Домашнее задание по прикладной статистике №1

Агаев Фархат

19 декабря 2020 г.

Задача №1

(a)

$$D_{KL}(f(y|\theta), f(y|\phi)) = 1/2I(\theta)(\phi - theta)^2 + O((\phi - \theta)^3)$$

$$H(f) = \mathbb{E}(\ln f)$$

$$D_{KL} = \int f_\theta(\ln f_\phi - \ln f_\theta) = \mathbb{E} \ln f_\theta$$

Сила Тейлора

$$\ln f_\phi =_{\phi \to \theta} \ln f_\theta + (\ln f_\theta)'(\phi - \theta) + 1/2(\ln f_\theta)''(\phi - \theta)^2 + O((\phi - \theta)^3) \Rightarrow$$

$$\mathbb{E} \ln f_\phi - \mathbb{E} \ln f_\theta = \mathbb{E}[(\ln f_\theta)'(\phi - \theta)] + 1/2\mathbb{E}[(\ln f_\theta)''(\phi - \theta)^2] + O((\phi - \theta)^3)$$

мы знаем, что $\mathbb{E}[(\ln f_\theta)''] = I(\theta)$ и
$\mathbb{E}[(\ln f_\theta)'(\phi - \theta)] = 0 \Rightarrow$

$$D_{KL}(f(y|\theta), f(y|\phi)) = 1/2I(\theta)(\phi - theta)^2 + O((\phi - \theta)^3)$$

(b) Из семинара нам известно что есть обратная зависимость у дисперсии МЛ оценок. Также из сема мы знаем, что что оси приблежаются к истинному параметру когда $I(\hat{\theta})$ растет, тогда $VAR(\hat{\theta})$ Падает.

Задача №2

(а) Найдите $\hat{q}_{ML}$.

Для начала найдем функцию правдоподобия

$$L(q|x) = \prod_{i=1}^{n} \frac{2x_i}{q} e^{\frac{-x_i^2}{q}} = \frac{2^n \sum x_i}{q^n} * e^{\frac{\sum -x_i^2}{q}}$$

Теперь найдем логарифмическую ф-ию правдоподобия

$$l = n\ln(2) - n\ln(q) + \ln(\sum x_i) + \frac{\sum -x_i^2}{q}$$

Далее найдем производную от l, то есть score ф-ию

$$s = -\frac{n}{q} + \frac{\sum x_i^2}{q^2}$$

Теперь приравняем к нули и найдем оценку

$$-\frac{n}{\hat{q}} + \frac{\sum x_i^2}{\hat{q}^2} = 0$$

$$\hat{q} = \frac{\sum x_i^2}{n}$$

(б) Из семинара мы знаем, что если g - гладкая, то

$$\widehat{g(\theta_{ML})} = g(\hat{\theta}_{ML})$$

$$\Rightarrow \widehat{(q^2)}_{ML} = \left(\frac{\sum x_i^2}{n}\right)^2$$

(в) Построить 95%-ый доверительный интервал для q. Из семинара мы знаем, что можем посчитать по формуле.

$$\hat{q}_{ML} - 1.96 * SE(\hat{q}_{ML}) \leq q \leq \hat{q}_{ML} + 1.96 * SE(\hat{q}_{ML})$$

Также мы знаем, что

$$SE(\hat{q}_{ML}) = \sqrt{(\mathbb{D}(\hat{q}_{ML}))}$$

$$\hat{I}(\hat{q}) = \frac{1}{\widehat{Var}(\hat{q})}$$

$$s' = \frac{n}{q^2} - \frac{2\sum x_i^2}{q^3}$$

$$\hat{I} = -\mathbb{E}_{\hat{q}}\left(\frac{n}{\hat{q}^2} - \frac{2\sum x_i^2}{\hat{q}^3}\right) = -\frac{n}{\hat{q}^2} + \frac{2}{\hat{q}^3}\mathbb{E}_{\hat{q}}(\sum x_i^2)$$

$$\mathbb{E}_q(\sum x_i^2) = \sum \int_0^{+\infty} x_i^2 \cdot \frac{2x_i}{q} \cdot e^{\frac{-x_i^2}{q}} dx_i = n \cdot \int_0^{+\infty} x_1^2 \cdot \frac{2x_1}{q} \cdot e^{\frac{-x_1^2}{q}} dx_1 =$$

$$= \left[y = \frac{x_1^2}{q}\right] = nq\int_0^{+\infty} ye^{-y}dy = nq\left(0 - \int_0^{-\infty} e^{-y}dy\right) = nq$$

$$\Rightarrow \hat{I} = -\frac{n}{\hat{q}^2} + \frac{2n}{\hat{q}^2} = \frac{n}{\hat{q}^2} \Rightarrow \mathbb{D}(\hat{q}) = \frac{\hat{q}^2}{n}$$

Ответ:

$$\hat{q}_{ML} - 1.96 * \frac{\hat{q}_{ML}}{\sqrt{n}}) \leq q \leq \hat{q}_{ML} + 1.96 * \frac{\hat{q}_{ML}}{\sqrt{n}}$$

Задача №3

(a)

$$L = \binom{150}{75\ 30\ 45} p_1^{75} p_2^{30} (1 - p_1 - p_2)^{45}$$

$$l = \ln \binom{75\ 30\ 45}{150} + 75 \ln p_1 + 30 \ln p_2 + 45 \ln(1 - p_1 - p_2)$$

$$l'_{p_1} = \frac{75}{p_1} - \frac{45}{1 - p_1 - p_2}$$

$$l'_{p_2} = \frac{30}{p_2} - \frac{45}{1 - p_1 - p_2}$$

$$\begin{cases} 75(1 - \hat{p}_1 - \hat{p}_2) = 45\hat{p}_1 \\ 30(1 - \hat{p}_1 - \hat{p}_2) = 45\hat{p}_2 \end{cases}$$

$$\begin{cases} 5(1 - \hat{p}_1 - \hat{p}_2) = 3\hat{p}_1 \\ 2(1 - \hat{p}_1 - \hat{p}_2) = 3\hat{p}_2 \end{cases}$$

$$\begin{cases} 8\hat{p}_1 + 5\hat{p}_2 = 5 \\ 2\hat{p}_1 + 5\hat{p}_2 = 2 \end{cases}$$

$$\begin{cases} \hat{p}_1 = 0.5 \\ \hat{p}_2 = 0.2 \end{cases}$$

Ответ наша ml оценка равана:

$$\hat{p}_{ML} = \begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix}$$

(b) Проверим гипотезу $H_0 : p_1 = 0.7$ против $H_A : p_1 \neq 0.7$для начала с помощью LR теста Вспомним, что

$$LR = 2(l_{ur} - l_r)$$

l_{ur} - значение логарифмической функции правдоподобия без ограничений (unrestricted), в нашем случае мы нашли $\hat{p}_{ML}$ просто подставим в функцию и посчитаем.

$$l_{ur} = \ln \binom{150}{75\ 30\ 45} + 75 \ln 0.5 + 30 \ln 0.2 + 45 \ln(1 - 0.5 - 0.2)$$

$$\approx 149.35 - 51.99 - 48.28 - 54.18 \approx -5.1$$

$$\frac{30}{\bar{p}_2} - \frac{45}{1-0.7-\bar{p}_2} = 0$$
$$\bar{p}_2 = \frac{9}{75} = 0.12$$

$$l_r = \ln\binom{150}{75\ 30\ 45} + 75\ln 0.7 + 30\ln 0.12 + 45\ln(1-0.7-0.12)$$
$$\approx 149.35 - 26.39 - 63.6 - 77.17 \approx -17.81$$
$$LR = 2(-5.1 + 17.81) = 25.42$$

Одно ограничение $\Rightarrow \chi^2_1 \Rightarrow$ сравниваем с 3.84 $\Rightarrow$ гипотеза отвергается

Теперь посчитаем LM тест

$$LM = \hat{S}_R^T \cdot \widehat{Var}(\hat{S}_R)^{-1} \cdot \hat{S}_R$$

Для этого для начала посчитаем матрицу Гессе

$$l''_{p_1p_1} = -\frac{75}{p_1^2} - \frac{45}{(1-p_1-p_2)^2}$$

$$l''_{p_1p_2} = -\frac{45}{(1-p_1-p_2)^2}$$

$$l''_{p_2p_2} = -\frac{30}{p_2^2} - \frac{45}{(1-p_1-p_2)^2}$$

Так как все константы, то мат ожидания будет тем же посчитаем сразу инфморацию фишера

$$I_r \approx \begin{pmatrix} 1542 & 1388 \\ 1388 & 1721 \end{pmatrix}$$

$$S_r = \begin{pmatrix} \frac{75}{0.7} - \frac{45}{1-0.7-0.12} \\ \frac{30}{0.12} - \frac{45}{1-0.7-0.12} \end{pmatrix} \approx \begin{pmatrix} -143 \\ 0 \end{pmatrix}$$

$$LM \approx \begin{pmatrix} -143 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1541 & 1388 \\ 1388 & 1721 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -143 \\ 0 \end{pmatrix} \approx 48.39$$

H_0 отвергается, так как LM $>$ 3.84

(c)

$$W = (\hat{\gamma}_{ur} - \gamma_0)^T \cdot \widehat{Var}(\hat{\gamma}_{ur})^{-1} \cdot (\hat{\gamma}_{ur} - \gamma_0) =$$

$$H_0 : \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$$

$$H_0 : \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \neq \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$$

$$= \begin{pmatrix} 0.1 & -0.4 \end{pmatrix} \cdot \begin{pmatrix} 800 & 500 \\ 500 & 1250 \end{pmatrix} \cdot \begin{pmatrix} 0.1 \\ -0.4 \end{pmatrix} = 168$$

тут у нас χ^2_2 так как два ограничения $\Rightarrow$ сравниваем с 6 и снова отвергаем гипотезу H_0

(d)

$$\hat{p}_{ML} = \begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix}$$

$$I(\hat{p}_{ML})^{-1} = \begin{pmatrix} \frac{1}{600} & -\frac{1}{1500} \\ -\frac{1}{1500} & \frac{2}{1875} \end{pmatrix} = Var(\hat{p}_{ML})$$

Теперь посчитаем

$$Var(\hat{p}_1 - hatp_2) = Var(\hat{p}_1) - 2cov(\hat{p}_1, \hat{p}_2) + Var(\hat{p}_2) \approx 0.004$$

Ответ:

$$0.3 - 1.96\sqrt{0.004} \leq p_1 - p_2 \leq 0.3 + 1.96\sqrt{0.004}$$

Задача 4

(a) Проверим в лоб линейную зависимость.

$$y_i = f(x_i, \beta_1, \beta_2) = \beta_1 e^{\beta_2 x_i} u_i$$

Проверим по β_1

$$f(x_i, (\alpha\gamma + \zeta\xi), \beta_2) = (\alpha\gamma + \zeta\xi) e^{\beta_2 x_i} u_i =$$

$$= \alpha(\gamma e^{\beta_2 x_i} u_i) + \zeta(\xi e^{\beta_2 x_i} u_i) = \alpha f(x_i, \gamma, \beta_2) + \zeta f(x_i, \xi, \beta_2) \Rightarrow \text{линейна}$$

Проверим по β_2

$$f(x_i, \beta_1, (\alpha\gamma + \zeta\xi)) = \beta_1 e^{(\alpha\gamma + \zeta\xi) x_i} u_i =$$

$$= \beta_1 e^{\alpha\gamma x_i} e^{\zeta\xi x_i} u_i = f(x_i, \beta_1, \alpha\gamma) * f(x_i, 1, \zeta\xi) \Rightarrow \text{нелинейна}$$

(b) Так тут сразу напрашивается замена, ибо мы знаем $\ln u_i \sim \mathcal{N}(0,1)$ сделаем замену, точнее возьмем логарифм от нашей сл. величины, а именно $\xi_i = ln(y_i)$

$$\xi_i = \ln(\beta_1) + \beta_2 x_i + \ln(u_i)$$

и теперь мы можем сказать, что $\xi_i \sim \mathcal{N}(ln(\beta_1) + \beta_2 x_i, 1)$

$$L = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{\sum_i^n \frac{(t_i - \ln\beta_1 - \beta_2 x_i)^2}{2}}$$

$$l = -\frac{n}{2}\ln(2\pi) + \sum_i^n \frac{(t_i - \ln\beta_1 - \beta_2 x_i)^2}{2}$$

$$l'_{\beta_1} = -\frac{2}{\beta_1}\sum_i^n \frac{(t_i - \ln\beta_1 - \beta_2 x_i)}{2}$$

$$l'_{\beta_2} = -2x_i \sum_i^n \frac{(t_i - \ln\beta_1 - \beta_2 x_i)}{2}$$

$$-\frac{2}{\hat{\beta}_1}\sum_i^n \frac{(t_i - \ln\hat{\beta}_1 - \hat{\beta}_2 x_i)}{2} = 0$$

$$-2x_i \sum_i^n \frac{(t_i - \ln\hat{\beta}_1 - \hat{\beta}_2 x_i)}{2} = 0$$

$$\sum_i^n (t_i - \ln\hat{\beta}_1 - \hat{\beta}_2 x_i) = 0$$

$$\sum_i^n (t_i - \ln\hat{\beta}_1 - \hat{\beta}_2 x_i) = 0$$

$$\hat{\beta}_1 = e^{\frac{\sum_i^n t_i + n\hat{\beta}_2 x_i}{n}}$$

$$\hat{\beta}_2 = \frac{\sum_i^n t_i - n\ln\hat{\beta}_1}{x_i}$$