**Задача.** Рассматривается модель линейной регрессии $Y_i = \beta x_i + \varepsilon_i$, $i = 1, \dots, n$, в которой случайные ошибки $\varepsilon_1, \dots, \varepsilon_n$ являются независимыми нормально распределенными случайными величинами с нулевым математическим ожиданием и дисперсией $\sigma^2$. Имеются следующие данные: $n = 100$, $\sum_{i=1}^n x_i^2 = 50$, $\sum_{i=1}^n x_i y_i = 112$, $\sum_{i=1}^n y_i^2 = 693$.

(a) Запишите логарифмическую функцию правдоподобия.

(b) При помощи метода максимального правдоподобия найдите оценки неизвестных параметров $\beta$ и $\sigma^2$.

(c) Найдите информационную матрицу Фишера $I_n(\theta)$, где $\theta = (\beta, \sigma^2)$.

(d) Найдите асимптотическое распределение вектора $\hat{\theta}_{ML} = (\hat{\beta}_{ML}, \hat{\sigma}^2_{ML})$.

(e) Найдите асимптотические распределения оценок $\hat{\beta}_{ML}$ и $\hat{\sigma}^2_{ML}$.

(f) Постройте двусторонний 95%-ый доверительный интервал для $\beta$.

(g) Постройте двусторонний 95%-ый доверительный интервал для $\sigma^2$.

(h) На уровне значимости 5% протестируйте гипотезу $H_0 : \beta = 2$ против альтернативы $H_1 : \beta \neq 2$.

(i) На уровне значимости 5% протестируйте гипотезу $H_0 : \sigma^2 = 1$ против альтернативы $H_1 : \sigma^2 \neq 1$.

**Решение.** (a) Поскольку $\varepsilon_i \sim N(0, \sigma^2)$, то $Y_i = \beta x_i + \varepsilon_i \sim N(\beta x_i, \sigma^2)$. Значит,

$$f_{Y_i}(y_i; \beta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - \beta x_i)^2}{2\sigma^2}}, \qquad i = 1, \dots, n.$$

Стало быть, функция правдоподобия имеет вид

$$\mathcal{L}(y_1, \dots, y_n; \beta, \sigma^2) = f_{Y_1, \dots, Y_n}(y_1, \dots, y_n; \beta, \sigma^2) \overset{Y_1, \dots, Y_n - \text{независимы}}{=}$$

$$= \prod_{i=1}^n f_{Y_i}(y_i; \beta, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i - \beta x_i)^2}{2\sigma^2}} =$$

$$= (2\pi)^{-n/2} \cdot (\sigma^2)^{-n/2} \cdot e^{-\frac{\sum_{i=1}^n (y_i - \beta x_i)^2}{2\sigma^2}},$$

а значит, логарифмическая функция правдоподобия

$$l(y_1, \dots, y_n; \beta, \sigma^2) = -\tfrac{n}{2} \ln 2\pi - \tfrac{n}{2} \ln \sigma^2 - \frac{\sum_{i=1}^n (y_i - \beta x_i)^2}{2\sigma^2}.$$

(b) Найдем точку максимума логарифмической функции правдоподобия:

$$\begin{cases} \dfrac{\partial l(y_1, \dots, y_n; \beta, \sigma^2)}{\partial \beta} = 0, \\ \dfrac{\partial l(y_1, \dots, y_n; \beta, \sigma^2)}{\partial \sigma^2} = 0; \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \dfrac{\partial l}{\partial \beta} = -\dfrac{\sum_{i=1}^n -2x_i(y_i - \beta x_i)}{2\sigma^2} = \dfrac{\sum_{i=1}^n x_i(y_i - \beta x_i)}{\sigma^2} = 0, \\ \dfrac{\partial l}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{\sum_{i=1}^n (y_i - \beta x_i)^2}{2\sigma^4} = 0. \end{cases}$$

Решением первого уравнения системы является

$$\hat{\beta}_{ML} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{112}{50} = 2.24\,.$$

Выражая из второго уравнения системы параметр $\sigma^2$ и подставляя в полученную формулу вместо параметра $\beta$ найденную выше оценку $\hat{\beta}_{ML}$, приходим к выражению для оценки параметра $\sigma^2$:

$$\hat{\sigma}_{ML}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{\beta}_{ML} x_i)^2 = \frac{1}{n}\left[\sum_{i=1}^{n} y_i^2 - 2\widehat{\beta}_{ML}\sum_{i=1}^{n} x_i y_i + \widehat{\beta}_{ML}^2 \sum_{i=1}^{n} x_i^2\right] =$$

$$= \frac{1}{100}\left[693 - 2\times 2.24\times 112 + 2.24^2\times 50\right] \approx 4.42\,.$$

(c) Информационная матрица Фишера может быть найдена по формуле:

$$I_n(\theta) = -\mathbb{E}\begin{bmatrix} \dfrac{\partial^2 l}{\partial \beta^2} & \dfrac{\partial^2 l}{\partial \beta \partial(\sigma^2)} \\ \dfrac{\partial^2 l}{\partial(\sigma^2)\partial\beta} & \dfrac{\partial^2 l}{(\partial(\sigma^2))^2} \end{bmatrix}.$$

Найдем вторые производные логарифмической функции правдоподобия:

$$\frac{\partial^2 l}{\partial \beta^2} = -\frac{\sum_{i=1}^{n} x_i^2}{\sigma^2},$$

$$\frac{\partial^2 l}{\partial \beta \partial(\sigma^2)} = \frac{\partial^2 l}{\partial(\sigma^2)\partial\beta} = -\frac{\sum_{i=1}^{n} x_i (y_i - \beta x_i)}{\sigma^4},$$

$$\frac{\partial^2 l}{(\partial(\sigma^2))^2} = \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}(y_i - \beta x_i)^2}{\sigma^6}.$$

Найдем теперь математические ожидания данных вторых производных:

$$\mathbb{E}\left[\frac{\partial^2 l}{\partial \beta^2}\right] = -\frac{\sum_{i=1}^{n} x_i^2}{\sigma^2},$$

$$\mathbb{E}\left[\frac{\partial^2 l}{\partial \beta \partial(\sigma^2)}\right] = -\frac{\sum_{i=1}^{n} x_i \mathbb{E}\left[Y_i - \beta x_i\right]}{\sigma^4} = -\frac{\sum_{i=1}^{n} x_i \mathbb{E}\left[\varepsilon_i\right]}{\sigma^4} = 0,$$

$$\mathbb{E}\left[\frac{\partial^2 l}{(\partial(\sigma^2))^2}\right] = \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}\mathbb{E}\left[(Y_i - \beta x_i)^2\right]}{\sigma^6} =$$

$$= \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}\mathbb{E}\left[\varepsilon_i^2\right]}{\sigma^6} = \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}\mathrm{D}(\varepsilon_i)}{\sigma^6} =$$

$$= \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}\sigma^2}{\sigma^6} = \frac{n}{2\sigma^4} - \frac{n\sigma^2}{\sigma^6} = -\frac{n}{2\sigma^4}.$$

Таким образом, получаем искомую информационную матрицу Фишера:

$$I_n(\theta) = \begin{bmatrix} \dfrac{\sum_{i=1}^{n} x_i^2}{\sigma^2} & 0 \\ 0 & \dfrac{n}{2\sigma^4} \end{bmatrix}.$$

(d) Известно, что при выполнении некоторых условий регулярности имеет место следующее свойство асимптотической нормальности оценок максимального правдоподобия:

$$\hat{\theta}_{ML} \overset{\text{as}}{\sim} N(\theta, I_n^{-1}(\hat{\theta}_{ML})).$$

Принимая во внимание данное соотношение, а также равенство

$$I_n^{-1}(\hat{\theta}_{ML}) = \begin{bmatrix} \dfrac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2} & 0 \\ 0 & \dfrac{2\hat{\sigma}_{ML}^4}{n} \end{bmatrix},$$

получаем, что

$$\begin{bmatrix} \hat{\beta}_{ML} \\ \hat{\sigma}_{ML}^2 \end{bmatrix} \overset{\text{as}}{\sim} N\left( \begin{bmatrix} \beta \\ \sigma^2 \end{bmatrix}, \begin{bmatrix} \dfrac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2} & 0 \\ 0 & \dfrac{2\hat{\sigma}_{ML}^4}{n} \end{bmatrix} \right).$$

(e) Непосредственно из пункта (d) вытекает, что

$$\hat{\beta}_{ML} \overset{\text{as}}{\sim} N\left( \beta, \frac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2} \right) \quad \text{и} \quad \hat{\sigma}_{ML}^2 \overset{\text{as}}{\sim} N\left( \sigma^2, \frac{2\hat{\sigma}_{ML}^4}{n} \right).$$

(f) Поскольку $\hat{\beta}_{ML} \overset{\text{as}}{\sim} N\left( \beta, \frac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2} \right)$, то $\dfrac{\hat{\beta}_{ML} - \beta}{\sqrt{\dfrac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2}}} \overset{\text{as}}{\sim} N(0,1)$. Значит, событие $-1.96 \le \dfrac{\hat{\beta}_{ML} - \beta}{\sqrt{\dfrac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2}}} \le 1.96$ имеет вероятность приблизительно 95%. Решая данное неравенство относительно $\beta$, получаем требуемый доверительный интервал:

$$\hat{\beta}_{ML} - 1.96\sqrt{\frac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2}} \le \beta \le \hat{\beta}_{ML} + 1.96\sqrt{\frac{\hat{\sigma}_{ML}^2}{\sum_{i=1}^n x_i^2}}.$$

Стало быть,

$$\underbrace{2.24 - 1.96\sqrt{\frac{4.42}{50}}}_{\approx 1.66} \le \beta \le \underbrace{2.24 + 1.96\sqrt{\frac{4.42}{50}}}_{\approx 2.82}.$$

(g) Поскольку $\hat{\sigma}_{ML}^2 \overset{\text{as}}{\sim} N\left( \sigma^2, \frac{2\hat{\sigma}_{ML}^4}{n} \right)$, то $\dfrac{\hat{\sigma}_{ML}^2 - \sigma^2}{\sqrt{\dfrac{2\hat{\sigma}_{ML}^4}{n}}} \overset{\text{as}}{\sim} N(0,1)$. Значит, событие $-1.96 \le \dfrac{\hat{\sigma}_{ML}^2 - \sigma^2}{\sqrt{\dfrac{2\hat{\sigma}_{ML}^4}{n}}} \le 1.96$ имеет вероятность приблизительно 95%. Решая данное неравенство относительно $\sigma^2$, получаем требуемый доверительный интервал:

$$\hat{\sigma}_{ML}^2 - 1.96\sqrt{\frac{2\hat{\sigma}_{ML}^4}{n}} \le \sigma^2 \le \hat{\sigma}_{ML}^2 + 1.96\sqrt{\frac{2\hat{\sigma}_{ML}^4}{n}}.$$

Стало быть,

$$\underbrace{4.42 - 1.96\sqrt{\frac{2 \times 4.42^2}{100}}}_{\approx 4.42} \le \sigma^2 \le \underbrace{4.42 + 1.96\sqrt{\frac{2 \times 4.42^2}{100}}}_{\approx 5.65}.$$

(h) В пункте (f) мы нашли 95%-ый доверительный интервал для параметра $\beta$. Он оказался равным $[1.66; 2.82]$. Учитывая, что гипотетическое значение параметра $\beta$ из гипотезы $H_0$ (равное 2) принадлежит данному доверительному интервалу, гипотеза $H_0$ не может быть отвергнута.

(i) В пункте (g) мы нашли 95%-ый доверительный интервал для параметра $\sigma^2$. Он оказался равным $[4.42; 5.65]$. Учитывая, что гипотетическое значение параметра $\sigma^2$ из гипотезы $H_0$ (равное 1) не принадлежит данному доверительному интервалу, гипотеза $H_0$ должна быть отвергнута. □