Конспекти лекцій зі статистичного висновування

Yehor KOROTENKO

16 March 2026

Зміст

Вступ

§1

1.1 Оцінювання

- 0.4 Поточний контроль +0.6 Іспит.
- Розподіл : 80% проміжний іспит, 20% Контрольна робота (26/01).

1.2 Статистична Модель

Означення 1.1 (Статистична Модель) – *Статистична модель — це ймовірнісний простір $(\Omega, \mathcal{A}, \mathcal{P})$ де $\mathcal{P}$ є сімейством розподілів ймовірностей $\{P_\theta; \theta \in \Theta\}$.*

- Якщо $\exists p \in \mathbb{N}^*, \Theta \subset \mathbb{R}^p$: параметрична модель.
- Інакше : непараметрична модель.

Приклад 1.2 (Сімейства розподілів) –

- *Розподіли Пуассона : $\mathcal{P} = \{P(\lambda); \lambda > 0\}$.*
- *Регулярна щільність :*
  $\mathcal{P} = \{\mathbb{P}; \mathbb{P}$ щільність якої допускає обмежену другу похідну$\}$.

◇

Означення 1.3 (Спостереження) – *Спостереження — це випадкова величина (в.в.), розподіл якої належить до $\{P_\theta, \theta \in \Theta\}$. Наше спостереження матиме структуру n-вибірок $X_1, ..., X_n$ н.о.р. (незалежних та однаково розподілених) із спільним розподілом $\in \{P_\theta, \theta \in \Theta\}$.*

Зауваження 1.4 – *$(X_1, ..., X_n)$ має розподіл $P_\theta^{\otimes n}$. Вибірка містить всю інформацію про P_θ, отже, про θ.* ◇

Означення 1.5 (Ідентифіковність) – *Модель є ідентифікованою тоді і тільки тоді (т.т.т.), якщо відображення $\theta \mapsto P_\theta$ є ін'єктивним.*

1.3 Оцінювачі

Гіпотеза : Спостерігаємо $X_1, ..., X_n$ i.i.d. зі спільним розподілом $\in \{P_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$ (ідентифікована параметрична модель). Нехай θ^* є справжнє невідоме значення таке, що $P_{X_i} = P_{\theta^*}$.

Означення 1.6 (Оцінювач) – *Оцінювач для θ — це функція вибірки $(X_1, ..., X_n)$, вимірна та незалежна від θ (обчислювана на основі даних).*

Позначення : $\hat{\theta} = \hat{\theta}_n = h(X_1, ..., X_n)$. Це випадкова величина.
Приклади : $\hat{\theta} = \overline{X}$, $\hat{\theta} = X_1 - X_3$, тощо.

Фундаментальні питання :

1. Як визначити хороший оцінювач ?
2. Як побудувати хороший оцінювач ?

1.4 Квадратичний ризик

Ідея : В середньому, $\hat{\theta}$ має бути близьким до θ. Ми розглядаємо $\mathbb{E}\big[\hat{\theta}-\theta\big]$.

Означення 1.7 (Зміщення) – *Зміщення $\hat{\theta}$ визначається як :*

$$B\big(\hat{\theta},\theta\big)=\mathbb{E}\big[\hat{\theta}\big]-\theta$$

Кажуть, що $\hat{\theta}$ є ***незміщеним*** *тоді і тільки тоді, якщо $B\big(\hat{\theta},\theta\big)=0$.*

Означення 1.8 (Квадратичний ризик / MSE) –

$$R\big(\hat{\theta},\theta\big)=\mathbb{E}\Big[\big(\hat{\theta}-\theta\big)^2\Big]$$

Це ***середньоквадратична помилка (MSE)*** *англійською.*

Кажуть, що $\hat{\theta}_1$ кращий за $\hat{\theta}_2$ тоді і тільки тоді, якщо $R\big(\hat{\theta}_1,\theta\big)\leq R\big(\hat{\theta}_2,\theta\big)$.

1.4.1 Приклад : Модель Пуассона

Нехай $X_1,...,X_n$ мають Пуассонівський розподіл P_θ , $\theta>0$. Шукаємо оцінку для $\theta=\mathbb{E}[X_i]$.

Запропонуємо : $\hat{\theta}=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.

Обчислення зміщення :

$$\begin{aligned}
B\big(\hat{\theta},\theta\big)&=\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]-\theta\\
&=\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i]-\theta\quad\text{(за лінійністю)}\\
&=\frac{1}{n}\cdot n\cdot\mathbb{E}[X_1]-\theta\\
&=\theta-\theta=0
\end{aligned}$$

Отже $\mathbb{E}\big[\overline{X}\big]=\theta$ є незміщеною оцінкою.

Обчислення ризику :

$$\begin{aligned}
R\big(\hat{\theta},\theta\big)&=\mathbb{E}\Big[\big(\overline{X}-\theta\big)^2\Big]=\mathbb{E}\Big[\big(\overline{X}-\mathbb{E}\big[\overline{X}\big]\big)^2\Big]\\
&=\mathrm{Var}\big(\overline{X}\big)=\mathrm{Var}\Big(\frac{1}{n}\sum X_i\Big)\\
&=\frac{1}{n^2}\sum\mathrm{Var}(X_i)\quad\text{(оскільки i.i.d)}\\
&=\frac{1}{n^2}\cdot n\cdot\mathrm{Var}(X_1)=\frac{\mathrm{Var}(X_1)}{n}=\frac{\theta}{n}
\end{aligned}$$

Теорема 1.9 (Розклад ризику на зміщення-дисперсію) –

$$R\big(\hat{\theta},\theta\big)=\big(B\big(\hat{\theta},\theta\big)\big)^2+Var\big(\hat{\theta}\big)$$

Доведення.

$$\begin{aligned}
R\left(\hat{\theta},\theta\right) &= \mathbb{E}\left[\left(\hat{\theta}-\theta\right)^2\right] \\
&= \mathbb{E}\left[\left(\hat{\theta}-\mathbb{E}\left[\hat{\theta}\right]+\mathbb{E}\left[\hat{\theta}\right]-\theta\right)^2\right] \\
&= \mathbb{E}\left[\left(\hat{\theta}-\mathbb{E}\left[\hat{\theta}\right]\right)^2\right]+\mathbb{E}\left[\left(\mathbb{E}\left[\hat{\theta}\right]-\theta\right)^2\right]+2\mathbb{E}\left[\left(\hat{\theta}-\mathbb{E}\left[\hat{\theta}\right]\right)\left(\mathbb{E}\left[\hat{\theta}\right]-\theta\right)\right] \\
&= \mathrm{Var}\left(\hat{\theta}\right)+\left(B\left(\hat{\theta},\theta\right)\right)^2+2\left(\mathbb{E}\left[\hat{\theta}\right]-\theta\right)\underbrace{\mathbb{E}\left[\hat{\theta}-\mathbb{E}\left[\hat{\theta}\right]\right]}_{0} \\
&= \mathrm{Var}\left(\hat{\theta}\right)+B\left(\hat{\theta},\theta\right)^2
\end{aligned}$$

□

1.5 Послідовність

Асимптотична властивість. Розглядаються лише послідовні оцінки.

Означення 1.10 (Послідовність) – *Нехай $(X_1, ..., X_n)$ н.о.р. з розподілом P_θ. Нехай $\hat{\theta}_n = h(X_1, ..., X_n)$. $\hat{\theta}_n$ є послідовною оцінкою (або збіжною) для θ тоді і тільки тоді, коли :*

$$\hat{\theta}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \theta$$

Зауваження 1.11 – *$\hat{\theta}_n$ є сильно послідовною тоді і тільки тоді, коли $\hat{\theta}_n \xrightarrow[n\to+\infty]{p.s.} \theta$.* ◇

1.5.1 Приклад : Повернення до моделі Пуассона

$\Theta = \mathbb{R}_+^*$, $\hat{\theta}_n = \overline{X}$.

- Можна покликатися на Закон великих чисел (ЗВЧ) : $\overline{X} \xrightarrow{\mathbb{P}} \mathbb{E}[X_i] = \theta$.
- Через квадратичний ризик :

$$R\left(\hat{\theta}_n,\theta\right) = \mathrm{Var}\left(\overline{X}\right) = \frac{\theta}{n} \underset{n\to+\infty}{\longrightarrow} 0$$

Згідно з нерівністю Чебишова-Біname :

$$P\left(|\hat{\theta}_n-\theta| > \varepsilon\right) \leq \frac{\mathbb{E}\left[\left(\hat{\theta}_n-\theta\right)^2\right]}{\varepsilon^2} = \frac{R\left(\hat{\theta}_n,\theta\right)}{\varepsilon^2} \to 0$$

1.5.2 Метод «Plug-in»

Нехай $(X_1, ..., X_n)$ — н.о.р. Пуассона(θ). Ми хочемо оцінити $\beta = P(X_i = 0) = e^{-\theta}$.

$$\hat{\beta} = e^{-\hat{\theta}} = e^{-\overline{X}}$$

$\hat{\beta}$ є слушним для оцінки β.

Лема 1.12 (Лема про неперервне відображення) – *Якщо $Z_n \xrightarrow{\mathbb{P}} Z$, тоді $h(Z_n) \xrightarrow{\mathbb{P}} h(Z)$ для будь-якої неперервної функції h.* ◇

Оцінювачі

§2

2.1 Параметрична структура

2.1.1 Параметрична статистична модель

Маємо спостереження $(X_1, ..., X_n)$, вибірку i.i.d випадкових величин (незалежних, однаково розподілених) із спільним законом P , що належить до параметризованої родини законів розподілу ймовірностей $\{P_{\theta,\theta\in\Theta\subset\mathbb{R}^p}\}$.

Зауваження 2.1 – *Якщо* $\Theta \subset$ *простір нескінченної розмірності* $\to$ *непараметрична модель.* ◇

Оцінити P означає оцінити $\theta \in \mathbb{R}^p$.

Приклад 2.2 – *Bernoulli* (θ), *Exp* (θ), $\mathcal{N}(\mu, \sigma^2)$, *закон із щільністю* $f_\theta(x) = \theta x^{\theta-1} 1_{x\in[0,1]}$ ◇

Позначення 2.3 – $E_{\theta_n}[h(X_1, ..., X_n)]$, $\Theta[h(X_1, ..., X_n)]$
Закон розподілу для $(X_1, ..., X_n) \to P_\theta^{\otimes n}$ ◇

Означення 2.4 (Оцінювач) –

$$\hat{\theta} = \hat{\theta}_n = h(X_1, ..., X_n)$$

Означення 2.5 (Якість) –

- *Ризик*

$$R\left(\hat{\theta}, \theta\right) = E_\theta\left[\left(\hat{\theta} - \theta\right)^2\right]$$

- *Зговірність*

$$\hat{\theta}_n \underset{n\to+\infty}{\overset{P}{\to}} \theta$$

Означення 2.6 (Ідентифіковна модель) –

$$\theta \to P_\theta \quad \textit{ін'єктивний}$$

2.2 Метод моментів

Означення 2.7 – *Називається* ***теоретичний момент*** *розподілу* X_i *порядку* k*:*

$$\mu_k = E\left[X_i^k\right], \quad k \geq 1$$

Означення 2.8 – *Називається* ***емпіричний момент*** *розподілу* X_i *порядку* k*:*

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

За законом великих чисел $\hat{\mu}_k \xrightarrow[n\to+\infty]{P} \mu_k$.

Метод моментів: якщо можна записати θ або $g(\theta)$ цільовий параметр як функцію від k перших теоретичних моментів.

$$\theta = \mathcal{L}(\mu_1, ..., \mu_k)$$

тоді оцінка

$$\hat{\theta} = \mathcal{L}(\hat{\mu}_1, ..., \mu_k)$$

отримується за допомогою методу.

Приклад 2.9 (Обчислення оцінок за допомогою методу моментів) –

- $X_i \sim Bernoulli(\theta)$ 0-1,

$$\theta = P(X_i = 1) = E[X_i] \to \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

- $X_i \sim Exp(\theta)$,, $f_\theta(x) = \theta e^{-\theta x}1_{x\geq 0}$, $E[X] = \frac{1}{\theta} \Longleftrightarrow \theta = \frac{1}{\mu_1}$, *за методом моментів,*

$$\hat{\theta} = \frac{1}{\hat{\mu}_1} = \frac{1}{\overline{X}}$$

$$\begin{aligned}\Theta(X_i) = \frac{1}{\theta^2} \Longleftrightarrow \theta^2 &= \frac{1}{E[X_i^2] - E[X_i]^2} \\ \Longleftrightarrow \theta &= \frac{1}{\sqrt{\mu_2 - \mu_1^2}} \\ \Longrightarrow \hat{\theta}_2 &= \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}\right)^2}}\end{aligned}$$

- $X_1, ..., X_n$ *i.i.d. з розподілу* P_θ *з щільністю*

$$f_\theta(x) = \theta x^{\theta-1}1_{x\in[0,1]}$$

$$E_\theta[X_i] = \theta\int_0^1 x^\theta \,\mathrm{d}x = \frac{\theta}{\theta+1}$$

Метод моментів:

$$(\theta+1)\mu_1 = \theta \Longleftrightarrow \theta(1-\mu_1) = \mu_1 \Longleftrightarrow \theta = \frac{E[X_i]}{1-E[X_i]}$$

$$\Longrightarrow \hat{\theta}_{\mathbb{M}} = \frac{\overline{X}}{1-\overline{X}},, P_\theta\left(\overline{X} = 1\right) = P_\theta(X_1 = X_2 = ... = X_n = 1) = 0$$

◇

2.3 Огляд L.A.C.

(L.A.C = лема неперервних відображень) $(X_n)_{n\geq 1}$ послідовність випадкових величин. Якщо X_n збігається до X, що можна сказати про $g(X_n)_{n\geq 1}$? Якщо g неперервна, L.A.C.

- якщо $X_n \xrightarrow{P} X$ то $g(X_n) \xrightarrow{P} g(X)$
- якщо $X_n \xrightarrow{\mathcal{L}} X$ то $g(X_n) \xrightarrow{\mathcal{L}} g(X)$

Зауваження 2.10 (Достатня умова) –

$$D_g = \{\textit{точки розриву } g\}$$

якщо $P(X \in D_g) = 0$, *то L.A.C. справедлива.* ◇

Приклад 2.11 –

$$g(x) = \frac{x}{1-x}$$

- *ЗВВН:* $\overline{X} \xrightarrow{P} E[X]$
- *L.A.C.:* $g(\overline{X}) = \hat{\theta}_n \xrightarrow[n\to+\infty]{P} g(E[X]) = \theta$

◇

L.A.C. для пар послідовностей випадкових величин:

- якщо $(X_n, Y_n) \xrightarrow{P} (X, Y)$, то $g(X_n, Y_n) \xrightarrow{P} g(X, Y)$, якщо $g : \mathbb{R}^2 \to \mathbb{R}$ або $\mathbb{R}^2$ неперервна
- якщо $(X_n, Y_n) \xrightarrow{\mathcal{L}} (X, Y)$, то $g(X_n, Y_n) \xrightarrow{\mathcal{L}} g(X, Y)$

Приклад 2.12 –

$$\hat{\theta}_2 = \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\overline{X}\right)^2}} \quad \textit{узгоджений?}$$

ЗВВН:

- $\overline{X} \xrightarrow{P} \mu_1$
- $\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{P} \mu_2$

отже

$$\begin{pmatrix} \overline{X} \\ \frac{1}{n}\sum_{i=1}^n X_i^2 \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$g(x,y) = \frac{1}{\sqrt{y-x^2}} \implies \hat{\theta}^{\mathbb{M}}$ *узгоджений для* θ, g *неперервна, крім* $\{(x,y) \in \mathbb{R}^2, y = x^2\}$ *з нульовою мірою.*

Але це невірно для збіжності за розподілом. ◇

Пропозиція 2.13 (Збіжність пар) –

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{P} \begin{pmatrix} X \\ Y \end{pmatrix} \quad \textit{т.т.к.} \quad \begin{cases} X_n \xrightarrow{P} X \\ Y_n \xrightarrow{P} Y \end{cases}$$

Доведення.

- $\Longrightarrow$ тоді L.A.C. $g(x, y) = x$ неперервна, отже $X_n \to X$ та $Y_n \to Y$
- $\Longleftarrow$ збіжність пари?

  $$\forall \varepsilon > 0, P(|X_n - X| + |Y_n - Y| > \varepsilon) \leq \underbrace{P\Big(|X_n - X| > \frac{\varepsilon}{2}\Big)}_{\to 0} + \underbrace{P\Big(|Y_n - Y| > \frac{\varepsilon}{2}\Big)}_{\to 0}$$

  <u>Ця обернена умова невірна для збіжності за розподілом!</u>

$\square$

2.3.1 Емпірична дисперсія

Якщо X_i мають математичне сподівання μ та дисперсію σ^2, то емпіричною дисперсією називається

$$\begin{aligned}\hat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 + \frac{1}{n}\sum_{i=1}^{n}\overline{X}^2 - \frac{2}{n}\sum_{i=1}^{n} X_i\overline{X} \\ &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 + \overline{X}^2 - 2\overline{X}\overline{X} = \tilde{\sigma}^2\end{aligned}$$

оцінювач моментів:

$$\sigma^2 = E[X_i^2] - E[X_i]^2$$

Теоретичні моменти замінюються емпіричними:

$$\to \widetilde{\sigma}^{2\mathbb{M}} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}\right)^2$$

Консистентність: $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X}\right)^2$,

$$\begin{cases}\overline{X} \xrightarrow{P} E[X] \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} E[X^2]\end{cases} \xrightarrow{\text{cv en proba}} \begin{pmatrix}\overline{X} \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2\end{pmatrix} \xrightarrow{\text{LAC}} \hat{\sigma}^2 \text{ що є консистентною оцінкою для } \operatorname{Var}(X) = E[X^2] - E[X]^2$$

Приклад 2.14 –
- *обчислити зміщення для* $\hat{\sigma}_n^2$
- *обчислити ризик для* $\hat{\sigma}_n^2$

$\diamond$

2.4 Метод максимальної правдоподібності

2.4.1 Дана модель

$(P_\theta)_{\theta\in\Theta}$ задається, якщо існує міра μ (позитивна σ -скінченна $\to X_i$ зі значеннями в E, $E = \cup E_n$ де $\mu(E_n)$ є скінченною) така, що $\forall\theta, P_\theta$ має густину відносно μ.

2.4.2 На практиці

- або E щонайбільше зліченний: μ = міра підрахунку. Якщо $\exists, \{a_1, a_2, ...\}$ такий, що $\sum_{k\geq 1} P_\theta(X_i = a_k) = 1$, то $\mu = \sum_{k\geq 1}\delta_{a_k}$ з $\delta_a(\{a\}) = 1$ як міра Дірака.

  Приклад 2.15 – *Bernoulli* (θ), $X_i = 1$, *ймовірності* $\theta \to \mu = \delta_0 + \delta_1$ *Запишемо*

$$f_\theta(x) = \underbrace{P_\theta(\{x\})}_{=1-\theta} - P_\theta(X_i = x) \quad \text{де} \quad x \in \{a_1, a_2, ...\}$$

◇

- або $E = \mathbb{R}^p$, тоді f_θ є звичайною щільністю

f_θ щільність P_θ

Означення 2.16 – *Функція $(X_1, ..., X_n)$ називається функцією правдоподібності вибірки*

$$\theta \to L_n(\theta) = \prod_{i=1}^{n} f_\theta(X_i) \quad \text{(випадкова величина)}$$

Означення 2.17 – *Оцінка максимальної правдоподібності $\hat{\theta}_{MV}$ визначається так:*

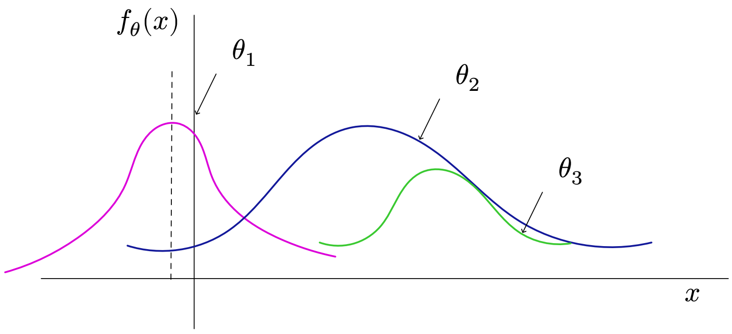
$$\forall \theta \in \Theta, L_n(\theta) \leq L_n\left(\hat{\theta}\right)$$

Ми часто працюємо з **лог-правдоподібністю**

$$\log L_n(\theta) = \sum_{i=1}^{n} \ln f_\theta(X_i) \quad \text{сума випадкових величин}$$

$$\log L_n\left(\hat{\theta}\right) = \sup_{\theta \in \Theta} \log L_n(\theta)$$

Зауваження 2.18 – $\hat{\theta}$ *є випадковою величиною*



◇

fdsa

Приклад 2.19 –

- $Bernoulli(\theta),\ f_\theta(x) = \theta^x (1-\theta)^{1-x},\ X_i$ 0-1

$$L_n(\theta) = \prod_{i=1}^{n} \theta^{X_i}(1-\theta)^{1-X_i} = \theta^{\sum_{i=1}^{n} X_i}(1-\theta)^{n-\sum_{i=1}^{n} X_i}$$

$$log\, L_n(\theta) = \left(\sum_{i=1}^{n} X_i\right) ln\, \theta + \left(n - \sum_{i=1}^{n} X_i\right) ln(1-\theta)$$

$$(log\, L_n)'(\theta) = \frac{\sum_{i=1}^{n} X_i}{\theta} - \frac{n - \sum_{i=1}^{n} X_i}{1-\theta} = \frac{\sum_{i=1}^{n} X_i - n\theta}{\theta(1-\theta)}\left(\overline{X} - \theta\right)$$

Рівняння правдоподібності:

$$(log\, L_n)'(\theta) = 0 \Longleftrightarrow (1-\theta)\sum_{i=1}^{n} X_i = \left(n - \sum_{i=1}^{n} X_i\right)\theta$$

$$\Longleftrightarrow \sum_{i=1}^{n} X_i = n\theta \Longrightarrow \theta = \frac{\sum_{i=1}^{n} X_i}{n}$$

критична точка, чи є вона максимумом?

Похідна змінює знак в $\overline{X} \rightarrow$ *ми дійсно маємо максимум* $\rightarrow \hat{\theta}^{MV} = \overline{X}$

2- , $(log\, L_n)''(\theta) < 0$ *для всіх* $\theta \Longrightarrow log\, L_n$ *є вгнутою* $\Longrightarrow$ *глобальний максимум*

$$(log\, L_n)''(\theta) = -\frac{\sum_{i=1}^{n} X_i}{\theta^2} - \frac{n - \sum_{i=1}^{n} X_i}{(1-\theta)^2} < 0,, \forall\theta$$

$\diamond$

Інформація Фішера, ефективність

§3

Нехай $(P_\theta)_{\theta\in\Theta}$, $\Theta \subset \mathbb{R}^p$ (ідентифіковане, задане). Позначимо f_θ щільність P_θ

$$\operatorname{Supp} f_\theta = \left\{x \in E,, f_{\theta(x)} > 0\right\}$$

Дано $(X_1, ..., X_n)$, н.о.р. з розподілом P_θ і $\theta \mapsto L(\theta) = \prod_{i=1}^n f_{\theta(X_i)}$ — функція правдоподібності вибірки. На $\operatorname{Supp} f_\theta$ можна обчислити

$$\log L_{n(\theta)} = \sum_{i=1}^n \log f_{\theta(X_i)}$$

$$\hat{\theta} = \operatorname{argmax}_{\theta\in\Theta} \log L_{n(\theta)}$$

Пропозиція 3.1 – *Якщо $\hat{\theta}$ ОМП*[1] *для θ, $g\left(\hat{\theta}\right)$ є ОМП для $g(\theta)$*

Мета: що може бути «кращим» оцінювачем? $\longrightarrow$ регулярна модель

3.1 Регулярна модель

Означення 3.2 – *Модель $(P_\theta)_{\theta\in\Theta}$ називається регулярною, якщо:*

1. *Θ є відкритою множиною і $\theta \mapsto f_{\theta(x)}$ є C^1*
2. *Supp f_θ не залежить від θ: $S = \left\{x,\ f_{\theta(x)} > 0\right\}$*
3. *Для будь-якого θ відображення*

$$x \mapsto \frac{\frac{\partial f_\theta}{\partial\theta}(x)}{f_{\theta(x)}} \mathbb{1}_{f_{\theta(x)}>0}$$

є інтегрованою (L, μ) і інтеграл

$$I(\theta) = \int_S \frac{\frac{\partial f_\theta}{\partial\theta}(x)}{f_{\theta(x)}} \mathbb{1}_{f_{\theta(x)}>0} dx$$

є неперервним на Θ.

Позначення 3.3 – *Позначаємо похідну від $f_{\theta(x)}$ відносно θ як: $\frac{\partial f_\theta}{\partial\theta}(x)$ Величина $I(\theta)$ називається **Інформацією Фішера моделі**.* ◇

Приклад 3.4 –

- *$f_{\theta(x)} = \theta e^{-x\theta}$ щільність відносно $\mu(dx) = \mathbb{1}_{x\geq 0} dx$*

  $\theta \mapsto \theta e^{-x\theta}$ *є C^∞ на $\Theta =]0, +\infty[$, Supp $f_\theta = \mathbb{R}_+$*

$$\frac{\partial f_\theta}{\partial\theta}(x) = (1 - x\theta)e^{-x\theta}$$

[1] **ОМП** = Оцінка Максимуму Правдоподібності

$$\frac{(1-x\theta)^2\left(e^{-x\theta}\right)^2}{\theta e^{-x\theta}} = \frac{(1-x\theta)^2}{\theta}e^{-x\theta}$$

$$\begin{aligned} I(\theta) &= \int_\theta^\infty \frac{(1-x\theta)^2}{\theta^2}\theta e^{-x\theta}dx \\ &= \frac{1}{\theta^2}E_\theta(1-X\theta)^2 \\ &= \frac{1}{\theta^2}[1-2\theta E(X)+\theta^2E(X^2)] = \frac{1}{\theta^2} \end{aligned}$$

неперервна на $]0,+\infty[$

$\diamond$

Приклад 3.5 – *Bernoulli*(θ), $x = 0.1$, $f_{\theta(0)} = 1-\theta$, $f_{\theta(1)} = \theta$, *щільність відносно* $\delta_0 + \delta_1$

Для будь-якого $x \in \{0,1\}$, $\theta \mapsto f_{\theta(x)}$ *є* C^1

$$\frac{\left(\frac{\partial f_{\theta(0)}}{\partial\theta}\right)^2}{f_{\theta(0)}} = \frac{1}{1-\theta}$$

$$\frac{\left(\frac{\partial f_{\theta(1)}}{\partial\theta}\right)^2}{f_{\theta(1)}} = \frac{1}{\theta} \Rightarrow I(\theta) = \frac{1}{1-\theta} + \frac{1}{\theta} = \frac{1}{\theta(1-\theta)}$$

неперервна на $]0,1[$

$\diamond$

Приклад 3.6 – $f_{\theta(x)} = \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x) = \frac{1}{\theta}\mathbb{1}_{[x,+\infty[(\theta)}$ *нерегулярна модель* $\diamond$

3.2 Оцінка та Інформація Фішера

$(X_1, ..., X_n)$ н.о.р. з розподілом P_θ, f_θ

Означення 3.7 – *Називаеться* ***оцінкою*** *або* ***вектором оцінки*** *похідна логарифмічної функції правдоподібності* $\frac{\partial}{\partial\theta} \log L_{n(\theta)} = S_{n(\theta)} = \sum_{i=1}^n \frac{\partial}{\partial\theta} \log f_\theta(X_i)$

Приклад 3.8 – $X_i \sim \mathcal{E}(\theta)$, $L_{n(\theta)} = \theta^n e^{-\theta\sum_i X_i}$, $\log L_n(\theta) = n\log\theta - \theta\sum_i X_i$, *отже* $S - n(\theta) = \frac{n}{\theta} - \sum_{i=1}^n X - i$ $\diamond$

Зауваження 3.9 –

$$E(S_n(\theta)) = E\left[n\left(\frac{1}{\theta} - \frac{\sum X_i}{n}\right)\right]$$

$\diamond$

Додаткова гіпотеза регулярності: (H) для будь-якого оцінювача $h(X)$ та будь-якого θ, наступні інтеграли існують і є рівними:

$$\frac{\partial}{\partial\theta}\int_S h(x)f_\theta(x)dx = \int_S h(x)\frac{\partial f_\theta}{\partial\theta}(x)dx$$

Зауваження 3.10 – *умова застосування теореми Лебега про диференціювання.*

$$h \sup_{\tilde{\theta}\in V_\theta}|\frac{\partial f_\theta}{\partial\theta}(x)| \in L_1(\mu)$$

◇

Пропозиція 3.11 – *За умови (H), оцінка є центрованою (P_θ), $n = 1$*

$$E_\theta\left[\frac{\partial}{\partial\theta}\log L_1(\theta)\right] = \int_S \frac{\partial}{\partial\theta}\log f_\theta(x)dx = \int_S \frac{\frac{\partial}{\partial\theta}f_{\theta(x)}}{\cancel{f_\theta(x)}}\cancel{f_\theta(x)}dx = \int_S \frac{\partial f_\theta}{\partial\theta}(x)dx \underbrace{=}_{(H)} \frac{\partial}{\partial\theta}\overbrace{\int f_\theta(x)dx}^{=1} = 0$$

Означення 3.12 – *Інформація Фішера, пов'язана з $(X_1, ..., X_n)$*

$$I_n(\theta) \underbrace{=}_{def} E_\theta\left[\left(\frac{\partial}{\partial\theta}\log L_{n(\theta)}\right)^2\right] \overbrace{=}^{\text{cor. de la prop 1}} = Var_\theta\left[\frac{\partial \log L_n(\theta)}{\partial\theta}\right]$$

$$(*)E_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right]^2 = \int_S \left(\frac{\frac{\partial}{\partial\theta}f_\theta(x)}{f_\theta(x)}\right)^2 f_\theta(x)dx = \int_S \frac{\left(\frac{\partial}{\partial\theta}f_\theta(x)\right)^2}{f_\theta(x)} = \text{"□□□□□ □ □□□□□□□□□ 1"}$$

Приклад 3.13 – $(X_1, ..., X_n) \sim \mathcal{E}(\theta)$, $\frac{\partial}{\partial\theta}\log L_n(\theta) = \frac{n}{\theta} - \sum_{i=1}^n X_i$

$$I_n(\theta) = E\left(\left(\frac{n}{\theta} - \sum X_i\right)^2\right) = n^2E\left[\left(\frac{1}{\theta} - \frac{\sum X_i}{n}\right)^2\right] = n^2\ Var\left(\overline{X}\right) = n^2\frac{1}{n}\frac{1}{\theta^2} = \frac{n}{\theta^2}$$

◇

Пропозиція 3.14 –

$$I_n(\theta) = nI(\theta)$$

справді,

$$I_n(\theta) = Var\left(\frac{\partial}{\partial\theta}\log L_n(\theta)\right) = Var\left(\sum_{i=1}^n \frac{\partial}{\partial\theta}\log f_\theta(X_i)\right) \underbrace{=}_{independance} = \sum_{i=1}^n Var\left(\frac{\partial}{\partial\theta}\log f_\theta(X_i)\right) =$$

$$= \underbrace{n\,Var\left(\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right)} = nI(\theta)$$

Приклад 3.15 – $(X_1, ..., X_n)$ *i.i.d* $\mathcal{P}(\theta)$, $f_\theta(x) = e^{-\theta}\frac{\theta^x}{x!}$

$$\log L_n(\theta) = -n\theta + \left(\sum X_i\right)\log\theta - \log\prod_{i=1}^n X_i!$$

$$\frac{\partial}{\partial\theta} log\, L_n(\theta) = -n + \frac{\sum X_i}{\theta} \Rightarrow I_n(\theta) = Var\left(\frac{\sum X_i}{\theta}\right) = \frac{1}{\theta^2} n\theta = \frac{n}{\theta}$$

◇

3.3 Інформація Фішера та друга похідна

Пропозиція 3.16 – *Якщо додати, що $\theta \mapsto f_\theta(x)$ є C^2 і що (H) справедливо для $\frac{\theta^2}{\partial\theta^2}$, то інформація Фішера також записується як*

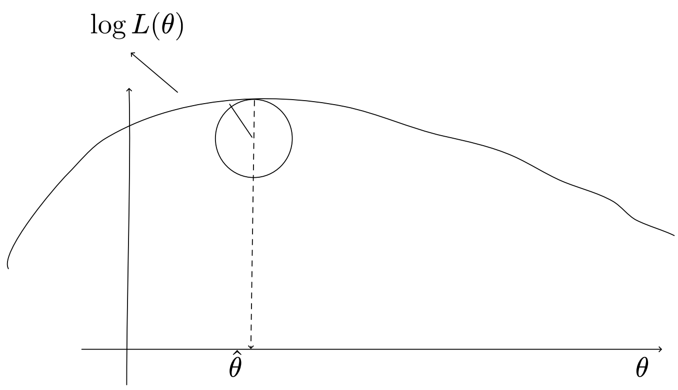
$$I_n(\theta) = -E_\theta\left[\frac{\partial^2\, log\, L_n(\theta)}{\partial\theta^2}\right]$$

якщо $\hat{\theta}$ є ОМВ, $I_{n(\hat{\theta})} > 0$

$n = 1$

$$\frac{\partial^2}{\partial\theta^2} log\, f_\theta(x) = \frac{\left(\frac{\partial^2 f_\theta(x)}{\partial\theta^2}\right)^2}{f_\theta(x)} - \frac{\left(\frac{\partial f_\theta}{\partial\theta}(x)\right)^2}{f_\theta^2(x)}$$

$$E\left[\frac{\partial^2}{\partial\theta^2} log\, f_\theta(X_1)\right] = \int_S \frac{\frac{\partial^2 f_{\theta(x)}}{\partial\theta^2}}{\cancel{f_\theta(x)}} \cancel{f_\theta(x)} dx - \underbrace{\int_S \frac{\left(\frac{\partial f_\theta(x)}{\partial}\theta\right)^2}{f_\theta^2(x)} dx}_{I(\theta)}$$



Якщо крива дуже «гостра» в ОМВ (тобто інформація Фішера велика), то ОМВ локалізується точно

3.4 Нерівність Крамера - Рао

Нехай $g(\theta)$ — параметр, що цікавить, де $g : \Theta \to \mathbb{R}$

Пропозиція 3.17 – *За припущеннями регулярної моделі, якщо для кожного θ $I(\theta) > 0$, тоді для будь-якого <u>незміщенного</u> $T = T(X_1, ..., X_n)$оцінювача, для якого $E_\theta T^2 < +\infty$, виконується*

$$\forall \theta \in \Theta,\ \underbrace{Var_\theta(T) \geq \frac{(g'(\theta)^2)}{I_n(\theta)}} = \frac{(g'(\theta))^2}{nI(\theta)}$$

Доведення.

$$\begin{aligned}
&\forall \theta \ \ E_{\theta(T)} = g(\theta) \\
&\Rightarrow \frac{\partial}{\partial \theta} E_\theta(T) = g'(\theta) \\
{\scriptstyle T=T(X_1)} &\Leftrightarrow \frac{\partial}{\partial \theta} \int_S T(x) f_\theta(x) dx = g'(\theta) \\
{\scriptstyle (H)} &\Leftrightarrow \int_S T(x) \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} f_\theta(x) dx = g'(\theta) \\
&\Leftrightarrow \int_S (T(x) - g(\theta)) \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} f_\theta(x) dx = g'(\theta)
\end{aligned}$$

$\square$

Нерівність Коші-Шварца для $\langle h_1, h_2 \rangle = \int h_1(x) h_2(x) f_\theta(x) dx$ з $h_1(X)$ та $h_2(X)$ центрованими

$$\left(\langle T(X) - g(\theta), \frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} \rangle_\theta \right)^2 = (g'(\theta))^2 \underbrace{\leq}_{\text{C.S}} \underbrace{\int (T(x) - g(\theta))^2 f_\theta(x) dx}_{= \mathrm{Var}_\theta(T)} \times \underbrace{\int \left(\frac{\frac{\partial}{\partial \theta} f_\theta(x)}{f_\theta(x)} \right)^2 f_\theta(x) dx}_{= I(\theta)}$$

Означення 3.18 – *Якщо T досягає рівності, тоді T називається* ***ефективним***.

Асимптотичне дослідження оцінювачів

§4

У параметричній моделі регулярній, якщо $\hat{\theta}_n$ оцінювач θ, тоді

$$\mathrm{Var}\left(\hat{\theta}_n\right) \geq \frac{1}{I_{n(\theta)}} = \frac{1}{nI(\theta)}$$

якщо $\mathrm{Var}\left(\hat{\theta}_n\right) = \frac{1}{nI(\theta)}$ неупереджений, $\hat{\theta}_n$ є ефективним ефективним

Асимптотика: $n \to +\infty$,

$$n \,\mathrm{Var}\left(\hat{\theta}_n\right) \underset{n\to+\infty}{\longrightarrow} \frac{1}{I(\theta)}$$

4.1 Збіжності

$(X_n)_{n\geq 0}$ послідовність дійсних випадкових величин $\left(\mathbb{R}^d\right)$

-
- Збіжність за розподілом: $X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X$ тоді й тільки тоді, коли $P(X_n \leq x) \longrightarrow P(X \leq x)$ в кожній точці неперервності x.

Лема 4.1 (lemme de Portmanteau) – *Еквівалентні характеристики:*

- *Для будь-якої неперервної обмеженої функції h,*

$$E[h(X_n)] \longrightarrow E[h(X)]$$

$\Rightarrow$ збіжність за розподілом є стійкою щодо переходу до неперервних функцій (LAC) АЛЕ загалом неправильно, що якщо $X_n \overset{\mathcal{L}}{\longrightarrow} X$ та $Y_n \overset{\mathcal{L}}{\longrightarrow} Y$, тоді $\binom{X_n}{Y_n} \overset{\mathcal{L}}{\longrightarrow} \binom{X}{Y}$

3 :

1. *якщо* $\begin{cases} \forall n, X_n \;\; et \;\; Y_n \;\; sont\ indépendantes \\ X \;\; et \;\; Y \;\; sont\ indépendantes \end{cases}$ *, тоді* $\begin{cases} convergence\ en\ loi\ de \;\; X_n \;\; et \;\; Y_n \\ convergence\ en\ loi\ du\ couple \;\; \binom{X_n}{Y_n} \end{cases}$

2.

$$\text{якщо } \begin{cases} X_n \overset{P}{\longrightarrow} X \\ Y_n \overset{p}{\longrightarrow} Y \end{cases} \Longrightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \overset{P}{\longrightarrow} \begin{pmatrix} X \\ Y \end{pmatrix} \Longrightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \overset{\mathcal{L}}{\longrightarrow} \begin{pmatrix} X \\ Y \end{pmatrix}$$

3. *(Лема Слуцького) (найважливіше)*

$$\text{якщо } \begin{cases} X_n \overset{\mathcal{L}}{\longrightarrow} X \\ Y_n \overset{\mathcal{L}}{\longrightarrow} c \end{cases} \text{, тоді } \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \overset{\mathcal{L}}{\longrightarrow} \begin{pmatrix} X \\ c \end{pmatrix}$$

застосовуючи LAX,

$$\begin{aligned} h(x,y) &= x+y & X_n+Y_n &\xrightarrow{\mathcal{L}} X+c \\ &= xy & X_nY_n &\xrightarrow{\mathcal{L}}\xrightarrow{\mathcal{L}} cX \\ &= \frac{x}{y} & \frac{X_n}{Y_n} &\xrightarrow{\mathcal{L}} \frac{X}{c} \end{aligned}$$

◇

4.2 Консистентність оцінок

Означення 4.2 – *$\hat{\theta}_n$ асимптотично незміщений тоді й лише тоді, коли*

$$Biais\left(\hat{\theta}_n,\theta\right) = E\left[\hat{\theta}_n\right] - \theta \underset{n\to+\infty}{\longrightarrow} 0$$

Зауваження 4.3 – *Збіжність за ймовірністю не передбачає збіжності математичних сподівань.*

Якщо $X_n \xrightarrow{P} X$, $|X_n| \le Y \in L'$, то за теоремою про доміновану збіжність $X_n \longrightarrow X$ в L_1 ◇

Приклад 4.4 – $\hat{\tau}_n = \frac{1}{n}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum X_i^2 - \left(\overline{X}\right)^2$ *оцінка моментів для* $\tau^2 = E[X^2)] - (E[X])^2$

$$Biais(\hat{\tau}_n^2,\tau^2) = -\frac{1}{n}\tau^2 \quad \textit{асимптотично незміщена}$$

Консистентність $\hat{\tau}_n^2$?

Інструменти для демонстрації консистентності:

- *LGN*
- *якщо $R\left(\hat{\theta}_n,\theta\right) \longrightarrow 0$ то $\hat{\theta}_n$ консистентний, оскільки збіжність $L^2 \Rightarrow$ збіжність за ймовірністю*
- *повернутися до визначення збіжності за ймовірністю*

- *якщо (X_i) є н.о.р., то (X_i^2) є н.о.р.*

$$E[X_i^2] < +\infty$$

  ЗВЧ: $\frac{1}{n}\sum_i X_i^2 \xrightarrow{P} E[X^2] = \tau^2 - \mu^2$
- $\overline{X} \xrightarrow{P} \mu$ *(ЗВЧ), з Лемою про неперервне відображення для $h(x) = x^2$: $\left(\overline{X}\right)^2 \xrightarrow{P} \mu^2$*
- *Отже* $\begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_i^2 \\ \frac{1}{n}\sum X_i \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \tau^2-\mu^2 \\ \mu \end{pmatrix}$
- *Лема про неперервне відображення $h(x,y) = x - y^2$*

Таким чином $\frac{1}{n}\sum X_i^2 - \left(\overline{X}\right)^2 \xrightarrow{P} \tau^2 + \mu^2 - \mu^2 = \tau^2$ ◇

4.3 Асимптотична нормальність

$\hat{\theta}_n$ для θ.

→ Питання: яка швидкість збіжності $\hat{\theta}_n$ до θ ?

$(X_1, ..., X_n)$ н.о.р., з математичним сподіванням θ, $\hat{\theta} = \overline{X}$ з дисперсією $\tau^2(\theta)$

<u>ЦГТ</u> $\sqrt{n}\left(\overline{X} - \theta\right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2(\theta))$ <u>незалежно від розподілу X_i</u>

Означення 4.5 – $\hat{\theta}_n$ *є <u>асимптотично нормальною</u> оцінкою тоді і тільки тоді, якщо*

- *швидкість збіжності за* $\sqrt{n}$
- *збіжність за розподілом*
- *граничний розподіл є нормальним*

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, \tau^2(\theta))$$

Приклад 4.6 – *Чи є* $\hat{\tau}_n^2$ *асимптотично нормальною ?*

$(X_1, ..., X_n)$ *н.о.р. з математичним сподіванням* μ*, дисперсією* τ^2

$$\hat{\tau}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 + \left(\overline{X} - \mu\right)^2 + \underbrace{\frac{2}{n}\sum_{i=1}^{n}(X_i - \mu)\left(\mu - \overline{X}\right)}_{=2\left(\mu - \overline{X}\right)\left(\overline{X} - \mu\right)}$$

- *ЦГТ: якщо* (X_i) *н.о.р., тоді* $(X_i - \mu)^2$ *є н.о.р. з математичним сподіванням* τ^2*,*

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - \tau^2\right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, u_4 - \tau^4)$$

$$Var(X_i - \mu)^2 = E\left[(X_i - \mu)^4\right] - \mu^4 = \mu_4 - \tau^4$$

- *ЦГТ:* $\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2)$
- $\sqrt{n}(\hat{\tau}_n^2 - \tau^2) = \sqrt{n}\left(\frac{1}{n}\sum_i (X_i - \mu)^2 - \tau^2\right) - \underbrace{\sqrt{n}\left(\overline{X} - \mu\right)^2}_{\sqrt{n}\underbrace{\left(\overline{X} - \mu\right)}_{\xrightarrow{\mathcal{L}}\mathcal{N}(0,\tau^2)} \times \underbrace{\left(\overline{X} - \mu\right)}_{\xrightarrow{\mathcal{L},P} 0}}$

$$\left.\begin{array}{l}\overline{X} - \mu \xrightarrow{\mathcal{L}} 0 \\ \sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{\mathcal{L}} U \sim \mathcal{N}(0,1)\end{array}\right\} \overset{\textit{lemme de Slutsky}}{\Longrightarrow} \sqrt{n}\left(\overline{X} - \mu\right)^2 \xrightarrow[P]{\mathcal{L}} 0$$

$$\sqrt{n}(\hat{\tau}_n^2 - \tau^2) \xrightarrow[n \to +\infty]{\mathcal{L}} Z + 0$$

Отже, $\hat{\tau}_n^2$ *є асимптотично нормальною оцінкою* ◇

Зауваження 4.7 – $\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0, \tau^2) \Longleftrightarrow \frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{\tau} \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0,1)$ *Застосування леми Слуцького: якщо* $\hat{\tau}^2$ *є консистентною оцінкою* τ^2*, тоді ми все ще маємо*

$$\frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{\hat{\tau}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

◇

Доведення.

$$\frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\hat{\tau}} = \underbrace{\left(\sqrt{n}\frac{\left(\hat{\theta}-\theta\right)}{\tau}\right)}_{\xrightarrow{\mathcal{L}} Z\sim\mathcal{N}(0,1)} \times \underbrace{\left(\frac{\tau}{\hat{\tau}}\right)}_{\xrightarrow{P} 1}$$

$$\xrightarrow{\mathcal{L}} 1\times Z \text{ за Слуцьким та консистентністю } \hat{\tau}$$

□

4.4 δ-метод

$\hat{\theta}$ асимптотично нормальна оцінка: який асимптотичний розподіл $g(\theta)$?

Лема 4.8 (дельта-метод) – *Нехай Z_n — послідовність дійсних випадкових величин, таких, що*

$$\sqrt{n}(Z_n-\mu) \xrightarrow{\mathcal{L}} Z\sim\mathcal{N}(0,\tau^2)$$

Нехай g — диференційовна функція, $g'(\mu)\neq 0$. За цих припущень маємо

$$\sqrt{n}[g(Z_n)-g(\mu)] \xrightarrow[n\to+\infty]{\mathcal{L}} \tilde{Z}\sim\mathcal{N}\left(0,\left(g'(\mu)\right)^2\tau^2\right)$$

$$g(x)=g(\mu)+g'(\mu)(x-\mu)+(x-\mu)R(x-\mu) \quad \textit{де} \quad R(y)\underset{y\to 0}{\to} 0$$

$$\sqrt{n}(g(Z_n)-g(\mu)) = \underbrace{g'(\mu)\underbrace{\sqrt{n}(Z_n-\mu)}_{\xrightarrow{\mathcal{L}} Z\sim\mathcal{N}(0,\tau^2)}}_{\longrightarrow\mathcal{N}\left(0,(g'(\mu))^2\tau^2\right)} + \underbrace{\underbrace{(\sqrt{n})(Z_n-\mu)}_{\xrightarrow{\mathcal{L}}\mathcal{N}(0,\tau^2)}\underbrace{R(Z_n-\mu)}_{\xrightarrow{P}0?}}$$

Чи маємо $Z_n \xrightarrow{P} \mu$?

$$\begin{aligned}
P(|X_n-\mu|>\varepsilon) &= P\left(\frac{\sqrt{n}|Z_n-\mu|}{\tau}>\frac{\sqrt{n}\varepsilon}{\tau}\right)\\
&= P\left(\frac{\sqrt{n}(Z_n-\mu)}{\tau}>\frac{\sqrt{n}\varepsilon}{\tau}\right)+P\left(\frac{\sqrt{n}(Z_n-\mu)}{\tau}<-\frac{\sqrt{n}\varepsilon}{\tau}\right)\\
&\sim= 1-\Phi_n\left(\frac{\sqrt{n}\varepsilon}{\tau}\right)+\Phi_n\left(-\frac{\sqrt{n}\varepsilon}{\tau}\right)=2\left(1-\Phi\left(\frac{\sqrt{n}\varepsilon}{\tau}\right)\right)
\end{aligned}$$

◇

Емпірична функція розподілу

§5

$(X_1, ..., X_n)$ — i.i.d. вибірка з дійсними значеннями, що підпорядковується невідомому закону F.

$$\forall x \in \mathbb{R}, F(x) = P(X_1 \leq x) = E\left[\mathbb{1}_{X_1 \leq x}\right]$$

Означення 5.1 – *Функція емпіричного розподілу, пов'язана з $(X_1, ..., X_n)$, визначається як:*

$$\hat{F}_n : \mathbb{R} \longrightarrow [0, 1]$$
$$x \mapsto \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$$

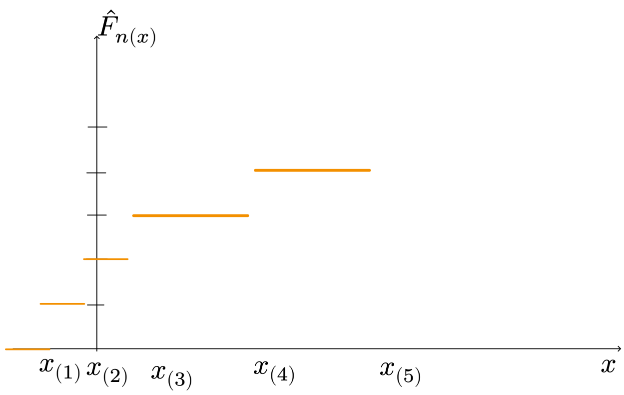
$\forall x \in \mathbb{R}, \hat{F}_n(x)$ є випадковою величиною, оцінювачем для $F(x)$.

Означення 5.2 – *Емпіричний закон розподілу $P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$ є дискретним рівномірним законом розподілу на $\{X_1, ..., X_n\}$.*

Графічне представлення

Умовно $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)} \text{ упорядковані значення}$$



fdsa

Пропозиція 5.3 (Безпосередні властивості) –

- $n\hat{F}_n(x) = \sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$ *підпорядковується біноміальному розподілу*$(n, F(x))$

- $R\left(\hat{F}_n(x), F(x)\right) = 0 + \frac{1}{n^2}\ Var\left(\sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}\right) \underbrace{=}_{indep} \frac{1}{n}F(x)(1-F(x)) \underset{n\to+\infty}{\longrightarrow} 0$ *отже* $\forall x \in \mathbb{R}, \hat{F}_n(x) \xrightarrow{P} F(x)$
- *або ж ЗВЧ:* $\hat{F}_n(x)$ *є консистентним оцінювачем* $F(x)$.
- *Ми маємо результат рівномірної збіжності:*

$$\sup_{x\in\mathbb{R}}|\hat{F}_n(x) - F(x)| \underset{n\to+\infty}{\overset{P}{\longrightarrow}} 0 \quad (\text{Теорема Гливенка-Кантеллі})$$

.[2]

- *Чи є* $\hat{F}_n(x)$ *асимптотично нормальним?*

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$$

ЦГТ: X_i *є н.о.р., тому* $\left\{\mathbb{1}_{X_i \leq x} = Y_i\right\}$ *є н.о.р.*

$$\forall x, F(x) \in]0,1[, \quad \sqrt{n}\left(\hat{F}_n(x) - F(x)\right) \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0, F(x)(1-F(x)))$$

$$\Leftrightarrow \frac{\hat{F}_n(x) - F(x)}{\sqrt{\frac{F(x)(1-F(x))}{n}}} \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1)$$

5.1 Емпіричне оцінювання «plug-in» або метод підстановки, параметр інтересу $\theta = c(F)$, емпіричний метод визначає $\hat{\theta}$, емпіричну оцінку, заміняючи F на $\hat{F}_n \to \hat{\theta}_n = c\left(\hat{F}_n\right)$.

Приклад 5.4 – $\theta = E_F(X) \to \hat{\theta}_n = E_{\hat{F}_n}(X) = \sum_{i=1}^{n} X_i \times \frac{1}{n} = \overline{X}$ *якщо* X_i *різні*

$$\theta = Var_F(X) \to \hat{\theta}_n = Var_{\hat{F}_n}(X) = \frac{1}{n}\sum\left(X_i - \overline{X}\right)^2$$

◇

5.2 Узагальнена обернена

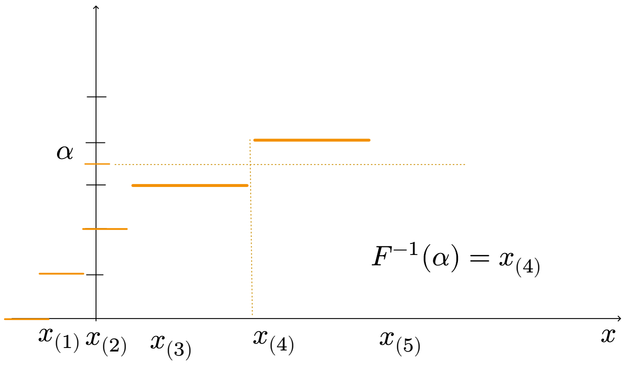
Означення 5.5 – *Узагальнена обернена функція для* F *визначається як:*

$$F^{-1} : [0,1] \longrightarrow \mathbb{R}$$

$$\forall \alpha \in [0,1], F^{-1}(\alpha) = \inf\{x \in \mathbb{R}, F(x) \geq \alpha\}$$

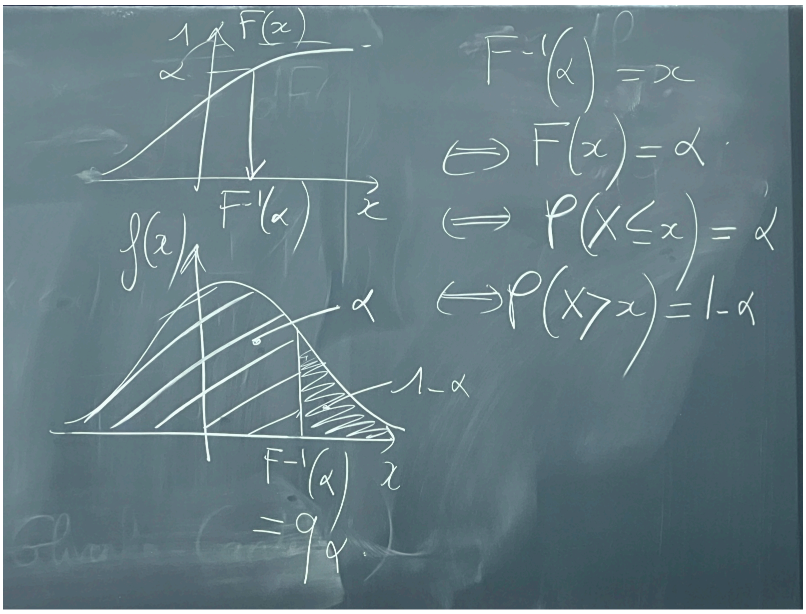
Якщо F *є строго зростаючою, то* $\inf x$ *такий, що* $F(x) \geq a \Leftrightarrow x \geq F^{-1}(\alpha)$, *якщо* F *є функцією дискретного розподілу.*

[2] ТМ Гливенка-Кантеллі: https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_de_Glivenko-Cantelli



Приклад 5.6 –

$$F^{-1}(\alpha) = x \Leftrightarrow F(x) = \alpha \Leftrightarrow P(X \leq x) = \alpha \Leftrightarrow P(X > x) = 1 - \alpha$$



◇

Словник:

- F^{-1} також називається <u>квантильною функцією</u>
- $F^{-1}(\alpha) =$ квантиль порядку α, розподілу F
- $F^{-1}(\frac{1}{4}) =$ 1-
- $F^{-1}(\frac{1}{2}) =$ медіана
- $F^{-1}(\frac{3}{4}) =$ 3-

Лема 5.7 – *Нехай U — випадкова змінна на $[0, 1]$, F — ф.р., тоді $F^{-1}(U)$ є випадковою змінною з розподілом F* ◇

- Якщо F бієктивна:

$$P(F^{-1}(U) \leq x) \underbrace{=}_{F \text{ bijective}} P(U \leq F(x)) \underbrace{=}_{\text{car } P(U \leq x) = x \text{ sur } [0,1]} F(x)$$

- Якщо F дискретна: F^{-1} узагальнена обернена: $F^{-1}(y) \leq x \Leftrightarrow y \leq F(x)$

5.3 Емпіричний квантиль

Означення 5.8 – *Визначається* <u>*емпіричний квантиль*</u> *(sample quantile) порядку α, як квантиль $\hat{F}_n$:*

$$\hat{q}_{n,\alpha} = \hat{F}_n^{-1}(\alpha) = inf\Big\{x, \hat{F}_n(x) \geq \alpha\Big\}$$

Пропозиція 5.9 –

- *Можна показати, що $\hat{q}_{n,\alpha} = X_{([n\alpha])}$ де $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ є впорядкованою вибіркою з $(X_i)_{1 \leq i < n}$*

$$[u] = \textit{найменше ціле число} \geq u$$

Приклад 5.10 – $\alpha = \frac{1}{2}$, $[\frac{n}{2}]$,

$$\begin{cases} si\ n = 2k \quad medianne = \hat{q}_{n,\frac{1}{2}} = X_{(k)} \\ si\ n = 2k+1 \quad medianne = \hat{q}_{n,\frac{1}{2}} = X_{(k+1)} \end{cases}$$

- *Узгодженість*

  якщо $\alpha \in]0,1[$, якщо F є строго зростаючою в околі α

◇

Інтервали довіри

§6

6.1 Визначення

$(X_1, ..., X_n)$ незалежні та однаково розподілені з розподілом $P \in \{P_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$, нас цікавить $\theta \in \mathbb{R}$ або $g(\theta) : \mathbb{R}^p \longrightarrow \mathbb{R}$.

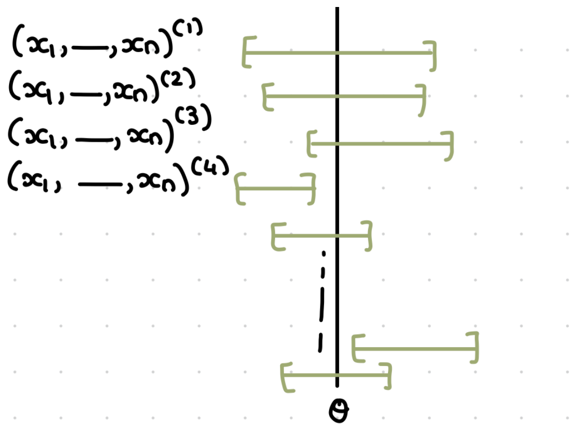
Довірчий інтервал для θ, з рівнем довіри $1 - \alpha, \alpha \in]0, 1[$ це інтервал, межі якого є <u>випадковими</u>, функціями вибірки і НЕ залежить від невідомих параметрів моделі, і такий, що

$$P([B\inf(X_1, ..., X_n); B\sup(X_1, ..., X_n)] \ni \theta) \geq 1 - \alpha$$

[3]

- ДІ можна обчислити з даних
- якщо нерівність є рівністю = рівень довіри є <u>точним</u>.
- якщо ми маємо $P(\theta \in [B\inf, B\sup]) \underset{n \to +\infty}{\longrightarrow} 1 - \alpha$, рівень є <u>асимптотичним</u>.
- зазвичай $\alpha = 1\%, 5\%$

6.2 Інтерпретація



$IC = [B\inf(X_1, ..., X_n), B\sup(X_1, ..., X_n)]$ математична формула, яка гарантує рівень $1 - \alpha$. Спостерігаємо $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$, реалізація випадкової вибірки. Обчислюємо $IC = [2.3; 5.1]$ з рівнем довіри 95% ($\alpha = 5\%$).

, 100 ( )
5 інтервалів, які не містять θ.

$$P(\theta \in [B\inf, B\sup]) = 1 - \alpha$$

$\cancel{P(\theta \in [2.3, 5.1]) = 95\%}$ оскільки θ число

6.3 Півотальний метод

$(X_1, ..., X_n)$ незалежні та однаково розподілені з математичним сподіванням $\theta \in \mathbb{R}$ та дисперсією $\sigma^2(\theta)$. Нехай $\hat{\theta}$ є асимптотично нормальною:

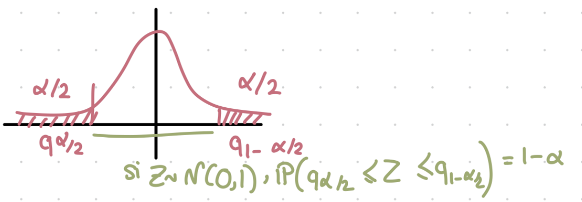
$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \underset{n \to +\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0, \sigma^2(\theta))$$

$$\Leftrightarrow \frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{\sigma(\theta)} \underset{n \to +\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0, 1)$$

За визначенням гаусових квантилів, $q_\alpha = \Phi^{-1}(\alpha)$, де Φ — функція розподілу $\mathcal{N}(0, 1)$

[3] $B\inf$ для нижньої межі та $B\sup$ для верхньої межі

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\sigma(\theta)} \leq q_{1-\frac{\alpha}{2}}\right) \underset{n\to+\infty}{\longrightarrow} 1-\alpha$$



- **півот або півотальна статистика** $= \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\hat{\sigma}}$ — стандартизована статистика, отримана з $\hat{\theta}$, де $\sigma^2(\theta)$ оцінюється за допомогою $\hat{\sigma}^2$, консистентного для оцінки $\sigma^2(\theta)$.

  Якщо це так,

$$\underbrace{\frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\sigma(\theta)}}_{\underset{\hat{\theta}\text{ as. normal}}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1)} \times \underbrace{\frac{\sigma^2(\theta)}{\hat{\sigma}^2}}_{\underset{\text{estimateur consistant}}{\overset{P}{\longrightarrow}} 1} \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1) \text{ par lemme de Slutsky}$$

- З цього випливає, що

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\hat{\sigma}}(\theta) \leq q_{1-\frac{\alpha}{2}}\right) \underset{n\to+\infty}{\longrightarrow} 1-\alpha$$

$$P\left(\hat{\theta} - \frac{1}{\sqrt{n}}\hat{\sigma}q_{1-\frac{\alpha}{2}} \leq \theta \leq \hat{\theta} - \frac{1}{\sqrt{n}}\hat{\sigma}q_{\frac{\alpha}{2}}\right) \longrightarrow 1-\alpha$$

Доповнення (перед іспитом)

§7

1. Повернення до асимптотичної нормальності
2. Приклад
3. Асимптотичний опорний елемент
4. 2

7.1 Асимптотичні властивості послідовності оцінок $\left(\hat{\theta}_n\right)_{n\geq 1}$

- Збіжність $\hat{\theta}_n \xrightarrow{P} \theta$
- Асимптотична нормальність, якщо існує $\sigma^2 > 0$

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \underset{n\to+\infty}{\xrightarrow{\mathcal{L}}} \mathcal{N}(0, \sigma^2)$$

Загалом, якщо існує $v_n \underset{n\to+\infty}{\longrightarrow} +\infty$

$$v_{n\left(\hat{\theta}_n - \theta\right)} \xrightarrow{\mathcal{L}} Y$$

Кажуть, що $\hat{\theta}_n$ збігається зі швидкістю $\frac{1}{v_n}$

Зауваження 7.1 – *Якщо $\hat{\theta}_n$ асимптотично нормальна $\Rightarrow \hat{\theta}_n$ збіжна*

$$\hat{\theta}_n - \theta = \underbrace{\frac{1}{\sqrt{n}}}_{\to 0} \underbrace{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0,\sigma^2)} \xrightarrow[\text{Slutsky}]{\mathcal{L}\ ou\ P} 0$$

$$U_n = \frac{1}{\sqrt{n}} \to 0$$

◇

δ-метод

$$\sqrt{n}(X_n - 1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

$$\sqrt{n}(X_n - 1) \overset{\mathcal{L}}{\approx} Z \sim \mathcal{N}(0, 1)$$

$$X_n \overset{\mathcal{L}}{\approx} 1 + \frac{1}{\sqrt{n}} Z$$

Якщо g диференційовна в 1,

$$g(1 + h) = g(1) + hg(1)$$

$$g(X_n) \approx g(1) + \frac{1}{\sqrt{n}} g'(1) Z$$

$$\sqrt{n}(g(X_n) - g(1)) \approx g'(1) Z$$

δ-метод

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0,1)$$

g диференційовна в θ

$$g(x) = g(\theta) + g'(\theta)[(x-\theta) + r(x)] \text{ де } r(x) \underset{x\to 0}{\to} 0$$

$\hat{\theta}_n \xrightarrow{P} \theta$ отже (LAC) $r\left(\hat{\theta}_n\right) \longrightarrow r(\theta) = 0$

$$g\left(\hat{\theta}_n\right) = g(\theta) + \left(\hat{\theta}_n - \theta\right)\left[g'(\theta) + r\left(\hat{\theta}_n\right)\right]$$

$$\sqrt{n}\left(g\left(\hat{\theta}_n\right) - g(\theta)\right) = \underbrace{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}_{\xrightarrow{\mathcal{L}} Z}\left[\underbrace{g'(\theta) + r\left(\hat{\theta}_n\right)}_{\xrightarrow{P} g'(\theta)}\right] \underset{\text{Slutsky}}{\Rightarrow} \sqrt{n}\left(g\left(\hat{\theta}_n\right) - g(\theta)\right) \xrightarrow{\mathcal{L}} g'(\theta)Z \sim \mathcal{N}\left(0, (g'(\theta))^2\right)$$

Приклад 7.2 – *$X_1, .., X_n$ з функцією щільності $f(x) = \frac{1}{\mu}e^{-\frac{x}{\mu}}, x \geq 0, \mu = E[X_i] > 0$ μ оцінений за допомогою $\hat{\mu} = \overline{X}$ ефективний? $\log L_n(\mu) = -n \log \mu - \frac{1}{\mu}\sum_{i=1}^n X_i$*

$$Var(\hat{\mu}) = \frac{1}{n^2}\, Var\left(\sum_i X_i\right) \underbrace{=}_{ind\acute{e}p} \frac{1}{n^2}\sum_i\, Var(X_i) \underbrace{=}_{i.i.d.} \frac{1}{n}\, Var(X_i) = \frac{\mu^2}{n}, E[\hat{\mu}] = \mu$$

$$\frac{\partial}{\partial \mu}(\log L_n)(\mu) = -\frac{n}{\mu} + \frac{1}{\mu^2}\sum (X_i)$$

$$\begin{aligned} I_{n(\mu)} &= Var\left(-\frac{n}{\mu} + \frac{1}{\mu^2}\sum X_i\right) \\ &= \frac{1}{\mu^4}\, Var\left(\sum X_i\right) \\ &= \frac{n}{\mu^4}\, Var(X_i) \end{aligned}$$

$$I_{n(\mu)} = \frac{n}{\mu^2}$$

$\hat{\mu}$ незміщений і $Var(\hat{\mu}) = \frac{1}{I_{n(\mu)}}$. Отже, $\hat{\mu}$ є ефективним.

ЦГТ: $\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu^2) \Rightarrow Var(\hat{\mu}_n) = \frac{\mu^2}{n}$ = дисперсія асимптотичного гауссового розподілу

$\frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\mu}$ має асимптотичний розподіл $\mathcal{N}(0,1)$

- *інша параметризація: $(X_1, ..., X_n)$ i.i.d. $f(x) = \theta e^{-\theta x}, x \geq 0$*

$$EX_i = \frac{1}{\theta},\, Var\, X_i = \frac{1}{\theta^2}$$

$$\log L_n(\theta) = n \log \theta - \theta \sum_{i=1}^n X_i$$

$$\frac{\partial}{\partial \theta}(\log L_n)(\theta) = \frac{n}{\theta} - \sum X_i \hookrightarrow \hat{\theta}^{MV} = \frac{1}{\overline{X}}$$

$$\hookrightarrow I_{n(\theta)} = Var\left(\frac{n}{\theta} - \sum X_i\right) = Var\left(\sum X_i\right) = \frac{n}{\theta^2}$$

Зауваження 7.3 – . TD1: $n\overline{X} \sim \Gamma(n, \theta)$

$$E\left[\frac{1}{n\overline{X}}\right] = \frac{\theta}{n-1} \text{ та Var}\left(\frac{1}{\left(n\overline{X}\right)^2}\right) = \frac{\theta^2}{(n-1)(n-2)}$$

◇

$$E\left[\frac{1}{\overline{X}}\right] = n\frac{\theta}{n-1}$$

$$\hookrightarrow \tilde{\theta} = \frac{n-1}{n}\hat{\theta} \textit{ незміщений}$$

$$\begin{aligned}
Var\left(\tilde{\theta}\right) &= \left(\frac{n-1}{n}\right)^2 Var\left(\frac{1}{\overline{X}}\right) = \frac{(n-1)^2}{n^2}\left[E\left[\frac{1}{\left(\overline{X}\right)^2}\right] - \left(E\left[\frac{1}{\overline{X}}\right]\right)^2\right] \\
&= \frac{\cancel{(n-1)^2}}{\cancel{n^2}} \times \frac{\cancel{n^2}\theta^2}{\cancel{(n-1)}(n-2)} - \frac{(n-1)^2}{n^2}\frac{n^2}{(n-1)^2}\theta^2 \\
&= \theta^2\frac{n-1}{n-2} - \theta^2 = \frac{\theta^2}{n-2} \underbrace{>}_{BCR} \frac{1}{I_{n(\theta)}} \textit{ неефективний}
\end{aligned}$$

$$\sqrt{n}\left(\overline{X} - \frac{1}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\theta^2}\right)$$

$\hat{\theta}$ є асимптотично ефективним $\overline{X}$ асимптотично нормальний (ЦГТ). $g(x) = \frac{1}{x}$ на $]0, +\infty[$, $g'(x) = -\frac{1}{x^2} \neq 0$, дельта-метод:

$$\sqrt{n}\left(\frac{1}{\overline{X}} - \theta\right) \xrightarrow{\mathcal{L}} \underbrace{g'\left(\frac{1}{\theta}\right)}_{=\theta^2}\mathcal{N}\left(0, \frac{1}{\theta^2}\right) = \mathcal{N}\left(0, \frac{\theta^4}{\theta^2} = \theta^2\right)$$

◇

7.2 Півот (асимптотичний) або півотна статистика

Означення 7.4 – *Статистика, розподіл якої не залежить від невідомих параметрів*

Приклад 7.5 – $X_1, ..., X_n$ *н.о.р. Бернуллі*(θ) *з* $\theta \in]0, 1[$*:*

$$\sqrt{n}\left(\overline{X} - \theta\right) \xrightarrow[TLC]{\mathcal{L}} \mathcal{N}(0, \theta(1-\theta))$$

$$\Leftrightarrow \underbrace{\sqrt{n}\frac{\overline{X} - \theta}{\sqrt{\theta(1-\theta)}}}_{\textit{pivot ou stat. pivotale}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

метод півота для ДІ: Ми оцінюємо $\sqrt{\theta(1-\theta)}$ за допомогою $\sqrt{\hat{\theta}\left(1-\hat{\theta}\right)}$ «plug-in» за допомогою функції $g(x) = \sqrt{x(1-x)}$ де $x \in]0, 1[$, $\sqrt{\hat{\theta}\left(1-\hat{\theta}\right)}$ є консистентною оцінкою для $\sqrt{\theta(1-\theta)}$

$$\sqrt{n}\frac{\hat{\theta}-\theta}{\sqrt{\hat{\theta}\left(1-\hat{\theta}\right)}} = \underbrace{\sqrt{n}\frac{\hat{\theta}-\theta}{\sqrt{\theta(1-\theta)}}}_{\underset{TLC}{\xrightarrow{\mathcal{L}}}\mathcal{N}(0,1)} \times \underbrace{\frac{\sqrt{\theta(1-\theta)}}{\sqrt{\hat{\theta}\left(1-\hat{\theta}\right)}}}_{\underset{consistant}{\xrightarrow{P}}1}$$

◇

Приклад 7.6 – $(X_1, ..., X_n)$ *з густиною* $\theta > 0$. $f_\theta(x) = \frac{3}{\theta}x^2\ exp\left(-\frac{x^3}{\theta}\right)\mathbb{1}_{x\geq 0}$

ОМВ?

$$log\,L_n(\theta) = n\left(log^3 - log\,\theta\right) + \sum_{i=1}^{n} log\left(X_i^2\right) - \frac{1}{\theta}\sum_{i=1}^{n} X_i^3$$

$$(log\,L_n)'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum X_i^3 \Rightarrow \hat{\theta} = \frac{\sum X_i^3}{n}$$

$$(log\,L_n)''(\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum X_i^3; (log\,L_n)''\left(\hat{\theta}\right) = \frac{n}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3}n\hat{\theta} = -\frac{n}{\hat{\theta}^2} < 0 + \text{унікальність}$$

$\Rightarrow$ *глобальний максимум*

ЦГТ:

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \theta^2)$$

$$\underbrace{\Leftrightarrow}_{pivot\ asymptotique} \quad \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\theta} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

$$\underbrace{\Rightarrow}_{Slutsky} \quad \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\hat{\theta}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

$q_{\frac{\alpha}{2}}$ *та* $q_{1-\frac{\alpha}{2}}$ *квантилі для* $\mathcal{N}(0,1)$

$$P\left(q_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\hat{\theta}-\theta\right)}{\hat{\theta}} \leq 1_{1-\frac{\alpha}{2}}\right) \underset{n\to+\infty}{\longrightarrow} 1-\alpha$$

$$P\left(q_{\frac{\alpha}{2}}\frac{\hat{\theta}}{\sqrt{n}} \leq \hat{\theta}-\theta \leq q_{1-\frac{\alpha}{2}}\frac{\hat{\theta}}{\sqrt{n}}\right) \longrightarrow 1-\alpha$$

$$P\left(\underbrace{\hat{\theta} - q_{1-\frac{\alpha}{2}}\frac{\hat{\theta}}{\sqrt{n}} \leq \theta \leq \hat{\theta} - q_{\frac{\alpha}{2}}\frac{\hat{\theta}}{\sqrt{n}}}_{\Rightarrow IC(\theta)\ de\ niveau\ asymptotique\ (1-\alpha)}\right) \longrightarrow 1-\alpha$$

◇

Оцінювання в гаусових вибірках

§8

1. Нормальний розподіл та похідні розподіли
2. Розподіл емпіричних оцінок
3. Довірчі інтервали параметрів
4. Вправа

8.1 Нормальний закон та похідні закони

Означення 8.1 – *Z називається гаусовим (нормальним) центрованим та зведеним, якщо його закон має щільність*

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

Позначається $Z \sim \mathcal{N}(0, 1)$.

X називається нормально розподіленою з параметрами $\mu \in \mathbb{R}$ та $\sigma^2 > 0$ т.і.т.т.

$$X = \mu + \sigma Z$$

, що позначається

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Інші характеристики нормального закону:

- за допомогою щільності

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

- за допомогою твірної функції моментів

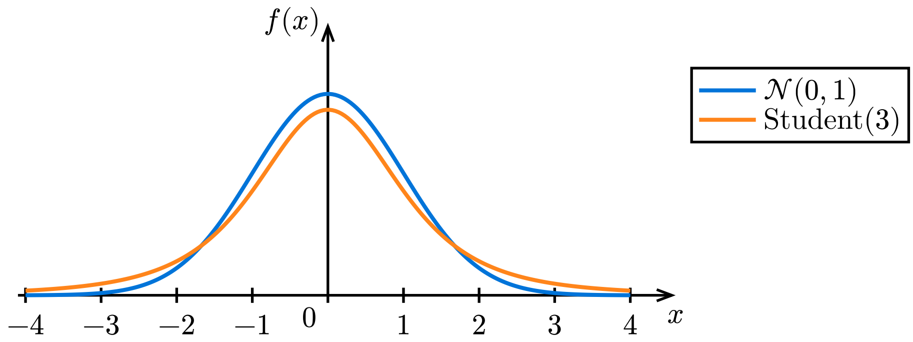
$$M(t) = E[e^{tX}] = e^{t\mu + \frac{1}{2}\sigma^2 t^2}, \forall t \in \mathbb{R}$$

Зауваження 8.2 –

- $\sigma^2 = 0 \to X = \mu$ *майже напевно*
- *якщо* $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ *і* $\lambda \in \mathbb{R}$, *тоді* $\lambda X_1 + X_2 \sim \mathcal{N}(\lambda\mu_1 + \mu_2, \lambda^2\sigma_1^2 + \sigma_2^2)$

◇

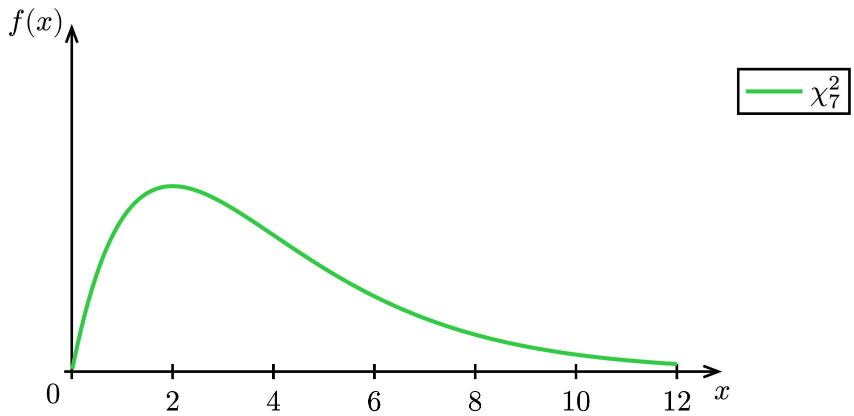
Центральні моменти: щільність симетрична відносно μ



центральні моменти: $E\big[(X-\mu)^k\big]$

- всі центральні моменти непарного порядку дорівнюють нулю
- $\mu_{2k} = \frac{(2k)!}{2^k k!}\sigma^{2k}$
  - $E\big[(X-\mu)^4\big] = 3\sigma^4$
  - $\mathrm{Var}(X) = E\big[(X-\mu)^2\big] = \sigma^2$

Означення 8.3 – $(X_1, \ldots, X_d)$ — *незалежна однаково розподілена вибірка* $\mathcal{N}(0,1)$. *Закон розподілу* $X_1^2 + X_2^2 + \ldots + X_d^2$ *називається законом* χ^2 *(хі-квадрат) з* d *ступенями свободи (сс).*



Наслідок 8.4 –

- *якщо* Y *має розподіл* $\chi^2(d)$, *то* $E[Y] = d$, $Var(Y) = 2d$

$$Var(X_1^2 + \ldots + X_d^2) \underbrace{=}_{indep} d \underbrace{Var(X_i^2)}_{EX_i^h = E[X_i^2]^2 = 3-1=2}$$

- *носій* $\mathbb{R}_+$
- $M(t) = (1-2t)^{-\frac{d}{2}}, \ \left(t < \frac{1}{2}\right)$

◇

Означення 8.5 – *якщо* $X \sim \mathcal{N}(0,1)$ *та* $Y \sim \chi^2(d)$ *незалежні, то розподіл* $Z = \frac{X}{\sqrt{\frac{Y}{d}}}$ *називається розподілом Стьюдента з* d *ступенями свободи.*

Зауваження 8.6 – *якщо* $d \to +\infty$, *розподіл Стьюдента збігається до розподілу* $\mathcal{N}(0,1)$

$$\frac{Y}{d} \underset{def}{\overset{\mathcal{L}}{=}} \frac{1}{d} \sum_{i=1}^{d} U_i^2 \text{ де } U_i \sim \mathcal{N}(0,1) \text{ незалежні між собою від } X$$

$$\xrightarrow[LGN]{P} E(U_i^2) = 1$$

отже (ЦГТ)

$$g(x) = \sqrt{x} \frac{1}{\sqrt{\frac{Y}{d}}} \xrightarrow{P} 1$$

за лемою Слуцького $Z \xrightarrow{\mathcal{L}} 1 \cdot X \sim \mathcal{N}(0,1)$ ◇

Введемо $(X_1, ..., X_n)$ н.о.р. $\mathcal{N}(\mu, \sigma^2)$ де μ та σ^2 невідомі параметри.

- $\hookrightarrow \mu = E[X_i] \rightsquigarrow \hat{\mu} = \overline{X}$
- $\hookrightarrow \sigma^2 = \mathrm{Var}(X_i) \rightsquigarrow \widehat{\sigma^2} = \frac{1}{n} \sum \left(X_i - \overline{X}\right)^2$

Нехай $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n \left(X - i - \overline{X}\right)^2$ <u>незміщений</u>

8.2 Закон емпіричних оцінок

Теорема 8.7 (закон розподілу $\hat{\mu}$ та $\hat{\sigma}^2$) –

- $\overline{X}$ *та* $\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$ *є випадковими змінними <u>незалежними</u>*
- $\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$
- $\frac{1}{\sigma^2} \sum_{i=1}^n \left(X_i - \overline{X}\right)^2 \sim \chi^2(n-1) \Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$ *та* $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$
- $\frac{\overline{X}-\mu}{\frac{S_n}{\sqrt{n}}} \sim Student(n-1)$
- $\overline{X}$ *та* $\overbrace{\left(\overline{X}, X_1 - \overline{X}, ..., X_n - \overline{X}\right)}^{T}$ *є незалежними*

Доведення.

$$
\begin{aligned}
M(u, t_1, ..., t_n) &= E\left[e^{u\overline{X}+t_1\left(X_1-\overline{X}\right)+...+t_n\left(X_n-\overline{X}\right)}\right] \\
&= E\left[e^{\left(\frac{u}{n}+\left(\underbrace{\frac{t_1+t_2+...+t_n}{n}}_{\bar{t}}\right)\right)X_1} ... e^{\left(\frac{u}{n}+\bar{t}\right)X_n}\right] \\
&= E\left[\prod_{i=1}^{n} e^{\left(\frac{u}{n}+t_i-\bar{t}\right)X_i}\right] \\
X_i \text{ незалежні} &= \prod_{i=1}^{n} \underbrace{E\left[e^{\left(u_n+t_i-\bar{t}\right)X_i}\right]}_{M\left(u_n+t_i-\bar{t}\right)} \\
&= \prod_{i=1}^{n} e^{\mu\left(\frac{u}{n}+t_i-\bar{t}\right)+\frac{\sigma^2}{2}\left(u_n+t_i-\bar{t}\right)^2} \\
&= e^{\sum_{i=1}^{n} \mu\left(\frac{u}{n}+t_i-\bar{t}\right)+\frac{\sigma^2}{2}\left(u_n+t_i-\bar{t}\right)^2} \\
&= e^{\mu u+\mu\overbrace{\sum_i\left(t_i-\bar{t}\right)}^{0}+\frac{\sigma^2}{2}\sum_i\left(\frac{u^2}{n^2}+\left(t_i-\bar{t}\right)^2+2\frac{u}{n}\left(t_i-\bar{t}\right)\right)} \\
&= e^{\mu u+\frac{\sigma^2}{2}\left(\frac{u^2}{n}+\sum_i\left(t_i-\bar{t}\right)^2\right)} \\
&= \underbrace{e^{\mu u+\frac{\sigma^2 u^2}{2n}}}_{M_{\overline{X}}(u)}\underbrace{e^{\frac{\sigma^2}{2}\sum_i\left(t_i-\bar{t}\right)^2}}_{M_T(t_1,...,t_n)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\mu\right)^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2+\frac{n}{\sigma^2}\left(\overline{X}-\mu\right)^2+\frac{2}{\sigma^2}\sum_i\left(X_i-\overline{X}\right)\left(\overline{X}-\mu\right) \\
&= \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2+\frac{n}{\sigma^2}\left(\overline{X}-\mu\right)^2+\frac{2}{\sigma^2}\left(\overline{X}-\mu\right)\underbrace{\sum_i\left(X_i-\overline{X}\right)}_{=0} \\
&= \underbrace{\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}_{=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)} + \underbrace{\frac{n}{\sigma^2}\left(\overline{X}-\mu\right)^2}_{=\left(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2\sim\chi^2(1)}
\end{aligned}
$$

$$
\text{par indép} \Rightarrow M_{\chi^2(n)}(t) = M_?(t)M_{\chi^2(1)}(t) \Rightarrow M_?(t) = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}} = (1-2t)^{-\frac{(n-1)}{2}}
$$

що характеризує розподіл $\chi^2(n-1)$

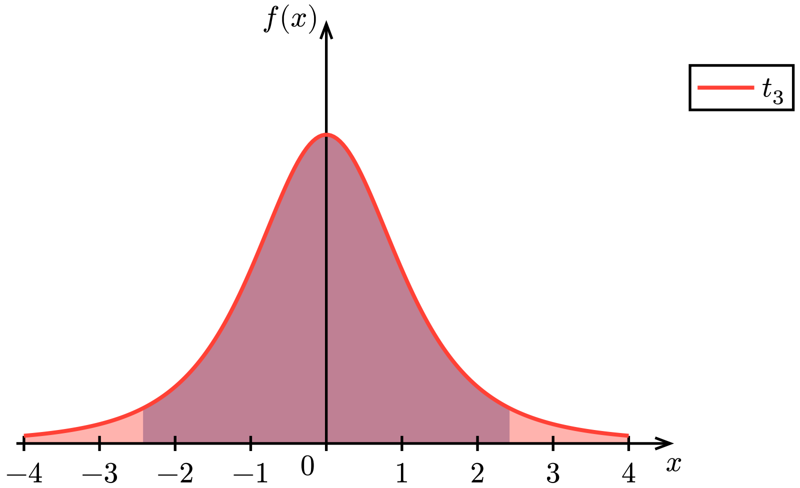
$$
\frac{\overline{X}-\mu}{\frac{S_n}{\sqrt{n}}} = \frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{S_n}{\sqrt{n}}\times\frac{\sqrt{n}}{\sigma}} = \frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{S_n^2}{\sigma^2}}}
$$

$$
\frac{S_n^2}{\sigma^2} = \left(\frac{\sum\left(X_i-\overline{X}\right)^2}{\sigma^2}\right) \sim \chi^2(n-1)
$$

отже $\overline{X}+S_n^2$ незалежні $\underbrace{\Rightarrow}_{\text{def Student}}$ Student$(n-1)$ □

8.3 Довірчий інтервал для параметрів

Опорна величина. $\frac{\overline{X}-\mu}{\frac{S_n}{\sqrt{n}}} \underset{\text{loi exacte}}{\sim} \text{Student}(n-1)$



$$P\Big(q_{\frac{\alpha}{2}}t(n-1)\Big) \leq \frac{\overline{X}-\mu}{\frac{S_n}{\sqrt{n}}} \leq q_{1-\frac{\alpha}{2}}t(n-1))$$

$$\Leftrightarrow P\left(\overline{X} - \frac{S_n}{\sqrt{n}}q_{1-\frac{\alpha}{2}}t(n-1) \leq \mu \leq \overline{X} + \frac{S_n}{\sqrt{n}}q_{1-\frac{\alpha}{2}}t(n-1)\right) = 1-\alpha$$

ДІ (σ^2), $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$

$$P\left(q_{\frac{\alpha}{2}}\chi^2(n-1) \leq \frac{n\hat{\sigma}^2}{\sigma^2} \leq q_{1-\frac{\alpha}{2}}\chi^2(n-1)\right) = 1-\alpha$$

$$= P\left(\frac{n\hat{\sigma}^2}{q_{1-\frac{\alpha}{2}}\chi^2(n-1)} \leq \sigma^2 \leq \frac{n\hat{\sigma}^2}{q_{\frac{\alpha}{2}}\chi^2(n-1)}\right) = \alpha - 1$$

$\rightsquigarrow$ ДІ $= \left[\frac{n\hat{\sigma}^2}{q_{1-\frac{\alpha}{2}}\chi^2(n-1)}, \frac{n\hat{\sigma}^2}{q_{\frac{\alpha}{2}}\chi^2(n-1)}\right]$

Зауваження 8.8 – $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1)$ *та* $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$

$$\frac{\overline{X}-\mu}{\frac{S_n}{\sqrt{n}}} \sim Student(n-1)$$

◇

8.4 Вправа

- Покажіть, що $\left(\hat{\mu}, \widehat{\sigma^2}\right)$ є ОМВ для μ та σ^2
- $R(S_n^2, \sigma^2) > R(\widehat{\sigma_n}^2, \sigma^2)$ де R представляє ризик