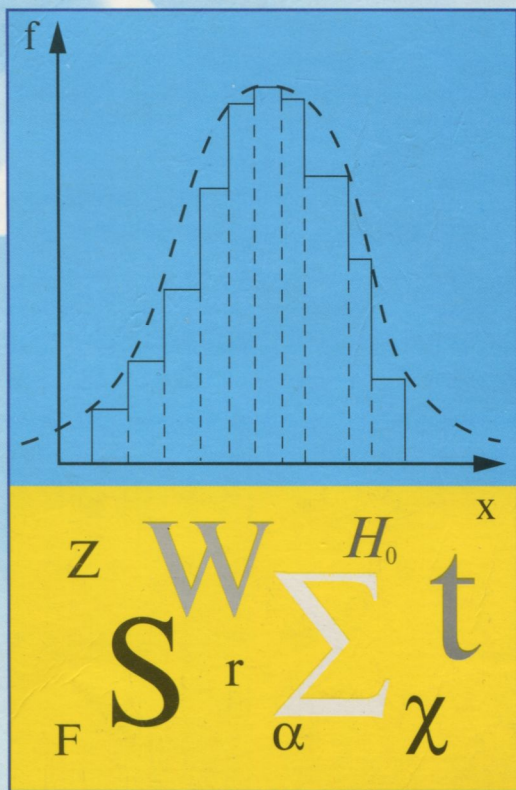


519.2
М34

М.А. Мартиненко
О.М. Нецадим
О.І. Радзієвська
В.М. Сафонов

МАТЕМАТИЧНА СТАТИСТИКА



Гриф надано Міністерством освіти і науки України (Лист від 26.08.2005 р.
№ 14/18.2-1187)

Рецензенти:

О.М. Станжицький, доктор фіз.-мат. наук, професор Національного університету ім. Т.Г. Шевченка;

Т.І. Олешко, професор Національного авіаційного університету

Мартиненко М.А., Нещадим О.М., Радзівська О.І., Сафонов В.М.

Математична статистика: Навчальний посібник. – К.: Четверта
хвиля, 2005. – 208 с.

ISBN 966-529-117-3

Розглядаються основи математичної статистики згідно з програмою дисципліни “Теорія ймовірностей і математична статистика”. Теоретичний матеріал подається у формі лекцій і супроводжується прикладами типових задач із економічної практики. Наприкінці кожної лекції наведено перелік основних запитань для самоконтролю знань і достатню кількість завдань для проведення практичних занять. Багато задач подано із детальними розв’язками. У посібнику вміщено завдання для індивідуальної роботи студентів та модульного контролю. Розрахований на студентів економічних спеціальностей вищих навчальних закладів усіх форм навчання.

ББК 966-529-117-3

ISBN 966-529-117-3

© Мартиненко М.А., Нещадим О.М.,
Радзівська О.І., Сафонов В.М., 2005
© Четверта хвиля, 2005

ПЕРЕДМОВА

В основу навчального посібника покладено багаторічний досвід авторів, якого вони набули при викладанні дисципліни “Теорія ймовірностей і математична статистика” в Національному університеті харчових технологій та інших навчальних закладах України. Цей посібник відповідає сучасним вимогам навчальної програми дисципліни для підготовки фахівців-економістів. За задумом авторів посібник має сприяти підвищенню рівня фундаментальної математичної підготовки студентів, посиленню прикладного економічного спрямування набутих знань.

Матеріал посібника подається у вигляді лекцій та практичних занять і охоплює основні теми розділу “Математична статистика”, а саме: вибірки та їх представлення, числові характеристики статистичного розподілу вибірки, статистичне оцінювання параметрів генеральної сукупності, перевірка достовірності статистичних гіпотез, елементи дисперсійного аналізу, основи кореляційного і регресійного аналізу. При написанні посібника автори намагалися досягти максимальної доступності матеріалу для студентів, зберігаючи необхідний рівень математичної строгості. Успішному засвоєнню теоретичного матеріалу має сприяти велика кількість розглянутих прикладів. Значну кількість задач запропоновано в кінці кожної лекції для проведення практичних занять та самостійного розв’язування. Необхідні для розв’язання задач основні таблиці математичної статистики наведено у додатках.

У кінці посібника представлені індивідуальні контрольні завдання та розгорнутий предметний покажчик основних термінів.

Автори

ВСТУП

Економічні і соціально-економічні процеси перебувають під дією різноманітних факторів, яким притаманний елемент випадковості. У значній кількості випадків закономірності можуть бути виявлені при цілеспрямованому статистичному дослідженні масових явищ або процесів, що включає збір даних, їх систематизацію і впорядкування та статистичний аналіз. Оскільки на досліджуване явище впливають багато факторів, то кожен окремий прояв цього явища буде відмінним від іншого.

При достатньо великій кількості спостережень випадкові дії значною мірою нейтралізуються, і одержаний результат стає практично не випадковим, передбачуваним. Цей принцип лежить в основі практичного використання ймовірносних і математико-статистичних методів дослідження.

Основним змістом математичної статистики є вивчення методів збору, систематизації обробки результатів спостережень з метою виявлення статистичних закономірностей певної сукупності. Математична статистика опирається на теорію ймовірностей.

Оскільки суцільне спостереження за елементами всієї (генеральної) сукупності практично неможливе, то методи математичної статистики ґрунтуються на дослідженні обмеженої кількості елементів сукупності (вибірок). Опираючись на результати теорії ймовірностей, математична статистика дає можливість не тільки визначити значення шуканих характеристик, але й оцінити точність одержаних за вибіркою висновків.

Застосування методів математичної статистики в економіці дає змогу будувати економіко-математичні моделі і оцінювати їхні параметри, перевіряти твердження (гіпотези) про властивості економічних характеристик і форми їх взаємозв'язку. Одержані результати є основою для економічного аналізу і прогнозування, дають можливість приймати обґрунтовані економічні рішення в умовах невизначеності.

Лекція 1. Вибірki та їх представлення

Предмет і методи математичної статистики.

Генеральна і вибіркові сукупності.

Статистичний розподіл вибірки.

Емпірична функція розподілу.

Графічне зображення статистичних розподілів.

1.1. Предмет і методи математичної статистики

Математична статистика займається встановленням закономірностей масових явищ і процесів на основі результатів спостережень або експериментів над ними. Для виявлення таких закономірностей використовуються математичні методи:

- збирання і систематизації статистичних даних;
- наукового аналізу цих даних.

Математична статистика опирається на теорію ймовірностей. Застосування математичної статистики в економіці дозволяє будувати економічні моделі і оцінювати їхні параметри, перевіряти твердження (гіпотези) про властивості економічних показників і форми їх взаємозв'язку. Одержані результати є основою для економічного аналізу і прогнозування, дають можливість приймати обґрунтовані економічні рішення у бізнесі в умовах невизначеності.

1.2. Генеральна і вибіркові сукупності

Нехай потрібно дослідити сукупність однорідних об'єктів відносно деякої якісної чи кількісної ознаки. Якісні ознаки характеризують деякі властивості або стан об'єкта, наприклад, стать, професія, якість продукції тощо. Кількісні ознаки одержують в результаті вимірювання або спостереження і виражаються числами, наприклад, маса, об'єм, прибуток тощо. Кількісні ознаки позначаються через X , Y , Z , ...; вони можуть бути дискретними або неперервними, тобто набувати будь-яких числових значень в деяких межах.

Можливі два способи дослідження ознаки сукупності: суцільне спостереження і вибіркове спостереження. При суцільному спостереженні вивчається кожний об'єкт сукупності, а при вибірковому – деякої частини, відібраної із всієї сукупності. При цьому вся сукупність об'єктів, яка підлягає дослідженню, називається *генеральною*, а та її частина, яка попала на перевірку або дослідження, *вибірковою*.

вою сукупністю або вибіркою. Число об'єктів у генеральній сукупності і у вибірці називається їх *обсягами*. За результатами вивчення ознаки у вибірці робиться висновок про властивості генеральної сукупності.

Дослідження вибіркової сукупності дає такі переваги:

- 1) практичність (із-за великого обсягу неможливо охопити всю генеральну сукупність);
- 2) виграш у часі;
- 3) зменшення затрат праці;
- 4) достатньо високу достовірність;
- 5) економія продукції.

Інформація про генеральну сукупність, одержана на основі вибірки, завжди буде мати деяку похибку, оскільки ґрунтується на вивченні тільки частини об'єктів. Тому дуже важливо так організувати вибіркоче спостереження, щоб згадана інформація була найбільш повною і давала правильне уявлення про генеральну сукупність. Неправильно сформована вибірка дасть викривлене уявлення про сукупність.

Розрізняють два способи відбору елементів генеральної сукупності у вибірку: *випадковий* і *невипадковий*. Випадковий відбір означає, що кожен об'єкт сукупності має рівний шанс потрапити до вибірки. Якщо сукупність містить групу об'єктів, то випадковий відбір повинен забезпечити представництво у вибірці об'єктів кожної із цих груп. Такий відбір називають *типовим*. Наприклад, продукція виробляється різними цехами, на різних конвеєрах. Тому при вивченні якості виробів підприємства доцільно робити вибірку із продукції кожного цеху і кожного конвеєра. Так, при наявності трьох цехів і двох ліній у кожному, продукцію доцільно розбити на шість груп.

Випадковий відбір, як правило, проводять за допомогою спеціальних таблиць випадкових чисел. Для цього елементи сукупності нумеруються і з таблиці випадкових чисел, відкритої навмання на довільній сторінці, виписуються підряд номери елементів, які повинні ввійти у вибірку. Останнім часом використовують комп'ютер, який генерує випадкові числа.

Серед невідповідних вибірок виділяють два види вибірок: *серійний* і *механічний*. Серійним називають відбір, при якому об'єкти вибираються із генеральної сукупності не по одному, а серіями, на

які попередньо поділена сукупність. При механічному відборі елементи генеральної сукупності відбираються за наперед встановленим правилом. Наприклад, до 5%-вої вибірки входить кожен двадцятий член генеральної сукупності.

Зауважимо, що при дослідженні таких вибірок не завжди можна застосувати ймовірнісні методи.

1.3. Статистичний розподіл вибірки

Нехай X – випадкова величина, яка уособлює певну кількісну ознаку об'єктів генеральної сукупності. Із цієї сукупності одержано вибірку ознаки $X : X_1, X_2, \dots, X_n$, де $X_i, (i = \overline{1, n})$ – незалежні однаково розподілені випадкові величини, розподіл кожної із яких співпадає із розподілом X . Різні значення, які приймає кількісна ознака, наприклад, $X = x_i, (i = \overline{1, k})$ називаються *варіантами* або *реалізаціями*. Послідовність варіант, розміщених у зростаючому порядку

$$x_1 < x_2 < \dots < x_k$$

називають *варіаційним рядом*. Якщо варіанта x_i зустрічається n_i раз у вибірці, то число n_i називається *частотою* варіанти $x_i, (i = \overline{1, k})$. При цьому $\sum_{i=1}^k n_i = n$ – обсяг вибірки. Відношення $\frac{n_i}{n} = w_i$

називається *відносною частотою* варіанти, причому $\sum_{i=1}^k w_i = 1$. В залежності від того, яких значень можуть набувати варіанти, варіаційні ряди поділяються на *дискретні* та *неперервні (інтервальні)*.

Означення. Перелік варіант варіаційного ряду і відповідних їм частот або відносних частот називають *статистичним розподілом* вибірки:

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

x_i	x_1	x_2	...	x_k
w_i	w_1	w_2	...	w_k

Приклад 1.1. Проведено обстеження величини річного продукту (млн. грн.) на десяти підприємствах харчової промисловості: 2, 5, 1, 2, 2, 5, 1, 2, 1, 2.

Побудувати статистичний розподіл вибірки.

Розв'язання. Досліджується ознака X – величина річного прибутку (млн. грн.) на підприємстві харчової промисловості. Маємо варіаційний ряд: $x_1 = 1, x_2 = 2, x_3 = 5$. Будуємо статистичні розподіли частот

x_i	1	2	5
n_i	3	5	2

$$n = \sum_{i=1}^3 n_i = 3 + 5 + 2 = 10;$$

і відносних частот

x_i	1	2	5
w_i	0,3	0,5	0,2

$$n = \sum_{i=1}^3 w_i = 1.$$

Для неперервного варіаційного ряду статистичний розподіл задається таблицею

Інтервал	$[\alpha_0; \alpha_1)$	$[\alpha_1; \alpha_2)$...	$[\alpha_{i-1}; \alpha_i)$
n_i	n_1	n_2	...	n_k
w_i	w_1	w_2	...	w_k

тут n_i, w_i – частота і відносна частота появи ознаки X в інтервалі $[\alpha_{i-1}; \alpha_i)$. Довжину $h_i = \alpha_i - \alpha_{i-1}$ цього інтервалу називають *кроком*.

При інтервальному статистичному розподілі доцільно використовувати поняття щільності частоти $f_i = \frac{n_i}{h_i}$ або щільності відносної частоти $g_i = \frac{w_i}{h_i}$, які характеризують число елементів або відносне

їх число, яке припадає на одиницю довжини h_i інтервалу.

Для зручності розрахунків за звичай крок обирається однаковим для даної вибірки. На практиці число інтервалів визначається за формулами: $k = [1 + 3,32 \lg n]$ (Стерджес) або $k = [5 \ln n]$ (Брукс-Карузертс), або $k = [\sqrt{n}]$, тут $[a]$ – ціла частина числа a .

Шляхом заміни $x_i = \frac{\alpha_{i-1} + \alpha_i}{2}$, ($i = \overline{1, k}$) можна перейти від інтервального статистичного розподілу до дискретного.

1.4. Емпірична функція розподілу

Нехай досліджується кількісна ознака X генеральної сукупності. Розглянемо варіаційний ряд x_1, x_2, \dots, x_k вибірки обсягу n , для якого відомий статистичний розподіл. Однією з характеристик вибірки є *емпірична функція розподілу* $F_n(x)$, яка визначає для кожного значення x відносну частоту події $X < x$:

$$F_n(x) = W(X < x) = \frac{\mu(x)}{n} \quad (1.1)$$

де $\mu(x)$ – кількість варіант вибірки, менших за варіанту $X = x$.

На відміну від емпіричної функції розподілу вибірки, функцію розподілу $F(x)$ генеральної сукупності називають теоретичною функцією розподілу. Відмінність між цими функціями полягає в тому, що $F(x)$ визначає ймовірність події $X < x$, а $F_n(x)$ – відносну частоту цієї ж події. При великих значеннях n $F_n(x)$ мало відрізняється від $F(x)$.

Приклад 1.2. Побудувати емпіричну функцію розподілу за даним статистичним розподілом вибірки:

X	2	5	9
n_i	12	18	30

Розв'язання. Визначаємо обсяг вибірки: $n = 12 + 18 + 30 = 60$.

1. Оскільки $x_1 = 2$ – найменша із варіант, то при $x \leq 2$ $\mu(x) = 0$ і $F_n(x) = 0$.

2. Значення випадкової величини $X < 5$ зокрема $x_1 = 2$ спостерігалось 12 разів, тому $\mu(x) = 12$ і $F_n(x) = \frac{12}{60} = 0,2$ при $2 < x \leq 5$.

3. Якщо x набуває значення в інтервалі $5 < x \leq 9$, то нерівність $X < x$ виконується для варіант $X = 2, X = 5$ і $\mu(x) = 12 + 18 = 30$, звідки $F_n(x) = \frac{30}{60} = 0,5$.

4. Оскільки $X = 9$ – найбільша варіанта, то $F_n(x) = 1$ при $x > 9$.

Отже, маємо емпіричну функцію

$$F_n(x) = \begin{cases} 0, & \text{при } x \leq 2; \\ 0,2, & \text{при } 2 < x \leq 5; \\ 0,5, & \text{при } 5 < x \leq 9; \\ 1, & \text{при } x > 9 \end{cases}$$

і її графік (рис. 1.1).

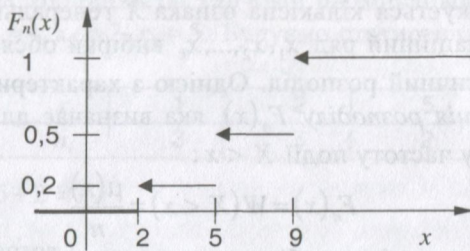


Рис. 1.1.

1.5. Графічне зображення статистичних розподілів

Графічна ілюстрація статистичних даних надає їм наглядності і часто приводить до спрощення аналізу цих даних. В залежності від вигляду варіаційного ряду і поставленої задачі будують різні графіки статистичних розподілів, зокрема *полігон*, *гістограму*, *кумулятивну криву*.

Розглянемо дискретний статистичний розподіл вибірки. Полігоном частот (або відносних частот) називають ламану лінію, відрізки якої з'єднують точки з координатами (x_i, n_i) (або (x_i, w_i)), $i = \overline{1, k}$.

У випадку неперервного статистичного розподілу гістограмою частот (або відносних частот) називають східчасту фігуру у вигляді послідовності прямокутників з основами $h_i = \alpha_i - \alpha_{i-1}$ і висотами f_i (або g_i). Площа S_i i -го прямокутника $S_i = h_i \cdot f_i = n_i$. Звідси, площа

гістограми частот дорівнює обсягу вибірки: $\sum_{i=1}^k S_i = n$. Аналогічно,

для гістограми відносних частот маємо $S_i = h_i \cdot g_i = w_i$ і $\sum_{i=1}^k S_i = n$.

Кумулятою дискретного статистичного розподілу називають графік функції $F_n(x)$. Кумулятивна крива неперервного статистичного розподілу – ламана лінія, яка сполучає точки з координатами $(\alpha_i; F_n(\alpha_i))$, $i = \overline{0, k}$.

Приклад 1.3. На кондитерській фабриці автомат розливає рідкий шоколад у форми для одержання шоколадних плиток. Для контролю навмання відібрано 20 плиток із партії готової продукції. Результати обстеження наведені в таблиці:

Маса, г	96-98	98-100	100-102	102-104
К-ть плиток	4	8	6	2

Побудувати: 1) полігон і гістограму відносних частот; 2) кумулятивну криву.

Розв'язання. Для кількісної ознаки X – маса шоколаду у формах, маємо неперервний статистичний розподіл.

Обчислимо відносні частоти $w_i = \frac{n_i}{n}$ появи ознаки X на відповідних інтервалах; маємо ряд розподілу:

Інтервал	96-98	98-100	100-102	102-104
w_i	0,2	0,4	0,3	0,1

Довжина кожного інтервалу $h_i = \alpha_i - \alpha_{i-1} = 2$, $i = \overline{1,4}$. Визначаємо щільності g_i : $g_1 = \frac{0,2}{2} = 0,1$; $g_2 = \frac{0,4}{2} = 0,2$; $g_3 = \frac{0,3}{2} = 0,15$; $g_4 = \frac{0,1}{2} = 0,05$.

Гістограма відносних частот має вигляд (рис. 1.2).

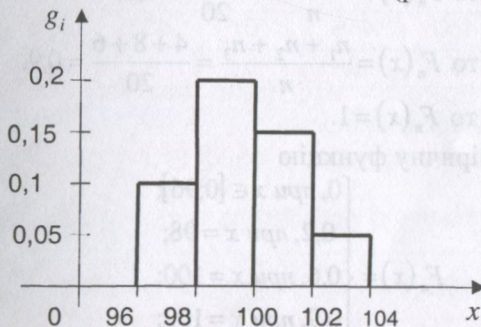


Рис. 1.2.

Перейдемо від інтервального до дискретного статистичного розподілу відносних частот

X	97	99	101	103
w_i	0,2	0,4	0,3	0,1

Побудуємо полігон відносних частот (рис. 1.3).

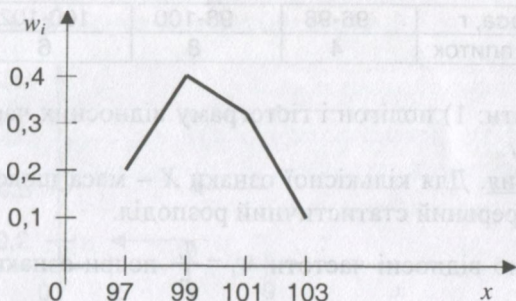


Рис. 1.3.

а) Якщо $x \in [0; 96]$, то $F_n(x) = \frac{\mu(x)}{n} = 0$, оскільки $\mu(x) = 0$ – кількість плиток, маса яких менша за x , дорівнює нулю.

б) Якщо $x = 98$, то $F_n(x) = \frac{n_1}{n} = \frac{4}{20} = 0,2$.

в) Якщо $x = 100$, то $F_n(x) = \frac{n_1 + n_2}{n} = \frac{4 + 8}{20} = 0,6$.

г) Якщо $x = 102$, то $F_n(x) = \frac{n_1 + n_2 + n_3}{n} = \frac{4 + 8 + 6}{20} = 0,9$.

д) Якщо $x \geq 104$, то $F_n(x) = 1$.

Отже, маємо емпіричну функцію

$$F_n(x) = \begin{cases} 0, & \text{при } x \in [0; 96]; \\ 0,2, & \text{при } x = 98; \\ 0,6, & \text{при } x = 100; \\ 0,9, & \text{при } x = 102; \\ 1, & \text{при } x \geq 104. \end{cases}$$

Її графіком буде кумулятивна крива (рис. 1.4).

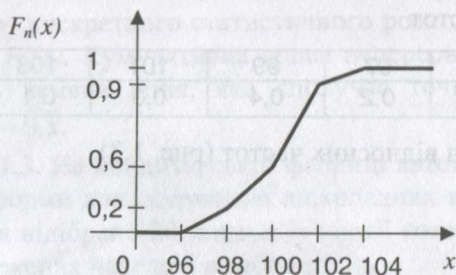


Рис. 1.4.

Запитання для самоконтролю

1. Дайте поняття генеральної та вибіркової сукупностей.
2. Які переваги дає дослідження вибіркової сукупності?
3. Що називають варіантами, варіаційним рядом?
4. Що таке частота, відносна частота варіантів?
5. Дайте означення дискретного статистичного розподілу вибірки.
6. Який статистичний розподіл вибірки називається інтервальним?
7. Як можна визначити кількість інтервалів для неперервного статистичного розподілу вибірки?
8. Що таке полігон частот, відносних частот?
9. Що називають гістограмою частот, відносних частот?
10. Що називається емпіричною функцією розподілу?

x_i	450	455	461	470	475
n_i	2	8	6	4	2

x_i	450	455	461	470	475
w_i	0,2	0,32	0,24	0,16	0,08

Практичне заняття № 1

Статистичний розподіл вибірки та його графічне зображення

№ 1.1. Результати вибіркового обстеження рівня заробітної плати (у гривнях) працівників одного підрозділу державного підприємства становлять: 450, 461, 450, 456, 470, 450, 456, 470, 456, 461, 475, 456, 461, 456, 470, 450, 456, 475, 450, 461, 456, 461, 470, 456, 461. Побудувати за цими даними дискретний статистичний розподіл частот і відносних частот.

Розв'язання. Досліджувана ознака X – рівень заробітної плати (у гривнях) працівників державного підприємства. Тоді варіаційний ряд має вигляд: $x_1 = 450$, $x_2 = 456$, $x_3 = 461$, $x_4 = 470$, $x_5 = 475$. Значення $x_1 = 450$ спостерігалось 5 разів, тобто частота варіанти x_1 складає $n_1 = 5$. Так само визначаємо частоти $n_2 = 8$, $n_3 = 6$, $n_4 = 4$ і $n_5 = 2$ для відповідних варіант x_2 , x_3 , x_4 , x_5 . Будуємо дискретний статистичний розподіл частот:

x_i	450	456	461	470	475
n_i	5	8	6	4	2

При цьому маємо обсяг вибірки

$$n = n_1 + n_2 + n_3 + n_4 + n_5 = 5 + 8 + 6 + 4 + 2 = 25.$$

Далі відшукаємо відносні частоти: $w_1 = \frac{n_1}{n} = \frac{5}{25} = 0,2$;

$$w_2 = \frac{n_2}{n} = \frac{8}{25} = 0,32; w_3 = \frac{n_3}{n} = \frac{6}{25} = 0,24; w_4 = \frac{n_4}{n} = \frac{4}{25} = 0,16;$$

$w_5 = \frac{n_5}{n} = \frac{2}{25} = 0,08$. Таким чином, приходимо до розподілу відносних частот:

x_i	450	456	461	470	475
w_i	0,2	0,32	0,24	0,16	0,08

№ 1.2. Результати дослідження річного обсягу споживання рибної продукції (у кг на душу населення) в регіонах складають: 11,5; 9,5; 12; 10; 14; 11; 10; 9; 10,5; 9,5; 10,5; 11,5; 10,5; 12; 10,5; 10; 9; 10; 12,5; 11; 10,5; 9; 12; 14; 12,5. Побудувати за цими даними інтерва-

льний статистичний розподіл частот і відносних частот з кроком $h=1$.

Розв'язання. Тут досліджувана ознака X – річний обсяг споживання рибної продукції (у кг) на душу населення. Оскільки найменше значення ознаки становить $X=9$, а найбільше її значення є $X=14$, то, враховуючи що $h=1$, маємо інтервальний варіаційний ряд: [9; 10), [10; 11), [11; 12), [12; 13), [13; 14].

Видно, що інтервалу [9; 10) належать значення ознаки: 9 і 9,5, які відповідно спостерігались 3 і 2 рази. Отже, частота $n_1 = 3 + 2 = 5$. Аналогічно знаходимо частоти $n_2 = 4 + 5 = 9$, $n_3 = 2 + 2 = 4$, $n_4 = 3 + 2 = 5$, $n_5 = 2$. Дістаємо інтервальний статистичний розподіл частот:

Інтервал [α_{i-1} ; α_i)	9–10	10–11	11–12	12–13	13–14
n_i	5	9	4	5	2

Обчислюємо відносні частоти $w_i = \frac{n_i}{n}$ появи ознаки X на відповідних інтервалах:

$$w_1 = \frac{5}{5+9+4+5+2} = \frac{5}{25} = 0,2; w_2 = \frac{9}{25} = 0,36; w_3 = \frac{4}{25} = 0,16;$$

$$w_4 = \frac{5}{25} = 0,2; w_5 = \frac{2}{25} = 0,08.$$

Тоді маємо такий розподіл відносних частот:

Інтервал [α_{i-1} ; α_i)	9–10	10–11	11–12	12–13	13–14
w_i	0,2	0,36	0,16	0,2	0,08

№ 1.3. Наведено дані про виробничий стаж (в роках) працівників цеху:

4	1	10	9	7	2	8	5
9	12	5	2	8	9	7	13
11	15	8	7	12	5	3	9
19	7	17	6	8	15	12	1
13	6	14	10	5	11	18	4

Побудувати за цими даними інтервальний статистичний розподіл частот і відносних частот, визначивши кількість інтервалів за формулою Стерджеса.

Розв'язання. Досліджується ознака X – виробничий стаж працівників цеху. Для обсягу вибірки $n = 40$ за формулою Стерджеса визначаємо кількість інтервалів: $k = [1 + 3,32 \lg 40] = [6,32] = 6$. Оскільки $x_{\min} = 1$ – найменше значення ознаки, а $x_{\max} = 19$ – найбільше, то знаходимо величину кроку: $h = \frac{x_{\max} - x_{\min}}{k} = \frac{19 - 1}{6} = \frac{18}{6} = 3$. Маємо інтервальний варіаційний ряд: [1;4), [4;7), [7;10), [10;13), [13;16), [16;19]. Підраховуємо частоти $n_i, i = \overline{1,6}$ і будуюмо інтервальний статистичний розподіл:

Інтервал $[\alpha_{i-1}; \alpha_i)$	1-4	4-7	7-10	10-13	13-16	16-19
n_i	5	8	12	7	5	3

Обчислюємо відносні частоти w_i :

$$w_1 = \frac{5}{40} = 0,125; w_2 = \frac{8}{40} = 0,2; w_3 = \frac{12}{40} = 0,3; w_4 = \frac{7}{40} = 0,175;$$

$$w_5 = \frac{5}{40} = 0,125; w_6 = \frac{3}{40} = 0,075$$

Отже, статистичний розподіл відносних частот має вигляд:

Інтервал $[\alpha_{i-1}; \alpha_i)$	1-4	4-7	7-10	10-13	13-16	16-19
w_i	0,125	0,2	0,3	0,175	0,125	0,075

№ 1.4. Задана вибірка у вигляді розподілу частот:

x_i	2	5	6	8	10
n_i	10	15	25	30	20

- Потрібно: 1) знайти розподіл відносних частот;
2) накреслити полігон частот і відносних частот;
3) визначити емпіричну функцію і зобразити її графічно.

Розв'язання. 1. Обсяг вибірки становить $n = 10 + 15 + 25 + 30 + 20 = 100$.

Обчислюємо відносні частоти: $w_1 = \frac{n_1}{n} = \frac{10}{100} = 0,1; w_2 = \frac{n_2}{n} = \frac{15}{100} = 0,15;$

$$w_3 = \frac{n_3}{n} = \frac{25}{100} = 0,25; w_4 = \frac{n_4}{n} = \frac{30}{100} = 0,3; w_5 = \frac{n_5}{n} = \frac{20}{100} = 0,2.$$

Шуканий розподіл відносних частот має вигляд:

x_i	2	5	6	8	10
w_i	0,1	0,15	0,25	0,3	0,2

2. Відкладемо на осі абсцис варіанти x_i , на осі ординат – відповідні їм частоти n_i . Сполучивши точки $(x_i; n_i)$ відрізками прямих, матимемо полігон частот (рис. 1.5):

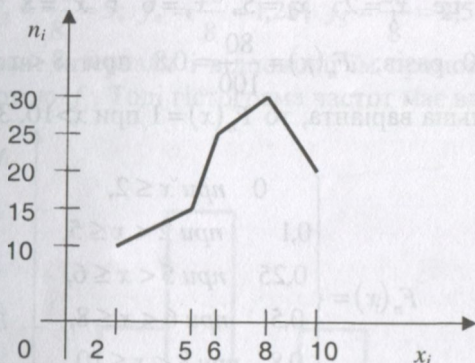


Рис. 1.5.

Відкладемо знов на осі абсцис варіанти x_i , а на осі ординат тепер – відповідні їм відносні частоти w_i . Сполучивши точки $(x_i; w_i)$ відрізками прямих, отримаємо полігон відносних частот (рис. 1.6):

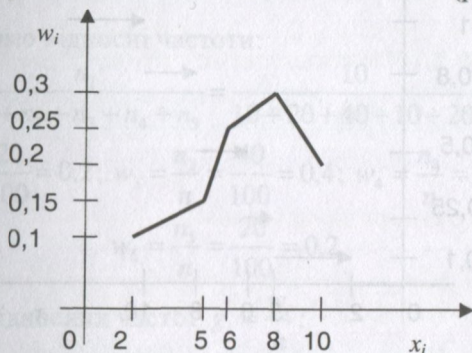


Рис. 1.6.

3. Скориставшись формулою (1.1), знайдемо емпіричну функцію розподілу. Найменша варіанта дорівнює двом. Отже, $F_n(x) = 0$ при $x \leq 2$. Значення $X < 5$, а саме $x_1 = 2$ спостерігалось 10 разів. Таким

чином, $F_n(x) = \frac{10}{100} = 0,1$ при $2 < x \leq 5$. Значення $X < 6$, тобто $x_1 = 2$ і

$x_2 = 5$ спостерігались $10+15=25$ разів. Тоді $F_n(x) = \frac{25}{100} = 0,25$ при $5 < x \leq 6$. Значення $X < 8$, тобто $x_1 = 2$, $x_2 = 5$ та $x_3 = 6$ спостерігались $10+15+25=50$ разів і $F_n(x) = \frac{50}{100} = 0,5$ при $6 < x \leq 8$. Нарешті, значення $X < 10$, а це $x_1 = 2$, $x_2 = 5$, $x_3 = 6$ і $x_4 = 8$ спостерігались $10+15+25+30=80$ разів; $F_n(x) = \frac{80}{100} = 0,8$ при $8 < x \leq 10$. Оскільки $x_5 = 10$ – найбільша варіанта, то $F_n(x) = 1$ при $x > 10$. Запишемо емпіричну функцію:

$$F_n(x) = \begin{cases} 0 & \text{при } x \leq 2, \\ 0,1 & \text{при } 2 < x \leq 5, \\ 0,25 & \text{при } 5 < x \leq 6, \\ 0,5 & \text{при } 6 < x \leq 8, \\ 0,8 & \text{при } 8 < x \leq 10, \\ 1 & \text{при } x > 10. \end{cases}$$

Побудуємо графік емпіричної функції (рис. 1.7):

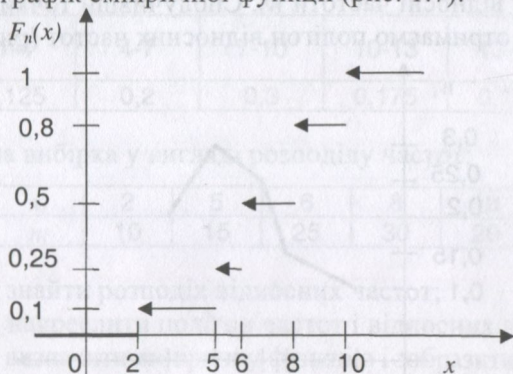


Рис. 1.7.

№ 1.5. За заданим інтервальним статистичним розподілом вибірки

Інтервал $[\alpha_{i-1}; \alpha_i)$	0–8	8–16	16–24	24–32	32–40
n_i	10	20	40	10	20

Потрібно: 1) побудувати гістограму частот і відносних частот;
2) знайти емпіричну функцію та зобразити її графічно.

Розв'язання. 1. Знайдемо довжини заданих інтервалів

$$h = \alpha_i - \alpha_{i-1} = 8 \quad \text{та} \quad \text{щільності частот} \quad f_i = \frac{n_i}{h}: \quad f_1 = \frac{10}{8} = 1,25;$$

$$f_2 = \frac{20}{8} = 2,5; \quad f_3 = \frac{40}{8} = 5; \quad f_4 = \frac{10}{8} = 1,25; \quad f_5 = \frac{20}{8} = 2,5.$$

Побудуємо на осі абсцис вказані інтервали і відповідні їм прямокутники з основою $h=8$ та висотою f_i . Тоді гістограма частот має вигляд (рис. 1.8):

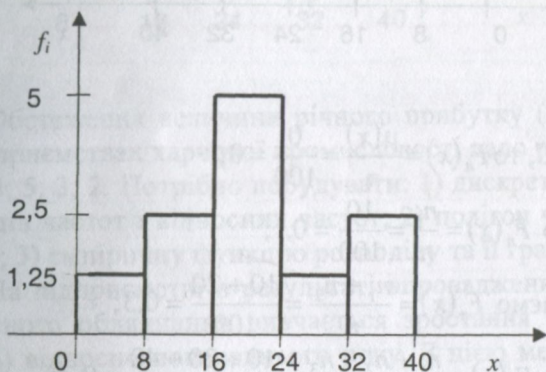


Рис. 1.8.

Тепер визначимо відносні частоти:

$$w_1 = \frac{n_1}{n_1 + n_2 + n_3 + n_4 + n_5} = \frac{10}{10 + 20 + 40 + 10 + 20} = 0,1;$$

$$w_2 = \frac{n_2}{n} = \frac{20}{100} = 0,2; \quad w_3 = \frac{n_3}{n} = \frac{40}{100} = 0,4; \quad w_4 = \frac{n_4}{n} = \frac{10}{100} = 0,1;$$

$$w_5 = \frac{n_5}{n} = \frac{20}{100} = 0,2$$

та щільності відносних частот $g_i = \frac{w_i}{h}$:

$$g_1 = \frac{0,1}{8} = 0,0125; \quad g_2 = \frac{0,2}{8} = 0,025; \quad g_3 = \frac{0,4}{8} = 0,05; \quad g_4 = \frac{0,1}{8} = 0,0125;$$

$$g_5 = \frac{0,2}{8} = 0,025.$$

Аналогічно тому, як це робилося для знаходження гістограми частот, будуємо прямокутники з основами $h=8$ і висотами g_i та дістаємо гістограму відносних частот (рис. 1.9):

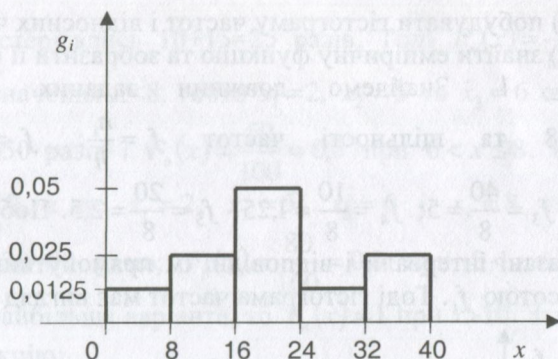


Рис. 1.9.

2. Якщо $x \leq 0$, то $F_n(x) = \frac{\mu(x)}{n} = \frac{0}{100} = 0$;

при $x=8$ маємо $F_n(x) = \frac{n_1}{n} = \frac{10}{100} = 0,1$;

при $x=16$ дістаємо $F_n(x) = \frac{n_1 + n_2}{n} = \frac{10 + 20}{100} = 0,3$;

при $x=24$ буде $F_n(x) = \frac{n_1 + n_2 + n_3}{n} = \frac{10 + 20 + 40}{100} = 0,7$;

якщо ж $x=32$ то $F_n(x) = \frac{n_1 + n_2 + n_3 + n_4}{n} = \frac{10 + 20 + 40 + 10}{100} = 0,8$ і, на-

решті, $F_n(x) = \frac{n_1 + n_2 + n_3 + n_4 + n_5}{n} = \frac{10 + 20 + 40 + 10 + 20}{100} = 1$ у разі,

коли $x \geq 40$.

Отже, будемо емпіричну функцію:

x	$x \leq 0$	8	16	24	32	$x \geq 40$
$F_n(x)$	0	0,1	0,3	0,7	0,8	1

та кумулятивну криву, з'єднаючи зазначені точки відрізками прямих (рис. 1.10):

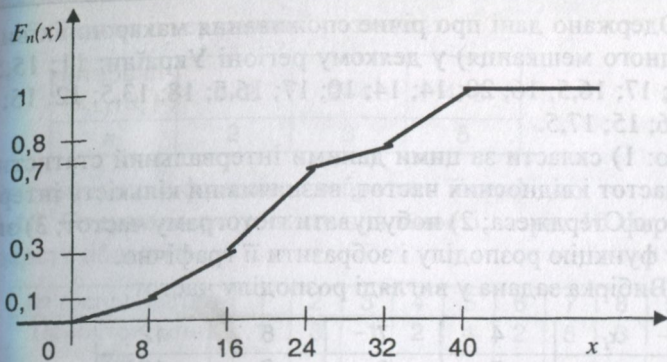


Рис. 1.10.

№ 1.6. Обстеження величини річного прибутку (в млн. грн.) на десяти підприємствах харчової промисловості дало результати: 3; 5; 3; 5; 3; 2; 3; 5; 3; 2. Потрібно побудувати: 1) дискретний статистичний розподіл частот і відносних частот; 2) полігон частот і відносних частот; 3) емпіричну функцію розподілу та її графік.

№ 1.7. На підприємстві в результаті впровадження у поточному році сучасного обладнання вивчається зростання продуктивності праці (в %) відносно попереднього року. З цією метою обстежено показники праці 20 робітників: 2; 3,8; 5,2; 0; 3,5; 4; 2,6; 1; 7,2; 2,8; 5; 1,5; 4,4; 5,8; 2,5; 6,5; 3,4; 0,5; 4,5; 3,6. Потрібно побудувати: 1) інтервальний розподіл частот і відносних частот з кроком $h=2$; 2) гістограму частот та відносних частот; 3) емпіричну функцію і кумулятивну криву.

№ 1.8. Наведені дані про продуктивність праці (в шт./год.) робітників підприємства:

12	17	15	24	10	5	17	14	19	20
9	13	18	14	20	18	12	22	16	17
16	19	7	19	16	22	15	20	8	23
20	26	21	11	20	13	19	9	29	15

Потрібно: 1) за цими даними побудувати інтервальний статистичний розподіл частот і відносних частот, приймаючи число інтервалів рівним $[\sqrt{n}]$; 2) побудувати гістограму відносних частот; 3) знайти емпіричну функцію розподілу і зобразити її графічно.

№ 1.9. Одержано дані про річне споживання макаронних виробів (в кг на одного мешканця) у деякому регіоні України: 11; 15,5; 19; 8; 13; 14,5; 17; 16,5; 16; 20; 14; 14; 10; 17; 16,5; 18; 13,5; 12; 15; 17,5; 20; 11,5; 16; 15; 17,5.

Потрібно: 1) скласти за цими даними інтервальний статистичний розподіл частот і відносних частот, визначивши кількість інтервалів за формулою Стерджеса; 2) побудувати гістограму частот; 3) знайти емпіричну функцію розподілу і зобразити її графічно.

№ 1.10. Вибірка задана у вигляді розподілу частот:

x_i	4	7	8	12
n_i	5	2	3	10

Знайти розподіл відносних частот та побудувати полігон частот і відносних частот.

№ 1.11. Задано розподіл частот:

x_i	2	5	7	8
n_i	1	3	2	4

Знайти емпіричну функцію та зобразити її графічно.

№ 1.12. Задано розподіл частот:

x_i	2	5	3	6
n_i	1	2	5	2

Побудувати полігон частот і відносних частот та відшукати емпіричну функцію і накреслити її графік.

№ 1.13. Вибірка задана у вигляді інтервального розподілу частот:

Інтервал $[\alpha_{i-1}; \alpha_i)$	2-6	6-10	10-14	14-18
n_i	6	10	4	5

Побудувати гістограму частот і відносних частот та знайти емпіричну функцію і накреслити кумулятивну криву.

№ 1.14. Задано інтервальний розподіл частот:

Інтервал $[\alpha_{i-1}; \alpha_i)$	0-5	5-10	10-15	15-20
n_i	3	7	4	6

Знайти інтервальний розподіл відносних частот і побудувати гістограми частот та відносних частот.

№ 1.15. Задано інтервальный розподіл:

Інтервал $[\alpha_{i-1}; \alpha_i)$	0–2	2–4	4–6
n_i	2	3	5

Знайти емпіричну функцію і зобразити її графічно.

№ 1.16. Одержано результати діяльності за рік 10 фермерських господарств області:

№ господарства	1	2	3	4	5	6	7	8	9	10
Прибуток, млн. грн.	1	3	-1	2	4	2	5	3	-1	4

Побудувати дискретний статистичний розподіл частот величини прибутку фермерських господарств та полігон частот і відносних частот. Знайти емпіричну функцію розподілу, накреслити її графік.

№ 1.17. Після шліфування деталі були піддані контрольним вимірюванням, результати яких наведено в таблиці:

x_i , мм	3,8	3,9	4	4,1	4,2
n_i	10	15	60	10	5

Побудувати полігон частот і відносних частот. Знайти емпіричну функцію розподілу і зобразити її графічно.

№ 1.18. Виміряно зріст у 100 осіб і результати наведені як інтервальный розподіл:

Інтервал, см $[\alpha_{i-1}; \alpha_i)$	168–172	172–176	176–180	180–184	184–188
n_i	10	30	40	15	5

Побудувати гістограму частот і відносних частот. Знайти емпіричну функцію розподілу та накреслити кумулятивну криву.