Оценки максимального правдоподобия

Пусть имеется выборка из генеральной совокупности с функцией распределения $f(x, \theta)$, θ - скалярный или векторный параметр

Оценка T параметра θ называется оценкой максимального правдоподобия, если она доставляет максимум функции правдоподобия

$L(X, T(X)) = \max\limits_{\theta \in \Theta} L(X, \theta)$

Пусть $f(x, \theta)$ - плотность $F(x, \theta)$, тогда $L(x, \theta) = \prod\limits_{i=1}^{n} f(x_i, \theta)$, где x - возможное значение вектора выборки

Если при любом возможном значении выборки функция L дифференцируема по параметру, то чтобы получить оценку максимального правдоподобия надо приравнять производные к нулю и решить уравнения $\frac{\partial L(x,\theta)}{\partial \theta_i} = 0,\ i = 1,\ \dots\ , n$ иногда вместо этого решают аналогичные уравнения $\frac{\partial \ln L(x,\theta)}{\partial \theta_i} = 0$

$\widehat{\theta}$ - обозначение оценки максимального правдоподобия

Свойства оценки максимального правдоподобия:

1) Если семейство допускает эффективную оценку для некоторой функции $\tau(\theta)$:

Пусть $\tau(\theta)$ - функция от параметра, $\widehat{\tau} := \tau(\widehat{\theta})$ - оценка функции от параметра

$T^*(X) - \tau(\theta) = a(\theta) U(X, \theta)$

$T^*(X) - \tau(\theta) = a(\theta) \frac{\partial \ln L(X,\theta)}{\partial \theta}$, если θ скаляр

$T^*(X) - \widehat{\tau} = 0 \Rightarrow T^* = \widehat{\tau}$

2) Если семейство допускает достаточную статистику, то оценка максимального правдоподобия является функцией от достаточной статистики:

$L(X, \theta) = g(T(X), \theta) h(X)$ - надо найти экстремум $g(t, \theta)$

Для каждого фиксированного значения t $g(t, \theta)$ достигает максимума при некотором $\theta_{\max}$. Для каждого t это значение своё, $\theta_{\max} = \theta_{\max}(t)$

$\widehat{\theta} = \theta_{\max}(T(X))$, где T - достаточная статистика

Примеры

1) $Bi(1, \theta),\ T^* = \overline{X}$ - эффективная оценка, также является оценкой максимального правдоподобия, проверим

$L(X, \theta) = \theta^{\sum\limits_{i=1}^{n} X_i} (1-\theta)^{n - \sum\limits_{i=1}^{n} X_i} = \theta^{nT^*} (1-\theta)^{n - nT^*}$

$\ln L(X, \theta) = nT \ln \theta + (n - nT) \ln(1 - \theta)$

Дифференцируем и подставляем $\theta = \widehat{\theta}$

$0 = \frac{nT}{\widehat{\theta}} - \frac{n - nT}{1 - \widehat{\theta}}$

$nT(1 - \widehat{\theta}) - (n - nT)\widehat{\theta} = 0$

$nT - nT\widehat{\theta} - n\widehat{\theta} + nT\widehat{\theta} = 0$

$nT = n\widehat{\theta}$

$\widehat{\theta} = T$

Нашли экстремум, но что если это минимум, а не максимум? Смотрим вторую производную

$\frac{T}{\widehat{\theta}^2} - (n - T)\frac{1}{(\widehat{\theta} - 1)^2} < 0$, значит нашли максимум

2) $\overline{Bi}(r, \theta)$, оценка максимального правдоподобия такая же, так как $\overline{X}$ - эффективная оценка

3) $N(\theta, \sigma^2)$ - аналогично

4) $N(\mu, \theta^2),\ \frac{1}{n}\sum\limits_{i=1}^{n} (X_i - \mu)^2$ - достаточная статистика

$L(X, \theta) = \frac{1}{(2\pi)^{n/2}\theta^n} \exp(-\frac{1}{2\theta^2} \sum\limits_{i=1}^{n} (X_i - \mu)^2)$

$\frac{\partial \ln L}{\partial \theta} = \frac{\partial}{\partial \theta}(-n \ln \theta - \frac{1}{2\theta^2} nT) = -\frac{n}{\theta} + \frac{1}{\theta^3} nT = 0$

$\widehat{\theta}^2 = T,\ \widehat{\theta} = \sqrt{T}$

5) Многомерный параметр $N(\theta_1, \theta_2^2)$

$f(x, \theta_1, \theta_2^2) = \frac{1}{\sqrt{2\pi}\theta_2}\exp(-\frac{1}{2\theta_2^2}(x-\theta_1)^2)$

$\ln L(X, \theta) = -\frac{n}{2}\ln 2\pi - n\ln\theta_2 - \frac{1}{2\theta_2^2}\sum\limits_{i=1}^{n}(X_i - \theta_1)^2$

$\frac{\partial \ln L}{\partial \theta_1} = \frac{1}{\theta_2^2}\sum\limits_{i=1}^{n}(X_i - \theta_1) = 0$

$\widehat{\theta}_1 = \overline{X}$

$\frac{\partial \ln L}{\partial \theta_2} = -\frac{n}{\theta_2} + \frac{1}{\theta_2^3}\sum\limits_{i=1}^{n}(X_i - \theta_1)^2 = -\frac{n}{\theta_2} + \frac{1}{\widehat{\theta}_2^3}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2 = 0$

$\widehat{\theta}_2^2 = S^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X}$

$\widehat{\theta}_2 = S$

$\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2) = (\overline{X}, S)$

Проверяем, что это максимум, а не минимум. Ищем 2-ой дифференциал

$\frac{\partial^2 \ln L}{\partial \theta_1^2} = -\frac{n}{\theta_2^2}$

$\frac{\partial^2 \ln L}{\partial \theta_2 \partial \theta_1} = -\frac{2}{\theta_2^3}\sum\limits_{i=1}^{n}(X_i - \theta_1)$ при $\theta_1 = \overline{X}$ равна нулю

$\frac{\partial^2 \ln L}{\partial \theta_2^2} = \frac{n}{\theta_2^2} - \frac{3}{\theta_2^4}\sum\limits_{i=1}^{n}(X_i - \theta_1)^2$

Подставляем значения оценки максимального правдоподобия и строим матрицу

$\begin{pmatrix} -\frac{n}{S^2} & 0 \\ 0 & \frac{n}{S^2} - \frac{3n}{S^2} = -\frac{2n}{S^2} \end{pmatrix}$ - матрица квадратичной формы 2-ого дифференциала, она отрицательно определённая

Что делать в случае нерегулярного семейства - примеры

1) $R((0, \theta)),\ \theta > 0$

$f(x, \theta) = \frac{e(\theta - x)}{\theta}$

$L(X, \theta) = \frac{1}{\theta^n}e(\theta - X_{(n)})$

Максимальное принимаемое значение при фиксированном θ - $\frac{1}{\theta^n}$ при условии $\theta \geq X_{(n)}$, ещё раз ищем максимум, но теперь θ произвольное

$\widehat{\theta}$ должно быть минимальным, но не нарушать условие $\theta \geq X_{(n)}$

$\widehat{\theta} = X_{(n)}$

2) $R((\theta_1, \theta_2)),\ \theta_1 < \theta_2$

$L(X, \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n}e(\theta_2 - X_{(n)})e(X_{(1)} - \theta_1)$

Максимальное принимаемое значение - $\frac{1}{(\theta_2 - \theta_1)^n}$, условия - $\theta_2 \geq X_{(n)},\ \theta_1 \leq X_{(1)}$

$\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2) = (X_{(1)}, X_{(n)})$

Принцип инвариантности для оценки максимального правдоподобия

Пусть θ - многомерный параметр, $\theta \in \Theta$

$\widetilde{q} : \Theta \to \mathbb{Q}$ - взаимоднозначное преобразование

$f(x, \theta)$ - исходная плотность

$f(x, \widetilde{q}^{-1}(q))$, ($q$ - параметр)

Тогда $\widehat{\theta} = \widetilde{q}^{-1}(\widehat{q}),\ \widehat{q} = \widetilde{q}(\widehat{\theta})$ - принцип инвариантности - при преобразовании оценка максимального правдоподобия переходит в оценку максимального правдоподобия

Пример - $N(\theta_1, \theta_2^2)$

$F(x) = \Phi(\frac{x - \theta_1}{\theta_2}) = \tau(\theta_1, \theta_2)$, x - фиксированный, нужно построить оценку максимального правдоподобия для $\tau(\theta_1, \theta_2)$

Пусть $q_1 = \Phi(\frac{x - \theta_1}{\theta_2}),\ q_2 = \theta_2$

$q_1 = \Phi(\frac{x - \theta_1}{q_1}),\ \frac{x - \theta_1}{q_2} = x_{q_1}$ - квантиль порядка q_1 для стандартного нормального закона (квантиль x_p - это решение уравнения $f(x) = p$)

$x - \theta_1 = x_{q_1} q_2$, θ_1 находится однозначно
$\widehat{q}_1 = \Phi(\frac{x-\widehat{\theta}_1}{\widehat{\theta}_2}) = \Phi(\frac{x-\overline{X}}{S})$, $\widehat{q}_2 = \widehat{\theta}_2$