

Лекція 2

Статистичні оцінки випадкової величини

Вимоги до статистичних оцінок

Для того щоб статистичні оцінки давали «хороше» наближення параметрів, що оцінюються, вони повинні задовольняти певним вимогам.

1. **Незміщеність.** Незміщеною називають статистичну оцінку θ^* , математичне сподівання якої дорівнює параметру, що оцінюється, при будь-якому об'ємі вибірки: $M(\theta^*) = \theta$. В протилежному випадку оцінка є зміщеною.
2. **Ефективність.** Ефективною називають статистичну оцінку, яка при заданому об'ємі вибірки має найменшу можливу дисперсію: $D(\theta^*) = D_{\min}$
3. **Спроможність** (для вибірок великого об'єму). Оцінка називається спроможною, якщо із зростанням об'єму вибірки дисперсія оцінки прямує до нуля: $D(\theta^*) \xrightarrow{N \rightarrow \infty} 0$

Питання того, що оцінка задовольняє чи не задовольняє необхідним вимогам є важливим в задачах, коли на основі оцінок вибірки необхідно дати оцінку генеральної сукупності.

Далі, щоб розрізняти оцінки вибірки та генеральної сукупності будемо використовувати відповідно індекси «В» та «Г».

Точкові статистичні оцінки

Статистична оцінка математичного сподівання.

Точковою оцінкою математичного сподівання величини X є середньо зважена величина:

$$m_x = \bar{x} = \frac{\sum_{i=1}^N n_i x_i}{N},$$

де n_i – частоти відповідних значень випадкової величини x_i ; N – об'єм сукупності.

Причому середнє значення є незміщеною оцінкою, тому $\bar{x}_B = \bar{x}_G$ (отримавши оцінку математичного сподівання вибірки можна оцінити математичне сподівання генеральної сукупності).

Якщо сукупність розбита на групи і обчислені середні значення кожної

групи x_j , J – кількість груп, то $\bar{x} = \frac{\sum_{j=1}^J \bar{x}_j}{J}$.

Статистична оцінка математичного сподівання визначає те значення випадкової величини, яке в результаті експериментальних вимірювань приймається як істине.

Приклад. За допомогою верстату або декількох верстатів виготовлена партія виробів. На процес виробництва впливають фактори зовнішнього середовища та характеристики заготовок. Якщо верстатів декілька, то вони можуть мати свої виробничі особливості. Необхідно визначити масу (або

довжину) виробу для маркування. Для цього з партії була організована вибірка і результати вимірювань елементів вибірки представлені варіаційним рядом:

x_i	2560	2600	2620	2650	2700
n_i	2	3	10	4	1

$$N = \sum_{i=1}^5 n_i = 20,$$

$$\bar{x}_G = \bar{x}_B = \frac{\sum_{i=1}^5 n_i x_i}{N} = \frac{2 \cdot 2500 + 3 \cdot 2600 + 10 \cdot 2620 + 4 \cdot 2650 + 2700}{20} = 2615$$

В теорії наукового експерименту особливу роль відіграє нормальний закон розподілу. Якщо наприклад необхідно визначити значення деякого параметру, його вимірювання виконують певну кількість разів. Результати вимірювань повинні мати нормальний закон розподілу, що означає, що є деяке значення біля якого скупчуються дані. Якщо закон розподілу нормальний то в якості «істинного» значення береться середнє.

Статистична оцінка дисперсії та середньоквадратичного відхилення

Для оцінки розсіювання значення навколо середнього застосовується дисперсія:

$$D = \frac{\sum_{i=1}^n n_i (x_i - \bar{x})^2}{N}.$$

Розсіювання характеризує похибку але якщо ознака X є розмірною величиною, то дисперсія має розмірність у квадраті. Для оцінки похибки в тих самих одиницях вимірювання застосовують середньоквадратичне відхилення:

$$\sigma = \sqrt{D}.$$

Обчислення дисперсії можна спростити, використовуючи наступну формулу:

$$D = \bar{x}^2 - [\bar{x}]^2,$$

де $[\bar{x}]^2$ – середнє значення у квадраті, $\bar{x}^2$ – середнє значення квадратів: $\sum_{i=1}^n n_i x_i^2 / N$

Зауваження. Дисперсія є зміщеною оцінкою, тому для оцінки дисперсії генеральної сукупності за вибірковою дисперсією необхідно робити поправку на кількість степенів свободи:

$$D_G = \frac{N}{N-1} D_B.$$

В якості оцінки генеральної дисперсії застосовують виправлену дисперсію:

$$s^2 = \frac{\sum_{i=1}^n n_i (x_i - \bar{x})^2}{N-1}.$$

Для випадку, коли сукупність розділена на групи виділяють:

- **групову** дисперсію – це дисперсія ознаки, що належить групі:

$$D_{j22} = \frac{\sum_{i=1}^n n_i (x_i - \bar{x}_j)^2}{N_j} \text{ де } j - \text{ номер групи;}$$

- **внутрішньо групову** дисперсію – середнє арифметичне дисперсій, зважене

$$\text{за об'ємами груп: } D_{\text{внгр}} = \frac{\sum_{j=1}^J N_j D_{j22}}{N}, ;$$

- **між групову** дисперсію – це дисперсія групових середніх відносно

$$\text{загального середнього: } D_{\text{міжгр}} = \frac{\sum_{j=1}^J N_j (\bar{x}_j - \bar{x})^2}{N} \$$$

- **загальну** дисперсію – дисперсія всієї сукупності відносно загального

$$\text{середнього: } D_{\text{заг}} = \frac{\sum_{i=1}^n n_i (x_i - \bar{x})^2}{N} \text{ причому } D_{\text{заг}} = D_{\text{внгр}} + D_{\text{міжгр}}$$

Приклад. Задана вибірка:

x_i	0	1	2	3	4
n_i	5	2	1	1	1

Знайти дисперсію вибірки та оцінити дисперсію та середньоквадратичне відхилення генеральної сукупності.

$$\bar{x}_n = \frac{0 \cdot 5 + 1 \cdot 2 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1}{10} = \frac{11}{10} = 1,1$$

Обчислимо середнє:

Знайдемо вибіркoву дисперсію:

$$D_n = \frac{5 \cdot (0 - 1,1)^2 + 2 \cdot (1 - 1,1)^2 + 1 \cdot (2 - 1,1)^2 + 1 \cdot (3 - 1,1)^2 + 1 \cdot (4 - 1,1)^2}{10} = \frac{18,9}{10} = 1,89$$

$$s^2 = \frac{10}{9} \cdot 1,89 = 2,1$$

Знайдемо оцінку генеральної дисперсії

Знайдемо середньоквадратичне відхилення: $\sigma = s = \sqrt{2,1} = 1,37$.

Приклад. Нехай сукупність розбита на дві групи:

1- група	x_i	2	4	5	2- група	x_i	3	8
	n_i	1	7	2		n_i	2	3

Знайти групoві та загальну дисперсії.

Знайдемо середні значення груп та загальне середнє:

$$\bar{x}_1 = \frac{2 \cdot 1 + 4 \cdot 7 + 5 \cdot 2}{10} = 4, \quad \bar{x}_2 = \frac{2 \cdot 3 + 3 \cdot 8}{5} = 6, \quad \bar{x} = \frac{\bar{x}_1 + \bar{x}_2}{2} = \frac{4 + 6}{2} = 5$$

Обчислимо дисперсію кожної групи:

$$D_1 = \frac{1 \cdot (2 - 4)^2 + 7 \cdot (4 - 4)^2 + 2 \cdot (5 - 4)^2}{10} = 0,6, \quad D_2 = \frac{2 \cdot (3 - 6)^2 + 3 \cdot (8 - 6)^2}{5} = 6$$

Знайдемо внутрішню групову дисперсію:

$$D_{\text{внгр}} = \frac{N_1 D_1 + N_2 D_2}{N} = \frac{10 * 0,6 + 5 * 6}{15} = \frac{12}{5} = 2,4$$

Обчислимо між групову дисперсію: $D_{\text{міжгр}} = \frac{10 * (4 - 5)^2 + 5 * (6 - 5)^2}{15} = 1.$

Загальна дисперсія: $D_{\text{зае}} = D_{\text{внгр}} + D_{\text{міжгр}} = 2,4 + 1 = 3,4.$

Інші точкові оцінки варіаційного ряду

Модю M_0 називають варіанту, яка має найбільшу частоту.

Особливості застосування моди:

1) якщо всі значення варіаційного ряду мають однакову частоту, то кажуть, що цей варіаційний ряд не має моди;

2) якщо дві сусідніх варіанти мають однакову домінуючу частоту, то мода обчислюється як середнє арифметичне цих варіант;

3) якщо дві несусідні варіанти мають однакову домінуючу частоту, то такий варіаційний ряд називається бімодальним;

4) якщо таких варіант більше двох, то ряд полімодальний.

Медіаною m_0 називають варіанту, яка ділить варіаційний ряд на дві частини, рівні за кількістю варіант. Якщо число варіант

непарне, тобто $n = 2k + 1$, то $m_0 = x_{k+1}$

парне, тобто $n = 2k$, то $m_0 = (x_k + x_{k+1}) / 2$.

Застосування властивості медіани: при проектуванні оптимального положення зупинок громадського транспорту; при проектуванні складських приміщень; при спорудженні бензозаправок і т. д.

Якщо середнє, мода і медіана майже співпадають то можна зробити припущення, що закон розподілу є нормальним.

Розмахом варіювання R називають різницю між найбільшою і найменшою варіантами: $R = x_{\text{max}} - x_{\text{min}}$

Середнім абсолютним відхиленням θ називають середнє арифметичне абсолютних відхилень: $\theta = \sum n_i |x_i - \bar{x}| / n$

Коефіцієнтом варіації V називають виражене у % відношення вибіркового середнього квадратичного відхилення до вибіркової середньої: $V = \sigma_s / \bar{x}_s * 100\%$.

Коефіцієнт варіації служить для порівняння величин розсіювання по відношенню до вибіркової середньої двох варіаційних рядів: той з лав має більшу розсіювання по відношенню до вибіркової середньої, у якого коефіцієнт варіації більше.

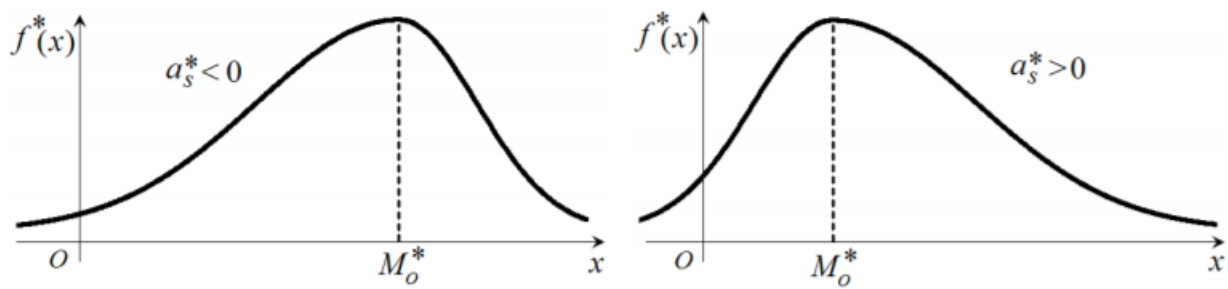
Показник асиметрії. Розраховується за формулою:

$$S = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3,$$

де s – стандартне відхилення за вибіркою;

n – розмір вибірки;

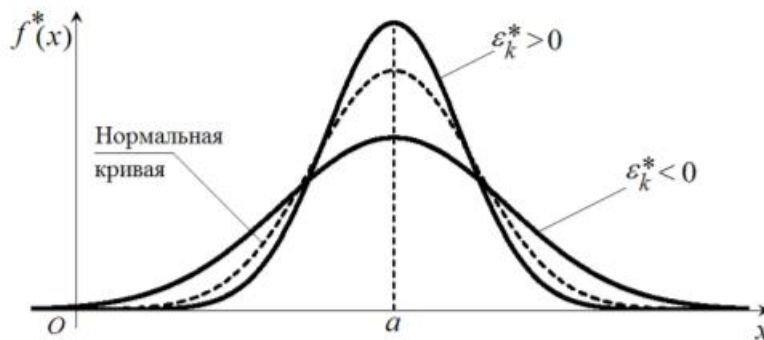
Якщо $S < 0$, то це означає, що переважають дані з великими значеннями, а якщо $S > 0$, то більше даних з меншими значеннями, ніж середнє арифметичне.



Право- та лівостороння асиметрія

Показник ексцесу. Ексцес характеризує відносне загострення або згладжування розподілу в порівнянні з нормальним розподілом. Показник ексцесу розраховується за формулою:

$$K = \left(\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 \right) - \frac{3(n-1)^2}{(n-2)(n-3)},$$



Ексцес

Квантиль. α -квантиль (квантиль порядку α) – це числова характеристика закону розподілу випадкової величини. α -квантиль - таке число, що дана випадкова величина потрапляє лівіше його з ймовірністю, що не перевищує α .

Найчастіше застосовують квартилі, які ділять ряд на 4 рівні частини, і децилі, які поділяють ряд на 10 рівних частин.

Розрізняють квартиль нижній (Q_1), що відокремлює 1/4 частина сукупності з найменшими значеннями ознаки, і квартиль верхній (Q_3), що відтинає 1/4 частина з найбільшими значеннями ознаки. Це означає, що 25% одиниць сукупності будуть менше за величиною Q_1 ; 25% одиниць будуть укладені між Q_1 і Q_2 ; 25% - між Q_2 і Q_3 і інші 25% перевершують Q_3 . Середнім квартилем Q_2 є медіана.

Для розрахунку квартилей інтервального варіаційного ряду використовують формули:

$$Q_1 = x_{Q_1} + i \frac{0,25 \sum f - S_{Q_1-1}}{f_{Q_1}}, \quad Q_3 = x_{Q_3} + i \frac{0,25 \sum f - S_{Q_3-1}}{f_{Q_3}}.$$

де x_{Q_1} – нижня межа інтервалу, що містить нижній квартиль (інтервал визначається по накопиченій частоті, першої перевищує 25%);

x_{Q_3} – нижня межа інтервалу, що містить верхній квартиль (інтервал визначається по накопиченій частоті, першої перевищує 75%);

S_{Q1-1} – накопичена частота інтервалу, що передує інтервалу, який містить нижній кuartиль;

S_{Q3-1} – теж для верхнього кuartиля;

f_{Q1} – частота інтервалу, що містить нижній кuartиль;

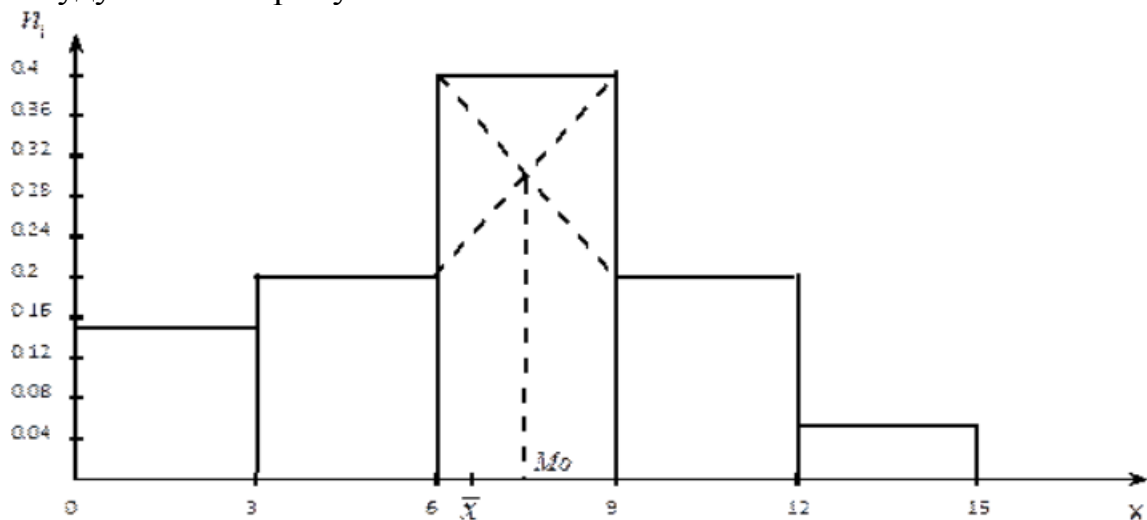
f_{Q3} – теж для верхнього кuartиля.

Приклад. Побудуйте графіки для варіаційного ряду. На графіку покажіть моду, медіану, середню, кuartилі.

Вік дитини x_i	0-3	3-6	6-9	9-12	12-15
Кількість дітей (частка) n_i	0,15	0,2	0,4	0,2	0,05

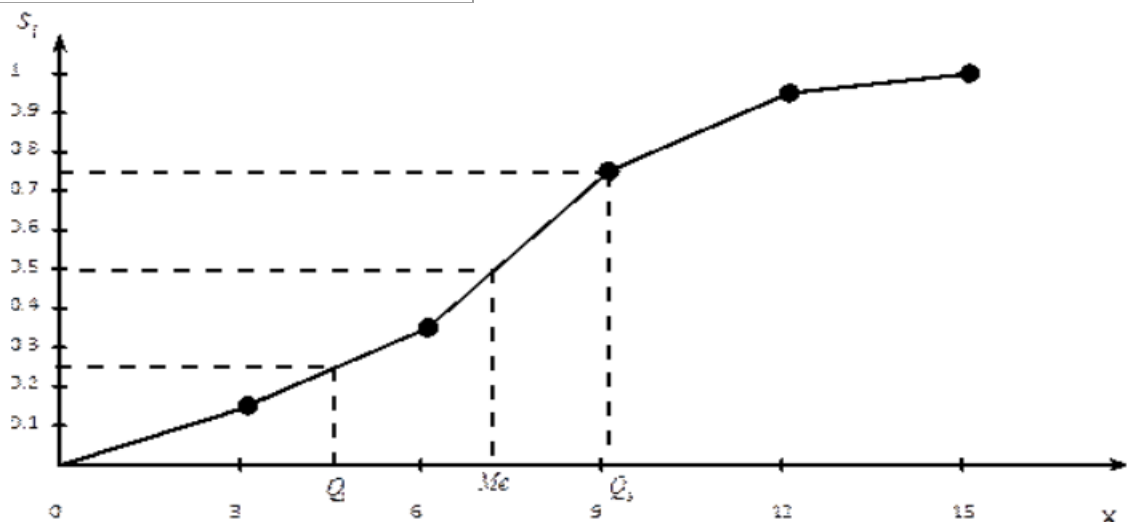
Обчислимо середнє:
$$\bar{x} = \frac{1.5 \cdot 0.15 + 4.5 \cdot 0.2 + 7.5 \cdot 0.4 + 10.5 \cdot 0.2 + 13.5 \cdot 0.05}{1} = 6.9 \text{ лет}$$

Побудуємо гістограму



Побудуємо кумуляту:

x_i	0-3	3-6	6-9	9-12	12-15
$n_{\text{накопич}}$	0.15	0.35	0.75	0.95	1



На отриманому графіку накопичених частот з останньої точки проведемо лінію перпендикулярну до осі X вона так само є максимальною висотою. Поділимо її на 4 частини. Через отримані точки будуємо паралельну осі лінії яка повинна перетинати висоту до осі X і кумуляту. Від місця перетину кумуляти опускаємо перпендикуляри. Утворені точки є квартилі і медіана (другий квартиль).

$$Q_1 \approx 4.5 \text{ лет}; Me \approx 7.3 \text{ лет}; Q_3 \approx 9.1 \text{ лет}$$

Інтервальні оцінки

Точність оцінки, надійність, довірчий інтервал

Точність оцінка характеризується позитивним числом, яке характеризує величину розбіжності між оцінками вибірки і генеральної сукупності:

$$|\theta - \theta^d| < \delta, \delta > 0$$

Надійністю (довірчою ймовірністю) оцінки 0 по θ^* називають ймовірність γ , з якої здійснюється нерівність $|\theta - \theta^d| < \delta$

$$P[|\theta^d - \delta < \theta < \theta^d + \delta| = \gamma$$

Як параметри надійності найбільш часто використовують величини, близькі до одиниці: 0,95; 0,99 і 0,999.

Довірчим називають інтервал $(\theta^d - \delta, \theta^d + \delta)$, який покриває невідомий параметр із заданою надійністю γ .

Інтегральна оцінка невідомого математичного очікування нормально розподіленої генеральної сукупності.

Існують два основні методи побудови довірчих інтервалів: байесовський метод і метод довірчих інтервалів, запропонований Нейманом. Застосовуючи метод побудови довірчих інтервалів, заснований на формулі Байеса, виходять із припущення, що оцінюваний параметр сам випадковий. Передбачається також, що відомо апіорно розподіл параметра. Цей метод часто непридатний, тому що оцінюється величина на практиці є просто невідомою постійною, а не випадковою величиною. Крім того, її розподіл буває також невідомим. Від цих недоліків вільний метод довірчих інтервалів. Розглянемо приклади побудови довірчих інтервалів в раді випадків.

1. Довірчий інтервал для математичного очікування при відомій дисперсії.

Нехай за вибіркою досить великого обсягу $n > 30$, і при заданій довірчій ймовірності необхідно визначити довірчий інтервал для математичного очікування, для оцінки якого використовується середнє арифметичне (середнє вибіркове).

Закон розподілу оцінки математичного очікування близький до нормального (розподіл суми незалежних випадкових величин з кінцевою дисперсією асимптотично нормальний). Якщо зажадати абсолютну надійність оцінки математичного очікування, то межі довірчого інтервалу будуть нескінченними $(-\infty, +\infty)$.

Будемо вимагати, щоб $P(|\bar{X} - a| < \delta) = \gamma$, де δ – точність. Користуючись обчисленням ймовірності заданого відхилення отримаємо:

$P(|\bar{X} - a| < \delta) = 2\Phi(\delta/\sigma) = 2\Phi(t)$, де $\Phi(t)$ – функція Лапласа $t = \delta\sqrt{n}/\sigma$. Формула обчислення довірчого інтервалу буде мати вигляд:

$$\bar{x} - t\sigma/\sqrt{n} < a < \bar{x} + t\sigma/\sqrt{n}$$

Для побудови довірчого інтервалу за таблицею функції Лапласа задавшись рівнянням $2\Phi(t) = \gamma$ знаходимо значення t . Далі знаходимо точність оцінки δ обчислюємо середнє $\bar{x}$ і обчислюємо границі інтервалу.

Зауваження. Для малих вибірок заміна закону розподілу нормальним приводить до грубих помилок.

2. Довірчий інтервал для математичного очікування при невідомій дисперсії: $\bar{x} - t_\gamma s/\sqrt{n} < a < \bar{x} + t_\gamma s/\sqrt{n}$, де t_γ – величина з розподілом Стюдента (знаходиться з таблиці розподілу Стюдента за заданими значеннями γ та n) s – виправлене середньо квадратичне відхилення.

Ці оцінки використовують для визначення найменшого об'єму вибірки (яку найменшу кількість вимірювань чи іспитів необхідно провести) щоб отримати значення із заданою точністю δ :

$$n = (t\sigma/\delta)^2 \quad (n = (t_\gamma s/\delta)^2)$$

Необхідно пам'ятати, що n натуральне число, тому результат округлюється до найближчого більшого цілого.

Довірчий інтервал для оцінки середньо квадратичного відхилення σ нормального розподілу.

Для отримання оцінки будемо вимагати виконання умови: $P(|\sigma - s| < \delta) = \gamma$.

Нехай $\delta/s = q$, тоді отримаємо:

$$s(1-q) < \sigma < s(1+q)$$

де q за заданими значеннями γ та n знаходиться з таблиці χ^2 розподілу.

Цю оцінку використовують в якості оцінки точності вимірювань.

Оцінка ймовірності (біноміальний розподіл) за відносною частотою

Для отримання оцінки будемо вимагати виконання умови: $P(|W - p| < \delta) = \gamma = 2\Phi(\delta/\sigma)$, де $W = m/n$ – відносна частота (точкова оцінка ймовірності). Довірчий інтервал ймовірності буде мати вигляд:

$$p_1 < p < p_2$$

$$p_1 = \frac{n}{t^2 + n} \left[W + \frac{t^2}{2n} - t \sqrt{\frac{W(1-W)}{n} + \left(\frac{t}{2n}\right)^2} \right],$$

$$p_2 = \frac{n}{t^2 + n} \left[W + \frac{t^2}{2n} + t \sqrt{\frac{W(1-W)}{n} + \left(\frac{t}{2n}\right)^2} \right]$$

де $\Phi(t) = \gamma/2$.

Зауваження. При великих об'ємах вибірки (сотні) можна скористатись спрощеною формулою:

$$p1 = W - t \sqrt{\frac{W(1-W)}{n}}, \quad p1 = W + t \sqrt{\frac{W(1-W)}{n}}$$

Поняття моментів

Для обчислення зведених показників вибірок використовують емпіричні моменти, аналогічні відповідним теоретичним моментів.

Вибіркове середнє і дисперсія варіаційного ряду є частковими випадками більш загального поняття – моментів варіаційного ряду.

Вибірковий початковий момент k -го порядку визначається формулою:

$$v_k^* = \sum_{i=1}^r x_i^k p_i^*$$

Зокрема: в вибіркове середнє є вибірковим початковим моментом першого порядку.

Вибірковий центральний момент k -ого порядку знаходиться за формулою:

$$\mu_k^* = \sum_{i=1}^r (x_i - \bar{x}_e)^k \cdot p_i^*$$

Зокрема:

$$\mu_1^* = 0;$$

$$\mu_2^* = D_e = v_2^* - (v_1^*)^2, \text{ т. е. виборочная дисперсия является } \\ \text{выборочным центральным моментом 2-го порядка;}$$

$$\mu_3^* = v_3^* - 3 v_1^* v_2^* + 2 (v_1^*)^3;$$

$$\mu_4^* = v_4^* - 4 v_1^* v_3^* + 6 (v_1^*)^2 \cdot v_2^* - 3 (v_1^*)^4.$$

Метод моментів

Можна вважати, що початкові і центральні емпіричні моменти є заможними оцінками відповідних теоретичних моментів. На цьому припущенні базується метод моментів, який заснований на прирівнювання теоретичних моментів розглянутого розподілу і відповідних емпіричних моментів того ж порядку. При цьому розрізняють випадки розподілів з одним параметром і з двома параметрами.

1. Оцінка одного параметра. Нехай задана щільність розподілу $f(x, \theta)$ з одним параметром. Згідно з методом моментів прирівнюємо, наприклад, відповідні початкові моменти першого порядку, тобто середню вибірки $\bar{x}$ і математичне очікування розподілу $M[X]$.

Тут досить одного рівняння щодо цього параметра $M(X) = \bar{x}$ (1)

Оскільки математичне очікування є функцією параметра θ

$$M(X) = \int_{-\infty}^{+\infty} x f(x, \theta) dx = \varphi(\theta) \quad (2)$$

співвідношення (1) можна розглядати як рівняння з одним невідомим, яке визначає точкову оцінку параметра θ , що є функцією.

Приклад. Методом моментів за вибіркою $x_1, x_2, \dots, x_n$ знайти точкову оцінку невідомого параметра λ показового розподілу з відомою функцією щільності розподілу $f(x) = \lambda e^{-\lambda x} \quad x \geq 0$.

Знайшовши інтеграл за формулою (2) отримаємо: $M(X) = 1/\lambda$. Отже оцінкою невідомого параметра буде $\lambda^* = 1/\bar{x}$.

2. Оцінка двох параметрів. Нехай задана функція щільності розподілу $f(x, \theta_1, \theta_2)$. Для знаходження невідомих параметрів потрібно мати два рівняння, тому прирівнюємо один одному відповідно початкові теоретичні та емпіричні моменти першого і другого порядків:

$$M(X) = \bar{x}, \quad D(X) = D_b.$$

Для ефективного використання методу згадаємо основні функції розподілу.

Закон розподілу	Щільність розподілу ймовірності	Математичне сподівання	Дисперсія	Позначення
Рівномірний	$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$	$M(x) = \frac{a+b}{2}$	$D(x) = \frac{(b-a)^2}{12}$	$R(a, b)$
Нормальний	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$	$M(x) = a$	$D(x) = \sigma^2$	$N(a, \sigma)$
Експоненціальний	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$M(x) = 1/\lambda$	$D(x) = 1/\lambda^2$	$E(\lambda)$
Пуассона	$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$	$M(x) = \lambda$	$D(x) = \lambda$	$P(\lambda)$
Біноміальний	$p_n(k) = C_n^k p^k q^{n-k}$	$M(x) = np$	$D(x) = npq$	$B(n, p)$

Метод найбільшої правдоподібності

1. Дискретні випадкові величини. Нехай X – дискретна випадкова величина, яка прийняла значення $x_1, x_2, \dots, x_n$ в результаті n випробувань. Нехай відомий закон розподілу випадкової величини X , але невідомий параметр θ , що його визначає. Потрібно знайти точкову оцінку цього параметра.

Позначимо ймовірність того, що в результаті випробування величина X прийме значення x_i , через $p(x_i, \theta)$.

Функція аргументу θ

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

де $x_1, x_2, \dots, x_n$ – фіксовані числа, називається **функцією правдоподібності** дискретної випадкової величини X .

В якості точкової оцінки параметра θ приймається значення θ^* , при якому функція досягає максимуму. Таку оцінку θ^* називають оцінкою найбільшої правдоподібності.

Функцію

$$\ln L(x_1, x_2, \dots, x_n, \theta) = \ln[p(x_1, \theta) \cdot p(x_2, \theta) \cdot \dots \cdot p(x_n, \theta)],$$

називають логарифмічною функцією правдоподібності.

Точка максимуму у обох функцій одна і та ж, але замість функції L , зручніше аналізувати функцію $\ln L$.

Приклад. Методом найбільшої правдоподібності знайти оцінку невідомого параметра p біноміального розподілу

$$P_n(m) = C_n^m p^m (1-p)^{n-m},$$

якщо в n_1 іспитаннях подія A наступила m_1 раз, а в n_2 випробуваннях – m_2 раз.

Розв'язок. Складемо функцію правдоподібності, де $p = \theta$,

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

Логарифмічна функція правдоподібності має вигляд:

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

Потім знаходимо похідну по p і, прирівнюючи її до нуля, отримуємо

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

$$p = \frac{m_1 + m_2}{n_1 + n_2}.$$

Розв'язок цього рівняння щодо p має вигляд:

Неважко переконатися, що друга похідна функції $\ln L < 0$, отже отримане значення p є точкою максимуму логарифмічною функції правдоподібності, а, значить, цю величину потрібно прийняти в якості оцінки найбільшої правдоподібності невідомого параметра p^* біноміального розподілу.

2. Неперервні випадкові величини. Нехай X – неперервна випадкова величина, яка прийняла значення $x_1, x_2, \dots, x_n$ в результаті n випробувань. Нехай відома функція розподілу щільності $f(x, \theta)$, але невідомий її параметр θ . Потрібно знайти точкову оцінку цього параметра.

Функція аргументу θ

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

де $x_1, x_2, \dots, x_n$ – фіксовані числа, називається **функцією правдоподібності** неперервної випадкової величини X .

Метод оцінки найбільшої правдоподібності невідомого параметра θ такий же, як і в попередньому випадку: пошук точки максимуму функції L .

Якщо у відомій функції розподілу $f(x, \theta_1, \theta_2)$ невідомі два параметра, то функція правдоподібності має вигляд

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

і тоді необхідно знайти точку максимуму функції двох змінних.

Відповідні логарифмічні функції найбільшої правдоподібності вводяться так само, як і в випадку дискретної випадкової величини.

Приклад. Методом найбільшої правдоподібності знайти оцінку невідомого параметра λ показового розподілу з прикладу.

Розв'язок. Складемо функцію правдоподібності. У нашому випадку

$$\theta = \lambda \Rightarrow L = (\lambda e^{-\lambda x_1}) \cdot (\lambda e^{-\lambda x_2}) \cdot \dots \cdot (\lambda e^{-\lambda x_n}) = \lambda^n e^{-\lambda(x_1 + x_2 + \dots + x_n)}.$$

Звідси знайдемо логарифмічну функцію правдоподібності

$$\ln L = n \ln \lambda - \lambda(x_1 + x_2 + \dots + x_n).$$

Похідну цієї функції по λ прирівнюємо до нуля. отримуємо

$$\lambda = \frac{n}{x_1 + x_2 + \dots + x_n} = \frac{1}{\bar{x}_g},$$