Метод Максимального Правдоподобия

Пусть $X_1, \ldots, X_n$ — выборка из распределения F_θ, $\theta \in \Theta$ (Θ — множество возможных значений параметра)

Будем называть обобщённой плотностью $f_\theta(x)$ функцию

$$f_\theta(x) = \begin{cases} P_\theta(X_1 = x), & \textit{если распределение } F_\theta \textit{ дискретно} \\ f_{X_1}(x;\theta), & \textit{если распределение } F_\theta \textit{ абсолютно непрерывно} \end{cases}$$

Функция правдоподобия $L(\vec{X};\theta) = f_\theta(X_1)\cdot\ldots f_\theta(X_n) = \prod_{i=1}^{n} f_\theta(X_i)$

(функция правдоподобия является случайной величиной при фиксированном ϑ)

$\ln L(\vec{X};\theta) = \sum_{i=1}^{n} \ln f_\theta(X_i)$ — логарифмическая функция правдоподобия

Оценка максимального правдоподобия (ОМП)

Значение θ, при котором функция правдоподобия достигает своего максимума

$$\hat{\theta} = \arg(\max_{\theta} L(\vec{X};\theta)) = \arg(\max_{\theta} \ln L(\vec{X};\theta))$$

Всюду далее предполагается, что *модель регулярна*, то есть допустимо дифференцировать по параметру θ интегралы от функций на выборочном пространстве и менять порядок дифференцирования и интегрирования.

Свойства ОМП

1. Если эффективная оценка параметра θ существует, то она совпадает с ОМП

2. Если $g(\theta)$ – взаимно-однозначная функция, то ОМП $g(\theta)$ совпадает с $g(\hat{\theta})$: $\hat{g}(\theta) = g(\hat{\theta})$ (свойство инвариантности)

Асимптотические свойства ОМП

$$\hat{\theta} = T(X_1, \ldots, X_n)$$

1. Состоятельность $\hat{\theta} \xrightarrow{P} \theta$ *при* $n \to \infty \; \forall \theta \in \Theta$
2. Асимптотическая несмещённость: $E(\hat{\theta}) \to \theta$ *при* $n \to \infty \; \forall \theta \in \Theta$
3. Асимптотическая нормальность (если логарифмическая функция правдоподобия трижды дифференцируема и математическое ожидание её третьих производных ограничено для всех значений θ):

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{F} N(0, i^{-1}(\theta)), \quad \text{при} \quad n \to \infty$$

где

$$i(\theta) = E\left(\left(\frac{\partial \ln f_\theta(x)}{\partial \theta}\right)^2\right)$$

Информация Фишера

$$I(\theta) = E\left(\left(\frac{\partial \ln L(\vec{X};\theta)}{\partial \theta}\right)^2\right)$$

информация о параметре,
содержащаяся во всей выборке

$$i(\theta) = E\left(\left(\frac{\partial \ln f_\theta(x)}{\partial \theta}\right)^2\right)$$

информация о параметре,
содержащаяся в одном наблюдении

Свойства информации Фишера

$$I(\theta) = -E\left(\frac{\partial^2 \ln L(\vec{X};\theta)}{\partial \theta^2}\right), \quad I(\theta) = D\left(\frac{\partial \ln L(\vec{X};\theta)}{\partial \theta}\right), \quad I(\theta) = ni(\theta)$$

Неравенство Рао-Крамера

Если $\hat{\theta} = T(X_1, \ldots, X_n)$ оценка со смещением $b(\theta) = E(\hat{\theta}) - \theta$ и функция $b(\theta)$ дифференцируема, то

$$E(\hat{\theta} - \theta)^2 \geq \frac{\left(1 + b'(\theta)\right)^2}{I(\theta)}$$

<u>Замечание</u> Для несмещённых оценок

$$D(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

Критерий эффективности (критерий Рао-Крамера)

Если в классе K_b оценок со смещением $b(\theta)$ существует оценка, для которой нижняя граница неравенства Рао-Крамера достигается, то она является эффективной в данном классе.

$$E(\hat{\theta}-\theta)^2 = \frac{\left(1+b'(\theta)\right)^2}{I(\theta)} \quad \Rightarrow \quad E(\hat{\theta}-\theta)^2 \le E(\tilde{\theta}-\theta)^2,$$

$$\forall \tilde{\theta} : E(\tilde{\theta}) - \theta = b(\theta), \quad \forall \theta$$

Класс экспоненциальных моделей

$$f_\theta(x) = e^{\{A(\theta)B(x)+C(\theta)+D(x)\}}$$

Примеры экспоненциальных моделей:

нерегулярная!

Непрерывные $N(\theta,\sigma^2)$, $N(\mu,\theta)$, $E(\theta)$, $U(0,\theta)$

нормальное показательное равномерное

Дискретные $\Pi(\theta)$, $Bi(m,\theta)$, $\overline{B}i(r,\theta)$,

Пуассона биномиальное отрицательное биномиальное

Примеры НЕ экспоненциальных моделей: распределение Коши

Свойства экспоненциальных моделей

$$f_\theta(x) = e^{\{A(\theta)B(x)+C(\theta)+D(x)\}}$$

Для регулярной экспоненциальной модели **существует эффективная оценка** T* для параметрической функции g(θ), имеющий вид:

$$g(\theta) = -\frac{C'(\theta)}{A'(\theta)} \quad \text{при этом} \quad T^* = \frac{1}{n}\sum_{i=1}^{n} B(X_i) \quad \text{и} \quad D(T^*) = \frac{g'(\theta)}{n \cdot A'(\theta)}$$

Верно и обратное: если для параметрической функции g(θ) существует эффективная оценка, то модель экспоненциальна.

Примеры эффективных оценок

Модель	Оценка θ	Дисперсия оценки
$N(\theta,\sigma^2)$	$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	σ^2/n
$N(\mu,\theta)$	$\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$	$2\theta^4/n$
$Bi(m,\theta)$	$\frac{\bar{X}}{m}=\frac{1}{m\cdot n}\sum_{i=1}^{n}X_i$	$\theta(1-\theta)/(m\cdot n)$
$\Pi(\theta)$	$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	θ/n

?

Что не так с равномерной моделью U(0,θ)?

Асимптотические свойства оценок

Некоторые полезные факты:

Теорема Слуцкого

Пусть последовательности случайных величин ξ_n, η_n, определённых на одном и том же вероятностном пространстве, таковы, что

$$\xi_n \xrightarrow{P} c, \quad \eta_n \xrightarrow{F} \eta \quad (n \to \infty),$$

Где $c \in R$, $\eta -$ случайная величина. Тогда

$$\xi_n + \eta_n \xrightarrow{F} c + \eta, \quad \xi_n \cdot \eta_n \xrightarrow{F} c \cdot \eta, \quad g(\xi_n, \eta_n) \xrightarrow{F} g(c, \eta),$$

где $g(x, y) -$ непрерывная функция.

Асимптотическая нормальность

Оценка $\hat{\theta} = T(X_1, \ldots, X_n)$, асимптотически нормальна, если существует

такое $\sigma^2(\theta)$, что

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \to \infty]{F} N(0, \sigma^2(\theta))$$

<u>Замечание</u>. Если оценка асимптотически нормальна, то она

состоятельна: $$(\hat{\theta} - \theta) = \frac{1}{\sqrt{n}} \cdot \sqrt{n}(\hat{\theta} - \theta) \xrightarrow{F} 0 \cdot z = 0,$$

где $z \sim N(0, \sigma^2(\theta))$.

В таком случае говорят о $\sqrt{n}$ – состоятельности.

Асимптотическая дисперсия

Величина $\dfrac{\sigma^2(\theta)}{n}$ — асимптотическая дисперсия оценки $\hat{\theta} = T(X_1, \dots, X_n)$,

если $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \to \infty]{F} N(0, \sigma^2(\theta))$ (оценка асимптотически нормальна)

<u>Замечание</u>. Асимптотическая дисперсия ОМП совпадает с нижней границей в неравенстве Рао-Крамера

Асимптотическая эффективность

Оценка асимптотически эффективна, если её асимптотическая дисперсия совпадает с нижней границей в неравенстве Рао-Крамера

Асимптотические свойства ОМП (продолжение)

4. Асимптотическая эффективность

Так как ОМП асимптотически нормальны

$$\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{F} N(0, i^{-1}(\theta)), \quad при \quad n \to \infty,$$

они являются асимптотически эффективными, причём их асимптотическая дисперсия обратно пропорциональна информации Фишера:

$$As.Var(\hat{\theta}) = \frac{\sigma^2(\theta)}{n} = I^{-1}(\theta) = \frac{1}{ni(\theta)}$$

Дельта-метод

Пусть оценка $\hat{\theta} = T(X_1, \ldots, X_n)$ параметра θ состоятельна и асимптотически нормальна, то есть

$$\hat{\theta} \xrightarrow{P} \theta, \quad \sqrt{n}(\hat{\theta} - \theta) \xrightarrow{F} N(0, \sigma^2(\theta)) \quad \forall \theta \quad \textit{при} \quad n \to \infty,$$

Пусть, далее, функция g дифференцируема и $g' \neq 0$.

Тогда $g(\hat{\theta})$ –состоятельная и асимптотически нормальная оценка $g(\theta)$:

$$g(\hat{\theta}) \xrightarrow{P} g(\theta), \quad \sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{F} N(0, (g'(\theta))^2 \sigma^2(\theta))$$

<u>Замечание</u>. Если $\sigma^2(\theta), g'(\theta) -$ непрерывны, то

$$\frac{g(\hat{\theta}) - g(\theta)}{g'(\theta)\sigma(\theta) / \sqrt{n}} \xrightarrow[n \to \infty]{F} N(0,1)$$

Пример. По выборке $X_1, \ldots, X_n$ из показательного распределения с функцией плотности

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

ММП найти оценку параметра λ, математического ожидания, вероятности $P(X_1 > 1)$ и вычислить их асимптотические дисперсии.