Конспектировал: Антон Захаренков.

Вспомним обозначения

- θ_0 — истинное значение параметра
- θ — произвольное значение параметра
- $\hat{\theta}$ — оценка ML

Производная логарифмической функции потерь: $s(\theta) = \operatorname{grad} l(\theta) = \operatorname{grad} \ln L(\theta)$

$$\operatorname{Var} s(\theta_0) = \begin{cases} \mathrm{E}(s(\theta_0) - s(\theta_0)^T) \\ \mathrm{E}\left(-\frac{\partial^2 l}{\partial\theta\,\partial\theta^T}\Big|_{\theta_0}\right) \end{cases} \tag{1}$$

Утверждение. $\operatorname{Var} s(\theta_0)$ и $\operatorname{Var}(\hat{\theta})$ — взаимообратны.

Пример
$y_1, y_2, \ldots, y_n$ — выборка независимых случайных величин из распределения $\operatorname{Exp}(\lambda)$

1. Найдите $\hat{\lambda}$
2. Найдите $\operatorname{Var}(\hat{\lambda})$
3. Найдите $\operatorname{Var}(s(\lambda_0))$
4. Найдите $\operatorname{Var}(\hat{\lambda}) \cdot \operatorname{Var}(s(\lambda_0))$

Решение
Экспоненциальное распределение:

$$f(y_0) = \begin{cases} \lambda e^{-\lambda y_i}, & y_i \geq 0 \\ 0, & y_i < 0 \end{cases} \tag{2}$$

Функция правдоподобия:

$$L(y_1, \ldots, y_n, \theta) = \lambda^n e^{-\lambda(y_1 + \ldots + y_n)} \tag{3}$$

Логарифмируем:

$$l(y_1, \ldots, y_n, \theta) = n \ln \lambda - \lambda(y_1 + \ldots + y_n) \tag{4}$$

Дифференцируем:

$$s(\lambda) = \frac{n}{\lambda} - (y_1 + \ldots + y_n) \tag{5}$$

$$\hat{\lambda} = \frac{n}{\sum y_i} \tag{6}$$

$$\operatorname{Var} s(\lambda_0) = \mathrm{E}\left(-\frac{\partial^2 l}{\partial\lambda\,\partial\lambda^T}\Big|_{\lambda_0}\right) = \mathrm{E}\left(\frac{n}{\lambda_0^2}\right) = \frac{n}{\lambda_0^2} \tag{7}$$

Примерно найдем $\text{Var}(\hat{\lambda})$ с помощью дельта-метода.

По ЦПТ:

$$\frac{\sum y_i}{n} \overset{acc}{\sim} \text{N}\left(\frac{1}{\lambda_0}; \frac{1}{n\lambda_0^2}\right) \tag{8}$$

Но:

$$\hat{\lambda} = \frac{1}{\left(\frac{\sum y_i}{n}\right)} = h\left(\frac{\sum y_i}{n}\right) \tag{9}$$

Разложим h(t) в ряд Тейлора в окрестности $\frac{1}{\lambda}$:

$$h(t) = \frac{1}{t} = \frac{1}{t_0} - \frac{1}{t_0^2}(t - t_0) \tag{10}$$

Тогда:

$$\hat{\lambda} \approx \lambda_0 - \lambda_0^2\left(\frac{\sum y_i}{n} - \frac{1}{\lambda_0}\right) \tag{11}$$

Дисперсия:

$$\text{Var}(\hat{\lambda}) = \lambda_0^4 \cdot \frac{1}{n\lambda_0^2} = \frac{\lambda_0^2}{n} \tag{12}$$

Видим, что:

$$\text{Var}(\hat{\lambda}) \cdot \text{Var}(s(\lambda_0)) \approx \frac{\lambda_0^2}{n} \cdot \frac{n}{\lambda_0^2} \to 1 \tag{13}$$

Упражнение

Рассмотрим модель множественной регрессии, $y = X\beta + u$, где регрессоры детерминистические, а $u \sim \mathcal{N}(0; \sigma^2 \cdot I)$. Величина σ^2 известна.

1. Найдите $\hat{\beta}_{ML}$.
2. Найдите $\text{Var}(\hat{\beta}_{ML})$.
3. Найдите $s(\beta_0)$.
4. Найдите $\text{Var}(s(\beta_0))$.
5. Найдите $\text{Var}(\hat{\beta}) \cdot \text{Var}(s(\beta_0))$.

Решение

По условию:

$$y \sim \mathcal{N}(X\beta; \sigma^2 \cdot I) \tag{14}$$

Функция правдоподобия:

$$L(y, \beta) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \cdot exp(-\frac{1}{2}(y - X\beta)^T(\sigma^2 \cdot I)^{-1}(y - X\beta)) \tag{15}$$

Логарифмируем:

$$l(y, \beta) = -\frac{n}{2}\ \ln(2\pi\sigma^2) - \frac{1}{2}(y - X\beta)^T\frac{1}{\sigma^2}(y - X\beta) \to \underset{\beta}{max} \tag{16}$$

Оценка ML совпадает с оценкой МНК:

$$\hat{\beta}_{ML} = \hat{\beta}_{MHK} = (X^T X)^{-1} X^T y \tag{17}$$

$$\operatorname{Var}(\hat{\beta}_{ML}) = \sigma^2 (X^T X)^{-1} \tag{18}$$

Посчитаем дифференциал dl:

$$dl = -\frac{1}{2\sigma^2}[-(Xd\beta)^T(y - X\beta) + (y - X\beta)^T(-Xd\beta)] = \frac{(y - X\beta)^T X d\beta}{\sigma^2} = s(\beta)^T d\beta \tag{19}$$

Отсюда:

$$s(\beta_0) = \frac{1}{\sigma^2} X^T (y - X\beta_0) \tag{20}$$

$$\operatorname{Var}(s(\beta_0)) = \frac{1}{\sigma^4} X^T \operatorname{Var}(y - X\beta_0) X = \frac{X^T X}{\sigma^2} \tag{21}$$

Видим, что:

$$\operatorname{Var}(\hat{\beta}) \cdot \operatorname{Var}(s(\beta_0) = 1 \tag{22}$$

Замечание
При [некоторые условия регулярности]:

1. $\operatorname{plim} \hat{\theta}_{ML} = \theta_0$
2. $\hat{\theta}_{ML} \sim \mathcal{N}(\theta; \operatorname{Var}(s(\theta_0))^{-1})$
3. $\sqrt{n}\,(\hat{\theta}_{ML} - \theta_0) \xrightarrow{d} \mathcal{N}(0; n \cdot \operatorname{Var}(s(\theta_0))^{-1})$

Откуда:

Если $\hat{a} \sim \mathcal{N}(a; \operatorname{Var}(\hat{a})$, то
$(\hat{a} - a)^T \operatorname{Var}(\hat{a})^{-1} (\hat{a} - a) \sim \chi^2_k$

Проверка гипотезы о равенстве параметров заданным значениям

- $H_0 : \theta = \theta_0$
- H_A : хотя бы одно из равенств нарушено

Если H_0:

$$W = (\hat{\theta}_{UR} - \theta_0)^T \operatorname{Var}(\hat{\theta}_{UR})^{-1} (\hat{\theta}_{UR} - \theta_0) \sim \chi^2_k \tag{23}$$

$$LM = (\hat{\theta}_R)^T \operatorname{Var}(s(\hat{\theta}_R))^{-1} (\hat{\theta}_R) \sim \chi^2_k \tag{24}$$

$$LR = 2 \cdot (l(\hat{\theta}_{UR}) - l(\hat{\theta}_R)) \sim \chi^2_k \tag{25}$$

Как состоятельно оценить $\mathrm{Var}(s(\theta_0)? = \mathrm{E}\left(-\frac{\partial^2 l}{\partial\theta\,\partial\theta^T}\Big|_{\theta_0}\right)$

1. Подставить $\hat{\theta}$ вместо θ_0
2. Можно даже не брать E

Пример
$y_1, ..., y_n$ — выборка независимых случайных величин из распределения Exp(λ), n=100, $\sum y_i = 400$

- $H_0 : \lambda = 0.3$
- $H_A : \lambda \neq 0.3$

Проверить с помощью:

1. Wald
2. LM
3. LR

Решение

$$\hat{\lambda}_{ML} = \frac{n}{\sum y_i} = \frac{100}{400} = 0.25 \tag{26}$$

$$\mathrm{Var}(\hat{\lambda}_{ML})^{-1} \approx \mathrm{Var}(s(\lambda_0)) = \frac{4}{\lambda_0^2} = \frac{100}{0.25^2} = 1600 \tag{27}$$

$$s(\lambda_0) = \frac{n}{\lambda_0} - \sum y_i = 1000.3 - 400 = -\frac{200}{3} \tag{28}$$

$$\mathrm{Var}\, s(\lambda_0) = \frac{n}{\lambda_0^2} = \frac{100}{0.3^2} \tag{29}$$

$$W = (0.25 - 0.3) \cdot 1600 \cdot (0.25 - 03) = 4 < \chi_1^2 = 6.63 \Rightarrow H_0 - notrejected \tag{30}$$

$$LM = (-\frac{200}{3}) \cdot \frac{1}{\frac{100}{0.3^2}} \cdot (-\frac{200}{3}) = 4 < \chi_1^2 = 6.63 \Rightarrow H_0 - notrejected \tag{31}$$

Упражнение
В пруду водятся

	Вероятность обнаружения	Было выловлено
щуки	p	10
русалки	r	30
водяные	v	20
караси	1-p-r-v	140

причем каждое выловленное животное не влияет на распределение последующих.
Проверяется гипотеза:

- $H_0 : \begin{cases} r = v \\ p = 0.1 \end{cases}$
- H_A : хотя бы что-то невыполнено

Решение
Пусть

$$H_0 = \begin{cases} r - v = \alpha = 0 \\ p - 0.1 = \beta = 0 \end{cases} \tag{32}$$

Правдоподобие:

$$L = const \cdot (0.1 + \beta)^{10} \cdot (v + \alpha)^{30} \cdot v^{20} \cdot (0.9 - \beta - 2v - \alpha)^{140} \tag{33}$$

$$l = const + 10 \ln(0.1 + \beta) + 30 \ln(v + \alpha) + 20 \ln v + 140 \ln(0.9 - \beta - 2v - \alpha) \tag{34}$$

Пусть $\theta = \left(\begin{smallmatrix} \alpha \\ \beta \\ v \end{smallmatrix}\right)$:

$$s(\theta) = \begin{pmatrix} 30 \dfrac{1}{v + \alpha} - 140 \dfrac{1}{0.9 - \beta - 2v - \alpha} \\ \ldots \\ \ldots \end{pmatrix} \tag{35}$$

$$\widehat{\begin{pmatrix} p \\ r \\ v \end{pmatrix}}_{UR} = \begin{pmatrix} 0.1 + \hat{\beta} \\ \hat{v} + \hat{\lambda} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} \frac{10}{200} \\ \frac{30}{100} \\ \frac{20}{200} \end{pmatrix} = \begin{pmatrix} 0.05 \\ -0.05 \\ 0.1 \end{pmatrix} \tag{36}$$

$$\widehat{\begin{pmatrix} p \\ r \\ v \end{pmatrix}}_{R} = \begin{pmatrix} 0.1 \\ \hat{v} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} 0.1 \\ \frac{45}{380} \\ \frac{45}{380} \end{pmatrix} \tag{37}$$