

Элементы математической статистики

Задача математической статистики – состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов

Метод наибольшего правдоподобия

Кроме метода моментов существуют и другие методы точечной оценки неизвестных параметров распределения.

Например метод наибольшего правдоподобия

А. Дискретные случайные величины. Пусть X — дискретная случайная величина, которая в результате n испытаний приняла значения x_1, x_2, \dots, x_n . Допустим, что вид закона распределения величины X задан, но неизвестен параметр θ , которым определяется этот закон. Требуется найти его точечную оценку.

Обозначим вероятность того, что в результате испытания величина X примет значение $x_i (i=1, 2, \dots, n)$, через $p(x_i; \theta)$.

Функцией правдоподобия дискретной случайной величины X называют функцию аргумента θ :

$$L(x_1, x_2, \dots, x_n; \theta) = p(x_1; \theta) p(x_2; \theta) \dots p(x_n; \theta),$$

где x_1, x_2, \dots, x_n — фиксированные числа.

Метод наибольшего правдоподобия

В качестве точечной оценки параметра θ принимают такое его значение $\theta^* = \theta^*(x_1, x_2, \dots, x_n)$, при котором функция правдоподобия достигает максимума. Оценку θ^* называют *оценкой наибольшего правдоподобия*.

Функции L и $\ln L$ достигают максимума при одном и том же значении θ , поэтому вместо отыскания максимума функции L ищут (что удобнее) максимум функции $\ln L$.

Логарифмической функцией правдоподобия называют функцию $\ln L$. Как известно, точку максимума функции $\ln L$ аргумента θ можно искать, например, так:

1) найти производную $\frac{d \ln L}{d\theta}$;

2) приравнять производную нулю и найти критическую точку — корень полученного уравнения (его называют *уравнением правдоподобия*);

3) найти вторую производную $\frac{d^2 \ln L}{d\theta^2}$; если вторая производная при $\theta = \theta^*$ отрицательна, то θ^* — точка максимума.

Найденную точку максимума θ^* принимают в качестве оценки наибольшего правдоподобия параметра θ .

Метод наибольшего правдоподобия

Метод наибольшего правдоподобия имеет ряд достоинств: оценки наибольшего правдоподобия, вообще говоря, состоятельны (но они могут быть смещенными), распределены асимптотически нормально (при больших значениях n приближенно нормальны) и имеют наименьшую дисперсию по сравнению с другими асимптотически нормальными оценками; если для оцениваемого параметра θ существует эффективная оценка θ^* , то уравнение правдоподобия имеет единственное решение θ^* ; этот метод наиболее полно использует данные выборки об оцениваемом параметре, поэтому он особенно полезен в случае малых выборок.

Недостаток метода состоит в том, что он часто требует сложных вычислений.

З а м е ч а н и е 1. Функция правдоподобия—функция от аргумента θ ; оценка наибольшего правдоподобия—функция от независимых аргументов x_1, x_2, \dots, x_n .

З а м е ч а н и е 2. Оценка наибольшего правдоподобия не всегда совпадает с оценкой, найденной методом моментов.

Пример

Пример 1. Найти методом наибольшего правдоподобия оценку параметра λ распределения Пуассона

$$P_m(X = x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!},$$

где m — число произведенных испытаний, x_i — число появлений события в i -м ($i = 1, 2, \dots, n$) опыте (опыт состоит из m испытаний).

Решение. Составим функцию правдоподобия, учитывая, что $\theta = \lambda$:

$$\begin{aligned} L &= p(x_1; \lambda) p(x_2; \lambda) \dots p(x_n; \lambda) = \\ &= \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \dots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} = \frac{\lambda^{\sum x_i} \cdot e^{-n\lambda}}{x_1! x_2! \dots x_n!}. \end{aligned}$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = (\sum x_i) \ln \lambda - n\lambda - \ln(x_1! x_2! \dots x_n!).$$

Найдем первую производную по λ :

$$\frac{d \ln L}{d\lambda} = \frac{\sum x_i}{\lambda} - n.$$

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

$$(\sum x_i / \lambda) - n = 0.$$

Найдем критическую точку, для чего решим полученное уравнение относительно λ :

$$\lambda = \sum x_i / n = \bar{x}_B.$$

Найдем вторую производную по λ :

$$\frac{d^2 \ln L}{d\lambda^2} = - \frac{\sum x_i}{\lambda^2}.$$

Легко видеть, что при $\lambda = \bar{x}_B$ вторая производная отрицательна; следовательно, $\lambda = \bar{x}_B$ — точка максимума и, значит, в качестве оценки наибольшего правдоподобия параметра λ распределения Пуассона надо принять выборочную среднюю $\lambda^* = \bar{x}_B$.

Пример

Пример 2. Найти методом наибольшего правдоподобия оценку параметра p биномиального распределения

$$p_n(k) = C_n^k p^k (1-p)^{n-k},$$

если в n_1 независимых испытаниях событие A появилось $x_1 = m_1$ раз и в n_2 независимых испытаниях событие A появилось $x_2 = m_2$ раз.

Решение. Составим функцию правдоподобия, учитывая, что $\theta = p$:

$$L = P_{n_1}(m_1) P_{n_2}(m_2) = C_{n_1}^{m_1} C_{n_2}^{m_2} p^{m_1+m_2} (1-p)^{[(n_1+n_2)-(m_1+m_2)]}.$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = \ln(C_{n_1}^{m_1} C_{n_2}^{m_2}) + (m_1 + m_2) \ln p + [(n_1 + n_2) - (m_1 + m_2)] \ln(1-p).$$

Найдем первую производную по p :

$$\frac{d \ln L}{dp} = \frac{m_1 + m_2}{p} - \frac{(n_1 + n_2) - (m_1 + m_2)}{1-p}.$$

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

$$\frac{m_1 + m_2}{p} - \frac{(n_1 + n_2) - (m_1 + m_2)}{1-p} = 0.$$

Найдем критическую точку, для чего решим полученное уравнение относительно p :

$$p = (m_1 + m_2) / (n_1 + n_2).$$

Найдем вторую производную по p :

$$\frac{d^2 \ln L}{dp^2} = -\frac{m_1 + m_2}{p^2} + \frac{(n_1 + n_2) - (m_1 + m_2)}{(1-p)^2}.$$

Легко убедиться, что при $p = (m_1 + m_2) / (n_1 + n_2)$ вторая производная отрицательна; следовательно, $p = (m_1 + m_2) / (n_1 + n_2)$ — точка максимума и, значит, ее надо принять в качестве оценки наибольшего правдоподобия неизвестной вероятности p биномиального распределения:

$$p^* = (m_1 + m_2) / (n_1 + n_2).$$

Метод наибольшего правдоподобия

Б. Непрерывные случайные величины. Пусть X — непрерывная случайная величина, которая в результате n испытаний приняла значения x_1, x_2, \dots, x_n . Допустим, что вид плотности распределения $f(x)$ задан, но не известен параметр θ , которым определяется эта функция.

Функцией правдоподобия непрерывной случайной величины X называют функцию аргумента θ :

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta),$$

где x_1, x_2, \dots, x_n — фиксированные числа.

Оценку наибольшего правдоподобия неизвестного параметра распределения непрерывной случайной величины ищут так же, как в случае дискретной величины.

З а м е ч а н и е. Если плотность распределения $f(x)$ непрерывной случайной величины X определяется двумя неизвестными параметрами θ_1 и θ_2 , то функция правдоподобия является функцией двух независимых аргументов θ_1 и θ_2 :

$$L = f(x_1; \theta_1, \theta_2) f(x_2; \theta_1, \theta_2) \dots f(x_n; \theta_1, \theta_2),$$

где x_1, x_2, \dots, x_n — наблюдавшиеся значения X . Далее находят логарифмическую функцию правдоподобия и для отыскания ее максимума составляют и решают систему

$$\begin{cases} \frac{\partial \ln L}{\partial \theta_1} = 0, \\ \frac{\partial \ln L}{\partial \theta_2} = 0. \end{cases}$$

Пример

Пример 3. Найти методом наибольшего правдоподобия оценку параметра λ показательного распределения

$$f(x) = \lambda e^{-\lambda x} \quad (0 < x < \infty),$$

если в результате n испытаний случайная величина X , распределенная по показательному закону, приняла значения x_1, x_2, \dots, x_n .
Решение. Составим функцию правдоподобия, учитывая, что $\theta = \lambda$:

$$L = f(x_1; \lambda) f(x_2; \lambda) \dots f(x_n; \lambda) = (\lambda e^{-\lambda x_1}) (\lambda e^{-\lambda x_2}) \dots (\lambda e^{-\lambda x_n}).$$

Отсюда

$$L = \lambda^n e^{-\lambda \sum x_i}.$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = n \ln \lambda - \lambda \sum x_i.$$

Найдем первую производную по λ :

$$\frac{d \ln L}{d \lambda} = \frac{n}{\lambda} - \sum x_i.$$

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

$$(n/\lambda) - \sum x_i = 0.$$

Найдем критическую точку, для чего решим полученное уравнение относительно λ :

$$\lambda = n / \sum x_i = 1 / (\sum x_i / n) = 1 / \bar{x}_B.$$

Найдем вторую производную по λ :

$$\frac{d^2 \ln L}{d \lambda^2} = -\frac{n}{\lambda^2}.$$

Легко видеть, что при $\lambda = 1/\bar{x}_B$ вторая производная отрицательна; следовательно, $\lambda = 1/\bar{x}_B$ — точка максимума и, значит, в качестве оценки наибольшего правдоподобия параметра λ показательного распределения надо принять величину, обратную выборочной средней: $\lambda^* = 1/\bar{x}_B$.

Пример

Пример 4. Найти методом наибольшего правдоподобия оценки параметров a и σ нормального распределения

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2},$$

если в результате n испытаний величина X приняла значения x_1, x_2, \dots, x_n .

Решение. Составим функцию правдоподобия, учитывая, что $\theta_1 = a$ и $\theta_2 = \sigma$:

$$L = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_1-a)^2/2\sigma^2} \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_2-a)^2/2\sigma^2} \dots \times \\ \dots \times \frac{1}{\sigma \sqrt{2\pi}} e^{-(x_n-a)^2/2\sigma^2}$$

Отсюда

$$L = \frac{1}{\sigma^n (\sqrt{2\pi})^n} e^{-(\sum (x_i - a)^2)/2\sigma^2}$$

Найдем логарифмическую функцию правдоподобия:

$$\ln L = -n \ln \sigma + \ln \frac{1}{(\sqrt{2\pi})^n} - \frac{\sum (x_i - a)^2}{2\sigma^2}.$$

Найдем частные производные по a и по σ :

$$\frac{\partial \ln L}{\partial a} = \frac{\sum x_i - na}{\sigma^2}; \quad \frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum (x_i - a)^2}{\sigma^3}.$$

Приравняв частные производные нулю и решив полученную систему двух уравнений относительно a и σ^2 , получим:

$$a = \sum x_i / n = \bar{x}_B; \quad \sigma^2 = (\sum (x_i - \bar{x}_B)^2) / n = D_B.$$

Итак, искомые оценки наибольшего правдоподобия: $a^* = \bar{x}_B$; $\sigma^* = \sqrt{D_B}$. Заметим, что первая оценка несмещенная, а вторая смещенная.

Условные варианты

Предположим, что варианты выборки расположены в возрастающем порядке, т.е. в виде вариационного ряда

Равноотстоящими называют варианты, которые образуют арифметическую прогрессию с разностью h .

Условными называют варианты, определяемые равенством

$$u_i = (x_i - C) / h$$

где C – ложный ноль (новое начало отсчета), h – шаг (новая единица масштаба).

Упрощенные методы расчета сводных характеристик выборки основаны на

З а м е ч а н и е 1. В качестве ложного нуля можно принять любую варианту. Максимальная простота вычислений достигается, если выбрать в качестве ложного нуля варианту, которая расположена примерно в середине вариационного ряда (часто такая варианта имеет наибольшую частоту).

З а м е ч а н и е 2. Варианте, которая принята в качестве ложного нуля, соответствует условная варианта, равная нулю.

пример

Пример. Найти условные варианты статистического распределения:

варианты . . .	23,6	28,6	33,6	38,6	43,6
частоты . . .	5	20	50	15	10

Решение. Выберем в качестве ложного нуля варианту 33,6 (эта варианта расположена в середине вариационного ряда).

Найдем шаг:

$$h = 28,6 - 23,6 = 5.$$

Найдем условную варианту:

$$u_1 = (x_1 - C)/h = (23,6 - 33,6)/5 = -2.$$

Аналогично получим: $u_2 = -1$, $u_3 = 0$, $u_4 = 1$, $u_5 = 2$. Мы видим, что условные варианты — небольшие целые числа. Разумеется, оперировать с ними проще, чем с первоначальными вариантами.

Обычные, начальные и центральные эмпирические моменты

Для вычислений сводных характеристик выборки удобно пользоваться эмпирическими моментами, определения которых аналогичны определениям соответствующих теоретических моментов.

В отличие от теоретических эмпирические моменты вычисляют по данным наблюдений.

Обычным эмпирическим моментом порядка k называют среднее значение k -тых степеней разности $x-C$:

$$M'_k = (\sum n_i (x_i - C)^k) / n,$$

где x_i — наблюдаемая варианта, n_i — частота варианты, $n = \sum n_i$ — объем выборки, C — произвольное постоянное число (ложный нуль).

Обычные, начальные и центральные эмпирические моменты

Начальным эмпирическим моментом порядка k называют обычный момент порядка k при $C = 0$

$$M_k = (\sum n_i x_i^k) / n.$$

В частности,

$$M_1 = (\sum n_i x_i) / n = \bar{x}_B,$$

т. е. начальный эмпирический момент первого порядка равен выборочной средней.

Центральным эмпирическим моментом порядка k называют обычный момент порядка k при $C = \bar{x}_B$

$$m_k = (\sum n_i (x_i - \bar{x}_B)^k) / n.$$

В частности,

$$m_2 = (\sum n_i (x_i - \bar{x}_B)^2) / n = D_B, \quad (**)$$

т. е. центральный эмпирический момент второго порядка равен выборочной дисперсии.

$$\left. \begin{aligned} m_2 &= M'_2 - (M'_1)^2, \\ m_3 &= M'_3 - 3M'_2 M'_1 + 2(M'_1)^3, \\ m_4 &= M'_4 - 4M'_3 M'_1 + 6M'_2 (M'_1)^2 - 3(M'_1)^4. \end{aligned} \right\} \begin{array}{l} (***) \\ \end{array}$$

Условные эмпирические моменты. Отыскание центральных моментов по условным

Вычисление центральных моментов требует довольно громоздких вычислений. Чтобы упростить расчеты, можно заменить первоначальные варианты условными.

Условным эмпирическим моментом порядка k называют начальный момент порядка k , вычисленной для условных вариантов

$$M_k^* = \frac{\sum n_i u_i^k}{n} = \frac{\sum n_i \left(\frac{x_i - C}{h} \right)^k}{n}.$$

В частности,

$$M_1^* = \frac{\sum n_i \left(\frac{x_i - C}{h} \right)}{n} = \frac{1}{h} \left[\frac{\sum n_i x_i}{n} - C \frac{\sum n_i}{n} \right] = \frac{1}{h} (\bar{x}_B - C).$$

Отсюда

$$\bar{x}_B = M_1^* h + C. \quad (*)$$

Для того чтобы найти выборочную среднюю, достаточно вычислить условный момент первого порядка, умножить его на h и результату прибавить ложный ноль C

Условные эмпирические моменты. Отыскание центральных моментов по условным

Выразим обычные моменты через условные:

$$M_k^* = \frac{1}{h^k} \frac{\sum n_i (x_i - C)^k}{n} = \frac{M_k^i}{h^k}.$$

Отсюда

$$M_k^i = M_k^* h^k.$$

Для того чтобы найти обычный момент порядка k , достаточно условный момент того же порядка умножить на h^k .

Зная обычные моменты можем найти центральные моменты:

$$\left. \begin{aligned} m_2 &= [M_2^* - (M_1^*)^2] h^2, & (**) \\ m_3 &= [M_3^* - 3M_2^* M_1^* + 2(M_1^*)^3] h^3, \\ m_4 &= [M_4^* - 4M_3^* M_1^* + 6M_2^* (M_1^*)^2 - 3(M_1^*)^4] h^4. \end{aligned} \right\} (***)$$

$$D_B = [M_2^* - (M_1^*)^2] h^2. \quad (****)$$

Задание на практическое занятие

490. Случайная величина X (число появлений события A в m независимых испытаниях) подчинена биномиальному закону распределения с неизвестным параметром p . Ниже приведено эмпирическое распределение числа появлений события A в 1000 испытаний (в первой строке указано число x_i появлений события в одном опыте из $m = 10$ испытаний, во второй строке приведена частота n_i — число опытов, в которых наблюдалось x_i появлений события A):

x_i	0	1	2	3	4	5	6	7
n_i	2	3	10	22	26	20	12	5

Найти методом наибольшего правдоподобия точечную оценку неизвестного параметра p биномиального распределения.

Задание на практическое занятие

491. Случайная величина X (число появлений события A в m независимых испытаниях) подчинена закону распределения Пуассона с неизвестным параметром λ :

$$P_m(X = x_i) = \lambda^{x_i} \cdot e^{-\lambda} / x_i!,$$

где m — число испытаний в одном опыте, x_i — число появлений события в i -м опыте ($i = 1, 2, \dots, n$).

Найти методом наибольшего правдоподобия по выборке x_1, x_2, \dots, x_n точечную оценку неизвестного параметра λ распределения Пуассона.

492. Случайная величина X (число поврежденных стеклянных изделий в одном контейнере) распределена по закону Пуассона с неизвестным параметром λ . Ниже приведено эмпирическое распределение числа поврежденных изделий в 500 контейнерах (в первой строке указано количество x_i поврежденных изделий в одном контейнере, во второй строке приведена частота n_i — число контейнеров, содержащих x_i поврежденных изделий):

x_i	0	1	2	3	4	5	6	7
n_i	199	169	87	31	9	3	1	1

Найти методом наибольшего правдоподобия точечную оценку неизвестного параметра λ распределения Пуассона.

Задание на практическое занятие

494. Случайная величина X (время безотказной работы элемента) имеет показательное распределение $f(x) = \lambda e^{-\lambda x}$ ($x \geq 0$). Ниже приведено эмпирическое распределение среднего времени работы 1000 элементов (в первой строке указано среднее время x_i безотказной работы одного элемента в часах; во второй строке указана частота n_i — количество элементов, проработавших в среднем x_i часов):

x_i	5	15	25	35	45	55	65
n_i	365	245	150	100	70	45	25

Найти методом наибольшего правдоподобия точечную оценку неизвестного параметра λ показательного распределения.

Задание на практическое занятие

469. Найти исправленную выборочную дисперсию по данному распределению выборки объема $n = 20$:

x_i	0,1	0,5	0,7	0,9
n_i	6	12	1	1

У к а з а н и е. Перейти к условным вариантам $u_i = 10x_i$.

470. Найти исправленную выборочную дисперсию по данному распределению выборки объема $n = 10$:

x_i	23,5	26,1	28,2	30,4
n_i	2	3	4	1

У к а з а н и е. Перейти к условным вариантам $u_i = 10x_i - 268$.