Метод максимального правдоподобия

ММП позволяет получить по крайней мере асимптотически несмещенные и эффективные оценки параметров распределения, которые имеют нормальный закон распределения.

В основе ММП лежит понятие функции правдоподобия выборки.

Определение. Пусть имеем случайную величину Y, которая имеет функцию плотности вероятностей $P_y(t, a_1,a_2,\dots,a_k)$ и случайную выборку наблюдений за поведением этой величины $Y(y_1,y_2,\dots,y_n)$. Тогда функцией правдоподобия выборки $Y(y_1,y_2,\dots,y_n)$ называется функция L, зависящая от аргументов $a=\{a_1,a_2,\dots,a_k\}$, а от элементов выборки как от параметров и определяется равенством:

$$L(a_1,a_2,\dots,a_k, y_1,y_2,\dots,y_n)=P_y(y_1, a)\, P_y(y_2, a)\dots P_y(y_n, a)$$

Метод максимального правдоподобия

Функция правдоподобия:

$L(a_1,a_2,\ldots,a_k, y_1,y_2,\ldots,y_n)=P_y(y_1, a)\, P_y(y_2, a)\ldots P_y(y_n, a)$

Основные свойства функции правдоподобия.

1. Правая часть равенства имеет смысл значения закона распределения выборки при случайных значениях аргументов $t_1=y_1$, $t_2=y_2,\ldots$, $t_n=y_n$. Следовательно, функция правдоподобия L также случайная величина при любых значениях аргументов $a=\{a_1,a_2,\ldots,a_k\}$.

2. Все значения функции правдоподобия $L \geq 0$.

Эти свойства являются следствием свойств выборки.

Метод максимального правдоподобия

- **Идея метода.**

В качестве оценки неизвестного параметра принимается такое, которое обеспечивает максимум функции правдоподобия при всех возможных значениях случайной величины Y. Математически это выражается так:

$$\tilde{a}_j = \mathbf{argmax(L(a_1,a_2,\ldots,a_k,\ y_1,y_2,\ldots,y_n)}$$

Очевидно, что оценка $\tilde{a}_j$ зависит от случайной выборки, следовательно, $\tilde{a}_j = f(\mathbf{y_1,y_2,\ldots,y_n})$, где f есть процедура вычисления оценки $\tilde{a}_j$ по результатам выборки.

Метод максимального правдоподобия

Алгоритм решения задачи с помощью ММП.

Предполагается:

1. Вид закона распределения известен;
2. Функция плотности вероятности гладкая во всей области определения.

Последовательность решения:

1. Составляется функция правдоподобия.
2. Вычисляется логарифм функции правдоподобия.
3. Оценки параметров получаются в результате решения системы уравнений вида:
   $\partial \ln(L)/\partial a_i = 0$; $i=1,2,3,\ldots,k$
4. Проверяется условие максимума функции правдоподобия.

Метод максимального правдоподобия

Пример 1. Рассмотрим случайную величину, индикатор случайного события.

Закон распределения этой величины: $P_y(t,p)=p^t(1-p)^{(1-t)}$, где t=0, 1. p-параметр закона распределения. $M(Y)=p$, $\sigma^2(Y)=p(1-p)$.

Имеем выборку наблюдений Y=**{$y_1,y_2,\ldots,y_n,p$}.**

Решение.

1. Составляем функцию правдоподобия:

$$\mathbf{L(y_1,y_2,\ldots,y_n,p)=p^{y1}(1-p)^{(1-y1)}\,p^{y2}(1-p)^{(1-y2)}\ldots\,p^{yk}(1-p)^{(1-yk)}}$$

$$\mathbf{=p^{\Sigma yi}(1-p)^{\Sigma(1-yi)}}$$

2. Вычисляем логарифм функции правдоподобия:

$$\mathbf{\ln(L)=\Sigma y_i\ln(p) + \Sigma(1-y_i)\ln(1-p)}$$

Метод максимального правдоподобия

Пример 1. (продолжение)

3. Составляем уравнение для вычисления оценки параметра «p».

$$\frac{\partial \ln(L)}{\partial p} = \frac{\sum y_i}{p} - \frac{n - \sum y_i}{(1-p)} = 0$$

Откуда $$p = \frac{\sum y_i}{n}$$

4. Проверяем условие максимума функции L.

$$\frac{\partial^2 \ln(L)}{d\,p^2} = -\frac{\sum y_i}{p^2} - \frac{n - \sum y_i}{(1-p)^2} \geq 0$$

Метод максимального правдоподобия

Проверка свойств полученной оценки.

1. Несмещенность:

$M(\Sigma y_i/n)=(1/n)\Sigma M(y_i)=(1/n)(np) = p$

Математическое ожидание оценки равно его теоретическому значению.

Вывод: получена несмещенная оценка на выборке ограниченного объема!

Метод максимального правдоподобия

1. Неравенство Рао-Крамера.

Метод проверки условия эффективности.

Оно позволяет оценить нижнюю границу точности, с которой можно несмещенно оценить неизвестные параметры.

Нижняя граница соответствует минимальной дисперсии оценки. Следовательно, если дисперсия полученной оценки равна нижней границе, то эта оценка удовлетворяет условию эффективности.

Теорема. Для любой ковариационной матрицы любой несмещенной оценки вектора параметров «а» неравенство Рао-Крамера имеет вид:

$$\mathbf{Cov(\tilde{a},\tilde{a}) \geq I^{-1}}$$

где: I – квадратная матрица, информационная матрица Фишера:

$$I = M\left(-\frac{\partial^2 \ln(L)}{\partial a_i \partial a_j}\right)$$

Если число оцениваемых параметров равно 1, то матрица Фишера вырождается в число, которое называют информационным количеством Фишера.

Метод максимального правдоподобия

Пример 1. (Продолжение)

Вычислим информационное количество Фишера.

$$I = M\left(-\frac{d^2 \ln(L)}{d\,p^2}\right) = M\left(\frac{\sum y_i}{p^2}\right) + M\left(\frac{n - \sum y_i}{p(1-p)}\right) = \frac{n}{p(1-p)}$$

Тогда I^{-1} есть

$$I^{-1} = \frac{p(1-p)}{n} \qquad (6.2)$$

Величина (6.2) равна минимальному значению дисперсии среди всех возможных несмещенных процедур оценки параметра «p». Т.е., если дисперсия полученной оценки равна значению (6.2), то эта оценка эффективная.

Метод максимального правдоподобия

Вычисляем дисперсию оценки $\quad p = \frac{\sum y_i}{n}$

$$\sigma^2\left(\frac{\sum y_i}{n}\right) = \frac{1}{n^2}\sigma^2\left(\sum y_i\right) = \frac{1}{n^2}\sum \sigma^2(y_i) = \frac{n}{n^2}\sigma^2 y_i = \frac{p(1-p)}{n}$$

Полученное значение совпадает с выражением (6.2), следовательно мы получили несмещенную и эффективную оценку параметра «p» на выборке ограниченного объема.

Метод максимального правдоподобия

Пример 2. Получить ММП оценки случайной величины, имеющей нормальный закон распределения.

Имеем выборку Y={y_1,y_2,…,y_n}. Переменная Y имеет нормальный закон распределения:

$$p_y(t,a,\sigma^2) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} e^{-\frac{(t-a)^2}{2\sigma^2}}$$

Тогда функция правдоподобия выборки примет вид:

$$L(Y,a,\sigma^2) = \prod \frac{1}{(2\pi)^{\frac{1}{2}}(\sigma^2)^{\frac{1}{2}}} e^{-\frac{(y_i-a)^2}{2\sigma^2}} = \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{\sum(y_i-a)^2}{\sigma^2}}$$

Здесь символ «П» обозначает оператор произведения выражения по всем i

Метод максимального правдоподобия

Решение. Для удобства введем $s=\sigma^2$

2. Логарифм функции правдоподобия:

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(s) - \sum \frac{(y_i - a)^2}{2s}$$

3. Составляем систему уравнений относительно параметров «a» и «s»

Откуда получаем:

$$\frac{\partial \ln(L)}{\partial a} = -2\sum \frac{(y_i - a)}{2s} = 0$$

$$\frac{\partial \ln(L)}{\partial s} = -\frac{n}{2s} + \frac{\sum (y_i - a)^2}{2s^2} = 0$$

$$\tilde{a} = \frac{\sum y_i}{n} \qquad (6.3)$$

$$\tilde{s} = \frac{\sum (y_i - a)^2}{n} \qquad (6.4)$$

Метод максимального правдоподобия

Проверка свойств полученных оценок.

1. **Несмещенность.**

$$M(\tilde{a}) = M\left(\frac{\sum y_i}{n}\right) = \frac{1}{n}\sum M(y_i) = \frac{n}{n}a = a$$

$$M(\sigma^2) = M\left(\sum \frac{(y_i - a)^2}{n}\right) = \frac{1}{n}\sum M(y_i - a)^2 = \frac{n-1}{n}\sigma^2$$

Выводы: 1. Оценка параметра «а» является несмещенной на ограниченной выборке;

2. Оценка параметра σ^2 состоятельная, т.е. несмещенная при $n \Rightarrow \infty$.

Метод максимального правдоподобия

Проверка свойств оценок (Продолжение).

2. Эффективность.

Вычисляется информационная матрица Фишера.

1. $M\left(-\dfrac{\partial^2 \ln(L)}{\partial a^2}\right) = -M\left(\dfrac{\sum 1}{\sigma^2}\right) = \dfrac{n}{\sigma^2}$

2. $M\left(-\dfrac{\partial^2 \ln(L)}{\partial(\sigma^2)^2}\right) = -M\left(\dfrac{n}{2\sigma^4} - \dfrac{\sum(y_i - a)^2 2\sigma^2}{2(\sigma^2)^2}\right) = -\dfrac{n}{2\sigma^4} + \dfrac{M(\sum(y_i - a)^2)}{(\sigma^2)^3} =$

$= -\dfrac{n}{2(\sigma^2)^2} + \dfrac{n\sigma^2}{(\sigma^2)^3} = \dfrac{n}{2\sigma^4}$

3. $M\left(-\dfrac{\partial^2 \ln(L)}{\partial a \partial(\sigma^2)}\right) = M\left(\dfrac{-\sum(y_i - a)}{(\sigma^2)^2}\right) = \dfrac{\sum(M(y_i)) - na}{(\sigma^2)^2} = 0$

$$I = \begin{pmatrix} \dfrac{n}{\sigma^2} & 0 \\ 0 & \dfrac{n}{2\sigma^4} \end{pmatrix}$$

$$I^{-1} = \begin{pmatrix} \dfrac{\sigma^2}{n} & 0 \\ 0 & \dfrac{2\sigma^4}{n} \end{pmatrix}$$

Метод максимального правдоподобия

Эффективность оценок (Продолжение).

Вычислим дисперсии оценок (6.3) и (6.4)

$$\sigma^2\left(\frac{\sum y_i}{n}\right) = \frac{1}{n^2} n \sigma^2(y_i) = \frac{\sigma^2}{n}$$

$$\sigma^2\left(\frac{\sum (y_i - a)^2}{n}\right) = \frac{(n-1)}{n}\sigma^2$$

Сравнение полученных результатов с элементами обратной матрицы Фишера, показывает, что дисперсия оценки параметра «а» совпадает с минимально возможной дисперсией, а дисперсия оценки параметра «σ^2» является состоятельной.