3.2.3 Метод максимального правдоподобия

Метод предложен Р. Фишером.

Этот метод наиболее полно использует данные выборки.

Оценки, полученные с помощью этого метода, называются ***оценками максимального правдоподобия.***

Оценки максимального правдоподобия состоятельны, но могут быть смещенными.

Функция правдоподобия

Пусть X – СВ, которая в результате n испытаний приняла значения $x_1,\ x_2,\ \ldots,\ x_n$.

Предположим:

вид закона распределения X известен, но не известны параметры этого закона.

Обозначим:

$$\bar{\theta} = (\theta_1, \theta_2, \ldots, \theta_p) -$$

параметры, подлежащие оцениванию.

Функция правдоподобия

- ***Функция правдоподобия дискретной СВ.***

Пусть X – дискретная СВ.

Обозначим:

$$p(x_i, \bar{\theta}) = P(X = x_i), \quad i = 1, 2, \ldots, n.$$

Функция

$$L(x_1, x_2, \ldots, x_n, \bar{\theta}) = p(x_1, \bar{\theta}) \cdot p(x_2, \bar{\theta}) \cdot \ldots \cdot p(x_n, \bar{\theta})$$

называется ***функцией правдоподобия дискретной СВ X.***

Функция правдоподобия

- ***Функция правдоподобия непрерывной СВ.***

Пусть X – непрерывная СВ.
Обозначим:

$f(x, \bar{\theta})$ – плотность распределения X.

Функция

$$L(x_1, x_2, \ldots, x_n, \bar{\theta}) = f(x_1, \bar{\theta}) \cdot f(x_2, \bar{\theta}) \cdot \ldots \cdot f(x_n, \bar{\theta})$$

называется ***функцией правдоподобия непрерывной СВ X.***

Функция правдоподобия

Функция правдоподобия (и дискретной, и непрерывной СВ) является функцией данных выборки ➔ сама является СВ.

Общий принцип получения оценок параметров

В качестве оценок максимального правдоподобия принимать тот набор значений параметров $\bar{\theta}^*$, при котором функция правдоподобия достигает максимума.

Пояснение:

подбираются значения параметров, обеспечивающие максимальную вероятность появления в выборке значений $x_1, x_2, \ldots, x_n$.

Логарифмическая функция правдоподобия

Для нахождения максимума функции правдоподобия необходимо выполнить ее дифференцирование по всем параметрам.

При большом количестве множителей это неудобно.

Удобнее использовать логарифмическое дифференцирование.

Функция

$$\ln L(x_1, x_2, \ldots, x_n, \bar{\theta})$$

называется ***логарифмической функцией правдоподобия СВ X.***

Максимумы функций L и $\ln L$ совпадают

Алгоритм получения оценок максимального правдоподобия по данным выборки

1. По данным выборки составить логарифмическую функцию правдоподобия.
2. Найти критические точки функции – корни уравнения

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \quad i = 1, 2, \ldots, n. \qquad \textbf{(3.27)}$$

Уравнения правдоподобия

3. Проверить выполнение достаточных условий максимума.

Получение оценок максимального правдоподобия по данным выборки

<u>Пример.</u>

По данным выборки $x_1, x_2, \ldots, x_n$ найти оценку наибольшего правдоподобия параметра a закона Пуассона

$$P_k(X = x_i) = \frac{a^{x_i}}{x_i!} \cdot e^{-a}, \quad i = 1, 2, \ldots, n,$$

где x_i – число появлений некоторого события в i–й серии испытаний;

k – число испытаний в каждой серии.

Получение оценок максимального правдоподобия по данным выборки

<u>Пример</u> (продолжение).

Будем использовать описанный выше алгоритм, применяя его к дискретной СВ X, распределенной по закону Пуассона с одним оцениваемым параметром.

Функция правдоподобия имеет вид:

$$L(x_1, x_2, \ldots, x_n, a) = \frac{a^{x_1}}{x_1!} \cdot e^{-a} \cdot \frac{a^{x_2}}{x_2!} \cdot e^{-a} \cdot \ldots \cdot \frac{a^{x_n}}{x_1!} \cdot e^{-a} =$$

$$= \frac{a^{x_1 + x_2 + \ldots + x_n}}{x_1! \cdot x_2! \cdot \ldots \cdot x_n!} \cdot e^{-na}.$$

Получение оценок максимального правдоподобия по данным выборки

<u>Пример</u> (продолжение).

Логарифмическая функция правдоподобия:

$$\ln L(x_1, x_2, \ldots, x_n, a) = \ln \frac{a^{x_1 + x_2 + \ldots + x_n}}{x_1! \cdot x_2! \cdot \ldots \cdot x_n!} \cdot e^{-na} =$$

$$= (x_1 + x_2 + \ldots + x_n) \ln a - na - \ln(x_1! \cdot x_2! \cdot \ldots \cdot x_n!)\ .$$

$$\frac{d \ln L}{da} = \frac{x_1 + x_2 + \ldots + x_n}{a} - n - 0 = \frac{x_1 + x_2 + \ldots + x_n}{a} - n\ .$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Уравнение правдоподобия имеет вид

$$\frac{x_1 + x_2 + \ldots + x_n}{a} - n = 0,$$

откуда

$$a^* = \frac{x_1 + x_2 + \ldots + x_n}{n} = \bar{x}.$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Для проверки достаточного условия максимума в точке a^* найдем вторую производную:

$$\frac{d^2 \ln L}{da^2} = -\frac{x_1 + x_2 + \ldots + x_n}{a^2}.$$

Поскольку для закона Пуассона $x_i \geq 0$, очевидно, что

$$\frac{d^2 \ln L}{da^2} < 0.$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Следовательно, оценка

$$a^* = \frac{x_1 + x_2 + \ldots + x_n}{n} = \bar{x} \qquad (3.28)$$

является оценкой максимального правдоподобия.

Получение оценок максимального правдоподобия по данным выборки

Пример.

По данным выборки $x_1, x_2, \ldots, x_n$ найти оценки наибольшего правдоподобия параметров m и σ нормального закона распределения

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Будем использовать описанный выше алгоритм, применяя его к непрерывной СВ X, распределенной по нормальному закону с двумя оцениваемыми параметрами.

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Функция правдоподобия имеет вид:

$$L(x_1,\ x_2,\ \ldots,\ x_n,\ m,\ \sigma)\ =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_1-m)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_2-m)^2}{2\sigma^2}} \cdot\ \ldots\ \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x_n-m)^2}{2\sigma^2}} =$$

$$= \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} \cdot e^{-\frac{(x_1-m)^2+(x_2-m)^2+\ \ldots\ +(x_n-m)^2}{2\sigma^2}}.$$

Логарифмическая функция правдоподобия:

$$\ln L\,(x_1,\ x_2,\ \ldots,\ x_n,\ m,\ \sigma)\ =$$

$$= -\,n\ln\sigma\ -\ \frac{n}{2}\ln 2\,\pi\ -\ \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2\ .$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Частные производные по параметрам равны

$$\frac{\partial \ln L}{\partial m} = -\frac{1}{2\sigma^2}\cdot 2\cdot(-1)\sum_{i=1}^{n}(x_i-m) = \frac{\sum_{i=1}^{n} x_i - nm}{\sigma^2},$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-m)^2\,.$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Получим систему уравнений:

$$\begin{cases} \dfrac{1}{\sigma^2}\left(\displaystyle\sum_{i=1}^{n} x_i - nm\right) = 0, & \Rightarrow\quad m^* = \dfrac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \\ -\dfrac{n}{\sigma} + \dfrac{1}{\sigma^3}\displaystyle\sum_{i=1}^{n}(x_i-m)^2 = 0, & \Rightarrow\quad \left(\sigma^*\right)^2 = \dfrac{\sum_{i=1}^{n}\left(x_i - m^*\right)^2}{n} = \sigma_B^2 \end{cases}$$

$$\Downarrow$$

$$\sigma^* = \sigma_B$$

Получение оценок максимального правдоподобия по данным выборки

Пример (продолжение).

Для проверки достаточных условий максимума необходимо проверить выполнение

$$\frac{\partial^2 \ln L}{\partial m^2}\cdot\frac{\partial^2 \ln L}{\partial \sigma^2} - \left(\frac{\partial^2 \ln L}{\partial m\,\partial\sigma}\right)^2 > 0 \quad \text{при} \quad \frac{\partial^2 \ln L^2}{\partial m^2} < 0.$$

Нетрудно убедиться, что при

$$m^* = \bar{x}, \qquad \sigma^* = \sigma_B \qquad (3.29)$$

эти условия выполнены.

Следовательно, оценки (3.29) являются оценками максимального правдоподобия.

Получение оценок максимального правдоподобия по данным выборки

<u>Пример</u> (продолжение).

Частные производные по параметрам равны

$$\frac{\partial \ln L}{\partial m} = -\frac{1}{2\sigma^2}\cdot 2\cdot(-1)\sum_{i=1}^{n}(x_i - m) = \frac{\sum_{i=1}^{n} x_i - nm}{\sigma^2},$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - m)^2.$$

3.2.4 Метод моментов

Метод предложен К. Пирсоном.

Основан на утверждении:

начальные и центральные эмпирические моменты являются состоятельными оценками соответственно начальных и центральных теоретических моментов того же порядка.

Достоинство метода – сравнительная простота.

Теоретические моменты (напоминание)

Начальным моментом порядка k СВ X называется математическое ожидание СВ X^k:

$$\nu_k = M\left(X^k\right).$$

В частности,

$$\nu_1 = M(X), \quad \nu_2 = M\left(X^2\right).$$

Центральным моментом порядка k СВ X называется математическое ожидание СВ $(X - M(X))^k$:

$$\mu_k = M\left((X - M(X))^k\right).$$

В частности,

$$\mu_1 = M(X - M(X)) = 0,$$

$$\mu_2 = M\left((X - M(X))^2\right) = D(X).$$

Эмпирические моменты

Пусть X – СВ, которая в результате n испытаний приняла значения $x_1, x_2, \ldots, x_n$.

Начальным эмпирическим моментом порядка k называется среднее значение k–х степеней вариант:

$$M_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k. \qquad (3.30)$$

В частности,

$$M_1 = \bar{x}.$$

Эмпирические моменты

Центральным эмпирическим моментом порядка k называется величина

$$m_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^k. \qquad (3.31)$$

В частности,

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \sigma_B^2.$$

Получение оценок методом моментов по данным выборки

Метод моментов точечной оценки неизвестных параметров заданного распределения состоит в приравнивании теоретических моментов рассматриваемого распределения соответствующим эмпирическим моментам того же порядка.

Получение оценок методом моментов по данным выборки

Пример.

По данным выборки $x_1, x_2, \ldots, x_n$ найти методом моментов точечную оценку параметра λ показательного распределения

$$f(x) = \lambda \cdot e^{-\lambda \cdot x} \quad (x \geq 0).$$

Приравняем начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка:

$$\nu_1 = M_1.$$

Получение оценок методом моментов по данным выборки

Пример (продолжение).

Для показательного распределения

$$\nu_1 = M(X) = \frac{1}{\lambda},$$

поэтому получим:

$$\frac{1}{\lambda} = \bar{x},$$

откуда

$$\lambda^* = \frac{1}{\bar{x}}. \qquad (3.32)$$

Получение оценок методом моментов по данным выборки

Пример.

По данным выборки $x_1, x_2, \ldots, x_n$ найти методом моментов точечные оценки параметров m и σ нормального закона распределения

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Получение оценок методом моментов по данным выборки

Пример (продолжение).

Приравняем начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка и центральный теоретический момент второго порядка к центральному эмпирическому моменту второго порядка:

$$\begin{cases} \nu_1 = M_1, \\ \mu_2 = m_2. \end{cases}$$

Получение оценок методом моментов по данным выборки

Для нормального распределения

$$\nu_1 = M(X) = m, \quad \mu_2 = D(X) = \sigma^2,$$

поэтому получим:

$$\begin{cases} m = \bar{x}, \\ \sigma^2 = \sigma_B^2, \end{cases}$$

откуда

$$m^* = \bar{x}, \quad \sigma^* = \sigma_B. \qquad (3.33)$$

Сравните с оценками максимального правдоподобия (3.29)