БПМИ182 ПСМО
Данг Куинь Ньы **Домашняя работа 1** 11 октября 2020 г.

Задача 1

а) Будем использовать следующие условия регулярности:

1. Для любой статистики $k(x)$ с конечным вторым моментом выполнено:

$$\frac{\partial}{\partial\theta}\mathbb{E}_\theta k(x) = \mathbb{E}_\theta(k(x)\frac{\partial}{\partial\theta}\ln f_\theta(x))$$

2. Функция правдоподобия (и логарифм от неё) везде дифференцируема по θ.

Доказательство:

Пусть $f(y \mid \theta)$ - функция вероятности. Рассмотрим $D_{KL}[f(y \mid \theta) \mid\mid f(y \mid \phi)]$.

$$\begin{aligned} D_{KL}[f(y \mid \theta) \mid\mid f(y \mid \phi)] &= \mathbb{E}_\theta\big[\ln f(y \mid \theta) - \ln f(y \mid \phi)\big] \\ &= \sum_{y\in Y} f(y \mid \theta)\ln f(y \mid \theta) - \sum_{y\in Y} f(y \mid \theta)\ln f(y \mid \phi) \end{aligned} \tag{1}$$

Разложим в ряд Тейлора функцию вероятности (используя то, что рассматривается $\phi \to \theta$):

$$\ln f(y \mid \phi) = \ln f(y \mid \theta) + (\phi-\theta)\frac{\partial}{\partial\theta}\ln f(y \mid \theta) + (\phi-\theta)^2\frac{\partial^2}{2\partial\theta^2}\ln f(y \mid \theta) + O((\phi-\theta)^3) \tag{2}$$

Подставим ряд (2) в (1):

$$\begin{aligned} D_{KL}[f(y \mid \theta) \mid\mid f(y \mid \phi)] &= \sum_{y\in Y} f(y \mid \theta)\ln f(y \mid \theta) - \sum_{y\in Y} f(y \mid \theta)\Big(\ln f(y \mid \theta) \\ &+ (\phi-\theta)\frac{\partial}{\partial\theta}\ln f(y \mid \theta) + (\phi-\theta)^2\frac{\partial^2}{2\partial\theta^2}\ln f(y \mid \theta) + O((\phi-\theta)^3)\Big) \\ &= \underbrace{\sum_{y\in Y} f(y \mid \theta)\ln f(y \mid \theta) - \sum_{y\in Y} f(y \mid \theta)\ln f(y \mid \theta)}_{=0} \\ &- (\phi-\theta)\underbrace{\sum_{y\in Y} f(y \mid \theta)\frac{\partial}{\partial\theta}\ln f(y \mid \theta)}_{[3]} \\ &- \frac{(\phi-\theta)^2}{2}\underbrace{\sum_{y\in Y} f(y \mid \theta)\frac{\partial^2}{\partial\theta^2}\ln f(y \mid \theta)}_{[4]} + O((\phi-\theta)^3)\Big) \end{aligned}$$

Преобразуем (3) и (4), используя условия регулярности:

$$\begin{aligned}\sum_{y\in Y} f(y\mid\theta)\frac{\partial}{\partial\theta}\ln f(y\mid\theta) &= \sum_{y\in Y}\frac{f(y\mid\theta)}{f(y\mid\theta)}\frac{\partial}{\partial\theta}f(y\mid\theta)\\ &= \sum_{y\in Y}\frac{\partial}{\partial\theta}f(y\mid\theta) \overset{\text{рег-ть}}{=} \frac{\partial}{\partial\theta}\sum_{y\in Y}f(y\mid\theta) = 0\end{aligned} \tag{3}$$

$$\begin{aligned}\sum_{y\in Y} f(y\mid\theta)\frac{\partial^2}{\partial\theta^2}\ln f(y\mid\theta) &= \sum_{y\in Y} f(y\mid\theta)\frac{\partial}{\partial\theta}\Big(\frac{1}{f(y\mid\theta)}\frac{\partial}{\partial\theta}f(y\mid\theta)\Big)\\ &= \sum_{y\in Y} f(y\mid\theta)\Big(-\big(\frac{1}{f(y\mid\theta)}\frac{\partial}{\partial\theta}f(y\mid\theta)\big)^2 + \frac{1}{f(y\mid\theta)}\frac{\partial^2}{\partial\theta^2}f(y\mid\theta)\Big)\\ &= \sum_{y\in Y}\frac{\partial^2}{\partial\theta^2}f(y\mid\theta) - \sum_{y\in Y}f(y\mid\theta)\big(\frac{1}{f(y\mid\theta)}\frac{\partial}{\partial\theta}f(y\mid\theta)\big)^2\\ &\overset{\text{рег-ть}}{=} \underbrace{\frac{\partial^2}{\partial\theta^2}\sum_{y\in Y}f(y\mid\theta)}_{=0} - \underbrace{\sum_{y\in Y}f(y\mid\theta)\big(\frac{\partial}{\partial\theta}\ln f(y\mid\theta)\big)^2}_{=\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\ln f(y\mid\theta)\right)^2\overset{\text{def}}{=}I(\theta))}\\ &= -I(\theta)\end{aligned} \tag{4}$$

Вернёмся к исходному выражению и подставим результаты (3) и (4):

$$D_{KL}[f(y\mid\theta)\mid\mid f(y\mid\phi)] = 0-0-\frac{(\phi-\theta)^2}{2}(-I(\theta))+O((\phi-\theta)^3) = \frac{(\phi-\theta)^2 I(\theta)}{2}+O((\phi-\theta)^3)$$

Used source: Stackexchange: Connection between Fisher metric and the relative entropy

b) Покажем, что при большей информации Фишера ML-оценки лежат ближе к истинному параметру:

Рассмотрим две модели с истинными значениями θ_a и θ_b и их оценками θ_a^*, θ_b^* соответственно, такие что дивергенции у них одинаковы, а информация Фишера первой модели больше другой:

$$\frac{1}{2}I(\theta_a)\Delta\theta_a^2+O(\Delta\theta_a^3) = \frac{1}{2}I(\theta_b)\Delta\theta_b^2+O(\Delta\theta_b^3) \iff \frac{1}{2}(I(\theta_a)\Delta\theta_a^2-I(\theta_b)\Delta\theta_b^2) = O(\Delta\theta_b^3-\Delta\theta_a^3)$$

Допустим, $\Delta\theta_a > \Delta\theta_b$. Но тогда получаем, что правая часть больше нуля, а левая меньше, что противоречит равнеству.

Таким образом, получаем, что при большей информации Фишера ($I(\theta_a) > I(\theta_b)$) расстояние между оценкой и истинным значением меньше ($\Delta\theta_a < \Delta\theta_b$).

Задача 2

a) Рассмотрим функцию правдоподобия:

$$f(X_1, X_2, ..., X_n \mid q) = \left(\frac{2}{q}\right)^n \cdot X_1 \cdot X_2 \cdot ... \cdot X_n \cdot e^{\frac{-\sum\limits_{i=1}^{n} X_i^2}{q}} I\{X_1, ..., X_n > 0\}$$

$$\ell(X, q) = \ln f = n \ln \frac{2}{q} + \sum_{i=1}^{n} \ln X_i - \frac{1}{q} \sum_{i=1}^{n} X_i^2, \ \ X_i > 0$$

Найдём $\hat{q}_{ML}$:

$$\ell'_q = -\frac{n}{q} + \frac{\sum\limits_{i=1}^{n} X_i^2}{q^2}$$

$$-\frac{n}{\hat{q}} + \frac{\sum\limits_{i=1}^{n} X_i^2}{\hat{q}^2} = 0 \iff \frac{n}{\hat{q}} = \frac{\sum\limits_{i=1}^{n} X_i^2}{\hat{q}^2} \iff \hat{q}_{ML} = \overline{X^2}$$

Проверим, является ли найденное значение точкой максимума:

$$\ell''_{qq} = \frac{n}{q^2} - 2\frac{\sum\limits_{i=1}^{n} X_i^2}{q^3} \qquad \ell''_{qq}(\hat{q}) = -\frac{n}{(\overline{X^2})^2} < 0 \implies \hat{q} \text{ - точка максимума}$$

b) Поскольку функция $g(x) = x^2$ - гладкая, можем использовать следующее утверждение из семинара: $\widehat{g(q)} = g(\hat{q})$. Отсюда:

$$\widehat{q^2}_{ML} = \hat{q}^2_{ML} = (\overline{X^2})^2$$

c) Найдём информацию Фишера по определению $I(q) = \mathbb{E}(-\ell''_{qq})$:

$$\ell''_{qq} = \frac{n}{q^2} - \frac{2}{q^3} \sum_{i=1}^{n} X_i^2$$

$$\mathbb{E}(-\ell''_{qq}) = \mathbb{E}\Big(-\frac{n}{q^2} + \frac{2}{q^3}\sum_{i=1}^{n} X_i^2\Big) = \frac{2}{q^3}\mathbb{E}\sum_{i=1}^{n} X_i^2 - \frac{n}{q^2}$$

Найдём $\mathbb{E}X_i^2$:

$$\mathbb{E}X_i^2 = \int\limits_{\mathbb{R}} f(x \mid q)x^2\,dx = \int_0^{+\infty} \frac{2x}{q}e^{-\frac{x^2}{q}}x^2\,dx = \begin{bmatrix} x^2/q = t \\ 2x\,dx/q = dt \end{bmatrix} = \int\limits_0^{+\infty} qte^{-t}\,dt = q\cdot\Gamma(2) = q$$

Вернёмся к информации Фишера:

$$I(q) = \mathbb{E}(-\ell''_{qq}) = \frac{2}{q^3}\cdot nq - \frac{n}{q^2} = \frac{n}{q^2}$$

Из этого получим дисперсию:

$$\mathrm{Var}(\hat{q}) = I^{-1}(\hat{q}) = \frac{\hat{q}^2}{n} = \frac{(\overline{X^2})^2}{n}$$

Получаем 95%-ый доверительный интервал для q:

$$\left[\overline{X^2} - 1.96\cdot\frac{\overline{X^2}}{\sqrt{n}}; \overline{X^2} + 1.96\cdot\frac{\overline{X^2}}{\sqrt{n}}\right]$$

Ответ: a) $\hat{q} = \overline{X^2}$; b) $\widehat{q^2} = (\overline{X^2})^2$; c) $\left[\overline{X^2} - 1.96\cdot\frac{\overline{X^2}}{\sqrt{n}}; \overline{X^2} + 1.96\cdot\frac{\overline{X^2}}{\sqrt{n}}\right]$

Задача 3

$$\begin{cases} y_1 = 75 \text{ - кол-во палочек из вишни} \\ y_2 = 30 \text{ - кол-во палочек из дуба} \\ y_3 = 45 \text{ - кол-во палочек из вяза} \end{cases}, n = 150$$

a) Рассмотрим функцию правдоподобия:

$$f(y,p) = p_1^{y_1}p_2^{y_2}(1-p_1-p_2)^{y_3}\binom{n}{y_1, y_2, y_3}$$

$$f(y,p) = p_1^{75}p_2^{30}(1-p_1-p_2)^{45}\binom{150}{75, 30, 45}$$

$$\ell(y,p) = \ln f = 75\ln p_1 + 30\ln p_2 + 45\ln(1-p_1-p_2) + C$$

Найдём p_{ML}:

$$\begin{cases} \ell'_{p_1} = \frac{75}{p_1} - \frac{45}{1-p_1-p_2} \\ \ell'_{p_2} = \frac{30}{p_2} - \frac{45}{1-p_1-p_2} \end{cases} \longrightarrow \begin{cases} \frac{75}{\hat{p_1}} - \frac{45}{1-\hat{p_1}-\hat{p_2}} = 0 \\ \frac{30}{\hat{p_2}} - \frac{45}{1-\hat{p_1}-\hat{p_2}} = 0 \end{cases} \iff \begin{cases} \hat{p_1} = 0.5 \\ \hat{p_2} = 0.2 \end{cases}$$

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

Проверим, являются ли найденные значения точкой максимума:

$$\ell''_{p_1p_1} = -\frac{y_1}{p_1^2} - \frac{y_3}{p_3^2} \qquad \ell''_{p_2p_2} = -\frac{y_2}{p_2^2} - \frac{y_3}{p_3^2} \qquad \ell''_{p_1p_2} = -\frac{y_3}{p_3^2}$$

$$H(\hat{p}) = \begin{pmatrix} -800 & -500 \\ -500 & -700 \end{pmatrix} \text{ - отрицательно определена } \implies \hat{p} \text{ - точка максимума}$$

b) 1. Проверим гипотезу по LR тесту: $LR = 2(\max \ell_{UR} - \max \ell_R)$

Найдём максимум функции ℓ при H_0:

$$\ell_R(y,p) = 75\ln 0.7 + 30\ln p_2 + 45\ln(1-0.7-p_2) + C$$

$$\ell_R{}'_{p_2} = \frac{30}{p_2} - \frac{45}{0.3-p_2} \longrightarrow \frac{30}{\hat{p_2}_R} - \frac{45}{0.3-\hat{p_2}_R} = 0 \iff \hat{p_2}_R = 0.12$$

Проверим, являются ли найденные значения точкой максимума, рассмотрев матрицу Гессе (вторые производные посчитаны в предыдущем пункте):

$$H(\hat{p}_R) = \begin{pmatrix} -\frac{680000}{441} & -\frac{12500}{9} \\ -\frac{12500}{9} & -\frac{31250}{9} \end{pmatrix} \text{ - отрицательно определена } \implies \hat{p} \text{ - точка максимума}$$

Подставим найденные оценки (для ℓ_{UR} оценка $\hat{p}_{ML}$ найдена в предыдущем пункте):

$$LR = 2\big((75\ln 0.5 + 30\ln 0.2 + 45\ln 0.3) - (75\ln 0.7 + 30\ln 0.12 + 45\ln 0.18)\big) = 26.153$$

$$\chi_{\text{crit}} = 3.84 < 26.153 \implies \text{ гипотеза опровергается}$$

2. Проверим гипотезу по LM тесту: $LM = \hat{s}_R^T(\widehat{\text{Var}}(\hat{s}_R)^{-1})\hat{s}_R$

Из предыдущего пункта имеем оценки параметров при ограничении:

$$\hat{p}_R = \begin{pmatrix} 0.7 \\ 0.12 \end{pmatrix}$$

Найдём информацию Фишера:

$$\ell''_{p_1p_1} = -\frac{y_1}{p_1^2} - \frac{y_3}{p_3^2} \qquad \ell''_{p_2p_2} = -\frac{y_2}{p_2^2} - \frac{y_3}{p_3^2} \qquad \ell''_{p_1p_2} = -\frac{y_3}{p_3^2}$$

$$I(p) = \mathbb{E}(-H) = \mathbb{E}\begin{pmatrix} \frac{y_1}{p_1^2} + \frac{y_3}{p_3^2} & \frac{y_3}{p_3^2} \\ \frac{y_3}{p_3^2} & \frac{y_2}{p_2^2} + \frac{y_3}{p_3^2} \end{pmatrix} = \begin{pmatrix} \frac{np_1}{p_1^2} + \frac{np_3}{p_3^2} & \frac{np_3}{p_3^2} \\ \frac{np_3}{p_3^2} & \frac{np_2}{p_2^2} + \frac{np_3}{p_3^2} \end{pmatrix} = n\begin{pmatrix} \frac{1}{p_1} + \frac{1}{p_3} & \frac{1}{p_3} \\ \frac{1}{p_3} & \frac{1}{p_2} + \frac{1}{p_3} \end{pmatrix}$$

Найдём оценку информации Фишера:

$$I_R = 150\begin{pmatrix} \frac{1}{0.7} + \frac{1}{0.18} & \frac{1}{0.18} \\ \frac{1}{0.18} & \frac{1}{0.12} + \frac{1}{0.18} \end{pmatrix} = 1500\begin{pmatrix} \frac{44}{63} & \frac{5}{9} \\ \frac{5}{9} & \frac{25}{18} \end{pmatrix}$$

$$\widehat{\text{Var}(\hat{s}_R)}^{-1} = I_R^{-1} = \frac{1}{1500} \cdot \frac{3}{250}\begin{pmatrix} 175 & -70 \\ -70 & 88 \end{pmatrix}$$

Подставим оценки в score function:

$$\hat{s}_R = \begin{pmatrix} \frac{75}{0.7} - \frac{45}{0.18} \\ \frac{30}{0.12} - \frac{45}{0.18} \end{pmatrix} = \begin{pmatrix} -\frac{1000}{7} \\ 0 \end{pmatrix}$$

Найдём значение теста:

$$\begin{pmatrix} -\frac{1000}{7} & 0 \end{pmatrix} \cdot \frac{3}{250 \cdot 1500}\begin{pmatrix} 175 & -70 \\ -70 & 88 \end{pmatrix} \cdot \begin{pmatrix} -\frac{1000}{7} \\ 0 \end{pmatrix} = \frac{200}{7} = 28.57$$

$$\chi_{\text{crit}} = 3.84 < 28.57 \implies \text{гипотеза опровергается}$$

c) Проверим гипотезу по W тесту: $W = (\hat{p}_{UR} - p_0)^T\widehat{\text{Var}}^{-1}(\hat{p}_{UR})(\hat{p}_{UR} - p_0)$

Из предыдущего пункта имеем:

$$I = n\begin{pmatrix} \frac{1}{p_1} + \frac{1}{p_3} & \frac{1}{p_3} \\ \frac{1}{p_3} & \frac{1}{p_2} + \frac{1}{p_3} \end{pmatrix}$$

Из пункта (a) имеем оценки: $\hat{p}_{UR} = \begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix}$, подставим в информацию:

$$\hat{Var}^{-1}(\hat{p}_{UR}) = \hat{I}_{UR} = 150 \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ \frac{10}{3} & \frac{25}{3} \end{pmatrix}$$

$$W = 150 \left(\begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix} - \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} \right)^T \cdot \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ \frac{10}{3} & \frac{25}{3} \end{pmatrix} \cdot \left(\begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix} - \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} \right) = 168$$

$$\chi_{crit} = 5.99 < 168 \implies \text{ гипотеза опровергается}$$

d) Из пунктов (a) и (c) имеем:

$$\hat{p}_{ML} = \begin{pmatrix} 0.5 \\ 0.2 \end{pmatrix} \qquad \hat{I}_{UR} = 150 \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ \frac{10}{3} & \frac{25}{3} \end{pmatrix}$$

$$\hat{I}_{UR}^{-1} = \widehat{\mathrm{Var}}(\hat{p}_{UR}) = \begin{pmatrix} \widehat{\mathrm{Var}}(\hat{p_1}) & \widehat{\mathrm{Cov}}(\hat{p_1}, \hat{p_2}) \\ \widehat{\mathrm{Cov}}(\hat{p_1}, \hat{p_2}) & \widehat{\mathrm{Var}}(\hat{p_2}) \end{pmatrix} = \frac{1}{15000} \begin{pmatrix} 25 & -10 \\ -10 & 16 \end{pmatrix}$$

$$\widehat{\mathrm{Var}(\hat{p_1})} = \frac{25}{15000} \qquad \widehat{\mathrm{Var}}(\hat{p_2}) = \frac{16}{15000} \qquad \widehat{\mathrm{Cov}}(\hat{p_1}, \hat{p_2}) = -\frac{10}{15000}$$

Найдём дисперсию и оценку $p_1 - p_2$:

$$\widehat{\mathrm{Var}}(\hat{p_1} - \hat{p_2}) = \widehat{\mathrm{Var}}(\hat{p_1}) - 2\widehat{\mathrm{Cov}}(\hat{p_1}, \hat{p_2}) + \widehat{\mathrm{Var}}(\hat{p_2}) = \frac{25}{15000} - 2 \cdot \left(-\frac{10}{15000} \right) + \frac{16}{15000} = \frac{61}{15000}$$

$$\widehat{p_1 - p_2} = \hat{p_1} - \hat{p_2} = 0.5 - 0.2 = 0.3$$

Получаем 95%-ый доверительный интервал:

$$\left[0.3 - 1.96\sqrt{\frac{61}{15000}}; 0.3 + 1.96\sqrt{\frac{61}{15000}} \right] = [0.175; 0.425]$$

Ответ:

a) $\hat{p_1} = 0.5$, $\hat{p_2} = 0.2$

b) $LR = 26.153$, гипотеза опровергается; $LM = 28.57$, гипотеза опровергается

c) $W = 168$, гипотеза опровергается

d) $[0.175; 0.425]$

Задача 4

$$y_i = \beta_1 e^{\beta_2 x_i} u_i,\ \ u_i \sim \mathcal{N}(0,1)$$

a) Покажем, что данная зависимость линейна по β_1:

Зафиксируем все переменные, кроме β_1 и рассмотрим зависимость как функцию от β_1: $f(\beta_1) = \beta_1 e^{\beta_2 x_i} u_i$

- $f(a+b) = (a+b)e^{\beta_2 x_i} u_i = ae^{\beta_2 x_i} u_i + be^{\beta_2 x_i} u_i = f(a) + f(b)$
- $f(ka) = (ka)e^{\beta_2 x_i} u_i = k\beta_1 e^{\beta_2 x_i} u_i = kf(a)$

Покажем, что данная зависимость нелинейна по β_2:

Зафиксируем все переменные, кроме β_2 и рассмотрим зависимость как функцию от β_2: $g(\beta_2) = \beta_1 e^{\beta_2 x_i} u_i$

- $g(ka) = \beta_1 e^{kax_i} u_i \neq k\beta_1 e^{ax_i} u_i = kg(a)$ при $k \neq 1$, $x_i, a \neq 0$

b) Рассмотрим $\ln y_i$:

$$\ln y_i = \ln \beta_1 + \beta_2 x_i + \ln u_i$$

Из того, что $\ln u_i \sim \mathcal{N}(0,1)$, получаем: $\ln y_i \sim \mathcal{N}(\ln \beta_1 + \beta_2 x_i, 1)$. Рассмотрим её функцию распределения:

$$f(\ln y) = \frac{1}{(2\pi)^{n/2}} \exp\Big(-\frac{1}{2}\sum_{i=1}^{n} (\ln y_i - \ln \beta_1 - \beta_2 x_i)^2 \Big)$$

$$\ell(\ln y) = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{n} (\ln y_i - \ln \beta_1 - \beta_2 x_i)^2$$

Найдём оценки β_1, β_2 ММП:

$$\ell'_{\beta_1} = \frac{1}{\beta_1}\sum_{i=1}^{n} (\ln y_i - \ln \beta_1 - \beta_2 x_i) \qquad \ell'_{\beta_2} = \sum_{i=1}^{n} (\ln y_i x_i - x_i \ln \beta_1 - \beta_2 x_i^2)$$

$$\begin{cases} \sum\limits_{i=1}^{n} (\ln y_i - \ln \hat\beta_1 - \hat\beta_2 x_i) = 0 \\ \sum\limits_{i=1}^{n} (\ln y_i x_i - x_i \ln \hat\beta_1 - \hat\beta_2 x_i^2) = 0 \end{cases} \iff \begin{cases} \ln \hat\beta_1 = \overline{\ln y} - \hat\beta_2 \overline{x} \\ \sum\limits_{i=1}^{n} x_i \ln y_i - \overline{\ln y} \sum\limits_{i=1}^{n} x_i = \hat\beta_2 \Big(-\overline{x}\sum\limits_{i=1}^{n} x_i + \sum\limits_{i=1}^{n} x_i^2 \Big) \end{cases}$$

$$\iff \begin{cases} \ln\hat{\beta}_1 = \overline{\ln y} - \frac{\overline{x}\sum\limits_{i=1}^n x_i(\ln y_i - \overline{\ln y})}{\sum\limits_{i=1}^n x_i(x_i - \overline{x})} \\ \hat{\beta}_2 = \frac{\sum\limits_{i=1}^n x_i(\ln y_i - \overline{\ln y})}{\sum\limits_{i=1}^n x_i(x_i - \overline{x})} \end{cases}$$

Проверим, являются ли найденные значения точкой максимума:

$$\ell''_{\beta_1\beta_1} = -\frac{1}{\beta_1^2}\sum_{i=1}^n(\ln y_i - \ln\beta_1 - \beta_2 x_i + 1) \qquad \ell''_{\beta_2\beta_2} = -\sum_{i=1}^n x_i^2 \qquad \ell''_{\beta_1\beta_2} = -\frac{1}{\beta_1}\sum_{i=1}^n x_i$$

$$H(\hat{\beta}) = \begin{pmatrix} -\frac{n}{\hat{\beta}_1{}^2} & -\frac{1}{\hat{\beta}_1}\sum\limits_{i=1}^n x_i \\ -\frac{1}{\hat{\beta}_1}\sum\limits_{i=1}^n x_i & -\sum\limits_{i=1}^n x_i^2 \end{pmatrix}$$

$-\frac{n}{\hat{\beta}_1{}^2} < 0$, покажем, что $\det H(\hat{\beta}) > 0$:

$$\det H(\hat{\beta}) = \frac{n}{\hat{\beta}_1{}^2}\sum_{i=1}^n x_i^2 - \frac{1}{\hat{\beta}_1{}^2}\Big(\sum_{i=1}^n x_i\Big)^2$$

Из неравенства КБШ:

$$\Big(\sum_{i=1}^n x_i\Big)^2 < \Big(\sum_{i=1}^n 1\Big)\cdot\Big(\sum_{i=1}^n x_i^2\Big) = n\sum_{i=1}^n x_i^2$$

Таким образом, матрица H отрицательно определена, из чего следует, что $\hat{\beta}$ - точка максимума.

Ответ:

a) Зависимость линейна по β_1, нелинейна по β_2

b)

$$\hat{\beta}_1 = \exp\left(\overline{\ln y} - \frac{\overline{x}\sum\limits_{i=1}^n x_i(\ln y_i - \overline{\ln y})}{\sum\limits_{i=1}^n x_i(x_i - \overline{x})}\right) \qquad \hat{\beta}_2 = \frac{\sum\limits_{i=1}^n x_i(\ln y_i - \overline{\ln y})}{\sum\limits_{i=1}^n x_i(x_i - \overline{x})}$$