

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Є. О. СЕМЕНОВ, В. Ф. РАЙКО, О. І. ІЛЬІНСЬКА

## **Числові методи аналізу з охорони праці**

**Практикум для студентів освітньо-кваліфікаційного рівня  
бакалавр спеціальності 263 «Цивільна безпека»**

Харків  
НТУ «ХП»  
2022

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Є. О. СЕМЕНОВ, В. Ф. РАЙКО, О. І. ІЛЬІНСЬКА

## **Числові методи аналізу з охорони праці**

**Практикум для студентів освітньо-кваліфікаційного рівня  
бакалавр спеціальності 263 «Цивільна безпека»**

Затверджено  
редакційно-видавничою радою  
університету,  
протокол № 1 від 28.01.2022 р.

Харків  
НТУ «ХП»  
2022

**УДК 331.45:519.22(075)**

**С30**

**Рецензенти:**

**Шкоп А.О.** – канд. техн. наук, директор ТОВ «НВП «Машинобудівник»;

**Михайлова Є.О.** – канд. техн. наук, доцент кафедри природоохоронних технологій, екології та безпеки життєдіяльності, ХНЕУ ім. С. Кузнеця

**Автори:**

**Є. О. Семенов** к.т.н., доц., **В. Ф. Райко** к.т.н., проф.,  
**О. І. Ільїнська** к.т.н.

**С30** «Числові методи аналізу з охорони праці» Практикум для студентів освітньо-кваліфікаційного рівня «бакалавр» спеціальності 263 «Цивільна безпека» / Є. О. Семенов, В. Ф. Райко, О. І. Ільїнська – Харків: НТУ «ХП», 2022. – 122 с.

**УДК 331.45:519.22(075)**

© Семенов Є. О., Райко В. Ф., Ільїнська О. І., 2022

© НТУ «ХП», 2022

## ЗМІСТ

ВСТУП	4
Практичне заняття 1	
<b>Аналіз результатів статистичного спостереження</b>	5
Практичне заняття 2	
<b>Використання статистичного зведення та групування.</b>	
<b>Статичні таблиці в обробці даних</b>	15
Практичне заняття 3	
<b>Розрахунок абсолютних, відносних та середніх величин</b>	32
Практичне заняття 4	
<b>Розрахунок характеристик варіаційних рядів</b>	42
Практичне заняття 5	
<b>Аналіз форми розподілу варіаційних рядів</b>	57
Практичне заняття 6	
<b>Визначення стандартної та граничної помилки випадкової вибірки, меж довірчих інтервалів та необхідної кількості вибірки</b>	68
Практичне заняття 7	
<b>Визначення стандартної та граничної помилки механічної, типової, серійної та малої вибірки. Визначення необхідної кількості спостережень</b>	82
Практичне заняття 8	
<b>Вирівнювання варіаційних рядів</b>	102

## ВСТУП

Для виробничої діяльності підприємств характерні промислові (виробничі, професійні) ризики, пов'язані з недостатньою кваліфікацією працівників, незадовільним станом організації охорони праці у виробничих підрозділах підприємства, несправністю інструменту і устаткування (верстатів, приладів, складних технічних комплексів тощо), а також ризики внаслідок порушення експлуатації виробничих будівель, інженерних споруд, технологічних процесів. Вивчення, оцінка і узагальнення причин зменшення і запобігання професійних ризиків у галузі безпеки праці мають велике значення у економічній діяльності малого і середнього бізнесу.

Передумовою для створення ефективної системи управління безпекою здоров'я працюючих на підприємстві є оцінка ризиків. Узагальнений аналіз дозволить роботодавцю розробити відповідний комплекс заходів для кожного структурного підрозділу, конкретного робочого місця не тільки для виявлення шкідливих і небезпечних виробничих факторів, а і приведення умов праці у відповідність до діючого законодавства.

Даний практикум має на меті формування у майбутніх фахівців вміння практичного застосування отриманих знань з курсу «Числові методи аналізу даних з охорони праці» і використання їх для розробки корегуючих заходів по менеджменту безпеки праці економічних структур, методів систематизації, опрацювання й аналізу масових експериментальних чисельних даних по розробці стратегії підприємств у забезпеченні промислової безпеки і здоров'я працюючих.

## **Практичне заняття 1**

### **АНАЛІЗ РЕЗУЛЬТАТІВ СТАТИСТИЧНОГО СПОСТЕРЕЖЕННЯ**

**Мета роботи:** Отримання навичок аналізу результатів статистичного спостереження.

#### **Загальні положення**

*Статистичне спостереження* – це спланована, науково організована реєстрація масових даних про соціальні, економічні та інші явища і процеси.

*Статистичні дані* – це масові системні кількісні характеристики соціально-економічних явищ та процесів.

Статистичне спостереження може бути первинним або вторинним.

*Первинне спостереження* – це реєстрація даних, що надходять безпосередньо від об'єкта спостереження.

*Вторинне спостереження* – це збирання раніше зареєстрованих та оброблених даних.

Статистична інформація повинна відповідати певним вимогам:

- вірогідність (відповідність реальному стану);
- повнота за обсягом та змістом;
- своєчасність;
- порівнянність за часом або у просторі;
- доступність.

Джерелами статистичної інформації є: звітність, спеціально організоване статистичне спостереження, реєстри.

За терміном охоплення звіти бувають: поточні звіти (місяць, чверть року, півроку) та річні звіти.

Розрізняють типову і спеціалізовану звітність.

Типова звітність має єдину форму і зміст для всіх підприємств і організацій незалежно від форми власності та відомчого підпорядкування.

Спеціалізована звітність властива тим підприємствам чи окремим

виробництвам, що мають свої специфічні властивості.

Залежно від рівня затвердження та призначення звітність поділяється на зовнішню та внутрішню.

За періодичністю подання звітність буває поточна (тижнева, місячна, квартальна), яка охоплює показники поточної діяльності суб'єктів, і річна.

Звітність характеризується такими властивостями, як обов'язковість, систематичність, вірогідність.

*Обов'язковість* – подання звітів обов'язкове для усіх зареєстрованих суб'єктів діяльності з додержанням уніфікованої форми, затвердженого переліку показників, із зазначенням реквізитів підзвітного об'єкту.

*Систематичність* передбачає регулярне, своєчасне складання та подання звітності в затверджені терміни.

*Вірогідність* – дані, наведені у звітності, мають відповідати дійсності і виключати будь-які викривлення.

Спеціально організоване статистичне спостереження охоплює ті сторони суспільного життя, які не відобразились у звітності. До них належать: переписи, одноразові обліки, спеціальні обстеження, опитування.

Переписи проводяться періодично або одноразово і дають повну характеристику масового явища станом на якусь дату або певний момент часу.

*Обліки* – суцільні спостереження масових даних, які ґрунтуються на даних огляду, опитування та документальних записів.

*Спеціальні спостереження* – переважно несуть цільне обстеження масових явищ згідно з певною тематикою, що виходить з певною тематикою, що виходить за межі звітності.

*Опитування* – це, як правило, несцільне спостереження з метою вивчення думок, мотивів, оцінок, що реєструються зі слів респондентів.

*Статистичний реєстр* – є третьою формою статистичного спостереження. Це список або перелік одиниць певного об'єкта спостереження зі значенням необхідних ознак, який складається та оновлюється під час постійного відстежування.

За часом реєстрації даних статистичне спостереження поділяється на поточне, періодичне та одноразове.

За способами одержання статистичних даних спостереження поділяється на безпосереднє, документальне і опитування респондентів.

*Опитування респондентів* – це таке спостереження, при якому відповіді на питання формуляра записують зі слів респондента. Воно буває експедиційне, самореєстрація, кореспондентське і анкетне.

За повнотою охоплення одиниць сукупності спостереження поділяють на суцільне і несучільне.

При суцільному спостереженні обстеженню і реєстрації підлягають усі без винятку елементи сукупності, а при несучільному спостереженні обліку підлягають не всі елементи, а тільки їх частина.

Несучільне спостереження поділяють на такі види: спостереження основного масиву, вибіркоче, монографічне, анкетне, моніторинг.

Спостереження основного масиву охоплює переважно частину елементів сукупності, обсяг значень істотної ознаки у яких визначає розмір явища, тобто при обстеженні відкидають певну кількість одиниць, які не можуть істотно вплинути на характеристику сукупності в цілому.

При вибіркочому спостереженні також обстежують певну відібрану частину.

Монографічне спостереження передбачає детальне обстеження лише окремих типових елементів сукупності.

*Моніторинг* – це спеціально організоване систематичне спостереження за станом певного середовища.

*Об'єкт спостереження* – сукупність явищ, що підлягають обстеженню. Може застосовуватись набори кількісних та якісних обмежувальних ознак.

*Статистична сукупність* – це множина елементів, яка складається головним чином під дією загальних причин та умов, які пов'язані загальними характерними рисами, загальними важливими ознаками, загальністю процесів розвитку.

Окремі елементи статистичної сукупності зветься *одиницями сукупності*.



*Одиниця сукупності* – первинний елемент об’єкта спостереження, що є носієм ознак, які підлягають реєстрації.

*Ознака* – це властивість, яка відбиває сутність, характер, особливість одиниці статистичної сукупності. За формою виразу ознаки розрізняють: описові (атрибутивні) та кількісні, котрі у свою чергу бувають дискретні та неперервні.

*Одиниця спостереження* – це первинна одиниця статистичної сукупності, від якої одержують інформацію.

### **Приклади розв’язання задач**

#### **Приклад 1.**

Визначте об’єкт спостереження, одиницю спостереження та одиницю сукупності обстежень:

а) оцінка якості підготовки студентів з фахових дисциплін у державних та недержавних технічних навчальних закладах;

б) облік наявності касових апаратів у комерційних торговельних пунктах міста.

#### **Розв’язання**

а) Об’єкт спостереження – якість підготовки студентів із фахових дисциплін у державних та недержавних технічних навчальних закладах.

Одиниця сукупності – студент у державних та недержавних технічних навчальних закладах.

Одиниця спостереження – навчальний відділ кожного державного та недержавного технічного навчального закладу.

б) Об’єкт спостереження – наявність касових апаратів у комерційних торгових пунктах міста.

Одиниця сукупності – комерційний торговий пункт міста.

Одиниця спостереження – податкова інспекція міста.

#### **Приклад 2.**

Вкажіть тип ознаки та можливі варіанти її кількісного вираження для наведених запитань «Анкети осіб, які звернулися до служби зайнятості».

1. Ваша стать:
  - 1) чоловіча; 2) жіноча.
2. Ваш вік (кількість років).
3. Ваша соціальна група до моменту втрати роботи?
  - 1) робітник; 2) службовець; 3) кооператор; 4) фермер, орендар;
- 5) пенсіонер; 6) військовослужбовець; 7) студент; 8) підприємець.
4. Чи робили Ви спробу знайти роботу?
  - 1) так; 2) ні; 3) не відповів.
5. Як довго Ви не маєте постійної роботи?
6. Якою за характером має бути Ваша робота? (можна зазначити не більше 2-х відповідей)
  - 1) переважно фізична;
  - 2) організаторська;
  - 3) виконавча; 4) творча, наукова;
  - 5) підприємницька.
7. Як Ви оцінюєте рівень своєї фахової кваліфікації?
  - 1) високий; 2) вищий за середній; 3) середній; 4) нижче середнього; 5) низький.

### **Розв'язання**

- 1 запитання – Описова ознака.
- 2 запитання – Кількісна ознака, неперервна.
- 3 запитання – Описова ознака.
- 4 запитання – Описова ознака.
- 5 запитання – Кількісна ознака, неперервна.
- 6 запитання – Описова ознака.
- 7 запитання – Описова ознака.

### **Приклад 3.**

Для кожної групи працівників визначте ранги цінностей трудової діяльності (від I до V, де I – максимальний ранг, а V – відповідно мінімальний ранг). Зробіть висновки. Вихідні дані до задачі наведені в табл. 1.

Таблиця 1 – Вихідні дані

Цінності трудової діяльності	Відсоток до числа опитаних		
	Майстри	Інженери	Робітники
Активна участь в управлінні виробництвом			
Можливість службової кар'єри			
Можливість добре заробити			
Престиж професії			
Творчий зміст праці			

### Розв'язання

Майстри надають перевагу активній участі в управлінням виробництвом. Інженери – віддають переваг престижу професії. Робітники надають перевагу можливості добре заробити. Результати наведемо у вигляді табл. 2.

Таблиця 2 – Результати розв'язання задачі

Цінності трудової діяльності	Відсоток до числа опитаних		
	Майстри	Інженери	Робітники
Активна участь в управлінні виробництвом	I	V	IV
Можливість службової кар'єри	II	II	V
Можливість добре заробити	III	IV	I
Престиж професії	IV	I	III
Творчий зміст праці	V	III	II

### Приклад 4.

За даними наведеними в табл. 3 необхідно визначити форму та вид статистичних спостережень. Належність до тієї чи іншої форми або виду позначити знаком “+”.

Таблиця 3 – Види та форми спостережень

Статичне спостереження	Форма			Вид				
	Звітність	Спеціально організоване спостереження	Реєстр	За ступенем охоплення одиниць сукупності		За часом реєстрації фактів		
				Суцільне	Несуцільне	Поточне	Періодичне	Одноразове
Анкетне опитування пасажирів авіакомпанії, яка обслуговує щотижневі рейси								
Перелік усіх релігійних громад України								
Місячний звіт машинобудівних підприємств області про виробництво продукції машинобудування								
Обстеження фінансової діяльності інвестиційної компанії								
Опитування окремих робітників фірми								

### Розв'язання

Результати визначення наведені в табл. 4.

Таблиця 4 – Результати визначення виду та форми спостережень

Статичне спостереження	Форма			Вид				
	Звітність	Спеціально організоване спостереження	Реєстр	За ступенем охоплення одиниць сукупності		За часом реєстрації фактів		
				Суцільне	Несуцільне	Поточне	Періодичне	Одноразове
Анкетне опитування пасажирів авіакомпанії, яка обслуговує щотижневі рейси		+			+		+	
Перелік усіх релігійних громад України			+	+			+	
Місячний звіт машинобудівних підприємств області про виробництво продукції машинобудування	+			+			+	
Обстеження фінансової діяльності інвестиційної компанії		+			+			+
Опитування окремих робітників фірми		+			+			+

### Завдання для самостійної роботи

#### Завдання 1.

Необхідно визначити об'єкт спостереження, одиницю спостереження та одиницю сукупності обстежень для вказаного обстеження. Обстеження: перепис виробничих площ у державних підприємствах промисловості.

#### Завдання 2.

Вкажіть тип ознаки та можливі варіанти її кількісного відображення для наведених запитань «Анкети студентів».

*Анкета студента*

1. Ваш вік (вказати кількість років).
2. Чи відповідає, на Вашу думку, сучасним вимогам фахова підготовка інженерів з охорони праці у нашому навчальному закладі?
  - 1) так, відповідає; 2) ні, не відповідає.
3. Зазначте свої життєві плани:
  - 1) здобути вищу освіту;
  - 2) забезпечити матеріальний добробут сім'ї;
  - 3) знайти цікаву роботу.
4. Ви задоволені своїм вибором професії?
  - 1) задоволений; 2) ставлюсь байдуже; 3) незадоволений; 4) не відповів.

**Завдання 3.**

За наведеними даними в табл. 5 визначте форму та вид статистичних спостережень.

Таблиця 5 – Види та форми спостережень

Статистичне спостереження	За ступенем охоплення одиниць сукупності	За часом реєстрації даних
Облік числа нещасних випадків на виробництвах		
Облік чисельності новонароджених		
Обстеження бюджетів незахищених верств населення		
Реєстрація кількості випадків побутового травматизму		
Опитування молодих сімей регіону з питань планування сім'ї		

**Контрольні питання**

1. Що таке статистичного спостереження?
2. Які бувають різновиди статистичного спостереження?
3. Назвіть джерела статистичної інформації.
4. Вкажіть які вимоги висуваються до статистичної інформації.

5. Зазначте якими властивостями характеризується звітність.
6. Вкажіть які види у типи звітності бувають.
7. Які форми спостережень належать до спеціально організованого статистичного спостереження?
8. Що таке статистичний реєстр?
9. Зазначте як поділяються статистичні спостереження за часом реєстрації даних та за повнотою охоплення одиниць сукупності.
10. Вкажіть різницю між вибіркоvim обстеженням, обстеженням основного масиву, монографічним обстеженням.
11. У чому різниця між одиницею спостереження та одиницею сукупності?

### **Література**

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Громько Г. Л. Общая теория статистики: практикум \ Г. Л. Громько. Москва : Инфра-М, 2008. 240 с.
5. Минашкин В. Г. Теория статистики : учебно-методический комплекс / В. Г. Минашкин [и др.]. Москва : Издательство Центр ЕАОИ, 2008. 296 с.

## **Практичне заняття 2**

### **ВИКОРИСТАННЯ СТАТИСТИЧНОГО ЗВЕДЕННЯ ТА ГРУПУВАННЯ. СТАТИЧНІ ТАБЛИЦІ В ОБРОБЦІ ДАНИХ**

**Мета роботи:** Отримання навичок первинної обробки результатів спостережень та їх представлення.

#### **Загальні положення**

Зібраний статистичний матеріал піддається логічному і арифметичному контролю (перевірки смислової узгодженості відомостей первинного документа і перевірці лічильної узгодженості). Потім приступають до статистичного зведення.

*Статистичне зведення* – систематизація одиничних фактів, що дозволяє перейти до узагальнюючих показників, що належать до всієї досліджуваної сукупності і її частин, і здійснювати аналіз і прогнозування досліджуваних явищ і процесів.

Зведення визначає загальний розмір досліджуваного явища за заданими показниками, представляючи загальні підсумки по досліджуваній сукупності в цілому без будь-якої попередньої систематизації зібраного матеріалу.

Статистичні дані в широкому її розумінні передбачає систематизацію та групування даних, характеристику освічених груп системою показників, підрахунок відповідних підсумків та подання результатів зведення у вигляді таблиць, графіків.

За складністю побудови статистичні зведення поділяються на *прості* та *групові*.

*Просте статистичне зведення* не передбачає розподіл на групи даних статистичного спостереження. При цьому зведені визначають тільки загальний підсумок за статистичною сукупністю, або загальне значення показника для усієї сукупності.

*Групове статистичне зведення* – це розподіл елементів сукупності на групи та підгрупи за принципом схожості та відмінності певних ознак. Групи характеризуються за допомогою системи



показників. Завершується групове статистичне зведення побудовою статистичної таблиці.

Є загальноприйнята методологія стандартного розподілу сукупності на групи – це *класифікація*. Класифікації мають рівень стандартів. У класифікаціях чітко визначені групувальні ознаки.

З метою розв'язання конкретних задач, які не підпадають під дію класифікаторів, проводяться *нестандартні групування* за певними ознаками. Залежно від мети дослідження та складності масового процесу, групувальних ознак може бути одна, дві та більше.

*Групування* – це процес утворення однорідних груп на основі розчленування статистичної сукупності на частини або об'єднання досліджуваних одиниць у приватні сукупності за істотними для них ознаками.

Ознаки, за якими проводиться розподіл одиниць сукупності на групи, називаються *групувальними ознаками*, або основою групування.

За допомогою методу групувань вирішуються завдання:

- виділення соціально-економічних типів явищ;
- вивчення структури явища і структурних зрушень, що відбуваються в ньому;
- виявлення зв'язку і залежності між явищами.

Для вирішення цих завдань застосовують відповідно типологічні, структурні і аналітичні угруповання.

Дана класифікація видів статистичних групувань має умовний характер, оскільки на практиці вони застосовуються в комплексі.

*Типологічне групування* – це розчленування різнорідної сукупності на окремі якісно однорідні групи і виявлення на цій основі соціально-економічних типів явищ. При використанні методу типологічних групувань важливе значення має правильний вибір групувальної ознаки. При атрибутивній ознаці з незначною різноманітністю значень число груп визначається властивостями досліджуваного явища (наприклад, групування підприємств за формами власності). Виділення типів на основі кількісної ознаки полягає у визначенні груп з урахуванням значень досліджуваних ознак.

*Структурне (варіаційне) групування* – це групування за

кількісними ознаками. Головною рисою цього групування є однорідність сукупності. Різновидом структурного групування є *ряд розподілу* – це структурне групування за однією ознакою. Характеристиками ряду розподілу є *варіанти* та *частоти (частки)*.

*Варіанти* – це конкретні значення груповальної ознаки.

*Частоти* – кількості елементів сукупності, які відповідають окремій варіанті. Відносні частоти (% до підсумку) називають **частками**.

Типологічні та структурні групування характеризують сукупність за складом, на відміну від аналітичного групування, яке з'ясовує взаємодію між одиницями сукупності.

Для вивчення зв'язку між окремими ознаками явища використовуються **аналітичні групування**.

Аналітичне групування виявляє наявність та напрямок зв'язку між двома ознаками: *факторною* та *результативною*. Сукупність поділяється на групи за факторною ознакою і в кожній групі визначається середній рівень результативної ознаки. Кількість груп при атрибутивній (описовій) ознаці певною мірою визначається кількістю найменувань ознаки. Якщо ознака кількісна, тоді кількість груп, на які розподіляється сукупність даних, залежить від ступеня варіації груповальної ознаки.

Утворення груп за двома і більше ознаками називається **комбінованими групуваннями**.

### **Побудова статистичних групувань**

Для побудови статистичних групувань необхідно:

- виконати вибір груповальної ознаки;
- зробити вибір кількості груп;
- визначити інтервал групування.

1. Вибір груповальної ознаки.

*Вибір груповальної ознаки* – це вибір ознаки, за якою здійснюється розбиття сукупності на окремі групи. Як ознаки необхідно використовувати істотні обґрунтовані ознаки.

За формою вираження груповальної ознаки бувають *атрибутивні*

(що не мають кількісного вираження, наприклад, професія) і *кількісні* (наприклад, число філій, величина доходу).

Кількісні ознаки можуть бути *дискретними* (переривчастими, значення яких виражаються тільки цілими числами, наприклад, число філій) і *безперервними* (приймають як цілі, так і дробові значення, наприклад, величина доходу).

За характером зміни ознаки поділяються на: *варіаційні*, які мають можливість приймати різні значення, та *альтернативні*, що мають можливість приймати одно з двох протилежних значень.

За роллю ознаки у взаємозв'язку різних явищ ознаки розрізняють: *факторні*, що впливають на результат соціально-економічного процесу, і *результативні*, що характеризують наслідки процесу та залежать від факторних ознак.

Від вибору групувальної ознаки залежить розв'язання питання про утворення груп.

## 2. Вибір кількості груп.

Після вибору групувальної ознаки постає питання про кількість груп, на які буде розподілена досліджувана сукупність, і про межі груп. Розв'язання даного питання залежить від конкретних умов і завдань. На цьому етапі встановлюють величину і границі кожного інтервалу. Оскільки характер реально існуючих сукупностей та їх розподіл досить різноманітні, то існують різні методичні підходи у вирішенні питання про кількість груп. Загальним принципом, з якого треба виходити, є характер матеріалу та чисельність досліджуваної сукупності. Характерні особливості розподілу не виявляються, якщо при невеликій сукупності одиниць спостереження взяти велике або дуже мале число груп. До цього питання існують різні підходи.

Якщо групувальна ознака описова (атрибутивна), так кількість груп, на які поділяється статистична сукупність, дорівнює кількості різновидів цієї ознаки.

Якщо групувальна ознака альтернативна, так кількість груп, на які поділяється статистична сукупність, дорівнює двом.

Групувальна ознака може змінюватися дискретно, тобто *перервно* і *безперервно*. Якщо мінливість ознаки має дискретний характер, число

груп варіаційного ряду, як правило, визначається числом цих дискретних значень (якщо їх небагато). Наприклад, групування підприємств за наявністю виробничих бригад – 1, 2, 3 і т.д.

При мінливості ознаки безперервного характеру звертають увагу на ранжований ряд (ряд, в якому значення ознаки знаходяться у порядку зростання чи спадання). Якщо зростання рівнів групувальної ознаки відбувається з плавними переходами, перевага віддається *рівним інтервалам*. У разі стрибкоподібних змін групувальної ознаки будують групи з *нерівними інтервалами*. Границі у таких випадках встановлюють, як правило, в точках різких переходів.

Питання визначення кількості груп в умовах порівняно поступових змін групувальної ознаки (у ранжованому ряді) може вирішуватися різними методичними підходами. Орієнтовно число інтервалів (груп) можна визначити шляхом добування квадратного кореня з обсягу досліджуваної сукупності. При цьому число інтервалів не повинно бути меншим 5 і більшим 20.

Якщо сукупність невелика за обсягом, інтервальний ряд будують таким чином, щоб у крайні групи (першу і третю) потрапило по 25% одиниць сукупності, а в середню – 50%. У цьому випадку групування складається з трьох нерівних інтервалів. Наприклад, сукупність з 28 підприємств матиме розподіл:

I група – 7 одиниць, II група – 14 одиниць, III група – 7 одиниць.

Якщо сукупність складається з великої кількості одиниць і розподіл одиниць за групувальними ознаками близький до нормального, використовують формулу Стерджеса (Стьорджеса):

$$n = 1 + 3,322 \cdot \lg N, \quad (1)$$

де  $n$  – кількість груп;  $N$  – кількість елементів сукупності.

Застосовуючи цю формулу, будемо мати для сукупності розміром 10-100 одиниць –  $4 \div 7$  груп, а для сукупності 100-1000 одиниць –  $7 \div 10$  груп.

Рекомендації В.П. Левинського є найвдалішими. Він пропонує

своєрідні нормативи числа інтервалів, зумовлені обсягами досліджуваної сукупності (табл. 1).

Якщо число одиниць спостереження налічується до 40, число інтервалів становитиме 3 або 5. Розподіл сукупності на 4 групи небажаний, адже в такому випадку втрачається середня група (інтервал).

Таблиця 1 – Рекомендоване число груп для різної кількості спостережень

Кількість одиниць спостережень	Рекомендоване число інтервалів (груп)
40–60	6–8
60–100	8–10
100–200	10–12
200–500	12–17

В такому разі дослідникові надається можливість певного вибору числа груп залежно від характеру сукупності.

### 3. Визначення інтервалу групування.

*Інтервал групування* – це значення варіюючої ознаки, що лежить в певних межах. Під величиною інтервалу розуміють різницю між максимальним і мінімальним значеннями ознаки в групі. При цьому максимальне значення ознаки в групі називається *верхньою межею інтервалу*, а мінімальне – *нижньою межею*.

Залежно від ступеня коливання групувальної ознаки, характеру розподілу статистичної сукупності встановлюються інтервали рівні або нерівні.

Якщо варіація ознаки відбувається в порівняно вузьких межах і розподіл носить рівномірний характер, то будують групування з рівними інтервалами, величина інтервалу визначається за формулою:

$$h = \frac{x_{\max} - x_{\min}}{n}, \quad (2)$$

де  $h$  – величина інтервалу,  
 $x_{\max}$  – максимальне значення,

$x_{min}$  – мінімальне значення,  
 $n$  – кількість груп сукупності.

В тих випадках, коли невелика частина сукупності значно віддалена за розміром групувальної ознаки від сукупності основного масиву, за  $x_{max}$  приймається максимальна варіанта основного масиву.

Існують рекомендації щодо встановлення величини інтервалу групувань з деякими поправками до попередньої формули. У цьому випадку формула набуває вигляду:

$$h = \frac{x_{max} - x_{min} + 1}{n}, \quad (3)$$

У випадках, коли максимальне і мінімальне значення у ранжированому ряду групувальних ознак значно відрізняється від решти показників, за  $x_{max}$  приймається суміжне наступне значення ознаки  $x_{max+1}$ , а за  $x_{min}$  суміжне попереднє її значення  $x_{min-1}$ .

Якщо розрахована величина рівного інтервалу становить дробове число, його заокруглюють до цілого, цим самим розширюючи границі, якими охоплює інтервал розмаху коливання значень групувальної ознаки.

Важливим моментом у практичному використанні результатів групувань слід вважати процес перевірки їх на вірогідність. Це питання потребує детального розгляду окремо.

### Вторинне групування

Вторинне групування виконується на підставі групувань, що вже маються. Побудова вторинних групувань на підґрунті первинних можливо у два способи:

- перегрупування за розміром інтервалів первинного групування;
- перегрупування за питомою вагою окремих груп у загальному обсязі групування.

Треба бути обережним у визначені границь інтервалу і не включати цю границю двічі. Якщо інтервал є відкритим, так його розмір

приймається рівним сусідньому (сумісному) закритому інтервалу.

### Статистичні таблиці

Завершується зведення чи групування статистичною таблицею, яка надає зведену якісну і кількісну характеристику статистичної сукупності.

Статистична таблиця за своїм логічним змістом розглядається як «статистичне речення». «Підмет» – об'єкт дослідження, «присудок» – система показників, що характеризують об'єкт дослідження.

Залежно від структури об'єкта дослідження (підмета) статистичні таблиці поділяються на:

- прості – перелік елементів сукупності, територіальний або хронологічний ряд;
- групові – підмет поділений на групи за однією ознакою;
- комбінаційні – підмет (об'єкт дослідження) поділений на групи за двома або більшою кількістю ознак.

Іноді до класифікації статистичної таблиці, автори додають її характеристику за структурою присудка.

### Приклади розв'язання задач

#### Приклад 1.

Визначити інтервал групування робітників машинобудівного виробництва за рівнем заробітної платні, якщо загальна чисельність робітників складає 120 чоловік, а мінімальний та максимальний дохід відповідно дорівнює 500 і 6500 грошових одиниць.

#### Розв'язання

Знайдемо спочатку кількість груп у сукупності. Оскільки елементи сукупності є кількісними і дискретними, то для знаходження кількості груп при групуванні скористаємось формулою Стерджеса.

Кількість груп за формулою (1) дорівнює:

$$n = 1 + 3,322 \cdot \lg(120) = 7,907 \approx 8,0.$$

Знайдемо величину інтервалу ( $h$ ):

$$h = \frac{x_{\max} - x_{\min}}{n} = \frac{6500 - 500}{8} = 750 \text{ грош. одиниць.}$$

Результати групування наведені в табл. 1:

Таблиця 1 – Результати групування

№ групи	Величина інтервалу групування, грош. одиниць
1	500÷1250
2	1250÷2000
3	2000÷2750
4	2750÷3500
5	3500÷4250
6	4250÷5000
7	5000÷5750
8	5750÷6500

### Приклад 2.

Забої вугільних шахт виробничого об'єднання відрізняються за потужністю та нахилом залягання пластів (1, 2, 3 категорії). Значення цих параметрів для 20 забоїв у табл. 2.

Складіть групування вугільних забоїв:

а) за потужністю пластів, виділивши три групи з рівними інтервалами;

б) за потужністю та нахилу залягання пластів.

Результати групувань надайте у формі статистичних таблиць, зробіть їх аналіз, висновки.



Таблиця 2 – Забої вугільних шахт

№ з/п	Потужність вугільного пласта, см	Нахил залягання пласта	№ з/п	Потужність вугільного пласта, см	Нахил залягання пласта
1	152	2	11	145	2
2	120	3	12	90	2
3	184	1	13	131	3
4	70	3	14	177	2
5	158	2	15	118	3
6	212	1	16	220	2
7	126	2	17	128	3
8	170	1	18	136	2
9	125	2	19	107	2
10	98	2	20	95	3

### Розв'язання

а) Ми маємо провести групування, виділивши три групи з рівними інтервалами. Для цього скористаємось формулою (2) для визначення ширини інтервалу.

Максимальне значення ознаки дорівнює 220 см, мінімальне – 70 см, кількість груп за умовою дорівнює трьом.

Маємо:

$$h = \frac{x_{\max} - x_{\min}}{n} = \frac{220 - 70}{3} = 50.$$

Будуємо інтервали:

$$x_{\min} = 70; 70 + 50 = 120$$

$$120; 120 + 50 = 170$$

$$170; 170 + 50 = 220 = x_{\max}.$$

Якщо одна і та сама величина зустрічається двічі (як верхня межа одного інтервалу і як нижня межа іншого інтервалу), одиниця, що має це значення, відноситься до тієї групи, де ця величина виступає як верхня межа. Так у нашому прикладі потужність вугільного пласта 120 см відноситься до першої групи, а потужність нахилу 170 см – до другої

групи. Підраховуємо, скільки разів зустрічається ознака в кожній групі. Результати групування наведемо у вигляді табл. 3.

Таблиця 3 – Розподіл забоїв вугільних шахт за потужністю пластів

Потужність вугільного пласта, см	Кількість
70-120	7
120-170	9
170-220	4
Разом	20

Згідно табл. 3 більшість забоїв вугільних шахт має потужність вугільного пласта від 120 см до 170 см – 9 пластів.

б) Складемо групування за потужністю та нахилу залягання пластів.

Розподіл забоїв вугільних шахт за потужністю та нахилом залягання пластів наведений в табл. 4.

Таблиця 4 – Розподіл забоїв вугільних шахт за потужністю та нахилом залягання пластів

Потужність вугільного пласта, см	Нахил залягання пласта			Разом
	1	2	3	
70-120	–	3	4	7
120-170	1	6	2	9
170-220	2	2	–	4
Разом	3	11	6	20

### Приклад 3.

За даними табл. 5 складіть аналітичне групування, яке описує залежність якості роботи ткацьких верстатів від їх технічного стану (“0” – пройшов планово-профілактичний ремонт; “1” – потребує ремонту).

Визначте ефект впливу технічного стану станку на якість його роботи. Результати викладіть у формі статистичної таблиці, зробіть висновки.

Таблиця 5 – Розподіл верстатів за технічним станом

№ з/п	Технічний стан верстату	Число обривів нитки на 100 м <sup>2</sup> тканини	№ з/п	Технічний стан верстату	Число обривів нитки на 100 м <sup>2</sup> тканини
1	1	73	11	1	71
2	1	76	12	0	66
3	0	68	13	0	65
4	0	64	14	1	76
5	1	70	15	1	72
6	0	67	16	0	69
7	1	79	17	1	71
8	1	75	18	0	72
9	0	73	19	1	70
10	1	69	20	1	73

### Розв'язання

*1 варіант розв'язання.*

Виконаємо групування верстатів за технічним станом. При цьому групуванні верстати розділяться на дві групи: ті, що пройшли планово-профілактичний ремонт та ті, що потребують ремонту.

Також проведемо групування верстатів за числом обривів нитки, виділивши три групи з нерівними інтервалами.

Максимальне значення ознаки дорівнює 79 обривів на 100 м<sup>2</sup> тканини, мінімальне – 64 на 100 м<sup>2</sup> тканини.

Побудуємо три групи з нерівними інтервалами: 64–68; 69–74; 75–79. Результати групування наведемо у вигляді табл. 6.

Таблиця 6 – Залежність якості роботи ткацьких верстатів від їх технічного стану

Технічний стан верстату	Число обривів нитки			Разом
	64–68	69–74	75–79	
Потребує ремонту (1)	0	8	4	12
Пройшов ремонт (0)	5	3	0	8
Всього	5	11	4	20

*2 варіант розв'язання.*

Наведемо результати групування у вигляді табл. 7.

Таблиця 7 – Залежність якості роботи ткацьких верстатів від їх технічного стану

Технічний стан верстату	Кількість верстатів	Число обривів нитки на 100 м <sup>2</sup> тканини
Потребує ремонту (1)	8	69–79
Пройшов ремонт (0)	12	64–73

В групі верстатів, що пройшла планово-профілактичний ремонт – максимальне число обривів нитки на 100 м<sup>2</sup> тканини складає 73, а в групі де верстати потребують ремонту максимальне значення обривів дорівнює 79.

Висновок: Обрив нитки скорочується в результаті проведення планово-попереджувального ремонту.

### Завдання для самостійної роботи

#### Завдання 1.

Визначити інтервал групування робітників виробництв за рівнем заробітної платні, якщо загальна чисельність робітників ( $N$ ) складає \_\_\_\_ чоловік, а мінімальний ( $x_{\min}$ ) та максимальний дохід ( $x_{\max}$ ) відповідно дорівнює \_\_\_\_\_ і \_\_\_\_\_ грош. одиниць. Вихідні дані наведені в табл. 8.

Таблиця 8 – Вихідні дані

Показник	Варіант									
	1	2	3	4	5	6	7	8	9	10
$N$	140	80	80	130	115	120	65	80	130	145
$x_{\min}$	600	600	500	500	700	1000	900	2000	5000	6000
$x_{\max}$	6200	6550	6800	5300	5420	9000	7900	6900	9900	9800

**Завдання 2.**

За минулий тиждень службою працевлаштування району було влаштовано 24 особи з числа зареєстрованих безробітних. За наведеними даними про стать (0 – чоловіки, 1 – жінки) та тривалістю перерви в роботі (міс.) згрупуйте працевлаштованих (табл. 9):

- а) за тривалістю перерви в роботі;
- б) за статтю та тривалістю перерви в роботі.

Результати групувань надайте в формі статистичних таблиць.

Таблиця 9 – Розподіл працевлаштування

Порядковий номер працевлаштованого	Стать	Тривалість перерви в роботі, міс.	Порядковий номер працевлаштованого	Стать	Тривалість перерви в роботі, міс.
1	0	2	13	0	2
2	1	4	14	0	3
3	0	3	15	1	4
4	0	1	16	0	1
5	1	3	17	1	3
6	0	2	18	0	2
7	1	3	19	0	3
8	1	2	20	0	2
9	0	1	21	1	2
10	0	3	22	1	4
11	1	2	23	0	1
12	0	2	24	1	3

**Завдання 3.**

Перегрупуйте наведені в табл. 10 дані про розподіл будівельно-монтажних робіт, виділивши три групи: малі – до 20 млн. грош. од.; середні – 20-100 млн. грош. од.; крупні – 100 млн грош. од. і більше. Результати вторинного групування надайте в формі статистичної таблиці, зробіть аналіз та висновки.

Таблиця 10 – Розподіл будівельно-монтажних робіт

Обсяг будівельно-монтажних робіт, млн. грош. од.	% до підсумку	
	Кількість організацій	Обсяг робіт
До 10	8	2
10-19	10	7
20-39	24	12
40-79	32	19
80-100	6	8
101-149	15	20
150 і більше	5	32
Разом	100	100

**Завдання 4.**

За наведеними в табл. 11 даними про порушення технологічної дисципліни та збитки (втрати), пов'язані з браком продукції на 22 виробничих ділянках, складіть:

а) комбінаційний розподіл виробничих ділянок за цими ознаками, утворивши три групи з рівними інтервалами (за результатами групування зробіть висновок про наявність та напрямок зв'язку між ознаками);

б) аналітичне групування, яке б показало залежність втрат від порушення технологічної дисципліни. Визначте ефекти впливу порушень технологічної дисципліни на втрати від браку продукції.

Результати групувань викладіть в формі статистичних таблиць, зробіть їх аналіз.

Таблиця 11 – Дані про порушення технологічної дисципліни

№ ділянки	Відсоток порушень технологічної дисципліни	Втрати від браку продукції, тис. грош. одиниць	№ ділянки	Відсоток порушень технологічної дисципліни	Втрати від браку продукції, тис. грош. одиниць
1	2	3	4	5	6
1	1,2	1,0	12	1,7	1,5
2	2,0	1,6	13	2,1	1,7
3	1,4	1,2	14	1,3	1,4
4	1,9	1,5	15	2,0	1,8

1	2	3	4	5	6
5	1,6	1,4	16	2,3	1,6
6	2,4	1,9	17	2,5	2,0
7	1,8	1,4	18	2,7	2,1
8	2,6	2,1	19	2,6	2,0
9	2,0	1,7	20	1,7	1,4
10	1,5	1,2	21	1,5	1,3
11	1,2	0,9	22	2,1	1,6

### Контрольні питання

1. Вкажіть, що являє собою зведення статистичних даних і у чому полягає його мета?
2. Вкажіть, чим просте статистичне зведення відрізняється від групового статистичного зведення?
3. Зазначте коли застосовуються класифікації в статистиці, та коли застосовуються нестандартні групування за певними ознаками?
4. Чим прості групування відрізняються від комбінаційних?
5. Надайте характеристику структурним, типологічним та аналітичним групуванням. Вкажіть, які групування у яких випадках використовують.
6. Які бувають групувальні ознаки?
7. Як визначається кількість груп при дискретній, неперервній і альтернативній ознаці?
8. Що є інтервалом групування, які бувають інтервали?
9. Вкажіть коли і для чого застосовується вторинне групування.
10. Що таке статистична таблиця і з чого вона складається?
11. Як поділяються статистичні таблиці залежно від структури об'єкта дослідження. Чим один вид відрізняється від іншого?

### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.

3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Громько Г. Л. Общая теория статистики: практикум \ Г. Л. Громько. Москва : Инфра-М, 2008. 240 с.
5. Минашкин В. Г. Теория статистики : учебно-методический комплекс / В. Г. Минашкин [и др.]. Москва : Издательство Центр ЕАОИ, 2008. 296 с.



### Практичне заняття 3

## РОЗРАХУНОК АБСОЛЮТНИХ, ВІДНОСНИХ ТА СЕРЕДНІХ ВЕЛИЧИН

**Мета роботи:** Отримання навичок розрахунку абсолютних, відносних та середніх величин, що використовуються в охороні праці.

### Загальні положення

Величини, які використовуються під час аналізу травматизму поділяються на абсолютні, відносні та середні.

*Абсолютні статистичні величини* – кількісні показники, які характеризують розміри соціально-економічних явищ – обсяги сукупності.

Абсолютні величини завжди іменовані числа, тобто мають одиниці виміру. Абсолютні показники вимірюються у натуральних, трудових та вартісних одиницях вимірювання.

Важливе місце займають протиставлення та порівняння даних за допомогою відносних величин.

**Відносні статистичні величини** характеризують кількісні співвідношення різнойменних чи однойменних показників. Кожна відносна величина, то є дріб, чисельником якого є порівняна величина, а знаменником – база порівняння.

Відносна величина показує, у скільки разів порівняна величина більша базисної, або яку частку вона становить відносно базисної, або скільки одиниць однієї величини припадає на 100, 1000, ... одиниць іншої.

Відносні статистичні величини можуть бути виражені іменованими числами, якщо обчислюється відносна величина з різнойменних абсолютних величин.

За своїм пізнавальним значенням відносні величини поділяються на відносні *величини інтенсивності, динаміки, порівняння, структури, координації*.

*Середня величина* – це узагальнююча міра варіюючої ознаки, що

характеризує її рівень у розрахунках на одиницю сукупності.

Середні величини застосовуються як правило за певних умов.

### Умови застосування середніх величин

Перша умова: індивідуальні величини повинні відноситися до якісно однорідної сукупності та кількість їх має бути достатньо великою.

Друга умова: розподіл сукупності на якісно однорідні групи, тобто поєднання методу середніх величин із методом групування.

Третя умова: застосування загальних та групових середніх при умові обчислення їх із якісно однорідної сукупності.

У статистичних розрахунках використовуються середні величини: арифметична, гармонічна, квадратична, геометрична. Кожна з середніх величин може бути надана у простій формі (за первинними, тобто не згрупованими даними) і у зваженій формі (за вторинними, згрупованими даними).

**Середня арифметична** обчислюється в тих випадках, коли обсяг осереднюємої ознаки утворюється як сума її значень у окремих одиниць досліджуваної статистичної сукупності.

Розрізняють середню арифметичну *просту* та *зважену*.

*Середня арифметична проста* застосовується у випадках, коли є окремі значення ознаки, тобто дані не згруповані. Вона дорівнює сумі окремих значень ознаки, поділеній на число цих значень:

$$\bar{x}_{ap.} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

*Середня арифметична зважена* в дискретному ряді розподілу застосовується у випадках, коли дані представлені у вигляді рядів розподілу або групувань. Одні і ті ж значення ознаки повторюються кілька разів:

$$\bar{x}_{ар.} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_n n_n}{n_1 + n_2 + \dots + n_n} = \frac{\sum x \cdot n}{\sum n}, \quad (2)$$

де  $n$  – число однакових значень ознаки в рядах розподілу, тобто частота, або вага.

Середня арифметична зважена залежить не тільки від значень ознаки, а й від частот, тобто від складу сукупності, від її структури.

**Середня гармонійна** – це величина, зворотна середньої арифметичної. Коли статистична інформація не містить частот за окремими варіантами, а представлена як їх добуток, застосовується формула *середньої гармонійної зваженої*.

В тому випадку, коли обсяги явищ за кожною ознакою рівні, застосовується *середня гармонійна проста*.

$$\text{Середня гармонійна проста: } x_{г\text{арм}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum \frac{1}{x_i}} \quad (3)$$

$$\text{Середня гармонійна зважена: } x_{г\text{арм}} = \frac{B_1 + B_2 + \dots + B_n}{\frac{B_1}{x_1} + \frac{B_2}{x_2} + \dots + \frac{B_n}{x_n}} = \frac{\sum B_i}{\sum \frac{B_i}{x_i}} \quad (4)$$

**Середня геометрична** – це величина, яка використовується як середня з відносин. Цією середньої зручно користуватися, коли приділяється уваги не абсолютним різницям, а відносинам двох чисел, тобто коли індивідуальні значення ознаки – відносні величини. Наприклад, середня геометрична використовується при розрахунку середнього коефіцієнта зростання.

$$\text{Середня геометрична проста: } \bar{x}_{г\text{еом.}} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{\prod x_i} \quad (5)$$

Існують також інші середні величини.

## Приклади розв'язання задач

### Приклад 1.

За наведеними статистичними даними розрахувати середній коефіцієнт частоти травматизму ( $K_c$ ) за 10 років, якщо отримані дані містяться у вигляді ряду чисел: 5,4; 4,3; 4,4; 4,5; 4,7; 4,8; 4,6; 4,2; 4,9; 4,1.

### Розв'язання

Оскільки наявні окремі значення признаку, дані не згруповані, зустрічаються 1 раз, то використаємо формулу середньої арифметичної простої:

$$\bar{x} = \frac{\sum x}{n} = \frac{5,4 + 4,3 + 4,4 + 4,5 + 4,7 + 4,8 + 4,6 + 4,2 + 4,9 + 4,1}{10} = 4,59 \approx 4,6$$

### Приклад 2.

Визначити середній коефіцієнт важкості травматизму по підприємствах регіону за наступними вихідними даними, що наведені у табл. 1.

Таблиця 1 – Вихідні дані для розрахунку

Коефіцієнт травматизму ( $K_v$ )	важкості	Кількість підприємств з однаковим коефіцієнтом важкості травматизму
2		1
3		5
4		8
5		4
6		2

### Розв'язання

Дані представлені у вигляді дискретного ряду розподілу, одні і ті ж значення групувальної ознаки повторюються кілька разів. Тому застосуємо формулу середньої арифметичної зваженої. Для розрахунку заповнимо стовпчик  $x \cdot n$ , та розрахуємо суму за стовпчиком. Результати наведемо у вигляді табл. 2.

Таблиця 2 – Результати розрахунків

Коефіцієнт важкості травматизму ( $Kв$ ), ( $x$ )	Кількість підприємств з однаковим коефіцієнтом важкості травматизму, ( $n$ )	$x \cdot n$
2	1	2
3	5	15
4	8	32
5	4	20
6	2	12
Всього	20	81

$$\bar{x} = \frac{\sum x \cdot n}{\sum n} = \frac{81}{20} = 4,05$$

### Приклад 3.

Троє працівників в цеху протягом одного робочого дня зайняті виготовленням однакових деталей. Один робочий використовує на виготовлення деталі в 30 хвилин, другий – 15 хвилин, третій – 10 хвилин. Визначити середні витрати часу на виготовлення однієї деталі в цеху.

### Розв'язання

Середня гармонійна використовується коли відомі дані про чисельник і невідомі про знаменник.

Витрати робочого часу на виготовлення протягом часу (8 годин роботи) складають:  $3 \cdot 8 \cdot 60 = 1440$  хвилин, або 480 хвилин на одного робітника.

Кількість деталей, що виготовляє один робітник:

– перший –  $480/30 = 16$  деталей;

– другий –  $480/15 = 32$  деталі;

– третій –  $480/10 = 48$  деталей.

В середньому на виготовлення 1 деталі витрачається:

$$x_{\text{гарм.}} = \frac{B_1 + B_2 + \dots + B_n}{\frac{B_1}{x_1} + \frac{B_2}{x_2} + \dots + \frac{B_n}{x_n}} = \frac{\sum B}{\sum \frac{B_i}{x_i}} = \frac{3 \cdot 8 \cdot 60}{\frac{480}{30} + \frac{480}{15} + \frac{480}{10}} = \frac{1440}{96} = 15 \text{ хвилини}$$

### Завдання для самостійної роботи

#### Завдання 1.

За статистичними даними розрахувати середній коефіцієнт частоти травматизму ( $Kч$ ) за даними, що наведені в табл. 3.

Таблиця 3 – Вихідні дані

Варіант	Розподіл коефіцієнту частоти травматизму ( $Kч$ ) за роками									
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1	4,4	3,3	3,4	3,5	3,7	3,8	3,6	3,2	3,9	3,1
2	7,4	6,3	6,4	6,5	6,7	6,8	6,6	6,2	6,9	6,1
3	6,4	5,3	5,4	5,5	5,7	5,8	5,6	5,2	5,9	5,1
4	3,4	2,3	2,4	2,5	2,7	2,8	2,6	2,2	2,9	2,1
5	3,4	7,3	7,4	7,5	7,8	7,7	7,6	7,2	7,1	7,9
6	1,3	2,4	1,4	1,5	1,7	1,8	1,6	1,2	1,9	1,1
7	3,4	2,3	4,4	2,5	2,7	4,3	2,6	2,2	2,9	3,1
8	3,4	1,3	1,2	1,5	1,7	3,3	1,6	4,2	1,9	5,1
9	3,4	2,3	2,4	2,5	6,1	2,8	2,6	2,7	2,9	4,1
10	4,4	3,3	3,4	3,5	7,1	3,8	3,6	3,7	3,9	5,1

#### Завдання 2.

Визначити середній коефіцієнт важкості травматизму по підприємствах регіону за наступними вихідними даними (табл. 4).

Таблиця 4 – Вихідні дані

Варіант	Коефіцієнт важкості травматизму ( $K\theta$ )				
	Кількість підприємств				
1	2	3	4	5	6
	2	3	4	5	6
2	2	3	4	5	6
	6	5	4	3	2
3	2	3	4	5	6
	1	3	2	4	6
4	2	3	4	5	6
	4	4	2	2	8
5	2	3	4	5	6
	3	5	4	2	2
6	1	2	3	4	5
	2	3	4	5	2
7	1	2	3	4	5
	6	5	4	3	2
8	1	2	3	4	5
	3	4	2	4	3
9	1	2	3	4	5
	4	4	2	2	8
10	1	2	3	4	5
	3	5	4	2	2

**Завдання 3.**

Троє працівників цеху протягом одного робочого дня зайняті виготовленням однакових деталей. Один робочий використовує на виготовлення деталі \_\_\_ хвилин, другий – \_\_\_ хвилин, третій – \_\_\_ хвилин. Визначити середні витрати часу на виготовлення однієї деталі. Дані наведені в табл. 5.

Таблиця 5 – Вихідні дані

Варіант	Витрати часу на виготовлення однієї деталі трьома робітниками, хвилин		
	1 робітник	2 робітник	3 робітник
1	24	20	15
2	30	15	5
3	24	15	10
4	20	15	8
5	15	12	10
6	5	30	15
7	25	30	15
8	20	30	8
9	5	10	30
10	40	10	20

**Завдання 4.**

За трьома підприємствам відома кількість коштів, що виділяється на охорону праці (у перерахунку на одного працівника) та загальна кількість грошей, що виділяється кожним керівником на підприємстві. Розрахувати середню кількість коштів, що виділяються на одного працівника на трьох підприємствах в цілому за наведеними даними згідно табл. 6.

Таблиця 6 – Вихідні дані

Варіант	Кількість коштів, що виділяються на 1 працівника на підприємстві, грош. одиниці		
	Загальна кількість коштів які виділяються на охорону праці на підприємстві керівником, грош. одиниці		
	Підприємство А	Підприємство Б	Підприємство В
1	2	3	4
1	28,5	28,62	28,75
	23240	29800	14900
2	25,7	26,68	27,73
	22247	25000	15000



1	2	3	4
3	27,7	24,68	23,73
	2000	26000	16000
4	25,7	26,68	24,73
	20000	26000	13000
5	18,7	22,68	38,73
	22247	25000	15000
6	12,7	22,68	28,73
	22247	25000	15000
7	10,5	20,4	33,0
	22200	28000	17000
8	10,5	20,4	30,0
	20000	23000	10000
9	7,0	8,0	15,0
	10000	17000	12000
10	3,0	8,3	20,1
	5000	7500	11000

### Контрольні питання

1. Зазначте, що характеризують абсолютні величини. В яких умовах вони застосовуються?
2. Що характеризують відносні величини? Назвіть види відносних величин.
3. Вкажіть, що таке середня величина.
4. Наведіть умови застосування середніх величин.
5. Які види середніх величин використовуються?
6. В яких випадках використовуються середня арифметична проста та зважена?
7. Наведіть формули для розрахунків середньої арифметичної простої та зваженої?

8. В яких випадках використовується середня гармонічна проста та зважена?

9. Наведіть формули для розрахунків середньої гармонічної простої та зваженої.

10. В яких випадках використовується середня геометрична?

11. Наведіть формулу для розрахунку середньої геометричної простої.

### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Громько Г. Л. Общая теория статистики: практикум \ Г. Л. Громько. Москва : Инфра-М, 2008. 240 с.
5. Минашкин В. Г. Теория статистики : учебно-методический комплекс / В. Г. Минашкин [и др.]. Москва : Издательство Центр ЕАОИ, 2008. 296 с.

## Практичне заняття 4

### РОЗРАХУНОК ХАРАКТЕРИСТИК ВАРІАЦІЙНИХ РЯДІВ

**Мета роботи:** Отримання навичок графічного представлення варіаційного ряду та розрахунку основних показників варіації.

#### Загальні положення

Різниця індивідуальних значень ознаки всередині досліджуваної сукупності називається *варіацією* ознаки. Середня величина – це абстрактна узагальнююча характеристика ознаки досліджуваної сукупності, але вона не показує будови сукупності. Для характеристики сукупностей і обчислених середніх величин важливо знати, яка варіація ознаки ховається за середніми. У деяких випадках окремі значення ознаки близько прилягають до середньої арифметичної і мало від неї відрізняються, в таких випадках середня добре представляє всю сукупність. В інших випадках, навпаки, окремі значення далеко відстоять від середньої, і середня погано представляє сукупність. Коливання окремих значень, ступінь їх близькості до середньої характеризують показники варіації.

*Ряд розподілу* – це впорядкований розподіл одиниць сукупності на групи за певною ознакою. Іншими словами, це групування, в якому для характеристики груп застосовується чисельність групи.

*Атрибутивні ряди розподілу* – ряди розподілу, побудовані за якісними ознаками.

*Варіаційні ряди розподілу* – ряди розподілу, побудовані за кількісними ознаками. Варіаційний ряд складається з двох елементів: варіанти і частоти.

*Варіанта* (позначається  $x$ ) – окреме значення варіюючої ознаки, яке вона приймає в ряді розподілу.

*Частота* (позначається  $f$ ) – чисельність окремих варіант, тобто частота повторення кожної варіанти.

Частота, виражена в частках одиниці або у відсотках до підсумку, називається *частість* (позначається  $\omega$ ).

За способом побудови варіаційні ряди бувають *дискретні* та *інтервальні*.

*Дискретний варіаційний ряд* характеризує розподіл одиниць сукупності за дискретною ознакою, що приймає тільки цілі значення. Для його побудови слід перерахувати всі варіанти значень ознаки і підрахувати частоту повторення.

При графічному зображенні дискретних варіаційних рядів використовується **полігон розподілу**, або **полігон частот**. Для його побудови в прямокутній системі координат на осі абсцис в однаковому масштабі відкладаються ранжоване значення варіюючої ознаки, а на осі ординат наноситься шкала для вираження величини частот. Отримані на перетині абсцис і ординат точки з'єднуються прямими лініями, в результаті чого отримують ламану лінію.

*Інтервальний варіаційний ряд* будується в разі безперервної варіації ознаки у одиниць сукупності (величина може приймати в певних межах будь-які значення, що відрізняються один від одного на як завгодно малу величину), а також в разі, коли число варіантів дискретного ознаки досить велике. Для графічного зображення інтервального варіаційного ряду застосовується **гістограма**.

При побудові гістограми на осі абсцис відкладаються величини інтервалів, а частоти зображуються прямокутниками, побудованими на відповідних інтервалах. Висота стовпчиків повинна бути пропорційна частотам. В результаті ми отримаємо графік, на якому ряд розподілу зображений у вигляді суміжних один з одним стовпчиків.

В ряді випадків для зображення варіаційних рядів (як дискретним, так і інтервальним) використовується кумулятивна крива (або **кумулята**). Для її побудови треба розрахувати накопичені частоти або частоті. Накопичені частоти (позначаються  $S$ ) показують, скільки одиниць сукупності мають значення ознаки не більше, ніж розглядається, і визначаються послідовним підсумовуванням частот інтервалів. При побудові кумуляти інтервального ряду розподілу нижній межі першого інтервалу відповідає частота, що дорівнює нулю, а верхньої межі – частота даного інтервалу.

### Абсолютні та середні показники варіації

Найбільш простий показник варіації – **розмах варіації**, який визначається як різниця між найбільшим ( $x_{\max}$ ) і найменшим ( $x_{\min}$ ) значеннями варіант:

$$R = x_{\max} - x_{\min} \quad (1)$$

Цей показник простий в обчисленні та вказує на загальні розміри варіації, але він не дає уявлення про ступінь коливання всередині сукупності, тому що вловлює тільки крайні відхилення.

Різниця всіх одиниць досліджуваної сукупності враховує **середнє лінійне відхилення**. Середнє лінійне відхилення є середня арифметична з відхилень індивідуальних значень від середньої (без урахування знака цих відхилень):

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} \text{ – для незгрупованих даних} \quad (2)$$

$$\bar{d} = \frac{\sum |x_i - \bar{x}| \cdot f}{\sum f} \text{ – для згрупованих даних} \quad (3)$$

На практиці міру варіації більш об'єктивно відображає показник дисперсії.

*Дисперсія* – це середня арифметична квадратів відхилень кожного значення ознаки від середньої арифметичної. Іншими словами, це середній квадрат відхилень. Дисперсія обчислюється за формулою:

$$\sigma^2 = \frac{\sum |x_i - \bar{x}|^2}{n} \text{ – для незгрупованих даних} \quad (4)$$

$$\sigma^2 = \frac{\sum |x_i - \bar{x}|^2 \cdot f}{\sum f} \text{ – для згрупованих даних} \quad (5)$$

Корінь квадратний з дисперсії являє собою *середньоквадратичне відхилення*. Його розраховують за формулою:

$$\sigma = \sqrt{\frac{\sum |x_i - \bar{x}|^2}{n}} \text{ – для незгрупованих даних} \quad (6)$$

$$\sigma = \sqrt{\frac{\sum |x_i - \bar{x}|^2 \cdot f}{\sum f}} \text{ – для згрупованих даних} \quad (7)$$

Перевагою цього показника є те, що він виражається в тих же одиницях виміру, що і ознака.

Дисперсія і середньоквадратичне відхилення є основними узагальнюючими показниками варіації. Чим менше середнє квадратичне відхилення, тим краще середня арифметична відображає всю подану сукупність.

### **Відносні показники варіації**

Відносні показники варіації дозволяють порівнювати характер розсіювання в різних сукупностях, наприклад, при порівнянні різнойменних сукупностей, при різних значеннях середньої. Розрахунок відносних показників варіації здійснюють як відношення абсолютного показника варіації до середньої арифметичної. Як правило, вони розраховуються у відсотках.

**Коефіцієнт осциляції** відображає відносну коливання крайніх значень навколо середньої:

$$V_R = \frac{R}{x} \times 100\% \quad (8)$$

**Лінійний коефіцієнт варіації** характеризує частку усередненого значення абсолютних відхилень від середньої величини:

$$V_{\bar{d}} = \frac{\bar{d}}{x} \times 100\% \quad (9)$$

**Квадратичний коефіцієнт варіації** – найбільш поширений показник коливання, використовуваний для оцінки типовості середньої:

$$V_{\sigma} = \frac{\sigma}{x} \times 100\% \quad (10)$$

За величиною  $V_{\sigma}$  розрізняють такі значення відносних коливань:

$V_{\sigma} < 10\%$  – незначне коливання;

$V_{\sigma}$  – від 10 % до 30 % – середнє коливання;

$V_{\sigma} > 30\%$  – велике коливання.

Чим більше розкид значень ознаки навколо середньої, тим більше коефіцієнт варіації і тим менш середня представляє сукупність. Як правило, вважають, що якщо  $V_{\sigma} > 33\%$ , то це говорить про великий коливання ознаки в сукупності, і сукупність неоднорідна.

### Приклади розв'язання задач

#### Приклад 1.

Двадцять підприємств однієї галузі в області мають різну кількість філій. Кількість філій в області у різних підприємств дорівнює: 2; 4; 3; 5; 4; 4; 6; 5; 4; 3; 4; 3; 4; 5; 3; 4; 6; 3; 5; 4. Побудувати ряд розподілу підприємств за кількістю філій в області за наявними даними. Дати графічне зображення ряду розподілу.

#### Розв'язання

Варіація ознаки носить дискретний характер, число варіант дискретної ознаки невелике, значення ознаки в окремих одиниць сукупності повторюються. Тому будується дискретний ряд розподілу. Для його побудови слід записати значення признака (кількість філій) в порядку збільшення, перерахувати всі варіанти значень ознаки і підрахувати частоту повторення.

Для побудови дискретного ряду розподілу, виконаємо необхідні розрахунки у вигляді табл. 1.

Частість  $\omega$  розраховується як відношення відповідної частоти до загальної суми частот:

$$\omega = \frac{f_i}{\sum_{i=1}^n f_i} \quad (11)$$

Таблиця 1 – Результати розрахунку

Кількість філій, (варіанта $x$ )	Число підприємств (частота, $f$ )	Частість, $\omega$	Накопичена частота, $S$
2	1	1/20=0,05	1
3	5	5/20=0,25	1 + 5 = 6
4	8	8/20=0,40	6 + 8 = 14
5	4	4/20=0,20	14 + 4 = 18
6	2	2/20=0,10	18 + 2 = 20
Всього	20	1,00	

За отриманим дискретним рядом розподілу будується полігон частот:  $x - f$  та кумулята  $x - S$ . Полігон частот побудовано на рис. 1.

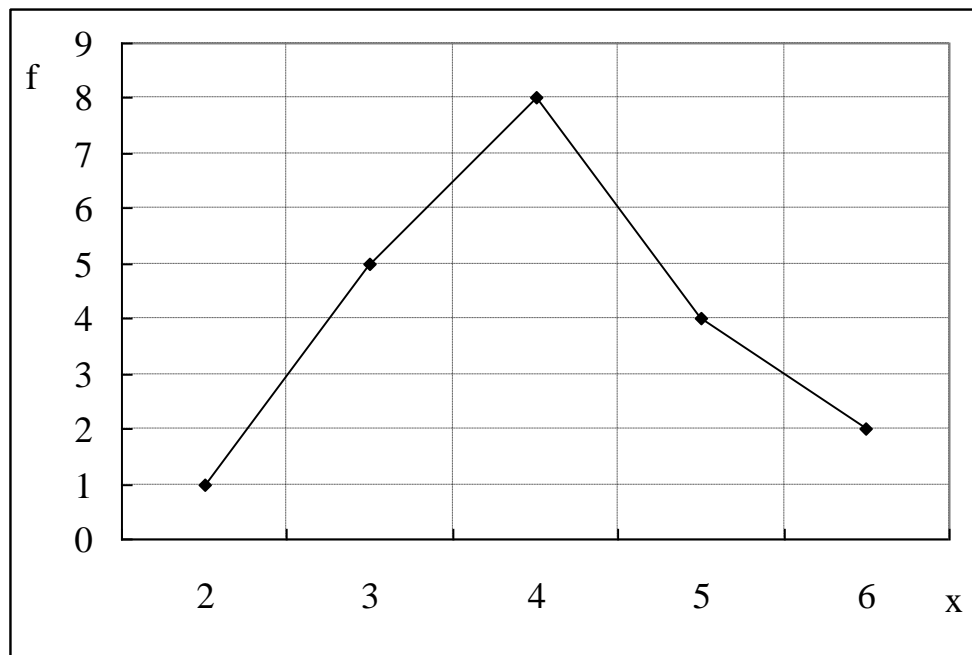


Рисунок 1 – Полігон частот

Для побудови кумуляти слід розрахувати накопичені частоти  $S$ . Накопичена частота першої варіанти дорівнює частоті першої варіанти, тобто всього 1 підприємство в області має не більше двох філій.



Накопичена частота другої варіанти дорівнює сумі частот першої і другої варіанти (або сумі накопиченої частоти першої варіанти і частоти другої варіанти), тобто не більш трьох філій мають 6 підприємств в області: у п'яти з них по 3 філії, у одного – 2 філії. Решта накопичених частот визначаються аналогічно. Накопичена частота останньої варіанти дорівнює сумі всіх частот ряду: всі підприємства в області мають не більше 6 філій. Побудована кумулята показана на рис. 2.

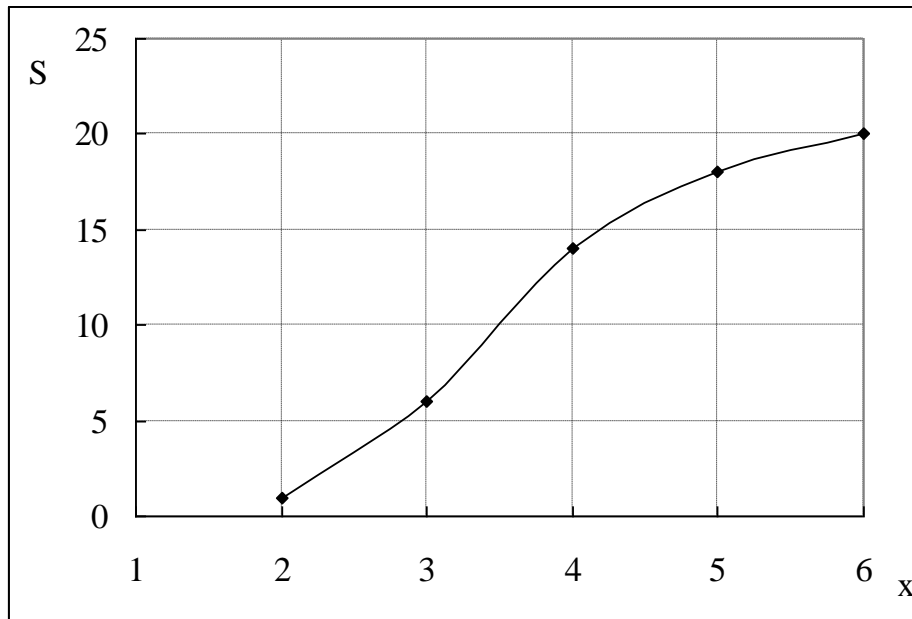


Рисунок 2 – Кумулята

### Приклад 2.

Є такі дані про розмір коштів двадцяти металургійних підприємств, що виділяються на охорону праці протягом року: 3,7; 4,3; 6,7; 5,6; 5,1; 8,1; 4,6; 5,7; 6,4; 5,9; 5,2; 6,2; 6,3; 7,2; 7,9; 5,8; 4,9; 7,6; 7,0; 6,9 млн. грош. одиниць. Побудувати ряд розподілу за наявними даними. Дати графічне зображення ряду розподілу.

### Розв'язання

Варіація ознаки носить безперервний характер, значення ознаки в окремих одиниць сукупності не повторюються. Тому будується інтервальний ряд розподілу. Для його побудови слід визначити кількість інтервалів і величину інтервалу.

Оскільки кількість інтервалів заздалегідь не задано, визначимо її

за формулою Стюрджеса:

$$n = 1 + 3,322 \cdot \lg 20 = 1 + 3,322 \cdot 1,3 = 5,3$$

Дробове число, що характеризує кількість інтервалів, бажано округляти в меншу сторону. Приймаємо  $n = 5$ .

Оскільки розмах варіації  $R$  невеликий, то утворимо групи з рівним інтервалами. Визначимо величину інтервалу:

$$h = \frac{X_{\max} - X_{\min}}{n} = (8,1 - 3,7) / 5 = 0,88$$

Число, що характеризує величину інтервалу, округляється з тією ж точністю, що і вихідні дані. У нашому випадку слід округлити до 0,1:  $h = 0,9$ .

Побудуємо інтервальний ряд розподілу у вигляді табл. 2.

Таблиця 2 – Розподіл підприємств за розміром коштів що виділяються на охорону праці

№ групи	Групи за розміром коштів ( $x$ )	Число підприємств ( $f$ )	Частість ( $\omega$ )	Накоплена частота ( $S$ )
1	3,7 – 4,6	3	3/20=0,15	3
2	4,6 – 5,5	3	3/20=0,15	6
3	5,5 – 6,4	7	7/20=0,35	13
4	6,4 – 7,3	4	4/20=0,2	17
5	7,3 – 8,2	3	3/20=0,15	20
Всього		20	1	

При підрахунку частот скористаємося принципом «включно», згідно з яким одиниця сукупності, що має значення ознаки, рівне кордоні двох суміжних груп (наприклад, підприємство з розміром коштів 4,6 млн. грош. одиниць), включається в інтервал, де він служить верхньою межею (підприємство з розміром коштів 4,6 млн. грн. включимо в групу з розміром коштів від 3,7 до 4,6 млн. грош. одиниць).

Розрахунок частостей і накопичених частот виконується аналогічно розрахунку в дискретних рядах розподілу. За отриманими значеннями частот для інтервального ряду будується гістограма розподілу, що зображена на рис. 3.

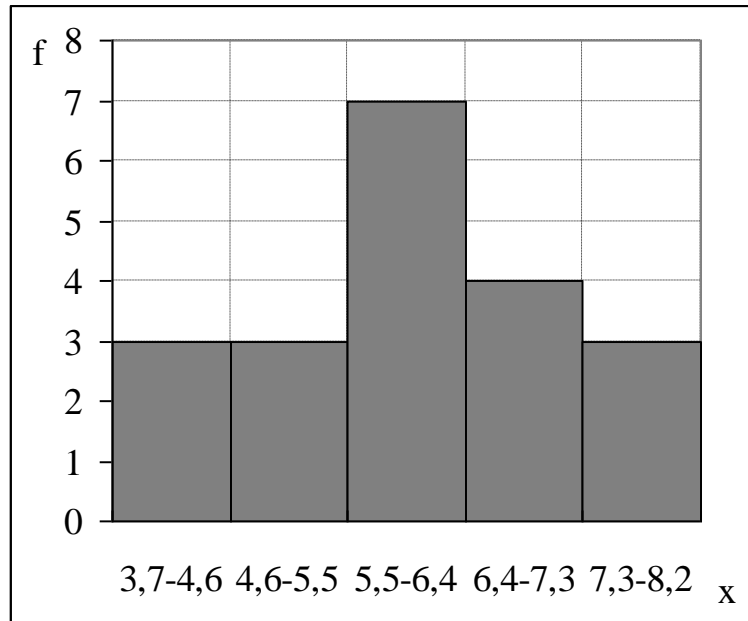


Рисунок 3 – Розподіл коштів, що виділяються на охорону праці за підприємствами

За накопиченими частотами побудуємо кумуляту (рис. 4).

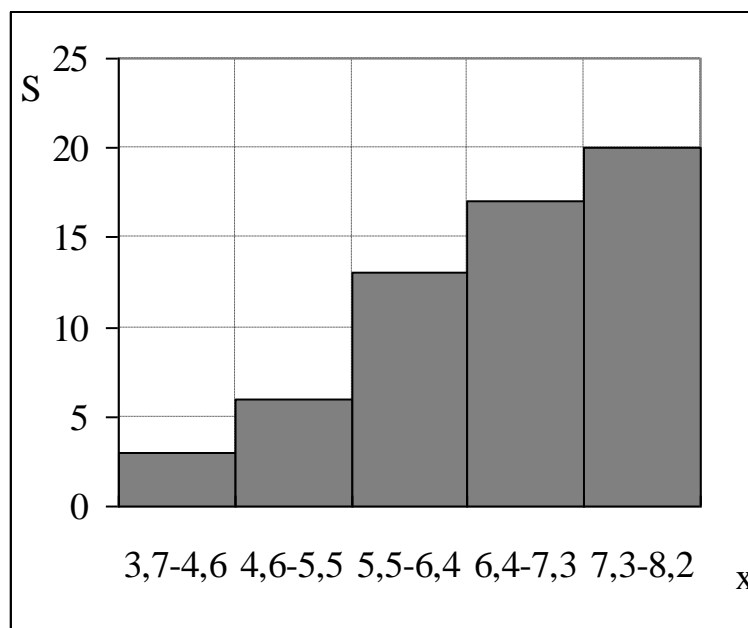


Рисунок 4 – Кумулята збільшення частот груп підприємств

**Приклад 3.**

За наведеними у вигляді ряду даними значень коефіцієнта частоти травматизму ( $Kч$ ) розрахувати абсолютні та відносні показники варіації.  $Kч$  для 10 статистичних одиниць дорівнює: 4,0; 4,1; 4,2; 4,3; 4,4; 4,5; 4,6; 4,7; 4,8; 4,9. Зробити висновок про однорідність сукупності.

**Розв'язання**

Абсолютні показники варіації:

$$R = x_{max} - x_{min} = 4,9 - 4,0 = 0,9.$$

Для розрахунку інших показників варіації виконаємо додаткові розрахунки у вигляді таблиці. Для цього спочатку знайдемо середнє арифметичне:

$$\bar{x} = \frac{\sum_{i=1}^n x}{n} = \frac{(4,0 + 4,1 + 4,2 + 4,3 + 4,4 + 4,5 + 4,6 + 4,7 + 4,8 + 4,9)}{10} = 4,45 \approx 4,5$$

Результати розрахунків наведені в табл. 3

Таблиця 3 – Результати розрахунків

Коефіцієнт частоти травматизму ( $Kч$ )	$ x_i - \bar{x} $	$ x_i - \bar{x} ^2$
4,0	0,5	0,25
4,1	0,4	0,16
4,2	0,3	0,09
4,3	0,2	0,04
4,4	0,1	0,01
4,5	0	0
4,6	0,1	0,01
4,7	0,2	0,04
4,8	0,3	0,09
4,9	0,4	0,16
Всього	2,5	0,85

Оскільки є окремі значення ознаки, дані не згруповані, застосуємо невиважені формули показників варіації.

Абсолютні показники варіації:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{2,5}{10} = 0,25$$

$$\sigma^2 = \frac{\sum |x_i - \bar{x}|^2}{n} = \frac{0,85}{10} = 0,085$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{0,085} \approx 0,292$$

Відносні показники варіації:

$$V_R = \frac{R}{x} \times 100\% = \frac{0,9}{4,5} * 100\% = 20,0\%$$

$$V_{\bar{d}} = \frac{\bar{d}}{x} \times 100\% = \frac{0,25}{4,5} * 100\% \approx 5,56\%$$

$$V_{\sigma} = \frac{\sigma}{x} \times 100\% = \frac{0,292}{4,5} * 100\% \approx 6,48\%$$

В результаті розрахунків коливання ознаки в сукупності невелике, оскільки  $V_{\sigma} = 6,48\% < 10\%$ .

Оскільки  $V_{\sigma} < 33\%$ , то сукупність можна вважати однорідною за цією ознакою.

### Завдання для самостійної роботи

#### Завдання 1.

Двадцять підприємств однієї галузі в Харківській області мають різну кількість філій. Кількість філій в області у різних підприємств дорівнює згідно табл. 4.

Побудувати ряд розподілу підприємств за кількістю філій в області за наявними даними. Дати графічне зображення ряду розподілу.

Таблиця 4 – Розподіл підприємств за кількістю філій

№ підприємства	Кількість філій в Харківській області									
	Варіант									
	1	2	3	4	5	6	7	8	9	10
1	4	6	3	3	2	3	3	2	4	6
2	5	2	2	2	2	5	3	5	6	5
3	3	3	6	2	3	4	4	5	5	3
4	5	6	6	2	2	5	5	6	5	3
5	4	5	5	5	6	4	2	4	4	6
6	4	4	4	4	4	4	5	4	2	4
7	6	6	6	6	5	5	6	2	2	3
8	5	5	5	5	5	6	3	2	4	3
9	4	4	4	4	5	5	5	4	3	4
10	3	3	3	3	2	3	5	3	2	3
11	4	4	2	2	3	4	5	3	2	3
12	2	2	4	4	4	2	3	3	3	1
13	4	4	6	6	6	4	4	5	3	3
14	5	5	4	4	4	6	5	6	6	2
15	3	3	6	6	6	3	5	5	4	6
16	4	4	4	4	4	4	4	5	5	4
17	3	6	3	3	3	2	6	5	5	5
18	6	3	3	3	3	6	5	5	5	6
19	5	4	4	4	4	5	4	4	3	5
20	2	5	3	3	3	2	2	5	3	1

**Завдання 2.**

Дані про розмір коштів двадцяти машинобудівних підприємств, що були виділені на заходи з охорони праці протягом 2010 року наведені в табл. 5.

Побудувати ряд розподілу за наявними даними. Дати графічне зображення ряду розподілу.

Таблиця 5 – Вихідні дані

№ підприємства	Розмір коштів, що виділяються на охорону праці, млн. грош. одиниць									
	Варіант									
	1	2	3	4	5	6	7	8	9	10
1	3,1	3,0	1,0	1,0	2,0	3,1	3,0	1,0	1,0	1,0
2	3,3	3,2	1,4	1,2	2,5	3,3	3,2	1,4	1,2	1,2
3	3,5	3,4	1,6	1,4	3,0	3,5	3,4	1,6	1,4	1,4
4	3,8	3,6	2,0	1,6	3,5	3,8	3,6	2,0	1,6	1,6
5	4,0	4,1	2,4	1,8	4,0	4,0	4,1	2,4	1,8	2,0
6	4,2	4,2	2,6	2,0	4,5	4,2	4,2	2,6	2,0	2,5
7	4,5	4,4	3,0	2,2	5,0	4,5	4,4	3,0	3,0	3,0
8	4,8	4,6	3,4	2,4	5,3	4,8	4,6	3,4	3,4	3,5
9	5,0	5,3	3,8	2,6	5,6	5,0	5,3	3,3	3,8	4,0
10	5,3	5,4	4,2	3,0	5,9	5,3	5,4	3,6	4,2	4,5
11	5,8	5,6	4,6	3,3	6,2	5,6	5,8	3,9	4,6	5,0
12	5,9	6,0	5,0	3,6	6,4	6,0	5,9	4,2	5,0	5,3
13	6,0	6,5	5,4	3,9	6,8	6,5	6,0	4,5	5,4	5,5
14	7,2	6,7	5,8	4,2	7,0	6,7	7,2	4,8	5,8	5,8
15	7,5	8,0	6,2	4,5	7,2	8,0	7,5	5,2	6,2	6,0
16	7,7	8,1	6,7	4,8	7,5	8,1	7,7	5,5	6,7	6,4
17	7,8	7,0	7,2	5,2	7,9	7,0	7,8	6,0	7,2	6,5
18	7,9	7,5	7,6	5,5	8,0	7,5	7,9	7,0	7,6	6,8
19	8,1	7,9	8,4	6,0	8,1	7,9	8,2	7,5	8,4	7,0
20	8,2	7,7	8,2	7,0	7,7	7,7	8,1	8,0	8,2	7,2

**Завдання 3.**

За наявними даними коефіцієнта частити травматизму ( $Kч$ ), що наведені в табл. 6 розрахувати абсолютні та відносні показники варіації для даного варіаційного ряду.

Таблиця 6 – Вихідні дані

Варіант	Розподіл коефіцієнта частоти травматизму ( $Kч$ ) за роками									
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9
2	2,0	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9
3	1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9
4	1,0	1,2	1,4	1,6	1,8	2,0	2,2	2,4	2,6	2,8
5	2,0	2,2	2,4	2,6	2,8	3	3,2	3,4	3,6	3,8
6	1,0	1,3	1,6	1,9	2,0	2,1	2,2	2,5	2,6	3,0
7	1,0	1,3	1,4	1,5	2,0	2,1	2,2	2,3	2,6	2,8
8	2,0	2,3	2,5	2,6	2,7	3,0	3,2	3,4	3,6	4,0
9	2,0	2,4	2,8	3,2	3,6	3,8	4,0	4,2	4,4	4,6
10	1,7	2,0	2,4	2,8	3,0	3,2	3,4	3,6	3,8	4,0

### Контрольні питання

1. Вкажіть що таке варіація ознаки та від чого залежить її розмах.
2. Наведіть формулу для розрахунку середнього лінійного відхилення.
3. Наведіть формулу для розрахунку середнього квадратичного відхиленням.
4. Наведіть формулу для розрахунку середнього квадрата відхилення (дисперсії).
5. Які існують коефіцієнти варіації?
6. Що відображає коефіцієнт осциляції та як він розраховується?
7. Що характеризує лінійний коефіцієнт варіації та як він розраховується?
8. Для чього використовується квадратичний коефіцієнт варіації та як він розраховується?
9. Як побудувати полігон частот?
10. Як побудувати кумуляту?
11. Як побудувати гістограму?
12. Про що свідчить значення коефіцієнта квадратичного коефіцієнта варіації  $V_\sigma < 33\%$ ?



### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Громько Г. Л. Общая теория статистики: практикум \ Г. Л. Громько. Москва : Инфра-М, 2008. 240 с.
5. Минашкин В. Г. Теория статистики : учебно-методический комплекс / В. Г. Минашкин [и др.]. Москва : Издательство Центр ЕАОИ, 2008. 296 с.

## Практичне заняття 5

### АНАЛІЗ ФОРМИ РОЗПОДІЛУ ВАРІАЦІЙНИХ РЯДІВ

**Мета роботи:** Отримання навичок аналізу форми розподілу варіаційних рядів.

#### Загальні положення

В варіаційному ряді в співвідношенні варіант і частот проявляється *закономірність розподілу*, яка описується різними статистичними показниками:

- частотні показники;
- центр розподілу;
- центр варіації;
- форма розподілу.

*Частотними показниками* будь-якого ряду розподілу є абсолютна чисельність  $i$ -ї групи – частота  $f_i$ , відносна частота – частість  $d_i$ , де

$$\sum_{i=1}^m f_i = n; \sum_{i=1}^m d_i = 1.$$

*Кумулятивна (накопичена) частота  $S_i$  (частість  $S_d$ )* характеризує обсяг сукупності зі значеннями варіантів, що не перевищують  $X_i$ . Кумулятивні частотні показники утворюються послідовним підсумовуванням абсолютних або відносних частот, наприклад:  $S_1 = f_1$ ;  $S_2 = f_1 + f_2$ ;  $S_3 = f_1 + f_2 + f_3$ .

До показників центра розподілу відносять *середню, моду й медіану*.

*Мода* – це значення ознаки, що найчастіше зустрічається в досліджуваній сукупності, тобто варіанта, що у ряді розподілу має найбільшу частоту (частість).

У *дискретному ряді* мода визначається візуально за максимальною частотою або частістю.

Мода в *інтервальному ряді* може бути розрахована аналітично за формулою:

$$M_o = x_0 + h_0 \cdot \frac{f_{M_0} - f_{M_{0-1}}}{(f_{M_0} - f_{M_{0-1}}) + (f_{M_0} - f_{M_{0+1}})}, \quad (1)$$

де  $x_0, h_0$  – нижня границя й ширина модального інтервалу відповідно;

$f_{M_0}, f_{M_{0-1}}, f_{M_{0+1}}$  – частоти модального, попереднього перед модальним і наступного за модальним інтервалів відповідно.

*Медіана* – це значення ознаки в сукупності, що ділить ранжируваний ряд навпіл: половина варіант має значення, менші за медіани, а половина – значення, більші за медіани.

У *дискретному ряді* медіаною буде значення ознаки, для якого кумулятивна частота  $S_i$  дорівнює або перевищує половину обсягу сукупності  $\sum_{i=1}^m f_i$  або кумулятивна частість  $S_d \geq 0,5$ .

В *інтервальному ряді* таким способом визначається медіанний інтервал. Конкретне значення медіани обчислюється за формулою:

$$Me = x_0 + h_e \cdot \frac{0,5 \sum_{i=1}^m f_i - S_{Me-1}}{f_{Me}}, \quad (2)$$

де  $x_0$  – нижня границя медіанного інтервалу;

$h_e$  – значення медіанного інтервалу;

$f_{Me}$  – частота медіанного інтервалу;

$S_{Me-1}$  – накопичена частота в інтервалі, що передує медіанному інтервалу.

Для аналізу варіаційних рядів використовуються такі характеристики, як *моменти розподілу*. *Момент розподілу  $k$ -го порядку* – це середня арифметична  $k$ -го степеня відхилень окремих значень варіант від якої-небудь постійної величини  $A$ :

$$M_k = \frac{\sum (X_i - A)^k f_i}{\sum f_i}. \quad (3)$$

Якщо  $A = 0$ , то момент – початковий, якщо  $A = \bar{X}$ , то момент – центральний.

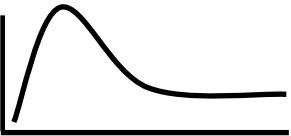
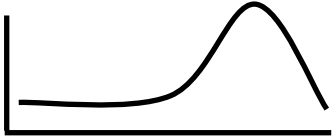
У ряді випадків, коли необхідно порівнювати моменти для різних рядів, розраховують *нормовані моменти*. *Нормовані моменти* – це відповідні моменти, що діляться на середньоквадратичні відхилення (СКВ) в  $k$ -му степені.

Варіаційний ряд, крім характеристик центра та розкиду, може бути описаний показниками, що визначають форму розподілу. До таких показників належать *показники асиметрії та ексцесу*.

### Асиметрія

Показники асиметрії подано у табл. 1.

Таблиця 1 – Показники асиметрії

Показники асиметрії $A$	Правостороння асиметрія $Mo < Me < \bar{X}$	Лівостороння асиметрія $\bar{X} < Me < Mo$
		
1. Показник Пірсона: $\frac{\bar{X} - Mo}{\sigma}$	додатне значення	від'ємне значення
2. Показник Ліндберга: $W^* - 50$	від'ємне значення	додатне значення
3. Нормований момент 3-го порядку: $\frac{\mu_3}{\sigma^3}$	додатне значення	від'ємне значення

\*  $W$  – це частка значень ознаки, що перевищують  $\bar{X}$ .

Найчастіше застосовують такий показник асиметрії, як *нормований центральний момент 3-го порядку*:

$$A_s = \frac{\mu_3}{\sigma^3}, \quad (4)$$

$$\text{де } \mu_3 = \frac{\sum (X_i - \bar{X})^3 f_i}{\sum f_i}.$$

Вважається, що при:

- \_  $A_s < 0,25$  – асиметрія незначна;
- \_  $0,25 < A_s < 0,5$  – асиметрія середня;
- \_  $A_s > 0,5$  – асиметрія значна.

Для симетричного розподілу мода, медіана й середня арифметична збігаються, показник асиметрії дорівнює нулю.

### Екссес

*Екссес* – це показник гостро- або пласковершинності розподілу.

Виділяють такі показники екссесу:

1. Показник Ліндберга:  $E_X = P - 38,29$ , де  $P$  – це частка варіант, що потрапляють в інтервал  $X \pm \sigma/2$  (у відсотках).

2. Показник, заснований на нормованому центральному моменті 4-го порядку:  $E_x = \frac{\mu_4}{\sigma^4} - 3$ .

### Приклади розв'язання задач

#### Приклад 1.

Розподіл родин міста за числом дітей характеризується такими даними (табл. 2):

Таблиця 2 – Розподіл родин міста за числом дітей

Число дітей у родині	0	1	2	3	4	5	Разом
Число родин, % до підсумку	10	26	29	17	13	5	100

Визначте коефіцієнти асиметрії й екссесу, використовуючи

центральні моменти перших чотирьох порядків. Зробіть висновки.

### Розв'язання

Складемо розрахункову таблицю (табл. 3).

Таблиця 3 – Розрахункова таблиця

Число дітей у родині $x$	Число родин $f_i$ , % до підсумку	$x_i \cdot f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2 \cdot f_i$	$(x_i - \bar{x})^3 \cdot f_i$	$(x_i - \bar{x})^4 \cdot f_i$
0	10	0	-2,12	44,944	-95,2813	201,9963
1	26	26	-1,12	32,6144	-36,5281	40,9115
2	29	58	-0,12	0,4176	-0,05011	0,006013
3	17	51	0,88	13,1648	11,58502	10,19482
4	13	52	1,88	45,9472	86,38074	162,3958
5	5	25	2,88	41,472	119,4394	343,9854
Разом	100	212	–	178,56	85,5456	759,4898

1) Знайдемо середнє арифметичне використовуючи графу 3:

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{212}{100} = 2,12.$$

2) Обчислимо дисперсію, тобто центральний момент другого порядку (графу 5):

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 \cdot f}{\sum f} = \frac{178,56}{100} = 1,7856.$$

3) Знайдемо середнє квадратичне відхилення (стандарт):

$$\sigma = \sqrt{\sigma^2} = \sqrt{1,7856} = 1,336.$$

4) Визначимо центральний момент третього порядку використовуючи графу 6:

$$\mu_3 = \frac{\sum (x - \bar{x})^3 \cdot f}{\sum f} = \frac{85,5456}{100} = 0,855.$$

5) Визначимо нормований момент третього порядку:

$$A_s = \frac{\mu_3}{\sigma^3} = \frac{0,855}{1,336^3} = 0,3585 > 0 \text{ – правостороння середня асиметрія.}$$

б) Знайдемо моду для знаходження коефіцієнта асиметрії Пірсона:

$$Mo = 2.$$

7) Розрахуємо коефіцієнт асиметрії Пірсона:

$$A = \frac{\bar{x} - Mo}{\sigma} = \frac{2,12 - 2}{1,336} = 0,09 > 0 \text{ – невелика правостороння асиметрія.}$$

8) Обчислимо центральний момент четвертого порядку користуючись графою 7:

$$M_4 = \frac{\sum (x - \bar{x})^4 \cdot f}{\sum f} = \frac{759,5}{100} = 7,595.$$

9) Визначимо нормований момент четвертого порядку:

$$A_4 = \frac{M_4}{\sigma^4} = \frac{7,595}{1,336^4} = 2,38.$$

10) Знайдемо ексцес розподілу:

$$E_x = A_4 - 3 = 2,38 - 3 = -0,62.$$

Оскільки  $E_x < 0$ , то розподіл пласковершинний.

### Приклад 2.

Визначити коефіцієнт асиметрії й ексцесу, а також нормовані моменти третього й четвертого порядку за даними про розподіл працюючих підприємства за кількістю вироблених деталей протягом дня згідно табл. 4.

Таблиця 4 – Розподіл працюючих машинобудівного підприємства за кількістю вироблених деталей протягом дня

Кількість вироблених деталей $x$ , штук	Число працівників
50–60	7
60–70	15
70–80	6
80–90	4
Разом	32

### Розв'язання

Складаємо таблицю (табл. 5).

Таблиця 5 – Розрахункова таблиця

Кількість вироблених деталей $x$ , штук	Число працівників $f_i$	Середина інтервалу $x_i$	$x_i \cdot f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2 \cdot f_i$	$(x_i - \bar{x})^3 \cdot f_i$	$(x_i - \bar{x})^4 \cdot f_i$
50–60	7	55	385	–12,2	1039,746	–12671,91	154438,8
60–70	15	65	975	–2,2	71,77734	–157,0129	343,466
70–80	6	75	450	7,8	366,2109	2861,023	22351,74
80–90	4	85	340	17,8	1269,141	22606,57	402679,5
Разом	32	–	2150	–	2746,88	12638,67	579814,54

1) Обчислимо середнє арифметичне (графа 3):



$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{2150}{32} = 67,19 \text{ деталей.}$$

2) Знайдемо дисперсію, тобто центральний момент другого порядку (графа 5):

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{2746,88}{32} = 85,84.$$

3) Знайдемо середнє квадратичне відхилення (стандарт):

$$\sigma = \sqrt{\sigma^2} = \sqrt{85,84} = 9,265.$$

4) Визначимо центральний момент третього порядку (графа 6):

$$\mu_3 = \frac{\sum (x - \bar{x})^3 f}{\sum f} = \frac{12638,67}{32} = 394,96.$$

5) Визначимо нормований момент третього порядку:

$$A_s = \frac{\mu_3}{\sigma^3} = \frac{394,96}{9,265^3} = 0,497 > 0 \text{ – правостороння асиметрія.}$$

6) Обчислимо моду для знаходження коефіцієнта асиметрії Пірсона:

$$Mo = 60 + 10 \cdot \frac{15 - 7}{2 \cdot 15 - 7 - 6} = 64,71 \text{ деталей.}$$

7) Знайдемо коефіцієнт асиметрії Пірсона:

$$A = \frac{\bar{x} - Mo}{\sigma} = \frac{67,19 - 64,71}{9,265} = 0,268.$$

У цьому випадку асиметрія невелика й скошеність правостороння.

8) Визначимо центральний момент четвертого порядку (графа 7):

$$M_4 = \frac{\sum (x - \bar{x})^4 f}{\sum f} = \frac{579814,56}{32} = 18119,2.$$

9) Знайдемо нормований момент четвертого порядку:

$$A_4 = \frac{M_4}{\sigma^4} = \frac{18119,2}{9,265^4} = 2,459.$$

10) Визначимо ексцес розподілу:

$$E_x = A_4 - 3 = 2,459 - 3 = -0,541.$$

Оскільки  $E_x < 0$ , то розподіл пласковершинний.

### **Завдання для самостійної роботи**

#### **Завдання 1.**

Розподіл працівників машинобудівного підприємства за кількістю вироблених деталей протягом робочої години характеризується даними згідно табл. 6.

Визначте коефіцієнти асиметрії й ексцесу, використовуючи центральні моменти перших чотирьох порядків. Зробіть висновки.

Таблиця 6 – Розподіл працівників машинобудівного підприємства за кількістю вироблених деталей

Кількість вироблених деталей, штук	Кількість робітників підприємства									
	Варіант									
	1	2	3	4	5	6	7	8	9	10
1						5				
2						9				
3						10			4	
4						12			9	
5						9		3	11	
6						7		5	8	
7						3		8	7	
8	5		4		5			13	6	
9	6		5		10			10	5	
10	8		6		12		4	7		
11	10		15		9		5	4		
12	13		20		9		8			
13	10		6		6		15			4
14	3	5	4	4	4		20			5
15		10		5			9			6
16		12		7			4			20
17		9		8						15
18		8		13						6
19		5		10						4
20		1		3						

### Завдання 2.

Розподіл працівників машинобудівного підприємства за віком наведений згідно табл. 7.

Таблиця 7 – Розподіл працівників машинобудівного підприємства за віком

Вік робітника $x$ , років	Кількість робітників підприємства									
	Варіант									
	1	2	3	4	5	6	7	8	9	10
20–30	2	3	4	10	7	3	5	4	3	7
30–40	8	12	8	15	9	8	9	10	8	10
40–50	15	11	15	12	15	20	14	16	12	20
50–60	10	9	12	10	10	12	10	12	7	13
60–70	5	5	5	3	4	7	7	8	5	5

Визначте коефіцієнти асиметрії й ексцесу, використовуючи центральні моменти перших чотирьох порядків. Зробіть висновки.

### Контрольні питання

1. Що таке “мода” і як вона визначається в дискретному ряді?
2. Як визначається “мода” в інтервальному ряді?
3. Що таке “медіана” і як вона визначається в дискретному ряді?
4. Як визначається “медіана” в інтервальному ряді?
5. Що таке “момент розподілу  $k$ -го порядку” і як він розраховується?
6. Що таке нормовані моменти і коли як вони розраховуються?
7. Які показники асиметрії варіаційного ряду існують?
8. Як розраховується такий показник асиметрії, як нормований центральний момент 3-го порядку?
9. Як розраховується коефіцієнт асиметрії Пірсона?
10. Що таке “ексцес” і які показники ексцесу виділяють?

### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.

## **Практичне заняття 6**

### **ВИЗНАЧЕННЯ СТАНДАРТНОЇ ТА ГРАНИЧНОЇ ПОМИЛКИ ВИПАДКОВОЇ ВИБІРКИ, МЕЖ ДОВІРЧИХ ІНТЕРВАЛІВ ТА НЕОБХІДНОЇ КІЛЬКОСТІ ВИБІРКИ**

**Мета роботи:** Отримання навичок визначення стандартної та граничної помилки випадкової вибірки, визначення меж довірчих інтервалів та необхідної кількості вибірки.

#### **Загальні положення**

Статистичне дослідження може здійснюватися за даними несучільного спостереження, основна мета якого полягає в отриманні характеристик досліджуваної сукупності за обстеженою її частиною. Одним з найбільш поширених методів, які застосовують несучільне спостереження, є вибірковий метод.

Під *вибірковим* розуміється метод статистичного дослідження, при якому узагальнюючі показники досліджуваної сукупності встановлюються за деякою її частиною на основі положень випадкового відбору. При вибірковому методі обстеження піддається порівняно невелика частина всієї досліджуваної сукупності (зазвичай до 5-10 %, рідше до 15-25 %). Статистична сукупність, з якої проводиться відбір частини одиниць, які підлягають вивченню, називається *генеральною сукупністю*. Відібрана з генеральної сукупності деяка частина одиниць, піддається обстеженню, називається *вибірковою сукупністю* (або *вибіркою*).

Значення вибіркового методу полягає в тому, що при меншій чисельності обстежуваних одиниць проведення дослідження здійснюється з меншими витратами і в коротші терміни, підвищуючи оперативність статистичної інформації.

Оскільки досліджувана статистична сукупність складається з одиниць з варуючими ознаками, то склад вибіркової сукупності може в тій чи іншій мірі відрізнятись від складу генеральної сукупності. Це об'єктивно викликає розбіжність між характеристиками вибірки та

генеральної сукупності та становить помилку вибірки. Вона залежить від ряду факторів:

- ступеня варіації досліджуваної ознаки,
- чисельності вибірки,
- методів відбору одиниць у вибіркочну сукупність,
- прийнятого рівня достовірності результату дослідження.

Способи визначення помилки вибірки при різних прийомах формування вибіркової сукупності і поширення характеристик вибірки на генеральну сукупність складають основний зміст статистичної методології вибіркового методу.

### **Характеристики вибіркової сукупності і їх поширення на генеральну сукупність**

При використанні вибіркового методу в соціально-економічних дослідженнях зазвичай застосовують два основних види узагальнюючих показників: відносну величину альтернативної ознаки і середню величину кількісної ознаки.

Відносна величина альтернативної ознаки характеризує частку (питома вага) одиниць у статистичній сукупності, які відрізняються від всіх інших одиниць цієї сукупності тільки наявністю (відсутністю) досліджуваної ознаки.

*Середня величина кількісної ознаки* – це узагальнююча характеристика варуючої ознаки, яка має різні значення в окремих одиницях статистичної сукупності. Наприклад, середня вага виробу, середня кількість вироблених деталей працівником і т.д.

У генеральній сукупності частка одиниць, які мають досліджувану ознаку, називається *генеральною часткою* (позначається  $P$ ), а середня величина варуючої ознаки – *генеральна середня* (позначається  $\bar{x}$ ).

У вибірковій сукупності частку досліджуваної ознаки називають *вибірковою часткою*  $\omega$ , а середню величину в вибірці – *вибірковою середньою*  $\tilde{x}$ .

Вибіркова частка визначається з відношення одиниць, які мають досліджувану ознаку  $t$ , до загальної чисельності одиниць вибіркової

сукупності:

$$w = \frac{m}{n} \quad (1)$$

Основне завдання вибіркового дослідження – на основі характеристик вибіркової сукупності  $\omega$  та  $\tilde{x}$  отримати достовірні судження про показники частки  $P$  і середньої  $\bar{x}$  у генеральній сукупності.

Можливі розбіжності між характеристиками вибіркової і генеральної сукупностей вимірюються середньою помилкою вибірки  $\mu$ . У математичній статистиці доводиться, що значення  $\mu$  визначаються за формулою:

$$\mu = \sqrt{\frac{\sigma_{gen}^2}{n}}, \quad (2)$$

де  $\sigma_{gen}^2$  – генеральна дисперсія.

Але при проведенні вибірових обстежень вона, як правило, невідома. На практиці для визначення  $\mu$  зазвичай використовується дисперсія вибіркової сукупності  $\sigma^2$ .

При цьому для показника частки альтернативної ознаки дисперсія визначається за формулою дисперсії альтернативної ознаки, тобто:

$$\sigma_{\omega}^2 = \omega \cdot (1 - \omega) \quad (3)$$

Слід мати на увазі, що наведена вище формула розрахунку середньої помилки вибірки  $\mu$  застосовується лише при *повторному відборі*, коли кожна одиниця, що потрапила у вибірку після фіксації значення досліджуваної ознаки повинна бути повернута в генеральну сукупність, де їй знову представляється можливість потрапити до вибірки. Але на практиці вибірові обстеження проводяться зазвичай за схемою *безповторного відбору*, при якому повторне потрапляння до

вибірки одних і тих же одиниць виключено.

Оскільки при безповторному відборі чисельність генеральної сукупності  $N$  в ході вибірки скорочується, то в формулу розрахунку  $\mu$  включають додатковий множник:  $\left(1 - \frac{n}{N}\right)$ .

Формула середньої помилки вибірки приймає наступний вигляд:

$$\mu = \sqrt{\frac{\sigma^2}{n} \cdot \left(1 - \frac{n}{N}\right)} \text{ – загальний вигляд,} \quad (4)$$

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1 - \omega)}{n} \cdot \left(1 - \frac{n}{N}\right)} \text{ – для вибіркової частки,} \quad (5)$$

$$\mu_x = \sqrt{\frac{\sigma_x^2}{n} \cdot \left(1 - \frac{n}{N}\right)} \text{ – для вибіркової середньої величини.} \quad (6)$$

Показники генеральної сукупності за показниками вибірки визначаються за формулами:

$$P = \omega \pm t \cdot \mu_{\omega} \quad (7)$$

$$\bar{x} = \tilde{x} \pm t \cdot \mu_x, \quad (8)$$

де  $t$  – коефіцієнт довіри, який залежить від імовірності ( $P$ ), з якою гарантується значення граничної помилки вибірки.

Гранична помилка вибірки  $\Delta$  дорівнює:

$$\Delta_{\omega} = t \cdot \mu_{\omega} \quad (9)$$

$$\Delta_x = t \cdot \mu_x \quad (10)$$

Значення коефіцієнта довіри  $t$  для заданої імовірності  $P$  визначається за таблицею значень функції  $\varphi(t)$ , яка виражається інтегралом імовірності Лапласа, та відображає залежність між  $t$  та імовірністю  $P$ .

На практиці користуються готовими таблицями цієї функції



(Додаток А), фрагмент якої наведений в табл. 1:

Таблиця 1 – Значення коефіцієнта довіри  $t$  для різних ступенів імовірності ( $P$ ) при умові нормально розподіленої сукупності

Коефіцієнт довіри ( $t$ )	Імовірність ( $P$ )	Коефіцієнт довіри ( $t$ )	Імовірність ( $P$ )
0,0	0,0000	2,2	0,9722
1,0	0,6827	2,5	0,9876
1,2	0,7699	2,6	0,9907
1,5	0,8664	2,8	0,9949
1,8	0,9281	2,9	0,9963
2,0	0,9545	3,0	0,9973

### Оптимальна чисельність вибірки

При організації вибіркового спостереження насамперед слід мати на увазі, що розмір помилки вибірки насамперед залежить від чисельності вибірки  $n$ . Зменшення середньої помилки вибірки завжди пов'язане зі збільшенням обсягу вибірки, але не в прямій пропорції. З формули розрахунку середньої помилки вибірки  $\mu$  слід, що  $\mu$  обернено пропорційно, тобто при збільшенні вибірки в 4 рази її помилки зменшуються лише вдвічі.

Розглянемо формулу граничної помилки вибірки для випадку повторного відбору одиниць сукупності:

$$\Delta_x = t \cdot \mu_x = t \cdot \sqrt{\frac{\sigma^2}{n}} = \frac{t \cdot \sigma}{\sqrt{n}}, \quad (11)$$

звідки чисельність необхідної виборки дорівнює:

$$n = \frac{t^2 \cdot \sigma_B^2}{\Delta_x^2}. \quad (12)$$

Необхідна чисельність вибірки для частки у випадку повторного

способу відбору одиниць:

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega)}{\Delta_w^2} \quad (13)$$

Необхідна чисельність вибірки для безповторного відбору визначається аналогічно:

$$n = \frac{t^2 \cdot \sigma_B^2 \cdot N}{\Delta_x^2 \cdot N + t^2 \cdot \sigma_B^2} \quad (14)$$

Необхідна чисельність вибірки для частки:

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega) \cdot N}{\Delta_\omega^2 \cdot N + t^2 \cdot \omega \cdot (1 - \omega)} \quad (15)$$

Використовувана в формулах величина  $\Delta_x$  – це абсолютна величина граничної помилки вибірки. На практиці нерідко задається величина не абсолютною граничної помилки, а величина відносної помилки виражена у відсотках до середньої:

$$\Delta_{\text{відносна}} = \frac{\Delta_x}{x} \cdot 100\% , \quad (16)$$

звідки

$$\Delta_x = \frac{\Delta_{\text{відносна}} \cdot \bar{x}}{100\%} . \quad (17)$$

Для оцінки невідомої величини  $\sigma^2$  (дисперсії в генеральній сукупності) застосовуються такі методи:

- пробне обстеження невеликого обсягу;
- використання даних минулих вибірових обстежень, що проводилися в аналогічних цілях;
- якщо розподіл ознаки в генеральній сукупності можна віднести

до нормального закону розподілу, то  $\sigma \approx R/6$ , де  $R$  – розмах варіації.

### Приклади розв'язання задач

#### Приклад 1.

За даними 5 %-го вибіркового обстеження, верстатів за терміном служби розподіляються згідно табл. 2:

Таблиця 2 – Розподіл верстатів за терміном служби

Термін служби, років	До 4	4÷8	8÷12	12 і більше	Разом
Кількість верстатів	25	40	20	15	100

Визначити середній термін служби верстатів та довірчий інтервал для середньої з імовірністю ( $P$ ) 0,9545.

#### Розв'язання

Дано:  $\frac{n}{N} = 0,05 = 5\%$ ,  $n = 100$ .

$$N = \frac{100}{0,05} = 2000$$

Знайдемо середню та граничну помилки вибірки. Для розрахунку середньої помилки безповторної вибірки  $\mu$  необхідно знайти дисперсію.

Дисперсію розрахуємо за формулою:  $\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{\sum f_i}$ .

Для розрахунку дисперсії ми повинні розрахувати середню арифметичну. Знайдемо середню арифметичну, як середню арифметичну зважену:

$$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f_i} = \frac{2 \cdot 25 + 6 \cdot 40 + 10 \cdot 20 + 14 \cdot 15}{100} = 7$$

Знайдемо дисперсію:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f_i}{\sum f_i} = \frac{(2-7)^2 \cdot 25 + (6-7)^2 \cdot 40 + (10-7)^2 \cdot 20 + (14-7)^2 \cdot 15}{100} = 15,8$$

Тоді середня помилка вибірки  $\mu$  дорівнює:

$$\mu_x = \sqrt{\frac{\sigma^2}{n} \cdot \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{15,8}{100} \cdot \left(1 - \frac{100}{2000}\right)} = 0,387 \approx 0,39$$

Довірчій імовірності  $F(x) = 0,9545$  за таблицями значень інтеграла імовірностей (Додаток А) відповідає квантиль  $t = 2,0$ .

Тоді гранична помилка вибірки дорівнює:

$$\begin{aligned} \Delta \bar{x} &= t \cdot \mu = 2 \cdot 0,39 = 0,78 \\ x_1 &= 7 - 0,78 = 6,22; \quad x_2 = 7 + 0,78 = 7,78. \\ 6,22 &\leq \bar{x}_0 \leq 7,78 \end{aligned}$$

Таким чином, з імовірністю 0,9545 можна стверджувати, що середній термін служби верстатів в генеральній сукупності знаходиться в межах 6,22÷7,78 років.

### Приклад 2.

З метою визначення середньої фактичної тривалості робочого дня в державній установі з чисельністю службовців 470 осіб в вересні 2019 року була проведена 20 %-ва випадкова вибірка. За результатами спостереження виявилося, що у 10 % обстежених втрати часу досягали більше 45 хв. за день. З імовірністю 0,683 встановить межі в яких знаходиться генеральна доля службовців з втратами робочого часу більше 45 хв. в день.

### Розв'язання

Визначимо обсяг вибіркової сукупності:

$$n = 470 \cdot 0,20 = 94 \text{ чоловіки.}$$

Вибіркова частка  $\omega$  за умовою дорівнює 10 %. При імовірності  $P = 0,683$  відповідно  $t = 1$ , обчислимо граничну помилку вибіркової частки для випадкової вибірки виконаної безповторним способом:

$$\Delta_w = t \cdot \sqrt{\frac{\omega \cdot (1-\omega)}{n} \cdot \left(1 - \frac{n}{N}\right)} = 1 \cdot \sqrt{\frac{0,1 \cdot (1-0,1)}{94} \cdot \left(1 - \frac{94}{470}\right)} \approx 0,02768 \text{ або } 2,8 \%$$

Межі частки ознаки у генеральній сукупності:

$$10 - 2,8 \leq W \leq 10 + 2,8$$

$$7,2 \leq W \leq 12,8$$

Таким чином, з імовірністю 0,683 можна стверджувати, що доля працівників установи з втратами робочого часу більше 45 хвилин в день знаходиться в межах від 7,2 до 12,8 %.

### Приклад 3.

Проектується опитування працівників машинобудівних підприємств з приводу задоволення умовами праці на робочих місцях. Визначте обсяг вибіркової сукупності, щоб з імовірністю  $P = 0,9545$  відносна похибка вибірки не перевищила 10 %. За результатами попереднього анонімного опитування 80 % опитуваних працівників вважали, що умови їх діяльності на робочих місцях є несприятливими.

### Розв'язання

Довірчій імовірності  $P = 0,9545$  за таблицями значень інтеграла імовірностей (Додаток А) відповідає квантиль  $t = 2,0$ .

Скористаємось формулою для безповторної вибірки:

$$\Delta = t \cdot \sqrt{\frac{\sigma^2}{n} \cdot \left(1 - \frac{n}{N}\right)} \text{ при } n \text{ значно менше } N$$

$$\Delta = t \cdot \sqrt{\frac{\sigma^2}{n}}, \quad n = \frac{t^2 \cdot \sigma_B^2}{\Delta_x^2}$$

Знайдемо дисперсію вибірки для альтернативної ознаки:

$$q = 0,8; p = 1 - q = 1 - 0,8 = 0,2;$$

$\sigma^2 = p \cdot q = 0,2 \cdot 0,8 = 0,16$  – за результатами попереднього опитування.

$$\text{Тоді, } n = \frac{t^2 \cdot \sigma_B^2}{\Delta_x^2} = \frac{2^2 \cdot 0,16}{0,1^2} = 64 \text{ працівника.}$$

### Завдання для самостійної роботи

#### Завдання 1.

За даними 5 %-го вибіркового обстеження, верстати за терміном служби розподіляються згідно табл. 3. Визначити середній термін служби верстатів та довірчий інтервал для середньої з імовірністю  $P = \underline{\hspace{1cm}}$ . Вихідні дані для розрахунку наведено в табл. 3:

Таблиця 3 – Вихідні дані для розрахунку

Варіант	Термін служби, років				Імовірність ( $P$ )
	Кількість верстатів				
	До 4	4÷8	8÷12	12 і більше	
1	35	40	25	25	0,9545
2	30	40	20	35	0,9876
3	35	30	15	20	0,9545
4	25	30	10	15	0,8664
5	35	25	10	20	0,6827
6	30	35	30	30	0,9545
7	20	50	20	35	0,9876
8	25	20	25	30	0,9545
9	25	20	20	15	0,8664
10	20	15	10	35	0,6827

#### Завдання 2.

З метою визначення середньої фактичної тривалості робочого дня

в державній установі з чисельністю службовців  $N = \_$  осіб в червні 2020 року була проведена  $V = \_ \%$ -ва випадкова вибірка. За результатами спостереження виявилось, що у  $\omega = \_ \%$  обстежених втрати часу досягали більше 33 хв. в день. З ймовірністю  $P = \_$ , встановіть межі в знаходиться яких генеральна доля службовців з втратами робочого часу більше 33 хв. в день. Вихідні дані для розрахунку наведено в табл. 4:

Таблиця 4 – Вихідні дані для розрахунку

Варіант	Чисельність службовців ( $N$ ), чоловік	Об'єм проведеної випадкової вибірки ( $V$ ), %	Вибіркова частка ( $\omega$ ), %	Імовірність ( $P$ )
1	480	25	15	0,8664
2	500	25	12	0,683
3	480	20	15	0,683
4	480	30	15	0,9545
5	460	20	10	0,8664
6	400	25	11	0,6827
7	420	30	12	0,7699
8	380	20	13	0,8664
9	440	25	14	0,9281
10	510	30	12	0,9545

### Завдання 3.

Проектується опитування працівників підприємств з приводу задоволення умовами праці на робочих місцях. Визначте обсяг вибіркової сукупності, щоб з ймовірністю  $P$  \_\_\_\_\_ відносна похибка вибірки не перевищила \_\_\_\_\_ %. За результатами попереднього анонімного опитування \_\_\_\_\_ % опитуваних працівників вважали, що умови їх діяльності на робочих місцях є несприятливими. Вихідні дані для розрахунку дивись в табл. 5.

Таблиця 6 – Вихідні дані для розрахунку

Варіант	Відносна похибка вибірки, %	Робочі місця з несприятливими умовами, %	Імовірність ( $P$ )
1	5,0	40	0,8664
2	6,5	40	0,9281
3	7,0	30	0,9545
4	7,5	30	0,9722
5	8,0	25	0,9876
6	8,5	35	0,9907
7	9,0	50	0,9949
8	9,5	20	0,9963
9	9,0	20	0,9973
10	8,5	15	0,6827

### Контрольні питання

1. Що розуміють під поняттям “вибірковий метод”? В чому його сутність та мета?
2. Назвіть переваги та недоліки вибіркового методу.
3. Чому при вибіркового спостереженні принципово присутні помилки? Назвіть їх різновиди.
4. В чому різниця між повторним та безповторним методами відбору?
5. Зазначте від чого залежить розбіжність між характеристиками вибірки та генеральної сукупності?
6. Наведіть формули для обчислення граничної помилки вибірки для середньої величини і для частки.
7. Як обчислюється середня помилка вибірки для випадкового відбору повторним і безповторним методом?
8. Від чого залежить значення коефіцієнта довіри  $t$ ?
9. Чим гарантується незсувність оцінки дисперсії вибіркової сукупності?
10. Як обчислюється оптимальна чисельність вибірки для випадкового відбору повторним методом?
11. Як обчислюється оптимальна чисельність вибірки для



випадкового відбору безповторним методом?

12. Як розраховується величина відносної помилки при вибірковому методі?

13. В яких випадках є можливість користуватися формулою стандартної помилки при повторному відборі для без повторного відбору?

### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Громько Г. Л. Общая теория статистики: практикум \ Г. Л. Громько. Москва : Инфра-М, 2008. 240 с.
5. Минашкин В. Г. Теория статистики : учебно-методический комплекс / В. Г. Минашкин [и др.]. Москва : Издательство Центр ЕАОИ, 2008. 296 с.

Значення інтеграла імовірностей  $F(t) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt$

t	Соті долі									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0797	0876	0955	1034	1114	1193	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2961	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5345	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	6827	6875	6923	6970	7017	7063	7109	7154	7199	7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7995	8030
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9089
1,7	9109	9127	9146	9164	9182	9199	9216	9233	9249	9265
1,8	9281	9281	9297	9312	9327	9342	9357	9385	9399	9412
1,9	9425	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	9545	9556	9566	9576	9586	9596	9606	9615	9625	9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9755	9762	9768	9774	9780
2,3	9785	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9840	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9909	9912	9915	9917	9920	9924	9926	9927	9929
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9950	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	9973	9974	9975	99755	9976	9977	9978	99786	9979	9980

*Примітка.* В таблиці наведені мантиси значень функції (0,...).

**Практичне заняття 7**  
**ВИЗНАЧЕННЯ СТАНДАРТНОЇ ТА ГРАНИЧНОЇ ПОМИЛКИ**  
**МЕХАНІЧНОЇ, ТИПОВОЇ, СЕРІЙНОЇ ТА МАЛОЇ ВИБІРКИ.**  
**ВИЗНАЧЕННЯ НЕОБХІДНОЇ КІЛЬКОСТІ СПОСТЕРЕЖЕНЬ**

**Мета роботи:** Отримання навичок визначення стандартної та граничної помилки механічної, типової, серійної та малих вибірок. Визначення необхідної кількості спостережень.

**Загальні положення**

*Механічна (систематична) вибірка*

Механічний (систематичний) відбір передбачає, що основою вибірки є упорядкована чисельність елементів досліджуваної сукупності, при якому вибір елементів здійснюється через рівні інтервали. Розмір інтервалу обчислюють діленням обсягу сукупності  $N$  на передбачений обсяг вибірки  $n$ .

При цьому початковий елемент визначається як випадкове число у першому інтервалі елементів сукупності, наприклад за таблицею випадкових чисел. Другий елемент залежить від першого і кроку інтервалу, який визначається як відношення  $N/n$ .

Наприклад, якщо початкове число – випадкове число 8, а крок інтервалу 20, то другий елемент становить  $8 + 20 = 28$ , третій  $28 + 20 = 48$ . Таким чином, вибіркова сукупність механічно розбивається на рівні групи, і з кожної такої групи у вибірку попадає лише один елемент.

*Помилки вибірки* при механічному відборі одиниць обчислюють за формулами простої випадкової вибірки.

Середня помилка вибірки при механічному відборі розраховується за формулою випадкового неповторного відбору:

$$\mu_{\omega} = \sqrt{\frac{\omega \cdot (1 - \omega)}{n} \cdot \left(1 - \frac{n}{N}\right)} \quad \text{– для вибіркової частки,} \quad (1)$$

$$\mu_x = \sqrt{\frac{\sigma_x^2}{n} \cdot \left(1 - \frac{n}{N}\right)} - \text{для вибіркової середньої величини.} \quad (2)$$

Гранична помилка вибірки  $\Delta$  розраховується так само як і для випадкової вибірки за формулою:

$$\Delta_\omega = t \cdot \mu_\omega \quad (3)$$

$$\Delta_x = t \cdot \mu_x \quad (4)$$

Значення  $t$  при для заданої імовірності  $P$  визначається за таблицею значень функції  $\varphi(t)$ , яка виражається інтегральною формулою Лапласа, та відображає залежність між  $t$  та імовірністю  $P$ .

Чисельність необхідної виборки для випадку повторного відбору дорівнює:

$$n = \frac{t^2 \cdot \sigma_B^2}{\Delta_x^2}. \quad (5)$$

Необхідна чисельність вибірки для частки у випадку повторного способу відбору одиниць:

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega)}{\Delta_w^2} \quad (6)$$

Необхідна чисельність вибірки для безповторного відбору визначається:

$$n = \frac{t^2 \cdot \sigma_B^2 \cdot N}{\Delta_x^2 \cdot N + t^2 \cdot \sigma_B^2} \quad (7)$$

Необхідна чисельність вибірки для частки:

$$n = \frac{t^2 \cdot \omega \cdot (1 - \omega) \cdot N}{\Delta_{\omega}^2 \cdot N + t^2 \cdot \omega \cdot (1 - \omega)} = \frac{0,25 \cdot t^2 \cdot N}{\Delta^2 \cdot N + 0,25 \cdot t^2} \quad (8)$$

### Типова (районована) вибірка

При типовій (районованій) вибірці генеральна сукупність розбивається на однородні типові групи за будь-яким признаком або на райони. З кожної типової групи або району в випадковому порядку вибираються одиниці вибіркової сукупності.

Відбір одиниць з типів може виконуватись трьома методами:

- пропорційно чисельності одиниць типових груп,
- непропорційно чисельності одиниць типових груп,
- пропорційно мінливості в групах.

Типова вибірка з пропорційним відбором одиниць з типових груп зустрічається найчастіше.

Об'єм вибірки при цьому визначається за формулою:

$$n_i = n \cdot \frac{N_i}{N}, \quad (9)$$

де  $n_i$  – об'єм вибірки з типової групи;

$n$  – загальний об'єм вибірки;

$N_i$  – об'єм типових груп;

$N$  – об'єм генеральної сукупності.

Наведемо для цього способу відбору формули розрахунку середньої квадратичної помилки вибірки при неповторному відборі усередині районів:

$$\text{а) для середньої} \quad \mu_x = \sqrt{\frac{\bar{\sigma}_B^2}{n} \cdot \left(1 - \frac{n}{N}\right)}, \quad (10)$$

де  $\bar{\sigma}_B^2$  – середня з дисперсій районів вибірки  $\bar{\sigma}_B^2 = \frac{\sum \sigma_i^2 \cdot n_i}{n}$ ;

$$\text{б) для частки} \quad \mu_x = \sqrt{\frac{w \cdot (1-w)}{n} \cdot \left(1 - \frac{n}{N}\right)}, \quad (11)$$

де  $\overline{w \cdot (1-w)}$  – середня з внутрішньо групових дисперсій.

Гранична помилка вибірки  $\Delta$  розраховується так само як і для випадкової вибірки за формулою (3) або (4), з тією відмінністю, що середня помилка вибірки розраховується за відповідними формулами (10) та (11).

Визначення необхідної чисельності вибірки при безповторному відборі усередині районів здійснюється за формулами:

$$\text{а) для середньої} \quad n = \frac{t^2 \cdot \overline{\sigma_B^2} \cdot N}{\Delta^2 \cdot N + t^2 \cdot \overline{\sigma_B^2}}; \quad (12)$$

$$\text{б) для частки} \quad n = \frac{t^2 \cdot \overline{w \cdot (1-w)} \cdot N}{\Delta^2 \cdot N + t^2 \cdot \overline{w \cdot (1-w)}}. \quad (13)$$

Об'єм вибірки з типових груп, районів при відборі, пропорційному чисельності одиниць типових груп, районів, визначається за формулою:

$$n_i = n \cdot \frac{N_i}{N} \quad (14)$$

#### *Серійна (гніздова) вибірка*

При серійній вибірці генеральну сукупність ділять на однакові за об'ємом групи – серії. До вибіркової сукупності отбирають серії. Всередині серій виконується суцільне спостереження одиниць, які потрапили в серію.

При без повторному відборі серій середня помилка вибіркової середньої визначається за формулою:

$$\mu = \sqrt{\frac{\delta^2}{r} \cdot \left(1 - \frac{r}{R}\right)}; \quad (15)$$

де  $\delta^2$  – міжсерійна дисперсія середніх;

$r$  – кількість відібраних серій;

$R$  – загальна кількість серій в генеральній сукупності.

Міжсерійна дисперсія розраховується за формулами:

а) для середньої 
$$\delta^2 = \frac{\sum (\tilde{x}_i - \tilde{x}_0)^2}{r}; \quad (16)$$

б) для частки 
$$\delta^2 = \frac{\sum (\omega_i - \bar{\omega})^2}{r}, \quad (17)$$

де  $\tilde{x}_i$  – середні в серіях;

$\tilde{x}_0$  – загальна середня для серій;

$\omega_i$  – частки в серіях (групах);

$\bar{\omega}$  – середня частка ознаки для всієї вибіркової сукупності.

Гранична помилка вибірки  $\Delta$  розраховується так само як і для випадкової вибірки за формулою (3) або (4), з тією відмінністю, що середня помилка вибірки розраховується за відповідними формулами (15–17).

Чисельність вибірки:

$$r = \frac{t^2 \cdot \delta^2 \cdot R}{\Delta^2 \cdot R + t^2 \cdot \delta^2}. \quad (18)$$

### *Малі вибірки*

Оцінка параметра вибірки вважається сталою, якщо при зростанні обсягу вибірки з певною мірою ( $P \rightarrow 1$ ) значення оцінки наближається до свого теоретичного значення. Сталість оцінки гарантує дослідникові

зростання точності з ростом обсягу вибірки:

$$n \rightarrow N, \bar{x}_0 \rightarrow \bar{x}$$

Незсувність оцінки параметра вибірки позначає відсутність систематичної (тенденційної) помилки в оцінці параметра. Незсувність оцінки дисперсії вибіркової сукупності гарантується введенням у знаменник коригуючого множника Шепарда ( $n - 1$ ):

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (19)$$

Множником Шепарда можливо знехтувати при  $n > 20$ . Але на практиці часто приходиться мати справу з малими вибірками де кількість елементів сукупності  $n < 30$ . В таких випадках для розрахунку помилки вибірки неможливо користуватися теоремами закону великих чисел, оскільки на вибіркочну середню великий вплив має значення кожної із випадково відібраних одиниць і їх розподіл може значно відрізнятись від нормального закону розподілу.

Для малих вибірок діє закон розподілу Стюдента, у якому квантіль розподілу « $t$ » (довірче число) залежить не тільки від імовірності, з якою обчислюється гранична похибка, а також від кількості елементів вибірки. Для  $t = 1$  залежність відсотка імовірностей від кількості елементів сукупності надано у табл. 1.

Таблиця 1 – Залежність відсотка імовірностей від кількості елементів сукупності

Кількість елементів сукупності, $n$	2	4	5	10	20	$\geq 30$
Відсоток імовірності	0,5	0,608	0,626	0,656	0,67	0,683

Тому при вибірках невеликого обсягу методи оцінювання результатів вибіркового спостереження мають свою особливість порівняно з методами великих вибірок.



Середня квадратична помилка для кількостей ознак малої вибірки  $\mu_{M.B.}$  визначається за формулою:

$$\mu_{M.B.} = \sqrt{\frac{\sigma_{M.B.}^2}{n-1}}, \quad (20)$$

де  $\sigma_{M.B.}^2$  – дисперсія малої вибірки

При знаходженні імовірності допуску тієї чи іншої похибки або визначенні довірчих меж інтервалів досліджуваного показника в генеральній сукупності користуються таблицями імовірності не Гауса, а Ст'юдента, де імовірність  $P = S(t, n)$  визначається в залежності від об'єму вибірки і коефіцієнта довіри  $t$ .

При цьому потрібно мати на увазі, що знайдене значенні за довідковими таблицями  $S(t, n)$  характеризує імовірність того, що фактичний (розрахований) коефіцієнта довіри ( $t_{факт}$ ) не більше заданого значення  $t$  (табличне), тобто  $t_{факт} < t$ , і графічно ця імовірність дорівнює площі, що обмежена кривою розподілу Ст'юдента та віссю абсцис в інтервалі від  $-\infty$  до  $t$ .

Виходить, що  $[1 - S(t, n)]$  характеризує імовірність того, що  $t_{факт} > t$ , тобто  $t_{факт}$  вийде за межі  $t$  в області позитивних значень. Якщо розглядати  $|t_{факт}|$ , то імовірність виходу за задані межу  $t$  в обидві сторони буде дорівнювати:

$$P(|t_{факт}| > t) = 2 \cdot [1 - S(t, n)] \quad (21)$$

При малій вибірці імовірність попадання середнього значення досліджуваної ознаки в генеральній сукупності в межах  $\tilde{x} \pm \Delta_{M.B.}$  визначається як  $2 \cdot S(t, n) - 1$ , тобто:

$$P(|\tilde{x} - \bar{x}| \leq \Delta_{M.B.}) = 2 \cdot S(t, n) - 1 \quad (22)$$

Розподіл Ст'юдента слід використовувати тільки для оцінки помилок вибірки, взятої із генеральної сукупності з нормальним законом розподілення ознаки.

### Приклади розв'язання задач

#### Приклад 1.

Проведено вибіркове дослідження партії заготовок деталей. При механічному неповторному відборі 2,5 % виробів отримані дані про розподіл зразків за вагою (табл. 2).

Таблиця 2 – Розподіл зразків за вагою

Вага виробу, г.	Число виробів
до 1000	22
1000-1025	77
1025-1050	183
1050-1075	85
1075-1100	23
понад 1100	10
Всього	400

При вимогах, що до нестандартної продукції відносяться заготовки вагою до 1000 г. та понад 1100 г. визначити межі значення питомої ваги стандартної продукції та середньої ваги виробу для всієї партії з імовірністю 0,954.

#### Розв'язання

За умовами  $n = 400$ . Знайдемо  $N$ :

$$N = 400 * 100\% / 2,5\% = 16000 \text{ шт.}$$

Встановимо узагальнюючі показники вибіркової сукупності.

Розрахунок вибіркової долі  $w$ .

Число стандартних одиниць в вибірці  $m$ :

$$m = 400 - (22 + 10) = 368$$

$$w = \frac{m}{n} = \frac{368}{400} = 0,92, \text{ тобто питома вага стандартних виробів в вибірці}$$

становить 92 %.

Розрахунок вибіркової середньої  $\tilde{x}$ .

Розрахуємо  $\tilde{x}$  за формулою середньої зваженої:

$$\tilde{x} = \frac{\sum x \cdot f}{\sum f}.$$

Для цього визначимо середини інтервалів. Середини крайніх (відкритих) інтервалів визначимо, виходячі з гіпотези однаково заповнених інтервалів, тобто приймаємо межі першого інтервала від 975 до 1000 г., останнього – від 1100 до 1125 г. Наступні проміжні розрахунки зведемо в таблицю (табл. 3)

Таблиця 3 – Розрахункові показники

Вага виробу, г.	Число виробів	Середина інтервала	$x \cdot f$	$x_i - \tilde{x}$	$(x_i - \tilde{x})^2$	$(x_i - \tilde{x})^2 \cdot f$
до 1000	22	987,5	21725	-52,5	2756,25	60637,5
1000-1025	77	1012,5	77962,5	-27,5	756,25	58231,25
1025-1050	183	1037,5	189862,5	-2,5	6,25	1143,75
1050-1075	85	1062,5	90312,5	22,5	506,25	43031,25
1075-1100	23	1087,5	25012,5	47,5	2256,25	51893,75
понад 1100	10	1112,5	11125	72,5	5256,25	52562,5
Всього	400		416000			267500

Середня вага виробу в вибірці складає:

$$\tilde{x} = \frac{\sum x \cdot f}{\sum f} = \frac{416000}{400} = 1040 \text{ г.}$$

Встановимо середні помилки вибірки для узагальнюючих

характеристик вибіркової сукупності, користуючись формулами для безповторного відбору.

Для вибіркової долі:

$$\mu_w = \sqrt{\frac{w \cdot (1-w)}{n} \cdot \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,92 \cdot (1-0,92)}{400} \cdot \left(1 - \frac{400}{16000}\right)} = 0,0134,$$

тобто, середня помилка вибірки для долі стандартної продукції складає 1,34 %.

Для вибіркової середньої:

Спочатку потрібно розрахувати дисперсію ( $\sigma^2$ ):

$$\sigma^2 = \frac{\sum (x_i - \tilde{x})^2 \cdot f}{\sum f} = \frac{267500}{400} = 668,75$$

$$\mu_x = \sqrt{\frac{\sigma_x^2}{n} \cdot \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{668,75}{400} \cdot \left(1 - \frac{400}{16000}\right)} = \sqrt{1,63} \approx 1,28 \text{ г.}$$

тобто, середня помилка вибірки для середньої величини складає 1,28 г.

Встановимо граничні значення для характеристик генеральної сукупності, враховуючи, що імовірності 0,9545 відповідає значення коефіцієнта довіри  $t = 2$  (Додаток А):

Для генеральної долі

$$P = w \pm t \cdot \mu_w = 92 \pm 2 \cdot 1,34 \% \text{ або } 89,32 \% \leq P \leq 94,68 \%$$

Для генеральної середньої

$$\bar{x} = \tilde{x} \pm t \cdot \mu_x = 1040 \pm 2 \cdot 1,28 \text{ г. або } 1037,44 \text{ г.} \leq \bar{x} \leq 1042,56 \text{ г.}$$

Таким чином, з імовірністю 95,45 % доля стандартних виробів в партії знаходиться в межах від 89,32 % до 94,68 %, а середня вага виробу – в межах від 1037,44 до 1042,56 г.

**Приклад 2.**

Для виявлення частки простоїв через несвоєчасне надходження напівфабрикатів була проведена фотографія робочого дня 10 % працівників чотирьох різних цехів. Вибір працівників всередині цехів виконувався методом механічного відбору. В результаті виборки були отримані наступні дані (табл. 4):

Таблиця 4 – Результати вибірки

Цех	Число робітників в вибірці	Питома вага простоїв через несвоєчасне надходження напівфабрикатів, %
№ 1	20	5
№ 2	36	10
№ 3	14	15
№ 4	30	2

З імовірністю 0,9545 потрібно визначити межі, в яких знаходиться частка простоїв на заводі через несвоєчасне надходження напівфабрикатів.

**Розв'язання**

Розрахуємо частку простоїв через несвоєчасне надходження напівфабрикатів у вибірці:

$$\bar{w} = \frac{\sum w_i \cdot n_i}{\sum n}$$

$$\bar{w} = \frac{\sum w_i \cdot n_i}{\sum n} = \frac{5 \cdot 20 + 36 \cdot 10 + 14 \cdot 15 + 30 \cdot 2}{(20 + 36 + 14 + 30)} = 7,3 \%$$

Розрахуємо дисперсії типових груп:

$$\text{Для I групи: } \sigma_I^2 = \omega_1 \cdot (1 - \omega_1) = 5 \cdot 95 = 475;$$

$$\text{Для II групи: } \sigma_{II}^2 = \omega_2 \cdot (1 - \omega_2) = 10 \cdot 90 = 900;$$

$$\text{Для III групи: } \sigma_{III}^2 = \omega_3 \cdot (1 - \omega_3) = 15 \cdot 85 = 1275;$$

$$\text{Для IV групи: } \sigma_{IV}^2 = \omega_4 \cdot (1 - \omega_4) = 2 \cdot 98 = 196.$$

Середня з групових дисперсій визначається за формулою:

$$\overline{w \cdot (1-w)} = \frac{\sum \omega_i \cdot (1-\omega_i) \cdot n_i}{\sum n_i};$$

$$\overline{w \cdot (1-w)} = \frac{475 \cdot 20 + 900 \cdot 36 + 1275 \cdot 14 + 196 \cdot 30}{100} = 656,3$$

Визначимо середню помилку вибіркової частки:

$$\mu_w = \sqrt{\frac{656,3}{100} \cdot \left(1 + \frac{100}{1000}\right)} = 2,43 \%$$

Гранична помилка вибірки для долі з імовірністю 0,9545:

$$\Delta = t \cdot \mu_w = 2 \cdot 2,43 = 4,86 \%$$

З імовірністю 0,9545 можна стверджувати, що частка простоїв робочих через несвоєчасне надходження напівфабрикатів знаходиться в межах  $7,3 - 4,86\% \leq P \leq 7,3 + 4,86\%$  або  $2,44\% \leq P \leq 12,16\%$ .

### Приклад 3.

У механічному цеху машинобудівного заводу А є 10 бригад по 20 робітників в кожній бригаді. Для встановлення кваліфікації робітників цеху проектується серійна вибірка методом механічного відбору. Яка кількість бригад необхідно відібрати щоб з імовірністю 0,9545 помилка вибірки не перевищувала 1,0, якщо на основі попередніх обстежень відомо, що середня з внутрішньо групових дисперсій дорівнює 0,9?

### Розв'язання

Розрахуємо необхідну кількість бригад на основі формули чисельності серійної неповторної вибірки:

$$r = \frac{t^2 \cdot \delta^2 \cdot R}{\Delta^2 \cdot R + t^2 \cdot \delta^2}.$$

Імовірності 0,9545 відповідає значення коефіцієнта довіри  $t = 2,0$  (Додаток А). Тоді кількість бригад, яку необхідно відібрати дорівнює:

$$r = \frac{t^2 \cdot \delta^2 \cdot R}{\Delta^2 \cdot R + t^2 \cdot \delta^2} = \frac{2^2 \cdot 0,9 \cdot 10}{1^2 \cdot 10 + 2^2 \cdot 0,9} = 2,647 \approx 3 \text{ бригади.}$$

#### Приклад 4.

З метою контролю за дотриманням норм витрат металу проведено вибіркоче спостереження партії заготовок деталей, які надходять з ливарного цеху. Вага відібраних 10 зразків становила, г: 520, 565, 550, 540, 525, 485, 515, 595, 500, 555.

Визначте: а) середню вагу перевірених заготовок деталей, граничну помилку та довірчі межі для середньої з імовірністю 0,95;

б) чи відповідають фактичні витрати металу на виготовлення деталей встановленій нормі – 525 г.?

Висновки зробіть з тією ж імовірністю.

#### Розв'язання

Середня вага – це вибіркова середня арифметична проста:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1}{10} \cdot (520 + 565 + 550 + 540 + 525 + 485 + 515 + 595 + 500 + 555) = 535 \text{ г.}$$

Довірчій імовірності  $P = 0,95$  ( $\alpha = 1 - P = 1 - 0,95 = 0,05$ ) і

$k = n - 1 = 10 - 1 = 9$  за таблицями критичних точок розподілу t-Стьюдента відповідає квантиль  $t = 1,8331 \approx 1,83$  (Додаток Б).

Розрахуємо дисперсію, скориставшись знайденої середньою:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = 960$$

Гранична помилка вибірки з вказаною імовірністю:

$$\Delta \bar{x} = t \cdot \sqrt{\frac{\sigma^2}{n-1}} = 1,83 \cdot \sqrt{\frac{960}{10-1}} = 18,9 \text{ г.}$$

Таким чином, довірчий інтервал складає:

$$\bar{x} - \Delta \bar{x} \leq \bar{x}_0 \leq \bar{x} + \Delta \bar{x}$$

де  $\bar{x}$  – вибіркова серія,  $\bar{x}_0$  – середня генеральної сукупності.

$$535 - 18,9 \leq \bar{x}_0 \leq 535 + 18,9$$

$$516,1 \leq \bar{x}_0 \leq 553,9$$

Фактичні витрати металу відповідають встановленій нормі (525 г), так як це значення належить довірчому інтервалу:  $516,1 \leq 525,0 \leq 553,9$ .

### Завдання для самостійної роботи

#### Завдання 1.

З партії готової продукції на машинобудівному заводі в порядку механічної вибірки перевірено  $n = \underline{\hspace{1cm}}$  одиниць продукції на тривалість роботи, яка склала  $T = \underline{\hspace{1cm}}$  годин при середньоквадратичному відхиленні  $\sigma = \underline{\hspace{1cm}}$  годин (табл. 5).

Таблиця 5 – Вихідні дані для розрахунку

Варіант	Кількість одиниць $n$	Тривалість роботи $T$	Середньоквадратичне відхилення $\sigma$	Імовірність $P$
1	2	3	4	5
1	50	840	60	0,9545
2	45	830	50	0,9876
3	40	820	40	0,7699



4	50	810	60	0,8664
1	2	3	4	5
5	40	800	50	0,6827
6	35	790	40	0,9281
7	55	780	60	0,9876
8	40	770	50	0,9545
9	60	760	40	0,8664
10	45	750	60	0,6827

Визначити:

а) середню помилку ( $\mu$ ) вибіркової середньої тривалості роботи продукції машинобудування;

б) з імовірністю  $P = \underline{\hspace{2cm}}$  довірчі межі тривалості роботи в генеральній сукупності.

### Завдання 2.

Середня тривалість горіння, встановлена шляхом досліджень  $n = \underline{\hspace{2cm}}$  випадково відібраних електричних лампочок дорівнює  $T = \underline{\hspace{2cm}}$  годин, при середньоквадратичному відхиленні  $\sigma = \underline{\hspace{2cm}}$  годин (табл. 6).

З якою імовірністю можна стверджувати, що допущена при цьому гранична помилка вибірки (тобто розходження між вибірковою та генеральною середньою) не перевищить  $\Delta = \underline{\hspace{2cm}}$  годин?

Таблиця 6 – Вихідні дані для розрахунку

Варіант	Кількість одиниць $n$	Тривалість роботи $T$ , годин	Середньоквадратичне відхилення $\sigma$ , годин	Гранична помилка вибірки $\Delta$ , годин
1	2	3	4	5
1	10	1280	18	12
2	9	1250	17,3	11
3	11	1200	15,8	10
4	12	1150	15	9
5	13	1120	16	11
6	14	1100	13	8

7	15	1300	12	9
1	2	3	4	5
8	14	1220	14	10
9	12	1240	12,6	11
10	10	1050	12	10

Для розрахунків необхідно скористатись довідковою таблицею  $S(t, n)$  в додатку В.

### Контрольні питання

1. Як розраховується середня помилка вибірки при механічному відборі для вибіркової частки та середньої величини?

2. Як розраховується гранична помилка вибірки при механічному відборі для вибіркової частки та середньої величини?

3. Як розраховується необхідна чисельність вибірки при механічному відборі для повторного та безповторного способу?

4. Як розраховується середня помилка вибірки при типовому відборі для вибіркової частки та середньої величини при безповторному відборі усередині районів?

5. Як розраховується необхідна чисельність вибірки при типовому відборі для повторного та безповторного способу?

6. Як визначається необхідна чисельність типової вибірки при безповторному відборі усередині районів для вибіркової частки та середньої величини?

7. Як розраховується середня помилка вибіркової середньої при безповторному відборі серій для серійної вибірки?

8. Наведіть формулу розрахунку чисельності вибірки для серійної вибірки.

9. Вкажіть чим гарантується незсувність оцінки дисперсії вибіркової сукупності.

10. Наведіть формулу для розрахунку середньої квадратичної помилки для кількостей ознак малої вибірки.

10. Від чого залежить імовірність ( $P$ ) в розподілі Стюдента?

### Література

1. Казарезов А. Я. Задачі з теорії статистики : навчальний посібник \ А. Я. Казарезов, І. В. Прядко, Г. О. Бурдельна. Миколаїв: ЧДУ ім. П. Могили, 2012. 248 с.
2. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
3. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
4. Таблиці функцій та критичних точок розподілів. Розділи: Теорія ймовірностей. Математична статистика. Математичні методи в психології. / Укладач: М.М. Горонескуль. Харків : УЦЗУ, 2009. 90 с.

Значення інтеграла імовірностей  $F(t) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt$

t	Соті долі									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0797	0876	0955	1034	1114	1193	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2961	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5345	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	6827	6875	6923	6970	7017	7063	7109	7154	7199	7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7995	8030
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9089
1,7	9109	9127	9146	9164	9182	9199	9216	9233	9249	9265
1,8	9281	9281	9297	9312	9327	9342	9357	9385	9399	9412
1,9	9425	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	9545	9556	9566	9576	9586	9596	9606	9615	9625	9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9755	9762	9768	9774	9780
2,3	9785	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9840	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9909	9912	9915	9917	9920	9924	9926	9927	9929
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9950	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	9973	9974	9975	99755	9976	9977	9978	99786	9979	9980

*Примітка.* В таблиці наведені мантиси значень функції (0,...).

**Критичні точки розподілу t-Стюдента**

Число ступенів вільності, $k$ ( $df$ )	Рівень значущості $\alpha$ (двобічна критична область)		
	<b>0,10</b>	<b>0,05</b>	<b>0,02</b>
1	6,31	12,70	31,82
2	2,92	4,30	6,97
3	2,35	3,18	4,54
4	2,13	2,78	3,75
5	2,01	2,57	3,37
6	1,94	2,45	3,14
7	1,89	2,36	3,00
8	1,86	2,31	2,90
9	1,83	2,26	2,82
10	1,81	2,23	2,76
11	1,80	2,20	2,72
12	1,78	2,18	2,68
13	1,77	2,16	2,65
14	1,76	2,14	2,62
15	1,75	2,13	2,60
16	1,75	2,12	2,58
17	1,74	2,11	2,57
18	1,73	2,10	2,55
19	1,73	2,09	2,54
20	1,73	2,09	2,53
21	1,72	2,08	2,52
22	1,72	2,07	2,51
23	1,71	2,07	2,50
24	1,71	2,06	2,49
25	1,71	2,06	2,49
26	1,71	2,06	2,48
27	1,71	2,05	2,47
28	1,70	2,05	2,46
29	1,70	2,05	2,46
30	1,70	2,05	2,46
	<b>0,05</b>	<b>0,025</b>	<b>0,01</b>
	Рівень значущості $\alpha$ (однобічна критична область)		

**$S(t, n)$  в розподілі Стюдента**

$t \backslash n$	1	2	3	4	5	6–7	8–10	11–15	16–25	25–30	$\infty$
0,0	500	500	500	500	500	500	500	500	500	500	500,00
0,1	532	535	537	537	535	538	539	539	539	539	539,827
0,2	563	570	573	574	575	576	578	578	578	578	579,259
0,3	593	606	608	610	612	613	615	616	616	616	617,911
0,4	621	636	642	645	647	649	651	652	653	654	655,421
0,5	648	667	674	678	681	683	685	687	689	689	691,462
0,6	672	695	705	710	713	715	718	721	722	724	725,746
0,7	694	723	733	739	742	746	749	752	754	756	758,036
0,8	715	746	759	766	770	774	778	781	783	785	788,144
0,9	733	768	783	790	795	800	804	808	811	813	815,939
1,0	750	789	804	813	818	823	828	832	835	838	841,344
1,1	765	807	824	834	839	844	850	854	858	860	864,333
1,2	779	824	842	852	858	864	870	874	878	881	884,930
1,3	791	838	858	868	875	881	887	892	896	899	906,199
1,4	803	852	872	883	890	896	902	907	912	915	919,243
1,5	813	864	885	896	903	909	916	921	925	928	933,192
1,6	822	875	896	908	915	921	928	933	937	940	945,200
1,7	831	884	906	918	925	932	938	943	948	951	955,434
1,8	839	893	915	927	934	941	947	952	956	959	946,069
1,9	846	901	923	935	942	948	955	960	964	967	971,283
2,0	852	908	930	942	949	955	962	967	970	973	977,249
2,1	858	915	937	948	955	961	967	972	976	978	982,135
2,2	864	921	942	954	960	966	972	977	980	982	986,096
2,3	870	926	948	958	965	971	977	981	984	986	989,275
2,4	874	931	952	963	969	975	980	984	987	989	991,802
2,5	879	935	956	966	973	978	983	987	989	991	993,790
2,6	883	939	960	970	976	981	986	989	991	993	995,338
2,7	887	943	963	973	979	983	988	991	993	995	996,533
2,8	891	946	966	976	981	985	990	993	995	996	997,444
2,9	894	949	969	978	983	987	991	994	996	997	998,134
3,0	898	952	971	980	985	989	993	995	997	997	998,650

*Примітка.* В таблиці наведені мантиси значень функції  $(0, \dots)$ .

## Практичне заняття 8 ВИРІВНЮВАННЯ ВАРІАЦІЙНИХ РЯДІВ

**Мета роботи:** Отримання навичок варівнювання варіаційних рядів по кривій нормального розподілу та розподілу Пуассона.

### Загальні положення

Під вирівнюванням варіаційних рядів розуміють заміну емпіричного розподілу близьким до нього за характером теоретичним (імовірносним) розподілом, який має визначений аналітичний вираз.

Дуже часто зустрічається саме нормальний розподіл випадкових величин або розподіл Га́усса.

### *Нормальний розподіл випадкових величин (розподіл Га́усса)*

Графік нормального розподілу має форму колоколоподібної кривої, симетричної відносно  $\bar{x}$ , кінці якої асимптотично наближаються до вісі абсцисс (рис. 1).

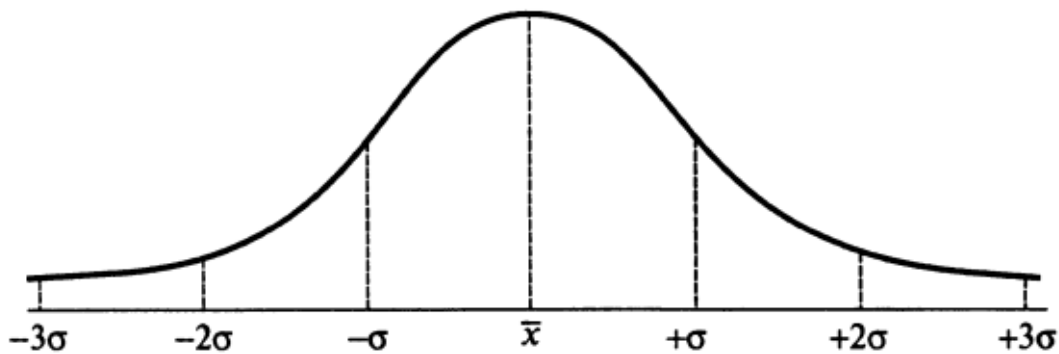


Рисунок 1 – Крива нормального розподілу

Крива має точки перегину, абсциси яких знаходяться на відстані  $\sigma$  від центра симетрії і розраховуються рівнянням:

$$y = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}, \quad (1)$$

де  $y$  – ордината кривої нормального розподілу;

$t = \frac{x - \bar{x}}{\sigma}$  – нормоване відхилення, тобто відхилення окремих

варіант від середньої  $\bar{x}$ , виражене в одиницях середнього квадратичного відхилення  $\sigma$ .

При вирівнюванні варіаційного ряду за кривою нормального розподілу теоретичні частоти ряду визначається за формулою:

$$f' = \frac{N \cdot h}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \text{ або } f' = \frac{N \cdot h}{\sigma} \cdot \varphi(t), \quad (2)$$

де  $N = \sum f$  – сума всіх частот варіаційного ряду;

$h$  – величина інтервалу в групах (класах);

$\sigma$  – середнє квадратичне відхилення;

$t = \frac{x - \bar{x}}{\sigma}$  – нормоване відхилення варіант від середньої арифметичної  $\bar{x}$ .

Величина  $\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}$  є табульованою і її легко визначити за таблицею Додатка А як функцію  $t$ , тобто  $\varphi(t)$ .

Як видно з формул (1–2) основними параметрами кривої нормального розподілу є  $\bar{x}$  та  $\sigma$ .

### ***Розподіл Пуассона***

В деяких випадках варіаційний ряд представляє собою розподіл за дискретним признаком, де при збільшенні значень признака  $x$  частоти різко зменшуються і середня арифметична ряду дорівнює або близька по значенню до дисперсії, тобто  $\bar{x} = \sigma^2$ . Такий ряд можна згладити за кривою Пуассона (рис. 2).



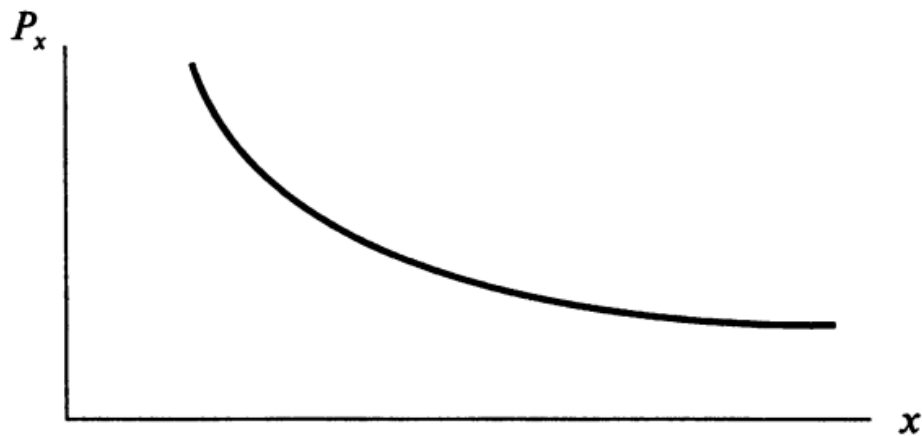


Рисунок 2 – Крива Пуассона

Крива Пуассона має аналітичний вираз:

$$P_x = \frac{a^x \cdot e^{-a}}{x!}, \quad (3)$$

де  $P_x$  – імовірність настання окремих значень  $x$ ;

$a = \bar{x}$  – середня арифметична ряду.

Теоретичні частоти при вирівнюванні емпіричних даних за кривою Пуассона визначаються за формулою:

$$f' = N \cdot P_x \quad (4)$$

де  $f'$  – теоретичні частоти;  $N$  – загальне число одиниць ряду.

Після вирівнювання рядів, тобто знаходження теоретичних частот, виникає необхідність перевірити, випадкові чи ні істотні розходження між емпіричними та теоретичними частотами. Тим самим перевірити правильність видвинутої гіпотези про наявність того чи іншого характеру розподілу в емпіричному ряді.

Для оцінки близькості емпіричних ( $f$ ) та теоретичних ( $f'$ ) частот можна застосувати критерії згоди: Пірсона ( $\chi^2$ ), Романовського ( $K_p$ ),

Колмогорова ( $\lambda$ ).

### **Критерій Пірсона ( $\chi^2$ )**

Цей критерій являє собою відношення квадратів розбіжностей між  $f$  і  $f'$  до теоретичних частот:

$$\chi^2 = \sum \frac{(f - f')^2}{f'} \quad (5)$$

Фактичне значення  $\chi^2$  порівнюють з критичним. Критичне значення визначається за допомогою спеціальних таблиць в залежності від приймаемого рівня значущості  $\alpha$  та числа ступенів свободи  $\nu$  (Добавок Б).

Рівень значущості  $\alpha$  – імовірність допущення помилки першого роду, тобто імовірність відкинути правильну гіпотезу про закон розподілу. Зазвичай приймається 5% або 1% ( $\alpha = 0,05$  або  $\alpha = 0,01$ ). Тобто в 5 % або в 1 % можи бути відкинута правильна гіпотеза.

Число ступенів свободи  $\nu$  – розраховується як число груп  $m$  мінус число зв'язків  $z$ :

$$\nu = m - z \quad (6)$$

Під кількістю зв'язків розуміють число показників емпіричного ряду, що використовуються при розрахунку теоретичних частот, тобто показників, що зв'язують емпіричні та теоретичні частоти  $(\bar{x}, \sigma, f_i)$ .

При вирівнюванні за кривою нормального розподілу  $\nu = m - 3$ . При вирівнюванні за кривою Пуассона  $\nu = m - 2$ , оскільки при побудові частот використовується два зв'язки, що обмежують  $(\bar{x}, f_i)$ .

У випадку повного збігу теоретичного і емпіричного розподілу  $\chi^2 = 0$ , в іншому випадку  $\chi^2 > 0$ .

Якщо  $\chi_{розр}^2 > \chi_{табл}^2$ , то при заданому рівні значущості  $\alpha$  і числі ступенів свободи  $\nu$  гіпотезу про неістотність (випадковість) розбіжностей відхиляємо.

Якщо  $\chi^2_{розр} \leq \chi^2_{табл}$ , робимо висновок, що емпіричний ряд гарно узгоджується з гіпотезою про пропонуємий розподіл і з імовірністю  $(1-\alpha)$  можна стверджувати, що розбіжність між емпіричними та теоретичними частотами випадкова.

Якщо вирівнювання виконується не за частотами, а за частостями  $\omega_i = \frac{f_i}{\sum f_i}$ , то  $\chi^2$  розраховується за формулою:

$$\chi^2 = N \cdot \sum \frac{(\omega_i - \omega'_i)^2}{\omega'_i}, \quad (6)$$

де  $\omega_i, \omega'_i$  – частоти відповідно емпіричного та теоретичного розподілів;

$N = \sum f_i$  – загальна сума частот (абсолютна чисельність одиниць розподілу).

### ***Критерій Романовського ( $K_p$ )***

$$K_p = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}} \quad (7)$$

Якщо  $K_p < 3$ , то розходження між емпіричними і теоретичними частотами можна вважати випадковими, якщо  $K_p > 3$ , то не є випадковими, і відповідно теоретичний розподіл не може служити моделлю для досліджуваного емпіричного розподілу.

### ***Критерій Колмогорова***

Даний критерій заснований на визначені максимальної розбіжності між накопченими частотами або частостями емпіричних та теоретичних розподілів:

$$\lambda = \frac{D}{\sqrt{N}} \quad (8)$$

де  $D$  – максимальна різниця між накопиченими частотами;  
 $N$  – сума всіх частот.

Розрахував значення  $\lambda$  за таблицею  $P(\lambda)$  (Додаток В) визначають імовірність, з якою можна стверджувати, що відхилення емпіричних частот від теоретичних випадкові. Імовірність  $P(\lambda)$  може змінюватись від 0 до 1. При  $P(\lambda)=1$  відбувається повне співпадання частот, при  $P(\lambda)=0$  – повна розбіжність. Вважається, що якщо  $\lambda$  приймає значення до 0,3, то  $P(\lambda)=1$ .

### Приклади розв'язання задач

#### Приклад 1.

Розподіл міцності 200 проб ниток наведений в табл. 1.

Таблиця 1 – Розподіл міцності ниток

Міцність, г	Число проб ( $f$ )
120–130	1
130–140	8
140–150	27
150–160	58
160–170	56
170–180	34
180–190	14
190–200	2
Всього	200

Виходячі з гіпотези про нормальний розподіл результатів дослідження згладити ряд за кривою нормального розподілу (тобто розрахувати теоретичні частоти) та оцінити близькість емпіричних та теоретичних частот за допомогою критеріїв згоди: Пірсона ( $\chi^2$ ),

Романовського ( $K_p$ ), Колмогорова ( $\lambda$ ). Побудувати графічне зображення розподілів.

### Розв'язання

Для знаходження теоретичних частот використаємо формулу (2).

Розрахуємо  $t$  та  $\sigma$  – основні параметри кривої нормального розподілу. Розрахунки виконаємо у вигляді табл. 2.

1) Для розрахунку середньої скористаємось формулою середньої арифметичної зваженої:

$$\bar{x}_{ар.} = \frac{\sum x \cdot f}{\sum f}.$$

За розрахунками:  $\bar{x} = 161,4$  г.

2)  $\sigma$  розрахуємо як  $\sqrt{\sigma^2}$ .

Дисперсію знайдемо за загальною формулою:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 \cdot f}{\sum f}.$$

За розрахунками:  $\sigma = 13$  г.

3) Находимо відхилення окремих варіант від середньої (графа 4):

$$x - \bar{x}.$$

4) Находимо нормоване відхилення від середньої (графа 5):

$$t = \frac{x - \bar{x}}{\sigma}.$$

5) Знаючи величину  $t$  за Додатком А находимо  $\varphi(t)$  (графа 6).

6) Оскільки маємо справу з рядом в якому рівні інтервали, то розрахуємо постійний множник:

$$\frac{N \cdot h}{\sigma}$$

$$\text{Маємо: } \frac{N \cdot h}{\sigma} = \frac{200 \cdot 10}{13} = 153,85 \approx 154.$$

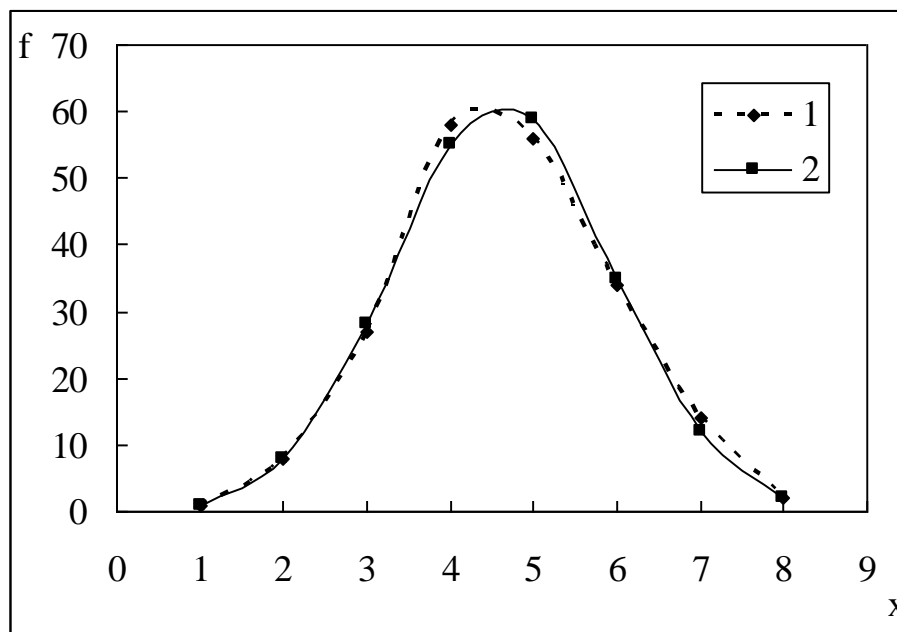
7) Помножуючи послідовно 154 на  $\varphi(t)$  і округляючи результати до цілих чисел, знаходимо теоретичні частоти за групами (графа 7).

Таблиця 2 – Результати розрахунку

Міцність, г	Число проб $f$	Середина інтервалу $x$	$x - \bar{x}$	$\frac{x - \bar{x}}{\sigma} = t$	$\varphi(t)$	$154 \cdot \varphi(t) = f'$
120–130	1	125	-36,4	-2,80	0,0079	1
130–140	8	135	-26,4	-2,03	0,0508	8
140–150	27	145	-16,4	-1,26	0,1804	28
150–160	58	155	-6,4	-0,49	0,3538	55
160–170	56	165	3,6	0,28	0,3836	59
170–180	34	175	13,6	1,05	0,2299	35
180–190	14	185	23,6	1,82	0,0761	12
190–200	2	195	33,6	2,58	0,0143	2
Всього	200					200

Як видно з табл. 2 теоретичні частоти  $f'$  близькі до емпіричних  $f$ , хоча окремі відхилення мають місце.

Побудуємо графічне зображення емпіричного розподілу та теоретичного (імовірнісного) розподілу вирівняного за кривою нормального розподілу (розподілу Гауса) (рис. 3).



1 – емпіричний розподіл; 2 – теоретичний (імовірнісний) розподіл  
вирівняний за кривою нормального розподілу  
Рисунок 3 – Графічне зображення розподілів

Для судження про випадковість або суттєвість розходжень теоретичних  $f'$  та емпіричних  $f$  частот використовуємо критерії згоди:

1) Розрахунок критерія Пірсона показаний в табл. 3

Таблиця 3 – Розрахунок критерія Пірсона

$f$	$f'$	$(f - f')$	$(f - f')^2$	$\frac{(f - f')^2}{f'}$
1	1	0	0	0
8	8	0	0	0
27	28	-1	1	0,04
58	55	3	9	0,16
56	59	-3	9	0,15
34	35	-1	1	0,03
14	12	2	4	0,33
2	2	0	0	0
200	200			0,71

В прикладі ряд має 8 груп (класів) варіантів, і значить 8 частот. Тому число ступенів свободи для частот (при вирівнюванні за кривою нормального розподілу) дорівнює:  $\nu = 8 - 3 = 5$ . Прийmemo для розрахунків найбільш часто використовуемий рівень значущості  $\alpha = 0,05$  та скористаємось Додатком Б.

За таблицею значень  $\chi^2$ -критерія Пірсона для числа ступенів свободи  $\nu = 5$  визначимо, що  $\chi^2_{\text{табл.}} = 11,07$ .

Оскільки отримане в задачі фактичне значення  $\chi^2_{\text{факт.}} = 0,71$  менше табличного, то можна рахувати, що розходження між емпіричними і теоретичними частотами випадковими та зазначена гіпотеза о близькості емпіричного розподілу до нормального не спростовується.

2. Використаємо критерій Романовського.

$$K_p = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}} = \frac{|0,71 - 5|}{\sqrt{10}} = 1,4$$

Оскільки  $K_p = 1,4 < 3$ , то гіпотеза не спростовується і розходження між емпіричними і теоретичними частотами неістотні.

3. Перевіримо нашу гіпотезу за допомогою Критерія Колмогорова.

Запишемо накопичені частоти емпіричного і теоретичного розподілу і знайдемо максимальний розрив між ними.

Таблиця 4 – Розрахунок критерія Колмогорова

$f$	$f'$	Накопичені частоти		$ S - S' $
		Емпіричні (S)	Теоретичні ( $S'$ )	
1	2	3	4	5
1	1	1	1	0
8	8	9	9	0
27	28	36	37	1
58	55	94	92	2
56	59	150	151	1



1	2	3	4	5
34	35	184	186	2
14	12	198	198	0
2	2	200	200	0

Максимальний розрив  $D=2$ ,  $N = \sum f = 200$ , звідціля:

$$\lambda = \frac{D}{\sqrt{N}} = \frac{2}{\sqrt{200}} = 0,14.$$

За таблицею  $P(\lambda)$  (Додаток В) находимо, для  $\lambda = 0,14$   $P = 1,00$ .

Отже, цілком можна припустити, що гіпотеза про нормальний розподіл не спростовується, а розбіжність між  $f$  та  $f'$  носить випадковий характер.

### Приклад 2.

Протягом робочого тижня виконувалось спостереження за працею 50 верстатів та реєструвались поломки, що потребують зупинки обладнання для його регулювання. Результати спостережень наведені в табл. 5

Таблиця 5 – Результати спостережень за верстатами

Число несправностей ( $x$ )	0	1	2	3	4	5
Кількість верстатів ( $f$ )	14	16	10	7	2	1

Необхідно:

1) розрахувати імовірності і теоретичні частоти (число верстатів), враховуючи, що розподіл останніх за числом несправностей підпорядковується закону Пуассона;

2) Оцінити близькість емпіричних та теоретичних частот за допомогою критеріїв згоди Пірсона ( $\chi^2$ ), Романовського ( $K_p$ ), Колмогорова ( $\lambda$ ).

3) Побудувати графічне зображення розподілів.

### Розв'язання

1) Розрахуємо середню кількість несправностей за формулою середньої арифметичної зваженої:

$$\bar{x} = a = \frac{\sum x \cdot f}{\sum f} = \frac{0 \cdot 14 + 1 \cdot 16 + 2 \cdot 10 + 3 \cdot 7 + 4 \cdot 2 + 5 \cdot 1}{14 + 16 + 10 + 7 + 2 + 1} = \frac{70}{50} = 1,4$$

2) Знайдемо значення  $e^{-a} = e^{-1,4}$ :

$$e^{-1,4} = 0,2466$$

3) Підставивши в формулу (3) при знайденому значенні  $a = 1,4$  значення  $x = 0; 1; 2; 3; 4; 5$  отримаємо значення імовірностей ( $P_x$ ) наступних випадків несправностей від 0 до 5.

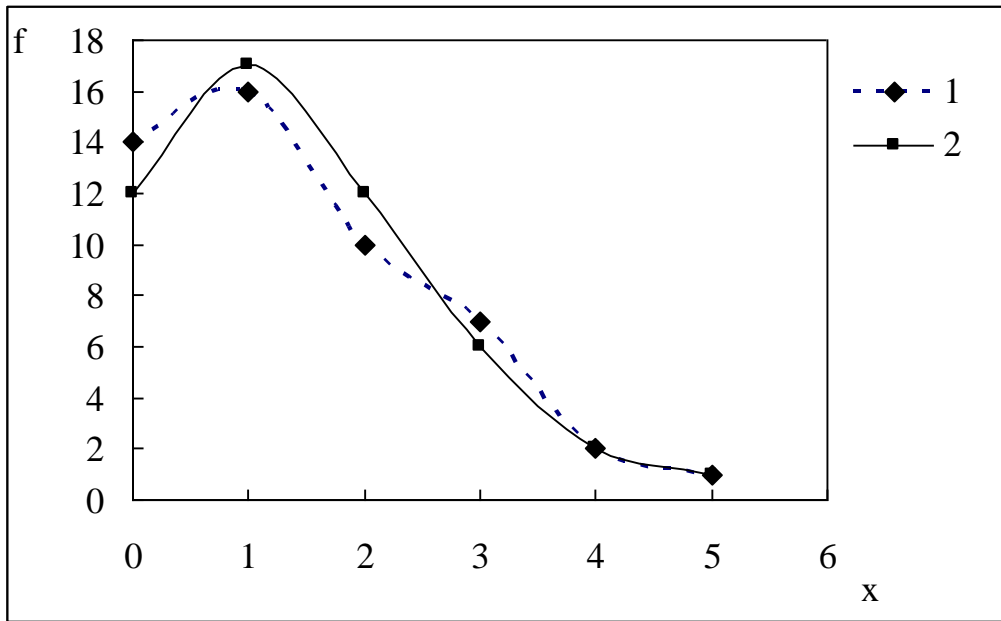
4) Помноживши останні на 50 (загальне число одиниць розподілу), отримаємо теоретичне число несправностей, тобто  $f' = N \cdot P_x$ .

Значення  $P_x$  та  $f'_x$  (округлені до цілих чисел) наведені в табл. 6.

Таблиця 6 – Значення  $P_x$  та  $f'_x$

x	$P_x$	$f'_x$ (теоретична частота) = $50 \cdot P_x$
0	0,2466	12
1	0,3452	17
2	0,2417	12
3	0,1128	6
4	0,0395	2
5	0,0111	1
	Всього	50

Побудуємо графічне зображення емпіричного розподілу та теоретичного (імовірнісного) розподілу вирівняного за кривою Пуассона (рис. 4).



1 – емпіричний розподіл; 2 – теоретичний (імовірнісний) розподіл  
вирівняний за кривою Пуассона

Рисунок 4 – Графічне зображення розподілів

Для оцінки близькості емпіричних та теоретичних частот скористаємось критеріями Пірсона ( $\chi^2$ ), Романовського ( $K_p$ ), Колмогорова ( $\lambda$ ).

1) Розрахунок критерія Пірсона показано в табл. 7.

Таблиця 7 – Розрахунок критерія Пірсона

$f$	$f'$	$(f - f')$	$(f - f')^2$	$\frac{(f - f')^2}{f'}$
14	12	2	4	0,33
16	17	-1	1	0,06
10	12	-2	4	0,33
7	6	1	1	0,17
2	2	0	0	0
1	1	0	0	0
Всього				0,89

В нашому прикладі ряд має 6 груп варіантів, і значить 6 частот.

Тому число ступенів свободи для частот (при вирівнюванні за кривою Пуассона) дорівнює:  $\nu = 6 - 2 = 4$ . Прийmemo для розрахунків рівень значущості  $\alpha = 0,05$  та скористаємось Додатком Б.

За таблицею значень  $\chi^2$ -критерія Пірсона для числа ступенів свободи  $\nu = 4$  визначимо, що  $\chi_{табл.}^2 = 9,49$ .

Оскільки отримане табличне значення в задачі значно більше фактичного значення  $\chi_{факт.}^2 = 0,89$ , то має все підґрунтя рахувати розходження між емпіричними і теоретичними частотами випадковими, а гіпотезу про те, що розподіл числа верстатів за кількістю несправностей протягом тижня за підпорядковується розподілу Пуассона неспростовною.

## 2. Використаємо критерій Романовського.

$$K_p = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}} = \frac{|0,89 - 4|}{\sqrt{8}} = 1,1$$

Оскільки  $K_p = 1,1 < 3$ , то розходження між емпіричними і теоретичними частотами можна вважати випадковими.

## 3. Перевіримо нашу гіпотезу за допомогою Критерія Колмогорова.

Запишемо накопичені частоти емпіричного і теоретичного розподілу і знайдемо максимальний розрив між ними (табл. 8).

Таблиця 8 – Розрахунок критерія Колмогорова

$f$	$f'$	Накопичені частоти		$ S - S' $
		Емпіричні (S)	Теоретичні ( $S'$ )	
14	12	14	12	2
16	17	30	29	1
10	12	40	41	1
7	6	47	47	0
2	2	49	49	0
1	1	50	50	0

Максимальний розрив  $D=2$ ,  $N = \sum f = 50$ , звідціля:

$$\lambda = \frac{D}{\sqrt{N}} = \frac{2}{\sqrt{50}} \approx 0,3.$$

За таблицею  $P(\lambda)$  (Додаток В) находимо, для  $\lambda = 0,3$   $P = 1,00$ .

Отже, всі три критерії оцінюють розходження між емпіричними і теоретичним частотами як випадкові, не спростовуючи тим самим гіпотезу про те, що розподіл кількості верстатів за числом несправностей протягом тижня підпорядковується закону Пуассона.

### Завдання для самостійної роботи

#### Завдання 1.

Нижче наведено розподіл робітників підприємства за показником виконання денної норми виготовлення деталей (табл. 9).

Таблиця 9 – Результати перевірки виконання денної норми виготовлення деталей робітниками підприємства

Варіант	Розподіл результатів перевірки								
	Відсоток виконання норми, %	94-96	96-98	98-100	100-102	102-104	104-106	106-108	108-110
1	Число робітників	2	7	13	26	19	10	2	1
2		1	5	12	28	20	9	2	1
3		2	6	20	27	12	9	2	2
4		3	9	19	29	15	8	3	1
5		1	5	18	30	16	11	2	2
6		2	6	17	31	17	9	2	1
7		3	7	18	32	18	8	1	1
8		2	8	19	34	19	9	3	1
9		1	9	14	33	20	8	2	1
10		2	7	13	35	18	11	3	1

1. Перевірити, можна рахувати чи ні даний розподіл нормальним.

2. Перевірити, випадкові або ні розходження між емпіричними і теоретичними частотами, використовуючи критерії Пірсона ( $\chi^2$ ), Романовського ( $K_p$ ), Колмогорова ( $\lambda$ ).

3. Побудувати графічне зображення розподілів.

### Завдання 2.

Для контролю якості продукції робітників машинобудівного підприємства, що виготовляють однотипну продукцію, перевірені вироби виготовлюємі кожним робітником. Було виявлено, що кількість бракованих деталей лежить в межах від 0 до 4 штук. Результати перевірки наступні згідно табл. 10.

Таблиця 10 – Результати перевірки якості продукції робітників

Варіант	Число робітників (f)	Кількість бракованих деталей з числа перевірених (x)				
		0	1	2	3	4
1	110	59	26	4	1	
2	100	50	30	3	2	
3	105	55	32	2	1	
4	102	56	28	3	1	
5	101	51	25	2	1	
6	97	50	27	4	2	
7	96	59	26	3	1	
8	94	54	21	5	1	
9	95	57	23	4	1	
10	91	49	32	6	2	

1. Розрахувати середню кількість бракованих виробів на одного робітника.

2. Розрахувати теоретичні частоти виходячі з гіпотези розподілу Пуассона.

3. Перевірити, випадкові або ні розходження між емпіричними і теоретичними частотами, використовуючи критерії Пірсона ( $\chi^2$ ), Романовського ( $K_p$ ) та Колмогорова ( $\lambda$ ). Прийняти рівень значущості  $\alpha = 0,05$ .

4. Побудувати графічне зображення розподілів.

### Контрольні питання

1. Що розуміють під вирівнюванням варіаційних рядів?
2. Наведіть графік кривої нормального розподілу та поясніть його.
3. Вкажіть як визначається теоретичні частоти ряду при вирівнюванні варіаційного ряду за кривою нормального розподілу.
4. Наведіть графік розподілу Пуассона та поясніть його.
5. Зазначте як розраховуються теоретичні частоти при вирівнюванні емпіричних даних за кривою Пуассона.
6. Які критерії згоди використовують для оцінки близькості емпіричних ( $f$ ) та теоретичних ( $f'$ ) частот?
7. Як розраховується критерій згоди Пірсона ( $\chi^2$ )?
8. Як розраховується критерій згоди Романовського ( $K_p$ )?
9. Як розраховується критерій згоди Колмогорова ( $\lambda$ )?
- 10) Зазначте як визначається ступінь свободи для критерія згоди Пірсона ( $\chi^2$ )?

### Література

1. Єріна А. М. Теорія статистики : практикум \ А. М. Єріна, З. О. Пальян. Київ : Знання, 2004. 255 с.
2. Практикум по теории статистики: учебное пособие / под ред. проф. Р. А. Шмойловой. Москва : Финансы и статистика, 2003. 416 с.
3. Таблиці функцій та критичних точок розподілів. Розділи: Теорія ймовірностей. Математична статистика. Математичні методи в психології. / Укладач: М.М. Горонескуль. Харків : УЦЗУ, 2009. 90 с.

Значення функції  $\varphi(t) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}$

t	Соті долі									
	0	1	2	3	4	5	6	7	8	9
0,0	3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
4,0	0001	0001	0001	0000	0000	0000	0000	0000	0000	0000

*Примітка.* В таблиці наведені мантиси значень функції (0,...).



**Значення  $\chi^2$ -критерію Пірсона при рівні значущості 0,10, 0,05, 0,01  
та числі ступенів свободи  $\nu$**

$df(\nu)$	0,10	0,05	0,01
1	2,71	3,84	6,63
2	4,61	5,99	9,21
3	6,25	7,81	11,34
4	7,78	9,49	13,28
5	9,24	11,07	15,09
6	10,64	12,59	16,81
7	12,02	14,07	18,48
8	13,36	15,51	20,09
9	14,68	16,92	21,67
10	15,99	18,31	23,21
11	17,28	19,68	24,72
12	18,55	21,03	26,22
13	19,81	22,36	27,69
14	21,06	23,68	29,14
15	22,31	25,00	30,58
16	23,54	26,30	32,00
17	24,77	27,59	33,41
18	25,99	28,87	34,81
19	27,20	30,14	36,19
20	28,41	31,41	37,57
21	29,62	32,67	38,93
22	30,81	33,92	40,29
23	32,01	34,17	41,64
24	33,20	36,42	42,98
25	34,38	37,65	44,31
26	35,56	38,89	45,64
27	36,74	40,11	46,96
28	37,92	41,34	48,28
29	39,09	42,56	49,59
30	40,26	43,77	50,89
40	51,80	55,76	63,69
50	63,17	67,50	76,15
60	74,40	79,08	88,38

100	118,50	124,34	135,81
-----	--------	--------	--------

Значення функції  $P(\lambda)$ 

$\lambda$	$P$
0,30	1,0000
0,35	0,9997
0,40	0,9972
0,45	0,9874
0,50	0,9639
0,55	0,9228
0,60	0,8643
0,65	0,7920
0,70	0,7112
0,75	0,6272
0,80	0,5441
0,85	0,4653
0,90	0,3927
0,95	0,3275
1,00	0,2700
1,10	0,1777
1,20	0,1122
1,30	0,0681
1,40	0,0397
1,50	0,0222
1,60	0,0120
1,70	0,0062
1,80	0,0032
1,90	0,0015
2,00	0,0007
2,10	0,0003
2,20	0,0001
2,30	0,0001
2,40	0,0000
2,50	0,0000

**Навчальне видання**

СЕМЕНОВ Євгеній Олександрович  
РАЙКО Валентина Федорівна  
ЇЛЬІНСЬКА Ольга Ігорівна

**Числові методи аналізу з охорони праці**

Практикум для студентів  
освітньо-кваліфікаційного рівня  
бакалавр спеціальності 263 «Цивільна безпека»

Відповідальний за випуск проф. Березуцький В. В.  
Роботу до видання рекомендувала проф. Самойленко Н. М.  
В авторській редакції

План 2022 р., поз. 14

Підп. до друку 2022. Формат 60×84 1/12. Папір офсет.  
Друк – різнографія. Гарнітура Minion Pro. Ум. друк. арк. 7,625.  
Наклад 50 прим. Зам. № . Ціна договірна.

---

Видавничий центр НТУ «ХП».  
Свідоцтво про державну реєстрацію ДК № 5478 від 21.08.2017 р.  
61002, Харків, вул. Кирпичова, 2

---

Самостійне електронне видання