Лекція 2

Інтервальний статистичний розподіл вибірки. Методи знаходження оцінок. Метод підстановки і метод Пірсона. Метод максимальної правдоподібності і метод найменших квадратів.

2.1. Інтервальний статистичний розподіл вибірки та його характеристики

Якщо варіанти вибірки відрізняються одна від одної на як завгодно малу величину, або коли вибірка містить досить багато варіант, то часто застосовують групування варіант. Для цього всю ширину вибірки розбивають на інтервали завдовшки k і данні спостережень подають у вигляді таблиці частот, де вказують часткові інтервали $\left(x_i, x_{i+1}\right)$ і частоти n_i — кількість варіант, які потрапили в цей інтервал. Число інтервалів m слід брати не дуже великим, щоб після групування ряд не був громіздким, і не дуже малим, щоб не втратити особливості ознаки за якою досліджується ряд.

Рекомендована кількість інтервалів $m = 1 + 3.3221 \cdot \lg n$,

а ширина інтервалу:

$$k = \frac{x_{max} - x_{min}}{1 + 3,3221 \cdot lg\, n},$$

де $x_{max} - x_{min}$ – **розмах** вибірки, тобто, різниця між найбільшим і найменшим значенням варіанти.

Початком першого інтервалу рекомендується обрати величину

$$x_{поч} = x_{min} - k / 2.$$

Всі обчисленні зводять в таблицю і називають **інтервальним статистичним розподілом** вибірки.

При побудові полігону для інтервального розподілу, в якості значення

$i-$ої ознаки використовують середину інтервалу. Але частіше для графічного дослідження використовують гістограми.

Приклад 2.1. Тестується нова програма для розпізнавання прихованих зображень. Було протестовано 100 картинок із прихованим зображенням. Одержали наступні данні щодо часу виявлення (у секундах) для кожного зі 100 прихованих зображень:

$$\underbrace{104{,}2;\ 97{,}8;\ 97{,}0;\ 101{,}7;\ \ldots;\ 142{,}0;\ 141{,}0;\ 122{,}1}_{100\ \text{значень}}.$$

Зробити статистичне дослідження особливостей роботи пропонованої програми в контексті швидкості виявлення прихованого зображення.

Розв'язання.

1) Ранжуємо вибірку: $x_{min} = \underbrace{97{,}0;\ 97{,}2;\ \ldots;\ 141{,}0;\ 142{,}0}_{n\ =\ 100\ \text{значень}} = x_{max}$

2) Розіб'ємо варіанти на окремі інтервали, тобто проведемо їх групування.

Кількість інтервалів $m = 1 + 3{,}322\, lg\, 100 = 7{,}644 \approx 8$ інтервалів.

Ширина інтервалу:

$$k = (142{,}0 - 97{,}0)/(1 + 3{,}322 \lg 100) = 5{,}89(\text{сек}).$$

Візьмемо $k = 6{,}0(\text{сек})$. $x_{поч} = 97{,}0 - 6{,}0/2 = 94{,}0(\text{сек})$.

Згрупований ранжований ряд зведемо в таблицю (таблиця 2.1).

Одержаний варіаційний ряд дозволяє легко виявити закономірності роботи програми дешифрування. Накопичену частоту (частість) для кожного інтервалу знаходять за допомогою сумування частот (частостей) всіх попередніх інтервалів, включно з даним. Наприклад, для $x = 124$ накопичена частота $n_5^{нак}$ =3 + 7 + 11 + 20 + 28 = 69, тобто 69 зображень були дешифровані менше ніж за 124 хвилини.

3) Графічні характеристики інтервального розподілу вибірки:

На рисунках 2.1 і 2.2 зображені полігон , гістограма і кумулята для інтервального (табл. 2.1) варіаційного ряду.

Таблиця 2.1

i	Час роботи програми де-шифрування	Середини інтервалів, x_i	Частота (кількість дешифрованих зображень) n_i	Частість (частота дешифрованих зображень) $\omega_i = \frac{n_i}{n}$	Накопичена частота $n_i^{нак}$	Накопичена частість $\omega_i^{нак} = \frac{n_i^{нак}}{n}$
1	94,0-100,0	97	3	0,003	3	0,03
2	100,0-106,0	103	7	0,007	10	0,10
3	106,0-112,0	109	11	0,11	21	0,21
4	112,0-118,0	115	20	0,20	41	0,41
5	118,0-124,0	121	28	0,28	69	0,69
6	124,0-130,0	127	19	0,19	88	0,88
7	130,0-136,0	133	10	0,1	98	0,98
8	136,0-142,0	139	2	0,02	100	1,0
	$\sum$		100	1,00	-	-

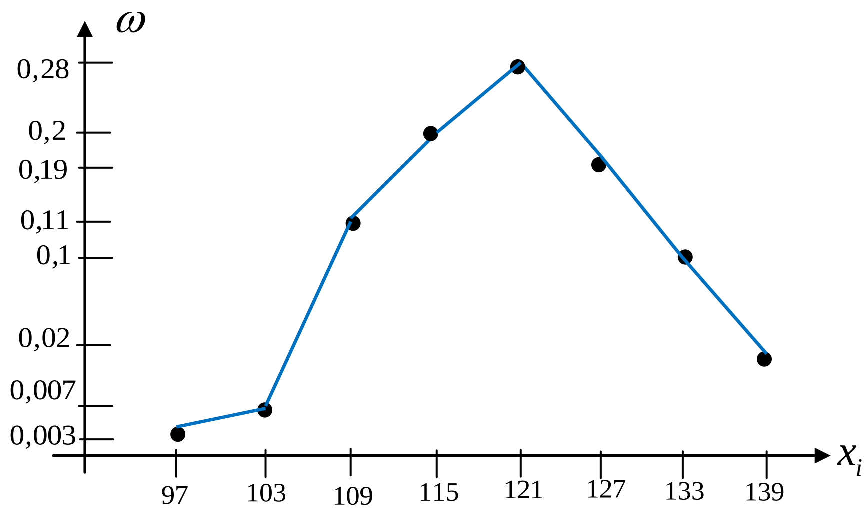

Рис. 2.1. Полігон для інтервального варіаційного ряду

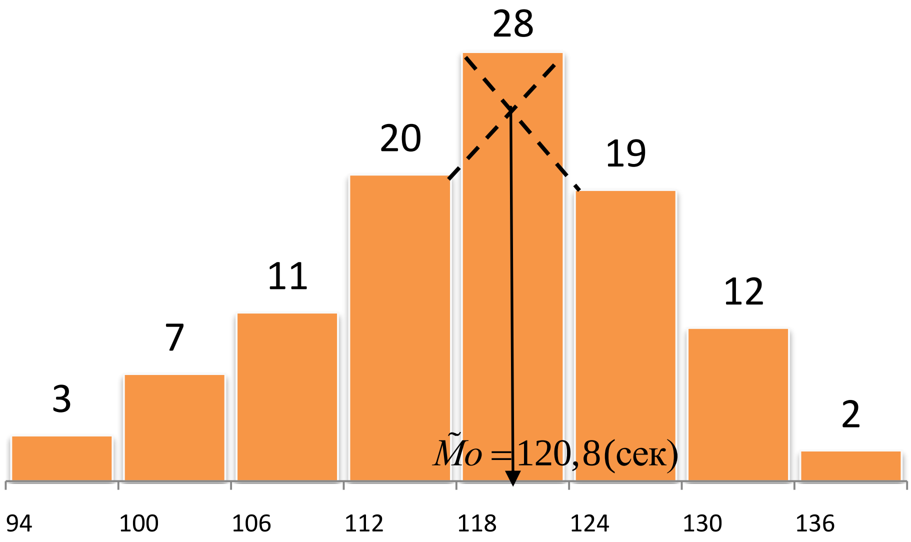

Рис. 2.2. Гістограма для інтервального варіаційного ряду

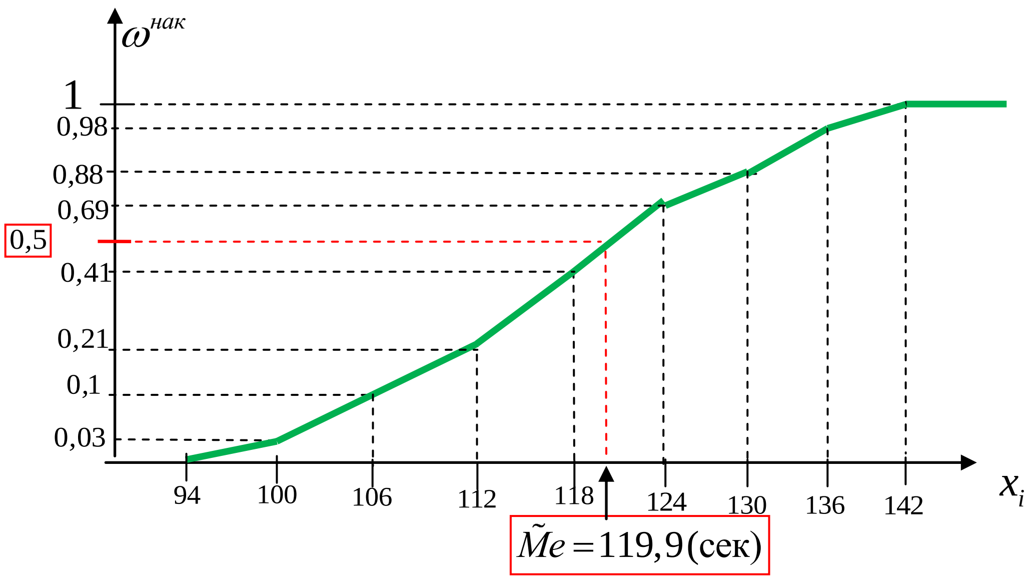

Рис. 2.3. Кумулята для інтервального варіаційного ряду

4) На практиці, в більшості випадків, достатньо знати тільки зведені характеристики варіаційних рядів: середні; характеристики мінливості (варіації) та ін. Розрахунок статистичних характеристик є наступним, після групування даних, етапом спостережень.

а) Знайдемо середній час роботи програми дешифрування зображення за даними табл. 2.1.

За формулою (1.5) для середньої арифметичної інтервального варіаційного ряду

$$\bar{x} = \frac{97 \cdot 3 + 103 \cdot 7 + \ldots + 133 \cdot 10 + 139 \cdot 2}{100} = 119,2\left(\text{сек}\right)$$

де числа 97, 103,…, 133, 139 – середини відповідних інтервалів.

б) Знайдемо медіану і моду розподілу часу роботи програми дешифрування зображення за даними табл. 2.1.

На рис. 2.3 проведемо горизонтальну пряму $y = 0,5$ (або $n = 50$), що відповідає накопиченій частості $\omega_x^{нак} = F^*(x) = 0,5$ (або накопиченій частоті $n_x^{нак} = 50$), до перетину з графіком кумуляти розподілу. Абсциса точки перетину і буде медіаною варіаційного ряду:

$$\tilde{M}e = 119,9 (\text{сек}).$$

На гістограмі розподілу (рис. 2.2) знаходимо прямокутник з найбільшою частістю. Сполучаючи відрізками прямих вершини цього прямокутника із відповідними вершинами двох сусідніх прямокутників (див. рис. 2.2), отримаємо точку перетину цих відрізків (діагоналей), абсциса якої і буде модою варіаційного ряду: $\tilde{M}o = 120,8 (\text{сек})$.

в) Обчислимо дисперсію, середнє квадратичне відхилення і коефіцієнт варіації розподілу часу роботи програми для дешифрування зображення за даними табл. 2.1.

Раніше було одержано значення $\overline{x} = 119,2 (\text{сек})$.

З означення дисперсії:

$$s^2 = \frac{(97-119,2)^2 \cdot 3 + (103-119,2)^2 \cdot 7 + ... + (139-119,2)^2 \cdot 2}{100} = 87,48.$$

Середнє квадратичне відхилення $s = \sqrt{87,48} = 9,35 (сек)$;

коефіцієнт варіації $\tilde{v} = (9,35 / 119,2)100 = 7,8$ (%).

г) Обчислимо коефіцієнт асиметрії та ексцес розподілу часу роботи програми для дешифрування зображення за даними таблиці 2.1.

Коефіцієнт асиметрії та ексцес варіаційного ряду, який наведено в таблиці 2.1, обчислюємо за формулами 1.10 і 1.11(Лекція 1):

$$\tilde{A} = \frac{(97-119,2)^3 \cdot 3 + (103-119,2)^3 \cdot 7 + ... + (139-119,2)^3 \cdot 2}{100 \cdot 9,35^3} = -0,302;$$

$$\tilde{E} = \frac{(97-119,2)^4 \cdot 3 + (103-119,2)^4 \cdot 7 + ... + (139-119,2)^4 \cdot 2}{100 \cdot 9,35^4} - 3 = -0,286$$

Оскільки коефіцієнт асиметрії від'ємний і близький до нуля, розподіл часу роботи програми для дешифрування зображення має незначну лівосторонню асиметрію, а оскільки ексцес близький до нуля, то розподіл за крутизною наближається до нормального.

Середнє арифметичне $\bar{x}$, дисперсія s^2 та інші характеристики варіаційного ряду є статистичними аналогами математичного сподівання $M(X)$, дисперсії σ^2 та відповідних характеристик випадкової величини *X*.►

2.2. Статистичні оцінки параметрів розподілу. Точкові оцінки

Види статистичних оцінок

Для дослідження кількісної ознаки генеральної сукупності потрібно знати наближені значення математичного сподівання $M(X)$, дисперсії $D(X)$, середнього квадратичного відхилення $\sigma(x)$, початкові та центральні моменти випадкової величини X. Іноді з деяких міркувань вдається встановити закон розподілу X. Тоді потрібно вміти оцінювати параметри цього закону розподілу. За ступенем охоплення одиниць досліджуваної сукупності статистичне спостереження може бути **суцільним** і **несуцільним** (або **вибірковим**). Суцільне спостереження передбачає обстеження всіх без винятку одиниць генеральної сукупності. Вибіркове спостереження дає висновок про всю сукупність одиниць при дослідженні тільки її частини.

Сукупність методів математичної статистики, які використовуються для обгрунтувань і висновків при проведенні вибіркового спостереження, називається **вибірковим методом**.

Розглянемо традиційні підходи до оцінок невідомих параметрів – точкове і інтервальне оцінювання.

Точкове оцінювання

Розглядає незміщені оцінки; за міру точності різних оцінок приймається величина їх дисперсій; основна увага приділяється методам побудови оптимальних оцінок.

Нехай із генеральної сукупності одержано вибірку $x_1, x_2, \ldots, x_n$ яка залежить від параметра θ , значення якого нам невідоме, і нехай $\overline{\theta} = f\left(x_1, x_2, \ldots, x_n\right)$, де $f : R^n \to R$ — деяка відома функція елементів цієї вибірки (деяка статистика). Числа $x_1, x_2, \ldots, x_n$ відомі, тому $\overline{\theta}$ для цієї вибірки є конкретним числом, яке будемо вважати певним наближенням до невідомого параметра θ.

$\overline{\theta}$, як функція елементів вибірки, є випадковою величиною. Інша вибірка такого самого обсягу і тій самій статистиці f, визначає інше число $\overline{\theta}$. Важливим є те, наскільки великою є похибка $\overline{\theta} - \theta$. Ця величина називається **зміщенням статистичної оцінки** $\overline{\theta}$.

Наближене значення шуканого параметра генеральної сукупності, яке знайдене за даними вибіркової сукупності, називається **вибірковою оцінкою параметра**.

Якщо числа $\overline{\theta}_k$, $\left(k = 1,2,\ldots,m\right)$, будуть більші значення θ, тоді оцінка $\overline{\theta}$ дає наближене значення з надлишком $\left(M\left(\overline{\theta}\right) > \theta\right)$. Якщо числа $\overline{\theta}_k$, $\left(k = 1,2,\ldots,m\right)$, будуть менші значення θ, тоді оцінка $\overline{\theta}$ дає наближене значення з недостачею $\left(M\left(\overline{\theta}\right) < \theta\right)$. Статистичну оцінку $\overline{\theta}$ параметра θ називають незміщеною, якщо $M\left(\overline{\theta}\right) = \theta$. Якщо $M\left(\overline{\theta}\right) \neq \theta$, то точкова оцінка параметра θ називається зміщеною. Різниця $\overline{\theta} - \theta$ називається зміщенням статистичної статистичної оцінки $\overline{\theta}$.

Теорема 2.1. Незміщенною оцінкою генеральної середньої (математичного сподівання) є вибіркова середня: $\bar{x}_B = \frac{1}{n}\sum_i x_i n_i$, де x_i - значення варіанти вибірки, n_i - її частота, $n = \sum_i n_i$ - обсяг вибірки.

Доведення. За означенням

$$M\left(\bar{x}_B\right) = M\left(\frac{1}{n}\sum_i x_i n_i\right) = \frac{1}{n}\sum_i M\left(x_i n_i\right) = \frac{1}{n}\sum_i n_i M\left(x_i\right).$$

Оскільки, $M\left(x_i\right) = \bar{x}_Г = a$, то

$$M\left(\bar{x}_B\right) = \frac{1}{n}\sum_i n_i M\left(x_i\right) = \frac{1}{n}\sum_i a n_i = \frac{a}{n}\sum_i n_i = a.$$

Отже, $M\left(\bar{x}_B\right) = a$, тобто оцінка $\bar{x}_B$ для математичного сподівання випадкової величини X є незміщеною. **Доведено**.

Точкова статистична оцінка $\bar{\theta}$ є **грунтовною**, якщо у разі необмеженого збільшення обсягу вибірки $\bar{\theta}$ наближається до оцінюваного параметра.

Якщо оцінка $\bar{\theta}_n$ параметра θ є незміщеною, а її дисперсія $D\left(\bar{\theta}_n\right) \to 0$ при $n \to \infty$, то оцінка $\bar{\theta}_n$ є і грунтовною. Це твердження випливає із другої нерівності Чебишева: $P\left(\left|\bar{\theta}_n - \theta\right| < \varepsilon\right) \geq 1 - \frac{D\left(\bar{\theta}_n\right)}{\varepsilon^2}$.

Теорема 2.2. Вибіркова середня є незміщеною і грунтовною оцінкою генеральної середньої $\bar{x}_Г$.

Доведення. Оскільки елементи вибірки є результатами незалежних спостережень над випадковою величиною X з математичним сподіванням a і дисперсією σ^2, то $M(x_1) = M(x_2) = ... = M(x_n) = a$ і

$$D(x_1) = D(x_2) = ... = D(x_n) = \sigma^2.$$

Знайдемо дисперсію вибіркової середньої:

$$D(\overline{x}_B) = D\left(\frac{1}{n}\sum_i x_i\right) = \frac{1}{n^2}\sum_i D(x_i) = \frac{1}{n^2}\sum_i \sigma^2 = \frac{n}{n^2}\sigma^2 = \frac{\sigma^2}{n}.$$

При $n \to \infty$ $D(\overline{x}_B) \to 0$. Оцінка незміщена і грунтовна. **Доведено**.

Простішим методом статистичного оцінювання є метод **підстановки** або аналогії. Він полягає в тому, що в якості оцінки тієї чи іншої числової характеристики (середнього арифметичного, дисперсії, середнього квадратичного відхилення,…) генеральної сукупності беруть відповідну характеристику розподілу вибірки – вибіркову характеристику.

Приклад 2.2. Нехай $x_1, x_2, ..., x_n$ - вибірка із генеральної сукупності зі скінченним математичним сподіванням $M(x) = a$ і дисперсією $D(x) = \sigma^2$. Використовуючи метод підстановки, знайти оцінку параметра a. Перевірити її незміщеність та грунтовність.

Розв'язання. В якості оцінки $\overline{a}$ математичного сподівання потрібно взяти математичне сподівання розподілу вибірки – вибіркову середню. Тоді,

$$\overline{a} = \overline{x}_B = \frac{1}{n}\sum_{i=1}^{n} x_i .$$

За теоремою 2.1. ця оцінка є незміщеною. Доведемо її грунтовність.

Нехай $D(x_i) = \sigma^2,\ i = 1, 2, ..., n$, причому x_i - незалежні в сукупності випадкові величини. Тоді, $D(\overline{x}_B) = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{\sigma^2}{n}$.

Оскільки при $n \to \infty$, $D\left(\overline{x}_B\right) \to 0$, то $\overline{x}_B$ є грунтовною оцінкою математичного сподівання a генеральної сукупності. **Доведено**.

Незміщеність оцінки $\overline{\theta}$ параметра θ означає, що при багаторазовій заміні θ на $\overline{\theta}$ середнє значення $\overline{\theta} - \theta$ дорівнюватиме нулю.

Для оцінювання невідомого параметра θ розподілу можна будувати багато незміщенних оцінок $\overline{\theta}$, тобто таких, що $M\left(\overline{\theta}\right) = \theta$. Потрібно із множини такіх оцінок вибрати ті, для яких дисперсія буде найменшою.

Незміщена оцінка $\overline{\theta}$ параметра θ називається **ефективною** (найкращою незміщеною оцінкою), якщо при заданому обсязі вибірки вона має мінімальну дисперсію. Нехай θ_n^* - ефективна оцінка параметра θ, а θ_n - деяка інша оцінка цього параметра.

Відношення $R = \dfrac{D\left(\theta_n^*\right)}{D\left(\theta_n\right)}$ називається **ефективністю оцінки**. $0 < R \leq 1$. Чим більш ефективна оцінка, тим R ближче до 1. Для нормального розподілу ефективною оцінкою дисперсії є вибіркова дисперсія D_B.

Грунтовна оцінка θ_n, дисперсія якої при $n \to \infty$ задовольняє умову $\lim\limits_{n\to\infty} \dfrac{D\left(\theta_n^*\right)}{D\left(\theta_n\right)} = 1$ називається **асимптотично ефективною оцінкою**.

Складна задача відшукання ефективних оцінок базується на наступній теоремі.

Теорема 2.3. Якщо випадкова вибірка складається з n незалежних спостережень над випадковою величиною X, яка має щільність ймовірності $\varphi\left(x,\theta\right)$ (а для дискретних величин – функцію розподілу ймовірностей), то при деяких достатньо загальних припущеннях відносно виду цієї оцінки і функції $\varphi(x,\theta)$ дисперсія ефективної оцінки θ_n^* параметра θ визначається формулою:

$$D\left(\theta_n^*\right)=\frac{1}{n\int\limits_{\Omega}\left(\frac{d\,ln\varphi}{d\theta}\right)^2\varphi(x,\theta)dx}$$ для неперервних величин і формулою

$$D\left(\theta_n^*\right)=\frac{1}{n\sum\limits_{x}\left(\frac{d\,ln\varphi}{d\theta}\right)^2\varphi(x,\theta)}$$ для дискретних випадкових величин.

Приклад 2.3. Знайти дисперсію ефективної оцінки параметра a, якщо генеральна сукупність із якої зроблено вибірку має нормальний розподіл з параметрами $N\left(a,\sigma^2\right)$.

Розв'язання.

Оскільки закон розподілу є нормальним, то $\varphi(x,\theta)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2\sigma^2}}$.

Для визначення дисперсії знайдемо

$$ln\varphi(x,\theta)=ln\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2\sigma^2}}\right)=ln\frac{1}{\sigma\sqrt{2\pi}}-\frac{(x-\theta)^2}{2\sigma^2}.$$

Тепер знайдемо похідну по θ:

$$\frac{d}{d\theta}ln\varphi(x,\theta)=-\frac{2(x-\theta)}{2\sigma^2}(-1)=\frac{x-\theta}{\sigma^2}.$$

Знайдені значення підставляємо у формулу для дисперсії:

$$D\left(\theta_n^*\right)=\frac{1}{n\int\limits_{-\infty}^{\infty}\left(\frac{x-\theta}{\sigma^2}\right)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2\sigma^2}}dx}$$

$$\int_{-\infty}^{\infty}\left(\frac{x-\theta}{\sigma^2}\right)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\theta)^2}{2\sigma^2}}dx=\frac{1}{\sigma^3\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{x-\theta}{\sigma}\right)^2e^{-\frac{(x-\theta)^2}{2\sigma^2}}dx=\begin{bmatrix}\textit{Заміна}\\ \dfrac{x-\theta}{\sigma}=t\\ dx=\sigma dt\end{bmatrix}=$$

$$=\frac{\sigma}{\sigma^3\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-\frac{t^2}{2}}dt=\left[\textit{інтегруємо частинами}\right]=\frac{1}{\sigma^2\sqrt{2\pi}}\left(-te^{-\frac{t^2}{2}}\right)\Bigg|_{-\infty}^{\infty}+$$

$$+\frac{1}{\sigma^2\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\left[\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}\right]=\frac{1}{\sigma^2\sqrt{2\pi}}\cdot\sqrt{2\pi}=\frac{1}{\sigma^2}.$$

Отже, $D\left(\theta_n^*\right)=\dfrac{\sigma^2}{n}$. Це означає, що вибіркова середня є ефективною оцінкою значень генеральної середньої a у нормальній генеральній сукупності.

Приклад 2.4. Показати, що вибіркова дисперсія є зміщеною оцінкою генеральної сукупності.

Розв'язання.

Нехай з генеральної сукупності отримано вибірку $x_1, x_2, \ldots, x_n$ і нехай розподіл генеральної сукупності залежить від параметра θ, значення якого невідоме.

Будемо вважати, що генеральна сукупність X має математичне сподівання a і дисперсію σ^2. Тоді

$$D_B=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}_B\right)^2=\frac{1}{n}\sum_{i=1}^{n}\left(\left(x_i-a\right)-\left(\overline{x}_B-a\right)\right)^2=$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-a\right)^2-2\frac{1}{n}\left(\overline{x}_B-a\right)\sum_{i=1}^{n}\left(x_i-a\right)+\frac{1}{n}n\left(\overline{x}_B-a\right)^2=$$

За формулою середньої арифметичної:

$$\frac{2}{n}\left(\overline{x}_B-a\right)\left(\sum_{i=1}^{n}x_i-na\right)=\frac{2}{n}\left(\overline{x}_B-a\right)\left(n\overline{x}_B-na\right)=2\left(\overline{x}_B-a\right)^2.$$

Тоді,

$$D_B = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2 - 2\left(\bar{x}_B - a\right)^2 + \left(\bar{x}_B - a\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2 - \left(\bar{x}_B - a\right)^2.$$

Звідки випливає, що

$$M\left(D_B\right) = M\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2 - \left(\bar{x}_B - a\right)^2\right) = M\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2\right) - M\left(\bar{x}_B - a\right)^2 =$$

$$= D_{\Gamma} - \frac{D\left(\bar{x}_B\right)}{n} = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2 .$$

Отже, вибіркова дисперсія є зміщеною оцінкою генеральної сукупності.►

Виправлену вибіркову дисперсію позначимо $\bar{D}_B$ або s^2:

$$s^2 = \frac{n}{n-1}D_B = \frac{1}{n-1}\sum_{i=1}^{k} n_i\left(x_i - \bar{x}_B\right)^2 .$$

Приклад 2.5. Знайти незміщені оцінки генеральної середньої та генеральної дисперсії за вибіркою:

37,35,36,35,36,37,36,34,35,36,36,35,34,36,36,38,36,35,35,35,34,36,35,37,36,36,37,36,36,38,34,36,35,36,35,36,34,37,36,36,35,37,34,36,35,36,37,35,38,37.

Розв'язання.

Обсяг вибірки: $n = 50$.

Статистиний розподіл:

x_i	34	35	36	37	38
n_i	6	13	20	8	3

Незміщена оцінка генеральної середньої:

$$\bar{x}_B = \frac{1}{n}\sum_{i=1}^{5} x_i n_i = \frac{1}{50}\left(34\cdot 6 + 35\cdot 13 + 36\cdot 20 + 37\cdot 8 + 38\cdot 3\right) = 35{,}78$$

Незміщена оцінка генеральної дисперсії:

$$S^2=\frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{5}x_i^2n_i-\left(\overline{x}_B\right)^2\right)=1{,}11.$$ ►

Приклад 2.6. Знайти ефективну оцінку генерального середнього $\overline{x}_Г$ (математичного сподівання a) повторної вибірки для нормально розподіленої генеральної сукупності.

Розв'язання. У випадку нормального закону розподілу щільність ймовірностей $\varphi(x,a)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$.

Тоді $$ln\varphi(x,a)=-ln\sqrt{2\pi\sigma^2}-\frac{(x-a)^2}{2\sigma^2}$$ і

$$\left[\left(ln\varphi(x,a)\right)_a'\right]^2=\left(0+\frac{x-a}{\sigma^2}\right)^2=\frac{(x-a)^2}{\sigma^4}.$$

Кількість інформації Фішера

$$I(a)=M\left[\left(ln\varphi(X,a)\right)_a'\right]^2=M\left[\frac{(x-a)^2}{\sigma^4}\right]=\frac{D(X)}{\sigma^4}=\frac{\sigma^2}{\sigma^4}=\frac{1}{\sigma^2}.$$

Мінімально можлива оцінка дисперсії $min\,D(\theta_n)=\frac{1}{nI(a)}=\frac{\sigma^2}{n}$.

Отже, вибіркове середнє $\overline{x}_B$ повторної вибірки для нормально розподіленої генеральної сукупності є ефективною оцінкою генеральної середньої $\overline{x}_Г$. Ефективною ж оцінкою генеральної дисперсії σ^2 є статистика

$$s^2=\frac{1}{n}\sum_{i=1}^{m}\left(x_i-\overline{x}_Г\right)^2n_i,$$

але для її знаходження необхідно знати генеральне середнє $\overline{x}_Г$, яке в більшості випадків застосування вибіркового методу невідоме.►

Методи знаходження точкових оцінок

1. Метод підстановок

Цей метод полягає в тому, що в якості оцінки тієї чи іншої числової характеристики (середнього, дисперсії, середнього квадратичного відхилення,…) генеральної сукупності беруть відповідну характеристику вибірки – вибіркову характеристику.

Приклад 2.7. Нехай $x_1, x_2, \ldots, x_n$ - вибірка із генеральної сукупності із скінченним математичним сподіванням a і дисперсією σ^2. Перевірити незміщеність і грунтовність оцінок математичного сподівання методом підстановки.

Розв'язання. За методом підстановки в якості оцінки математичного сподівання генеральної сукупності $M(x) = a$ потрібно взяти математичне сподівання розподілу вибірки – вибіркову середню: $\overline{x}_B = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Для перевірки незміщеності і грунтовності оцінки розглянемо цю статистику як функцію вибіркового вектора $(x_1, x_2, \ldots, x_n)$. За означенням вибіркового вектора маємо: $M(x_i) = a;\ D(x_i) = \sigma^2;\ i = 1, 2, \ldots, n$. Зауважимо, що x_i - незалежні в сукупності випадкові величини. Тому

$$M(\overline{x}_B) = M\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} M(x_i) = \frac{1}{n} na = a;$$

$$D(\overline{x}_B) = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(x_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

З першого співвідношення маємо незміщеність оцінки, а з другого – грунтовність, оскільки $D(\overline{x}_B) \to 0$ при $n \to \infty$ ►

Метод моментів (Пірсона)

Цей метод полягає в прирівнювані теоретичних моментів розподілу відповідним емпіричним моментам того самого порядка.
Нехай щільність ймовірностей розподілу $\varphi(x,\theta)$ генеральної сукупності визначається одним невідомим параметром θ. Для одержання точкової оцінки досить мати одне рівняння. Прирівняємо початковий момент першого порядку $\mu_1 = M(x)$ початковому емпіричному моменту першого порядка:

$\mu_1^{емп} = \frac{1}{n}\sum_{i=1}^{к} x_i n_i = \bar{x}_B$. Одержимо рівняння $M(x) = \mu_1^{емп} = \bar{x}_B$ з одним невідомим θ. Розв'язавши це рівняння одержимо оцінку параметра θ, яка буде функцією від $\bar{x}_B$, тобто функцією від $x_1, x_2, \ldots, x_n$.

Приклад 2.8. Випадкова величина X — число бракованих деталей у партії виробів, має розподіл Пуассона. У таблиці наведений розподіл бракованих виробів в 200 партіях. Знайти методом моментів точкову оцінку параметра λ розподілу Пуассона.

x_i	0	1	2	3	4
n_i	143	37	16	3	1

Розв’язання.

Розподіл Пуассона має вид: $P_m(x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$, де m - кількість випробувань, проведених в одному досліді, x_i - число настання події в *i*-ому досліді. Для знаходження оцінки параметра λ розглянемо рівняння $M(x) = \bar{x}_B$. Оскільки для розподілу Пуассона $M(x) = \lambda$, то одержимо рівняння:

$$\hat{\lambda} = \bar{x}_B = \frac{1}{n}\sum_{i=1}^{5} n_i x_i = \frac{1}{200}(143\cdot 0 + 1\cdot 37 + 2\cdot 16 + 3\cdot 3 + 4\cdot 1) = 0{,}41.$$

Отже, $\hat{\lambda}=0{,}41$ ►

Приклад 2.9. Випадкова величина X розподілена за показниковим законом із щільністю ймовірності $\varphi(x,\theta)=\theta e^{-\theta x}$, $x\geq 0$. За вибіркою $x_1,x_2,\ldots,x_n$ оцінити параметр θ.

Розв'язання. Для оцінки за методом Пірсона, потрібно розв'язати рівняння:

$$M_1=M_1^*=\bar{x}_B.$$

Початковий момент першого порядку :

$$M_1=\int_0^{\infty}x\varphi(x,\theta)\,dx=\theta\int_0^{\infty}xe^{-\theta x}dx=[\textit{інтегрування частинами}]=$$

$$=\theta\left(\left(-x\frac{1}{\theta}e^{-\theta x}\right)\Bigg|_0^{\infty}+\int_0^{\infty}\frac{1}{\theta}e^{-\theta x}dx\right)=\theta\left(-\frac{1}{\theta^2}e^{-\theta x}\Bigg|_0^{\infty}\right)=\frac{1}{\theta}$$

Підставимо це значення у початкову рівність: $\frac{1}{\theta^*}=\bar{x}_B$, $\quad\theta^*=\frac{1}{\bar{x}_B}$.►

Приклад 2.10. Випадкова величина X (час роботи елемента) має показниковий розподіл $\varphi(x,\theta)=\theta e^{-\theta x}\ (x\geq 0)$. В таблиці наведений емпіричний розподіл середнього часу роботи $n=200$ елементів, де у першоу рядку вказано середній час роботи, а в другому - кількість елементів, які працювали вказану кількість годин.

x_i	1,5	5,5	9,5	13,5	17,5
n_i	147	38	11	3	1

Методом Пірсона знайти точкову оцінку невідомого параметра θ .

Розв'язання. В прикладі 2.6. доведено, що ефективна оцінка параметра $\theta^*=\frac{1}{\bar{x}_B}$. За заданою таблицею обчислюємо:

$$\bar{x}_B=\frac{1}{n}\sum_{i=1}^{5}n_ix_i=\frac{1}{200}(1{,}5\cdot 147+5{,}5\cdot 38+9{,}5\cdot 11+13{,}5\cdot 3+17{,}5\cdot 1)=2{,}96$$

Тоді, $\theta^* = \dfrac{1}{2{,}96} = 0{,}338$.►

Нехай щільність розподілу визначається двома параметрами $\varphi\left(x,\theta_1,\theta_2\right)$.

За методом моментів Пірсона матимемо систему $\begin{cases} M\left(x\right) = \overline{x}_B \\ D\left(x\right) = D_B \end{cases}$. Розв'язавши цю систему відносно θ_1 і θ_2, знайдемо оцінки цих параметрів.

Приклад 2.11. Нехай $x_1, x_2, \ldots, x_n$ - вибірка з рівномірно розподіленої генеральної сукупності X на відрізку $\left[a,b\right]$. Оцінити параметри a та b за методом моментів.

Розв'язання. Для рівномірно розподіленої випадкової величини:

$$\varphi\left(x,a,b\right) = \begin{cases} 0, & x \notin \left[a,b\right] \\ \dfrac{1}{b-a}, & x \in \left[a,b\right] \end{cases}.$$

За теоремою про математичне сподівання і дисперсію для рівномірного розподілу, маємо:

$$\begin{cases} \dfrac{a+b}{2} = \overline{x}_B \\ \dfrac{\left(b-a\right)^2}{12} = D_B \end{cases} \Rightarrow \begin{cases} a = 2\overline{x}_B - b \\ \left(b-\overline{x}_B\right)^2 = 3D_B \end{cases} \Rightarrow \begin{cases} a = 2\overline{x}_B - b \\ b - \overline{x}_B = \sqrt{3D_B} \end{cases} \Rightarrow \quad ►$$

$$a = \overline{x}_B - \sqrt{3D_B}; \; b = \overline{x}_B + \sqrt{3D_B}$$

Метод максимальної правдоподібності

Цей метод розроблений Р.Фішером (англійський математик) і займає центральне місце в теорії статистичного оцінювання параметрів.

Дискретні випадкові величини

Нехай X – дискретна випадкова величина, яка в результаті n випробувань набула значень $x_1, x_2, \dots, x_n$. Відомий закон розподілу X, але невідоме значення параметра θ цього розподілу. Потрібно знайти точкову оцінку цього параметра. Позначимо через $p(x_i, \theta)$ ймовірність того, що в результаті випробування величина X прийме значення x_i $(i = 1, 2, \dots, n)$.

Функцією правдоподібності дискретної випадкової величини X називають функцію аргумента 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.

Оцінкою максимальної правдоподібності параметра θ називається таке його значення $\theta^* = \theta^*(x_1, \dots, x_n)$, при якому функція правдоподібності набуває максимального значення.

Наприклад, коли ознака генеральної сукупності розподілена за законом Пуассона, то функція максимальної правдоподібності матиме наступний вид:

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

$$= \frac{\theta^{x_1} e^{-\theta}}{x_1!} \cdot \frac{\theta^{x_2} e^{-\theta}}{x_2!} \cdot \dots \cdot \frac{\theta^{x_n} e^{-\theta}}{x_n!} = \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{x_1! x_2! \dots x_n!}$$

Оскільки функції L і $\ln L$ досягають максимума при одному і тому ж значенні θ, то замість функції L розглядають функцію $\ln L$ і знаходять її максимум.

Знайдене значення θ^* функції $\ln L$ приймають за точкову оцінку параметра θ.

Функцію $\ln L$ називають **логарифмічною функцією правдоподібності**.

Досліджують функцію на екстремум за звичайною схемою математичного аналіза.

Отримані оцінки є грунтовними, але можуть бути зміщенними. Якщо для параметра θ існує ефективна оцінка θ^*, то рівняння правдоподібності має єдиний розв'язок. До недоліків цього метода можна віднести складність обчислень, крім того, оцінка знайдена з функції правдоподібності не завжди співпадає із оцінкою, знайденою методом моментів.

Приклад 2.12. В схемі Бернуллі необхідно оцінити невідому ймовірність p появи події A в одному випробуванні. Яке найбільш правдоподібне значення p, якщо число настання події A дорівнює x?

Розв'язання. Запишемо функцію правдоподібності: $L(x,\theta)=C_n^x\theta^x(1-\theta)^{n-x}$.

C_n^x - це константа, яку при відшуканні максимума можна не враховувати.

Досліджуватимемо на максимум функцію:

$L(x,\theta)=\theta^x(1-\theta)^{n-x}$; $\ln L(x,\theta)=x\ln\theta+(n-x)\ln(1-\theta)$.

$$\frac{d\ln L}{d\theta}=\frac{x}{\theta}-\frac{(n-x)}{1-\theta}=\frac{x-n\theta}{\theta(1-\theta)}=0.$$

Звідси, $x-n\theta=0$; $\theta=\frac{x}{n}$. При $\theta<\frac{x}{n}$, $\frac{d\ln L}{d\theta}>0$. При $\theta>\frac{x}{n}$, $\frac{d\ln L}{d\theta}<0$.

Отже, $\theta=\frac{x}{n}$ - точка максимума, тому $\theta^*=\frac{x}{n}$. Ця оцінка є незміщеною, оскільки її математичне сподівання дорівнює p. Оцінка є грунтовною, оскільки $\frac{x}{n}\to p$ при $n\to\infty$.

Приклад 2.13. За вибіркою $x_1,x_2,\ldots,x_n$ знайти точкову оцінку параметра p геометричного розподілу методом максимальної правдоподібності.

Розв'язання. Геометричний розподіл: $P(x=x_i)=(1-p)^{x_i-1}p$, де x_i - число випробувань, проведених до настання події, p - ймовірність настання події в одному випробуванні.

Функція правдоподібності:

$$L(x_1,x_2,\ldots,x_n,\theta)=(1-p)^{x_1-1}p\cdot(1-p)^{x_2-1}p\cdot\ldots\cdot(1-p)^{x_n-1}p=(1-p)^{\sum x_i-1}p^n.$$

$$ln\,L = \left(\sum x_i - n\right) ln\left(1-p\right) + n\,ln\,p.$$

$$\frac{d\,ln\,L}{dp} = -\frac{\left(\sum x_i - n\right)}{1-p} + \frac{n}{p} = \frac{n - p\sum x_i}{p\left(1-p\right)} = 0.$$

$$n - p\sum x_i = 0;\ \ p = \frac{n}{\sum x_i} = \frac{1}{\frac{1}{n}\sum x_i} = \frac{1}{\overline{x}_B}.$$

Дослідимо на екстремум логарифмічну функцію правдоподібності за другою похідною при $p = \frac{1}{\overline{x}_B}$.

$$\frac{d^2\,ln\,L}{dp^2} = -\frac{\left(\sum x_i - n\right)}{\left(1-p\right)^2} - \frac{n}{p^2} = -\frac{p^2\sum x_i + n\left(1-2p\right)}{p^2\left(1-p\right)^2} = -\frac{n\left(p^2\overline{x}_B + 1 - 2p\right)}{p^2\left(1-p\right)^2}.$$

При $p = \frac{1}{\overline{x}_B}$: $\frac{d^2\,ln\,L}{dp^2} = -\frac{n\left(\frac{1}{\overline{x}_B} + 1 - \frac{2}{\overline{x}_B}\right)}{p^2\left(1-p\right)^2} = -\frac{n\left(1 - \frac{1}{\overline{x}_B}\right)}{p^2\left(1-p\right)^2} < 0$, оскільки $\overline{x}_B > 1$.

Отже, $p = \frac{1}{\overline{x}_B}$ - точка максимума. Тоді, $\theta^* = \frac{1}{\overline{x}_B}$.►

Неперервні випадкові величини

Нехай X – неперервна випадкова величина, яка в процесі випробувань набуває значень $x_1, x_2, \ldots, x_n$. Відома щільність ймовірностей $\varphi\left(x, \theta\right)$ розподілу X, але невідомий параметр θ цього розподілу. Потрібно знайти точкову оцінку цього параметра.

Функція правдоподібності неперервної випадкової величини

$L\left(x_1, x_2, \ldots, x_n, \theta\right) = \varphi\left(x_1, \theta\right) \cdot \varphi\left(x_2, \theta\right) \cdot \ldots \cdot \varphi\left(x_n, \theta\right)$, де $x_1, x_2, \ldots, x_n$ - фіксовані значення. Оцінку шукають аналогічно дискретному випадку.

Приклад 2.14. За вибіркою $x_1, x_2, \ldots, x_n$ знайти точкову оцінку параметра θ показникового розподілу методом максимальної правдоподібності.

Розв'язання. Щільність ймовірностей показникового розподілу:

$\varphi(x,\theta)=\theta e^{-\theta x}$.

Функція правдоподібності:

$$L(x_1,x_2,...,x_n,\theta)=\left(\theta e^{-\theta x_1}\right)\cdot\left(\theta e^{-\theta x_2}\right)\cdot...\cdot\left(\theta e^{-\theta x_n}\right)=\theta^n e^{-\theta\sum x_i}.$$

Логарифмічна функція правдоподіьності:

$ln L=n ln\theta-\theta\sum x_i$; $\dfrac{d\,ln L}{d\theta}=\dfrac{n}{\theta}-\sum x_i=0$.

Стаціонарна точка $\theta=\dfrac{n}{\sum x_i}=\dfrac{1}{\frac{1}{n}\sum x_i}=\dfrac{1}{\bar{x}_B}$.

Досліджуємо на екстремум за допомогою другої похідної:

$\dfrac{d^2\,ln L}{d\theta^2}=-\dfrac{n}{\theta^2}<0$. Отже, $\theta^*=\dfrac{1}{\bar{x}_B}$ - точка максимума.▶

Якщо щільність розподілу визначається двома параметрами $\varphi(x,\theta_1,\theta_2)$, то дослідження логарифмічної функції правдоподібності виконують спираючись на теорію функцій багатьох змінних.

Приклад 2. 15. Знайти оцінки параметрів нормально розподіленої випадкової величини методом максимальної правдоподібності.

Розв'язання. Щільність ймовірностей нормального розподілу:

$\varphi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$. Нехай $x_1,x_2,...,x_n$ - вибірка з нормальної сукупності.

Функція правдоподібності:

$$L(x,a,\sigma^2)=\varphi(x_1,a,\sigma^2)\cdot\varphi(x_2,a,\sigma^2)\cdot...\cdot\varphi(x_n,a,\sigma^2)=$$

$$=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_1-a)^2}{2\sigma^2}}\cdot\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_2-a)^2}{2\sigma^2}}\cdot...\cdot\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_n-a)^2}{2\sigma^2}}=\frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n}e^{-\frac{\sum(x_i-a)^2}{2\sigma^2}}.$$

Логарифмічна функція правдоподіьності:

$$ln\,L=-\frac{n}{2}ln\left(2\pi\right)-\frac{n}{2}ln\left(\sigma^2\right)-\frac{\sum\limits_{i=1}^{n}\left(x_i-a\right)^2}{2\sigma^2}.$$

Для дослідження знаходимо частинні похідні:

$$\begin{cases}\dfrac{\partial\,ln\,L\left(x,a,\sigma^2\right)}{\partial a}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}\left(x_i-a\right)=0\\ \dfrac{\partial\,ln\,L\left(x,a,\sigma^2\right)}{\partial\left(\sigma^2\right)}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\left(\sigma^2\right)^2}\sum\limits_{i=1}^{n}\left(x_i-a\right)^2=0\end{cases}$$

З першого рівняння: $\sum\limits_{i=1}^{n}\left(x_i-a\right)=0;\ \sum\limits_{i=1}^{n}x_i-na=0\Rightarrow a=\frac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}_B$.

Підставимо це значення в лруге рівняння:

$\frac{1}{2\left(\sigma^2\right)^2}\sum\limits_{i=1}^{n}\left(x_i-\bar{x}_B\right)^2=\frac{n}{2\sigma^2}\Leftrightarrow\frac{1}{n\sigma^2}\sum\limits_{i=1}^{n}\left(x_i-\bar{x}_B\right)^2=1$. Тоді,

$$\sigma^2=\frac{1}{n}\sum\limits_{i=1}^{n}\left(x_i-\bar{x}_B\right)^2=D_B.$$

Дослідимо на екстремум:

$$\frac{\partial^2\,ln\,L}{\partial a^2}=-\frac{n}{\sigma^2};\qquad \frac{\partial^2\,ln\,L}{\partial a\partial\left(\sigma^2\right)}=-\frac{1}{\left(\sigma^2\right)^2}\sum\limits_{i=1}^{n}\left(x_i-a\right);$$

$$\frac{\partial^2\,ln\,L}{\partial^2\left(\sigma^2\right)}=-\frac{n}{2\left(\sigma^2\right)^2}-\frac{\sum\limits_{i=1}^{n}\left(x_i-a\right)^2}{\left(\sigma^2\right)^3}$$

Підставимо стаціонарну точку $\left(\bar{x}_B,D_B\right)$ у вирази для других похідних:

$$\frac{\partial^2\,ln\,L}{\partial a^2}=-\frac{n}{D_B}<0;\qquad \frac{\partial^2\,ln\,L}{\partial a\partial\left(\sigma^2\right)}=-\frac{\sum\limits_{i=1}^{n}\left(x_i-\bar{x}_B\right)}{D_B^2}=0;$$

$$\frac{\partial^2\,ln\,L}{\partial^2\left(\sigma^2\right)}=\frac{n}{2D_B^2}-\frac{\sum\limits_{i=1}^{n}\left(x_i-\bar{x}_B\right)^2}{D_B^3}$$

$\Delta = -\dfrac{n}{D_B}\left(-\dfrac{n}{2D_B^2}\right) = \dfrac{n^2}{2D_B^3} > 0$. Отже, функція в цій точці має максимум.▶

Метод найменших квадратів дослідження вибірки

Для відшукання наближених залежностей між величинами, які вивчаються за вибіркою, доцільно використовувати метод найменших квадратів.
Нехай відома функціональна залежність між випадковими величинами X та Y: $Y = f\left(x, a_1, a_2, ..., a_m\right)$. І нехай за результатами спостережень отримали таку таблицю:

x	x_1	x_2	x_3	$\dots$	x_i	$\dots$	x_n
y	y_1	y_2	y_3	$\dots$	y_i	$\dots$	y_n

Для знаходження оцінок параметрів $a_1, a_2, ..., a_m$ функціональної залежності за даними вибірки застосуємо метод найменших квадратів: найімовірніші значення параметрів мають давати мінімум функції

$S = \sum_{i=1}^{n}\left(y_i - f\left(x, a_1, a_2, ..., a_m\right)\right)^2$. Необхідна умова існування екстремума:

$\dfrac{\partial S}{\partial a_1} = 0; \ \dfrac{\partial S}{\partial a_2} = 0; ...; \dfrac{\partial S}{\partial a_m} = 0$, якщо вони існують.

Знаходження функціональної залежності між випадковими величинами X та Y за даними вибірки називають **вирівнюванням емпіричних даних вздовж кривої** $y = f\left(x, a_1, a_2, ..., a_m\right)$.

2.3. Оцінка параметрів лінійної функції

Нехай між випадковими величинами X та Y існує лінійна залежність: $Y = aX + b$. Параметри a та b невідомі. За методом найменших квадратів шукатимемо мінімум функції. Ця функція диференційована. Необхідні умови існування екстремума: $\frac{\partial S}{\partial a}=0;\ \frac{\partial S}{\partial b}=0$.

$$\begin{cases}\sum\limits_{i=1}^{n}\left(y_i-ax_i-b\right)x_i=0\\ \sum\limits_{i=1}^{n}\left(y_i-ax_i-b\right)=0\end{cases}\Leftrightarrow\begin{cases}a\sum\limits_{i=1}^{n}x_i+nb=\sum\limits_{i=1}^{n}y_i\\ a\sum\limits_{i=1}^{n}x_i^2+b\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}x_iy_i\end{cases}$$

. Ця система називається **нормальною системою** метода найменших квадратів.

За правилом Крамера для лінійних систем:

$$\Delta=\begin{vmatrix}\sum\limits_{i=1}^{n}x_i & n\\ \sum\limits_{i=1}^{n}x_i^2 & \sum\limits_{i=1}^{n}x_i\end{vmatrix}=\left(\sum_{i=1}^{n}x_i\right)^2-n\sum_{i=1}^{n}x_i^2\ ;$$

$$\Delta_a=\begin{vmatrix}\sum\limits_{i=1}^{n}y_i & n\\ \sum\limits_{i=1}^{n}x_i\,y_i & \sum\limits_{i=1}^{n}x_i\end{vmatrix}=\left(\sum_{i=1}^{n}x_i\cdot\sum_{i=1}^{n}y_i\right)-n\sum_{i=1}^{n}x_i\,y_i\,;$$

$$\Delta_b=\begin{vmatrix}\sum\limits_{i=1}^{n}x_i & \sum\limits_{i=1}^{n}y_i\\ \sum\limits_{i=1}^{n}x_i^2 & \sum\limits_{i=1}^{n}x_iy_i\end{vmatrix}=\left(\sum_{i=1}^{n}x_i\cdot\sum_{i=1}^{n}x_iy_i\right)-\sum_{i=1}^{n}x_i^2\sum_{i=1}^{n}y_i\,.$$

$$a=\frac{\sum\limits_{i=1}^{n}x_i\sum\limits_{i=1}^{n}y_i-n\sum\limits_{i=1}^{n}x_iy_i}{\left(\sum\limits_{i=1}^{n}x_i\right)^2-n\sum\limits_{i=1}^{n}x_i^2};\quad b=\frac{\sum\limits_{i=1}^{n}x_i\sum\limits_{i=1}^{n}x_iy_i-\sum\limits_{i=1}^{n}x_i^2\sum\limits_{i=1}^{n}y_i}{\left(\sum\limits_{i=1}^{n}x_i\right)^2-n\sum\limits_{i=1}^{n}x_i^2}.$$

За означенням $\frac{1}{n}\sum\limits_{i=1}^{n}x_i=\overline{x}_a$. Підставимо це значення у вирази для параметрів:

$$a=\frac{\sum\limits_{i=1}^{n}\left(x_i-\overline{x}_a\right)y_i}{\sum\limits_{i=1}^{n}\left(x_i-\overline{x}_a\right)^2};\quad b=\frac{1}{n}\sum_{i=1}^{n}y_i-a\overline{x}_a.$$

Перевіримо, що функція $S\left(a,b\right)$ в знайденній точці $\left(a,b\right)$ матиме мінімум.

$$\frac{\partial^2 S}{\partial a^2}=2\sum_{i=1}^{n}x_i^2;\quad \frac{\partial^2 S}{\partial a\partial b}=2\sum_{i=1}^{n}x_i;\quad \frac{\partial^2 S}{\partial b^2}=2n.$$

$$\Delta\bigg|_{(a,b)}=\frac{\partial^2 S}{\partial a^2}\cdot\frac{\partial^2 S}{\partial b^2}-\left(\frac{\partial^2 S}{\partial a\partial b}\right)^2=4n\sum_{i=1}^{n}x_i^2-\left(2\sum_{i=1}^{n}x_i\right)^2=2\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i-x_j\right)^2>0$$

$\left(x_i\neq x_j\right)$. Отже, функція $S\left(a,b\right)$ в точці $\left(a,b\right)$ має мінімум.

Приклад 2.16. Скласти рівняння прямої, яка проходить найближче до точок заданої вибірки методом найменших квадратів.

x_i	-1	0	1	3	5	7
y_i	5,0	4,3	4,0	2,6	2,0	0,6

Розв'язання.

$$\sum_{i=1}^{n}x_i=-1+0+1+3+5+7=15$$

$$\sum_{i=1}^{n}y_i=5{,}0+4{,}3+4{,}0+2{,}6+2{,}0+0{,}6=18{,}5$$

$$\sum_{i=1}^{n}x_iy_i=-1\cdot5+0\cdot4{,}3+1\cdot4+3\cdot2{,}6+5\cdot2+7\cdot0{,}6=21{,}0$$

$$\sum_{i=1}^{n}x_i^2=1+0+1+9+25+49=85$$

Нормальна система: $\begin{cases}15a+6b=18{,}5\\85a+15b=21\end{cases}\Leftrightarrow a=-0{,}532;\ b=4{,}412$.

Отже, $y=-0{,}537x+4{,}412$ ►

Оцінка параметрів параболічної функціональної залежності

Нехай між випадковими величинами X та Y існує лінійна залежність: $Y = aX^2 + bX + c$. Параметри a, b та c невідомі. За методом найменших квадратів на основі даних випробувань шукатимемо мінімум функції для оцінки невідомих параметрів.

$$S(a,b,c) = \sum_{i=1}^{n} \left(y_i - ax^2 - bx - c \right)^2$$

Ця функція диференційована. Необхідні умови існування екстремума:

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum_{i=1}^{n} \left(y_i - ax_i^2 - bx_i - c \right) x_i^2 = 0 \\ \dfrac{\partial S}{\partial b} = -2\sum_{i=1}^{n} \left(y_i - ax_i^2 - bx_i - c \right) x_i = 0 . \\ \dfrac{\partial S}{\partial c} = -2\sum_{i=1}^{n} \left(y_i - ax_i^2 - bx_i - c \right) = 0 \end{cases}$$

Запишемо систему в іншому виді:

$$\begin{cases} a\sum_{i=1}^{n} x_i^4 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i^2 y_i \\ a\sum_{i=1}^{n} x_i^3 + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i + cn = \sum_{i=1}^{n} y_i \end{cases}$$

Цю систему можна розв’язати будь-яким методом: матричним, за правилом Крамера, за методом Гаусса. Але, якщо значення вибірки рівновіддалені і центровані, то розв'язок значно спрощується і знаходиться за формулами:

$$a = \frac{n\sum_{i=1}^{n} x_i^2 y_i - \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^2 \right)^2}; \;\; b = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}; \;\; c = \frac{\sum_{i=1}^{n} x_i^4 \cdot \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} x_i^2 y_i}{n\sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^2 \right)^2}.$$

Приклад 2.17. Скласти рівняння параболи $y = ax^2 + bx + c$, яка проходить найближче до точок заданої вибірки методом найменших квадратів.

x_i	1	2	3	4	5	6	7
y_i	0,5	0,5	1	2	3	5	8

Розв'язання. Складемо розрахункову таблицю:

	x_i	x_i^2	x_i^3	x_i^4	y_i	$x_i\,y_i$	$x_i^2\,y_i$
1	1	1	1	1	0,5	0,5	0,5
2	2	4	8	16	0,5	1	2
3	3	9	27	81	1	3	9
4	4	16	64	256	2	8	32
5	5	25	125	625	3	15	75
6	6	36	216	1296	5	30	180
7	7	49	343	2401	8	56	392
Σ	28	140	784	4676	20	113,5	690,5

Підставляємо ці значення в загальні формули для оцінки параметрів:

$$\begin{cases} 4676a + 784b + 140c = 690{,}5 \\ 784a + 140b + 28c = 113{,}5 \\ 140a + 28b + 7c = 20 \end{cases}$$

За правилом Крамера:

$$\Delta = \begin{vmatrix} 4676 & 784 & 140 \\ 784 & 140 & 28 \\ 140 & 28 & 7 \end{vmatrix} = 16464; \quad \Delta_a = \begin{vmatrix} 690{,}5 & 784 & 140 \\ 113{,}5 & 140 & 28 \\ 20 & 28 & 7 \end{vmatrix} = 4410;$$

$$\Delta_b = \begin{vmatrix} 4676 & 690{,}5 & 140 \\ 784 & 113{,}5 & 28 \\ 140 & 20 & 7 \end{vmatrix} = -15582; \quad \Delta_c = \begin{vmatrix} 4676 & 784 & 690{,}5 \\ 784 & 140 & 113{,}5 \\ 140 & 28 & 20 \end{vmatrix} = 21168$$

Отже, $a=\dfrac{\Delta_a}{\Delta}=0,268;\quad b=\dfrac{\Delta_b}{\Delta}=-0,946;\quad c=\dfrac{\Delta_c}{\Delta}=1,286$.

Тоді парабола, яка проходить найближче до емпіричних точок має рівняння:

$y=0,268x^2-0,946x+1,286$ ►

Оцінка параметрів гіперболічної функціональної залежності

Нехай між випадковими величинами X та Y існує гіперболічна залежність:

$Y=\dfrac{a}{X}+b$. Параметри a та b невідомі. За методом найменших квадратів на основі вибірки шукатимемо мінімум функції для оцінки невідомих параметрів.

$$S(a,b)=\sum_{i=1}^{n}\left(y_i-\frac{a}{x}-b\right)^2 .$$

$$\begin{cases}\dfrac{\partial S}{\partial a}=2\displaystyle\sum_{i=1}^{n}\frac{1}{x_i}\left(y_i-\frac{a}{x_i}-b\right)=0\\ \dfrac{\partial S}{\partial b}=2\displaystyle\sum_{i=1}^{n}\left(y_i-\frac{a}{x_i}-b\right)=0\end{cases};\qquad \begin{cases}a\displaystyle\sum_{i=1}^{n}\frac{1}{x_i^2}+b\sum_{i=1}^{n}\frac{1}{x_i}=\sum_{i=1}^{n}\frac{y_i}{x_i}\\ a\displaystyle\sum_{i=1}^{n}\frac{1}{x_i}+bn=\sum_{i=1}^{n}y_i\end{cases}.$$

Приклад 2.18. Скласти рівняння гіперболи $y=\dfrac{a}{x}+b$, яка проходить найближче до точок заданої вибірки методом найменших квадратів.

x_i	1	2	3	4	5	6	7	8
y_i	12,2	6,8	5,2	4,6	3,9	3,7	3,5	3,2

Розв'язання. Складемо розрахункову таблицю:

x_i	y_i	$1/x_i$	$1/x_i^2$	$x_i\ /\ y_i$
1	12,2	1,000	1	12,200
2	6,8	0,500	0,250	3,400
3	5,2	0,333	0,111	1,7333

4	4,6	0,250	0,0625	1,150
5	3,9	0,200	0,040	0,780
6	3,7	0,167	0,028	0,617
7	3,5	0,143	0,020	0,500
8	3,2	0,125	0,016	0,400
Σ	43,1	2,718	1,5275	20,7803

Підставляємо ці значення в формули для оцінки параметрів:

$\begin{cases} 1,52a + 2,718b = 20,78 \\ 2,71a + 8b = 43,1 \end{cases} \Leftrightarrow a = 10,15;\ b = 1,94$. Отже, рівняння гіперболи:

$$y = \frac{10,15}{x} + 1,94$$ ►

Оцінка параметрів експоненціальної функціональної залежності

Нехай між випадковими величинами X та Y існує експоненціальна залежність: $Y = ae^{bX}$. Параметри a та b невідомі. За методом найменших квадратів на основі вибірки шукатимемо мінімум логарифмічної функції для оцінки невідомих параметрів.

Приклад 2.19. Для визначення параметрів a, b у формулі функції $y = ae^{bx}$ були виміряні наступні данні:

x_i	-1	-0,75	-0,5	-0,25	0	0,25	0,5	0,75	1
y_i	7,14	4,46	2,81	1,68	0,99	0,5	0,36	0,33	0,40

Знайти оцінки параметрів методом найменших квадратів.

Розв'язання. Прологарифмуємо задану функцію:

$$ln\, y = bx + ln\, a = bx + c \ \ (c = ln\, a).$$

$$S\left(b,c\right)\approx\sum_{i=1}^{n}\left(ln\,y_i-bx_i-c\right)^2 y_i\ .$$

Шукаємо мінімум цієї функції:

$$\begin{cases}\dfrac{\partial S}{\partial b}=-2\sum\limits_{i=1}^{n}\left(ln\,y_i-bx_i-c\right)y_ix_i\ =0\\ \dfrac{\partial S}{\partial c}=-2\sum\limits_{i=1}^{n}\left(ln\,y_i-bx_i-c\right)y_i=0\end{cases};\ \begin{cases}b\sum\limits_{i=1}^{9}y_ix_i^2+c\sum\limits_{i=1}^{9}x_iy_i=\sum\limits_{i=1}^{n}x_iy_i\,ln\,y_i\\ b\sum\limits_{i=1}^{n}y_ix_i\ +c\sum\limits_{i=1}^{n}y_i=\sum\limits_{i=1}^{n}y_i\,ln\,y_i\end{cases}.$$

Обчисливши коефіцієнти системи за даними таблиці, дістанемо:

$$\begin{cases}11{,}08b-11{,}36c=-21{,}63\\ -11{,}36b+18{,}67c=23{,}02\end{cases}\Leftrightarrow b=-1{,}83;\ \ c=0{,}12\,.\,c=ln\,a=ln\,0{,}12\approx1{,}128$$

Отже, $y=ae^{bx}\ =1{,}13e^{-1{,}83x}$ ▶