

ЛАБОРАТОРНА РОБОТА №2-1.**ПОБУДОВА ВИБІРКОВИХ ФУНКЦІЙ РОЗПОДІЛУ**

Мета: навчитися будувати частотні розподіли випадкових величин

ТЕОРЕТИЧНІ ВІДОМОСТІ

Розглянутий раніше розподіл ймовірностей випадкової величини (ВВ) опирється на знання закону розподілу ВВ. Для практичних задач таке знання – рідкість. Тут закон розподілу за звичай невідомий, або відомий із деякою точністю до деяких невідомих параметрів. Зокрема, неможливо розрахувати точне значення відповідних ймовірностей, оскільки неможливо визначити кількість спільних і позитивних результатів. Тому вводиться статистичне означення ймовірності. За цим означенням ймовірність рівна відношенню числа випробовувань з позитивним результатом до загального числа спроб у випробовуванні. Така ймовірність називається **статистичною частотою**.

Зв'язок між **емпіричною функцією розподілу** і функцією розподілу (теоретичною функцією розподілу) така ж, як зв'язок між частотою події і її ймовірністю.

Для побудови вибіркової функції розподілу весь діапазон зміни випадкової величини X (вибірки) розбивають на ряд інтервалів (кишень) однакової ширини. Число інтервалів за звичай вибирають не менше 3 і не більше 15. Потім визначають число значень випадкової величини X , що потрапляють в кожен інтервал (абсолютна частота, частота інтервалов).

Частота інтервалів – число, що показує скільки разів значення, що потрапляють в кожен інтервал групування, зустрічається у вибірці. Поділивши ці числа на загальну кількість спостережень (n), знаходять **відносну частоту (частість)** попадання випадкової величини X в задані інтервали.

За знайденими відносними частотами будують гістограми вибірових функцій розподілу. **Гістограма розподілу частот** – це графічне представлення вибірки, де по осі абсцис (OX) відкладені величини інтервалів, а по осі ординат (OY) – величини частот, що потрапляють в даний класовий інтервал. При збільшенні до безкінечності розміру вибірки вибірові функції розподілу перетворюються в теоретичні: гістограма перетворюється в графік густини розподілу.

Накопичена частота інтервалів – це число, одержане послідовним додаванням частот в напрямку від першого інтервала до останнього, до того інтервалу включно, для якого визначається накопичена частота.

В Excel для побудови вибірових функцій розподілу використовується спеціальна функція **ЧАСТОТА** і процедура **Гистограмма** із пакету аналізу.

Функція **ЧАСТОТА** (*массив данных, двоичный массив*)

обчислює частоти появи випадкової величини в інтервалах значень і виводить їх як масив цифр, де

- *массив_данных* — це масив або адреси множини даних, для яких обчислюються частоти;
- *двоичный_массив* — це масив інтервалів, за якими групуються значення вибірки.

Процедура **Гистограмма** із **Пакета анализа** виводить результати вибіркового розподілу у вигляді таблиці і графіку. Параметри діалогового вікна **Гистограмма**:

- **Входной диапазон** - діапазон досліджуваних даних (вибірка);
- **Интервал карманов** - діапазон комірок або набір граничних значень, що визначають вибрані інтервали («кишені»). Ці значення повині бути введені в порядку зростання. Якщо діапазон «кишень» не був введений, то набір інтервалів, рівномірно розподілених між мінімальним і максимальним значеннями даних, буде створено автоматично.

выходной диапазон призначений для введення зсилки на ліву верхню комірку вихідного діапазону.

- перемикач **Интегральный процент** дозволяє установити режим включення в гістограму графіка інтегральних процентів.
- перемикач **Вывод графика** дозволяє встановити режим автоматичного створення встроеної діаграми на аркуші, що містить вихідний діапазон.

Приклад 1. Побудувати емпіричний розподіл вступних балів студентів для наступної вибірки: 64, 57, 63, 62, 58, 61, 63, 70, 60, 61, 65, 62, 62, 40, 164, 61, 59, 59, 63, 61.

Розв'язання

1. В клітинку A1 введіть слово **Спостереження**, а в діапазон A2:A21 — значення вступного балу студентів (див. рис. 1).

2. В клітинку B1 введіть назву інтервалів *Бал*. В діапазон B2:B8 введіть граничне значення інтервалів(40, 45, 50, 55, 60, 65, 70).

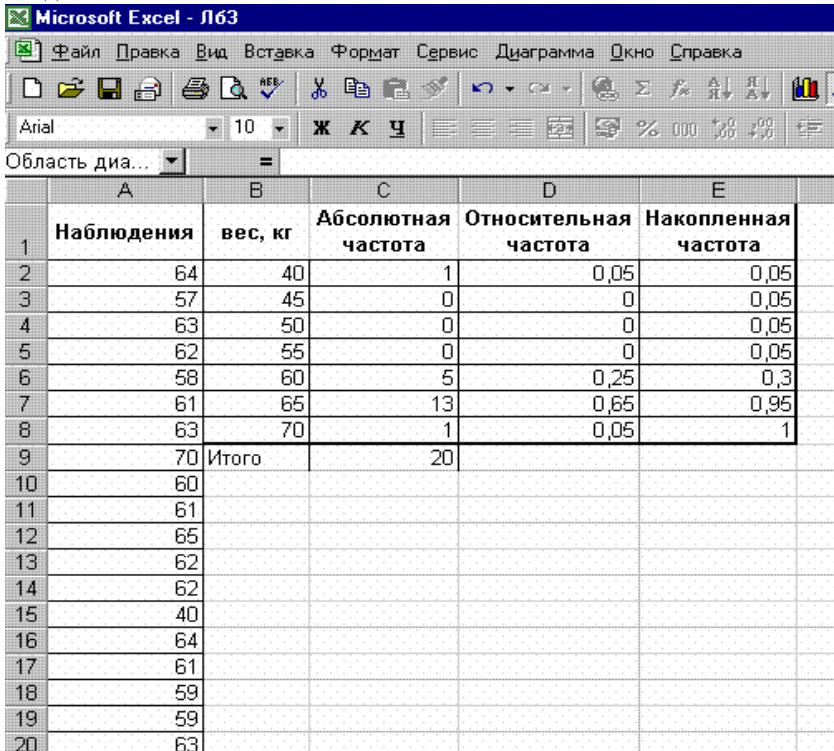
3. Введіть заголовки створюваної таблиці: в клітинки C1 — **Абсолютні частоти**, в клітинку D1 — **Відносні частоти**, в комірці E1 — **Нагромаджені частоти**. (див. рис. 1).

4. За допомогою функції **Частота** завніть стопець абсолютних частот, для цього виділіть блок комірок C2:C8. З панелі інструментов **Головна** викличте **Мастер функцій** (кнопка $f(x)$). У діалоговому вікні виберіть категорію **Статистические** і функцію **ЧАСТОТА**, після чого натисніть кнопку **ОК**. Вказівником миші в робоче поле **Массив_данных** введіть діапазон даних спостереження (A2:A8). В робоче поле **Двоичный_массив** мишею введіть діапазон інтервалів (B2:B8). Зліва на клавіатурі послідовально натисніть комбінацію клавіш **Ctrl+Shift+Enter**. В стовпці C повинен появиися масив абсолютних частот (див. рис.1).

5. В клітинці C9 знайдіть загальну кількість спостережень. Активізуйте клітинку **C9**, на панелі інструментов **Головна** натисніть кнопку **Автосумма**. Переконайтеся, що діапазон додавання вказано правильно і натисніть клавішу **Enter**.

6. Заповніть стовпчик **відносних частот**. В комірку введіть формулу для обчислення відносної частоти: $=C2/SC\$9$. Натисніть клавішу **Enter**. Протягуємо (за правий нижній кут при натиснутій лівій кнопці миші) копіюючи введену формулу в діапазон і одержуємо масив відносних частот.

7. Заповніть стовпчик **нагромаджених частот**. В клітинку D2 скопіюйте значення відносної частоти із комірки E2. В комірку D3 введіть формулу: $=E2+D3$. Натисніть клавішу **Enter**. Протягуємо (за правий нижній кут при натиснутій лівій кнопці миші) скопіюйте введену формулу в діапазон D3:D8. Одержимо масив нагромаджених частот.



	A	B	C	D	E
	Наблюдения	вес, кг	Абсолютная частота	Относительная частота	Накопленная частота
1				0,05	0,05
2	64	40	1	0	0,05
3	57	45	0	0	0,05
4	63	50	0	0	0,05
5	62	55	0	0	0,05
6	58	60	5	0,25	0,3
7	61	65	13	0,65	0,95
8	63	70	1	0,05	1
9	70	Итого	20		
10	60				
11	61				
12	65				
13	62				
14	62				
15	40				
16	64				
17	61				
18	59				
19	59				
20	63				

Рис. 1. Результат обчислень (приклад 1)

8. Побудуйте діаграму відносних і нагромаджених частот. На панелі **Головна** виберіть **Вставка діаграми**. В діалоговому вікні виберіть закладку **Нестандартні** і тип діаграми **График/гистограмма**. Після редаування діаграма буде мати такий вигляд, як нарис. 2.

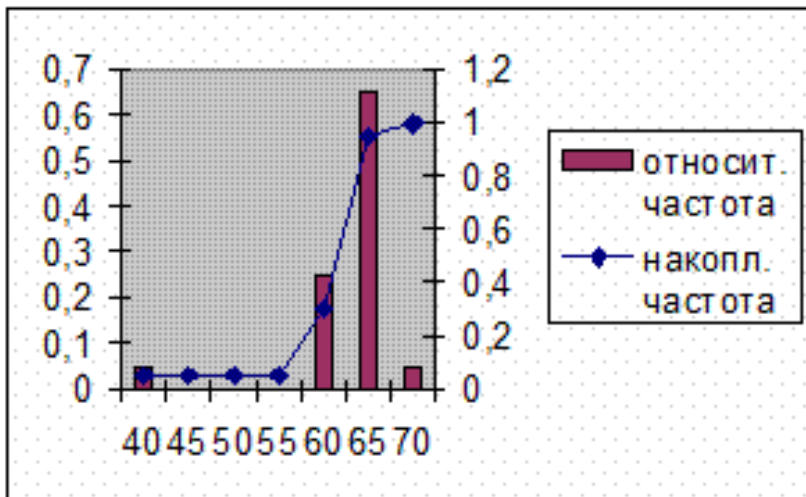


Рис. 2 Діаграма відносних і нагромаджених частот (приклад 1)

ПРАКТИЧНЕ ЗАВДАННЯ

1. Для даних із прикладу 1 побудувати вибірккові функції розподілу, скориставшись процедурою **Гистограма** із пакету **Анализ**.

2. Побудувати вибірккові функції розподілу (відносні і нагромаджені частоти) для вступного бала 20 студентів: 181, 169, 178, 178, 171, 179, 172, 181, 179, 168, 174, 167, 169, 171, 179, 181, 181, 183, 172, 176.

3. Знайдіть розподіл за абсолютними частотами для наступних результатів тестування в балах: 79, 85, 78, 85, 83, 81, 95, 88, 97, 85 (використайте границі інтервалів 75, 82, 90).

4. Розглянемо будь-який із критеріїв оцінки якості педагога-професіонала, наприклад, «успішне розв'язання завдань навчання і виховання». Відповідь на це питання анкети типу «так», «ні» достатньо примітивний. Щоби зменшити відносну похибку такого вимірювання, необхідно збільшити число можливих відповідей на конкретне критеріальне питання. В табл. 1 подані можливі варіанти відповідей.

Позначимо цей параметр через x . Тоді в процесі відповіді на питання величина x набуде дискретного значення x , що належить визначеному інтервалу значень. Поставимо у відповідність кожній із відповідей певне числове значення параметра x (див. табл. 1).

Табл. 1 Критеріальне запитання: успішне розв'язання задач навчання і виховання

№ п/п	Варіанти відповідей	X
1	Абсолютно неуспішно	0,1
2	Неуспішно	0,2
3	Успішно в дуже малій мірі	0,3
4	В певній мірі успішно, але ще багато недоліків	0,4
5	В середньому успішно, але недоліки ще є	0,5
6	Успішно з деякими обмеженнями	0,6
7	Успішно, але хотілось би покращити результат	0,7
8	Достатньо успішно	0,8
9	Дуже успішно	0,9
10	Абсолютно успішно	1

При проведенні анкетування в кожній окремій анкеті параметр x приймає випадкове значення, але тільки в межах числового інтервалу від 0,1 до 1. Тоді в результаті вимірювань ми одержуємо нерангований ряд випадкових значень (див. табл. 2).

Таблиця 2. Результати опитування ста учителів

0,6	0,7	1	0,6	0,2	0,8	0,3	0,5	0,9	0,3
0,5	0,1	0,4	0,5	0,5	0,4	0,4	0,6	0,5	0,4
0,6	0,9	0,7	0,9	0,8	0,5	0,5	0,6	0,8	0,4
0,4	0,4	0,8	0,7	0,6	0,6	0,7	0,8	0,5	0,6
0,7	0,6	0,7	0,3	0,2	0,7	0,5	0,3	0,4	0,5
0,9	0,7	0,6	0,5	0,7	0,6	0,2	0,8	0,8	0,3
0,7	0,5	0,7	0,6	0,2	0,5	0,8	0,3	0,7	0,8
0,7	0,6	0,6	0,8	0,4	0,6	0,6	0,6	0,9	0,7
0,7	0,5	0,7	0,6	0,9	0,4	0,8	0,7	0,5	0,8
0,8	0,9	0,4	0,3	0,4	0,6	0,4	0,5	0,3	0,5

Згрупуйте одержану вибірку, розрахуйте середнє значення вибірки, стандартне відхилення, абсолютну і відносну частоту появи параметра, а також побудуйте графік густини ймовірності

$$f(x) = \frac{W(x)}{\sigma \sqrt{2\pi}}$$

де $W(x)$ – відносна частота події

σ - стандартне відхилення;

$$\pi = 3,14.$$

Побудуйте графік функції $f(x)$ і порівняйте його з нормальним розподілом Гаусса.