Упражнение 9.1. *Для распределений, указанных в Таблице 9.2 (см. стр. 182), представить функцию правдоподобия $\mathcal{L}(x_1,\dots,x_N|\theta)$ в виде*

$$\mathcal{L}(x_1,\dots,x_N|\theta) = A(\theta)\sum_{n=1}^{N} B(x_n) + NC(\theta) + \sum_{n=1}^{N} E(x_n),$$

вычислить $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)}$ и убедиться, что

$$\tau(\theta_N^*) \stackrel{\text{def}}{=} T(x_1,\dots,x_N) = \frac{1}{N}\sum_{n=1}^{N} B(x_n), \quad T(x) = B(x) \quad \textit{при} \quad N=1, \tag{1}$$

где θ_N^* –точка максимума θ функции правдоподобия $\mathcal{L}(x_1,\dots x_N|\theta)$, т.е. $\partial_\theta\mathcal{L}(x_1,\dots,x_N|\theta_N^*) = 0$, удовлетворяющая условию несмещенности $\mathbb{E}_\theta\, T(x_1,\dots,x_N) = \tau(\theta)$.

Вычислить информацию Фишера $I(\theta) = -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x|\theta)$ и дисперсию $\mathbb{D}_\theta$ оценки θ_N^*: $\mathbb{D}_\theta = \mathbb{E}_\theta\, T(x_1,\dots,x_N) = \mathbb{E}_\theta\,\tau(\theta_N^*)$.

Пример 1. Пусть $x_n \in \mathcal{N}(\theta,\sigma)$ –независимые нормально распределенные случайные величины с неизвестным средним θ и известной дисперсией σ^2, $n\in\{1,N\}$. Таким образом, $p_{\theta,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\theta)^2}{2\sigma^2}}$. Согласно формулам (9.6)–(9.7) (см. стр. 179),

$$\mathcal{L}(x_1,\dots,x_N|\theta) = \sum_{n=1}^{N}\mathcal{L}(x_n|\theta) = -\sum_{n=1}^{N}\frac{(x_n-\theta)^2}{2\sigma^2}$$

$$= \frac{\theta}{\sigma^2}\sum_n x_n - N\,\frac{\theta^2}{2\sigma^2} - \sum_n\frac{x_n^2}{2\sigma^2} = A(\theta)\sum_n B(x_n) + NC(\theta) + \sum_n E(x_n).$$

Положим

$$A(\theta) = \frac{\theta}{\sigma^2}, \quad B(x) = x, \quad C(\theta) = -\frac{\theta^2}{2\sigma^2}, \quad E(x) = -\frac{x^2}{2\sigma^2}.$$

В этом случае имеем $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \theta$. Поэтому из (1) следует, что выборочная оценка θ_N^* равна (см. Следствие 9.3 на стр. 181)

$$\boxed{\theta_N^* = \frac{1}{N}\sum_n B(x_n) = \frac{1}{N}\sum_n x_n}$$

Нетрудно видеть, что θ_N^* – аргумент максимума функции правдоподобия $\mathcal{L}(x,\dots,x_N|\theta)$, т.е. в этой точке $\partial_\theta\mathcal{L}(x,\dots,x_N|\theta) = 0$.

Из равенства $-\partial_\theta^2\mathcal{L}(x_1|\theta) = \sigma^{-2}$ следует

$$I_1(\theta) \stackrel{\text{def}}{=} -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1|\theta) = \mathbb{E}_\theta\,\sigma^{-2} = \sigma^{-2}.$$

Кроме того, $\mathbb{E}_\theta\, T(x_1,\dots,x_N) = \mathbb{E}_\theta x = \theta$, поэтому оценка θ_N^* является несмещенной, т.е. $\mathbb{E}_\theta\,\theta_N^* = \theta$ и $b(\theta) = 0$. Далее, поскольку $\mathcal{L}(x_1,\dots,x_N|\theta) = \sum_n\mathcal{L}(x_n|\theta)$, то

$$I_N(\theta) = -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1,\dots,x_N|\theta) = N\,I(\theta) = (\mathbb{D}_\theta T(x_1,\dots,x_N))^{-1},$$

где $\mathbb{D}_\theta T(x_1,\dots,x_N) = \frac{1}{N^2}\,\mathbb{D}_\theta(\sum_n x_n) = \frac{1}{N^2}\,\sum_n\mathbb{D}_\theta(x_n) = \frac{\sigma^2}{N}$.

Пример 2. Пусть $x_n \in \mathcal{N}(\mu,\theta)$ –независимые нормально распределенные случайные величины с известным средним μ и неизвестной дисперсией θ. В этом случае $p_{\mu,\theta}(x) = \frac{1}{\sqrt{2\pi\theta^2}}\, e^{-\frac{(x-\mu)^2}{2\theta^2}}$. Положим $\mathcal{L}((x_1,\dots,x_N|\theta) = \sum_n \ln p_{\mu,\theta}(x_n) + \text{const}$,

$$\mathcal{L}(x_1,\dots,x_N|\theta) = -\sum_n\frac{(x_n-\mu)^2}{2\theta^2} - \ln\theta = A(\theta)\sum_n B(x_n) + NC(\theta),$$

где

$$A(\theta) = -\frac{1}{2\theta^2}, \quad B(x) = (x-\mu)^2, \quad C(\theta) = -\ln\theta, \quad E(x) = 0.$$

Тогда $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \theta^2$ и аргументом максимума функции правдоподобия является точка

$$\boxed{\theta_N^2 = \frac{1}{N}\sum_n B(x_n) = \frac{1}{N}\sum_n (x_n-\mu)^2}$$

При условии, что среднее значение μ известно, а выборочные значения x_n независимы, оценка дисперсии случайной величины не зависит от μ и является несмещенной: $\mathbb{E}_\theta\,(\theta_N^*)^2 = \theta^2$.

Нетрудно видеть, что $I_1(\theta) \stackrel{\rm def}{=} -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1|\theta) = 6\,\mathbb{E}_\theta\,\frac{(x_1-\mu)^2}{2\theta^4} - \theta^{-2} = 2\,\theta^{-2}$. Поскольку $\mathcal{L}(x_1,\ldots,x_N|\theta) = \sum_n \mathcal{L}(x_n|\theta)$, то

$$I_N(\theta) = -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1,\ldots,x_N|\theta) = N\,I(\theta) = 2N\,\theta^{-2},$$

а, с другой стороны,

$$\mathbb{D}_\theta T(x_1,\ldots,x_N) = \frac{1}{N^2}\left\{\mathbb{E}_\theta\,\big(\sum_n (x_n-\mu)^2\big)^2 - \big(\mathbb{E}_\theta\,\sum_n (x_n-\mu)^2\big)^2\right\} =$$

$$\frac{1}{N^2}\left\{N\mathbb{E}_\theta\,(x_1-\mu)^4 + N(N-1)\big(\mathbb{E}_\theta\,(x_1-\mu)^2\big)^2 - N^2\big(\mathbb{E}_\theta\,(x_1-\mu)^2\big)^2\right\} = \frac{1}{N}\,(3\theta^4-\theta^4) = \frac{2}{N}\,\theta^4.$$

Сравнивая полученные выше выражения для дисперсии и информации Фишера, мы видим, что симметрия между ними в этом примере отсутствует: $I_N(\theta) \neq (D_\theta T_N)^{-1}$.

Пример 3. Пусть $x_n \in \Gamma(\alpha,\theta)$ –независимые случайные величины, имеющие гамма-распределение с плотностью $p_{\alpha,\theta}(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\theta}}}{\theta^\alpha\,\Gamma(\alpha)}$, $x \in \mathbb{R}_+$. Тогда

$$\mathcal{L}(x|\theta) = \ln p_{\theta,\sigma}(x) + \text{const} = -\frac{x}{\theta} - \alpha\ln\theta + (\alpha-1)\ln x \stackrel{\rm def}{=} A(\theta)\,B(x) + C(\theta) + E(x).$$

Нетрудно видеть. что $\mathbb{E}_\theta\,x = \int_0^\infty x\,p_{\alpha,\theta}(x)\,dx = \alpha\,\theta$. Следовательно,

$$A(\theta) = -\theta^{-1}, \quad B(x) = T(x) = x, \quad C(\theta) = -\alpha\ln\theta, \quad E(x) = (\alpha-1)\ln x$$

и $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \alpha\theta$. Равенство $\mathbb{E}_\theta\,T(x) = \alpha\,\theta \stackrel{\rm def}{=} \tau(\theta)$ указывает на несмещенность оценки $\alpha\,\theta_N^* = T(x_1,\ldots,x_N) = \frac{1}{N}\sum_n x_n$. Итак, $b(\theta) = 0$ и аргумент максимума функции правдоподобия равен

$$\boxed{\theta_N^* = \frac{1}{\alpha N}\sum_n x_n}$$

Аналогичным образом находим $I(\theta) \stackrel{\rm def}{=} -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1|\theta) = \frac{2}{\theta^3}\,\mathbb{E}_\theta\,x - \frac{\alpha}{\theta^2} = \frac{\alpha}{\theta^2}$. Поэтому $I_N(\theta) = N\,\frac{\alpha}{\theta^2}$ и

$$\mathbb{D}_\theta T(x_1,\ldots,x_N) = \frac{1}{N^2}\,\mathbb{E}_\theta\left(\sum_n (x_n^2 - \alpha^2\,\theta^2)\right) = \frac{\theta^2}{N}\,\mathbb{E}_\theta\,(x_1^2-\alpha^2) = \frac{\theta^2}{N}\,(\alpha(\alpha+1)-\alpha^2) = \frac{\theta^2\,\alpha}{N}.$$

В отличие от примера 2, линейное преобразование коэффициентов $A(\theta)$, $B(\theta)$ позволяет установить симметрию между дисперсией и информацией Фишера. Для этого положим

$$A(\theta) = -\alpha\theta^{-1}, \quad B(x) = T(x) = \frac{x}{\alpha}, \quad C(\theta) = -\alpha\ln\theta, \quad E(x) = (\alpha-1)\ln x.$$

В этом случае $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \theta$ и нетрудно видеть, что $\mathbb{E}_\theta T(x) = \theta$. Аналогично,

$$\mathbb{E}_\theta T^2(x) = \frac{\alpha(\alpha+1)}{\alpha^2}\,\theta^2, \quad \mathbb{D}_\theta T(x) = \frac{(\alpha(\alpha+1)-\alpha^2)}{\alpha^2}\,\theta^2 = \frac{\theta^2}{\alpha} = I_1^{-1}(\theta).$$

Нетрудно видеть, что такое преобразование не меняет положение экстремума функции правдоподобия и, следовательно, полученная выше оценка θ_N^* не меняется.

Пример 4. Пусть $x_n \in \Pi(\theta)$ –независимые случайные величины, имеющие дискретное пуассоновское распределение $p_\theta(x) = \frac{\theta^x e^{-\theta}}{x!}$, $x \in \{0, 1, \dots\}$. Тогда

$$\mathcal{L}(x|\theta) = x\ln\theta - \theta - \ln x! + \mathrm{const} \stackrel{\text{def}}{=} A(\theta)\,B(x) + C(\theta) + E(x).$$

Положим

$$A(\theta) = \ln\theta, \quad B(x) = T(x) = x, \quad C(\theta) = -\theta, \quad E(x) = -\ln x!.$$

В этом случае $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \theta$ и нетрудно видеть, что $\mathbb{E}_\theta T(x) = \sum_{n=0}^{\infty} n\frac{\theta^n e^{-\theta}}{n!} = \theta$, $\mathbb{D}_\theta T(x) = \sum_{n=0}^{\infty} n^2\frac{\theta^n e^{-\theta}}{n!} - \theta^2 = \theta$, $I_1(\theta) = -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1|\theta) = \theta^{-2}\mathbb{E}_\theta\,x = \theta^{-1} = (\mathbb{D}_\theta T(x))^{-1}$. Следовательно, аргумент максимума функции правдоподобия равен

$$\boxed{\theta_N^* = T(x_1, \dots, x_N) = \frac{1}{N}\sum_{n=0}^{N} x_n}$$

и между дисперсией оценки параметра статистики $T_N \stackrel{\text{def}}{=} T(x_1, \dots, x_N) = \frac{1}{N}\sum_{n=0}^{N} x_n$ и информацией Фишера выполнено аналогичное соотношение $\mathbb{D}_\theta T_N = I_N^{-1}(\theta)$.

Пример 5. Пусть $x_n \in \mathrm{Bi}(M, \theta)$ –независимые случайные величины, имеющие дискретное биномиальное распределение $p_{M,\theta}(x) = C_M^x\theta^x(1-\theta)^{M-x}$, $x \in \{0, 1, \dots, M\}$. Тогда

$$\mathcal{L}(x|\theta) = x\ln\frac{\theta}{1-\theta} + M\ln(1-\theta) + \ln C_M^x + \mathrm{const} \stackrel{\text{def}}{=} A(\theta)\,B(x) + C(\theta) + E(x).$$

Для симметризации свойств пары дисперсия-информация положим

$$A(\theta) = M\ln\frac{\theta}{1-\theta}, \quad B(x) = T(x) = \frac{x}{M}, \quad C(\theta) = M\ln(1-\theta), \quad E(x) = \ln C_M^x.$$

В этом случае $A'(\theta) = M\frac{1}{\theta(1-\theta)}$, $C'(\theta) = M\frac{1}{1-\theta}$, $\tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \theta$ и нетрудно видеть, что $\mathbb{E}_\theta T(x) = \frac{1}{M}\sum_{m=0}^{M} m\,C_M^m\theta^m(1-\theta)^{M-m} = \frac{1}{M}\mathbb{E}_\theta\,m = \theta$ (см. Таблицу 1.1 на стр. 27),

$$\mathbb{D}_\theta T(x) = \frac{1}{M^2}\sum_{n=0}^{M} m^2\,C_M^m\theta^m(1-\theta)^{M-m} - \theta^2 = \frac{M(M-1)\theta^2 + M\theta}{M^2} - \theta^2 = \frac{\theta(1-\theta)}{M},$$

$$I_1(\theta) = -\mathbb{E}_\theta\,\partial_\theta^2\mathcal{L}(x_1|\theta) = \theta^{-2}\mathbb{E}_\theta\,m + (1-\theta)^{-2}\mathbb{E}_\theta\,(M-m) = M\theta^{-1} + M(1-\theta)^{-1} = \frac{M}{\theta(1-\theta)}.$$

Таким образом, $\mathbb{D}_\theta T(x) = I_1^{-1}(\theta)$ и $\theta_N^* = T(x_1, \dots, x_N)$ для статистики $T_N \stackrel{\text{def}}{=} T(x_1, \dots, x_N)$ аргумент максимума функции правдоподобия определяется из условия

$$\boxed{\theta_N^* = T(x_1, \dots, x_N) = \frac{1}{MN}\sum_{n=1}^{N} x_n}$$

и в этой точке выполнено равенство $\mathbb{D}_\theta T_N = I_N^{-1}(\theta)$.

Пример 6. Пусть $x_n \in P(1,\theta)$ –независимые случайные величины, имеющие распределение Парето $p(x|\theta) = \frac{\theta-1}{x^\theta}$, $x \in [1,\infty)$. Тогда

$$\mathcal{L}(x_1,\ldots,x_N|\theta) = -\theta \sum_n \ln x_n + N\ln(\theta-1) \stackrel{\text{def}}{=} A(\theta)\sum_n B(x_n) + NC(\theta) + \sum_n E(x_n).$$

Для симметризации свойств пары дисперсия-информация положим

$$A(\theta) = -\theta, \quad B(x) = T(x) = \ln x, \quad C(\theta) = \ln(\theta-1), \quad E(x) = 0, \quad \tau(\theta) = -\frac{C'(\theta)}{A'(\theta)} = \frac{1}{\theta-1}.$$

Аргумент максимума функции правдоподобия $\mathcal{L}(x_1,\ldots,x_N|\theta)$ определяется из соотношения

$$\theta_N^* : \quad \frac{N}{\theta_N^*-1} = \sum_{n=1}^N \ln x_n.$$

Таким образом,

$$\boxed{\frac{1}{\theta_N^*-1} = \frac{1}{N}\sum_{n=1}^N \ln x_n}$$

причем эта оценка является несмещенной. Действительно, вычисляя интеграл по частям, получим

$$\mathbb{E}_\theta\,\frac{1}{\theta_N^*-1} = \mathbb{E}_\theta\,\ln x = \int_1^\infty \ln x\,\frac{\theta-1}{x^\theta}\,dx = \int_1^\infty \ln x\,d\frac{1}{x^{\theta-1}} = \int_1^\infty \frac{1}{x^\theta}\,dx = \frac{1}{\theta-1}.$$

Нетрудно видеть, что для параметра $(\theta-1)^{-1}$ информация Фишера при $N=1$ равна

$$I_1(\theta) = -\mathbb{E}_\theta\,\partial^2_{(\theta-1)^{-1}}\mathcal{L}(x|\theta) = (\theta-1)^2,$$

а дисперсия оценки $\frac{1}{\theta_1-1} = \ln x_1$ принимает значение обратное информации Фишера:

$$\mathbb{D}_\theta = \int_1^\infty (\ln x)^2\,\frac{\theta-1}{x^\theta}\,dx - \frac{1}{(\theta-1)^2} = \frac{2}{(\theta-1)^2} - \frac{1}{(\theta-1)^2} = \frac{1}{(\theta-1)^2}.$$