Экспоненциальное семейство распределений

$$f(x|\theta) = \frac{h(x)}{Z(\theta)} \exp\left(\boldsymbol{\theta}^T \mathbf{T}(x)\right)$$

Например, распределение Пуассона

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Неравенство Рао-Крамера

Дисперсия несмещенной оценки $\hat{\theta}$ параметра θ ограничена снизу:

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

где $I(\theta)$ – информация Фишера

$$I(\theta) = n\mathbb{E}\left[\left(\frac{\partial l(X;\theta)}{\partial\theta}\right)^2\right]$$

По теореме факторизации Неймана-Фишера статистика $T(X)$ является достаточной для оценки параметра θ тогда и только тогда, когда правдоподобие выборки можно представить как

$$p(x|\theta) = h(x)g(\theta, T(X))$$

На примере Пуассоновского распределения

$$p(x_1, ..., x_n|\lambda) = \frac{e^{-n\lambda}\lambda^{x_1+\cdots+x_n}}{x_1!\ldots x_n!} = \frac{e^{-n\lambda+(x_1+\cdots+x_n)\log\lambda}}{x_1!\ldots x_n!}$$

Распишем правдоподобие распределения

$$\log p(x_1, \ldots, x_n|\theta) = -n\log Z(\theta) + \sum_{i=1}^{n}\log h(x_i) + \sum_{i=1}^{n}\theta^T T(x_i)$$

$$\frac{\partial \log p(x_1, \ldots, x_n|\theta)}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\left(-\log Z(\theta) + \theta^T\left(\frac{1}{n}\sum_{i=1}^{n} T(x_i)\right)\right)$$
$$= -\frac{1}{Z(\theta)}\frac{\partial(Z(\theta))}{\partial \theta_i} + \frac{1}{n}\sum_{j=1}^{n} T_i(x_j)$$

Распишем нормировочный множитель

$$Z(\theta) = \int h(x)e^{\theta^T T(x)}dx$$
$$\frac{\partial Z(\theta)}{\theta_i} = \int T_i(x)h(x)e^{\theta^T T(x)}dx$$

Тогда

$$\frac{1}{Z(\theta)}\frac{\partial(Z(\theta))}{\partial \theta_i} = \int \frac{1}{Z(\theta)}T_i(x)h(x)e^{\theta^T T(x)}dx = \mathbf{E}(T_i(x))$$

То есть ноль производной функции правдоподобия достигается в точке

$$\mathbf{E}(T_i(x)) = \frac{1}{n}\sum_{i=1}^{n} T_i(x_i)$$

Задача 1

Дано $X \sim N(\mu, \sigma^2)$, $Y \sim \Gamma(\alpha, \beta)$. Вычислите матрицы Фишера для вектора параметров (μ, σ) и (α, β)

$$p_Y(x) = \frac{x^{\alpha-1}e^{-\beta x}\beta^\alpha}{\Gamma(\alpha)}$$

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\frac{\partial \log p_X(x)}{\partial \mu} = -\frac{\mu - x}{\sigma^2}$$

$$\frac{\partial^2 \log p_X(x)}{\partial \mu^2} = -\frac{1}{\sigma^2}$$

$$\frac{\partial^2 \log p_X(x)}{\partial \mu \partial \sigma} = \frac{2(\mu - x)}{\sigma^3}$$

$$\frac{\partial \log p_X(x)}{\partial \sigma} = -\frac{1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}$$

$$\frac{\partial^2 \log p_X(x)}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{3(x-\mu)^2}{\sigma^4}$$

Вычислив матожидания получим матрицу

$$I(\mu, \sigma) = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix}$$

Задача 2

Докажите, что $\dfrac{X}{n}$ является эффективной оценкой параметра p распределения $Bin(n, p)$

$$D\left(\frac{X}{n}\right) = \frac{1}{n^2} D(X) = \frac{p(1-p)}{n}$$

Вычислим информацию Фишера:

$$\begin{aligned} I(p) &= \mathbb{E}\left(\frac{\partial \log(C_n^X p^X (1-p)^{n-X})}{\partial p}\right)^2 = \mathbb{E}\left(\frac{X}{p} - \frac{n-X}{1-p}\right)^2 \\ &= \mathbb{E}\left(\frac{X-np}{p(1-p)}\right)^2 = \frac{n^2}{p^2(1-p)^2}\mathbb{E}\left(\frac{X}{n} - p\right) \\ &= \frac{n^2}{p^2(1-p)^2} D\left(\frac{X}{n}\right) = \frac{n}{p(1-p)} = \frac{1}{D\left(\frac{X}{n}\right)} \end{aligned}$$

Задача 3

Пусть $\mu \in \mathbb{R}, \sigma > 0$ неизвестны и $X_1, \dots X_n \sim N(\mu, \sigma^2)$. Пусть

$$T(X) = (T_1(X), T_2(X))' = (\overline{X}, S^2)' = \left(\frac{1}{n}\sum_{i=1}^n X_i, \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2\right)'$$

Покажите, что $T(X)$ является достаточной статистикой для оценки параметра (μ, σ^2)

$$\begin{aligned}
p(x|\mu,\sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right) \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2\right) \\
&= (2\pi)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2 - n\log(\sigma)\right)
\end{aligned}$$

Распишем выражение под экспонентой

$$\begin{aligned}
\sum_{i=1}^n (x_i - \mu)^2 &= \sum_{i=1}^n (x_i - \overline{x} + \overline{x} - \mu)^2 \\
&= \sum_{i=1}^n (x_i - \overline{x})^2 + \sum_{i=1}^n (\overline{x} - \mu)^2 + 2(\overline{x} - \mu)\sum_{i=1}^n (x_i - \overline{x}) \\
&= (n-1)s^2 + n(\overline{x} - \mu)^2 = (n-1)T_2(x) + n(T_1(x) - \mu)^2
\end{aligned}$$

В итоге получаем

$$p(x|\mu,\sigma) = (2\pi)^{-n/2} \exp\left(-\frac{n}{2\sigma^2}(T_1(x) - \mu)^2 - \frac{n-1}{2\sigma^2}T_2(x) - n\log(\sigma)\right)$$