

Міністерство освіти і науки, молоді та спорту України  
Державний вищий навчальний заклад  
«Національний гірничий університет»

Пілов П.І.  
Анісімов М.Т.  
Анісімов В.М.

**Методи математичної статистики та теорії  
ймовірностей в збагаченні корисних копалин**

Навчальний посібник

Дніпропетровськ  
ДВНЗ «НГУ»

2012

УДК 519.21: 519.22: 622.7

ББК 22.172: 22.171: 33.4

П 32

Надано гриф Міністерством освіти і науки, молоді та спорту України як навчальний посібник для студентів вищих навчальних закладів, які навчаються за напрямом підготовки 6.050303 Переробка корисних копалин 18.10.2011р.

№ 1/11 - 9756

Рецензенти:

Т.А. Олійник, д-р техн. наук, професор (Криворізький технічний університет, завідувач кафедри збагачення корисних копалин)

О.Д. Полулях, д-р техн. наук ( ДП УкрНДІВуглезбагачення, начальник науково – дослідного відділу удосконалення технологій та схем)

П.І. Пілов, М.Т. Анісімов, В.М. Анісімов

П Методи математичної статистики та теорії ймовірностей в збагаченні корисних копалин: навч. пос./ П.І. Пілов, М.Т. Анісімов, В.М. Анісімов. Державний вищий навчальний заклад «Національний гірничий університет», 2012. – 125 с.

ISBN

Зміст видання відповідає освітньо – професійній програмі підготовки бакалаврів, спеціалістів і магістрів з напрямку підготовки «Переробка корисних копалин» (Галузевий стандарт вищої освіти України ГСВО ОПП - 04), зокрема програмі дисципліни «Математичні методи в збагаченні корисних копалин».

Викладені положення теорії ймовірностей і математичної статистики, які є єдино можливими для вирішення фахових питань галузі і які використовуються у основних навчальних дисциплінах спеціальності збагачення корисних копалин у такому об'ємі, як це необхідно для практичного їх застосування. А саме, відносно законів розподілу випадкових величин (чинників і характеристик корисних копалин, їх представлення та методи обробки), теорії оцінок чинників та методів визначення їх спроможності, кореляційного та дисперсійного аналізу, теорії гіпотез і практичного використання теорії випадкових функцій і процесів змін характеристик, теорії планованого експерименту.

УДК 519.21: 519.22: 622.7

ББК 22.172: 22.171: 33.4

ISBN

© П.І. Пілов, М.Т. Анісімов,  
В.М. Анісімов,

© ДВНЗ «НГУ», 2012

## Зміст

Вступ.....	7
1. Елементи теорії ймовірностей. Випадкові події .....	8
1.1. Об'єкт теорії ймовірностей і математичної статистики.....	8
1.2. Події і випробування.....	8
1.2.1. Класифікація випадкових подій.....	8
1.2.2. Ймовірність.....	9
1.3. Теорема складання ймовірностей.....	10
1.3.1. Теорема складання ймовірностей несумісних подій.....	10
1.3.2. Повна група подій і сума їх ймовірностей.....	11
1.3.3. Протилежні події і сума їх ймовірностей.....	12
1.3.4. Ймовірність та принцип практичної неможливості малоймовірних подій.....	12
1.4. Незалежні і залежні події, визначення їх ймовірностей.....	14
1.4.1. Теорема множення ймовірностей незалежних подій.....	14
1.4.2. Ймовірність появи хоча би однієї події.....	16
1.4.3. Умовні події і визначення їх ймовірностей.....	17
1.4.4. Теорема множення ймовірностей залежних подій.....	17
1.5. Сумісне застосування формул суми і добутку ймовірностей для сумісних подій.....	19
1.5.1. Теорема складання ймовірностей сумісних подій.....	19
1.5.2. Формула повної ймовірності.....	20
1.5.3. Ймовірність гіпотез. Формули Бейєса.....	21
1.6. Повторення випробувань і визначення ймовірностей.....	23
1.6.1. Формула Бернуллі.....	23
1.6.2. Локальна теорема Лапласа.....	24
1.6.3. Інтегральна теорема Лапласа .....	25
1.6.4. Ймовірність відхилення відносної частоти від постійної ймовірності в незалежних випробуваннях.....	25
1.7. Випадкові величини і їх особливості.....	27
1.7.1. Дискретні і безперервні випадкові величини.....	27
1.7.2. Задання випадкових величин.....	28
1.7.3. Числові характеристики дискретних випадкових величин... ..	31
1.8. Способи задання і числові характеристики безперервних випадкових величин.....	32
1.9. Система випадкових величин.....	33
1.9.1. Способи задання двовимірних випадкових величин.....	33
1.9.2. Числові характеристики системи двох випадкових величин.....	36
1.9.3. Залежні і незалежні випадкові величини.....	36
1.9.4. Числові характеристики системи двох випадкових величин. Кореляційний момент. Коефіцієнт кореляції.....	37
1.9.5. Корельованість і залежність випадкових величин.....	39
2. Математична статистика.....	40
2.1. Призначення математичної статистики.....	40

2.1.1. Генеральна і вибіркова сукупності.....	40
2.1.2. Повторна і безповторна вибірки. Репрезентативна вибірка..	41
2.1.3. Способи відбору.....	42
2.1.4. Статистичний розподіл вибірки.....	43
2.1.5. Емпірична функція розподілу.....	45
2.1.6. Полігон і гістограма.....	46
2.2. Статистичні оцінки параметрів розподілу.....	48
2.2.1. Незміщені, ефективні і спроможні оцінки.....	49
2.2.2. Генеральна середня.....	49
2.2.3. Вибіркова середня.....	50
2.2.4. Оцінка генеральною середньою по вибірковій середній.	
Стійкість вибірових середніх.....	50
2.2.5. Групова і загальна середні.....	51
2.2.6. Відхилення від загальної середньої і його властивість.....	52
2.2.7. Генеральна дисперсія.....	52
2.2.8. Вибіркова дисперсія.....	53
2.2.9. Групова, внутрішньогрупова, міжгрупова і загальна дисперсії.....	54
2.2.10. Складання дисперсій.....	56
2.2.11. Оцінка генеральної дисперсії по виправленій вибірковій... 2.2.12. Точність оцінки, довірча ймовірність (надійність).	56
Довірчий інтервал.....	57
2.2.13. Визначення довірчого інтервалу для оцінки математичного очікування чинника, що має нормальний закон розподілу при відомому $\sigma$ .....	58
2.2.14. Довірчі інтервали для оцінки математичного очікування чинника, що має нормальний закон розподілу при невідомому $\sigma$ .....	60
2.2.15. Оцінка дійсного значення вимірюваної величини.....	61
2.2.16. Довірчі інтервали для оцінки середнього квадратичного відхилення $\sigma$ чинника, що має нормальний закон розподілу.....	61
2.2.17. Оцінки точності вимірювань.....	62
2.2.18. Інші характеристики варіаційного ряду.....	62
2.2.19. Закономірності розподілу дискретних випадкових величин.....	63
2.2.20. Емпіричні і вирівнюючі частоти.....	64
2.2.21. Безперервний розподіл.....	65
2.2.22. Побудова кривої нормального закону розподілу чинника з використанням практичних даних.....	65
2.2.23. Оцінка відхилення емпіричного розподілу від нормального. Асиметрія і ексцес.....	67
2.3. Основи кореляційного аналізу.....	69
2.3.1. Функціональна, статистична і кореляційна залежності.....	70
2.3.2. Умовні середні. Кореляційна залежність.....	70
2.3.3. Основні завдання теорії кореляції.....	72

2.3.4. Розрахунок параметрів вибіркового рівняння прямої лінії регресії.....	72
2.3.5. Розрахунок параметрів вибіркового рівняння прямої лінії регресії за незгрупованими даними.....	72
2.3.6. Відшукання параметрів вибіркового рівняння прямої лінії регресії за згрупованими даними.....	75
2.3.7. Властивості вибіркового коефіцієнта кореляції.....	76
2.3.8. Вибіркове кореляційне відношення.....	77
2.3.9. Випадки криволінійної кореляції.....	77
2.3.10. Множинна кореляція.....	77
2.4. Статистична перевірка статистичних гіпотез.....	78
2.4.1. Статистична гіпотеза. Нульова і конкуруюча, проста і складна гіпотези.....	78
2.4.2. Помилки першого і другого роду.....	79
2.4.3. Статистичний критерій перевірки нульової гіпотези. Спостережуване значення критерію.....	80
2.4.4. Критична область. Область ухвалення гіпотези, критичні крапки.....	80
2.4.5. Відшукання правосторонньої критичної області.....	82
2.4.6. Відшукання лівобічної і двосторонньої критичних областей.....	82
2.4.7. Вибір критичної області. Потужність критерію.....	82
2.4.8. Порівняння двох дисперсій деяких ознак з нормальним законом розподілу.....	83
2.4.9. Порівняння виправленої вибіркової дисперсії деякого параметра з нормальним законом розподілу з його генеральною дисперсією.....	84
2.4.10. Порівняння двох середніх значень деяких параметрів, що мають нормальний закон розподілу, дисперсії яких відомі.....	86
2.4.11. Порівняння двох середніх значень деяких параметрів нормально розподілених, дисперсії яких невідомі і однакові.....	87
2.4.13. Перевірка гіпотези про нормальний розподіл генеральної сукупності деякого параметра.....	88
3. Випадкові функції.....	88
3.1. Визначення випадкової функції випадкового процесу.....	88
3.2. Взаємозв'язок випадкової функції і системи випадкових величин. Закон розподілу випадкової функції.....	92
3.3. Характеристики випадкових функцій.....	93
3.4. Властивості кореляційної функції.....	98
3.5. Практичне визначення характеристик випадкової функції з дослідів.....	99
3.6. Поняття про стаціонарний випадковий процес.....	100
3.7. Спектральне розкладання стаціонарної випадкової функції.....	109
4. Планування експерименту.....	110
4.1. Основи методу.....	111

4.1.1. Повний експеримент.....	111
4.1.2. Визначення числа дослідів, які необхідно провести при планованому експерименті.....	111
4.1.3. План експерименту.....	113
4.2. Проведення експериментів.....	113
4.2.1. Обробка результатів експерименту.....	115
4.2.2. Оцінка результатів планованого експерименту.....	116
4.3. Дробовий експеримент.....	117
Додатки.....	118
Література.....	126
Предметний покажчик.....	126

## Вступ

Збагачення корисних копалин це одна з важливих і достатньо розвинених галузей держави. Головною метою збагачення є підготовка сировинних ресурсів до їх подальшого використання. Збагачувальні підприємства фабрики і комбінати переробляють значні кількості, що обчислюються десятками і сотнями мільйонів тон на рік природньої сировини. Схеми збагачення є складними як по використовуваних процесах, так і по насиченості устаткуванням, що зумовлює високу собівартість збагачення.

На жаль, в процесі збагачення мають місце значні небажані втрати цінних компонентів корисних копалин. Тому в збагаченні завжди нагальним є питання про зниження собівартості переробки і виключення втрат мінералів. Зменшення витрат досягається, як правило, вдосконаленням технологічних схем і устаткування, в цьому плані в збагаченні досягнутий деякий рівень, причому визначальними в розвитку були експериментальні дослідження і практично не використані теоретичні. Такий стан можна пояснити тим, що самі збагачувальні процеси багаточинникові, а їх математичний опис є складним, а також тим, що відсутні достатньо відпрацьовані методики математичної формалізації в даній галузі.

Розвиток обчислювальної техніки і поява достатньо простих і потужних систем програмування дозволяє вирішувати практично будь-які завдання даної галузі, розробляти математичні залежності процесів і об'єктів, здійснювати на них всебічне дослідження стосовно їх вдосконалення. Формалізація технологій в гірничо-збагачувальній галузі вимагає глибоких знань різних розділів математики.

Унаслідок того, що в видобутку і збагаченні переважне число чинників мають випадковий характер, в цьому навчальному посібнику викладені елементи теорії ймовірності і математичної статистики так, як це необхідно для виконання інженерних розрахунків, розробки математичних моделей, проведення теоретичних та практичних досліджень з реальними об'єктами і з математичними моделями, оцінки формалізованих замінників і результатів досліджень, отриманих за допомогою моделей і прямих дослідів, а також практичне використання результатів.



## **1. Елементи теорії ймовірностей. Випадкові події**

### **1.1. Об'єкт теорії ймовірностей і математичної статистики**

Об'єктом теорії ймовірності і математичної статистики є реально існуючі і спостережувані нами явища - події.

Наприклад, «холодно», «закінчилася зима», «вступ до вузу», «отримання роботи» і таке інше.

У всій сукупності явища можна класифікувати з погляду їх формалізованого представлення на три види: достовірні, неможливі і випадкові.

*Достовірними називають події*, які мають місце (обов'язково відбуваються) за наявності певної сукупності умов. Наприклад, при нормальному атмосферному тиску і плюсовій температурі вода знаходиться в рідкому стані. Тут атмосферний тиск і температура – це сукупність умов.

*Неможливими називають події*, які не відбуваються (не наступають) за наявності певної сукупності умов. Наприклад, подія – «вода знаходиться в твердому стані» дійсно не наступить, якщо збережеться сукупність умов, вказаних для достовірних подій. .

*Випадковими називають події*, які за наявності певної сукупності умов можуть відбуватись - наступити або не відбуватись – не наступити. Наприклад, якщо кинути монету, то вона може впасти однією стороною, або іншою.

### **1.2. Події і випробування**

Здійснення певної сукупності умов, при якій випадкова подія з'явиться або не з'явиться, називають «випробуванням». В такому формулюванні вміщується процес і результат випробування.

Наприклад, на підприємстві здійснюється переробка корисних копалин по певній технологічній схемі з отриманням кінцевих продуктів заданої якості. В даному випадку переробка корисної копалини по певній технологічній схемі – це випробування. Отриманням кінцевих продуктів заданої якості – подія - результат.

#### **1.2.1. Класифікація випадкових подій**

Серед подій достовірних, неможливих і випадкових з погляду їх формалізованого представлення особливе місце займають випадкові. Виходячи з формулювання стосовно визначення випадкових подій, виникають питання – наскільки вони випадкові, чи можна ввести чисельну міру поняття «випадкові». Для вирішення цього питання у відповідності до їх суті прийнята класифікація випадкових подій на сумісні і несумісні, єдино можливі і рівно можливі, залежні і незалежні.

*Сумісними вважають випадкові події*, якщо одна з них не виключає появи інших випадкових подій в одному і тому ж випробуванні.

Наприклад, якщо переробка здійснюється на двох ідентичних технологічних лініях, то «досягнення заданої якості» на обох лініях є сумісною подією.

*Несумісними вважають випадкові події*, якщо настання однієї з них виключає появу інших випадкових подій в одному і тому ж випробуванні.

Наприклад, одночасне «досягнення якості» і «не досягнення якості» при переробці корисної копалини на одній технологічній лінії події несумісні.

*Єдино можливими* вважають всі випадкові події, які можуть наступити в результаті випробувань.

Наприклад, якщо розглядати попередній приклад, то події «якість досягнута» і «якість відсутня» є єдино можливими.

*Рівноможливими вважають випадкові події*, для яких немає підстав вважати, що будь-яка з них ймовірніша ніж інші в одному і тому ж випробуванні.

Наприклад, якщо працюють дві ідентичні збагачувальні технологічні секції, то немає підстав вважати, що «досягнення якості» на одній з них має більші переваги, ніж на іншій.

### 1.2.2. Ймовірність

Чисельною мірою поняття «випадковість» для випадкових подій є ймовірність.

Визначення ймовірності пов'язане з випробуваннями, тобто здійсненням певної сукупності умов і можливих результатів, для яких розглядається деяке явище (подія).

Наприклад, у пробі знаходиться десять зразків породи і два зразки вугілля. Навмання, тобто, не дивлячись, виймається один зразок.

В результаті такого випробування можливі результати: з'явиться породний зразок, з'явиться вугільний зразок.

Введемо позначення  $\Pi_1, \Pi_2, \dots, \Pi_{10}$  - породні зразки, а  $Y_1, Y_2$  - вугільні.

Виймається один зразок з дванадцяти – це *випробування*, в результаті такого випробування можливі *результати*: з'явиться породний зразок, з'явиться вугільний зразок. Породний зразок братиметься з 10, а вугільний - з 2 зразків. Кількість зразків, з яких здійснюється відбір, що *сприяє* його появи, позначимо –  $m_1$  для породних зразків і  $m_2$  - для вугільних зразків. Знайдемо відношення кількості сприяючих результатів для вугільного і породного зразків до загальної кількості зразків, що складають сукупність -  $n = m_1 + m_2, n = 10 + 2 = 12$ . Вказане відношення позначають  $P$  і називають ймовірністю. Поряд з позначенням ймовірності -  $P$  в дужках вказується та подія, для якої вказана ймовірність визначається. У нашому прикладі позначення  $P(\Pi)$  – ймовірність появи породного зразка,  $P(Y)$  – ймовірність появи вугільного зразка.

Вказані відношення та їх числові результати:

$$\text{для породи} \quad P(\Pi) = \frac{m_1}{n} = \frac{10}{12},$$

$$\text{для вугілля} \quad P(Y) = \frac{m_2}{n} = \frac{2}{12}.$$

Грунтуючись на результатах наведеного прикладу, можна зробити формулювання – визначення поняття «ймовірність».

*Ймовірністю деякої події* називають відношення числа результатів випробування, що сприяють появі даної події до загального числа результатів, передбачуваних до виконання випробувань.

З визначення ймовірності виходить, що ймовірність - це число, яке може приймати значення від нуля до одиниці, причому, якщо ймовірність дорівнює нулю, то подія «неможлива», якщо ймовірність дорівнює одиниці, то подія «достовірна».

### **Питання**

1. Наведіть приклади подій.
2. Наведіть приклади достовірних подій.
3. Наведіть приклади неможливих подій.
4. Наведіть приклади випадкових подій.
5. Наведіть приклади несумісних подій.
6. Наведіть приклади сумісних подій.
7. Наведіть приклади єдиноможливих подій.
8. Наведіть приклади рівноможливих подій.

### **Завдання**

1. У відділ технічного контролю для обробки поступила партія з п'яти проб під номерами: 10, 11, 12, 13, 14. Знайти ймовірність того, що обробці піддаватиметься проба з номером 13.

Відповідь  $P = 1/120$ .

2. Для умов попереднього прикладу знайти ймовірність обробки проби з номером 10 за умови, що одна з проб оброблена.

Відповідь  $P = 1/24$ .

3. На складі є 5 редукторів з різними передавальними числами, на які відсутні паспортні дані. Для установки на устаткування отримано один редуктор. Потрібно визначити ймовірність того, що передавальне число взятого редуктора таке, як потрібно.

Відповідь  $P = 1/5$ .

4. Для умов першого прикладу знайти ймовірність того, що першою оброблятиметься проба з номером 10.

Відповідь  $P = 1/120$ .

5. Для умов першого прикладу знайти ймовірність того, що другою оброблятиметься проба з номером 11, якщо першою оброблена проба з номером 10.

Відповідь  $P = 1/24$ .

### **1.3. Теорема складання ймовірностей подій**

Приймаючи до уваги те, що ймовірності подій оцінюються числами, то загальновідомо, що з числами можна виконувати певні арифметичні або логічні дії. У подальшому розглядаються такі дії у відповідності до специфічності подій.

#### **1.3.1. Теорема складання ймовірностей несумісних подій**

Спочатку розглянемо особливості подій

Сумою двох подій  $A$  і  $B$  називають подію, що полягає в появі події  $A$  або події  $B$ , або обох цих подій.

Наприклад, якщо технологічна секція збагачення працює дві зміни, то можливі події:  $A$  — досягнення якості продукту збагачення в першу зміну,  $B$  — досягнення якості в другу зміну,  $A$  або  $B$  — досягнення якості в першу зміну, або в другу, або в обох змінах. В даному випадку несумісність полягає у тому, що технологічні зміни роз'єднані у часі.

Зокрема, якщо дві події  $A$  і  $B$  несумісні, то  $A$  або  $B$  — подія, що полягає в появі однієї з цих подій, байдуже якої. Логічне «або» для будь-якої події, замінюється арифметичним «+», тобто, для наведеного прикладу можемо записати суму несумісних подій  $A$  або  $B$  -  $A + B$ .

Сумою декількох подій називають подію, яка полягає в появі хоч би однієї з цих подій.

Наприклад, подія  $A + U + B$  полягає в появі однієї з наступних подій:  $A$ ,  $B$ ,  $U$ ,  $A$  або  $B$ ,  $A$  або  $U$ ,  $B$  або  $U$ ,  $A$  або  $B$  або  $U$ .

Якщо для вищерозглянутих подій визначити їх ймовірності, то можна визначити їх взаємодію, використавши теорему складання ймовірностей.

Теорема. Ймовірність появи однієї з двох несумісних подій, байдуже якої, дорівнює сумі ймовірностей цих подій:

$$P(A + B) = P(A) + P(B).$$

Доведення теореми. Введемо позначення:  $n$  — загальне число можливих елементарних результатів випробування;  $m_1$  — число результатів, що сприяють події  $A$ ;  $m_2$  - число результатів, що сприяють події  $B$ . Загальне число елементарних результатів, що сприяють появі події  $A$  або події  $B$ , дорівнює  $m_1 + m_2$ .

Використавши теорему визначення ймовірностей, знайдемо:

$$P(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n}$$

З попереднього викладу відомо, що:

$$P(A) = \frac{m_1}{n}, \quad P(B) = \frac{m_2}{n},$$

підставивши ці значення в попередній вираз, отримаємо:

$$P(A + B) = P(A) + P(B).$$

Якщо події попарно несумісні, то ймовірність появи однієї з декількох цих попарно несумісних, байдуже яких, дорівнює сумі ймовірностей цих подій:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Доказ. Розглянемо три події  $A$ ,  $B$  і  $C$ . Так як ці події попарно несумісні, то поява однієї з них  $A$ ,  $B$  і  $C$  рівносильно появі однієї з двох подій  $A + B$  і  $C$ , тому:

$$P(A+B+C) = P[(A+B)+C] = P(A+B)+P(C) = P(A)+P(B)+P(C).$$

### 1.3.2. Повна група подій і сума їх ймовірностей

Повною групою називають сукупність єдино можливих подій випробування.

Приклад. В результаті роботи технологічної секції протягом двох змін стосовно одержання кількості і якості кінцевих продуктів можливі наступні результати.  $A_1$  – план за якістю і кількістю виконаний в одну зміну,  $A_2$  - план за якістю і кількістю, виконаний в обох змінах,  $A_3$  - план за якістю і кількістю, не виконаний в двох змінах.

Теорема. Сума ймовірностей подій  $A_1, A_2, \dots, A_n$ , створюючих повну групу, дорівнює одиниці:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

До цього приходимо таким чином. Оскільки поява однієї з подій повної групи достовірна, тобто одна з подій обов'язково наступить, а ймовірність достовірної події дорівнює одиниці, то

$$P(A_1 + A_2 + \dots + A_n) = 1.$$

Будь-які дві події повної групи несумісні, тому можна застосувати теорему складання:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

отже

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

### 1.3.3. Протилежні події і сума їх ймовірностей

*Протилежними називають* дві єдино можливі події, які створюють повну групу. Якщо одна з двох протилежних подій позначена через  $A$ , то іншу – протилежну, позначають  $\bar{A}$ .

Наприклад, відвідання і пропуск заняття студентом - протилежні події. Якщо  $A$  - відвідання, то  $\bar{A}$  - пропуск.

Теорема. Сума ймовірності протилежних подій дорівнює одиниці:

$$P(A) + P(\bar{A}) = 1.$$

Доказ. Протилежні події створюють повну групу, а сума ймовірності подій, створюючих повну групу, дорівнює одиниці.

Якщо ймовірність одного з двох протилежних подій позначена через  $p$ , а ймовірність іншої позначають через  $q$ , то, у відповідності з теоремою:

$$p + q = 1$$

Приклад. Ймовірність того, що якість продукту збагачення відповідає нормі  $p = 0,9$ , знайти ймовірність того, що якість не одержана.

Рішення. Події «якість продукту збагачення відповідає нормі» і «якість не одержана» — протилежні, тому скористаємося попередньою формулою, заздалегідь позначивши  $p$  - якість продукту збагачення відповідає нормі,  $q$  - якість не одержана

$$q = 1 - p = 1 - 0,9 = 0,1.$$

### 1.3.4. Ймовірність та принцип практичної неможливості малоїмовірних подій

При вирішенні багатьох практичних питань зустрічаються події, ймовірності яких дуже малі - близькі до нуля. Чи правомірно припустити, що малоїмовірна подія  $A$  в одиничному випробуванні не відбудеться. Такого висновку зробити не можна, оскільки не виключено, що подія  $A$  наступить.

З практики повсякденного життя відомо, що малоймовірна подія в одиничному випробуванні в більшості випадків не настає. На підставі цього прийнятий «принцип практичної неможливості малоймовірних подій», суть якої полягає у тому, то практично можна вважати, що в одиничному випробуванні ця подія не наступить, якщо випадкова подія має дуже малу ймовірність.

Використання вказаного принципу потребує належної уваги. Наприклад, якщо ймовірність того, що при проведенні хірургічної операції ймовірність припущення помилки складе 0,01, то було б неприпустимим проводити операцію. Якщо ж ймовірність того, що відмова двигуна технологічного апарату складе 0,01, то на таку передмову можна не звертати увагу – двигун можна запускати.

Прийняте припущення (принцип) застосовується таким чином. Достатньо малу ймовірність, при якій подію можна рахувати практично неможливою, називають рівнем значущості. У кожному конкретному випадку приймають рівні значущості - 0,01 і 0,05. Рівні значущості представляють у процентному відношенні

Розглянутий принцип дозволяє робити прогнози не тільки про події, що мають малу ймовірність, але і про події, ймовірність яких близька до одиниці. Дійсно, якщо подія  $A$  має ймовірність близьку до нуля, то ймовірність протилежної події  $\bar{A}$  близька до одиниці. З іншого боку, неоява події  $A$  означає настання протилежної події  $\bar{A}$ . Таким чином, з принципу неможливості малоймовірних подій є наступне : якщо випадкова подія має ймовірність дуже близьку до одиниці, то практично можна вважати, що в одиничному випробуванні ця подія наступить і навпаки.

### **Питання**

1. Сформулюйте письмово теорему складання ймовірностей несумісних подій і напишіть формулу.
2. Приведіть приклад подій, що складають повну групу, вкажіть теорему, за допомогою якої проводяться обчислення ймовірності такої групи
3. Приведіть приклад протилежних подій і теорему для обчислення ймовірності їх суми.
4. Вкажіть особливість принципу «практичної неможливості малоймовірних подій» і практичний приклад його застосування.

### **Завдання**

1. Проводиться розділення матеріалу по крупності часток на три частини. Ймовірність попадання матеріалу в продукт заданої крупності часток дорівнює 0.85, в продукт, крупність часток якого більше заданого, дорівнює 0.1 і в продукт з найбільшою крупністю - 0.05. Потрібно знайти ймовірність того, що матеріал потрапить або в найкрупнішу, або в середню крупність.

Відп.  $P = 0.15$ .

2. Для умов попереднього прикладу потрібно знайти ймовірність того, що матеріал потрапить або в задану, або в середню крупність

Відп.  $P = 0.95$

3. Технологічна секція працює протягом п'яти змін. Потрібно знайти ймовірність того, що хоч би протягом однієї зміни з п'яти не буде досягнута якість, якщо ймовірність досягнення якості дорівнює 0.8.

Відп.  $P = 3.2 \cdot 10^{-4}$ .

#### **1.4. Незалежні і залежні події, визначення їх ймовірностей**

Окрім сумісних і несумісних, єдино можливих і рівно можливих випадкові події можуть бути подіями залежними і незалежними.

Дві події називають незалежними, якщо ймовірність однієї з них не залежить від появи або не появи іншої.

Приклад 1. Протягом однієї зміни працюють дві технологічні секції збагачення. Ймовірність отримання якості продукту збагачення першою секцією не залежить від виконання завдання другою і навпаки, тому дані події для двох секцій є незалежними.

Приклад 2. Протягом однієї зміни працює одна технологічна секція. Події дотримання регламентної технологічної норми і не дотримання події - несумісні, протилежні і їх ймовірність залежить один від одного. Такі події називають залежними.

Декілька подій називають попарно незалежними, якщо кожні дві з них незалежні.

Дві події називають залежними, якщо ймовірність появи однієї з них залежить від настання або ненастання іншої події.

##### **1.4.1. Теорема множення ймовірностей незалежних подій**

Результатом випробування двох подій  $A$  і  $B$  називають подію  $AB$ , що полягає в сумісній появі (поєднанні) цих подій.

Наприклад, переробка корисної копалини, що складається з частинок вугілля і породи, полягає в її розділенні по питомій густині за допомогою двох гідроциклонів по Дгр. на два продукти. Умовимось;  $A$  – подія, яка полягає у тому, що відповідно продукт - частинки вугілля і породи менші Дгр. незалежно на якому гідроциклоні отримано,  $B_1$ , продукт, одержаний на першому гідроциклоні. Постає питання - знайти ймовірність вилучення частинок вугілля і породи необхідної крупності, якщо він одержаний на першому гідроциклоні. Вказана подія можлива, якщо з'являться події  $A$  і  $B_1$ . За умови поєднання подій логічним «і» ймовірність результату визначається з використанням теореми множення, тобто:  $A$  і  $B_1 = A \cdot B_1$

Теорема. Ймовірність сумісної появи двох незалежних подій дорівнює добутку ймовірності цих подій.

Формула для обчислення ймовірності добутку двох незалежних подій наступна:

$$P(AB_1) = P(A) \times P(B_1)$$

Доказ теореми визначається з наступного. Введемо позначення  $n$  — число можливих елементарних результатів випробування, в яких подія  $A$  настає або не настає;  $n_1$  — число результатів, що сприяють події  $A$ ;  $m$  — число можливих

елементарних результатів випробування, в яких подія  $B_1$  настає або не настає;  $m_1$  - число результатів, що сприяють події  $B_1$ .

Загальне число можливих елементарних результатів випробування дорівнює добутку  $n \cdot m$ , оскільки кожен з  $n$  результатів, в яких подія  $A$  настає або не настає, може поєднуватися з кожним з  $m$  результатів, в яких подія  $B_1$  з'являється або не з'являється.

Числа результатів, що сприяють сумісній появі подій  $A$  і  $B_1$ , дорівнює добутку  $n_1 m_1$ . Враховуючи вищесказане, можна записати:

$$P(AB_1) = \frac{n_1 m_1}{nm} = \frac{n_1}{n} \times \frac{m_1}{m},$$

з урахуванням того, що

$$P(A) = \frac{n_1}{n} \text{ а } P(B_1) = \frac{m_1}{m}$$

отримаємо:

$$P(AB_1) = P(A) \times P(B_1)$$

Ймовірність сумісної появи декількох подій, незалежних в сукупності, рівна добутку ймовірностей цих подій:

$$P(A_1 A_2 \dots A_n) = P(A_1) \cdot P(A_2) \dots P(A_n)$$

Доведення теореми здійснюється аналогічно, як і для двох незалежних подій.

Приклад. На складі є 2 партії електричних двигунів, що містять по 10 одиниць. У першій партії 8, а в другій 7 двигунів з необхідним числом обертів. З кожної партії беруть по одному двигуну. Знайти ймовірність того, що обидва двигуни матимуть необхідне число обертів.

Рішення. Ймовірність того, що перший двигун має необхідне число обертів (подія  $A$ ) рівна:

$$P(A) = \frac{n_1}{n} = \frac{8}{10} = 0.8$$

Ймовірність того, що з другого ящика вийнятий двигун з необхідним числом обертів (подія  $B$ ) рівна:

$$P(B) = \frac{m_1}{m} = \frac{7}{10} = 0.7$$

Оскільки події  $A$  і  $B$  незалежні в сукупності, то шукана ймовірність (за теоремою множення) визначиться:

$$P(AB) = P(A) P(B) = 0,8 \cdot 0,7 = 0,56.$$

Зустрічаються задачі, в яких необхідно для незалежних подій застосовувати спільно теорему множення і складання, яким чином вирішуються такі задачі наведено у наступному прикладі.

Приклад. Ймовірність появи кожної з трьох незалежних подій  $A_1$ ,  $A_2$  і  $A_3$  відповідно рівні  $p_1$ ,  $p_2$ ,  $p_3$ . Знайти ймовірність появи тільки однієї з цих подій.

Рішення. Поява тільки  $A_1$  першої події рівносильно появі події  $A_1 \bar{A}_2 \bar{A}_3$  (з'явилася перша і не з'явилися друга і третя події).

Позначимо:

$\Pi_1$  - з'явилася тільки подія  $A_1$ , тобто  $\Pi_1 = A_1 \bar{A}_2 \bar{A}_3$



$\Pi_2$  - з'явилася тільки подія  $A_2$ , тобто  $\Pi_2 = A_2 \bar{A}_1 \bar{A}_3$  ;

$\Pi_3$  - з'явилася тільки подія  $A_3$ , тобто  $\Pi_3 = A_3 \bar{A}_1 \bar{A}_2$  .

Так, щоб знайти ймовірність появи тільки однієї з подій  $A_1$ ,  $A_2$  і  $A_3$ , знайдемо ймовірність  $P(\Pi_1 + \Pi_2 + \Pi_3)$  появи однієї, байдуже якої з подій  $\Pi_1$ ,  $\Pi_2$ ,  $\Pi_3$ . Оскільки ці події несумісні, то застосована теорема складання:

$$P(\Pi_1 + \Pi_2 + \Pi_3) = P(\Pi_1) + P(\Pi_2) + P(\Pi_3)$$

Знайдемо ймовірність кожної з подій  $\Pi_1$ ,  $\Pi_2$ ,  $\Pi_3$ .

Події  $A_1$ ,  $A_2$  і  $A_3$  незалежні, отже, незалежні події  $A_1 \bar{A}_2 \bar{A}_3$ ,  $A_2 \bar{A}_1 \bar{A}_3$  і  $A_3 \bar{A}_1 \bar{A}_2$ , тому до них застосована теорема множення:

$$P(\Pi_1) = P(A_1 \bar{A}_2 \bar{A}_3) = P(A_1)P(\bar{A}_2)P(\bar{A}_3) ;$$

$$P(\Pi_2) = P(A_2 \bar{A}_1 \bar{A}_3) = P(A_2)P(\bar{A}_1)P(\bar{A}_3) ;$$

$$P(\Pi_3) = P(A_3 \bar{A}_1 \bar{A}_2) = P(A_3)P(\bar{A}_1)P(\bar{A}_2)$$

Підставивши цю ймовірність в попередній вираз, знайдемо невідому ймовірність появи тільки однієї з подій  $A_1$ ,  $A_2$  або  $A_3$  .

$$P(\Pi_1 + \Pi_2 + \Pi_3) = P(A_1)P(\bar{A}_2)P(\bar{A}_3) + P(A_2)P(\bar{A}_1)P(\bar{A}_3) + P(A_3)P(\bar{A}_1)P(\bar{A}_2)$$

#### 1.4.2. Ймовірність появи хоч би однієї події

Якщо в результаті випробування може з'явитися  $n$  подій незалежних в сукупності, або деякі з них, тільки одна або жодної, причому ймовірності появи кожної з подій відомі, то ймовірність того, що наступить хоч би одна з цих подій визначається по теоремі.

Теорема. Ймовірність появи хоч би однієї з подій  $A_1, A_2 \dots A_n$  незалежних в сукупності, дорівнює різниці між одиницею і добутком ймовірності протилежних подій  $\bar{A}_1 \bar{A}_2 \bar{A}_3 (q_1 q_2 \dots q_n)$

$$P(A) = 1 - q_1 q_2 \dots q_n$$

Доведення даної теореми виходить з того, що події  $A$  і  $\bar{A}_1 \bar{A}_2 \bar{A}_3$  протилежні, отже, сума їх ймовірності дорівнює одиниці, отже

$$P(A) = 1 - q_1 q_2 q_3$$

Якщо події  $A_1, A_2, \dots, A_n$  мають однаковою ймовірність, рівну  $p$ , то ймовірність появи хоч би однієї з цих подій

$$P(A) = 1 - q^n$$

Приклад. . Ймовірність пропуску занять студентом в один день складає  $p = 0.1$ . Яка ймовірність того, що протягом п'яти днів студент зробить хоч би один пропуск.

Рішення. Вважаємо, що ймовірність пропуску заняття студентом не залежить від попереднього і подальшого днів. Тобто події – пропуск заняття в один день незалежні. Позначимо  $A$  – подія «пропуск студентом одного заняття протягом п'яти днів»,  $q$  – ймовірність протилежної події.

Знайдемо ймовірність  $q$  не пропуску занять студентом.

$$q = 1 - p = 1 - 0.1 = 0.9$$

Унаслідок того, що  $p = 0.1$  є постійною, то при рішенні скористаємося формулою

$$P(A) = 1 - q^n$$

де  $n$  – число днів.

Підставивши значення у формулу, отримаємо:

$$P(A) = 1 - 0.90^5 = 0.55$$

### 1.4.3. Умовні події і визначення їх ймовірностей

Умовимось, що події  $A$  і  $B$  залежні. З визначення залежних подій виходить, що ймовірність однієї з подій залежить від появи або не появи іншої. Тому, якщо визначати ймовірність, наприклад, події  $B$ , то необхідно знати, чи напустила подія  $A$ .

*Умовною ймовірністю* називають ймовірність події  $B$ , обчислену в припущенні, що подія  $A$  вже напустила, таку ймовірність позначають  $P_A(B)$ .

### 1.4.4 . Теорема множення ймовірностей залежних подій

Якщо події  $A$  і  $B$  залежні, причому ймовірності  $P(A)$  і  $P_A(B)$  відомі. Ймовірність поєднання цих подій, тобто ймовірність того, що з'явиться подія  $A$  і подія  $B$  знаходиться за допомогою теореми.

Теорема. Ймовірність сумісної появи двох залежних подій дорівнює добутку ймовірності однієї з них на умовну ймовірність іншої, розраховану в припущенні, що перша подія вже напустила.

$$P(AB) = P(A) P_A(B)$$

Для підтвердження введемо позначення:

$n$  — число можливих елементарних результатів випробування, в яких подія  $A$  настає або не настає;  $n_1$  — число результатів, що сприяють події  $A$ ;  $m$  — число елементарних результатів випробування, в яких настає подія  $B$ , в припущенні, що подія  $A$  вже напустила, тобто результат, що сприяє появі події  $AB$ .

Ймовірність сумісної появи подій  $A$  і  $B$  рівна

$$P(AB) = \frac{m}{n} = \frac{n_1}{n} \frac{m}{n_1}$$

Враховуючи те, що

$$P(A) = \frac{n_1}{n}, \quad P_A(B) = \frac{m}{n_1}$$

отримаємо

$$P(AB) = P(A) P_A(B)$$

Ймовірність сумісної появи декількох залежних подій дорівнює добутку ймовірності однієї з них на умовну ймовірність всіх останніх, причому ймовірність кожної наступної події розраховується в припущенні, що всі попередні події вже з'явилися:

$$P(A_1 A_2 A_3 \dots A_n) = P(A_1) P_{A_1}(A_2) P_{A_1 A_2}(A_3) \dots P_{A_1 A_2 A_3 \dots A_{n-1}}(A_n).$$

Приклад. У технологічній схемі задіяно дві операції, на одній встановлено 3 апарати, а на іншій 7, упродовж доби сталося дві зупинки апаратів. Знайти ймовірність того, що першою сталась зупинка апаратів першої операції, а другою — зупинка апаратів другої операції.

Рішення. Ймовірність того, що першою була відмова апаратів першої операції така:

$$P(A) = \frac{3}{10} = 0.3$$

Ймовірність того, що відмова апаратів другої операції (подія В), буде такою:

$$P_A(B) = \frac{7}{10} = 0.7$$

Ймовірність появи події P(AB) по теоремі множення ймовірностей залежних подій визначиться:

$$P(AB) = P(A) P_A(B) = \frac{3}{10} \frac{7}{10} = \frac{21}{100} = 0.21$$

### Питання

1. Приведіть приклад незалежних подій і запишіть формулу, по якій виконується множення ймовірностей.
2. Приведіть приклад залежних подій і запишіть формулу, по якій виконується множення ймовірностей.
3. У чому полягає суть теореми про розрахунок ймовірності появи хоча би однієї події.
4. У чому полягає суть умовної ймовірності, приведіть приклад і запишіть її за допомогою умовних зображень.
5. Приведіть приклад протилежних подій і теорему для розрахунку їх суми.

### Завдання

1. Ймовірність отримання кількості продуктів в зміну рівна  $p = 0,9$ . Відпрацьовано 3 зміни. Знайти ймовірність того, що у всіх трьох змінах кількість отримана.

Відп.  $P = 0.729$ .

2. У технологічній схемі працює три зневоднюючі центрифуги, ймовірність відмови однієї центрифуги дорівнює 0.05. Знайти ймовірність того, що в даний момент не працює хоча би одна центрифуга.

Відп.  $P = 0.14$

3. Для визначення якості із загальної маси вугілля відібрали шість кусків, з них чотири вугільних і два порідних. При обробці навмання по черзі виймається два куски. Потрібно визначити ймовірність появи вугільного куска при другому вийманні, якщо при першому був породний кусок.

Відп.  $P = 4/5$ .

4. Для умов попереднього прикладу потрібно визначити ймовірність появи вугільного куска при другому вийманні, якщо при першому був вугільний.

Відп.  $P = 3/5$ .

5. У складі рядового вугілля з десяти вагонів три мають якість по зольності 39.0 %, останні 45.0 %. Знайти ймовірність того, що при розвантаженні перший вагон має якість 39.0 %, а другий 45.0 %.

Відп.  $P = 7/30$

## 1.5. Сумісне застосування формул суми і добутку ймовірностей для сумісних подій

### 1.5.1. Теорема складання ймовірностей сумісних подій

Дві події називають сумісними, якщо поява однієї з них не виключає появи іншої в одному і тому ж випробуванні.

Ймовірність таких подій обчислюється за допомогою теореми.

Теорема. Ймовірність появи хоч би однієї з двох сумісних подій дорівнює сумі ймовірності цих подій без ймовірності їх сумісної появи:

$$P(A+B) = P(A) + P(B) - P(AB)$$

Доказ. Оскільки події  $A$  і  $B$  по умові сумісні, то подія  $A + B$  наступить, якщо наступить одна з наступних трьох несумісних подій:

$$A\bar{B}, \bar{A}B \text{ або } AB.$$

Використавши теорему складання ймовірностей несумісних подій, одержимо рівняння:

$$P(A+B) = P(A\bar{B}) + P(\bar{A}B) + P(AB) \quad (1)$$

Подія  $A$  відбудеться, якщо наступить одна з двох несумісних подій:  $(A\bar{B})$ ,  $(AB)$ .

По теоремі складання ймовірності несумісних подій маємо:

$$P(A) = P(A\bar{B}) + P(AB) \quad (2)$$

Звідси:

$$P(A\bar{B}) = P(A) - P(AB) \quad (3)$$

Аналогічно матимемо:

$$P(\bar{A}B) = P(B) - P(AB) \quad (4)$$

Підставивши знайдені значення з формул (3) і (4) в (1), остаточно отримаємо:

$$P(A+B) = P(A) + P(B) - P(A) \cdot P(B)$$

Події  $A$  і  $B$  можуть бути як незалежними, так і залежними, для залежних подій :

$$P(A+B) = P(A) + P(B) - P(A) \cdot P_A(B)$$

Якщо події  $A$  і  $B$  несумісні, то їх сумісна поява є неможливою і, отже,  $P(AB) = 0$  і в цьому випадку формула має вигляд  $P(A+B) = P(A) + P(B)$ .

Приклад. Ймовірність отримання якості і кількості продуктів збагачення в зміну на ідентичних технологічних лініях дорівнює:  $p_1 = 0,9$ ;  $p_2 = 0,95$ . Знайти ймовірність отримання якості за зміну хоч би на одній з ліній.

Рішення. Ймовірність отримання якості і кількості на першій  $A$  і другій  $B$  лініях події незалежні.

Ймовірність події  $AB$  (якість і кількість отримана на обох лініях):

$$P(AB) = P(A) \cdot P(B) = 0,9 \cdot 0,95 = 0,85$$

Ймовірність отримання якості і кількості продукту хоч би на одній з ліній:

$$P(A+B) = P(A) + P(B) - P(A) \cdot P(B) = (0,9 + 0,95) - 0,85 = 0,995$$

Оскільки події  $A$  і  $B$  незалежні, то при рішенні можна було скористатися формулою:

$$P = 1 - q$$

Знайдемо ймовірність подій, протилежних подіям  $A$  і  $B$ :

$$q_A = 1 - 0,9 = 0,1, \quad q_B = 1 - 0,95 = 0,05$$

Ймовірність того, що одночасно хоч би на одній лінії буде отримано якість і кількість, дорівнює:

$$P = 1 - q_A \cdot q_B = 1 - 0.005 = 0.995$$

### 1.5.2. Формула повної ймовірності

В практичній роботі виникають задачі, в яких потрібно знайти ймовірність події А за умови появи однієї з несумісних подій  $B_1, B_2, \dots, B_n$ , які утворюють повну групу. Якщо відомі ймовірності цих подій і умовна ймовірність  $P_{B_1}(A)$ ,  $P_{B_2}(A)$ , ...,  $P_{B_n}(A)$  події А, то в цьому випадку необхідно скористатися формулою повної ймовірності.

Теорема. Ймовірність події А, яка може наступити лише за умови появи однієї з несумісних подій  $B_1, B_2, \dots, B_n$  створюючих повну групу, дорівнює сумі добутків ймовірності кожної з цих подій на відповідну умовну ймовірність події А

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \dots + P(B_n) \cdot P_{B_n}(A)$$

Цю формулу називають «формулою повної ймовірності».

Доказ. По умові подія А може наступити, якщо наступить одна з несумісних подій  $B_1, B_2, \dots, B_n$ . Тобто, поява події А означає здійснення однієї, байдуже якої, з несумісних подій  $B_1A, B_2A, \dots$ . Для розрахунку ймовірності події А скористаємося теоремою складання, отримаємо:

$$P(A) = P(B_1A) + P(B_2A) + \dots + P(B_nA)$$

Кожен з додатків обчислимо за формулою множення ймовірності залежних подій:

$$P(B_1A) = P(B_1) \cdot P_{B_1}(A);$$

$$P(B_2A) = P(B_2) \cdot P_{B_2}(A);$$

$$P(B_nA) = P(B_n) \cdot P_{B_n}(A).$$

Підставивши праві частини цієї рівності в початкове рівняння, отримаємо формулу повної ймовірності:

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \dots + P(B_n) \cdot P_{B_n}(A)$$

Приклад. Є два набори лабораторного хімічного посуду. Ймовірність того, що вироби першого набору стандартні, дорівнює 0,7, а другого — 0,9. Знайти ймовірність того, що узятий навмання посуд з набору — стандартний.

Рішення. Позначимо через А подію — взято стандартний посуд.

Посуд можна брати або з першого набору - подія  $B_1$ , або з другого - подія  $B_2$ .

Ймовірність того, що посуд буде вийнятий з першого набору:

$$P(B_1) = \frac{1}{2}$$

Ймовірність того, що деталь буде вийнята з другого набору:

$$P(B_2) = \frac{1}{2}$$

Умовна ймовірність того, що з першого набору взяли стандартний посуд:

$$P_{B_1}(A) = 0.7$$

Умовна ймовірність того, що з другого набору витягнули стандартний посуд:

$$P_{B_2}(A) = 0.9$$

Формула для визначення ймовірності появи стандартного посуду буде такою:

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) = 0.5 \cdot 0.7 + 0.5 \cdot 0.9 = 0.8$$

### 1.5.3. Ймовірність гіпотез. Формули Бейєса

Якщо потрібно визначити ймовірність події  $A$  за умови появи однієї з несумісних подій  $B_1, B_2, \dots, B_n$ , створюючих повну групу і, якщо заздалегідь невідомо, яка з цих подій наступить, їх називають гіпотезами. Ймовірність появи події  $A$  визначається по формулі повної ймовірності:

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \dots + P(B_n) \cdot P_{B_n}(A)$$

Припустимо, що проведене випробування, в результаті якого з'явилася подія  $A$ . Постає питання, як змінилися у зв'язку з тим, що подія  $A$  вже наступила, умовні ймовірності. Іншими словами, потрібно знайти умовну ймовірність:  $P_A(B_1), P_A(B_2), P_A(B_n)$ .

Спочатку знайдемо умовну ймовірність  $P_A(B_1)$ , скориставшись формулою множення:

$$P_A(B_1) = P(A) \cdot P_A(B_1) = P(B_1) \cdot P_{B_1}(A)$$

Звідси:

$$P_A(B_1) = \frac{P(B_1) \cdot P_{B_1}(A)}{P(A)}$$

Замінивши тут  $P(A)$ , яке знайдене за формулою повної ймовірності, отримаємо:

$$P_A(B_1) = \frac{P(B_1) \cdot P_{B_1}(A)}{P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + \dots + P(B_n) \cdot P_{B_n}(A)}$$

Аналогічно виводяться формули визначення умовних ймовірностей решти гіпотез.

Отримані формули називають по імені англійського математика Бейєса. Формули дозволяють переоцінити умовну ймовірність після того, як стає відомим результат випробування, внаслідок якого з'явилася подія.

Приклад. Визначення якості товарної продукції здійснюється з проб двома лаборантами. Ймовірність того, що проба потрапляє до першого лаборанта рівна 0,7, а до другого — 0,3. Ймовірність того, що аналіз буде проведений правильно і якість відповідатиме нормі першим лаборантом рівна 0,96, а другим — 0,98. Виконаний аналіз був проведений правильно, а якість виявилася потрібною (нормативною). Знайти ймовірність того, що аналіз проводив перший лаборант.

Рішення. Позначимо подію  $A$ , що полягає в тому, що якість відповідає нормі. В цьому випадку правомочні два припущення:

- 1) аналіз проводив перший лаборант (гіпотеза  $B_1$ );

2) аналіз проводив другий лаборант (гіпотеза  $B_2$ ).

Шукану ймовірність того, що аналіз проводив перший лаборант, знайдемо:

$$P_A(B_1) = \frac{P(B_1) \cdot P_{B_1}(A)}{P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A)}$$
$$P_A(B_1) = \frac{0.7 \times 0.96}{0.7 \times 0.96 + 0.3 \times 0.98} = 0.6$$

Відносно умови задачі маємо:

$P(B_1) = 0,7$  - ймовірність того, що аналіз проводить перший лаборант;

$P(B_2) = 0,3$  - ймовірність того, що аналіз проводить другий лаборант;

$P_{B_1}(A) = 0,96$  - ймовірність того, що якість визначення якості відповідає нормі визначення першим лаборантом;

$P_{B_2}(A) = 0,98$  - ймовірність того, що якість визначення якості відповідає нормі визначення другим лаборантом.

Підставивши значення в попередню формулу, отримаємо шукану ймовірність

$$P_A(B_1) = 0.66$$

Як видно, до випробування ймовірність гіпотези  $P_A(B_1)$  дорівнювала 0,7, а після того, як став відомий результат випробування, ймовірність цієї гіпотези змінилася і стала рівною 0,66. Таким чином, використання формули Бейеса дозволило переоцінити ймовірність даної гіпотези.

### Питання

1. Приведіть суть теореми складання ймовірностей сумісних подій.
2. У яких випадках слід застосовувати формулу повної ймовірності.
3. Призначення формул Бейеса, приведіть конкретну ситуацію їх застосування.

### Завдання

1. Два студенти склали іспит. Ймовірність успішної здачі першим студентом дорівнює 0,7, а другим - 0,6. Знайти ймовірність того, що хоч би один із студентів склав іспит.

Відп. 0,88.

2. Ймовірності отримання якості на технологічній лінії по змінах відповідно дорівнюють 0,8; 0,85; 0,9; 0,99. Знайти ймовірність отримання якості на технологічній лінії в якій завгодно зміні.

Відп. 0,875

3. На вуглезбагачувальній фабриці працюють дві ідентичні технологічні лінії для отримання товарної продукції. Ймовірність отримання якості для першої лінії - 0.91, для другої – 0.97. Потрібно знайти ймовірність того, що отриманий товарний продукт байдуже на якій лінії відповідає вимогам за якістю.

Відп.  $P = 0.94$ .

4. Технологічні лінії спрацювали, якість отримана. Потрібно визначити, як у зв'язку з цим зміниться ймовірність отримання якості технологічними секціями, в розрахунках прийняти умови попереднього завдання.

## 1.6. Повторення випробувань і визначення ймовірностей

### 1.6.1. Формула Бернуллі

Якщо проводиться декілька дослідів (випробувань), причому ймовірність деякої події  $A$ , тобто результат в кожному з дослідів не залежить від результатів інших випробувань, то такі випробування називають незалежними щодо події  $A$ .

У різних незалежних випробуваннях подія  $A$  може мати або різну, або одну і ту ж ймовірність. Розглянемо незалежні випробування, в яких подія  $A$  має одну і ту ж ймовірність.

Хай проводиться  $n$  незалежних випробувань, в кожному з яких подія  $A$  може з'явитися або не з'явиться. Вважатимемо, що ймовірність події  $A$  в кожному випробуванні одна і та ж і дорівнює  $p$ . Отже, ймовірність ненастання події  $\bar{A}$  в кожному випробуванні також постійна і дорівнює  $q = 1-p$ .

Як знайти ймовірність того, що при  $n$  випробуваннях подія  $A$  наступить рівно  $k$  раз і не наступить  $n-k$  разів.

Без врахування послідовності появи події буде наступне. Наприклад, якщо мова йде про появу події  $A$  двічі в трьох випробуваннях, то можливі наступні складні події:  $AA\bar{A}$ ,  $A\bar{A}A$ , і  $\bar{A}AA$ . Вираз  $AA\bar{A}$  означає, що в першому і другому випробуваннях подія  $A$  наступила, а в третьому випробуванні вона не з'явилася, тобто наступила протилежна подія  $\bar{A}$ ; відповідне значення мають і інші записи.

Таку ймовірність позначають  $P_n(k)$  і читають так: «ймовірність того, що в  $n$  випробуваннях подія з'явиться рівно  $k$  раз і, отже, не наступить  $n-k$  раз». Згідно теореми множення ймовірності незалежних подій:

$$P(A\bar{A}) = p^k q^{n-k}$$

Число складних подій  $A\bar{A}$  може бути таким, скільки можна скласти поєднань з  $n$  елементів по  $k$  елементів, тобто  $C_n^k$ . Оскільки ці складні події несумісні, то по теоремі складання ймовірності несумісних подій невідома ймовірність дорівнює сумі ймовірностей всіх можливих складних подій. Оскільки ймовірність всіх цих складних подій однакова, то ймовірність (появи  $k$  раз події  $A$  в  $n$  випробуваннях) дорівнює ймовірності однієї складної події, помноженій на їх число:

$$P_n(k) = C_n^k p^k q^{n-k}$$

або

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Отриманий вираз називають формулою Бернуллі.

Приклад. Ймовірність того, що витрата електроенергії на технологічній секції в зміну не перевищить встановленої норми, дорівнює  $p = 0,7$ . Знайти



ймовірність того, що в подальших 5 змінах витрата електроенергії протягом 4 змін не перевищить норми.

Рішення. Ймовірність нормативної витрати електроенергії для кожних з п'яти змін постійна і дорівнює  $p = 0,7$ . Отже, ймовірність перевитрати електроенергії в кожну добу також постійна і дорівнює  $q = 1-p = 1-0,7 = 0,3$ .

Ймовірність того, що в подальших 5 змінах витрата електроенергії протягом 4 змін не перевищить, визначиться:

$$P_5(4) = \frac{5!}{4!(5-4)!} p^4 q^{5-4} = \frac{5 \cdot 4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3 \cdot 4} 0.7^4 0.3^1 = 0.36$$

### 1.6.2. Локальна теорема Лапласа

Користуватися формулою Бернуллі при великих значеннях  $n$  і  $k$  важко, оскільки при цьому потрібно виконувати дії над великими числами. В цьому випадку слід користуватися формулою Лапласа.

Теорема Лапласа. Якщо ймовірність  $p$  появи події  $A$  в кожному випробуванні постійна і відмінна від нуля і одиниці., то ймовірність  $P_n(k)$  того, що подія  $A$  з'явиться в  $n$  випробуваннях рівно  $k$  разів приблизно дорівнює тим точніше, чим більше  $n$  значення функції:

$$P_n(k) = \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}} \varphi(x),$$

де

$$x = \frac{k - np}{\sqrt{npq}}$$

Значення функції  $\varphi(x)$  відповідні позитивним значенням аргументу  $x$  знаходять за допомогою таблиці (Додаток 1). Для негативних значень аргументу користуються тими ж таблицями, оскільки функція парна, тобто  $\varphi(x) = -\varphi(x)$ .

Приклад 1. Знайти ймовірність того, що подія  $A$  відбудеться рівно 10 разів в 40 випробуваннях, якщо ймовірність появи цієї події в кожному випробуванні дорівнює 0,2.

Рішення. По умові  $n = 40$ ;  $k = 10$ ;  $p = 0,2$ ;  $q = 0,8$ . Знайдемо значення  $x$ :

$$x = \frac{k - np}{\sqrt{npq}} = \frac{10 - 40 \cdot 0.2}{\sqrt{40 \cdot 0.2 \cdot 0.8}} = 2 / 2.52 = 0.79$$

По таблиці (додаток 1) знаходимо  $\varphi(x) = 0.29$

Шукана ймовірність:

$$P_{40}(10) = \frac{1}{\sqrt{40 \cdot 0.2 \cdot 0.8}} 0.29 = 0.115$$

### 1.6.3. Інтегральна теорема Лапласа

При проведенні випробувань, що повторюються, в кожному з яких ймовірність появи події  $A$  постійна і знаходиться в межах ( $0 < p < 1$ ), потрібно визначати ймовірність  $P_n(k_1, k_2)$  того, що подія  $A$  з'явиться не менше  $k_1$  і не більш  $k_2$  разів. Вирішення таких завдань здійснюється за допомогою формули, отриманої по теоремі Лапласа.

Теорема. Якщо ймовірність настання події в кожному випробуванні постійна і відмінна від нуля і одиниці, то ймовірність того, що подія з'явиться у  $n$  випробуваннях від  $k_1$  до  $k_2$  разів приблизно дорівнює виразу:

$$P_n(k_1, k_2) \cong \frac{1}{\sqrt{2\pi}} \int_{x^1}^{x^2} e^{-\frac{z^2}{2}} dz \quad (*)$$

де

$$x^1 = \frac{k_1 - np}{\sqrt{npq}} \quad \text{а} \quad x^2 = \frac{k_2 - np}{\sqrt{npq}}$$

При вирішенні задач такого типу користуються спеціальною таблицею (Додаток 2) функції  $\Phi(x)$ , тому вираз (\*) перетворений до вигляду:

$$P_n(k_1, k_2) \cong \Phi(x^2) - \Phi(x^1)$$

Функція  $\Phi(x)$  не парна, тобто  $\Phi(-x) \neq -\Phi(x)$ . У таблиці приведені значення для  $x = 5$ , для великих значень ця величина практично не змінюється.

Приклад. Ймовірність відхилення в технологічному процесі якості відходів від встановленої норми дорівнює 0.2. Визначити ймовірність того, що протягом 90 змін якості відходів не буде досягнута - від 5 до 15 змін.

Рішення. Введемо позначення:  $p = 0.2$  - ймовірності не досягнення якості відходів,  $q = 0.8$  - ймовірність протилежної події, до - 5, до - 15,  $n = 90$ .

По інтегральній теоремі Лапласа:

$$P_{90}(5, 15) \cong \Phi(x^2) - \Phi(x^1)$$

$$x^1 = \frac{5 - 90 \times 0.2}{\sqrt{90 \times 0.2 \times 0.8}} = -3.4 \quad x^2 = \frac{15 - 90 \times 0.2}{\sqrt{90 \times 0.2 \times 0.8}} = -0.79$$

Підставимо значення у формулу:

$$P_{90}(5, 15) \cong \Phi(-0.79) - \Phi(-3.4)$$

Знайдемо значення функції:  $\Phi(-0.79) = 0.2825$ ,  $\Phi(-3.4) = 0.499$ .

Шукана ймовірність :

$$P_{90}(5, 15) = -0.2825 + 0.499 = 0.216$$

#### 1.6.4. Ймовірність відхилення відносної частоти від постійної ймовірності в незалежних випробуваннях

Проводиться  $n$  незалежних випробувань, в кожному з яких ймовірність появи події  $A$  постійна, а її значення належать інтервалу ( $0 < p < 1$ ). Потрібно знати ймовірність того, що відхилення відносної частоти  $\frac{m}{n}$  від постійної ймовірності  $p$  по абсолютній величині не перевищує заданого числа  $\varepsilon > 0$ . Другими словами, знайдемо ймовірність здійснення нерівності :

$$P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right)$$

Замінімо вираз в дужках рівносильною йому нерівністю:

$$-\varepsilon \leq \frac{m}{n} - p \leq \varepsilon \quad \text{або} \quad -\varepsilon \leq \frac{m - np}{n} \leq \varepsilon$$

Помножимо останню нерівність на позитивний множник

$\sqrt{\frac{n}{pq}}$ , отримаємо нерівність, рівносильну початковій:

$$-\varepsilon \sqrt{\frac{n}{pq}} \leq \frac{m - np}{\sqrt{npq}} \leq \varepsilon \sqrt{\frac{n}{pq}}$$

Якщо прирівняти:

$$x' = -\varepsilon \sqrt{\frac{n}{pq}} \quad \text{а,} \quad x'' = \varepsilon \sqrt{\frac{n}{pq}}, \quad \text{тоді}$$

$$P\left(-\varepsilon \sqrt{\frac{n}{pq}} \leq \frac{m - np}{\sqrt{npq}} \leq \varepsilon \sqrt{\frac{n}{pq}}\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon \sqrt{\frac{n}{pq}}}^{\varepsilon \sqrt{\frac{n}{pq}}} e^{-\frac{z^2}{2}} dz = 2\Phi\left(\varepsilon \sqrt{\frac{n}{pq}}\right)$$

У даному виразі замінимо нерівності, заключені в дужках, рівносильними їм початковими нерівностями, отримаємо:

$$P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) \cong 2\Phi\left(\varepsilon \sqrt{\frac{n}{pq}}\right)$$

Таким чином, ймовірність здійснення нерівності приблизно дорівнює значенню подвоєної функції Лапласа  $2\Phi(x)$ .

Приклад I. Ймовірність того, що технологічний апарат має вади, дорівнює  $p = 0,1$ . Знайти ймовірність того, що серед випадково відібраних 100 апаратів відносна частота появи нестандартного відхилиться від ймовірності  $p = 0,1$  по абсолютній величині не більш ніж на 0,03.

Рішення. По умові  $n = 100$ ;  $p = 0,1$ ;  $q = 0,9$ ;  $\varepsilon = 0.03$

Запишемо формулу, підставивши в неї значення:

$$P\left(\left|\frac{m}{100} - 0.1\right| \leq 0.03\right) \cong 2\Phi\left(0.03 \sqrt{\frac{100}{0.1 \cdot 0.9}}\right)$$

Після обчислень отримаємо:

$$P\left(\left|\frac{m}{100} - 0.1\right| \leq 0.03\right) \cong 2\Phi(1.01)$$

По таблиці додатку знаходимо:  $\Phi(1.01) = 0.3438$ , тоді  $2\Phi(1.01) = 0.6876$   
Тобто, шукана ймовірність приблизно дорівнює 0,6876.

### Питання

1. Приведіть приклад повторення випробувань для реального процесу.
2. Вкажіть особливість формули Бернуллі.
3. У чому особливість локальної теореми Лапласа.
4. Приведіть відмінність інтегральної теореми Лапласа від локальної.
5. Яким чином в реальних умовах можна використовувати особливість відхилення відносної частоти від постійної ймовірності.

### Завдання

1. У відділенні встановлений насос для перекачування шламів в сховище. Ймовірність відмови насоса дорівнює 0.1. Потрібно визначити ймовірність того, що протягом чотирьох змін насос вийде з ладу.

Відпідь. 0.0036.

2. Тривалість роботи піскового насосу складає 400 годин. Ймовірність його відмови дорівнює 0.1. Потрібно визначити ймовірність того, що час відмови складе 10 годин.

Відповідь. 0.00016.

3. Для умов попереднього прикладу визначити ймовірність того, що час відмови складе від 10 до 20 годин.

Відповідь. 0.000001.

### **1.7. Випадкові величини і їх особливості**

Якщо аналізувати роботу збагачувальної фабрики з приводу отримання кінцевого продукту заданої якості за дві зміни, то можливі наступні результати: у обох випадках якість не виконана, якість виконана в одну зміну, якість виконана в обох змінах. В даному випадку результати представляються числами: 0, 1, 2 і їх називають випадковими величинами. Всі можливі результати випробування називають повною групою.

Випадковими називають величини, які в результаті випробувань обов'язково набувають конкретного значення, проте яким це значення буде, ми не знаємо і заздалегідь визначити не можемо унаслідок дії деяких невідомих нам причин. Наприклад, для наведеного прикладу можуть з'явитись числа 0, 1 або 2, а яке саме - невідомо.

#### **1.7.1. Дискретні і безперервні випадкові величини**

В практичному використанні випадкові величини ділять на дискретні і безперервні. Дискретними називають величини, які приймають окремі через свою специфіку значення, наприклад, число студентів, які можуть з'явитися в зал для слухання лекції. У даному прикладі число студентів є величина випадкова внаслідок того, що ми заздалегідь визначити не можемо, яке число їх буде. Ми знаємо, що це число має бути цілим, тобто вся сукупність представляється дискретними цілими числами.

В інших прикладах числа можуть бути безперервними. Безперервними називають величини, які можуть набувати будь-яких значень на кінцевому проміжку або нескінченному ряду.

Температура навколишнього середовища в деякому середовищі. Якщо її вимірювати безперервно, то отримаємо деяку безперервну сукупність на числовій осі. Температуру можна вимірювати переривчасто, через певні інтервали часу, тоді температуру можна представити у вигляді дискретних чисел безперервного типу.

#### **1.7.2. Задання випадкових величин**

З попереднього виходить, що однією з числових мір для випадкових величин, тобто зявиться подія або ж ні, є ймовірність, а другою- це конкретні її значення.

Для того, щоб задати випадкову величину, необхідно вказати значення, які випадкова величина може приймати і вказати ймовірність цих значень, тобто зробити словесний опис. У практичній роботі це не завжди вдається. В зв'язку з цим відпрацьовані прості і достатньо надійні способи задання випадкових величин - це табличний закон розподілу випадкової величини, інтегральна і диференціальна функції розподілу.

Законом розподілу випадкової величини називають відповідність між можливими її значеннями і їх ймовірностями. Закон розподілу випадкової величини може бути заданий табличний, аналітично і у вигляді графіка.

*Таблична форма задання закону розподілу дискретної випадкової величини.*

У загальному вигляді табличний закон розподілу виглядає таким чином, див. табл. 1:

				Таблиця 1
X	x <sub>1</sub>	x <sub>2</sub>	...	x <sub>n</sub>
P	p <sub>1</sub>	p <sub>2</sub>	...	p <sub>n</sub>

де X і P- умовні позначення даної випадкової величини і відповідні її ймовірності; x<sub>i</sub> і p<sub>i</sub> – конкретні значення випадкової величини і відповідна їм ймовірність. У кожному конкретному випадку закон розподілу повинен представляти повну групу подій, тому для нього  $\sum_{i=1}^n p_i = 1.0$

Ймовірності конкретних значень випадкових величин мають деяку специфічність стосовно їх визначення, тобто які для їх розрахунку використовують математичні вирази і залежно від цього розрізняють закони розподілу випадкових величин, наприклад:

- біноміальний розподіл;
- розподіл Пуассона.

Біноміальним називають закон розподілу, в якому ймовірність визначається за допомогою формули Бернуллі:

$$P_n(k) = C_n^k p^k q^{n-k},$$

де k – чисельне значення кількості випадкової величини (число можливих результатів), n – число випробувань, C – число поєднань з n елементів по k, p – ймовірність появи даної події в одному досліді, q – ймовірність не появи даної події.

Приклад. Намічено, що технологічна секція збагачення працюватиме протягом двох змін з отриманням продукції заданої якості. Потрібно скласти закон розподілу випадкової величини – отримання заданої якості, якщо відомо, що ймовірність його досягнення на секції дорівнює 0.8.

Введемо позначення A - якість отриманої продукції в зміну. Визначимо можливі значення даної величини. В результаті роботи може статись так, що у обох випадках якість не виконана, якість виконана в одну зміну, якість виконана в обох змінах.

Формалізовано це виглядає так:

А 0 1 2

Знайдемо ймовірність значень випадкової величини, використовуючи формулу Бернуллі.

Знайдемо ймовірність протилежної події, скориставшись виразом  $p + q = 1$ , а також ймовірності величин:

$$q = 1 - p = 1 - 0.8 = 0.2$$

$$P_2(0) = C_2^0 p^0 q^{2-0} = 1 \cdot 1 \cdot 0.2^2 = 0.04$$

$$P_2(1) = C_2^1 p^1 q^{2-1} = 2 \cdot 0.8 \cdot 0.2^1 = 0.32$$

$$P_2(2) = C_2^2 p^2 q^{2-2} = 1 \cdot 0.8^2 \cdot 0.2^0 = 0.64$$

Тобто ймовірності появи вказаних результатів:

Р 0.04 0.32 0.64

У табличному вигляді закон розподілу, табл. 2

Таблиця 2

А	0	1	2
Р	0.04	0.32	0.64

Перевірка правильності обчислень. Оскільки всі значення випадкової величини складають повну групу, то сума їх ймовірностей повинна дорівнювати одиниці. Знайдемо цю суму:

$$0.04 + 0.032 + 0.64 = 1.0$$

Розподілом Пуассона називають закон розподілу, в якому ймовірність обчислюється за допомогою формули:

$$P_n(k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

де  $\lambda = np$ ,  $n$  – число випробувань,  $p$  – ймовірність появи випадкової величини в одному випробуванні. Характерним для цього закону розподілу є те, що  $\lambda$  є постійною.

Дану формулу застосовують в тому разі, якщо  $n$  велике значення, а  $p$  – мале.

Ще однією з форм завдання дискретної випадкової величини є інтегральна функція розподілу.

Інтегральною функцією розподілу дискретної випадкової величини називають функцію, яка для кожного значення випадкової величини визначає ймовірність того, що випадкова величина набуде значення меншого деякого наперед заданого. Виходячи з даного визначення, вираз для обчислення інтегральної функції такий:

$$F(X) = P(X < x_s)$$

Крім того, що інтегральна функція є однією з характеристик випадкової величини, з її допомогою можна задати випадкову величину аналітично в явному вигляді.

Приклад. Випадкова величина задана табличним законом розподілу – табл. 3. Потрібно написати інтегральну функцію розподілу даної величини в явному вигляді.

Таблиця 3

X	1.0	2.0	3.0
P	0.2	0.3	0.5

При написанні закону розподілу використовується визначення інтегральної функції.

При  $X \leq x_1$ , тобто  $0.0 < X \leq 1.0$ ., ймовірність для випадкової величини рівна (див. таблицю)  $P = 0.0$

При  $x_1 < X \leq x_2$ , тобто при  $1.0 < X \leq 2.0$  ймовірність для випадкової величини дорівнює (див. таблицю) 0.2

При  $x_2 < X \leq x_3$ , тобто  $2.0 < X \leq 3.0$  ймовірність для випадкової величини дорівнює (див. таблицю) 0.2 або 0.3, при логічному «або» ймовірність складається, тому  $P = 0.2 + 0.3 = 0.5$ .

При  $x_3 < X$ , тобто  $3.0 \leq X$  ймовірність для випадкової величини дорівнює (див. табл.3) 0.2 або 0.3 або 0.5, при логічному «або» ймовірність складається, тому  $P = 0.2 + 0.3 + 0.5 = 1.0$ .

Остаточний вираз інтегральній функції для даної випадкової величини :

$$P(X) = \begin{cases} 0.0 & \text{при } X \leq 1.0 \\ 0.2 & \text{при } 1.0 < X \leq 2.0 \\ 0.5 & \text{при } 2.0 < X \leq 3.0 \\ 1.0 & \text{при } 3.0 \leq X \end{cases}$$

Використавши одержану формулу, можна побудувати графічне зображення розподілу випадкової величини – рис. 1.

Рис. 1. Графічне зображення інтегральної функції розподілення випадкової величини

### 1.7.3. Числові характеристики дискретних випадкових величин

Іноді представляти випадкові величини законом розподілу не раціонально або не можливо, в цьому випадку застосовують деякі узагальнювальні величини, які їх характеризують. Ці величини називають числовими характеристиками.

До числових характеристик дискретних випадкових величин відносять:

- математичне очікування;
- дисперсію;
- середньоквадратичне відхилення.

*Математичне очікування дискретної випадкової величини* - це сума добутку всіх почергових її значень на їх відповідну ймовірність. Визначається математичне очікування такою формулою:

$$M(X) = \sum_{i=1}^n x_i p_i$$

де  $M(X)$  - математичне очікування випадкової величини  $X$ ,  $x_i$  – конкретні значення, які приймала випадкова величина  $X$ ,  $p_i$  – ймовірність кожного відповідного значення. Математичне очікування є узагальнювальною характеристикою всіх значень випадкової величини, проте, в певних випадках ця характеристика її характеризує не повною мірою. Наприклад, дві випадкові величини задано табличними законами розподілу – табл. 4 і табл. 5.

Таблиця 4

X	- 0.01	0	0.01
P	0.33	0.33	0.33

Таблиця 5

Y	- 100	0	100
P	0.33	0.33	0.33

Знайдемо математичні очікування цих величин:

$$M(X) = - 0.01 \cdot 0.33 + 0 \cdot 0.33 + 0.01 \cdot 0.33 = 0;$$

$$M(Y) = - 100 \cdot 0.33 + 0 \cdot 0.33 + 100 \cdot 0.33 = 0.$$

Математичні очікування цих величин однакові, а їх конкретні значення дуже розрізняються, причому якщо для  $X$  абсолютна величина між найбільшим і найменшим значеннями складає 0.02, то для  $Y$  ця відмінність дорівнює 200.0. Щоб врахувати таку розбіжність використовують іншу характеристику – дисперсію.

*Дисперсією дискретної випадкової величини* називають математичне очікування квадрата відхилення значень випадкової величини від її математичного очікування.

Визначається дисперсія таким виразом:

$$D(X) = M(X_i - M(X))^2$$

Для обчислення дисперсії зручно користуватися формулою, отриманою в результаті перетворення попередньої.

$$D(X) = M(X^2) - (M(X))^2$$

Зручнішою характеристикою відхилення є середньоквадратичне відхилення.

*Середнім квадратичним відхиленням випадкової величини* називають корінь квадратний з дисперсії.

Згідно наведеного вище визначення вираз для обчислення середнього квадратичного відхилення :

$$\sigma(x) = \sqrt{D(X)}$$



Якщо дисперсія має розмірність одиниці вимірювання випадкової величини в квадраті, то середньоквадратичне відхилення виражається такою ж розмірністю, як випадкова величина.

### 1.8. Способи задання і числові характеристики безперервних випадкових величин

Найзручніше задавати безперервну випадкову величину інтегральною функцією розподілу.

Застосовують декілька видів розподілу, проте найбільш загальним і часто вживаним є нормальний закон розподілу.

Інтегральна функція нормального закону розподілу :

$$F(X) = \frac{1}{\sqrt{2\pi} \sigma} \int e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

Для безперервної випадкової величини представляється можливим визначити диференціальну функцію розподілу, якщо відома інтегральна функція розподілу. Для наведеної вище інтегральної функції диференціальна така:

$$f(X) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Як для дискретної випадкової величини, так і для безперервної обчислюються числові характеристики:

- математичне очікування
- дисперсія
- середньоквадратичне відхилення.

Математичне очікування визначається виразом:

$$M(X) = \int_{-\infty}^{\infty} xf(x)dx .$$

Дисперсія розраховується такою формулою:

$$D(X) = \int_{-\infty}^{\infty} (x - M(x))^2 f(x)dx .$$

Середньоквадратичне відхилення, як і для дискретної випадкової величини, дорівнює:

$$\sigma(x) = \sqrt{D(x)}$$

### 1.9. Система випадкових величин

Розглянуте вище відносилось до конкретних випадкових величин, які називають одновимірними, їх значення визначаються одним числом. Окрім одновимірних існують величини, значення яких визначається двома, трьома і так далі числами, такі випадкові величини називають двох-, трьох- і так далі мірними.

#### 1.9.1. Способи задання двовимірних випадкових величин

Як і одновимірні, двовимірні випадкові величини задають табличним законом розподілу, при цьому таблиця має два входи. Один з входів

представляє одну випадкову величину – це перший рядок, а другий з входів - другу – це перший стовпець. У середині таблиці на перетині значень складових стовпців і стрічок записується ймовірність, з якою таке поєднання зустрічається. Приклад таблиці для двовимірної випадкової величини.

Таблиця 6

Y \ X	x <sub>1</sub>	X <sub>2</sub>	...	X <sub>m</sub>	p <sub>y</sub>
y <sub>1</sub>	p(x <sub>1</sub> y <sub>1</sub> )	p(x <sub>2</sub> y <sub>1</sub> )		p(x <sub>m</sub> y <sub>1</sub> )	p <sub>y1</sub>
y <sub>2</sub>	p(x <sub>1</sub> y <sub>2</sub> )	p(x <sub>2</sub> y <sub>2</sub> )		p(x <sub>m</sub> y <sub>2</sub> )	p <sub>y2</sub>
...					
y <sub>n</sub>	p(x <sub>1</sub> y <sub>n</sub> )	p(x <sub>2</sub> y <sub>n</sub> )		p(x <sub>m</sub> y <sub>n</sub> )	p <sub>ym</sub>
p <sub>x</sub>	p <sub>x1</sub>	p <sub>x2</sub>		p <sub>xn</sub>	p <sub>xn</sub> = p <sub>ym</sub>

Спільно X і Y утворюють систему, а окремо X і Y називають складовими і вони є одновимірними.

Дискретну двовимірну випадкову величину можна задавати інтегральною функцією розподілу, безперервну двовимірну величину - диференціальною функцією розподілу. Знаючи диференціальну функцію розподілу безперервної випадкової величини, можна знайти її інтегральну функцію. Для дискретної двовимірної випадкової величини визначаються безумовні і умовні закони розподілу.

У таблиці, якою представляється двовимірною випадковою величиною, *безумовні закони розподілу складових* записуються при заповненні таблиці.

Якщо подивитися на таблицю 6, то безумовний закон розподілу складової X - це перший стовпець, в якому записані її чисельні значення і останній стовпець, що містить ймовірність цих значень, а складової Y, це перший рядок, в якому записані її чисельні значення і останній рядок, що містить ймовірність цих значень.

Якщо в таблиці за будь-якими причинами відсутні останній рядок і стовпець, то їх дописують, а значення чисел, які в них слід записати, отримують підсумовуванням ймовірності відповідно по стовпцях або рядках.

*Умовні закони розподілу складових системи дискретних випадкових величин.*

Для одновимірних випадкових величин, якщо дві події A і B залежні, то їх умовна ймовірність відрізняється від безумовної.

Наприклад, якщо подія A напустила, то умовна ймовірність події B обчислюється за формулою:

$$P_A(B) = \frac{P(AB)}{P(A)}$$

Для того, щоб охарактеризувати залежність між складовими дискретної двовимірної випадкової величини використовується поняття умовного розподілу, який представляє значення складової і відповідну умовну ймовірність. Значення складових - це числа, які їх представляють на одному з входів таблиці.

Умовна ймовірність дискретної двовимірної випадкової величини в загальному вигляді:

$$P(x_i / y_j) \text{ або } P(y_j / x_i)$$

де  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, m$ .

Такий вигляд умовної ймовірності вказує на те, що перша в дужках складова набула конкретного значення, а для цих значень обчислюється умовна ймовірність другої складової.

Умовним законом розподілу складової, наприклад,  $X$ , якщо  $Y$  набула значення  $y_i$ , називають відповідність між значеннями складової  $X$  і їх умовною ймовірністю. У загальному вигляді умовний закон розподілу складової такий:

$$\begin{array}{cccc} X & 1 & 2 & \dots & 4 \\ P(x_i / y_1) & P(x_1 / y_1) & P(x_2 / y_1) & \dots & P(x_4 / y_1) \end{array}$$

Аналогічно виглядає умовний закон розподілу іншої складової.

Умовний закон розподілу зручно розраховувати, користуючись табличним законом розподілу двовимірної випадкової величини

Приклад. Двовимірна випадкова величина задана таблицею 7.

Знайти безумовні і умовні закони розподілу складових  $Y$  та  $X$ .

Знаходимо безумовні закони розподілу складових.

Складова  $X$  набуває значень  $x_1 = 2$ ,  $x_2 = 3$ , ймовірність цих значень знайдемо, підсумувавши ймовірність по рядках навпроти кожного відповідного значення  $x_1 = 2$  і  $x_2 = 3$ . Результат записується в клітинки додаткового стовпця.

Таблиця 7

Y	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$
X			
$x_1 = 2$	0.10	0.20	0.30
$x_2 = 3$	0.05	0.15	0.20

Складова  $Y$  набуває значень  $y_1 = 2$ ,  $y_2 = 4$ ,  $y_3 = 6$ , ймовірність цих значень знайдемо, підсумувавши ймовірність по стовпцях навпроти кожного відповідного значення  $y_1 = 2$  і  $y_2 = 4$  і  $y_3 = 6$ . Результат заноситься в клітинки додаткового рядка. Результат обчислення запишемо у вигляді табл. 8, в якій також вказується початкова умова.

Таблиця 8

Y	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	$P(x)$
X				
$x_1 = 2$	0.10	0.20	0.30	0.60
$x_2 = 3$	0.05	0.15	0.20	0.40
$P(y)$	0.15	0.35	0.50	1.00

Знайдемо умовні закони розподілу складових. Всього таких законів можна скласти 5. Кількість законів дорівнюватиме сумі індексів складових  $i$  і  $j$ .

Для знаходження умовного закону розподілу  $Y$  від  $X$  знайдемо умовну ймовірність, наприклад,  $Y$ , якщо  $X$  набуває значення  $x_1 = 2$ , для цього скористаємося формулою:

$$P(y_1/x_1=2) = \frac{P(y_1x_1)}{P(x_1)} = \frac{0.10}{0.60} = \frac{1}{6};$$

Якщо  $X$  набуває значення  $x_1 = 2$ , то умовна ймовірність визначається формулою:

$$P(y_2/x_1=2) = \frac{P(y_2x_1)}{P(x_1)} = \frac{0.20}{0.60} = \frac{1}{3};$$

Якщо  $X$  набуває значення  $x_1 = 2$ , то умовна ймовірність визначається так:

$$P(y_3/x_1=2) = \frac{P(y_3x_1)}{P(x_1)} = \frac{0.30}{0.60} = \frac{1}{2};$$

Запишемо умовний закон розподілу складової  $Y$ , якщо  $X$  набула значення  $x_1 = 2$ , у вигляді табл. 9.

Таблиця 9

	$Y$	$Y_1 = 0.1$	$Y_2 = 0.2$	$Y_3 = 0.3$
$X$				
$P(y_i/x_1 = 2)$		$1/6$	$1/2$	$1/3$

Для перевірки правильності обчислень знайдемо суму ймовірності  $P(y_i/x_1 = 2)$  складових  $y_i$ :

$$Sp = (1/6) + (1/2) + (1/3) = 1.0.$$

В даному випадку  $y_i$  складають повну групу подій, тому сума ймовірності дорівнює одиниці, це означає, що обчислення виконане правильно.

Аналогічно знаходиться решта умовних законів розподілу, що становлять  $X$  і  $Y$ . Рекомендується знайти їх самостійно.

Умовні закони розподілу складових системи безперервних випадкових величин знаходяться за допомогою диференціальної функції системи і диференціальної функції складових. Дане питання достатньо детально розгляньте в /1/ або /2/.

### 1.9.2. Числові характеристики системи двох випадкових величин

Для опису як одновимірних так і для двовимірних випадкових величин в особливих випадках вигідно скористатися числовими характеристиками, при цьому можливо визначити: безумовні математичні очікування, дисперсії, середні квадратичні відхилення складових; умовні математичні очікування, дисперсії, середні квадратичні відхилення складових. Окрім вказаних числових характеристик, двовимірні випадкові величини характеризують кореляційним моментом і коефіцієнтом кореляції.

### 1.9.3. Залежні і незалежні випадкові величини

Складові двох-, трьох- і  $n$  – мірної випадкової величини можуть бути залежні один від одного або незалежні. Для визначення даної властивості

використовують різні методи, а саме: дві випадкові величини (складові) є незалежними, якщо закон розподілу однієї з них не залежить від того, яких можливих значень набуває інша величина.

З цього визначення виходить, що умовні закони розподілу незалежних складових величин рівні їх безумовним розподілам.

Існують необхідні і достатні умови незалежності випадкових величин.

Необхідність і достатність вказаної умови обґрунтовується рівністю інтегральної функції системи  $(X, Y)$  добутку інтегральних функцій складових:

$$F(x, y) = F_1(x) \cdot F_2(y)$$

Якщо  $X$  і  $Y$  незалежні, то події  $X < x$  і  $Y < y$  також незалежні, отже, ймовірність поєднання цих подій дорівнює добутку їх ймовірностей. Тобто може наступити і та і інша при логічному «і» ймовірність подій помножується. Тому на основі визначення інтегральної функції можна визначити:

$$P(X < x, Y < y) = P(X < x) \cdot P(Y < y) \quad \text{або} \quad F(x, y) = F_1(x) \cdot F_2(y)$$

Виходячи від зворотнього, а саме, якщо  $F(x, y) = F_1(x) \cdot F_2(y)$ , то

$$P(X < x, Y < y) = P(X < x) \cdot P(Y < y)$$

Оце і є обґрунтуванням необхідності та достатності.

Наведене є справедливим, якщо обумовлення виконується з використанням диференціальних функцій складових:

$$f(x, y) = f_1(x)f_2(y)$$

Якщо  $X$  і  $Y$  — незалежні безперервні випадкові величини, тоді (на підставі попередньої теореми):

$$F(x, y) = F_1(x) \cdot F_2(y)$$

Диференціюючи цю рівність по  $x$ , потім по  $y$ , маємо:

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial F_1}{\partial x} \cdot \frac{\partial F_2}{\partial y}$$

або (за визначенням диференціальних функцій двовимірної і одновимірної величин):

$$f(x, y) = f_1(x)f_2(y)$$

Якщо:

$$f(x, y) = f_1(x)f_2(y)$$

Інтегруючи цю рівність по  $x$  і по  $y$ , отримаємо:

$$\int_{-\infty}^y \int_{-\infty}^x f(x, y) dx dy = \int_{-\infty}^x f_1(x) dx \int_{-\infty}^y f_2(y) dy \quad \text{або} \quad F(x, y) = F_1(x) \cdot F_2(y)$$

Звідси (на підставі попередньої теореми) виходить, що  $X$  і  $Y$  незалежні.

Оскільки приведені вище умови є необхідними і достатніми, то виникають нові визначення незалежних випадкових величин:

- дві випадкові величини називають незалежними, якщо інтегральна функція системи цих величин дорівнює добутку інтегральних функцій складових;

- дві безперервні випадкові величини називають незалежними, якщо диференціальна функція систем цих величин дорівнює добутку диференціальних функцій складових.

#### 1.9.4. Числові характеристики системи двох випадкових величин. Кореляційний момент. Коефіцієнт кореляції

Для опису системи двох випадкових величин окрім математичних очікувань і дисперсій складових користуються і іншими характеристиками, до яких належать кореляційний момент і коефіцієнт кореляції.

Кореляційним моментом  $\mu_{xy}$  випадкових величин  $X$  і  $Y$  називають математичне очікування добутку відхилень цих величин від математичного очікування:

$$\mu_{xy} = M[(X - M(X))(Y - M(Y))]$$

Для обчислення кореляційного моменту дискретних величин користуються формулою:

$$\mu_{xy} = \sum_{i=1}^n \sum_{j=1}^m (x_i - M(X))(y_j - M(Y))p(x_i, y_j),$$

а для безперервних величин:

$$\mu_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - M(X))(y - M(Y))f(x, y) dx dy$$

Кореляційний момент служить для характеристики зв'язку між величинами  $X$  і  $Y$ . Як буде показано нижче, кореляційний момент дорівнює нулю, якщо  $X$  і  $Y$  незалежні; отже, якщо кореляційний момент не дорівнює нулю, то  $X$  і  $Y$  — залежні випадкові величини.

Якщо в якості аргументу для встановлення міри залежності чинників застосовується кореляційний момент, то очевидно стає норма. Кореляційний момент двох незалежних випадкових величин  $X$  і  $Y$  дорівнює нулю.

Доказ. Оскільки  $X$  і  $Y$  незалежні випадкові величини, то їх відхилення  $X - M(X)$  і  $Y - M(Y)$  також незалежні. Користуючись властивостями математичного очікування (математичне очікування добутку незалежних випадкових величин рівно добутку математичних очікувань складових) і відхилення (математичне очікування відхилення дорівнює нулю), отримаємо:

$$\begin{aligned} \mu_{xy} &= M[(X - M(X))(Y - M(Y))] = \\ &= M[(X - M(X))]M[(Y - M(Y))] = 0 \end{aligned}$$

З визначення кореляційного моменту виходить, що він має розмірність рівну добутку розмірностей величин  $X$  і  $Y$ . Іншими словами величина кореляційного моменту залежить від одиниць вимірювання випадкових величин. З цієї причини для одних і тих же двох величин величина

кореляційного моменту матиме різні значення залежно від того, в яких одиницях були виміряні величини.

Хай, наприклад,  $X$  і  $Y$  були зміряні в кілограмах і  $\mu_{xy} = 2 \text{ кг}^2$ ; якщо виміряти  $X$  і  $Y$  в грамах, то  $\mu_{xy} = 2000 \text{ гр}^2$ . Це властивість кореляційного моменту як числової характеристики є її недоліком, оскільки порівняння кореляційних моментів різних систем випадкових величин стає неправомірним. Для усунення цього недоліку вводять нову числову характеристику — коефіцієнт кореляції.

Коефіцієнтом кореляції  $r_{xy}$  випадкових величин  $X$  і  $Y$  називають відношення кореляційного моменту до добутку середніх квадратичних відхилень цих величин:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}$$

Оскільки розмірність  $r_{xy}$  дорівнює добутку розмірностей величин  $X$  і  $Y$ , а  $\sigma_x$  має розмірність величини  $X$ ,  $\sigma_y$  має розмірність величини  $Y$ , то  $r_{xy}$  є безрозмірна величина. Таким чином, величина коефіцієнта кореляції не залежить від вибору одиниць вимірювання випадкових величин. У цьому полягає перевага коефіцієнта кореляції перед кореляційним моментом.

Очевидно, коефіцієнт кореляції незалежних випадкових величин дорівнює нулю (оскільки  $r_{xy} = 0$ ).

### 1.9.5. Корельованість і залежність випадкових величин

Дві випадкові величини  $X$  і  $Y$  називають корельованими, якщо їх кореляційний момент (або ж - коефіцієнт кореляції) відмінний від нуля;  $X$  і  $Y$  називають некорельованими величинами, якщо їх кореляційний момент дорівнює нулю.

Якщо величини корельовані, то вони також і залежні. Дійсно, припускаючи протилежне, ми повинні визначити, що  $\mu_{xy} = 0$ , а це протирічить умові, оскільки для корельованих величин  $\mu_{xy} \neq 0$ .

Зворотне припущення не завжди має місце, тобто якщо дві величини залежні, то вони можуть бути як корельованими, так і некорельованими. Іншими словами, кореляційний момент двох залежних величин може не дорівнювати нулю, але може і дорівнювати нулю.

Для безперервних випадкових величин доказ їх залежності можливий, якщо скористатися диференціальною функцією.

На практиці часто зустрічаються двовимірні випадкові величини, розподіл яких мають нормальний закон розподілу.

Нормальним законом розподілу на площині називають розподіл ймовірності двовимірної випадкової величини  $(X, Y)$ , якщо:

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - r_{xy}^2}} \cdot e^{-\frac{1}{2(1 - r_{xy}^2)} \left[ \frac{(x - a_1)^2}{\sigma_x^2} + \frac{(y - a_2)^2}{\sigma_y^2} - 2r_{xy} \frac{(x - a_1)(y - a_2)}{\sigma_x \sigma_y} \right]}$$

Нормальний закон на площині визначається п'ятьма параметрами:  $a_1$ ,  $a_2$ ,  $\sigma_x$ ,  $\sigma_y$  і  $r_{xy}$ . Ці параметри мають наступне значення:

$a_1, a_2$  - математичні очікування  $\sigma_x, \sigma_y$  - середні квадратичні відхилення,  $r_{xy}$  - коефіцієнт кореляції величин  $X$  і  $Y$ . Якщо складові двовимірної нормально розподіленої випадкової величини не корельовані, то вони і незалежні. Дійсно, якщо  $X$  і  $Y$  є не корельовані. Вважаючи, що в попередній формулі  $r_{xy} = 0$ , отримаємо:

$$f(x, y) = f_1(x)f_2(y)$$

Таким чином, якщо складові нормально розподіленої випадкової величини не корельовані, то диференціальна функція системи дорівнює добутку диференціальних функцій складових, а звідси і слідує незалежність складових. Справедливо і зворотнє твердження. Для нормально розподілених складових двовимірної випадкової величини поняття незалежності і некорельованості — рівносильні.

### Питання

1. У чому суть випадковості величин. Приведіть конкретний приклад.
2. Як можна задати випадкову величину.
3. Вкажіть особливості числових характеристик випадкових величин.
4. Дайте визначення двовимірної випадкової величини.
5. Як задаються двовимірні випадкові величини.
6. Вкажіть особливості кореляційного моменту і коефіцієнта кореляції і як вони використовуються практично.

### Завдання

1. Знайти закони розподілу складових дискретної двовимірної випадкової величини, заданої законом розподілу

X	$x_1 = 1$	$x_2 = 2$	$x_3 = 3$
Y			
$y_1 = 2$	0.05	0.3	0.05
$y_2 = 4$	0.01	0.4	0.19

Відп. X 1 2 3 Y 2 4  
N 0.06 0.7 0.24 N 0.4 0.6

2. Для умов даного прикладу знайти умовні закони розподілу: для  $X$ , якщо  $Y = 4$  і для  $Y$ , якщо  $X=2$ .
3. Для умов даного прикладу знайти кореляційний момент і коефіцієнт кореляції.

## 2. Математична статистика

### 2.1. Призначення математичної статистики

У реальній життєвій дійсності існує безліч задач, вирішити які за допомогою теорії ймовірностей неможливо. Сказане відноситься також стосовно визначення дійсних значень чинників і закономірностей їх розподілу та зв'язку - взаємодії.

Для вирішення таких питань створено математичний метод – математична статистика.



Головними складовими призначення математичної статистики є розробка способів збору для чинників - статистичних даних і їх групування, якщо даних багато, визначення значень оцінок невідомих чинників, створення методів аналізу та їх обробка для отримання наукових і практичних висновків стосовно їх взаємодії.

### **2.1.1. Генеральна і вибіркова сукупності**

*Генеральна і вибіркова сукупності* це узагальнювальні поняття, які використовуються багатогранно. Наприклад, якщо проводиться експеримент для встановлення закономірності взаємодії між параметрами або визначаються шляхом виміру відповідні їх характеристики, то цим самим ми намагаємося вивчити те, що насправді є в природі. Вказані дії, як правило, проводяться багато разів. Ту кількість проведених дослідів або вимірів прийнято називати вибіркою, а кількість, з якої витягується (отримана з експерименту або виміру) дана вибірка називається генеральною сукупністю. Наприклад, якщо встановлюється математична залежність розміру часток матеріалу на виході барабанного кульового млина від числа обертів барабана, його діаметру і інших параметрів, то можна провести нескінченну кількість експериментів. В даному випадку нескінченна кількість експериментів складатиме генеральну сукупність. Насправді проводять певну кількість експериментів і ця кількість складатиме вибірку. Інший приклад, є 100 вагонів вугілля, яке має відповідну зольність і цю зольність потрібно визначити. Для вирішення даного питання потрібно спалити все вугілля, частина, що при цьому залишилася, представлятиме золу. Таке рішення даного питання є абсурдним і на практиці цього не роблять. Із ста вагонів намічають певну кількість (5-10) і в них за певним правилом відбирають відповідну частину вугілля, спалюють її або іншим способом визначають зольність. В даному випадку 100 вагонів представлятимуть генеральну сукупність, а число відібраних – вибірковою сукупністю. Надалі досліді, виміри і об'єкти називатимемо узагальнено об'єктами.

Вибірковою сукупністю, або просто вибіркою, називають сукупність випадково відібраних об'єктів.

Генеральною сукупністю називають сукупність об'єктів, з яких проведена вибірка.

Об'ємом сукупності (вибіркової або генеральної) називають число об'єктів цієї сукупності.

Кількість об'єктів, які повинні відбиратися у вибірку, визначається за відповідними методиками або правилами, розробленими з врахуванням галузевих або іншими особливостей. Відібрані об'єкти досліджуються, тобто визначаються відповідні значення їх чинників.

### **2.1.2. Повторна і неповторна вибірки. Репрезентативна вибірка**

Об'єкти у вибірку сукупність можуть вибиратись по різному. Наприклад, якщо після того, як об'єкти вибірки відібрані і відпрацьовані, з ними можна поступати двояко: вони можуть бути повернені, або не повернені в

генеральну сукупність. Відповідно до сказаного, вибірки підрозділяють на повторні і неповторні.

Повторною називають вибірку, при застосуванні якої відібраний об'єкт (перед відбором наступного) повертається в генеральну сукупність.

Безповторною називають вибірку, при застосуванні якої відібраний об'єкт в генеральну сукупність не повертається.

Для того, щоб за даними вибірки можна було достатньо упевнено судити про відповідну ознаку генеральної сукупності, яка визначається, необхідно, щоб об'єкти вибірки правильно його представляли. Цю вимогу коротко формулюють так: вибірка має бути репрезентативною (представницькою).

Репрезентативність вибірки реалізується в тому випадку, якщо відбір об'єктів або його складових здійснюється випадково і кожен об'єкт або складова мають однакову ймовірність потрапити у вибірку.

### **2.1.3. Способи відбору**

На практиці, в різних галузях застосовуються різні способи відбору, причому всю їх сукупність ділять на два види:

1. Відбір, що не вимагає розчленовування генеральної сукупності на окремі частини, куди відносяться:

- а) простий випадковий неповторний відбір;
- б) простий випадковий повторний відбір.

2. Відбір, при якому генеральна сукупність розбивається на частини, куди відносяться:

- а) типовий відбір;
- б) механічний відбір;
- в) серійний відбір.

Простим випадковим називають такий відбір, при якому об'єкти відбирають з генеральної сукупності поодиноці. Здійснюють простий відбір різними способами. Наприклад, якщо потрібно взяти  $n$  об'єктів з генеральної сукупності об'єму  $N$ , то виписують номери об'єктів від 1 до  $N$  на картки, які ретельно перемішують і навмання виймають одну картку; об'єкти, що мають однакові номери з карткою, піддають обстеженню; потім картка повертається в пачку і процес повторюється, тобто картки перемішують, навмання виймають одну з них і так далі. Так поступають  $n$  разів; у результаті отримують просту випадкову повторну вибірку об'ємом  $n$ .

Якщо картки, що витягують, не повертати в пачку, то вибірка буде простою випадковою неповторною.

При великому об'ємі генеральної сукупності користуються готовими таблицями «випадкових чисел». Для того, щоб відібрати, наприклад  $N$  об'єктів з пронумерованою генеральною сукупністю, відкривають будь-яку сторінку таблиці випадкових чисел і виписують підряд  $N$  чисел; у вибірку потрапляють ті об'єкти, номери яких співпадають з виписаними випадковими числами.

Типовим називають відбір, при якому об'єкти відбирають не зі всієї генеральної сукупності, а з кожної її «типової» частини. Наприклад, якщо деталі виготовляють на декількох верстатах, то відбір проводять не зі всієї сукупності деталей, виготовлених всіма верстатами, а з продукції кожного

верстата окремо. Типовим відбором користуються тоді, коли обстежувана ознака помітно коливається в різних типових частинах генеральної сукупності. Наприклад, якщо продукція виготовлялась на декількох машинах, серед яких є більш і менш зношені, то тут типовий відбір є доречним.

Механічним називають відбір, при якому генеральна сукупність «механічно» ділиться на стільки груп, скільки об'єктів повинно увійти до вибірки і з кожної групи відбирається один об'єкт.

Наприклад, якщо потрібно відібрати 20% виготовлених верстатом деталей, то відбирають кожну п'яту деталь; якщо потрібно відібрати 5% деталей, то відбирають кожну двадцату деталь і так далі.

Серійним називають відбір, при якому об'єкти відбирають з генеральної сукупності не поодинці, а «серіями», які піддаються суцільному обстеженню. Наприклад, якщо вироби виготовляються великою групою верстатів-автоматів, то піддають суцільному обстеженню продукцію тільки декількох верстатів. Серійним відбором користуються тоді, коли обстежувана ознака коливається в різних серіях не суттєво.

На практиці часто застосовується комбінований відбір, при якому поєднуються вказані вище способи.

Іноді генеральну сукупність розбивають на серії однакового об'єму, потім простим випадковим відбором вибирають декілька серій, а з кожної серії простим випадковим відбором відбирають окремі об'єкти.

Репрезентативність, це поняття яке має відповідний числовий вимір. Яким чином визначається його значення буде показано у подальшому.

Передумовою дотримання репрезентативності являється те, яким чином одержують числові значення чинника, тобто як виконується дослідження значень складових вибірки.

У практичній роботі, що стосується збагачення корисних копалин, значення можуть одержувати з урахування певних обставин по різному.

Якщо відсутні технічні засоби безперервного або дискретного контролю вимірювання значень чинників, то їх визначають з результатів дослідження продуктів відбору. Наприклад, визначається значення зольності продуктів збагачення по операціях технологічної схеми або кінцевих продуктів. Від певних продуктів відбирають проби, обробляють їх відповідним чином і за допомогою відповідного методу визначають зольність. Таких досліджень роблять стільки, скільки чисел необхідно одержати у відповідну сукупність, наприклад, упродовж терміну роботи зміни.

Якщо значення чинників визначаються безперервно або дискретно за допомогою технічних засобів, то як правило значення представляються відповідною діаграмою. В такому випадку з діаграми необхідно взяти певну кількість чисел або пар чисел. В цьому випадку теж необхідно дотримуватись певної умови. Практично це робиться таким чином.

Якщо діаграма представлена безперервною лінією, що має певні коливання різної амплітуди та частоти, то візуально знаходять ділянку, на якій зміна знаку чинника має найменшу довжину. Намічену довжину ділять на дві рівні частини

і через такі інтервали знімають – виписують значення чинника. Приклад одержання значень з діаграми наведено на рис. 2.

#### 2.1.4. Статистичний розподіл вибірки

Після того, як відібрані об'єкти - обстежені щодо деякої ознаки або отримані результати дослідів або вимірів, здійснюється їх обробка і представлення. Для обробки і представлення статистичних даних розроблені спеціальні методи, в даному випадку представляються загальні поняття.

Як в теорії ймовірності, так і в математичній статистиці застосовують відповідні методи представлення випадкових величин - це статистичний розподіл і емпіричні функції.

Нехай з генеральної сукупності одержана вибірка, причому  $x_1$  спостерігалось  $n_1$  разів,  $x_2 - n_2$  разів,  $x_k - n_k$  разів і  $\sum n_i = N$  об'єм вибірки. Одержані значення  $x_i$  називають *варіантами*, а послідовність варіантів, записаних в зростаючому порядку — *варіаційним рядом*. Числа, що вказують кількість спостережень називають *частотами*, а їх відношення до об'єму вибірки  $n_i / N$  — *відносними частотами*.



Рис. 2. Діаграма змін значень чинника

Статичним розподілом вибірки називають перелік варіантів і відповідних їм частот або відносних частот. Статистичний розподіл можна задати також у вигляді послідовності інтервалів і відповідних їм частот (розміром відповідного інтервалу, приймають суму частот, що потрапили в цей інтервал).

У теорії ймовірності під розподілом розуміють відповідність між можливими значеннями випадкової величини і їх ймовірністю, а в математичній статистиці — відповідність між спостережуваними варіантами - числами і їх частотами, або відносними частотами.

Приклад. Для визначення зольності намічено і випробовано 20 вагонів, при цьому в п'яти зольність була 35 %, у десяти - 38 % і в п'яти - 40 %. Якщо

позначити зольність  $A^d$  – це варіанти, а число вагонів –  $n_i$  – частоти, то отримане наступне:

$A^d$	35	38	40
$N$	5	10	5

В даному випадку записано розподіл значень якості у відповідності до частоти її появи.

Представимо розподіл з використанням відносних частот, для чого розділимо відповідні частоти на об'єм вибірки:

$$w_1 = \frac{5}{20} = 0.25; \quad w_2 = \frac{10}{20} = 0.5; \quad w_3 = \frac{5}{20} = 0.25$$

Напишемо розподіл значення зольності для відносних частот:

$A^d$	35	38	40
$N$	0.25	0.5	0.25

Контроль:  $0,25 + 0,5 + 0,25 = 1$ .

### 2.1.5. Емпірична функція розподілу

У деяких випадках окрім вказаного способу представлення випадкових величин потрібно знати – встановити закономірності їх співвідношення.

У цьому випадку застосовують емпіричну функцію розподілу.

Розглянемо наступну ситуацію.

Нехай відомий статистичний розподіл частот кількісної ознаки  $X$ . Введемо позначення:  $n_x$  — число спостережень, при яких спостерігалось значення ознаки менше  $x$ ,  $n$  — загальне число спостережень (об'єм вибірки).

Відносна частота події  $X < x$  дорівнює  $\frac{n_x}{n}$ . Якщо  $x$  змінюватиметься, то змінюватиметься і відносна частота, тобто відносна частота є функція від  $x$ . Оскільки ця функція знаходиться емпіричним (дослідним) шляхом, то її називають емпіричною.

Емпіричною функцією розподілу (функцією розподілу вибірки) називають функцію  $F'(x)$ , яка для кожного значення  $x$  визначає відносну частоту події  $X < x$ . Згідно визначенню вираз для обчислення емпіричної функції такий:

$$F'(x) = \frac{n_x}{n}$$

У теорії ймовірності цю функцію називають теоретичною функцією розподілу. Відмінність між емпіричною і теоретичною функціями полягає в тому, що теоретична функція  $F(x)$  визначає *ймовірність* події  $X < x$ , а емпірична функція  $F'(x)$  визначає *відносну частоту* цієї ж події. Емпірична функція розподілу вибірки використовується для наближеного представлення теоретичної (інтегральної) функції розподілу генеральної сукупності.

Даний висновок підтверджується тим, що  $F'(x)$  присутні всі властивості  $F(x)$ .

1. Значення емпіричної функції належать відрізку  $[0,1]$ ;
2.  $F'(x)$  - неубуваюча функція;
3. Якщо  $x_l$  - найменша варіанта, то  $F'(x) = 0$  при  $x \leq x_l$ , якщо  $x_k$  - найбільша варіанта, то  $F'(x) = 1$  при  $x > x_k$ .

Приклад. Побудувати емпіричну функцію по даному розподілу вибірки:

A <sup>d</sup>	35	38	40
N	0.25	0.5	0.25

Рішення. Найменша варіанту рівна  $x_1 = 35$ , а значень менше цього немає, отже,  $F'(x) = 0$  при  $x_1 \leq 35$ .

Якщо узяти наступну варіанту  $x_2 \leq 38$ , то менше її перша варіанту, а саме  $x_1 = 35$ , яка спостерігалася 0.25 разів;

отже

$$F'(x) = 0.25 \text{ при } 35 < x \leq 38$$

Наступна варіанта  $x_3 = 40$ , а менших є дві  $x_1 = 35$ , що спостерігалася 0.25 разів і  $x_2 = 38$  з відносною частотою 0.5 разів. Внаслідок того, що варіанти є несумісними, то за умови  $x \leq 40$  буде  $x_1 = 35$ , або  $x_2 = 38$ . Для логічного «або» відносна частота дорівнюватиме сумі частот даних варіантів, отже

$$F'(x) = 0.25 + 0.5 = 0.75 \text{ при } 38 < x \leq 40$$

Оскільки  $x_3 = 40$  найбільша варіанту, а відносна частота її появи дорівнює 0.25, то в даний інтервал потрапляє будь-яка з тих, що є і через їх несумісність емпірична функція буде дорівнювати:

$$F'(x) = 0.25 + 0.5 + 0.25 = 1.0 \text{ при } x > 40$$

Остаточний вираз в явній формі для шуканої емпіричної функції :

$$F'(x) = \left. \begin{array}{l} 0 \\ 0,22 \\ 0,75 \\ 1.0 \end{array} \right\} \begin{array}{l} \text{при } x \leq 35, \\ \text{при } 35 < x \leq 38, \\ \text{при } 38 < x \leq 40, \\ \text{при } x > 40. \end{array}$$

Графік цієї функції зображений на рис. 3.

### 2.1.6. Полігон і гистограма

Для більш легкого сприймання значень чинника будують різні графіки статистичного розподілу, зокрема, полігон і гистограму.

Полігоном частот називають ламану, відрізки якої з'єднують крапки  $(x_1, n_1)$ ,  $(x_2, n_2)$ , ...,  $(x_k, n_k)$ . Для побудови полігону частот на осі абсцис відкладають варіанти  $x_i$ , а на осі ординат відповідні ним частоти  $n_i$ . Крапки  $(x_i, n_i)$  з'єднують відрізками прямих і отримують полігон частот.

Полігоном відносних частот називають ламану, відрізки якої з'єднують крапки  $(x_1, W_1), (x_2, W_2), \dots, (x_k, W_k)$ . Для побудови полігону відносних частот на осі абсцис відкладають варіанти  $x_i$ , а на осі ординат відповідні їм відносні частоти  $W_i$ . Крапки  $(x_1, W_1)$  з'єднують відрізками прямих і отримують полігон відносних частот.

На рис. 4 зображений полігон відносних частот наступного розподілу попереднього прикладу:

$A^d$	35	38	40
$N$	0.25	0.5	0.25

Рис. 4. Полігон відносних частот

Гістограмою частот називають ступінчасту фігуру, що складається з прямокутників, основами яких служать часткові інтервали довжиною  $h_i$ , а висоти дорівнюють відношенню  $\frac{n_i}{h_i}$  (щільність частоти).

Для побудови гістограми частот на осі абсцис відкладають часткові інтервали, а над ними проводять відрізки, паралельні осі абсцис на відстані

$$\frac{n_i}{h_i}$$

Площа  $i$ -го часткового прямокутника дорівнює  $h \frac{n_i}{h} = n_i$  сумі частот варіант  $i$ -го інтервалу; таким чином, площа гістограми частот дорівнює сумі всіх частот, тобто об'єму вибірки.

Гістограмою відносних частот називають ступінчасту фігуру, що складається з прямокутників, основами яких служать часткові інтервали довжиною  $h$ , а висоти дорівнюють відношенню  $\frac{W_i}{h}$  (щільність відносних частот).

Для побудови гістограми відносних частот на осі абсцис відкладають часткові інтервали, а над ними проводять відрізки, паралельні осі абсцис на відрізок  $\frac{W_i}{h}$ . Приклад гістограми показано на рисунку 5.

### Завдання

1. Побудувати графік емпіричної функції розподілу

$x_i$	5	7	9	11
$n_i$	2	3	8	7

2. Побудувати полігони частот і відносних частот розподілу

$x_i$	3	5	7	9
$n_i$	10	31	33	12

3. Для попереднього прикладу розрахувати і записати емпіричну функцію розподілу в явному вигляді.

Рис. 5. Гістограма частот

### 2.2. Статистичні оцінки параметрів розподілу

Для певних умов оперувати вказаними представленнями випадкових чинників незручно, тому, зважаючи на це у математичній статистиці користуються узагальнюючими поняттями – оцінками.

Нехай потрібно визначити значення - оцінку деякої ознаки . Припустимо, що з теоретичних міркувань або практичного досвіду вдається встановити, який саме розподіл має ознака. В цьому випадку задача зводиться до визначення таких характеристик - (параметрів) чинника, за допомогою яких визначається цей розподіл. Наприклад, якщо відомо, що ознака, яка вивчається, розподілена в генеральній сукупності згідно нормального закону, то необхідно знайти математичне очікування і середнє квадратичне відхилення, оскільки по цих двох параметрах можна написати нормальний закон розподілу. Якщо визначити за результатами практичних даних дійсне математичне очікування і середнє квадратичне відхилення неможливо, то їх значення знаходять по практичним даним, а ці знайдені значення являються оцінками дійсних і невідомих нам значень чинника. У цьому випадку одержані – знайдені значення називають статистичними оцінками. Якщо відомо, що ознака має закон розподілу, наприклад, розподіл Пуассона, то необхідно знайти оцінку - параметр  $\lambda$ , яким цей розподіл визначається.

Зазвичай, в цьому випадку користуються даними досліджень або вимірів (спостережень), які є вибіркою з генеральної сукупності, наприклад, значення  $x_1, x_2, \dots, x_n$  деякої ознаки, отримані в результаті  $n$  спостережень (тут і далі



спостереження передбачаються незалежними). Розглядаючи сукупність  $x_1, x_2, \dots, x_n$  як незалежні випадкові величини  $X_1, X_2, \dots, X_n$ , можна сказати, що знайти статистичну оцінку невідомого параметра теоретичного розподілу — це означає знайти функцію (математичне очікування, дисперсію, середнє квадратичне відхилення і ін.) від спостережуваних випадкових величин, яка і дає наближене значення оцінюваного параметра. У загальному вигляді це можна записати:

$$T \approx O = f(X),$$

де  $X$  — ознака (випадкова величина),  $O$  — оцінка,  $T$  — дійсне значення ознаки.

Отже, статистичною оцінкою невідомого параметра теоретичного розподілу називають функцію, за якою визначається оцінка, скориставшись значеннями спостережуваних випадкових величин.

### 2.2.1. Незміщені, ефективні і спроможні оцінки

Для того, щоб статистичні оцінки якомога точніше оцінювали невідомі - дійсні параметри, вони повинні задовольняти певним вимогам.

Нехай  $O$  - статистична оцінка невідомого параметра,  $T$  — його дійсне теоретичне значення. По вибірці об'єму  $n$  знайдена оцінка  $O_1$ . У наступному дослідженні, тобто витягуючи з генеральної сукупності іншу вибірку того ж об'єму і за її даними, знайдемо оцінку  $O_2$ . Повторюючи дослід багато разів, отримуємо числа  $O_1, O_2, \dots, O_n$ , які будуть різні між собою. Таким чином, оцінку  $O$  можна розглядати як випадкову величину, а числа  $O_1, O_2, \dots, O_n$  - як її можливі значення.

Якщо оцінка  $O$  дає наближене значення  $T$  з лишком; тоді кожне знайдене за даними вибірок чисел  $O_1, O_2, \dots, O_n$  буде більше дійсного значення  $T$ . В цьому випадку математичне очікування (середнє значення) випадкової величини  $O$  буде більше, ніж  $T$ , тобто  $M(O) > T$ . Очевидно, що якщо  $O$  дає оцінку з недоліком, то  $M(O) < T$ .

Щоб математичне очікування оцінки  $O$  дорівнювало оцінюваному параметру, повинна виконуватися вимога  $M(O) = T$ .

Незміщеною називають статистичну оцінку  $O$ , математичне очікування якої дорівнює оцінюваному параметру  $T$  при будь-якому об'ємі вибірки, тобто  $M(O) = T$ .

Зміщеною називають оцінку, математичне очікування якої не дорівнює оцінюваному параметру.

Проте, незміщена оцінка не завжди дає хороше наближення оцінюваного параметра. Значення  $O$  можуть бути сильно розсіяні навколо свого середнього значення, тобто дисперсія  $D(O)$  може бути значною. В цьому випадку знайдена за даними однієї вибірки оцінка, наприклад  $O_i$ , може виявитися дуже віддаленою від середнього значення  $O$ , а значить, і від самого оцінюваного параметра  $T$ ; прийнявши  $O_i$  як наближене значення  $T$ , ми допустили б велику помилку. Якщо ж вимагати, щоб дисперсія  $D(O)$  була малою, то допустити велику помилку буде унеможливлено. З цієї причини до статистичної оцінки пред'являється вимога ефективності.

Ефективною називають статистичну оцінку, яка при заданому об'ємі вибірки  $n$  має найменшу можливу дисперсію.

При розгляді вибірок великого об'єму до статистичних оцінок пред'являється вимога спроможності.

Спроможною називають статистичну оцінку, яка при  $n \rightarrow \infty$  прагне по ймовірності до оцінюваного параметра. Наприклад, якщо дисперсія незміщеної оцінки при  $n \rightarrow \infty$  прагне до нуля, то така оцінка є і спроможною.

### 2.2.2. Генеральна середня

Нехай вивчається дискретна генеральна сукупність щодо деякої ознаки  $X$ .

Генеральною середньою називають середнє арифметичне значення ознаки генеральної сукупності.

Якщо всі значення  $x_1, x_2, \dots, x_n$  ознаки генеральної сукупності об'єму  $N$  різні, то вона розраховується такою формулою:

$$\bar{x}_2 = \frac{x_1 + x_2 + \dots + x_N}{N}$$

Якщо ж значення ознаки  $x_1, x_2, \dots, x_n$  мають відповідно частоти  $n_1, n_2, \dots, n_n$ , причому  $n_1 + n_2 + \dots + n_n = N$ , то формула така:

$$\bar{x}_2 = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{N}$$

### 2.2.3. Вибіркова середня

Якщо генеральна сукупність велика, то з неї вибирають вибірку.

Нехай для вивчення генеральної сукупності відносної кількісної ознаки  $X$  одержана вибірка об'єму  $n$ .

Вибірковою середньою називають середнє арифметичне значення ознаки вибіркової сукупності.

Якщо всі значення  $x_1, x_2, \dots, x_n$  ознаки вибірки об'єму  $n$  різні, то оцінку розраховують:

$$\bar{x}_g = \frac{x_1 + x_2 + \dots + x_N}{N}$$

Якщо ж значення ознаки  $x_1, x_2, \dots, x_n$  мають відповідно частоти  $n_1, n_2, \dots, n_n$ , причому  $n_1 + n_2 + \dots + n_n = N$ , формула наступна:

$$\bar{x}_g = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{N}$$

Тобто вибіркова середня є середня зважена значень ознаки з вагами, рівними відповідним частотам.

### 2.2.4. Оцінка генеральної середньої по вибірковій середній. Стійкість вибірових середніх

Нехай з генеральної сукупності (в результаті незалежних спостережень над кількісною ознакою  $X$ ) одержана повторна вибірка об'єму  $n$  із значеннями

ознаки  $x_1, x_2, \dots, x_n$ . Можливо ці значення ознаки різні. Хай генеральна середня  $\bar{x}_g$  невідома і потрібно оцінити її за даними вибірки. У якості оцінки генеральної середньої приймають вибіркovo середню:

$$\bar{x}_g = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Розглядаючи  $\bar{x}_g$  як випадкову величину отриману по незалежних однаково розподілених випадкових величинах  $x_1, x_2, \dots, x_n$ , є підстава припустити, що вони мають однакові числові характеристики, зокрема однакове математичне очікування, яке позначимо через  $a$ . Оскільки математичне очікування середнього арифметичного однаково розподілених випадкових величин дорівнює математичному очікуванню кожній з величин, то:

$$M(\bar{X}_g) = M\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = a$$

Враховуючи те, що кожна з величин  $X_1, X_2, \dots, X_n$  має той же розподіл, що і генеральна сукупність (яку ми також розглядаємо як випадкову величину), припускаємо, що і числові характеристики цих величин і генеральної сукупності однакові. Зокрема, математичне очікування кожної з величин дорівнює математичному очікуванню ознаки  $X$  генеральної сукупності, тобто:

$$M(\bar{X}_g) = a$$

Замінивши в цій формулі математичне очікування  $a$  через  $\bar{x}_g$ , остаточно отримуємо:

$$M(\bar{X}_g) = \bar{x}_g$$

Отже вибіркoва середня є незміщеною оцінкою генеральної середньої.

Вибіркова середня є і спроможною оцінкою генеральною середньою.

Якщо збільшувати об'єм вибірки  $n$ , то вибіркoва середня прагне по ймовірності до генеральної середньої, а це означає, що вибіркoва середня є спроможною оцінкою генеральної середньої.

З цього виходить також, що якщо по декількох вибірках достатньо великого об'єму з однієї і тієї ж генеральної сукупності будуть знайдені вибіркові середні, то вони будуть приблизно рівні між собою. У цьому полягає властивість стійкості вибіркових середніх.

Близькість вибіркових середніх до генеральних залежить від об'єму вибірки: чим об'єм вибірки більший, тим менше вибіркoва середня відрізняється від генеральної. Наприклад, якщо з однієї сукупності відібрано 1% об'єктів, а з іншої сукупності відібрано 10% об'єктів, причому об'єм першої вибірки виявився більшим, ніж другий, то перша вибіркoва середня менше відрізнятиметься від відповідної генеральної середньої, ніж друга.

### 2.2.5. Групова і загальна середні

Якщо об'єм вибірки виявився великим, то визначати статистичні оцінки дійсних чинників складно, у цьому випадку всю сукупність розбивають на окремі групи і певні математичні дії проводять для виділених груп.

Припустимо, що всі наявні значення  $X$  генеральної або вибіркової сукупності розбиті на декілька груп. Розглядаючи кожну групу як самостійну сукупність, можна знайти її середню арифметичну.

Груповою середньою називають середнє арифметичне значення ознаки, що належить групі.

Загальною середньою  $\bar{x}$  називають середнє арифметичне значення ознаки, що належить всій сукупності.

Знаючи групові середні і об'єми груп, можна знайти загальну середню, яка дорівнює середньому арифметичному групових середніх, зважених по об'ємах груп.

Приклад. Знайти загальну середню сукупності, що складається з наступних груп:

Група	перша		друга	
Значення ознаки	1	3	2	6
Частота	4	8	10	20
Об'єм	4 + 8 = 12		10 + 20 = 30	

Знайдемо групові середні:

$$\bar{x}_1 = \frac{1 \cdot 4 + 3 \cdot 8}{12} = 3$$

$$\bar{x}_2 = \frac{2 \cdot 10 + 6 \cdot 20}{30} = 4.6$$

Знайдемо загальну середню по групових середніх:

$$\bar{x} = \frac{3 \cdot 12 + 4.6 \cdot 30}{42} = 4.19$$

### 2.2.6. Відхилення від загальної середньої і його властивість

Відхиленням називають різницю  $x_i - \bar{x}$  між значеннями ознаки і загальною середньою.

Теорема. Сума добутку відхилень на відповідні частоти дорівнює нулю

$$\sum n_i(x_i - \bar{x}) = 0$$

Доказ:

$$\sum n_i(x_i - \bar{x}) = \sum n_i x_i - \sum n_i \bar{x} = n\bar{x} - n\bar{x} = 0$$

Приклад. Дано розподіл кількісної ознаки  $X$ :

$x_i$	1	2	4
$n_i$	2	4	8

Рішення. Знайдемо загальну середню:

$$\bar{x} = \frac{1 \cdot 2 + 2 \cdot 4 + 4 \cdot 8}{14} = 3.0$$

Знайдемо суму добутків відхилень на відповідні їм частоти:

$$\sum n_i(x_i - \bar{x}) = 2(1-3.0) + 4(2-3.0) + 8(4-3.0) = 0$$

### 2.2.7. Генеральна дисперсія

Для того, щоб охарактеризувати розсіяння значень ознаки  $X$  генеральної сукупності навколо свого середнього значення, використовують узагальнюючу характеристику — генеральну дисперсію.

Генеральною дисперсією  $D$  називають середнє арифметичне квадратів відхилень значень ознаки генеральної сукупності від їх середнього значення.

Якщо всі значення  $x_1, x_2, \dots, x_n$  ознаки генеральної сукупності об'єму  $N$  різні, то дисперсія обчислюється такою формулою:

$$D_e = \frac{\sum_{i=1}^N (x_i - \bar{x}_e)^2}{N}$$

Якщо ж значення ознаки  $x_1, x_2, \dots, x_n$  мають відповідно частоти  $N_1, N_2, \dots, N_k$ , то формула наступна:

$$D_e = \frac{\sum_{i=1}^N N_i (x_i - \bar{x}_e)^2}{N}$$

Тобто генеральна дисперсія є середня зважена квадратів відхилень з вагами, рівними відповідним частотам.

Приклад. Генеральна сукупність задана табличним законом розподілу:

$x_i$	1	2	3	4
$N_i$	2	3	4	1

Знайти генеральну дисперсію.

Рішення. Знайдемо генеральну середню:

$$\bar{x}_e = \frac{2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + 1 \cdot 4}{10} = 2.4$$

Знайдемо генеральну дисперсію:

$$D_e = \frac{2(1 - 2.4)^2 + 3(2 - 2.4)^2 + 4(3 - 2.4)^2 + 1(4 - 2.4)^2}{10}$$

Окрім дисперсії для характеристики розсіяння значень ознаки генеральної сукупності навколо свого середнього значення користуються узагальнюючою характеристикою — середнім квадратичним відхиленням.

Генеральним середнім квадратичним відхиленням (стандартом) називають квадратний корінь з генеральної дисперсії:

$$\sigma_e = \sqrt{D_e}$$

### 2.2.8. Вибіркова дисперсія

Для того, щоб охарактеризувати розсіяння одержаних значень ознаки вибірки навколо свого середнього значення, вводять узагальнюючу характеристику — вибіркову дисперсію.

Вибіркова дисперсія являється оцінкою генеральної.

Вибірковою дисперсією  $D_v$  називають середнє арифметичне квадратів відхилень спостережуваних значень ознаки від їх середнього значення  $\bar{x}_e$ .

Якщо всі значення  $x_1, x_2, \dots, x_n$  ознаки вибірки об'єму  $n$  різні, то формула для обчислення дисперсії має вигляд:

$$D_v = \frac{\sum_{i=1}^n (x_i - \bar{x}_e)^2}{n}$$

Якщо значення ознаки  $x_1, x_2, \dots, x_n$  мають відповідно частоти  $N_1, N_2, \dots, N_k$ , то формула така:

$$D_{\sigma} = \frac{\sum_{i=1}^N N(x_i - \bar{x}_{\sigma})^2}{N}$$

Приклад. Вибіркова сукупність задана таблицею розподілу

$x_i$	2	3	4
$\pi_i$	5	10	5

Знайти вибірку дисперсію.

Рішення. Знайдемо вибірку середню:

$$\bar{x}_{\sigma} = \frac{5 \cdot 2 + 10 \cdot 3 + 5 \cdot 4}{5 + 10 + 5} = 3$$

Знайдемо вибірку дисперсію:

$$D_{\sigma} = \frac{5(2-3)^2 + 10(3-3)^2 + 5(4-3)^2}{20} = 0.5$$

Окрім дисперсії для характеристики розсіяння значень ознаки вибіркової сукупності навколо свого середнього значення користуються узагальнюючою характеристикою — середнім квадратичним відхиленням.

Вибірковим середнім квадратичним відхиленням (стандартом) називають квадратний корінь з вибіркової дисперсії:

$$s_{\hat{a}} = \sqrt{S_{\hat{a}}}$$

Обчислення дисперсії, байдуже, вибіркової або генеральної, зручніше здійснювати, використовуючи теорему.

Теорема. Дисперсія дорівнює різниці між середнім квадратів ознаки мінус квадрат середньої:

$$S = \bar{x}^2 - [\bar{x}]^2$$

Використовуючи попередні викладення, доведіть теорему самостійно

Приклад. Знайти дисперсію по даному розподілу:

$x_i$	2	3	4
$\pi_i$	5	10	5

Рішення. Знайдемо загальну середню:

$$\bar{x}_{\sigma} = \frac{5 \cdot 2 + 10 \cdot 3 + 5 \cdot 4}{5 + 10 + 5} = 3$$

Знайдемо середню квадратів значень ознаки:

$$\bar{x}^2 = \frac{5 \cdot 2^2 + 10 \cdot 3^2 + 5 \cdot 4^2}{20} = 9.5$$

Шукана дисперсія

$$S = 9.5 - 3^2 = 0.5$$

### 2.2.9. Групова, внутрішньогрупова, міжгрупова і загальна дисперсії

Припустимо, що всі значення ознаки  $X$  сукупності, байдуже, генеральної або вибіркової, розбиті на  $k$  груп. Розглядаючи кожну групу як самостійну сукупність, можна знайти групову середню і дисперсію значень ознаки, що належать відповідній групі щодо групової середньої.

Груповою дисперсією називають дисперсію значень ознаки, що належать групі щодо групової середньої:

$$D_{jзр} = \frac{\sum_{i=1}^N N_i(x_i - \bar{x}_j)^2}{N_j},$$

де  $n_i$  - частота значення  $x_i$ ;  $j$  - номер групи;  $\bar{x}_j$  - групова середня групи  $j$ ;  
 $N_j = \sum n_i$  - об'єм групи  $j$ .

Приклад I. Знайти групові дисперсії сукупності, що складаються з наступних двох груп:

	Перша група				Друга група			
	$X_1$	1	3	5	$X_2$	2	3	4
	$N_1$	2	4	1	$N_1$	1	4	1

Знайдемо групові середні:

$$\bar{x}_1 = \frac{\sum n_i x_i}{\sum n_i} = \frac{2 \times 1 + 4 \times 3 + 1 \times 5}{7} = 2.71; \quad \bar{x}_2 = \frac{\sum n_i x_i}{\sum n_i} = \frac{1 \times 2 + 4 \times 3 + 1 \times 4}{6} = 3.0$$

Знайдемо шукані групові дисперсії:

$$D_{1зр} = \frac{\sum n_i (x_i - \bar{x}_1)^2}{N_1} = \frac{2(1 - 2.71)^2 + 4(3 - 2.71)^2 + 1(5 - 2.71)^2}{7} = 1.63$$

$$D_{2зр} = \frac{\sum n_i (x_i - \bar{x}_2)^2}{N_1} = \frac{1(2 - 3)^2 + 4(3 - 3)^2 + 1(4 - 3)^2}{6} = 0.33$$

Знаючи дисперсію кожної групи, знаходять їх середню арифметичну і таку дисперсію називають внутрішньогруповою.

Внутрішньогруповою дисперсією називають середнє арифметичне групових дисперсій зважену по об'ємам груп:

$$D_{зр} = \frac{\sum N_j D_{jзр}}{n},$$

де  $N_j$  - об'єм групи  $j$ ;  $n$  - об'єм всієї сукупності

Шукана внутрішньогрупова дисперсія для даного прикладу рівна:

$$D_{зр} = \frac{N_1 \cdot D_{1зр} + N_2 \cdot D_{2зр}}{N_1 + N_2} = \frac{7 \cdot 1.63 + 6 \cdot 0.33}{7 + 6} = 1.03$$

При відомих, групових і загальних середніх для складових знаходять дисперсію групових середніх щодо загальної середньої. Таку дисперсію називають міжгруповою дисперсією:

$$D_{мжгр} = \frac{\sum N_j (\bar{x}_j - \bar{x})^2}{n},$$

де  $\bar{x}_j$  - групова середня групи  $j$ ;  $N_j$  - об'єм групи  $j$ ;  $\bar{x}$  - загальна середня;  $n$  - об'єм всієї сукупності.

Для попереднього прикладу знайти міжгрупову дисперсію.

Спочатку знайдемо загальну середню:

$$\bar{x} = \frac{2 \cdot 1 + 3 \cdot 4 + 1 \cdot 5 + 1 \cdot 2 + 4 \cdot 3 + 1 \cdot 4}{13} = 2.84$$

Використовуючи обчислені вище величини  $\bar{x}_1 = 2.71$ ,  $\bar{x}_2 = 3.0$ , знайдемо шукану міжгрупову дисперсію:

$$D_{\text{мжгр}} = \frac{7(2.71 - 2.84)^2 + 6(3.0 - 2.84)^2}{13} = 0.011$$

Для дисперсії всієї сукупності вводиться поняття загальної дисперсії.

Загальною дисперсією називають дисперсію значень ознаки всієї сукупності щодо загальної середньої:

$$D_{\text{общ}} = \frac{\sum n_i(x_i - \bar{x})^2}{n},$$

де  $n_j$  - частота значення  $x_j$ ,  $\bar{x}$  - загальна середня;  $n$  — об'єм всієї сукупності.

Приклад 4. Знайти загальну дисперсію за даними приклада 1.

Знайдемо шукану загальну дисперсію, враховуючи, що загальна середня рівна  $\bar{x} = 2.8$

$$D_{\text{заг}} = \frac{2(1 - 2.8)^2 + 4(3 - 2.8)^2 + 1(5 - 2.8)^2 + 1(2 - 2.8)^2 + 4(3 - 2.8)^2 + 1(4 - 2.8)^2}{13} = 1.05$$

Зауваження. Знайдена загальна дисперсія дорівнює сумі внутрішньогрупової і міжгрупової дисперсій.

### 2.2.10. Складання дисперсій

Якщо сукупність складається з декількох груп, то загальна дисперсія дорівнює сумі внутрішньогрупової і міжгрупової дисперсій:

$$D_{\text{заг}} = D_{\text{вн}} + D_{\text{мжгр}}$$

Зауваження. Формула має теоретичне і практичне значення. Якщо в результаті спостережень отримано декілька груп значень ознаки, то для обчислення загальної дисперсії групи в єдину сукупність можна не об'єднувати. З іншого боку, якщо сукупність має великий об'єм, то доцільно розбити її на декілька груп, що полегшує розрахунки.

### 2.2.11. Оцінка генеральної дисперсії по виправленій вибірковій

Нехай з генеральної сукупності в результаті  $n$  незалежних спостережень над деяким параметром  $X$  витягута повторна вибірка об'єму  $n$ :

$$\begin{array}{l} X \quad x_1, x_2, \dots, x_n \\ N \quad n_1, n_2, \dots, n_k \end{array}$$

За цими даними вибірки оцінюють (приблизно знаходять) невідому генеральну дисперсію  $D_{\text{г}}$ . Якщо в якості оцінки генеральної дисперсії прийняти вибіркову дисперсію —  $D_{\text{в}}$ , то ця оцінка приводить до систематичних помилок, даючи занижене значення генеральної дисперсії. Пояснюється це тим, що вибіркова дисперсія є зміщеною оцінкою  $D_{\text{г}}$ , іншими словами, математичне очікування вибіркової дисперсії не дорівнює оцінюваній генеральній дисперсії, а саме:

$$M[D_{\text{в}}] \neq D_{\text{г}}$$



Використовуючи дану формулу, вибіркочну дисперсію «виправляють» так, щоб її математичне очікування дорівнювало генеральній дисперсії. Для цього помножують  $D_T$  на дріб  $\frac{n}{n-1}$ , в результаті отримують «виправлену дисперсію», яку зазвичай позначають через  $s^2$ :

$$s^2 = \frac{n}{n-1} D_g = \frac{n}{n-1} \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_g)^2}{n} = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_g)^2}{n-1}$$

Виправлена дисперсія є незміщеною оцінкою генеральної дисперсії.

Отже, в якості оцінки генеральної дисперсії приймають виправлену дисперсію:

$$s^2 = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_g)^2}{n-1}$$

Для оцінки середнього квадратичного відхилення генеральної сукупності використовують «виправлене» середнє квадратичне відхилення, яке дорівнює квадратному кореню з виправленої дисперсії:

$$s = \sqrt{\frac{\sum_{i=1}^k n_i (x_i - \bar{x}_g)^2}{n-1}}$$

Таким чином,  $s$  не є незміщеною оцінкою, тобто воно є «виправленою».

При достатньо великих значеннях  $n$  об'єму вибірки, вибіркочна і виправлена дисперсія розрізняються мало. На практиці користуються виправленою дисперсією, якщо  $n$  менше  $n < 30$ .

### 2.2.12. Точність оцінки, довірча ймовірність (надійність). Довірчий інтервал

Всі оцінки розглянуті до цих пір були точковими.

Точковою називають оцінку, яка визначається одним числом. Якщо сукупність представлена не великим об'ємом, то точкова оцінка може значно відрізнитися від оцінюваного параметра, тобто приводити до грубих помилок. З цієї причини при невеликому об'ємі вибірки слід користуватися інтервальними оцінками.

*Інтервальною* називають оцінку, яка визначається двома числами—кінцями інтервалу. Інтервальні оцінки пов'язані з такими поняттями як точність і надійність оцінок.

Припустимо, що знайдена за даними вибірки статистична характеристика  $S^*$ , яка використовується в якості оцінки невідомого (теоретичного) параметра  $S$ . Будем вважати  $S$  постійним числом. Безумовно, що  $S^*$  тим точніше оцінює параметр  $S$ , чим менше абсолютна величина різниці  $|S - S^*|$ . Тобто, якщо узяти деяке число  $> \delta$  і припустити, що  $|S - S^*| < \delta$ , то, чим менше прийняте число, тим оцінка  $S^*$  точніша. Таким чином, позитивне число характеризує точність оцінки.

На жаль, статистичними методами не можна категорично стверджувати, що оцінка  $S^*$  задовольняє нерівності  $|S - S^*| < \delta$ ; можна лише говорити про ймовірність, з якою ця нерівність здійснюється, таку ймовірність називають

довірчою ймовірністю або надійністю. Для довірчої ймовірності введено загальноприйняте позначення  $\gamma$ .

Надійністю  $\gamma$  (довірчою ймовірністю) оцінки невідомого параметра  $C$  оцінкою  $C^*$  називають ймовірність, з якою здійснюється нерівність  $|C - C^*| < \delta$ . Надійність оцінки задається, причому в якості  $\gamma$  беруть число, близьке до одиниці. Найчастіше задають надійність таких значень 0,95; 0,99 і 0,999.

Нехай ймовірність того, що  $|C - C^*| < \delta$  рівна  $\gamma$ . Формалізована ця умова така:

$$P(|C - C^*| < \delta) = \gamma$$

Замінивши нерівність  $|C - C^*| < \delta$  рівносильним йому подвійною нерівністю  $-\delta < C - C^* < +\delta$ , отримаємо:

$$P(C^* - \delta < C < C^* + \delta) = \gamma$$

Це співвідношення слід розуміти так:  $(C^* - \delta)$  і  $(C^* + \delta)$  – це числа, які визначають інтервал, в якому з ймовірністю  $\gamma$  знаходиться оцінюваний параметр  $C$ .

Довірчим називають інтервал  $-(C^* - \delta), (C^* + \delta)$ , в якому із заданою надійністю (ймовірністю) –  $\gamma$  знаходиться оцінюваний - невідомий параметр.

Числа  $-(C^* - \delta)$  і  $(C^* + \delta)$  інтервалу  $-(C^* - \delta), (C^* + \delta)$  є випадковими (їх називають довірчими межами). Дійсно, в різних вибірках одержують різні значення  $C^*$ . Отже, від вибірки до вибірки змінюватимуться і кінці довірчого інтервалу, тобто довірчі межі самі є випадковими величинами - функціями від  $C_1, C_2, \dots, C_n$

Оскільки випадковою величиною є не оцінюваний параметр  $C$ , а довірчий інтервал, то правильніше говорити не про ймовірність попадання  $C$  в довірчий інтервал, а про ймовірність того, що довірчий інтервал вміщує в собі  $C$ .

Застосування на практиці методу довірчих інтервалів має свої особливості, які обов'язково необхідно враховувати в практичній роботі, що буде викладено у подальшому.

### **2.2.13. Визначення довірчого інтервалу для оцінки математичного очікування чинника, що має нормальний закон розподілу при відомому $\sigma$**

Припустимо, що деяка ознака  $X$  генеральної сукупності розподілена нормально, причому середнє квадратичне відхилення  $\sigma$  цього розподілу відомо. Потрібно оцінити невідоме математичне очікування  $a$  генеральної сукупності по вибірковій середній  $\bar{x}$ . Знайдемо довірчі інтервали, що покривають параметр  $a$  з надійністю  $\gamma$ .

Якщо розглядати вибіркову середню  $\bar{x}$  як випадкову величину  $\bar{X}$  ( $\bar{x}$  змінюється від вибірки до вибірки) і вибіркові значення ознаки  $x_1, x_2, \dots, x_n$ , як однаково розподілені незалежні випадкові величини  $X_1, X_2, \dots, X_n$  (ці числа також змінюються від вибірки до вибірки), то математичне очікування кожної з цих величин дорівнює  $a$ , а середнє квадратичне відхилення -  $\sigma$ .

Якщо випадкова величина  $X$  розподілена по нормальному закону, то вибіркова середня, знайдена за незалежними спостереженнями, також

розподілена нормально. Числові характеристики такого чинника –  $X$ , що має нормальний закон розподілу, наступні:

$$M(\bar{X}) = a, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

Задаємось умовою, щоб виконувалося співвідношення:

$$P(|\bar{X} - a| < \delta) = \gamma,$$

де  $\gamma$  - задана надійність.

Користуючись інтегральною функцією Лапласа і прирівнявши  $\gamma = 2\Phi\left(\frac{\delta}{\sigma}\right)$ , отримаємо вираз:

$$P(|X - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right),$$

замінивши  $X$  через  $\bar{X}$ , а  $\sigma$  через  $\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$  отримаємо:

$$P(|\bar{X} - a| < \delta) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)$$

Позначимо  $\frac{\delta\sqrt{n}}{\sigma} = t$ , з урахуванням цього можна записати:

$$P(|\bar{X} - a| < \delta) = 2\Phi(t)$$

З передостанньої рівності знайдемо  $\delta$ , тоді останній вираз такий:

$$P(|\bar{X} - a| < t \frac{\sigma}{\sqrt{n}}) = 2\Phi(t)$$

Враховуючи те, що ймовірність здійснення нерівності  $\gamma$  відома, заздалегідь приймається, то останній вираз для оцінки невідомого математичного очікування  $a$  по вибірковому середньому наступний:

$$P(\bar{X} - t \frac{\sigma}{\sqrt{n}} < a < \bar{X} + t \frac{\sigma}{\sqrt{n}}) = 2\Phi(t) = \gamma$$

Суть отриманого співвідношення такий: з надійністю  $\gamma$  можна стверджувати, що довірчий інтервал покриває невідомий параметр  $a$ .

Вираз  $\delta = t \frac{\sigma}{\sqrt{n}}$  є точністю оцінки параметра  $a$ .

У практичній роботі розрахунок здійснюється визначенням значення параметра  $t$  з рівності  $2\Phi(t) = \gamma$ , а  $\Phi(t) = \gamma / 2$ . Залежно від величини  $\gamma / 2$  по таблиці функції Лапласа (додаток 2) знаходять аргумент  $t$ .

Приклад. Випадкова величина  $X$  має нормальне розподілення з відомим середнім квадратичним відхиленням  $\sigma = 2$ . Знайти довірчі інтервали для оцінки невідомого математичного очікування  $a$  по вибірковій середній, якщо об'єм вибірки  $n = 49$  і задана надійність оцінки  $\gamma = 0,95$ .

Рішення. Знайдемо  $t$  з співвідношення  $2\Phi(t) = 0,95$ , отримаємо  $\Phi(t) = 0,475$ . По таблиці (Додаток 2) знаходимо  $t = 1,96$ .

Знайдемо точність оцінки:

$$\delta = t \frac{\sigma}{\sqrt{n}} = \frac{1,96 \cdot 2}{\sqrt{49}} = 0,56$$

Довірчі інтервали такі:

$$(\bar{x} - 0,56; \bar{x} + 0,56)$$

Наприклад, якщо  $\bar{x} = 4,0$ , то довірчий інтервал має наступні довірчі межі:

$$\begin{aligned}\bar{x} - 0,56 &= 4,0 - 0,56 = 3,44; \\ \bar{x} + 0,56 &= 4,0 + 0,56 = 4,56.\end{aligned}$$

Таким чином, значення невідомого математичного очікування  $a$  чинника оцінюється його середнім значенням, яке одержано з даних вибірки і задовольняє нерівності

$$3,02 < a < 4,98.$$

Якщо потрібно оцінити математичне очікування з наперед заданою точністю  $\delta$  і надійністю  $\gamma$ , то можна встановити мінімальний об'єм вибірки, який забезпечить цю точність, при цьому використовується формула розрахунку точності:

$$n = \frac{t^2 \sigma^2}{\delta^2}$$

#### 2.2.14. Довірчі інтервали для оцінки математичного очікування чинника, що має нормальний закон розподілу при невідомому $\sigma$

Якщо ознака  $X$  генеральної сукупності розподілена нормально, причому середнє квадратичне відхилення  $\sigma$  невідомо, то як оцінити невідоме математичне очікування  $a$  за допомогою довірчих інтервалів.

В цьому випадку за даними вибірки можна побудувати випадкову величину, а її можливі значення позначають через  $t$ ,  $t = \frac{\bar{X} - a}{\frac{s}{\sqrt{n}}}$  яка має розподіл

Стюдента з  $k = n - 1$  степенями свободи; тут  $\bar{X}$  — вибіркова середня,  $s$  — «виправлене» середнє квадратичне відхилення,  $n$  — об'єм вибірки.

Якщо виконати відповідні перетворення, отримаємо вираз для знаходження інтервалів, коли середнє квадратичне відхилення невідоме:

$$P\left(\bar{X} - t_\gamma \frac{s}{\sqrt{n}} < a < \bar{X} + t_\gamma \frac{s}{\sqrt{n}}\right) = 2\Phi(t) = \gamma$$

По таблиці (Додаток 3) знаходимо значення  $t_\gamma$ .

#### 2.2.15. Оцінка дійсного значення вимірюваної величини

Проведено  $n$  незалежних рівноточних вимірювань деякої фізичної величини, дійсне значення  $a$  якої невідомо. Результати окремих вимірювань заздалегідь ми не знаємо, тому вважаємо їх випадковими величинами  $X_1, X_2, \dots, X_n$ . Ці величини незалежні (вимірювання незалежні), мають одне і те ж математичне очікування -  $a$ , однакові дисперсії  $s^2$  (якщо вимірювання рівноточні) і розподілені нормально (таке допущення підтверджується досвідом). Для вирішення поставленого питання доцільно використати отримані раніше формули. Іншими словами, дійсне значення вимірюваної величини можна оцінювати по середньому арифметичному результатів окремих вимірювань за допомогою довірчих інтервалів.

Приклад. Проведено 15 незалежних рівноточних вимірювань фізичної величини, по яких знайдене середнє арифметичне  $\bar{x} = 20.0$  і «виправлене» середнє квадратичне відхилення  $s = 2.0$ . Потрібно оцінити дійсне значення вимірюваної величини з надійністю  $\gamma = 0,95$ .

Рішення. Дійсне значення вимірюваної величини дорівнює її математичному очікуванню. Тому рішення зводиться до оцінки математичного очікування -  $a$  (при невідомому  $\sigma$ ) за допомогою довірчого інтервалу:

$$\bar{X} - t_{\gamma} \frac{s}{\sqrt{n}} \leq a \leq \bar{X} + t_{\gamma} \frac{s}{\sqrt{n}} = \gamma$$

Задаємось  $\gamma = 0,95$ .

Користуючись табл. (Додаток 3) при  $\gamma = 0,95$  і  $n = 15$ , знаходимо  $t = 2.15$ .

Знайдемо точність оцінки дійсного, невідомого нам значення оцінюваного чинника:

$$t_{\gamma} \frac{s}{\sqrt{n}} = \frac{2.15 \cdot 2}{\sqrt{16}} = 1.75$$

Знайдемо довірчі межі:

$$\bar{X} - t_{\gamma} \frac{s}{\sqrt{n}} = 20 - 1.75 = 18.25$$

$$\bar{X} + t_{\gamma} \frac{s}{\sqrt{n}} = 20 + 1.75 = 21.75$$

Отже, з надійністю 0,95 дійсне значення досліджуваної величини знаходиться в довірчому інтервалі

$$18.25 < a < 21.75$$

### 2.2.16. Довірчі інтервали для оцінки середнього квадратичного відхилення $\sigma$ чинника, що має нормальний закон розподілу

Припустимо, що ознака  $X$  генеральної сукупності розподілена нормально. Потрібно оцінити невідоме генеральне середнє квадратичне відхилення  $\sigma$  по «виправленому» вибірковому середньому квадратичному відхиленню  $s$ . Знайдемо довірчі інтервали, що покривають параметр  $\sigma$  із заданою надійністю  $\gamma$ . Для вирішення цього питання задаємось такою умовою, щоб виконувалося співвідношення  $P(|\sigma - s| < \delta) = \gamma$  або  $P(s - \delta < \sigma < s + \delta) = \gamma$

Для того, щоб можна було користуватися готовою таблицею, перетворимо подвійну нерівність  $s - \delta < \sigma < s + \delta$  в рівносильну їй.

Позначивши  $\frac{\delta}{s} = q$ , отримаємо  $s(1 - q) < \sigma < s(1 + q)$ .

У отриманому виразі  $s$  розраховується, а  $q$  знаходять по таблиці. Для знаходження  $q$  користуються випадковою величиною, яка розподілена згідно із законом «хі» хіквадрат -  $\chi^2$ :

$$\chi = \frac{s}{\sigma} \sqrt{n-1}$$

де  $n$  — об'єм вибірки.

Приклад. Ознака  $X$  генеральної сукупності розподілена нормально. По сукупності значень об'єму  $n = 15$  знайдене «виправлене» середнє квадратичне відхилення  $s = 0.4$ . Для вказаної умови знайдемо довірчий інтервал, що покриває генеральне середнє квадратичне відхилення  $\sigma$  з надійністю  $\gamma = 0,95$ .

По таблиці (Додаток 4) при  $\gamma = 0,95$  і  $n = 15$  знаходимо  $q = 0,46$ .

Довірчий інтервал буде таким:

$$0.4(1 - 0.46) < \sigma < 0.4(1 + 0.46) \text{ або } 0.216 < \sigma < 0.584$$

### 2.2.17. Оцінка точності вимірювань

У теорії помилок точність вимірювань (точність приладу) характеризують за допомогою середнього квадратичного відхилення  $\sigma$  випадкових помилок вимірювань. Для оцінки  $\sigma$  використовують «виправлене» середнє квадратичне відхилення  $s$ .

Оскільки результати вимірювань взаємно незалежні і мають одне і те ж математичне очікування (дійсне значення вимірюваної величини) і однакову дисперсію (у разі рівноточних вимірювань), то теорія, що викладена в попередньому параграфі, годиться для оцінки вимірювань..

Приклад. По 25 рівноточним вимірюванням знайдене «виправлене» середнє квадратичне відхилення  $s = 0,2$ . Знайти точність вимірювань з надійністю  $\gamma = 0,99$ .

Точність вимірювань характеризується середнім квадратичним відхиленням  $\sigma$  випадкових помилок, тому завдання зводиться до відшукування довірчого інтервалу. По таблиці (Додаток 4) по  $\gamma = 0,99$  і  $n = 25$  знайдемо  $q = 0.49$ .

Шуканий довірчий інтервал такий:

$$0.2(1 - 0.49) < \sigma < 0.2(1 + 0.49) \text{ або } 0.102 < \sigma < 0.298$$

### 2.2.18. Інші характеристики варіаційного ряду

Окрім вибіркової середньої і вибіркової дисперсії застосовуються і інші характеристики варіаційного ряду.

Мода. Модою  $M_0$  називають варіанту, яка має найбільшу частоту. Наприклад, для ряду

Варіанта	1	2	5	7
Частота	5	1	20	6

мода дорівнює 5.

Медіана. Медіаною  $m_e$  називають варіанту, яка ділить варіаційний ряд на дві частини, рівні по числу варіант.

Наприклад, для ряду 1 3 5 7 9 - медіана дорівнює 5;

для ряду 1 3 5 7 9 11 - медіана дорівнює 5,5.

Розмах. Розмахом варіювання  $R$  називають різницю між найбільшою і найменшою варіантами:

$$R = x_{\max} - x_{\min}$$

Наприклад, для ряду, 1 3 4 5 6 7 розмах дорівнює  $7 - 1 = 6$ .

Розмах є простою характеристикою розсіювання варіаційного ряду.

Середнє абсолютне відхилення. Середнім абсолютним відхиленням  $\theta$  називають середнє арифметичне абсолютних відхилень:

$$\theta = \frac{\sum n_i |x_i - \bar{x}_e|}{\sum n_i}$$

Наприклад, для ряду

$x_i$ ,	1	3	5	7
$n_i$	2	4	4	1

маємо

$$\bar{x}_e = \frac{2 \cdot 1 + 4 \cdot 3 + 4 \cdot 5 + 1 \cdot 7}{2 + 4 + 4 + 1} = 3.727$$

$$\theta = \frac{2|1 - 3.73| + 4|3 - 3.73| + 4|5 - 3.73| + 1|7 - 3.73|}{2 + 4 + 4 + 1} = 0.89$$

Середнє абсолютне відхилення служить для характеристики розсіяння варіаційного ряду.

Коефіцієнт варіації. Коефіцієнтом варіації  $V$  називають представлене у відсотках відношення вибіркового середнього квадратичного відхилення до вибіркової середньої:

$$V = \frac{\sigma_{\hat{a}}}{\bar{x}_{\hat{a}}} 100 \%$$

Коефіцієнт варіації служить для порівняння величин розсіяння двох варіаційних рядів. Той з рядів має більше розсіяння, у якого коефіцієнт варіації більший.

### 2.2.19. Закономірності розподілу дискретних випадкових величин

У практичній роботі з дискретними випадковими величинами закон розподілу яких невідомий, часто потрібно знайти параметри і вигляд закону розподілу. Вирішують таку задачу, користуючись експериментальними статистичними даними, причому зручно представити їх у вигляді закону розподілу.

### 2.2.20. Емпіричні і вирівнюючі частоти

Припустимо, що параметр  $X$ , який вивчається, набував значень з відповідними частотами, закон розподілу якого в загальному -табличному вигляді, наступний:

$X$	$x_1$	$x_2, \dots, x_n$
$N$	$n_1$	$n_2, \dots, n_k$

Емпіричними частотами називають фактично спостережувані частоти  $n_i$ .

Аналізуючи складові табличного представлення роблять висновок про вигляд розподілення чинника. Наприклад, маємо табличний закон розподілу:

$X$	$x_1 = 1$	$x_2 = 3$	$x_3 = 5$	$x_4 = 7$	$x_5 = 9$
$N$	$n_1 = 2$	$n_2 = 6$	$n_3 = 10$	$n_4 = 5$	$n_5 = 1$

Крайні варіанти  $x_1$  і  $x_5$  мають малі частоти, від яких до центральної варіанти  $x_3$  частоти зростають, це дає підставу припустити, що чинник - параметр має нормальний закон розподілу.

Якби всі варіанти мали приблизно рівні частоти, то можна припускати, що такий параметр має рівномірний закон розподілу.

Частоти, отримані в результаті спостережень, називають емпіричними частотами.

Якщо прийняте припущення вважати дійсним, то звичайно існують параметри, за допомогою яких можна представити, яким є теоретичний закон розподілу, для цього необхідно знайти теоретичні частоти.

Частоти, які отримують в результаті розрахунку, називають вирівнюючими або теоретичними.

Вирівнюючі частоти знаходять по виразу:

$$n'_i = nP_i,$$

де  $n$  - число випробувань,  $P_i$  - ймовірність спостережуваного значення ознаки, яка обчислюється при допущенні, що вона має передбачуваний розподіл.

Таким чином, вирівнююча частота значення ознаки дискретного розподілу дорівнює добутку числа випробувань на ймовірність цього спостережуваного значення.

Приклад. В результаті експерименту, що складається з  $n = 153$  випробувань, в кожному з яких реєструвалося число появ деякого параметра, отримано наступний емпіричний розподіл:

сп. знач. $x_i$	0	1	2	3	4	5	6	7
емп. частота $n_i$	20	65	30	22	12	2	1	1.

По величині емпіричних частот і їх розташуванню роблять висновок, що параметр, який вивчається (аналізується) розподілений за законом Пуассона. Потрібно знайти вирівнюючі частоти -  $n'_i$

Рішення. Відомо, що розподіл Пуассона визначається параметром  $\lambda$ , який дорівнює математичному очікуванню цього розподілу. В якості оцінки математичного очікування використовуємо вибірккову середню, це означає, що для оцінки  $\lambda$  також можна прийняти вибірккову середню. Знайдемо вибірккову середню по формулі:

$$\bar{x}_e = \sum_{i=1}^n \frac{n_i x_i}{n} = 1.7$$

Приймаємо  $\lambda = 1.7$

По формулі Пуассона визначаємо ймовірність частот:

$$P_{153}(k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!} = \frac{1.7^k \cdot e^{-1.7}}{k!},$$

де  $k$  – чисельне значення відповідної варіанти  $x_i$ .

$$P_{153}(0) = P_{153}(1) = P_{153}(2) = P_{153}(3) = P_{153}(4) = P_{153}(5) = P_{153}(6) = P_{153}(7) =$$

Користуючись цією формулою, знайдемо ймовірність  $P_{153}(x_i)$  при  $\lambda = 1.7$ ,  $x_i = 0, 1, 2, 3, 4, 5, 6, 7$  (для простоти запису індекс 153 далі опущений):  $P(0) = 0.12$ ,  $P(1) = 0.44$ ,  $P(2) = 0.18$ ,  $P(3) = 0.13$ ,  $P(4) = 0.065$ ,  $P(5) = 0.006$ ,  $P(6) = 0.0000$ ,  $P(7) = 0.0000$ .

Знайдемо вирівнюючі частоти по виразу:

$$n'_i = nP_i$$



Закон розподілу з емпіричними і вирівнюючими частотами наступний:

сп. знач.	$x_i$	0	1	2	3	4	5	6	7
емп. частота	$n_i$	20	65	30	22	12	2	1	1.
теорет. част	$n_i'$	18	68	28	20	10	1	0	0

Невелика розбіжність емпіричних і вирівнюючих частот підтверджує припущення, що даний розподіл підпорядкований закону Пуассона.

### 2.2.21 Безперервний розподіл

У разі безперервного розподілу ймовірність окремих можливих значень дорівнює нулю. Тому весь інтервал можливих значень ділять на  $k$  непересічних інтервалів і обчислюють ймовірність  $P_i$  попадання параметра в частковий інтервал, а потім розрахунок здійснюється так, як і для дискретного розподілу.

Зокрема, якщо є підстави припустити, що випадкова величина  $X$  розподілена нормально, то вирівнюючі частоти можуть бути знайдені по формулі:

$$n_i' = \frac{nh}{\sigma_{\epsilon}} \varphi(u_i),$$

де  $n$  — число випробувань,  $h$  — довжина часткового інтервалу,  $\sigma$  - вибіркове середнє квадратичне відхилення,  $u_i$  - умовна варіанта

$$u_i = \frac{x_i - \bar{x}_{\epsilon}}{\sigma_{\epsilon}},$$

де  $x_i$  - середина  $i$ -го часткового інтервалу

### 2.2.22. Побудова кривої нормального закону розподілу чинника з використанням практичних даних

Спосіб побудови нормальної кривої за даними спостережень полягає в наступному.

1) знаходять  $\bar{x}_{\epsilon}$  та  $\sigma_{\epsilon}$ ;

2) знаходять ординати (вирівнюючі частоти) теоретичної кривої по формулі:

$$n_i' = \frac{nh}{\sigma_{\epsilon}} \varphi(u_i)$$

Значення  $\varphi(u_i)$  знаходять по таблиці, додаток 1.

3) відкладають крапки  $(x_i, y_i)$  в прямокутній системі координат і з'єднують їх плавною кривою.

Близькість вирівнюючих частот до спостережуваних підтверджує правильність допущення про те, що обстежувана ознака розподілена нормально або не підтверджує.

Приклад. Побудувати нормальну криву по даному розподілу:

варіанту $x_i$	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5
частота $n_i$ ,	3	7	16	37	53	42	15	5	2.

Рішення. Знайдемо  $\bar{x}_{\epsilon} = 3.01$  і  $\sigma = 0.83$ .

Розрахунки доцільно виконувати, скориставшись відповідною таблицею, дивись табл. 10.

Таблиця 10

$x_i$	$n_i$	$x_i - \bar{x}_e$	$u_i = \frac{x_i - \bar{x}_e}{\sigma_e}$	$\varphi(u_i)$	$n_i'$
1.5	3	- 1.51	1.81	0.07	7.59
2.0	7	- 1.01	1.21	0.19	20.6
2.5	16	- 0.51	0.61	0.33	35.78
3.0	37	- 0.01	0.01	0.39	42.28
3.5	53	0.49	0.59	0.33	35.78
4.0	42	0.99	1.19	0.19	20.6
4.5	15	1.49	1.79	0.08	8.67
5.0	5	1.99	2.39	0.02	2.16
5.5	2	2.49	3.0	0.003	0.32
3.01	180				

Таким чином одержимо закон розподілу з емпіричними та вирівнюючи ми частотами.

$x_i$	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5
$n_i$	3	7	16	37	53	42	15	5	2
$n_i'$	7	20	35	42	35	20	8	2	0

На рис. 6 побудовані нормальна (теоретична) крива по вирівнюючих частотах - штрихпунктирна і полігон спостережуваних частот. Порівняння графіків показує, що теоретична крива задовільно відображає дані спостережень.

Для оцінки твердження по співпаданню частот використовують певні оцінки та критерії згоди.

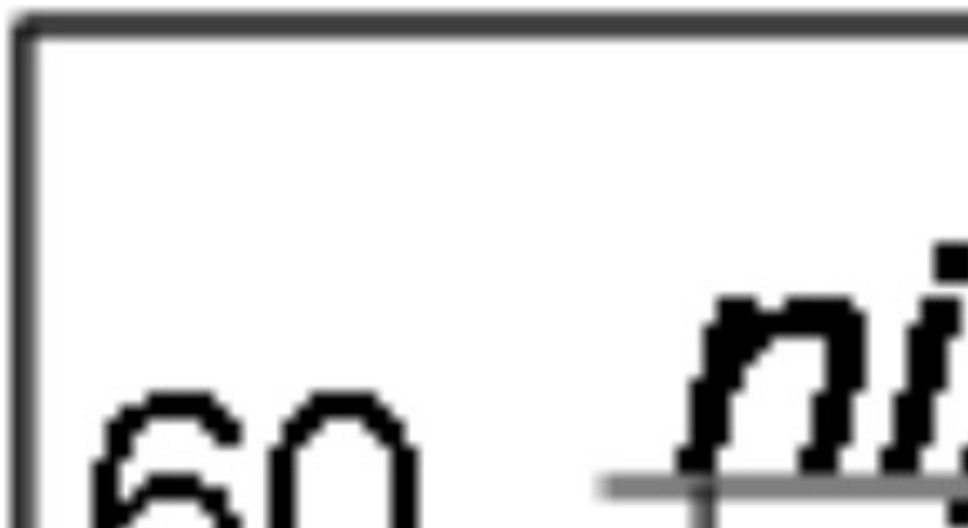


Рис. 6. Графічне представлення теоретичного та практичного розподілу випадкової величини

### 2.2.23. Оцінка відхилення емпіричного розподілу від нормального. Асиметрія і ексцес

Для оцінки відхилення емпіричного розподілу від нормального використовують характеристики - асиметрію і ексцес. Визначення цих характеристик аналогічне визначенням асиметрії і ексцесу теоретичного розподілу.

Асиметрія емпіричного розподілу визначається рівністю:

$$a_s = \frac{m_3}{\sigma^3},$$

де  $m_3$  - центральний емпіричний момент третього порядку.

Ексцес емпіричного розподілу визначається рівністю:

$$a_k = \frac{m_4}{\sigma^4} - 3$$

Обчислення моментів. Розрізняють звичайні емпіричні моменти і центральні.

Звичайним емпіричним моментом  $k$  - того порядку називають вираз:

$$M'_k = \sum \frac{n_i(x_i - c)^k}{n},$$

де  $x_i$  - спостережувана варіанту (число),  $n_i$  - частота, з якою зустрічається дане число,  $n$  - об'єм чисел,  $c$  - постійне число (початок відліку),  $k$  - номер моменту (порядок моменту).

Серед звичайних емпіричних моментів особливим є початковий емпіричний момент, який обчислюють за тією ж формулою, тільки  $c = 0$ .

Центральним емпіричним моментом  $k$  - того порядку називають звичайний момент  $k$  - того порядку, в якому  $c = \bar{x}_e$

$$m_k = \sum \frac{n_i(x_i - \bar{x}_e)^k}{n}$$

Надалі центральні емпіричні моменти виражають через звичайні:

$$m_1 = M'_1$$

$$m_2 = M'_2 - (M'_1)^2$$

$$m_3 = M'_3 - 3M'_2M'_1 + 2(M'_1)^3$$

$$m_4 = M'_4 - 4M'_3M'_1 + 6M'_2(M'_1)^2 - 3(M'_1)^4$$

Приклад. Знайти асиметрію і ексцес емпіричного розподілу:

варіанта	1,2	1,4	1,6	1,8	2,0	2,2	2,4
частота	2	3	8	13	25	20	10

$$\bar{x}_e = 1.98, \sigma = 0.28$$

$$M_1 = 1.98, M_2 = 4.021, M_3 = 8.28, M_4 = 17.33.$$

Розрахуємо центральні моменти

$$m_1 = 1.98, m_2 = 0.1, m_3 = 19.68, m_4 = 0.23$$

Асиметрія

$$a_s = \frac{19.68}{0.008} = 2460.$$

На рис. 7 представлено графічне зображення ліво- та правосторонньої асиметрії чинника, що має нормальний закон розподілу



Рис. 7. Правостороння та лівостороння асиметрії чинника, що має нормальний закон розподілу

На рис. 8 представлено графічне зображення ексцесу чинника, що має нормальний закон розподілу

$$\text{Ексцес } a_k = \frac{0.23}{0.0002} - 3 = 1150.$$

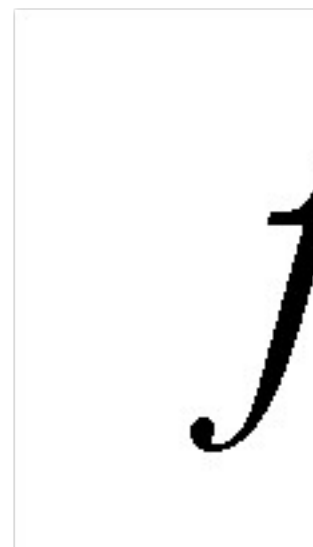


Рис. 8. Види ексцесу чинника, що має нормальний закон розподілу

### 2.3. Основи кореляційного аналізу

Кореляційний аналіз - один з доступних, ефективних і простих математичних методів для розробки математичних моделей апаратів і процесів.

Використання даного методу , отримання математичних моделей в гірничо – рудній і збагачувальній технологіях має свою специфіку в тому плані, що технологічні процеси, які підлягають математичному моделюванню, є багаточинниковими, більшість з яких технічними засобами не контролюються з

тієї причини, що до теперішнього часу необхідні засоби не розроблені і значення цих чинників з погляду їх контролю є випадковими. Даний метод створений для роботи саме з випадковими числами. На жаль, при підготовці фахівців - технологів для рудної і вугільної галузей даному методу не відводиться належної уваги.

Використання методу для розробки математичних моделей апаратів і процесів вимагає наявності статистичного матеріалу. У гірничому виробництві і в збагаченні практичних даних накопичена значна кількість, проте, з цим матеріалом потрібно уміти працювати. У даному розділі розглядаються основні, на наш погляд, питання.

Це, як правильно зібрати дані, щоб ці дані були статистичними. Основним при цьому є вимога до репрезентативності.

Як правильно обробити статистичні дані і їх представити, якими оцінками при цьому користуватися.

Як щонайкраще використовувати метод отримання рівняння – математичні моделі.

Як перевірити адекватність отриманої моделі об'єкту моделювання, який з критеріїв згоди слід застосовувати.

### **2.3.1. Функціональна, статистична і кореляційна залежності**

У багатьох задачах потрібно встановити і оцінити залежність випадкової величини від однієї або декількох інших випадкових або не випадкових величин. Розглянемо спочатку залежність  $Y$  від однієї випадкової величини  $X$ .

Дві випадкові величини можуть бути зв'язані функціональною залежністю, або залежністю іншого роду, яку називають статистичною, або бути незалежними.

За функціональну вважають таку залежність між величинами, при якій є однозначний математичний зв'язок.

Наприклад,  $y = x^2$ .

Якщо обидві величини або одна з них піддаються дії деяких випадкових чинників, то таку залежність називають статистичною.

Наприклад,  $Y$  залежить від випадкових чинників.

$A_1, A_2, A_3$ , а  $X$  від  $A_1, A_2, A_4$ , то між  $Y$  і  $X$  є статистична залежність, ці величини мають загальні випадкові чинники  $A_1$  і  $A_2$ .

Статистичною називають залежність, при якій зміна однієї з величин викликає зміну розподілу іншої.

Якщо при зміні однієї з величин змінюється середнє значення іншої, то таку статистичну залежність називають кореляційною.

Розглянемо чинники  $\gamma$  — вихід продукту,  $A^d$  — якість початкового матеріалу, що поступає в схему збагачення. На паралельних ідентичних збагачувальних технологічних секціях отримують різні виходи, тобто  $\gamma$  не є функцією від  $A^d$ . Це пояснюється впливом випадкових чинників, які властиві

даній технології. Проте, як показує досвід, вихід продукту є функцією від початкової зольності, тобто  $y$  пов'язаний з  $A^d$  кореляційною залежністю.

### 2.3.2. Умовні середні. Кореляційна залежність

Розглянемо визначення кореляційної залежності детальніше, для чого введемо поняття умовної середньої.

Припустимо, що вивчається зв'язок між випадковою величиною  $Y$  і випадковою величиною  $X$ , які задані табличним законом розподілу, дивись табл. 11.

Таблиця 11

Y X	$Y_1=1$	$Y_2=2$	$y_3=3$	$n_x$
$X_1=2$	1	2	1	$n_{x_2}=4$
$X_2=4$	2	3	1	$n_{x_4}=6$
$X_3=6$	1	3	2	$n_{x_6}=6$
$n_y$	$n_{y_1}=4$	$n_{y_2}=8$	$n_{y_3}=4$	$n_y=16$ $n_x=16$

Для знаходження умовної середньої  $\bar{Y}(y_i/x=2)$  запишемо умовний закон розподілу.

При  $x_1=2$  величина  $Y$  прийняла закон розподілу – значення – табл. 12:

Таблиця 12

Y	$Y_1=1$	$y_2=2$	$y_3=3$	$n_x$
$n(y_i/x_1=2)$	1	2	1	$n_{x_2}=4$

Знайдемо середнє арифметичне величини  $Y$ :

$$\bar{y}(x_1=2) = \frac{\sum_{i=1}^n n_i \cdot y_i}{n_{x_2}} = \frac{1 \cdot 1 + 2 \cdot 2 + 1 \cdot 3}{1 + 2 + 1} = 2.$$

Якщо при  $x_2=4$  величина  $Y$  набула значень – табл. 13:

Таблиця 13

Y	$Y_1=1$	$y_2=2$	$y_3=3$	$n_x$
$n(y_i/x_1=4)$	2	3	1	$n_{x_2}=6$

Знайдемо середнє арифметичне величини  $Y$ :

$$\bar{y}(x_1=4) = \frac{\sum_{i=1}^n n_i \cdot y_i}{n_{x_4}} = \frac{2 \cdot 1 + 3 \cdot 2 + 1 \cdot 3}{2 + 3 + 1} = 1.83.$$

Якщо при  $x_3=6$  величина  $Y$  набула значень – табл. 14.

Таблиця 14

Y	$Y_1=1$	$y_2=2$	$y_3=3$	$n_x$
$n(y_i/x_1=6)$	1	3	2	$n_{x_3}=6$

Знайдемо середнє арифметичне величини  $Y$ :

$$\bar{y}(x_1 = 6) = \frac{\sum_{i=1}^n n_i \cdot y_i}{n_{x_6}} = \frac{1 \times 1 + 3 \times 2 + 2 \times 3}{1 + 3 + 2} = 2.16.$$

Число  $\bar{y}_{xi}$  називають умовним середнім; риска над буквою  $y$  служить позначенням середнього арифметичного, а значення селектора  $i$  вказує, що розглядаються ті значення  $Y$ , які відповідають  $X_i$ .

Умовним середнім  $\bar{y}_{xi}$  називають середнє арифметичне значень  $Y$ , відповідних значенню  $X = x_i$ .

Якщо кожному значенню  $x$  відповідає одне значення умовної середньої, то очевидно умовна середня є функція від  $x$ ; в цьому випадку говорять, що випадкова величина  $Y$  залежить від  $X$  кореляційно.

Кореляційною залежністю  $Y$  від  $X$  називають функціональну залежність умовної середньої  $\bar{y}_{xi}$  від  $x$ :

$$\bar{y}_{xi} = f(x)$$

Дане рівняння називають рівнянням регресії  $Y$  на  $X$ ; функцію  $f(x)$  називають функцією регресії  $Y$  на  $X$ , а її графік — лінією регресії  $Y$  на  $X$ .

Вище приведене також справедливо, якщо розглядати залежність  $X$  від  $Y$

### 2.3.3. Основні задачі теорії кореляції

Першою вимогою теорії кореляції являється те, щоб правильно зібрати і представити статистичні дані.

Друга задача - встановити форму кореляційного зв'язку, тобто вигляд функції регресії (лінійну, нелінійну, квадратичну показову і т. д.). Найчастіше функції регресії виявляються лінійними. Якщо обидві функції регресії  $f(x)$  і  $f(y)$  лінійні, то кореляцію називають лінійною; інакше — нелінійною. При лінійній кореляції обидві лінії регресії є прямими лініями.

Третя задача — оцінити тісноту (силу - величину) кореляційного зв'язку. Величина зв'язку між  $Y$  і  $X$  оцінюється по величині розсіювання значень  $Y$  навколо умовного середнього  $\bar{y}_{xi}$  і навпаки. Велике розсіювання свідчить про слабку залежність  $Y$  від  $X$ , або взагалі про її відсутність. Мале розсіювання вказує наявність достатньо сильної залежності – можливо, навіть, що  $Y$  і  $X$  зв'язані функціонально.

Тіснота кореляційного зв'язку оцінюється коефіцієнтом кореляції -  $r$ , якщо зв'язок лінійний і кореляційним відношенням -  $\eta$ , якщо зв'язок не лінійний.

Основною задачею теорії кореляції є розрахунок коефіцієнтів рівняння регресії.

### 2.3.4. Розрахунок параметрів вибіркового рівняння прямої лінії регресії

Розрахунок виконують за умови, якщо правильно зібрані статистичні дані, тобто числа чинників і відповідним чином представлені, встановлено вигляд зв'язку між чинниками, то після цього розраховують складові закономірностей.

Найбільш ефективним методом для розрахунку параметрів рівняння регресії є «метод найменших квадратів». Залежно від того, яким чином представлені статистичні дані, накладаються деякі специфічні особливості використання методу. Розрізняють використання методу, тобто розрахунку коефіцієнтів рівняння регресії за незгрупованими даними і за згрупованими даними.

### 2.3.5. Розрахунок параметрів вибіркового рівняння прямої лінії регресії за не згрупованими даними

Незгруповані дані це такі, які представлені відповідними парами чисел і які не мають повторень.

Припустимо, що закономірність між параметрами  $X$  і  $Y$  лінійна. В цьому випадку обидві лінії регресії будуть прямими.

Вигляд закономірності визначається з попередньо побудованого кореляційного поля.

Для розрахунку коефіцієнтів рівнянь цих прямих користуються статистичними даними - вибіркою, яку представляють у вигляді статистичного розподілу.

$$\begin{array}{l} X \quad x_1 \quad x_2 \dots x_n \\ Y \quad y_1 \quad y_2 \dots y_n \end{array}$$

Оскільки спостережувані пари чисел  $(x_i, y_i)$  розподілу можна розглядати як випадкову вибірку з генеральної сукупності  $(X, Y)$ , то всі величини і рівняння, знайдені за цими даними, називають вибірковими.

Знайти значення коефіцієнтів рівняння прямої лінії регресії  $Y$  на  $X$  можна навпаки.

При не згрупованих даних значення  $x$  ознаки  $X$  і відповідні їм значення  $y$  параметра  $Y$  спостерігалися по одному разу, тому використовується рівняння не для умовної середньої -  $\bar{y}_x = ax + b$ , а рівняння -  $Y = ax + b$ .

Кутовий коефіцієнт -  $a$  прямої лінії регресії  $Y$  на  $X$  прийнято називати вибірковим коефіцієнтом регресії  $Y$  на  $X$  і позначати через  $\rho_{yx}$ .

Тобто, шукатимемо значення коефіцієнтів вибіркового рівняння прямої лінії регресії  $Y$  на  $X$ :

$$Y = \rho_{yx}x + b$$

Відшукання значень коефіцієнтів виконуємо за умови, коефіцієнти  $\rho_{yx}$  і  $b$  повинні бути такими, щоб крапки для всієї сукупності  $(x_1, y_1)$   $(x_2, y_2)$ , ...,  $(x_n, y_n)$ , побудовані на відповідному полі за даними вибірки  $y_i$  на площині  $ХОУ$ , якомога ближче лежали біля прямої, побудованої по прийнятому - теоретичному рівнянню -  $Y = \rho_{yx}x + b$

Для того, щоб реалізувати поставлену вище умову, досить підібрати коефіцієнти  $\rho_{yx}$  і  $b$  такими, щоб сума квадратів відхилень між  $Y_i$  і  $y_i$  була мінімальною:

$$\sum_{i=1}^n (\hat{O}_i - o_i)^2 \rightarrow \text{мін.}$$

В цьому полягає суть методу найменших квадратів, що забезпечує одержання оптимального результату при використанні вказаного методу.



Оскільки кожне відхилення  $(Y_i - y_i)$  залежить від відшукуваних коефіцієнтів  $\rho_{yx}$  і  $b$ , то сума квадратів відхилень є функція  $F$  цих параметрів (опустимо індекси  $x$  у при коефіцієнті -  $\rho$ ). Даний висновок є у вигляді рівняння:

$$F(\rho, b) = \sum_{i=1}^n (Y_i - y_i)^2$$

Якщо в цей вираз замість  $Y_i$  підставити його значення, то отримаємо:

$$F(\rho, b) = \sum_{i=1}^n (\rho x_i + b - y_i)^2$$

Для того, щоб виконалася поставлена умова - відшукування мінімуму квадратів відхилень, знайдемо приватні похідні від останнього виразу і прирівняємо їх нулю, в результаті отримаємо систему:

$$\left. \begin{aligned} \frac{\partial F}{\partial \rho} &= 2 \sum_{i=1}^n (\rho x_i + b - y_i) x_i = 0 \\ \frac{\partial F}{\partial b} &= 2 \sum_{i=1}^n (\rho x_i + b - y_i) = 0 \end{aligned} \right\}$$

Виконавши елементарні перетворення, знайдемо шукані коефіцієнти  $\rho_{yx}$  і  $b$ :

$$\rho_{yx} = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$b = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}$$

Аналогічно можна знайти вибіркове рівняння прямої лінії регресії  $X$  на  $Y$ .

Приклад. Знайти коефіцієнти вибіркового рівняння прямої лінії регресії  $Y$  на  $X$  за статистичними даними

X	1.0	1,5	2,0	2,5	3,0;
Y	1,1	1,3	1,5	1,7	1,9

Для зручності обчислення представимо у вигляді таблиці – табл. 15, в практичній роботі краще скористатися готовими комп'ютерними програмами або розрахунок провести іншим чином за допомогою комп'ютерних технологій.

Таблиця 15

$x_i$	$y_i$	$x_i^2$	$x_i \cdot y_i$
1.0	1.1	1.0	1.1
1.5	1.3	2.25	1.95
2.0	1.5	4.0	3.0
2.5	1.7	6.25	4.25
3.0	1.9	9.0	5.7
$\sum x_i = 10.0$	$\sum y_i = 7.5$	$\sum x_i^2 = 22.0$	$\sum x_i \cdot y_i = 17.0$

$$\rho_{yx} = \frac{5 \cdot 17 - 10 \cdot 7.5}{5 \cdot 22 - 10^2} = 1.0$$

$$b = \frac{22 \cdot 7.5 - 10 \cdot 17}{5 \cdot 22 - 10^2} = -0.5$$

Шукане рівняння регресії з урахуванням значень знайдених коефіцієнтів:

$$Y = 1.0x - 0.5.$$

Для того, щоб отримати уявлення, наскільки добре обчислені по цьому рівнянню значення  $Y_i$  узгоджуються із спостережуваними значеннями  $y_i$ , знайдемо відхилення  $Y_i - y_i$ . Для наочності і зручності аналізу результати обчислень представимо у вигляді табл. 16.

Таблиця 16

$x_i$	$Y_i$	$y_i$	$Y_i - y_i$
1.0	0.5	1.1	- 0.6
1.5	1.0	1.3	-0.3
2.0	1.5	1.5	0.0
2.5	2.0	1.7	0.3
3.0	2.5	1.9	0.4

Для того, щоб оцінити, наскільки знайдене рівняння правильно відображає всю генеральну сукупність, необхідно скористатися відповідними критеріями згоди. Як це здійснюється, буде показано далі.

### 2.3.6. Відшукування параметрів вибіркового рівняння прямої лінії регресії за згрупованими даними

При великому числі статистичних даних, як окремі значення кожного з параметрів, так і їх пари, можуть зустрічатися по кілька разів. Тому дані зводять у групи і записують у вигляді таблиці, яку називають кореляційною – табл. 17.

Таблиця 17.

$Y$ $x_i$	$y_i=1$	$y_i=2$	$y_i=3$	$y_i=4$	$n_y$
$Y_1=2$	1	2	3	1	7
$y_2=4$	-	1	3	1	5
$y_3=6$	2	3	1	-	6
$n_x$	3	6	7	2	$n=18$

Така таблиця має особливість. У першому рядку таблиці вказані значення параметра  $X$ , а в першому стовпці - значення параметра  $Y$ . На перетині рядків і стовпців записуються частоти  $n_{xy}$  спостережуваних пар значень параметрів. Наприклад, частота 3 вказує, що пара чисел (6; 2) спостерігалася 3 рази. У останньому стовпці записані суми частот рядків. У останньому рядку записані суми частот стовпців. У клітинці, розташованій в нижньому правому кутку таблиці, є сума всіх частот (загальне число всіх спостережень  $n$ ).

Для визначення коефіцієнтів рівняння прямої лінії регресії  $Y$  на  $X$  за згрупованими даними скористаємося отриманою раніше системою рівнянь:

$$\left. \begin{aligned} (\sum x^2) \times \rho_{xy} + (\sum x) \times b &= \sum xy \\ (\sum x) \times \rho_{xy} + n \times b &= \sum y \end{aligned} \right\}$$

Перетворимо дану систему, скориставшись тотожностями:

$$\sum x = n\bar{x}; \text{ звідки}$$

$$\begin{aligned} \sum y &= n\bar{y}; \text{ звідки} \\ \sum x^2 &= n\bar{x}^2; \text{ звідки} \\ \sum xy &= \sum n_{xy}xy \end{aligned}$$

Підставивши праві частини тотожностей в попередню систему і скоротивши обидві частини другого рівняння на  $n$ , отримаємо:

$$\left. \begin{aligned} (n\bar{x}^2) \times \rho_{yx} + (n\bar{x}) \times b &= \sum n_{xy}xy \\ (\bar{x}) \times \rho_{yx} + b &= \bar{y} \end{aligned} \right\}$$

Вирішивши дану систему, знайдемо значення коефіцієнтів  $\rho_{yx}$  і  $b$ , і, отже, шукане рівняння в явному вигляді:

$$\bar{y}_x = \rho_{yx}x + b.$$

Коефіцієнти  $\rho_{yx}$  і  $b$  можна розрахувати за допомогою коефіцієнта кореляції, попередньо виконавши відповідні перетворення, а саме, друге рівняння системи визначають відносно  $-b$ , праву частину якого включають в останнє рівняння, з якого знаходять  $\rho_{yx}$ :

$$\rho_{yx} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n(\bar{x}^2 - (\bar{x})^2)}$$

Помножимо обидві частини рівності на дріб  $-\frac{\sigma_x}{\sigma_y}$  і враховуючи те, що  $x^2 - (\bar{x})^2 = \sigma_x^2$  отримаємо:

$$\rho_{yx} \cdot \frac{\sigma_x}{\sigma_y} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\sigma_x\sigma_y}$$

Права частина рівності є не що інше, як коефіцієнт кореляції, тому:

$$\rho_{yx} \cdot \frac{\sigma_x}{\sigma_y} = r_{xy}$$

отже:

$$\rho_{yx} = r_{xy} \frac{\sigma_y}{\sigma_x}.$$

З урахуванням вище розглянутих положень остаточне рівняння:

$$\bar{y}_x = r_{xy} \frac{\sigma_y}{\sigma_x} (x_i - \bar{x}) + \bar{y}$$

В даному випадку знаходиться рівняння для вибірових середніх. Аналогічно знаходиться рівняння параметра  $X$  на  $Y$ .

### 2.3.7. Властивості вибіркового коефіцієнта кореляції

Коефіцієнт кореляції може приймати наступні значення:

$$-1.0 \leq r_b \leq 1.0$$

Якщо вибіровий коефіцієнт кореляції дорівнює нулю і вибірові лінії регресії — прямі, то  $X$  і  $Y$  не зв'язані лінійною кореляційною залежністю, у той же час вони можуть бути зв'язані нелінійною кореляційною залежністю.

Якщо абсолютна величина вибіркового коефіцієнта кореляції дорівнює одиниці, то чинники зв'язані лінійною функціональною залежністю.

Із зростанням абсолютної величини вибіркового коефіцієнта кореляції лінійна кореляційна залежність стає тіснішою і при  $|r_b| = 1$  переходить у функціональну залежність.

Вибірковий коефіцієнт кореляції дорівнює середньому геометричному вибірових коефіцієнтів регресії:

$$\rho_{yx} \rho_{xy} = r_b^2.$$

Звідси:

$$r_b = \pm \sqrt{\rho_{yx} \rho_{xy}}.$$

### 2.3.8. Вибіркове кореляційне відношення

Тісноту лінійного кореляційного зв'язку між параметрами оцінюють вибіровим коефіцієнтом кореляції. Для оцінки тісноти нелінійного кореляційного зв'язку між чинниками  $X$  і  $Y$  необхідно використовувати вибірове кореляційне відношення –  $\eta_{xy}$ ;

Вибірковим кореляційним відношенням  $Y$  на  $X$  або  $X$  на  $Y$  називають відношення міжгрупового середнього квадратичного відхилення до загального середнього квадратичного відхилення відповідного параметра:

$$\eta_{yx} = \frac{\sigma_{i\ sae}}{\sigma_{pfi}},$$

або інакше:

$$\eta_{yx} = \frac{\sigma_{\bar{y}_x}}{\sigma_y}.$$

Якщо  $\eta_{xy} = 0$ , то параметри  $Y$  і  $X$  кореляційною залежністю не зв'язані.

Якщо  $\eta_{xy} < 1.0$ , то параметри  $Y$  і  $X$  зв'язані кореляційною залежністю.

Якщо  $\eta_{xy} = 1.0$ , то параметри  $Y$  і  $X$  зв'язані функціональною залежністю.

Особливості кореляційного відношення.

Кореляційне відношення дозволяє оцінити тісноту як лінійного, так і нелінійного зв'язку. Разом з тим, кореляційне відношення має недолік: воно не дозволяє судити, наскільки близько розташовані крапки, знайдені за даними спостережень до кривої певного вигляду, наприклад, до параболи, гіперболи і так далі. Це пояснюється тим, що кореляційне відношення у рівнянні регресії не входить, тобто форма зв'язку до уваги не береться.

### 2.3.9. Випадки криволінійної кореляції

Функції регресії  $f(x)$  або  $\varphi(y)$  можуть мати вигляд:

$$\bar{y}_x = ax^3 + bx^2 + cx + d \text{ (параболічна кореляція третього порядку);}$$

$$\bar{y}_x = ax^2 + vx + c \text{ (параболічна кореляція другого порядку);}$$

$$\bar{y}_x = \frac{g}{x} + h \text{ (гіперболічна кореляція).}$$

У разі криволінійної кореляції вирішують ті ж задачі, що і при лінійній.

Невідомі параметри рівняння регресії шукають методом найменших квадратів. Для оцінки тісноти криволінійної кореляції служать вибірові кореляційні відношення.

### 2.3.10 Множинна кореляція

Якщо встановлюється закономірність зв'язку між декількома ознаками, то кореляцію називають множинною.

Простішою в даному випадку є задача, якщо число чинників дорівнює трьом і зв'язок між ними лінійний:

$$z = ax + by + c$$

Вирішення питання виконують у такій послідовності:

- встановлюють закономірність співвідношення між чинниками, записують його у загально прийнятній формі і розраховують коефіцієнти цього рівняння регресії  $a$ ,  $b$  і  $c$  і таке інше;

- визначають тісноту зв'язку між складовими процесу – задачі;

- визначають загальну тісноту зв'язку з урахуванням всіх складових.

Перша задача вирішується методом найменших квадратів, при цьому зручніше користуватися рівняннями вигляду:

$$z - \bar{z} = a(x - \bar{x}) + b(y - \bar{y}),$$

де

$$a = \frac{r_{xy} - r_{yx}r_{xy} \sigma_x}{1 - r_{xy}^2} \sigma_y; \quad b = \frac{r_{yz} - r_{xz}r_{xy} \sigma_z}{1 - r_{xy}^2} \sigma_y.$$

Тіснота зв'язку між складовими  $Z$  і  $X$  (при постійному  $Y$  і  $Z$  і  $Y$  (при постійному  $X$ ) оцінюється відповідно приватними вибірковими коефіцієнтами кореляції:

$$r_{xz(y)} = \frac{r_{xz} - r_{xy}r_{yz}}{\sqrt{(1 - r_{xy}^2)(1 - r_{yz}^2)}}$$

$$r_{yz(x)} = \frac{r_{yz} - r_{xy}r_{xz}}{\sqrt{(1 - r_{xy}^2)(1 - r_{xz}^2)}}.$$

Ці коефіцієнти мають ті ж властивості і те ж саме значення, що і звичайний вибірковий коефіцієнт кореляції, тобто служать для оцінки лінійного зв'язку між ознаками.

Тіснота загального зв'язку ознаки  $Z$  з ознаками  $X$ ,  $Y$  оцінюється вибірковим сукупним коефіцієнтом кореляції;

$$R = \sqrt{\frac{r_{xz}^2 - 2r_{xy}r_{xz}r_{yz} + r_{yz}^2}{1 - r_{xy}^2}},$$

причому  $0 < R < 1$ .

## 2.4. Статистична перевірка статистичних гіпотез

### 2.4.1. Статистична гіпотеза. Нульова і конкуруюча, проста і складна гіпотези

Теорія гіпотез, а також процедура гіпотез складається з трьох етапів:

- прийняття гіпотез;
- перевірка гіпотез;
- помилки та наслідки використання теорії.

Часто при вирішенні практичних інженерних задач необхідно знати закон розподілу генеральної сукупності деякого параметра. Якщо закон розподілу невідомий, то роблять припущення, що він має відповідний вигляд, наприклад – А. Прийняте припущення можна вважати як гіпотезу: «генеральна сукупність ознаки розподілена згідно із законом А». Мова може йти не тільки про вигляд розподілу, але і про характеристики закону - параметри. Якщо є підстави припустити, що відповідна характеристика закону розподілу параметр -  $X$  дорівнює певному значенню  $X_0$ , висувають гіпотезу  $X = X_0$ . У цій гіпотезі мова йде про передбачувану величину характеристики параметра з відомим законом розподілу.

Можливі і інші гіпотези - про рівність параметрів двох або декількох розподілів, про незалежність вибірок та інші.

Статистичною називають гіпотезу про вид невідомого розподілу, або про характеристики параметра з відомим розподілом.

Наприклад, статистичними будуть гіпотези:

- генеральна сукупність параметра розподілена по нормальному закону;
- дисперсії двох нормально розподілених сукупностей параметрів рівні між собою.

У першій гіпотезі зроблено припущення про вид невідомого розподілу, в другій - про числові характеристики параметра двох відомих розподілів.

Гіпотеза «завтра буде дощ» не є статистичною, оскільки в ній не йде мова ні про закон розподілу, ні про параметри розподілу.

Безумовно, якщо висувається одна гіпотеза, то має місце протирічна їй гіпотеза. Якщо висунута гіпотеза буде знехтувана, то буде мати місце протирічна їй гіпотеза. Для цих гіпотез прийняті відмінності у визначенні.

Висунуту гіпотезу називають нульовою - основною  $\Gamma_0$

Протирічну їй конкуруючою - альтернативною  $\Gamma_1$ .

Наприклад, якщо нульова гіпотеза полягає в припущенні, що «математичне очікування -  $a$  деякого параметра з нормальним законом розподілу дорівнює 2», то конкуруюча гіпотеза може полягати в припущенні, «що  $a \neq 2$ ». Формалізовано це має такий вигляд:

$$\Gamma_0 : a = 2; \quad \Gamma_1 : a \neq 2;$$

Розрізняють гіпотези, які містять одне і більш за одне припущень, в зв'язку з цим їх ділять на прості і складні.

#### **2.4.2. Помилки першого і другого роду, що стосуються гіпотез**

Висунута гіпотеза може бути правильною або неправильною, тому виникає необхідність її перевірки. Оскільки перевірку проводять статистичними методами, її називають статистичною.

У результаті статистичної перевірки прийнятої гіпотези можливо допущення помилок, причому допущені помилки можуть бути двох родів.

Помилка першого роду полягає в тому, що буде знехтувана правильна гіпотеза.

Помилка другого роду полягає в тому, що буде прийнята неправильна гіпотеза.

Як правило, наслідки цих помилок різні.

Оцінкою статистичної перевірки приймається ймовірність виконання певного значення.

Ймовірність зробити помилку першого роду прийнято позначати через  $\alpha$ ; її називають рівнем значущості. Найчастіше рівень значущості приймають рівним 0,05 або 0,01. Якщо, наприклад, прийнятий рівень значущості рівний 0,05, то це означає, що в п'яти випадках із ста ми ризикуємо допустити помилку першого роду (відкинути правильну гіпотезу).

#### **2.4.3. Статистичний критерій перевірки нульової гіпотези. Спостережуване значення критерію**

Для перевірки нульової гіпотези використовують спеціально підібрану випадкову величину, точний або наближений розподіл якої відомий. Цю величину позначають відповідним чином, наприклад,  $Z$  - якщо вона розподілена нормально,  $F$  або  $v^2$ - згідно із законом Фішера - Снедекора,  $T$  - згідно із законом Стьюдента  $\chi^2$ - згідно із законом «хі квадрат» » і так далі.

Статистичним критерієм називають випадкове значення -  $K$  певної величини, що має певний розподіл і яке служить для перевірки нульової гіпотези.

Наприклад, якщо перевіряють гіпотезу про рівність дисперсій двох нормальних генеральних сукупностей деяких параметрів, то в якості критерія -  $K$  приймають відношення виправлених вибірових дисперсій:

$$F = \frac{s_1^2}{s_2^2}$$

Ця величина випадкова, тому що в різних дослідах дисперсії прийматимуть різні наперед невідомі значення і розподілена згідно із законом Фішера — Снедекора.

Для перевірки гіпотези за даними вибірок обчислюють значення величин, за якими виконується перевірка і знаходять значення критерія, тому значення критерія називають спостережуваним – вибіровим -  $K_{cn}$ . Наприклад, якщо по двох вибірках, що взяті з нормальних генеральних сукупностей, знайдені виправлені вибірові дисперсії  $S_1^2 = 6$  и  $S_2^2 = 3$ , то спостережуване значення критерію  $F$  буде таке:

$$F_{cn} = \frac{s_1^2}{s_2^2} = \frac{6}{3} = 2$$

#### **2.4.4. Критична область. Область ухвалення гіпотези. Критичні крапки**

Значення критерію знайдено, а у подальшому постає питання, як перевірити правильність прийнятої гіпотези, як скористатись цим значенням.

У попередньому сказано, що стосовно статистичної перевірки можуть прийматись різні гіпотези, для різних чинників або їхніх законів розподілу, для цього створена певна кількість критеріїв -  $K_{кр}$ , з якими порівнюють спостережуване значення.

У загальному розумінні критерій - це безліч всіх його можливих значень. Усі значення критерію розбивають на дві або три неперетинаючих (частини) підмножин – області, в яких критерій має певні властивості, стосовно яких області називають – «область ухвалення гіпотез», - «критична область, в якій намічену гіпотезу відкидають».

Областю ухвалення гіпотези (областю допустимих значень) називають сукупність значень критерію, при яких гіпотезу приймають.

Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу відкидають.

Принцип перевірки статистичних гіпотез реалізується таким чином : якщо спостережуване значення критерію належить критичній області — гіпотезу відкидають, якщо спостережуване значення критерію належить області ухвалення гіпотези — гіпотезу приймають.

Внаслідок того, що критерій  $K$  є одновимірною випадковою величиною, а його можливі значення належать деякому інтервалу, то критична область і область ухвалення гіпотези також є інтервалами і, безумовно, існують крапки, які їх розділяють.

Критичними крапками (межами)  $K_{кр}$  називають крапки, що відокремлюють критичну область від області прийняття гіпотези.

Якщо сукупність критерій розбивається на дві частини, то його називають одностороннім, причому розрізняють односторонню - правосторонню або лівосторонню області, одна з них є критичною, а друга - область прийняття гіпотез.

Якщо сукупність критерія розбивається на три частини, то її називають двосторонньою. Різновидності критеріїв представлені на рис. 9.

Правосторонньою називають критичну область, визначувану нерівністю  $K > K_{кр}$ , де  $K_{кр}$  — позитивне число (малюнок А).

Лівобічною називають критичну область, визначувану нерівністю  $K < K_{кр}$ , де  $K_{кр}$  — негативне число (малюнок - В).

Двосторонньою називають критичну область, визначувану нерівностями  $K < K_{кр1}$  ,  $K > K_{кр2}$ , де  $K_{кр2} > K_{кр1}$ . (малюнок С).

Рис. 9. Графічне представлення критеріїв



Якщо критичні крапки симетричні щодо нуля, двостороння критична область визначається нерівностями при  $|K_{кр}| > 0$ .

Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу відкидають:

$$K < -K_{кр}, K > K_{кр}$$

Значення критичної області критерію на малюнку виділені чорним кольором. Якщо значення спостережуваного критерію припаде на критичну область, то нульову гіпотезу відкидають, якщо ні, то її приймають.

#### 2.4.5. Відшукування правосторонньої критичної області

Правостороння критична область визначається нерівністю:

$$K_{сп} > K_{кр},$$

де  $K_{кр} > 0$

Для відшукування правосторонньої критичної області досить знайти критичну крапку. З цією метою задаються достатньо малою ймовірністю — рівнем значущості -  $\alpha$ . Потім шукають критичну крапку -  $K_{кр}$ , виходячи з вимоги, щоб за умови справедливості нульової гіпотези ймовірність того, що критерій  $\Gamma_0$  набуде значення більше  $K_{кр}$ , дорівнювала прийнятому рівню значущості:

$$P(K > K_{кр}) = \alpha$$

Для кожного критерію є відповідні таблиці, по яких і знаходять критичну крапку, що задовільняє цій вимозі.

#### 2.4.6. Відшукування лівобічної і двосторонньої критичних областей

Відшукування лівобічної і двосторонньої критичних областей так само зводиться до знаходження відповідних критичних крапок.

Лівобічна критична область визначається нерівністю  $K < K_{кр}$  ( $K_{кр} < 0$ ).

Критичну крапку знаходять, виходячи з вимоги, щоб при справедливості нульової гіпотези ймовірність того, що критерій прийме значення, менше  $K_{кр}$ , дорівнювала прийнятому рівню значущості:

$$P(K < K_{кр}) = \alpha$$

Двостороння критична область визначається нерівностями  $K < K_{кр1}$ ,  $K > K_{кр2}$ .

Критичні крапки знаходять, виходячи з вимоги, щоб при справедливості нульової гіпотези сума ймовірності того, що критерій прийме значення менше  $K_{кр1}$ , або більше  $K_{кр2}$ , дорівнювала прийнятому рівню значущості:

$$P(K > K_{кр1}) + P(K < K_{кр2}) = \alpha$$

Якщо розподіл критерію симетричний щодо нуля і обгрунтований, вибираються симетричні щодо нуля точки  $K_{кр}$ , то:

$$P(K < -K_{кр}) = P(K > K_{кр})$$

Враховуючи попередню рівність, отримаємо:

$$P(K > K_{кр}) = \frac{\alpha}{2}$$

Це співвідношення і служить для відшукування критичних точок двосторонньої критичної області.

Як було вказано раніше, критичні крапки знаходять по відповідних таблицях.

#### 2.4.7. Вибір критичної області. Потужність критерію

Критична область будується, виходячи з вимоги, щоб ймовірність попадання в неї критерію була рівна  $\alpha$ , за умови, що нульова гіпотеза справедлива. Практично поступають таким чином - визначається ймовірність попадання критерію в критичну область за умови, що нульова гіпотеза невірна і, отже, справедлива конкуруюча їй.

Потужністю критерію називають ймовірність попадання критерію в критичну область за умови, що справедлива конкуруюча гіпотеза. Іншими словами, потужністю критерію є ймовірність того, що нульова гіпотеза буде знехтувана, якщо вірна конкуруюча гіпотеза.

#### 2.4.8. Порівняння двох дисперсій деяких параметрів (ознак) з нормальним законом розподілу

Розглянемо ознаки  $X$  і  $Y$ , приймемо нульову гіпотезу «дисперсії вказаних чинників рівні між собою».

У відповідності до прийнятої нульової гіпотези запишемо вказану вимогу формалізовано:

$$\Gamma_0: M[S_x^2] = M[S_y^2]$$

В якості оцінки генеральної дисперсії для перевірки нульової гіпотези про рівність генеральних дисперсій, приймаємо виправлену дисперсію, а в якості критерія - відношення більшої виправленої дисперсії до меншої, і таке відношення позначаємо через випадкову величину  $F$ :

$$F = \frac{S_{\sigma}^2}{S_{\mu}^2}$$

Величина  $F$  за умови справедливості нульової гіпотези має розподіл Фішера – Снедекора з мірами свободи  $\kappa_1 = n_1 - 1$  і  $\kappa_2 = n_2 - 1$ , в даному випадку  $n_1$  - об'єм вибірки, по якій обчислювалася більша виправлена дисперсія, а  $n_2$  - менша.

Порівняння виконується залежно від прийнятої конкуруючої гіпотези. Наприклад, при нульовій гіпотезі:

$$\Gamma_0: M[S_x^2] = M[S_y^2]$$

Конкуруюча одностороння, а саме правостороння:

$$\Gamma_1: M[S_x^2] > M[S_y^2].$$

При рівні значущості  $\alpha$  ймовірність попадання критерію в цю область:

$$P[F > F_{кр}(\alpha, \kappa_1, \kappa_2)] = \alpha$$

Точка  $F_{кр}(\alpha, \kappa_1, \kappa_2)$  знаходиться по таблиці критичних точок розподілу Фішера – Снедекора, при цьому критична область визначиться:

$$F > F_{кр},$$

а область ухвалення нульової гіпотези:

$$F < F_{кр}$$

Приклад. По двох незалежних вибірках  $n_1 = 8$  і  $n_2 = 14$ , що одержані з сукупностей значень чинників  $X$  і  $Y$ , які мають нормальний закон розподілу, знайдені виправлені вибіркові дисперсії  $s_x^2 = 0.5$  і  $s_y^2 = 2.5$ . При рівні значущості  $\alpha = 0.05$  перевірити нульову гіпотезу  $\Gamma_0 : D(X) = D(Y)$  про рівність двох дисперсій, при конкуруючій гіпотезі  $\Gamma_1 : D(X) > D(Y)$ .

Знайдемо спостережуване значення критерію:

$$F_{cn} = \frac{2.5}{0.5} = 5.0,$$

оскільки конкуруюча гіпотеза  $\Gamma_1 : D(X) > D(Y)$ , то критична область правостороння.

По таблиці додатку 7 знаходимо критичну точку  $F_{кр}(\alpha = 0.05, \kappa_1 = 8 - 1, \kappa_2 = 14 - 1) = 4.44$ .

Оскільки  $F_{cn} > F_{кр}$ , нульова гіпотеза відкидається, тобто генеральні дисперсії не рівні між собою.

При нульовій гіпотезі  $\Gamma_0 : M[Sx^2] = M[Sy^2]$  і конкуруючій  $\Gamma_1 : M[Sx^2] \neq M[Sy^2]$  є двостороння критична область.

Якщо позначити через  $F_1$  ліву межу критичної області, а через  $F_2$  – праву, то можна записати співвідношення:

$$P(F < F_1) = \frac{\alpha}{2} \quad P(F > F_2) = \frac{\alpha}{2},$$

при цьому критичні крапки:

$$F < F_1, F > F_2$$

а область ухвалення нульової гіпотези:

$$F_1 < F < F_2$$

Для знаходження правої критичної крапки користуються таблицею розподілу Фішера – Снедекора з параметрами  $(\frac{\alpha}{2}, \kappa_1, \kappa_2)$ , якщо  $F_{cn} < F_{кр}$  - нульову гіпотезу приймають, інакше її відкидають. Зауваження в даному випадку приймається не  $\alpha$ , а  $\frac{\alpha}{2}$ .

Приклад. По двох незалежних вибірках  $n_1 = 9$  і  $n_2 = 12$ , що одержані з сукупностей значень чинників  $X$  і  $Y$ , які мають нормальний закон розподілу, знайдені виправлені вибіркові дисперсії  $s_x^2 = 1.5$  і  $s_y^2 = 3.5$ . При рівні значущості  $\alpha = 0.1$  перевірити нульову гіпотезу  $\Gamma_0 : D(X) = D(Y)$  про рівність двох дисперсій, при конкуруючій гіпотезі  $\Gamma_1 : D(X) \neq D(Y)$ .

Знайдемо спостережуване значення критерію:

$$F_{cn} = \frac{3.5}{1.5} = 2.33$$

Оскільки конкуруюча гіпотеза  $\Gamma_1 : D(X) \neq D(Y)$ , то критична область двостороння.

По таблиці додатку 7 знаходимо критичну точку:

$$F_{кр}(\frac{\alpha}{2} = 0.05, \kappa_1 = 9-1, \kappa_2 = 12 - 1) = 2.26.$$

Оскільки  $F_{сп} < F_{кр}$ , нульова гіпотеза приймається, тобто генеральні дисперсії рівні між собою.

#### 2.4.9. Порівняння виправленої вибіркової дисперсії деякого чинника, що має нормальний закон розподілу з його генеральною дисперсією

Нульову гіпотезу представляють:

$$\Gamma_0: M[S^2] = M[\sigma_0^2]$$

Для перевірки нульової гіпотези приймають випадкову величину:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}.$$

Критична область будується залежно від вигляду конкуруючої гіпотези.

Якщо конкуруюча гіпотеза  $\Gamma_1: \sigma^2 > \sigma_0^2$ , то будують правосторонню критичну область.

При прийнятому рівні значущості  $\alpha$  вираз для визначення попадання критерію в цю область такий:

$$P[\chi^2 > \chi_{кр}^2(a, k)] = \alpha,$$

при цьому правостороння критична область визначиться нерівністю:

$$\chi^2 > \chi_{кр}^2,$$

а область ухвалення нульової гіпотези:

$$\chi^2 < \chi_{кр}^2$$

Приклад. По незалежній вибірці  $n_1 = 13$ , що одержана з генеральної сукупності чинника  $X$ , який має нормальний закон розподілу, знайдена виправлена вибірка дисперсія  $Sx_2 = 1.5$ .

При рівні значущості  $a = 0.01$  перевірити нульову гіпотезу  $\Gamma_0: D_b = D_r = 0.5$ , якщо конкуруючою буде гіпотеза  $\Gamma_1: D_b > 0.5$ .

Знайдемо спостережуване значення критерію:

$$\chi_{сп}^2 = \frac{(n-1)s_e^2}{s_e^2} = \frac{(13-1)1.5}{0.5} = 23$$

Оскільки конкуруюча гіпотеза  $\Gamma_1: D_b > 0.5$ , то критична область правостороння.

По таблиці додатку 5 знаходимо критичну крапку:

$$\chi_{кр}(a = 0.01, k = 13-1) = 26.2.$$

Оскільки  $F_{сп} < F_{кр}$  - нульова гіпотеза приймається, тобто вибірка дисперсія правильно представляє генеральну.

При нульовій гіпотезі  $\Gamma_0: M[S_2] = M[\sigma_0^2]$  і конкуруючій  $\Gamma_1: \sigma^2 \neq \sigma_0^2$  будують двосторонню критичну область.

Ймовірність попадання в критичну двосторонню область визначається таким чином:

$$P(\chi^2 > \chi_{лев}^2) + P(\chi^2 < \chi_{прав}^2) = 1$$

При рівні значущості  $a = 0.01$  по таблиці знаходять ліву критичну крапку:

$$\chi_{лев}^2(1 - \frac{a}{2}, k), \text{ а правосторонню } \chi_{прав}^2(\frac{a}{2}, k).$$

Якщо  $\chi_{лев}^2 < \chi^2 < \chi_{прав}^2$  - немає підстав відкинути нульову гіпотезу.

Приклад. По незалежній вибірці  $n_l=13$ , що одержана з генеральної сукупності чинника  $X$ , який має нормальний закон розподілу, знайдена виправлена вибіркова дисперсія  $Sx^2=10.5$ .

При рівні значущості  $\alpha = 0.01$ , перевірити нульову гіпотезу  $\Gamma_0 : D_v = D_g = 10.0$ , при конкуруючій гіпотезі  $\Gamma_1 : D_v \neq 10.0$

Знайдемо спостережуване значення критерію:

$$\chi^2_{cn} = \frac{(n-1)s_g^2}{s_v^2} = \frac{(13-1)10.5}{10.0} = 12.6$$

Оскільки конкуруюча гіпотеза  $\Gamma_1 : D_v \neq 10.0$ , то критична область двостороння.

По таблиці додатку 5 знаходимо критичні крапки:

ліву  $\chi^2_{кр}(1 - \frac{\alpha}{2} = 0.99, k=13-1) = 4.11,$

праву  $\chi^2_{кр}(\frac{\alpha}{2} = 0.005, k=13-1) = 27.7$

Оскільки  $\chi^2_{л} < \chi^2_{сп} < \chi^2_{кр}$  - нульова гіпотеза приймається, тобто вибіркова дисперсія правильно представляє генеральну.

Якщо конкуруюча гіпотеза  $\Gamma_1 : D_v < D_g$ , то критичну крапку знаходять по виразу  $\chi^2_{кр}(1 - \alpha, k)$ .

#### 2.4.10. Порівняння двох середніх значень деяких параметрів, що мають нормальний закон розподілу, дисперсії яких відомі

З генеральних сукупностей значень чинників  $X$  і  $Y$  отримані вибірки об'ємів  $n$  і  $m$ , по яких знайдені їх вибіркові середні значення  $\bar{x}$ ,  $\bar{y}$ .

По вибіркових середніх при заданому рівні значущості  $\alpha$  перевіряється нульова гіпотеза, яка полягає в тому, що генеральні середні або математичні очікування даних значень сукупностей вказаних чинників рівні між собою, тобто:

$$\Gamma_0 : M[X] = M[Y]$$

Оскільки вибіркові середні є незміщеними оцінками генеральних середніх, то нульова гіпотеза :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}]$$

В якості критерія перевірки нульової гіпотези приймається випадкова величина:

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}}$$

Критичну область будують залежно від виду конкуруючої гіпотези.

Якщо нульова гіпотеза прийнята :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

то конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] \neq M[\bar{Y}].$$

В цьому випадку будується двостороння критична область при рівні значущості нульової гіпотези  $\alpha$

Внаслідок того, що випадкова величина розподілена нормально, а отже симетрично щодо нуля, то критичні крапки також симетричні щодо нуля. З цього виходить, що крапка  $Z_{кр}$ - права дорівнює крапці  $Z_{кр}$ -лівій, тобто, щоб знайти двосторонню критичну область досить знайти праву межу. Права межа знаходиться за допомогою функції Лапласа. При рівні значущості  $\alpha$  функція Лапласа :

$$\Phi(z_{кр}) = \frac{1-\alpha}{2}.$$

Значення аргументу функції Лапласа знаходять по таблиці, додаток.

Критична область визначається нерівністю:

$$|Z| > Z_{кр}.$$

Область ухвалення нульової гіпотези:

$$|Z| < Z_{кр}.$$

Якщо  $|Z_{сп}| < Z_{кр}$  – нульову гіпотезу приймають.

Якщо  $|Z_{сп}| > Z_{кр}$  – нульову гіпотезу відкидають.

Якщо нульова гіпотеза прийнята :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

а конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] > M[\bar{Y}]$$

В цьому випадку будують правосторонню критичну область.

Критична крапка знаходиться за допомогою функції Лапласа, додаток 2.

$$\Phi(z_{кр}) = \frac{1-2\alpha}{2}$$

Якщо  $|Z_{сп}| < Z_{кр}$  – нульову гіпотезу приймають.

Якщо  $|Z_{сп}| > Z_{кр}$  – нульову гіпотезу відкидають.

Якщо нульова гіпотеза прийнята :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

а конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] < M[\bar{Y}].$$

В цьому випадку будують лівобічну критичну область.

#### 2.4.11. Порівняння двох середніх значень деяких параметрів нормально розподілених, дисперсії яких невідомі і однакові

Якщо вибірка має великий об'єм, то спостережуване значення критерію обчислюють за формулою:

$$Z_{cn} = \frac{\bar{X} - \bar{y}}{\sqrt{\frac{D_s(X)}{n} + \frac{D_s(Y)}{m}}}$$

Критичну крапку знаходять з рівності:

$$\Phi(z_{кр}) = \frac{1-2\alpha}{2}$$

При  $Z_{сп} > Z_{кр}$ , нульову гіпотезу відкидають.

Якщо об'єм вибірки малий, то критичну область будують залежно від виду конкуруючої гіпотези.

Якщо нульова гіпотеза прийнята :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

то конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] \neq M[\bar{Y}].$$

Як критерій перевірки нульової гіпотези застосовують випадкову величину  $T$ , яка розподілена за законом розподілу Стюдента з  $k = n + m$  мірами свободи.

Якщо  $|T_{сп}| < t_{кр.дв}$  – нульову гіпотезу приймають.

Якщо  $|T_{сп}| > t_{кр.дв}$  – нульову гіпотезу відкидають.

Якщо нульова гіпотеза прийнята :

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

а конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] > M[\bar{Y}]$$

В цьому випадку будують правосторонню критичну область.

Якщо  $T_{сп} < t_{кр.пр}$  – нульову гіпотезу приймають.

Якщо  $T_{сп} > t_{кр.пр}$  – нульову гіпотезу відкидають.

Якщо нульова гіпотеза прийнята:

$$\Gamma_0 : M[\bar{X}] = M[\bar{Y}],$$

а конкуруюча їй:

$$\Gamma_1 : M[\bar{X}] < M[\bar{Y}]$$

В цьому випадку будують лівобічну критичну область.

Критична крапка знаходиться за допомогою  $t$  – розподіли Стюдента.

Спочатку знаходять  $t_{наб}$  як для правосторонньої критичної області, а потім порівнюють  $t_{сп}$  з  $t_{сп-пр}$ .

Якщо  $T_{сп} > -t_{кр.пр}$  – нульову гіпотезу приймають.

Якщо  $T_{сп} < -t_{кр.пр}$  – нульову гіпотезу відкидають.

#### 2.4.12. Перевірка гіпотези про нормальний розподіл генеральної сукупності деякого параметра

Перевірка здійснюється за допомогою критеріїв згоди про передбачуваний закон розподілу.

На практиці застосовують декілька критеріїв згоди:  $\chi^2$  (хі квадрат) Пірсона, Колмогорова і ін.

В якості критерія перевірки нульової гіпотези використовується випадкова величина з  $k = s - l - r$  мірами свободи, де  $s$  - число груп вибірки;  $r$  - число параметрів передбачуваного розподілу, які оцінені за даними вибірки:

$$\chi^2 = \sum \frac{(n_i - n_i')^2}{n_i'}$$

Якщо  $\chi^2_{сп} < \chi^2_{кр}$  – нульову гіпотезу приймають.

Якщо  $\chi^2_{сп} > \chi^2_{кр}$  – нульову гіпотезу відкидають.

### 3. Випадкові функції

У збагаченні корисних копалини при вирішенні багатьох практичних питань виникає потреба дослідження поведінки відповідних чинників в часі, або на деякому числовому інтервалі його значень. В цьому разі єдино доцільним є застосування теорії випадкових функцій. У подальшому

розглянуто деякі положення теорії випадкових функцій стосовно їх практичного використання.

### **3.1. Визначення випадкової функції випадкового процесу**

У теорії ймовірностей і математичній статистиці всі дії, наприклад, визначення вмісту певного мінералу корисної копалини у всіх випадках з випадковими величинами, виконуються при деякій фіксованій сукупності умов. Причому, не у всіх випадках фіксація сукупності умов може бути досягнута, тобто умови можуть безперервно змінюватись. Наприклад, змінюються продуктивність технологічної секції по збагаченню корисної копалини і вміст корисного мінералу в ній упродовж зміни, або при зміні деякого чинника розподільчого середовища.

Ті конкретні безперервні значення, які приймає випадкова величина-деякий чинник, при зміні часу або іншого чинника, прийнято називати випадковими.

З метою однозначного розуміння положень і мотивування рішень доцільним ввести однозначні поняття, «випадковий процес» та «випадкова функція». У наявних науково технічних джерелах така особливість згадується, однак при викладенні теоретичних положень не дотримується.

У подальшому викладенні матеріалу вказані терміни вміщують такий зміст.

Якщо розглядається безперервна зміна деякого чинника у часі, то використовується поняття - випадковий процес.

Якщо визначається співвідношення чинника при зміні іншого, який виступає у якості аргументу, то використовується термін – випадкова функція. При дослідженні таких випадкових функцій при зміні часу також використовується термін – випадковий процес. У цьому випадку досліджуваний чинник залежить не від одного аргументу, а від двох.

У посібнику викладені основні положення теорії випадкових функцій - процесів так, як це потрібно для загальної постановки і вирішення деяких інженерних практичних задач.

Одним з елементів теорії є поняття випадкової функції - процесу. У подальшому дублювання понять функція – процес, якщо це не конкретизується, використовуватись не буде.

Випадковими називають функції, які в результаті проведення випробування - дослідження можуть набирати того або іншого конкретного вигляду - значення, проте яким він буде, ми не знаємо і заздалегідь встановити не можемо.

Конкретний вигляд, що одержується функцією при її отриманні, називається реалізацією випадкової функції.

Якщо в результаті безлічі дослідів проводяться безперервні вимірювання значень деякого чинника, то отримують «сімейство» значень - реалізацій цієї функції.

Приклад випадкового процесу.



В процесі переробки вугілля, незмінного, стосовно деяких чинників на збагачувальній фабриці безперервно вимірюється зольність вхідного продукту упродовж декількох змін. З практичних спостережень відомо, що зольність не є постійною, тобто вона безперервно змінюється, тому як результат буде отримана безперервна характеристика параметра-зольності, яка є реалізацією випадкового процесу.

Якщо одержану реалізацію розділити на декілька з тривалістю, що відповідає одній зміні і одержані значення розмістити на одному графіку, тобто характеристики зольності декількох змін зобразити на загальному графіку, то отримаємо сімейство реалізацій, які внаслідок наявності помилки вимірювання і зміни зольності відрізнятимуться між собою і практично в точності ніколи не повторюватимуть одна одну, в цьому полягає суть випадковості процесів. Відповідно, якщо аргументом виступає не час, а інший технологічний чинник, то суть визначає випадковість функції.

Реалізація зміни зольності протягом часу вимірювань її значень показана на рис. 10, а сімейство реалізацій віднесених до однієї зміни - на рис. 11.

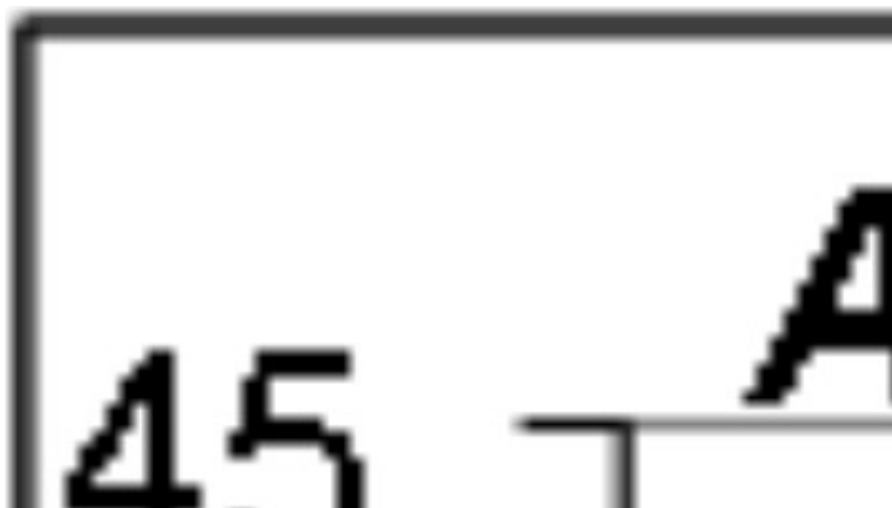


Рис.10. Реалізація випадкового процесу



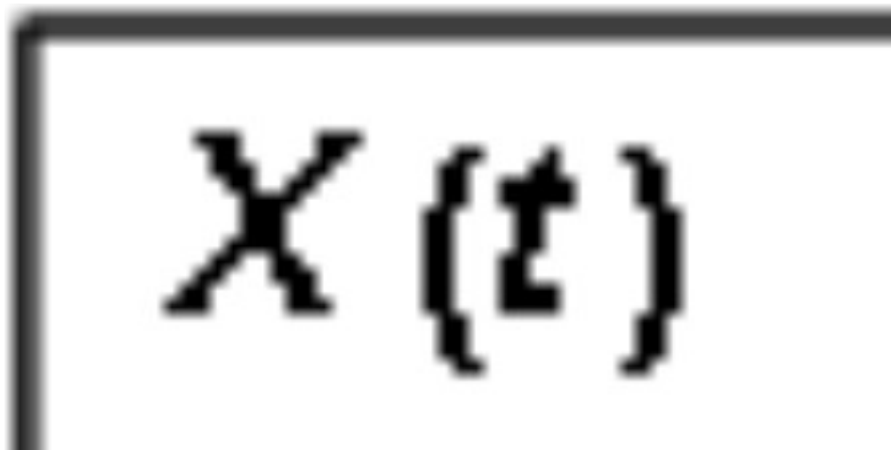
Рис. 11. Сімейство реалізацій випадкового процесу

У практичній роботі більшість задач, в яких розглядаються випадкові функції такі, що аргументом може бути не тільки час, але і будь-який інший один або декілька фізичних параметрів. Однак, незалежно від того, що являється аргументом у процесі дослідження випадкової функції, процедури визначення певних положень одні і ті самі.

Розглянемо випадкові функції певних чинників, що визначаються тільки одним аргументом, наприклад, часом -  $t$ , тобто розглядаємо випадковий процес.

Загальноприйнято, випадкові функції чинників позначати великими буквами ( $X(t)$ ,  $Y(t)$ ,  $U(t)$ ) на відміну від не випадкових функцій ( $x(t)$ ,  $y(t)$ ,  $u(t)$ ), тобто їх реалізацій.

Розглянемо деякий випадковий чинник - величину  $X(t)$ . Припустимо, що над ним проведено  $n$  незалежних дослідів з безперервним вимірюванням його значень, в результаті яких отримано  $n$  реалізацій, представлених на рис. 12. Позначимо їх у відповідності з номерами досліджень - одержання  $x_1(t)$ ,  $x_2(t)$ , ...,  $x_n(t)$ .



## Рис. 12. Реалізація випадкових функцій

Кожна отримана реалізація це конкретні значення випадкової функції - чинника, і в цьому випадку реалізація є не випадковою для випадкової функції. Тобто, якщо одержана конкретна реалізація випадкової функції, то вона вже не є невідомою, а також, випадковою. Слід усвідомити, що одна реалізація це певна конкретна оцінка дійсної оцінюваної нами випадкової функції або процесу. Наскільки точно одержана оцінка - оцінює оцінюваний чинник, що має властивість випадковості, це питання вирішується з застосуванням відповідних математичних процедур.

Дуже важливим етапом при дослідженні випадкових функцій являється особливість дослідження, тобто спосіб виконання.

Використовується декілька методів досліджень випадкових чинників.

При безперервному вимірюванні значень чинника за допомогою технічних засобів виконуються розрахунки значень його складових, що передбачені методом дослідження.

З досліджувального чинника заміряються відповідна кількість його значень і на підставі їх обробки визначаються особливості випадкового чинника – функції.

Другий метод являється більш простим і достатньо ефективним.

Заміряють значення певної реалізації випадкової функції в деякі фіксовані моменти часу. В результаті одержують стільки числових значень чинника, скільки моментів часу було прийнято. Одержані числа представляють випадкову функцію у вибрані моменти часу. Сукупність цих чисел можна розглядати як відповідну кількість значень випадкової величини.

Таким чином в практичній роботі один і той же безперервний випадковий чинник можна розглядати як випадковий (процес), функцію  $X(t)$ , або як випадкову величину, залежно від того, розглядаються його значення на всьому діапазоні зміни або у деяких фіксованих точках, наприклад, якщо ці точки представляють часові відрізки.

### **3.2. Взаємозв'язок випадкової функції і системи випадкових величин.**

#### **Закон розподілу випадкової функції (процесу)**

Випадкову функцію можна представляти як сімейство її реалізацій і однією узагальнювальною реалізацією, представленою жирною лінією. Розглянемо випадкову функцію  $X(t)$  на певному відрізку часу, рис. 13, представленою однією реалізацією

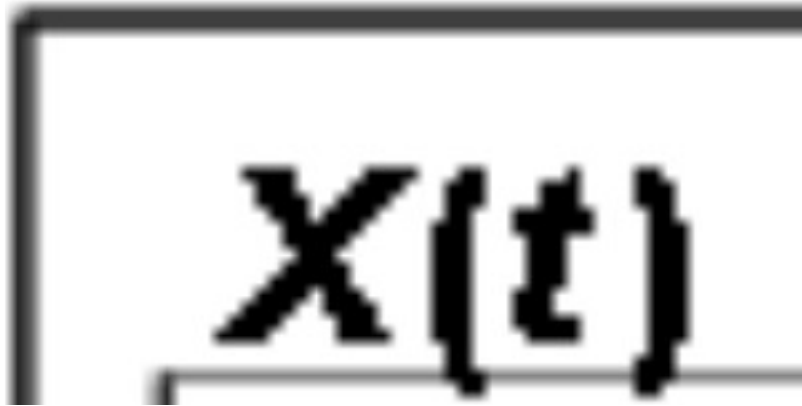


Рис. 13. Реалізація випадкової функції

Як вказано у попередньому, при фіксованому значенні аргументу  $t$  випадкову функцію можна представити відповідною кількістю значень випадкової величини.

Якщо вибрати деяку довжину інтервалу, після закінчення якого вимірювати значення випадкової функції представленою сімейством реалізацій, то можна її конкретні реалізації представити сімейством  $n$  випадкових величин, де  $X_1, X_2, \dots, X_n$ , кожна з яких має певну кількість значень у відповідності до кількості реалізацій:

$$x_1(t_i), x_2(t_i), \dots, x_n(t_i).$$

Такий розгляд випадкової функції називають поперечним перетином, а значення кожної реалізації в перетині представляють як сукупність випадкових величин. Безумовно, що чим більшим числом реалізацій представлена випадкова функція, тим точніше буде результат представляти досліджувану функцію. Графічне зображення такого розгляду показано на рис. 14.

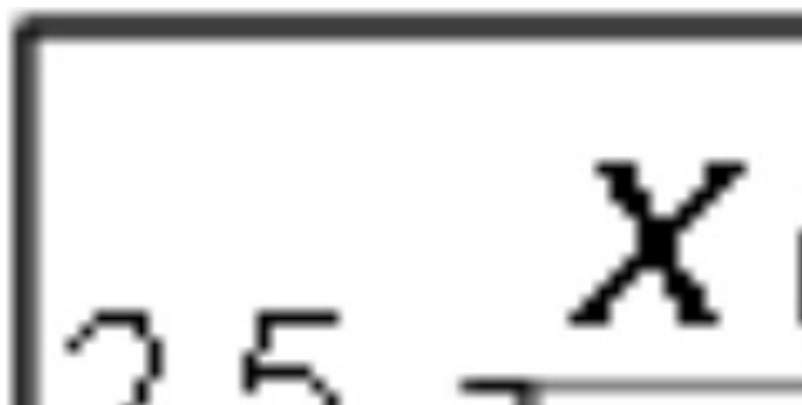


Рис. 14. Розтини випадкової функції

Якщо розглядати окрему реалізацію або узагальнюючу всієї сукупності у багатьох фіксованих крапках, наприклад, часу, то отримаємо конкретні

значення чинника в цих крапках, а ці числа також можна представити як систему випадкових величин. При достатньо високій частоті фіксованих значень аргументу значення випадкової функції дадуть достатньо точне уявлення про те, як вона змінюється. Такий розгляд випадкової функції називають поздовжнім перетином. Таким чином, розгляд випадкової функції можна з певним наближенням замінити розглядом системи випадкових величин.

Найбільш інформативним стосовно випадкової функції (процесу) є закон її розподілу, на підставі якого визначаються його характеристики, що її характеризують. Виходячи з наведеного вище аналізу розглянемо, яким може бути закон розподілу випадкової функції (процесу).

Відомо, що закон розподілу однієї випадкової величини є функція одного аргументу, двох величин — функція двох аргументів і так далі. Проте, на практиці користуватися ймовірностями, як функціями багатьох аргументів, незручно і трудомістко, навіть для систем трьох - чотирьох величин відмовляються від користування законами розподілу і розглядають тільки числові характеристики.

В даному випадку обмежимося розглядом відповідних характеристик випадкових функцій аналогічних числовим характеристикам випадкових величин.

### **3.3. Характеристики випадкових функцій**

Як випадкові величини так і випадкові функції мають свої числові характеристики. За допомогою таких характеристик вирішуються багато практичних питань.

На відміну від числових характеристик випадкових величин, які являються певними числами, характеристики випадкових функцій являються в загальному випадку не числами, а функціями.

В якості числових характеристик випадкових функцій використовують математичне очікування, дисперсію, середнєквадратичне відхилення, а також кореляційну функцію.

Математичне очікування випадкової функції  $X(t)$  визначається таким чином. Розглядається поперечний перетин випадкової функції  $X(t)$ , представленої сімейством реалізацій при фіксованому  $t$ , дивись рис. 15.

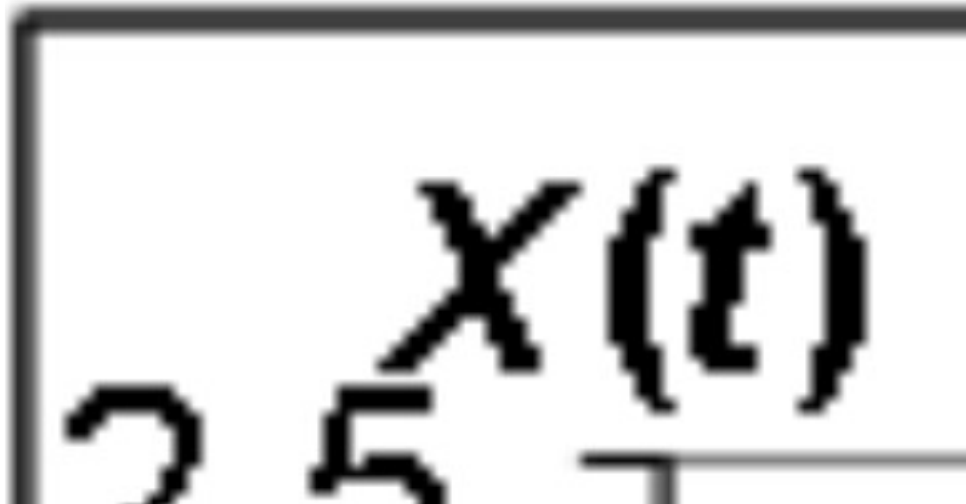


Рис. 15. Перетини випадкової функції

У цьому перетині реалізації мають певні значення, які можна розглядати як випадкову величину. Для одержаної випадкової величини у цьому розтині представляється можливим визначити її математичне очікування. Очевидно, в загальному випадку воно залежить від  $t$ , тобто є деякою функцією  $t$ :

$$m_x(t) = M[X(t)].$$

Таким чином, математичним очікуванням випадкової функції  $X(t)$  називають не випадкову функцію  $m_x(t)$ , яка при кожному значенні аргументу  $t$  дорівнює математичному очікуванню відповідного перетину випадкової функції.

По суті математичне очікування випадкової функції є деяка середня дійсна функція, біля якої відповідним чином розташовуються конкретні реалізації випадкової функції.

Безумовно, що у такому разі випадкова функція для всієї досліджуваної ділянки представляється набором математичних очікувань значень конкретних реалізацій.

На рис. 15 лініями, які коливаються, показані реалізації випадкової функції, прямою лінією - її математичне очікування.

Аналогічним чином визначається дисперсія випадкової функції. Дисперсією випадкової функції  $X(t)$  називають не випадкову функцію  $D_x(t)$ , значення якої для кожного  $t$  дорівнює дисперсії відповідного перетину випадкової функції:

$$D_x(t) = D[X(t)].$$

Дисперсія випадкової функції при кожному  $t$  характеризує відхилення можливих реалізацій випадкової функції щодо середнього значення. Тобто, дисперсія відображає «ступінь випадковості» випадкової функції або ступінь можливого відхилення окремих реалізацій від її усередненого значення.

Очевидно, дисперсія випадкової функції -  $D_x(t)$  завжди позитивна, однак одиниця її виміру не співпадає з одиницею виміру випадкової функції.

Добуваючи з дисперсії квадратний корінь, отримаємо функцію  $\sigma_x(t)$  - середньоквадратичне відхилення випадкової функції:

$$\sigma_x(t) = \sqrt{D_x(t)}$$

Математичне очікування, дисперсія, та середнє квадратичне відхилення являються важливими характеристиками випадкової функції, проте для опису особливостей випадкової функції цих характеристик недостатньо. Щоб переконатися в цьому, розглянемо дві випадкові функції  $X_1(t)$  і  $X_2(t)$ , зображеними сімействами реалізацій на рис. 16 і 17.

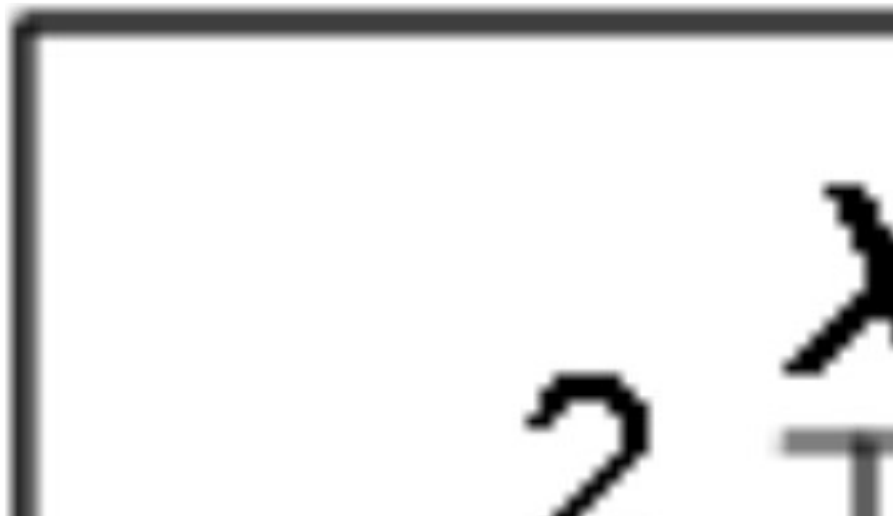


Рис.16. Випадкова функція

По візуальному спостереженню наведені випадкові функції мають приблизно однакові математичні очікування і дисперсії, проте характер змін цих випадкових функцій різний.

Реалізації випадкової функції рис. 16 змінюються часто. Якщо, наприклад, в крапці  $t_1$  реалізація випадкової функції прийняла значення, що перевищує середнє значення сукупності реалізацій, то ймовірно, що в точці  $t_2$ , що розташована на деякій відстані від першої крапки, вона набуде іншого значення. Для випадкової функції, рис. 17, характерно виражена постійна залежність між її значеннями при різних  $t$ .

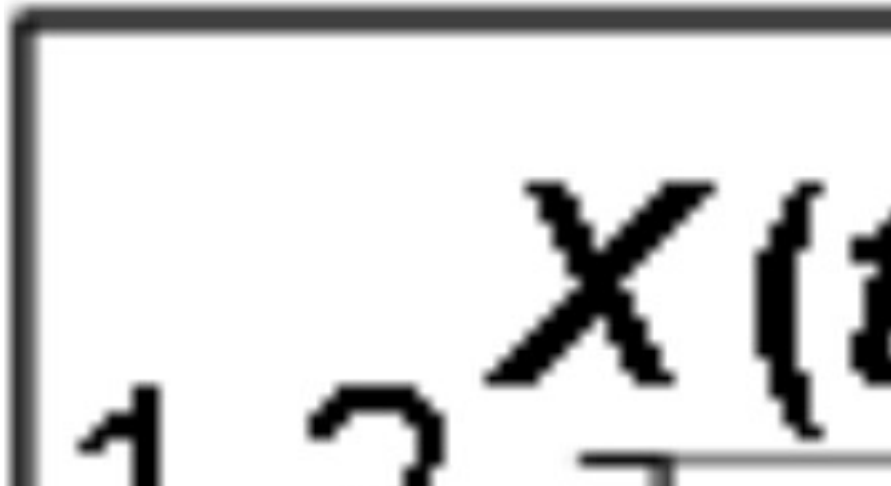


Рис. 17. Випадкова функція

Очевидно, внутрішня структура цих випадкових процесів абсолютно різна, але ця відмінність не обумовлюється ні математичним очікуванням, ні дисперсією, які отримують з поперечного розгляду реалізацій випадкової функції.

Для описання поведінки випадкової функції в часі, або на деякому безперервному числовому відрізку введена спеціальна характеристика, яку можна отримати з подовжнього та поперечного розгляду окремих реалізацій випадкової функції, або продовжнього розгляду функції представленої достатнім інтервалом аргументу. Цю характеристику називають кореляційною функцією (або — автокореляційною функцією).

Кореляційна функція визначається за величиною залежності між продовжними перетинами випадкової функції, що відносяться до різних значень  $t$ , тобто за її допомогою визначають та оцінюють характер коливань чинника.

Розглянемо два перетини реалізації випадкової функції -  $X(t)$ , тобто її математичні очікування, що відносяться до моментів часу  $t_1$  і  $t_2$ , які можна розглядати як дві випадкові величини  $X(t_1)$  і  $X(t_2)$ , рис. 18.

Очевидно, що при близьких значеннях  $t_1$  і  $t_2$  величини  $X(t_1)$  і  $X(t_2)$  зв'язані тісною залежністю, тобто їх значення мало відрізняються між собою. Якщо величина  $X(t_1)$  набула певного значення, то і величина  $X(t_2)$  набуде значення, близького до нього. Очевидно також, що при збільшенні інтервалу між перетинами  $t_1$  і  $t_2$  залежність величин  $X(t_1)$  і  $X(t_2)$  повинна зменшуватись. З цього слідує, що кореляційна функція відображає внутрішню структуру випадкової функції при зміні аргументу, в даному випадку -  $t$ .



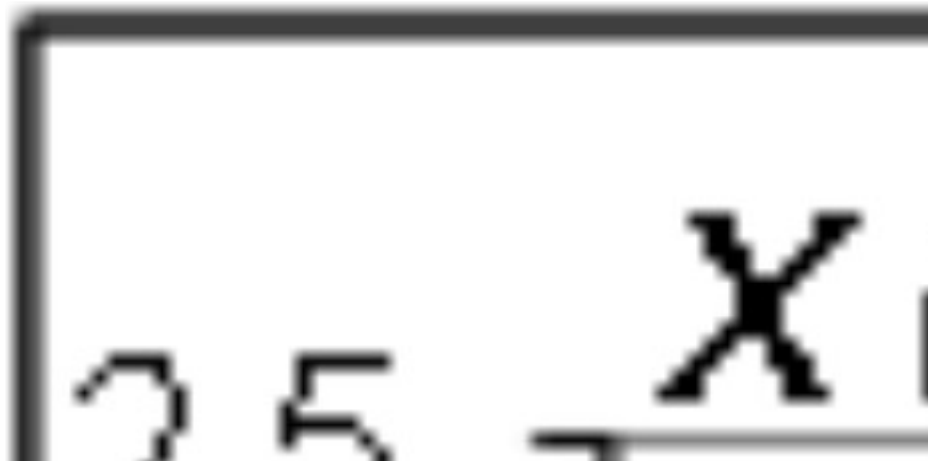


Рис. 18. Випадкова функція

Величину залежності величин  $X(t_1)$  і  $X(t_2)$  характеризують їх кореляційним моментом, очевидно, він є функцією двох аргументів  $t_1$  і  $t_2$ . Ця функція і називається кореляційною функцією.

Таким чином, кореляційною функцією випадкової функції  $X(t)$  називається не випадкова функція двох аргументів  $K_x(t, t_1)$ , яка при кожній парі значень  $t$  і  $t_1$  дорівнює кореляційному моменту відповідних перетинів випадкової функції:

$$K_x(t_1, t_2) = M[X(t_1) X(t_2)],$$

де

$$X(t_1) = X(t_1) - m_x(t_1), \quad X(t_2) = X(t_2) - m_x(t_2)$$

Визначення кореляційного моменту дано в попередньому викладі.

Для розглянутих випадкових функцій візуальний аналіз показує, що при однакових математичних очікуваннях і дисперсіях випадкові функції  $X(t_1)$  і  $X(t_2)$ , рис. 16 і 17 мають винятково різні кореляційні функції. Кореляційна функція випадкової функції рис. 16 швидко убуває у міру збільшення проміжку  $(\Delta t)$ ; навпаки, кореляційна функція випадкової функції, рис. 17, повільно убуває із збільшенням цього проміжку. При  $t_1 - t_2 = 0$  кореляційна функція являється не чим іншим, як дисперсією випадкової функції.

Оскільки кореляційний момент двох випадкових величин  $X(t_1)$  і  $X(t_2)$  не залежить від послідовності, в якій ці величини розглядаються, то кореляційна функція симетрична щодо своїх аргументів, тобто не міняється при зміні аргументів місцями:

$$K_x(t_1, t_2) = K_x(t_2, t_1)$$

Замість кореляційної функції  $K_x(t_1, t_2)$  можна користуватися нормованою кореляційною функцією щодо середніх квадратичних функцій, яка є коефіцієнтом кореляції величин  $X(t_1)$  і  $X(t_2)$ :

$$r_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1) \sigma_x(t_2)}$$

### 3.4. Властивості кореляційної функції

Нормована кореляційна функція аналогічна нормованій кореляційній матриці системи випадкових величин. Під процесом нормування розуміється відношення кореляційного моменту до дисперсії випадкової функції. При  $t_1 = t_2$  нормована кореляційна функція дорівнює одиниці:

$$r_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{[\sigma_x(t)]^2} = \frac{D_x(t)}{[\sigma_x(t)]^2} = 1.0$$

При постійному -  $t_1$  і змінному -  $t_2$  одержують ряд кореляційних моментів -  $R_i$ , виконуючи їх нормування до дисперсії одержують ряд значень нормованої кореляційної функції, яку називають автокореляційною функцією.

При додаванні до випадкової функції  $X(t)$  невідповідного додатку  $\varphi(t)$  отримують нову випадкову функцію:

$$Y(t) = X(t) + \varphi(t),$$

тобто при додаванні до випадкової функції невідповідного додатку до її математичного очікування додається теж невідповідний додаток і при цьому кореляційна функція випадкової функції не змінюється.

Множення випадкової функції  $X(t)$  на невідповідний множник  $\varphi(t)$ :

$$Y(t) = \varphi(t)X(t)$$

Виносячи невідповідну величину  $\varphi(t)$  за знак математичного очікування, маємо:

$$m_y(t) = M[\varphi(t)X(t)] = \varphi(t)m_x(t),$$

тобто при множенні випадкової функції на невідповідний множник її математичне очікування перемножується на той же множник.

При множенні випадкової функції на невідповідну функцію  $\varphi(t)$  її кореляційна функція перемножується на  $\varphi(t_1)\varphi(t_2)$ .

Якщо  $\varphi(t) = n$  постійна величина (не залежить від  $t$ , тобто виключається  $t$ ), кореляційна функція перемножується на  $n^2$ .

Користуючись властивостями характеристик випадкових функцій, можна значно спростити операції з ними. Коли потрібно досліджувати кореляційну функцію або дисперсію випадкової функції, доцільно заздалегідь перейти від неї до так званої центрованої функції:

$$\dot{X}(t) = X(t) - m_x(t)$$

Математичне очікування центрованої функції тотожно дорівнює нулю, а її кореляційна функція збігається з кореляційною функцією випадкової функції  $X(t)$  :

$$\dot{K}_x(t_1, t_2) = M[\dot{X}(t_1)\dot{X}(t_2)] \rightarrow K_x(t_1, t_2)$$

Іноді окрім центрування застосовують ще нормування випадкових функцій. Нормованою називають випадкову функцію вигляду:

$$x_n(t) = \frac{x^3(t)}{\sigma_n(t)}$$

Кореляційна функція нормованої випадкової функції  $X(t)$  дорівнює:

$$K_{xx}(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1)\sigma_x(t_2)} = r_x(t_1, t_2)$$

### 3.5. Практичне визначення характеристик випадкової функції

Визначаючи вище згадані характеристики потрібно дуже добре усвідомлювати, стосовно якої умови вони розраховуються, а саме для поперечного або для продовжного розтинів. Звичайно, що найбільш універсальним є дослідження випадкових функцій за певною кількістю поперечних перетинів. При такому дослідженні випадкових функцій, представлених сукупністю реалізацій, можна визначити для кожного перетину математичне очікування, дисперсію, середнє квадратичне відхилення та автокореляційну функцію. Для будь-якої реалізації також можна визначити вказані характеристики, якщо реалізацію розглядати по повздовжньому перетині.

Припустимо, що над випадковою функцією  $X(t)$ , представленою на рис. 19, проведено  $n$  незалежних дослідів (спостережень) і в результаті отримано  $n$  реалізацій випадкової функції.

Потрібно знайти оцінки для характеристик випадкової функції, а саме, її математичне очікування -  $m_x(t)$ , дисперсію -  $D_x(t)$ , кореляційну функцію -  $K_x(t_1, t_2)$ .

Для цього розглянемо ряд перетинів випадкової функції для моментів часу

$$t_1, t_2, \dots, t_m$$

і зафіксуємо значення, прийняті реалізаціями випадкової функції  $X(t)$  в намічені моменти часу. Кожному з моментів -  $t_i$  відповідатиме  $n$  значень випадкової функції.

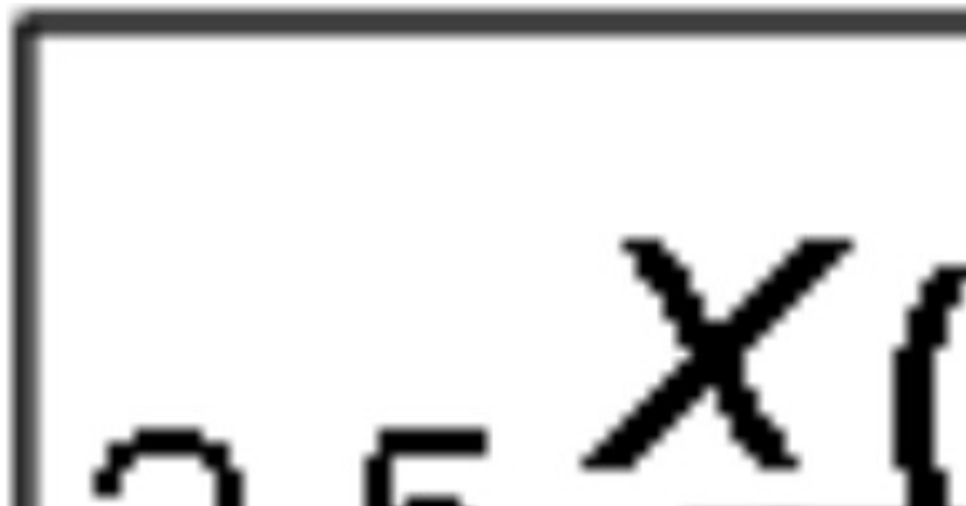


Рис. 19. Реалізації випадкової функції

Значення інтервалів задаються рівновіддаленими; величина інтервалу між сусідніми значеннями вибирається залежно від виду експериментальних кривих так, щоб по вибраних крапках можна було відтворити основний хід кривих.

Зареєстровані значення  $X(t)$  заносяться в таблицю 18, кожен рядок якої відповідає певній реалізації, а число стовпців дорівнює числу значень випадкової функції – числу перетинів.

Таблиця 18

$t$ $X(t)$	$t_1$	$t_2$	.	$t_m$
$X_1(t)$	$X_1(t_1)$	$X_1(t_2)$	.	$X_1(t_m)$
$X_2(t)$	$X_2(t_1)$	$X_2(t_2)$	.	$X_2(t_m)$
.	.	.	.	.
$X_n(t)$	$X_n(t_1)$	$X_n(t_2)$	.	$X_n(t_m)$

Отримані дані є не що інше, як результати  $n$  дослідів над системою  $m$  випадкових величин:

$$X(t_1), X(t_2), \dots, X(t_m)$$

і обробляються так, як це вказано у попередньому. Перш за все знаходяться оцінки випадкової функції;

математичні очікування по формулі:

$$m_x(t) = \frac{\sum_{i=1}^n x_i(t)}{n}$$

дисперсії:

$$D_x(t) = \frac{\sum_{i=1}^n [x_i(t) - m_x(t)]^2}{n - 1}$$

автокореляційні функції поперечних і поздовжніх перетинів:

$$K_x(t, t') = \frac{\sum_{i=1}^{n-1} [x_i(t) - m_x(t)][x_{i+1}(t') - m_x(t')]}{n - 1}$$

Отримані обчислення можна представити у вигляді графіків, побудувавши залежності  $m_x(t)$ ,  $D_x(t)$  і  $K_x(t, t')$ .

При необхідності отримані функції апроксимують певними аналітичними виразами.

### 3.6. Поняття про стаціонарний випадковий процес

На практиці часто зустрічаються випадкові процеси, що протікають в часі приблизно однорідно і мають вигляд безперервних випадкових коливань навколо деякого середнього значення, причому ні середня амплітуда, ні характер цих коливань не виявляють істотних змін з часом. Такі випадкові процеси називаються стаціонарними.

В якості стаціонарних випадкових процесів можна привести: 1) частоту коливань корпуса грохота в сталому режимі роботи; 2) розподіл щільності піщаного і зливних продуктів в гідроциклоні; 3) висоту підйому матеріалу в

барабанному млині при його обертанні; 4) зміну частотної характеристики, що виникає під час роботи барабанного млина та інші.

В якості стаціонарного процесу можна розглядати любий випадковий, якщо він продовжується в часі довго, а початок відліку можна вибрати будь-який момент часу.

Досліджуючи стаціонарний процес на будь-якій ділянці часу, ми повинні отримати одні і ті ж значення його характеристик.

За результатом візуального аналізу можна зробити висновок, що на рис. 20 приведена характеристика стаціонарного випадкового процесу, а на рис. 21 - нестаціонарного .

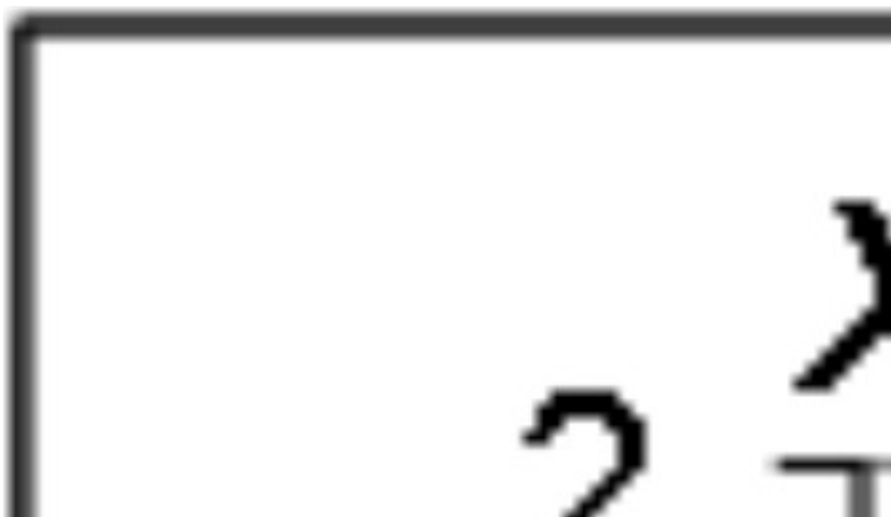


Рис. 20. Випадкова стаціонарна функція

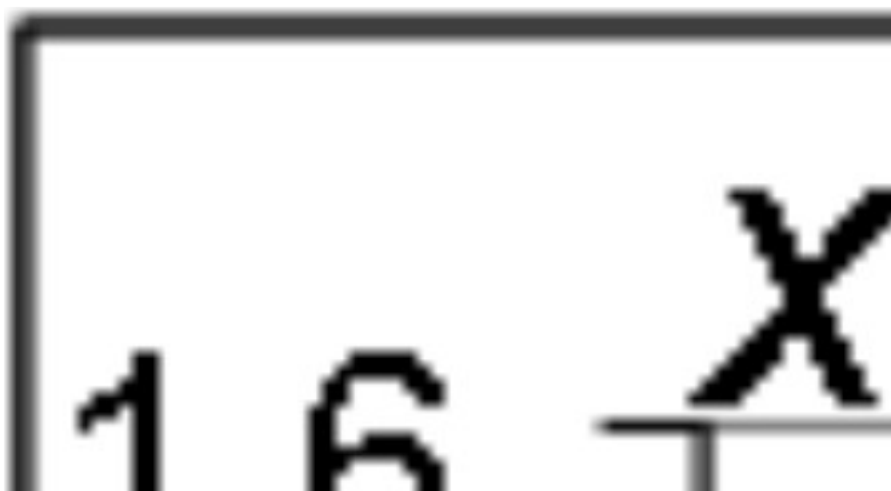


Рис. 21. Випадкова нестаціонарна функція

Не всі нестационарні випадкові процеси є істотно нестационарними на всьому тимчасовому інтервалі. Більшість нестационарних процесів являються такими, що на певних відрізках часу і за певних умов можуть бути прийняті за стаціонарні.

Кожен випадковий процес в будь-якій динамічній системі починається з нестационарної стадії - з так званого «перехідного процесу». Після загасання перехідного процесу система переходить на сталий режим і тоді випадкові процеси, що протікають в ній, можуть вважатися за стаціонарні.

Стаціонарні випадкові процеси часто зустрічаються у фізичних і технічних процесах. По своїй природі ці процеси простіші, ніж нестационарні, і описуються простішими характеристиками. Певні перетворення випадкових процесів здійснюються простіше, ніж нестационарних. У зв'язку з цим на практиці набула широкого поширення теорія стаціонарних випадкових процесів, або, точніше, теорія стаціонарних випадкових функцій (оскільки аргументом стаціонарної випадкової функції в загальному випадку може бути і не час).

Випадкова функція  $X(Y)$  називається стаціонарною, якщо всі її імовірні характеристики не залежать від  $Y$ , (точніше не міняються при будь-якій зміні аргумента, від якого вона залежить, по осі  $Y$ ).

Оскільки зміна стаціонарної випадкової функції може протікати однорідно за часом, то природно зажадати, щоб для стаціонарної випадкової функції математичне очікування було постійним:

$$m_x(t) = m_x = const.$$

Ця вимога не є істотною, ми знаємо, що від випадкової функції  $X(t)$  завжди можна перейти до центрованої випадкової функції  $X(t)$ , для якої математичне очікування дорівнює нулю і, отже, задовольняє вказаній умові. Таким чином, якщо випадковий процес не стаціонарний тільки відносно математичного очікування, це не заважає вивчати його як стаціонарний процес.

Друга умова, якій повинна задовольняти стаціонарна випадкова функція - це умова постійності дисперсії:

$$D_x(t) = D_x = const.$$

Визначимо, якій умові повинна задовольняти кореляційна функція стаціонарної випадкової функції. Розглянемо випадкову функцію  $X(t)$ , рис. 22.

Покладемо у виразі  $K_x(t_1, t_2)$   $t_2 = t_1 + \tau$  і розглянемо  $K_x(t_1, t_1 + \tau)$  - кореляційний момент двох перетинів випадкової функції, розділених інтервалом часу  $\tau$ .

Якщо випадковий процес  $X(t)$  дійсно стаціонарний, то кореляційний момент не повинен залежати від того, де саме на осі  $t$  ми взяли ділянку  $\tau$ , а повинен залежати тільки від довжини цієї ділянки. Наприклад, для ділянок 1 і 2 на рис. 22, що мають одну і ту ж довжину -  $\tau$  значення кореляційної функції  $K_x(t_1, t_2)$  і  $K_x(t_1, t_1 + \tau)$  мають бути однаковими. Взагалі, кореляційна функція стаціонарного випадкового процесу повинна залежати не від положення першого аргументу на осі абсцис, а тільки від проміжку  $\tau$  між першим і другим аргументами:

$$K_x(t, t + \tau) = k_x(\tau)$$



Рис. 22. Реалізація стаціонарної випадкової функції

Отже, кореляційна функція стаціонарного випадкового процесу є функція не двох, а одного аргументу. Це обставина у ряді випадків значно спрощує операції над стаціонарними випадковими функціями.

Таким чином умова, яка вимагає від стаціонарної випадкової функції постійності дисперсії, є окремим випадком, дійсно, вважаючи у формулі  $t + \tau = t$  маємо:

$$D_x(t) = K_x(t, t) = k_x(0) = const$$

Таким чином, вказана умова є однією з важливих, якій повинна задовольняти стаціонарна випадкова функція, тобто при постійності дисперсії випадкову функцію – процес можна вважати стаціонарним.

Тому, в якості стаціонарної випадкової функції приймаємо таку випадкову функцію, кореляційна функція якої залежить не від обох своїх аргументів  $t_1$  і  $t_2$ , а тільки від різниці  $\tau$  між ними.

З результатів практичного використання відомо, що кореляційна функція будь-якої випадкової функції симетрична:

$$k_x(t_1, t_2) = k_x(t_2, t_1)$$

Виходячи з цього для стаціонарного випадкового процесу, вважаючи -  $(t_2 - t_1) = \tau$ , маємо:

$$k_x(\tau) = k_x(-\tau)$$

Тобто кореляційна функція  $k_x(\tau)$  є парна функція свого аргументу. Тому кореляційну функцію визначають тільки для позитивних значень аргументу.

На практиці замість кореляційної функції  $k_x(\tau)$  часто користуються нормованою - автокореляційною кореляційною функцією:

$$\rho_x(\tau) = \frac{k_x(\tau)}{D_x},$$

де  $D_x = \kappa_x(\tau)$  - постійна дисперсія стаціонарного процесу. Функція  $\rho_x(\tau)$  є не що інше, як коефіцієнт кореляції між перетинами випадкової функції, розділеними інтервалом  $\tau$  за часом. Очевидно, що  $\rho_x(\tau) = 1$ .

Розглянемо два стаціонарних випадкових процеси і розрахуємо та побудуємо їх характеристики.

Приклад. Випадкова функція  $X(t)$  задана сукупністю 7 реалізацій, які одержані на протязі 7 годин роботи технологічної лінії. Реалізації показані на рис. 23.

Знайти її характеристики  $m_x(t)$ ,  $K_x(t, t')$ ,  $D_x(t)$  і нормовану кореляційну функцію  $r_x(t, t')$ . По візуальному спостереженню наведену випадкову функцію можна вважати стаціонарною. Для розрахунку характеристик необхідно визначити довжину інтервалу поперечних розтинів реалізацій, з яких будуть взяті їх значення. Довжина інтервалу визначається таким чином. Передивляються реалізації випадкової функції і серед них виділяють таку, яка має найменший інтервал зміни знака, приклад показано на рис. 24.

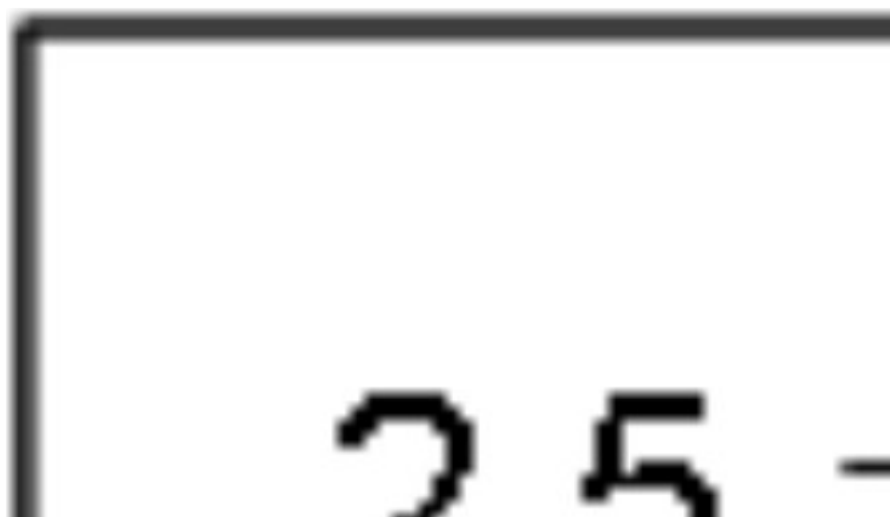


Рис. 23. Реалізації випадкової функції

Рис. 24. Визначення довжини інтервалу розтинів випадкової функції

Визначений інтервал позначками  $t_1$  і  $t_2$  ділиться на дві рівні половини:



$$\Delta_t = \frac{t_2 - t_1}{2}$$

У відповідності до одержаного значення  $\Delta_t$  позначають розтини, в яких з реалізацій знімають значення реалізацій випадкової функції.

У відповідності до виконаного визначення довжина інтервала дорівнює 1.0 г. Таким чином, випадкова функція буде представлена системою семи випадкових величин, розміщених у перетинах  $t = : 1.0; 2.0; 3.0; 4.0; 5.0; 6.0; 7.0$ . Позначимо ці перетини на графіку і знімаючи з графіка значення реалізацій випадкової функції в цих перетинах, отримуємо табл.. 19.

Таблиця 19

Номер реалізації	X (t)						
	0.0	1.0	2.0	3.0	4.0	5.0	6.0
1	1.0	1.1	1.2	1.18	1.05	1.23	1.25
2	1.04	1.26	1.08	0.6	0.85	0.95	1.1
3	1.11	1.25	0.85	0.9	1.02	1.08	1.08
4	0.52	1.35	1.37	1.2	0.95	0.9	1.0
5	1.21	0.9	1.0	1.2	1.3	1.33	1.27
6	1.3	0.9	0.8	1.08	1.15	1.0	0.8
7	0.9	0.75	1.03	0.8	0.85	0.8	0.9
$\Sigma$	7.08	7.51	7.33	8.13	7.17	7.29	7.4
m	1.01	1.07	1.04	1.16	1.02	1.04	1.06

В таблиці по строчках записані значення реалізацій у певних перетинах, а числа у стовпчиках представляють значення реалізацій у перетині.

За даними таблиці розраховуємо оцінки для характеристик випадкових величин  $X(0)$ ,  $X(0,4)$  ...,  $X(2,4)$ . Підсумовуючи значення по стовпцях і ділячи суму на число реалізацій  $n = 7$ , знайдемо приблизно залежність математичного очікування від часу, тобто запишемо закон розподілу математичного очікування - випадкової функції відносно зміни аргументу - часу:

$\Sigma$	7.08	7.51	7.33	8.13	7.17	7.29	7.4
m	1.01	1.07	1.04	1.16	1.02	1.04	1.06

На графіку, рис. 23, математичне очікування показане прямою лінією.

Розраховуємо дисперсії відповідних розтинів випадкової функції, для цього скористаємось формулою:

$$D(x) = m^2 - \left[ \frac{\Sigma}{n} \right]^2$$

Розрахунок виконуємо по стовбцях, тобто перетинах реалізацій випадкової функції, кожен елемент стовбця підносимо до степені і знаходимо середнє значення суми квадратів. Результат розрахунку в табл.. 20.

Таблиця 20

Номер реалізації	$X^2 (t)$						
	0.0	1.0	2.0	3.0	4.0	5.0	6.0
1	1.0	1.21	1.44	1.39	1.1	1.51	1.56
2	1.08	1.58	1.16	0.36	0.72	0.9	1.21
3	1.23	1.56	0.72	0.81	1.04	1.16	1.16
4	0.27	1.82	1.87	1.44	0.9	-0.81	1.0
5	1.46	0.81	1.0	1.44	1.69	1.76	1.61
6	1.69	0.81	0.64	1.16	1.32	1.0	0.64
7	0.81	0.56	1.06	0.64	0.72	0.64	0.81
$\Sigma$	7.54	8.35	7.89	7.24	7.49	7.78	7.99
$\bar{m}^2$	1.07	1.19	1.13	1.04	1.07	1.11	1.14

Значення чисел і результат розрахунку дисперсій в табл. 21.

Таблиця 21

Найменування оцінки	Оцінка						
	0.0	1.0	2.0	3.0	4.0	5.0	6.0
$\left[ \bar{m} \right]^2$	1.02	1.14	1.08	1.04	1.04	1.08	1.12
$\bar{m}^2$	1.07	1.19	1.13	1.03	1.07	1.11	1.14
$D(X)$	0.05	0.05	0.05	0.	0.03	0.03	0.02

Якщо з значень дисперсії добути корінь квадратний, то одержимо значення середніх квадратичних відхилень випадкового чинника у відповідних перетинах. Значення оцінок середнього квадратичного відхилення записані в табл. 22

Таблиця 22

$D(X)$	0.05	0.05	0.05	0.	0.03	0.03	0.02
$\sigma$	0.22	0.22	0.22	0	0.17	0.17	0.14

Розраховані оцінки одержані за даними поперечних перетинів окремих реалізацій випадкової функції.

Для розрахунку кореляційної функції необхідно провести її дослідження продовжних перетинів, тобто визначити характер - частоту коливань випадкової функції відносно деякого початкового значення.

Дослідження випадкової функції простіше виконувати, користуючись однією реалізацією, одержаною за достатньо великий проміжок часу. Або, якщо окремі реалізації з'єднати в одну по довжині часу, рівному сумі всіх реалізацій, то можна виконати її дослідження у продовжних перетинах, тобто замінити її відповідною випадковою величиною.

Якщо реалізації не з'єднувати, то дослідження стосовно кореляційної функції слід проводити окремо для кожної реалізації. Найбільш доцільним стосовно розуміння методу визначення кореляційної функції являється табличний метод. Форма таблиці для визначення кореляційної функції однієї реалізації наступна:

Таблиця 23

№ п/п	$x_i$	$x_i - \bar{X}$	$m=0$	$x_i - \bar{X}$	$m=1$	$m=2$	.	n-1
			$(x_i - \bar{X})^2$		$(x_i - \bar{X})$ $(x_{i+m} - \bar{X})$	$(x_i - \bar{X})$ $(x_{i+m} - \bar{X})$		
1	$x_1$	$x_1 - \bar{X}$	$(x_1 - \bar{X})^2$	$x_1 - \bar{X}$			.	
2	$x_2$	$x_2 - \bar{X}$	$(x_2 - \bar{X})^2$	$x_2 - \bar{X}$	$(x_1 - \bar{X})$ $(x_2 - \bar{X})$		.	
.	.	.	.	.	.	$(x_1 - \bar{X})$ $(x_3 - \bar{X})$	.	.
n	$x_n$	$x_n - \bar{X}$	$(x_n - \bar{X})^2$	$x_n - \bar{X}$	$(x_{n-1} - \bar{X})$ $(x_n - \bar{X})$	$(x_{n-1} - \bar{X})$ $(x_n - \bar{X})$	.	$(x_1 - \bar{X})$ $(x_n - \bar{X})$
N	$\bar{X}$						.	
R			$R_0$		$R_1$	$R_2$	.	$R_n$
r			$r_0$		$r_1$	$r_2$	.	$r_n$

Перший стовбець таблиці – це продовжна нумерація розтинів, другий - значення реалізації випадкової функції.

Кореляційна функція розраховується з використанням формули:

$$R_j = \frac{\sum_{i=1, j=0}^{N, M} (x_i - \bar{X}) \times (x_{i-m} - \bar{X})}{N - M}$$

В наведеній формулі  $\bar{X}$  - середнє значення випадкової функції – однієї реалізації для взятих поздовжних розтинів,  $x_i$  - значення випадкової функції – однієї реалізації для взятих поздовжних розтинів,  $i$  – порядковий номер значень випадкової функції, його значення приймаються  $i = 1, 2, 3, \dots, N$ ,  $j$  – числа, що вказують кількість інтервалів, для яких встановлюється кореляційна залежність, змінюється  $j = 0, 1, 2, 3, \dots, M$ ,  $M = N - 1$ .

Інші стовпці записані у відповідності до складових формули.

Значення автокореляційної функції -  $r_1$  одержані шляхом ділення кожного значення кореляційних функцій на  $R_0$ . В табл.. 24 показано фрагмент розрахунку кореляційної функції для однієї реалізації.

В табл.. 24 одержано значення автокореляційної функції з мінусовим знаком, тому подальший розрахунок можна не виконувати.

Аналогічно виконуються розрахунки для інших реалізацій. Пропонується виконати це самостійно.

Таблиця 24

№ Розтин ну	$x_i$	$x_i - \bar{X}$	m=0	$x_i - \bar{X}$	m=1	m=2	.	n-1
			$(x_i - \bar{X})^2$		$(x_i - \bar{X})$ $(x_{i+m} - \bar{X})$	$(x_i - \bar{X})$ $(x_{i+m} - \bar{X})$		$(x_i - \bar{X})$ $(x_{i+m} - \bar{X})$
0.	1.	- 0.14	0.02	- 0.14			.	
1.	1.1	- 0.04	0.002	- 0.04	0.0056		.	
2.	1.2	0.06	0.004	0.06	- 0.0024	- 0.0084	.	.
3.	1.18	0.04	0.002	0.04	0.0024	- 0.0016		
4.	1.05	- 0.11	0.01	- 0.11	- 0.0044	- 0.0066		
5.	1.23	0.09	0.008	0.09	- 0.0099	0.0036		
6.	1.25	0.11	0.01	0.11	0.0099	0.0121		
$\bar{X}$	1.14		0.056		0.0032	- 0.0003	.	
$R$			$R_{\sigma}=0.08$		0.000046	- 0.00005	.	$R_n$
$r$			$r_{\sigma}=1.$		0.0057	- 0.000625	.	$r_n$

Визначення загальних – усереднених характеристик – оцінок випадкової функції.

Проаналізуємо отримані дані щодо передбачуваної стаціонарності випадкової функції  $X(t)$ . Якщо судити з даних, отриманих в результаті обробки, то можна сказати, що випадкова функція  $X(t)$  не стаціонарна: її математичне очікування не постійне; дисперсія так само міняється з часом; значення нормованої кореляційної функції уздовж паралелей головної діагоналі також непостійні. Проте, зважаючи на обмежене число оброблених значень ( $n = 7$ ), будемо вважати, що випадкова функція стаціонарна.

Для того, щоб одержані оцінки достовірно відображали випадкову функцію, необхідно виконати розрахунки для всіх реалізацій і за одержаними результатами визначаються усереднені їх значення для математичного очікування:

$$\tilde{m}_x = \frac{\tilde{m}_x(0) + \tilde{m}_x() + \dots + \tilde{m}_x()}{7} = ,$$

для дисперсій:

$$\tilde{D}_x = \frac{\tilde{D}_x(0) + \tilde{D}_x() + \dots + \tilde{D}_x()}{7} = ,$$

Витягуючи корінь квадратний з дисперсії, знайдемо середнє квадратичне відхилення

$$\tilde{\sigma}_x$$

Що стосується кореляційної функції випадкової стаціонарної або не стаціонарної функції, то виконувати її усереднення по окремим реалізаціям недоречно. Причиною цього являється те, що значення реалізацій в перетинах різні, тому кореляційні функції також різні. Основне призначення кореляційної функції це визначення відстані між розтинами, при якій значення стають незалежними. Наприклад, при дослідженні родовища корисної копалини пробурюють свердловини і визначають чинники корисної копалини, зазвичай це вміст корисного мінералу. За одержаними значеннями відстані між

свердловинами та якості корисного мінерала розраховується кореляційна функція і будується її графік. При побудові графіка на вертикальній осі відкладаються значення кореляційної функції, а на горизонтальній - відстань між свердловинами. Якщо кореляційна функція перетне вісь абсцис у точці, яка розміщується ближче до нуля ніж відстань між свердловинами, то всі одержані оцінки стосовно родовища будуть неправильними. Тому при визначенні допустимого значення чинника за кореляційними функціями реалізацій потрібно поступати таким чином. На одному графіку будуються кореляційні функції поздовжних перетинів реалізацій і в якості оцінюваної приймається така, яка перетинає горизонтальну вісь найближче до нуля, дивись рис. 25.

Рис. 25. Графічне зображення кореляційної функції

### **3.7. Спектральне розкладання стаціонарної випадкової функції.**

#### **Спектр дисперсій**

Кореляційна функція являється характеристикою внутрішньої структури випадкової функції. В більшості випадків цієї характеристики не достатньо для вивчення випадкової функції. Як правило випадкова функція має певну кількість коливань різної амплітуди і частоти, тому виникає потреба визначити ,які співвідношення переважають у складі випадкової функції, ті або інші коливання - спектри частот. Для характеристики цього введено ще одне поняття – характеристика - спектральна щільність. Доречне уточнення: з попереднього викладення відомо, що кореляційна функція відображає в деякій мірі коливання випадкової функції. За результатами аналізу кореляційної функції вирішується багато практичних питань задач, наприклад, через які проміжки часу слід відбирати проби продуктів технологічної схеми, щоб одержані значення показників були достовірними – спроможними, при визначенні точності вимірювальних приладів, та багато інших. По суті кореляційна функція для випадкового процесу визначає її властивість зміни знаку. Вказана властивість базується на зміні значень дисперсії. Якщо при дослідженні випадкової функції на певному проміжку часу, або деякого чинника збільшувати кількість перетинів, то кількість дисперсій буде збільшуватись і вони будуть відрізнятись між собою тим більше, чим більша частота коливань випадкової функції. В

багатьох практичних задачах виникає потреба детального дослідження зміни частот коливань – зміни значень дисперсій.

Оскільки кореляційна функція представляє зміну дисперсій, то спочатку розглянемо дослідження спектральної щільності для кореляційної функції випадкової функції. Розподіл значень дисперсій по певним частотам називають спектром дисперсій.

Для визначення спектру дисперсій застосовують перетворення Фуре. При цьому математичний вираз розкладення такий:

$$K_x(t, t') = \sum_{k=1}^{\infty} (D_k \cos \omega_k t' \cos \omega_k t + D_k \sin \omega_k t' \sin \omega_k t)$$

У більшості практичного вирішення питань користуються нормованими числовими характеристиками, в цьому випадку правомірне співвідношення:

$$\left. \begin{aligned} r_x(\tau) &= \int_0^{\infty} s_x(\omega) \cos \omega t d\omega \\ s_x(\omega) &= \frac{2}{\pi} \int_0^{\infty} r_x(\tau) \cos \omega t d\tau \end{aligned} \right\}$$

Приклад спектрального представлення кореляційної функції і частотної характеристики показаний на рис. 26.



Рис. 26. Спектральна щільність випадкової функції

Іноді для вирішення питань перетворення стаціонарних випадкових функцій виконують в комплексній формі.

Якщо виконано перетворення кореляційної функції випадкової функції, то представляється можливим здійснити перетворення самої випадкової функції.

#### 4. Планування експерименту

У збагаченні корисних копалин і пов'язаній з нею гірничо - добувній промисловості, технологічні процеси і апарати, які можуть бути виділені в окремі об'єкти досліджень і управлінь, являються складними внаслідок їх багатofакторності і випадковості. Окрім того, що об'єкти багаточинникові, самі чинники відносяться до різних розмірних груп, а це значною мірою ускладнює проведення наукових досліджень, отже і здійснення їх керування - управління.

Якщо для певного об'єкту необхідно виконати його формалізацію і єдиним методом є такий, як використання експериментальних статистичних даних, то в цьому випадку необхідно провести експериментальні дослідження і такі дані отримати. В результаті таких досліджень визначають закономірності поведінки об'єкту при зміні параметрів. Значення параметрів в процесі проведення досліджень інтуїтивно встановлювати досить складно, в цьому випадку слід скористатися існуючою методикою проведення досліджень багаточинникових об'єктів. Методикою передбачається складання певного плану проведення експерименту, обробки отриманих даних і перевірки отриманого результату. Така методика отримала назву «Планований факторний експеримент (ФЕ)», є дві різновидності даної методики - «повний факторний експеримент» (ПФЕ) і «дробовий факторний експеримент» (ДФЕ).

#### **4.1. Основи методу**

У переважній більшості значення чинників - факторів, що визначають суть об'єкту, мають випадковий характер, тому в якості їх значень використовуються оцінки, а основною математичною базою методу є положення теорії ймовірності і математичної статистики.

##### **4.1.1. Повний факторний експеримент**

Експериментальне дослідження полягає в тому, що встановлюється певне значення одного з параметрів і фіксується, як зміниться значення іншого, якщо об'єкт функціонує. Суть планованого експерименту полягає в тому, що однозначно встановлюється, скільки брати параметрів, скільки значень кожного параметра, і які по величині ці значення мають бути.

##### **4.1.2. Визначення числа дослідів, які необхідно провести при планованому експерименті**

Кількість змін однакова для всіх параметрів. Якщо позначити через  $k$  кількість змін, а через  $n$  кількість параметрів, то кількість дослідів  $m$ , які необхідно поставити, буде дорівнювати:

$$m = k^n$$

Наприклад, якщо в дослідженні розглядається залежність деякого параметра  $Y$  від параметрів  $X_1$  і  $X_2$  ( $n = 2$ ), при цьому для кожного з параметрів  $X_1$  і  $X_2$  взяти по два ( $k = 2$ ) різних значення, то число дослідів складе:

$$m = k^n = 2^2 = 4$$

Якщо проводити дослідження параметра  $Y$  від параметрів  $X_1$ ,  $X_2$  і  $X_3$  ( $n = 3$ ), а кожен з трьох параметрів  $X_1$ ,  $X_2$  і  $X_3$  змінювати так само по два ( $k = 2$ ) рази, то число дослідів складе:

$$m = k^n = 2^3 = 8$$

При збільшенні числа змінних параметрів,  $m$  - число дослідів, які необхідно виконати, збільшуватиметься у функції  $k^n$

Встановлення конкретних значень параметрів, які змінюють в процесі експерименту.

Розглянемо конкретний приклад. Продуктивність барабанного кульового млина по готовому класу -  $Q_r$ , що подрібнює залізовміщуючу руду, залежить від наступних параметрів: кількості руди, яку подають на подрібнення -  $Q_p$ , кількості води, що подається в млин -  $Q_v$ , числа обертів барабана млина -  $\omega$ . Насправді, даний процес характеризується набагато більшою кількістю параметрів.

Для того, щоб встановити кількісно міру дії кожного з параметрів на продуктивність млина по готовому класу досить провести експериментальні дослідження таким чином: змінювати параметри;  $Q_p$ ,  $Q_v$  і  $\omega$  відповідним чином при цьому вимірювати величину параметра  $Q_r$ . При проведенні планованого експерименту кількість змін кожного з параметрів -  $k$  має бути однаковим, прийmemo, що  $k = 2$ . Число параметрів - чинників -  $n$ , в нашому випадку дорівнює трьом, це  $Q_p$ ,  $Q_v$  і  $\omega$ . Визначимо, яку кількість дослідів необхідно виконати при планованому експерименті:

$$m = k^n = 2^3 = 8$$

Ті два, в даному випадку, конкретні значення, які приймаються для кожного параметра, називають рівнями. Величину цих значень можна прийняти, скориставшись практичними даними роботи даного апарату, його технічною характеристикою або технологічним регламентом.

Спочатку встановлюються нульові значення чинників, які будуть змінюватись. З даних практичної роботи млина - МШР - 4500 x 5000 мм., що переробляє аналогічну сировину, маємо:

$$Q_p = 100 \text{ т/г}, Q_v = 40 \text{ т/г} \text{ з технічної характеристики } \omega = 16,6 \text{ об/хв.}$$

Дані значення чинників вважатимемо за нульові.

Рівні - значення чинників встановлюємо таким чином. В експерименті -  $Q_p$  змінимо на  $\Delta Q_p = 10$  т/г, що складе 10% від нульового значення у більшу і меншу сторони, тобто в досліді встановлюватимемо два значення.

$$Q_{pм} = Q_p - \Delta Q_p = 100 - 10 = 90 \text{ т/г} - \text{менший (нижній рівень) і}$$

$$Q_{pб} = Q_p + \Delta Q_p = 100 + 10 = 110 \text{ т/г} - \text{більший (верхній рівень).}$$

Таким же чином визначаються рівні для параметра  $Q_v$  і  $\omega$ , тобто на десять відсотків у більшу і меншу сторони.

$$Q_{vм} = Q_v - \Delta Q_v = 40 - 4 = 36 \text{ т/г} - \text{менший (нижній рівень) і}$$

$$Q_{vб} = Q_v + \Delta Q_v = 40 + 4 = 44 \text{ т/г} - \text{більший (верхній рівень)}$$

$$\omega_m = \omega - \Delta \omega = 16,6 - 1,66 = 14,94 \text{ об/хв і}$$

$$\omega_b = \omega + \Delta \omega = 16,6 + 1,66 = 18,26 \text{ об/хв}$$

У планованому експерименті встановлюється процедура проведення дослідів, тобто в якому порядку з наявних рівнів - конкретних чисел, виключаючи в поєднаннях однойменні параметри, брати числа - рівні різнойменних параметрів і проводити досвід.



### 4.1.3. План експерименту

Процедуру – план проведення дослідів, зручно представити у вигляді таблиці, заповнення якої має свої особливості, див. Табл. 25.

Таблиця 25

№ досліду	Рівні чинників		
	$Q_p$ , т/г	$Q_b$ , т/г	$\omega$ , об/хв
1	90	36	14.94
2	110	36	14.94
3	90	44	14.94
4	110	44	14.94
5	90	36	18.26
6	110	36	18.26
7	90	44	18.26
8	110	44	18.26

Особливістю таблиці планованого експерименту є те, що зміна значень – нижнього і верхнього рівнів першого записаного в стовпці таблиці параметра чередується. Чередування рівнів другого параметра, значення яких записані в другий стовпець таблиці, має удвічі меншу частоту в порівнянні з першим. Частота третього удвічі менша по порівнянню з частотою другого. Ця властивість виконується для будь-якої кількості чинників.

### 4.2. Проведення експериментів

Користуючись складеною таблицею 25, проводяться досліді. Для проведення технологічного процесу подрібнення встановлюються значення чинників такі, які записані в рядку для відповідного досліді і потім проводиться дослід. Тобто один рядок це один дослід. Досліді можна проводити в різній послідовності, не дотримуючись тієї, яка розміщена в таблиці планування, причому кожен дослід необхідно провести кілька разів. В цьому випадку встановлюють рандомізацію дослідів, тобто послідовність їх проведення. Для встановлення рандомізації можна скористатися таблицею випадкових чисел, з якої виписуються стільки чисел, скільки дослідів має бути проведено, а потім кожен дослід проводять в тій послідовності, який номер йому привласнений. Якщо кожен дослід повинен повторюватися три рази, то таблицю доповнюють трьома стовпцями для кодування номерів дослідів і трьома стовпцями для запису результатів експерименту. У нашому випадку таблиця виглядає наступним чином, табл. 26.

Таблиця 26

№ дослід- ду, і	Послідовність проведення, j			Рівні чинників			Значення вихідного чинника		
	N <sub>1</sub>	N <sub>2</sub>	N <sub>3</sub>	q <sub>p</sub> , т/Г	q <sub>в</sub> , т/Г	ω, м/с	Q <sub>г.</sub> , т/Г	Q <sub>г.</sub> , т/Г	Q <sub>г.</sub> , т/Г
1				90	36	14.9			
2				110	36	14.9			
3				90	44	14.9			
4				110	44	14.9			
5				90	36	18.2			
6				110	36	18.2			
7				90	44	18.2			
8				110	44	18.2			

При проведенні дослідів у всіх випадках слід враховувати особливості об'єкту досліджень. Наприклад, в нашому випадку відомо, що млин це апарат з певною інерційністю. Тому слід визначити тривалість часу знаходження матеріалу в млині і вимірювання кількості готового класу на його виході проводити тільки після закінчення встановленого часу. У інших випадках можуть бути інші особливості об'єкту.

У багатьох джерелах таблицю планованого експерименту представляють в так званих умовних рівнях. Форма таблиці в цьому випадку не міняється, такою ж залишається процедура проведення експерименту.

Перехід значень рівнів до умовних розглянемо для одного параметра, наприклад, кількості руди, що подається на подрібнення. Значення вказаного параметра:

$$Q_{рм} \quad Q_p \quad Q_{рб}, \quad \text{інакше} \quad 90 \quad 100 \quad 110$$

В даному випадку сукупністю чисел є варіаційний ряд з рівновіддаленими варіантами і кроком  $\Delta h = 10$ .

Значення умовних варіант розраховуються таким чином:

$$u_i = \frac{q_i - q_0}{\Delta h}$$

де  $u_i$  – умовна варіанту,  $q_i = Q$  – поточне значення варіанти, тобто нижній або верхній рівень,  $q_0 = Q$  – нульове значення – нульовий рівень, до якого приводяться варіанти, в якості нульового приймається  $q_0 = 100$ . У нашому випадку:

$$u_1 = \frac{q_1 - q_0}{\Delta h} = \frac{90 - 100}{10} = -1$$

$$u_1 = \frac{q_1 - q_0}{\Delta h} = \frac{110 - 100}{10} = +1$$

Приведений варіаційний ряд виглядатиме:

Оскільки в дослідах нульовий рівень не враховується, то з урахуванням значень верхнього і нижнього рівнів і з урахуванням того, що в кожному параметрі рівні мають значення рівне одиниці і їх можна не враховувати, то в таблиці планованого експерименту (табл. 27) записують тільки знаки таким чином:

Таблиця 27

№ досліджу	Рівні чинників			Вихідний параметр Q <sub>г.</sub> , т / г
	Q <sub>р</sub> , т/г	Q <sub>в</sub> , т/г	ω, об/хв	
1	-	-	-	
2	+	-	-	
3	-	+	-	
4	+	+	-	
5	-	-	+	
6	+	-	+	
7	-	+	+	
8	+	+	+	

#### 4.2.1. Обробка результатів експерименту

Отримані дані експерименту можуть використовуватися різним чином залежно від мети дослідження. Наприклад, якщо потрібно максимізувати вихід готового класу, то надалі процес слід проводити при складових параметрів (відповідного досліджу), при яких отриманий максимальний вихід.

Якщо потрібно формалізувати процес, то на підставі експериментальних даних можна отримати його математичну модель.

На жаль, до теперішнього часу відсутні однозначні рекомендації по вибору виду залежності, яка б щонайкраще відображала реальний об'єкт.

Для планованого експерименту трьохчинникового, який прийнято називати повним факторним експериментом (ПФЕ) математичну залежність можна отримати, використовуючи метод множинної кореляції. Або представити її складовими розкладання в ряд Тейлора.

Обробку експериментальних даних зручно виконувати, користуючись таблицею-табл. 28, в якій представлені рівні параметрів, результати дослідів і результати розрахунків.

У таблиці всі дані трьох передостанніх стовпців - це результати дослідів, а дані останнього стовпця розраховані для кожного рядка по виразу:

$$\bar{Q}_i = \frac{\sum_{k=1}^k Q_k}{k},$$

де  $k$  – число повторень кожного планованого досліджу.

Далі для стовпців ідентифікованих від  $X_0$  до  $X_7$  розраховують коефіцієнти  $p_i$ , скориставшись формулою:

$$\rho_i = \frac{\sum_{i=1}^N X_i \bar{Q}_i}{N}$$

У даному виразі під знаком суми знаходиться добуток відповідного знаку в рядку стовпця на середнє значення вихідного параметра, наприклад, для  $X_1$  вираз для визначення коефіцієнта:

$$\rho_i = \frac{\sum_{i=1}^N X_i \bar{Q}_i}{N} = \frac{(-)Q_1 + (+)Q_2 + (-)Q_3 + \dots + (+)Q_8}{8}$$

Після обчислення всіх коефіцієнтів можна представити об'єкт у вигляді повної математичної моделі:

$$Q = \rho_0 + \rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3 + \rho_{12} x_1 x_2 + \rho_{13} x_1 x_3 + \rho_{23} x_2 x_3 + \rho_{123} x_1 x_2 x_3,$$

або у вигляді лінійної моделі:

$$Q = \rho_0 + \rho_1 x_1 + \rho_2 x_2 + \rho_3 x_3,$$

Таблиця 28

№ досл іду	Послідов ність проведення			Чинники, що змінюються, та їх позначення								Вихідний чинник			Середнє значення $\bar{Q}_r$
				$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$Q_r$	$Q_r$	$Q_r$	
	$X_0$	$q_p$	$q_v$	$\omega$	$q_p q_v$	$q_p \omega$	$q_v \omega$	$q_p q_v \omega$							
1	2	5	4	+	-	-	-	+	+	+	-	$Q_{r2}$	$Q_{r5}$	$Q_{r4}$	$\bar{Q}_1$
2	8	3	1	+	+	-	-	-	-	+	+	$Q_{r8}$	$Q_{r3}$	$Q_{r1}$	$\bar{Q}_2$
3	6	7	5	+	-	+	-	-	+	-	+	$Q_{r6}$	$Q_{r7}$	$Q_{r5}$	$\bar{Q}_3$
4	1	4	3	+	+	+	-	+	-	-	-	$Q_{r1}$	$Q_{r4}$	$Q_{r3}$	$\bar{Q}_4$
5	3	2	8	+	-	-	+	+	-	-	+	$Q_{r3}$	$Q_{r2}$	$Q_{r8}$	$\bar{Q}_5$
6	5	1	6	+	+	-	+	-	+	-	-	$Q_{r5}$	$Q_{r1}$	$Q_{r6}$	$\bar{Q}_6$
7	7	8	2	+	-	+	+	-	-	+	-	$Q_{r7}$	$Q_{r8}$	$Q_{r2}$	$\bar{Q}_7$
8	4	6	7	+	+	+	+	+	+	+	+	$Q_{r4}$	$Q_{r6}$	$Q_{r7}$	$\bar{Q}_8$

#### 4.2.2. Оцінка результатів планованого експерименту

Кінцевою метою кожного експериментального дослідження є встановлення кількісної міри взаємного впливу чинників і якщо це можливо - представлення даного впливу у вигляді математичного виразу, який з певною достовірністю замінює об'єкт. Величину достовірності такої заміни слід оцінити, скориставшись відповідним критерієм згоди.

Разом з позитивною очевидністю правильної організації проведення експерименту у разі його попереднього планування не слід покладати на це великих надій. По-перше, при планованому експерименті можуть бути не вірно вибрані нульові рівні, а також інтервали варіювання і друге, варіювання на двох рівнях не завжди є достатнім. Об'єм робіт при планованому експерименті значний. І якщо в першому експерименті досліджувана область не відповідає оптимальному співвідношенню параметрів, то експеримент вимагає повторення і пошуку нульових рівнів з урахуванням отриманих результатів попереднього експерименту. Слід зазначити, що для збагачувальних процесів і апаратів

планований експеримент в реальному технологічному процесі проводити дуже складно, зважаючи на складність встановлення необхідних значень рівнів. Об'єм робіт в цьому випадку можливо скоротити, якщо скористатися так званим дробовим експериментом (ДФЕ), який є певною частиною ПФЕ.

### 4.3. Дробовий експеримент чинника

Ідея ДФЕ полягає в тому, що для дослідження з використанням двох чинників береться ПФЕ двохчинник типу  $2^2$  і в ньому добуток  $x_1x_2$  прирівнюється третьому чиннику. В цьому випадку дробовий планований експеримент для об'єкту з трьома варійованими чинниками представляється таблицею планування – табл. 29. При цьому таблиця ПФЕ ділиться на дві частини – табл. 28. А потім по одній з них проводиться експеримент. Таблиця ДФЕ складається таким чином.

Таблиця 29

Номер досліджу	$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
	$x_0$	$X_1$	$x_2$	$x_3$	$x_1x_2$	$x_1x_3$	$x_2x_3$	$x_1x_2x_3$
1	+	-	-	+	+	-	-	+
2	+	+	-	-	-	-	+	+
3	+	-	+	-	-	+	-	+
4	+	+	+	+	+	+	+	+

У клітинки стовпців ( $x_1, x_2, x_3$ ) таблиці ДФЕ рівні в приведених одиницях («+», «-») беруться з відповідних стовпців таблиці ПФЕ з тих рядків, в яких є один верхній рівень і з рядка, в якому всі рівні верхні. Значення в клітинках ( $X_4, X_5, X_6, X_7$ ) отримують шляхом перемножування відповідних знаків для рівнів даного рядка.

Проведення експерименту і обробка результатів здійснюється аналогічно тому, як це робиться для ПФЕ.

Додаток 1

Таблиця значень диференційної функції нормально розподіленої випадкової величини при  $\mu = 0$  і  $\sigma = 1$ ,  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

X	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	3723	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	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	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	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013

## Закінчення таблиці - Додаток 1

X	0	1	2	3	4	5	6	7	8	9
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблиця значень інтегральної функції Лапласа  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$

X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$
0,00	0,0000	0,45	0,1736	0,90	0,3159	1,35	0,4115
0,01	0,0040	0,46	0,1772	0,91	0,3186	1,36	0,4131
0,02	0,0080	0,47	0,1808	0,92	0,3212	1,37	0,4147
0,03	0,0120	0,48	0,1844	0,93	0,3238	1,38	0,4162
0,04	0,0160	0,49	0,1879	0,94	0,3264	1,39	0,4177
0,05	0,0199	0,50	0,1915	0,95	0,3289	1,40	0,4192
0,06	0,0239	0,51	0,1950	0,96	0,3315	1,41	0,4207
0,07	0,0279	0,52	0,1985	0,97	0,3340	1,42	0,4222
0,08	0,0319	0,53	0,2019	0,98	0,3365	1,43	0,4236
0,09	0,0359	0,54	0,2054	0,99	0,3389	1,44	0,4251
0,10	0,0398	0,55	0,2088	1,00	0,3413	1,45	0,4265
0,11	0,0438	0,56	0,2123	1,01	0,3438	1,46	0,4279
0,12	0,0478	0,57	0,2157	1,02	0,3461	1,47	0,4292
0,13	0,0517	0,58	0,2190	1,03	0,3485	1,48	0,4306
0,14	0,0557	0,59	0,2224	1,04	0,3508	1,49	0,4319
0,15	0,0596	0,60	0,2257	1,05	0,3531	1,50	0,4332
0,16	0,0636	0,61	0,2291	1,06	0,3554	1,51	0,4345
0,17	0,0675	0,62	0,2324	1,07	0,3577	1,52	0,4357
0,18	0,0714	0,63	0,2357	1,08	0,3599	1,53	0,4370
0,19	0,0753	0,64	0,2389	1,09	0,3621	1,54	0,4382
0,20	0,0793	0,65	0,2422	1,10	0,3643	1,55	0,4394
0,21	0,0832	0,66	0,2454	1,11	0,3665	1,56	0,4406
0,22	0,0871	0,67	0,2486	1,12	0,3686	1,57	0,4418
0,23	0,0910	0,68	0,2517	1,13	0,3708	1,58	0,4429
0,24	0,0948	0,69	0,2549	1,14	0,3729	1,59	0,4441
0,25	0,0987	0,70	0,2580	1,15	0,3749	1,60	0,4452
0,26	0,1026	0,71	0,2611	1,16	0,3770	1,61	0,4463
0,27	0,1064	0,72	0,2642	1,17	0,3790	1,62	0,4474
0,28	0,1103	0,73	0,2673	1,18	0,3810	1,63	0,4484
0,29	0,1141	0,74	0,2703	1,19	0,3830	1,64	0,4495
0,30	0,1179	0,75	0,2734	1,20	0,3849	1,65	0,4505
0,31	0,1217	0,76	0,2764	1,21	0,3869	1,66	0,4515
0,32	0,1255	0,77	0,2794	1,22	0,3883	1,67	0,4525
0,33	0,1293	0,78	0,2823	1,23	0,3907	1,68	0,4535
0,34	0,1331	0,79	0,2852	1,24	0,3925	1,69	0,4545
0,35	0,1368	0,80	0,2881	1,25	0,3944	1,70	0,4554
0,36	0,1406	0,81	0,2910	1,26	0,3962	1,71	0,4564
0,37	0,1443	0,82	0,2939	1,27	0,3980	1,72	0,4573



X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$
0,38	0,1480	0,83	0,2967	1,28	0,3997	1,73	0,4582
0,39	0,1517	0,84	0,2995	1,29	0,4015	1,74	0,4591
0,40	0,1554	0,85	0,3023	1,30	0,4032	1,75	0,4599
0,41	0,1591	0,86	0,3051	1,31	0,4049	1,76	0,4608
0,42	0,1628	0,87	0,3078	1,32	0,4066	1,77	0,4616
0,43	0,1664	0,88	0,3106	1,33	0,4082	1,78	0,4625
0,44	0,1700	0,89	0,3133	1,34	0,4099	1,79	0,4633
1,80	0,4641	2,00	0,4772	2,40	0,4918	2,80	0,4974
1,81	0,4649	2,02	0,4783	2,42	0,4922	2,82	0,4976
1,82	0,4656	2,04	0,4793	2,44	0,4927	2,84	0,4977
1,83	0,4664	2,06	0,4803	2,46	0,4931	2,86	0,4979
1,84	0,4671	2,08	0,4812	2,48	0,4934	2,88	0,4980
1,85	0,4678	2,10	0,4821	2,50	0,4938	2,90	0,4981
1,86	0,4686	2,12	0,4830	2,52	0,4941	2,92	0,4982
1,87	0,4693	2,14	0,4838	2,54	0,4945	2,94	0,4984
1,88	0,4699	2,16	0,4846	2,56	0,4948	2,96	0,4985
1,89	0,4706	2,18	0,4854	2,58	0,4951	2,98	0,4986
1,90	0,4713	2,20	0,4861	2,60	0,4953	3,00	0,49865
1,91	0,4719	2,22	0,4868	2,62	0,4956	3,20	0,49931
1,92	0,4726	2,24	0,4875	2,64	0,4959	3,40	0,49966
1,93	0,4732	2,26	0,4881	2,66	0,4961	3,60	0,499841
1,94	0,4738	2,28	0,4887	2,68	0,4963	3,80	0,409928
1,95	0,4744	2,30	0,4893	2,70	0,4965	4,00	0,499968
1,96	0,4250	2,32	0,4898	2,72	0,4967	4,50	0,499997
1,97	0,4756	2,34	0,4904	2,74	0,4969	5,00	0,499997
1,98	0,4761	2,36	0,4909	2,76	0,4971		
1,99	0,4767	2,38	0,4913	2,78	0,4973		

### Додаток 3

Таблиця значень  $t_\gamma = t(\gamma, n)$  для інтервальної оцінки математичного сподівання випадкової величини

$n$	$\gamma$	0,95	0,99	0,999	$n$	$\gamma$	0,95	0,99	0,999
5		2,78	4,60	8,61	20		2,093	2,861	3,883
6		2,57	4,03	6,86	25		2,064	2,797	3,745
7		2,45	3,71	5,96	30		2,045	2,756	3,659
8		2,37	3,50	5,41	35		2,032	2,729	3,600
9		2,31	2,36	5,04	40		2,023	2,708	3,558
10		2,26	3,25	4,78	45		2,016	2,692	3,527
11		2,23	3,17	4,59	50		2,009	2,679	3,502
12		2,20	3,11	4,44	60		2,001	2,662	3,464
13		2,18	3,06	4,32	70		1,996	2,649	3,439
14		2,16	3,01	4,22	80		1,001	2,640	3,418
15		2,15	2,98	4,14	90		1,987	2,633	3,403
16		2,13	2,95	4,07	100		1,98	2,627	3,392
17		2,12	2,92	4,02	120		1,980	2,617	3,374
18		2,11	2,90	3,97	$\infty$		1,960	2,576	3,291
19		2,10	2,88	3,92					

### Додаток 4

Таблиця значень  $q = q(\gamma, n)$  для інтервальної оцінки середньквдратичного відхилення

$n$	$\gamma$	0,95	0,99	0,999	$n$	$\gamma$	0,95	0,99	0,999
5		1,37	2,67	5,64	20		0,37	0,58	0,88
6		1,09	2,01	3,88	25		0,32	0,49	0,73
7		0,92	1,62	2,98	30		0,28	0,43	0,63
8		0,80	1,38	2,42	35		0,26	0,38	0,56
9		0,71	1,20	2,06	40		0,24	0,35	0,50
10		0,65	1,08	1,80	45		0,22	0,32	0,46
11		0,59	0,98	1,60	50		0,21	0,30	0,43
12		0,55	0,90	1,45	60		0,188	0,269	0,38
13		0,52	0,83	1,33	70		0,174	0,245	0,34
14		0,48	0,78	1,23	80		0,161	0,226	0,31
15		0,46	0,73	1,15	90		0,151	0,211	0,29
16		0,44	0,70	1,07	100		0,143	0,198	0,27
17		0,42	0,66	1,01	150		0,115	0,160	0,211
18		0,40	0,63	0,96	200		0,099	0,136	0,185
19		0,39	0,60	0,92	250		0,089	0,120	0,162

Критичні точки розподілу  $\chi^2$ 

Число степенів вільності $k$	Рівень значимості $\alpha$					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

## Критичні точки розподілу Стьюдента

Число степенів вільності $k$	Рівень значимості $\alpha$ (двостороння критична область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
$\infty$	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005

Критичні точки розподілу Фішера – Снедекора:

$k_1$  – число степенів вільності більшої дисперсії;

$k_2$  – число степенів вільності меншої дисперсії.

Рівень значення $\alpha = 0,01$												
$k_1$	1	2	3	4	5	6	7	8	9	10	11	12
$k_2$												
1	4052	4999	5403	5625	5764	5889	5928	5981	6022	6056	6082	6106
2	98,49	99,01	99,17	99,25	99,33	99,30	99,34	99,36	99,36	99,40	99,41	99,42
3	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,85	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45

## Література

1. Е.С.Вентцель. Теория вероятностей. – М.: Наука, 1969.-576 с.
2. В.Е.Гмурман. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1972. – 368 с.
3. Иванова В.М., Нешумова Л.А., Решетникова И.О. Математическая статистика. – М.: Высшая школа, 1981.-368 с.
4. П.І. Пілов, М.Т. Анісімов, В.М. Анісімов. Математичне моделювання процесів збагачення корисних копалин. – Дніпропетровськ.: НГУ, 2005. – 103 с.

## Предметний покажчик

Автокореляційна функція .....	94
Асиметрія .....	65
Випадкові величини .....	26
Випадкові функції .....	26
Величини одновимірні .....	26
Величини багатовимірні .....	32
Генеральна сукупність .....	39
Гістограма .....	46
Дисперсія .....	31
Довірча ймовірність .....	56
Закон Гауса .....	31
Закон Бернуллі .....	22
Залежність кореляційна .....	36, 68
Залежність лінійна .....	72
Залежність нелінійна.....	75
Інтегральна теорема Лапласа .....	23
Кореляція .....	36
Критерій згоди .....	78
Локальна теорема Лапласа .....	23
Математичне очікування .....	31
Матриця планування .....	112
Оцінка .....	46
Події .....	7
Реалізація .....	88
Репрезентативність .....	46
Середньоквадратичне відхилення .....	31
Сімейство реалізацій .....	88
Стаціонарність .....	98
Умовні варіанти .....	110

Навчальне видання

**Пілов** Петро Іванович

**Анісімов** Микола Тимофійович

**Анісімов** Владислав Миколайович

Методи математичної статистики та теорії  
ймовірностей в збагаченні корисних копалин

Навчальний посібник

Редактор

Підписано до друку 2011. Формат 30Х42/4  
Папір офсетний. Ризографія. Ум. друк. арк..11.2.  
Обл. – вид. арк.. 11.2. Тираж 100 прим. Зам.

Підписано до друку та видруковано  
у Державному вищому навчальному закладі  
«Національний гірничий університет».  
Свідоцтво про внесення до Державного реєстру  
ДК №

49027. Дніпропетровськ – 27,  
Просп.. К.Маркса, 19