

24. Элементы математической статистики-2

24.1. Предварительные сведения

Пусть x_1, \dots, x_n — выборочные значения, θ — оцениваемый параметр, $p(x; \theta)$ — плотность распределения генеральной совокупности или вероятность события $\{X = x\}$ для дискретной статистической модели.

Функция правдоподобия $L(x_1, \dots, x_n; \theta) = \prod_{k=1}^n p(x_k; \theta)$.

Логарифмическая функция правдоподобия $l(x_1, \dots, x_n; \theta) = \ln L(x_1, \dots, x_n; \theta) = \sum_{k=1}^n \ln p(x_k; \theta)$.

Статистическая модель называется *регулярной*, если интегралы $\int p(x; \theta) dx = 1$ и $\int \varphi(x) p(x; \theta) dx = g(\theta)$ можно дифференцировать по параметру θ .

Информация Фишера относительно семейства $p(x; \theta)$ равна

$$I(\theta) = \mathbb{E} \left(\frac{\partial \ln p(X; \theta)}{\partial \theta} \right)^2 = \int \left(\frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 p(x; \theta) dx.$$

Неравенство Рао-Крамера для регулярной модели имеет вид $\text{Var} \hat{\theta}_n \geq \frac{1}{nI(\theta)}$.

Если для регулярной модели неравенство Рао-Крамера превращается в *равенство*, то несмещенная оценка $\hat{\theta}_n$ является *эффективной*, т.е. обладает минимально возможной дисперсией.

24.2. Практическое занятие

1. Имеется случайная выборка объема n из генеральной совокупности с распределением

- а) $\text{Pois}(\lambda)$; б) $\mathcal{N}(a, 1)$; в) $\mathcal{N}(0, \sigma^2)$, оцениваемый параметр $\theta = \sigma^2$;
- г) $\text{Exp}(\lambda)$, оцениваемый параметр $\theta = \frac{1}{\lambda}$;
- д) $\text{Geom}(p)$, оцениваемый параметр $\theta = \frac{1-p}{p}$; напоминание: $\mathbb{P}(X = k) = pq^k, k = 0, 1, \dots$;
- е) $\text{Unif}[-N, N]$ (в методе моментов используйте абсолютный момент первого порядка, исследуйте, что будет если использовать моменты более высоких порядков),

где параметр распределения неизвестен. Оцените значение неизвестного параметра с помощью метода моментов и метода максимального правдоподобия. Являются ли полученные оценки несмещенными? Найдите информацию Фишера, проверьте неравенство Рао-Крамера и сделайте вывод об эффективности оценок.

24.3. Домашнее задание

2*. Плотность распределения $p(x; \theta)$ зависит от неизвестного параметра $\theta > 0$. При $0 \leq x \leq \theta$ график плотности имеет вид симметричной параболы, проходящей через точки $(0, 0)$ и $(\theta, 0)$ и с вершиной в точке с абсциссой $\theta/2$; иначе плотность равна нулю.

- 1) Запишите выражение для плотности $p(x; \theta)$ в аналитическом виде.

2) Найдите оценку $\hat{\theta}_n$ с помощью метода моментов и проверьте её несмещённость.

3) Запишите уравнение правдоподобия для $\theta > \max(x_1, \dots, x_n)$ и докажите, что оно имеет решение.

3*. Исследуйте возможность применения метода моментов и метода максимального правдоподобия для оценки параметра θ распределения с плотностью

$$p(x; \theta) = \frac{1}{\pi(1 + (x - \theta)^2)} \quad (\text{распределение Коши}).$$

С помощью компьютера сгенерируйте несколько независимых реализаций случайной величины с указанной плотностью и постройте графики функции правдоподобия и логарифмической функции правдоподобия. Какую статистику Вы бы предложили для оценки θ ?