}}{\sigma}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\xi_i - a}{\sigma} - \frac{\overline{\xi} - a}{\sigma}\right)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\eta_i - \overline{\eta})^2 = \frac{1}{n}\sum_{i=1}^{n}\eta_i^2 - (\overline{\eta})^2.$$

Тому $$\frac{nS^2}{\sigma^2} = \sum_{i=1}^{n}\eta_i^2 - (\sqrt{n}\,\overline{\eta})^2.$$

З іншого боку
$$\sqrt{n}\,\overline{\eta} = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix}.$$

Вектор $\left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$ має довжину 1.
Цей вектор завжди можливо добудувати до ортогональної матриці A, в якій він буде першим рядком. Тоді $\sqrt{n}\,\overline{\eta}$ буде співпадати з першою компонентою вектора $A(\eta_1, \dots, \eta_n)^T$ та за лемою Фішера маємо
$$\sum_{i=1}^{n}\eta_i^2 - (\sqrt{n}\,\overline{\eta})^2 \in \chi^2_{n-1}.$$
Із леми Фішера випливає також, що $\frac{nS^2}{\sigma^2}$ і $\sqrt{n}\,\overline{\eta} = \frac{\overline{\xi} - a}{\sigma}\sqrt{n}$ незалежні в.в.

Таким чином S^2 і $\bar{\xi}$ незалежні, як (9)
~~як~~ функції зазначених випадкових величин.

<u>Наслідок</u>. В умовах теореми

$$\frac{\bar{\xi}-\alpha}{S}\sqrt{n-1} \in T_{n-1}.$$

<u>Доведення</u>. У відповідності з визначенням розподілу Ст'юдента, T_{n-1} – це розподіл дробу

$$\frac{\eta}{\sqrt{\frac{\zeta}{n-1}}},$$

де η і ζ незалежні, $\eta \in \Phi_{0,1}$ і $\zeta \in \chi^2_{n-1}$. У відповідності з теоремою можемо взяти

дзета

$$\eta = \frac{\bar{\xi}-\alpha}{\sigma}\sqrt{n}, \quad \zeta = \frac{nS^2}{\sigma^2}. \ \square$$