Решения задач по математической статистике

КОНСТАНТИН ЗЮБИН

Задача 1 (КООП)

Пусть $X_1, \dots, X_n \sim \mathcal{N}(\theta_x, \theta_1^2)$ и $Y_1, \dots, Y_n \sim \mathcal{N}(\theta_y, \theta_2^2)$ — независимые случайные величины. Пусть нулевая гипотеза заключается в равенстве $\theta_1 = \theta_2$, а альтернатива — её отрицание.

Построим критерий обобщённого отношения правдоподобий.

Имеем в общем случае

$$L(\theta_1, \theta_2, \theta_x, \theta_y, X_1, \dots, X_n, Y_1, \dots, Y_n) = \prod_{i=1}^{n} \tfrac{1}{2\pi\theta_1\theta_2} e^{-\frac{(X_i-\theta_x)^2}{2\theta_1^2} - \frac{(Y_i-\theta_y)^2}{2\theta_2^2}}.$$

Тогда

$$\ln L(\theta_1, \theta_2, \theta_x, \theta_y, X_1, \dots, X_n, Y_1, \dots, Y_n) = -n\ln(\theta_1) - n\ln(\theta_2) - n\ln(2\pi) - \tfrac{1}{2\theta_1^2}\sum_{i=1}^{n}(X_i - \theta_x)^2 - \tfrac{1}{2\theta_2^2}\sum_{i=1}^{n}(Y_i - \theta_y)^2,$$

$$\tfrac{\partial}{\partial\theta_x} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = \tfrac{1}{\theta_1^2}\left(-n\theta_x + \sum_{i=1}^{n} X_i\right) = 0,$$

$$\tfrac{\partial}{\partial\theta_y} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = \tfrac{1}{\theta_2^2}\left(-n\theta_y + \sum_{i=1}^{n} Y_i\right) = 0,$$

$$\tfrac{\partial}{\partial\theta_1} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = -\tfrac{n}{\theta_1} + \tfrac{1}{\theta_1^3}\sum_{i=1}^{n}(X_i - \theta_x)^2 = 0,$$

$$\tfrac{\partial}{\partial\theta_2} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = -\tfrac{n}{\theta_2} + \tfrac{1}{\theta_2^3}\sum_{i=1}^{n}(Y_i - \theta_y)^2 = 0.$$

Поэтому максимум достигается в точке $\hat{\theta}_x = \overline{X}$, $\hat{\theta}_y = \overline{Y}$, $\hat{\theta}_1 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2}$ и $\hat{\theta}_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$.

При $\theta_1 = \theta_2 = \theta$ имеем

$$L(\theta, \theta_x, \theta_y, X_1, \dots, X_n, Y_1, \dots, Y_n) \prod_{i=1}^{n} \tfrac{1}{2\pi\theta^2} e^{-\frac{(X_i-\theta_x)^2}{2\theta^2} - \frac{(Y_i-\theta_y)^2}{2\theta^2}}.$$

Тогда

$$\ln L(\theta, \theta_x, \theta_y, X_1, \dots, X_n, Y_1, \dots, Y_n) = -2n\ln(\theta) - n\ln(2\pi) - \tfrac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - \theta_x)^2 - \tfrac{1}{2\theta^2}\sum_{i=1}^{n}(y_i - \theta_y)^2,$$

$$\tfrac{\partial}{\partial\theta_x} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = \tfrac{1}{\theta^2}\left(-n\theta_x + \sum_{i=1}^{n} X_i\right) = 0,$$

$$\tfrac{\partial}{\partial\theta_y} L(\theta_1, \theta_2, X_1, \dots, X_n, Y_1, \dots, Y_n) = \tfrac{1}{\theta^2}\left(-n\theta_y + \sum_{i=1}^{n} Y_i\right) = 0,$$

$$\frac{\partial}{\partial\theta}L(\theta,X_1,\ldots,X_n,Y_1,\ldots,Y_n)=-\frac{2n}{\theta}+\frac{1}{\theta^3}\sum_{i=1}^{n}(X_i-\theta_x)^2+\frac{1}{\theta^3}\sum_{i=1}^{n}(Y_i-\theta_y)^2=0.$$

Тогда максимум достигается в точек $\hat\theta_x=\overline{X},\hat\theta_y=\overline{Y}$ и $\hat\theta=\sqrt{\frac{1}{2n}\sum\limits_{i=1}^{n}\left((X_i-\overline{X})^2+(Y_i-\overline{Y})^2\right)}$.

Положим $S_X=\sum\limits_{i=1}^{n}(X_i-\overline{X}), S_Y=\sum\limits_{i=1}^{n}(Y_i-\overline{Y})$ и $S=\frac{1}{2n}\sum\limits_{i=1}^{n}\left((X_i-\overline{X})^2+(Y_i-\overline{Y})^2\right)$.

Имеем

$$\lambda=\frac{\sup_{\theta\in\Theta}L(\theta_1,\theta_2,\theta_x,\theta_y,X_1,\ldots,X_n,Y_1,\ldots,Y_n)}{\sup_{\theta\in\Theta_0}L(\theta,\theta_x,\theta_y,X_1,\ldots,X_n,Y_1,\ldots,Y_n)}=\frac{\frac{1}{S_X^{\frac{n}{2}}S_Y^{\frac{n}{2}}}\cdot\frac{1}{(2\pi)^n}e^{-n}}{\frac{1}{S^n}\cdot\frac{1}{(2\pi)^n}e^{-n}}=\frac{S^n}{\sqrt{S_X^nS_Y^n}}$$

и

$$\ln\lambda=n\ln S-\frac{n}{2}\ln S_X-\frac{n}{2}\ln S_Y.$$

Так как $\dim\Theta-\dim\Theta_0=4-3=1$, то $2\ln\lambda\underset{\theta\in\Theta_0}{\xrightarrow{d}}\xi\sim\chi_1^2$.

Тогда $\mathrm{P}_{H_0}(2n\ln S-n\ln S_X-n\ln S_Y>q_{1-\alpha})\underset{n\to+\infty}{\to}\alpha$, где $q_{1-\alpha}$ — квантиль уровня $1-\alpha$ распределения χ_1^2. Нулевая гипотеза отклоняется при $2n\ln S-n\ln S_X-n\ln S_Y>q_{1-\alpha}$.

Задача 1 (Vald)

Построим вальдовский критерий в рамках задачи 1а. Пусть $h(\theta_1,\theta_x,\theta_2,\theta_y)=\theta_1-\theta_2$. Нулевая гипотеза в этом случае обретает равносильную формулировку $h(\theta_1,\theta_2,\theta_x,\theta_y)=0$.

Выше мы вычислили точку максимума на Θ как точку с координатами $\hat\theta_x=\overline{X}$, $\hat\theta_y=\overline{Y}$, $\hat\theta_1=\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}=\sqrt{S_X}$ и $\hat\theta_2=\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(Y_i-\overline{Y})^2}=\sqrt{S_Y}$.

Вычислим информационную матрицу для $n=1$ в точке максимума:

$$I_1(\theta_1,\theta_x,\theta_2,\theta_y)=\mathrm{E}(-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln L)\Big|_{\theta=\hat\theta}=$$

$$=\left.\begin{pmatrix}\mathrm{E}\left(-\frac{n}{\theta_1^2}+\frac{3}{\theta_1^4}\sum\limits_{i=1}^{n}(X_i-\theta_x)^2\right) & \mathrm{E}\left(\frac{2}{\theta_1^3}\left(-n\theta_x+\sum\limits_{i=1}^{n}X_i\right)\right) & 0 & 0\\ \mathrm{E}\left(\frac{2}{\theta_1^3}\left(-n\theta_x+\sum\limits_{i=1}^{n}X_i\right)\right) & \mathrm{E}\frac{n}{\theta_1^2} & 0 & 0\\ 0 & 0 & \mathrm{E}\left(-\frac{n}{\theta_2^2}+\frac{3}{\theta_2^4}\sum\limits_{i=1}^{n}(Y_i-\theta_y)^2\right) & \mathrm{E}\left(\frac{2}{\theta_2^3}\left(-n\theta_y+\sum\limits_{i=1}^{n}Y_i\right)\right)\\ 0 & 0 & \mathrm{E}\left(\frac{2}{\theta_2^3}\left(-n\theta_y+\sum\limits_{i=1}^{n}Y_i\right)\right) & \mathrm{E}\frac{n}{\theta_2^2}\end{pmatrix}\right|_{\theta=\hat\theta,n=1}=$$

$$=\left.\begin{pmatrix}\frac{2n}{\theta_1^2} & 0 & 0 & 0\\ 0 & \frac{n}{\theta_1^2} & 0 & 0\\ 0 & 0 & \frac{2n}{\theta_2^2} & 0\\ 0 & 0 & 0 & \frac{n}{\theta_2^2}\end{pmatrix}\right|_{\theta=\hat\theta,n=1}=\begin{pmatrix}\frac{2}{S_X} & 0 & 0 & 0\\ 0 & \frac{1}{S_X} & 0 & 0\\ 0 & 0 & \frac{2}{S_Y} & 0\\ 0 & 0 & 0 & \frac{1}{S_Y}\end{pmatrix}.$$

Градиент h есть $(1,0,-1,0)$, поэтому

$$JI_1^{-1}(\hat\theta)J^T=\frac{S_X+S_Y}{2}.$$

Тогда для квантиля $q_{1-\alpha}$ получаем

$$\mathrm{P}_{H_0}(\frac{2n(\sqrt{S_Y}-\sqrt{S_X})^2}{S_X+S_Y}>q_{1-\alpha})\underset{n\to+\infty}{\to}\alpha.$$

Нулевая гипотеза отвергается при $\frac{2n(\sqrt{S_Y}-\sqrt{S_X})^2}{S_X+S_Y}>q_{1-\alpha}$.

Задача 2

Пусть фиксировано целое положительное число k и имеются k последовательностей независимых случайных величин $X_{k,1}, \ldots, X_{k,n_k} \sim \mathrm{Exp}(\frac{1}{\theta_k})$. Пусть нулевая гипотеза состоит в том, что $\theta_1 = \ldots = \theta_k$.

Построим score-test для проверки гипотезы.

Имеем

$$\ln L(\overrightarrow{\theta}, X_{ij}) = \ln \prod_{i=1}^{k} \prod_{j=1}^{n_i} \tfrac{1}{\theta_i} e^{-\frac{X_{ij}}{\theta_i}} I(X_{ij} > 0) = \sum_{i=1}^{k} \left(-n_i \ln \theta_i - \tfrac{1}{\theta_i} \sum_{j=1}^{n_i} X_{ij} \right).$$

Частная производная по θ_i не равна нулю только для i-го слагаемого в внешней сумме. Поэтому

$$\tfrac{\partial}{\partial \theta_i} \ln L(\overrightarrow{\theta}, X_{ij}) = \tfrac{\partial}{\partial \theta_i} \left(-n_i \ln \theta_i - \tfrac{1}{\theta_i} \sum_{j=1}^{n_i} X_{ij} \right) = \tfrac{-n_i}{\theta_i} + \tfrac{1}{\theta_i^2} \sum_{j=1}^{n_i} X_{ij}.$$

Найдём оценку максимального правдоподобия при условии гипотезы ($\theta = \theta_1 = \ldots = \theta_k$). Имеем

$$\sum_{i=1}^{k} \left(\tfrac{-n_i}{\theta} + \tfrac{1}{\theta^2} \sum_{j=1}^{n_i} X_{ij} \right) = 0.$$

Отсюда

$$\hat{\theta} = \frac{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} X_{ij}}{\sum\limits_{i=1}^{k} n_i}.$$

Вычислим информационную матрицу. Из формул для частных производных видим, что вторые частные производные по θ_i и θ_j равны 0 при $i \neq j$. Поэтому матрица диагональна. Вычислим значения элементов на диагонали. Тогда

$$-\tfrac{\partial^2}{\partial^2 \theta_i} \ln L(\theta, X_{ij}) = -\tfrac{n_i}{\theta_i^2} + \tfrac{2}{\theta_i^3} \sum_{j=1}^{n_i} X_{ij},$$

$$I_{ii}(\theta_i) = \mathrm{E} \left(-\tfrac{n_i}{\theta_i^2} + \tfrac{2}{\theta_i^3} \sum_{j=1}^{n_i} X_{ij} \right) = -\tfrac{n_i}{\theta_i^2} + \tfrac{2}{\theta_i^3} \cdot n_i \theta_i = \tfrac{n_i}{\theta_i^2}.$$

Отсюда

$$\sum_{i=1}^{k} \tfrac{\hat{\theta}^2}{n_i} \left(\tfrac{-n_i}{\hat{\theta}} + \tfrac{1}{\hat{\theta}^2} \sum_{j=1}^{n_i} X_{ij} \right)^2 = \tfrac{\hat{\theta}^2}{\hat{\theta}^4} \sum_{i=1}^{k} \tfrac{1}{n_i} \left(\sum_{j=1}^{n_i} X_{ij} - n_i \hat{\theta} \right)^2 = \frac{\sum\limits_{i=1}^{k} \frac{1}{n_i} \left(\sum\limits_{j=1}^{n_i} X_{ij} - n_i \hat{\theta} \right)^2}{\left(\frac{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} X_{ij}}{\sum\limits_{i=1}^{k} n_i} \right)^2}.$$

Эта случайная величина сходится по распределению к χ^2_{k-1}-случайной величине. Пусть $q_{1-\alpha}$ — квантиль уровня $1 - \alpha$. Тогда

$$\mathrm{P} \left(\frac{\sum\limits_{i=1}^{k} \frac{1}{n_i} \left(\sum\limits_{j=1}^{n_i} X_{ij} - n_i \hat{\theta} \right)^2}{\left(\frac{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} X_{ij}}{\sum\limits_{i=1}^{k} n_i} \right)^2} > q_{1-\alpha} \right) \underset{n \to +\infty}{\to} \alpha.$$

Задача 3а

Пусть имеются три набора независимых случайных величин $X_1, \ldots, X_n \sim \mathcal{N}(\theta_1, \sigma^2)$, $Y_1, \ldots, Y_n \sim \mathcal{N}(\theta_2, \sigma^2)$ и $Z_1, \ldots, Z_n \sim \mathcal{N}(\theta_3, \sigma^2)$, представляющие значения длин рёбер параллелепипеда.

Пусть V — фиксированное положительное число. Построим критерий Вальда для проверки гипотезы о том, что объём параллелепипеда равен V:

$$\theta_1\theta_2\theta_3 = V.$$

В пространстве всех допустимых значений $\Theta = \{(\theta_1, \theta_2, \theta_3, \sigma) \mid \theta_1, \theta_2, \theta_3, \sigma > 0\}$ равенство $\theta_1\theta_2\theta_3 = V$ определяет гиперповерхность, то есть подмногообразие размерности $3 - 1 = 2$.

Для построения критерия Вальда будем рассматривать функцию $h(x, y, z, t) = xyz - V$. Тогда $\operatorname{grad} h = (yz, zx, xy, 0)$.

Вычислим функцию правдоподобия:

$$L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \prod_{i=1}^{n} \tfrac{1}{\sqrt{8\pi^3}\sigma^3} e^{-\frac{(X_i-\theta_1)^2+(Y_i-\theta_2)^2+(Z_i-\theta_3)^2}{2\sigma^2}};$$

$$\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\tfrac{3n}{2}\ln(2\pi) - 3n\ln\sigma - \tfrac{1}{2\sigma^2}\sum_{i=1}^{n}\left((X_i-\theta_1)^2 + (Y_i-\theta_2)^2 + (Z_i-\theta_3)^2\right).$$

Для частных производных имеем:

$$\tfrac{\partial}{\partial\theta_1}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \tfrac{1}{\sigma^2}(-n\theta_1 + \sum_{i=1}^{n} X_i) = \tfrac{n}{\sigma^2}(\overline{X} - \theta_1),$$

$$\tfrac{\partial}{\partial\theta_2}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \tfrac{1}{\sigma^2}(-n\theta_2 + \sum_{i=1}^{n} Y_i) = \tfrac{n}{\sigma^2}(\overline{Y} - \theta_2),$$

$$\tfrac{\partial}{\partial\theta_3}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \tfrac{1}{\sigma^2}(-n\theta_3 + \sum_{i=1}^{n} Z_i) = \tfrac{n}{\sigma^2}(\overline{Z} - \theta_3),$$

$$\tfrac{\partial}{\partial\sigma}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\tfrac{n}{\sigma} + \tfrac{1}{\sigma^3}\sum_{i=1}^{n}\left((X_i-\theta_1)^2 + (Y_i-\theta_2)^2 + (Z_i-\theta_3)^2\right).$$

Тогда точка максимума есть точка с координатами $\hat{\theta}_1 = \overline{X}, \hat{\theta}_2 = \overline{Y}, \hat{\theta}_3 = \overline{Z}$ и $\hat{\sigma}^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\left((X_i-\overline{X})^2 + (Y_i-\overline{Y})^2 + (Z_i-\overline{Z})^2\right)$. Матождиания от минус частных производных второго порядка равны

$$-\operatorname{E}\tfrac{\partial^2}{\partial\theta_i^2}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\operatorname{E}\left(-\tfrac{n}{\sigma^2}\right) = -\tfrac{n}{\sigma^2};$$

$$-\operatorname{E}\tfrac{\partial^2}{\partial\theta_i\partial\theta_j}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\operatorname{E}0 = 0 \ (i \neq j);$$

$$-\operatorname{E}\tfrac{\partial^2}{\partial\sigma^2}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\operatorname{E}\left(\tfrac{n}{\sigma^2} - \tfrac{3}{\sigma^4}\sum_{i=1}^{n}\left((X_i-\theta_1)^2 + (Y_i-\theta_2)^2 + (Z_i-\theta_3)^2\right)\right) = -\tfrac{n}{\sigma^2} + \tfrac{9\sigma^2}{\sigma^3} = \tfrac{8n}{\sigma^2};$$

$$-\operatorname{E}\tfrac{\partial^2}{\partial\sigma\partial\theta_1}\ln L(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\operatorname{E}\tfrac{2}{\sigma^3}(n\theta_1 - \sum_{i=1}^{n} X_i) = 0.$$

Тогда информационная матрица в точке максимума равна

$$I_n(\hat{\theta}) = \begin{pmatrix} \frac{n}{\hat{\sigma}^2} & 0 & 0 & 0 \\ 0 & \frac{n}{\hat{\sigma}^2} & 0 & 0 \\ 0 & 0 & \frac{n}{\hat{\sigma}^2} & 0 \\ 0 & 0 & 0 & \frac{8n}{\hat{\sigma}^2} \end{pmatrix}.$$

Имеем

$$g = \operatorname{grad} h(\hat\theta) I_1(\hat\theta)^{-1} \operatorname{grad} h(\hat\theta)^t = \hat\sigma^2(\overline{X}^2\overline{Y}^2 + \overline{Y}^2\overline{Z}^2 + \overline{Z}^2\overline{X}^2).$$

Тогда случайная величина

$$\frac{nh^2}{g} = \frac{n(\overline{X}\overline{Y}\overline{Z} - V)^2}{\hat\sigma^2(\overline{X}^2\overline{Y}^2 + \overline{Y}^2\overline{Z}^2 + \overline{Z}^2\overline{X}^2)}$$

сходится по распределению к $\chi^2_{3-2} = \chi^2_1$.

Если $q_{1-\alpha}$ — квантиль уровня $1-\alpha$ соответствующего распределения, то

$$\mathrm{P}\left(\frac{n(\overline{X}\cdot\overline{Y}\cdot\overline{Z} - V)^2}{\hat\sigma^2(\overline{X}^2\overline{Y}^2 + \overline{Y}^2\overline{Z}^2 + \overline{Z}^2\overline{X}^2)} > q_{1-\alpha}\right) \underset{n\to+\infty}{\to} \alpha.$$

Задача 3б

В условиях задачи 3а построим score-test для проверки гипотезы $\theta_1 = \theta_2 = \theta_3$. Данная система задаёт подмногообразие размерности 2 в пространстве $\Theta = \{(\theta_1, \theta_2, \theta_3, \sigma) \mid \theta_1, \theta_2, \theta_3, \sigma > 0\}$.

Найдём точку максимума функции правдоподобия на этом подмногообразии. Имеем

$$L_{H_0}(\theta, \sigma, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \prod_{i=1}^n \frac{1}{\sqrt{8\pi^3}\sigma^3} e^{-\frac{(X_i-\theta)^2 + (Y_i-\theta)^2 + (Z_i-\theta)^2}{2\sigma^2}};$$

$$\ln L_{H_0}(\overrightarrow{\theta}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\frac{3n}{2}\ln(2\pi) - 3n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n \left((X_i-\theta)^2 + (Y_i-\theta)^2 + (Z_i-\theta)^2\right).$$

Вычислим частные производные:

$$\frac{\partial}{\partial\theta}\ln L_{H_0}(\theta, \sigma, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = \frac{1}{\sigma^2}(-3n\theta + \sum_{i=1}^n (X_i + Y_i + Z_i));$$

$$\frac{\partial}{\partial\sigma}\ln L_{H_0}(\theta, \sigma, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) = -\frac{3n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^n \left((X_i-\theta)^2 + (Y_i-\theta)^2 + (Z_i-\theta)^2\right).$$

Тогда координаты точки максимума равны $\hat\theta_{H_0} = \frac{1}{3}(\overline{X} + \overline{Y} + \overline{Z})$ и $\hat\sigma^2_{H_0} = \frac{1}{3n}\sum\limits_{i=1}^n \left((X_i - \hat\theta)^2 + (Y_i - \hat\theta)^2 + (Z_i - \hat\theta)^2\right)$.

Оценка

$$T = \frac{1}{n}\operatorname{grad} L(\hat\theta_{H_0}, \hat\theta_{H_0}, \hat\theta_{H_0}, \hat\sigma_{H_0}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z}) I_1(\hat\theta_{H_0}, \hat\theta_{H_0}, \hat\theta_{H_0}, \hat\sigma_{H_0})^{-1}\operatorname{grad} L(\hat\theta_{H_0}, \hat\theta_{H_0}, \hat\theta_{H_0}, \hat\sigma_{H_0}, \overrightarrow{X}, \overrightarrow{Y}, \overrightarrow{Z})^t =$$

$$= \frac{\hat\sigma^2_{H_0}}{n}\cdot\frac{n^2}{\hat\sigma^4_{H_0}}\left((\hat\theta_{H_0} - \overline{X})^2 + (\hat\theta_{H_0} - \overline{Y})^2 + (\hat\theta_{H_0} - \overline{Z})^2\right) +$$

$$+ \frac{\hat\sigma^2_{H_0}}{8n}\cdot\left(-\frac{n}{\hat\sigma_{H_0}} + \frac{1}{\hat\sigma^3_{H_0}}\sum_{i=1}^n \left((X_i - \hat\theta_{H_0})^2 + (Y_i - \hat\theta_{H_0})^2 + (Z_i - \hat\theta_{H_0})^2\right)\right)^2$$

сходится по распределению к ξ^2_2.

Если $q_{1-\alpha}$ — квантиль уровня $1-\alpha$, то имеем

$$\mathrm{P}(T > q_{1-\alpha}) \underset{n\to+\infty}{\to} \alpha.$$

Задача 4

Пусть $X_1, \ldots, X_n \sim \mathcal{N}(\theta_x, \theta^2)$ и $Y_1, \ldots, Y_n \sim \mathcal{N}(\theta_y, \theta^2)$ — независимые случайные величины. Построим критерий обобщённого отношения правдоподобий для проверки гипотезы $\theta_x = \theta_y$.

Вычислим функции правдоподобия в общем случае и при условии гипотезы.

$$L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = \prod_{i=1}^n \frac{1}{2\pi\theta^2} e^{-\frac{(X_i-\theta_x)^2 + (Y_i-\theta_y)^2}{2\theta^2}};$$

$$\ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = -n \ln 2\pi - 2n \ln \theta - \tfrac{1}{2\theta^2} \sum_{i=1}^{n} \left((X_i - \theta_x)^2 + (Y_i - \theta_y)^2 \right).$$

Вычислим частные производные:

$$\tfrac{\partial}{\partial \theta_x} \ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = \tfrac{1}{\theta^2} \left(-n\theta_x + \sum_{i=1}^{n} X_i \right);$$

$$\tfrac{\partial}{\partial \theta_y} \ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = \tfrac{1}{\theta^2} \left(-n\theta_y + \sum_{i=1}^{n} Y_i \right);$$

$$\tfrac{\partial}{\partial \theta} \ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = -\tfrac{2n}{\theta} + \tfrac{1}{\theta^3} \sum_{i=1}^{n} \left((X_i - \theta_x)^2 + (Y_i - \theta_y)^2 \right).$$

Максимум достигается в точке с координатами $\hat{\theta}_x = \overline{X}, \hat{\theta}_y = \overline{Y}$ и $\hat{\theta}^2 = \frac{1}{2n} \sum\limits_{i=1}^{n} \left((X_i - \overline{X})^2 + (Y_i - \overline{Y})^2 \right)$. В точке максимума значение равно

$$L(\hat{\theta}_x, \hat{\theta}_y, \hat{\theta}, \overrightarrow{X}, \overrightarrow{Y}) = \tfrac{ne^{-n}}{\pi} \left(\sum_{i=1}^{n} (X_i - \overline{X})^2 + (Y_i - \overline{Y})^2 \right)^{-1}.$$

В предположении гипотезы имеем

$$L_0(\theta_0, \theta, \overrightarrow{X}, \overrightarrow{Y}) = \prod_{i=1}^{n} \tfrac{1}{2\pi\theta^2} e^{-\frac{(X_i - \theta_0)^2 + (Y_i - \theta_0)^2}{2\theta^2}};$$

$$\ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = -n \ln 2\pi - 2n \ln \theta - \tfrac{1}{2\theta^2} \sum_{i=1}^{n} \left((X_i - \theta_0)^2 + (Y_i - \theta_0)^2 \right).$$

Снова вычислим частные производные:

$$\tfrac{\partial}{\partial \theta_0} \ln L(\theta_0, \theta, \overrightarrow{X}, \overrightarrow{Y}) = \tfrac{1}{\theta^2} \left(-2n\theta_0 + \sum_{i=1}^{n} (X_i + Y_i) \right);$$

$$\tfrac{\partial}{\partial \theta} \ln L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y}) = -\tfrac{2n}{\theta} + \tfrac{1}{\theta^3} \sum_{i=1}^{n} \left((X_i - \theta_0)^2 + (Y_i - \theta_0)^2 \right).$$

Максимум достигается в точке $\hat{\theta}_0 = \frac{\overline{X} + \overline{Y}}{2}$, $\hat{\theta}^2 = \frac{1}{2n} \sum\limits_{i=1}^{n} \left((X_i - \frac{\overline{X}+\overline{Y}}{2})^2 + (Y_i - \frac{\overline{X}+\overline{Y}}{2})^2 \right)$. Значение в этой точке равно

$$L(\hat{\theta}_0, \hat{\theta}, \overrightarrow{X}, \overrightarrow{Y}) = \tfrac{ne^{-n}}{\pi} \left(\sum_{i=1}^{n} (X_i - \tfrac{\overline{X}+\overline{Y}}{2})^2 + (Y_i - \tfrac{\overline{X}+\overline{Y}}{2})^2 \right)^{-1}.$$

Тогда для

$$\lambda = \tfrac{\sup_{\theta \in \Theta} L(\theta_x, \theta_y, \theta, \overrightarrow{X}, \overrightarrow{Y})}{\sup_{\theta \in \Theta_0} L(\theta_0, \theta, \overrightarrow{X}, \overrightarrow{Y})}$$

имеем

$$2 \ln \lambda = \ln \left(\sum_{i=1}^{n} (X_i - \tfrac{\overline{X}+\overline{Y}}{2})^2 + (Y_i - \tfrac{\overline{X}+\overline{Y}}{2})^2 \right) - \ln \left(\sum_{i=1}^{n} (X_i - \overline{X})^2 + (Y_i - \overline{Y})^2 \right) \xrightarrow[n \to +\infty]{d} \xi \sim \chi_1^2.$$

Если $q_{1-\alpha}$ — квантиль уровня $1-\alpha$, то

$$\mathrm{P}\left(\ln\left(\sum_{i=1}^{n}(X_i - \tfrac{\overline{X}+\overline{Y}}{2})^2 + (Y_i - \tfrac{\overline{X}+\overline{Y}}{2})^2\right) - \ln\left(\sum_{i=1}^{n}(X_i - \overline{X})^2 + (Y_i - \overline{Y})^2\right) > q_{1-\alpha}\right) \to \alpha.$$