# 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 $x_1, x_2, \ldots, x_n$. Допустим, что вид закона распределения величины $X$ задан, но неизвестен параметр $\theta$, которым определяется этот закон. Требуется найти его точечную оценку.

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

*Функцией правдоподобия дискретной случайной величины* $X$ называют функцию аргумента $\theta$:

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

где $x_1, x_2, \ldots, x_n$ — фиксированные числа.

# Метод наибольшего правдоподобия

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

Функции $L$ и $\ln L$ достигают максимума при одном и том же значении $\theta$, 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 $\theta$ 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 $\theta^*$ — точка максимума.

Найденную точку максимума $\theta^*$ принимают в качестве оценки наибольшего правдоподобия параметра $\theta$.

# 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 $\theta$ существует эффективная оценка $\theta^*$, то уравнение правдоподобия имеет единственное решение $\theta^*$; 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.

Замечание 1. Функция правдоподобия — функция от аргумента $\theta$; оценка наибольшего правдоподсбия — функция от независимых аргументов $x_1, x_2, \ldots, x_n$.

Замечание 2. Оценка наибольшего правдоподобия не всегда совпадает с оценкой, найденной методом моментов.

# Пример

**Пример 1.** Найти методом наибольшего правдоподобия оценку параметра $\lambda$ 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 ($i = 1, 2, \ldots, n$) опыте (опыт состоит из $m$ испытаний).

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

$$L = p(x_1; \lambda) p(x_2; \lambda) \ldots p(x_n; \lambda) =$$

$$= \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \ldots \frac{\lambda^{x_n} \cdot e^{-\lambda}}{x_n!} = \frac{\lambda^{\sum x_i} \cdot e^{-n\lambda}}{x_1! x_2! \ldots x_n!}.$$

Найдем логарифмическую функцию правдоподобия:

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

Найдем первую производную по $\lambda$:

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

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

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

Найдем критическую точку, для чего решим полученное уравнение относительно $\lambda$:

$$\lambda = \sum x_i/n = \overline{x}_в.$$

Найдем вторую производную по $\lambda$:

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

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

# 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:

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

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 $x_1, x_2, \ldots, x_n$. Допустим, что вид плотности распределения $f(x)$ задан, но не известен параметр $\theta$, которым определяется эта функция.

*Функцией правдоподобия непрерывной случайной величины* $X$ называют функцию аргумента $\theta$:

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

где $x_1, x_2, \ldots, x_n$ — фиксированные числа.

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

Замечание. Если плотность распределения $f(x)$ непрерывной случайной величины $X$ определяется двумя неизвестными параметрами $\theta_1$ и $\theta_2$, то функция правдоподобия является функцией двух независимых аргументов $\theta_1$ и $\theta_2$:

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

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

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

# Пример

**Пример 3.** Найти методом наибольшего правдоподобия оценку параметра $\lambda$ показательного распределения

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

если в результате $n$ испытаний случайная величина $X$, распределенная по показательному закону, приняла значения $x_1, x_2, \ldots, x_n$.

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

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

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 $\lambda$:

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

Напишем уравнение правдоподобия, для чего приравняем первую производную нулю:

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

Найдем критическую точку, для чего решим полученное уравнение относительно $\lambda$:

$$\lambda=n/\sum x_i=1/\left(\sum x_i/n\right)=1/\overline{x}_в.$$

Найдем вторую производную по $\lambda$:

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

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

# Пример

**Пример 4.** Найти методом наибольшего правдоподобия оценки параметров $a$ и $\sigma$ 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 $x_1, x_2, \ldots, x_n$.

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

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

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 $\sigma$:

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

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

$$a=\sum x_i/n=\overline{x}_в;\quad \sigma^2=\left(\sum(x_i-\overline{x}_в)^2\right)/n=D_в.$$

Итак, искомые оценки наибольшего правдоподобия: $a^*=\overline{x}_в$; $\sigma^*=\sqrt{D_в}$. 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

**Замечание 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).

**Замечание 2.** Варианте, которая принята в качестве ложного нуля, соответствует условная варианта, равная нулю.

# пример

**Пример.** Найти условные варианты статнстического 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 первоначальнымн 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 ***$x_i$-C***:

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

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 $C=0$

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

В частности,

$$M_1 = \left(\sum n_i x_i\right)/n = \overline{x}_в,$$

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

*Центральным эмпирическим моментом порядка* $k$ называют обычный момент порядка $k$ при $C=\overline{x}_в$

$$m_k = \left(\sum n_i (x_i - \overline{x}_в)^k\right)/n.$$

В частности,

$$m_2 = \left(\sum n_i (x_i - \overline{x}_в)^2\right)/n = D_в, \qquad (*)$$

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

$$m_2 = M'_2 - (M'_1)^2, \qquad (**)$$

$$\left.\begin{aligned} m_3 &= M'_3 - 3M'_2M'_1 + 2(M'_1)^3, \\ m_4 &= M'_4 - 4M'_3M'_1 + 6M'_2(M'_1)^2 - 3(M'_1)^4. \end{aligned}\right\} \qquad (***)$$

# 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 вариант

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

В частности,

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

Отсюда

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

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:

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

Отсюда

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

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:

$$m_2 = [M_2^* - (M_1^*)^2] h^2, \qquad (**)$$

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

$$D_в = [M_2^* - (M_1^*)^2] h^2. \qquad (****)$$

# 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 $\lambda$:

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

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

Найти методом наибольшего правдоподобия по выборке $x_1, x_2, \ldots, x_n$ точечную оценку неизвестного параметра $\lambda$ распределения Пуассона.

**492.** Случайная величина $X$ (число поврежденных стеклянных изделий в одном контейнере) распределена по закону Пуассона с неизвестным параметром $\lambda$. 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 $\lambda$ распределения Пуассона.

## Задание на практическое занятие

**494.** Случайная величина $X$ (время безотказной работы элемента) имеет показательное распределение $f(x) = \lambda e^{-\lambda x}$ $(x \geqslant 0)$. 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 $\lambda$ показательного рас пределения.

# 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 |

Указание. 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 |

Указание. Перейти к условным вариантам $u_i = 10x_i - 268$.