

Приклади розв'язування задач на побудову статистичного розподілу

1. Вибірка задана у вигляді розподілу частот:

x_i	2	5	7
n_i	1	3	6

Знайти розподіл частот.

Розв'язування

Зайдемо об'єм вибірки: $n = 1 + 3 + 6 = 10$. Знайдемо відносні частоти:

$$w_1 = \frac{1}{10}, w_2 = \frac{3}{10}, w_3 = \frac{6}{10}.$$

Залишимо шуканий розподіл відносних частот:

x_i	2	5	7
ω_i	0,1	0,3	0,6

Контроль $0,1 + 0,3 + 0,6 = 1$.

2. Знайти емпіричну функцію за даним розподілом вибірки:

x_i	1	4	6
n_i	10	15	25

Розв'язування

Зайдемо об'єм вибірки: $n = 10 + 15 + 25 = 50$. Найменша варіанта n_x - число значень ознаки, для яких $X \leq x$ дорівнює одиниці, тому $F(x) = 0$, якщо $x \leq 1$.

Значення $X < 4$, а саме $x_1 = 1, x_2 = 4$, зустрічається 10 разів, отже,

$$F(x) = \frac{10}{50} = 0,2, \quad 1 < x \leq 4.$$

Значення $X < 6$, а саме $x_1 = 1, x_2 = 4$, зустрічається $10 + 15 = 25$ разів, отже,

$$F(x) = \frac{25}{50} = 0,5, \quad 4 < x \leq 6.$$

Оскільки $X = 6$ - найбільша варіанта, то $F^*(x) = 1$, якщо $x > 6$.

Запишемо шукану емпіричну функцію:

$$F(x) = \begin{cases} 0, & x \leq 1, \\ 0,2, & 1 < x \leq 4, \\ 0,5, & 4 < x \leq 6, \\ 1, & x > 6. \end{cases}$$

3. Для вивчення попиту на певний розмір взуття власник магазину спостерігає розміри взуття, проданого протягом дня:

40, 35, 37, 39, 40, 41, 36, 42, 40, 39, 43, 43, 41, 38, 37, 36, 42, 40, 38.

Знайти статистичний розподіл цієї вибірки.

Розв'язування

Для побудови статистичного розподілу використаємо програмний продукт Excel і, зокрема, статистичну функцію ЧАСТОТА. Для цього у стовпчик A1-A20 введемо вхідні дані у стовпчик B1-B9 запишемо відповідні розміри від 35 до 43. Далі виділимо стовпчик C1-C9 і введемо функцію масиву ЧАСТОТА(A1:A20;B1-B9) і після натискання клавіш Ctrl+Shift+Enter у стовпчику C1-C9 отримаємо відповідні результати (див. Табл.10.1).

Таблиця 10.1.

Розмір взуття x_k	35	36	37	38	39	40	41	42	43
Кількість n_k	1	3	2	2	2	4	2	2	2

Контроль $1 + 3 + 2 + 2 + 2 + 4 + 2 + 2 + 2 = 20$, $n = 20$.

4. Для вибірки з таблиці 10.1 знайти розподіл відносних частот.

Розв'язування

Об'єм вибірки $n = 1 + 3 + 2 + 2 + 2 + 4 + 2 + 2 + 2 = 20$.

Відносними частотами будуть:

$$\omega_1 = \frac{1}{20} = 0,05, \omega_2 = \frac{3}{20} = 0,15, \omega_9 = \frac{2}{20} = 0,1.$$

Тому розподіл відносних частот цієї вибірки має вигляд (табл.10.2)

Таблиця 10.2

Розмір взуття x_k	35	36	37	38	39	40	41	42	43
Кількість ω_k	0,05	0,15	0,10	0,10	0,10	0,20	0,10	0,10	0,10

Контроль $0,05 + 0,15 + 3 \cdot 0,1 + 0,2 + 3 \cdot 0,1 = 1$

Зауваження. Властивості елементів ряду частот (див. формулу (10.1)) та ряду відносних частот (див. формулу (10.2)) використовують для контролю одержаного статистичного розподілу вибірки.

5. Знайти емпіричну функцію розподілу за статистичним розподілом вибірки

x_k	2	6	1
n_k	1	1	3
	2	8	0

та побудувати її графік.

Розв'язування

Об'єм цієї вибірки буде $n = 12 + 18 + 30 = 60$. Найменша варіанта дорівнює 2, тому $F^*(x) = 0$, для $x \leq 2$. Найбільша варіанта дорівнює 10,

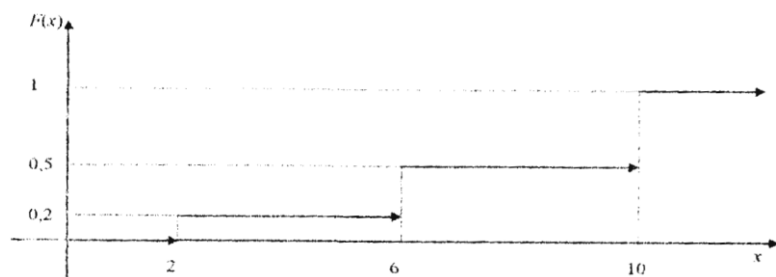
тому $F^*(x)=1$, для $x > 10$. Значення $x < 6$, тобто $X = \{x_1 = 2\}$, спостерігалася 12 разів, тому $F^*(x) = \frac{12}{60} = 0,2$, при $2 < x \leq 6$.

Значення $X < 10$, тобто $X = \{x_1 = 2 \text{ та } x_2 = 6\}$, спостерігалася $12 + 18 = 30$ разів, тому $F^*(x) = \frac{30}{60} = 0,5$, при $6 < x \leq 10$.

Таким чином, емпірична функція розподілу має вигляд:

$$F'(x) = \begin{cases} 0, & x \leq 2, \\ 0,2, & 2 < x \leq 6, \\ 0,5, & 6 < x \leq 10, \\ 1, & x > 10. \end{cases}$$

Графік цієї функції зображено на малюнку 10.2. Цей графік є наближенням графіка теоретичної функції розподілу $F(x)$.



Мал.10.2

4. У результаті спостереження одержали розподіл ознаки X вибірки у вигляді

Варіанти X_k	-2	0	1	2	3	5	7
Частоти n_k	4	5	7	8	6	2	1

Побудувати гістограму частот цього розподілу.

Розв'язування

У даному випадку найменше значення варіанти $x_{\min} = -2$, а найбільше значення — $x_{\max} = 7$, тому довжина проміжку $[x_{\min}, x_{\max}]$ дорівнює 9. Розіб'ємо цей відрізок на 4 рівних частини довжиною $h = 9/4 = 2,25$.

Для побудови гістограми без використання комп'ютера доцільно скласти таблицю 10.3: у перший рядок таблиці записують одержані відрізки, у другий рядок таблиці записують суму частот варіант, що належать відповідному відрізку, а у третій рядок записують висоти відповідних прямокутників. У даному випадку ця таблиця буде мати вигляд

Таблиця 10.3

Відрізки довжини $h = 2.25$	$[-2; 0.25]$	$[0.25; 2.5]$	$[2.5; 4.75]$	$[4.75; 7]$
Кількість значень n_k	9	15	6	3
Щільність частоти $\frac{n_k}{h}$	4	$20/3$	$8/3$	$4/3$

Відмітимо, що

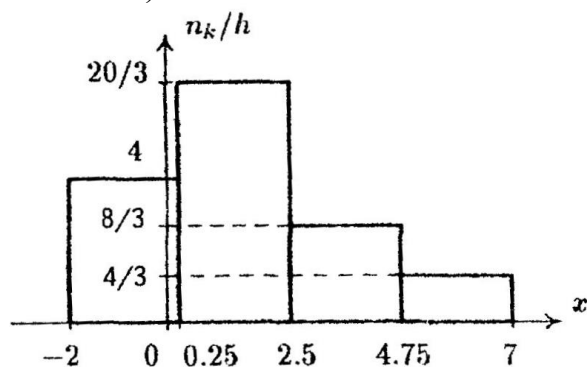
Відрізку $[-2; 0.25]$ належать 9 значень варіант (-2 чотири рази та 0 п'ять разів);

відрізку $[0.25; 2.5]$ належать 15 значень варіант (1 — сім разів та 2 — вісім разів);

відрізку $[2.5; 4.75]$ належать 6 значень варіант (3 — шість разів);

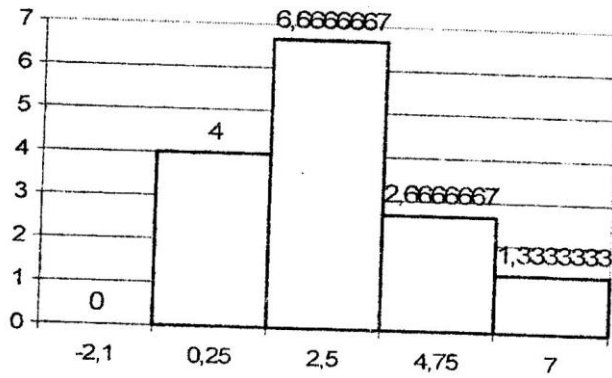
відрізку $[4.75; 7]$ належать 3 значення (5—2 рази та 7 — один раз).

За даними таблиці 10.3 будемо відповідну гістограму частот (див. мал. 10.3).



Мал.10.3.

Розглянемо побудову таблиці 10.3 і відповідної гістограми на комп'ютері. Для побудови статистичного розподілу використаємо програмний продукт Excel і, зокрема, статистичну функцію ЧАСТОТА. Для цього у стовпчик A1-A33 введемо вихідні дані з першого рядка вихідної таблиці у стовпчик B1—B5 запишемо праві границі відповідних відрізків(-2; 0,25; 2,5; 4,75; 7). Далі виділимо стовпчик C1-C6 і введемо функцію масиву ЧАСТОТА(A1:A33;B1:B5) і після натискання клавіш Ctrl+Shift+Enter у стовпчику C2—C5 отримаємо результат з другого рядка таблиці 10.3. Щоб отримати третій рядок, розділимо на h елементи стовпчика B. Результати занесемо у стовпчик D2-D5. Щоб побудувати гістограму, в меню ВСТАВКА виберемо пункт ДИАГРАММА. Далі виберемо тип ГИСТОГРАММА, потім ДИАПАЗОН D1:D4, потім на вкладці РЯД ПОДПИСИ НА ОСИ X B1:B5. Далі, корегуючи параметри діаграми, отримаємо гістограму, що зображена на малюнку 10.4.



Мал.10.4

Зауваження. Іноді для побудови гістограми перші два рядки таблиці розподілу вже задані. У цьому випадку треба підрахувати елементи третього рядка таблиці і побудувати відповідну гістограму.

б. Контрольні виміри радіуса 200 циліндрів дали наступні результати.

Таблиця 10.4

7,39	7,43	7,54	7,64	7,4	7,55	7,4	7,26	7,42	7,5
7,32	7,31	7,28	7,52	7,46	7,63	7,38	7,44	7,52	7,53
7,37	7,33	7,24	7,13	7,53	7,53	7,39	7,57	7,51	7,34
7,39	7,47	7,51	7,48	7,62	7,58	7,57	7,33	7,51	7,4
7,3	7,48	7,4	7,57	7,51	7,4	7,52	7,56	7,4	7,34
7,23	7,37	7,48	7,48	7,62	7,35	7,36	7,4	7,45	7,29
7,48	7,58	7,44	7,56	7,28	7,59	7,47	7,62	7,54	7,2
7,38	7,43	7,35	7,56	7,51	7,47	7,4	7,29	7,2	7,46
7,42	7,44	7,41	7,29	7,48	7,39	7,5	7,38	7,45	7,5
7,45	7,42	7,29	7,53	7,34	7,55	7,33	7,32	7,69	7,46
7,32	7,46	7,32	7,48	7,38	7,43	7,51	7,43	7,6	7,44
7,25	7,29	7,31	7,45	7,43	7,44	7,31	7,58	7,28	7,24
7,34	7,49	7,5	7,38	7,48	7,43	7,37	7,29	7,54	7,33
7,36	7,46	7,23	7,44	7,38	7,27	7,52	7,4	7,26	7,66
7,59	7,48	7,46	7,4	7,24	7,41	7,34	7,43	7,38	7,5
7,26	7,43	7,37	7,55	7,42	7,41	7,38	7,14	7,42	7,52
7,46	7,39	7,35	7,32	7,18	7,3	7,54	7,38	7,37	7,34
7,5	7,61	7,42	7,32	7,35	7,4	7,57	7,31	7,4	7,36
7,28	7,58	7,38	7,58	7,26	7,37	7,28	7,39	7,32	7,2
7,43	7,34	7,45	7,33	7,41	7,33	7,45	7,31	7,45	7,39

Побудувати гістограму розподілу частот вибірки та емпіричну функцію розподілу.

Розв'язування

Перенесемо таблицю 10.4 на лист MS Excel у діапазон A1:J20. За допомогою функції МАКС і МИН знайдем $R_{\max} = 7,69$ і $R_{\min} = 7,13$. Розіб'ємо діапазон на відрізки довжиною $h = 0,25$:

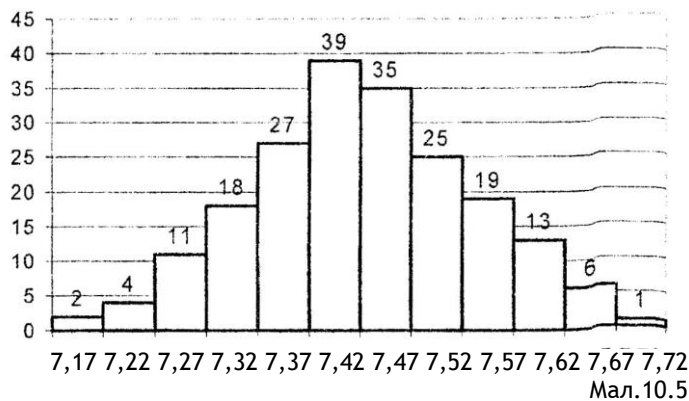
[7,12; 7,17], [7,17; 7,22], [7,22; 7,27], [7,27; 7,32], [7,32; 7,37], [7,37; 7,42], [7,42; 7,47], [7,47; 7,52], [7,52; 7,57], [7,57; 7,62], [7,62; 7,67], [7,67; 7,72].

Далі згрупуємо результати вимірів на цих відрізках. Для цього у стовпчик L1—L12 запишемо праві границі відповідних відрізків. Виділимо стовпчик M1—M12 і введемо функцію масиву ЧАСТОТА(A1:J20;L1:L12) і після натискання клавіш Ctrl+Shift+Enter у стовпчику M1—M12 отримаємо результати з другого рядка таблиці 10.5.

Таблиця 10.5

7,17	7,22	7,27	7,32	7,37	7,42	7,47	7,52	7,57	7,62	7,67	7,72
2	4	11	18	27	39	35	25	19	13	6	1

Відповідна гістограма зображена на малюнку 10.5



Зауваження. Часто розподіл варіант на відрізки доцільно починати з варіанти, яка найближча до середнього арифметичного значення варіант. У цій задачі середнє арифметичне дорівнює 7,41645, тому відрізки розподілу треба будувати так, щоб одна із середин цих відрізків була точкою 7,42 — одне із значень варіант.

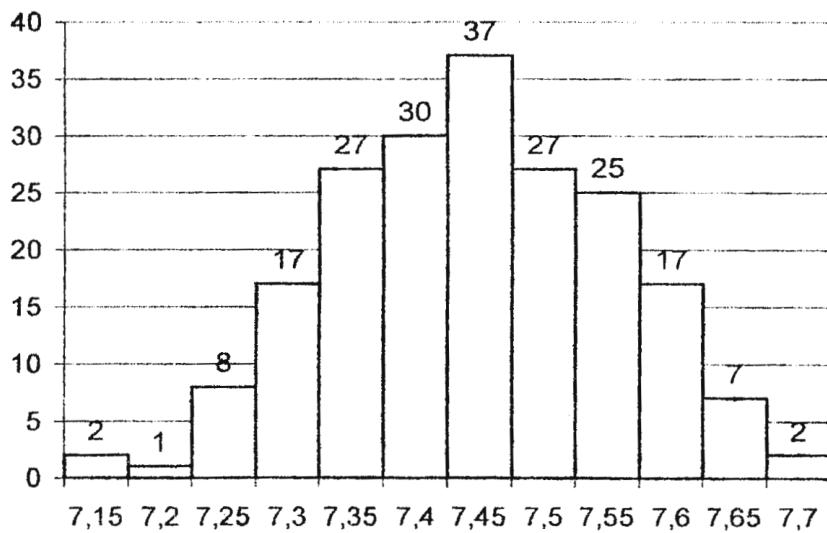
З огляду на це зауваження розіб'ємо діапазон на інші відрізки довжиною $h = 0,25$ (середина сьомого відрізка дорівнює 7,42):

[7,095; 7,145], [7,145; 7,195], [7,195; 7,245], [7,245; 7,295], [7,295; 7,345], [7,345; 7,395], [7,395; 7,445], [7,445; 7,495], [7,495; 7,545], [7,545; 7,595], [7,595; 7,645], [7,645; 7,695].

Відповідні частоти і гістограми наведені у таблиці 10.6 і малюнку 10.6.

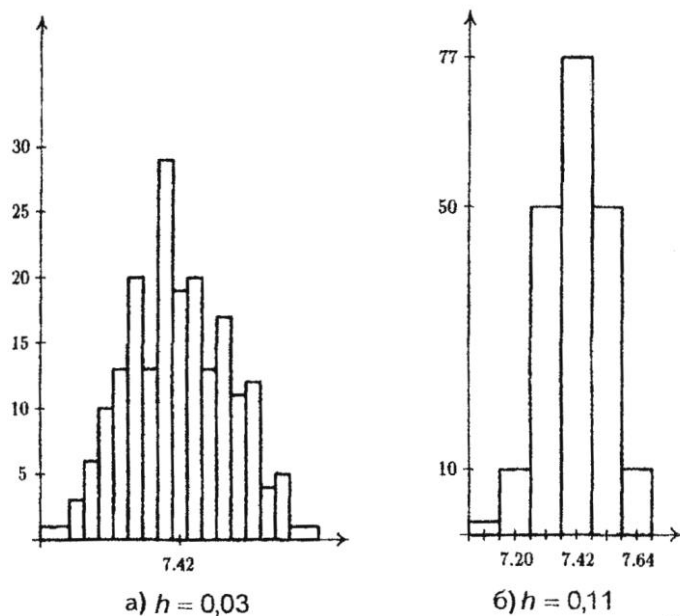
Таблиця 10.6

7,145	7,195	7,245	7,295	7,345	7,395	7,445	7,495	7,545	7,595	7,645	7,695
2	1	8	17	27	30	37	27	25	17	7	2



Мал.10.6

Зауваження. Від способу вибору довжини відрізків (числа h) залежить виразність гістограми. Якщо h досить мале, то гістограма містить багато випадкового. Якщо h досить велике, то в гістограмі зникають індивідуальні особливості вибірки. На малюнку 10.7 зображені гістограми для прикладу 6 при різних h .



Мал.10.7

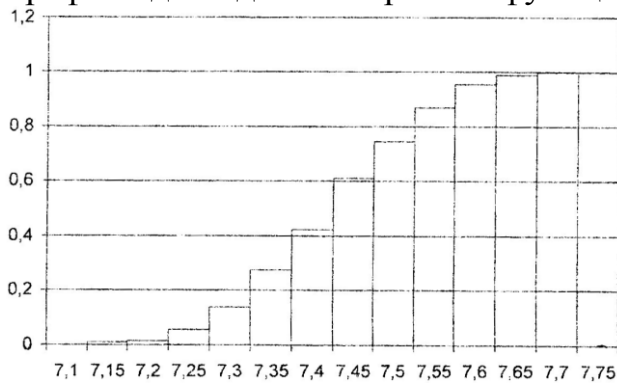
Зауваження. Вигляд гістограми дає уявлення про щільність ймовірностей випадкової величини, яку досліджують. Гістограма — це наближений графік щільності ймовірностей випадкової величини, тобто наближений графік диференціальної функції її розподілу.

Для побудови функції розподілу результати другого рядка табл. 10.6 будемо послідовно додавати, а потім розділимо на загальну кількість вимірювань (200), результати наведені у табл. 10.7.

Таблиця 10.7

2	3	11	28	55	85	122	149	174	191	198	200
0,01	0,015	0,055	0,14	0,275	0,425	0,61	0,745	0,87	0,955	0,99	1

Графік відповідної емпіричної функції розподілу наведено на малюнку 10.8.



Мал.10.8

Завдання для самостійного розв'язування задач на статистичний розподіл.

1. Для вивчення попиту на закупівлю газет власник кіоску спостерігає щоденну кількість проданих газет протягом 100 днів (табл. 10.8):

Таблиця 10.8

2 •	10	15	18	21	23	26	29	32	37
3	10	15	18	21	23	26	29	32	38
4	11	15	18	21	23	26	29	33	38
5	12	16	19	21	24	26	29	33	39
5	12	16	19	22	24	27	30	34	40
6	12	16	19	22	24	27	30	34	40
7	13	17	19	22	24	27	31	35	42
8	14	17	20	22	25	28	31	35	43
9	14	17	20	22	25	28	31	36	46
9	14	17	20	23	25	28	32	37	49

Знайти статистичний розподіл цієї вибірки з кроками $h = 10$ і $h = 5$, та побудувати гістограму частот і емпіричну функцію розподілу.

2. З булочок, що випускає хлібопекарня, зроблено вибірку. Зважування булочок, які попали у вибірку, дало такі результати (у грамах):

Таблиця 10.9

100,3	99,2	100,1	101,2	103,2	98,2	99,2	101,5
102,7	99,7	98,8	100,7	102,4	101,3	99,3	100,4
100,7	101,2	98,3	100,1	98,2	99,4	97,6	101,2
100,4	98,6	101,6	98,9	100,6	100,3	97,2	99,7
100,2	99,8	99,6	99,7	100,7	99,5	98,7	99,9

Скласти статистичний розподіл ваги X булочки у вибірці з кроком $h = 0,5$ та побудувати гістограму частот і емпіричну функцію розподілу.

3. Скласти таблицю статистичного розподілу розміру X чоловічого взуття, яке продане магазином протягом дня: 39,

41, 40, 40, 43, 41, 44, 42, 40, 42, 41, 41, 43, 42, 39, 42, 43, 41,
42, 41, 38, 42, 42, 41, 40, 41, 43, 41, 39, 40.

