$\boxed{1}$ 1) $x_1, \dots, x_n$ – выборка $\mathcal{N}(a, \sigma^2)$

$$L_x(\theta) = \prod_{i=1}^{n} p_\theta(x_i) \; ; \; l_x(\theta) = \sum_{i=1}^{n} \ln p_\theta(x_i)$$

АНО: $i^{-1}(\theta)$, где $i_{jk} = \frac{\partial l_x}{\partial \theta_j} \frac{\partial l_x}{\partial \theta_k}$

$$p = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

$$l_x(\theta) = \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-a)^2}{2\sigma^2}}\right) = \sum_{i=1}^{n} \left(\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(x_i-a)^2}{2\sigma^2}\right) =$$

$$= \sum_{i=1}^{n} \left(\ln\frac{1}{\sqrt{2\pi}} - \frac{1}{2}\ln\sigma^2 - \frac{(x_i-a)^2}{2\sigma^2}\right)$$

1) $\theta = (a, \sigma^2)$

$$\begin{pmatrix} \frac{\partial l_x(\theta)}{\partial a} \\ \frac{\partial l_x(\theta)}{\partial \sigma^2} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \frac{(x_i-a)}{\sigma^2} \\ \sum_{i=1}^{n} \left(-\frac{1}{\sigma^2} + \frac{(x_i-a)^2}{2\sigma^4}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow a = \hat{a}_{омп} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\sum_{i=1}^{n} \left(-\frac{1}{\sigma^2} + \frac{(x_i-a)^2}{2\sigma^4}\right) = -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^{n} (x_i-a)^2 = 0$$

$$\Rightarrow \sigma = \hat{\sigma}_{омп} = \frac{\sum_{i=1}^{n} (x_i - \hat{a}_{омп})}{n}$$

2) $\theta = \sigma^2$ $\quad a$ - известно

$$\frac{\partial l_x(\theta)}{\partial \sigma^2} = \sum_{i=1}^{n} \left(-\frac{1}{\sigma^2} + \frac{(x_i - a)^2}{2\sigma^4} \right) = 0 \Rightarrow$$

$$\sigma^2 = \hat{\sigma}^2_{\text{омп}} = \frac{\sum_{i=1}^{n} (x_i - a)^2}{n}$$

3) $\theta = a$ $\quad \sigma^2$ известно

$$\frac{\partial l_x(\theta)}{\partial a} = \sum_{i=1}^{n} \frac{x_i - a}{\sigma^2} = 0 \Rightarrow a = \hat{a}_{\text{омп}} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$i = E_a \left(\frac{\partial \sum_{i=1}^{n} \left(\ln \frac{1}{\sqrt{2\pi}} - \ln \sigma - \frac{(x_i - a)^2}{2\sigma^2} \right)}{\partial a} \right) =$$

$$= E_a \left(\sum_{i=1}^{n} \left(\frac{x_i - a}{\sigma^2} \right) \right)^2 = E_a \sum_{i,j=1}^{n} \frac{(x_i - a)(x_j - a)}{\sigma^4} =$$

$$= \sum_{i,j=1}^{n} \frac{E_a x_i x_j - a E_a (x_i + x_j) + a^2}{\sigma^4} = \sum_{i,j=1}^{n} \frac{\sigma^2 + a^2}{\sigma^4} =$$

$$= n^2 \frac{\sigma^2 + a^2}{\sigma^4} \Rightarrow \text{асимптот. дисп. } \frac{\sigma^4}{(\sigma^2 + a^2) n^2}$$

2) $x_1, \dots, x_n$ - выборка $\sim Pois(\theta)$

$$p(\theta) = \frac{\theta^k}{k!} e^{-\theta}$$

$$\frac{l_x(\theta)}{\partial \theta} = \sum_{i=1}^{n} \ln \left(\frac{\theta^{x_i}}{x_i!} e^{-\theta} \right) = \sum_{i=1}^{n} \left(x_i \ln \theta - \ln(x_i!) - \theta \right)$$

$$\frac{\partial l_x(\theta)}{\partial\theta} = \sum_{i=1}^{n}\left(\frac{x_i}{\theta} - 1\right) = 0 \Rightarrow \theta = \hat\theta_{ОМП} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$i = \mathbb{E}_\theta\left(\sum_{i=1}^{n}\left(\frac{x_i-\theta}{\theta}\right)\right)^2 = \sum_{i,j=1} \mathbb{E}_\theta \frac{(x_i-\theta)(x_j-\theta)}{\theta^2} = \sum_{i,j=1}^{n} \frac{\mathbb{E}_\theta x_i x_j - 2\theta^2+\theta^2}{\theta^2} =$$

$$= \sum_{i,j=1}^{n} \frac{\theta^2+\theta-\theta^2}{\theta^2} = \sum_{i,j=1}^{n} \frac{1}{\theta} = \frac{n^2}{\theta} \Rightarrow \text{ас. дисперсия } \frac{\theta}{n^2}$$

$\boxed{2}$ $x_1, \dots, x_n$ – выборка

1) $L(\theta) = \prod_{i=1}^{n} P_\theta(X_i = x_i) = \prod_{i=1}^{n} \theta_{x_i}$

$$l = \sum_{i=1}^{n} \ln \theta_{x_i}$$

вероятность принадлежности i-го объекта к классу j

$$CrossEntropy = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} \mathbb{I}(i=j)\log(p_{ij})$$

$$logloss = -\frac{1}{n}\sum_{i=1}^{n}\left(y_i \log(p_i) + (1-y_i)\log(1-p_i)\right)$$

Заметим, что при $k=2$ они совпадают

2) $$\left\{\begin{array}{l} \dfrac{\partial l}{\partial\theta_j} = \sum_{i=1}^{n} \dfrac{1}{\theta_j}\mathbb{I}(i=j) = 0 \quad \forall j \\ \sum_{i=1}^{k}\theta_i \Rightarrow \theta_k = 1 - \sum_{i=1}^{k-1}\theta_i \end{array}\right\} \Rightarrow$$

$$\hat\theta_{j\,ОМП} = \frac{\sum_{i=1}^{n}\mathbb{I}(X_i=j)}{n} \quad \forall j = \overline{1,k}$$

Оценка состоят потому что θ – соотношение вероятности и определенного признака. Это похоже на схему испытаний Бернулли, но со множественными исходами.

$$\underbrace{\frac{\sum_{i=1}^{n} \mathbb{I}(x_i = j)}{n} \xrightarrow{P} \theta_1}_{\text{ЗБЧ}} \Rightarrow \text{состоятельна}$$

$\boxed{3}$ $y = X\theta + \varepsilon$

1) $L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i - x_i\theta)^2}{2}}$

$$l(\theta) = \sum_{i=1}^{n} \left(\ln \frac{1}{\sqrt{2\pi}} - \frac{(y_i - x_i^T\theta)^2}{2} \right)$$

$$MSE = \| y - X\hat{\theta} \|^2$$

т.е. максимизация ф-ии правдоподобия эквив-на минимизации MSE, т.к. $l(\theta) \sim const - MSE$

2) $l(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^{n} (y_i - x_i^T\theta)^2$

т.к. максимизация экв-на MSE

$$\hat{\theta}_{MSE} = \arg\min_{\theta} \sum_{i=1}^{n} (y_i - x_i^T\theta)^2$$

$$\frac{\partial \| y - X\theta \|^2}{\partial \theta} = \left((y - X\theta)^T (y - X\theta) \right)'_\theta =$$

$$= 2X^T X \theta - 2X^T y = 0$$

$$\Rightarrow \hat{\theta}_{MSE} = (X^T X)^{-1} X^T y$$

и $\Rightarrow$ $$\hat{\sigma}_{ОМП} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i^T \hat{\theta}_{MSE})^2$$

3) Ожидаем $y_{new} = x_{new}^T \theta$

$\xRightarrow{\text{св-во MSE}}$ $$\hat{\theta}_{MSE} \sim \mathcal{N}(\theta, (X^T X)^{-1} \sigma^2)$$

$$x_{new}^T \hat{\theta}_{MSE} \sim \mathcal{N}(x_{new}^T \theta, x_{new}^T (X^T X)^{-1} x_{new} \sigma^2)$$

т.е. $$y_{new} \in \left(y_{new} \pm z_{\frac{1+\alpha}{2}} \sqrt{x_{new}^T (X^T X)^{-1} x_{new} \hat{\sigma}} \right)$$

где $$\hat{\sigma} = \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i^T \hat{\theta}_{MSE})^2$$ – сост. оценка σ