

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

**Лабораторний практикум
з навчальної дисципліни
"ТЕОРІЯ ЙМОВІРНОСТЕЙ
ТА МАТЕМАТИЧНА СТАТИСТИКА"**

Навчальний посібник

**Харків
ХНЕУ ім. С. Кузнеця
2016**

УДК 519.21(075)

ББК 22.171я7

Л 12

Рецензенти: д-р фіз.-мат. наук, професор кафедри математичних методів в економіці Харківського національного університету імені В. Н. Каразіна *А. А. Янцевич*; д-р техн. наук, професор, завідувач кафедри комп'ютерної математики і математичного моделювання Національного технічного університету "Харківський політехнічний інститут" *Л. М. Любчик*.

Рекомендовано до видання рішенням вченої ради Харківського національного економічного університету імені Семена Кузнеця.

Протокол № 7 від 15.02.2016 р.

Лабораторний практикум з навчальної дисципліни "Теорія Л 12 ймовірностей та математична статистика" : навчальний посібник / Е. Ю. Железнякова, І. Л. Лебедєва, Л. О. Норік, К. В. Степанова. – Харків : ХНЕУ ім. С. Кузнеця, 2016. – 184 с.

ISBN 978-966-676-633-8

Наведено лабораторні роботи, перелік яких повністю охоплює програму навчальної дисципліни та її зміст за модулями й темами. Стисло викладено основні теоретичні відомості з базових питань навчальної дисципліни та подано методичні рекомендації щодо розв'язання типових задач за допомогою електронних таблиць MS Excel 2010. За кожною лабораторною роботою подано завдання для самостійного розв'язання та питання для самоперевірки.

Рекомендовано для студентів усіх напрямів підготовки денної форми навчання.

УДК 519.21(075)

ББК 22.171я7

© Е. Ю. Железнякова, І. Л. Лебедєва,
Л. О. Норік, К. В. Степанова, 2016

© Харківський національний економічний
університет імені Семена Кузнеця, 2016

ISBN 978-966-676-633-8

Вступ

"Теорія ймовірностей та математична статистика" є базовою навчальною дисципліною природничо-наукового циклу. Її вивчення здійснюється згідно з навчальним планом підготовки фахівців освітньо-кваліфікаційного рівня "бакалавр" галузей знань 0305 "Економіка та підприємництво", 0306 "Менеджмент і адміністрування" денної форми навчання.

Метою вивчення навчальної дисципліни є формування цілісної системи теоретичних знань математичного апарату, необхідного для розв'язання теоретичних та практичних задач у професійній діяльності компетентного фахівця у сфері економіки та менеджменту, вміння аналітичного мислення та навичок застосування математичного апарату до формалізації реальних процесів і явищ, а також до розв'язання економічних задач.

Завдання навчальної дисципліни полягає в засвоєнні студентами теоретичної бази методології статистичних досліджень і отримання практичних навиків їх використання при розв'язанні прикладних задач економіки, набуття досвіду самостійної роботи з літературою як з математики, так і з прикладних питань економіки, а також застосування Інтернет-ресурсів.

Предметом даної навчальної дисципліни є фундаментальні положення теорії ймовірностей щодо дослідження випадкових подій і визначення їх ймовірності, основні закони розподілу дискретної на неперервної випадкових величин, їх числові характеристики, теорія масового обслуговування, а також принципи формування вибіркової сукупності, реєстрації, опису та аналізу результатів статистичних спостережень, основи теорії кореляції та регресійного аналізу.

У процесі опанування даної дисципліни студент отримує комплекс аналітично-дослідницьких знань, вмінь та навичок, а саме: набуває навичок обробки статистичних даних, оцінювання числових характеристик одновимірної та двовимірної випадкових величин, формулювання статистичних гіпотез та побудови статистичних критеріїв для їх перевірки.

Знання та навички, набуті студентом у результаті вивчення навчальної дисципліни, складають основу аналітично-дослідницьких **компетентностей** сучасного фахівця в галузі економіки та менеджменту у контексті Національної рамки кваліфікацій.

Серед них можна виділити такі:

вміння застосовувати набуті теоретичні знання до розв'язання ситуаційних вправ та реальних практичних задач;

навички самостійно визначати алгоритм математичних обчислень;

навички самостійно працювати із науково-методичною літературою;

вміння у випадку складного завдання використовувати метод розкладання від складного до простого, тобто зведення окремих складових вихідного завдання до більш простих;

вміння відокремлювати незалежні та залежні фактори, визначати числові характеристики розподілу випадкової величини, проводити аналіз результатів із метою уточнення типу та форми зв'язку між факторами;

вміння аналізувати отримані результати та з їх урахуванням робити обґрунтовані висновки на достатньо високому професійному рівні;

навички застосування програмного середовища *MS Excel* у дослідженні економічних процесів;

здатність автономно і відповідально виконувати завдання;

комунікації щодо взаємодії студентів у процесі спільного розв'язання проблем і прийняття узгоджених рішень.

Навчальна дисципліна є необхідним ланцюгом неперервної аналітико-математичної підготовки економістів. Вона передуює вивченню дисциплін економічного спрямування, які передбачають використання інструментарію математичної статистики для розв'язання суто економічних задач.

Застосування такої форми навчання як лабораторні роботи допомагає вдосконалити знання програмного середовища *MS Excel*, що в подальшому сприятиме використанню комп'ютерних технологій у професійній діяльності під час розв'язання економічних задач великої вимірності.

Лабораторна робота з навчальної дисципліни "Теорія ймовірностей та математична статистика" – це форма заняття, під час якого студент під керівництвом викладача здійснює виконання практичного завдання за допомогою комп'ютера відповідно до заданої теми. Проведенню лабораторної роботи з певної теми передують засвоєння основних теоретичних положень з цієї теми та формування практичних навичок.

Лабораторна робота виконується в комп'ютерному класі з використанням електронних таблиць *MS Excel 2010*.

Необхідною умовою успішного засвоєння навчального матеріалу з навчальної дисципліни є самостійна робота студентів із додатковою літературою та вміння застосовувати пакети прикладних програм у процесі

виконання лабораторних робіт та індивідуальних навчально-дослідних завдань за допомогою комп'ютера. Застосування програмного середовища *MS Excel* безпосередньо як потужного калькулятора, тобто для проведення розрахунків шляхом уведення потрібних формул до комірок електронних таблиць, сприяє більш глибокому засвоєнню теоретичного матеріалу. Крім того, використання вбудованих функцій і надбудови **Анализ данных**¹ дозволяє швидко й ефективно розв'язувати не тільки навчальні приклади, але й складні економічні задачі великої вимірності, що мають практичне значення.

Застосування пакета прикладних програм *MS Excel* у процесі виконання лабораторних робіт ґрунтується на тому, що на сьогодні програмне середовище *MS Excel* є однією з найбільш популярних і зручних програм, які призначені для роботи з електронними таблицями, отже, воно є звичним робочим середовищем для сучасного фахівця в галузі економіки та менеджменту. Додаток *MS Excel* входить до складу стандартного пакета офісних додатків *Microsoft Office*. Він надає користувачам широкі можливості щодо здійснення всіляких статистичних розрахунків та розв'язання оптимізаційних задач, а також щодо роботи з графічними інструментами. На сучасному етапі *MS Excel* є основним засобом, що допомагає офісним працівникам успішно створювати або формувати таблиці, здійснювати їх аналіз, управляти й обмінюватися даними, складати діаграми, виконувати обчислення будь-якої складності. Додаток *MS Excel* має власну мову програмування – об'єктно-орієнтовану мову *Visual Basic for Applications (VBA)*, що може застосовуватися користувачами в процесі розробки макросів для *MS Excel*.

У процесі підготовки лабораторних робіт використовувалась сучасна версія *MS Excel 2010* – динамічна програма, що дозволяє прораховувати різні варіанти розвитку економічних процесів, подаючи їх у зручному для читання вигляді за допомогою унікальних засобів візуалізації. Потужні засоби аналізу, що містить надбудова **Пакет анализа**, дозволяють виділяти й відстежувати важливі тенденції розвитку економічних процесів. Крім того, будь-які файли можна відправити в Інтернет, що дає можливість роботи над ними в режимі on-line разом з іншими користувачами.

¹ Тут і далі по роботі спеціальні терміни, що використовуються в програмному середовищі *MS Excel*, наведені мовою оригіналу і виділяються напівжирним шрифтом.

Для зарахування виконання лабораторної роботи студенту необхідно оформити індивідуальний звіт, у якому викладається економічний сенс задачі, що запропонована для самостійного розв'язання, наведено її математичну модель, знайдено розв'язок задачі за допомогою вбудованих функцій і надбудов *MS Excel*, а також здійснено аналіз отриманих результатів і наведено висновки. Крім того, студент повинен володіти теоретичними знаннями, що відповідають темі лабораторної роботи. Для їх перевірки та визначення ступеня самостійності виконання лабораторної роботи студент під час її захисту повинен відповісти на контрольні запитання з теми даної роботи. Приклад контрольних питань наведено наприкінці кожної лабораторної роботи. В електронному вигляді результати проведених обчислень необхідно подати викладачеві наприкінці заняття, а на початку наступного заняття – надати друкований варіант звіту про виконання лабораторної роботи.

На титульному аркуші звіту студент повинен указати номер лабораторної роботи та її назву, свої П. І. Б., групу та факультет, а також П. І. Б. викладача, під керівництвом якого була виконана лабораторна робота та прийнята до захисту.

Оцінювання знань студента за темою лабораторної роботи здійснюється за результатами її виконання та захисту, при цьому приділяється увага знанням теоретичного матеріалу, правильності висновків, повноті економічної інтерпретації набутих результатів та дотримання терміну виконання. Максимальний бал, який може отримати студент, визначається за технологічною картою дисципліни.

1. План лабораторних робіт

Перелік тем лабораторних робіт, їх зміст та література з кожної роботи наведені в таблиці.

Таблиця

План лабораторних робіт

Теми лабораторних робіт	Питання для опрацювання (за модулями та темами)	Кількість годин	Література
1	2	3	4
Змістовий модуль 1. Теорія ймовірностей			
Л. Р. 1. Класичне означення ймовірності та елементи комбінаторики. Основні теореми теорії ймовірностей	Знайомство з пакетом прикладних програм <i>MS Excel</i> . Застосування вбудованих функцій <i>MS Excel</i> до обчислення ймовірності випадкових подій за теоремами множення та додавання ймовірностей. Застосування формули повної ймовірності (апостеріорна ймовірність) та формули Байєса (апостеріорна ймовірність)	2	Основна: [1 – 3]. Додаткова: [4; 6; 8; 9]
Л. Р. 2. Дискретні випадкові величини: основні числові характеристики та їх властивості	Визначення основних та додаткових числових характеристик дискретної випадкової величини за означеннями і за допомогою вбудованих функцій <i>MS Excel</i> . Побудова закону розподілу суми, різниці і добутку випадкових величин, обчислення їх основних числових характеристик за означенням та з використанням властивостей математичного сподівання та дисперсії	2	Основна: [1 – 3]. Додаткова: [4; 6; 8; 9]
Л. Р. 3. Дискретна випадкова величина: біноміальний закон розподілу та його асимптотичні наближення	Побудова біноміального закону розподілу для певних значень p та n на базі моделі повторних випробувань за схемою Бернуллі, визначення основних числових характеристик розподілу. Асимптотичні теореми Муавра – Лапласа ($n \rightarrow \infty$) та Пуассона ($p \rightarrow 0$ при $n \rightarrow \infty$)	2	Основна: [1 – 3]. Додаткова: [4; 6; 8; 9]

Закінчення таблиці

1	2	3	4
Л. Р. 4. Багатовимірні випадкові величини	Побудова теоретичного закону розподілу багатовимірної випадкової величини (на прикладі двовимірної випадкової величини), визначення її основних числових характеристик за означенням та за допомогою вбудованих функцій <i>MS Excel</i>	2	Основна: [1 – 3]. Додаткова: [4; 6; 8; 9]. Інформаційні ресурси: [10; 11]
Змістовий модуль 2. Математична статистика			
Л. Р. 5. Основні поняття математичної статистики. Статистичне оцінювання параметрів розподілу	Побудова інтервального варіаційного ряду, надання його у вигляді гістограми та полігону. Визначення точкових та інтервальних оцінок основних числових характеристик розподілу за допомогою вбудованих функцій і надбудови Анализ данных <i>MS Excel</i>	2	Основна: [1 – 3]. Додаткова: [4; 5; 7 – 9]. Інформаційні ресурси: [10; 11]
Л. Р. 6. Перевірка статистичних гіпотез щодо закону розподілу	На прикладі нормального, рівномірного та експоненціального законів розподілу перевірка статистичної гіпотези щодо відповідності закону розподілу в генеральній сукупності певному виду	2	Основна: [1 – 3]. Додаткова: [4; 5; 7 – 9]. Інформаційні ресурси: [10; 11]
Л. Р. 7. Елементи теорії кореляції. Елементи регресійного аналізу	Визначення основних числових характеристик розподілу двовимірної випадкової величини за означенням і за допомогою вбудованих функцій. Побудова емпіричного рівняння регресії за допомогою вбудованих функцій і надбудов <i>MS Excel</i>	2	Основна: [1 – 3]. Додаткова: [4; 5; 7 – 9]. Інформаційні ресурси: [10; 11]
Л. Р. 8. Елементи дисперсійного аналізу	Побудова довірчого інтервалу лінії регресії за допомогою вбудованих функцій і надбудов <i>MS Excel</i> . Застосування дисперсійного аналізу до перевірки значущості кореляційного зв'язку між двома випадковими величинами	2	Основна: [1 – 3]. Додаткова: [4; 5; 7 – 9]. Інформаційні ресурси: [10; 11]

Вихідні дані завдання для розв'язання за кожною лабораторною роботою студент визначає за своїм номером у списку навчальної групи.

2. Зміст лабораторних робіт та інструкційна карта їх виконання

Змістовий модуль 1. Теорія ймовірностей

Лабораторна робота 1

Класичне означення ймовірності та елементи комбінаторики.

Основні теореми теорії ймовірностей.

Функції *MS Excel* для обробки даних

1.1. Мета роботи:

ознайомитися з основними принципами роботи з табличним процесором *MS Excel*, його вбудованими функціями та надбудовами;

вивчити формули комбінаторики та за їх допомогою обчислення ймовірностей випадкових подій за класичним означенням;

розв'язувати задачі, використовуючи теореми додавання та множення ймовірностей, обчислювати умовну ймовірність та застосовувати формули повної ймовірності й формули Байєса;

програмувати розв'язання на робочому аркуші *MS Excel*.

1.2. Базові принципи роботи з *MS Excel*

Найбільш поширеним способом реалізації обчислювальних операцій є автоматичне опрацювання даних з використанням електронних таблиць. Таким універсальним засобом, який зручно застосовувати для роботи з електронними таблицями, є табличний процесор *MS Excel*. Він має широкий спектр можливостей і може застосовуватися у ході дослідження економічних процесів великої вимірності.

Запуск *MS Excel* здійснюється або через головне меню за допомогою команд **Пуск** ⇒ **Все програми** ⇒ **Microsoft Office** ⇒ **Microsoft Office Excel 2010**, або з використанням попередньо створеного ярлика програми, який винесено на робочий стіл. Після запуску програми на моніторі з'являється чистий **Лист** робочої книги. **Робоча книга** – це файл, що має розширення *xls*. Кожний **Лист** робочої книги є окремою робочою таблицею, стовпці якої позначені латинськими літерами, а рядки – арабськими цифрами. Для посилання на комірку електронної таблиці, що

розташована, наприклад, на перетині стовпця **F** та **3**-го рядка, необхідно вказати її адресу, що визначається як **F3**.

Для виконання обчислень безпосередньо в комірку можна вводити **формулу**, за якою будуть виконуватись обрахування. Формула складається з однієї або кількох адрес комірок та чисел, а також знаків, що визначають математичні дії між ними. Кожна формула розпочинається зі знака **=**. Для автоматизації обрахувань існують вбудовані функції та надбудови, по суті, будь-яка **вбудована функція** – це формула, що побудована заздалегідь, або декілька таких формул, що пов'язані між собою. Кожна вбудована функція має назву. Для проведення обчислень із застосуванням вбудованої функції необхідно вказати назву цієї функції та її аргументи, які є вихідними даними для виконання обчислень. У такому разі для виконання обчислень у комірку, до якої буде виведено результат, записують: **=НАЗВА ФУНКЦІЇ(Аргументи)**. Аргументами функції можуть бути числа, текст, логічні значення, масиви, значення помилок чи посилання, а також формули, які, у свою чергу, можуть містити інші функції. Функції, що є аргументом іншої функції, називаються **вкладеними**. Так, *MS Excel* дозволяє використовувати до семи рівнів вкладеності функцій. У деяких випадках перелік аргументів функції може бути пустим. Наприклад, якщо під час обчислень необхідно ввести число π , то це можна здійснити, ввівши формулу: **=ПИ()**.

Якщо вбудована функція містить декілька аргументів, то їх необхідно відділяти крапкою з комою. Наприклад, функції **СУММ** або **ПРОИЗВЕД** можуть містити до 255 аргументів.

Деякі дії в *MS Excel* можна виконати значно швидше, якщо використовувати функціональні клавіші (табл. 1.1) та клавіатурні комбінації, так звані "гарячі" клавіші (табл. 1.2).

Таблиця 1.1

Функціональні клавіші

Клавіша	Функція	Клавіші			
		Shift	Ctrl	Ctrl + Shift	Alt + Shift
1	2	3	4	5	6
F1	Вивести довідку чи запустити майстер відповідей	Довідка типу "Що це?"	–	–	Вставити новий аркуш


1	2	3	4	5	6
F2	Перейти до виправлення вмісту комірки і рядків формул	Перейти до виправлення примітки комірки	Вивести вікно Сведения	–	–
F3	Вставити ім'я у формулу	Запустити майстер функцій	Присвоїти ім'я	Створити імена за текстом комірок	–
F4	Повторити останню дію; посилання на комірку перетворити на абсолютне	Повторити останній перехід чи пошук	Закрити вікно	–	–
F5	Виконати команду Перейти (меню Правка)	Виконати команду Найти (меню Правка)	Відновити вихідний розмір вікна	–	–
F6	Перейти в наступну ділянку вікна	Перейти в попередню ділянку вікна	Перейти в наступну книгу	Перейти в попередню книгу	–
F7	Виконати команду Орфографія (меню Сервіс)	–	Виконати команду Переместить (меню документа)	–	–
F8	Включити режим розширення виділеної ділянки	Включити режим переходу до наступної частини виділеної ділянки	Виконати команду Размер (віконне меню документа)	–	–
F9	Перерахувати всі аркуші в усіх відкритих книгах	Перерахувати точний аркуш	Згорнути вікно документа	–	–
F10	Перейти в рядок меню	Вивести контекстне меню	Розгорнути вікно документа	–	–
F11	Створити діаграму	–	–	–	–
F12	Виконати команду Сохранить как (меню Файл)	Виконати команду Сохранить (меню Файл)	Виконати команду Открыть (меню Файл)	Виконати команду Печать (меню Файл)	–

"Гарячі" клавіші

Операція	Комбінація клавіш
1	2
Клавіші для виправлення вмісту комірок чи рядка формул	
Увести набрані дані в комірку	Enter
Видалити набрані дані	Esc
Повторити останню дію	F4 або Ctrl + Y
Почати новий абзац у поточній комірці	Alt + Enter
Видалити символ або виділені символи ліворуч від курсору	Backspace
Вставити в комірку символ табуляції	Ctrl + Alt + Tab
Видалити символ або виділені символи праворуч від курсору	Delete
Перемістити курсор на один символ угору, вниз, ліворуч або праворуч	Клавіші зі стрілками
Перемістити курсор на початок рядка	Home
Перейти до виправлення примітки комірки	Shift + F2
Створити імена за текстом комірок	Ctrl + Shift + F3
Заповнити внизу	Ctrl + D
Заповнити справа	Ctrl + R
Заповнити виділені комірки набраним значенням	Ctrl + Enter
Увести дані в комірку та перейти до комірки, розташованої знизу	Enter
Увести дані в комірку та перейти до комірки, розташованої вгорі	Shift + Enter
Увести дані в комірку та перейти до комірки, розташованої справа	Tab
Увести дані в комірку та перейти до комірки, розташованої зліва	Shift + Tab
Почати формулу	=
Перейти в режим виправлення вмісту комірки	F2
Очистити рядок формул після зазначення комірки чи видалити в рядку формул символ зліва від курсору	Backspace
Вставити ім'я функції у формулу	Shift + F3
Присвоїти ім'я	Ctrl + F3
Перерахувати всі аркуші в усіх відкритих книгах	F9 чи Ctrl + =
Перерахувати поточний аркуш	Shift + F9
Виконати автопідсумовування	Alt + =
Увести поточну дату	Ctrl + ;

Закінчення табл. 1.2

1	2
Увести поточний час	Ctrl + Shift + :
Скасувати результати виправлення комірки чи рядки формул	Esc
Завершити виправлення комірки	Enter
Почати новий абзац	Alt + Enter
Вставити символ табуляції	Ctrl + Alt + Tab
Скопіювати вміст верхньої комірки в поточну комірку чи в рядок формул	Ctrl + Shift + "
Увести набрану формулу як формулу масиву	Ctrl + Shift + Enter
Скопіювати формулу верхньої комірки в поточну комірку чи в рядок формул	Ctrl + ' (апостроф)
Відобразити список автовведення	Alt + ↓ (Стрілка вниз)
Клавіші для форматування даних	
Виконати команду Стиль (меню Формат)	Alt + ' (апостроф)
Виконати команду Ячейка (меню Формат)	Ctrl + 1
Виконати форматування загальним числовим форматом	Ctrl + Shift + ~
Виконати форматування грошовим форматом із двома десятковими знаками після крапки (від'ємні числа відображаються в круглих дужках)	Ctrl + Shift + \$
Виконати форматування процентним форматом з відсутньою дробовою частиною	Ctrl + Shift + %
Виконати форматування науковим форматом із двома десятковими знаками після коми	Ctrl + Shift + ^
Виконати форматування форматом для дат з полями дня, місяця і року	Ctrl + Shift + #
Виконати форматування форматом для часу з полями годин і хвилин та індексами А. М. чи Р. М.	Ctrl + Shift + @
Виконати форматування форматом із двома десятковими знаками після коми	Ctrl + Shift + !
Вставити рамку структури	Ctrl + Shift + &
Видалити всі межі	Ctrl + Shift + _
Виконати чи видалити форматування жирним шрифтом	Ctrl + И
Виконати чи скасувати форматування курсивом	Ctrl + Ш
Підкреслити текст чи видалити лінію підкреслення	Ctrl + Г
Перекреслити текст чи видалити лінію перекреслення	Ctrl + 5
Сховати рядки	Ctrl + 9
Показати рядки	Ctrl + Shift + (
Сховати стовпці	Ctrl + 0 (нуль)
Показати стовпці	Ctrl + Shift +)

Для роботи з функціями в *MS Excel* є спеціальний засіб – **Мастер функций**. Звернутись до нього можна, натиснувши кнопку **Вставка функции** , що розташована на панелі інструментів **Главная**, або набравши комбінацію клавіш **Shift+F3**. Також можна набрати команду **Формулы** ⇒ **Вставить функцию**. Будь-яка з цих дій відкриває діалогове вікно **Мастер функций – шаг 1 из 2** (рис. 1.1).

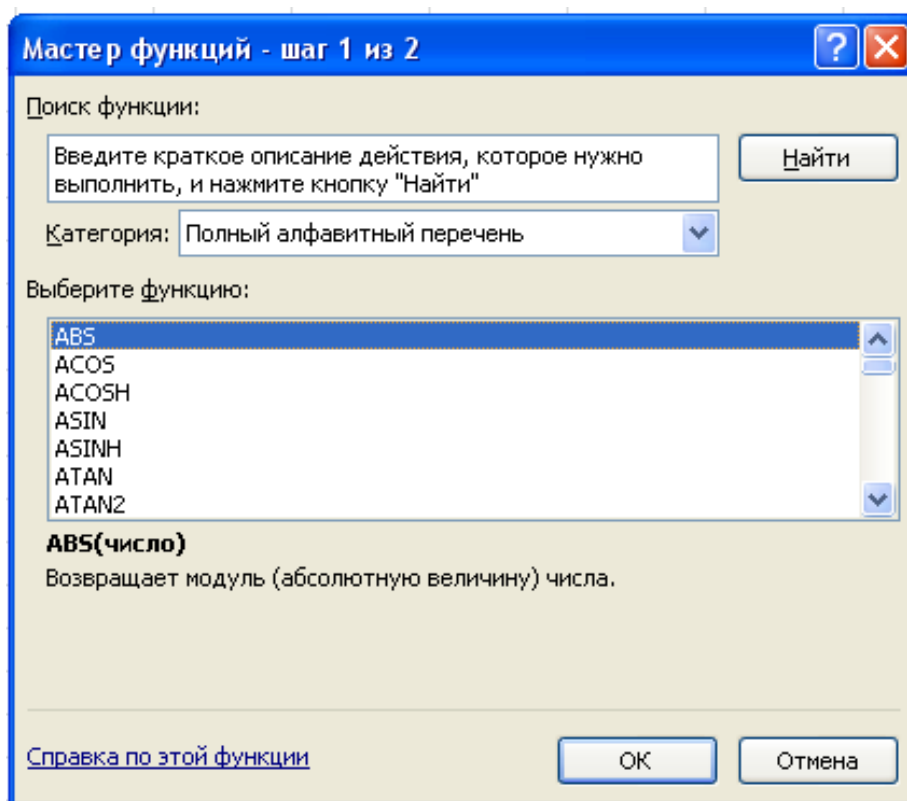


Рис. 1.1. Вікно панелі "Мастер функций"

Для зручності функції в *MS Excel* об'єднані за такими категоріями: логічні та текстові функції, функції перевірки властивостей та значень, функції посилання і масивів, функції дати і часу, функції керування базами даних і списками, функції надбудов та програмування, аналітичні функції оперативної обробки транзакцій *OLAP* (*Online Transaction Processing*), математичні, статистичні, інженерні, фінансові та інформаційні функції.

У діалоговому вікні **Мастер функций – шаг 1 из 2** міститься перелік категорій функцій (список **Категория**) та доступних функцій (список **Выберите функцию**). Список **Категория** містить також опцію **Полный алфавитный список функций**.

У процесі виконання лабораторних робіт найчастіше застосовуються вбудовані функції категорії **Математические** та **Статистические**.

Побудова формули починається зі знака **=**. Після введення імені функції в діалоговому вікні **Мастер функций – шаг 1 из 2** треба натиснути **ОК**, що приведе до появи діалогового вікна **Аргументы функции**. Після його заповнення в діалоговому вікні відображаються аргументи функції та результати обчислень. Щоб вивести результат розрахунків у обрану комірку листа *MS Excel*, натискаємо **ОК**.

Інформацію щодо функції можна отримати з її діалогового вікна **Аргументы функции**. Якщо цього виявиться недостатньо, то більш повну інформацію можна отримати, натиснувши напис **Справка по этой функции**, що міститься в лівому нижньому куті її діалогового вікна.

Ввести функцію у формулу можна безпосередньо з клавіатури. При цьому *MS Excel 2010* динамічно розкриває список функцій, що починаються так само, як і текст, який вводиться. Після введення функції цей список автоматично зникає. Як тільки ім'я функції набране і відкрита дужка, поруч із рядком формул з'являється підказка, де відображається повний синтаксис функції, тобто назва функції і список її аргументів. Причому той аргумент, який необхідно ввести на даний момент, визначено напівжирним шрифтом. Для набору імені функції можна застосовувати як малі, так і великі літери. Якщо ім'я набране правильно, то малі літери автоматично перетворюються на великі. Якщо ж цього не відбулось, то це означає, що в імені функції було зроблено помилку.

Деякі вбудовані функції повертають масиви значень. Щоб здійснити обчислення за допомогою формули масиву, потрібно для виведення результату обчислень виділити діапазон клітинок, що має таку саму кількість рядків і стовпців, що й очікуваний масив. Формули масиву створюються так само, як і інші формули, за винятком того, що в цьому випадку для її введення слід натискати комбінацію клавіш **Ctrl+Shift+Enter**. Електронна таблиця *MS Excel* має ряд вбудованих функцій для роботи з масивами, наприклад, функція **МУМНОЖ()**, що повертає добуток матриць.

1.3. Теоретичні положення

Вихідним поняттям теорії ймовірності є **випробування** (експеримент). Під випробуванням розуміють спосіб дослідження процесів та явищ за чітко визначених умов, які дозволяють відтворювати явище достатню кількість разів. **Випадкова подія** – це наслідок, який може відбутися або не відбутися як результат проведення випробування. Для позначення випадкових подій використовують великі літери латини: *A*, *B*, *C* та ін.,

або S . **Імовірність події** $P(S)$ є кількісною мірою об'єктивної можливості реалізації цієї події.

Сукупність усіх рівноможливих результатів випробування визначають як **простір елементарних подій** $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, де ω_i – елементарний наслідок випробування, тобто **елементарна подія**, що належить простору Ω . При цьому елементарні події попарно несумісні, тобто $\omega_i \cap \omega_j = \emptyset$, якщо $i \neq j$. Випадкову подію можна подати як підмножину простору елементарних подій, тобто $S \subseteq \Omega$. Елементарні події, що належать до цієї підмножини, називаються **сприятливими** щодо появи події S . Якщо підмножина елементарних подій, що утворює подію S , є пустою, то така подія S називається **неможливою** і позначається \emptyset . Якщо подія S містить усі елементарні події даного простору елементарних подій, тобто за заданих умов випробувань вона не може не відбутися, то така подія є **вірогідною** і позначається Ω .

Подією \bar{S} , що є **протилежною** до події S , є подія, для якої сприятливими є всі ті елементарні події, що доповнюють випадкову подію S до простору елементарних подій.

Події, що не мають спільних сприятливих елементарних подій, називаються **несумісними**. Випадкові попарно несумісні події $S_1; S_2; \dots; S_m$, що визначені на одному й тому ж просторі елементарних подій $\Omega = \{\omega_1; \omega_2; \dots; \omega_n\}$, утворюють у сукупності **повну групу несумісних випадкових подій**, якщо при реалізації будь-якої елементарної події $\omega_i \subset \Omega$ ($i = \overline{1, n}$) відбувається одна з випадкових подій S_j ($j = \overline{1, m}$).

Кожній події S із множини подій, що можуть бути побудовані на просторі елементарних подій $\Omega = \{\omega_1; \omega_2; \dots; \omega_n\}$, поставлено у відповідність невід'ємне дійсне число $P(S) \geq 0$, яке визначається як імовірність події S , при цьому для вірогідної події $P(\Omega) = 1$. За **класичним означенням** імовірність події $P(S)$ визначається співвідношенням:

$$P(S) = \frac{m}{n}, \quad (1.1)$$

де n – загальна кількість рівноможливих елементарних подій, які утворюють простір елементарних подій Ω ;

m – кількість сприятливих елементарних подій, тобто подій, що супроводжуються появою події S .

З аксіом теорії ймовірностей випливає, що ймовірність неможливої події дорівнює нулю: $P(\emptyset) = 0$.

Формула (1.1) застосовується в тому випадку, коли множина елементарних подій Ω є дискретною. Однак множина Ω може бути й неперервною. Якщо простір елементарних подій містить нескінченну множину елементів і йому можна поставити у відповідність певний геометричний простір, а ймовірність кожної події залежить тільки від міри цієї події, а не від її розташування, то говорять, що на цьому просторі визначена **геометрична ймовірність**. Тоді ймовірність кожної події визначається як відношення **міри** (*mes*) цієї події до міри простору елементарних подій:

$$P(S) = \frac{\text{mes } g}{\text{mes } G}, \quad (1.2)$$

де G – область, що відповідає множині елементарних подій Ω ;

g – область, що відповідає підмножині елементарних подій, які є сприятливими для появи події S .

Для ілюстрації основних понять теорії ймовірностей застосовують **діаграми Ейлера – Венна**, тобто геометричні схеми, що відображують дії з множинами (рис. 1.2).

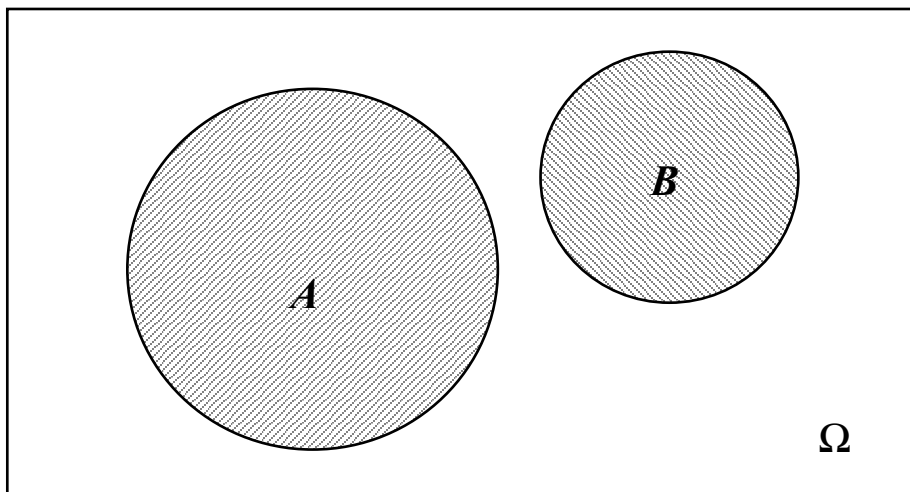


Рис. 1.2. Приклад діаграми Ейлера – Венна

Правило побудови діаграм Ейлера – Венна передбачає, що множині, яка відображає простір елементарних подій $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$,

ставлять у відповідність прямокутник, а підмножину елементарних подій, що є сприятливими для даної випадкової події, виділяють у вигляді круга (або іншої фігури) на цьому прямокутнику. Імовірність події визначається як відношення площі круга до площі всього прямокутника. Якщо випадкові події сумісні, то відповідні їм круги перетинаються. На рис. 1.1 наведено діаграму Ейлера – Венна, на якій відображені несумісні випадкові події A і B . Також можна стверджувати, що випадковій події A відповідає більша ймовірність, ніж випадковій події B .

У тому випадку, коли простір елементарних подій є дискретним, для обчислення ймовірностей зручно застосовувати формули комбінаторики. **Комбінаторика** – це розділ математики, який вивчає питання щодо кількості комбінацій, які можна побудувати за певним принципом, використовуючи задану кількість об'єктів.

Розглянемо множину, що містить n різних об'єктів. **Перестановкою** називається впорядкована множина (**кортеж**), що містить усі n об'єктів. Будь-яка перестановка відрізняється від іншої лише порядком розташування об'єктів. **Кількість перестановок** (із n), що отримують на множині обсягом n , визначається за формулою:

$$P_n = n!, \quad (1.3)$$

де $n!$ – факторіал числа n , він визначається як добуток усіх натуральних чисел, починаючи з 1 і закінчуючи n , тобто $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$.

Якщо з множини різних об'єктів обсягом n витягати по k об'єктів і будувати з них кортежі, для яких має значення не тільки склад об'єктів, але й порядок їх розміщення у кортежі, то такий кортеж називається **розміщенням**. **Кількість розміщень** (із n по k) визначається за формулою:

$$A_n^k = \frac{n!}{(n-k)!}. \quad (1.4)$$

Якщо під час побудови кортежу з k об'єктів, що належать множині обсягом n , порядок розташування об'єктів не має значення, то такий кортеж називається **сполученням**.

Кількість сполучень (із n по k) визначається за формулою:

$$C_n^k = \frac{n!}{k! \cdot (n-k)!}, \quad (1.5)$$

тобто кількість сполучень менша за кількість розміщень у $k!$ разів, що визначає кількість перестановок у множині обсягом k .

Імовірність події, яка визначається лише умовою випробування і не залежить від того, чи відбулась до цього будь-яка інша подія, називається **безумовною**.

Умовна ймовірність $P(B | A)$ – це ймовірність появи події B за умови, що подія A вже відбулася. Якщо ймовірність події не залежить від того, відбулась чи ні інша подія, то події називаються **незалежними**.

Теорема множення ймовірностей 1.1 (про ймовірність добутку подій). Імовірність одночасної появи двох подій дорівнює добутку ймовірності появи однієї з подій з умовною ймовірністю іншої події у припущенні, що перша подія вже відбулася:

$$P(AB) = P(A) \cdot P(B | A), \quad (1.6)$$

де $P(A)$ – ймовірність події A (безумовна ймовірність події A);

$P(B | A)$ – ймовірність появи події B за умови, що подія A вже відбулася (умовна ймовірність події B).

Теорема додавання ймовірностей 1.2 (про ймовірність суми подій). Імовірність появи хоча б однієї з двох сумісних подій дорівнює сумі їх ймовірностей без імовірності їх добутку:

$$P(A + B) = P(A) + P(B) - P(AB). \quad (1.7)$$

У процесі аналізу можливого розвитку подій та визначення ризиків, що пов'язані із впливом факторів, появу яких можна передбачити, а відносно ймовірностей появи цих факторів можна висловити певні припущення, застосовується **формула повної ймовірності**.

Нехай подія A настає лише за умови появи однієї з несумісних випадкових подій H_1, H_2, \dots, H_n , які утворюють повну групу подій. Ці події називаються **гіпотезами**.

Тоді ймовірність події A визначається за формулою:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A | H_i), \quad (1.8)$$

де $P(H_i)$ – ймовірність реалізації гіпотези H_i ;

$P(A | H_i)$ – умовна ймовірність події A , тобто ймовірність події A за умови, що реалізується саме гіпотеза H_i .

Оскільки гіпотези H_1, H_2, \dots, H_n утворюють повну групу несумісних випадкових подій, то відносно них виконуються такі умови:

$$H_i \cap H_j = \emptyset, \text{ якщо } i \neq j, \text{ та } \sum_{i=1}^n P(H_i) = 1. \quad (1.9)$$

Слід зауважити, що $P(H_i)$ є ймовірністю реалізації гіпотези H_i , яка визначається до випробування, тобто це **апріорна ймовірність**.

Нехай випадкова подія A , яка може мати місце як наслідок реалізації однієї з гіпотез H_1, H_2, \dots, H_n , вже відбулася, і необхідно визначити ймовірність того, що це мало місце саме завдяки реалізації гіпотези H_i . Отже, необхідно знайти **апостеріорну** ймовірність гіпотези H_i , тобто умовну ймовірність гіпотези, якщо відомо, що подія A вже відбулася.

Ймовірність гіпотези H_i за умови, що подія A вже відбулася, обчислюється за **формулою Байєса**:

$$P(H_i | A) = \frac{P(H_i) \cdot P(A | H_i)}{\sum_{i=1}^n P(H_i) \cdot P(A | H_i)}, \quad (1.10)$$

де $P(H_i | A)$ – апостеріорна ймовірність реалізації гіпотези H_i за умов, що подія A відбулася.

Відносно апостеріорних імовірностей виконується співвідношення:

$$\sum_{i=1}^n P(H_i | A) = 1. \quad (1.11)$$

1.4. Змістовна постановка задачі

За допомогою вбудованих функцій, що містить *MS Excel 2010*, необхідно виконати такі завдання:

1.1. Застосовуючи функції **ФАКТР**, **ПЕРЕСТ** та **ЧИСЛКОМБ**, обчислити значення виразу: $\frac{A_{15}^5 + C_{21}^4}{P_6}$.

1.2. За класичним означенням імовірності обчислити ймовірність випадкової події, яка полягає в тому, що при довільному наборі коду доступу, про який відомо, що він складається з чотирьох різних цифр, буде набрано правильний код.

1.3. Із застосуванням теорем додавання та множення ймовірностей обчислити ймовірність випадкової події, яка полягає в тому, що серед трьох студентів здадуть іспит саме два з них, якщо перший студент вивчив 80 % питань програми, другий – 90 %, а третій – 70 %. Навести діаграму Ейлера – Венна, що відповідає випадковій події.


1.4. Із застосуванням формули повної ймовірності визначити ймовірність отримання певного рівня прибутку за таких умов. Підприємець має можливість орендувати місце на одному з двох торговельних майданчиків: закритому чи відкритому. Якщо погода сонячна, то ймовірність досягти певного рівня прибутковості (успіх) складає 0,7 для відкритого майданчика та 0,6 – для закритого. У дощову погоду на відкритому майданчику покупців нема, а для закритого майданчика ймовірність отримати певний рівень прибутку зберігає те ж саме значення 0,6. Протягом року дощова погода спостерігається в середньому для половини днів. У сонячну погоду підприємець з однаковою ймовірністю обирає торгівлю на відкритому чи закритому майданчиках, а у дощову – тільки на закритому.

1.5. Із застосуванням формули Байєса за умовою завдання 1.4 визначити ймовірність того, що підприємець отримав прибуток саме у день, коли йшов дощ, якщо відомо, що в цей день був досягнутий певний рівень прибутку.

1.5. Приклад виконання лабораторної роботи 1

Завдання 1.1. Для обчислення значення виразу $\frac{A_{16}^5 + C_{15}^6}{P_{10}}$, що передбачає використання формул комбінаторики, застосуємо вбудовані

функції *MS Excel* **ФАКТР**(n), **ПЕРЕСТ**($n; m$) та **ЧИСЛКОМБ**($n; m$). Для обчислення значення заданого виразу визначимо кожний із його елементів окремо, тобто введемо їх значення в окремі комірки. Так, для виведення результату обчислення значення кількості розміщень вибираємо комірку **C3** і ставимо курсор до цієї комірки. Кількість розміщень A_{16}^5 визначається за допомогою функції **ПЕРЕСТ**, що належать до категорії **Статистические**.

Викликаємо діалогове вікно надбудови **Мастер функций – шаг 1 из 2**, натиснувши на кнопку , вибираємо необхідну функцію, переходимо до її діалогового вікна і вводимо її аргументи (рис. 1.3).

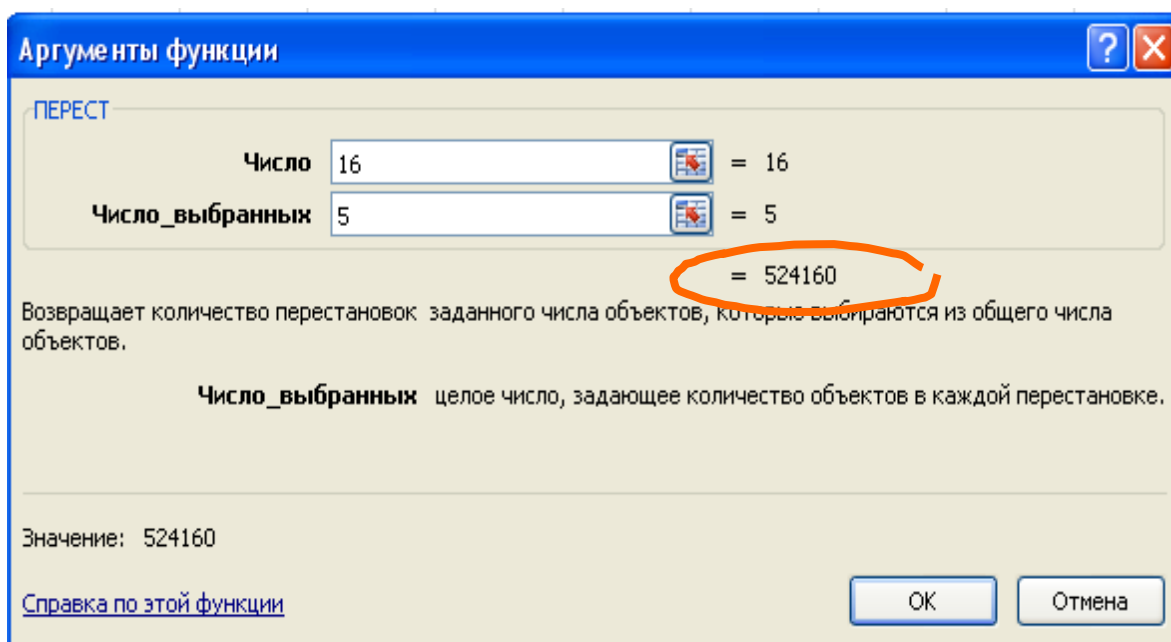



Рис. 1.3. Діалогове вікно функції ПЕРЕСТ

Після натискання **Ок** результат обчислень виводиться на екран. Для обчислення кількості сполучень C_{15}^6 поставимо курсор до комірки **C4**, натисканням кнопки  викликаємо діалогове вікно надбудови **Мастер функций – шаг 1 из 2** і вибираємо функцію **ЧИСЛКОМБ**, яка належить до категорії **Математические**. Після заповнення діалогового вікна **Аргументы функции** (рис. 1.3) натискаємо **Ок** і виводимо результат.

Результат обчислення кількості переставлень P_{10} введемо в комірку **C5**, застосувавши для здійснення розрахунків функцію **ФАКТР**, що належить до категорії **Математические**.

Тепер обчислюємо значення всього виразу, для чого в комірці **C7** набираємо формулу: **=(C3+C4)/C5**. Оскільки результат обчислень виявився дробовим, то для перетворення нескінченного періодичного дроби у простий дріб поставимо курсор у комірку **C7**, натискаємо праву клавішу маніпулятора "миша", обираємо режим **Формат ячеек** і в діалоговому вікні вибираємо **Число**. Потім у полі **Числовые форматы** вказуємо **Дробный**, а у полі **Тип** вказуємо (про всяк випадок) **Дробями до трех цифр** (рис. 1.4).

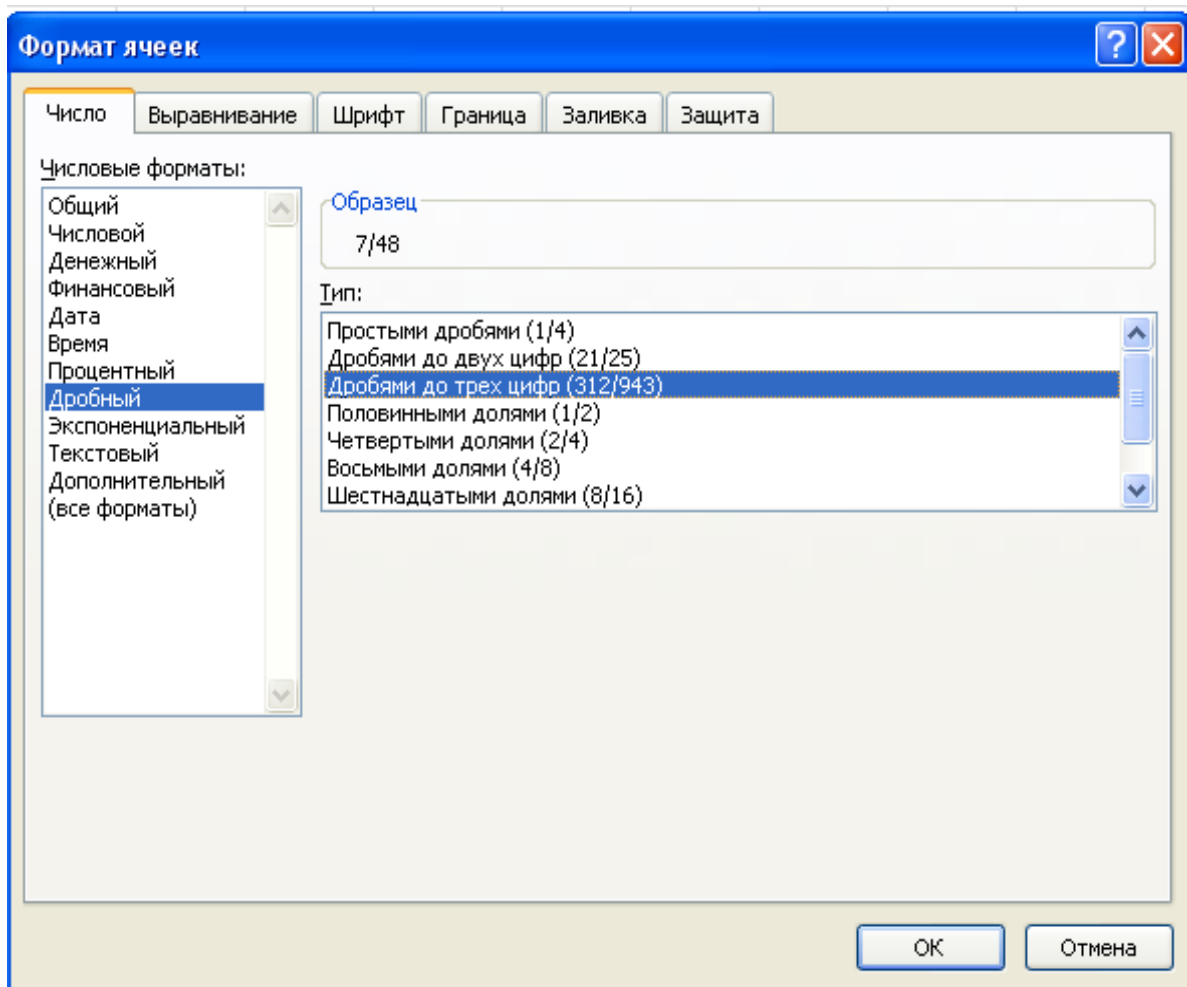


Рис. 1.4. Визначення формату комірки

За результатами обчислення отримуємо шукане значення:

$$\frac{A_{16}^5 + C_{15}^6}{P_{10}} = \frac{7}{48}.$$

Приклад оформлення результатів обчислень наведено на рис. 1.5.

	A	B	C	D	E
1	Завдання 1.1				
2					
3	Кількість розміщень із 16 по 5		524160		
4	Кількість сполучень із 15 по 6		5005		
5	Кількість переставлень із 10		3628800		
6					
7	Значення виразу		7/48		
8					
9					

Рис. 1.5. Приклад оформлення завдання 1.1

Завдання 1.2. Відомо, що код доступу складається з чотирьох різних цифр. Для визначення ймовірності випадкової події, яка полягає в тому, що за довільного набору цифр буде набрано правильний код, застосуємо класичне означення ймовірності. Оскільки код доступу тільки один, тобто лише одна елементарна подія є сприятливою, то $m=1$. Загальну кількість елементарних подій, що утворює простір елементарних подій, визначаємо як кількість розміщень із 10 (загальна кількість цифр) по 4 (кількість цифр у коді доступу). Набираємо **=ПЕРЕСТ(10; 7)** і отримуємо, що $n = 5040$. За формулою (1.1) маємо: $P(A) = \frac{1}{5040}$.

Завдання 1.3. Позначимо через A випадкову подію, яка полягає в тому, що успішно складуть іспит саме два будь-які студенти. Нехай A_1 , A_2 та A_3 – це успішне складання іспиту першим, другим та третім студентами окремо. Події A_1 , A_2 та A_3 є сумісними. Оскільки перший студент вивчив 80 % питань програми, другий – 90 %, а третій – 70 %, то при складанні іспиту ймовірності успіху для кожного з них окремо становлять, відповідно, $P(A_1) = 0,8$, $P(A_2) = 0,9$ та $P(A_3) = 0,7$.

Ймовірність події A , яка полягає в тому, що іспит складуть тільки два будь-які студенти, визначається співвідношенням:

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

де $P(\bar{A}_i)$ – ймовірність того, що i -й студент не складе іспит, тобто події, протилежної до A_i . Вона становить $P(\bar{A}_i) = 1 - P(A_i)$.

Оскільки всі три події, що утворюють подію A , є попарно несумісними, то ймовірність їх суми за аксіомою IV аксіоматики Колмогорова дорівнює сумі їхніх ймовірностей, тобто:

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

У свою чергу, події, що утворюють кожний з добутоків, є незалежними, отже, за теоремою множення ймовірностей маємо, що ймовірність добутку дорівнює добутку безумовних ймовірностей. Звідси остаточно отримуємо:

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

На діаграмі Ейлера – Венна (рис. 1.6) область, що відповідає випадковій події A , виділена штриховкою.

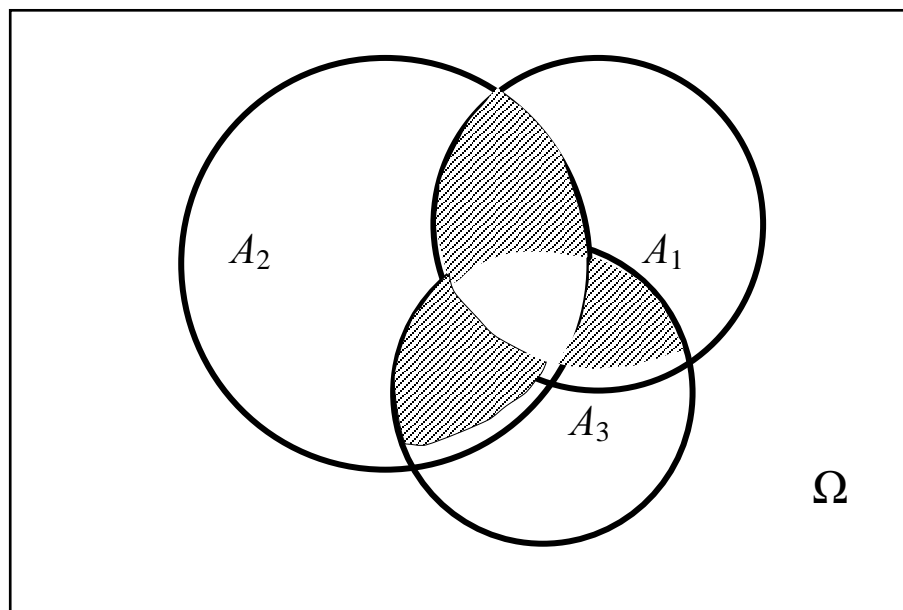


Рис. 1.6. Діаграма Венна – Ейлера, що визначає ймовірність випадкової події "іспит складуть саме два студента"

Для обчислення ймовірності події A скористаємося можливостями MS Excel. Для цього відкриваємо новий Лист і записуємо вихідні дані завдання в діапазоні комірок **D3:D5**, тобто ймовірності скласти іспит для кожного студента, у діапазоні комірок **D7:D9** виводимо результати обчислення ймовірностей протилежних подій, а до комірки **D11** вводимо формулу для обчислення ймовірності випадкової події A (рис. 1.7).

D11		fx = D3*D4*D9+D7*D4*D5+D3*D8*D5				
	A	B	C	D	E	F
1	Завдання 1.3					
2						
3	Імовірність події A1			0,8		
4	Імовірність події A2			0,9		
5	Імовірність події A3			0,7		
6						
7	Імовірність події, протилежної A1			0,2		
8	Імовірність події, протилежної A2			0,1		
9	Імовірність події, протилежної A3			0,3		
10						
11	Імовірність події A			0,398		
12						

Рис. 1.7. Обчислення за теоремами додавання та множення ймовірностей

Отже, ймовірність того, що саме два будь-яких студенти складуть іспит, дорівнює:

$$P(A) = 0,8 \cdot 0,9 \cdot (1 - 0,7) + (1 - 0,8) \cdot 0,9 \cdot 0,7 + 0,8 \cdot (1 - 0,9) \cdot 0,7 = 0,398.$$

Завдання 1.4. Розглянемо стани, в яких може перебувати система "Підприємець". На рис. 1.8 визначена множина станів, у яких може перебувати система, показані шляхи переходів та ймовірності, з якими ці переходи здійснюються. Маємо два послідовних рівні гіпотез, на кожному з яких виконуються умови формули (1.9). Першому рівню відповідають гіпотези відносно стану погоди – гарна (+) чи погана (-). Оскільки за умовою задачі половина днів протягом року є сонячними, то стани погоди є рівноймовірними, отже, ймовірність їх реалізації становить 0,5. Гіпотези другого рівня визначають тип майданчика – закритий чи відкритий. Ймовірності їх реалізації є умовними, оскільки залежать від стану погоди. Третій рівень гіпотез – отримання певного рівня прибутку (успіх) або його відсутність.

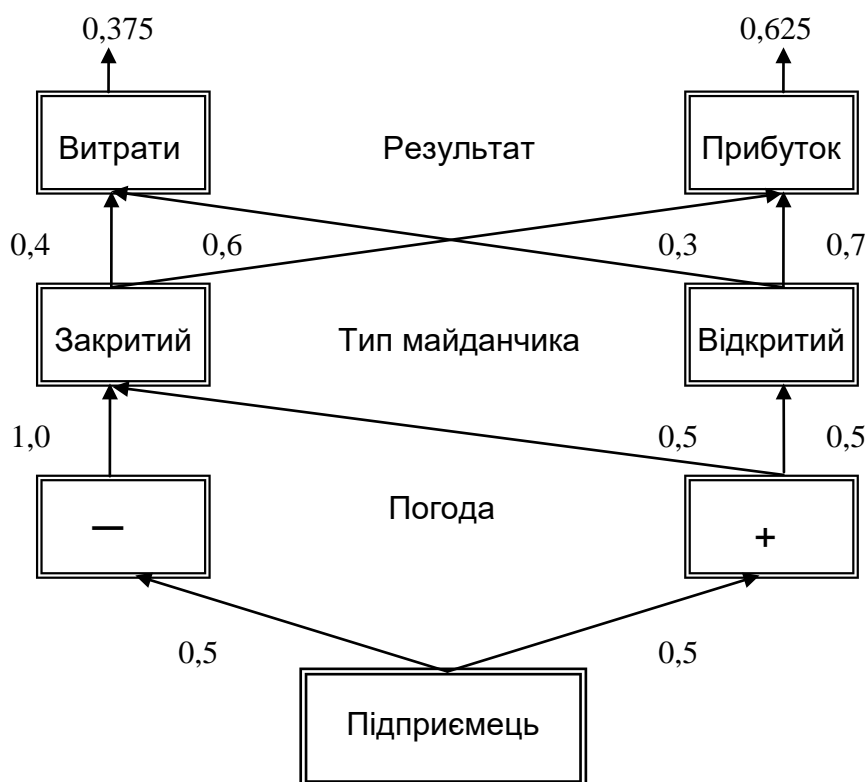


Рис. 1.8. Множина станів системи "Підприємець"

На робочому аркуші *MS Excel* у рядках запишемо всі шляхи, якими можна досягти стану "Успіх", та відповідні їм імовірності (рис. 1.9).

	A	B	C	D	E	F	G	H	I	J	K
1											
2		Вихідний стан		Погода		Тип майданчика		Кінцевий результат		Ймовірність	
3			0,5		0,5		0,6				
4		Підприємець	→	+	→	Закритий	→	Успіх	→	0,15	
5			0,5		0,5		0,7				
6		Підприємець	→	+	→	Відкритий	→	Успіх	→	0,175	
7			0,5		1		0,6				
8		Підприємець	→	-	→	Закритий	→	Успіх	→	=C7*E7*G7	
9											
10											

Рис. 1.9. Скриншот, який містить схему обчислення повної ймовірності

Стани системи позначаються прямокутниками з відповідною назвою, а перехід із одного стану в інший – стрілкою, над якою написана ймовірність даного переходу.

Кожний шлях досягнення успіху має вихідний стан "Підприємець" і кінцевий стан – "Успіх". Ці шляхи визначають за множиною станів системи, якщо рухатися вздовж стрілок. На рис. 1.9 у рядку формул наведена формула обчислення ймовірності успіху, якщо торгівля відбувається на закритому майданчику за умови поганої погоди.

Отже, відповідно до формули повної ймовірності (1.8) для підприємця ймовірність досягти успіху, тобто отримати певний рівень прибутку, становить:

$$P(A) = 0,5 \cdot 0,5 \cdot 0,6 + 0,5 \cdot 0,5 \cdot 0,7 + 0,5 \cdot 1 \cdot 0,6 = 0,625.$$

Завдання 1.5. Повернемося до умови завдання 1.4. Оскільки відомо, що подія A відбулася (успіх досягнуто) і необхідно визначити, чи обумовлена вона реалізацією гіпотези H_3 , то обчислення апостеріорної ймовірності здійснюється за формулою (1.10), де в знаменнику міститься повна ймовірність події A (це відповідь завдання 1.4), а в чисельнику – ймовірність успіху за поганої погоди, тобто третій доданок у формулі повної ймовірності (йому відповідає останній рядок на схемі рис. 1.9). Вводимо цю формулу до будь-якої комірки й отримуємо:

$$P(H_3 | A) = \frac{0,5 \cdot 1 \cdot 0,6}{0,625} = 0,48.$$

Отже, якщо успіх досягнуто, то він майже рівною мірою забезпечується торгівлею на відкритому чи закритому майданчиках.

1.6. Завдання для самостійної роботи

Варіант 1.1

1. Обчислити значення виразу: $\frac{A_{13}^2 + C_{21}^4}{P_{11}} - 4 \cdot C_{13}^4$.

2. У конверті серед 100 фотокарток знаходиться та, яку розшукують. З конверта навмання витягли 10 фотокарток. Знайти ймовірність того, що серед них виявиться потрібна картка.

3. Проводиться хімічний аналіз проб води із трьох джерел. Ймовірність того, що вода з першого джерела відповідає нормі, дорівнює 0,6; із другого – 0,8; із третього – 0,5.

Знайти ймовірність того, що в межах норми буде склад води саме з другого джерела, якщо відносно інших можна зробити будь-яке припущення.

4. Фірма має трьох постачальників, обсяг постачання для кожного з яких становить відповідно 30 %, 20 % та 50 %. Кожний із постачальників дотримується стандарту під час виготовлення продукції з імовірністю 0,8, 0,7 та 0,9, відповідно. Визначити ймовірність того, що одиниця продукції, яка була вибрана для перевірки (цей вибір здійснювався довільно), виявилася якісною.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо постачальника, якщо продукція виявилася неякісною.

Варіант 1.2

1. Обчислити значення виразу: $\frac{A_{20}^6 + C_{13}^{10}}{P_{15}} - 6,3 \cdot C_{23}^{15}$.

2. Група складається з 15 юнаків і 10 дівчат. Для чергування призначають 5 студентів. Визначити ймовірність того, що серед них буде саме 2 дівчини.

3. На зупинці можуть зупинятися автобуси № 17, 26 та 443. Імовірність того, що автобус № 17 приїде на зупинку за розкладом, дорівнює 0,8, автобус № 443 – 0,6, і автобус № 26 – 0,2. Знайти ймовірність того, що:
а) за розкладом приїде хоча б один автобус; б) за розкладом приїдуть автобуси як № 26, так і 17; в) усі три автобуси спізняться; г) спізниться один автобус.

4. Біля верстату знаходиться три деталі, до яких додали одну стандартну. Потім навмання взяли одну деталь. Визначити ймовірність того, що ця деталь буде стандартною, якщо відносно якості трьох вихідних деталей можна зробити будь-яке припущення.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо постачальника, якщо продукція виявилася неякісною.

Варіант № 1.3

1. Обчислити значення виразу: $\frac{A_{28}^{10} + C_{19}^6}{P_{23}} - 4 \cdot A_{10}^3$.

2. У партії 40 стандартних і 4 нестандартні деталі. Для контролю взяли навмання 8 деталей, які виявилися стандартними.

Знайти ймовірність того, що наступна взята навмання деталь буде також стандартною.

3. Три менеджери влаштовуються на роботу. Перший із них на 90 % відповідає вимогам роботодавця, другий – на 70 %, третій – на 80 %. За умов, що є три вільних місця і робітника вибирають випадково, знайти ймовірність того, що на роботу візьмуть хоча б одного менеджера.

4. Три експерти роблять висновок щодо можливостей здійснення певного проекту, і відповідь надається у формі "так" або "ні". Перший експерт володіє інформацією на 80 %, другий – на 90 %, третій – на 95 %. Визначити ймовірність того, що відповіді всіх трьох експертів співпадуть.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо правильності їх висновків за умов, що їх відповіді співпали.

Варіант 1.4

1. Обчислити значення виразу: $\frac{A_{23}^6 - C_{18}^3}{P_{10}} + \frac{1}{2} \cdot C_{20}^5$.

2. У партії із 15 деталей 7 стандартних. Знайти ймовірність того, що серед 5 узятих навмання деталей 3 стандартні.

3. На склад доставили 2 партії товару. У першій партії 76 % якісних виробів, у другій – 83 %. Навмання вибирають по одному виробу з кожної партії. Визначити ймовірність того, що серед них виявиться один якісний і один бракований.

4. У першій із шухляд знаходяться п'ять білих та сім чорних кульок, у другій – три білих та три чорних кульки, у третій – шість білих та шість чорних. З першої шухляди до другої переклали якусь кульку, після чого з другої до третьої теж переклали одну кульку, колір якої невідомий. Потім із третьої шухляди вийняли одну кульку. Визначити ймовірність того, що остання кулька є білою.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо кольору кульок, які переклали з першої шухляди до другої та з другої до третьої.

Варіант 1.5

1. Обчислити значення виразу: $\frac{A_{17}^9 + C_{20}^3}{P_{13}} - \frac{1}{5} \cdot A_7^2$.

2. Підручник має 208 сторінок. Визначити ймовірність того, що порядковий номер сторінки, яка відкрита навмання, буде закінчуватися цифрою 3.

3. Три стрільці роблять по одному пострілу в мішень. Імовірність влучення для першого стрільця дорівнює 0,5, для другого – 0,8 і для третього – 0,75. Визначте ймовірність того, що буде лише одне влучення.

4. Три консультанти роблять висновок щодо можливостей застосування певного проекту, і їх відповідь надається у формі "так" або "ні". Перший консультант володіє інформацією на 70 %, другий – на 80 %, третій – на 90 %. Визначити ймовірність того, що відповіді всіх трьох консультантів не співпадуть.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо правильності їх висновків за умов, що відповіді співпали.

Варіант 1.6

1. Обчислити значення виразу: $\frac{A_{19}^5 + C_{12}^3}{P_{20}} - \frac{1}{4} \cdot C_{25}^{10}$.

2. На 20 однакових жетонах написано 20 двозначних чисел від 11 до 30. Яка ймовірність того, що номер навмання взятого жетона буде кратним 4 або 7?

3. Із трьох гармат зробили залп по цілі. Імовірність влучення в ціль у разі одного пострілу з першого знаряддя дорівнює 0,9; з другого – 0,8, а з третього – 0,9. Знайти ймовірність того, що постріл хоча б із однієї гармати буде влучним.

4. Магазин має трьох постачальників, обсяг постачання для кожного з яких становить відповідно 10 %, 20 % та 70 %. Кожний із постачальників у процесі виготовлення продукції дотримується стандарту з імовірністю 0,8, 0,7 та 1,0, відповідно. Визначити ймовірність того, що продукція, яка була навмання вибрана для перевірки, виявилася неякісною.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо постачальників, від яких було отримано продукцію, якщо вона виявилася якісною.

Варіант 1.7

1. Обчислити значення виразу: $\frac{C_{13}^5 + A_{19}^{10}}{P_{14}} - 3 \cdot A_{12}^7$.

2. У пакунку 6 білих і 4 чорних мотків пряжі. Із пакунка навмання беруть 2 мотки. Яка ймовірність того, що вони одного кольору?

3. Робітник обслуговує три верстати. Імовірність того, що протягом години перший верстат не зупиниться, дорівнює 0,81; другий – 0,65; третій – 0,43. Визначити ймовірність того, що протягом години зупиняться будь-які два верстати.

4. Під час експлуатації протягом певного терміну система може зберігати робочий стан з імовірністю 0,7, досягти критичного стану з імовірністю 0,2 або перейти в аварійний стан з імовірністю 0,1. У разі досягнення критичного стану система самостійно або повертається до робочого стану (з імовірністю 0,6), або переходить у катастрофічний стан. У разі досягнення катастрофічного стану зовнішнє джерело впливу повертає систему до робочого стану (з імовірністю 1,0). Визначити ймовірність того, що після двох термінів експлуатації система перебуватиме в катастрофічному стані.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо шляхів перетворення системи, якщо виявилось, що після двох термінів експлуатації система перебуває в робочому стані.

Варіант 1.8

1. Обчислити значення виразу: $\frac{2A_{18}^9 - C_{24}^4}{P_{21}} + C_{12}^5$.

2. Механізм містить дві однакові деталі. Він не працюватиме, якщо обидві деталі будуть меншими за розміром, ніж стандартні. Складальник має 10 деталей, з яких 3 – менші за розміром. Знайти ймовірність того, що механізм працюватиме, якщо складальник бере потрібні для механізму дві деталі навмання.

3. Пристрій складається із трьох елементів, що працюють незалежно. Імовірність безвідмовної роботи першого елемента дорівнює 0,5, другого – 0,4, третього – 0,7. Знайти ймовірність того, що працюватиме безвідмовно другий елемент, а стан інших елементів є довільним.

4. У першій з урн знаходяться сім білих та сім чорних кульок, у другій – три білих та вісім чорних, у третій – чотири білих та шість чорних. З першої урни до другої переклали одну кульку, колір якої невідомий, після чого з другої урни до третьої теж переклали одну кульку, колір якої також є невідомим. Після цього з третьої урни вийняли одну кульку.

Визначити ймовірність того, що ця кулька буде білою.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо кольору кульок, які переклали з першої урни до другої.

Варіант 1.9

1. Обчислити значення виразу: $\frac{A_{16}^7 + 4C_{20}^{18}}{P_{12}} - A_{24}^6$.

2. На столі лежать 30 екзаменаційних білетів. Яка ймовірність того, що номер білета, який взято навмання, буде кратним 3 або 7?

3. Для повідомлення про пожежу встановили три автомати, що працюють незалежно. Ймовірність того, що в разі пожежі спрацює перший автомат, дорівнює 0,81, другий – 0,83, третій – 0,85. Знайти ймовірність того, що в разі пожежі надійде сигнал лише від одного автомата.

4. У шухляді зберігаються три лампи, до яких додали одну стандартну. Потім навмання взяли одну з ламп. Визначити ймовірність того, що ця лампа є нестандартною, якщо відносно якості трьох вихідних ламп можна зробити будь-яке припущення.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо кожного з припущень за умов, що випадково взята лампа виявилася стандартною.

Варіант 1.10

1. Обчислити значення виразу: $\frac{A_{12}^7 - C_{14}^5}{2P_7} - C_{30}^{27}$.

2. Студент знає 50 із 60 питань. Знайти ймовірність того, що студент відповість на два питання.

3. Від аеровокзалу вирушають три автобуси до трапа літака. Ймовірність своєчасного прибуття кожного автобуса дорівнює 0,8. Знайти ймовірність того, що два автобуси спізняться.

4. Завод має три цехи, обсяг робіт для кожного з яких становить відповідно 70 %, 20 % та 10 %. Кожний цех у процесі виготовлення продукції дотримується стандарту з ймовірністю 0,8, 0,7 та 0,9 відповідно. Визначити ймовірність того, що одиниця продукції, яка була вибрана для перевірки навмання, виявилася неякісною.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез постачальника, якщо продукція виявилася якісною.

Варіант 1.11

1. Обчислити значення виразу: $\frac{A_{25}^8 + 5C_{12}^7}{P_{14}} + A_{18}^8$.

2. У групі 12 студентів, серед яких 8 мають з економіки оцінки від 7 до 9 балів. За списком відібрано 9 студентів. Знайти ймовірність того, що серед них п'ятеро знають економіку на 7 – 9.

3. Контролер, перевіряючи якість 20 пал'ят, встановив, що серед них 16 відповідають першому ґатунку, а інші – другому. З цієї партії навмання взяли 3 пал'ята. Знайти ймовірність того, що серед них хоча б одне пал'ято буде першого ґатунку.

4. Фірма має трьох постачальників, обсяг постачання для кожного з яких становить відповідно 10, 40 та 50 %. Кожний із постачальників дотримується стандарту з імовірністю 0,6, 0,7 та 0,8, відповідно. Визначити ймовірність того, що одиниця продукції, яка була вибрана для перевірки, причому цей вибір здійснювався довільно, виявилася неякісною.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо постачальника, якщо продукція виявилася якісною.

Варіант 1.12

1. Обчислити значення виразу: $\frac{3A_{14}^5 - C_{11}^9}{P_{18}} + C_{17}^8$.

2. Пристрій складається з п'яти однотипних елементів, з яких два є зношеними. Для перевірки вибрали два з них. Знайти ймовірність того, що вибрали саме зношені елементи, якщо вибір здійснювався навмання.

3. Проводиться аналіз води із трьох джерел. Імовірність того, що хімічний склад води з першого джерела відповідає нормі, дорівнює 0,4, із другого – 0,9, із третього – 0,5. Знайти ймовірність того, що в межах норми є склад води тільки з третього джерела.

4. Є три урни. У першій урні знаходиться три білих та п'ять чорних кульок, у другій – дві білих та вісім чорних. З перших двох урн вийняли навмання по дві кульки і поклали до третьої, в якій до цього знаходилися три білих кульки. Потім з третьої урни навмання вийняли дві кульки. Визначити ймовірність того, що ці обидві кульки виявляться чорними.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез щодо кольору кульок, які перебрали з першої урни до третьої.

Варіант 1.13

1. Обчислити значення виразу: $\frac{A_{15}^6 + C_{19}^{15}}{P_{14}} - \frac{1}{2} A_{21}^4$.

2. У цеху працюють шість чоловіків і чотири жінки. За табельними номерами навмання відібрано семеро. Знайти ймовірність того, що серед відібраних працівників буде саме три жінки.

3. Три студенти складають іспит. Перший із них вивчив 85 % питань програми, другий – 80 %, третій – 90 %. Знайти ймовірність того, що лише один студент складе іспит.

4. Біля верстата знаходиться три деталі, з яких одна є нестандартною. До них додали дві деталі, якість яких невідома. Потім навмання взяли одну деталь. Визначити ймовірність того, що ця деталь буде стандартною, якщо відносно якості двох останніх деталей можна зробити будь-яке припущення.

5. За умовою завдання 4 визначити апостеріорні ймовірності кожної з гіпотез, якщо деталь, що взята для перевірки, виявилася стандартною.

Варіант 1.14

1. Обчислити значення виразу: $\frac{A_{16}^7 - C_{14}^{11}}{3P_7} - C_{20}^{17}$.

2. У шухляді міститься 10 однакових деталей, які мають номери від 1 до 10. Навмання вибрали шість деталей. Знайти ймовірність того, що серед них виявляться деталь № 1 та деталь № 2.

3. Із трьох урн, які містять білі та чорні кульки, витягують по одній кульці. Імовірність вилучення білої кульки з першої урни дорівнює 0,9, з другої – 0,8, з третьої – 0,75. Знайти ймовірність того, що з другої і третьої урн вилучать білі кульки.

4. Підприємство щотижня постачає товар фірмовим магазинам у співвідношенні 2:3:5. Імовірність того, що протягом тижня розпродано весь товар даного підприємства для першого магазину становить 0,75, для другого – 0,85, а для третього – 0,9. Визначити ймовірність того, що наприкінці тижня в магазині весь товар даного підприємства буде розпродано, якщо покупець вибирає один із трьох магазинів навмання.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо того, який саме магазин продасть увесь товар.

Варіант 1.15

1. Обчислити значення виразу: $\frac{2A_{22}^9 - C_{12}^7}{P_{17}} + A_{22}^8$.

2. У шухляді є 7 білих і 6 червоних кульок. Із шухляди навмання беруть 3 кульки. Знайдіть ймовірність того, що всі кульки будуть одного кольору.

3. Робітник обслуговує три пристрої. Імовірність того, що протягом години перший пристрій не зупиниться, дорівнює 0,81, другий – 0,65, третій – 0,43. Визначити ймовірність того, що протягом години зупиняться будь-які два пристрої.

4. Магазин здійснює перевезення вантажу трьома автопарками у співвідношенні 3:1:4. Імовірність дотримання терміну перевезення становить для першого автопарку 0,9, для другого – 0,85, для третього – 0,8. Визначити ймовірність того, що товар поставлено відповідно до терміну.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо перевізника, який порушив термін постачання, якщо товар не було перевезено в означений термін.

Варіант 1.16

1. Обчислити значення виразу: $\frac{2A_{21}^8 - C_{10}^7}{P_{10}} + A_{24}^8$.

2. Набір трицифрового номера білета, який виграє, виконується триразовим автоматичним викиданням із ящика одного за одним трьох жетонів із загальної кількості 9 жетонів, пронумерованих цифрами від 1 до 9. Знайти ймовірність того, що набраний номер не містить цифри 7.

3. Прилад складається з трьох вузлів, які працюють незалежно один від одного. Імовірності виходу з ладу вузлів на заданому часовому проміжку становлять відповідно 0,2, 0,3 і 0,4. Знайти ймовірність того, що протягом заданого часу прилад працюватиме.

4. Є три партії деталей, з яких перша містить 10 стандартних і 4 нестандартних деталей, друга – 14 стандартних і 4 нестандартних, третя – 16 стандартних і 5 нестандартних. Із партії, що вибрана навмання, беруть деталь. Знайти ймовірність того, що деталь виявиться стандартною.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо партії деталей, якщо деталь виявилася нестандартною.

Варіант 1.17

1. Обчислити значення виразу: $\frac{3A_{27}^9 - C_{17}^8}{P_9} - A_{22}^8$.

2. На складі є 10 двигунів заводу № 1 і вісім двигунів заводу № 2. Навмання взято чотири двигуни. Знайти ймовірність того, що серед них два двигуни заводу № 1 і два двигуни заводу № 2.

3. У партії із 20 деталей 15 стандартних, а решта – нестандартні. Навмання беруть 4 деталі. Знайти ймовірність того, що серед них принаймні 1 нестандартна.

4. Для посіву пшениці заготовлено насіння, серед якого 95 % 1-го ґатунку, 3 % – 2-го та 2 % – 3-го ґатунку. Ймовірність того, що з насінини виросте колосок, в якому не менш ніж 50 зерен, для насіння 1-го ґатунку становить 0,5, 2-го – 0,2 та 3-го – 0,1. Знайти ймовірність того, що навмання взятий колосок у разі такого посіву матиме не менш ніж 50 зерен.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо сорту насіння, якщо навмання взятий колосок має менше 50 зерен.

Варіант 1.18

1. Обчислити значення виразу: $\frac{2A_{28}^9 + C_{18}^7}{P_{16}} + A_{25}^9$.

2. Набір трицифрового номера виграшної облігації виконують триразовим викиданням з урни одного за одним трьох жетонів із п'яти, пронумерованих цифрами від 1 до 5. Знайти ймовірність того, що вибраний номер містить цифру 3.

3. У цеху є три резервні мотори, для кожного з яких ймовірність бути ввімкненим у даний момент дорівнює 0,3. Знайти ймовірність того, що в даний момент ввімкнено принаймні один мотор.

4. Є три партії зовні однакових деталей. У першій партії 20 стандартних і 5 нестандартних деталей, у другій – 15 стандартних і 3 нестандартні, у третій – 14 стандартних і 2 нестандартні деталі. Із навмання вибраної партії взяли деталь. Знайти ймовірність того, що вона виявилася стандартною.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо партії деталей, якщо вибрана деталь виявилася нестандартною.

Варіант 1.19

1. Обчислити значення виразу: $\frac{3A_{24}^9 - C_{19}^7}{P_{19}} + A_{26}^7$.

2. Партія з 10 деталей містить 4 браковані. Знайти ймовірність того, що з навмання взятих двох деталей будуть дві придатні.

3. Перевірка партії рису, розфасованого по 0,5 кг, дала такі результати: 20 % усіх пачок були по 498 г, 60 % – по 500 г, 20 % – по 502 г. Із партії навмання взяли дві пачки. Знайти ймовірність того, що обидві пачки мають однакову масу.

4. На конвеєр надходять деталі від трьох автоматів. Перший автомат дає 90 %, другий – 93 %, а третій – 95 % стандартної продукції. Протягом зміни від першого автомата надходить 60 деталей, від другого – 50, від третього – 40. Знайти ймовірність потрапляння на конвеєр нестандартної деталі.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо автомата, який виготовляє якісну продукцію, якщо на конвеєр потрапила стандартна деталь.

Варіант 1.20

1. Обчислити значення виразу: $\frac{2A_{24}^8 + C_{20}^7}{P_{18}} - A_{28}^8$.

2. За підсумком року акції десяти фірм мали прибуток, чотирьох фірм знецінились, а акції шести фірм – зберегли свою номінальну вартість. Знайти ймовірність того, що випадково куплені шість акцій різних фірм будуть прибутковими.

3. Із партії суконь дівчина має намір вибрати дешеву. Ймовірність того, що навмання взята сукня дешева, дорівнює 0,7. Знайти ймовірність того, що із трьох суконь тільки дві будуть дешевими.

4. На двох верстатах-автоматах виробляються однакові заготовки, які транспортером перекидаються в те саме місце. Продуктивність другого верстата в 1,5 рази більша, ніж першого. Перший верстат дає 5 % нестандартних заготовок, а другий – 93 % стандартних. Знайти ймовірність того, що взята навмання заготовка буде стандартна.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо верстата, якщо взята навмання заготовка виявилася стандартною.

Варіант 1.21

1. Обчислити значення виразу: $\frac{4A_{26}^7 - C_{21}^7}{P_{17}} - A_{25}^9$.

2. Партія складається з 20 виробів, з яких 8 виробів 1-го ґатунку, 6 – 2-го, 2 – 3-го ґатунку, а решта – браковані. Навмання беруть 4 вироби. Знайти ймовірність того, що серед них виявилось 2 вироби 1-го ґатунку, 1 – 2-го ґатунку і 1 – бракований.

3. Випущено 45 лотерейних білетів, серед яких 12 виграшних. Гравець придбав 3 білети. Знайти ймовірність того, що серед них буде принаймні один виграшний.

4. Імовірність повної сплати податків для першої групи підприємств $4/5$, для другої групи ця ймовірність задовольняє рівняння $9 - 9p = 4p^2$. До першої групи входить 70 підприємств, а до другої групи – 30. Знайти ймовірність повної сплати податків.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо групи підприємств, якщо виявлено неповну сплату податків.

Варіант 1.22

1. Обчислити значення виразу: $\frac{3A_{21}^8 + C_{22}^9}{P_{18}} + A_{24}^8$.

2. У партії із 16 деталей чотири нестандартні. Навмання беруть три деталі. Знайти ймовірність того, що серед них дві деталі будуть стандартними.

3. Перевірка партії рису, розфасованого по 0,5 кг, дала такі результати: 60 % усіх пачок мали вагу 500 г, 20 % – нижче за 496 г, 20 % – вище за 504 г. Із партії навмання взято дві пачки. Знайти ймовірність того, що маса обох пачок відхиляється від норми більше за 4 г.

4. Металеві заготовки для подальшої обробки надходять із двох цехів: 55 % – із 1-го, 45 % – із 2-го. При цьому продукція з 1-го цеху містить 3 %, а з 2-го – 5 % браку. Знайти ймовірність того, що заготовка, яка надійшла на оброблення, виявиться придатною.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо цеху, якщо заготовка, яка надійшла на оброблення, не придатна.

Варіант 1.23

1. Обчислити значення виразу: $\frac{2A_{22}^9 + C_{23}^{11}}{P_{21}} - A_{22}^{12}$.

2. Для молодіжної вечірки діджей наготував 20 компакт-дисків, 7 з яких з інструментальною музикою. Знайти ймовірність того, що з чотирьох навмання відібраних компакт-дисків три будуть з інструментальною музикою.

3. Механізм, що містить 4 однакові деталі, не працюватиме, якщо під час його складання буде взято 3 або більше деталей меншого розміру, ніж потрібно. У робітника є 15 деталей, серед яких 6 меншого розміру. Знайти ймовірність того, що механізм працюватиме, якщо робітник братиме деталі навмання.

4. Ймовірність виконання договору для першої фірми 0,5, для другої – 0,7, для третьої ця ймовірність складає 60 % від суми ймовірностей першої та другої фірми. Податкова інспекція довільним чином перевіряє кожну з фірм. Знайти ймовірність того, що під час перевірки буде виявлено порушення договору.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо фірми, якщо перевірка не виявила порушення договору.

Варіант 1.24

1. Обчислити значення виразу: $\frac{5A_{24}^9 + 2C_{15}^7}{P_{18}} + A_{24}^8$.

2. У конверті 10 акцій, серед яких три фірми А. Навмання відібрано 4 акції. Яка ймовірність того, що серед них буде одна акція фірми А?

3. Маємо дві партії деталей. У першій партії 7 придатних і 3 браковані деталі. У другій – 10 придатних і 4 браковані. Із кожної партії навмання беруть по одній деталі. Знайти ймовірність того, що обидві деталі придатні.

4. Вивчаються результати іспиту з теорії ймовірностей у трьох групах. У першій групі 30 студентів, з них 8 отримали відмінну оцінку, в другій – відповідно 28 і 6, а в третій – відповідно 32 і 8. Яка ймовірність того, що навмання вибраний студент отримав за іспит відмінну оцінку?

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо групи студентів, в якій студент отримав на заліку відмінну оцінку.

Варіант 1.25

1. Обчислити значення виразу: $\frac{4A_{25}^8 - C_{18}^9}{2P_7} - 3A_{21}^9$.

2. Серед 30 видів акцій будівельних організацій 19 стали прибутковими, 5 – збитковими, а 6 – залишилися без змін. Яка ймовірність того, що серед п'яти навмання придбаних акцій різних видів прибутковими виявляться саме три?

3. Перша партія складається з 8 придатних і 3 бракованих деталей, друга – з 12 придатних і 4 бракованих. Із кожної партії навмання беруть по одній деталі. Знайти ймовірність того, що одна деталь буде придатною, а інша – бракованою.

4. Фірма встановлює тюнери, 70 % яких виготовлено на заводі № 1, а решта – на заводі № 2. Ймовірність того, що тюнер працюватиме протягом гарантійного терміну для заводу № 2 дорівнює 0,9, тоді як для заводу № 1 – 0,8. Знайти ймовірність того, що навмання взятий тюнер не відпрацює гарантійного терміну.

5. За умовою завдання 4 визначити апостеріорні ймовірності гіпотез щодо заводу, який виготовив тюнер, якщо тюнер зламався.

1.7. Контрольні запитання

1. Наведіть формули для визначення перестановок, розміщень та сполучень за умов, що елементи множини не повторюються.

2. Поясніть, чим відрізняються розміщення від сполучень.

3. Які події називаються випадковими, неможливими, вірогідними? Наведіть приклади таких подій.

4. Що таке "простір елементарних подій"?

5. Дайте класичне означення ймовірності випадкової події.

6. Що таке "повна група несумісних подій"? Наведіть приклад випадкових подій, що утворюють повну групу подій.

7. Яка подія є протилежною до вихідної?

8. Які події називаються залежними? Наведіть приклади залежних та незалежних подій.

9. Дайте означення умовної та безумовної ймовірностей.

10. Дайте означення сумісних подій. Наведіть приклади сумісних випадкових подій та випадкової події, що є результатом їх перетину на просторі повної групи подій.

11. Сформулюйте теорему додавання ймовірностей.
12. Сформулюйте теорему множення ймовірностей.
13. Охарактеризуйте властивості гіпотез, які утворюють повну групу несумісних подій. Наведіть приклад гіпотез.
14. Поясніть, чому у формулі повної ймовірності ймовірність суми подій $P(\sum_{i=1}^n A \cdot H_i)$ визначається як сума їх ймовірностей.
15. Наведіть формулу Байєса для визначення апостеріорної ймовірності гіпотези.
16. Сформулюйте властивості апостеріорних ймовірностей щодо гіпотез, які висуваються при визначенні повної ймовірності події.

Лабораторна робота 2

Дискретні випадкові величини: основні числові характеристики та їх властивості

2.1. Мета роботи:

ознайомлення з поняттям дискретної випадкової величини;
формування навиків визначення основних числових характеристик випадкової величини за означенням та за допомогою вбудованих функцій *MS Excel*;

засвоєння принципів побудови закону розподілу дискретної випадкової величини, яка є алгебраїчною сумою або добутком двох випадкових величин, та визначення основних числових характеристик нової випадкової величини за числовими характеристиками її компонентів.

2.2. Теоретичні положення

Випадковою величиною називається така величина, яка в результаті випробування приймає одне із своїх можливих значень, яке невідоме заздалегідь і залежить від випадкових причин. Або коротко: **випадкова величина** – це кількісна ознака наслідку випадкового експерименту. Якщо мова йде взагалі про випадкову величину, то у теорії ймовірностей вона позначається малими літерами грецького алфавіту ξ або η . Якщо необхідно вказати їх можливі значення, то ці значення позначають малими літерами x_i або y_i . Те, що випадкова величина прийме одне із своїх значень, є **випадковою подією**.

Розрізняють дискретні та неперервні випадкові величини. **Дискретною випадковою величиною** (ДВВ) називається випадкова величина, яка має скінченну або нескінченну, але зліченну кількість значень, які змінюються стрибками. Якщо x_i та x_{i+1} є двома послідовними значеннями дискретної випадкової величини ξ , то вона не може приймати значення із проміжку $(x_i; x_{i+1})$. Або іншими словами, випадкова величина називається **дискретною**, якщо значення, які вона приймає, можна пронумерувати. Отже значення випадкової величини ξ утворюють множину $(x_1, x_2, x_3, \dots, x_n)$, де n – кількість можливих значень, яка є **зліченою**, тобто такою, елементи якої можна пронумерувати. Наприклад, дискретною є випадкова величина, яка визначає кількість студентів, що перебувають у даний момент часу на сайті персональних навчальних систем. Випадкова величина, множина значень якої заповнює суцільно деякий числовий (скінченний або нескінченний) проміжок, називається **неперервною випадковою величиною** (НВВ). Наприклад, неперервною випадковою величиною є вага коробки з цукерками у межах припустимих відхилень від її точної ваги.

Очевидно, що для повної характеристики дискретної випадкової величини недостатньо знати лише її значення. Необхідно також у відповідність цим значенням поставити ймовірності, з якими випадкова величина їх приймає. Саме відповідність між усіма можливими значеннями дискретної випадкової величини та їх ймовірностями називається **законом розподілу**. Закон розподілу, як і будь-яку функцію, можна задати за допомогою таблиці, графічно або аналітично. Найпростішою формою завдання закону розподілу ДВВ є **ряд розподілу**, що має вигляд таблиці (табл. 2.1), у першому рядку якої містяться всі її можливі значення x_i , а в другій – відповідні їм ймовірності $p_i = P(\xi = x_i)$.

Таблиця 2.1

Ряд розподілу дискретної випадкової величини

x_i	x_1	x_2	x_3	x_n
p_i	p_1	p_2	p_3	p_n

Під час побудови ряду розподілу необхідно мати на увазі, що

$$\sum_{i=1}^n p_i = 1, \quad \text{при цьому } 0 \leq p_i \leq 1, \quad (2.1)$$

оскільки випадкові події $\xi = x_i$, де $i = \overline{1, n}$, утворюють повну групу попарно несумісних подій.

У процесі графічного завдання закону розподілу ДВВ у декартовій системі координат відкладають по осі абсцис можливі значення випадкової величини, а по осі ординат – відповідні їм ймовірності. Точки з координатами $(x_i; P(\xi = x_i))$ з'єднують відрізками й отримують **многокутник розподілу**. Крім того, деякі закони розподілу можна задати **аналітично**, тобто за допомогою формули.

Сумою двох випадкових величин $\xi_1 = (x_1, x_2, \dots, x_n)$ та $\xi_2 = (y_1, y_2, \dots, y_m)$ називається випадкова величина $\eta = \xi_1 + \xi_2$, де $\eta = (z_1, z_2, \dots, z_k)$, значення якої обчислюються як $z_k = x_i + y_j$ ($i = \overline{1, n}$, $j = \overline{1, m}$). Імовірність кожного значення випадкової величини дорівнює добутку ймовірностей її доданків: $P(\eta = z_k) = P(\xi_1 = x_i) \cdot P(\xi_2 = y_j)$.

Аналогічно **різницею випадкових величин** $\xi_1 = (x_1, x_2, \dots, x_n)$ та $\xi_2 = (y_1, y_2, \dots, y_m)$ називається випадкова величина $\eta = \xi_1 - \xi_2$, де $\eta = (z_1, z_2, \dots, z_k)$, значення якої обчислюються як $z_k = x_i - y_j$ ($i = \overline{1, n}$, $j = \overline{1, m}$). Імовірність кожного значення випадкової величини $\eta = \xi_1 - \xi_2$ дорівнює добутку ймовірностей, з якими випадкові величини ξ_1 та ξ_2 приймають відповідні значення: $P(\eta = z_k) = P(\xi_1 = x_i) \cdot P(\xi_2 = y_j)$.

Добутком незалежних випадкових величин $\xi_1 = (x_1, x_2, \dots, x_n)$ та $\xi_2 = (y_1, y_2, \dots, y_m)$ називається випадкова величина $\eta = \xi_1 \cdot \xi_2$, де $\eta = (z_1, z_2, \dots, z_k)$, значення якої обчислюються як $z_k = x_i \cdot y_j$ ($i = \overline{1, n}$, $j = \overline{1, m}$). Імовірність кожного значення випадкової величини $\eta = \xi_1 \cdot \xi_2$ дорівнює добутку ймовірностей, що відповідають відповідним значенням цих множників: $P(\eta = z_k) = P(\xi_1 = x_i) \cdot P(\xi_2 = y_j)$.

Підкреслимо, що закон розподілу випадкової величини становить її вичерпну характеристику. Однак, під час розв'язання багатьох задач більш

зручно розглядати не закон розподілу, а його числові характеристики, які в стислій формі відображають найбільш суттєві особливості розподілу.

Залежно від роду інформації числові характеристики поділяють на такі групи:

характеристики положення випадкової величини на числовій осі (мода M_o , медіана M_e , математичне сподівання $M(\xi)$);

характеристики розпорошення випадкової величини навколо її середнього значення (дисперсія $D(\xi)$, середнє квадратичне відхилення $\sigma(\xi)$, коефіцієнт варіації $V(\xi)$);

характеристики форми многокутника розподілу (коефіцієнт асиметрії A_s , коефіцієнт ексцесу E_x).

У свою чергу, ці характеристики поділяють на основні та допоміжні. До числа **основних** (найбільш важливих) **числових характеристик** належать математичне сподівання, дисперсія та середнє квадратичне відхилення. Інші характеристики вважаються **допоміжними**.

Для дискретної випадкової величини, яка має скінчену множину можливих значень, **математичним сподіванням $M(\xi)$ випадкової величини ξ** називають суму добутків усіх можливих значень випадкової величини на ймовірності, з якими випадкова величина приймає ці значення:

$$M(X) = x_1 p_1 + x_2 p_2 + x_3 p_3 + \dots + x_n p_n = \sum_{i=1}^n x_i p_i. \quad (2.2)$$

Математичне сподівання випадкової величини також скорочено називають **математичним сподіванням**. Мовою моментів математичне сподівання є **початковим моментом 1-го порядку**.

Математичне сподівання має такі властивості:

математичне сподівання сталої величини дорівнює самій сталій, тобто $M(C) = C$, де $C = const$;

сталій множник можна виносити за знак математичного сподівання, тобто $M(C\xi) = C \cdot M(\xi)$, де $C = const$;

математичне сподівання відхилення випадкової величини від свого математичного сподівання дорівнює нулю (**центральна властивість математичного сподівання**), тобто $M(\xi - M(\xi)) \equiv 0$;

математичне сподівання алгебраїчної суми двох випадкових величин дорівнює алгебраїчній сумі математичних сподівань її доданків, тобто $M(\xi_1 \pm \xi_2) = M(\xi_1) \pm M(\xi_2)$;

математичне сподівання добутку незалежних випадкових величин дорівнює добутку їх математичних сподівань, тобто $M(\xi_1 \cdot \xi_2) = M(\xi_1) \cdot M(\xi_2)$.

Математичне сподівання випадкової величини визначається як середнє виважене значень випадкової величини, кожне з яких беруть з вагою, яка дорівнює ймовірності цього значення. Дисперсія і середнє квадратичне відхилення характеризують розпорошення значень випадкової величини навколо її математичного сподівання.

Дисперсією $D(\xi)$ випадкової величини ξ називають математичне сподівання квадрату відхилення цієї випадкової величини від її математичного сподівання:

$$D(\xi) = M(\xi - M(\xi))^2. \quad (2.3)$$

За означенням дисперсія дискретної випадкової величини обчислюється за формулою:

$$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + \dots + (x_n - M(X))^2 p_n = \sum_{i=1}^n (x_i - M(X))^2 p_i. \quad (2.3')$$

Отже, дисперсія є **центральним моментом 2-го порядку**.

Для обчислення дисперсії зручно застосовувати перетворену формулу:

$$D(\xi) = M(\xi^2) - M^2(\xi), \quad (2.4)$$

де $M(\xi^2)$ – математичне сподівання квадрата випадкової величини ξ , тобто **початковий момент 2-го порядку**.

Отже, для дискретної випадкової величини маємо:

$$D(X) = \sum_{i=1}^n x_i^2 \cdot p_i - (M(X))^2. \quad (2.4')$$

Дисперсія має такі властивості:

дисперсія завжди невід'ємна, тобто $D(\xi) \geq 0$;

дисперсія сталої величини дорівнює нулю, тобто $D(C) = 0$, де $C = const$;

сталій множник можна виносити за знак дисперсії, підносячи його до квадрата, тобто $D(C\xi) = C^2 \cdot D(\xi)$, де $C = const$;

дисперсія суми (різниці) двох незалежних випадкових величин дорівнює сумі їх дисперсій, тобто $D(\xi_1 \pm \xi_2) = D(\xi_1) + D(\xi_2)$;

дисперсія добутку двох незалежних величин визначається співвідношенням:

$$D(\xi_1 \cdot \xi_2) = D(\xi_1)D(\xi_2) + M^2(\xi_1)D(\xi_2) + M^2(\xi_2)D(\xi_1).$$

Оскільки дисперсія має розмірність квадрата випадкової величини, то доцільно розглядати таку характеристику розпорошення, розмірність якої збігається з розмірністю випадкової величини. Такою числовою характеристикою є середнє квадратичне відхилення, яке визначається як квадратний корінь із дисперсії:

$$\sigma(X) = \sqrt{D(X)}. \quad (2.5)$$

Зрозуміло, що $\sigma(X) > 0$ і середнє квадратичне відхилення для сталої величини дорівнює нулю, тобто $\sigma(C) = 0$, де $C = const$.

2.3. Змістовна постановка задачі

Задані закони розподілу двох незалежних випадкових величин ξ_1 і ξ_2 (табл. 2.2).

Таблиця 2.2

Закони розподілу випадкових величин ξ_1 і ξ_2

x_i	$P(\xi_1 = x_i)$		y_i	$P(\xi_2 = y_j)$
1	0,2		-1	0,3
2	0,5		1	0,5
3	0,3		3	0,2

За даними табл. 2.2 необхідно:

скласти закон розподілу випадкової величини $\eta = \xi_1 - 2\xi_2$;

визначити основні числові характеристики випадкових величин ξ_1 і ξ_2 ;

обчислити математичне сподівання, дисперсію випадкової величини $\eta = \xi_1 - 2\xi_2$ за їх означеннями та з урахуванням властивостей цих числових характеристик;

знайти закон розподілу випадкової величини $\eta = 3\xi_1 \cdot \xi_2$;

визначити математичне сподівання та дисперсію випадкової величини $\eta = 3\xi_1 \cdot \xi_2$ за означенням та зробити висновки щодо зв'язку математичного сподівання та дисперсії випадкової величини η з основними числовими характеристиками випадкових величин ξ_1 і ξ_2 .

2.4. Приклад виконання лабораторної роботи 2

На робочому аркуші *MS Excel* у вигляді двох вертикальних таблиць заносимо вихідні умови задачі (рис. 2.1).

	A	B	C	D	E	
1	x_i	p_i		y_i	p_i	
2	1	0,2		-1	0,3	
3	2	0,5		1	0,5	
4	3	0,3		3	0,2	
5						

Рис. 2.1. Скриншот умови задачі

Знайдемо всі можливі значення випадкової величини $\eta = \xi_1 - 2\xi_2$. Для цього створюємо комбінації значень випадкової величини ξ_1 зі значеннями величини ξ_2 за таким принципом.

Розрахунки значень шуканої випадкової величини $\eta = \xi_1 - 2\xi_2$ та їх ймовірностей зручно проводити в таблиці (рис. 2.2).

Так, у комірку **G2** вводимо формулу **=A\$2\$-2*D2**.

	A	B	C	D	E	F	G	H	I
1	x_i	p_i		y_i	p_i		$x_i - 2y_i$	p_i	
2	1	0,2		-1	0,3		3	0,06	
3	2	0,5		1	0,5		-1	0,10	
4	3	0,3		3	0,2		-5	0,04	
5							4	0,15	
6							0	0,25	
7							-4	0,10	
8							5	0,09	
9							1	0,15	
10							-3	0,06	
11							Сума	1,00	

Рис. 2.2. Скриншот обчислення закону розподілу випадкової величини $\eta = \xi_1 - 2\xi_2$

Посилання на комірку **A2** робимо абсолютним (натисканням клавіші **F4**), що дає змогу протягнувши цю формулу на комірки **G2:G4**, отримати всі можливі комбінації значення $x_1 = 1$ випадкової величини ξ_1 з усіма значеннями випадкової величини ξ_2 . Їх імовірності визначаємо у комірках **H2:H4**, протягнувши формулу **=B\$2*\$E2**.

Тепер у комірку **G5** вводимо формулу **=A\$3\$-2*D2**, а у комірку **H5** – формулу **=B\$3*\$E2** і протягуємо їх одночасно на рядки **5, 6** і **7**, завдяки чому отримуємо всі можливі комбінації значення $x_2 = 2$ випадкової величини ξ_1 з усіма значеннями випадкової величини ξ_2 і їх ймовірності. Далі пропонуємо самостійно обчислити всі можливі комбінації значення $x_3 = 3$ випадкової величини ξ_1 з усіма значеннями випадкової величини ξ_2 і їх ймовірності. Результат наведено на рис. 2.2.

Переконаємося, що ми визначили усі можливі значення випадкової величини $\eta = \xi_1 - 2\xi_2$ (сума їх імовірностей повинна дорівнювати 1). Для цього у вільну комірку вводимо формулу: **=СУММ(H2:H10)**. Натискаємо **Enter** і дійсно отримуємо 1. Отже, у комірках **G2:H10** виведено закон розподілу випадкової величини $\eta = \xi_1 - 2\xi_2$. Зверніть увагу, що у комірках **G2:G10** наведені 9 значень шуканої випадкової величини і кожне з них зустрічається лише раз. Після ранжування цих значень закон розподілу запишемо у вигляді таблиці (табл. 2.3).

Закон розподілу випадкової величини $\eta = \xi_1 - 2\xi_2$

$\eta = \xi_1 - 2\xi_2$	-5	-4	-3	-1	0	1	3	4	5
p_i	0,04	0,10	0,06	0,10	0,25	0,15	0,06	0,15	0,09

Слід зауважити, якщо якесь із значень шуканої випадкової величини повторюється, то ймовірності цих подій треба додавати (події несумісні, отже, за аксіомою III їх ймовірності додаються).

Перейдемо тепер до обчислення основних числових характеристик випадкової величини $\eta = \xi_1 - 2\xi_2$ безпосередньо за означенням. Для цього на робочому аркуші *MS Excel* побудуємо таблицю, у якій будемо проводити допоміжні обчислення (рис. 2.3).

J16	=J13^2*J14										
	A	B	C	D	E	F	G	H	I	J	K
12											
13	$\eta=Z_i$	-5	-4	-3	-1	0	1	3	4	5	Сума
14	p_i	0,04	0,10	0,06	0,10	0,25	0,15	0,06	0,15	0,09	1,00
15	z_i*p_i	-0,20	-0,40	-0,18	-0,10	0,00	0,15	0,18	0,60	0,45	0,50
16	$(z_i)^2*p_i$	1,00	1,60	0,54	0,10	0,00	0,15	0,54	2,40	2,25	8,58
17											

Рис. 2.3. Скриншот обчислення основних числових характеристик випадкової величини $\eta = \xi_1 - 2\xi_2$

Для обчислення математичного сподівання треба попарно перемножити значення випадкової величини (комірки **B13:J13**) та відповідні цим значенням імовірності (комірки **B14:J14**) і результати додати.

Введемо у комірку **B15** формулу: **=B13*B14** і розтягнемо її на комірки **B15:J15**. Тоді у комірці **K15** набираємо формулу **=СУММ(B15:J15)** і натисканням **Enter** виводимо значення математичного сподівання випадкової величини $\eta = \xi_1 - 2\xi_2$. Отримуємо результат: $M(\eta) = 0,5$.

Підрахунок дисперсії будемо здійснювати за перетвореною формулою. Для цього потрібно визначити початковий момент другого порядку, тобто добуток квадрату кожного значення випадкової величини на відповідну цим значенням імовірність. Отже, у комірку **B16** вводимо формулу: **=B13^2*B14** і розтягнемо її на комірки **B16:J16**.

Тепер у комірці **K15** набираємо формулу **=СУММ(B16:J16)** і натисканням **Enter** виводимо результат: $M(\eta^2) = 8,58$.

Обчислюємо дисперсію:

$$D(\eta) = M(\eta^2) - M^2(\eta) = 8,58 - 0,5^2 = 8,33.$$

У процесі визначення основних числових характеристик можна скористатись вбудованою функцією **СУММПРОИЗВ**. Для цього у будь-яку вільну комірку (хай для визначеності це буде комірка **K2**) вводимо формулу: **=СУММПРОИЗВ(B13:J13;B14:J14)** і натискаємо **Enter**. Отримуємо той самий результат: $M(\eta) = 0,5$. Або можна викликати діалогове вікно цієї функції і у полі Масив 1 дати посилання на комірки **B13:J13**, а у полі Масив 2 – посилання на комірки **B14:J14** (рис. 2.4).

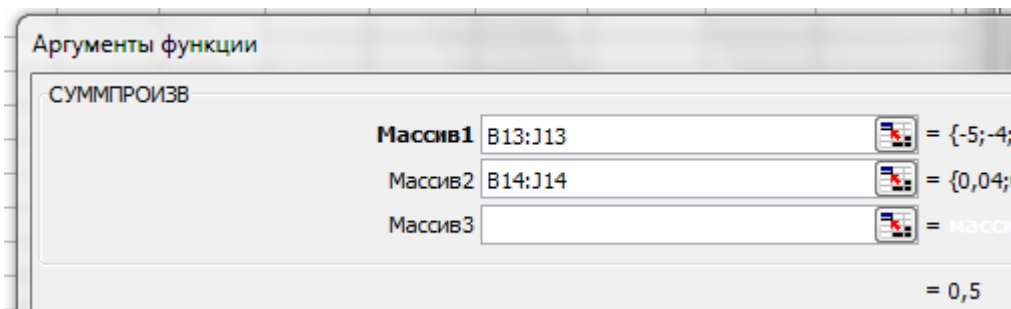


Рис. 2.4. Фрагмент діалогового вікна функцією **СУММПРОИЗВ** для обчислення математичного сподівання

Для підрахунку дисперсії теж застосуємо функцію **СУММПРОИЗВ**, ввівши у вільну комірку **K4** формулу:

$$=СУММПРОИЗВ(B13:J13; B13:J13;B14:J14)-K2^2.$$

Натискаємо **Enter** і отримуємо результат: $D(\eta) = 8,33$. Зауважимо, що аргументами функції **СУММПРОИЗВ** є діапазон комірок **B13:J13** (**масив 1** та **масив 2**) і діапазон комірок **B14:J14** (**масив 3**). Слід також відняти квадрат математичного сподівання, що міститься в комірці **K2**.

Для обчислення середнього квадратичного відхилення в комірку **K6** вводимо формулу: **=K4^0,5** і після натискання **Enter** отримуємо результат: $\sigma(\eta) = 2,89$. Отже, ми обчислили основні числові характеристики випадкової величини η , закон розподілу якої надано в табл. 2.3. Результати обчислень подано на рис. 2.5.

K10		fx =B8+2^2*E8									
	A	B	C	D	E	F	G	H	I	J	K
1	x_i	p_i		y_i	p_i		$x_i - 2y_i$	p_i			
2	1	0,2		-1	0,3		3	0,06		$M(\eta) = 0,5$	
3	2	0,5		1	0,5		-1	0,10			
4	3	0,3		3	0,2		-5	0,04		$D(\eta) = 8,33$	
5							4	0,15			
6	$M(\xi_1) = 2,1$			$M(\xi_2) = 0,8$			0	0,25		$\sigma(\eta) = 2,89$	
7							-4	0,10			
8	$D(\xi_1) = 0,49$			$D(\xi_2) = 1,96$			5	0,09		$M(\xi_1 - 2\xi_2) = 0,5$	
9							1	0,15			
10	$\sigma(\xi_1) = 0,7$			$\sigma(\xi_2) = 1,4$			-3	0,06		$D(\xi_1 - 2\xi_2) = 8,33$	
11							Сума	1,00			
12											

Рис. 2.5. Скриншот з результатами обчислень

Порівняємо значення основних числових характеристик випадкової величини η , які були отримані при обчисленні за їх означеннями, з тими, що можна отримати за властивостями математичного сподівання та дисперсії. Для цього спочатку обчислюємо основні числові характеристики випадкових величин ξ_1 і ξ_2 за допомогою функції **=СУММПРОИЗВ**. Так, у комірку **B6** вводимо формулу: **=СУММПРОИЗВ(A2:A4;B2:B4)** і після натискання **Enter** отримуємо значення математичного сподівання для випадкової величини ξ_1 : $M(\xi_1) = 2,1$. Дисперсію цієї випадкової величини виводимо у комірку **B8**, ввівши туди формулу:

$$=СУММПРОИЗВ(A2:A4;A2:A4;B2:B4)-B6^2.$$

Маємо результат: $D(\xi_1) = 0,49$. У комірці **B10** виводимо значення середнього квадратичного відхилення за допомогою формули: **=B8^0,5**.

Аналогічно проводимо обчислення числових характеристик випадкової величини ξ_2 (пропонуємо зробити це самостійно). У комірку **E6** виводимо значення математичного сподівання $M(\xi_2) = 0,8$, у комірку **E8** – значення дисперсії $D(\xi_2) = 1,96$ та у комірку **E10** – значення середнього квадратичного відхилення $\sigma(\xi_2) = 1,4$.

За властивостями математичного сподівання маємо:

$$M(\xi_1 - 2\xi_2) = M(\xi_1) - 2 \cdot M(\xi_2) = 2,1 - 2 \cdot 0,8 = 0,5.$$

Для виконання цих обчислень в електронній таблиці у комірку **K6** вводимо формулу: **=B6-2*E6**. Отже, ми отримали той самий результат. Дійсно $M(\eta) = M(\xi_1) - 2 \cdot M(\xi_2)$.

Тепер обчислюємо дисперсію випадкової величини $\xi_1 - 2\xi_2$, ввівши у комірку **K10** формулу: **=B8+2^2*E8**. Отримали такий результат: $D(\xi_1 - 4 \cdot \xi_2) = 8,33$, тобто $D(\eta) = D(\xi_1) + 4 \cdot D(\xi_2)$.

Переходимо тепер до розв'язання наступного завдання і побудуємо ряд розподілу випадкової величини $\eta = 3\xi_1 \cdot \xi_2$.

Будемо дотримуватись тих самих принципів і логіки розв'язання, які ми застосовували у попередньому завданні.

Продовжуємо виконання лабораторної роботи 2 на робочому аркуші **Лист 2 MS Excel**, де у вигляді таблиць розміщуємо вихідні дані завдання та результати обчислень основних числових характеристик випадкових величин ξ_1 і ξ_2 (див. рис. 2.1).

Знайдемо всі можливі значення випадкової величини $\eta = 3\xi_1 \cdot \xi_2$, утворюючи всі можливі комбінації значень випадкових величин ξ_1 і ξ_2 шляхом послідовного множення значень випадкової величини ξ_1 на кожне із значень випадкової величини ξ_2 з урахуванням коефіцієнта 3 (комірки **G2:G10**).

У комірках **H2:H10** виводимо ймовірності цих значень як добуток відповідних ймовірностей випадкових величин ξ_1 і ξ_2 (рис. 2.6). Перевірка суми ймовірностей значень випадкової величини за формулою **=СУММ(H2:H10)** показує, що знайдені всі значення.

Зверніть увагу, що деякі значення випадкової величини повторюються (див. **H2:H10**). Отже, шукана випадкова величина $\eta = 3\xi_1 \cdot \xi_2$ приймає такі значення: -9, -6, -3, 3, 6, 9, 18, 27.

Так, значення 9 відповідає двом елементарним подіям: $\omega_3(x_1 = 1; y_3 = 3)$ та $\omega_8(x_3 = 3; y_2 = 1)$. Оскільки ці події несумісні, то їх імовірності додаються.

Отже, маємо:

$$P(\eta = 9) = 0,04 + 0,15 = 0,19.$$

	A	B	C	D	E	F	G	H
1	x_i	p_i		y_i	p_i		$3x_i \cdot y_i$	p_i
2	1	0,2		-1	0,3		-3	0,06
3	2	0,5		1	0,5		3	0,10
4	3	0,3		3	0,2		9	0,04
5							-6	0,15
6							6	0,25
7							18	0,10
8							-9	0,09
9							9	0,15
10							27	0,06
11							Сума	1,00
12								

Рис. 2.6. Скриншот таблиці розрахунків для побудови закону розподілу випадкової величини $\eta = 3\xi_1 \cdot \xi_2$

Остаточно закон розподілу має такий вигляд (табл. 2.4).

Таблиця 2.4

Закон розподілу випадкової величини $\eta = 3\xi_1 \cdot \xi_2$

$\eta = 3\xi_1 \cdot \xi_2$	-9	-6	-3	3	6	9	18	27
p_i	0,09	0,15	0,06	0,10	0,25	0,19	0,10	0,09

Обчислимо основні числові характеристики випадкової величини $\eta = 3\xi_1 \cdot \xi_2$ за їх означеннями та порівняємо ці значення з тими, що обчислюються за властивостями основних числових характеристик (рис. 2.7).

Для обчислення математичного сподівання випадкової величини $\eta = 3\xi_1 \cdot \xi_2$ за його означенням введемо, наприклад, у комірку **K2** таку формулу: **=СУММПРОИЗВ(G2:G10;H2:H10)**.

Після натискання **Enter** виводиться результат $M(\eta) = 5,04$. Тепер обчислимо дисперсію цієї випадкової величини, ввівши у комірку **K4** таку формулу:

$$\text{=СУММПРОИЗВ(G2:G10;G2:G10;H2:H10)-K2^2.}$$

	A	B	C	D	E	F	G	H	I	J	K
1	x_i	p_i		y_i	p_i		$3x_i \cdot y_i$	p_i			
2	1	0,2		-1	0,3		-3	0,06		$M(\eta) = 5,04$	
3	2	0,5		1	0,5		3	0,10			
4	3	0,3		3	0,2		9	0,04		$D(\eta) = 89,2584$	
5							-6	0,15			
6	$M(\xi_1) = 2,1$			$M(\xi_2) = 0,8$			6	0,25		$\sigma(\eta) = 9,45$	
7							18	0,10			
8	$D(\xi_1) = 0,49$			$D(\xi_2) = 1,96$			-9	0,09		$M(3 \cdot \xi_1 \cdot \xi_2) = 5,04$	
9							9	0,15			
10	$\sigma(\xi_1) = 0,7$			$\sigma(\xi_2) = 1,4$			27	0,06		$D(3 \cdot \xi_1 \cdot \xi_2) = 89,2584$	
11							Сума	1,00			
12											

Рис. 2.7. Скриншот таблиці розрахунків для побудови закону розподілу випадкової величини $\eta = 3\xi_1 \cdot \xi_2$

Натискаємо **Enter** і отримуємо, що $D(\eta) = 89,2584$.

Порівняємо ці дані з результатами, які можна отримати, виходячи з властивостей основних числових характеристик. Нагадаємо, що основні числові характеристики незалежних випадкових величин ξ_1 і ξ_2 ми обчислювали у першій частині лабораторної роботи (див. рис. 2.3).

Скористаємося цими результатами.

Так, для математичного сподівання (його значення виводиться у комірці **K2** за формулою $=3 \cdot B6 \cdot E6$) маємо:

$$M(3 \cdot \xi_1 \cdot \xi_2) = 3 \cdot M(\xi_1) \cdot M(\xi_2) = 3 \cdot 2,1 \cdot 0,8 = 5,04.$$

Отже, ми отримали той самий результат, що і при обчисленні безпосередньо за означенням.

Для обчислення дисперсії за властивостями введемо у комірку **K10** формулу: $=3^2 \cdot (B8 \cdot E8 + B6^2 \cdot E8 + E6^2 \cdot D8)$.

Отримуємо результат: $D(3 \cdot \xi_1 \cdot \xi_2) = 89,2584$.

Дійсно,

$$D(3 \cdot \xi_1 \cdot \xi_2) = 9 \cdot (0,49 \cdot 1,96 + 2,1^2 \cdot 1,96 + 0,8^2 \cdot 0,49) = 89,2584.$$

2.5. Завдання для самостійної роботи

У середовищі *MS Excel* за законами розподілів незалежних випадкових величин ξ_1 і ξ_2 , які задані в табл. 2.5, необхідно:

побудувати закон розподілу випадкової величини $\eta = 3 \cdot \xi_1 - \xi_2$;

обчислити математичне сподівання та дисперсію випадкової величини $\eta = 3 \cdot \xi_1 - \xi_2$ за означенням цих числових характеристик та їх властивостями;

знайти закон розподілу випадкової величини $\eta = 2\xi_1 \cdot \xi_2$;

визначити математичне сподівання та дисперсію випадкової величини $\eta = 2\xi_1 \cdot \xi_2$ за означенням цих числових характеристик та їх властивостями.

Таблиця 2.5

Закони розподілу випадкових величин ξ_1 та ξ_2

Варіант 1				Варіант 2			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
1	0,2	2	0,3	1	0,8	0	0,5
4	0,4	4	0,3	4	0,1	1	0,2
5	0,3	8	0,4	6	0,1	3	0,2
7	0,1					5	0,1
Варіант 3				Варіант 4			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-2	0,3	-3	0,2	-2	0,1	-1	0,4
3	0,4	0	0,1	3	0,4	0	0,2
4	0,3	1	0,5	5	0,3	2	0,1
		2	0,2	6	0,2	4	0,3
Варіант 5				Варіант 6			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
3	0,4	1	0,3	-2	0,2	0	0,6
4	0,1	2	0,1	3	0,4	2	0,1
5	0,2	3	0,6	4	0,3	4	0,3
6	0,3			5	0,1		

Варіант 7				Варіант 8			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-4	0,1	-3	0,2	-2	0,3	-4	0,1
-2	0,4	0	0,1	2	0,1	-2	0,2
1	0,2	1	0,3	3	0,4	1	0,3
3	0,3	2	0,4	5	0,2	2	0,4
Варіант 9				Варіант 10			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-2	0,3	-3	0,3	3	0,3	-1	0,3
0	0,2	0	0,6	5	0,3	2	0,1
2	0,4	2	0,1	7	0,4	5	0,4
4	0,1					7	0,2
Варіант 11				Варіант 12			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-1	0,1	-4	0,2	-4	0,1	-3	0,4
2	0,2	2	0,7	0	0,2	0	0,1
3	0,2	3	0,1	1	0,6	1	0,3
5	0,5			3	0,1	2	0,2
Варіант 13				Варіант 14			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-2	0,2	2	0,3	1	0,2	2	0,2
1	0,5	3	0,1	2	0,7	4	0,1
2	0,3	4	0,4	3	0,1	6	0,3
		5	0,2			8	0,4
Варіант 15				Варіант 16			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-2	0,3	-4	0,1	-2	0,3	-1	0,3
0	0,2	3	0,2	0	0,2	0	0,5
1	0,4	1	0,3	3	0,1	2	0,2
3	0,1	5	0,4	5	0,4		

Варіант 17				Варіант 18			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
2	0,5	-1	0,3	2	0,3	-2	0,3
4	0,1	2	0,2	4	0,6	0	0,4
8	0,4	5	0,2	6	0,1	2	0,2
		7	0,3			4	0,1
Варіант 19				Варіант 20			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-3	0,4	-2	0,1	-4	0,2	-3	0,2
-1	0,2	0	0,5	-2	0,3	2	0,6
1	0,1	1	0,2	2	0,3	4	0,2
2	0,3	3	0,2	4	0,2		
Варіант 21				Варіант 22			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-1	0,3	2	0,2	2	0,6	1	0,2
1	0,3	3	0,5	7	0,3	3	0,4
2	0,2	4	0,3	6	0,1	4	0,3
4	0,2					6	0,1
Варіант 23				Варіант 24			
$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j	$\xi_1 = x_i$	p_i	$\xi_2 = y_j$	p_j
-2	0,5	-3	0,4	-4	0,3	1	0,2
2	0,3	1	0,2	-1	0,3	4	0,5
3	0,1	4	0,3	0	0,1	6	0,3
5	0,1	6	0,1	2	0,3		

2.6. Контрольні запитання

1. Дайте означення випадкової величини, наведіть приклади.
2. Які типи випадкових величин ви знаєте?
3. У чому полягає різниця між дискретною та неперервною випадковими величинами? Наведіть приклади.
4. Як можна задати дискретну випадкову величину?

5. Що таке ряд розподілу? Як він формується? Чому дорівнює сума ймовірностей у законі розподілу?
6. Що називають сумою та різницею двох випадкових величин?
7. Які випадкові величини є незалежними? Наведіть приклади.
8. Що називають добутком двох незалежних випадкових величин?
9. Як обчислюються значення добутку незалежних випадкових величин і відповідні їм ймовірності?
10. Назвіть основні числові характеристики дискретної випадкової величини.
11. Дайте означення математичного сподівання. За якою формулою його можна обчислити? Наведіть властивості математичного сподівання випадкової величини.
12. Дайте означення дисперсії випадкової величини. Які формули для обчислення дисперсії ви знаєте? Наведіть властивості дисперсії, виходячи з її означення.
13. Чому дорівнюють математичне сподівання і дисперсія сталої величини?
14. Що таке середнє квадратичне відхилення і за якою формулою воно обчислюється?
15. Навіщо одночасно використовують і дисперсію, і середнє квадратичне відхилення випадкової величини?
16. Чи однаковим буде закон розподілу дискретних випадкових величин: $\xi + \xi$ і $2 \cdot \xi$? Для обґрунтування відповіді побудуйте закони розподілу цих випадкових величин на конкретному прикладі.
17. Порівняйте значення $M(\xi + \xi)$ та $M(2 \cdot \xi)$, спираючись на властивості математичного сподівання.
18. Доведіть, що $D(\xi + \xi) \neq D(2 \cdot \xi)$.

Лабораторна робота 3

Дискретна випадкова величина: біноміальний закон розподілу та його асимптотичні наближення

3.1. Мета роботи:

ознайомлення з загальною задачею про повторені однорідні незалежні випробування;

дослідження біноміального закону розподілу залежно від його параметрів та визначення основних числових характеристик розподілу;
порівняння результатів обчислення ймовірностей за формулою Бернуллі й за асимптотичними формулами Муавра – Лапласа та Пуассона.

3.2. Теоретичні положення

Закон розподілу ймовірностей значень дискретної випадкової величини можна задавати не тільки у вигляді таблиці (ряд розподілу) або графічно (многокутник розподілу), але й аналітично, якщо відома функція, за якою обчислюється ймовірність кожного з можливих значень випадкової величини.

Нехай поява події у даному випробуванні не залежить від того, скільки разів ця подія з'являлась у попередніх випробуваннях, тобто наслідки випробувань є взаємно **незалежними**, і ймовірність появи випадкової події S для кожного випробування є сталою величиною $P(S) = p$, тобто випробування є **однорідними**. Схема незалежних і однорідних випробувань називається **схемою Бернуллі**.

Ймовірність того, що подія S з'явиться саме m разів у серії з n випробувань, визначається за **формулою Бернуллі**:

$$P_n(m) = C_n^m p^m q^{n-m} \quad \text{при } 0 \leq m \leq n, \quad (3.1)$$

де C_n^m – кількість сполучень із n по m ;

p – ймовірність появи події в окремому випробуванні;

q – ймовірність того, що в окремому випробуванні подія не з'явиться, тобто $q = 1 - p$.

Закон розподілу дискретної випадкової величини ξ , яка набуває своїх значень $0; 1; 2; 3; \dots; n$ з імовірністю, що визначається за формулою Бернуллі, називається **біноміальним**.

Основними числовими характеристиками розподілу випадкової величини є її математичне сподівання, яке для випадкової величини, що розподілена за біноміальним законом, дорівнює:

$$M(\xi) = np, \quad (3.2)$$

та її дисперсія, яка дорівнює: $D(\xi) = npq$. (3.3)

Відповідно, середнє квадратичне відхилення дорівнює:

$$\sigma(\xi) = \sqrt{npq}. \quad (3.4)$$

Найбільш імовірніша кількість появи події у серії із n випробувань, тобто мода M_0 , для випадкової величини, що розподілена за біноміальним законом, обчислюється за формулою:

$$np - q \leq M_0 \leq np + p \quad \text{при} \quad M_0 \in N \cup \{0\}. \quad (3.5)$$

Розподіл випадкової величини будь-якої природи можна задати за допомогою функції розподілу, яка визначає ймовірність того, що випадкова величина набуде значення, яке буде меншим за деяке значення x . Отже, за означенням:

$$F(x) = P(\xi < x). \quad (3.6)$$

Звідси ймовірність того, що значення, якого набуде випадкова величина, належатиме напіввідкритому інтервалу $[m_1; m_2)$, визначається як різниця значень функції розподілу, що відповідають межах інтервалу:

$$P_n(m_1 \leq m < m_2) = F(x = m_2) - F(x = m_1). \quad (3.7)$$

Для випадкової величини, що розподілена за біноміальним законом, маємо:

$$P_n(m_1 \leq m < m_2) = \sum_{m=m_1}^{m_2-1} C_n^m p^m q^{n-m}. \quad (3.8)$$

Кількість випробувань, які необхідно провести, щоб із заданою **довірчою ймовірністю** P (**надійністю**) гарантувати, що у цій серії випробувань за схемою Бернуллі подія з'явиться хоча б один раз, визначається співвідношенням:

$$n \geq \frac{\ln(1 - P)}{\ln(1 - p)} \quad \text{при} \quad n \in N. \quad (3.9)$$

Якщо кількість випробувань велика ($n \rightarrow \infty$), то застосування формул (3.1) і (3.8) потребує громіздких розрахунків. У цьому випадку доцільно застосовувати **асимптотичні теореми** теорії ймовірностей. Так, у відповідності з **локальною теоремою Муавра – Лапласа** ймовірність того, що випадкова величина набуде значення $\xi = m$, можна наближено обчислити за співвідношенням:

$$P_n(\xi = m) \approx \frac{1}{\sqrt{npq}} \varphi(x), \quad 0 < p < 1, \quad (3.10)$$

де $\varphi(x)$ – **функція Гаусса**, яка визначається як:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right). \quad (3.11)$$

Аргументом функції Гаусса є **нормована випадкова величина**:

$$x = \frac{m - np}{\sqrt{npq}}. \quad (3.12)$$

Для нормованої випадкової величини математичне сподівання дорівнює нулю, а дисперсія – одиниці. Отже, для випадкової величини, закон розподілу якої описується функцією Гаусса, маємо:

$$M(\zeta) = 0, \quad D(\zeta) = 1. \quad (3.13)$$

Локальна теорема Муавра – Лапласа може застосовуватись при $n > 100$ та $npq > 20$. Точність обчислення зростає зі зростанням кількості випробувань.

За великої кількості випробувань ($n > 100$ та $npq > 20$) для визначення ймовірності того, що значення випадкової величини належатиме напіввідкритому інтервалу $[m_1; m_2)$, відповідно до **інтегральної теореми Муавра – Лапласа** можна наближено обчислити за співвідношенням:

$$P_n(m_1 \leq m < m_2) \approx \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right), \quad (3.14)$$

де $\Phi(x)$ – **функція Лапласа**, що визначається як:

$$\Phi(x) = \int_0^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt. \quad (3.15)$$

Оскільки інтеграл $\int e^{-t^2/2} dt$ не визначається через елементарні функції, то значення функції Лапласа знаходять за спеціальною таблицею.

При тому ж самому значенні n обчислення за формулами (3.10) та (3.14) дають тим більш точні результати, чим точніше виконується співвідношення: $p \approx q$.

Якщо $p < 0,1$, то похибка наближеного обчислення з використанням теорем Муавра – Лапласа стає значною. Отже, якщо випробування здійснюються за схемою Бернуллі, то для наближеного обчислення ймовірностей значень випадкової величини, що розподілена за біноміальним законом, в умовах, коли $n \rightarrow \infty$ та $p \ll 1$, застосовується **формула Пуассона**:

$$P_n(m) \approx \frac{a^m e^{-a}}{m!}, \quad (3.16)$$

де a – параметр, який визначається як $a = np$.

Слід зазначити, що, на відміну від формули Муавра – Лапласа, яка не може застосовуватися для $p = 0$, формула Пуассона має сенс і для неможливих подій.

Закон розподілу дискретної випадкової величини, для якої ймовірність її певного значення визначається асимптотичною формулою (3.16), називається **розподілом Пуассона**, або **законом виключних подій**. Оскільки розподіл Пуассона є окремим випадком біноміального розподілу, то випадкова величина, що розподілена за законом виключних подій ($q \rightarrow 1$), має такі основні числові характеристики:

$$M(\xi) = np = a, \quad D(\xi) \approx np = a. \quad (3.17)$$

Оскільки для розподілу Пуассона можна вважати, що $M(\xi) \approx D(\xi)$, то цей закон розподілу є однопараметричним.

3.3. Змістовна постановка задачі

3.1. Імовірність того, що взяту напрокат автомашину повернуть неушкодженою, дорівнює 0,8. Вважаючи випадковою величиною кількість машин, що повернуть неушкодженими, визначити ймовірність того, що із семи орендованих машин неушкодженими будуть повернуті саме: а) $m = 0, 1, 2, \dots$ машин; б) менше п'яти машин; в) не більше шести й не менше чотирьох машин. Визначити основні числові характеристики випадкової величини. Побудувати многокутник розподілу. Дослідити, як зміниться многокутник розподілу, якщо: а) зменшити ймовірність до 0,7; б) збільшити кількість машин до 8.

3.2. Схожість зерна в середньому становить 60 %. Вважаючи випадковою величиною кількість пророслих зернин, знайти ймовірність того, що з 500 посіяних зернин зійде саме: а) 300; б) не менше 250 і не більше 350 зернин. Знайти обсяг вибірки, за яким із надійністю 95 % у вибірковій сукупності буде хоча б одна проросла зернина.

3.3. Завод виготовляє точні прилади. Партія деталей містить 0,01 % нестандартних. З цієї партії на контроль узяли 200 деталей. Вважаючи випадковою величиною кількість нестандартних деталей, визначити ймовірність того, що серед 200 деталей кількість нестандартних деталей, що виявить контроль, становитиме: а) $m = 0, 1, 2, \dots$ деталей; б) більше, ніж одна деталь.

3.4. Приклад виконання лабораторної роботи 3

Завдання 3.1. У даних випробуваннях випадковою величиною ξ є кількість машин, що повернуті неушкодженими. Оскільки те, буде чи ні ушкоджена певна машина, не залежить від того, чи буде ушкоджена інша машина і для всіх машин імовірність ушкодження вважається однаковою, то випробування слід вважати незалежними й однорідними, отже, реалізується схема Бернуллі.

На робочому аркуші *MS Excel* (рис. 3.1) побудуємо многокутник розподілу випадкової величини ξ , що розподілена за біноміальним розподілом. Параметрами розподілу є $p = 0,8$ та $n = 7$. Для цього запишемо найменування параметрів в одному стовпці (комірки **A4** та **A5**), а їх значення – поруч, в окремому стовпці (комірки **B4** та **B5**).

B7		=ЧИСЛКОМБ(\$B\$5;A7)*\$B\$4^A7*(1-\$B\$4)^(B\$5-A7)									
	A	B	C	D	E	F	G	H	I	J	K
1	Завдання 2.1										
2	Біноміальний закон розподілу										
3											
4		$p= 0,8$					$p= 0,7$			$p= 0,8$	
5		$n= 7$					$n= 7$			$n= 8$	
6	m	$P_n(\xi=m)$		X	$F(x)$		m	$P_n(\xi=m)$		m	$P_n(\xi=m)$
7	0	0,0000128		≤ 0	0		0	0,0002187		0	0,00000256
8	1	0,0003584		(0; 1]	0,0000128		1	0,0035721		1	0,00008192
9	2	0,0043008		(1; 2]	0,0003712		2	0,0250047		2	0,00114688
10	3	0,0286720		(2; 3]	0,0046720		3	0,0972405		3	0,00917504
11	4	0,1146880		(3; 4]	0,0333440		4	0,2268945		4	0,04587520
12	5	0,2752512		(4; 5]	0,1480320		5	0,3176523		5	0,14680064
13	6	0,3670016		(5; 6]	0,4232832		6	0,2470629		6	0,29360128
14	7	0,2097152		(6; 7]	0,7902848		7	0,0823543		7	0,33554432
15				>7	1,0000000					8	0,16777216
16											
17	$\Sigma P_n(\xi=m)=$	1,0000000					$\Sigma P_n(\xi=m)=$	1,0000000		$\Sigma P_n(\xi=m)=$	1,0000000
18											
19	$P(\xi=m < 5)=$	0,1480320		$M(\xi)=$	5,6						
20	$P(4 \leq \xi = m \leq 6)=$	0,7569408		$D(\xi)=$	1,12						
21											

Рис. 3.1. Скриншот, що містить розрахунки біноміального розподілу для різних значень його параметрів

Тепер у вигляді стовпчика (комірки **A7:A14**) запишемо значення, яких може набувати випадкова величина, тобто $m = 0, 1, 2, \dots, 7$. Відповідні їм імовірності обчислюємо за співвідношенням (3.1). Для цього в комірку **B7** вводимо формулу:

$$=ЧИСЛКОМБ(\$B\$5;A7)*\$B\$4^A7*(1-\$B\$4)^(B\$5-A7).$$

Зверніть увагу, що в цій формулі посилання на комірки, що містять параметри розподілу, є **абсолютними**, тобто адреса комірки не змінюється при переміщенні формули у іншу комірку. Для цього символ \$ ставиться перед найменуванням стовпця і номером рядка, що здійснюється натисканням клавіші **F4**.

Інші посилання у формулі є **відносними**, отже, при переміщенні формули у іншу комірку цього ж рядка або стовпчика відповідно змінюється це посилання. Розтягнувши цю формулу вздовж стовпчика комірок **B7:B14**, отримаємо ймовірності, з якими випадкова величина набуває своїх значень. Переконаємось, що ми отримали саме закон розподілу.

Для цього визначаємо суму ймовірностей усіх значень випадкової величини. Вводимо в комірку **B17** формулу: **=СУММ(B7:B17)**. Те, що ця сума дорівнює 1 (значення), означає, що дійсно ми побудували закон розподілу випадкової величини.

Для обчислення ймовірностей влучення значення випадкової величини у певний інтервал побудуємо функцію розподілу (3.6), для чого для дискретної випадкової величини можна скористатись рекурентним співвідношенням: $F(x_{i+1}) = F(x_i) + P_n(x_i)$. Зрозуміло, що при $x \leq 0$ маємо $F(x) = 0$. Це значення записуємо в комірку **E7**. Значення комірки **E8** обчислюємо, ввівши туди формулу: **=E7+B7**. Тепер розтягуємо цю формулу на комірки **E8:E15**. Для комірки **E15** отримуємо значення функції розподілу, що відповідає $x > 7$, як і повинно бути, воно дорівнює 1.

Ймовірність того, що значення, якого набуде випадкова величина ξ , буде меншим за п'ять, дорівнює значенню функції розподілу, аргумент якої відповідає умові: $x \in (4; 5]$. Відповідно, у комірку **B19** записуємо посилання на комірку **E12**.

Ймовірність того, що значення, якого набуде випадкова величина ξ , належатиме певному інтервалу, визначимо за формулою (3.7). Для цього перетворимо умову: $P_7(4 \leq \xi = x \leq 6) = P_7(4 \leq \xi = x < 7)$. Тепер у комірку **B10** вводимо формулу: **=E14-E11**. Також можна скористатись аксіомою про суму ймовірностей несумісних подій.

За формулою (3.8) маємо: $P_7(4 \leq \xi = x \leq 6) = P_7(4) + P_7(5) + P_7(6)$, тобто ми отримали той же результат.

Обчислимо основні числові характеристики випадкової величини безпосередньо за означеннями. Для цього застосуємо вбудовані функції *MS Excel*. Оскільки дані згруповані, то для визначення математичного сподівання за означенням застосовуємо функцію **СУММПРОИЗВ**, що належить до категорії **Математические**. Вводимо її в комірку **E19** і вказуємо аргументи, що записані в комірках **A7:A14** та **B7:B14**.

Отримане значення співпадає з тим, що обчислюється за формулою (3.2) для біноміального закону розподілу (див. рис. 3.1).

Якщо обчислювати дисперсію випадкової величини як різницю між математичним сподіванням квадрату випадкової величини і квадратом її математичного сподівання, то для визначення зменшеного можна застосувати функцію **СУММПРОИЗВ**.

Для цього в комірку **E20** вводимо формулу:

$$=СУММПРОИЗВ((A7:A14)^2;B7:B14)-E19^2.$$

Результат співпадає з тим, що можна отримати за формулою (3.3).

Для побудови многокутника розподілу скористаємося надбудовою **Мастер диаграмм**. Виділяємо комірки **B7:B14** і вибираємо шлях **Вставка** \Rightarrow **Точечная**. Діаграма виводиться на екран натисканням відповідної іконки. Сам графік має зараз назву **Ряд 1**. Тепер правою клавішею миші натискаємо на будь-яку точку на діаграмі і в діалоговому вікні **Ряд 1** вибираємо **Выбрать данные**, переходимо в діалогове вікно **Выбор источника данных**. Тепер для ряду 1 вибираємо **Изменить**, переходимо у вікно **Изменение ряда** і заповнюємо його поля (рис. 3.2).

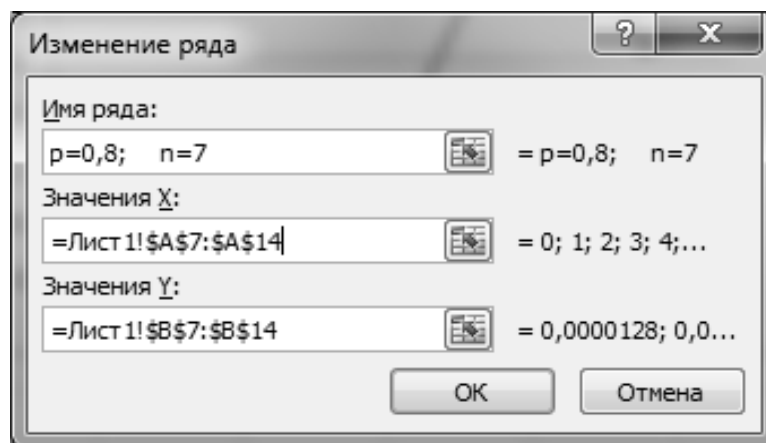


Рис. 3.2. Діалогове вікно "Изменение ряда"

Значения Y вже вказані, вони є результатами обчислення ймовірностей (комірки **B7:B14**). Записуємо **Имя ряда**, вказуємо **Значения X**, яких може набувати випадкова величина (комірки **A7:A14**). Натисканням **ОК** повертаємось до вікна **Выбор источника данных** і там знову натискаємо **ОК**, внаслідок чого на екран виводиться многокутник розподілу.

Аналогічним чином організуємо обчислення побудови біноміального закону розподілу випадкової величини ξ для значення параметрів $p = 0,7$ та $n = 7$, а також $p = 0,8$ та $n = 8$ (див. рис. 3.1). Для того щоб побудувати многокутники, що відповідають цим розподілам, правою клавішею миші натискаємо на будь-яку точку на діаграмі і в діалоговому вікні, яке щойно відкрилось, вибираємо **Выбрать данные**, переходимо в діалогове вікно **Выбор источника данных** і натискаємо **Добавить**. У новому вікні

Изменение ряда вказуємо його ім'я, тобто $p = 0,7$ та $n = 7$. Поле **Значения X** містить посилання на комірки **G7:G14**, а **Значения Y** – на комірки **H7:H14**. Аналогічно додаємо ряд $p = 0,8$ та $n = 8$, для якого поле **Значения X** містить посилання на комірки **J7:J14**, а **Значения Y** – на комірки **K7:K14**. Імена рядів автоматично перенесені в **Легенду** (рис. 3.3).

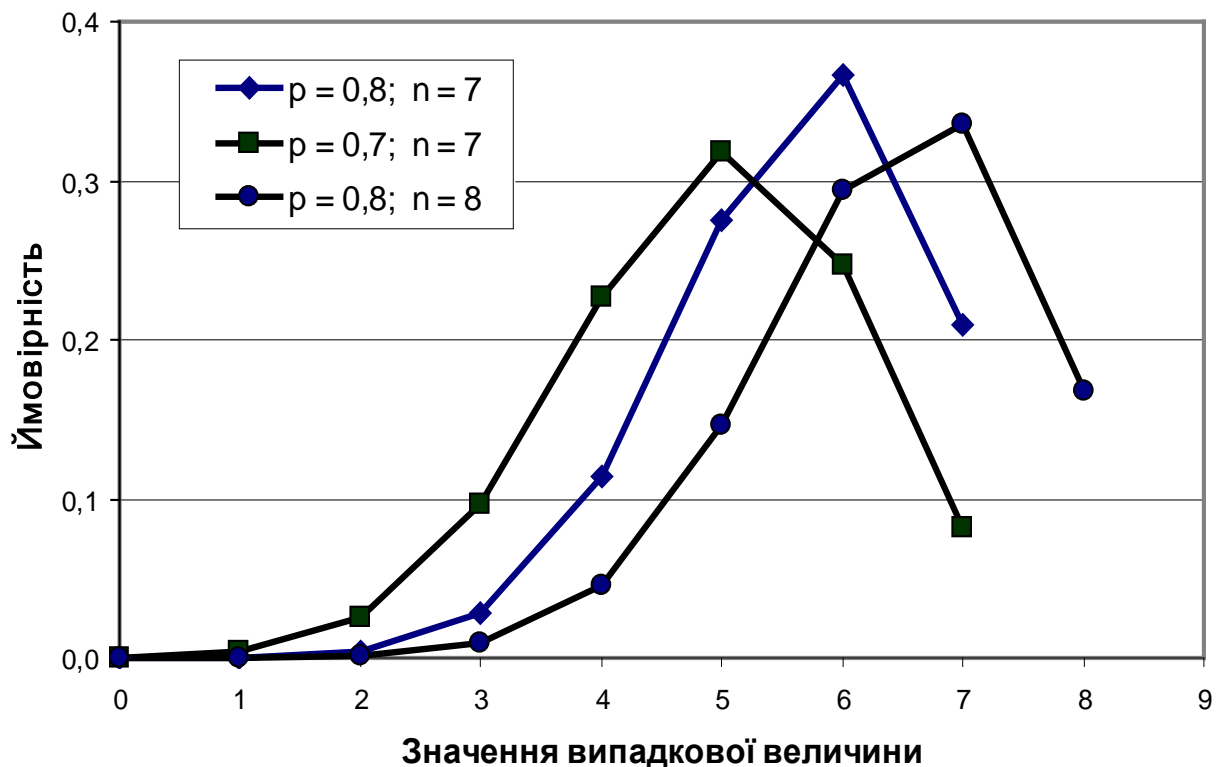


Рис. 3.3. Многокутник біноміального розподілу залежно від значення його параметрів

З аналізу рис. 3.3 видно, що всі розподіли, які досліджувались, є одномодальними.

Значення моди зсунуте в бік більших значень випадкової величини, тобто розподіл асиметричний (негативна асиметрія). Це пов'язано з тим, що для цих умов поява події є більш імовірною, ніж її відсутність. Так, одному й тому ж значенню $n = 7$ при $p = 0,7$ відповідає $M_0 = 5$, а при $p = 0,8$ маємо $M_0 = 6$ і цій моді відповідає більша ймовірність. Зі збільшенням кількості випробувань (за однакової ймовірності появи події в одному випробуванні) спостерігається зсув моди в бік більших значень випадкової величини.

Завдання 3.2. Оскільки реалізується схема Бернуллі, тобто випробування є однорідними й незалежними, то випадкова величина ξ (кількість пророслих зернин у вибірковій сукупності) розподілена за біноміальним законом. Однак кількість випробувань велика ($n = 500$), а ймовірність появи події в окремому випробуванні $p = 0,6$. Отже, для обчислення ймовірностей, з якими випадкова величина набуває певних значень, можна застосовувати асимптотичні теореми Муавра – Лапласа.

За допомогою *MS Excel* визначити ймовірність певного значення випадкової величини можна за допомогою функції **НОРМ.РАСП**, що належить до категорії **Статистические**. Скористаємося формулою (3.10). Для цього спочатку для функції Гаусса обчислимо її аргумент за

формулою (3.12): $x = \frac{300 - 500 \cdot 0,6}{\sqrt{500 \cdot 0,6 \cdot 0,4}} = 0$. Тепер для обчислення значення

функції Гаусса (3.11), що відповідає цьому значенню аргументу, викликаємо функцію **НОРМ.РАСП**. У полях **Среднее** і **Стандартное_откл** її діалогового вікна вказуємо параметри функції Гаусса за співвідношенням (3.13), а у полі **Интегральная** ставимо 0 (**ложь**), й отримуємо значення функції Гаусса (рис. 3.4).

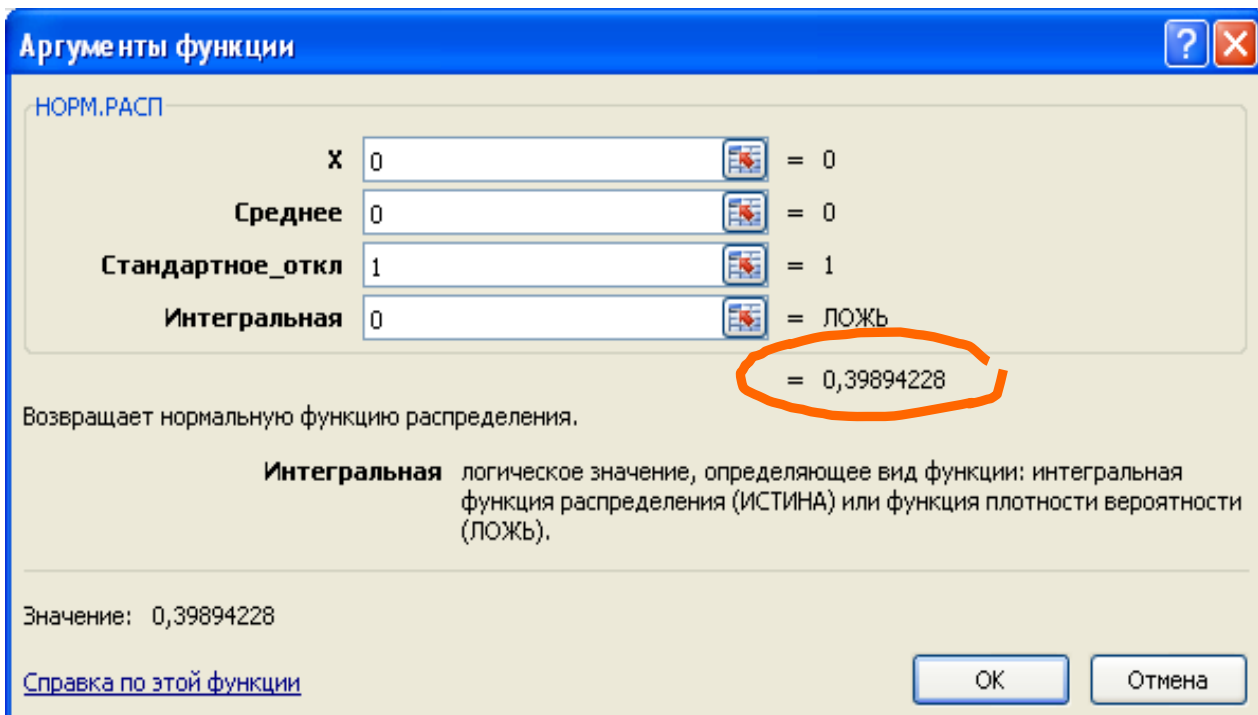


Рис. 3.4. Заповнення полів діалогового вікна **НОРМ.РАСП** при визначенні функції Гаусса

Для обчислення ймовірності того, що з 500 посіяних зернин зійде саме 300 зернин, отримане значення функції Гаусса треба поділити на середнє квадратичне відхилення випадкової величини ξ від її математичного сподівання, тобто на величину $\sqrt{500 \cdot 0,6 \cdot 0,4}$. Відповідно, отримуємо результат: $P_{500}(\xi = 300) \approx 0,0364$.

Обчислення можна скоротити, якщо застосувати ту ж саму вбудовану функцію **НОРМ.РАСП**, вказавши у її діалоговому вікні значення **X**, яке відповідає значенню випадкової величини $m = 300$ і такі параметри біноміального розподілу, як математичне сподівання у полі вводу **Среднее** та середнє квадратичне відхилення у полі вводу **Стандартное_откл**, а у полі **Интегральная** ставимо 0 (рис. 3.5). Як бачимо, ми безпосередньо отримали значення ймовірності.

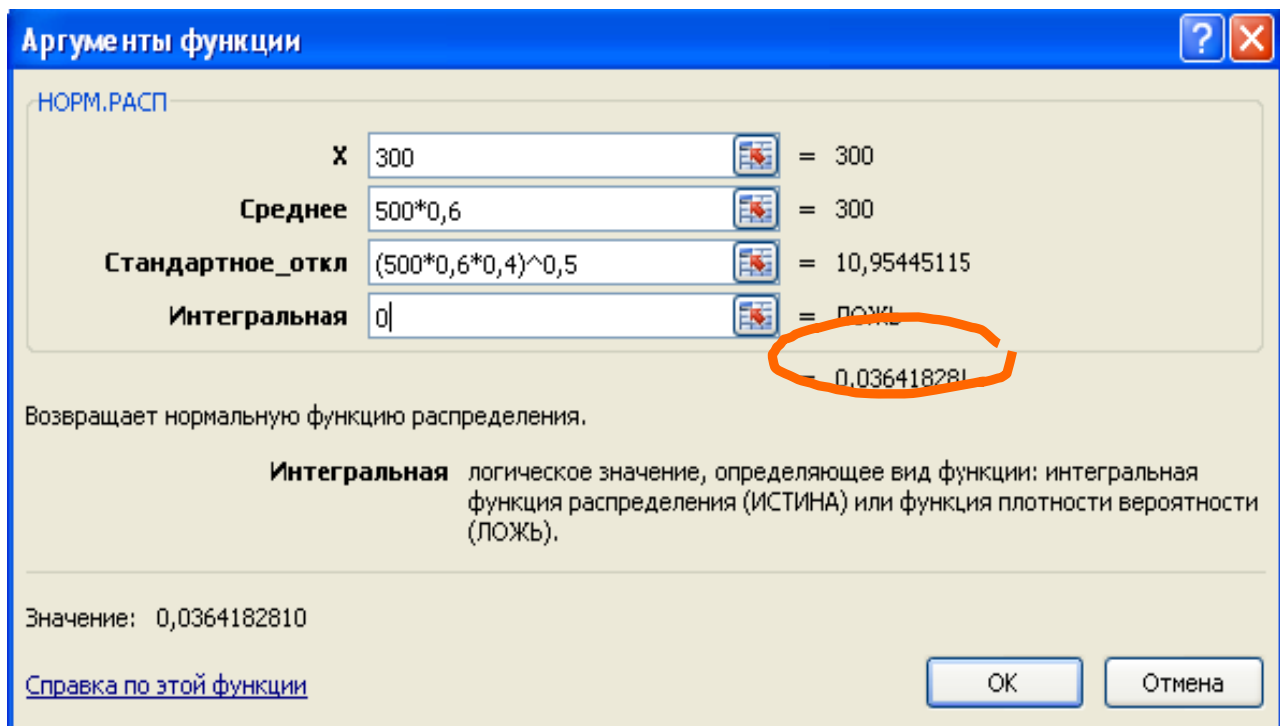


Рис. 3.5. Заповнення полів діалогового вікна **НОРМ.РАСП** при визначенні функції щільності ймовірностей

Ймовірності того, що з 500 зернин зійде не менше 250 і не більше 350, обчислюємо за інтегральною теоремою Муавра – Лапласа (3.14). Для цього спочатку обчислюємо аргументи функції Лапласа, що відповідають значенням випадкової величини $\xi = 250$ та $\xi = 350$.

Потім викликаємо функцію **НОРМ.РАСП** і отримані значення x_1 та x_2 вказуємо у полі **Х**. У полі **Среднее** вказуємо 0, у полі **Стандартное_откл** пишемо 1 (як і на рис. 3.4), а в полі **Интегральная** вказуємо 1, підтверджуючи тим самим, що ми визначаємо саме інтегральну функцію. Імовірність влучення в інтервал визначається як різниця значень функції Лапласа на кінцях інтервалу.

Отже, отримуємо: $P_{500}(250 \leq m \leq 350) \approx 0,999975$.

Тобто ймовірність того, що значення випадкової величини не влучить у вказаний інтервал, є подією малої ймовірності.

За співвідношенням (3.9) визначаємо обсяг вибіркової сукупності, за яким із надійністю 95 % у вибірковій сукупності буде хоча б одна непроросла зернина. Для цього в довільну комірку аркуша MS Excel записуємо формулу: **=ln(1-0,95)/ln(1-0,6)**. Натискаємо **Enter** і отримуємо, що $n \geq 3,27$. З урахуванням того, що $n \in \mathbb{N}$, обсяг вибірки повинен бути не меншим, ніж 4. Тоді вибірка сукупність міститиме хоча б одну пророслу зернину.

Завдання 3.3. Оскільки реалізується схема Бернуллі, тобто випробування є однорідними й незалежними, то випадкова величина ξ (кількість стандартних деталей у вибірковій сукупності) розподілена за біноміальним законом. Однак кількість випробувань велика ($n = 200$), а ймовірність появи події в окремому випробуванні мала ($p = 0,001$), відповідно, для обчислення ймовірностей, з якими випадкова величина набуває своїх значень, застосуємо асимптотичну формулу Пуассона (3.16).

Приклад того, як можна на робочому аркуші MS Excel організувати обчислення, наведено на рис. 3.6.

Для наближеного обчислення ймовірності формула (3.16) безпосередньо записувалась у комірку, як показано в рядку формул на рис. 3.6. У стовпчику **Е** обчислювались значення функції розподілу. Хоча випадкова величина ξ може набувати значень $m = 0, 1, 2, \dots, 200$, найбільше значення $\xi = m$, для якого обчислювалась ймовірність у даному прикладі, становило $m = 4$, оскільки функція розподілу вже досягла значення $F(x) \approx 0,9999$. Відповідно, $P_{200}(m > 4) \approx 0,0001$, отже, випадкова подія, яка полягає в тому, що значення випадкової величини належатиме відріzkу $[5; 200]$, є малої ймовірності.

H8		fx = \$H\$4*G8/EXP(\$H\$4)/ФАКТР(G8)								
	A	B	C	D	E	F	G	H	I	
1	Завдання 2.3									
2	Асимптотична формула Пуассона									
3										
4	p= 0,001								a= 0,2	
5	n= 200									
6	За формулою Бернуллі					За формулою Пуассона				
7	m	$P_n(\xi=m)$		X	F(x)		m	$P_n(\xi=m)$		
8	0	0,8186488		≤0	0,0000000		0	0,8187308		
9	1	0,1638937		(0; 1]	0,8186488		1	0,1637462		
10	2	0,0163237		(1; 2]	0,9825425		2	0,0163746		
11	3	0,0010784		(2; 3]	0,9988662		3	0,0010916		
12	4	0,0000532		(3; 4]	0,9999447		4	0,0000546		
13										
14				M(ξ)= 0,2		P(ξ=m>1)= 0,0175231				
15				D(ξ)= 0,1998						
16										

Рис. 3.6. Обчислення ймовірностей для закону виключних подій

Визначимо ймовірність того, що випадкова величина набуде значення, яке більше за 1, через ймовірність протилежної події:

$$P_{200}(\xi = m > 1) = 1 - (P_{200}(m = 0) + P_{200}(m = 1)) \approx 0,0175.$$

Слід зазначити, що для умов задачі різниця між точним значенням ймовірності, яке обчислювалось за формулою Бернуллі (комірки **B8:B14**), і наближеним значенням, що обчислювалось за формулою Пуассона (комірки **H8:H14**) відчувається лише починаючи з четвертого знака після десяткової коми (див. рис. 3.6).

3.5. Завдання для самостійної роботи

Варіант 3.1

1. В університеті 20 % студентів є відмінниками. Для перевірки знань відібрали 8 студентів. Визначити ймовірність того, що серед них буде: а) саме $m = 0, 1, 2, \dots$ відмінників; б) не менше п'яти, але не більше семи відмінників. Скільки студентів треба відібрати для перевірки, щоб найімовірніше число відмінників дорівнювало шести? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики

випадкової величини і многокутник розподілу, якщо: а) збільшити ймовірність до 0,3; б) зменшити обсяг вибірки до 7.

2. Ймовірність порушення стандарту в процесі виготовлення певної продукції дорівнює 0,3. Визначити ймовірність того, що з 400 виробів кількість непридатних: а) дорівнює 240; б) не менше 210; в) не менше 210 і не більше 270. Визначити кількість випробувань, які необхідно провести, щоб із упевненістю 95 % можна було стверджувати, що подія відбулась хоча б один раз.

3. Середня кількість заявок, що надходять на підприємство побутового обслуговування за 1 годину, дорівнює чотирьом. Знайти ймовірність того, що за годину надійде $m = 0, 1, 2, \dots$ заявок. Визначити ймовірність того, що за годину надійде: а) більше трьох заявок; б) менше чотирьох заявок; в) не більше шести й не менше двох заявок.

Варіант 3.2

1. Ймовірність того, що абонент правильно набере номер телефону, в середньому дорівнює 0,8. Знайти ймовірність того, що з восьми викликів абоненти правильно наберуть саме $m = 0, 1, 2, \dots$ номерів. Знайти ймовірність того, що абоненти наберуть правильно: а) не менше п'яти номерів; б) не більше п'яти номерів; в) не менше трьох і не більше шести номерів. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу в разі: а) зменшення ймовірності появи події до 0,7; б) збільшення кількості випробувань до 9.

2. Ймовірність появи події в окремому випробуванні дорівнює 0,8. Проводиться 200 незалежних випробувань. Знайти ймовірність того, що подія з'явиться: а) 160 разів; б) не менше 150 разів; в) не менше 150 і не більше 170 разів. Визначити кількість випробувань, які необхідно провести, щоб із упевненістю 90 % можна було стверджувати, що подія відбулась хоча б один раз.

3. Ймовірність того, що витрати електроенергії протягом однієї доби не перевищуватимуть встановленої норми, дорівнює 0,05. Знайти ймовірність того, що витрати електроенергії не перевищать норми протягом такої кількості діб: а) $m = 0, 1, 2, \dots$; б) протягом не більше чотирьох діб; в) протягом менше п'яти діб, але більше трьох.

Варіант 3.3

1. Імовірність виграшу в новорічній лотереї дорівнює 0,3. Знайти ймовірність того, що з десяти білетів виграш припаде на $m = 0, 1, 2, \dots$ білетів. Визначити ймовірність: а) виграшу не більше, ніж за двома білетами; б) найімовірнішої кількості виграшних білетів; в) виграшу не менше, ніж за одним і не більше, ніж за трьома білетами. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) збільшити ймовірність до 0,6; б) зменшити обсяг вибірки до 9.

2. Завод випускає електричні лампочки, серед яких 75 % становить продукція першого ґатунку. У вибірку навмання взяли 200 лампочок. Обчислити ймовірність того, що кількість лампочок першого ґатунку у вибірковій сукупності буде: а) не менше, ніж 180; б) не більше, ніж 120. Визначити, яким повинен бути обсяг вибіркової сукупності для того, щоб із надійністю 95 % можна було стверджувати, що в даній вибірковій сукупності є хоча б одна нестандартна лампочка.

3. Через телефонну станцію протягом години проходить в середньому 180 викликів. Знайти ймовірність того, що протягом хвилини буде саме $m = 0, 1, 2, \dots$ викликів. Визначити ймовірність того, що кожну хвилину виникає: а) не менше трьох викликів; б) не менше двох і не більше шести викликів.

Варіант 3.4

1. Завод випускає 80 % виробів першого ґатунку. Знайти ймовірність того, що серед десяти виробів, які були взяті навмання, першого ґатунку буде: а) саме $m = 0, 1, 2, \dots$ виробів; б) хоча б один виріб; в) не менше восьми виробів. Визначити найімовірніше число виробів першого ґатунку і його ймовірність. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,7; б) зменшити обсяг вибірки до 9.

2. Частина об'єктів, що має певну ознаку, в генеральній сукупності дорівнює 75 %. До вибірки взято 400 об'єктів. Обчислити найімовірнішу кількість об'єктів, для яких притаманна ця ознака, та ймовірність цієї кількості. Визначити ймовірність того, що у вибірковій сукупності кількість об'єктів із даною ознакою становить: а) менше, ніж 280; б) не більше, ніж 320; в) не менше, ніж 280 і не більше, ніж 320 об'єктів.

3. Середня кількість літаків, що прибувають в аеропорт за 1 хв, дорівнює 2. Знайти ймовірність того, що кількість літаків, які прибули протягом 2 хв, дорівнює $m = 0, 1, 2, \dots$. Визначити ймовірність того, що протягом 2 хв прибуде: а) менше 4-х літаків; б) більше 5-и літаків; в) хоча б два літаки.

Варіант 3.5

1. Магазин отримав партію електроламп, з яких 40 % є стоватними. Визначити ймовірність того, що серед 12 ламп, які вибрали навмання, кількість шестидесятиватних буде: а) саме $m = 0, 1, 2, \dots$; б) більше шести; в) не менше чотирьох; г) не більше восьми й не менше трьох. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) збільшити ймовірність до 0,6; б) зменшити обсяг вибірки до 9.

2. Відділ технічного контролю перевіряє 1 000 деталей на відповідність стандарту. Імовірність того, що деталь виявиться стандартною, дорівнює 0,75. Визначити найімовірнішу кількість стандартних серед деталей, що проходять перевірку, та ймовірність найімовірнішого числа. Знайти ймовірність того, що кількість стандартних деталей буде: а) не меншою ніж 750 деталей; б) не більшою ніж 900 деталей. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 90 % можна було стверджувати, що в даній вибірці є хоча б одна нестандартна деталь.

3. Імовірність того, що виріб не витримає випробування, дорівнює 0,001. Знайти ймовірність того, що серед 5 000 виробів не витримають випробування: а) саме $m = 0, 1, 2, \dots$; б) більше, ніж один виріб; в) менше чотирьох виробів.

Варіант 3.6

1. В університеті 60 % усіх студентів займаються в спортивних секціях. Для перевірки відібрали 10 студентів. Визначити ймовірність того, що серед них у секціях займаються: а) $m = 0, 1, 2, \dots$ студентів; б) не менше 3-х і не більше 8-и студентів; в) більше 4-х студентів. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,4; б) збільшити обсяг вибірки до 11.

2. У генеральній сукупності частка об'єктів, що мають пролонгований термін придатності, дорівнює 0,7. У вибірку взято 300 об'єктів.

Визначити найімовірніше число об'єктів, що мають цю ознаку, та ймовірність найімовірнішого числа. Визначити ймовірність, з якою можна стверджувати, що у вибірковій сукупності кількість об'єктів із пролонгованим терміном придатності становить: а) менше, ніж 200; б) не менше ніж 200 і не більше ніж 250. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 95 % можна було стверджувати, що в даній вибірці є хоча б один об'єкт, якому не притаманна дана ознака.

3. Середня кількість кораблів, що заходять у порт протягом години, дорівнює трьом. Знайти ймовірність того, що протягом години в порт зайде: а) саме $m = 0, 1, 2, \dots$ кораблів; б) більше п'яти кораблів; в) менше чотирьох кораблів; г) не більше шести й не менше двох кораблів.

Варіант 3.7

1. Оптова база обслуговує 10 магазинів, від кожного з яких може надійти протягом дня замовлення на товар з імовірністю 0,4. Визначити ймовірність того, що протягом дня на базу надійде: а) $m = 0, 1, 2, \dots$ замовлень; б) більше чотирьох замовлень; в) не менше двох і не більше чотирьох замовлень. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,3; б) збільшити обсяг вибірки до 11.

2. У великій партії деталей 65 % відповідає вищому ґатунку. З цієї партії було взято 200 деталей. Визначити найімовірнішу кількість деталей вищого ґатунку і ймовірність цієї найімовірнішої кількості. Обчислити ймовірність, з якою кількість деталей вищого ґатунку в даній вибірці буде: а) меншою від 140; б) не меншою від 140 і не більшою від 160. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 90 % можна було стверджувати, що в даній вибірці є хоча б одна неякісна деталь.

3. Принцип роботи лічильника Гейгера передбачає, що прилад реєструє частки, що вилітають з джерела радіоактивності, з імовірністю 0,0001. Якщо за час спостереження із джерела вилетіло 30 000 часток, визначити ймовірність того, що лічильник зареєструє: а) $m = 0, 1, 2, \dots$ часток; б) більше трьох часток; в) менше п'яти часток; г) не більше семи часток і не менше трьох часток.

Варіант 3.8

1. Автобаза має 8 машин, імовірність виходу на лінію для кожної з яких дорівнює 0,8. Знайти ймовірність виходу на лінію найближчого дня

$m = 0, 1, 2, \dots$ машин. Визначити, яка ймовірність нормальної роботи автобаси в найближчий день, якщо для цього необхідно мати на лінії не менше 6 автомашин. Обчислити ймовірність виходу на лінію: а) не більше 5 автомашин; б) не менше 5 і не більше 7 автомашин. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,7; б) збільшити кількість машин до 9.

2. Кількість робітників, що перевиконують норму на виробництві, дорівнює 20 %. У вибірку взято 200 робітників. Визначити ймовірність, з якою можна стверджувати, що перевиконують норму: а) не менше 50 робітників; б) не менше 40 і не більше 100 робітників. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 95 % можна було стверджувати, що в даній вибірці є хоча б один робітник, який не перевиконує норми.

3. Середня кількість літаків, які прибувають в аеропорт щохвилини, дорівнює 2. Знайти ймовірність того, що протягом 2 хв в аеропорт прибудуть: а) $m = 0, 1, 2, \dots$ літаків; б) менше чотирьох літаків; в) не більше п'яти літаків; г) хоча б два літаки.

Варіант 3.9

1. Схожість насіння становить 80 %. Знайти ймовірність того, що з 12 посіяних насінин зійде: а) $m = 0, 1, 2, \dots$ насінин; б) не менше трьох насінин; в) не більше п'яти й не менше трьох насінин; г) менше семи насінин. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,7; б) зменшити обсяг вибірки до 11.

2. При застосуванні певного технологічного процесу 80 % від усієї продукції становитиме продукція першого ґатунку. Обчислити ймовірність того, що в партії з 300 виробів кількість продукції першого ґатунку становитиме: а) 250 виробів; б) не менше 230 і не більше 250 виробів. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 90 % можна було стверджувати, що в даній вибірці є хоча б один неякісний виріб.

3. Середня кількість заявок, що надходять на підприємство побутового обслуговування протягом години, дорівнює п'яти. Знайти ймовірність того, що за годину надійде: а) саме $m = 0, 1, 2, \dots$ заявок; б) більше трьох заявок; в) не більше шести й не менше двох заявок.

Варіант 3.10

1. Імовірність того, що телевізор потребуватиме ремонту протягом гарантійного терміну, дорівнює 0,2. Яка ймовірність того, що з 10 телевізорів протягом гарантійного терміну потребуватимуть ремонту: а) саме $m = 0, 1, 2, \dots$ телевізорів; б) хоча б один телевізор; в) менше трьох телевізорів? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу в разі: а) збільшення ймовірності до 0,3; б) зменшення кількості випробувань до 9.

2. Імовірність порушення стандарту в процесі застосування певної технології обробки деталей дорівнює 0,3. Визначити ймовірність того, що з 800 готових деталей кількість непридатних становитиме: а) 240 деталей; б) не менше 210 деталей; в) не менше 210 і не більше 420 деталей. Визначити, яким повинен бути обсяг вибірки, щоб із надійністю 95 % можна було стверджувати, що в даній вибірці є хоча б один неякісний виріб.

3. Імовірність того, що в певному автогосподарстві одна автомашина зазнає аварії протягом місяця, дорівнює 0,01. Автогосподарство має 500 автомашин. Визначити ймовірність того, що протягом місяця зазнають аварії: а) $m = 0, 1, 2, \dots$ автомашин; б) не менше дев'яти автомашин; в) не більше п'яти й не менше двох автомашин.

Варіант 3.11

1. Робітник обслуговує 8 однотипних верстатів. Імовірність того, що протягом тижня один верстат потребуватиме ремонту, дорівнює 0,3. Знайти ймовірність того, що за тиждень уваги робітника потребуватимуть: а) $m = 0, 1, 2, \dots$ верстатів; б) більше трьох, але менше шести верстатів; в) не більше трьох верстатів. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу в разі: а) збільшення ймовірності до 0,4; б) збільшення кількості верстатів до 9.

2. Імовірність зняття симптомів грипу завдяки прийому певних ліків дорівнює 0,6. Ліки приймали 100 хворих на грип. Знайти ймовірність того, що кількість хворих, яким ці ліки допомогли: а) дорівнює 60; б) не менше 50; в) не менше 40 і не більше 70. Визначити, яку кількість хворих треба відібрати, щоб із надійністю 95 % можна було стверджувати, що серед них ліки допомогли хоча б одному хворому.

3. На прядильній фабриці робітниця обслуговує 1 000 веретен. У процесі обертання веретена пряжа рветься у випадкові моменти часу. Вважаючи, що ймовірність обриву пряжі на кожному з веретен протягом деякого проміжку часу дорівнює 0,004, знайти ймовірність того, що одночасно відбудеться: а) $m = 0, 1, 2, \dots$ обривів; б) не більше п'яти обривів. Визначити найімовірніше число одночасних обривів і обчислити його ймовірність.

Варіант 3.12

1. Ймовірність відмови приладу під час випробування дорівнює 0,25. Проведено вісім випробувань. Визначити ймовірність того, що у восьми випробуваннях кількість відмов становитиме: а) $m = 0, 1, 2, \dots$; б) більше шести; в) не більше шести й не менше трьох. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу в разі: а) збільшення ймовірності до 0,4; б) збільшення кількості верстатів до 9.

2. Завод випускає електричні лампочки, серед яких 70 % становлять енергозберіжні лампи. У вибірку взято 200 лампочок. Визначити найімовірнішу кількість енергозберіжних ламп та ймовірність цього найімовірнішого числа. Визначити ймовірність того, що кількість звичайних електричних ламп становитиме: а) не менше, ніж 140; б) не більше, ніж 180; в) не менше, ніж 130 і не більше, ніж 180. Знайти обсяг вибірки, за яким із надійністю 90 % у вибірковій сукупності буде хоча б одна звичайна електрична лампа.

3. Вироби заводу містять 5 % браку. Знайти ймовірність того, що серед 800 узятих навмання виробів бракованих виробів буде: а) саме $m = 0, 1, 2, \dots$; б) менше чотирьох; в) хоча б один. Визначити найбільш ймовірну кількість бракованих виробів і знайти її ймовірність.

Варіант 3.13

1. В університеті з усіх студентів 30 % відмінників. Для перевірки залишкових знань відібрали 15 студентів. Визначити ймовірність того, що серед них кількість відмінників становитиме: а) $m = 0, 1, 2, \dots$; б) не менше п'яти й не більше восьми; в) не менше чотирьох. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу в разі: а) збільшення ймовірності до 0,4; б) зменшення кількості студентів у вибірковій сукупності до 10.

2. Схожість насіння певної рослини становить 80 %. Знайти ймовірність того, що з 900 посіяних насінин зійде: а) 720; б) не менше 700; в) не більше 740; г) не менше 700 і не більше 740. Визначити, якою повинна бути кількість насіння, щоб із надійністю 95 % можна було стверджувати, що зійде хоча б одна зернина.

3. Імовірність виграшу в лотереї дорівнює 0,01. Визначити ймовірність того, що зі ста білетів виграш припаде: а) на $m = 0, 1, 2, \dots$ білетів; не більше ніж на два білети; в) більш ніж на три білети.

Варіант 3.14

1. Імовірність того, що взяту напрокат автомашину повернуть неушкодженою, дорівнює 0,7. Знайти ймовірність того, що з дванадцяти орендованих автомашин неушкодженими будуть повернуті: а) саме $m = 0, 1, 2, \dots$ автомашин; б) не більше десяти й не менше шести автомашин; в) менше семи автомашин. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,6; б) зменшити кількість автомашин до 11.

2. Схожість досліджуваної рослини становить 70 %. Знайти ймовірність того, що з 700 посіяних зернин зійде: а) 490; б) менше 500; в) не менше 390 і не більше 590. Знайти обсяг вибірки, за яким із надійністю 90 % у вибірковій сукупності буде хоча б одна непроросла зернина.

3. 3 партії деталей, серед яких 1 % нестандартних, на контроль відбирається 500 деталей. Визначити ймовірність того, що кількість нестандартних деталей, що виявить контроль, становитиме: а) $m = 0, 1, 2, \dots$ деталей; б) не більше трьох деталей; в) не менше однієї й не більше п'яти деталей. Визначити найбільш ймовірну кількість бракованих деталей у вибірковій сукупності і знайти її ймовірність.

Варіант 3.15

1. У новорічну ніч ймовірність того, що оператор забезпечить мобільний зв'язок між абонентами, дорівнює 0,8. Знайти ймовірність того, що з восьми викликів буде успішно здійснено: а) $m = 0, 1, 2, \dots$ з'єднань; б) менше п'яти з'єднань; в) не менше трьох і не більше шести з'єднань. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність до 0,7; б) збільшити кількість викликів до 9.

2. У великій партії деталей лише 70 % відповідає вищому ґатунку. З партії взято 300 деталей. Обчислити ймовірність того, що кількість деталей вищого ґатунку буде: а) не більше ніж 260; б) більше від 210; в) не менше 240 і не більше 260. Визначити, яким повинен бути обсяг вибірки, щоб із упевненістю 90 % можна було стверджувати, що серед деталей буде хоча б одна, що не відповідає вищому ґатунку.

3. Під час вхідного контролю партії деталей виявилось в середньому 1 % бракованих. Визначити ймовірність того, що серед 300 деталей виявиться: а) рівно $m = 0, 1, 2, \dots$ бракованих; б) не більше чотирьох бракованих; в) не менше двох і не більше п'яти бракованих. Знайти найімовірніше число бракованих деталей та ймовірність цього числа.

Варіант 3.16

1. У цеху є 12 верстатів. Імовірність того, що верстат працює, дорівнює 0,8. Яка ймовірність того, що серед них працюють: а) $m = 0, 1, 2, \dots$ верстатів; б) не менше п'яти й не більше восьми верстатів? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини і многокутник розподілу, якщо: а) зменшити ймовірність з 0,8 до 0,5; б) зменшити кількість верстатів з 12 до 8.

2. Імовірність виготовлення бракованого свердла дорівнює 0,2. Свердла пакуються в коробки по 100 штук. Знайти ймовірність того, що в коробці буде: а) 65 бракованих свердел; б) не більш ніж 75 бракованих свердел; в) не менш ніж 35 бракованих свердел. Визначити обсяг вибірки, для якого з упевненістю 95 % можна було стверджувати, що хоча б одне свердло буде бракованим.

3. Із партії обсягом 500 валків взято 50 з метою контролю їхнього діаметра. Із попередніх досліджень відомо, що в середньому 2 % валків браковані. Знайти ймовірність того, що серед 50 вибраних валків буде $m = 0, 1, 2, \dots$ бракованих. Визначити ймовірність того, що серед 50 вибраних валків бракованих буде: а) більше трьох; б) менше п'яти; в) більше двох, але не більше п'яти.

Варіант 3.17

1. Робітник обслуговує 10 однотипних верстатів. Імовірність того, що верстат потребуватиме обслуговування протягом певного часу t , дорівнює 0,3. Яка ймовірність того, що за час t : а) потрібно буде

обслуговувати $m = 0, 1, 2, \dots$ верстатів; б) кількість вимог на обслуговування буде більше трьох, але менше семи? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини в разі: а) зменшення ймовірності з 0,3 до 0,25; б) збільшення кількості верстатів з 10 до 15.

2. Ймовірність відмови датчика протягом місяця дорівнює 0,1. Щомісяця перевіряють 1 000 датчиків. Визначити ймовірність того, що серед вибраних на перевірку буде: а) 650 датчиків з відмовою; б) не більш ніж 750 датчиків з відмовою; в) не менш ніж 300 датчиків з відмовою. Обчислити необхідний обсяг вибірки, щоб із упевненістю 95 % можна було стверджувати, що серед датчиків хоча б один є з відмовою.

3. У певній місцевості на кожні 100 кавунів припадає в середньому один, маса якого не менша за 10 кг. Знайти ймовірність того, що в партії із 70 кавунів буде $m = 0, 1, 2, \dots$ кавуни масою не менше 10 кг. Визначити ймовірність того, що серед 70 вибраних кавунів кількість кавунів, маса яких не менша за 10 кг, буде: а) більше двох; б) менше двох; в) не більше чотирьох й не менше одного.

Варіант 3.18

1. За один цикл верстат-автомат виготовляє 10 деталей. Ймовірність того, що довільна деталь бракована, становить 0,01. Яка ймовірність того, що серед них виявиться бракованих деталей: а) $m = 0, 1, 2, \dots$; б) не менше двох і не більше семи? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини, якщо: а) збільшити ймовірність з 0,01 до 0,05; б) збільшити кількість деталей з 10 до 18.

2. Телевізійна компанія, вирішуючи питання щодо доцільності придбання права на трансляцію чемпіонату міста з футболу, провела опитування серед фанатів. Визначено, що кожні 20 зі 100 фанатів, які не мають кабельного зв'язку, бажають стати абонентами телевізійної компанії. Знайти ймовірність того, що серед вибраних 200 фанатів абонентами телевізійної компанії стане: а) 105 фанатів; б) не більш ніж 175 фанатів; в) не менш ніж 135 фанатів. Визначити необхідну кількість вболівальників футболу, щоб із упевненістю 90 % можна було стверджувати, що хоча б один фанат стане абонентом телевізійної компанії.

3. Перевізник укладає договір на перевезення цінного вантажу. У процесі транспортування ймовірність пошкодити виріб дорівнює 0,05. Обчислити ймовірність того, що при цьому із 60 виробів буде саме $m = 0, 1, 2, \dots$ пошкоджених. Визначити ймовірність того, що серед 60 вибраних виробів пошкодженими будуть: а) не більше чотирьох; б) менше восьми; в) не менше двох, але не більше семи.

Варіант 3.19

1. Податкова служба визначила, що 50 % всіх особистих декларацій про прибуток містить принаймні одну помилку. Випадковим чином відібрали 10 декларацій. Яка ймовірність того, що серед них з помилками: а) $m = 0, 1, 2, \dots$; б) не менше трьох й не більше восьми? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини та многокутник розподілу, якщо: а) збільшити ймовірність до 60 %; б) збільшити кількість декларацій з 10 до 20.

2. У процесі виробництва детекторів брехні вимагається, щоб вони могли відрізнити правильні відповіді від неправильних з надійністю 85 %. Детектори тестують, використовуючи для цього 50 запитань. Знайти ймовірність того, що правильно визначених відповідей буде: а) 35; б) не більш ніж 45; в) не менш ніж 15. Визначити необхідну кількість запитань, щоб із упевненістю 95 % можна було стверджувати, що хоча б одна хибна відповідь помилково буде визнана правильною.

3. Банк видає кредитні картки VISA. Було встановлено, що 2 % усіх рахунків оплачується повністю за їх допомогою. З попереднього року вибрали навмання 50 рахунків. Яка ймовірність того, що серед них буде $m = 0, 1, 2, \dots$ оплачених за картою VISA? Визначити ймовірність того, що серед 50 відібраних рахунків оплачених цими картками буде: а) більше трьох; б) менше десяти; в) не більше семи й не менше одного рахунку.

Варіант 3.20

1. Митниця дає офіційну оцінку того, що 20 % усіх осіб, що повертаються з-за кордону, не декларує весь товар, який оподатковується. Якщо випадково відібрати 16 осіб, які повертаються з-за кордону, то яка ймовірність того, що: а) $m = 0, 1, 2, \dots$ осіб не задекларує товар, який оподатковується; б) не менше трьох із них не задекларує товар, який оподатковується? Побудувати многокутник розподілу. Дослідити, як змінюються

числові характеристики випадкової величини та многокутник розподілу, якщо: а) зменшити ймовірність до 10 %; б) збільшити кількість відібраних осіб з 16 до 30.

2. Штампування металевих клем для з'єднувальних пластин дає 20 % браку. Визначити ймовірність того, що в партії із 600 клем буде: а) саме 100 клем, що не відповідають стандарту; б) не більше 225 клем, що не відповідають стандарту; в) не менше 350 клем, що відповідають стандарту. Визначити необхідну кількість клем, які необхідно взяти, щоб із упевненістю 95 % можна було стверджувати, що хоча б одна клема відповідає стандарту.

3. Кількість довгих волокон у партії бавовни складає 0,6 % від їх загальної кількості. Для контролю взяли навмання пучок з 500 волокон. Яка ймовірність того, що в обраному пучку волокон буде $m = 0, 1, 2, \dots$ довгих волокон? Визначити ймовірність того, що в обраному пучку кількість довгих волокон буде: а) більше п'яти; б) менше десяти; в) не більше одного.

Варіант 3.21

1. На прикордонній заставі провели статистичні дослідження, які свідчать, що 20 % машин, які прибувають з-за кордону, є легковими. До в'їзду на КПП прибуло 10 машин. Яка ймовірність того, що серед них: а) $m = 0, 1, 2, \dots$ машин будуть легковими; б) не менше двох машин будуть легковими? Побудувати многокутник розподілу кількості легкових машин. Дослідити, як зміняться числові характеристики випадкової величини та многокутник розподілу, якщо: а) збільшити ймовірність до 30 %; б) збільшити кількість машин з 10 до 20.

2. У процесі автоматичного пресування карболітових деталей за певною технологією $2/3$ від загальної кількості деталей не містить браку. Знайти ймовірність того, що із 450 взятих навмання деталей кількість деталей без браку: а) дорівнює 100; б) не перевищує 250; в) знаходиться в межах від 280 до 320. Визначити необхідну кількість деталей, яку треба взяти, щоб із упевненістю 95 % можна було стверджувати, що хоча б одна деталь без браку.

3. У скрині знаходяться 100 карток з номерами від 1 до 100. Навмання 200 разів з цієї скрині виймають одну картку, фіксують її номер і повертають у скриню. Яка ймовірність того, що в цих випробуваннях картка з номером "7" з'явиться $m = 0, 1, 2, \dots$ разів? Визначити

ймовірність того, що в серії з 200 випробувань картку з номером "7" витягли: а) більше семи разів; б) менше п'яти раз; в) не більше десяти й не менше двох раз.

Варіант 3.22

1. У великій партії деталей у середньому з кожних 10 деталей 9 є стандартними. Знайти ймовірність того, що з 15 деталей стандартними будуть: а) $m = 0, 1, 2, \dots$; б) не менше п'яти. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини та многокутник розподілу за умов, що: а) ймовірність того, що деталь виявиться стандартною, зменшити до 0,8; б) кількість деталей у вибірковій сукупності збільшити від 15 до 25.

2. Відсоток схожості насіння складає 80 %. Знайти ймовірність того, що з 800 посіяних зернин зійдуть: а) 560; б) не більш ніж 600; в) більше 640, але менше 750 зернин. Визначити кількість зернин, яку потрібно взяти, щоб із упевненістю 95 % стверджувати, що хоча б одне зерно зійде.

3. Із партії друкарських видань взято 100 видань для проведення перевірки щодо порушення якості друку. Із попередніх досліджень відомо, що в середньому 0,01 % видань містять такі порушення. Яка ймовірність того, що серед 100 вибраних видань саме $m = 0, 1, 2, \dots$ будуть містити порушення? Визначити ймовірність того, що серед 100 відібраних видань з порушеннями буде: а) більше трьох; б) менше п'яти; в) не більше десяти й не менше двох видань.

Варіант 3.23

1. Податкова інспекція визначила, що серед усіх особистих декларацій про прибуток 60 % містять помилку. Визначити ймовірність того, що серед 15 декларацій, які відібрані випадково, без помилок виявиться: а) $m = 0, 1, 2, \dots$; б) не менше двох й не більше десяти декларацій. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини та многокутник розподілу, якщо: а) зменшити ймовірність до 50 %; б) збільшити кількість декларацій з 15 до 25.

2. При застосуванні певної технології виготовлення свердел імовірність браку дорівнює 0,7. Свердла пакуються в коробки по 200 штук. Знайти ймовірність того, що якісних свердел у коробці буде: а) 150; б) не більше 175; в) не менше 105. Визначити, скільки необхідно взяти свердел, щоб

із упевненістю 95 % можна було стверджувати, що серед них хоча б одне свердло буде якісним.

3. Банк видає пільгові кредити приватним клієнтам. Було встановлено, що 1 % усіх кредитів клієнти сплачують до закінчення терміну кредитування. Для аналізу за результатами попереднього року вибрали навмання 100 кредитних угод. Визначити ймовірність того, що серед них буде $m = 0, 1, 2, \dots$ таких, що були сплачені достроково. Яка ймовірність того, що серед 100 кредитних угод приватних клієнтів достроково буде оплачених: а) більше трьох; б) не менш ніж п'ять; в) не більше семи й не менше двох?

Варіант 3.24

1. Робітник обслуговує 14 однотипних верстатів. Імовірність того, що верстат потребуватиме обслуговування протягом певного часу t , дорівнює 0,4. Яка ймовірність того, що за час t : а) обслуговування потребуватимуть $m = 0, 1, 2, \dots$ верстатів; б) кількість вимог на обслуговування буде надходити не менш ніж чотирьох, але не більш ніж десяти верстатів? Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини та многокутник розподілу, якщо: а) збільшити ймовірність з 0,4 до 0,7; б) збільшити загальну кількість верстатів з 14 до 20.

2. Виробник кондиціонерів вимагає, щоб вони могли розрізняти температурний режим приміщень з надійністю 75 %. Для тестування вибрано 450 кондиціонерів. Знайти ймовірність того, що правильно визначать температурний режим приміщень: а) 350; б) не більш як 270; в) не менш ніж 150 кондиціонерів. Визначити, яку кількість кондиціонерів необхідно відібрати у вибірку сукупність, щоб із упевненістю 95 % можна було стверджувати, що хоча б один кондиціонер визначить температурний режим приміщень правильно.

3. При вирощуванні гарбузів із насіння вищого ґатунку в середньому на кожні 200 гарбузів припадає два таких, що маса кожного перевищує 10 кг. Знайти ймовірність того, що в партії із 100 гарбузів буде $m = 0, 1, 2, \dots$ гарбузів, маса яких перевищує 10 кг. Визначити ймовірність того, що серед 100 гарбузів масу завбільшки 10 кг матимуть: а) більше чотирьох гарбузів; б) менше семи гарбузів; в) не більше дев'яти й не менше двох гарбузів.

Варіант 3.25

1. У партії деталей двох подібних форматів кількість деталей, великих за розміром, утричі перевищує кількість менших за розміром. Деталі скидані до купи. Знайти ймовірність того, що серед навмання взятих з цієї купи 12 деталей будуть: а) $m = 0, 1, 2, \dots$ деталей великого розміру; б) не менше двох й не більше семи великих деталей. Побудувати многокутник розподілу. Дослідити, як змінюються числові характеристики випадкової величини та многокутник розподілу, якщо: а) збільшити ймовірність вибору деталі великого формату до 0,8; б) збільшити кількість деталей з 12 до 16.

2. Протягом періоду налаштування нової технології ймовірність браку виробництва складає 15 %. Для детальної діагностики було вибрано 500 деталей, виготовлених за новою технологією. Знайти ймовірність того, що серед цих деталей бракованих буде: а) 350 деталей; б) не більше як 420 деталей; в) не менше ніж 100. Визначити необхідну кількість деталей, щоб із упевненістю 95 % можна було стверджувати, що хоча б одна деталь буде якісною.

3. Із партії друкарських видань взято 200 примірників (по одному примірнику кожного видання) з метою контролю на наявність помилок. Із попередніх досліджень відомо, що в середньому 0,5 % видань містять помилки. Знайти ймовірність того, що серед 200 примірників, відібраних для перевірки з різних видань, з помилками буде $m = 0, 1, 2, \dots$ видань. Визначити ймовірність того, що з 200 примірників, які вибрані для перевірки, помилки будуть містити: а) більше ніж сім примірників; б) менше трьох примірників; в) не більше ніж десять, але не менше ніж два примірники?

3.6. Контрольні запитання

1. Поясніть, яким умовам відповідають випробування, що здійснюються за схемою Бернуллі.

2. Наведіть формулу Бернуллі і поясніть, за яких умов вона може застосовуватись.

3. Чи має обмеження формула Бернуллі за кількістю випробувань?

4. Які числові характеристики розподілу випадкової величини вважаються основними? Наведіть їх означення для дискретної випадкової величини.

5. Який закон розподілу називають біноміальним? Наведіть приклад такого розподілу.

6. Дайте означення функції розподілу. Наведіть приклад графічного надання функції розподілу для дискретної випадкової величини.

7. Наведіть формули для обчислення основних числових характеристик випадкової величини, що розподілена за біноміальним законом.
8. Які додаткові числові характеристики розподілу випадкової величини ви знаєте? Наведіть їх означення.
9. Що таке початковий та центральний моменти k -го порядку?
10. Як визначити додаткові числові характеристики розподілу, користуючись мовою моментів?
11. Що таке многокутник розподілу?
12. Дайте означення поняттям "довірча ймовірність" та "надійність". Як вони пов'язані між собою?
13. Наведіть формулу для визначення кількості випробувань, які необхідно провести, щоб із заданою надійністю забезпечити появу певної випадкової події.
14. Які рівні надійності найчастіше застосовуються для визначення необхідної кількості випробувань, що забезпечує появу певної події? Поясніть, чому саме обирають такі рівні.
15. Поясніть, чому формули Муавра – Лапласа та формула Пуассона називаються асимптотичними.
16. Сформулюйте локальну теорему Муавра – Лапласа. Які існують обмеження для її застосування?
17. Сформулюйте інтегральну теорему Муавра – Лапласа.
18. Порівняйте можливості застосування локальної формули Муавра – Лапласа та формули Пуассона.
19. Наведіть формулу Пуассона.
20. Наведіть пояснення, чому закон, згідно з яким для визначення ймовірності застосовується формула Пуассона, називається законом виключних подій.
21. Поясніть, чому закон розподілу Пуассона вважається однопараметричним.
22. Як впливає параметр розподілу Пуассона на вигляд многокутника його розподілу?

Лабораторна робота 4

Двовимірна дискретна випадкова величина, її закон розподілу та числові характеристики

4.1. Мета роботи:

ознайомлення на прикладі двовимірної дискретної випадкової величини з принципами обчислення основних числових характеристик багатовимірних випадкових величин;

побудова умовних законів розподілу компонентів двовимірної випадкової величини;
 визначення рівняння регресії;
 дослідження щільності кореляційного зв'язку за коефіцієнтом кореляції.

4.2. Теоретичні положення

Якщо на просторі елементарних подій Ω задані дві випадкових величини $\xi = \xi(\omega)$ та $\eta = \eta(\omega)$ такі, що кожній елементарній події ω ставиться у відповідність впорядкована пара значень $\xi = x_i$ та $\eta = y_j$, то впорядкована пара $(\xi; \eta)$ двох одновимірних величин ξ та η називається **двовимірною випадковою величиною**. Одновимірні випадкові величини ξ та η є її **компонентами**.

Законом розподілу двовимірної випадкової величини називається відповідність між можливими значеннями цієї величини, тобто парами чисел (x_i, y_j) , та ймовірністю $p_{ij} = P(\xi = x_i; \eta = y_j)$, з якою випадкова величина $(\xi; \eta)$ набуває цих значень. Події $\xi = x_i, \eta = y_j$ утворюють

повну групу подій, отже, $\sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1$, де n – кількість значень, яких може набувати випадкова величина ξ , m – кількість значень, яких може набувати випадкова величина η . Закон розподілу дискретної випадкової величини можна задати у вигляді таблиці (**кореляційна таблиця**) або аналітично.

Відповідність між значеннями одновимірної випадкової величини $\xi = x_i$ ($i = \overline{1, n}$) та ймовірностями $P(\xi = x_i) = \sum_{j=1}^m p_{ij}$ є законом розподілу компонента ξ двовимірної випадкової величини $(\xi; \eta)$, а між значеннями одновимірної випадкової величини $\eta = y_j$ ($j = \overline{1, m}$) та ймовірностями $P(\eta = y_j) = \sum_{i=1}^n p_{ij}$ – законом розподілу її компонента η .

Кожна з одновимірних випадкових величин характеризується її математичним сподіванням:

$$M(\xi) = \sum_{i=1}^n x_i \sum_{j=1}^m p_{ij}, \quad M(\eta) = \sum_{j=1}^m y_j \sum_{i=1}^n p_{ij}, \quad (4.1)$$

та дисперсією:

$$D(\xi) = \sum_{i=1}^n x_i^2 \sum_{j=1}^m p_{ij} - (M(\xi))^2, \quad D(\eta) = \sum_{j=1}^m y_j^2 \sum_{i=1}^n p_{ij} - (M(\eta))^2. \quad (4.2)$$

Для характеристики щільності кореляційного зв'язку між компонентами двовимірної випадкової величини "мовою моментів" застосовується **кореляційний момент**, який визначається як математичне сподівання добутку відхилень компонентів випадкової величини від їх математичних сподівань:

$$\mu_{\xi\eta} = M((\xi - M(\xi)) \cdot (\eta - M(\eta))), \quad (4.3)$$

або

$$\mu_{\xi\eta} = M(\xi \cdot \eta) - M(\xi) \cdot M(\eta). \quad (4.3')$$

Відповідно, для дискретної двовимірної випадкової величини має місце співвідношення:

$$\mu_{\xi\eta} = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_{ij} - M(\xi) \cdot M(\eta). \quad (4.4)$$

Основними числовими характеристиками двовимірної випадкової величини $(\xi; \eta)$ вважаються математичні сподівання $M(\xi)$ та $M(\eta)$ і дисперсії $D(\xi)$ та $D(\eta)$ кожного з її компонентів, а також **коефіцієнт кореляції**:

$$r_{\xi\eta} = \frac{\mu_{\xi\eta}}{\sqrt{D(\xi) \cdot D(\eta)}}. \quad (4.5)$$

Розподіл одного з компонентів випадкової величини $(\xi; \eta)$ за умови, що її інший компонент набуває певного значення, називається **умовним розподілом**.

Так, **умовна ймовірність** $P(x_i | \eta = y_j)$ визначає ймовірність, з якою випадкова величина ξ набуває значення x_i ($i = \overline{1, n}$) за умови, що випадкова величина η має стале значення y_j ($j = \overline{1, m}$):

$$P(x_i | \eta = y_j) = \frac{p(x_i; y_j)}{P(\eta = y_j)}, \quad \text{де } i = \overline{1, n}, \quad j = \overline{1, m}. \quad (4.6)$$

Аналогічно умовна ймовірність $P(y_j | \xi = x_i)$ – це ймовірність, з якою випадкова величина η набуває значення y_j ($j = \overline{1, m}$) за умови, що випадкова величина ξ матиме стале значення x_i ($i = \overline{1, n}$):

$$P(y_j | \xi = x_i) = \frac{p(x_i; y_j)}{P(\xi = x_i)}, \quad \text{де } i = \overline{1, n}, \quad j = \overline{1, m}. \quad (4.7)$$

Відповідність між значеннями, яких набуває один із компонентів двовимірної випадкової величини, та їх умовними ймовірностями є умовним законом розподілу.

Так, умовний закон розподілу одного з компонентів випадкової величини $(\xi; \eta)$ записано в рядках кореляційної таблиці, для іншого – в її стовпцях.

Умовне математичне сподівання випадкової величини ξ за умови, що випадкова величина η набуває значення y_j ($j = \overline{1, m}$), визначається формулою:

$$M(\xi | \eta = y_j) = \frac{\sum_{i=1}^n x_i p_{ij}}{\sum_{i=1}^n p_{ij}}, \quad \text{де } j = \overline{1, m}. \quad (4.8)$$

Умовне математичне сподівання випадкової величини ξ є функцією від значення y_j , якого набуває випадкова величина η , тобто:

$$M(\xi | \eta = y) = g_1(y). \quad (4.9)$$

Функція $g_1(y)$ називається **регресією випадкової величини ξ на випадкову величину η** .

Аналогічно умовне математичне сподівання випадкової величини η за умови, що випадкова величина ξ набуває значення x_i ($i = \overline{1, n}$), визначається формулою:

$$M(\eta | \xi = x_i) = \frac{\sum_{j=1}^m y_j P_{ij}}{\sum_{j=1}^m P_{ij}}, \quad \text{де } i = \overline{1, n}, \quad (4.10)$$

а функція

$$M(\eta | \xi = x) = g_2(x) \quad (4.11)$$

є **регресією випадкової величини η на випадкову величину ξ** .

Графіки, що побудовані за функціями $g_1(y)$ та $g_2(x)$, називаються відповідними **лініями регресії**. Кут між лініями регресії відображає щільність кореляційного зв'язку.

Якщо двовимірна випадкова величина $(\xi; \eta)$ є такою, що обидва її компоненти ξ та η розподілені за нормальним законом, то регресія є лінійною й умовне математичне сподівання визначається за відповідним лінійним рівнянням:

$$M(\xi | \eta = y) = a_0 + a_1 y \quad \text{або} \quad M(\eta | \xi = x) = b_0 + b_1 x. \quad (4.12)$$

Так, при визначенні параметрів **рівняння регресії** випадкової величини η на ξ через основні числові характеристики двовимірної випадкової величини $(\xi; \eta)$ маємо:

$$M(\eta | \xi = x) = M(\eta) + r_{\xi\eta} \frac{\sigma_\eta}{\sigma_\xi} \cdot (x - M(\xi)).$$

Отже, коефіцієнт регресії η на ξ дорівнює:

$$\rho_{\eta/\xi} = b_1 = r_{\xi\eta} \frac{\sigma_\eta}{\sigma_\xi}, \quad (4.13)$$

а вільний член обчислюється за формулою:

$$b_0 = M(\eta) - \rho_{\eta/\xi} \cdot M(\xi). \quad (4.14)$$

4.3. Змістовна постановка задачі

Чотири монети кидають одночасно. Необхідно виконати такі завдання:

4.1. Побудувати закон розподілу випадкової величини $(\xi; \eta)$, компонентом ξ якої є кількість "орлів" на двох монетах, компонентом η – кількість "орлів" на всіх чотирьох монетах.

4.2. Побудувати умовний закон розподілу випадкової величини η , що відповідає умові $\xi = 0$.

4.3. Обчислити умовні математичні сподівання компонентів ξ та η і за цими результатами побудувати лінії регресії η на ξ та ξ на η .

4.4. Визначити основні числові характеристики двовимірної випадкової величини і зробити висновок щодо щільності кореляційного зв'язку.

4.5. Обчислити параметри рівняння регресії η на ξ .

4.4. Приклад виконання лабораторної роботи 4

Завдання 4.1. У процесі випробувань компонент ξ випадкової величини $(\xi; \eta)$ набуває значення $x_i \in \{0; 1; 2\}$, а компонент η – значення $y_j \in \{0; 1; 2; 3; 4\}$. Позначимо випадкову подію, яка полягає в тому, що на монеті під номером k ($k = \overline{1,4}$) випаде "орел", через A_k . Імовірність такої події не залежить від номера монети і дорівнює 0,5. При обчисленні ймовірності $p_{ij} = P(\xi = x_i; \eta = y_j)$ слід спиратись на теорему множення та додавання ймовірностей незалежних подій.

Вважатимемо, що результат випробування на першій і другій монетах нам відомий, а результат на третій і четвертій монетах – ні.

Нехай $\xi = x_1 = 0$ та $\eta = y_1 = 0$, тоді:

$$p_{11} = P(\overline{A}_1 \cdot \overline{A}_2 \cdot \overline{A}_3 \cdot \overline{A}_4) = P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot P(\overline{A}_3) \cdot P(\overline{A}_4) = (0,5)^4 = 1/16.$$

Якщо $\xi = x_1 = 0$ та $\eta = y_2 = 1$, то:

$$p_{12} = P(\overline{A}_1 \cdot \overline{A}_2 \cdot A_3 \cdot \overline{A}_4) + P(\overline{A}_1 \cdot \overline{A}_2 \cdot \overline{A}_3 \cdot A_4) = A_2^1 \cdot \left(\frac{1}{2}\right)^4 = \frac{2}{16},$$

де A_2^1 – кількість розміщень одного "орла" на двох монетах (третій і четвертій), результат для яких вважається невідомим.

Аналогічно обчислюємо ймовірності інших значень двовимірної випадкової величини. Обчислення здійснюємо за допомогою MS Excel і оформлюємо у вигляді кореляційної таблиці (рис. 4.1).

D3		fx = ПЕРЕСТ(2;1)*(1/2)^2*ПЕРЕСТ(2;1)*(1/2)^2					
	A	B	C	D	E	F	G
1	$\xi \backslash \eta$	0	1	2	3	4	$P(\xi=x_i)$
2	0	1/16	1/8	1/16	0	0	1/4
3	1	0	1/4	1/4	1/8	0	1/2
4	2	0	0	1/16	1/8	1/16	1/4
5	$P(\eta=y_j)$	1/16	1/4	3/8	1/4	1/16	1

Рис. 4.1. Скриншот, що містить розрахунки закону розподілу двовимірної випадкової величини

Комірки **A1:F4** описують закон розподілу двовимірної випадкової величини. Комірки **B5:F5**, значення яких обчислюються як суми за відповідними стовпцями, містять ймовірності, з якими одновимірна випадкова величина η набуває своїх значень. Отже, комірки **B1: F1** та **B5:F5** – це закон розподілу одновимірної випадкової величини η . Значення комірок **G2:G4** обчислюються як суми за відповідними рядками, отже, комірки **A2:A4** та **G2:G4** – закон розподілу одновимірної випадкової величини ξ . До комірки **G4** виводиться сума ймовірностей всіх значень випадкової величини $(\xi; \eta)$. Вона дорівнює 1, тобто це дійсно є законом розподілу.

Завдання 4.2. Для побудови умовного закону розподілу випадкової величини η , що відповідає умові $\xi = 0$, за формулами (4.7) необхідно поділити ймовірності, значення яких записані в комірках **B2:F2**, на їхню суму, що записана в комірці **G2**. Нагадуємо, що посилання на комірку **G2** повинне бути абсолютним. Отримані результати наведено в табл. 4.1.

Таблиця 4.1

Умовний закон розподілу випадкової величини η при $\xi = 0$

η	0	1	2	3	4
$P(y_j \xi = 0)$	1/4	1/2	1/4	0	0

Завдання 4.3. Для обчислення умовних математичних сподівань доповнимо кореляційну таблицю, що наведена на рис. 4.1, ще одним стовпчиком (комірки **H1:H4**) та ще одним рядком (комірки **A6:F6**).

Для обчислення умовних математичних сподівань випадкової величини η за співвідношенням (4.8) введемо, наприклад, у комірку **H2** формулу, як показано на рис. 4.2, і розтягнемо її на комірки **H2:H4**.

	A	B	C	D	E	F	G	H
1	$\xi \backslash \eta$	0	1	2	3	4	$P(\xi=x_i)$	$M(\eta \xi=x_i)$
2	0	1/16	1/8	1/16	0	0	1/4	1
3	1	0	1/8	1/4	1/8	0	1/2	2
4	2	0	0	1/16	1/8	1/16	1/4	3
5	$P(\eta=y_j)$	1/16	1/4	3/8	1/4	1/16	1	
6	$M(\xi \eta=y_j)$	0	1/2	1	1 1/2	2		

Рис. 4.2. Скриншот, що містить розрахунки умовних математичних сподівань

При побудові ліній регресії за допомогою надбудови **Мастер діаграмм** застосовуємо тип діаграми **Точечный**, при цьому вісь **OX** відповідає випадковій величині ξ , вісь **OY** – випадковій величині η . Результат наведено на рис. 4.3.

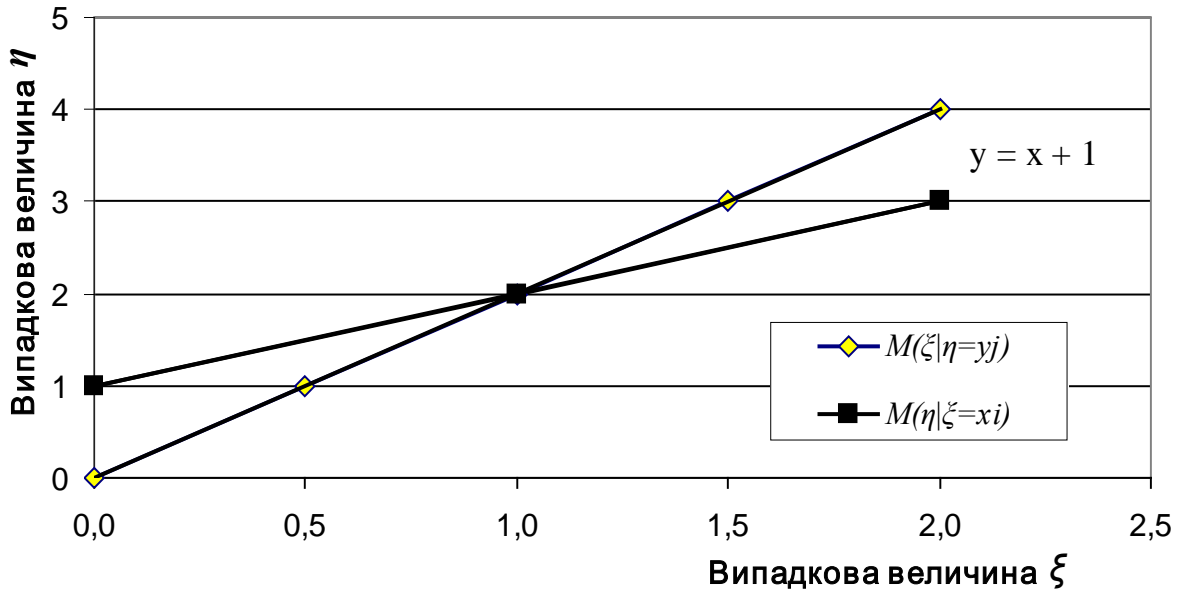


Рис. 4.3. Лінії регресії η на ξ та ξ на η

Завдання 4.4. Відкриємо новий аркуш *MS Excel*. Для визначення основних числових характеристик двовимірної випадкової величини візьмемо за основу кореляційну таблицю і доповнимо її двома рядками та двома стовпцями для проведення проміжних обчислень у процесі визначення основних числових характеристик компонентів двовимірної випадкової величини та ще одним рядком для обчислення коваріаційного моменту. Обчислення організовані таким чином, як це наведено на рис. 4.4.

		f _x = (G8-G6*H5)/(H7*16)^0,5							
	$\xi \backslash \eta$	0	1	2	3	4	$P(\xi=x_i)$	$\sum \eta_j \cdot p_{ij}$	$\sum \eta_j^2 \cdot p_{ij}$
1	0	1/16	1/8	1/16	0	0	1/4	1/4	3/8
2	1	0	1/8	1/4	1/8	0	1/2	1	2 1/4
3	2	0	0	1/16	1/8	1/16	1/4	3/4	2 3/8
4	$P(\eta=y_j)$	1/16	1/4	3/8	1/4	1/16	1	2	5
5	$\sum \xi_i \cdot p_{ij}$	0	1/8	3/8	3/8	1/8	1		1
6	$\sum \xi_i^2 \cdot p_{ij}$	0,0	0,1	0,5	0,6	0,3	1,5	1/2	
7	$\sum \xi_i \cdot \eta_j \cdot p_{ij}$	0	0,125	0,75	1,125	0,5	2,5		0,7071

Рис. 4.4. Скриншот, що містить розрахунки основних числових характеристик випадкової величини $(\xi; \eta)$

Під час обчислення математичного сподівання випадкової величини ξ в комірку **B6** вводимо формулу: **=СУММПРОИЗВ(\$A\$2:\$A\$4;B2:B4)** і розтягуємо її вздовж рядка. У комірку **G6** виводимо значення математичного сподівання як суму комірок **B6:G6**. Отже, $M(\xi) = 1$.

Для обчислення математичного сподівання квадрату випадкової величини ξ (першого доданка в перетвореній формулі дисперсії) в комірку **B7** вводимо формулу: **=СУММПРОИЗВ((\$A\$2:\$A\$4)^2;B2:B4)** і розтягуємо її вздовж рядка. У комірку **G7** виводимо значення математичного сподівання ξ^2 як суму комірок **B7:F7**. Отже, $M(\xi^2) = 2,5$. Тепер у комірку **H7** вводимо формулу: **=G7-G6^2** і отримуємо значення дисперсії: $D(\xi) = 0,5$.

Аналогічно в комірку **H2** для допоміжних обчислень вводимо формулу: **=СУММПРОИЗВ(\$B\$1:\$F\$1;B2:F2)**, розтягуємо її на комірки **H2:H4** і в комірці **H5** виводимо їхню суму. Отримуємо, що $M(\eta) = 2$. Для допоміжних обчислень при визначенні $M(\eta^2)$ в комірку **I2** вводимо формулу: **=СУММПРОИЗВ(\$B\$1:\$F\$1^2;B2:F2)**, розтягуємо її на комірки **I2:I4** і в комірці **I5** виводимо їхню суму. Отже, $M(\eta^2) = 5$. Для обчислення дисперсії у комірку **I6** вводимо формулу: **=I5-H5^2**. Звідси маємо $D(\eta) = 1$.

Тепер перейдемо до обчислення математичного сподівання добутку випадкових величин ξ та η . Для допоміжних обчислень у комірку **B8** вводимо формулу: **=B1*СУММПРОИЗВ(\$A\$2:\$A\$4;B2:B4)**, розтягуємо її на комірки **B8:F8**, а в комірці **F8** записуємо їх суму. Отже, $M(\xi \cdot \eta) = 2,5$. Залишилося визначити коефіцієнт кореляції. Ввівши в комірку **I8** формулу: **=(G8-G6*H5)/(H7*I6)^0,5**, отримуємо $r = 0,7071$. Оскільки $|r| > 0,35$, то кореляційний зв'язок між випадковими величинами ξ та η слід вважати суттєвим.

Завдання 4.5. Для обчислення параметрів рівняння регресії η на ξ підставимо значення основних числових характеристик випадкової величини $(\xi; \eta)$ у співвідношення (4.13) і (4.14). Так, для обчислення коефіцієнта регресії η на ξ в довільну комірку робочого аркуша вводимо формулу: **=I8*(I6/H7)^0,5**. Отримуємо, що $\rho_{\eta/\xi} = 1$. Далі визначаємо вільний член рівняння: $b_0 = 2 - 1 \cdot 1 = 1$.

Отже, маємо рівняння регресії η на ξ :

$$M(\eta | \xi = x) = 1 + x.$$

Повернемося до графіків лінії регресії (див. рис. 3.3). Зайдемо правою клавішею миші на будь-яку точку лінії регресії η на ξ і введемо на екран рівняння лінійного тренда. Ми отримали те ж саме рівняння.

4.5. Завдання для самостійної роботи

За допомогою вбудованих функцій та надбудов MS Excel для двовимірної випадкової величини $(\xi; \eta)$ необхідно: 1) побудувати її закон розподілу у вигляді кореляційної таблиці; 2) побудувати умовний закон розподілу одновимірної випадкової величини η за умови, що $\xi = 1$; 3) визначити умовні математичні сподівання для всіх можливих для даного прикладу умовних законів розподілу і за цими результатами побудувати спряжені лінії регресії η на ξ та ξ на η ; 4) визначити основні числові характеристики двовимірної випадкової величини і за коефіцієнтом кореляції зробити висновок щодо щільності кореляційного зв'язку між компонентами двовимірної випадкової величини $(\xi; \eta)$; 5) обчислити параметри рівняння регресії η на ξ .

Варіант 4.1

Випробування полягає у тому, що кидають п'ять монет одночасно, і спостерігають за появою "орлів". Розглядається двовимірний випадковий вектор $(\xi; \eta)$, компонентом ξ якої є кількість "орлів" на трьох монетах, а компонентом η – кількість "орлів" на всіх п'яти монетах.

Варіант 4.2

Є шість монет, які кидають одночасно. Розглядається двовимірний випадковий вектор $(\xi; \eta)$, компонентом ξ якої є кількість "орлів" на трьох монетах, а компонентом η – кількість "орлів" на шести монетах.

Варіант 4.3

У процесі перевезення ймовірність пошкодження для кожного виробу дорівнює 0,3. Після перевезення п'яти виробів вибірково перевіряють якість двох з них.

Випадкова величина ξ – кількість пошкоджених виробів серед тих, що перевіряли, η – загальна кількість пошкоджених виробів.

Варіант 4.4

Робітник обслуговує п'ять верстатів, два з яких розташовані в першому цеху, а три – у другому. Верстати працюють автономно. Імовірність того, що протягом однієї години будь-який верстат буде потребувати уваги робітника, для кожного з верстатів однакова і становить 0,2. Випадковою величиною ξ є кількість верстатів другого цеху, які будуть потребувати уваги робітника, η – загальна кількість верстатів, які потребуватимуть уваги робітника.

Варіант 4.5

Велика партія деталей містить 80 % деталей першого ґатунку. Деталі упаковані в ящики по 20 одиниць у кожному. Для вибіркової перевірки з одного ящика взяли навмання п'ять деталей. Випадковою величиною ξ є кількість деталей першого ґатунку серед деталей, вибраних для перевірки, а η – кількість деталей 1-го ґатунку серед деталей, що знаходяться в ящику.

Варіант 4.6

За даними митного поста 20 % осіб не декларують у повному обсязі товар, який підлягає оподаткуванню. Для перевірки з числа восьми пасажирів мікроавтобусу випадковим чином було відібрано п'ять осіб. Випадкова величина ξ – кількість осіб, що не задекларували товар у повному обсязі, серед тих, які були відібрані для перевірки, η – кількість осіб, що не задекларували весь товар, з числа пасажирів мікроавтобусу.

Варіант 4.7

Прилад складається з шести елементів, імовірність несправності для кожного з них дорівнює $1/9$. Для перевірки випадковим чином було відібрано два елементи. Випадкова величина ξ – кількість несправних елементів серед відібраних для перевірки, η – кількість несправних елементів серед усіх елементів, з яких складається прилад.

Варіант 4.8

У процесі виготовлення макету за певною технологією ймовірність відхилення від норми становить 0,2. Серед п'яти макетів, що були

виготовлені на замовлення, замовник відібрав два для ретельної перевірки. Випадкова величина ξ – кількість якісних макетів серед тих, що були відібрані для перевірки, η – кількість якісних макетів серед загальної кількості макетів.

Варіант № 4.9

Велика партія деталей містить 10 % деталей 2-го ґатунку. Деталі упаковані в ящики по 10 одиниць у кожному. Для перевірки якості з одного ящика взяли навмання 3 деталі. Випадкова величина ξ – кількість деталей 1-го ґатунку серед деталей, що відібрані для перевірки, а η – загальна кількість деталей 1-го ґатунку серед деталей, що знаходяться в ящику.

Варіант 4.10

Є вісім монет, які кидають одночасно. Розглядається двовимірний випадковий вектор $(\xi; \eta)$, компонентом ξ якої є кількість "орлів" на трьох монетах, а компонентом η – кількість "орлів" на восьми монетах.

Варіант 4.11

Під час проведення змагань з парашутного спорту на точність приземлення парашутист здійснює п'ять стрибків і з імовірністю 0,7 після кожного влучає в залікову зону (коло). Випадкова величина ξ – кількість влучень у залікову зону після двох стрибків, η – загальна кількість влучень у залікову зону.

Варіант 4.12

Кульки, з яких 60 % мають білий колір, а інші – чорний, упаковані в коробки по 10 кульок у кожній. З коробки виймають три кульки. Випадкова величина ξ – кількість чорних кульок серед витягнутих, η – кількість червоних кульок серед тих, що знаходяться у коробці.

Варіант 4.13

Робітник обслуговує шість верстатів, два з яких розташовані в першому цеху, а чотири – в другому. Верстати працюють автономно. Імовірність того, що протягом однієї години верстат потребуватиме налагодження, становить 0,1. Випадкова величина ξ – кількість верстатів у першому цеху, що потребуватимуть налагодження, η – загальна кількість верстатів, які потребуватимуть налагодження.

Варіант 4.14

Гральний кубик підкидається п'ять разів. Випадкова величина ξ – сума очок, які випали, η – кількість п'ятірок, що випали.

Варіант 4.15

Статистика свідчить, що 7 % працездатних людей – безробітні. Навмання вибрано 4 особи. Випадкова величина ξ – кількість безробітних серед цих 4-х осіб, η – кількість працюючих серед цих 4-х осіб.

Варіант 4.16

Два лучника роблять по 2 постріли. Імовірність влучення для першого лучника дорівнює 0,5, для другого – 0,6. Випадкова величина ξ – кількість влучень першим лучником, η – загальна кількість влучень.

Варіант № 4.17

У магазині діє акція: знижка на другий товар у чеку. У середньому цією знижкою користується кожний четвертий покупець. Протягом години товари придбали вісім покупців. Менеджер перевірів чеки чотирьох з них. Випадкова величина ξ – кількість чеків, в які внесено другий товар, серед чотирьох, що перевіряв менеджер, η – загальна кількість чеків, в які було внесено другий товар протягом години.

Варіант 4.18

Прилад складається з 5-ти елементів, імовірність несправності кожного з них дорівнює $1/9$. Випадкова величина ξ – кількість несправних елементів, η – кількість елементів, придатних до експлуатації.

Варіант № 4.19

Гральний кубик підкидається п'ять разів. Випадкова величина ξ – сума очок, які випали, η – кількість четвірок, що випали.

Варіант 4.20

На шляху руху автомобіля розташовані 6 світлофорів, кожен з яких з імовірністю 0,5 дозволяє або забороняє рух автомобіля. Чотири світлофори розташовані вздовж дороги, а два – на прилеглій вулиці. Випадкова величина ξ – кількість світлофорів, які автомобіль проїхав по вулиці без зупинки, η – загальна кількість світлофорів, які автомобіль проїхав без зупинки.

Варіант 4.21

За даними податкової інспекції 20 % осіб подають декларацію протягом першого тижня. З десяти декларацій осіб, що працюють у певній фірмі, відібрано п'ять. Випадкова величина ξ – кількість осіб (із цих п'яти), що подали декларації протягом першого тижня, а η – загальна кількість осіб, що задекларували свої статки протягом першого тижня.

Варіант 4.22

Є сім монет, які кидають одночасно й спостерігають за результатами на трьох з них. Випадкова величина ξ – кількість "орлів" на трьох монетах, а η – кількість "орлів" на всіх семи монетах.

Варіант 4.23

Два цехи обладнані однаковими верстатами, в одному з них знаходяться три верстати, у другому – п'ять. Імовірність безвідмовної роботи верстата протягом зміни становить 0,7. Випадкова величина ξ – кількість верстатів у першому цеху, які безперервно працюватимуть протягом зміни, η – кількість таких верстатів у другому цеху.

Варіант 4.24

Імовірність пошкодження виробу під час перевезення дорівнює 0,2. Випадкова величина ξ – кількість пошкоджених виробів серед двох, узятих для контролю, η – кількість пошкоджених виробів серед усіх семи виробів, які були перевезені.

Варіант 4.25

Два приятелі придбали по 5 лотерейних білетів. За статистикою ймовірність виграшу для кожного з білетів дорівнює 0,2. Випадкова величина ξ – кількість виграшних білетів у одного з приятелів, η – загальна кількість виграшних білетів у обох приятелів.

4.6. Контрольні запитання

1. Дайте означення багатовимірної випадкової величини. Наведіть приклад двовимірної випадкової величини.
2. Якими бувають багатовимірні величини?

3. Наведіть способи задавання закону розподілу двовимірної дискретної випадкової величини й охарактеризуйте їх особливості.

4. Що таке умовний розподіл?

5. Як побудувати умовний розподіл за кореляційною таблицею двовимірної випадкової величини?

6. Опишіть, як за кореляційною таблицею, що визначає розподіл дискретної двовимірної величини, отримати розподіл кожного з її компонентів.

7. Назвіть основні числові характеристики розподілу двовимірної випадкової величини. Поясніть, що саме вони характеризують.

8. Наведіть означення кореляційного зв'язку.

9. Які ви знаєте числові характеристики, які визначають щільність кореляційного зв'язку?

10. У чому полягають переваги використання коефіцієнта кореляції як характеристики щільності кореляційного зв'язку порівняно з кореляційним моментом?

11. Чи може коефіцієнт кореляції бути від'ємним? Якщо так, то що це означає?

12. Як визначити параметри рівняння лінійної регресії через основні числові характеристики двовимірної випадкової величини?

13. Поясніть, яке взаємне розташування можуть мати спряжені лінії регресії η на ξ та ξ на η , якщо ξ та η є компонентами двовимірної випадкової величини $(\xi; \eta)$.

14. Яку інформацію можна отримати щодо тісноти і напряму кореляційного зв'язку з дослідження взаємного розташування спряжених ліній регресії?

15. Чи є некорельовані випадкові величини незалежними?

Змістовий модуль 2. Математична статистика

Лабораторна робота 5

Основні поняття математичної статистики.

Статистичне оцінювання параметрів розподілу

5.1. Мета роботи:

ознайомлення з вибіркоvim методом статистичних досліджень;
засвоєння принципів побудови інтервального варіаційного ряду,
гістограми й кумуляти;

визначення точкових та інтервальних оцінок параметрів емпіричного розподілу;

застосовування таких вбудованих функцій *MS Excel*, як **МИН**, **МАКС**, **СРЗНАЧ**, **ДИСП.В**, **ДОВЕРИТ.НОРМ**, **СТАНДОТКЛОН**, **НОРМ.ОБР**, а також процедурами **Гистограмма** та **Описательная статистика** надбудови **Анализ данных** для визначення числових характеристик випадкової величини.

5.2. Теоретичні положення

Генеральною сукупністю є множина всіх значень, яких набуває (або може набувати) в процесі випробувань випадкова величина X . Оскільки об'єкти поєднані у генеральну сукупність за певною ознакою, це передбачає, що вплив економічних факторів, дослідження яких і ставиться за мету, у межах сукупності підпорядковується єдиному закону. Якщо вимірюють лише одну характеристику об'єкта, то її значення утворюють **одновимірну випадкову величину**. Обсяг генеральної сукупності може бути занадто великим або процес дослідження передбачає руйнування об'єкта, тому для проведення дослідження застосовується **вибірковий метод**.

Множина значень випадкової величини, яких вона набуває під час обмеженої кількості вимірювань, називається **вибірковою сукупністю**, або **вибіркою**. Кількість вимірювань n становить **обсяг вибірки**. Значення x_i , якого набуває досліджувана випадкова величина X у вибірковій сукупності, називається **варіантою**.

За результатами обчислення вибірових даних роблять висновок щодо характеристик випадкової величини в генеральній сукупності, тому

вибірка повинна бути **репрезентативною**, тобто правильно відображати властивості об'єкта дослідження.

Відповідність між впорядкованою послідовністю варіант x_i і частотами m_i , з якими ці варіанти зустрічаються у вибірковій сукупності, називається **дискретним варіаційним рядом**. Цей ряд є **статистичним**, або **емпіричним розподілом** дискретної випадкової величини, а його графічним відображенням є **полігон частот**.

Для неперервної випадкової величини або в тому випадку, коли випадкова величина є дискретною, але має велику кількість варіант, то користуватися дискретним варіаційним рядом стає незручно. У цьому випадку емпіричний закон розподілу надають у вигляді **інтервального варіаційного ряду**, де результати вимірювань об'єднуються за інтервалами однакової довжини, яка називається **кроком**, а сам інтервал у термінах *MS Excel* називається **карманом** (табл. 5.1).

Таблиця 5.1

Інтервальний варіаційний ряд

$[x_i; x_{i+1})$	$[x_1; x_2)$	$[x_2; x_3)$	$[x_3; x_4)$...	$[x_k; x_{k+1})$
m_i	m_1	m_2	m_3	...	m_k

Процедура визначення оптимальної кількості інтервалів базується на застосуванні формули Стерджесса:

$$k = 1 + 3,322 \cdot \lg n, \quad (5.1)$$

де k – кількість інтервалів; n – обсяг вибірки.

Кількість інтервалів, що обчислюється за формулою (5.1), округляється до цілого числа. Залежно від кількості вимірювань це число знаходиться в межах від 8 до 12.

Для визначення довжини інтервалу, тобто кроку (h), необхідно різницю між найбільшим (x_{\max}) та найменшим (x_{\min}) значеннями випадкової величини, тобто **розмах** (R), поділити на кількість інтервалів.

Отже, попередню оцінку довжини інтервалу можна отримати за формулою:

$$h = \frac{x_{\max} - x_{\min}}{k}. \quad (5.2)$$

Значення, яке отримують за формулою (5.2), округлюють до найбільш зручного, якщо це необхідно. У результаті округлення шагу нижньою границею першого інтервалу, тобто x_1 , можна зсунути відносно x_{\min} у бік менших значень, а верхню границю останнього інтервалу, тобто x_{k+1} , можна зсунути відносно x_{\max} у бік більших значень. Однак цей зсув не повинен перевищувати $0,5h$.

Графічним відображенням дискретного варіаційного ряду є **полігон частот**, тобто ламана лінія, для побудови якої точки $(x_i; m_i)$ з'єднують відрізками прямої. Для отримання **полігону відносних частот** ламану лінію будують на точках $(x_i; w_i)$. Графічним відображенням інтервального варіаційного ряду є **гістограма**. Для того, щоб отримати **гістограму частот**, на кожному інтервалі $[x_i; x_{i+1})$ інтервального ряду, як на основі, будують прямокутники, висота кожного з яких дорівнює m_i/h (або, відповідно, w_i/h для **гістограми відносних частот**). Отже, площа i -го прямокутника гістограми частот становить $h \cdot m_i/h = m_i$, звідси загальна площа всіх прямокутників такої гістограми чисельно дорівнює обсягу вибіркової сукупності n . Відповідно загальна площа всіх прямокутників гістограми відносних частот (частостей) дорівнює одиниці. Гістограма відносних частот відображає оцінку функції щільності розподілу ймовірностей випадкової величини. За виглядом полігону або гістограми можна висловити припущення щодо закону розподілу випадкової величини в генеральній сукупності.

Числові характеристики, що обчислені за вибірковими даними, є **статистичними оцінками** $\hat{\Theta}$ відповідних числових характеристик Θ випадкової величини. Так, статистичною оцінкою математичного сподівання випадкової величини X є **вибіркова середня** \bar{x} (центр вибіркової сукупності), яка є статистичним початковим моментом 1-го порядку.

Для згрупованих даних вона обчислюється як середнє виважене варіант x_i , кожна з яких береться з вагою, що відповідає її відносній частоті:

$$\bar{x} = \sum_{i=1}^k x_i w_i, \quad (5.3)$$

де k – кількість варіант у дискретному варіаційному ряді або кількість інтервалів в інтервальному варіаційному ряді;

x_i – значення варіанти для дискретного варіаційного ряду;

w_i – **відносна частота**, або **частіть**, яка визначається як відношення частоти m_i до обсягу вибірки $w_i = \frac{m_i}{n}$ ($i = \overline{1, n}$).

Для інтервального варіаційного ряду формула (5.3) набуває вигляду:

$$\bar{x} = \sum_{i=1}^k \bar{x}_i w_i, \quad (5.3^1)$$

де \bar{x}_i – середина i -го інтервалу, яка визначається як $\bar{x}_i = 0,5 \cdot (x_i + x_{i+1})$.

Вибіркова середня є **незсунутою оцінкою** математичного сподівання, тобто математичне сподівання вибіркової середньої дорівнює математичному сподіванню випадкової величини.

Вибіркова дисперсія \hat{D} є статистичним центральним моментом 2-го порядку і визначається як середнє виважене квадратів відхилення варіант від вибіркової середньої випадкової величини, кожне з яких береться з вагою, що відповідає відносній частоті варіанти:

$$\hat{D} = \sum_{i=1}^k (x_i - \bar{x})^2 \cdot w_i. \quad (5.4)$$

Для інтервального варіаційного ряду формула (5.4) набуває вигляду:

$$\hat{D} = \sum_{i=1}^k (\bar{x}_i - \bar{x})^2 \cdot w_i, \quad (5.4^1)$$

Вибіркова дисперсія є зсунутою оцінкою дисперсії випадкової величини генеральної сукупності. Для обчислення **виправленої вибіркової дисперсії** S_x^2 користуються формулою:

$$S_x^2 = \frac{n}{n-1} \hat{D}. \quad (5.5)$$

Відповідно **виправлене середнє квадратичне відхилення** визначається як $S_x = \sqrt{S_x^2}$. Саме воно є незсунутою оцінкою середнього квадратичного відхилення теоретичного розподілу, що характеризує розпорошення випадкової величини навколо центра сукупності.

Вибіркова середня та виправлене середнє квадратичне відхилення є **точковими оцінками** відповідних числових характеристик теоретичного розподілу, оскільки кожна з них визначена одним числом $\hat{\Theta}$ (точкою). Однак оцінки числових характеристик генеральної сукупності самі є випадковими величинами, тому окрім точкових розглядають ще **інтервальні оцінки**, тобто **довірчий інтервал** $(\hat{\Theta}_1; \hat{\Theta}_2)$, до якого з **довірчою імовірністю** P належатиме невідоме значення числової характеристики теоретичного розподілу.

Імовірність того, що числова характеристика теоретичного розподілу накриває певний довірчий інтервал, називається **надійністю оцінки** γ , яка, у свою чергу, пов'язана з **рівнем значущості** α : $\alpha = 1 - \gamma$. За означенням довірчого інтервалу маємо:

$$P(\hat{\Theta}_1 < \Theta < \hat{\Theta}_2) = 1 - \alpha. \quad (5.6)$$

У припущенні нормального закону розподілу в генеральній сукупності довірчий інтервал для математичного сподівання a визначається як:

$$\bar{x} - t_\alpha \cdot S_x / \sqrt{n} < a < \bar{x} + t_\alpha \cdot S_x / \sqrt{n}, \quad (5.7)$$

де t_α – квантіль розподілу Стюдента за рівнем α .

Чинник $t = t(\alpha, n)$ визначається залежно від рівня значущості α та кількості ступенів вільності $df = n - 1$ за довідковою таблицею розподілу Стюдента.

Якщо обсяг вибірки великий, то в якості t_α у формулі (5.7) застосовується аргумент функції Лапласа, який визначається за довідковою

таблицею відповідно до вибраного рівня значущості α зі співвідношення $\Phi(t) = 0,5 \cdot (1 - \alpha)$.

З довірчою ймовірністю P , що дорівнює γ , похибка оцінки ε математичного сподівання за вибірковою середньою не перевищуватиме величини $\varepsilon = t_\alpha \cdot S_x / \sqrt{n}$.

Довірчий інтервал, який накриває значення дисперсії генеральної сукупності з надійністю $\gamma = 1 - \alpha$, можна обчислити за формулою:

$$S_x^2 \cdot \frac{n-1}{\chi_1^2} < \sigma^2 < S_x^2 \cdot \frac{n-1}{\chi_2^2}, \quad (5.8)$$

де χ_1^2 – значення критерію χ^2 для ймовірності $\alpha/2$ та кількості ступенів вільності $n - 1$;

χ_2^2 – значення критерію χ^2 для ймовірності $1 - \alpha/2$ та кількості ступенів вільності $n - 1$.

Критичні точки розподілу χ^2 визначаються за відповідною довідковою таблицею.

Емпіричною функцією розподілу називається функція $\hat{F}(x)$, яка для будь-якого значення x визначається як відносна частота події, яка полягає в тому, що випадкова величина набуває значень, які менші за x . Згідно з цим означенням $\hat{F}(x) = m_x / n$, де m_x – загальна кількість варіант, значення яких менші за x . Отже, емпірична функція розподілу визначається шляхом послідовного додавання відносних частот, що відповідають варіантам, меншим за x . Графіком цієї функції є **кумулята**.

5.3. Змістовна постановка задачі

Вихідні дані, що містяться в таблиці (табл. 5.2), описують місячну продуктивність вуглевидобувної бригади. Ця вибіркова сукупність містить 168 варіант і є репрезентативною.

Необхідно за даними вибіркової сукупності:

побудувати емпіричний закон розподілу у вигляді інтервального варіаційного ряду;

побудувати гістограму частот та кумуляту для емпіричної функції розподілу $\hat{F}(x)$;

обчислити вибірккову середню \bar{x} , дисперсію та виправлене середнє квадратичне відхилення S_x ;

знайти довірчий інтервал для математичного сподівання та середнього квадратичного відхилення, до якого відповідні основні числові характеристики теоретичного розподілу належатимуть з надійністю 95 %.

Таблиця 5.2

Довжина горизонтальної проходки (м) у процесі вуглевидобування в умовах, близьких за категорією гірських робіт

56	139	54	160	133	125	196	188	93	123	148	105
216	99	197	186	119	129	145	140	63	78	124	111
402	196	199	105	105	50	331	88	141	83	130	185
118	134	72	101	125	104	243	165	77	90	115	188
53	161	117	115	76	344	58	133	142	114	143	141
121	71	147	106	89	320	61	75	150	127	176	179
161	117	132	118	193	102	122	81	107	235	223	99
70	86	108	140	119	60	169	138	112	230	244	66
137	106	151	350	55	100	124	127	222	309	116	154
147	127	91	92	197	72	103	128	178	156	66	81
112	470	182	176	141	106	115	176	301	144	159	85
126	233	135	60	72	87	107	136	250	65	149	126
290	51	101	89	116	131	200	187	118	104	137	114
203	117	59	194	138	126	120	173	184	80	107	110

5.4. Приклад виконання лабораторної роботи 5

Слід зауважити, що великий обсяг вибіркової сукупності ($n = 168$) свідчить про ретельність досліджень. Це підвищує надійність висновків, які можна зробити завдяки аналізу емпіричних даних. Спочатку на робочому аркуші *MS Excel* записуємо вихідні дані. Оскільки статистика наведена для одновимірної випадкової величини, то значення варіант набирають в одному стовпчику (комірки **A1:A168**), інакше *MS Excel* сприйматиме випадкову величину як багатовимірну.

Побудуємо інтервальний варіаційний ряд. Це можна зробити, застосувавши надбудову **Анализ данных**. Ця надбудова дозволяє також

одночасно отримати гістограму і кумуляти. Слід зазначити, що у ході застосування процедури **Гистограмма**, що міститься в пакеті **Анализ данных**, нема потреби в явному задаванні меж інтервалів, оскільки ця процедура сама здатна їх визначити.

Для застосування процедури **Гистограмма** вибираємо на **Ленте** вкладку **Данные** і визначаємо шлях **Анализ данных** ⇒ **Гистограмма**. У діалоговому вікні цієї процедури (рис. 5.1) вказуємо **Входной интервал**: (посилання на комірки **A1:A168**, де містяться значення одномірної випадкової величини), **Выходной интервал**: (посилання на будь-яку вільну комірку на робочому аркуші) та прапорцями відзначаємо потрібні нам опції **Интегральный процент** і **Вывод графика**. Поле **Интервал карманов** залишаємо вільним, отже, межі інтервалів будуть визначатись автоматично.

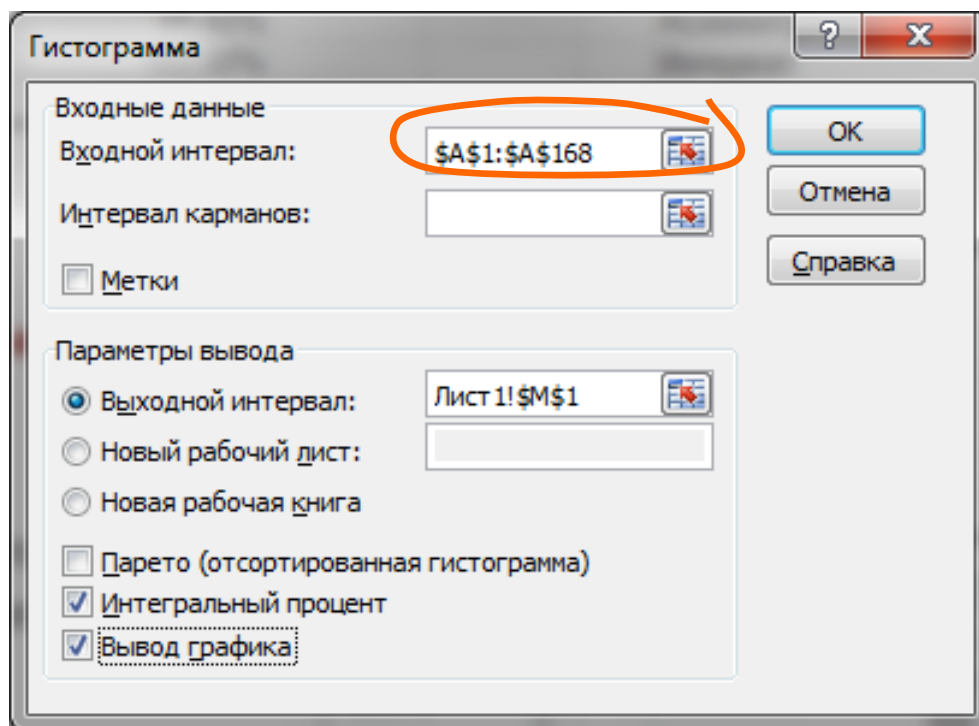


Рис. 5.1. Діалогове вікно процедури **Гистограмма**

Натисканням клавіші **ОК** виводяться таблиця з результатами систематизації (табл. 5.3), гістограма й кумулята (рис. 5.2). Видно, що при цьому розмах автоматично було поділено на 12 інтервалів, і в даному прикладі крок становить 37,75.

У першому стовпчику табл. 5.3, що має назву **Карман**, містяться значення верхньої межі кожного інтервалу. Другий стовпчик, який має

назву **Частота**, містить кількість значень випадкової величини у вибраній інтервал, які належать даному інтервалу. Ці два стовпчики й утворюють емпіричний розподіл випадкової величини X .

Таблиця 5.3

Результат систематизації вибіркової сукупності

<i>Карман</i>	<i>Частота</i>	<i>Интегральный %</i>
50	1	0,60 %
85	28	17,26 %
120	48	45,83 %
155	44	72,02 %
190	20	83,93 %
225	12	91,07 %
260	6	94,64 %
295	1	95,24 %
330	3	97,02 %
365	3	98,81 %
400	0	98,81 %
435	1	99,40 %
Еще	1	100,00 %

Слід зауважити, що під час визначення частот потрапляння значень випадкової величини до певного інтервалу програма *MS Excel* автоматично враховує такі умови:

$$x \leq 50; 85 < x \leq 120; 120 < x \leq 155; \dots; 400 < x \leq 435; x > 435.$$

За даними перших двох стовпчиків побудована гістограма. Третій стовпчик, що має назву **Интегральный %**, містить відносну кількість значень (у відсотках) випадкової величини, які менші за нижню межу відповідного інтервалу, за цими даними побудована кумулята (рис. 5.2).

Гістограма, яка виводиться після виконання відповідної процедури, не відповідає двом вимогам. По-перше, ширина стовпчиків менша ніж крок. Цей недолік легко виправляється. Для цього правою клавішею миші заходимо на будь-який стовпчик гістограми і у діалоговому вікні, що відкрилося, вибираємо опцію **Формат ряда данных** (рис. 5.3).

Встановлюємо **Боковой зазор: Без зазора** (рис. 5.3) і натискаємо **ОК**.

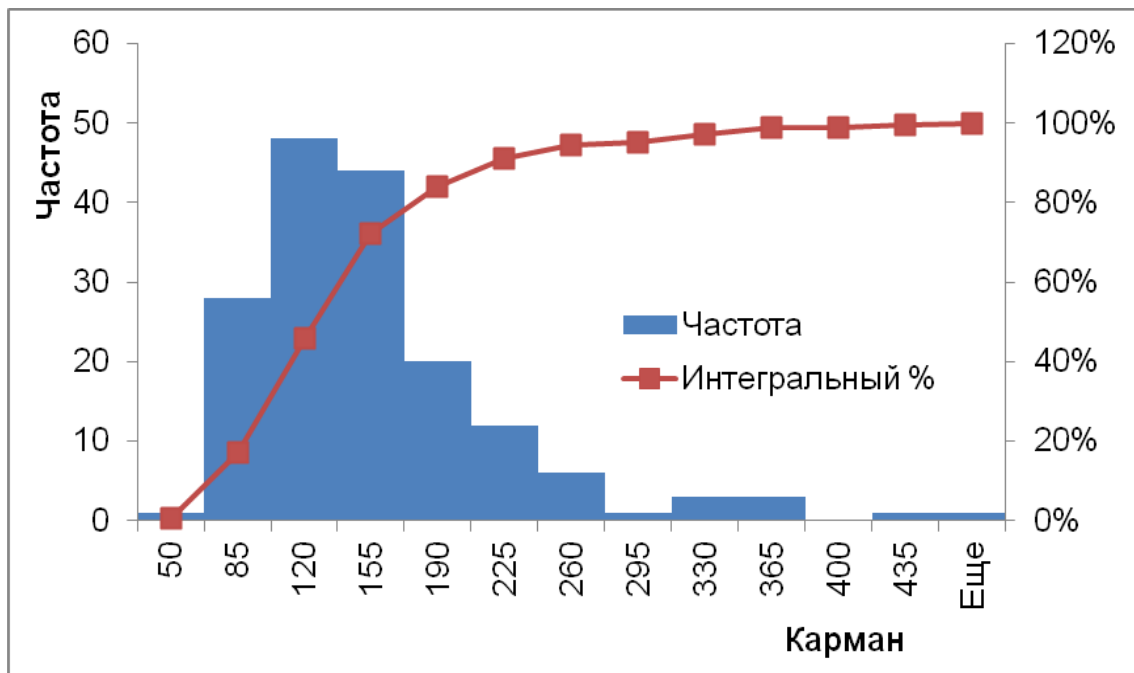


Рис. 5.2. Результати процедури *Гистограмма*

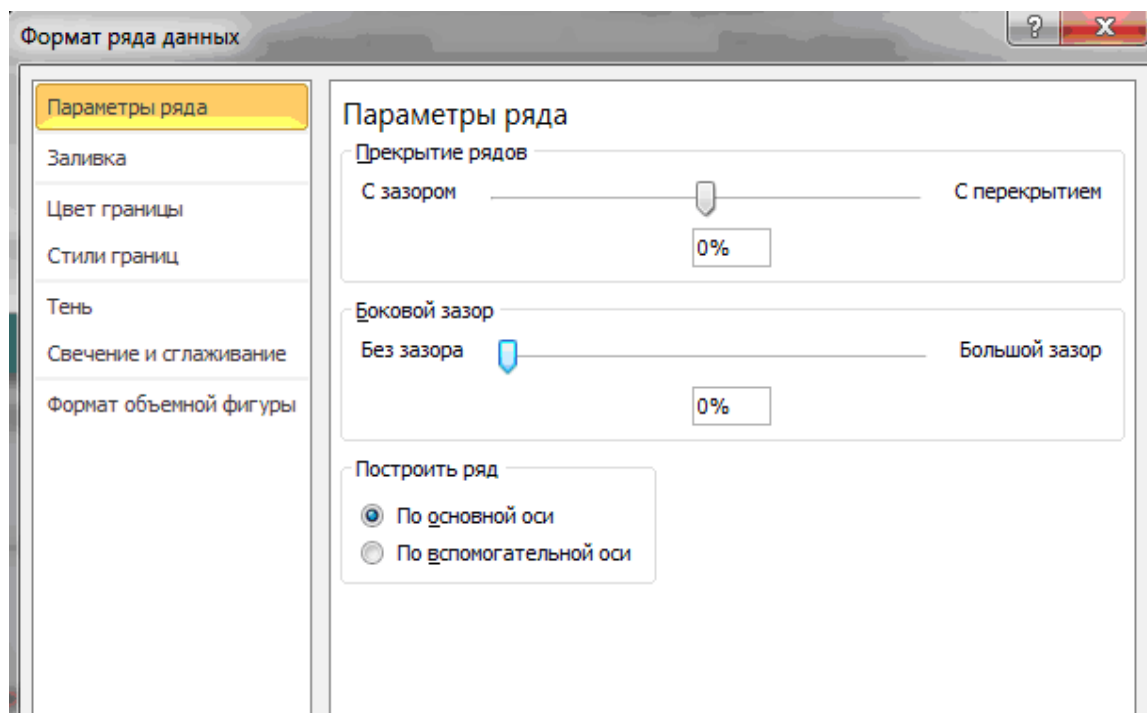


Рис. 5.3. Діалогове вікно *Формат ряда данных*

Але є ще один недолік: площа під гистограмою частот повинна чисельно дорівнювати обсягу вибірки. Однак при побудові гистограми за допомогою процедури **Гистограмма** ця умова порушується. Для того, щоб умова нормування виконувалася, необхідно замість частот

використовувати відношення частот до довжини інтервалу (або до кроку, якщо інтервали є рівними за довжиною). Це можна здійснити за допомогою додаткових обчислень.

Для явного задавання меж інтервалів знаходимо верхнє і нижнє значення варіант. Це можна зробити кількома способами. Наприклад, вибираємо на **Ленте** вкладку **Главная**, виділяємо комірки **A1:A168**, натискаємо кнопку **Сортировка и фильтр** і здійснюємо операцію **Сортировка по возрастанию**.

Отже, знаходимо найменше та найбільше значення випадкової величини: $x_{\min} = 50$, $x_{\max} = 470$, які потрібні для визначення меж інтервалів (**карманов**).

Визначити межі розмаху випадкової величини можна і без упорядкування варіант, якщо застосувати вбудовані функції **МИН** і **МАКС**. З цією метою встановлюємо курсор у будь-яку вільну комірку, і у списку функцій, що належать до категорії **Статистические**, вибираємо **МАКС**. У її діалоговому вікні в полі **Число 1** надаємо посилання на діапазон вихідних даних, тобто **=МАКС(A1:A168)** і натисканням **ОК** виводимо результат: $x_{\max} = 470$. Аналогічно за допомогою формули **=МИН(A1:A168)** визначаємо мінімальне значення варіант випадкової величини: $x_{\min} = 50$.

Тепер визначаємо розмах варіант як різницю між їх максимальним та мінімальним значеннями: $R = x_{\max} - x_{\min} = 420$.

Якщо прийняти кількість інтервалів $k = 10$, то довжина інтервалу становитиме $h = 420/10 = 42$. Але такий крок є не дуже зручним. Спробуємо винайти інший. Якщо нижню межу варіант зсунути до $x = 0$, а верхню – до $x = 500$, то, поділивши новий розмах $R = 500 - 0 = 500$ на $k = 10$, отримаємо довжину кроку $h = 50$. Такий шаг зручний у застосуванні, отже, приймемо його. Визначимо межі інтервалів, які вказуємо на робочому аркуші в окремому стовпчику (нехай для визначеності це будуть комірки **D1:D10**). Для побудови гістограми викликаємо її діалогове вікно і на цей раз вказуємо межі інтервалів, для чого у полі **Интервал карманов:** даємо посилання на діапазон комірок **D1:D10** (рис. 5.4).

Після натискання **ОК** на робочому аркуші виводяться таблиця результатів групування (табл. 5.4), а також гістограма і кумулята (рис. 5.5). Зрозуміло, що ми отримали трохи інший результат, ніж при автоматичному визначенні меж інтервалів.

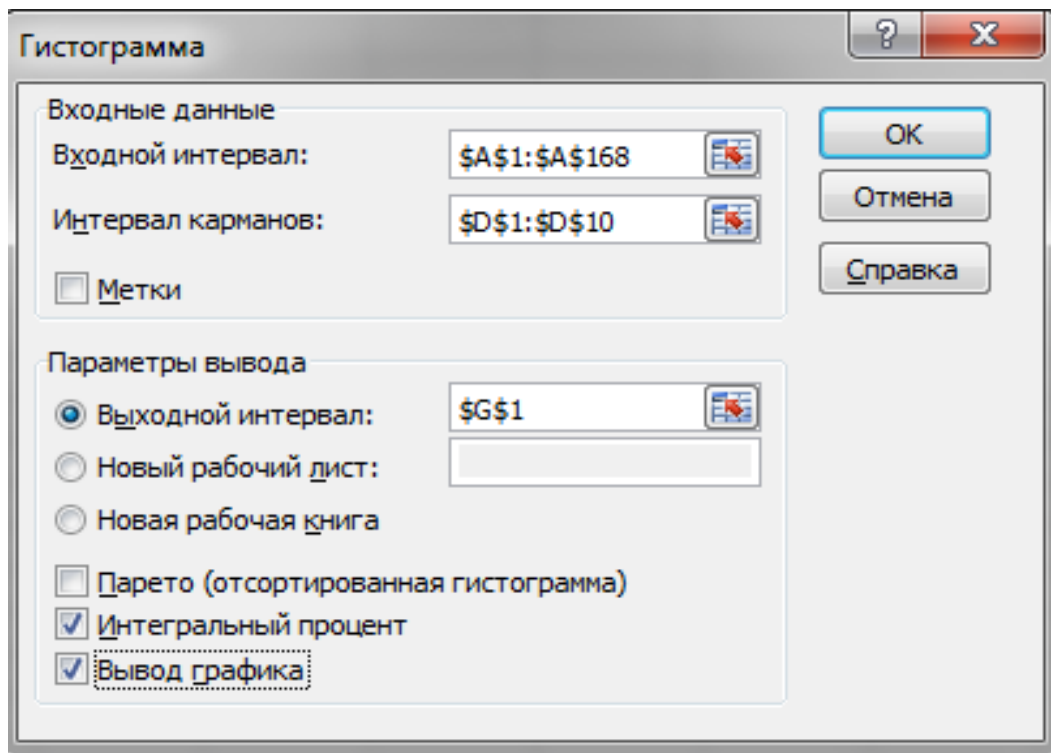


Рис. 5.4. **Діалогове вікно процедури Гистограмма**

Таблица 5.4

Групування даних за визначеними інтервалами

<i>Карман</i>	<i>Частота</i>	<i>Интегральный %</i>
50	2	0,60 %
100	40	24,40 %
150	77	70,83 %
200	30	88,69 %
250	10	94,64 %
300	1	95,24 %
350	6	98,81 %
400	0	98,81 %
450	1	99,40 %
500	1	100,00 %
Еще	0	100,00 %

Оскільки ми отримали трохи інший інтервальний варіаційний ряд, то і залежно від прийнятого емпіричного закону розподілу матимемо трохи інші статистичні оцінки його числових характеристик, однак тип гістограми не змінився.

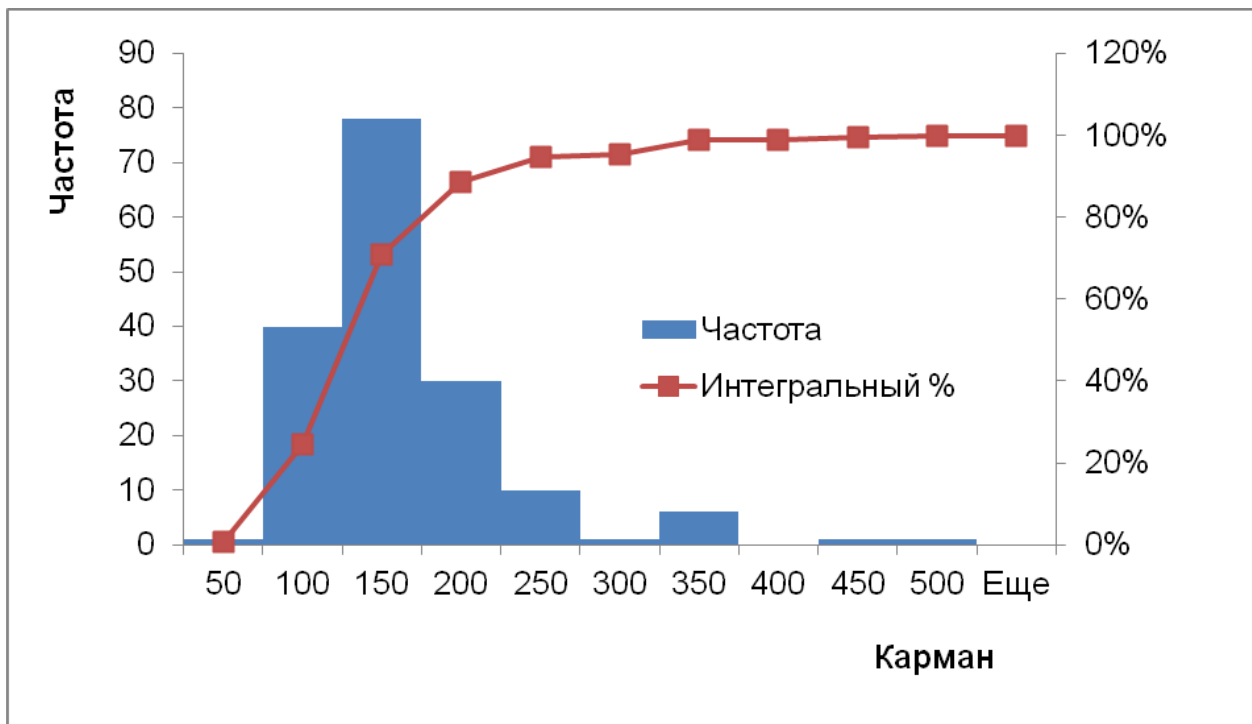


Рис. 5.5. Графічне надання результатів статистичної обробки вихідних даних за допомогою процедури *Гистограмма*

Проведемо аналіз результатів групування. Оскільки гістограма має один максимум, а також частоти монотонно спадають як у разі збільшення, так і в разі зменшення значення випадкової величини відносно максимуму, то вибірккову сукупність можна вважати такою, що наближена до однорідної, та перейти до оцінки параметрів вибіркової сукупності, тобто до описової статистики.

Електронні таблиці *MS Excel* дозволяють визначати числові характеристики емпіричного розподілу кількома способами, наприклад, використовуючи функцію *MS Excel* або за допомогою процедури **Описательная статистика**, або шляхом організації відповідних розрахунків.

Спочатку на **Ленте** вибираємо вкладку **Данные**, а потім визначаємо шлях **Анализ данных** ⇒ **Описательная статистика**. Після заповнення діалогового вікна цієї процедури (рис. 5.6) на робочому аркуші у вихідному інтервалі виводиться таблиця з результатами обчислення статистичних параметрів та їх назвами (табл. 5.5). Результати обчислення, що виводяться на екран, містяться в перших двох стовпчиках табл. 5.5, у третьому ми додали коментар, що пояснює зміст оцінок. Зауважимо, що кожна з цих характеристик окремо можна вивести за допомогою вбудованих функцій.

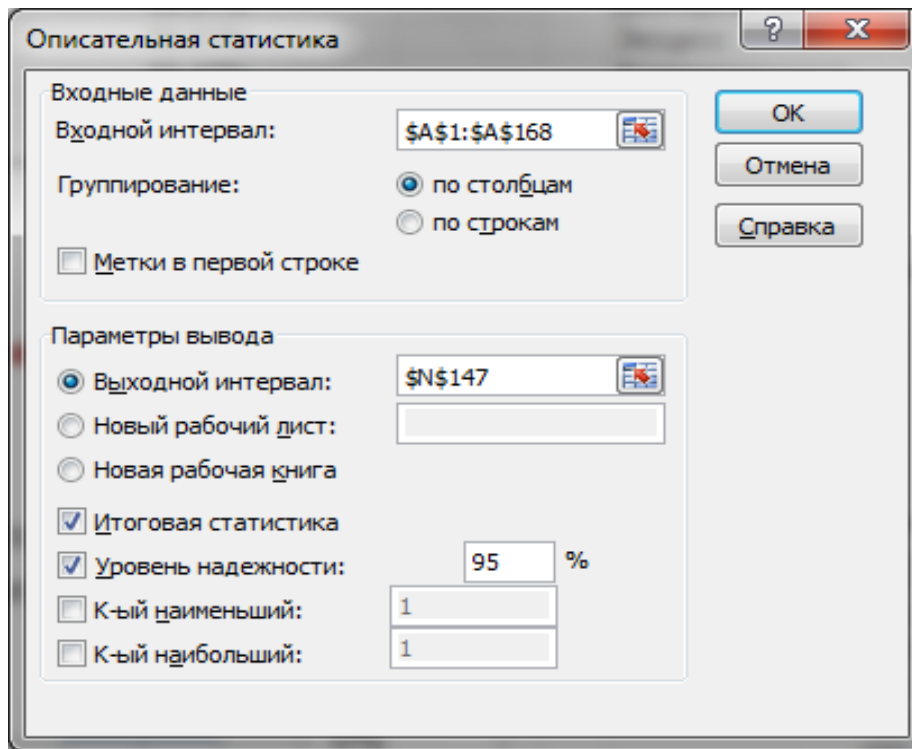


Рис. 5.6. Діалогове вікно процедури *Описательная статистика*

Таблица 5.5

**Підсумкова таблиця, що відповідає виконанню процедури
*Описательная статистика***

Столбец 1		Коментар
Среднее	139,7381	Среднее
Стандартная ошибка	5,219274	Стандартная ошибка
Медиана	125,5	Медиана
Мода	118	Мода
Стандартное отклонение	67,64952	Стандартное отклонение
Дисперсия выборки	4576,458	Дисперсия выборки
Эксцесс	4,827949	Эксцесс
Асимметричность	1,828735	Асимметричность
Интервал	420	Интервал
Минимум	50	Минимум
Максимум	470	Максимум
Сумма	23476	Сумма
Счет	168	Счет

Слід звернути увагу на те, що **стандартна похибка** під час визначення середньої вибіркової сукупності обчислюється за формулою:

$$SE = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1) \cdot n}}, \quad \text{або} \quad SE = \frac{S_x}{\sqrt{n}}. \quad (5.9)$$

Пропонуємо самостійно визначити оцінки основних і допоміжних числових характеристик розподілу за допомогою відповідних вбудованих функцій, що належать до категорії **Статистические**.

Наприклад, за допомогою функції **СРЗНАЧ**, аргументами якої є значення випадкової величини, що утворюють вибірку сукупність (нагадаємо, що вони наведені у комірках **A1:A168**), можна визначити вибірку середню, а за допомогою функції **ДИСП.В** – оцінити дисперсію вибірки.

Пропонуємо самостійно визначити ці та інші точкові оцінки числових характеристик розподілу за допомогою вбудованих функцій **СРЗНАЧ**, **ДИСП.В**, **ДОВЕРИТ.НОРМ**, або **ДОВЕРИТ.СТЬЮДЕНТ**, **МЕДИАНА**, **МОДА.НСК**, **СТАНДОТКЛОН.В**, **СКОС**, **ЭКССЕСС** і порівняти їх із результатами застосування процедури **Описательная статистика**. Зрозуміло, що ми отримуємо ті ж самі результати. Однак для інтервальних законів розподілу ми отримуємо дещо інші показники.

Для визначення довірчого інтервалу, до якого з надійністю $\gamma = 95\%$ належатиме математичне сподівання теоретичного закону розподілу, застосовуємо формулу (5.7), скориставшись для цього даними табл. 5.5. Так, ми отримали, що $\bar{x} = 139,7381$ та $SE = 5,2193$.

За допомогою функції **НОРМ.ОБР**, яка виводить значення інтегральної функції нормального розподілу, знаходимо аргумент функції Лапласа, що відповідає її значенню $0,5 \cdot 0,95 = 0,475$.

Для цього в діалоговому вікні аргументу цієї функції (рис. 5.7) у полі **Вероятность** вказуємо значення $0,5 + 0,475 = 0,945$, у полі **Среднее** ставимо 0, у полі **Стандартное_откл** – значення 1. Отримуємо, що $t_\alpha = 1,96$.

Тепер обчислюємо похибку оцінки математичного сподівання, що відповідає заданій надійності: $\varepsilon = 1,96 \cdot 5,2193 \approx 10,2296$.

Отже, з довірчою ймовірністю $P = 0,95$ можна стверджувати, що математичне сподівання випадкової величини в генеральній сукупності не менше ніж 136,48 та не більше ніж 149,97.

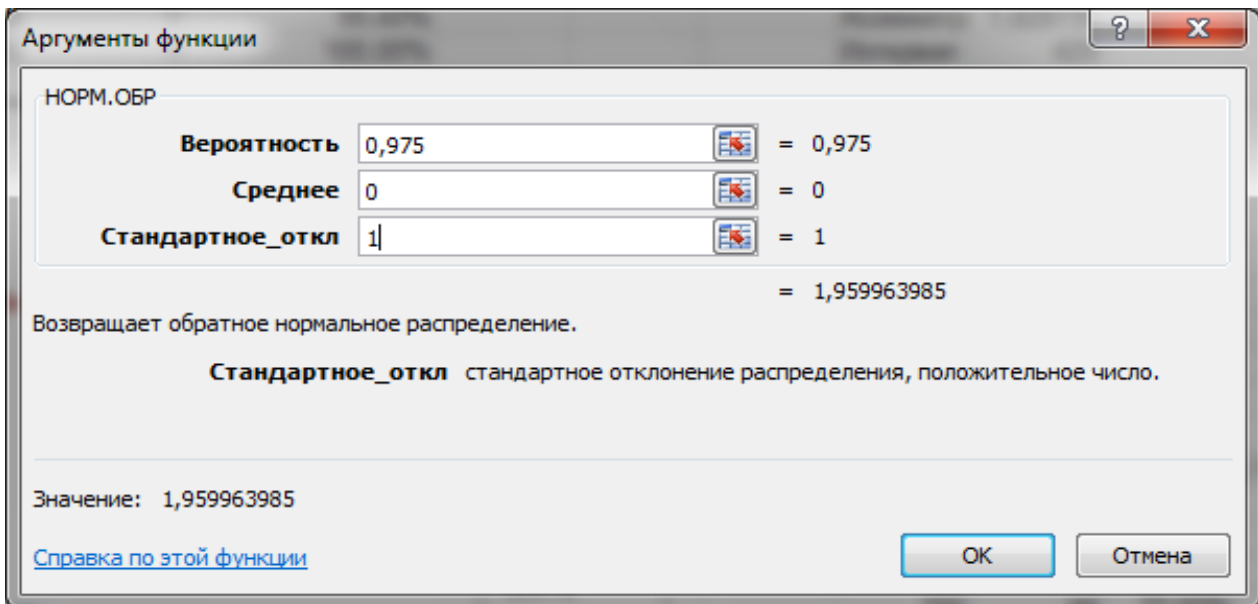


Рис. 5.7. Диалоговое окно функции НОРМ.ОБР

Порівняємо ці значення з результатами, які можна отримати за допомогою функції **ДОВЕРИТ.НОРМ**. Викликаємо цю функцію і у діалоговому вікні вказуємо її аргументи (рис. 5.8).

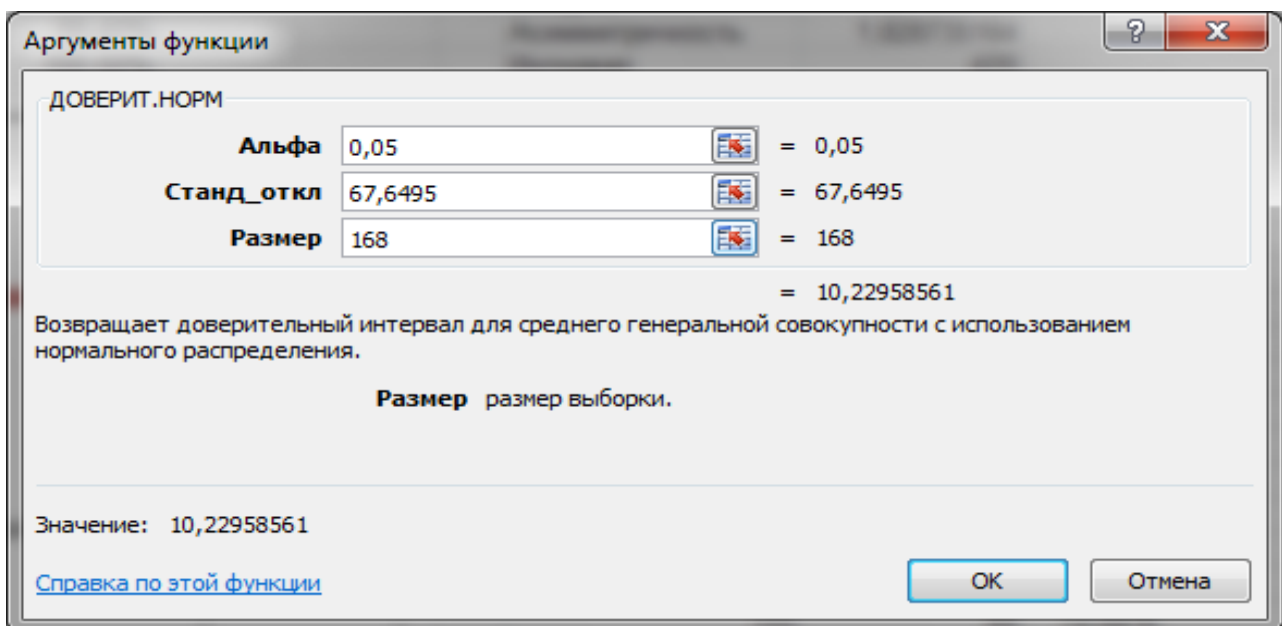


Рис. 5.8. Диалоговое окно функции ДОВЕРИТ.НОРМ

Як бачимо, ми отримали той самий результат.

Тепер визначимо довірчий інтервал, до якого з надійністю $\gamma = 95\%$ належатиме дисперсія теоретичного закону розподілу. Для застосування

формули (5.8) спочатку треба знайти ліву та праву критичні точки розподілу χ^2 . Скористаємося вбудованою функцією **ХИ2.ОБР.ПХ**.

На рис. 5.9 наведено діалогове вікно цієї функції, у полях якого вказані аргументи для обчислення правої границі критичної області.

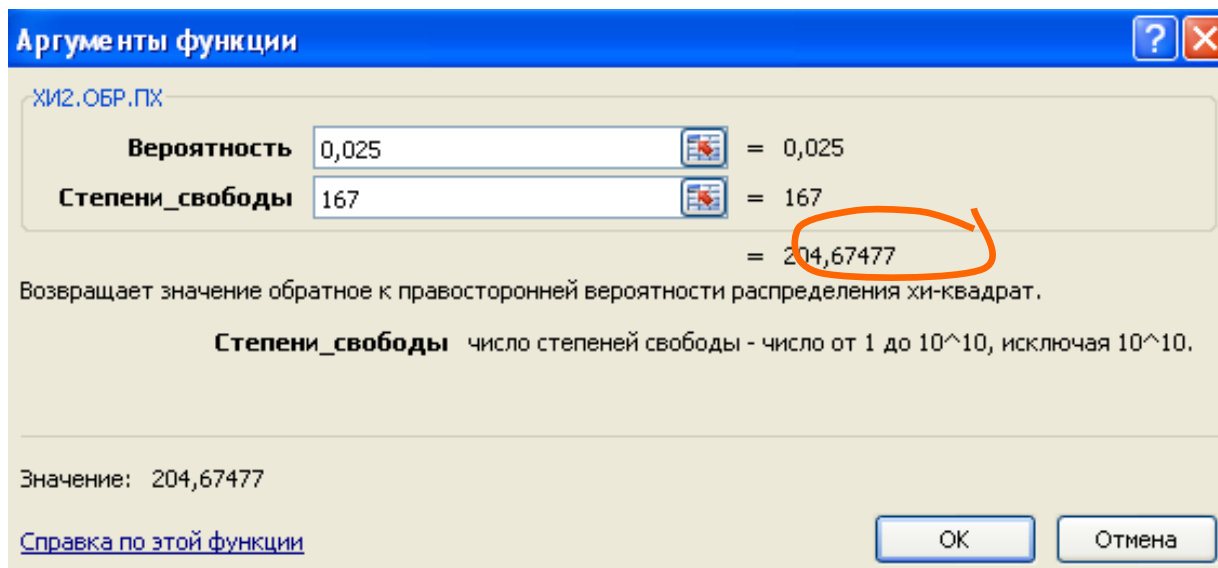


Рис. 5.9. Діалогове вікно функції **ХИ2.ОБР.ПХ**

У цьому ж діалоговому вікні отримуємо результат застосування цієї функції, який можна вивести у вільну комірку робочого аркуша, натиснувши **Enter**.

Отримуємо результат: $\chi_1^2 = 204,67$.

Аналогічно обчислюємо $\chi_2^2 = 133,11$.

Звідси для дисперсії отримуємо довірчий інтервал:

$$67,65^2 \cdot \frac{167}{204,67} < D(X) < 67,65^2 \cdot \frac{167}{133,11}.$$

Після обчислень для середньоквадратичного відхилення маємо:

$$61,11 < \sigma < 75,77.$$

5.5. Завдання для самостійної роботи

У табл. 5.6 наведені результати вимірювання випадкової величини X .

За допомогою вбудованих функцій та надбудов *MS Excel* необхідно виконати такі завдання:

визначити емпіричний закон розподілу випадкової величини X та побудувати гістограму частот та кумуляту;

обчислити статистичні оцінки числових характеристик випадкової величини X ;

знайти довірчий інтервал, до якого з надійністю 95 % належить математичне сподівання теоретичного закону розподілу.

Таблиця 5.6

Вихідні дані

№ варіанта	Значення, які приймає випадкова величина у вибірковій сукупності																																																																																																																	
1	2																																																																																																																	
5.1.	5,5	5,9	5,8	5,7	6,0	5,5	5,0	6,5	5,7	5,6	6,2	5,4	5,4	5,0	5,7	5,5	6,2	5,4	5,6	4,2	5,6	5,4	5,3	4,7	4,5	5,0	5,1	5,5	5,6	5,1	6,3	5,4	5,9	5,6	6,4	5,7	5,0	5,9	5,7	6,0	6,7	4,1	6,2	5,8	5,5	6,2	6,2	5,8	5,8	5,6	5,5	4,7	5,8	4,9	6,1	6,2	6,0	4,3	5,9	6,9	4,5	6,8	5,6	5,3	4,7	5,5	5,0	5,3	5,2	4,8	5,5	5,9	6,5	5,2	5,7	6,0	5,4	5,8	5,7	5,6	4,8	5,8	6,2	5,3	4,5	5,8	5,6	5,4	6,4	6,0	6,1	5,9	6,4	6,3	5,4	5,8	6,1	6,3	5,7	6,2	5,8	5,5	4,7	4,5	5,0	5,1	4,3	5,5	4,7	5,8	5,7	5,0	4,1	6,5.
5.2.	4,5	6,1	5,7	5,3	5,8	5,5	5,7	5,5	5,7	5,9	5,7	4,6	5,1	6,0	6,4	6,4	5,1	6,5	4,7	5,6	6,3	6,5	5,0	5,4	5,8	6,3	4,9	5,3	5,6	6,0	5,7	5,4	6,0	5,5	6,4	5,7	4,8	5,5	5,7	6,6	5,0	5,8	7,0	5,9	4,7	6,2	5,9	6,2	5,0	5,5	5,2	5,1	5,8	6,9	5,7	5,8	6,2	5,3	4,3	5,9	5,8	5,7	5,3	4,3	6,7	6,4	5,6	6,1	5,9	5,2	5,0	4,9	5,6	6,0	5,8	5,2	5,9	5,5	6,0	5,3	5,9	6,3	6,8	5,2	5,4	6,5	6,8	6,5	5,4	5,6	5,8	6,3	7,3	4,7	7,1	6,3	5,6	4,2	5,6	5,5	5,6	6,3	6,5	5,0	5,4	5,8	6,3	4,9	5,5	6,7	5,6	6,1	5,9	5,1.
5.3.	4,0	5,8	5,4	4,9	4,2	4,7	4,5	6,4	7,3	5,5	6,7	5,9	4,0	6,8	5,0	4,7	6,5	6,0	6,3	6,5	6,0	4,6	7,9	4,3	7,4	7,1	7,5	4,6	6,8	7,3	6,9	5,4	5,7	4,3	7,5	4,4	4,5	4,9	5,0	5,2	5,5	5,9	4,5	5,4	4,5	6,5	7,3	6,1	7,3	5,7	6,2	6,6	7,1	7,5	3,9	6,7	6,6	7,0	6,8	6,4	5,4	4,4	4,8	3,9	5,3	5,5	5,8	6,6	5,6	5,0	4,8	4,7	7,2	5,1	6,5	5,3	5,2	4,2	7,0	6,4	3,9	4,4	6,2	4,1	6,4	5,0	6,7	6,0	5,4	5,6	5,8	6,3	4,5	4,9	4,0	4,1	7,5	4,7	4,5	6,4	7,3	5,5	6,7	5,9	4,0	6,1	7,1	7,5	3,9	6,7	5,3	8,0	5,1	3,8.
5.4.	3,5	3,7	3,8	3,7	3,6	3,0	3,5	3,3	3,6	4,3	3,1	3,6	4,0	3,5	4,0	3,6	3,9	3,8	4,2	3,5	3,7	3,8	3,7	3,6	3,5	4,3	3,8	3,5	3,7	3,9	4,9	3,5	4,3	3,8	4,5	3,9	3,5	3,5	4,1	3,8	3,5	4,3	3,7	3,8	3,5	3,7	3,9	3,5	4,5	3,6	3,8	4,7	3,7	4,7	3,9	3,5	3,8	3,5	4,0	3,6	3,9	3,5	4,2	3,7	3,9	3,7	4,0	3,7	3,5	4,8	3,5	4,2	3,5	3,8	3,5	4,4	3,9	4,6	3,6	3,5	3,9	3,8	3,6	3,7	3,9	3,6	3,5	4,0	3,5	3,9	3,5	3,8	3,5	3,8	4,4	3,6	3,5	3,8	3,5	3,7	3,9	3,5	4,5	3,6	3,8	4,7	3,7	3,5	4,0	4,6	4,1	3,8	3,5	3,2.

1	2
5.5.	5,7 3,8 5,8 6,9 6,0 7,9 6,0 3,9 4,9 5,6 4,9 5,9 5,1 4,0 6,5 6,8 6,5 7,3 6,0 5,1 4,4 5,6 6,1 5,4 6,0 3,9 5,4 4,0 3,7 4,5 5,0 7,9 7,3 6,4 7,0 5,6 6,4 4,5 7,0 4,2 4,3 6,5 6,1 7,8 5,4 3,7 5,5 3,5 6,8 6,5 4,7 6,5 7,8 7,2 6,8 7,9 4,8 4,0 4,3 4,4 7,0 4,5 6,0 5,5 4,6 6,1 5,0 4,0 4,3 6,9 7,4 4,9 7,9 5,5 4,5 3,9 4,0 7,4 7,9 3,9 5,4 6,2 5,9 5,6 6,5 5,2 6,5 6,8 7,0 3,6 6,2 5,2 4,4 3,5 4,5 6,9 7,3 6,1 3,5 5,2 5,9 4,3 4,8 7,1 4,5 6,0 5,5 4,6 6,1 5,2 4,0 4,3 6,9 7,4.
5.6.	7,0 5,4 4,5 6,5 6,1 6,6 5,6 7,3 6,7 4,2 4,9 3,4 4,8 5,9 5,3 5,2 5,3 5,4 4,5 6,0 5,1 7,0 6,0 6,4 5,5 5,8 6,9 5,6 5,2 6,5 6,7 4,2 5,0 4,5 5,0 4,0 5,4 5,9 4,4 5,1 5,3 5,5 5,9 6,2 5,3 4,0 5,5 4,7 5,0 5,7 6,4 5,8 4,6 4,5 6,0 5,3 4,6 5,4 6,0 7,5 5,4 6,3 4,6 6,8 7,5 4,9 6,7 4,5 5,1 4,9 4,0 5,9 6,8 4,1 6,5 5,8 4,6 3,8 6,5 5,8 5,3 3,7 6,0 6,4 3,9 6,0 5,0 6,4 5,5 5,8 6,2 6,3 5,2 5,9 5,8 5,7 5,4 5,1 6,8 7,5 4,9 6,7 4,5 5,1 6,1 4,8 7,1 6,8.
5.7.	3,2 6,7 5,8 5,3 5,7 5,4 4,7 5,5 5,3 6,7 5,5 5,4 4,3 6,5 5,7 6,5 5,0 7,0 4,7 5,0 5,5 6,5 7,3 7,9 4,9 5,4 3,6 5,4 5,7 5,2 7,0 4,5 5,5 5,8 4,9 5,7 5,5 4,8 5,5 6,0 7,0 6,4 5,1 6,8 5,8 5,3 4,5 7,8 5,6 4,4 3,1 6,3 4,4 6,3 4,9 6,5 4,7 5,0 4,6 5,4 4,9 5,4 5,2 6,5 5,3 6,0 5,3 5,6 5,8 4,8 6,7 5,9 5,7 6,3 5,8 6,3 4,4 6,4 4,9 6,1 5,9 5,4 5,2 4,7 5,4 5,2 4,3 3,7 5,4 4,7 5,4 6,1 5,3 6,5 6,9 5,0 6,1 4,7 7,2. 5,4 4,9 6,1 5,9 5,4 5,2 4,7 5,4 6,1.
5.8.	3,0 6,9 5,1 4,9 5,3 7,0 3,4 5,0 5,1 6,4 3,9 7,2 5,5 4,3 5,4 4,9 3,9 4,8 5,8 4,2 6,2 4,7 5,4 5,6 6,6 4,1 5,5 5,8 6,0 5,9 4,5 6,4 5,0 6,5 4,5 5,5 5,0 3,7 5,5 6,5 6,8 3,9 5,8 6,3 5,9 6,1 5,5 5,8 4,0 4,5 7,9 3,8 4,0 5,3 4,3 5,4 5,9 4,6 5,9 6,4 4,7 6,4 4,0 6,0 5,5 5,7 5,0 6,0 3,7 5,1 5,0 4,8 5,4 5,3 6,3 4,9 6,4 5,6 5,4 4,9 5,5 4,1 4,7 5,3 4,5 5,5 6,0 4,5 5,0 4,4 4,3 5,4 5,0 5,3 5,8 4,7 5,4 5,6 6,6 4,1 5,5 5,8 4,6 5,0 5,4 5,1 4,7 5,2.
5.9.	3,5 3,7 4,4 4,2 4,8 3,9 4,0 5,1 4,3 4,0 6,8 7,4 4,2 3,5 6,3 3,8 4,2 3,6 5,5 3,8 3,8 3,5 3,9 4,2 3,8 3,9 5,9 4,4 3,9 4,9 4,5 3,8 4,8 3,7 4,8 3,4 3,6 3,9 3,8 6,4 5,0 3,8 5,5 4,5 4,2 3,5 4,0 3,5 6,5 5,8 4,3 3,7 4,0 3,9 4,1 4,4 3,7 4,4 3,5 3,8 3,7 3,9 4,0 4,5 7,9 3,8 4,0 5,3 4,6 5,9 3,9 4,7 3,4 4,0 3,6 3,8 3,7 5,0 3,5 6,0 3,9 4,1 5,7 5,1 3,7 5,0 5,4 5,3 4,5 3,8 4,5 5,0 3,8 5,5 4,5 3,9 4,5 5,0 3,9 4,3 3,8 5,0 4,1 4,7 5,0 4,1 3,7 4,1.
5.10.	6,0 5,4 5,0 6,0 3,6 5,4 6,2 4,0 5,4 5,6 5,2 5,0 6,4 5,0 4,2 5,0 5,3 5,2 4,0 4,4 5,3 5,3 5,0 5,8 5,6 5,8 5,2 5,4 5,8 5,7 6,1 5,4 5,3 5,0 4,5 4,3 7,1 5,0 7,4 5,6 6,0 4,7 5,4 4,7 5,8 6,0 6,2 6,8 4,9 5,0 6,6 5,5 5,9 6,5 4,5 5,4 6,9 5,6 4,5 6,5 4,4 4,9 6,1 4,9 5,9 5,4 4,3 5,6 5,4 4,9 5,1 5,6 4,4 4,8 5,2 5,6 5,6 5,4 4,7 5,2 4,9 5,4 5,8 5,4 4,9 5,4 5,9 6,0 4,9 7,2 4,8 5,8 4,9 6,1 4,9 5,9 5,4 4,3 5,6 5,5 7,4 5,8 5,4 4,7 4,8 5,4 5,0 6,3.

1	2
5.11.	3,9 4,4 3,8 3,9 5,8 4,2 4,3 3,5 4,8 3,5 5,1 3,7 3,6 4,0 3,6 5,0 4,0 3,5 3,5 4,0 3,5 3,4 3,4 3,8 6,8 4,8 3,9 4,5 4,4 3,5 4,8 3,6 4,2 3,5 3,5 3,9 5,1 3,7 3,6 3,9 4,0 3,6 3,7 3,8 3,4 3,8 4,8 3,9 5,0 4,0 3,5 3,5 4,0 4,5 4,4 3,5 4,8 3,6 3,5 4,2 3,8 3,8 3,7 3,5 4,5 7,3 4,6 3,5 4,8 4,0 3,2 5,0 3,7 5,8 4,6 3,8 6,2 3,8 3,6 3,8 4,0 3,1 3,4 3,8 5,5 3,7 4,3 4,5 4,3 3,5 4,0 3,8 4,8 3,9 5,0 4,0 3,5 3,5 4,0 3,6 3,2 4,0 3,4 3,9 3,2 3,8 7,0 6,7.
5.12.	5,4 4,5 4,9 4,2 4,0 4,7 4,5 4,1 4,5 4,6 4,0 4,6 4,5 4,2 3,5 3,9 4,0 4,7 4,2 3,9 5,6 3,8 4,5 3,9 3,8 5,7 6,2 3,6 3,4 3,7 4,2 3,7 4,4 3,4 3,8 3,9 3,6 3,5 3,9 5,1 3,7 4,0 3,9 4,4 4,8 4,6 5,0 4,0 3,5 4,2 5,7 4,2 4,8 3,6 3,4 3,5 3,7 3,5 3,7 3,9 3,8 3,7 4,0 3,5 3,8 3,7 4,3 3,8 3,5 4,5 3,6 4,2 5,5 4,2 3,5 4,3 4,7 3,5 5,8 4,2 3,6 4,6 4,2 3,6 4,4 5,9 3,7 4,3 4,5 3,6 6,5 3,5 3,8 3,5 4,6 4,2 4,0 5,3 3,7 6,0 4,4 3,6 3,5 3,9 5,1 3,7 4,0 3,9.
5.13.	6,1 6,6 6,2 6,3 6,1 6,4 5,9 6,2 6,3 5,7 6,2 6,0 6,4 6,2 6,3 6,4 6,1 6,6 6,0 6,2 6,3 6,0 6,5 6,1 5,9 6,2 6,3 6,1 6,0 6,1 6,2 6,1 6,3 6,1 6,0 6,2 6,8 6,2 6,3 5,7 6,0 5,8 6,3 5,3 6,2 5,5 6,4 6,1 6,0 6,2 6,3 6,0 6,0 6,1 6,0 6,2 6,3 6,2 6,1 6,5 6,1 6,0 6,2 6,2 6,1 6,0 6,2 6,4 6,1 6,0 6,2 6,3 5,9 6,2 6,2 6,2 5,8 6,0 5,8 6,3 6,1 6,0 6,2 6,3 6,2 6,1 5,6 6,0 6,2 6,3 6,7 6,2 6,1 5,9 6,2 6,4 5,9 5,8 5,9 6,1 6,7.
5.14.	4,5 5,5 4,7 5,4 5,5 6,0 4,8 6,8 5,3 5,6 5,7 5,8 7,5 7,1 5,2 5,7 6,3 6,7 5,1 5,7 6,4 6,6 5,0 5,0 5,7 5,8 6,3 6,4 5,8 4,9 5,4 6,7 5,9 6,4 6,2 6,4 5,7 5,5 5,7 4,6 5,5 6,0 6,4 6,4 5,1 6,5 4,7 6,3 6,3 6,5 6,8 5,0 4,9 3,2 7,2 4,5 6,0 5,7 6,3 5,5 5,4 6,7 4,3 6,5 6,3 6,1 5,7 7,7 6,5 5,0 7,0 4,7 5,0 5,5 4,3 6,3 7,9 5,7 5,2 7,0 6,8 5,6 5,8 4,0 6,0 5,5 5,7 6,0 5,5 6,0 7,3 6,4 5,4 6,8 6,5 5,6 5,7 5,8 7,5 7,1 5,2 5,7 6,3 5,0 6,1 7,0 4,7 5,8.
5.15.	6,5 4,4 4,9 5,4 5,7 6,0 6,5 4,8 3,6 5,4 5,3 4,0 6,3 6,0 5,3 5,0 5,7 5,6 5,0 6,7 5,1 6,3 7,9 6,5 6,0 4,1 6,1 4,5 5,1 5,6 6,4 5,6 6,8 4,9 5,0 5,6 4,7 5,4 5,2 5,7 5,8 5,4 4,9 5,6 5,0 4,9 4,5 5,4 5,6 5,8 5,9 4,5 4,0 5,4 5,0 6,5 4,7 5,0 4,4 5,5 6,5 5,4 6,0 4,8 7,5 6,0 4,6 5,6 7,0 5,4 4,1 5,7 4,5 7,3 5,5 5,0 7,4 5,4 7,0 5,4 4,0 5,4 6,0 4,7 5,0 5,4 5,3 6,1 7,2 5,5 5,1 5,4 4,8 3,6 5,4 4,0 6,3 6,0 5,3 6,1 5,7 6,2 4,7 5,6 4,5 5,5 5,7 6,8.
5.16.	6,2 7,2 8,1 6,4 7,8 6,1 6,2 8,8 6,2 6,2 6,3 6,2 6,8 7,7 6,7 6,2 6,0 6,0 6,1 6,2 6,1 6,5 6,1 6,3 6,1 6,0 6,1 6,7 6,1 6,4 5,9 6,6 6,1 7,2 6,2 7,8 7,2 6,4 5,9 6,7 6,2 6,7 7,2 6,2 6,4 7,8 5,9 7,2 6,2 6,2 7,2 6,2 6,4 6,2 8,2 6,3 7,4 6,3 6,4 6,3 6,1 7,4 6,3 7,4 6,3 7,1 6,0 6,4 6,3 6,4 6,3 6,4 6,5 7,5 6,5 6,8 6,1 6,0 6,5 6,6 6,2 6,2 6,7 7,6 6,4 6,8 6,1 6,1 6,8 6,9 6,1 7,2 6,2 6,4 7,8 5,9 7,2 7,5 6,2 7,2 6,2 8,3 7,1 6,1 7,1 6,5 6,6 8,0.

1	2
5.17.	3,4 3,1 3,2 3,1 4,7 3,2 3,1 2,9 3,0 2,5 3,7 3,2 3,1 2,8 4,0 3,1 3,2 2,5 2,8 4,6 4,5 3,1 3,2 2,4 3,4 2,3 3,0 4,7 2,6 2,9 4,7 2,6 3,0 4,9 2,9 3,0 2,9 3,9 4,0 3,1 3,2 3,1 3,2 4,2 3,1 5,3 4,6 3,2 3,6 3,2 3,7 3,2 3,5 3,2 3,5 3,4 3,2 3,8 4,4 3,4 3,2 3,9 3,4 3,2 3,6 2,2 3,2 3,3 3,4 3,2 3,3 3,4 3,2 3,7 3,4 3,2 3,3 3,4 4,0 3,2 4,2 3,3 3,2 3,1 3,4 3,2 3,3 3,2 3,5 3,6 3,1 3,5 4,4 3,6 3,2 3,1 3,5 4,3 3,6 3,2 4,7 3,4 3,2 3,4 4,0 3,2 4,2 4,5.
5.18.	4,8 5,7 5,3 5,2 5,0 5,2 6,6 5,7 5,4 5,5 5,7 5,5 5,9 5,6 5,8 5,4 5,3 5,8 5,4 5,2 5,7 5,1 5,9 5,2 5,7 5,5 5,0 5,4 5,0 5,6 5,8 5,9 5,6 5,3 5,6 5,9 5,3 5,8 5,9 5,0 5,8 5,1 5,3 5,7 5,6 5,4 5,7 5,8 5,4 5,6 6,5 5,4 5,6 5,8 5,4 5,5 5,1 5,5 5,3 5,4 5,7 5,2 5,0 5,6 5,5 5,0 5,7 5,8 5,9 5,5 5,3 5,6 5,7 5,5 5,4 5,3 5,5 5,9 5,4 5,2 5,3 5,2 5,9 5,2 4,9 5,3 5,2 5,0 5,5 6,2 5,9 5,0 5,2 4,6 4,6 5,8 5,9 5,6 5,3 5,6 5,9 5,4 4,2 5,6 5,3 5,7 5,9 4,8.
5.19.	5,2 5,7 4,4 5,5 5,4 5,9 7,0 5,1 4,8 6,0 5,5 5,6 5,2 5,3 6,1 5,6 5,2 5,8 5,3 5,5 6,4 6,3 5,9 5,8 4,6 4,7 4,9 5,0 5,6 4,9 5,5 5,1 5,5 5,7 6,7 5,4 5,0 5,7 5,4 5,9 5,7 5,1 5,0 6,4 5,1 7,2 5,2 6,1 5,6 5,3 5,2 5,3 4,3 5,2 5,9 5,3 5,2 5,8 5,3 5,2 5,3 5,6 5,2 5,4 4,6 5,3 6,0 5,2 5,4 5,5 5,4 5,5 5,3 5,8 5,4 5,6 5,4 6,6 5,4 6,2 5,6 5,4 5,5 5,4 5,5 5,3 5,4 5,2 5,1 5,0 6,4 5,1 7,2 5,2 6,1 5,6 5,3 5,2 5,7 5,9 5,2 5,8 6,0 5,8 5,3 5,1 6,5 7,4.
5.20.	3,4 3,1 2,7 2,8 3,1 3,2 2,4 3,1 3,4 3,8 4,2 2,7 2,2 3,2 2,3 2,5 2,4 3,0 3,4 3,2 3,8 2,5 2,9 2,3 2,9 3,6 2,5 2,4 2,2 4,7 3,6 2,1 3,2 3,4 4,7 2,8 3,3 3,7 3,9 2,7 3,5 2,3 3,7 3,2 3,6 4,3 3,3 3,7 2,2 3,2 1,5 2,2 2,3 1,8 3,6 3,3 3,7 3,3 2,7 2,8 3,8 3,0 4,2 1,8 3,2 2,3 3,9 3,3 3,6 4,4 2,5 1,7 3,6 2,7 3,9 3,2 3,9 2,0 3,0 3,4 4,5 2,4 5,2 2,8 2,5 2,8 1,9 2,8 2,2 2,8 3,0 3,8 2,7 1,9 3,6 2,8 3,1 2,8 2,9 2,7 3,5 2,3 3,7 3,2 3,6 4,3 3,3 3,9.
5.21.	2,0 2,8 2,0 2,3 3,0 2,0 2,3 2,2 3,1 2,4 2,5 2,7 2,5 2,4 2,6 2,7 2,4 2,2 2,1 2,0 2,8 2,0 1,7 2,9 2,0 2,4 2,9 3,3 2,4 2,5 2,5 2,4 2,6 2,3 2,1 2,7 2,1 2,3 2,0 2,9 2,5 2,0 2,7 2,0 2,1 2,2 2,7 2,9 2,0 3,2 2,2 2,0 2,1 2,4 2,5 2,7 2,1 2,6 2,4 2,1 2,0 2,4 2,1 2,4 2,3 2,0 2,3 2,0 1,8 2,1 3,0 2,2 2,7 2,0 1,6 2,1 2,3 2,0 1,9 2,4 2,1 2,8 2,3 2,0 3,0 2,5 2,1 2,2 2,4 2,0 2,1 1,9 2,1 2,0 2,1 2,4 2,3 2,0 2,3 1,8 2,1 3,0 2,2 2,1 2,2 2,0 2,7 1,8.
5.22.	64 92 18 13 28 132 17 140 18 35 116 135 20 53 90 40 85 15 18 42 83 38 36 24 12 170 92 132 14 68 115 185 89 19 73 50 18 37 11 67 50 20 12 82 31 129 10 242 32 19 128 195 10 40 44 25 76 82 15 14 35 92 25 22 21 364 14 128 85 19 109 141 83 20 80 92 85 13 36 29 12 20 18 26 22 132 89 240 36 84 116 202 172 57 19 22 21 81 127 12 26 35 30 145 98 164 67 132 210 148 192 210 184 140 132 94 78 34.

1	2
5.23.	6,8 4,4 4,9 5,4 5,7 6,9 6,5 4,8 3,6 5,4 5,3 4,0 6,3 6,0 5,3 5,0 5,7 5,6 5,0 5,7 5,1 6,3 7,9 6,5 6,0 4,1 6,1 4,5 5,1 5,6 6,4 5,6 6,8 4,9 5,0 5,6 4,7 5,4 4,2 5,7 5,8 6,4 4,9 5,6 5,0 4,9 4,5 5,4 5,7 5,8 5,9 4,5 4,0 5,4 5,0 6,5 4,3 5,6 4,4 5,5 6,5 5,4 6,4 4,8 7,5 6,8 4,6 5,6 7,0 4,4 5,1 5,7 4,5 7,3 5,5 5,0 7,1 6,4 7,0 5,4 4,0 5,4 6,0 4,7 6,0 5,4 5,3 6,1 7,2 5,5 5,1 5,4 4,8 3,6 5,4 4,2 5,3 6,0 5,3 6,1 5,7 6,2 4,7 5,6 4,5 5,5 5,7 6,8.
5.24.	4,9 5,0 5,8 5,1 5,3 5,7 5,6 5,4 5,7 5,8 5,4 5,6 6,5 5,4 5,6 5,8 6,4 5,5 5,1 5,5 5,3 5,9 5,7 5,2 5,0 5,6 5,5 5,0 5,7 5,8 5,9 5,5 5,3 5,6 5,7 5,5 5,4 4,3 4,8 5,7 6,3 5,2 5,5 5,2 6,6 5,7 5,9 5,5 5,7 5,5 5,9 5,6 5,8 5,4 5,3 5,8 5,4 6,2 5,7 5,1 5,9 5,8 5,7 5,5 5,0 4,4 5,0 5,6 5,8 5,9 5,6 5,3 5,6 4,9 5,3 5,8 4,5 5,9 6,4 5,2 5,3 6,2 5,9 5,2 4,9 5,3 5,2 4,6 4,7 6,6 5,9 5,0 5,2 4,6 4,9 4,8 5,9 5,6 5,3 5,6 4,9 5,4 4,2 5,6 6,3 5,7 5,9 5,8.
5.25.	6,2 5,4 5,0 6,2 3,6 5,4 6,2 4,0 5,4 5,6 5,2 5,0 6,4 5,0 4,2 5,0 5,3 5,2 4,0 5,8 5,2 5,8 5,0 5,8 4,6 5,8 5,2 5,4 5,8 5,7 6,1 5,4 5,3 5,0 4,5 4,3 7,1 5,0 4,4 5,6 4,0 7,7 5,4 4,7 5,8 6,0 6,2 6,8 4,9 5,0 6,6 5,5 5,9 6,5 4,5 5,4 6,9 5,3 4,5 6,5 4,8 4,9 6,5 4,2 5,9 5,4 4,3 5,6 5,4 4,9 5,1 5,6 5,8 4,8 5,2 5,6 6,1 5,4 4,7 5,2 4,9 5,4 5,8 5,4 4,9 5,4 5,9 6,0 4,9 7,2 4,8 5,9 4,9 5,1 4,9 5,9 5,8 5,3 5,6 5,5 7,4 5,8 5,4 6,7 4,8 5,4 7,0 4,3.

5.6. Контрольні запитання

1. Дайте означення генеральної та вибіркової сукупностей.
2. Як формується вибірка сукупність? Яким вимогам вона повинна відповідати?
3. Що таке варіанта? Дайте означення варіаційного ряду.
4. Дайте означення статистичного розподілу, визначіть способи його задавання та наведіть приклади.
5. Поясніть, чому числові характеристики статистичного розподілу мають сенс оцінок.
6. Дайте означення точкових оцінок основних числових характеристик розподілу випадкової величини в генеральній сукупності.
7. Визначте вимоги, яким повинні відповідати статистичні оцінки.
8. Що таке зсунуті та незсунуті статистичні оцінки? Навіщо вводиться таке поняття, як виправлена дисперсія?

9. Що таке інтервальні оцінки невідомих параметрів розподілу випадкової величини? Поясніть, чому виникає така задача.

10. Наведіть формули для визначення довірчого інтервалу для математичного сподівання та дисперсії випадкової величини, що розподілена за нормальним законом.

11. Дайте означення емпіричних моментів.

12. Наведіть приклади використання початкових та центральних моментів для оцінювання числових характеристик розподілу.

13. Дайте означення емпіричної функції розподілу, визначте основні властивості цієї функції.

14. Яку інформацію можна отримати за допомогою додаткових числових характеристик розподілу: моди, медіани, коефіцієнта варіації?

15. Як оцінити додаткові числові характеристики розподілу за допомогою гістограми або полігону та кумуляти?

16. Чим відрізняється визначення основних числових характеристик розподілу за незгрупованими та згрупованими емпіричними даними?

17. Дайте означення дискретного та інтервального варіаційних рядів. Наведіть приклад, коли використання інтервального варіаційного ряду є доцільним для надання розподілу випадкової величини.

18. Дайте визначення обсягу вибіркової сукупності, що з наперед заданою точністю та надійністю дозволить оцінити математичне сподівання випадкової величини, що розподілена за нормальним законом.

Лабораторна робота 6

Перевірка за критерієм Пірсона статистичної гіпотези щодо нормального закону розподілу випадкової величини в генеральній сукупності

6.1. Мета роботи:

ознайомитись з принципами перевірки статистичних гіпотез;

навчитися перевіряти за допомогою критерію узгодження Пірсона нульову гіпотезу щодо відповідності закону розподілу випадкової величини у генеральній сукупності нормальному закону;

використовувати такі вбудовані функції *MS Excel*, як **НОРМ.РАСП**, **ХИ2.ОБР.ПХ**.

6.2. Теоретичні положення

Під гіпотезою розуміють певне наукове припущення про властивості досліджуваних явищ, яке потребує перевірки. **Статистичною гіпотезою** називається припущення відносно параметрів або форми розподілу генеральної сукупності, яке перевіряється за даними спостереження, тобто вибіркової сукупності. У ході перевірки статистичної гіпотези необхідно встановити, чи узгоджуються дані спостереження з цією гіпотезою, тобто чи можна відмінності між теоретичними припущеннями і результатами спостережень віднести до випадковості, чи ці відмінності обумовлені впливом певних причин, які діють систематично. У результаті перевірки статистична гіпотеза або приймається, або відхиляється.

Найбільш часто перевіряють припущення про те, що отримана за вибіркою величина несуттєво відрізняється від гіпотетичної (теоретично можливої) або встановленої величини в генеральній сукупності. Для перевірки цього положення висувається гіпотеза про те, що різниця між фактичним і гіпотетичним показниками дорівнює нулю. У зв'язку з цим гіпотезу, що перевіряється, називають **нульовою** і позначають H_0 .

Нульову гіпотезу ще називають **основною гіпотезою**. Підкреслимо, що саме вона і підлягає перевірці. У кожному випадку нульовій гіпотезі протиставляється **альтернативна (конкурентна) гіпотеза**, яка заперечує нульову гіпотезу. По відношенню до нульової гіпотези можна побудувати декілька альтернативних гіпотез, тому вони позначаються H_i ($i = \overline{1, n}$), де i – номер гіпотези. За формою побудови гіпотези можуть бути простими й складними.

Простою називають гіпотезу, яка містить тільки одне припущення, **складною** – гіпотезу, яка містить два і більше припущення.

Результатом перевірки статистичної гіпотези є висновок щодо прийнятності нульової гіпотези. У результаті перевірки нульова гіпотеза або приймається, або відхиляється, і тоді приймається альтернативна гіпотеза. Отже, перевірка статистичної гіпотези пов'язана з ризиком прийняти помилкове рішення.

Розрізняють помилки двох родів. **Помилка першого роду** полягає в тому, що правильну нульову гіпотезу помилково відхиляють. **Помилка другого роду** полягає в тому, що приймають нульову гіпотезу, хоча вона є хибною, а правильною є альтернативна гіпотеза.

Правильні або неправильні рішення можуть бути отримані у двох випадках (табл. 6.1).

Таблиця 6.1

Можливі результати перевірки нульової гіпотези

Результат перевірки H_0	Гіпотеза, яка є правильною	
	H_0	H_1
H_0 відхиляється	Помилка першого роду	Правильне рішення
H_0 приймається	Правильне рішення	Помилка другого роду

Імовірність зробити помилку першого роду (тобто невиправдане відхилення H_0) визначається як **рівень значущості** і позначається α . Імовірність зробити помилку другого роду (прийняття неправильної гіпотези) позначається β . Якщо кількість вибірок велика, то частка хибних висновків дорівнює α (якщо правильна H_0) або β (якщо правильна H_1).

Рівень значущості встановлюється самим дослідником залежно від характеру і важливості завдань, що розв'язуються (за так званим принципом практичної впевненості). Рівень значущості становить ту мінімальну ймовірність, починаючи з якої можна визнати подію практично неможливою. Користуються стандартними значеннями. Найбільш часто α встановлюють на рівні 0,05 або 0,01. Встановлюючи певний рівень значущості, дослідник контролює ймовірність помилки першого роду: чим він нижчий, тим частіше H_0 буде визнаватися правильною. Однак зниження рівня значущості веде до появи помилок другого роду. У більшості випадків єдиним шляхом одночасного зменшення ймовірності появи помилок обох родів є збільшення обсягу вибірки.

Для перевірки нульової гіпотези і прийняття висновку щодо її сумісності з емпіричними даними використовують **статистичні критерії**, які становлять зведення правил, за якими гіпотеза H_0 або приймається, або відхиляється. Для перевірки кожного виду гіпотез розроблені свої статистичні критерії, застосування яких обумовлено змістом задачі.

Слід підкреслити, що узгодженість емпіричних даних з нульовою гіпотезою за статистичним критерієм ще не доводить її справедливості. Тому в процесі формулювання остаточних висновків правильно говорити

про те, що дані спостереження не суперечать нульовій гіпотезі, отже, не дають підстав її відхилити.

На множині можливих значень статистичного критерію виділяють дві відокремлені підмножини, одна з яких містить допустимі значення критерію, а інша – недопустимі. Перша підмножина називається областю допустимих значень, а друга – критичною областю.

Критичною областю називають ті значення критерію, за яких нульова гіпотеза відхиляється.

Областю допустимих значень називають сукупність значень критерію за яких нульова гіпотеза приймається.

Отже, якщо фактичне значення критерію потрапляє в критичну область, то H_0 відхиляють, якщо ж фактичне значення критерію належить області допустимих значень, то H_0 нема підстав відхилити.

Точки, які відділяють критичну область від області допустимих значень, називаються **критичними точками розподілу**. Для кожного статистичного критерію існують спеціальні таблиці, за якими знаходять його критичні точки, і знайдене табличне значення критерію порівнюють з його емпіричним значенням.

Критерій перевірки гіпотези повинен бути підібраний так, щоб ризик допущення помилок був мінімальним. При цьому дуже важливо визначати ймовірність того, що не буде допущена помилка другого роду. Ця ймовірність характеризує чутливість критерію до помилок другого роду і отримала назву потужності критерію.

Потужністю критерію є ймовірність відхилення гіпотези H_0 , коли правильною є альтернативна гіпотеза H_1 , тобто величина $1 - \beta$.

Перевірка статистичної гіпотези передбачає послідовне виконання таких етапів:

1. Оцінка вихідної інформації і опис статистичної моделі вибіркової сукупності.
2. Формулювання нульової і альтернативної гіпотез.
3. Вибір найбільш потужного критерію для перевірки нульової гіпотези.
4. Встановлення рівня значущості, за допомогою якого контролюється помилка першого роду.
5. Розрахунок емпіричного значення статистичного критерію.
6. Визначення критичної області й області узгодження з нульовою гіпотезою, тобто встановлення гранично припустимого значення критерію.

7. Зіставлення емпіричного і критичного значень критерію і формулювання висновків щодо результатів перевірки нульової гіпотези.

Вибір конкретного методу перевірки гіпотези залежить від характеру досліджуваної гіпотези, властивостей вихідної інформації та інших умов.

Однією з найважливіших задач математичної статистики є визначення теоретичного закону розподілу випадкової величини, яка характеризує ознаку, що досліджується, за даними вибіркової сукупності.

Розв'язання цієї задачі передбачає визначення вигляду і параметрів закону розподілу. У загальному випадку закон розподілу випадкової величини в генеральній сукупності до початку досліджень є невідомим, але певні припущення відносно його характеру можна зробити за виглядом гістограми. Після визначення числових характеристик вибіркової сукупності обчислюють **вирівнюючі частоти** \hat{m}_i , які в загальному випадку відрізняються від емпіричних частот m_i , що отримані за результатами дослідження вибіркової сукупності. Згідно з нульовою гіпотезою $H_0: m_i = \hat{m}_i$, тобто відхилення емпіричних частот від вирівнюючих обумовлене випадковим розпорошенням. Оскільки вирівнюючі частоти визначаються, виходячи з певного закону розподілу, то перевірка нульової гіпотези є одночасно і перевіркою припущення відносно теоретичного закону розподілу. Перевірка гіпотези H_0 здійснюється за допомогою **критерію Пірсона**, або **критерій** χ^2 . Зручність його полягає в тому, що він може застосовуватись не тільки для перевірки гіпотези про відповідність нормальному закону розподілу, але й при інших законах розподілу в генеральній сукупності.

Критерій узгодженості Пірсона визначається за формулою:

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}, \quad (6.1)$$

де k – кількість інтервалів, що містить інтервальний варіаційний ряд статистичного розподілу випадкової величини.

Критичні значення статистичного критерію визначаються за довідковою таблицею. Критерій узгодженості Пірсона є випадковою

величиною, що має розподіл $\chi^2 = \chi^2(\alpha; s)$, і залежить від рівня значущості α і кількості ступенів свободи:

$$s = k - r - 1, \quad (6.2)$$

де r – кількість параметрів, якими визначається закон розподілу ймовірностей випадкової величини у генеральній сукупності згідно з нульовою гіпотезою.

Для нормального закону розподілу $r=2$, оскільки цей закон визначається двома параметрами: a – математичне сподівання та σ – середнє квадратичне відхилення. При визначенні критичної області критерію Пірсона приймають два рівня значущості: $\alpha = 0,05$ та $\alpha = 0,01$.

Якщо емпіричне значення χ^2 менше ніж критичне значення, яке визначалось при $\alpha = 0,05$ ($\chi_{0,05}^2$) для заданої кількості ступенів свободи, то нульову гіпотезу нема підстав відхилити. Отже, область $\chi^2 < \chi_{0,05}^2$ є довірчим інтервалом для критерію Пірсона. Його розширення призводить до збільшення ймовірності помилки другого роду.

Якщо ж емпіричне значення критерію χ^2 перевищує критичне, яке визначене при $\alpha = 0,01$ ($\chi_{0,01}^2$), то нульову гіпотезу треба відкинути на користь альтернативної.

У тому випадку, коли емпіричне значення критерію χ^2 належить до області невизначеності $\chi_{0,05}^2 < \chi^2 < \chi_{0,01}^2$, то необхідні додаткові дослідження. У процесі визначення емпіричного значення критерію χ^2 результати проміжних розрахунків зручно здійснювати в таблиці, що укладається за таким зразком (табл. 6.2).

Таблиця 6.2

Визначення вирівнюючих частот розподілу випадкової величини у припущенні нормального закону із застосуванням функції Гаусса

№	m_i	\bar{x}_i	t_i	$\varphi(t_i)$	\hat{f}_i	\hat{P}_i	\hat{m}_i	$\frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	2	3	4	5	6	7	8	9

Табл. 6.2 заповнюється таким чином:

стовпчик 1 – номер інтервалу інтервального варіаційного ряду;

стовпчик 2 – емпірична частота, що відповідає даному інтервалу (m_i);

стовпчик 3 – середина відповідного інтервалу (\bar{x}_i);

стовпчик 4 – значення стандартизованої випадкової величини t_i , що обчислюється за значенням випадкової величини, що відповідає

середині даного інтервалу: $t_i = \frac{\bar{x}_i - \bar{x}}{\sigma}$;

стовпчик 5 – значення функції Гаусса $\varphi(t_i)$ за аргументом t_i (обчислюються за довідковою таблицею або за допомогою функції **НОРМ.РАСП**);

стовпчик 6 – значення диференціальної функції розподілу

$$\hat{f}_i = \frac{\varphi(t_i)}{\sigma};$$

стовпчик 7 – ймовірності (\hat{P}_i) потрапляння випадкової величини до i -го інтервалу, яка обчислюється за формулою $\hat{P}_i = \hat{f}_i \cdot h_i$. Контрольна сума цього стовпчика не повинна відрізнятись від одиниці більш, ніж на 5 %;

стовпчик 8 – вирівнюючі частоти (\hat{m}_i), які обчислюються за формулою: $\hat{m}_i = \hat{P}_i \cdot n$;

стовпчик 9 – частка квадрата відхилення емпіричної частоти від вирівнюючої у відношенні до вирівнюючої. Емпіричне значення критерію Пірсона обчислюється як сума рядків цього стовпчика.

Існує ще один спосіб обчислення вирівнюючих частот, який передбачає використання інтегральної функції розподілу.

Для цього обчислення зручно організувати у вигляді таблиці, зразок якої наведено в табл. 6.3.

Табл. 6.3 заповнюється таким чином:

стовпчик 1 – номер інтервалу інтервального варіаційного ряду;

стовпчик 2 – емпірична частота, що відповідає даному інтервалу (m_i);

стовпчик 3 – ліва межа даного інтервалу (x_i).

Визначення вирівнюючих частот розподілу випадкової величини у припущенні нормального закону із застосуванням функції Лапласа

№	m_i	x_i	t_i	$\Phi(t_i)$	$\hat{F}(x_i)$	\hat{P}_i	\hat{m}_i	$\frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	2	3	4	5	6	7	8	9

Стовпчик 4 табл. 6.3 – значення стандартизованої випадкової величини t_i , що обчислюється за значенням випадкової величини, яке

співпадає з нижньою границею i -го інтервалу $t_i = \frac{x_i - \bar{x}}{\sigma}$;

стовпчик 5 – значення функції Лапласа $\Phi(t_i)$ за аргументом t_i (обчислюються за довідковою таблицею або за допомогою функції **НОРМ.РАСП**);

стовпчик 6 – значення функції розподілу $\hat{F}(x_i) = 0,5 + \Phi(t_i)$;

стовпчик 7 – ймовірності \hat{P}_i потрапляння випадкової величини до i -го інтервалу, яка обчислюється за формулою: $\hat{P}_i = \hat{F}_{i+1} - \hat{F}_i$;

стовпчик 8 – вирівнюючі частоти (\hat{m}_i), які обчислюються за формулою: $\hat{m}_i = \hat{P}_i \cdot n$;

стовпчик 9 – частка квадрата відхилення емпіричної частоти від вирівнюючої у відношенні до вирівнюючої.

6.3. Змістовна постановка задачі

Для наданого емпіричного розподілу, який задано інтервальним варіаційним рядом (табл. 6.4) необхідно:

висловити припущення щодо закону розподілу випадкової величини у генеральній сукупності;

за допомогою критерію χ^2 перевірити гіпотезу про рівність емпіричних частот і вирівнюючих частот, що обчислені у припущенні нормального закону розподілу на рівні значущості $\alpha = 0,05$.

Інтервальний варіаційний ряд

$[x_i; x_{i+1})$	[2,7; 3,3)	[3,3; 3,9)	[3,9; 4,5)	[4,5; 5,1)	[5,1; 5,7)	[5,7; 6,3)	[6,3; 6,9)	[6,9; 7,5)	[7,5; 8,1)
m_i	1	7	17	24	23	16	9	2	1

6.4. Приклад виконання лабораторної роботи 6

На робочому аркуші *MS Excel* за зразком табл. 6.2 записуємо вихідні дані і проводимо допоміжні обчислення, які необхідні для побудови гістограми й оцінювання основних числових характеристик емпіричного розподілу (рис. 6.1).

	A	B	C	D	E	F	G	H
1	№	$[x_i;$	$x_{i+1})$	x_i	m_i	$m_i/(n \cdot h)$	$x_i \cdot m_i$	$x_i^2 \cdot m_i$
2	1	2,7	3,3	3,0	1	0,0167	3,0	9
3	2	3,3	3,9	3,6	7	0,1167	25,2	90,72
4	3	3,9	4,5	4,2	17	0,2833	71,4	299,88
5	4	4,5	5,1	4,8	24	0,4000	115,2	552,96
6	5	5,1	5,7	5,4	23	0,3833	124,2	670,68
7	6	5,7	6,3	6	16	0,2667	96,0	576
8	7	6,3	6,9	6,6	9	0,1500	59,4	392,04
9	8	6,9	7,5	7,2	2	0,0333	14,4	103,68
10	9	7,5	8,1	7,8	1	0,0167	7,8	60,84
11	Сума				100		516,6	2755,8
12								

Рис. 6.1. Результати проміжних обчислень

Таблиця заповнюється таким чином:

стовпчик **A** – номер інтервалу інтервального варіаційного ряду;

стовпчик **B** – ліва межа x_i відповідного інтервалу;

стовпчик **C** – права межа x_{i+1} відповідного інтервалу;

стовпчик **D** – значення випадкової величини, що відповідає середині i -го інтервалу;

стовпчик **E** – емпірична частота, що відповідає даному інтервалу;

стовпчик **F** – висота стовпчиків гістограми, що відповідають даному інтервалу й обчислюється за формулою: $m_i / (n \cdot h)$;

стовпчик **G** – проміжні обчислення для визначення вибіркової середньої (значення у кожній комірці обчислюється за формулою $\bar{x}_i \cdot m_i$);

стовпчик **H** – проміжні обчислення для визначення вибіркової дисперсії (кожна комірка обчислюється за формулою $(\bar{x}_i)^2 \cdot m_i$).

За даними таблиці на рис. 6.1 обчислюємо вибіркoву середню за формулою для згрупованих даних:

$$\bar{x} = \frac{1}{\sum m_i} \cdot \sum x_i \cdot m_i.$$

Отже, маємо:

$$\bar{x} = \frac{1}{100} \cdot 516,6 = 5,166.$$

Виправлену дисперсію обчислюємо за формулою для згрупованих даних:

$$S^2 = \frac{n}{n-1} \left(\overline{x^2} - (\bar{x})^2 \right).$$

Отже, отримуємо:

$$S^2 = \frac{100}{100-1} \cdot \left(\frac{2\,755,8}{100} - 5,166^2 \right) = 0,8792.$$

Тепер обчислюємо значення виправленого середнього квадратичного відхилення: $S = 0,9377$.

За даними, що наведені у стовпчиках **D** та **F** (див. рис. 6.1) побудуємо гістограму відносних частот. Для цього виділяємо комірки **F2:F10**, що містять значення висот стовпчиків гістограми частостей, на **Ленте** переходимо до вкладки **Вставка**, викликаємо опцію **Гистограмма** і серед запропонованих типів гістограм обираємо **Гистограмма с группировкой** і натискаємо на її іконку, унаслідок чого отримуємо такий результат (рис. 6.2).

У цій гістограмі необхідно змінити підпис осі OX. Для цього правою клавішею миші заходимо на будь-який стовпчик гістограми і у діалоговому вікні, що з'явилось, натискаємо рядок **Выбрать данные**.

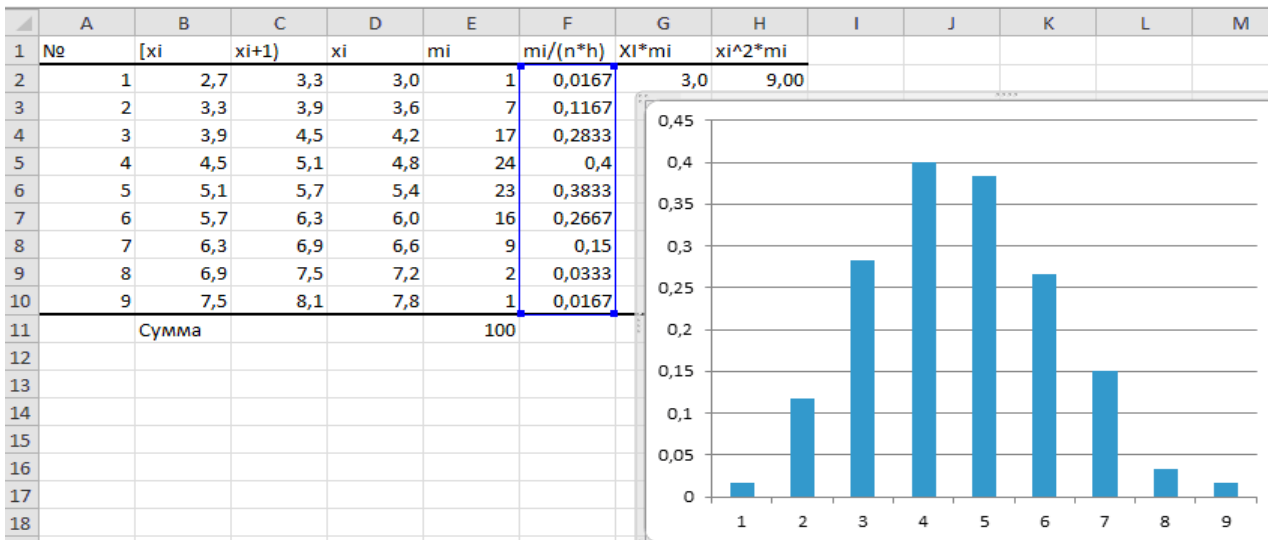


Рис. 6.2. Скриншот з результатами застосування опції *Гистограмма*

Відкривається діалогове вікно **Выбор источника данных**. У його лівій частині **Элементы легенды (ряды)** присутній тільки **Ряд 1** (рис. 6.3).

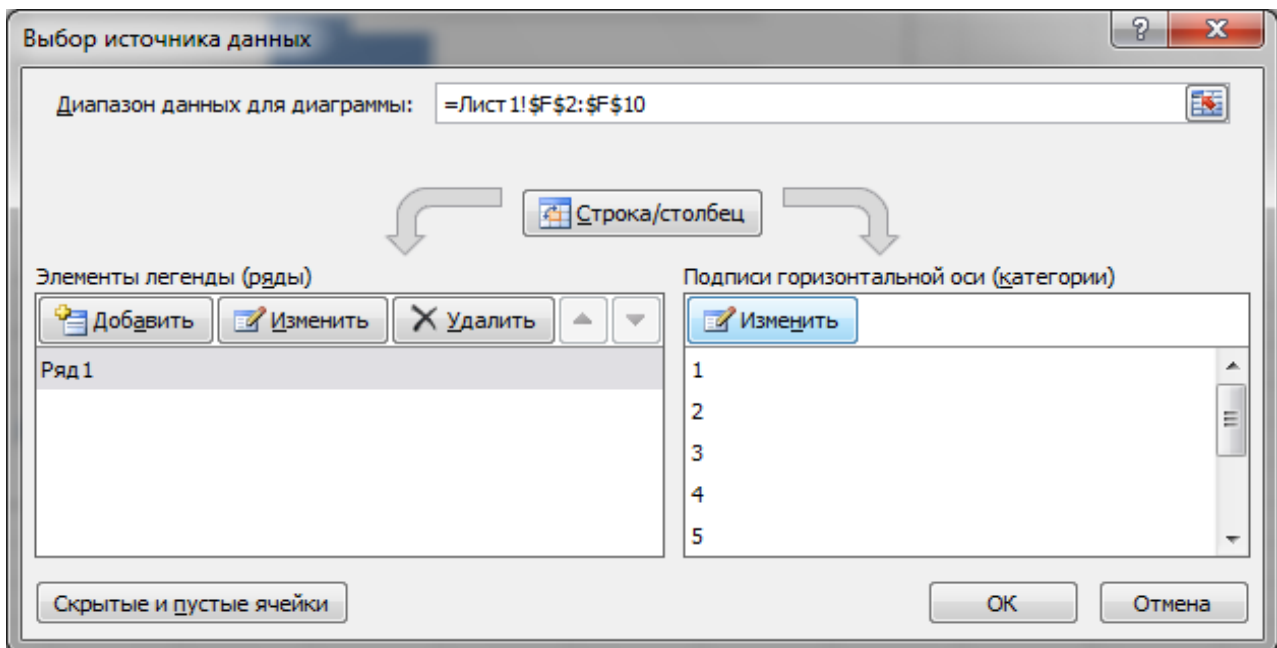


Рис. 6.3. Діалогове вікно *Выбор источника данных*

У правій частині цього вікна знаходимо **Подписи горизонтальной оси (категории)**. Для **Ряда 1** натискаємо **Изменить** і отримуємо діалогове вікно **Подписи оси**, у якому вводимо посилання на комірки **D2:D10** (рис. 6.4).

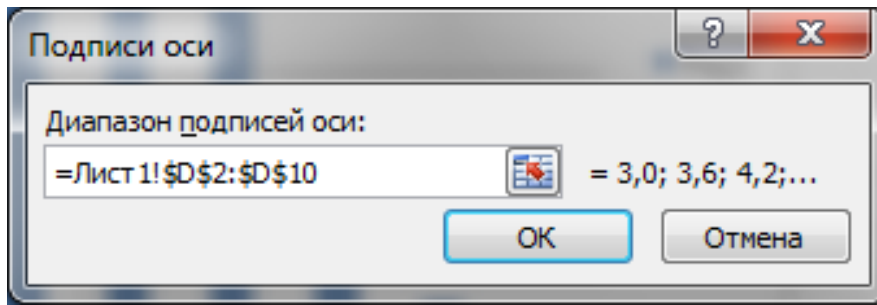


Рис. 6.4. Введення посилання на значення інтервальних середніх

Натисканням **ОК** повертаємося у діалогове вікно **Выбор источника данных** і знов натискаємо **ОК**. Залишається прибрати проміжки між стовпчиками гістограми. Для цього знов правою клавішею миші заходимо на будь-який стовпчик гістограми, вибираємо **Формат точки данных** і у **Параметрах ряда, Боковой зазор** вказуємо **Без зазора**. Отже, ми отримали гістограму відносних частот (рис. 6.5).

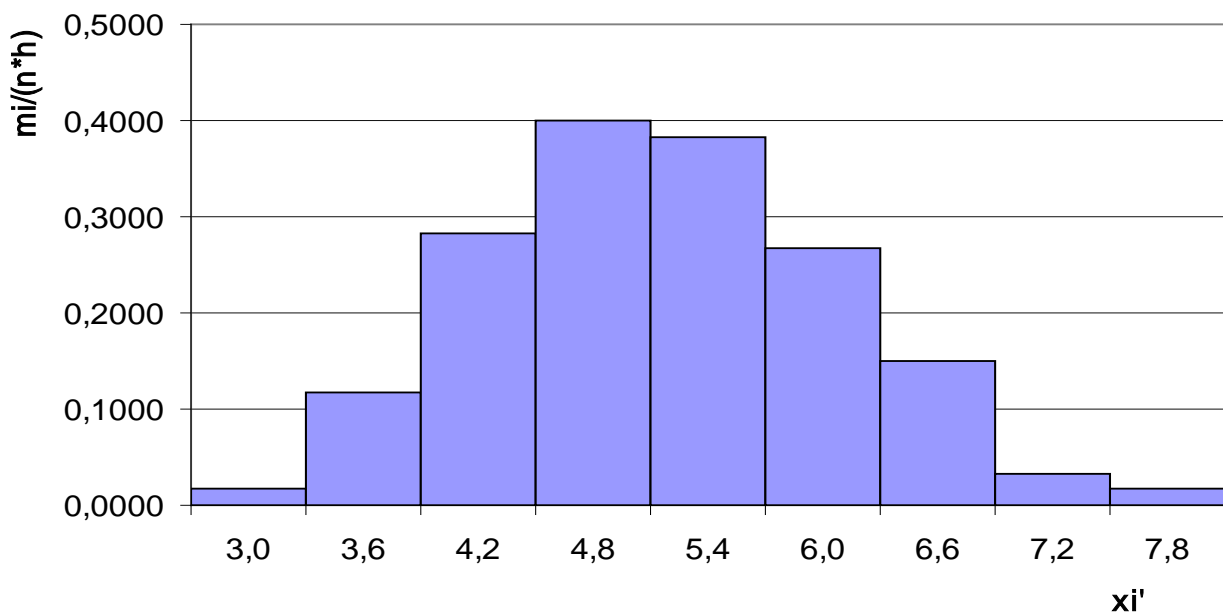


Рис. 6.5. Гістограма відносних частот

За загальним виглядом гістограми емпіричного розподілу можна зробити припущення, що розподіл випадкової величини у генеральній сукупності відповідає нормальному закону. Перевіряємо основну гіпотезу $H_0: m_i = \hat{m}_i$ за умови, що вирівнюючі частоти \hat{m}_i обчислюються в припущенні нормального закону розподілу у генеральній сукупності.

Альтернативною є гіпотеза $H_1: m_i \neq \hat{m}_i$.

Перевірка здійснюється за критерієм Пірсона. Для обчислення вирівнюючих частот скористуємося методом, що базується на побудові диференціальної функції розподілу (див. табл. 6.2). Для визначення емпіричного значення критерію згоди χ^2 обчислення будемо проводити у таблиці (табл. 6.5).

Таблиця 6.5

Визначення вирівнюючих частот розподілу випадкової величини у припущенні нормального закону

№	\bar{x}_i	m_i	t_i	$\varphi(t_i)$	\hat{f}_i	\hat{P}_i	\hat{m}_i	$\frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
1	3	1	-2,3100	0,0277	0,0295	0,0177	1,7714	0,3359
2	3,6	7	-1,6701	0,0989	0,1055	0,0633	6,3291	0,0711
3	4,2	17	-1,0302	0,2347	0,2503	0,1502	15,0157	0,2622
4	4,8	24	-0,3903	0,3697	0,3943	0,2366	23,6551	0,0050
5	5,4	23	0,2496	0,3867	0,4124	0,2474	24,7449	0,1230
6	6,0	16	0,8894	0,2686	0,2865	0,1719	17,1880	0,0821
7	6,6	9	1,5293	0,1239	0,1321	0,0793	7,9277	0,1451
8	7,2	2	2,1692	0,0379	0,0405	0,0243	2,4280	0,0754
9	7,8	1	2,8091	0,0077	0,0082	0,0049	0,4938	0,5190
Сума		100				0,9955	99,5536	1,6190

Значення функції Гаусса $\varphi(t_i)$ знаходимо за допомогою вбудованої функції **НОРМ.РАСП.**

Для цього у діалоговому вікні цієї функції (рис. 6.6) у рядку **X** вводимо посилання на комірку, де міститься відповідне значення t_i , вказуємо **Среднее**, яке для аргументу функції Гаусса дорівнює 0, **Стандартное_откл.**, що дорівнює 1, а у рядку **Интегральная** ставимо 0, оскільки ми визначаємо диференціальну функцію розподілу.

Слід зазначити, що застосування допоміжної змінної t_i для оцінювання диференціальної функції розподілу випадкової величини X обумовлено тим, що в довідкових таблицях наведені значення функції Гаусса. Отже, для того, щоб скористатись цими таблицями необхідно побудувати випадкову величину, для якої середня дорівнює 0, а середнє квадратичне відхилення має значення 1. Саме такою є стандартизована випадкова величина T .

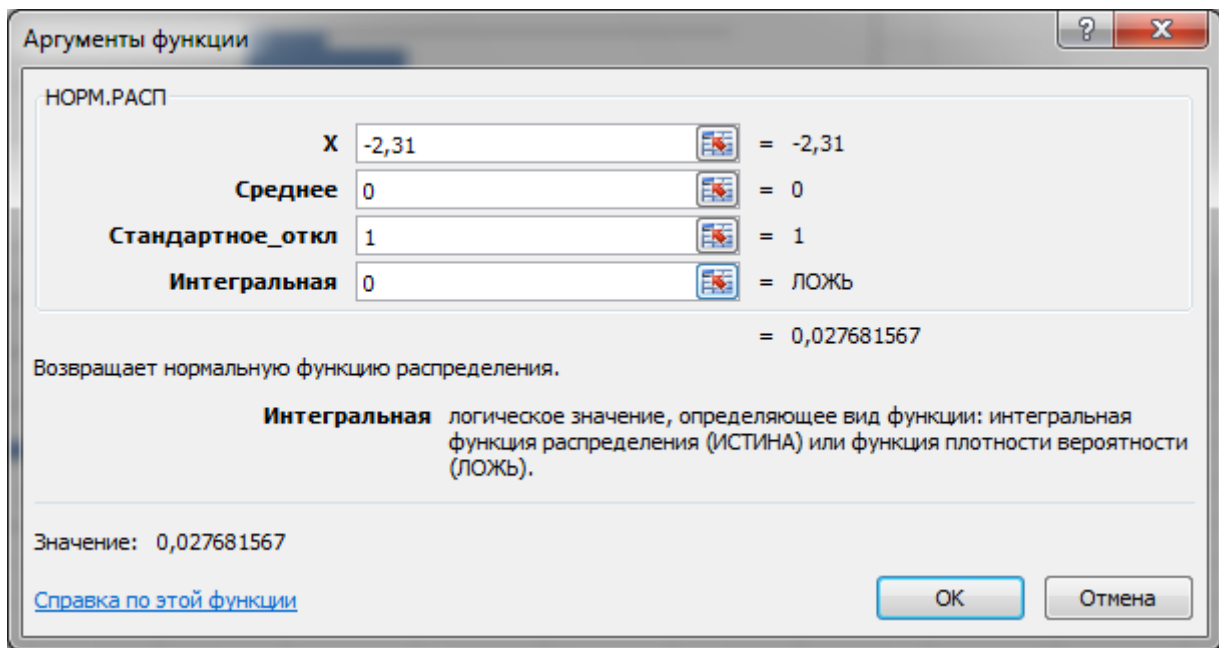


Рис. 6.6. Діалогове вікно функції НОРМ.РАСП

Якщо обчислення проводяться у програмному середовищі *MS Excel*, то застосування вбудованої функції **НОРМ.РАСП**, у діалогове вікно якої вводяться числові характеристики випадкової величини X , дозволяє безпосередньо обчислювати диференціальну функцію розподілу \hat{f}_i . Пропонуємо переконатись у цьому самостійно.

Порівняємо графік диференціальної функції \hat{f}_i , яка побудована у припущенні нормального закону розподілу, з гістограмою, що отримана за емпіричними даними. Для цього натискаємо правою клавiшею миші на будь-який стовпчик гістограми і в діалоговому вікні, що відкрилося, натискаємо **Выбрать данные**.

Відкривається діалогове вікно **Выбор источника данных**, у лівій частині якого присутній тільки **Ряд 1** (див. рис. 6.3). Натискаємо клавiшу **Добавить**, внаслідок чого з'являється діалогове вікно **Изменение ряда** і у його полі **Значения:** вводимо посилання на комірки, де міститься інформація про значення функції \hat{f}_i . Після натискання **OK** отримуємо дві гістограми на одному полі. Слід мати на увазі, що графік функції \hat{f}_i є плавною кривою, тому правою клавiшею миші заходимо на стовпчик нової гістограми і у діалоговому вікні обираємо **Изменить тип диаграммы для ряда...**, внаслідок чого відкривається вікно **Изменение типа диаграммы**. Змінюємо тип діаграми на **Точечная**.

Результат наведено на рис. 6.7.

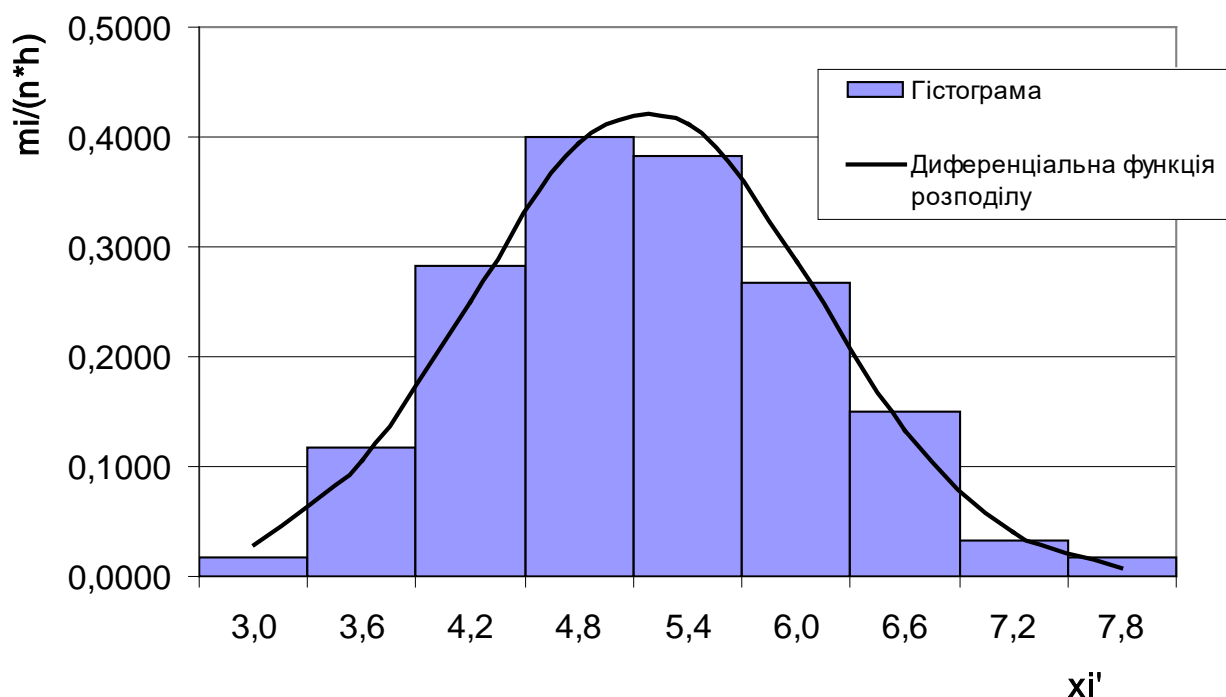


Рис. 6.7. Порівняння гістограми і диференціальної функції розподілу, яка обчислювалась у припущенні нормального закону розподілу

З рис. 6.7 видно, що обидва графіки добре співпадають. Про узгодженість розрахунків також свідчить те, що сума ймовірностей становить 0,9955 (нагадаємо, що для вірогідної події повинно бути 1), а сума вирівнюючих частот дорівнює 99,55, тобто відхиляється від обсягу вибірки ($n = 100$) менше, ніж на 5 %.

Тепер перевіримо узгодженості між емпіричними і вирівнюючими частотами (останні обчислювались у припущенні нормального закону розподілу) за допомогою критерію χ^2 . Обчислення емпіричного значення критерію Пірсона дає такий результат: $\chi^2 = 1,619$.

Для визначення критичного значення критерію Пірсона обчислюємо кількість ступенів вільності $s = 9 - 2 - 1 = 6$ і за допомогою вбудованої функції **ХИ2.ОБР.ПХ**, діалогове вікно якої наведено на рис. 6.8, знаходимо критичне значення критерію Пірсона для рівня значущості 0,05, тобто ймовірність приналежності емпіричного значення критерію його критичній області за умов, що нульова гіпотеза є правильною, дорівнює 0,95.

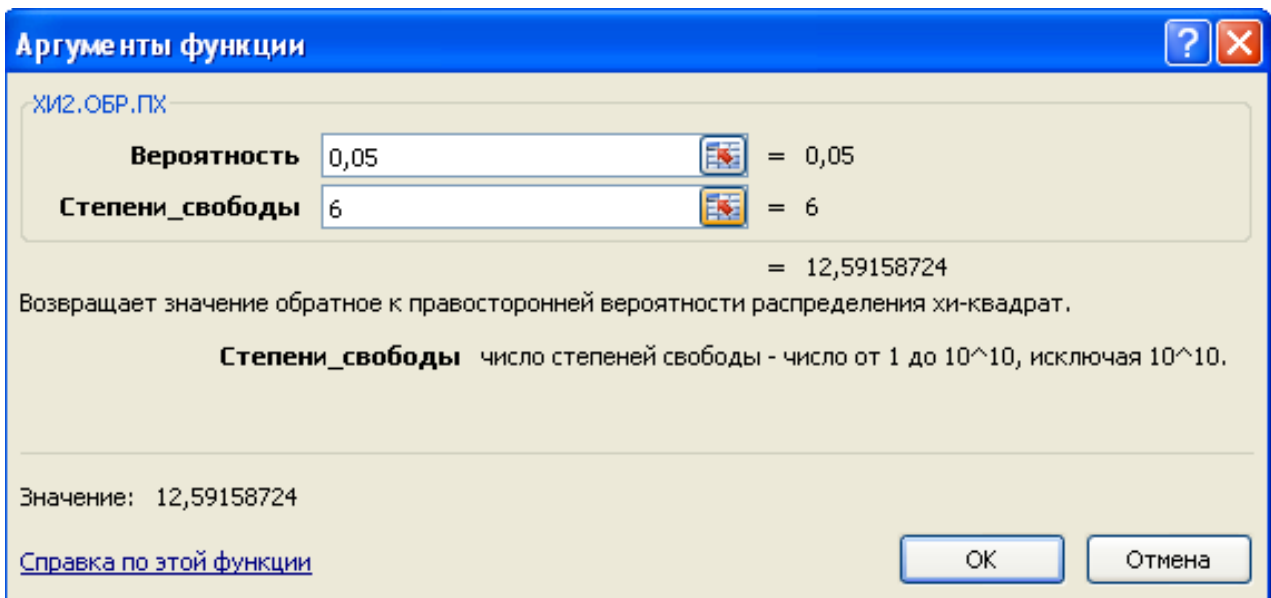


Рис. 6.8. Діалогове вікно функції ХИ2.ОБР.ПХ

Маємо, що $\chi^2(0,05; 6) = 12,59$. Оскільки емпіричне значення χ^2 менше, ніж теоретичне для рівня значущості $\alpha = 0,05$, то нульову гіпотезу щодо статистичної незначущості розбіжностей між емпіричними та вирівнюючими частотами нема підстави відхилити. Отже, розбіжності між емпіричними і вирівнюючими частотами можна пояснити випадковим розпорошенням. Оскільки вирівнюючі частоти були визначені у припущенні, що закон розподілу випадкової величини у генеральній сукупності є нормальним, то припущення про нормальний розподіл генеральної сукупності нема підстав відхилити.

Порівняємо результати обчислення вирівнюючих частот, що вже були отримані шляхом визначення диференціальної функції розподілу, з тими, які можна отримати за допомогою іншого методу обчислення, а саме, шляхом визначення функції розподілу. Враховуючи те, що для заданого емпіричного розподілу частоти першого і останнього інтервалів ($m_1 = 1$ та $m_9 = 1$) менші від 5, то при використанні критерію Пірсона доцільно об'єднати вказані інтервали з сусідніми. Якщо в процесі обчислення функції розподілу використовується довідкова таблиця значень функції Лапласа, то необхідно вводити допоміжну змінну t_i , яка є стандартизованою (аналогічно тому, як це було зроблено задля можливості застосовувати функцію Гаусса), і визначати за цією змінною як за аргументом значення функції Лапласа $\Phi(t_i)$, а потім оцінювати значення функції розподілу випадкової величини X за співвідношенням:

$\hat{F}(x_i) = 0,5 + \Phi(t_i)$. У цьому випадку обчислення проводять у таблицях на зразок табл. 6.3.

Під час застосування вбудованої функції **НОРМ.РАСП** ми отримуємо значення функції розподілу або для стандартизованої змінної t_i , або безпосередньо для випадкової величини X залежно від того, числові характеристики якої з випадкових величин будемо вводити у діалогове вікно цієї функції. Так, на рис. 6.9 наведено приклад заповнення діалогового вікна функції **НОРМ.РАСП** при визначенні функції розподілу випадкової величини X без застосування допоміжної змінної.

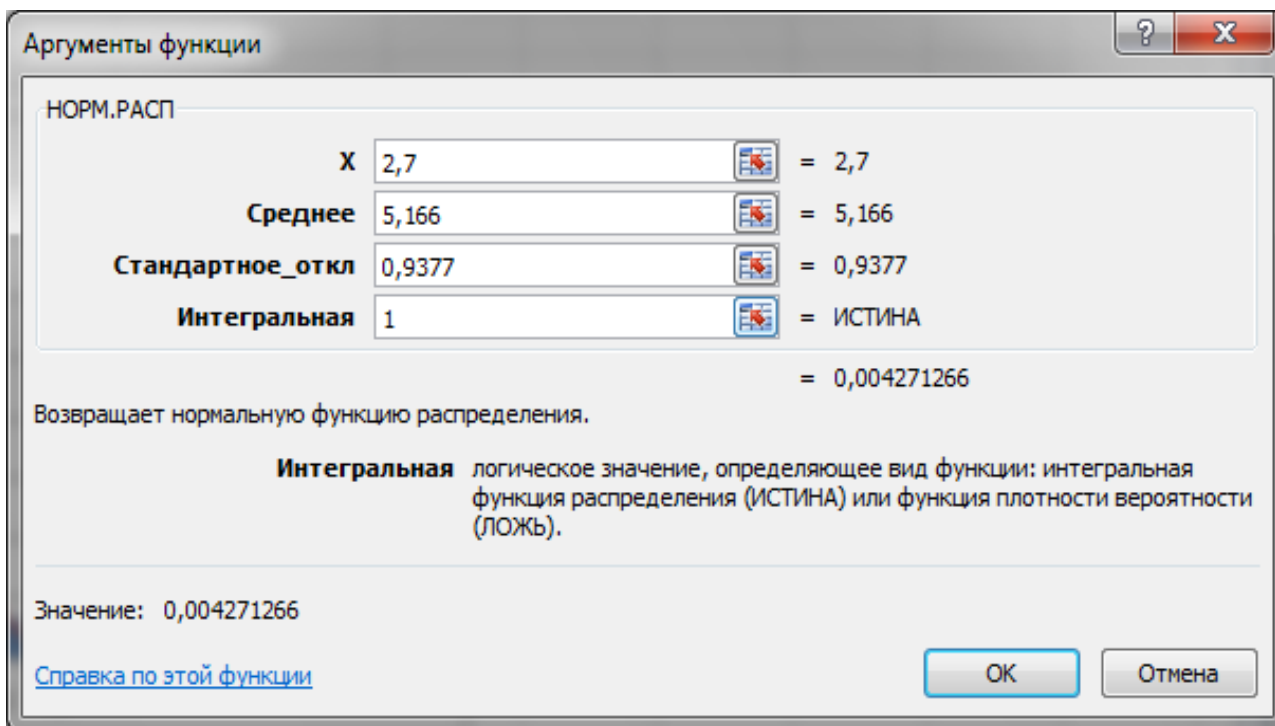


Рис. 6.9. Діалогове вікно функції **НОРМ.РАСП**

Результати розрахунків вирівнюючих частот, у процесі здійснення яких застосовувалось безпосереднє обчислення оцінки функції розподілу з використанням вбудованої функції **НОРМ.РАСП**, наведені в табл. 6.6.

Отже, маємо емпіричне значення критерію Пірсона: $\chi^2 = 1,07$.

Визнаємо кількість ступенів вільності $s = 9 - 2 - 1 = 6$ і за допомогою функції **ХИ2.ОБР.ПХ** знаходимо критичне значення критерію Пірсона за умови, що рівень значущості дорівнює 0,05.

У діалоговому вікні функції **ХИ2.ОБР.ПХ** вказуємо її аргументи і отримуємо: $\chi^2(0,05; 6) = 12,59$.

**Визначення вирівнюючих частот розподілу випадкової величини
у припущенні нормального закону з використанням функції розподілу**

Межі інтервалів	m_i	Номер інтервалу	Верхня межа інтервалу	$\hat{F}(x_i)$	\hat{P}_i	\hat{m}_i	$\frac{(m_i - \hat{m}_i)^2}{\hat{m}_i}$
	0	1	2,7	0,0043	0,0043	0,43	0,43
[2,7; 3,3)	1	2	3,9	0,0885	0,0842	8,42	0,012
[3,3; 3,9)	7						
[3,9; 4,5)	17	3	4,5	0,2388	0,1503	15,03	0,26
[4,5; 5,1)	24	4	5,1	0,4719	0,2332	23,32	0,02
[5,1; 5,7)	23	5	5,7	0,7155	0,2436	24,36	0,08
[5,7; 6,3)	16	6	6,3	0,8867	0,1712	17,12	0,07
[6,3; 6,9)	9	7	6,9	0,9678	0,0811	8,11	0,10
[6,9; 7,5)	2	8	8,1	0,9991	0,0313	3,13	0,01
[7,5; 8,1)	1						
	0	9	$+\infty$	1	0,0009	0,09	0,09
$\chi^2 = \sum_{i=1}^k \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$							1,07

Оскільки $\chi^2 < \chi_{0,05}^2$, то різниця між емпіричними і вирівнюючими частотами є статистично незначущою.

Цей метод обчислення підтверджує, що нема підстав відхилити припущення про нормальний закон розподілу випадкової величини.

6.5. Завдання для самостійної роботи

За заданим статистичним розподілом вибірки висунути гіпотезу про закон розподілу ознаки в генеральній сукупності та при рівні значущості $\alpha = 0,05$ перевірити її правильність.

6.1. У пологовому будинку проводяться вимірювання ваги новонароджених дітей. Результати таких вимірювань, що здійснювались протягом тижня, подано у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, кг	[1,6; 1,8)	[1,8; 2,0)	[2,0; 2,2)	[2,2; 2,4)	[2,4; 2,6)	[2,6; 2,8)	[2,8; 3,0)	[3,0; 3,2)	[3,2; 3,4)	[3,4; 3,6)	[3,6; 3,8)
m_i	5	7	11	18	26	34	29	21	15	9	4

6.2. Проводились вимірювання щоденного заробітку працівників комунального підприємства. Результати вимірювання подано у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, грн.	[80; 100)	[100; 120)	[120; 140)	[140; 160)	[160; 180)	[180; 200)	[200; 220)	[220; 240)	[240; 260)	[260; 280)
m_i	5	9	15	20	25	30	22	10	4	2

6.3. У ході проведення медичного огляду юнаків віком 17 – 20 років, що вступали до вищого навчального закладу, проводилося вимірювання їх зросту. Результати вимірювань наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, см	[148; 154)	[154; 160)	[160; 166)	[166; 172)	[172; 178)	[178; 184)	[184; 190)	[190; 196)	[196; 202)
m_i	8	16	24	30	42	34	21	9	2

6.4. Вимірювався знос автошин протягом місяця. Результати вимірювання наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1}),$ мм	[0; 0,02)	[0,02; 0,04)	[0,04; 0,06)	[0,06; 0,08)	[0,08; 0,10)	[0,10; 0,12)	[0,12; 0,14)	[0,14; 0,16)
m_i	48	42	34	26	18	10	6	4

6.5. Вимірювався час неперервної роботи приладу до його виходу з ладу. Результати надані як інтервальний статистичний розподіл:

$[x_i; x_{i+1}),$ год.	[22; 22,5)	[22,5; 23)	[23; 23,5)	[23,5; 24)	[24; 24,5)	[24,5; 25)	[25; 25,5)	[25,5; 26)	[26; 26,5)	[26,5; 27)	[27; 27,5)
m_i	4	12	16	24	36	28	22	18	16	8	4

6.6. Вимірювалась кількість опадів протягом весняно-польових робіт у північних регіонах України. Результати вимірювання наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1}),$ мм	[440; 452)	[452; 464)	[464; 476)	[476; 488)	[488; 500)	[500; 512)	[512; 524)	[524; 536)	[536; 548)	[548; 560)
m_i	24	18	16	14	12	10	8	6	4	2

6.7. Вимірювалась плінність кадрів на підприємствах певної галузі виробництва протягом року. Результати вимірювання показано як інтервальний статистичний розподіл:

$[x_i; x_{i+1}),$ %	[0,12; 0,18)	[0,18; 0,24)	[0,24; 0,30)	[0,30; 0,36)	[0,36; 0,42)	[0,42; 0,48)	[0,48; 0,54)	[0,54; 0,60)	[0,60; 0,66)
m_i	46	38	32	28	24	18	16	8	6

6.8. Проводився експеримент з вимірювання потужності відбиття від поверхні моря сигналу радіолокаційного приймача. Результати вимірювання надані у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, Вт	[0; 0,05)	[0,05; 0,1)	[0,1; 0,15)	[0,15; 0,2)	[0,2; 0,25)	[0,25; 0,3)	[0,3; 0,35)	[0,35; 0,4)	[0,4; 0,45)
m_i	88	64	58	42	30	22	18	6	4

6.9. Вимірювання показників виконання річного плану підприємствами певної галузі наведено у вигляді статистичного розподілу:

$[x_i; x_{i+1})$, %	[10; 20)	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)	[70; 80)	[80; 90)	[90; 100)	[100; 120)
m_i	2	5	13	16	25	12	10	5	3	1

6.10. Дані про врожайність рису, який зібрано з поливних ділянок фермерських господарств, наведена у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, ц/га	[8; 8,5)	[8,5; 9)	[9; 9,5)	[9,5; 10)	[10; 10,5)	[10,5; 11)	[11; 11,5)	[11,5; 12)	[12; 12,5)	[12,5; 13)
m_i	3	8	16	24	32	28	18	10	5	2

6.11. Проводилися вимірювання жирності молока від різних корів. Результати вимірювання наведені у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, %	[3,45; 3,55)	[3,55; 3,65)	[3,65; 3,75)	[3,75; 3,85)	[3,85; 3,95)	[3,95; 4,05)	[4,05; 4,15)	[4,15; 4,25)	[4,25; 4,35)	[4,35; 4,45)
m_i	10	16	22	30	34	20	14	10	6	4

6.12. Вимірювання опору елементів, виготовлених шляхом використання нової технології, дали результати, які наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, Ом	[0,32; 0,36)	[0,36; 0,40)	[0,40; 0,44)	[0,44; 0,48)	[0,48; 0,52)	[0,52; 0,56)	[0,56; 0,60)	[0,60; 0,64)	[0,64; 0,68)	[0,68; 0,72)
m_i	40	36	30	24	20	18	16	12	8	2

6.13. Результати вимірювання граничного навантаження на сталевий болт наведено інтервальним статистичним розподілом:

$[x_i; x_{i+1})$, кг/мм ²	[4,5; 5,5)	[5,5; 6,5)	[6,5; 7,5)	[7,5; 8,5)	[8,5; 9,5)	[9,5; 10,5)	[10,5; 11,5)	[11,5; 12,5)	[12,5; 13,5)
m_i	40	32	28	24	20	18	16	12	4

6.14. Результати вимірювання рівня води навесні під час повені наведені у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, см	[24; 48)	[48; 72)	[72; 96)	[96; 120)	[120; 144)	[144; 168)	[168; 192)	[192; 216)
m_i	5	4	6	12	16	7	4	1

6.15. Результати досліджень кількості проданих пар чоловічого взуття від його вартості надано у вигляді інтервального розподілу:

$[x_i; x_{i+1})$, грн	[75; 125)	[125; 175)	[175; 225)	[225; 275)	[275; 325)	[325; 375)	[375; 425)	[425; 475)	[475; 525)	[525; 575)
m_i	2	12	18	24	38	21	19	12	10	8

6.16. Урожайність цукрових буряків вимірювалась у певному районі південного регіону України. Результати вимірювання подано як інтервальний статистичний розподіл:

$[x_i; x_{i+1})$, ц/га	[340; 342,6)	[342,6; 345,2)	[345,2; 347,8)	[347,8; 350,4)	[350,4; 353)	[353; 355,6)	[355,6; 358,2)	[358,2; 360,8)
m_i	12	18	26	38	40	26	16	6

6.17. Вимірювалось відхилення діаметра валика від його номінального розміру. Результати вимірювання наведено як інтервальний статистичний розподіл:

$[x_i; x_{i+1})$, МК	[0,228; 0,248)	[0,248; 0,268)	[0,268; 0,288)	[0,288; 0,308)	[0,308; 0,328)	[0,328; 0,348)	[0,348; 0,368)	[0,368; 0,388)	[0,388; 0,408)
m_i	6	16	21	36	42	32	22	12	8

6.18. Вимірювання показників виконання річного плану підприємствами певної галузі наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, %	[10; 20)	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)	[70; 80)	[80; 90)	[90; 100)	[100; 120)
m_i	2	5	13	16	25	12	10	5	3	1

6.19. Урожайність рису, який зібрано з поливних ділянок фермерських господарств, наведено у вигляді інтервального розподілу:

$[x_i; x_{i+1})$, ц/га	[9,5; 10)	[10; 10,5)	[10,5; 11)	[11; 11,5)	[11,5; 12)	[12; 12,5)	[12,5; 13)	[13; 13,5)
m_i	5	16	24	32	40	43	47	50

6.20. Вимірювання діаметрів кульок, виготовлених верстатом-автоматом, наведено як інтервальний статистичний розподіл:

$[x_i; x_{i+1})$, мм	[6; 6,08)	[6,08; 6,16)	[6,16; 6,24)	[6,24; 6,32)	[6,32; 6,4)	[6,4; 6,48)	[6,48; 6,56)	[6,56; 6,64)
m_i	8	18	24	32	28	21	15	6

6.21. Вимірювався знос автошин за місяць. Результати вимірювання наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, мм	[0; 0,02)	[0,02; 0,04)	[0,04; 0,06)	[0,06; 0,08)	[0,08; 0,1)	[0,1; 0,12)	[0,12; 0,14)	[0,14; 0,16)
m_i	48	42	34	26	18	10	6	4

6.22. Вимірювалась в'язкість нафти, видобутої із свердловин. Результати вимірювання наведено як інтервальний статистичний розподіл:

$[x_i; x_{i+1})$, ум. од.	[4,2; 4,28)	[4,28; 4,36)	[4,36; 4,44)	[4,44; 4,52)	[4,52; 4,6)	[4,6; 4,68)	[4,68; 4,76)	[4,76; 4,84)	[4,84; 4,92)	[4,92; 5)
m_i	2	6	10	14	16	8	6	4	2	1

6.23. Вимірювання опору елементів, виготовлених шляхом використання нової технології, дали результати, які наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, Ом	[0,32; 0,36)	[0,36; 0,4)	[0,4; 0,44)	[0,44; 0,48)	[0,48; 0,52)	[0,52; 0,56)	[0,56; 0,6)	[0,6; 0,64)	[0,64; 0,68)	[0,68; 0,72)
m_i	40	36	30	24	20	18	16	12	8	2

6.24. Вимірювався час неперервного горіння електролампочок, виготовлених фірмою, до виходу їх з ладу. Результати вимірювання наведено як інтервальний статистичний розподіл:

$[x_i; x_{i+1})$, год	[22; 22,5)	[22,5; 23)	[23; 23,5)	[23,5; 24)	[24; 24,5)	[24,5; 25)	[25; 25,5)	[25,5; 26)	[26; 26,5)	[26,5; 27)	[27; 27,5)
m_i	4	12	16	24	36	28	22	18	16	8	4

6.25. Вимірювалась продуктивність праці за зміну робітників однакового профілю спеціальності й однакового віку певної галузі. Результати вимірювання наведено у вигляді інтервального статистичного розподілу:

$[x_i; x_{i+1})$, грн	[28,4; 29,2)	[29,2; 30)	[30; 30,8)	[30,8; 31,6)	[31,6; 32,4)	[32,4; 33,2)	[33,2; 34)	[34; 34,8)	[34,8; 35,6)	[35,6; 36,4)
m_i	60	48	36	24	18	14	12	10	4	2

6.6. Контрольні запитання

1. Дайте означення статистичної гіпотези.
2. Що таке нульова й альтернативна статистичні гіпотези?
3. Що називають простою та складною статистичними гіпотезами?
4. Що називається статистичним критерієм?
5. Що називають критичною областю статистичного критерію?
6. Які висновки можна зробити, порівнюючи емпіричне значення критерію з критичним значенням?
7. Що таке одностороння та двостороння критичні області?
8. Що таке рівень значущості α ?
9. Які значення статистичного критерію обмежують критичну область?
10. Чи можна підтвердити нульову гіпотезу?
11. Що таке помилка першого роду? Наведіть приклади.
12. Що таке помилка другого роду? Наведіть приклади.
13. Що визначає потужність критерію?

14. Що називають емпіричними частотами?
15. Що називають теоретичними частотами?
16. Що стверджує статистична гіпотеза, яка розглядається при порівнянні теоретичного та емпіричного розподілів?
17. Запишіть формули для обчислення теоретичних частот за допомогою функції Гаусса, якщо припускається, що ознака X генеральної сукупності має нормальний закон розподілу.
18. Запишіть формули для обчислення теоретичних частот за допомогою інтегральної функції розподілу, якщо припускається, що ознака X генеральної сукупності має нормальний закон розподілу.
19. Охарактеризуйте критерій узгодженості Пірсона.
20. У чому полягають переваги критерію Пірсона порівняно з іншими статистичними критеріями щодо перевірки гіпотези про закон розподілу випадкової величини?
21. Опишіть загальну методику перевірки правильності гіпотези щодо закону розподілу ознаки в генеральній сукупності.

Лабораторна робота 7

Елементи теорії кореляції.

Елементи регресійного аналізу

7.1. Мета роботи:

закріпити основні теоретичні положення кореляційного та регресійного аналізів;

оволодіти технікою використання функції **ЛИНЕЙН** та надбудови **Анализ данных MS Excel** для побудови рівняння регресії за **незгрупованими даними** у припущенні лінійної форми кореляційного зв'язку;

навчитися використовувати статистичні функції **ЛИНЕЙН** та **СТЮДЕНТ.ОБР.2Х** для перевірки значущості параметрів рівняння регресії.

7.2. Теоретичні положення

Основні задачі кореляційного та регресійного аналізів полягають у дослідженні взаємозв'язку між двома або більшою кількістю випадкових

величин, визначенні форми цього зв'язку, обчисленні статистичних оцінок параметрів моделі та перевірка статистичних гіпотез щодо значущості цих моделей та адекватності моделі в цілому.

Вважається, що випадкові величини пов'язані між собою **статистичною**, або **стохастичною залежністю**, якщо під впливом зміни однієї з випадкових величин змінюється закон розподілу іншої випадкової величини або його числові характеристики.

Якщо незалежною змінною є випадкова величина X (**екзогенний**, або **зовнішній** фактор) і під її впливом змінюється випадкова величина Y (**ендогенний**, або **внутрішній** фактор), то X називається **фактор-аргументом**, або **регресором**, тоді Y називається **функціональним фактором**. У загальному випадку проста регресійна модель має вигляд:

$$Y = \hat{Y} + \varepsilon, \quad (7.1)$$

де Y – матриця-стовпець (або вектор) спостережень, що відповідають функціональному фактору:

$$Y = (y_1; y_2; \dots; y_n)^T;$$

\hat{Y} – матриця-стовпець (або вектор) значень функціонального фактора, які обчислюються відповідно до прийнятої моделі:

$$\hat{Y} = f(X);$$

$f(X)$ – функція, за допомогою якої здійснюється апроксимація стохастичної залежності, тобто **тренд**, або тенденція, зміни функціонального фактора, що обумовлені зміною зовнішнього фактора;

X – матриця-стовпець (або вектор) спостережень, що відповідає незалежній змінній:

$$X = (x_1; x_2; \dots; x_n)^T;$$

n – загальна кількість спостережень (обсяг вибіркової сукупності);

ε – матриця-стовпець (або вектор) помилок (похибок) моделі:

$$\varepsilon = (e_1; e_2; \dots; e_n)^T.$$

Однією з передумов застосування кореляційного та регресійного аналізів є нормальний розподіл кожної з випадкових величин сукупності.

Окремим випадком статистичної залежності є **кореляційна залежність**, коли під впливом мінливості незалежної величини X змінюється положення центра угруповання значень функціонального фактора Y , тобто його умовне середнє.

Умовним середнім $\overline{y_{x_i}}$ є середнє арифметичне значень випадкової величини Y , що відповідають певному значенню x_i фактор-аргументу.

У випадку **парної кореляції** (або **однофакторної регресії**) розглядається вплив тільки одного фактор-аргументу на функціональний фактор.

Регресійний аналіз полягає у визначенні загального вигляду рівняння регресії, оцінці невідомих параметрів цього рівняння та їх значущості. Побудові рівняння регресії передують припущення щодо функції, за допомогою якої описують кореляційний зв'язок між факторами.

Однією з передумов застосування кореляційного та регресійного аналізів є нормальний розподіл кожної з випадкових величин сукупності.

Для оцінювання параметрів рівняння регресії проводять n незалежних випробувань (спостережень), внаслідок чого отримують n пар чисел $(x_i; y_j)$, які утворюють вибірку сукупності. Множина емпіричних точок, що зображені на координатній площині Oxy , утворюють **еліпс розпорошення**.

З'єднавши емпіричні точки ламаною, отримаємо **емпіричну лінію регресії**, за виглядом якої можна висловити припущення про наявність і форму кореляційного зв'язку.

Під **наближенням**, або **апроксимацією** мають на увазі заміну вихідних даних, що описують економічні явища або процеси на такі, що близькі до них, але більш прості й, відповідно, більш зручні в обчисленні. Апроксимація є одним з методів побудови економіко-математичних моделей із застосуванням кореляційного та регресійного аналізів

Визначення параметрів рівняння регресії здійснюється за **методом найменших квадратів** (МНК). Оскільки помилки моделі залежать від параметрів моделі, то до розгляду вводиться функція, яка є сумою квадратів відхилень емпіричних даних від тих, що обчислені за рівнянням регресії. Суть МНК полягає у тому, що у припущенні певної форми кореляційного зв'язку параметри рівняння регресії визначаються, виходячи з умови, що суми квадратів відхилень емпіричних даних від тих,

що обчислені за рівнянням регресії, повинні приймати найменше значення. Найбільш простою і водночас найбільш поширеною є лінійна модель кореляційного зв'язку.

Апроксимацією кореляційної залежності у випадку парної лінійної регресії є функція, що має вигляд:

$$\hat{y} = b_0 + b_1 x, \quad (7.2)$$

де b_1 – кутовий коефіцієнт лінії тренда, який є оцінкою **коефіцієнта регресії** Y на X у генеральній сукупності;

b_0 – точкова оцінка вільного члена рівняння регресії, що відповідає генеральній сукупності.

За необхідною ознакою екстремуму функції суми квадратів похибок моделі всі її похідні першого порядку дорівнюють нулю.

Звідси отримують систему нормальних рівнянь відносно параметрів моделі.

Розв'язок цієї системи дає такі співвідношення для оцінок параметрів лінійного рівняння регресії Y на X за МНК:

$$b_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}; \quad (7.3)$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}. \quad (7.4)$$

Параметр b_1 можна також обчислювати за формулою:

$$b_1 = r \cdot \frac{\sigma_Y}{\sigma_X}, \quad (7.5)$$

де r – **вибірковий коефіцієнт кореляції** (Пірсона), який є точковою оцінкою **коефіцієнта кореляції** ρ генеральної сукупності.

У свою чергу, вибірковий коефіцієнт кореляції визначають як

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_X \cdot \sigma_Y}, \quad (7.6)$$

де \overline{xy} – початковий емпіричний момент першого порядку добутку випадкових величин:

$$\overline{xy} = \frac{1}{n} \cdot \sum_{i=1}^n x_i y_i.$$

Коефіцієнт кореляції генеральної сукупності, статистичною оцінкою якого є вибірковий коефіцієнт кореляції, можна записати як відношення **коваріації випадкових величин** X та Y до їх стандартних відхилень:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \cdot \sigma_Y}. \quad (7.7)$$

Коваріація випадкової величини $\text{cov}(X, Y)$ є **центральним емпіричним моментом першого порядку добутку випадкових величин** X та Y :

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - M(X))(y_i - M(Y)). \quad (7.8)$$

Коваріація випадкової величини є оцінкою кореляційного моменту випадкових величин X та Y , або **коефіцієнта коваріації**, який характеризує щільність лінійного кореляційного зв'язку. Так, для некорельованих випадкових величин $\text{cov}(X, Y) = 0$.

Дисперсії випадкових величин X та Y у генеральній сукупності

$$\sigma_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - M(X))^2, \quad \sigma_Y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - M(Y))^2$$

мають також назву **варіації** і позначаються, відповідно, як $\text{var}(X)$ та $\sigma_Y^2 = \text{var}(Y)$.

Отже, формулу (7.4) можна записати таким чином:

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}}. \quad (7.8^l)$$

Саме така форма запису прийнята в *MS Excel*.

Коефіцієнт кореляції на відміну від коефіцієнта регресії не має вимірності. Оскільки коефіцієнт кореляції (Пірсона) є однаковим для спряжених рівнянь лінійної регресії, то для його позначення індекси не застосовуються, тобто $r_{XY} = r_{YX} = r$.

Коефіцієнт кореляції визначає, на яку частку від свого середнього квадратичного відхилення змінюється випадкова величина Y за умов, що X змінюється на величину свого середнього квадратичного відхилення. Коефіцієнт кореляції характеризує щільність саме лінійного кореляційного зв'язку. Якщо $|r| \leq 0,35$, то кореляційний зв'язок вважається статистично незначущим. При $|r| > 0,35$ необхідно враховувати наявність кореляційного зв'язку між факторами, що розглядаються. При $|r| > 0,7$ кореляційний зв'язок вважається щільним.

Аналогічно визначається рівняння регресії X на Y :

$$\hat{x} = a_0 + a_1 \cdot y, \quad (7.9)$$

де a_1 – кутовий коефіцієнт лінії тренду, який є точковою оцінкою коефіцієнта регресії X на Y ;

a_0 – точкова оцінка вільного члена рівняння регресії, що відповідає генеральній сукупності.

Відповідно, за методом найменших квадратів маємо:

$$a_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{y^2 - (\bar{y})^2}; \quad (7.10)$$

$$a_0 = \bar{x} - a_1 \cdot \bar{y}. \quad (7.11)$$

Кут між лініями регресії також є показником щільності кореляційного зв'язку. Так, при $r = 1$ лінії регресії співпадають. Якщо $r = 0$, тобто кореляція відсутня, то лінії регресії взаємно перпендикулярні, при цьому лінія регресії X на Y паралельна осі Oy , а лінія регресії Y на X паралельна осі Ox . Точка їх перетину має координати $(\bar{x}; \bar{y})$.

Наступним кроком після побудови рівняння регресії є перевірка значущості кожного з його параметрів окремо.

Для рівняння регресії Y на X перевірці підлягає основна статистична гіпотеза: $b_i = 0$, де $i = \overline{0, 1}$, при альтернативній: $b_i \neq 0$.

Перевірка здійснюється за критерієм Стюдента. Емпіричне значення критерію обчислюється за формулою:

$$t = \frac{b_i}{\hat{S}_{b_i}}, \quad i = \overline{0, 1} \quad (7.12)$$

де \hat{S}_{b_i} – скорегована помилка i -го параметру рівняння регресії.

7.3. Змістовна постановка задачі

Інвестування вважається засобом страхування від інфляції, якщо ставка відсотка не нижча за рівень інфляції або знаходиться на тому ж рівні. Перевіримо, чи є інвестування у золото страховкою від інфляції. Для цього скористаємося даними, що наведені у табл. 7.1.

Таблиця 7.1

Ціни на золото та індекс споживчих цін (CPI) у Сполучених Штатах за період 1977 – 1991 рр.

№	Рік	Ціна на золото в Нью-Йорку, долар за тройську унцію	Індекс споживчих цін CPI
1	1977	147,98	60,6
2	1978	193,44	65,2
3	1979	307,62	72,6
4	1980	612,51	82,4
5	1981	459,61	90,9
6	1982	376,01	96,5
7	1983	423,83	99,6
8	1984	360,29	103,9
9	1985	317,30	107,6
10	1986	367,87	109,6
11	1987	446,50	113,6
12	1988	436,93	118,3
13	1989	381,28	124,0
14	1990	348,08	130,7
15	1991	362,04	136,2

7.4. Приклад виконання лабораторної роботи 7

Завантажимо дані табл. 7.1 на перший аркуш книги *MS Excel* таким чином, що дані стовпця 3, який розглядатимемо як внутрішній фактор Y , утворюють стовпець А, а дані стовпця 4, тобто зовнішній фактор X , утворюють стовпець В, клітини **A1** і **B1**, відповідно, містять "мітки", що позначають певний фактор.

Для визначення параметрів рівняння регресії використаємо вбудовану функцію **ЛИНЕЙН**, що належить до категорії **Статистические**. Для цього на робочому аркуші поруч з даними слід виділити поле розміром 5×2 (наприклад, комірки **D1:E5**). Після цього натискаємо клавішу "=", викликаємо функцію **ЛИНЕЙН** та заповнюємо поля у вікні **Аргументы функции**. При цьому до поля **Конст** та поля **Статистика** слід записати одиниці, тоді програма буде виводити рівняння регресії з урахуванням вільного члена.

На рис. 7.1 наведено вигляд діалогового вікна функції **ЛИНЕЙН**, поля якого заповнені відповідно до умов прикладу.

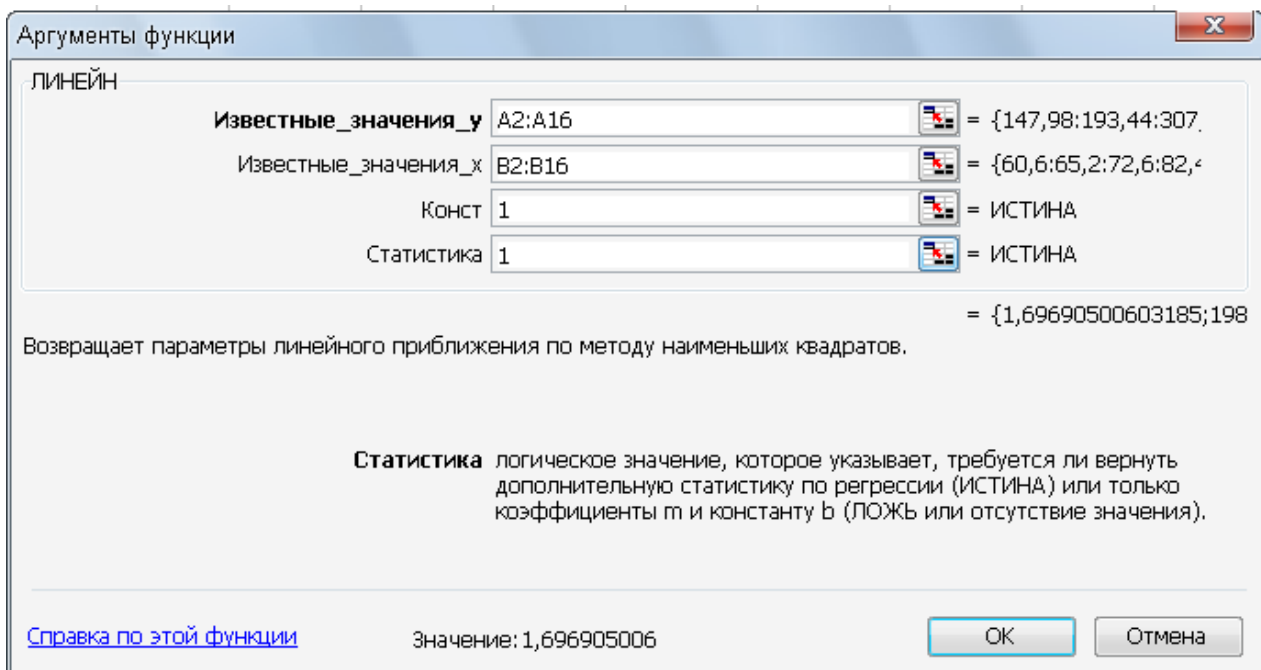


Рис. 7.1. Діалогове вікно функції **ЛИНЕЙН**

Для виведення результатів обчислення замість "OK" слід натиснути "**Ctrl+Shift+Enter**", і у наслідок цього у виділеному полі з'явиться таблиця даних. Результати застосування функції **ЛИНЕЙН** (вони виведені у комірках **D1:E5**) та їх розшифровку подано в табл. 7.2.

Результати застосування функції ЛИНЕЙН

$b_1 =$	1,6969	$b_0 =$	198,4052
$\widehat{S}_{b_1} =$	1,2326	$\widehat{S}_{b_0} =$	127,2120
$R^2 =$	0,1272	$SE_Y =$	106,2035
$F =$	1,8953	$df =$	13
$SSR =$	21377,09	$SSE =$	146629,30

Дані, що виводяться в таблиці як результат застосування функції **ЛИНЕЙН**, мають такий зміст:

1) перший рядок таблиці містить коефіцієнти регресії, при цьому справа – константа b_0 , а зліва – коефіцієнт регресії b_1 ;

2) у другому рядку – відповідні коефіцієнтам рівняння регресії їх скореговані помилки;

3) у третьому рядку, зліва – коефіцієнт детермінації, справа – стандартна помилка розрахункових значень;

4) у четвертому рядку, зліва – спостережене значення критерію Фішера, справа – число степенів вільності;

5) у п'ятому рядку, зліва – сума квадратів відхилень розрахункових значень, справа – сума квадратів відхилень залишкової похибки.

Отже, відповідно до формули (7.2) за результатами використання функції **ЛИНЕЙН** отримуємо рівняння регресії:

$$\hat{y} = 198,4052 + 1,6969 \cdot x.$$

За даними функції **ЛИНЕЙН** маємо, що коефіцієнт кореляції між факторами X та Y дорівнює:

$$r = \sqrt{R^2} = \sqrt{0,12724} = 0,3567,$$

тобто майже на межі статистичної значущості.

Отже, залишається питання, чи можна вважати вплив фактора X на фактор Y статистично суттєвим. За критерієм Стюдента перевіримо статистичну значущість коефіцієнта регресії b_1 . Перевірці підлягає

гіпотеза $H_0: b_1 = 0$ при альтернативній гіпотезі $H_1: b_1 \neq 0$. Підставимо дані табл. 7.2 у формулу (7.12) і отримаємо емпіричне значення критерію Стьюдента:

$$t_{b_1} = \frac{1,6969}{1,2326} = 1,5596.$$

Порівняємо його з критичним значенням розподілу Стьюдента, для визначення якого скористаємось функцією **СТЬДРАСПОБР**, що належить до категорії **Статистические** (рис. 7.2).

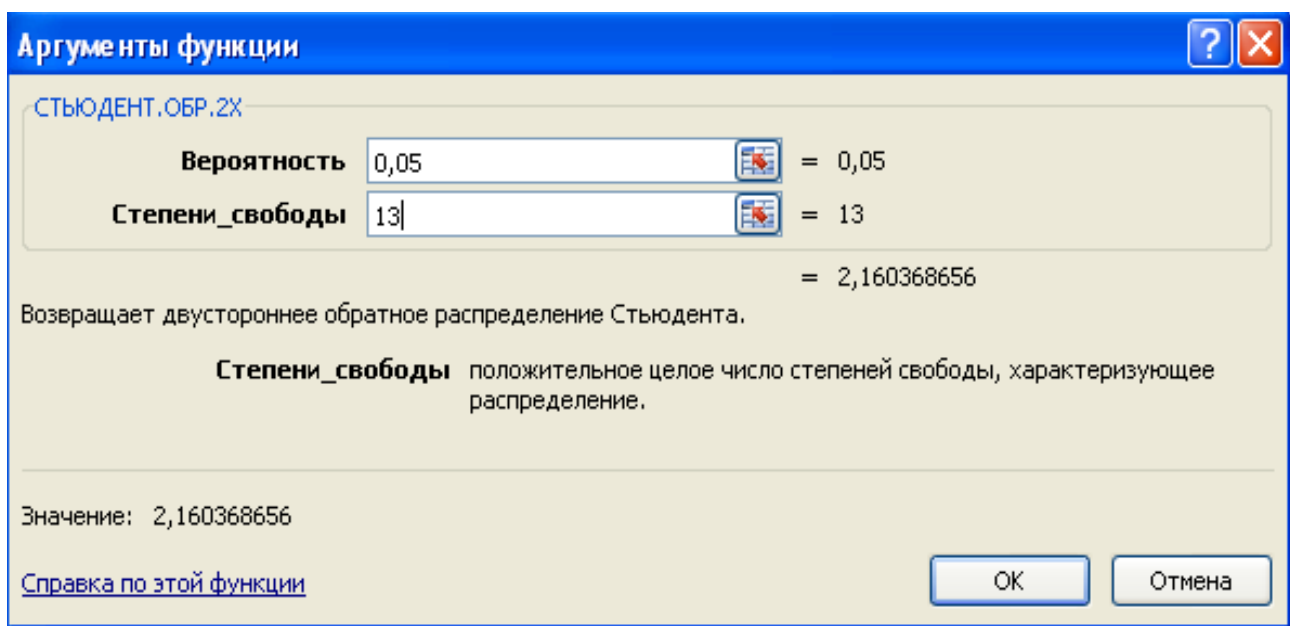


Рис. 7.2. Діалогове вікно функції **СТЬЮДЕНТ.ОБР.2X**

Аргументами функції є рівень значущості $\alpha = 0,05$ (**Вероятность**) та кількість ступенів вільності (**Степени_свободы**) $df = n - 2$. У нашому прикладі $n = 15$.

Отже, $t_{0,05}(13) = 2,16$. Звідки маємо, що емпіричне значення критерію Стьюдента для коефіцієнта регресії менше за критичне для рівня надійності 95 %:

$$t_{b_0} = 1,3767 < t_{0,05}(13) = 2,1604.$$

Отже, гіпотезу H_0 , за якою значення коефіцієнта кореляції є статистично несуттєвим, немає причин відхилити.

Це означає, що за статистичними даними індекс споживчих цін не впливає на ціну золота, тобто кореляційна залежність між цими факторами є статистично незначущою.

Перевіримо значущість вільного члена рівняння регресії b_0 .

Перевірці підлягає гіпотеза $H_0: b_0 = 0$ при альтернативній гіпотезі $H_1: b_0 \neq 0$.

Підставимо дані табл. 7.2 у формулу (7.12) й отримаємо емпіричне значення критерію Стюдента:

$$t_{b_0} = \frac{198,4052}{127,2120} = 1,3767.$$

Оскільки для b_0 емпіричне значення критерію Стюдента менше критичного, тобто $t_{b_0} = 1,5596 < t_{0,05}(13) = 2,1604$, то з надійністю 95 % гіпотезу H_0 , за якою вільний член рівняння регресії є статистично несуттєвим, немає причин відхилити.

Отже, остаточно робимо висновок, що за даними табл. 7.2 інвестиції у золото не є надійним захистом від інфляції.

7.5. Завдання для самостійної роботи

За допомогою функції **ЛИНЕЙН** побудувати модель однофакторної лінійної регресії Y на X та перевірити значущість коефіцієнта регресії.

У всіх варіантах Y – середньомісячна продуктивність лави.

У варіантах 7.1; 7.5; 7.9 та 7.13 (табл. 7.3) змінна X відповідає довжині лави.

У варіантах 7.2; 7.6; 7.10; 7.13; 7.15; 7.17; 7.22; 7.23; 7.25 змінна X – це продуктивність лави.

У варіантах 7.3; 7.7; 7.11 та 7.14 X – термін служби комбайна.

У інших варіантах змінна X – середня швидкість подачі вугілля.

Таблиця 7.3

Варіанти завдань

Варіант 7.1		Варіант 7.2		Варіант 7.3		Варіант 7.4		Варіант 7.5	
1		2		3		4		5	
Y	X	Y	X	Y	X	Y	X	Y	X
23,55	348	23,56	2,18	23,56	9	21,56	0,79	10,76	150

Продовження табл. 7.3

1		2		3		4		5	
7,64	94	7,64	2,25	7,64	4	8,44	0,66	20,14	163
14,35	195	14,53	1,54	14,83	14	14,34	1,81	9,87	160
18,92	300	18,98	1,54	18,92	14	18,28	1,21	30,21	175
33,61	164	33,61	2,55	33,61	6	33,68	2,54	9,22	180
28,26	208	28,26	2,55	28,26	6	28,22	2,49	28,55	209
15,68	248	15,63	1,83	15,63	10	15,63	1,33	25,76	160
28,7	201	28,47	2,37	28,47	14	28,47	1,57	13,44	170
11,84	144	11,88	2,29	11,88	10	11,84	1,24	35,99	160
7,38	153	7,31	2,38	7,31	6	7,81	0,78	14,36	173
13,27	185	13,27	1,63	13,27	15	13,07	1,82	9,55	157
20,73	290	20,73	1,89	20,73	6	20,73	1,23	3,72	145
8,29	140	8,24	1,43	8,29	9	8,29	1,46	11,35	158
13,21	160	13,21	1,85	13,23	3	13,13	1,05	13,51	131
9,54	230	9,57	1,95	9,57	7	9,47	1,16	13,72	168
6,66	150	6,66	1,58	6,66	3	6,66	0,93	10,84	125
15,17	163	15,17	2,13	15,07	18	15,01	0,74	13,35	216
32,36	334	32,36	2,58	32,46	3	32,46	1,54	7,56	115
15,82	144	15,82	1,71	15,92	18	15,82	1,28	9,51	84
7,94	135	7,94	1,32	7,99	21	7,94	0,74	13,61	165
8,05	150	8,05	2,14	8,55	10	8,05	0,83	21,64	227
16,27	194	16,27	2,12	16,27	14	16,24	0,76	25,38	250
11,16	160	11,06	1,61	11,28	12	11,16	0,67	12,23	170
10,81	270	10,81	1,53	11,81	6	10,91	0,93	12,11	205
13,15	147	13,15	2,58	13,15	21	13,15	0,87	10,27	271
Варіант 7.6		Варіант 7.7		Варіант 7.8		Варіант 7.9		Варіант 7.10	
1		2		3		4		5	
Y	X	Y	X	Y	X	Y	X	Y	X
10,76	2,71	10,06	10	10,76	0,68	11,09	148	11,09	1,26
20,74	2,16	20,74	5	20,74	1,78	13,53	135	13,23	2,02
9,81	1,69	9,81	11	9,87	2,08	5,74	119	5,64	1,34
30,29	2,76	30,29	2	30,29	1,96	4,96	122	4,96	0,42
9,22	1,42	9,22	22	9,22	1,54	10,64	155	10,64	1,77
28,58	2,76	18,85	10	28,85	2,06	15,56	156	15,56	2,82
25,75	2,09	25,26	8	25,76	1,92	5,18	117	5,18	1,22
13,44	1,58	13,48	2	13,84	2,27	4,81	122	4,01	0,38
35,99	2,54	32,92	4	35,99	2,51	7,62	130	7,62	1,43
14,38	2,23	14,86	12	14,36	1,59	8,61	140	8,71	1,48
9,55	1,37	9,85	11	9,55	1,02	9,83	145	9,35	1,41

Продовження табл. 7.3

1		2		3		4		5	
3,72	2,21	3,74	12	3,72	0,98	10,39	150	9,92	1,42
11,35	1,86	11,38	12	11,35	1,14	12,22	170	12,29	1,83
13,51	2,24	12,51	13	13,51	0,70	12,45	175	12,40	1,96
13,76	2,41	13,62	11	13,76	1,25	16,35	138	16,35	2,82
10,89	1,42	10,84	18	10,21	0,96	22,18	120	22,08	2,29
13,35	2,33	13,85	10	13,35	0,94	22,78	140	22,78	2,03
7,59	1,26	8,56	10	7,54	0,94	20,71	137	20,71	3,84
9,51	1,62	9,19	12	9,59	0,87	30,16	225	30,16	4,15
13,63	1,49	13,13	8	12,63	0,96	15,24	220	15,24	2,56
21,64	2,64	22,64	13	21,64	2,27	12,65	265	12,65	1,49
25,38	2,57	25,38	7	25,38	1,25	16,76	237	16,96	2,55
12,29	1,92	12,39	3	11,93	1,11	10,85	215	10,85	1,28
12,11	1,93	12,82	4	11,62	1,78	23,43	200	28,48	3,25
10,24	1,96	10,73	11	10,17	0,98	13,85	226	13,85	1,57
7,03	1,42	7,45	8	7,35	0,57	11,83	150	12,23	1,64
16,51	1,66	17,14	15	16,54	1,29	15,72	180	16,27	2,81
9,95	1,29	9,52	2	9,92	1,05	22,85	220	22,65	1,96
24,18	2,71	28,81	17	21,16	1,20	18,91	217	15,61	1,25
Варіант 7.11		Варіант 7.12		Варіант 7.13		Варіант 7.14		Варіант 7.15	
1		2		3		4		5	
Y	X	Y	X	Y	X	Y	X	Y	X
11,09	14	11,09	1,04	16,62	2,15	16,62	11	16,62	2,96
13,53	8	13,53	1,06	4,31	1,22	4,31	29	4,31	0,72
5,64	28	5,64	0,75	4,04	2,25	4,04	28	4,14	0,81
4,91	23	4,96	1,02	11,36	1,63	11,23	20	11,43	1,05
10,66	19	10,64	1,21	14,07	2,14	14,07	4	14,07	1,54
15,56	5	15,56	1,18	21,47	2,35	21,42	2	21,42	2,81
5,18	27	5,18	1,03	14,04	2,31	14,06	1,8	14,64	1,85
4,01	20	4,01	1,05	26,01	2,18	26,01	4	26,01	3,65
7,92	26	7,92	1,38	6,15	1,36	6,58	27	6,15	0,49
8,71	25	8,71	1,76	11,54	1,95	11,05	11	11,54	0,97
9,63	17	9,23	0,93	14,51	2,21	14,57	2	14,57	0,92
9,39	29	9,32	1,13	14,23	2,34	14,12	5	14,52	1,05
12,29	16	12,29	1,21	12,13	2,27	12,51	14	12,13	1,02
12,45	5	12,45	1,24	7,85	1,35	7,18	18	7,85	1,12
16,35	9	16,38	1,74	10,52	1,61	10,52	11	10,92	1,87
22,08	6	22,18	1,55	18,67	2,06	18,67	4	18,67	2,18
22,78	4	22,78	1,58	11,47	1,08	11,47	10	11,47	0,64

Продовження табл. 7.3

1		2		3		4		5	
20,28	18	20,71	1,64	14,85	2,25	14,87	3	14,87	1,08
30,16	8	30,61	1,23	5,42	1,27	5,42	13	5,42	1,82
15,29	6	15,24	1,26	28,54	1,93	28,54	3	28,54	2,91
12,65	5	12,65	1,13	25,41	2,14	25,41	3	25,41	1,42
16,76	8	16,76	0,96	9,23	1,35	9,20	12	9,20	1,32
10,88	21	10,88	0,68	28,88	2,35	28,85	11	28,85	3,55
23,48	9	23,48	1,25	26,64	2,01	26,64	7	26,64	3,34
13,59	7	13,85	0,87	19,86	1,93	19,86	12	19,86	1,69
11,38	6	11,83	1,27	25,63	1,64	25,34	10	25,63	3,39
15,27	23	15,72	1,34	19,21	2,15	19,21	7	19,21	1,33
22,65	18	22,68	1,42	16,62	2,08	16,66	11	16,62	2,36
31,61	14	32,25	1,33	6,12	1,54	5,44	11	4,42	1,95
24,53	13	13,58	1,26	4,62	1,27	4,81	26	4,51	1,71
Варіант 7.16		Варіант 7.17		Варіант 7.18		Варіант 7.19		Варіант 7.20	
1		2		3		4		5	
Y	X	Y	X	Y	X	Y	X	Y	X
12,19	14	11,89	2,04	16,62	3,58	16,26	11	16,26	3,36
14,53	8	13,53	2,06	4,25	1,22	4,32	29	4,51	0,57
8,64	23	5,64	0,75	4,42	0,25	4,04	28	4,28	0,81
5,91	28	4,96	1,02	11,34	2,65	11,43	12	11,83	1,05
11,45	19	10,66	1,81	14,07	2,14	14,07	4	12,07	1,54
15,56	12	15,56	2,89	21,42	4,35	21,47	2	21,47	3,21
6,18	27	5,18	1,03	14,04	2,31	14,04	6	14,04	1,85
5,51	20	6,08	1,05	26,01	2,15	26,01	4	26,01	3,65
7,92	26	17,92	2,38	6,75	1,36	6,15	27	6,52	1,09
8,71	25	8,76	1,76	11,04	3,95	11,04	11	11,04	0,97
9,38	21	10,23	1,83	14,51	2,35	14,57	2	14,51	2,92
10,26	19	9,32	1,13	14,23	3,45	14,23	5	14,23	1,05
12,29	16	12,24	1,91	12,13	2,27	12,53	14	12,13	2,02
11,45	17	13,85	2,24	7,15	1,35	7,85	25	7,85	0,12
17,35	9	16,28	2,74	10,52	2,61	10,52	11	10,52	2,87
22,18	6	24,18	3,55	18,67	4265	18,67	4	18,67	2,18
21,74	5	22,18	3,58	11,47	1,08	11,47	10	11,47	1,64
22,18	8	21,88	3,64	14,85	2,25	14,75	3	14,85	1,08
30,61	3	30,61	5,23	5,42	0,87	5,42	30	5,44	0,82
15,94	6	15,29	1,26	28,54	5,93	28,54	2	28,54	4,45
12,65	9	12,65	1,13	25,41	5,14	25,41	3	25,48	4,42
17,96	6	15,96	2,96	9,23	1,35	9,203	19	9,25	1,32

Закінчення табл. 7.3

1		2		3		4		5	
9,58	21	12,85	1,68	28,85	6,68	28,85	2	28,85	4,55
21,43	9	23,83	3,25	26,64	5,01	26,64	4	26,64	4,34
13,85	12	14,19	2,87	19,86	4,93	19,86	12	19,86	1,69
12,88	6	12,38	1,27	25,34	4,65	25,34	7	25,64	4,39
15,72	13	16,27	1,84	19,21	3,15	19,21	5	19,21	3,33
22,65	8	21,85	4,42	16,62	3,08	16,66	11	16,26	2,96
Варіант 7.21		Варіант 7.22		Варіант 7.23		Варіант 7.24		Варіант 7.25	
Y	X	Y	X	Y	X	Y	X	Y	X
11,09	7	11,09	3,36	16,62	1,15	16,62	14	16,62	1,04
13,53	9	13,53	5,57	4,31	4,22	4,31	8	4,31	1,06
5,64	2	5,64	0,81	4,04	4,25	4,04	3	4,14	0,75
4,91	2	4,96	1,05	11,36	1,63	11,23	18	11,43	1,02
10,66	8	10,64	4,54	14,07	2,14	14,07	19	14,07	1,21
15,56	11	15,56	6,21	21,47	0,35	21,42	22	21,42	1,18
5,18	5	5,18	4,85	14,04	2,31	14,06	27	14,64	1,03
4,01	4	4,01	3,65	26,01	1,18	26,01	30	26,01	1,05
7,92	8	7,92	8,09	6,15	4,36	6,58	6	6,15	1,38
8,71	10	8,71	6,97	11,54	1,95	11,05	15	11,54	1,76
9,39	9	9,32	5,05	14,23	2,34	14,12	19	14,52	1,13
12,29	14	12,29	7,02	12,13	2,27	12,51	16	12,13	1,21
16,35	18	16,38	9,87	10,52	1,61	20,52	19	10,92	1,74
22,08	24	22,18	14,18	18,67	2,06	18,67	16	18,67	1,55
20,28	23	20,71	10,08	14,85	2,25	14,87	10	14,87	1,64
30,16	29	30,61	18,82	5,42	4,27	5,42	3	5,42	1,23
12,65	13	12,65	10,42	25,41	1,14	25,41	29	25,41	1,13
16,76	14	16,76	10,32	9,23	4,35	9,20	6	9,20	0,96
10,88	11	10,88	4,55	28,88	2,35	28,85	21	28,85	0,68
23,48	21	23,48	14,34	26,64	1,01	26,64	29	26,64	1,25
13,59	12	13,85	2,69	19,86	1,93	19,86	12	19,86	0,87
11,38	10	11,83	3,39	25,63	1,14	25,34	26	25,63	1,27
15,27	9	15,72	4,33	29,21	0,15	19,21	13	19,21	1,34
22,65	19	22,68	12,96	16,62	1,08	16,66	18	16,62	1,42
31,61	28	32,25	15,36	26,12	0,54	5,44	4	4,42	1,33
24,53	20	13,58	3,57	24,62	1,27	4,81	8	4,51	1,26

7.6. Контрольні запитання

1. Який зв'язок називається кореляційним?
2. Поясніть, чи завжди наявність кореляційного зв'язку означає наявність причинно-наслідкового зв'язку.

3. Як визначити функцію, за якої здійснюють апроксимацію?
4. Поясніть, чому одночасно можуть розглядатись спряжені лінії регресії Y на X та X на Y .
5. Запишіть модель парної лінійної регресії і наведіть тлумачення її параметрів.
6. Що таке похибка моделі? Як вона визначається?
7. У чому полягає метод найменших квадратів?
8. Які статистичні гіпотези розглядаються при перевірці значущості параметрів моделі?
9. Які висновки можна зробити на розташуванням спряжених ліній регресії?

Лабораторна робота 8

Елементи дисперсійного аналізу

8.1. Мета заняття:

набути навички застосування дисперсійного аналізу та перевірити значущість рівняння регресії в цілому (за критерієм Фішера), тобто визначити, наскільки мінливість функціонального фактора Y , що пов'язана з впливом випадкової величини X (регресора), перевищує мінливість функціонального фактора, що пов'язана з випадковою похибкою;

визначити межі довірчого інтервалу, до якого лінія регресії належатиме з довірчою ймовірністю 95 %, тобто межі можливих змін лінії регресії (від вибірки до вибірки) за значенням залишкової похибки лінійної регресійної моделі.

8.2. Теоретичні відомості.

Вибіркова дисперсія функціонального фактора Y пов'язана з загальною сумою квадратів відхилень, або SST , яку можна подати як суму двох взаємно незалежних доданків:

$$SST = SSR + SSE, \quad (8.1)$$

де SSR – сума квадратів відхилення значень функціонального фактора від його середньої, що обумовлені регресією (**Regression Square Sum**):

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2;$$

SSE – сума квадратів відхилення, що пов'язана з помилками моделі (**Error Square Sum**):

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2;$$

SST – загальна сума квадратів відхилень значень функціонального фактора від його середньої (**Total Square Sum**):

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2.$$

Відповідно до рівняння (8.1) можна записати співвідношення для кількості степенів свободи (**degrees of freedom**):

$$df_T = df_R + df_E, \quad (8.2)$$

де df_T – загальна кількість степенів свободи: $df_T = n - 1$;

df_R – кількість степенів свободи, що пов'язана з регресією: $df_R = 1$;

df_E – кількість степенів свободи, що пов'язана з помилками моделі.

Відповідно: $df_E = n - 2$.

Процедура дисперсійного аналізу полягає у визначенні співвідношення між систематичною складовою (або SSR) загальної дисперсії (або SST) та її випадковою складовою (або SSE) у вимірюваних даних.

Кожний доданок рівняння (8.1) має сенс частки загальної дисперсії:

$$1 = \frac{SSR}{SST} + \frac{SSE}{SST}. \quad (8.3)$$

Так, відношення $\frac{SSR}{SST} = R^2$ визначає частку загальної мінливості

функціонального фактора, яка згідно з регресійною моделлю обумовлена впливом фактор-аргумента X , і має назву **коефіцієнта детермінації**. Отже, коефіцієнт детермінації визначає, яку частину загальної мінливості функціонального фактора можна пояснити в межах моделі.

Корінь квадратний з коефіцієнта детермінації називається **кореляційним відношенням**:

$$\eta = \sqrt{\frac{SSR}{SST}}, \quad 0 \leq \eta \leq 1.$$

Кореляційне відношення є характеристикою щільності кореляційного зв'язку незалежно від форми цього зв'язку, тоді як коефіцієнт кореляції Пірсона є характеристикою щільності лінійного кореляційного зв'язку. Якщо випадкові величини Y та X незалежні, то $\eta = 0$. Навпаки, якщо випадкові величини Y та X пов'язані лінійним кореляційним зв'язком і розпорошення випадкової величини Y під впливом неврахованих факторів відсутнє, тобто загальна дисперсія випадкової величини Y визначається тільки впливом зміни випадкової величини X , то $\eta = 1$. У випадку лінійного кореляційного зв'язку $\eta = |r|$.

Вибіркове кореляційне відношення, як і вибірковий коефіцієнт кореляції, є статистичною оцінкою відповідної числової характеристики генеральної сукупності.

Відношення $\frac{SSE}{SST} = 1 - R^2$ визначає частку від загальної мінливості функціонального фактора, яка обумовлена впливом факторів, що не враховуються в межах обраної моделі.

Для перевірки значущості рівняння регресії в цілому, тобто впливовості фактор-аргумента X , який у межах даної моделі розглядається як регресор, на значення функціонального фактора Y , використовується **критерій Фішера** F . Цей критерій визначає, на скільки мінливість функціонального фактора Y , що пов'язана з впливом фактор-аргумента X , перевищує мінливість фактора Y під впливом факторів, які не розглядаються у межах даної моделі. Отже, за означенням:

$$F = \frac{MSR}{MSE} = \frac{SSR / df_R}{SSE / df_E}, \quad \text{за умов, що } MSR > MSE, \quad (8.4)$$

де MSR – питома сума квадратів, що пов'язана з регресією (**Regression Mean Square**);

MSE – питома сума квадратів, що обумовлена помилками (**Error Mean Square**).

Для моделі парної регресії маємо:

$$F = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / 1}{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / (n-2)}, \text{ або } F = \frac{r^2}{1-r^2} \frac{n-2}{1}. \quad (8.5)$$

Критичні точки для критерію Фішера визначаються за допомогою таблиці значень F -розподілу з урахуванням кількості степенів свободи, які дорівнюють $df_R = 1$ (для чисельника) та $df_E = n - 2$ (для знаменника).

Якщо емпіричне значення критерію не перевищує критичного значення для рівня значущості $\alpha = 0,05$, тобто $F_{\text{емп.}} \leq F_{0,05}(1; n-2)$, то з надійністю 95 % гіпотезу $H_0: r = 0$ нема підстав відхилити, тобто детермінована мінливість функціонального фактора (під впливом фактор-аргумента) не перевищує його мінливості під впливом факторів, які не враховані моделлю. Звідси випливає, що кореляційний зв'язок є статистично незначущим.

Якщо емпіричне значення критерію перевищує $F_{0,05}(1; n-2)$, то гіпотеза H_0 відхиляється на користь альтернативної $H_1: r \neq 0$, отже, кореляційний зв'язок є значущим.

Оскільки параметри вибіркового рівняння регресії є статистичними оцінками відповідних характеристик генеральної сукупності, то лінія регресії, що побудована за цим рівнянням, теж є статистичною оцінкою. Аналогом інтервальної оцінки лінії регресії є **довірча смуга**, до якої з заданою надійністю γ (довірчою ймовірністю $P = 1 - \alpha$) належатиме у генеральній сукупності лінія регресії. У випадку лінійної регресії напівширина довірчої смуги обчислюється за формулою:

$$\Delta \hat{y}(x) = t_{\alpha} \cdot \hat{\sigma}_e \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \quad (8.6)$$

Верхня та нижня межі довірчої смуги лінії регресії Y на X визначаються як $\hat{y} + \Delta\hat{y}$ та $\hat{y} - \Delta\hat{y}$, відповідно. Рівняння (8.5) є рівнянням гіперболи в ортогональній системі координат, початок якої знаходиться в центрі вибіркової сукупності, тобто у точці з координатами $(\bar{x}; \bar{y})$, а уявною віссю гіперболи є графік лінії регресії. При обчисленні меж довірчої смуги з використанням даних, що виводяться за допомогою вбудованою функції **ЛИНЕЙН** MS Excel, зручно також використовувати перетворену формулу (8.6):

$$\Delta\hat{y}(x) = t_{\alpha} \cdot \sigma_Y \sqrt{\frac{1-r^2}{n-2}} \cdot \sqrt{1 + \frac{(x-\bar{x})^2}{\sigma_X^2}}, \quad (8.6^I)$$

де σ_X та σ_Y – середні квадратичні відхилення випадкових величин X та Y , відповідно.

Оскільки величина стандартної помилки

$$\hat{\sigma}_e = \sqrt{\frac{SSE}{df_E}} = \sqrt{\frac{SSE}{n-2}} = \sqrt{MSE}$$

виводиться у таблиці як результат застосування функції **ЛИНЕЙН** (третій рядок правого стовпчика), то формулу (8.6) також зручно застосовувати у вигляді:

$$\Delta\hat{y}(x) = t_{\alpha} \cdot \sqrt{\frac{MSE}{n}} \cdot \sqrt{1 + \frac{(x-\bar{x})^2}{\sigma_X^2}}. \quad (8.6^{II})$$

8.3. Змістовна постановка задачі

У табл. 8.1 наведені дані про термін служби обладнання X та середньомісячну продуктивність праці при розробці вугільної лави Y .

За вихідними даними необхідно:

побудувати рівняння регресії у припущенні лінійної форми кореляційного зв'язку;

за критерієм Фішера перевірити значущість кореляційного зв'язку;

побудувати довірчий інтервал, до якого лінія регресії належатиме з надійністю 95 %.

Таблиця 8.1

Вихідні дані для кореляційного аналізу

№	x_i	y_i	№	x_i	y_i	№	x_i	y_i
1	2	30,291	8	10	28,585	15	12	3,720
2	2	13,440	9	10	13,350	16	12	11,350
3	4	35,990	10	10	7,596	17	12	9,519
4	5	20,714	11	11	9,817	18	13	13,510
5	8	25,756	12	11	9,555	19	13	21,64
6	8	13,613	13	11	13,762	20	18	10,894
7	10	10,076	14	12	14,386	21	22	9,220

8.4. Приклад виконання лабораторної роботи 8

На робочому аркуші *MS Excel* розміщуємо таблицю, де вихідні дані, що містять значення кожного із факторів, записуються у вигляді стовпчиків (стовпці 2 і 3 табл. 8.2). Отже, вибірка сукупність містить 21 вимірювання.

Таблиця 8.2

Вихідні дані та результати розрахунків для побудови лінії регресії та її довірчого інтервалу

№	x_i	y_i	\hat{y}_i	$(x_i - \bar{x})^2$	$\Delta\hat{y}_i$	$\hat{y}_i - \Delta\hat{y}_i$	$\hat{y}_i + \Delta\hat{y}_i$
	2	3	4	5	6	7	8
1	2	30,291	23,599	68,653	6,8439	16,7555	30,4433
2	2	13,440	23,599	68,653	6,8439	16,7555	30,4433
3	4	35,990	21,659	39,510	5,6221	16,0371	27,2812
4	5	20,714	20,689	27,939	5,0557	15,6333	25,7447
5	8	25,756	17,779	5,224	3,6998	14,0788	21,4784
6	8	13,613	17,779	5,224	3,6998	14,0788	21,4784
7	10	10,076	15,838	0,082	3,3167	12,5216	19,1550
8	10	28,585	15,838	0,082	3,3167	12,5216	19,1550
9	10	13,350	15,838	0,082	3,3167	12,5216	19,1550
10	10	7,596	15,838	0,082	3,3167	12,5216	19,1550
11	11	9,817	14,868	0,510	3,3503	11,5179	18,2185
12	11	9,555	14,868	0,510	3,3503	11,5179	18,2185
13	11	13,762	14,868	0,510	3,3503	11,5179	18,2185
14	12	14,386	13,898	2,939	3,5347	10,3634	17,4327
15	12	3,720	13,898	2,939	3,5347	10,3634	17,4327

1	2	3	4	5	6	7	8
16	12	11,350	13,898	2,939	3,5347	10,3634	17,4327
17	12	9,519	13,898	2,939	3,5347	10,3634	17,4327
18	13	13,510	12,928	7,367	3,8482	9,0798	16,7761
19	13	21,64	12,928	7,367	3,8482	9,0798	16,7761
20	18	10,894	8,077	59,510	6,4854	1,5918	14,5627
21	22	9,220	4,197	137,224	9,0927	-4,8960	13,2894
Σ	216	326,784		440,286			

Нижче (під табл. 8.2) виділяємо прямокутну сукупність розміром 5x2, натискаємо клавішу "=" і викликаємо функцію **ЛИНЕЙН** (табл. 8.3). У її діалоговому вікні заповнюємо поля таким чином, як було визначено у лабораторній роботі 7.

Після натискання **Ctrl+Shift+Enter** виводиться інформація щодо характеристик розподілу двовимірної випадкової величини (X, Y).

Таблиця 8.3

Результат застосування функції ЛИНЕЙН

-0,97014	25,53968
0,345406	3,888879
0,293383	7,247652
7,888689	19
414,3807	998,0408

За першим рядком табл. 8.3 маємо рівняння регресії:

$$\hat{y} = 25,54 - 0,97 \cdot x. \quad (8.6)$$

Відповідно до рівняння (8.6) у стовпчику 4 табл. 8.2 обчислюємо теоретичні значення \hat{y}_i , що відповідають кожному емпіричному значенню фактор-аргумента X . Це зручно робити за допомогою формули, що містить абсолютні посилання на ті комірки (табл. 8.3), де виведені значення параметрів моделі.

Тепер обчислюємо напівширину довірчого інтервалу. За табл. 8.3 маємо, що $\sqrt{MSE} = 7,247652$.

Дійсно, за даними тієї ж табл. 8.3 маємо:

$$\sqrt{MSE} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{998,0408}{19}} = 7,247652.$$

За допомогою функцій **СРЗНАЧ** та **ДИСП.В** обчислюємо вибірку середню та дисперсію випадкової величини X : $\bar{x} = 10,29$ та $\sigma_X^2 = 20,97$, або $\bar{x} = \frac{216}{21} = 10,29$ та $\sigma_X^2 = \frac{440,286}{21} = 20,97$ (за табл. 8.2).

До 5-го стовпчика (див. табл. 8.2) заносимо відхилення емпіричного значення випадкової величини X від її середньої і за цими даними обчислюємо напівширину довірчого інтервалу $\Delta\hat{y}_i$ (за допомогою функції **СТЮДЕНТ.ОБР** знаходимо, що $t_{0,05}(19) = 2,093024$) та записуємо ці значення до 6-го стовпчика табл. 8.2.

Для побудови гіперболи, гілки якої є межами довірчого інтервалу, до якого лінія регресії генеральної сукупності належатиме з надійністю 95 %, заповнюємо 7-ий та 8-ий стовпчики табл. 8.2, записавши туди значення $\hat{y}_i - \Delta\hat{y}_i$ та $\hat{y}_i + \Delta\hat{y}_i$, відповідно.

Далі будуємо всі лінії, про які йшла мова вище. Для цього вказуємо шлях: **Вставка** \Rightarrow **Діаграми** \Rightarrow **Точечная** \Rightarrow **Точечная с гладкими кривыми**. Після цього область діаграми буде створено (рис. 8.1).

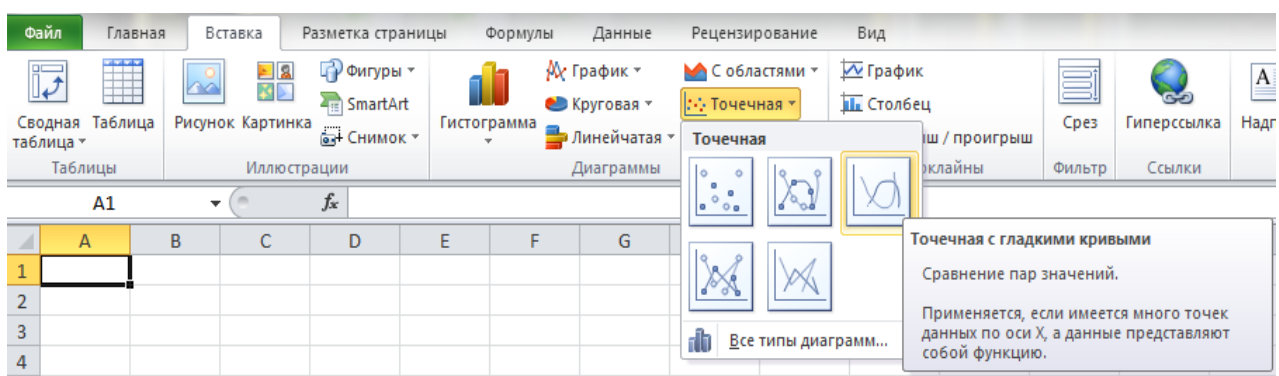


Рис. 8.1. Вставка діаграми відповідного типу

Для виведення даних необхідно натиснути на праву кнопку миші, вибрати з контекстного меню **Выбрать данные** (рис. 8.2) і натиснути **Enter**.

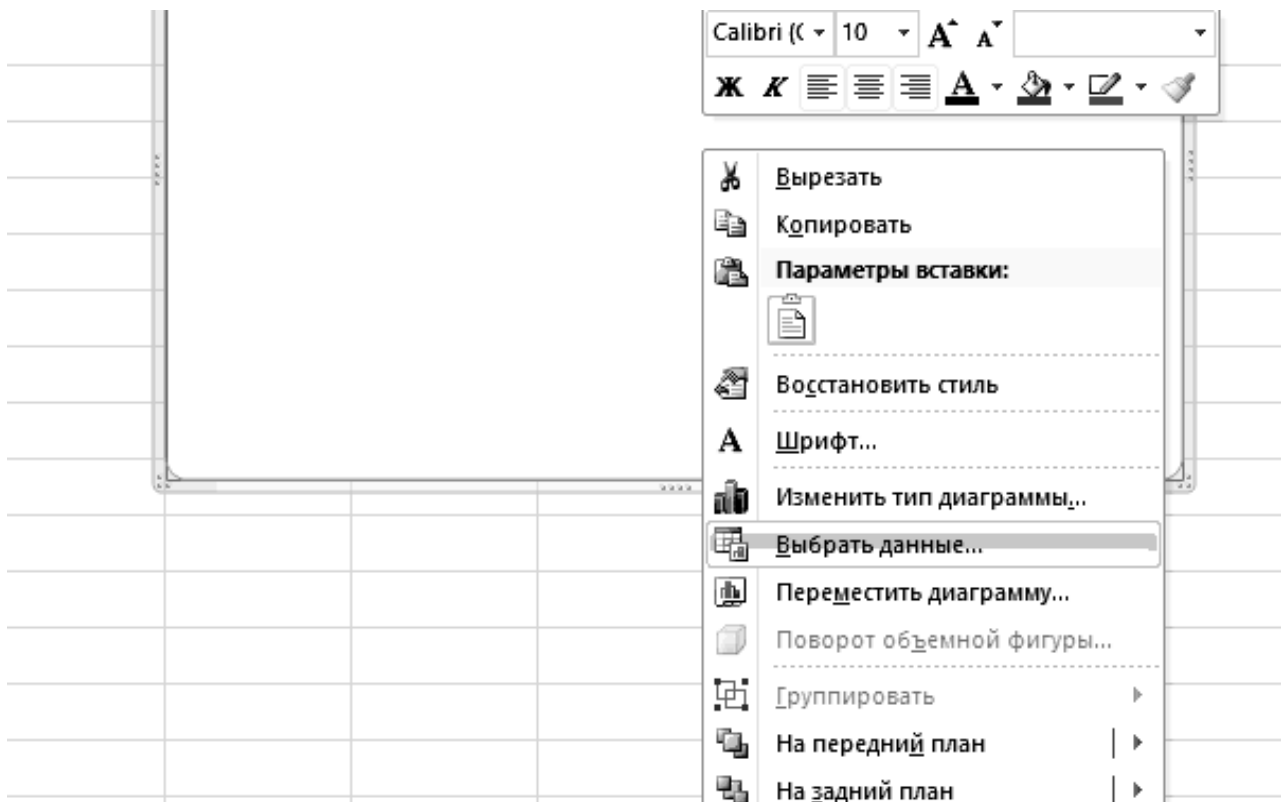


Рис. 8.2. Вибір даних

Унаслідок цього на екрані з'являється діалогове вікно **Выбор источника данных** (рис. 8.3), у якому необхідно натиснути **Добавить**.

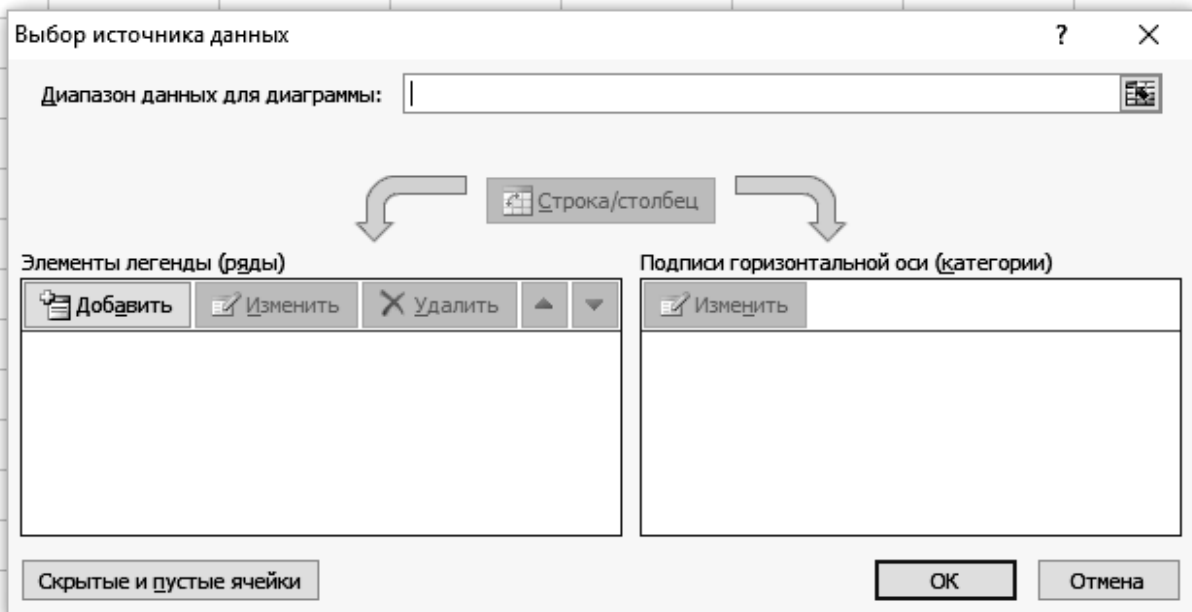


Рис. 8.3. Діалогове вікно *Выбор источника данных*

Далі у діалоговому вікні **Изменение ряда**, що з'явилося на екрані, заповнюємо його поля (рис. 8.4).

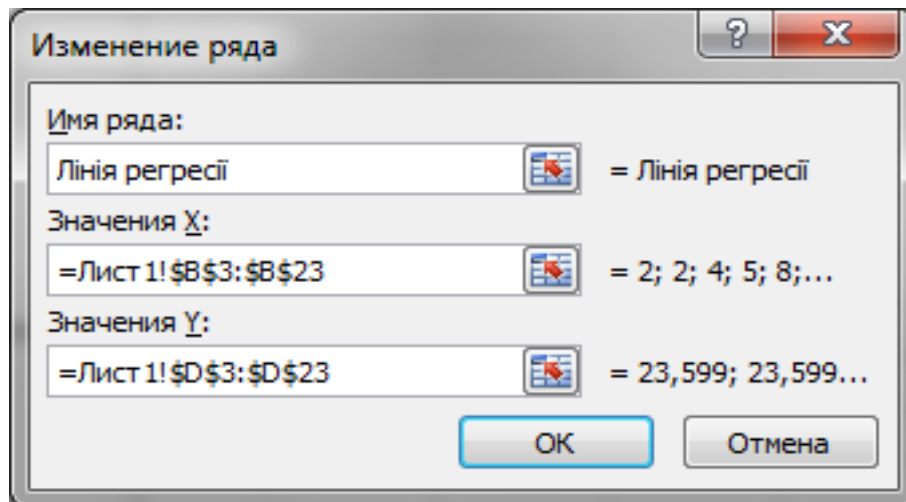


Рис. 8.4. Діалогове вікно **Изменение ряда**

Так, у поле **Значения X** вносимо вміст 2-го стовпчика, а в поле **Значения Y** – вміст 4-го стовпчика, **Имя ряда** – назву "Линия регрессии" (теоретична лінія регресії). Натискаємо **OK** і на діаграмі з'являється графік теоретичної лінії регресії. Після цього знову викликаємо діалогове вікно **Выбор источника данных**, яке тепер має вигляд (рис. 8.5).

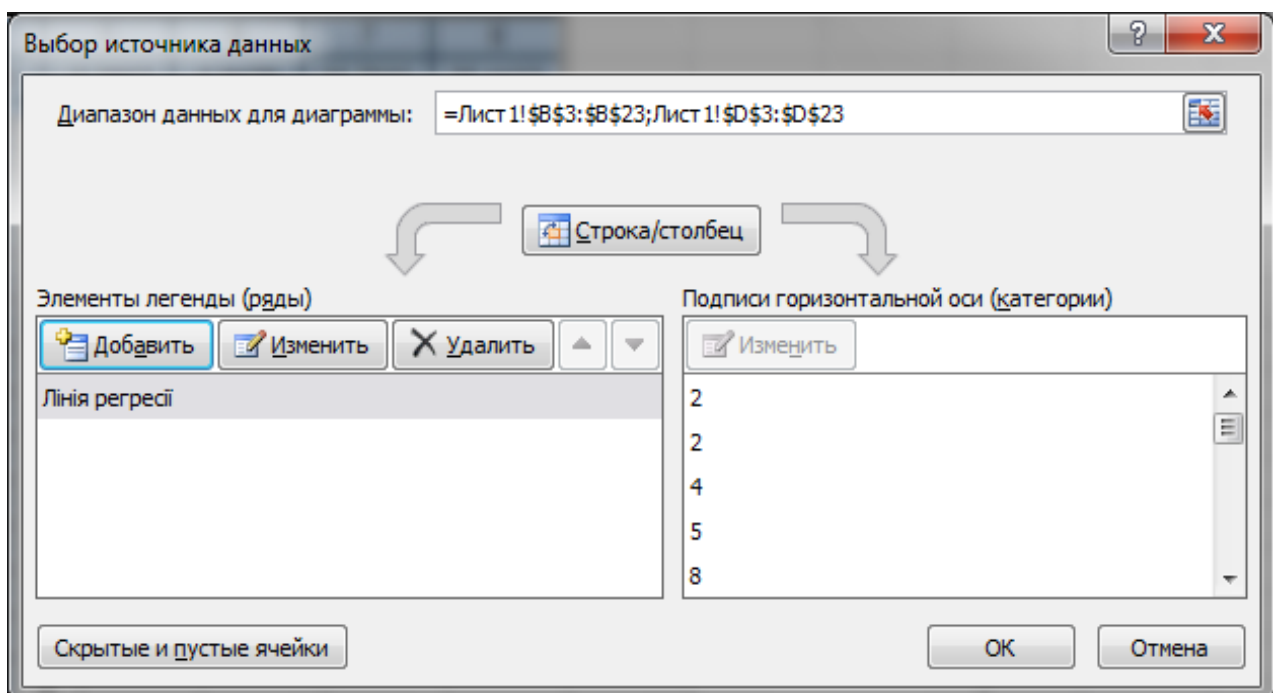


Рис. 8.5. Діалогове вікно **Выбор источника данных**

Набираємо послідовність команд: **Добавить** \Rightarrow **Ряд 2** \Rightarrow **Имя ряда** – Нижня межа (нижня гілка гіперболи) \Rightarrow **Значения X** – вміст 2-го стовпчика \Rightarrow **Значения Y** – вміст 7-го стовпчика.

Аналогічно додаємо графік верхньої гілки гіперболи: **Добавить** \Rightarrow **Ряд 3** \Rightarrow **Имя ряда** – Верхня межа \Rightarrow **Значения X** – вміст 2-го стовпчика \Rightarrow **Значения Y** – вміст 8-го стовпчика.

Тепер нанесемо вихідні дані. Для цього у вікні **Выбор источника данных** додаємо ще один рядок: **Добавить** \Rightarrow **Ряд 4** \Rightarrow **Имя ряда** – вихідні дані \Rightarrow **Значения X** – вміст 2-го стовпчика, тобто хмара розпорошення значень \Rightarrow **Значения Y** – вміст 3-го стовпчика. Для зміни типу діаграми заходимо правою клавішею миші на будь-яку точку ряду вихідних даних та з контекстного меню вибираємо **Изменить тип диаграммы для ряда**.

Далі у діалоговому вікні **Изменение типа диаграммы** вибираємо **Точечная с маркерами** і натискаємо **ОК**. Побудову графіків завершено (рис. 8.6). Отже, ми отримали лінію регресії $\hat{y} = b_0 + b_1 \cdot x$ та довірчий інтервал, до якого вона належить з імовірністю 0,95, а також емпіричні точки, за якими була побудована модель парної лінійної регресії.

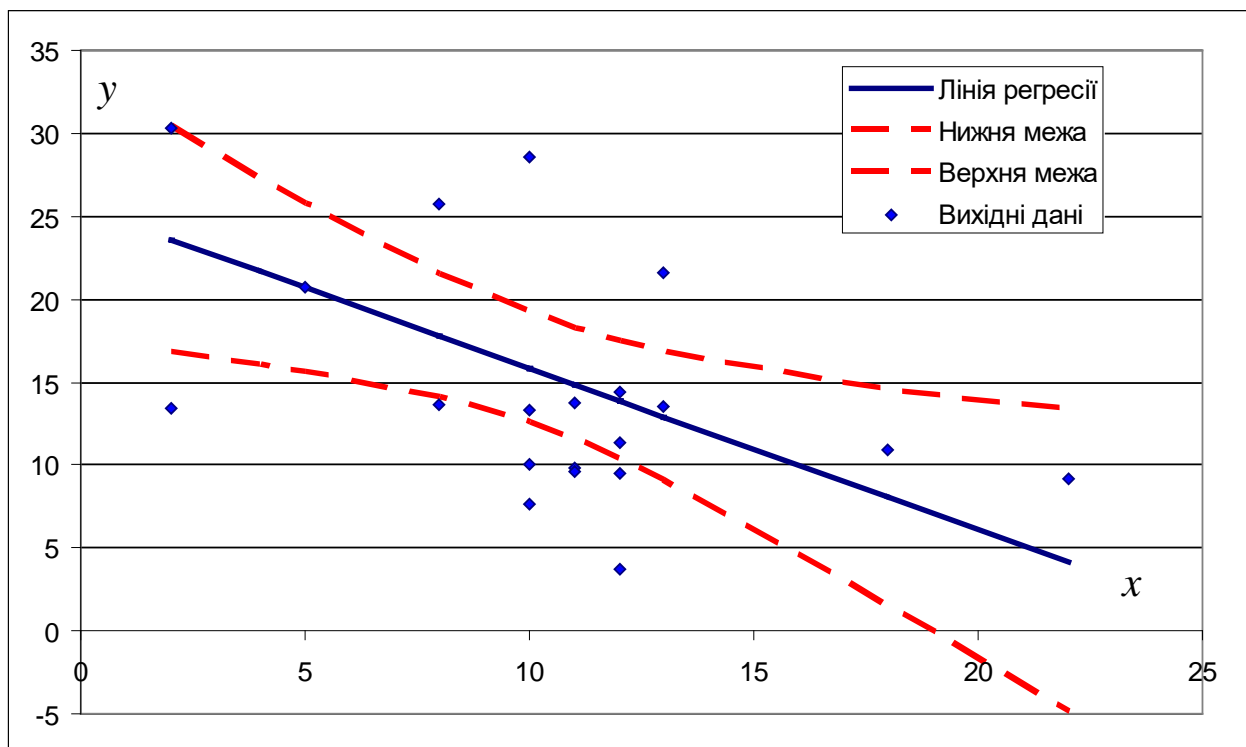


Рис. 8.6. Теоретична лінія регресії та довірчий інтервал, до якого лінія регресії належатиме з надійністю 95 %

Наявність довірчої смуги дозволяє встановити межі використання моделі. Найбільш надійні результати застосування моделі (обрахунки за цією моделлю дають невелику довірчу похибку) одержуємо в такому околі центра групування даних: $(\bar{x} - S_x, \bar{x} + S_x)$. Із збільшенням відстані від центра довірча похибка швидко зростає (гілки гіперболи розходяться).

Перевіримо значущість моделі в цілому за критерієм Фішера. Відповідно до гіпотези $H_0: R^2 = 0$. Альтернативна гіпотеза $H_1: R^2 > 0$.

Отже, ми перевіряємо значущість впливу фактор-аргумента X на функціональний фактор Y .

Повернемось до результатів застосування функції **ЛИНЕЙН**, що наведені в табл. 8.3. У четвертому рядку лівого стовпчика беремо емпіричне значення критерію Фішера $F = 7,888689$.

Аналогічний результат також можна отримати за даними п'ятого рядка табл. 8.3, якщо застосувати формулу (8.4):

$$F = \frac{SSR / df_R}{SSE / df_E} = \frac{414,3807 / 1}{998,0408 / 19} = 7,888689.$$

Критичні точки критерію Фішера визначаємо за допомогою функції **Ф.ОБР**, що належить до категорії **Статистические**. Викликаємо цю функцію і заповнюємо поля її діалогового вікна (рис. 8.7).

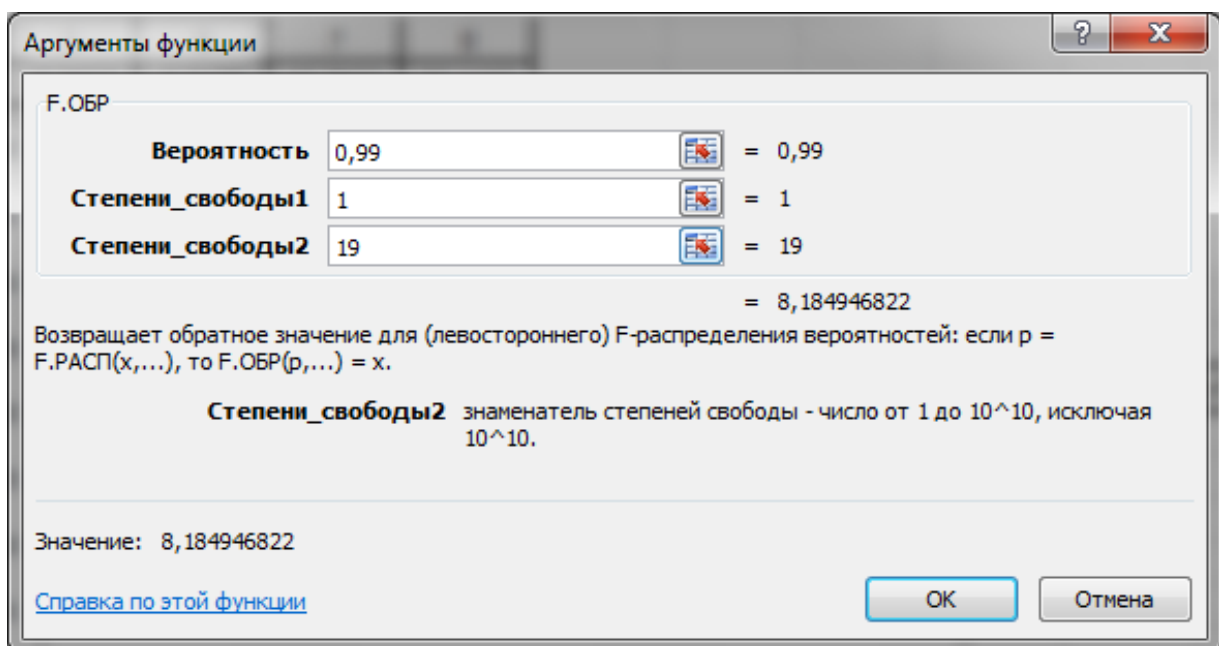


Рис. 8.7. Діалогове вікно **Аргументы функции Ф.ОБР**

Слід мати на увазі, що у поле **Вероятность** для рівня значущості $\alpha = 0,01$ необхідно вводити 0,99. Отже, з урахуванням кількості степенів свободи формула має вигляд: **F.ОБР(0,99;1;19)** і $F_{0,01}(1; 19) = 8,184947$. Для рівня значущості $\alpha = 0,05$ маємо $F_{0,05}(1; 19) = 4,38075$. Результати надані в табл. 8.4.

Таблица 8.4

Результати обчислення критичних значень критерію Фішера

$F_{0,01}(1,19)=$	F.ОБР (0,99;1;19)=	8,184947
$F_{0,05}(1,19)=$	F.ОБР (0,95;1;19)=	4,38075

Порівняємо емпіричне значення F -критерію з критичними точками розподілу Фішера. Маємо, що $F_{0,05} < F_{емп} < F_{0,01}$, отже, емпіричне значення критерію належить до області невизначеності. Оскільки помилки другого роду мають більш погані наслідки, то нульову гіпотезу, згідно з якою $R^2 = 0$, краще відкинути, тобто кореляційний зв'язок доцільно вважати статистично значущим.

Перевіряємо результати за допомогою надбудови **Анализ данных** (рис. 8.8). Вибираємо шлях **Данные** \Rightarrow **Анализ данных** \Rightarrow **Регрессия**.

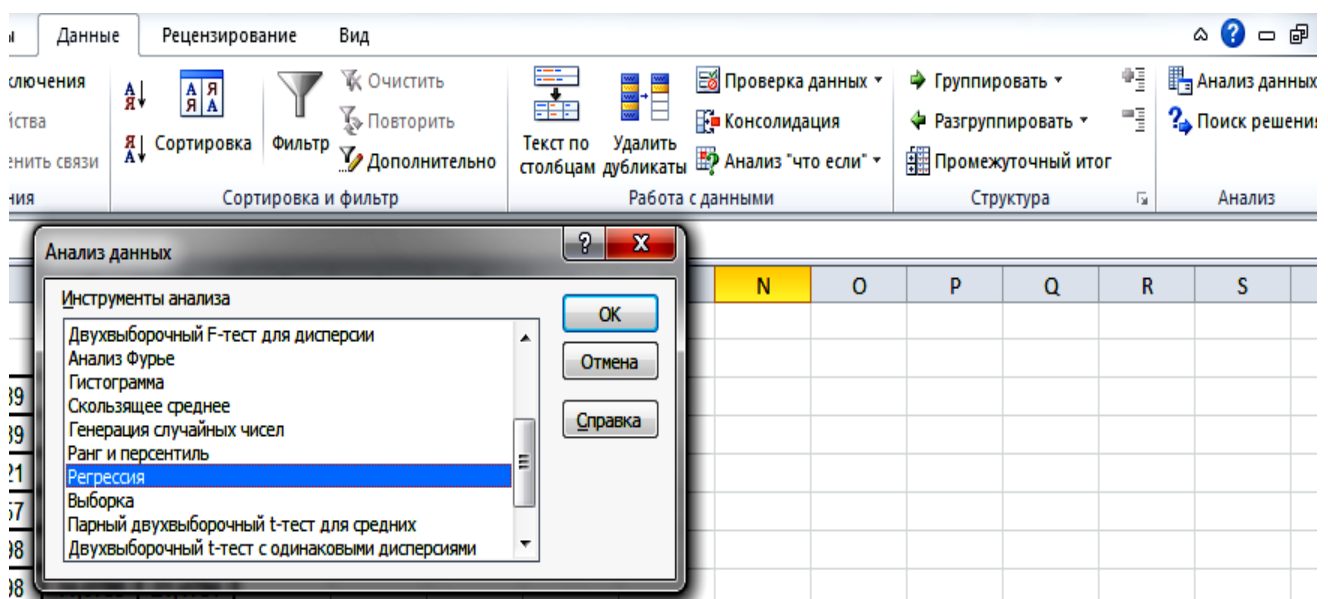


Рис. 8.8. Виклик діалогового вікна процедури **Регрессия**

Заповнюємо її діалогове вікно, відповідаючи на запитання (рис. 8.9).

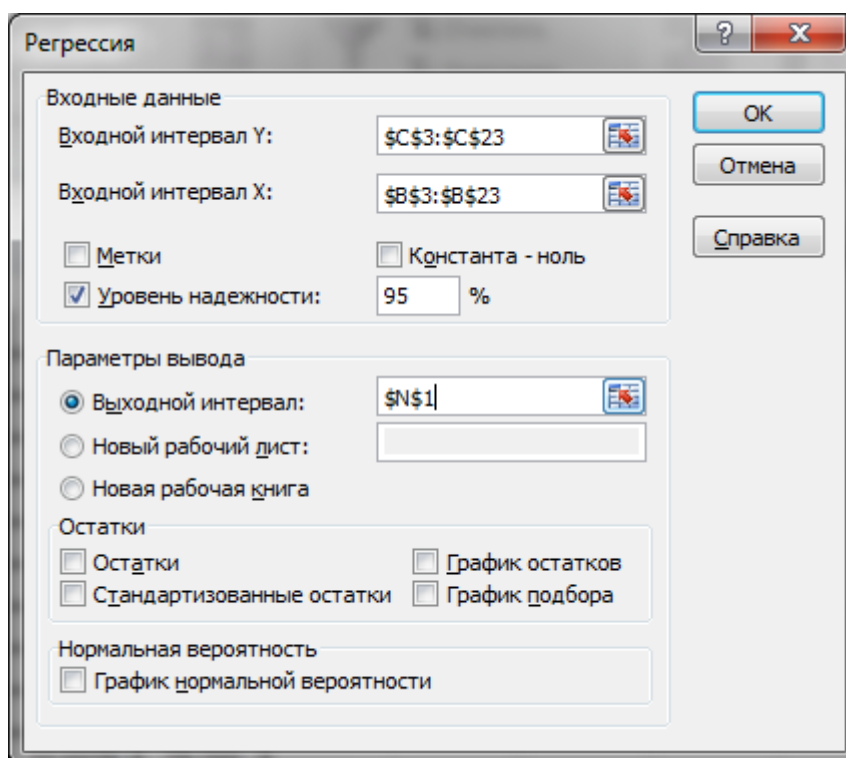


Рис. 8.9. **Діалогове вікно Регрессия надбудови Анализ данных**

Результати кореляційного, регресійного та дисперсійного аналізів виводяться у вигляді окремих таблиць під загальною назвою **ВЫВОД ИТОГОВ**. Пропонуємо самостійно порівняти ці дані з результатами попередніх досліджень і зробити висновки.

8.5. Завдання для самостійної роботи

За допомогою вбудованих функцій і надбудов *MS Excel* необхідно:
побудувати рівняння регресії у припущенні лінійної форми кореляційного зв'язку;
за критерієм Фішера перевірити значущість моделі в цілому;
побудувати довірчий інтервал, до якого лінія регресії належатиме з надійністю 95 %.

Варіанти завдань для самостійного розв'язання наведені в лабораторній роботі 7.

8.6. Контрольні запитання

1. Що досліджує дисперсійний аналіз?
2. Що таке довірна смуга для вибіркової лінії регресії?

3. Як визначаються межі довірчої смуги, до якої теоретична лінія регресії повинна належати з наперед заданою імовірністю?
4. Поясніть, як за шириною довірчої смуги оцінити значущість кореляційного зв'язку.
5. Спираючись на дані щодо ширини довірчої смуги, охарактеризуйте точність розрахунків за рівнянням лінійної регресії.
6. Розкрийте поняття значущості регресійної моделі.
7. Наведіть формули для обчислення суми квадратів регресії (SSR) та суми квадратів похибок (SSE). Що таке стандартна похибка?
8. Про що свідчить значення коефіцієнта детермінації?
9. Що таке кореляційне відношення?
10. Як співвідносяться між собою коефіцієнт кореляції Пірсона та кореляційне відношення?
11. Наведіть статистичні гіпотези щодо перевірки значущості моделі лінійної регресії за допомогою критерію Фішера.
12. Як визначається критичне значення критерію Фішера (рівень значущості, кількість ступенів вільності)?

3. Рекомендована література

3.1. Основна

1. Малярець Л. М. Математика для економістів. Теорія ймовірностей та математична статистика : навч. посіб. У 3-х ч., ч.3 / Л. М. Малярець, І. Л. Лебедева, Л. Д. Широкоград. – Х. : Вид. ХНЕУ, 2011. – 568 с.

2. Малярець Л. М. Практикум з теорії ймовірностей та математичної статистики в Excel : навч.-практ. посіб. / Л. М. Малярець, І. Л. Лебедева, Е. Ю Железнякова. – Х. : Вид. ХНЕУ, 2007. – 160 с.

3. Теорія ймовірностей та математична статистика : навч. посіб. / Л. М. Малярець, І. Л. Лебедева, Е. Ю. Железнякова та ін. – Х. : Вид. ХНЕУ, 2010. – 404 с.

3.2. Додаткова

4. Барковський В. В. Теорія ймовірностей та математична статистика / В. В. Барковський, Н. В. Барковська, О. К. Лопатін. – 5-е вид. – К. : Центр учбової літератури, 2010. – 424 с.

5. Елисеєва І. І. Теорія статистики с основами теорії вероятностей : учеб. пособ. / І. І. Елисеєва, В. С. Князевский. – М. : ЮНИТИ-ДАНА, 2001. – 446 с.

6. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посіб. У 2 ч. – Ч. I Теорія ймовірностей / В. І. Жлуктенко, С. І. Наконечний. – К. : КНЕУ, 2000. – 304 с.

7. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посіб. У 2 ч. – Ч. II Математична статистика / В. І. Жлуктенко, С. І. Наконечний, С. С. Савіна. – К. : КНЕУ, 2001. – 336 с.

8. Кремер Н. Ш. Теория вероятностей и математическая статистика / Н. Ш. Кремер. – М. : ЮНИТИ-ДАНА, 2000. – 544 с.

9. Теория вероятностей и математическая статистика : учеб. пособ. / В. А. Ватулин, Г. И. Ивченко, Ю. И. Медведев и др. – 2-е изд., испр. – М. : Дрофа, 2003. – 328 с.

3.3. Інформаційні ресурси

10. Доклад ЮНЕСКО по науке: на пути к 2030 году [Электронный ресурс]. – Режим доступа : <http://unesdoc.unesco.org/images/0023/002354/235407r.pdf>

11. Регіони України: статистичний збірник / за ред. О. Г. Осауленка. – К. : Державна служба статистики України, 2013. – 783 с.

Зміст

Вступ.....	3
1. План лабораторних робіт	7
2. Зміст лабораторних робіт та інструкційна карта їх виконання.....	9
Змістовий модуль 1. Теорія ймовірностей.....	9
Лабораторна робота 1. Класичне означення ймовірності та елементи комбінаторики. Основні теореми теорії ймовірностей. Функції <i>MS Excel</i> для обробки даних.....	9
Лабораторна робота 2. Дискретні випадкові величини: основні числові характеристики та їх властивості.....	42
Лабораторна робота 3. Дискретна випадкова величина: біноміальний закон розподілу та його асимптотичні наближення.....	59
Лабораторна робота 4. Двовимірна дискретна випадкова величина, її закон розподілу та числові характеристики	88
Змістовий модуль 2. Математична статистика.....	104
Лабораторна робота 5. Основні поняття математичної статистики. Статистичне оцінювання параметрів розподілу	104
Лабораторна робота 6. Перевірка за критерієм Пірсона статистичної гіпотези щодо нормального закону розподілу випадкової величини в генеральній сукупності	126
Лабораторна робота 7. Елементи теорії кореляції. Елементи регресійного аналізу	151
Лабораторна робота 8. Елементи дисперсійного аналізу	166
3. Рекомендована література.....	181
3.1. Основна	181
3.2. Додаткова	181
3.3. Інформаційні ресурси.....	181

НАВЧАЛЬНЕ ВИДАННЯ

Железнякова Еліна Юріївна
Лебедєва Ірина Леонідівна
Норік Лариса Олексіївна
Стєпанова Катерина Вадимівна

**Лабораторний практикум
з навчальної дисципліни
"ТЕОРІЯ ЙМОВІРНОСТЕЙ
ТА МАТЕМАТИЧНА СТАТИСТИКА"**

Навчальний посібник

Відповідальний за видання *Л. М. Малярець*

Відповідальний редактор *М. М. Оленіч*

Редактор *О. Г. Лященко*

Коректор *О. Г. Лященко*

План 2016 р. Поз. № 5-НП.

Підп. до друку 26.09.2016 р. Формат 60 x 90 1/16. Папір офсетний. Друк цифровий.

Ум. друк. арк. 11,5. Обл.-вид. арк. 14,38. Тираж 400 пр. Зам. № 164.

Видавець і виготовлювач – ХНЕУ ім. С. Кузнеця, 61166, м. Харків, просп. Науки, 9-А

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру

ДК № 4853 від 20.02.2015 р.