

В. В. БАРКОВСЬКИЙ
Н. В. БАРКОВСЬКА
О. К. ЛОПАТІН

**ТЕОРІЯ
ІМОВІРНОСТЕЙ ТА
МАТЕМАТИЧНА
СТАТИСТИКА**

5-те видання

НАВЧАЛЬНИЙ ПОСІБНИК

Київ
“Центр учбової літератури”
2010

УДК 519.2(075.8)
ББК 22.17я73
Б 25

Рецензенти:

Королюк Володимир Семенович — академік Національної академії наук України,
доктор фізико-математичних наук;

Віноградська Алла Вікторівна — кандидат фізико-математичних наук.

Барковський В.В., Барковська Н.В., Лопатін О.К.

Б 25 Теорія ймовірностей та математична статистика. 5-те видання. — Київ: Центр
учбової літератури, 2010. — 424 с.

ISBN 978-966-364-992-4

Навчальний посібник написаний відповідно до навчальної програми дисципліни «Вища математика для економістів» (Частина 2: «Теорія ймовірностей та математична статистика») для підготовки бакалаврів, спеціалістів та магістрів з економіки.

Посібник містить основні поняття, методи, теореми та формули; багато розв'язаних типових задач; необхідні таблиці; завдання для самостійної роботи студентів та поточного контролю якості одержаних знань та навичок; зразки контрольних робіт, індивідуального семестрового завдання, тести.

Рекомендований для студентів економічних спеціальностей вищих навчальних закладів та коледжів, які вивчають дисципліну «Теорія ймовірностей та математична статистика».

УДК 519.2(075.8)
ББК 22.17я73

ISBN 978-966-364-992-4

© Барковський В.В., Барковська Н.В.,
Лопатін О.К., 2010.

© Центр учбової літератури, 2010.

ЗМІСТ

Передмова до першого видання	9
Передмова до другого видання	11

Частина I. ТЕОРІЯ ІМОВІРНОСТЕЙ

Розділ 1. Основні поняття теорії імовірностей та комбінаторики	15
1.1. Предмет теорії імовірностей	15
1.2. Коротка історична довідка	16
1.3. Алгебра випадкових подій	17
1.4. Означення та властивості імовірності та частоти	22
1.5. Основні поняття та принципи комбінаторики	26
<i>Питання для самоперевірки та вправи</i>	<i>31</i>
Розділ 2. Основні теореми теорії імовірностей	38
2.1. Додавання імовірностей несумісних подій	38
2.2. Залежні та незалежні події, умовні імовірності	40
2.3. Множення імовірностей	41
2.4. Імовірність появи хоча б однієї випадкової події	44
2.5. Теорема додавання імовірностей сумісних подій	45
2.6. Надійність системи	46
2.7. Формули повної імовірності та Байєса	48
<i>Питання для самоперевірки та вправи</i>	<i>52</i>
Розділ 3. Послідовності випробувань	61
3.1. Схема та формула Бернуллі	61
3.2. Граничні теореми у схемі Бернуллі	65
3.3. Послідовність випробувань із різними імовірностями	70
3.4. Теорема Бернуллі	72
3.5. Проста течія подій	75
<i>Питання для самоперевірки та вправи</i>	<i>77</i>
Розділ 4. Випадкові величини	83
4.1. Види випадкових величин та способи їх задання	83
4.2. Закони розподілу та числові характеристики дискретних випадкових величин	89
4.2.1. Способи задання та закони розподілу	89

4.2.2. Числові характеристики	93
4.3. Числові характеристики законів розподілу неперервних випадкових величин	102
4.3.1. Числові характеристики неперервних випадкових величин	102
4.3.2. Закони розподілу НВВ та їх числові характеристики	105
4.4. Закон великих чисел та центральна гранична теорема	113
4.4.1. Загальні поняття	113
4.4.2. Нерівність Чебишова	114
4.4.3. Важливі граничні теореми	116
4.5. Закони розподілу та числові характеристики двовимірних випадкових величин	120
4.5.1. Загальні поняття	120
4.5.2. Закон розподілу імовірностей дискретної двовимірної випадкової величини	121
4.5.3. Неперервна двовимірна випадкова величина	125
4.5.4. Залежні та незалежні випадкові величини	126
4.5.5. Числові характеристики двовимірної випадкової величини ...	126
4.6. Функції випадкової величини та їх характеристики	128
4.6.1. Поняття функції	128
4.6.2. Закон розподілу та числові характеристики функції дискретного випадкового аргументу	129
4.6.3. Закон розподілу та числові характеристики функції неперервного випадкового аргументу	130
<i>Питання для самоперевірки та вправи</i>	134

Частина II. МАТЕМАТИЧНА СТАТИСТИКА

Розділ 5. Основні поняття та статистичний розподіл	153
5.1. Предмет математичної статистики та коротка історична довідка	153
5.2. Генеральна та вибіркова сукупності	154
5.3. Джерела даних у статистиці	156
5.4. Способи відбору	159
5.5. Проста випадкова вибірка	160
5.5.1. Умови здійснення простої випадкової вибірки	160
5.5.2. Випадкові вибіркові числа	161
5.5.3. Здійснення простої випадкової вибірки з використанням випадкових чисел	162
5.6. Організація даних: статистичний розподіл вибірки	165
5.6.1. Впорядкування даних	165
5.6.2. Розподіл частот	166
5.6.3. Згрупований розподіл накопиченої частоти	170

5.6.4. Розподіл відносної частоти (частоти) вибірки	171
5.6.5. Ряди розподілу частоти із змінною шириною класів інтервалів	173
5.6.6. Згрупований розподіл щільності частоти і щільності відносної частоти (частоти)	174
5.6.7. Загальна схема побудови згрупованого розподілу частот	175
5.7. Емпірична функція розподілу та її властивості	176
5.8. Графічне зображення статистичних розподілів	179
5.8.1. Незгруповані дані: полігони частот та частостей, гістограма	179
5.8.2. Згруповані дані: гістограма і полігон частот	188
5.8.3. Імовірнісний зміст гістограми та полігону частот	193
5.8.4. Полігони накопичених частот та частостей. Знаходження їх медіани	194
5.8.5. Графік емпіричної функції розподілу	197
<i>Питання для самоперевірки та вправи</i>	200
Розділ 6. Статистичні оцінки параметрів розподілу	203
6.1. Основні вимоги до статистичних оцінок	203
6.2. Числові характеристики вибіркової сукупності	206
6.2.1. Вибіркові характеристики	206
6.2.2. Обчислення вибіркових характеристик методом добутків	210
6.2.3. Статистичні моменти розподілу	213
6.2.4. Приклади знаходження статистик вибірки	214
6.3. Точкові та інтервальні оцінки	218
6.3.1. Загальні поняття	218
6.3.2. Довірчий інтервал для оцінки математичного сподівання нормального розподілу	219
6.4. Обробка вибірки методом найменших квадратів	223
6.4.1. Основні поняття	223
6.4.2. Оцінка параметрів лінійної функції	224
6.4.3. Оцінка параметрів параболічної функціональної залежності ...	226
<i>Питання для самоперевірки та вправи</i>	229
Розділ 7. Статистична перевірка гіпотез	233
7.1. Статистичні гіпотези та їх різновиди	233
7.2. Похибки перевірки гіпотез	235
7.3. Критерії узгодження для перевірки гіпотез	236
7.3.1. Статистичний критерій перевірки основної гіпотези	236
7.3.2. Критична область	237
7.3.3. Знаходження критичних областей	238
7.3.4. Порядок дій при перевірці статистичних гіпотез	239

7.4. Деякі критерії перевірки статистичних гіпотез	240
7.4.1. Перевірка гіпотези про рівність математичних сподівань нормальних генеральних сукупностей	240
7.4.2. Критерій дисперсійного аналізу	242
7.4.3. Критерій узгодження Пірсона (χ^2)	244
<i>Питання для самоперевірки та вправи</i>	248

Частина III. КОМП'ЮТЕРНИЙ АНАЛІЗ

Розділ 8. Розв'язання задач засобами Excel	253
8.1. Комп'ютерне розв'язання задач теорії імовірностей та математичної статистики у бізнесі	253
8.1.1. Огляд основних задач, що вирішені за допомогою Excel	253
8.1.2. Електронна таблиця Excel: елементи вікна та умовні скорочення	255
8.1.3. Статистичні функції Excel	258
8.2. Здійснення простої випадкової вибірки за допомогою функції «Выборка»	261
<i>Вправи для самостійної роботи</i>	266
8.3. Організація даних вибіркової сукупності	267
8.3.1. Побудова варіаційного ряду	267
8.3.2. Побудова розподілів частот за допомогою функції ЧАСТОТА	269
8.3.3. Побудова розподілів частот та накопичених частот і частостей	277
<i>Вправи для самостійної роботи</i>	282
8.4. Графічне представлення розподілів вибіркової сукупності	285
8.4.1. Графічне представлення розподілів за допомогою функції «Гистограмма»	285
8.4.2. Графічне представлення розподілів за допомогою «Мастера диаграмм»	298
<i>Вправи для самостійної роботи</i>	311
8.5. Знаходження числових характеристик вибірових даних	313
8.5.1. Знаходження числових характеристик за допомогою статистичних функцій	313
8.5.2. Знаходження числових характеристик вибірки у вигляді згрупованого розподілу частоти	318
<i>Вправи для самостійної роботи</i>	329
8.6. Комбінаторика	331
8.6.1. Обчислення факторіалу числа за допомогою функції ФАКТР	331
8.6.2. Обчислення сполучень за допомогою функції ЧИСЛОКОМБ	333
8.6.3. Обчислення розміщень за допомогою функції ПЕРЕСТ	335

8.7. Обчислення, пов'язані з законами розподілу випадкових величин	337
8.7.1. Біноміальний розподіл	337
8.7.2. Розподіл Пуассона	343
8.7.3. Нормальний розподіл	346
8.8. Побудова довірчого інтервалу для оцінки математичного сподівання ..	355
8.9. Алгоритми Excel корисні в статистичних обчисленнях	358
8.9.1. Введення даних у таблиці Excel	358
8.9.2. Редагування графіків за допомогою майстра діаграм	363
8.9.3. Виведення цифрових даних із заданою точністю	379
8.10. Висновки	381
Додатки	383
Д.1. Завдання для контролю навичок розв'язування задач з теорії імовірностей	383
Д.2. Завдання для індивідуальної семестрової роботи	389
Д.3. Тест для перевірки знань та навичок	391
Д.4. Таблиці	406
Значення функції e^{-x}	406
Значення локальної функції Лапласа	408
Значення інтегральної функції Лапласа	411
Розподіл Пуассона	411
Значення критерію Пірсона (χ^2)	414
Критичні точки розподілу Стьюдента	415
Критичні точки розподілу Фішера-Снедекора	417
2220 випадкових чисел	419
Список літератури	421

Передмова до першого видання

Теорія імовірностей та математична статистика — самостійні математичні науки, які є теоретичною основою викладання багатьох економічних, соціологічних та спеціальних дисциплін, вони використовуються в багатьох галузях науки та техніки, а також розумними бізнесменами та керівниками різного рівня.

Тому ці теорії становлять обов'язкову частину математичного циклу дисциплін усіх економічних вищих навчальних закладів.

Більшість наявних підручників з теорії імовірностей та математичної статистики призначено для студентів вищих технічних навчальних закладів або математичних та фізичних спеціальностей університетів. Вони розраховані на досить велику кількість учбових годин та високий рівень знань та навичок студентів не тільки елементарної математики, але й вищої математики, включаючи інтегрування функцій багатьох змінних.

Мета цього навчального посібника — ознайомити студентів економічних вищих навчальних закладів (а також усіх зацікавлених фахівців) з основними поняттями, методами, теоремами та формулами теорії імовірностей та математичної статистики, допомогти їм набути первинні навички застосування теоретичного матеріалу в багатьох випадках.

Посібник розроблено відповідно типової учбової програми дисципліни «Теорія імовірностей та математична статистика», затвердженої Міністерством освіти України для економічних вищих навчальних закладів. Він містить багато розв'язаних прикладів, завдань для самостійної роботи студентів та поточного контролю якості одержаних знань та навичок, необхідні таблиці, зразки планової контрольної роботи, комплексного семестрового завдання та тести перевірки остаточних знань студентів.

Розподіл матеріалу на частини, розділи та підрозділи дозволяє виділяти головне і зосереджує на ньому увагу читача.

Нумерацію формул, теорем, прикладів виконано окремо для кожного підрозділу для зручності читача.

Автори звертають увагу читача на те, що вивчати матеріал цього посібника треба послідовно з обов'язковим розв'язуванням

задач: вправи доцільно виконувати після ознайомлення з теоретичним матеріалом та прикладами, наведеними у навчальному посібнику.

Автори впевнені, що посібник буде корисним не тільки студентам економічних вищих навчальних закладів стаціонарної та заочної форм навчання, але й фінансистам, актуаріям, керівникам різного рівня, бізнесменам, співробітникам податкової інспекції та страхових компаній, соціологам та політологам.

Передмова до другого видання

В друге видання додана третя частина «Комп'ютерний аналіз» та розширені розділи 5 та 6. Мета включення Частини III у підручник полягає у побудові основ сучасного мислення фахівця у галузі бізнесу та управління.

У цій частині засобами Excel повністю вирішені наступні важливі задачі теорії імовірності та математичної статистики:

- побудова простої випадкової вибірки заданого об'єму по заданій скінченній вибірці;
- впорядкування і організація даних у вигляді розподілів частот, накопичених частот і т.д.;
- подання отриманих розподілів у графічному вигляді — гістограм, полігонів частот і т.д.;
- обчислення числових характеристик вибірки — середнього значення, дисперсії і середнє квадратичного відхилення;
- обчислення, пов'язані з комбінаторикою;
- обчислення, пов'язані з біноміальним розподілом, розподілом Пуассона, нормальним розподілом;
- побудова довірчого інтервалу для оцінки математичного сподівання нормального розподілу.

Окремо розглянуто алгоритми Excel, які корисні в статистичних обчисленнях: введення даних, редагування графіків за допомогою майстра діаграм, оформлення результатів та підготовка до друку.

Засвоєння цього матеріалу надзвичайно важливо для розуміння всього курсу теорії імовірності та математичної статистики. У той же час, саме розглянуті задачі є найбільш трудомісткими через велику кількість обчислень.

Вибір електронної таблиці Excel (пакет Microsoft Office 2003, Windows XP, русифікована версія) зумовлено тим, що вона широко поширена і, отже, доступна. Excel має дружній інтерфейс. Це означає, що потужні засоби Excel використовувати легко та просто. Excel доступний для початківця і цілком задовольнить професіонала.

Як показує досвід Національної академії управління, засвоєння студентом практичних навиків використання комп'ютерів сприяє

більш глибокому засвоєнню теорії і надає студентові найбільш сучасні і потужні засоби вирішення будь-яких складних задач теорії імовірностей та математичної статистики у бізнесі або в соціологічних дослідженнях.

Викладання матеріалу у цій частині розраховано на те, щоб кожний бажачий міг засвоїти ці алгоритми самостійно.

Всі алгоритми викладені дуже детально. Наприкінці параграфів, що розглядають первинну обробку даних, наведено вправи для самостійного виконання. Необхідною умовою для засвоєння алгоритмів є доступ до комп'ютера, де встановлено Excel, та послідовне і самостійне виконання всіх алгоритмів, що наведені у тексті.

Автори виражають щире подяку викладачам кафедри комп'ютерної техніки та програмування Національної академії управління, зокрема завідувачу кафедрою доценту, к.т.н. В.В.Кривцову, доценту, к.т.н. В.П.Корневу, викладачу В.Й.Ніколайчуку, за обговорення роботи і цінні зауваження, які були враховані при підготовці рукопису.

В текст другого видання були внесені деякі уточнення і виправлення. Автори вдячні всім читачам, що висловили свої побажання та зауваження.

Частина I

**ТЕОРІЯ
ІМОВІРНОСТЕЙ**

*Науку цю у давнині
з азартних ігор почали,
віками потім розвивали,
розумні люди на Землі,
щоб Ви Її застосували.*

Розділ 1

ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІМОВІРНОСТЕЙ ТА КОМБІНАТОРИКИ

*Фундамент кожної науки —
її важливі поняття,
основа красеня-будівлі,
пригодної на все життя.*

1.1. Предмет теорії імовірностей

Розглянемо деякий дослід, у результаті якого може з'явитись або не з'явитись подія A . Прикладами такого дослідження можуть бути:

- а) дослід — виготовлення певного виробу, подія A — стандартність цього виробу;
- б) дослід — кидання монети, подія A — випав герб;
- в) дослід — стрільба п'ятьма пострілами у мішень, подія A — вибито 30 очок;
- г) дослід — введення програми у комп'ютер, подія A — безпомилковий ввід.

Загальним для усіх дослідів є те, що кожен із них може реалізуватись у певних умовах скільки завгодно разів. Такі дослідження називають *випробуваннями*.

Події бувають *достовірні*, *випадкові* та *неможливі*.

Достовірною називають таку подію, яка при розглянутих умовах обов'язково трапиться.

Неможливою називають таку подію, яка при розглянутих умовах не може трапитись.

Випадковою називають таку подію, яка при умовах, що розглядаються, може трапитися, а може й не трапитися.

Наприклад, якщо в урні є лише білі кулі, то добування білої кулі з урни — достовірне подія, а добування з цієї урни кулі іншого кольору — неможлива подія.

Якщо кинути монету на площину, то поява герба буде випадковою подією, тому що замість герба може з'явитися надпис.

Випадкові події позначають великими літерами, наприклад

$$A, B, C, D, X, Y, A_1, A_2, \dots, A_n.$$

Кожна випадкова подія є наслідком багатьох випадкових або невідомих нам причин, які впливають на подію. Тому неможливо завбачити наслідок одночинного випробування.

Але якщо розглядати випадкову подію багато разів при однакових умовах, то можна виявити певну закономірність її появи або не появи. Таку закономірність називають **імовірною закономірністю масових однорідних випадкових подій**.

У теорії імовірностей під **масовими однорідними випадковими** подіями розуміють такі події, які здійснюються багатократно при однакових умовах або багато однакових подій.

Наприклад, кинути одну монету 1000 разів або 1000 однакових монет кинути один раз в теорії імовірностей вважають однаковими подіями.

Предметом теорії імовірностей є вивчення імовірностних закономірностей масових однорідних випадкових подій.

Основні поняття, методи, теореми та формули теорії імовірностей ефективно використовуються в науці, техніці, економіці, у теоріях надійності та масового обслуговування, у плануванні та організації виробництва, у страховій та податковій справах, у соціології та політології, у демографії та охороні здоров'я.

1.2. Коротка історична довідка

Перші роботи, в яких виникли основні поняття теорії імовірностей, з'явились у XV–XVI ст. як спроба побудови теорії азартних ігор і належать таким видатним ученим, як Б. Спіноза, Дж. Кардано, Галілео Галілей.

Слідуючий етап (кінець XVII – початок XVIII ст.) розвитку теорії імовірностей пов'язаний з роботами Б.Паскаля, П.Ферма, Х.Гюйгенса, К.Гаусса, Я.Бернуллі та Н.Бернуллі, С.Пуассона, А.Муавра, П.Лапласа, Т.Байеса.

Я.Бернуллі зробив перші теоретичні обґрунтування накопичених раніше фактів.

В XIX ст. теорію імовірностей почали, успішно застосовувати у страховій справі, артилерії, статистиці.

Лише наприкінці XIX ст. П.Л.Чебишов та його учні А.А.Марков та А.М.Ляпунов перетворили теорію імовірностей у математичну науку.

Подальшим розвитком теорії імовірностей та випадкових процесів зобов'язані таким математикам, як С.Н.Берштейн, А.М.Колмогоров, Б.В.Гніденко, А.В.Скороход, В.С.Королюк, Ю.Нейман, І.І.Гіхман, І.М.Коваленко.

1.3. Алгебра випадкових подій

Спочатку ознайомимось із різновидами випадкових подій.

Означення 1. *Події називають **несумісними**, якщо поява однієї з них виключає появу інших подій в одному і тому ж випробуванні.*

Приклад 1. Серед однорідних деталей у ящику є стандартні та нестандартні. Навмання беруть із ящика одну деталь.

Події

A — взята стандартна деталь,

B — взята нестандартна деталь

несумісні тому, що взята лише одна деталь, яка не може бути одночасно стандартною та нестандартною.

Означення 2. *Події називають **сумісними**, якщо поява однієї з них не виключає можливості появи інших (не обов'язково одночасно).*

Приклад 2. Два стрільця стріляють у мішень.

Події

A_1 — перший стрілок влучив у мішень,

A_2 — другий стрілок влучив у мішень

будуть сумісними випадковими подіями.

Означення 3. *Випадкові події A_1, A_2, \dots, A_n утворюють **повну групу подій**, якщо внаслідок випробування хоча б одна з них з'явиться обов'язково.*

Приклад 3. Кидають шестигранний кубик. Позначимо події так A_1 — випала грань 1; A_2 — випала грань 2; A_3 — випала грань 3; A_4 — випала грань 4; A_5 — випала грань 5; A_6 — випала грань 6. Події A_1, A_2, \dots, A_6 утворюють повну групу.

У прикладі 2 події A_1 та A_2 не утворюють повної групи. Але якщо позначити A_0 подію, що ніхто із стрільців не влучив у мішень, тоді події A_0, A_1 та A_2 утворюють повну групу.

Означення 4. Події називають **рівноможливими**, якщо немає причин стверджувати, що будь-яка з них можливіша за інші.

Приклад 4. Події — поява 1, 2, 3, 4, 5 або 6 очок при киданні шестигранного кубика — рівноможливі при умові, що центр його ваги не зміщений.

Означення 5. Дві несумісні події, які утворюють повну групу, називають **протилежними**.

Подія, протилежна події A , позначається \bar{A} .

Приклад 5. Якщо позначити через A подію, що при стрільбі по мішені вибито 8 очок, то подія \bar{A} — при стрільбі по мішені вибито будь-яке інше число очок.

Тепер розглянемо важливе поняття **простору елементарних наслідків**.

Нехай виконується деякий експеримент, який має елементи випадковості. Кожне випробування може мати різні наслідки.

Так, при киданні монети можуть бути два можливих наслідки: герб або надпис.

При киданні грального кубика можуть бути шість можливих наслідків.

У випробуванні «постріл у мішень» можна розглядати такі наслідки, як влучення у мішень, або кількість вибитих очок, або координати точки влучення.

Отже, що приймати за наслідок випробування, залежить від умови задачі.

Означення 6. **Елементарними наслідками** називають такі події, які неможливо розділити на більш прості.

Множину усіх можливих елементарних наслідків називають **простором елементарних наслідків**.

Простір елементарних наслідків може містити *скінчену, злічену* або *незлічену* множину елементів.

У ролі елементарних наслідків можна розглядати точки n -вимірного простору, відрізок деякої лінії, точки поверхні S або об'єму V трьохвимірного простору, функцію однієї або багатьох змінних.

У більшості випадків що розглядаються, припускають, що елементарні наслідки рівноможливі.

Приклад 6.

а) При двократному киданні монети простір елементарних наслідків містить 4 точки

$$\{(G, G), (G, H), (H, G), (H, H)\},$$

де G — означає появу герба, H — появу надпису.

б) Нехай по мішені стріляють одиночними пострілами до першого влучення. Можливі такі елементарні події

ω_1 {влучення при першому пострілі},

ω_2 {влучення при другому пострілі},

ω_3 {влучення при третьому пострілі}

і т.д.

У цьому випадку простір елементарних наслідків може мати нескінченну кількість точок, які можна шляхом нумерації перелічити. Тому простір елементарних наслідків буде зліченим.

в) При виробництві кінескопів виникають неоднакові умови технологічного процесу, тому час роботи кінескопа відрізняється від його номінального значення, тобто буде випадковою подією.

Простір елементарних наслідків у цьому випадку буде нескінченною незліченою множиною, елементи якої неможливо пронумерувати.

Тепер ознайомимось з алгеброю випадкових подій.

Нехай A та B — випадкові події.

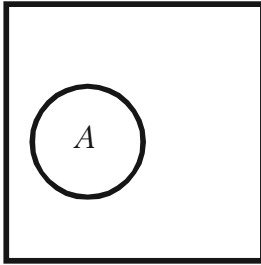
Об'єднанням (сумою) випадкових подій $A \cup B$ (або $A + B$) називають таку випадкову подію, яка полягає у появі подій

A або B

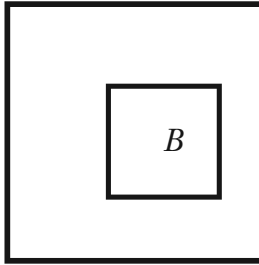
або

A та B .

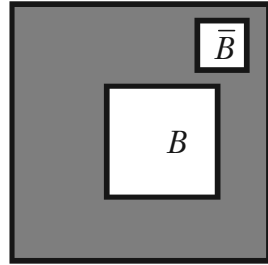
Якщо A та B — несумісні, то $A \cup B$ означає появу події A або події B .



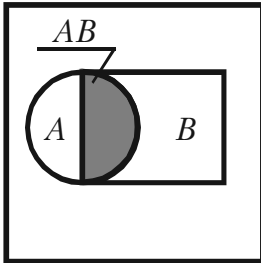
а)



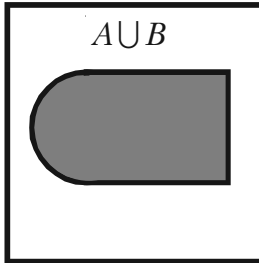
б)



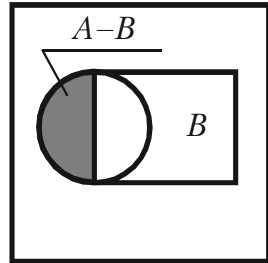
в)



г)



д)



е)

Мал. 1а та 1б. Подія A та подія B .

Мал. 1в. Подія B та протилежна їй \bar{B} .

Мал. 1г. Заштрихована площа — добуток подій AB .

Мал. 1д. Заштрихована площа — сума подій $A \cup B$.

Мал. 1е. Заштрихована площа — різниця подій $A - B$.

Аналогічно визначають об'єднання (суму) більшої кількості випадкових подій.

Означення 7. Об'єднанням (сумою) випадкових подій $A_1 \cup A_2 \cup \dots \cup A_n$ називають таку випадкову подію, яка полягає в появі хоча б однієї з цих подій.

Якщо події попарно несумісні, то їх сума полягає в тому, що повинна з'явитися подія A_1 або A_2 ... або A_n . Нескінченну суму випадкових подій позначають

$$\bigcup_{k=1}^{\infty} A_k.$$

Приклад 7. Стрілок робить один постріл у мішень, поділену на три області. Позначимо

подія A_1 — влучення в першу область;

подія A_2 — влучення у другу область;

подія A_3 — влучення в третю область;

подія A_4 — немає влучення у мішень;

подія B — влучення в першу або другу області;

подія D — влучення хоча б в одну область мішені.

Тоді маємо Тоді маємо $B = A_1 \cup A_2$; $D = A_1 \cup A_2 \cup A_3$.

Відмітимо, що події A_1 , A_2 , A_3 та A_4 — несумісні.

Означення 8. *Різницею $B - A$ (або $B \setminus A$) двох випадкових подій B , A називають усі наслідки, які полягають у тому, що подія A не з'являється.*

Добутком (перетином) $A \cap B$ (або $A \cdot B$) випадкових подій A , B називають таку випадкову подію, яка полягає у появі подій A та B одночасно.

Якщо A та B — несумісні, то добуток $A \cap B$ є множина, яка не має жодного елемента. Така множина називається *порожньою* (порожньою) і позначається \emptyset .

Таким чином, у разі несумісності подій A , B маємо

$$A \cap B = A \cdot B = \emptyset.$$

Означення 9. *Добутком (перетином) скінченної кількості випадкових подій A_1, A_2, \dots, A_n , називають таку випадкову подію, яка полягає у сумісній появі усіх цих подій одночасно.*

Подія $\bigcap_{k=1}^n A_k$ означає, що розглядаються усі події A_k ($k = 1, 2, \dots, n$) одночасно.

Вказані співвідношення між подіями є звичайними співвідношеннями між множинами, які можна представити графічно (див. мал. 1).

Приклад 8. Стрілець стріляє двічі по мішені. Описати простір елементарних наслідків. Записати подію, яка полягає в тому, що: а) стрілець влучив у мішень принаймні один раз (подія C); б) стрілець влучив рівно один раз (подія D); в) стрілець не влучив у мішень (подія F).

Розв'язання.

Позначимо

подія A — влучення при першому пострілі,

подія B — влучення при другому пострілі.

Простір елементарних наслідків складається з чотирьох подій

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

а) Якщо стрілець влучив у мішень принаймні один раз, то це означає, що він влучив або при першому пострілі $A\bar{B}$, або при другому пострілі $\bar{A}B$, або при обох AB .

Тобто,

$$C = A\bar{B} \cup \bar{A}B \cup AB.$$

б) Рівно одне влучення може бути тільки тоді, коли стрілець при першому пострілі влучив, а при другому — ні, або при першому пострілі не влучив, а при другому — влучив.

Тому,

$$D = A\bar{B} \cup \bar{A}B.$$

в) Якщо стрілець не влучив у мішень, то це означає, що він не влучив при обох пострілах,

Тобто,

$$F = \bar{A}\bar{B}.$$

1.4. Означення та властивості імовірності та частоти

Для порівняння випадкових подій за ступенем їх можливості треба кожному подію пов'язати з певним числом, яке повинно бути тим більше, чим більш можлива подія. Таке число p називають **імовірністю події**. Існує декілька означень імовірності. Ознайомимось із ними.

Означення 1. *Імовірність події є численна міра ступеня об'єктивної можливості цієї події.*

Розділ 1. Основні поняття теорії імовірностей та комбінаторики

Це означення імовірності визначає філософську суть імовірності, але не вказує закону знаходження імовірності будь-якої події.

Означення 2 (класичне). *Імовірність події A дорівнює відношенню числа елементарних наслідків, які сприяють появі події A , до загального числа усіх єдиноможливих та рівноможливих елементарних наслідків.*

Імовірність події A позначають $P(A)$. За Означенням 2

$$P(A) = \frac{m}{n}, \quad (1)$$

де m — число елементарних наслідків, що сприяють події A ,
 n — число усіх єдиноможливих та рівноможливих наслідків.

Приклад 1. В урни 6 однакових за розміром куль: 2 червоні, 3 сині, 1 біла. Знайти імовірність появи червоної кулі, якщо беруть одну кулю з урни навмання.

Розв'язання. Нехай подія A — навмання взята червона куля. З урни можна взяти будь-яку кулю із шести, тому усіх можливих наслідків 6 ($n = 6$). Для появи червоної кулі сприяти будуть лише 2 кулі, тому $m = 2$. За формулою (1) одержуємо

$$P(A) = \frac{2}{6} = \frac{1}{3}.$$

Зауваження 1. При розв'язанні багатьох задач знаходження чисел m та n має певні труднощі, запобігти яким допомагають принципи та формули комбінаторики, з якими ознайомимось нижче.

Зауваження 2. Класичне означення імовірності має місце лише тоді, коли m та n скінченні, усі елементарні наслідки рівноможливі (саме таке становище у більшості азартних ігор, що здійснюються без шахрайства).

Якщо множина елементарних наслідків нескінченна або елементарні наслідки не рівноможливі, то формулою (1) користуватись не можна.

Якщо множина усіх елементарних наслідків нескінченна і, як наслідок, займає деяку область G , а події A сприяє лише частина $g \in G$, то обчислення імовірності події A виконують згідно геометричного означення імовірності.

Означення 3 (геометричне). Імовірність випадкової події A дорівнює відношенню міри g до міри G

$$P(A) = \frac{m(g)}{m(G)}. \quad (2)$$

Зауваження 3. Якщо область G – проміжок, поверхня, або просторове тіло, g – частина G , тоді мірою G та g буде довжина, площа або об'єм відповідної області. Якщо G та g проміжки часу, тоді їх мірою буде час.

У загальному випадку міру області визначають аксіомами.

Приклад 2. Два туристичних пароплава повинні причалити до одного причалу. Час прибуття обох пароплавів рівноможливий на протязі доби.

Визначити імовірність того, що одному з пароплавів доведеться чекати звільнення причалу, якщо час стоянки першого пароплава дорівнює одній годині, а другого – двом годинам.

Розв'язання. Нехай X та Y – час прибуття пароплавів.

Можливі значення X та Y :

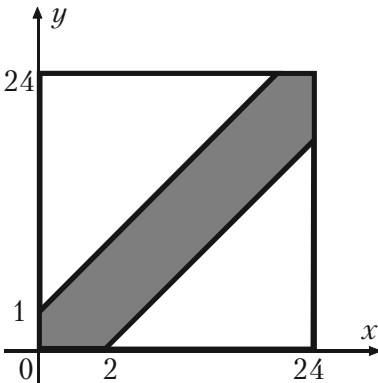
$$0 \leq X \leq 24; 0 \leq Y \leq 24.$$

Сприятливі значення:

$$Y - X \leq 1; X - Y \leq 2.$$

Побудуємо цю область (див. мал. 2).

Відношення площі заштрихованої фігури $m(g)$ до площі квадрата, сторона якого дорівнює 24, згідно формули (2) дорівнює шуканій імовірності



Мал. 2

$$p = \frac{24 \cdot 24 - 0,5 \cdot 23 \cdot 23 - 0,5 \cdot 22 \cdot 22}{24 \cdot 24} = 0,121.$$

Означення 4. Відносною частотою або частотою події A називають відношення числа випробувань, у яких подія A з'явилась, до числа фактично виконаних випробувань.

Відносну частоту події A позначають $W(A)$ або $p_n(A)$. Отже,

$$p_n(A) = W(A) = \frac{m}{n},$$

де m — кількість випробувань, у яких з'явилась подія A ,
 n — кількість усіх випробувань.

Приклад 3. Відділ технічного контролю серед 100 виробів виявив 8 нестандартних. Чому дорівнює відносна частота появи нестандартних виробів?

Розв'язання. Позначимо через A таку подію, як поява нестандартного виробу. Тоді за означенням частоти події A одержимо

$$W(A) = \frac{8}{100} = 0,08.$$

Зауваження 4. Підкреслимо, що імовірність $P(A)$ події A обчислюється до випробування, а частота $W(A)$ обчислюється після випробування.

Частота має властивість стійкості: при великій кількості випробувань частота змінюється дуже мало, коливаючись біля деякого постійного числа — імовірності появи цієї події, тобто

$$p(A) = \lim_{n \rightarrow \infty} p_n(A).$$

Означення 5. Статистична імовірність — це відносна частота (частота) або число, близьке до неї.

Тепер розглянемо **основні властивості імовірності**, використовуючи формулу (1) класичного означення імовірності події A .

1. Якщо подія A достовірна, то її імовірність дорівнює одиниці, тобто $p(A) = 1$.

2. Якщо подія A неможлива, то її імовірність дорівнює нулеві, тобто $p(A) = 0$.

3. Якщо подія A випадкова, то її імовірність задовольняє співвідношення

$$0 < P(A) < 1. \quad (3)$$

Дійсно, при розгляданих умовах достовірна подія обов'язково з'явиться, як наслідок, усі можливі елементарні наслідки сприяють події A , тобто $n = m$ і з формули (1) одержимо

$$p(A) = \frac{m}{n} = \frac{n}{n} = 1.$$

Якщо при умовах, що розглядаються, подія A неможлива, тоді серед усіх можливих наслідків немає тих, що сприяють події A , тобто $m = 0$, і за формулою (1) одержимо $p(A) = 0$.

Якщо подія A випадкова, то серед усіх n можливих наслідків є m наслідків, що сприяють події A , $0 < m < n$. Тому, згідно формули (1) одержимо співвідношення (3).

Зауваження 5. *Остання властивість імовірності випадкових подій використовується для здійснення самоконтроля при розв'язанні багатьох задач теорії імовірностей.*

1.5. Основні поняття та принципи комбінаторики

Часто для знаходження чисел m та n , що входять у класичне означення імовірності події, потрібно знати кількість різноманітних сполук, які можна одержати з n елементарних наслідків.

Класифікація та властивості таких сполук, а також формули для обчислення кількості різних сполук розроблені математиками і містяться у розділі «Комбінаторика» курсу алгебри.

Ознайомимось із основними поняттями та формулами комбінаторики.

Означення 1. *Різні групи, складені з будь-яких елементів, що відрізняються елементами або порядком цих елементів, називають **сполуками** або **комбінаціями** цих елементів.*

Приклад 1. Із цифр 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 можна скласти багато різних сполук по 2, 3, 4, ..., цифр. Деякі з них будуть відрізнятися кількістю цифр, а деякі відрізнятимуться лише порядком цифр. Наприклад 123, 8056, 96, 312, 123, 231, 321.

Усі можливі сполуки доцільно класифікувати. Сполуки бувають трьох видів:

- переставлення;
- розміщення;
- сполучення.

Означення 2. Сполуки з n елементів, що відрізняються лише порядком елементів, називають **переставленням** цих елементів.

Кількість переставлень з n елементів позначають P_n і знаходять за формулою

$$P_n = \prod_{k=1}^n k = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n = n! \quad (1)$$

Позначення $n!$ вимовляють « n факторіал».

За означенням $0! = 1$.

Приклад 2. Скільки п'ятизначних чисел можна записати, використовуючи п'ять різних цифр (крім нуля)?

Розв'язання. Сполуки, що утворюють з п'яти різних цифр п'ятизначні числа, можуть відрізнитися лише порядком цифр, тому такі сполуки будуть переставленням з 5 елементів. Згідно формули (1) їх кількість буде

$$P_5 = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120.$$

Означення 3. **Розміщенням з n елементів по m** називають такі комбінації, які складаються з m елементів, взятих з даних n елементів ($m < n$) і відрізняються як порядком, так і елементами.

Кількість розміщень з n елементів по m позначають A_n^m і знаходять за формулою

$$A_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}. \quad (2)$$

Приклад 3. Студенти другого курсу згідно учбового плану вивчають 10 дисциплін. На один день можна планувати заняття з 4 дисциплін. Скількома способами можна скласти розклад занять на один день?

Розв'язання. Усі можливі розклади занять на один день — це сполуки з 10 по 4, які можуть відрізнитися дисциплінами або їх порядком, тобто ці сполуки — розміщення. Кількість таких розміщень згідно формули (2) буде

$$A_{10}^4 = 10 \cdot 9 \cdot 8 \cdot 7 = 5040.$$

Означення 4. Сполученням з n елементів по m називають комбінації, що складаються з m елементів, взятих з даних n елементів і які відрізняються хоч би одним елементом.

Кількість сполучень з n елементів по m позначають C_n^m знаходять за формулою

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Зауваження 1. Переставлення можна розглядати як частковий випадок розміщень

$$P_n = A_n^n = n!.$$

Між кількістю переставлень, розміщень та сполучень існує простий зв'язок

$$C_n^m = \frac{A_n^m}{P_m}.$$

Часто доцільно використовувати такі властивості сполучень:

- а) $C_n^m = C_n^{n-m}$;
- б) $C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n$;
- г) $C_n^n = 1$;
- д) $C_n^1 = n$.

Приклад 4. У ящику 10 виробів, з яких 2 нестандартні. Навмання беруть 6 виробів. Яка імовірність того, що усі взяті вироби будуть стандартними?

Розв'язання. Позначимо подію A — взято 6 стандартних виробів. Згідно умови задачі, немає значення, в якому порядку беруть 6 виробів, тобто це будуть сполучення.

Тому кількість усіх можливих елементарних наслідків буде

$$n = C_{10}^6.$$

Події A сприяють лише сполуки по 6 виробів з 8 стандартних у будь-якому порядку, тобто $m = C_8^6$.

Отже згідно класичному означенню імовірності події A маємо

$$P(A) = \frac{C_8^6}{C_{10}^6} = \frac{8!}{6! \cdot 2!} : \frac{10!}{6! \cdot 4!} = \frac{8! \cdot 6! \cdot 4!}{6! \cdot 2! \cdot 10!} = \frac{2}{15}.$$

Тепер ознайомимось з основними принципами комбінаторики.

Принцип суми. *Якщо множина A містить n елементів, а множина B містить m елементів і $A \cap B = \emptyset$, тоді множина $A \cup B$ містить $n + m$ елементів.*

Доведення. Здійснюється простим рахуванням елементів множини $A \cup B$.

Спочатку рахуємо усі елементи множини A . Вони одержать номери від 1 до n . Серед них немає елементів множини B , тому що $A \cap B = \emptyset$.

Тепер будемо рахувати елементи множини B . Вони одержать номери від $n + 1$ до $n + m$, оскільки множина B за умовою має m елементів.

Таким рахуванням усі елементи множини $A \cup B$ будуть вичерпані. Вони одержать номери від 1 до $n + m$, тому $A \cup B$ містить $n + m$ елементів.

Зауваження 2. *Принцип суми має місце для суми k множин, тобто для*

$$A_1 \cup A_2 \cup \dots \cup A_k = \bigcup_{i=1}^k A_i.$$

Принцип добутку. *Якщо множина A містить n елементів, а множина B містить m елементів, то множина C усіх можливих пар (a_i, b_k) ($i = 1, 2, \dots, n$; $k = 1, 2, \dots, m$) містить $n \cdot m$ елементів.*

Доведення. Множину C розіб'ємо на підмножини

$$C_1 = \{(a_1, b_k), k = 1, 2, \dots, m\},$$

$$C_2 = \{(a_2, b_k), k = 1, 2, \dots, m\},$$

.....

$$C_i = \{(a_i, b_k), k = 1, 2, \dots, m\},$$

.....

$$C_n = \{(a_n, b_k), k = 1, 2, \dots, m\}.$$

Оскільки C_1 складається лише з пар, що містять a_1 , а множина C_2 складається лише з пар, що містять a_2 , то $C_1 \cap C_2 = \emptyset$.

Аналогічно одержуємо, що $C_i \cap C_l = \emptyset$, коли $i \neq l$.

Тепер доведемо, що

$$C = \bigcup_{i=1}^n C_i.$$

Дійсно, нехай (a_i, b_k) будь-яка пара. Вона входить в C згідно визначенню множини C . Вона також входить і в множину C_i тому, що $(a_i, b_k) \in C_i$. Кожна підмножина C_i множини C містить m елементів, тому згідно визначенню принцип а суми число елементів у їх об'єднанні дорівнює mn .

Приклад 5. У кошику 4 яблука першого сорту та 5 яблук другого сорту. Навмання беруть 2 яблука. Знайти імовірність того, що будуть взяті яблука різних сортів.

Розв'язання. Нехай подія A — навмання взяті 2 яблука різних сортів.

Всього яблук 9, з них сполучень по 2 буде C_9^2 , тобто кількість усіх можливих наслідків $n = C_9^2$.

Події A будуть сприяти сполуки, утворені з пар, елементами яких будуть яблука різних сортів. Згідно принципу добутку кількість таких пар буде дорівнювати $m = C_4^1 \cdot C_5^1$.

Використовуючи класичне означення імовірності, одержимо шукану імовірність події A

$$P(A) = \frac{m}{n} = \frac{C_4^1 \cdot C_5^1}{C_9^2} = \frac{4!}{3!} \cdot \frac{5!}{4!} \cdot \frac{9!}{2! \cdot 7!} = \frac{4! \cdot 5! \cdot 2! \cdot 7!}{3! \cdot 4! \cdot 9!} = \frac{5}{54}.$$

Питання для самоперевірки та вправи

1. Які події називають достовірними, неможливими, випадковими?
2. Що є предметом теорії імовірностей?
3. Як визначають та позначають суму, добуток випадкових подій, протилежну подію, повну групу подій?
4. Які події називають сумісними, несумісними, рівноможливими?
5. Як визначають та в яких випадках використовують класичне та геометричне означення імовірності?
6. Як визначають та позначають частоту випадкової події A ?
7. Які основні властивості імовірності та частоти?
8. Що є предметом комбінаторики?
9. Які комбінації називають переставленням, розміщенням, сполученням? Як позначають та обчислюють кількість цих сполук?
10. Як формулюють основні принципи комбінаторики?

Розв'язати задачі

1. Дві особи стріляють у мішень по одному разу. Подія A — у мішень влучив перший стрілець. Подія B — у мішень влучив другий стрілець. Виразити через A та B такі події

C — два влучення у мішень;

D — хоч би одне влучення у мішень;

E — лише одне влучення у мішень.

Відповідь:

$$C = A \cdot B; D = A \cdot B \cup A \cdot \bar{B} \cup \bar{A} \cdot B; E = A \cdot \bar{B} \cup \bar{A} \cdot B.$$

2. Маємо такі події

A — навмання взята деталь першого сорту;

B — навмання взята деталь другого сорту;

C — навмання взята деталь третього сорту.

Що означають події $A \cup B$; $\overline{A \cup C}$; $A \cap C$; $(A \cap B) \cup C$?

3. На складі зберігають 500 акумуляторів. Відомо, що після року зберігання 20 із них будуть непридатними. Треба знайти імовірність того, що навмання взятий після року зберігання акумулятор буде справним, якщо відомо, що після 6 місяців зберігання було вилучено 5 несправних акумуляторів.

Відповідь: $\frac{32}{33}$.

4. Лотерея випущена на загальну суму n гривень. Вартість одного квитка r гривень. Виграють m квитків. Знайти імовірність виграшу на один квиток.

Відповідь: $\frac{rm}{n}$.

5. Підприємство одержало партію виробів, серед яких n доброякісних та m бракованих. Для контролю навмання взято s виробів, перші з яких k виявились доброякісними. Знайти імовірність того, що наступний перевірений виріб буде доброякісним.

Відповідь: $\frac{n-k}{n+m-s}$.

6. Монету кинуто двічі. Знайти імовірність того, що хоча б один раз вона впаде гербом вгору.

Відповідь: $\frac{3}{4}$.

7. У книжковій шафі стоять підручники з математики, фізики та хімії. Учень навмання бере два підручника. Скласти повну групу подій, які при цьому можуть відбутися.

Відповідь: Нехай події

A_1 — учень взяв обидва підручники з математики;

A_2 — обидва підручники з фізики;

A_3 — обидва підручники з хімії;

A_4 — один підручник з математики, а другий з фізики;

A_5 — один підручник з математики, а другий з хімії;

A_6 — один підручник з фізики, а другий з хімії, тоді події A_1, A_2, A_3, A_4, A_5 та A_6 утворюють повну групу.

8. Монету кидають два рази. Вказати кількість випадкових подій, які при цьому утворюють повну групу.

Відповідь: 4.

9. Дві особи домовились зустрітись у певному місці між 18 та 19 годинами. Той, хто прийде першим, чекає другого протягом 20 хвилин, після чого йде.

Розділ 1. Основні поняття теорії імовірностей та комбінаторики

Чому дорівнює імовірність зустрічі осіб за умови, що кожен із них може прийти навмання у будь-який момент між 18 та 19 годин, незалежно один від одного?

Відповідь:

Якщо взяти за одиницю масштабу 1 хвилину, тоді усі можливі наслідки будуть точками квадрата, сторона якого дорівнює 60.

Сприятли будуть наслідки, що задовольняють умову $|x - y| \leq 20$.

Тому

$$p = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

10. На площині проведені паралельні прямі, відстані між якими поперемінно рівні 1,5 та 8 см. Визначити імовірність того, що навмання кинутий на цю площину круг радіусу 2,5 см не буде перетинати жодної прямої.

Відповідь: $\frac{3}{9,5}$.

11. Два студента домовились зустрітись у певному місці між 17 та 18 годинами. Той, хто прийде першим, повинен чекати другого на протязі 15 хвилин, після чого йде.

Знайти імовірність зустрічі, якщо час приходу на місце зустрічі кожного студента незалежний і рівноможливий на протязі вказаної години.

Відповідь: $\frac{7}{16} = 0,4375$.

12. На площині проведені паралельні прямі, відстань між якими 8 см. Визначити імовірність того, що навмання кинутий на цю площину круг радіуса 3 см не буде перетинати жодну пряму.

Відповідь: $\frac{1}{4} = 0,25$.

13. У філії банку працюють 15 співробітників, троє з яких не мають потрібної кваліфікації. Скільки можна скласти списків

а) по 8 співробітників;

б) по 6 кваліфікованих співробітників;

в) по 9 співробітників, два з яких не мають потрібної кваліфікації?

Відповідь: а) C_{15}^8 ; б) C_{12}^6 ; в) $C_{12}^7 \cdot C_3^2$.

14. Правління підприємства складається з 9 осіб. Скільки можна скласти варіантів обрання з їх числа трьох керівників — президента, директора та комерційного директора?

Відповідь: $A_9^3 = 504$.

15. У малому підприємстві працюють 4 жінки та 5 чоловіків. Випадковим способом дві особи запізнались. Знайти імовірність того, що одна з цих осіб — жінка, а друга — чоловік.

Відповідь: $\frac{C_4^1 \cdot C_5^4}{C_9^2} = \frac{5}{9}$.

16. До профкому вибрано 7 осіб. З них потрібно вибрати голову профкому та його заступника. Скількома способами це можна зробити?

Відповідь: 42.

17. Скільки п'ятизначних чисел можна скласти з цифр 1, 2, 3, 4, 5 так, щоб жодна з них не повторювалась?

Відповідь: 120.

18. Скількома способами з 10 осіб можна вибрати комісію у складі чотирьох осіб?

Відповідь: 210.

19. З цифр 1, 2, 3, 4, 5 складають різні тризначні та чотиризначні числа. Скільки таких чисел можна скласти, якщо повторення цифр у числах заборонені?

Відповідь: 180.

20. Скільки парних чотиризначних чисел, що складаються з цифр 2, 3, 5, 7 можна одержати, якщо повторення цифр у числах заборонені?

Відповідь: $3! = 6$.

21. Знайти A_7^4 .

Відповідь: 840.

Розділ 1. Основні поняття теорії імовірностей та комбінаторики

22. Скількома способами можна розмістити 12 гостей за столом, біля якого стоять 12 стільців?

Відповідь: $12! = 479001600$.

23. З 10 можливих кандидатів обирають старосту, профорга та культорга. Скільки може бути різних комбінацій вибору?

Відповідь: $A_{10}^3 = 720$.

24. Партію з N виробів перевіряє контролер шляхом відбору навмання n виробів і визначення їх якості. Якщо серед вибраних контролером виробів немає жодного бракованого, то уся партія виробів приймається, у протилежному випадку усі вироби посилають на додаткову перевірку.

Яка імовірність того, що партія виробів, яка містить M бракованих виробів, буде прийнята контролером?

Відповідь: $\frac{C_{N-M}^m}{C_N^m}$.

25. 25 екзаменаційних білетів містять по 2 питання, що не повторюються. Студент може відповісти лише на 45 питань. Яка імовірність того, що взятий студентом екзаменаційний білет містить відомі йому питання?

Відповідь: $\frac{C_{45}^2}{C_{50}^2} \approx 0,8$.

26. Мале підприємство одержало 20 радіоприймачів, з яких 5 бракованих. Навмання для перевірки взяли три приймача.

Яка імовірність того, що серед взятих приймачів будуть

- а) тільки стандартні приймачі;
- б) тільки браковані приймачі;
- в) один бракований та два стандартних?

Відповідь:

а) $\frac{C_{15}^3}{C_{20}^3} = \frac{91}{228}$; б) $\frac{C_5^3}{C_{20}^3} = \frac{1}{114}$; в) $\frac{C_{15}^2 \cdot C_5^1}{C_{20}^3} = \frac{105}{228}$.

27. В урні 15 червоних, 9 синіх та 6 зелених куль однакового розміру. Навмання беруть 6 куль. Яка імовірність того, що будуть узяті 1 зелена, 2 синіх та 3 червоних кулі?

Відповідь: $\frac{C_6^1 \cdot C_9^2 \cdot C_{15}^3}{C_{30}^6} = \frac{24}{145} \approx 0,17.$

28. Студент забув останні три цифри потрібного телефону, але він пам'ятає, що всі три цифри різні, тому набирає їх навмання. Знайти імовірність того, що набрані цифри вірні.

Відповідь: $\frac{1}{A_{10}^3} = \frac{1}{720}.$

29. Перед новим роком фірма одержала ящик стандартних виробів. Покупець навмання бере з ящика один виріб.

Яка імовірність того, що взятий виріб буде

а) стандартним,

б) бракованим?

Відповідь: а) 1; б) 0.

30. Серед 25 студентів групи, в якій є 10 дівчат, розігрують 5 квитків на концерт. Визначити імовірність того, що білети виграють дві дівчини.

Відповідь: $\frac{C_{15}^3 \cdot C_{10}^2}{C_{25}^5} = \frac{195}{253} \approx 0,771.$

31. Відділ контролю виявив 5 бракованих виробів з випадково відібраних 100 однакових виробів. Знайти відносну частоту появи бракованих виробів.

Відповідь: $W(A) = 0,05.$

32. Відносна частота стандартних радіоприймачів деякої фірми 0,9. Знайти кількість стандартних радіоприймачів серед перевірених 200.

Відповідь: 180.

33. Проведено 20 пострілів по мішені і зареєстровано 18 влучень. Знайти частоту влучення у мішень.

Відповідь: 0,9.

34. При перевірці 30 пар взуття виявлено 6 бракованих пар. Чому дорівнює частота

а) стандартного взуття;

б) бракованого взуття?

Відповідь: а) 0,8; б) 0,2.

35. Серед 250 деталей, виготовлених на автоматичному верстаті, виявилось 5 деталей, що не відповідають стандарту. Визначити частоту появи бракованих деталей.

Відповідь: 0,02.

36. Контролер, перевіряючи якість 500 виробів, встановив, що 10 із них другого сорту, а інші — першого. Визначити частоту виробів першого сорту, другого сорту, а також суму цих частот.

Відповідь: $\frac{49}{50}$; $\frac{1}{50}$; 1.

Розділ 2

ОСНОВНІ ТЕОРЕМИ ТЕОРІЇ ІМОВІРНОСТЕЙ

*На фундаменті міцному
будем класти поверхи,
перегородки та сходинки,
що їх з'єднають на віки.*

2.1. Додавання імовірностей несумісних подій

Теорема 1. *Імовірність об'єднання двох випадкових несумісних подій дорівнює сумі їх імовірностей*

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

Доведення. Нехай число усіх можливих елементарних наслідків появи подій A та B дорівнює n ; m_1 та m_2 — числа наслідків, що сприяють подіям A та B відповідно. Тоді події $A \cup B$ будуть сприяти $m_1 + m_2$ наслідків. Отже, за класичним означенням імовірності, маємо

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

тобто твердження теореми доведено.

Зовсім аналогічно можна довести наступне твердження.

Теорема 2. *Якщо випадкові події A_1, A_2, \dots, A_n попарно несумісні, то імовірність появи хоча б однієї з цих подій дорівнює сумі їх імовірностей*

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

Приклад 1. Імовірність влучення стрілкою у першу область мішені дорівнює 0,45, у другу область — 0,35, у третю — 0,15. Знайти імовірність того, що при одному пострілі стрілок влучить у першу або другу області мішені.

Розділ 2. Основні теореми теорії імовірностей

Розв'язання. Позначимо за подію A_1 — влучення у першу область мішені; за подію A_2 — влучення у другу область мішені.

При одному пострілі події A_1 та A_2 несумісні. Тому імовірність влучення в першу або другу області мішені буде

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) = 0,45 + 0,35 = 0,8.$$

Теорема 3. Сума імовірностей повної групи випадкових подій дорівнює одиниці

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

Доведення. Якщо випадкові події A_1, A_2, \dots, A_n утворюють повну групу, то вони попарно несумісні, а їх об'єднання буде достовірною подією. За Теоремою 2 маємо

$$P\left(\bigcup_{k=1}^n A_k\right) = P(A_1) + P(A_2) + \dots + P(A_n). \quad (2)$$

Імовірність достовірної події дорівнює одиниці, тому

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1. \quad (3)$$

Ліві частини рівностей (2) та (3) однакові, тому праві частини будуть рівними, тобто має місце рівність (1). Теорема доведена.

Наслідок. Дві протилежні події A та \bar{A} утворюють повну групу, тому має місце рівність

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

з якої одержуємо формулу

$$P(\bar{A}) = 1 - P(A) \quad (4)$$

знаходження імовірності протилежної події.

Приклад 2. Імовірність одержати повідомлення від певної особи на протязі доби дорівнює 0,25. Знайти імовірність того, що повідомлення на протязі доби від цієї особи не буде одержано.

Розв'язання. Позначимо за подію A — повідомлення від цієї особи на протязі доби поступить. За умовою задачі має місце співвідношення $P(A) = 0,25$. Протилежна подія \bar{A} означає, що на

протязі доби від цієї особи повідомлення не поступить. За формулою (4) одержимо

$$P(\bar{A}) = 1 - 0,25 = 0,75.$$

У страховій справі треба вираховувати, наприклад, таку задачу.

Приклад 3. За статистичними показниками держави можна зробити висновок, що 68% чоловіків, які досягли 60-ліття, досягають також і 70-ліття. Яка імовірність того, що 60-річний чоловік не досягне свого 70-річчя?

Розв'язання. Якщо подія A — 60-річний чоловік досягає свого 70-ліття, то протилежна подія \bar{A} — 60-річний чоловік не досягає свого 70-ліття. За умовою задачі $P(A) = 0,68$, тому за формулою (4) одержимо

$$P(\bar{A}) = 1 - 0,68 = 0,32.$$

Отже, використовуючи статистичні дані держави, можна обчислити імовірність того, що 32% 60-річних чоловіків помре на протязі 10 років.

2.2. Залежні та незалежні події, умовні імовірності

Означення 1. *Випадкові події A та B називають залежними, якщо імовірність появи однієї з них залежить від появи або не появи другої події.*

Якщо імовірність появи однієї події не залежить від появи або не появи другої, то такі події називають незалежними.

Означення 2. *Імовірність події B , обчислена при умові появи події A , називають умовною імовірністю події B і позначають $P(B|A)$ або $P_A(B)$.*

Приклад 1. В урні 10 куль: 3 білих і 7 чорних. Навмання беруть дві кулі. Нехай подія A — взята біла куля; подія B — взята чорна куля.

Розділ 2. Основні теореми теорії імовірностей

Якщо кулю, яку взяли першою, повертають до урни, то імовірність появи другої кулі не залежить від того, яка взята перша куля.

Якщо перша куля не повертається до урни, то імовірність другої події залежить від результату першого випробування.

Якщо першою взяли білу кулю, то в урні залишилося 2 білих кулі та 7 чорних, тому

$$P_A(B) = \frac{7}{9}.$$

Якщо першою взяли чорну кулю, то в урні залишилося 3 білих кулі та 6 чорних куль, тому

$$P_B(B) = \frac{6}{9} = \frac{1}{3}.$$

Отже імовірність події B залежить від появи або не появи події A .

Зауваження. Якщо події A та B незалежні, то умовна імовірність дорівнює безумовній імовірності, тобто

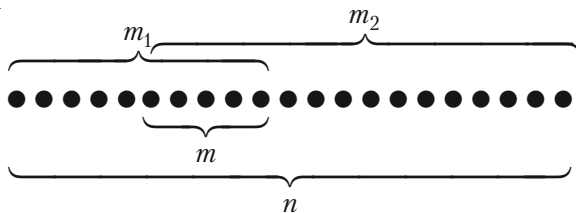
$$P_A(B) = P(B).$$

2.3. Множення імовірностей

Теорема 4. Імовірність сумісної появи двох випадкових подій A та B дорівнює добутку імовірностей однієї з цих подій та умовної імовірності другої події при умові, що перша подія з'явилась

$$P(A \cdot B) = P(A) \cdot P_A(B) = P(B) \cdot P_B(A). \quad (1)$$

Доведення. Усі елементарні наслідки зобразимо у вигляді точок (мал. 3).



Мал. 3.

Нехай появи події A сприяють m_1 наслідків, а появи події B — m_2 наслідків. Усіх можливих наслідків n , а події $A \cdot B$ будуть сприяти m наслідків.

Так як

$$P(A \cdot B) = \frac{m}{n}, \quad P(A) = \frac{m_1}{n}, \quad P_A(B) = \frac{m}{m_1},$$

то

$$P(A \cdot B) = \frac{m_1}{n} \cdot \frac{m}{m_1} = P(A) \cdot P_A(B),$$

що і треба було довести.

Співвідношення (1) називають формулою множення імовірностей залежних випадкових подій.

Наслідок. У випадку незалежних випадкових подій A та B формула (1) приймає вигляд

$$P(A \cdot B) = P(A) \cdot P(B) \quad (2)$$

і називається *формулою множення імовірностей незалежних подій.*

У випадку скінченної кількості незалежних випадкових подій формула (2) приймає вигляд

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

Приклад 1. У деякому людському суспільстві 70% палять, 40% хворіють на рак легенів та 25% палять та мають рак легенів. Знайти імовірність того, що навдачу взята особа з цього суспільства:

- а) не палить, але має рак легенів;
- б) палить, але не має раку легенів;
- в) ніколи не палить і не має раку легенів;
- г) палить і має рак легенів;
- д) або палить або має рак легенів.

Розв'язання. Позначимо події: A — особа палить; B — особа хворіє на рак легенів. Тоді за умовою задачі маємо

$$P(A) = 0,7, \quad P(B) = 0,4, \quad P(A \cdot B) = 0,25 \Rightarrow P(\bar{A}) = 0,3,$$

$$P(\bar{B}) = 0,6;$$

а) $P(\bar{A} \cdot B) = 0,3 \cdot 0,4 = 0,12$;

б) $P(A \cdot \bar{B}) = 0,7 \cdot 0,6 = 0,42$;

в) $P(\bar{A} \cdot \bar{B}) = 0,3 \cdot 0,6 = 0,18$;

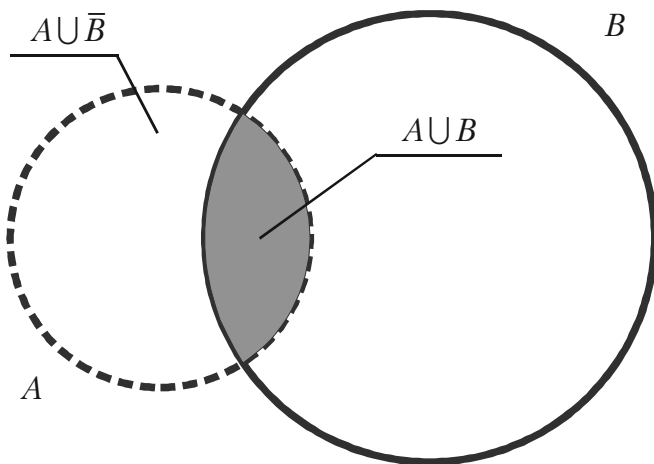
г) $P(A \cup B) = P(A) + P(B) - P(\bar{A} \cup \bar{B}) = 0,7 + 0,4 - 0,18 = 0,92$;

д) $P(A \cdot \bar{B} \cup \bar{A} \cdot B) = P(A \cdot \bar{B}) + P(\bar{A} \cdot B) = 0,42 + 0,12 = 0,54$.

Приклад 2. Навести ілюстративну діаграму властивості

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

Відповідь. Див. мал. 4.



Мал. 4.

2.4. Імовірність появи хоча б однієї випадкової події

Нехай є n сумісних випадкових подій A_1, A_2, \dots, A_n . Позначимо через A подію, яка полягає в тому, що з'явиться хоча б одна з цих подій. Тоді подія \bar{A} полягає в тому, що події $\bar{A} = \bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n$. Події A та \bar{A} утворюють повну групу подій, тому

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

Звідси одержуємо

$$P(A) = 1 - P(\bar{A}_1 \cdot \bar{A}_2 \dots \bar{A}_n). \quad (1)$$

За цією формулою треба обчислювати імовірність появи хоча б однієї випадкової події з n сумісних подій.

Приклад 1. Імовірність влучення у мішень першого стрілка дорівнює 0,7, другого стрілка — 0,8, а третього стрілка — 0,9. Знайти імовірність влучення у мішень хоча б одного стрілка.

Розв'язання. Позначимо події

- A_1 — у мішень влучив перший стрілок;
- A_2 — у мішень влучив другий стрілок;
- A_3 — у мішень влучив третій стрілок;
- A — у мішень влучив хоча б один стрілок.

За умовою задачі події A_1, A_2 та A_3 незалежні, тому події \bar{A}_1, \bar{A}_2 та \bar{A}_3 також незалежні.

Згідно формули (1) та формули множення імовірностей незалежних подій маємо

$$P(A) = 1 - P(\bar{A}_1 \cdot \bar{A}_2 \cdot \bar{A}_3) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot P(\bar{A}_3). \quad (2)$$

Так як

$$P(\bar{A}_1) = 1 - 0,7 = 0,3; \quad P(\bar{A}_2) = 1 - 0,8 = 0,2; \quad P(\bar{A}_3) = 1 - 0,9 = 0,1,$$

то з формули (2) одержимо

$$P(A) = 1 - 0,3 \cdot 0,2 \cdot 0,1 = 1 - 0,006 = 0,994.$$

2.5. Теорема додавання імовірностей сумісних подій

Теорема 5. *Якщо випадкові події A та B сумісні, то імовірність їх об'єднання дорівнює сумі їх імовірностей без імовірності їх сумісної появи, тобто*

$$P(A \cup B) = P(A) + P(B) - P(A \cdot B). \quad (1)$$

Доведення. Згідно з умовою теореми події A та B сумісні, тому $A \cup B$ з'явиться, якщо з'явиться одна з трьох несумісних подій $A \cdot \bar{B}$, $\bar{A} \cdot B$ або $A \cdot B$.

Згідно з теоремою додавання імовірностей несумісних подій одержимо

$$P(A \cup B) = P(A \cdot \bar{B}) + P(\bar{A} \cdot B) + P(A \cdot B). \quad (2)$$

Подія A з'явиться, якщо з'явиться одна з двох несумісних подій $A \cdot \bar{B}$ або $A \cdot B$.

Згідно з теоремою додавання імовірностей несумісних подій

$$P(A) = P(A \cdot \bar{B}) + P(A \cdot B) \Rightarrow P(A \cdot \bar{B}) = P(A) - P(A \cdot B). \quad (3)$$

Аналогічно одержимо

$$P(B) = P(\bar{A} \cdot B) + P(A \cdot B) \Rightarrow P(\bar{A} \cdot B) = P(B) - P(A \cdot B). \quad (4)$$

Підставимо (3) та (4) у формулу (2), тоді одержимо рівність (1), яку треба було довести.

Зауваження. *Якщо події A та B незалежні, то формула (1) приймає вигляд*

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

Для залежних випадкових подій маємо

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

Приклад 1. У залежності від наявності сировини підприємство може виробити та відправити замовникам щодобово кількість певної продукції від 1 до 100. Яка імовірність того, що одержану кількість продукції можна розподілити без залишку.

- а) трьом замовникам;
- б) чотирьом замовникам;
- в) дванадцяти замовникам;
- г) трьом або чотирьом замовникам?

Розв'язання. Позначимо події

A – одержана кількість виробів ділиться на 3 без залишку;

B – одержана кількість виробів ділиться на 4 без залишку.

Використовуючи класичне означення імовірності, знаходимо

$$\text{а) } P(A) = \frac{33}{100}; \quad \text{б) } P(B) = \frac{25}{100}; \quad \text{в) } P(A \cdot B) = \frac{8}{100}.$$

Події A та B – сумісні, тому за формулою (1) одержимо

$$\text{г) } P(A \cup B) = P(A) + P(B) - P(A \cdot B) = \frac{33}{100} + \frac{25}{100} - \frac{8}{100} = \frac{1}{2}.$$

2.6. Надійність системи

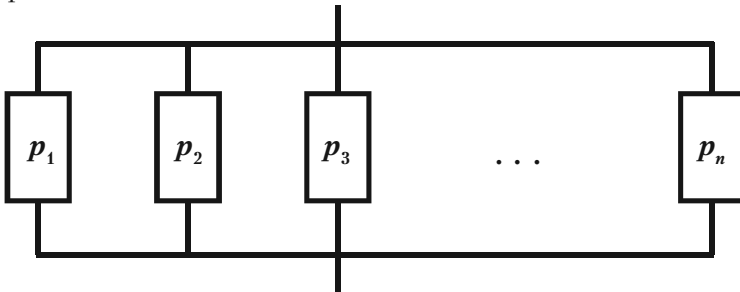
Означення 1. *Надійністю системи називають імовірність її безвідмовної роботи в певний час t (гарантійний термін).*

Системи складаються з елементів, поєднаних послідовно



Мал. 5.

або паралельно



Мал. 6.

Розділ 2. Основні теореми теорії імовірностей

При обчисленні надійності систем необхідно виразити надійність системи через надійність елементів та блоків.

Надійність елементів вважається відомою, бо вона пов'язана з технологією їх виготовлення.

Позначимо за p_k надійність k -того елемента, q_k – імовірність виходу з строю за час t k -того елемента, P – надійність блоку.

Розглянемо блок, усі елементи якого незалежні і з'єднані послідовно (див. мал. 5).

Такий блок буде працювати безвідмовно лише в той час, коли усі елементи працюють безвідмовно. Згідно теореми множення імовірностей незалежних подій імовірність P безвідмовної роботи такого блоку буде

$$P = p_1 \cdot p_2 \cdots p_n. \quad (1)$$

Тепер розглянемо блок, елементи якого з'єднані паралельно (див. мал. 6).

Такий блок буде працювати безвідмовно, якщо хоч один елемент не вийде зі строю. Тому імовірність P безвідмовної роботи буде

$$P = 1 - q_1 \cdot q_2 \cdots q_n. \quad (2)$$

Будь-яку складну систему можна розглядати як послідовне або паралельне з'єднання блоків, надійність яких обчислюють за формулами (1) та (2).

Приклад 1. Прилад складено з двох блоків, з'єднаних послідовно і незалежно працюючих. Імовірність відмови блоків дорівнює 0,05 та 0,08. Знайти імовірність відмови приладу.

Розв'язання. Відмова приладу є подія протилежна його безвідмовної роботи. Імовірності безвідмовної роботи блоків будуть

$$p_1 = 1 - 0,05 = 0,95; \quad p_2 = 1 - 0,08 = 0,92.$$

Імовірність безвідмовної роботи приладу буде згідно формули (1)

$$P_1 = 0,95 \cdot 0,92 = 0,874.$$

Тому імовірність відмови приладу буде

$$P = 1 - 0,874 = 0,126.$$

2.7. Формули повної імовірності та Байєса

Теорема 6. *Якщо випадкова подія A може з'явитись лише сумісно з однією із несумісних між собою подій B_1, B_2, \dots, B_n , що утворюють повну групу, тоді імовірність події A обчислюється за формулою*

$$P(A) = \sum_{k=1}^n P(B_k)P_{B_k}(A). \quad (1)$$

Доведення. За умовою теореми поява події A означає появу однієї з подій AB_1, AB_2, \dots, AB_n , тобто

$$A = AB_1 \cup AB_2 \cup \dots \cup AB_n.$$

Події B_1, B_2, \dots, B_n несумісні, тому й події AB_1, AB_2, \dots, AB_n також несумісні. Згідно з теоремою додавання імовірностей несумісних подій маємо

$$P(A) = P(AB_1 \cup AB_2 \cup \dots \cup AB_n) = \sum_{k=1}^n P(AB_k). \quad (2)$$

Події A та B_k – залежні, тому для обчислення $P(AB_k)$ можна використати теорему множення імовірностей залежних подій, тобто

$$P(AB_k) = P(B_k)P_{B_k}(A). \quad (3)$$

Підставимо (3) у формулу (2) і одержимо рівність (1), яку треба було довести.

Формулу (1) називають формулою повної імовірності.

Приклад 1. У першому ящику 20 деталей, з яких 15 стандартних. У другому ящику 10 деталей, з яких 9 стандартних. З другого ящика беруть навмання одну деталь і перекладають її до першого ящика. Знайти імовірність того, що взята після цього навмання деталь з першого ящика стандартна.

Розв'язання. Позначимо такі події: A – з першого ящика взято стандартну деталь; B_1 – з другого ящика переклали до першого стандартну деталь; B_2 – з другого ящика переклали до першого нестандартну деталь.

Розділ 2. Основні теореми теорії імовірностей

Згідно з умовою задачі, з першого ящика можна взяти деталь лише після того, як здійсниться подія B_1 або подія B_2 .

Події B_1 та B_2 несумісні, а подія A може з'явитись лише сумісно з однією із них. Тому для знаходження імовірності події A можна використати формулу повної імовірності (1), яка у даному випадку прийме вигляд

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A). \quad (4)$$

Знайдемо потрібні імовірності

$$P(B_1) = \frac{9}{10}; \quad P(B_2) = \frac{1}{10}; \quad P_{B_1}(A) = \frac{16}{21}; \quad P_{B_2}(A) = \frac{15}{21}.$$

Підставимо ці значення у формулу (4) і одержимо

$$P(A) = \frac{9}{10} \cdot \frac{16}{21} + \frac{1}{10} \cdot \frac{15}{21} = \frac{144 + 15}{210} = \frac{159}{210} = \frac{53}{70}.$$

Тепер ознайомимось з формулами Байєса.

В умовах Теорема 1 невідомо, з якою подією із несумісних подій B_1, B_2, \dots, B_n з'явиться подія A . Тому кожна з подій B_1, B_2, \dots, B_n можна вважати гіпотезою. Тоді $P_A(B_k)$ – імовірність k -тої гіпотези.

Якщо випробування проведено і в результаті його подія A з'явилась, то умовна імовірність $P_A(B_k)$ може не дорівнювати $P(B_k)$.

Порівняння імовірностей $P(B_k)$ та $P_A(B_k)$ дозволяє переоцінити імовірність гіпотези при умові, що подія A з'явилась.

Для одержання умовної імовірності використовуємо теорему множення імовірностей залежних подій

$$P(AB_k) = P(B_k)P_{B_k}(A) = P(A)P_A(B_k) \Rightarrow$$
$$P_A(B_k) = \frac{P(B_k)P_{B_k}(A)}{P(A)}. \quad (5)$$

Підставимо у формулу (5) замість $P(A)$ її значення з формули повної імовірності. Одержимо

$$P_A(B_k) = \frac{P(B_k)P_{B_k}(A)}{\sum_{k=1}^n P(B_k)P_{B_k}(A)}, \quad k = 1, 2, \dots, n. \quad (6)$$

Формули (6) називають **формулами Байєса**. Вони дозволяють переоцінювати імовірності гіпотез. Це важливо при контролі або ревізіях.

Приклад 2. Деталі, виготовлені цехом заводу, попадають для перевірки їх стандартності до одного з двох контролерів. Імовірність того, що деталь попаде до першого контролера, дорівнює 0,6, а до другого — 0,4. Імовірність того, що придатна деталь буде признана стандартною першим контролером, дорівнює 0,94, а другим — 0,98.

Придатна деталь при перевірці признана стандартною. Знайти імовірність того, що деталь перевіряв перший контролер.

Розв'язання. Позначимо такі події: A — придатна деталь признана стандартною; B_1 — деталь перевіряв перший контролер; B_2 — деталь перевіряв другий контролер. За умовою прикладу

$$P(B_1) = 0,6; \quad P(B_2) = 0,4; \quad P_{B_1}(A) = 0,94; \quad P_{B_2}(A) = 0,98.$$

За формулою Байєса (6) при $k = 1$ одержимо

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A)} = \frac{0,6 \cdot 0,94}{0,6 \cdot 0,94 + 0,4 \cdot 0,98} = 0,59.$$

Відмітимо, що до появи події A імовірність $P(B_1) = 0,6$, а після появи події A імовірність перевірки деталі першим контролером $P_A(B_1) = 0,59$ поменшала.

Приклад 3. Імовірність знищення літака з одного пострілу для першої гармати дорівнює 0,2, а для другої гармати — 0,1. Кожна гармата робить по одному пострілу, причому було одне влучення у літак. Яка імовірність того, що влучила перша гармата?

Розв'язання. Позначимо такі події: A — знищення літака з одного пострілу першою гарматою; B — знищення літака з одного пост-

Розділ 2. Основні теореми теорії імовірностей

рілу другою гарматою; C – одне влучення у літак. Маємо чотири гіпотези

$$H_1 = A \cdot B; \quad H_2 = A \cdot \bar{B}; \quad H_3 = \bar{A} \cdot B; \quad H_4 = \bar{A} \cdot \bar{B}$$

які утворюють повну групу подій. Імовірностями цих гіпотез будуть

$$P(H_1) = 0,2 \cdot 0,1 = 0,02; \quad P(H_2) = 0,2 \cdot 0,9 = 0,18;$$

$$P(H_3) = 0,8 \cdot 0,1 = 0,08; \quad P(H_4) = 0,8 \cdot 0,9 = 0,72.$$

Так як сума

$$H_1 + H_2 + H_3 + H_4$$

є достовірною подією, то

$$P(H_1) + P(H_2) + P(H_3) + P(H_4) = 1.$$

Умовні імовірності події C будуть

$$P_{H_1}(C) = 0; \quad P_{H_2}(C) = 1; \quad P_{H_3}(C) = 1; \quad P_{H_4}(C) = 0.$$

Тепер за формулою Байєса знаходимо шукану імовірність

$$P_C(H_2) = \frac{0,18 \cdot 1}{0,18 \cdot 1 + 0,08 \cdot 1} = 0,7.$$

Питання для самоперевірки та вправи

1. Як формулюють і якими формулами записують теореми додавання імовірностей сумісних та несумісних подій?
2. Які випадкові події називають незалежними?
3. Як визначають та позначають умовну імовірність?
4. Як формулюють і якими формулами записують теореми множення імовірностей залежних та незалежних випадкових подій?
5. За якою формулою можна обчислити імовірність появи хоча б однієї з n сумісних подій?
6. Що називають надійністю системи? Як знайти імовірність безвідмовної роботи блоку, усі елементи якого незалежні і з'єднані послідовно, паралельно? Як треба розглядати та досліджувати надійність будь-якої складної системи?
7. Яким умовам повинна задовольняти подія, щоб її імовірність можна було знаходити за формулою повної імовірності? Який вигляд має ця формула?
8. Коли застосовують формули Байєса та як їх записують?

Розв'язати задачі

1. Стрілок влучить у «десятку» з імовірністю 0,05, в «дев'ятку» – з імовірністю 0,20, у «вісімку» – з імовірністю 0,6.

Зроблено один постріл. Яка імовірність того, що а) вибито не менше 8 очок; б) вибито більше 8 очок?

Відповідь: а) 0,85; б) 0,25.

2. Транспортування вантажу для замовника виконують два автотранспортних підприємства, кожне з яких повинно виділити для цього по одній вантажівці. Імовірність виходу на лінію вантажівки з першого автотранспорту дорівнює 0,7, а з другого – 0,6. Знайти імовірність того, що замовник одержить потрібні вантажівки.

Відповідь: 0,88.

3. В урні 10 білих, 15 чорних, 20 блакитних та 25 червоних куль однакового розміру. Навмання взято одну кулю. Знайти імовірність того, що ця куля буде

а) біла або чорна; б) блакитна або червона; в) біла, чорна або блакитна.

Відповідь: а) $\frac{5}{14}$; б) $\frac{9}{14}$; в) $\frac{9}{14}$.

4. У білому ящику 12 червоних та 6 блакитних куль. У чорному ящику – 15 червоних та 10 блакитних куль. Кидають гральний кубик.

Якщо випаде кількість очок, кратна 3, то навання беруть кулю з білого ящика.

Якщо випаде будь-яка інша кількість очок, то беруть кулю з чорного ящика.

Якою буде імовірність узяти червону кулю?

Відповідь: $\frac{28}{45}$.

5. Із колоди карт (32 карти) навання взято одну. Яка імовірність того, що це дама, якщо відомо, що взято карту червоної масті?

Відповідь: $\frac{1}{8}$.

6. Кожна з літер слова «інтеграл» записана на окремому аркуші. Аркуші перемішані. Яка імовірність того, що

а) з'явиться слово «гра» при витягуванні трьох аркушів (у порядку їх появи);

б) з перших чотирьох узятих літер можна скласти слово «трал»?

Відповідь: а) $\frac{1}{336}$; б) $\frac{1}{70}$.

7. У коробці 5 однакових виробів, але 3 з них забарвлені. Навання беруть 2 вироби. Знайти імовірність того, що серед них

а) два забарвлених вироби;

б) хоча б один забарвлений виріб.

Відповідь: а) 0,3; б) 0,9.

8. Перевіряють половину виготовлених виробів. Умови прийомки дозволяють не більше 2% браку. Визначити імовірність того, що 100 виробів, які мають 5% браку, будуть прийнятими.

Відповідь: $\frac{C_{95}^{50}}{C_{100}^{50}} + \frac{C_5^1 \cdot C_{95}^{49}}{C_{100}^{50}} \approx 0,18$.

9. Обчислити імовірність того, що 50 виробів, серед яких 3 нестандартних, будуть прийнятими при перевірці навмання вибраної половини виробів. Умовами прийому дозволено не більше 4% нестандартних виробів.

$$\text{Відповідь: } \frac{C_{47}^{25}}{C_{50}^{25}} + \frac{C_3^1 \cdot C_{47}^{24}}{C_{50}^{25}} = 0,5.$$

10. Прилад складено з 4 блоків. Імовірність справної роботи кожного блоку в проміжок T годин дорівнює 0,9. Якою є імовірність справної роботи приладу в цей проміжок часу T , якщо вихід з ладу кожного блоку призводить до виходу з ладу всього приладу, та відомо, що блоки виходять з ладу незалежно один від одного?

$$\text{Відповідь: } 0,6561.$$

11. Два стрілки незалежно один від одного стріляють в одну мішень. Імовірність влучення першим стрільком дорівнює 0,3, а другим — 0,4. При будь-якому влученні мішень зіпсується. Знайти імовірність того, що мішень буде зіпсовано.

$$\text{Відповідь: } 0,58.$$

12. Імовірність того, що напруга в електричній мережі стане вищою за номінальну, дорівнює p_1 . Коли напруга збільшується, імовірність аварії приладу-споживача електроенергії дорівнює p_2 . Визначити імовірність аварії приладу внаслідок збільшення напруги в мережі.

$$\text{Відповідь: } p_1 \cdot p_2.$$

13. В урні 7 білих та 3 чорних кульки. Послідовно з урни витягнули дві кульки (не повертаючи їх до урни). Знайти імовірність того, що обидві витягнуті кульки — білі.

$$\text{Відповідь: } \frac{7}{15}.$$

14. Чому дорівнює імовірність того, що при киданні трьох гральних кубиків число 5 випаде хоч би на одному з них?

$$\text{Відповідь: } \frac{91}{216}.$$

15. Працівник обслуговує три станка, що працюють незалежно один від одного. Імовірність того, що на протязі години перший станок не вимагатиме уваги працівника, дорівнює 0,9; а для другого та третього станків — 0,8 та 0,85 відповідно. Якою є імовірність того, що на протязі години

- а) три станка не вимагатимуть уваги працівника;
- б) усі три станки потребують уваги працівника;
- в) хоч один станок потребує уваги працівника?

Відповідь: а) 0,612; б) 0,003; в) 0,388.

16. На складі 70% кінескопів, виготовлених заводом N_1 а решта – виготовлені заводом N_2 . Імовірність того, що кінескоп, виготовлений заводом N_1 , витримає гарантійний термін, дорівнює 0,8. Ця ж імовірність для кінескопа заводу N_2 дорівнює 0,9. Знайти імовірність того, що навмання взятий кінескоп

- а) не витримає гарантійного терміну;
- б) витримає гарантійний термін.

Відповідь: а) 0,17; б) 0,83.

17. Для сіяння пшениці заготовлено насіння 95% першого сорту, 3% другого сорту, 2% третього сорту. Імовірність того, що із насіння виросте колосся, яке містить 50 насінин, становить для першого сорту 0,5, для другого сорту — 0,2, а для третього — 0,1. Знайти імовірність того, що навмання взяте колосся буде мати не менше 50 насінин.

Відповідь: 0,483.

18. Знайти імовірність $P(A)$, якщо

$$P(A \cdot B) = 0,72; P(A \cdot \bar{B}) = 0,18.$$

Відповідь: 0,9.

19. У трьох ящиках маємо однакові деталі з різних заводів:

- у першому – 20 стандартних та 5 нестандартних деталей;
- у другому – 15 стандартних та 3 нестандартних;
- у третьому – 14 стандартних та 2 нестандартних.

Із навмання взятого ящика навмання взята деталь, яка виявилася стандартною. Знайти імовірність того, що ця деталь взято

- а) з першого ящика; б) із другого ящика; в) із третього ящика.

Відповідь: а) 0,319; б) 0,332; в) 0,349.

20. Система має два незалежно працюючих елемента. Імовірність їх відмови дорівнює 0,05 та 0,08 відповідно. Знайти імовірність відмови системи, якщо для цього достатньо відмови хоча б одного з елементів.

Відповідь: 0,126.

21. Імовірність хоча б одного влучення в ціль при трьох пострілах дорівнює 0,875. Знайти імовірність влучення при одному пострілі.

Відповідь: 0,5.

22. Із 100 виробів 5 є бракованими. Якою є імовірність взяти навмання два стандартні вироби?

Відповідь: $\approx 0,9$.

23. У обчислювальному центрі є 6 та 4 комп'ютерів *IBMAT386* та *IBMXT286* відповідно. Імовірність виходу з ладу під час роботи для 386-ї моделі дорівнює 0,05, а для 286-ї моделі — 0,2. Студент виконує лабораторну роботу на обраному навмання комп'ютері. Знайти імовірність того, що під час роботи комп'ютер не вийде з ладу.

Відповідь: 0,89.

24. У першій урні 10 куль, з них 8 білих. У другій урні 20 куль, з них 4 білих. Із кожної урни навмання взято по одній кулі, а потім із двох обраних навмання взято одну. Знайти імовірність того, що остання куля буде білою.

Відповідь: 0,5.

25. У лікарню поступають (у середньому) 50% хворих на грип, 30% хворих на ангіну та 20% хворих на запалення легенів. Імовірність повного одужання від грипу дорівнює 0,7, від ангіни та запалення легенів — 0,8 та 0,9 відповідно. Виписано хворого, який повністю одужав. Знайти імовірність того, що він був хворий на грип.

Відповідь: $\frac{5}{11}$.

26. На полиці у випадковому порядку стоять 15 підручників, причому 5 із них у твердій обкладинці. З полиці навмання беруть 3 підручника. Знайти імовірність того, що хоча б один із них буде з твердою обкладинкою.

Відповідь: $\frac{67}{91}$.

27. Із колоди карт (32 карти) навмання беруть послідовно три карти. Знайти імовірність того, що це будуть сімка, дама, туз.

Відповідь: $\frac{1}{465}$.

28. Одного із трьох стрілків викликають на лінію вогню і він виконує два постріла. Імовірність влучення у мішень при одному пострілі першим стрілком дорівнює 0,3, другим – 0,5, а третім – 0,8. У мішені немає влучень. Знайти імовірність того, що постріли виконані першим стрілком.

Відповідь: $\frac{0,49}{0,78} \approx 0,6282$.

29. На трьох автоматичних лініях виготовляють однакові деталі, причому 30% на першій лінії, 25% – на другій та 45% – на третій. Імовірність виготовлення стандартної деталі першою лінією дорівнює 0,99, другою – 0,988, третьою – 0,98. Виготовлені на протязі доби деталі надходять до складу. Визначити імовірність того, що навмання взята деталь не відповідає стандарту.

Відповідь: 0,015.

30. У двох урнах є кулі, що відрізняються лише кольором, причому у першій урні 5 білих, 11 чорних та 8 червоних куль, а у другій урні 10 білих, 8 чорних та 6 червоних. Із обох урн навмання виймають по одній кулі. Яка імовірність того, що обидві кулі будуть одного кольору?

Відповідь: $\frac{31}{96}$.

31. Із колоди (36 карт) навмання беруть дві карти. Знайти імовірність того, що хоча б одна з цих карт буде пікової масті.

Відповідь: $\frac{31}{70}$.

32. Імовірність того, що подія відбудеться хоча б один раз у трьох незалежних випробуваннях, дорівнює 0,936. Знайти імовірність появи події в одному випробуванні.

Відповідь: 0,4.

33. Із трьох гармат по одній цілі зроблено по одному пострілу. Імовірність влучення із першої гармати дорівнює 0,9, із другої — 0,8, а із третьої — 0,7. Знайти імовірність двох влучень у ціль.

Відповідь: 0,398.

34. Відомо, що імовірність влучення даним стрільком у мішень хоча б один раз із трьох пострілів дорівнює 0,75. Знайти імовірність того, що він із двох пострілів обидва рази влучив у мішень.

Відповідь: $(0,25)^{\frac{2}{3}}$.

35. Імовірність того, що потрібна студентові сума грошей є в одного з його чотирьох друзів дорівнює 0,6; 0,7; 0,8 та 0,9 відповідно. Знайти імовірність того, що необхідна сума грошей є не менше ніж у одного із друзів.

Відповідь: $1 - P(0) = 0,9976$.

36. У крамницю для продажу надійшла продукція з трьох фабрик, відносними частками яких є: 1 — 50%; 2 — 30%; 3 — 20%. Брак продукції цих фабрик становить 2%; 3% та 5% відповідно. Знайти імовірність того, що навання куплений у магазині виріб цих фабрик буде якісним.

Відповідь: 0,971.

37. Знайти імовірність того, що при киданні трьох гральних кубиків число 3 випаде хоча б на одному з них.

Відповідь: $\frac{91}{216}$.

38. У майстерні на верстатах a , b , c виробляють 25%, 35% та 40% усіх деталей, причому вони мають 15%; 12% та 6% браку відповідно. Знайти імовірність того, що навання взята деталь — бракована.

Відповідь: 0,1035.

39. Маємо два однакових ящики. У першому з них 8 пар взуття 41-го розміру та 6 пар 42-го розміру, а в другому — 10 пар 41-го

Розділ 2. Основні теореми теорії імовірностей

розміру та 4 пари 42-го розміру. З вибраного навання ящика витягли одну пару взуття 42-го розміру. Знайти імовірність того, що ця пара взята з першого ящика.

Відповідь: 0,6.

40. У рибалки є три улюблених місця, куди він приходить з однаковою імовірністю. Імовірність кльову на першому місці дорівнює $\frac{1}{3}$, на другому – $\frac{1}{2}$, на третьому – $\frac{1}{4}$. Рибалка закинув вудку

у навання вибраному місці і риба клюнула. Знайти імовірність того, що рибалка закинув вудку у першому місці.

Відповідь: $\frac{4}{13}$.

41. У першій урні 2 білих та 4 чорних кулі, а у другій урні – 3 білих та 1 чорна кулі. Із першої урни переклали у другу одну кулю. Знайти імовірність того, що куля, вийнята із другої урни після перекладання, буде білою.

Відповідь: 0,4.

42. Підприємство виготовляє 98% стандартних виробів, причому 90% з них – першого сорту. Знайти імовірність того, що взятий навання виріб буде першого сорту.

Відповідь: 0,882.

43. Робітник отримав 16 деталей, які виготовлені заводом N_1 , та 4 деталі заводом N_2 . Він навання бере дві деталі. Знайти імовірність того, що хоча б одна з них виготовлена заводом N_1 .

Відповідь: $1 - \frac{C_4^2}{C_{20}^2} = \frac{92}{95}$.

44. У магазин надійшло 30% телевізорів із першого заводу, серед яких 20% бракованих; 20% – із другого заводу, серед яких 10% бракованих; 50% із третього заводу, який має лише 5% браку. Яка імовірність купити справний телевізор недосвідченим покупцем?

Відповідь: 0,895.

45. Імовірність відмови мікросхеми дорівнює 0,4. Для підвищення надійності приладу мікросхему дублюють, тобто замість однієї

беруть n . Яким повинно бути n , щоб імовірність безвідмовної роботи дорівнювала 0,99?

$$\text{Відповідь: } n \geq \frac{\lg 0,01}{\lg 0,4} \approx 5.$$

46. Подія A може з'явитись при умові появи однієї із несумісних подій B_1, B_2, B_3 , які утворюють повну групу подій. Були знайдені умовні імовірності

$$P_A(B_1) = 0,6 \quad \text{та} \quad P_A(B_2) = 0,3$$

після появи події A .

Чому дорівнює умовна імовірність $P_A(B_3)$ гіпотези B_3 ?

$$\text{Відповідь: } P_A(B_3) = 1 - (0,6 + 0,3) = 0,1.$$

47. Серед студентів IV курсу $\frac{2}{5}$ одружені та $\frac{3}{5}$ не одружені.

Серед неодружених студентів $\frac{1}{2}$ молодші 22 років, а $\frac{2}{3}$ одружених старші 22 років. Знайти імовірність того, що довільно обраний студент цього курсу

а) старший 22 років; б) одружений та старший 22 років.

$$\text{Відповідь: а) } \frac{3}{5}; \quad \text{б) } \frac{4}{15}.$$

Розділ 3

ПОСЛІДОВНОСТІ ВИПРОБУВАНЬ

*Послідовно, по сходинках
із граничних теорем
піднімаємось мов у ліфті
до центральних теорем.*

3.1. Схема та формула Бернуллі

У багатьох задачах теорії імовірностей, статистики та повсякденної практики треба досліджувати послідовність (серію) n випробувань. Наприклад, випробування «кинуто 1000 однакових монет» можна розглядати як послідовність 1000 більш простих випробувань — «кинуто одна монета». При киданні 1000 монет імовірність появи герба або надпису на одній монеті не залежить від того, що з'явиться на інших монетах. Тому можна казати, що у цьому випадку випробування повторюються 1000 разів незалежним чином.

Означення 1. *Якщо усі n випробувань проводити в однакових умовах і імовірність появи події A в усіх випробуваннях однакова та не залежить від появи або неяви A в інших випробуваннях, то таку послідовність незалежних випробувань називають **схемою Бернуллі**.*

Нехай випадкова подія A може з'явитись у кожному випробуванні з імовірністю $P(A) = p$ або не з'явитись з імовірністю

$$q = P(\bar{A}) = 1 - p.$$

Поставимо задачу: знайти імовірність того, що при n випробуваннях подія A з'явиться m разів і не з'явиться $n - m$ разів. Шукану імовірність позначимо $P_n(m)$.

Спочатку розглянемо появу події A три рази в чотирьох випробуваннях. Можливі такі події

$$AAAA, AA\bar{A}, A\bar{A}A, \bar{A}AAA,$$

тобто їх $4 = C_4^3$.

Якщо подія A з'явилася 2 рази в 4 випробуваннях, то можливі такі події

$$AA\bar{A}\bar{A}, A\bar{A}A\bar{A}, \bar{A}\bar{A}AA, \bar{A}A\bar{A}\bar{A}, \bar{A}\bar{A}\bar{A}A, \bar{A}\bar{A}A\bar{A},$$

$$\text{їх } 6 = C_4^2.$$

У загальному випадку, коли подія A з'являється m разів у n випробуваннях, таких складних подій буде

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Обчислимо імовірність однієї складної події, наприклад,

$$\underbrace{A \cdot A \cdots A}_m \cdot \underbrace{\bar{A} \cdot \bar{A} \cdots \bar{A}}_{n-m}.$$

Імовірність сумісної появи n незалежних подій дорівнює добутку імовірностей цих подій згідно з теоремою множення імовірностей, тобто

$$\begin{aligned} P &= \underbrace{(A \cdot A \cdots A)}_m \cdot \underbrace{(\bar{A} \cdot \bar{A} \cdots \bar{A})}_{n-m} = P(\underbrace{A \cdot A \cdots A}_m) \cdot P(\underbrace{\bar{A} \cdot \bar{A} \cdots \bar{A}}_{n-m}) = \\ &= P^m(A)P^{n-m}(\bar{A}) = p^m q^{n-m}. \end{aligned}$$

Кількість таких складних подій C_n^m і вони несумісні. Тому, згідно з теоремою додавання імовірностей несумісних подій, маємо

$$P_n(m) = C_n^m p^m q^{n-m}. \quad (1)$$

Формулу (1) називають **формулою Бернуллі**. Вона дозволяє знаходити імовірність появи події A m разів при n випробуваннях, які утворюють схему Бернуллі.

Зауваження 1. Імовірність появи події A в n випробуваннях схеми Бернуллі менше m разів знаходять за формулою

$$P_n(k < m) = P_n(0) + P_n(1) + \cdots + P_n(m-1).$$

Імовірність появи події A не менше m разів можна знайти за формулою

$$P_n(k \geq m) = P_n(m) + P_n(m+1) + \cdots + P_n(n),$$

Розділ 3. Послідовності випробувань

або за формулою

$$P_n(k \geq m) = 1 - \sum_{k=0}^{m-1} P_n(k).$$

Імовірність появи події A хоча б один раз у n випробуваннях доцільно знаходити за формулою

$$P_n(1 \leq m \leq n) = 1 - q^n.$$

Зауваження 2. У багатьох випадках треба знаходити найбільші імовірне значення m_0 числа m появ події A . Це значення m визначається співвідношеннями

$$np - q \leq m_0 \leq np + p \quad \text{або} \quad (n+1)p - 1 \leq m_0 \leq (n+1)p.$$

Число m_0 повинно бути цілим. Якщо $(n+1)p$ – ціле число, тоді найбільше значення імовірність має при двох числах

$$m_1 = (n+1)p - 1 \quad \text{та} \quad m_2 = (n+1)p.$$

Зауваження 3. Якщо імовірність появи події A в кожному випробуванні дорівнює p , то кількість n випробувань, які необхідно здійснити, щоб з імовірністю P можна було стверджувати, що подія A з'явиться хоча б один раз, знаходять за формулою

$$n > \frac{\ln(1-P)}{\ln(1-p)}. \quad (2)$$

Приклад 1. Прилад складено з 10 блоків, надійність кожного з них 0,8. Блоки можуть виходити з ладу незалежно один від одного. Знайти імовірність того, що

- а) відмовлять два блоки;
- б) відмовить хоча б один блок;
- в) відмовлять не менше двох блоків.

Розв'язання. Позначимо за подією A відмову блока. Тоді імовірність події A за умовою прикладу буде

$$P(A) = p = 1 - 0,8 = 0,2, \quad \text{тому} \quad q = 1 - p = 1 - 0,2 = 0,8.$$

Згідно з умовою задачі $n = 10$. Використовуючи формулу Бернуллі та Зауваження 1, одержимо

$$\text{а) } P_{10}(2) = C_{10}^2 p^2 q^8 = C_{10}^2 (0,2)^2 (0,8)^8 = 0,202;$$

$$\text{б) } P_{10}(1 < m \leq 10) = 1 - P_{10}(0) = 1 - C_{10}^0 (0,2)^0 (0,8)^{10} = 0,8926;$$

$$\begin{aligned} \text{в) } P_{10}(2 \leq m \leq 10) &= 1 - (P_{10}(0) + P_{10}(1)) = \\ &= 1 - (C_{10}^0 (0,2)^0 (0,8)^{10} + C_{10}^1 (0,2)^1 (0,8)^9) = 0,6244. \end{aligned}$$

Приклад 2. За одну годину автомат виготовляє 20 деталей. За скільки годин імовірність виготовлення хоча б однієї бракованої деталі буде не менше 0,952, якщо імовірність браку будь-якої деталі дорівнює 0,01?

Розв'язання. Застосовуючи формулу (2), знайдемо спочатку таку кількість виготовлених деталей, щоб з імовірністю $p = 0,952$ можна було стверджувати про наявність хоча б однієї бракованої деталі, якщо імовірність браку за умовою $p = 0,01$

$$n \geq \frac{\ln(1-0,952)}{\ln(1-0,01)} = \frac{\ln 0,048}{\ln 0,99} \approx 300.$$

Отже, за час $t = \frac{300}{20} = 15$ (годин) автомат з імовірністю 0,952

виготовить хоча б одну браковану деталь.

Приклад 3. При новому технологічному процесі 80% усієї виготовленої продукції має найвищу якість. Знайти найбільш імовірне число виготовлених виробів найвищої якості серед 250 виготовлених виробів.

Розв'язання. Позначимо шукане число m_0 .

Згідно Зауваження 2

$$np - q \leq m_0 \leq np + q.$$

За умовою прикладу $n = 250$, $p = 0,8$, $q = 0,2$, тому

$$199,8 \leq m_0 \leq 200,8.$$

Але m_0 повинно бути цілим числом, тому то $m_0 = 200$.

3.2. Граничні теореми у схемі Бернуллі

Знаходження імовірностей $P_n(m)$ та $P_n(m_1 \leq m \leq m_2)$ за формулою Бернуллі ускладнюється при досить великих значеннях n та при малих p або q . У таких випадках часто можна використовувати замість формули Бернуллі наближені асимптотичні формули.

Вкажемо без доведення три граничні теореми, які містять наближені формули для імовірностей

$$P_n(m) \text{ та } P_n(m_1 \leq m \leq m_2).$$

Теорема 1 (Теорема Пуассона). *Якщо $n \rightarrow \infty$ і $p \rightarrow 0$ так, що $np \rightarrow \lambda$, $0 < \lambda < \infty$, то*

$$P_n(m) = C_n^m p^m q^{n-m} \rightarrow \frac{\lambda^m}{m!} e^{-\lambda}$$

для будь-якого постійного $m = 0, 1, 2, \dots$.

Наслідок. Імовірність появи події A m разів у n випробуваннях схеми Бернуллі можна знаходити за наближеною формулою Пуассона

$$P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad (1)$$

де $\lambda = np$.

Формулу (1) доцільно застосовувати при великих n та малих p .

Приклад 1. Підручник надруковано тиражем 100000 екземплярів. Імовірність невірної брошурування підручника дорівнює 0,0001. Знайти імовірність того, що тираж має 5 бракованих підручників.

Розв'язання. Брошурування кожного підручника можна розглядати як випробування. Випробування незалежні і мають однакову імовірність невірної брошурування, тому задача вкладається у схему Бернуллі. Згідно з умовою задачі

$$n = 100000 \text{ досить велике; } p = 0,0001 \text{ мала; } m = 5.$$

Застосовуючи формулу Пуассона (1), одержимо

$$P_{100000}(5) = \frac{10^5}{5!} e^{-10} = 0,0375.$$

Для наведення ще двох граничних теорем треба спочатку визначити **локальну** та **інтегральну функції Лапласа** та ознайомитись з їх основними властивостями.

Означення 1. *Локальною функцією Лапласа називають функцію вигляду*

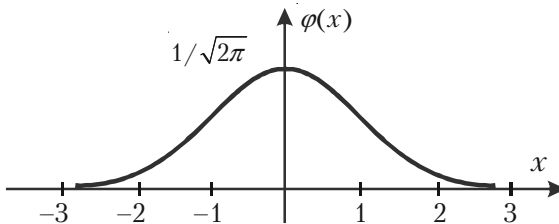
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Ця функція часто використовується, тому її значення для різних x наведені в підручниках та посібниках із теорії імовірностей. Вона табульована для додатних x .

Основні властивості локальної функції Лапласа

1. Функція Лапласа $\varphi(x)$ парна, тобто $\varphi(-x) = \varphi(x)$.
2. Функція $\varphi(x)$ визначена для усіх $x \in (-\infty, \infty)$.
3. $\varphi(x) \rightarrow 0$, коли $x \rightarrow \pm\infty$.
4. $\varphi_{\max} = \varphi(0) = \frac{1}{\sqrt{2\pi}}$.

Графік локальної функції Лапласа має вигляд, зображений на мал. 7.



Мал. 7.

Означення 2. Інтегральною функцією Лапласа називають функцію

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Легко бачити, що між локальною функцією $\varphi(x)$ та інтегральною функцією $\Phi(x)$ існує простий зв'язок

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt.$$

Основні властивості інтегральної функції Лапласа

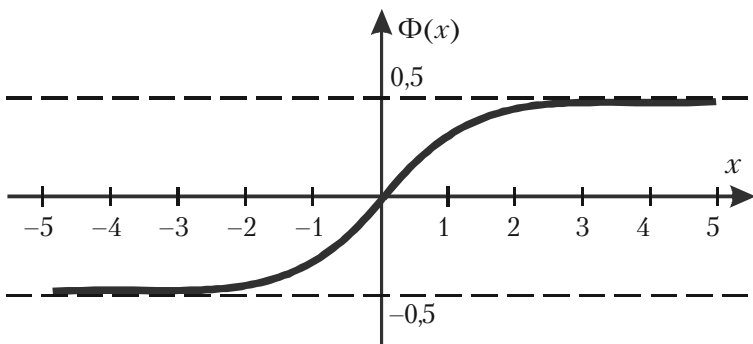
1. Інтегральна функція Лапласа є непарною функцією

$$\Phi(-x) = -\Phi(x).$$

2. $\Phi(0) = 0$.

3. $\Phi(x) = 0,5$ для $x \geq 5$.

Графік інтегральної функції Лапласа зображено на мал. 8.



Мал. 8.

Інтегральна функція Лапласа $\Phi(x)$ табульована для $x \in [0, 5]$.

Теорема 2 (локальна теорема Муавра-Лапласа). Якщо у схемі Бернуллі кількість випробувань n достатньо велика, а імовірність p появи події A в усіх випробуваннях однакова, то імовірність появи події A m разів може бути знайдена за наближеною формулою

$$P_n(m) = \frac{1}{\sqrt{npq}} \varphi(x_m), \quad (2)$$

де $x_m = \frac{m - np}{\sqrt{npq}}$.

Зауваження. Формулу (2) доцільно використовувати при $n > 100$ та $npq > 20$.

Приклад 2. Гральний кубик кидають 800 разів. Яка імовірність того, що кількість очок, кратна трьом, з'явиться 267 разів.

Розв'язання. У даному випадку n та m досить великі. Тому для знаходження $P_{800}(267)$ можна використати формулу (2). Маємо

$$P(A) = p = \frac{2}{6}, \quad q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3},$$

$$x_{267} = \frac{m - np}{\sqrt{npq}} = \frac{267 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = \frac{801 - 800}{40} = \frac{1}{40} = 0,025.$$

Отже, за формулою (2) одержимо

$$P_{800}(267) = \frac{3}{40} \cdot \varphi(0,025) = \frac{3}{40} \cdot 0,3988 = 0,03.$$

Значення $\varphi(0,025)$ взято з таблиці так

x	$\varphi(x)$	x	$\varphi(x)$
0,02	0,398862		
		$\Rightarrow 0,025$	$\frac{0,398862 + 0,398783}{2} \approx 0,3988$
0,03	0,398783		

Теорема 3 (Інтегральна теорема Муавра-Лапласа). *Якщо у схемі Бернуллі в кожному із n незалежних випробувань подія A може з'явитися з постійною імовірністю p , тоді імовірність появи події A не менш m_1 та не більш m_2 разів може бути знайдена за формулою*

$$P_n(m_1 \leq m \leq m_2) = \Phi(x_2) - \Phi(x_1), \quad (3)$$

де $\Phi(x)$ – інтегральна функція Лапласа,

$$x_2 = \frac{m_2 - np}{\sqrt{npq}}, \quad x_1 = \frac{m_1 - np}{\sqrt{npq}}. \quad (4)$$

Приклад 3. Гральний кубик кидають 800 разів. Яка імовірність того, що кількість очок, кратна трьом, з'явиться не менше 260 та не більше 274 разів?

Розв'язання. Для знаходження імовірності

$$P_{800}(260 \leq m \leq 274)$$

використаємо формули (4) та (3). Маємо

$$x_2 = \frac{274 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = \frac{22}{40} = 0,55, \quad x_1 = \frac{260 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = -0,5,$$

$$P_{800}(260 \leq m \leq 274) = \Phi(0,55) + \Phi(0,5) = 0,298840 + 0,191482 = 0,490322.$$

Значення інтегральної функції Лапласа взято з таблиці і використана властивість непарності $\Phi(-0,5) = -\Phi(0,5)$ функції $\Phi(x)$.

3.3. Послідовність випробувань із різними імовірностями

У схемі Бернуллі імовірність появи події A в усіх випробуваннях однакова. Але у практичній діяльності іноді зустрічаються і такі випадки, коли у n незалежних випробуваннях імовірності появи події A різні, наприклад, вони дорівнюють p_1, p_2, \dots, p_n .

Тоді імовірності не появи події A також будуть різними

$$q_1 = 1 - p_1, q_2 = 1 - p_2, \dots, q_n = 1 - p_n.$$

У цьому випадку не можна обчислювати за формулою Бернуллі імовірність появи події A m разів у n випробуваннях, а треба використовувати твірну функцію

$$\varphi_n(z) = \prod_{k=1}^n (q_k + p_k z). \quad (1)$$

Правило. *Шукана імовірність $P_n(m)$ дорівнює коефіцієнту, що стоїть при z^m .*

Приклад 1. Імовірність відмови кожного з 4 приладів у 4 незалежних випробуваннях різні і дорівнюють

$$p_1 = 0,1, p_2 = 0,2, p_3 = 0,3, p_4 = 0,4.$$

Знайти імовірність того, що внаслідок випробувань

- а) не відмовить жоден прилад;
- б) відмовлять один, два, три, чотири прилади;
- в) відмовить хоча б один прилад;
- г) відмовлять не менше двох приладів.

Розв'язання. Імовірності відмови приладів у випробуваннях різні, тому застосовуємо твірну функцію (1), яка у даному випадку матиме вигляд

$$\varphi_n(z) = (0,9 + 0,1z)(0,8 + 0,2z)(0,7 + 0,3z)(0,6 + 0,4z).$$

Розкриємо дужки та зведемо подібні члени. Тоді матимемо

$$\varphi_n(z) = 0,302 + 0,46z + 0,205z^2 + 0,031z^3 + 0,002z^4.$$

Відповідно Правилу звідси одержуємо відповіді на питання прикладу

а) $P_4(0) = 0,302$;

б) $P_4(1) = 0,46$; $P_4(2) = 0,205$; $P_3(1) = 0,031$; $P_4(4) = 0,002$;

в) $P_4(1 \leq m \leq 4) = 1 - P_4(0) = 0,698$;

г) $P_4(m \geq 2) = 1 - (P_4(0) + P_4(1)) = 1 - (0,302 + 0,46) = 0,238$.

Приклад 2. Працівник обслуговує три станка, що працюють незалежно один від одного. Імовірність того, що на протязі години перший станок не вимагатиме уваги працівника, дорівнює 0,9, а для другого та третього станків — 0,8 та 0,85, відповідно. Якою є імовірність того, що на протязі години

а) жоден станок не потребуватиме уваги працівника;

б) усі три станки потребують уваги працівника;

в) хоча б один станок потребує уваги працівника?

Розв'язання. Цей приклад можна розв'язати з використанням теорем множення та додавання імовірностей (дивись вправу 15 Розділу 2). Розв'яжемо тепер цей приклад з використанням твірної функції, яка у даному випадку прийме вигляд

$$\begin{aligned}\varphi_n(z) &= \prod_{k=1}^3 (q_k + p_k z) = (0,1 + 0,9z)(0,2 + 0,8z)(0,15 + 0,85z) = \\ &= 0,003 + 0,056z + 0,0329z^2 + 0,612z^3.\end{aligned}$$

Отже, коефіцієнт при z^k ($k = 0, 1, 2, 3$) дорівнює імовірності того, що на протязі години уваги працівника не потребують k станків. Тому одержуємо відповіді на питання цього прикладу:

а) імовірність того, що усі три станка не потребують уваги працівника, дорівнює коефіцієнту при z^3 , тобто

$$P_3(3) = 0,612;$$

б) $P_3(0) = 0,003$;

в) $P_3(1 \leq m \leq 3) = 1 - P_3(3) = 1 - 0,612 = 0,388$.

3.4. Теорема Бернуллі

Теорема Бернуллі встановлює зв'язок теорії імовірностей з її практичними застосуваннями. Вона була доведена Я.Бернуллі в кінці XVII ст., а опублікована у 1713 р.

Теорема 4 (Я.Бернуллі). Якщо у n незалежних випробуваннях імовірність p появи події A однакова і подія A з'явилася m разів, то для будь-якого додатнього числа a має місце рівність

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| \geq a\right) = 0, \quad (1)$$

тобто границя імовірності відхилення відносної частоти $\frac{m}{n}$ події A від її імовірності на величину, що більше або дорівнює a , дорівнює нулеві.

Згідно означенню границі рівність (1) означає, що

$$P\left(\left|\frac{m}{n} - p\right| \geq a\right) = \alpha \text{ — нескінченно мала величина.}$$

Та це означає, що подія

$$\left|\frac{m}{n} - p\right| \geq a$$

практично неможлива. Але тоді протилежна подія

$$\left|\frac{m}{n} - p\right| < a$$

практично достовірна для будь-якого додатнього числа a .

Наслідок з теореми Я.Бернуллі

Рівність

$$\left|\frac{m}{n} - p\right| = 0$$

може відрізнятись від практично достовірної події

$$\left| \frac{m}{n} - p \right| < a,$$

$a > 0$ на нескінченно малу величину.

Це означає, що $\frac{m}{n} \rightarrow p$, тобто відносна частота (частість)

$W(A) = \frac{m}{n}$ події A відрізняється від імовірності p події A на нескінченно малу величину, яку практично можна не враховувати.

Інше формулювання та доведення теореми Бернуллі дивись в підрозділі 4.4.3 Розділу 4.

Зауваження. Формулу (1) можна записати з використанням інтегральної функції Лапласа $\Phi(x)$ у вигляді

$$\lim_{x \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) = 2\Phi\left(\varepsilon \sqrt{\frac{n}{pq}}\right).$$

Звідси одержуємо важливу формулу

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

яка дозволяє розв'язувати багато задач.

Приклад 1. Імовірність появи події в кожному із 625 незалежних випробувань дорівнює 0,8. Знайти імовірність того, що частість появи події відхиляється від імовірності за абсолютною величиною не більше ніж на 0,04.

Розв'язання. За умовою прикладу $n = 625$, $p = 0,8$, $q = 1 - 0,8 = 0,2$,

$$\varepsilon = 0,04. \text{ Треба знайти } P\left(\left|\frac{m}{625} - 0,8\right| \leq 0,04\right).$$

За формулою (2) маємо

$$P\left(\left|\frac{m}{625} - 0,8\right| \leq 0,04\right) \approx 2\Phi\left(0,04 \sqrt{\frac{625}{0,8 \cdot 0,2}}\right) = 2\Phi(2,5).$$

З таблиці значень функції Лапласа $\Phi(x)$ знаходимо $\Phi(2,5) = 0,4938$. Отже,

$$2\Phi(2,5) = 2 \cdot 0,4938 = 0,9876 \Rightarrow P\left(\left|\frac{m}{625} - 0,8\right| \leq 0,04\right) = 0,9876.$$

Таким чином, шукана імовірність наближено дорівнює 0,9876.

Приклад 2. Імовірність появи події в кожному із незалежних випробувань дорівнює 0,5. Знайти число випробувань n , при якому з імовірністю 0,7698 можна чекати, що частота появи події відхиляється від її імовірності за абсолютною величиною не більше ніж на 0,02.

Розв'язання. За умовою задачі $p = 0,5$, $q = 0,5$, $\varepsilon = 0,02$,

$$P\left(\left|\frac{m}{n} - 0,5\right| \leq 0,02\right) = 0,7698.$$

Застосуємо формулу (2). Тоді згідно умови одержимо

$$2\Phi\left(0,02\sqrt{\frac{n}{0,5 \cdot 0,5}}\right) = 0,7698 \Rightarrow \Phi(0,04\sqrt{n}) = 0,3849.$$

Із таблиці значень інтегральної функції Лапласа знайдемо

$$0,3849 = \Phi(1,2) \Rightarrow 0,04\sqrt{n} = 1,2 \Rightarrow \sqrt{n} = 30 \Rightarrow n = 900.$$

Отже, шукана кількість випробувань $n = 900$.

Приклад 3. Відділ технічного контролю перевіряє стандартність 900 виробів. Імовірність того, що виріб стандартний, дорівнює 0,9. Знайти з імовірністю 0,9544 межі інтервалу, що містить число m стандартних виробів серед перевірених.

Розв'язання. За умовою $n = 900$, $p = 0,9$, $q = 0,1$,

$$2\Phi\left(\varepsilon\sqrt{\frac{900}{0,9 \cdot 0,1}}\right) = 0,9544 \Rightarrow \Phi(100\varepsilon) = 0,4772.$$

З таблиці значень інтегральної функції Лапласа знаходимо

$$0,4772 = \Phi(2), 100\varepsilon = 2 \Rightarrow \varepsilon = 0,02.$$

Отже, з імовірністю 0,9544 відхилення частоти кількості стандартних виробів від імовірності 0,9 задовольняє нерівність

$$\left| \frac{m}{900} - 0,9 \right| \leq 0,02 \Rightarrow 0,88 \leq \frac{m}{900} \leq 0,92.$$

З останніх співвідношень випливає, що шукане число m стандартних виробів серед 900 перевірених з імовірністю 0,9544 належить інтервалу $792 \leq m \leq 828$.

3.5. Проста течія подій

Означення 1. *Течією подій називають послідовність таких подій, які з'являються у випадкові моменти часу.*

Наприклад, заява до диспетчерського пункту з викликом таксі.

Означення 2. *Течія подій називається пуассонівською, якщо вона:*

1. **Стаціонарна**, тобто залежить від кількості k появ події та часу t і не залежить від моменту свого початку.

2. **Має властивість відсутності післядії**, тобто імовірність появи події не залежить від появи або не появи події раніше та не впливає на найближче майбутнє.

3. **Ординарна**, тобто імовірність появи більше однієї події в малій проміжок часу є величина нескінченно мала у порівнянні з імовірністю появи події один раз у цей проміжок часу.

Означення 3. *Середнє число λ появ події A в одиницю часу називають інтенсивністю течії.*

Теорема 5. *Якщо течія подій пуассонівська, то імовірність появи події A k разів за час t можна знайти за формулою*

$$P_t(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (1)$$

де λ – інтенсивність течії.

Зауваження 1. *Формулу (1) іноді звать математичною моделлю простої течії подій.*

Приклад 1. Середня кількість замовлень, що поступають до комбінату побутового обслуговування кожну годину, дорівнює 3. Знайти імовірність того, що за дві години поступлять

- а) 5 замовлень;
- б) менше 5 замовлень;
- в) не менше 5 замовлень.

Розв'язання. Маємо просту течію подій з інтенсивністю $\lambda = 3$. За формулою (1) одержуємо

$$\text{а) } P_2(5) = \frac{(3 \cdot 2)^5}{5!} e^{-3 \cdot 2} = \frac{(6)^5}{5!} e^{-6};$$

$$\text{б) } P_2(k < 5) = P_2(0) + P_2(1) + P_2(2) + P_2(3) + P_2(4) = 115e^{-6};$$

$$\text{в) } P_2(k \geq 5) = 1 - P_2(k < 5) = 1 - 115e^{-6}.$$

Зауваження 2. Прикладами простої течії подій можуть бути: поява викликів на АТС, на пункти швидкої медичної допомоги, прибуття літаків до аеропорту або клієнтів у підприємство побутового обслуговування, серія відмов елементів або блоків приладів та таке інше.

Питання для самоперевірки та вправи

1. Яка послідовність випробувань утворює схему Бернуллі?
2. Яку формулу звуть формулою Бернуллі і що вона дозволяє обчислювати?
3. За якими формулами знаходять імовірність появи події A менше m або не менше за m разів у n випробуваннях схеми Бернуллі?
4. За якою формулою знаходять імовірність появи події A хоча б один раз у n випробуваннях?
5. Як можна знайти найбільш імовірне значення числа появ події A у схемі Бернуллі?
6. Як можна знайти кількість випробувань у схемі Бернуллі, яка дозволяє з імовірністю P стверджувати, що подія A з'явиться хоча б один раз?
7. У яких випадках доцільно використовувати граничні теореми у схемі Бернуллі?
8. Коли доцільно застосовувати формули Пуассона, локальну або інтегральну формули Муавра-Лапласа?
9. Як визначаються і які мають властивості локальна та інтегральна функції Лапласа?
10. Як знаходять $P_n(m)$ у випадку послідовності випробувань із різними імовірностями?
11. Як формулюється теорема Бернуллі і який вона має наслідок?
12. Який існує зв'язок між твердженням теореми Бернуллі та інтегральною функцією Лапласа? Які задачі дозволяє розв'язувати цей зв'язок?
13. За якою формулою знаходять імовірність появи у випадку простої течії?

Розв'язати задачі

1. Є 12 стандартних та 4 нестандартних деталі. Навмання беруть 3 з них (з поверненням). Знайти імовірність того, що серед взятих деталей: а) усі три стандартні; б) не більше однієї стандартної; в) хоча б одна нестандартна.

Відповідь: а) $P_3(3) = 0,422$; б) $P_3(0) + P_3(1) = 0,156$;

в) $1 - P_3(3) = 0,578$.

2. У виробництві деякої продукції третій сорт складає 20%. Знайти імовірність того, що з 5 навмання взятих виробів цієї продукції не менше ніж три будуть третього сорту.

Відповідь: $P_5(3) + P_5(4) + P_5(5) \approx 0,012$.

3. На складі є вироби двох сортів, причому виробів другого сорту в 1,5 рази більше, ніж виробів першого сорту. Знайти імовірність того, що серед трьох навмання взятих виробів хоча б один першого сорту.

Відповідь: $1 - P_3(0) \approx 0,784$.

4. Імовірність виготовлення стандартного виробу дорівнює 0,95. Яка імовірність того, що серед 10 виробів не більше одного нестандартного?

Відповідь: $P_{10}(0) + P_{10}(1) = (0,95)^{10} + 0,315 = 0,9134$.

5. За даними технічного контролю 2% виготовлених заводом автоматичних верстатів потребують додаткового регулювання. Чому дорівнює імовірність того, що із 6 виготовлених верстатів 4 потребують додаткового регулювання?

Відповідь: 0,0000023.

6. Імовірність влучення стрільком у десятку дорівнює 0,6. Чому дорівнює імовірність того, що при 8 пострілах буде 6 влучень у десятку?

Відповідь: 0,209.

7. Імовірність влучення у мішень при одному пострілі дорівнює 0,4. Зроблено 6 незалежних пострілів. Знайти імовірність того, що буде хоча б одне влучення у мішень.

Відповідь: $1 - (0,6)^6 = 0,953344$.

8. Імовірність влучення у мішень при одному пострілі дорівнює 0,4. Зроблено 7 незалежних пострілів. Знайти імовірність хоча б одного влучення у мішень.

Відповідь: 0,972.

9. Імовірність появи події A хоча б один раз у 5 незалежних випробуваннях дорівнює 0,9. Яка імовірність появи події A в одному випробуванні, якщо при кожному випробуванні ця імовірність однакова?

Відповідь: $\approx 0,37$.

10. Імовірність появи події A в кожному з 15 незалежних випробувань дорівнює 0,3. Визначити імовірність появи події A не менше двох разів.

Відповідь: $\approx 0,965$.

11. Скільки разів треба кинути гральний кубик, щоб найімовірніше число появи трійки дорівнювало 55?

Відповідь: $n = 330$.

12. Чому дорівнює імовірність появи події в кожному випробуванні, якщо найімовірніше число появ події в 160 випробуваннях дорівнює 40?

Відповідь: $p = 0,25$.

13. Імовірність виготовлення нестандартної деталі дорівнює 0,05. Скільки треба виготовити деталей, щоб найімовірніше число нестандартних деталей було 55?

Відповідь: 1119.

14. Імовірність виготовлення виробу відмінної якості дорівнює 0,9. Виготовлено 50 виробів. Чому дорівнює найімовірніше число виробів відмінної якості та імовірність такого числа виробів?

Відповідь: $m_0 = 45$; $p \approx 0,18$.

15. База обслуговує 8 магазинів. Щодня вимоги на товари можуть поступити з імовірністю 0,6. Знайти найімовірніше число вимог, які можуть поступити у будь-який день.

Відповідь: $m_0 = 5$.

16. Прилад складено з 1000 незалежно працюючих елементів. Імовірність відмови будь-якого елемента за час T дорівнює 0,002. Знайти імовірність того, що за час T відмовлять 3 елементи.

Відповідь: $P(x = 3) = 0,18e^{-2} = 0,13534$.

17-20. Імовірність виробництва бракованого виробу дорівнює 0,008. Знайти імовірність того, що серед 1000 виробів бракованих буде не більше ніж 17) 3; 18) 4; 19) 5; 20) 10.

Відповідь: 17) $26,3e^{-8}$; 18) $297e^{-8}$; 19) $P(x \leq 4) + P(x = 5)$; 20) 0,719.

21. Завод відправив до бази 500 виробів. Імовірність пошкодження виробу при транспортуванні дорівнює 0,002. Знайти

імовірність того, що при транспортуванні буде пошкоджено виробів:
а) три; б) менше трьох; в) більше за три.

Відповідь: а) 0,0613; б) 0,9197; в) 0,019.

22. Засівний фонд має 92% насіння першого сорту. Навмання взято 150 насінин. Знайти імовірність того, що серед цих насінин 140 є першого сорту.

Відповідь: $P_{150}(140) \approx 0,1004$.

23. Виробництво дає 10% браку. Яка імовірність того, що із 200 навмання вибраних виробів 15 будуть бракованими?

Відповідь: $P_{200}(15) = \frac{0,236}{\sqrt{2\pi}} e^{-\frac{1}{2}(1,18)^2} = 0,047$.

24. Ткаля обслуговує 1000 веретен. Імовірність обриву нитки на одному веретені протягом однієї хвилини дорівнює 0,005. Знайти імовірність того, що протягом однієї хвилини обрив станеться на 7 веретенах.

Відповідь: $P_{1000}(7) = \frac{5^7}{7!} e^{-5}$.

25. Процент проростання пшеничного насіння становить 95%. Знайти імовірність того, що з 2000 посіяних насінин проросте від 1880 до 1920.

Відповідь: $\Phi(2,06) - \Phi(-2,06) = 0,9606$.

26. Електростанція обслуговує мережу з 10000 ламп; імовірність включення кожної з них 0,6. Визначити імовірність одночасного включення від 5900 до 6100 ламп.

Відповідь: $P(5900 \leq x \leq 6100) = 0,9624$.

27. Імовірність присутності студента на лекції дорівнює 0,8. Знайти імовірність того, що із 100 студентів на лекції будуть а) не менше 75 та не більше 90; б) не менше 75; в) не більше 74.

Відповідь:

а) $P_{100}(75 \leq m \leq 90) = 0,4938 + 0,3944 = 0,8882$;

б) $P_{100}(75 \leq m \leq 100) = \Phi(5) - \Phi(-1,25) = 0,5 + 0,3944 = 0,8944$;

в) $P_{100}(0 \leq m \leq 74) = 0,1056$.

28. Імовірність появи додатнього результату в кожному із n випробувань дорівнює 0,9. Скільки треба здійснити випробувань, щоб з імовірністю 0,98 можна було чекати, що не менше 150 випробувань будуть мати додатний результат?

Відповідь: $n = 177$.

29. Досліджують 500 проб руди. Імовірність промислового вмісту заліза у кожній пробі дорівнює 0,7. Знайти імовірність того, що кількість проб з промисловим вмістом заліза буде між 300 та 370.

Відповідь: $P_{500}(300 \leq m \leq 370) \approx 0,975$.

30. Прилад складено із трьох незалежно працюючих елементів. Імовірності безвідмовної роботи елементів (за час t) дорівнюють 0,7, 0,8 та 0,9 відповідно. Знайти імовірність того, що за час t будуть працювати безвідмовно а) усі елементи; б) два елементи; в) один елемент; г) жоден із елементів.

Відповідь: а) $P_3(3) = 0,504$; б) $P_3(2) = 0,398$; в) $P_3(1) = 0,092$; г) $P_3(0) = 0,006$.

31. Дві гармати стріляють у ціль. Імовірність влучення для першої гармати дорівнює 0,8, а для другої — 0,9. Знайти імовірність: а) двох влучень; б) не менше одного влучення.

Відповідь: а) $P_2(2) = 0,72$; б) $P_2(0) + P_2(1) = 0,28$.

32. Імовірність появи події в кожному з 10 000 незалежних випробувань дорівнює 0,75. Знайти таке число $\varepsilon > 0$, щоб з імовірністю 0,979 абсолютна величина відхилення частоти події від її імовірності була не більша за ε .

Відповідь: $\varepsilon = 0,01$.

33. Верстат-автомат виготовляє стандартний виріб з імовірністю 0,9. Із продукції верстату складають партію виробів. Скільки виробів повинен виготовити верстат, щоб з імовірністю 0,9973 можна було стверджувати, що відхилення відносної частоти появи нестандартного виробу від імовірності буде менше ніж 0,03? Визначити також можливу кількість нестандартних виробів, що містить ця партія.

Відповідь: $n = 900$, $63 \leq m \leq 117$.

34. 4 елементи обчислювального устрою працюють незалежно. Імовірність відмови першого елемента за час t дорівнює 0,2, другого — 0,25, третього — 0,3, четвертого — 0,4. Знайти імовірність того, що за час t відмовлять: а) 4 елемента; б) 3 елемента; в) 2 елемента; г) 1 елемент; д) жоден елемент; е) не більше двох елементів.

Відповідь: а) $P_4(4) = 0,006$; б) $P_4(3) = 0,065$; в) $P_4(2) = 0,254$; г) $P_4(1) = 0,423$; д) $P_4(0) = 0,252$; е) $P_4(0) + P_4(1) + P_4(2) = 0,929$.

35. Подія A може з'явитись у випробуванні з імовірністю p . Провели n випробувань при однакових умовах незалежним чином. Чому дорівнює імовірність того, що будуть виконані нерівності

$$np - 2\sqrt{npq} \leq m \leq np + 2\sqrt{npq},$$

де m — число випробувань, у яких подія A з'явилася.

Відповідь: $2\Phi(2) = 0,9545$.

36. При умовах вправи 35 знайти імовірність того, що виконується нерівність $|np - m| \leq 4\sqrt{npq}$.

Відповідь: $2\Phi(4) = 0,999936$.

37. Середня кількість обривів нитки на ткацькому верстаті в одну хвилину дорівнює 3. Знайти імовірність того, що за три хвилини буде: а) 5 обривів нитки; б) менше 5 обривів нитки; в) не менше 5 обривів нитки.

Відповідь: а) $P_3(5) = 0,605$; б) $P_3(k < 5) = 0,0548$;

в) $P_3(k \geq 5) = 0,9452$.

38. Імовірність виграшу однієї облігації за весь час займу дорівнює 0,6. Знайти імовірність того, що у власника 5 облігацій: а) виграш випаде на дві облігації; б) виграш випаде більш ніж на дві облігації;

Відповідь: а) $P_5(2) = 0,2304$; б) 0,3072.

39. Середня кількість замовлень таксі, що надходять до диспетчерського пункту щохвилини, дорівнює 3, знайти імовірність того, що за дві хвилини надійде а) 4 замовлення; б) менше за 4 замовлення; в) не менше 4 замовлень.

Відповідь: а) 0,135; б) 0,1525; в) 0,8475.

Розділ 4

ВИПАДКОВІ ВЕЛИЧИНИ

*Для комфорту у будівлі —
випадкові величини,
їх розподілу закони
вимагають пам'ятать:
як знаходять інтеграли,
похідні як треба брать.*

*Головні характеристики
випадкових величин
є основою статистики.
Тому треба без лінця
вивчити їх до кінця.*

4.1. Види випадкових величин та способи їх задання

При дослідженні багатьох проблем виникають такі випадкові події, наслідком яких є поява деякого числа, заздалегідь невідомого. Тому такі числові значення – випадкові.

Прикладом такої події є: кількість очок, що випадає при киданні грального кубика; кількість студентів, які прийдуть на лекцію; кількість цукрового буряка, який чекають одержати з одного гектара.

Випадковою величиною називають таку величину, яка в наслідок випробування може прийняти лише одне числове значення, заздалегідь невідоме і обумовлене випадковими причинами.

Випадкові величини доцільно позначати великими літерами X , Y , Z , а їх можливі значення – відповідними малими літерами з індексами. Наприклад,

$$X: x_1, x_2, \dots, x_n; \quad Z: z_1, z_1, \dots, z_m.$$

Випадкові величини бувають **дискретними** та **неперервними**.

Означення 1. *Дискретною випадковою величиною (ДВВ) називають таку величину, яка може приймати відокремлені ізольовані одне від одного числові значення (їх можна пронумерувати) з відповідними імовірностями.*

Приклад 1. Кількість влучень у мішень при трьох пострілах буде $X : 0, 1, 2, 3$. Отже, X може приймати чотири ізольовані числові значення з різними імовірностями. Тому X – дискретна випадкова величина.

Кількість викликів таксі Y на диспетчерському пункті також буде дискретною випадковою величиною, але при $t \rightarrow \infty$ значення Y також зростають, тобто їх кількість прямує до нескінченності $Y : 0, 1, 2, \dots, n, \dots$.

Означення 2. *Неперервною випадковою величиною (НВВ) називають величину, яка може приймати будь-яке числове значення з деякого скінченного або нескінченного інтервалу (a, b) . Кількість можливих значень такої величини є нескінченна.*

Приклад 2. Величина похибки, яка може бути при вимірюванні відстані; час безвідмовної роботи приладу; зріст людини; розміри деталі, яку виготовляє станок-автомат.

Приклад 3. Розглянемо випадкові величини: кількість очок, X та Y , що можуть з'явитись при киданні правильного грального кубика та неправильного грального кубика. їх можливі значення

$$X : 1, 2, 3, 4, 5, 6; \quad Y : 1, 2, 3, 4, 5, 6$$

однакові.

Імовірність появи будь-якого значення x_k дорівнює $\frac{1}{6}$, однакова для усіх можливих значень X , а імовірності появи можливих значень Y будуть різними. Отже, випадкові величини X та Y не рівні тому, що при $x_k = y_k$ маємо $P(x_k) \neq P(y_k)$, $k = 1, 2, 3, 4, 5, 6$.

Таким чином, для повної характеристики випадкової величини треба вказати не тільки усі її можливі значення, але й закон, за яким знаходять імовірності кожного значення

$$p_k = P(X = x_k) = f(x_k) \quad \text{або} \quad P(X) = f(X).$$

Означення 3. *Законом розподілу випадкової величини називають таке співвідношення, яке встановлює зв'язок між можливими значеннями випадкової величини і відповідними їм імовірностями.*

У випадку дискретної випадкової величини X функціональну залежність можна задавати **таблицно, аналітично** або **графічно**.

У економічних дисциплінах усі ці способи задання ДВВ мають інші назви, тому ознайомимось з ними більш детально у наступному розділі.

У випадку неперервної випадкової величини для її повної характеристики вводять інтегральну та диференціальну функції розподілу.

Означення 4. *Інтегральною функцією розподілу (функцією розподілу) називають імовірність того, що випадкова величина X прийме значення, менше x .*

Функцію розподілу позначають $F(x)$. Таким чином,

$$F(x) = P(X < x).$$

Якщо НВВ X може приймати будь-яке значення з (a, b) , то

$$P(a < X < b) = F(b) - F(a), \quad (1)$$

тобто імовірність прийняття величиною X значень з (a, b) дорівнює приросту функції розподілу.

Формулу (1) часто називають **основною формулою теорії імовірностей**.

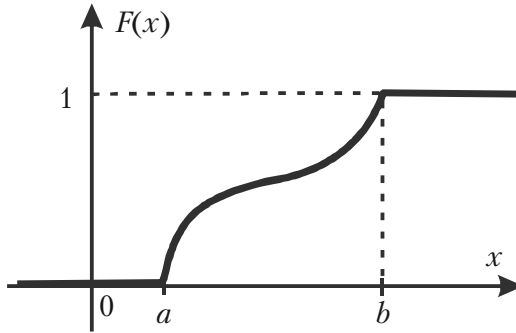
Зауваження. *Неперервна випадкова величина X , що приймає значення у проміжку (a, b) , має незлічену кількість можливих значень, тому набуття X певних значень $X = a$ або $X = b$ буде майже неможливою подією. Це означає, що $P(X = a)$ та $P(X = b)$ будуть нескінченно малими величинами, які у практичних розрахунках можна не враховувати. Тому мають місце рівності*

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

Означення інтегральної функції розподілу та властивості імовірності P дозволяють одержати такі властивості функції розподілу:

- 1) $0 \leq F(x) \leq 1$;
- 2) $F(x)$ – зростаюча функція, тобто $F(x_2) > F(x_1)$, якщо $x_2 > x_1$;
- 3) $F(x) = 0$ при $x \leq a$; $F(x) = 1 = 1$ при $x \geq b$.

Графік функції розподілу $F(x)$ може мати вигляд, зображений, наприклад, на мал. 9.



Мал. 9.

Означення 5. Диференціальною функцією розподілу або щільністю імовірностей неперервної випадкової величини називають похідну першого порядку від її інтегральної функції розподілу і позначають

$$f(x) = F'(x). \quad (2)$$

Назва «щільність імовірностей» випливає з рівності

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x) - P(X < x)}{\Delta x}.$$

Із формули (2) випливає, що функція розподілу $F(x)$ є первісною для диференціальної функції розподілу $f(x)$.

Теорема 1. Імовірність того, що неперервна випадкова величина X прийме значення з інтервалу (a, b) , можна знайти за формулою

$$P(a < X < b) = \int_a^b f(x) dx. \quad (3)$$

Доведення. Інтегральна функція розподілу $F(x)$ — первісна для $f(x)$, тому згідно з формулою Ньютона-Лейбніца маємо

$$\int_a^b f(x) dx = F(b) - F(a). \quad (4)$$

Праві частини рівностей (1) та (4) рівні, тому і ліві їх частини рівні, тобто має місце рівність (3), яку й треба було довести.

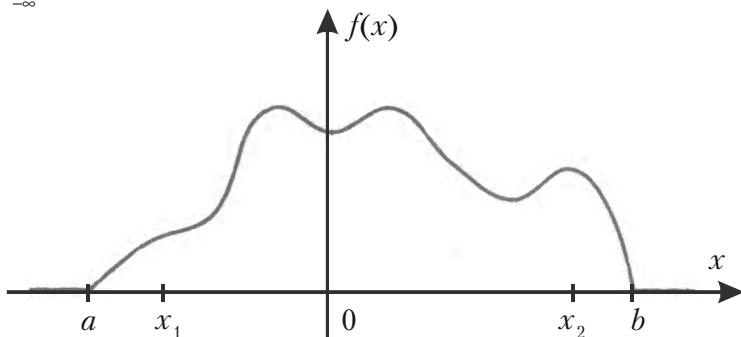
Розділ 4. Випадкові величини

Наслідок. Якщо диференціальна функція розподілу (щільність імовірності) $f(x)$ відома, то інтегральну функцію розподілу $F(x)$ можна знайти за формулою

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (5)$$

Диференціальна функція розподілу НВВ $X \in (a, b)$ має такі властивості:

- 1) $f(x) \geq 0$ тому, що вона є похідною зростаючої функції $F(x)$;
- 2) $f(x) = 0$ при $x < a$ та $x \geq b$ тому, що є похідною $F(x) = 0$ при $x < a$ та $F(x) = 1$ при $x \geq b$;
- 3) $\int_{-\infty}^{+\infty} f(x) dx = 1$ тому, що подія $\{-\infty < X < \infty\}$ – достовірна.



Мал. 10

Графік щільності імовірності $f(x)$ називають **кривою розподілу**. Він може мати вигляд, зображений, наприклад, на мал. 10.

Приклад 4. Випадкова величина має щільність імовірностей

$$f(x) = \frac{a}{1+x^2}. \text{ Визначити параметр } a \text{ та функцію розподілу.}$$

Розв'язання. Параметр a знайдено використовуючи властивість 3 диференціальної функції розподілу

$$\begin{aligned}
 1 &= \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{adx}{1+x^2} = a \cdot \operatorname{arctg} x \Big|_{-\infty}^{\infty} = \\
 &= a \left[\lim_{b \rightarrow \infty} \operatorname{arctg} b - \lim_{c \rightarrow -\infty} \operatorname{arctg} c \right] = a \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right] = a \cdot \pi.
 \end{aligned}$$

Отже, одержали

$$1 = a \cdot \pi \Rightarrow a = \frac{1}{\pi}.$$

Функцію розподілу знайдемо за формулою (5)

$$\begin{aligned}
 F(x) &= \int_{-\infty}^x f(x) dx = \int_{-\infty}^x \frac{\frac{1}{\pi} dx}{1+x^2} = \frac{1}{\pi} \operatorname{arctg} x \Big|_{-\infty}^x = \\
 &= \frac{1}{\pi} \left(\operatorname{arctg} x - \lim_{c \rightarrow -\infty} \operatorname{arctg} c \right) = \frac{1}{\pi} \operatorname{arctg} x - \frac{1}{\pi} \cdot \left(-\frac{\pi}{2} \right) \Rightarrow \\
 &F(x) = \frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}.
 \end{aligned}$$

Приклад 5. Випадкова величина X задана функцією розподілу

$$F(x) = x^2 - 4x + 4. \quad (6)$$

Визначити область значень випадкової величини X та імовірність того, що $X \geq 2,3$.

Розв'язання. Згідно властивостям функції розподілу маємо

$$0 \leq F(x) \leq 1,$$

тому повинні виконуватись умови

$$0 \leq x^2 - 4x + 4 \leq 1.$$

Якщо область значень випадкової величини $[a, b]$, то $F(a) = 0$ та $F(b) = 1$. Підставимо в (6) замість x a та b , тоді одержимо

$$a^2 - 4a + 4 = 0 \Rightarrow (a - 2)^2 = 0 \Rightarrow a = 2;$$

$$b^2 - 4b + 4 = 1 \Rightarrow b^2 - 4b + 3 = 0 \Rightarrow b_1 = 3, b_2 = 1.$$

Але в проміжку $[a, b]$, $b > a$, тому $b = 3$. Отже, областю значень НВВ X буде $[2, 3]$.

Тепер знайдемо імовірність $P(X \geq 2,3)$. Подія $X < 2,3$ буде протилежною, тому

$$P(X \geq 2,3) = 1 - P(X < 2,3) = 1 - F(2,3). \quad (7)$$

З рівності (6) одержуємо

$$F(2,3) = (2,3)^2 - 4 \cdot 2,3 + 4 = 5,29 - 9,2 + 4 = 0,09.$$

Тепер за формулою (7) знаходимо

$$P(X \geq 2,3) = 1 - 0,09 = 0,91.$$

4.2. Закони розподілу та числові характеристики дискретних випадкових величин

4.2.1. Способи задання та закони розподілу

Нехай випадкова дискретна величина X приймає значення x_1, x_2, \dots, x_n з відповідними імовірностями p_1, p_2, \dots, p_n .

Задати закон розподілу такої випадкової величини – це задати рівність $p_k = P(X = x_k)$, яку можна розглядати як функцію.

Тому закон розподілу X можна задати **аналітично, таблично, графічно**. Функція розподілу для дискретної випадкової величини має вигляд

$$F(x) = P(X < x) = \sum_{x_i < x} p(x_i).$$

Найчастіше використовують табличний спосіб задання ДВВ, який називають **рядом розподілу** і зображують у вигляді

X	x_1	x_2	\dots	x_n
$P(X)$	p_1	p_2	\dots	p_n

У першому рядку записані усі можливі значення X , а у другому рядку – відповідні імовірності, які мають властивість

$$\sum_{k=1}^n p_k = 1.$$

Приклад 1. Умовами лотереї передбачено: один виграш – 100 гривень, два – 50 гривень, вісім – 10 гривень, дев'ятнадцять – 1 гривня. Знайти закон розподілу суми виграшу власником одного лотерейного білету, якщо продано 1000 білетів.

Розв'язання. Будемо шукати закон розподілу суми виграшу X у вигляді ряду розподілу. Тоді

X	100	50	10	1	0
$P(X)$	0,001	0,002	0,008	0,019	0,97

де $p(0) = 1 - (0,001 + 0,002 + 0,008 + 0,019) = 1 - 0,03 = 0,97$.

Зауваження 1. Якщо випадкова дискретна величина може приймати нескінчену кількість значень, то її ряд розподілу (таблиця) буде мати нескінчену кількість елементів у кожному рядку, причому ряд

$\sum_{k=1}^{\infty} p_k$ повинен бути збіжним, а його сума повинна дорівнювати одиниці.

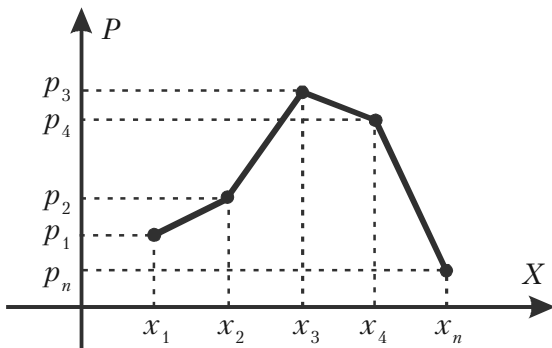
Графічний спосіб. Візьмемо прямокутну систему координат. На осі абсцис будемо відкладати можливі значення ДВВ, а на осі ординат – відповідні значення імовірності. Одержимо точки з координатами (x_1, p_1) , (x_2, p_2) , \dots , (x_n, p_n) .

Поєднавши ці точки прямими, одержимо графік (див. мал. 11) у вигляді **многокутника розподілу випадкової дискретної величини**.

Значення ДВВ, імовірність якого найбільша, називають **модой**.

На малюнку 11 мода – x_3 .

Аналитичний спосіб задання випадкової дискретної величини базується на заданні певної функції, за якою можна знайти імовірність p відповідного значення x_k , тобто $p_k = f(x_k)$, $k = 1, 2, \dots, n$.



Мал. 11.

Вкажемо деякі найважливіші закони розподілу ДВВ та задачі, в яких вони зустрічаються.

1. Біноміальний закон розподілу. Цей закон має вигляд

$$P(X = m) = C_n^m p^m (1 - p)^{n-m}, \quad m = 0, 1, 2, \dots, n$$

і використовується у схемі Бернуллі, тобто у випадку n незалежних повторних випробувань, в кожному з яких деяка подія з'являється з імовірністю p .

2. Закон розподілу Пуассона. ДВВ X приймає злічену множину значень ($m = 0, 1, 2, \dots$) з імовірностями

$$P(X = m) = \frac{a^m}{m!} e^{-a} \quad (a > 0).$$

Цей розподіл використовують у задачах статистичного контролю якості, в теорії надійності, теорії масового обслуговування, для обчислення: кількості вимог на виплату страхових сум за рік, кількості дефектів однакових виробів.

Якщо у схемі незалежних повторних випробувань n досить велике, а p або $1 - p$ прямує до нуля, то біноміальний розподіл апроксимує розподіл Пуассона, параметр якого $a = np$, причому при $p \leq 0,1$ або $p \geq 0,9$ ця апроксимація дає добрі результати незалежно від величини n .

Зауваження 2. Якщо у формулі Пуассона покласти $a = \lambda t$, де λ – інтенсивність течії випадкових подій в одиницю часу, то формула прийме вигляд

$$P_t(X = m) = \frac{(\lambda t^m)}{m!} e^{-\lambda t} \quad (m = 0, 1, 2, \dots).$$

3. Геометричний розподіл. Цей розподіл має вигляд

$$P(X = m) = pq^{m-1},$$

де $p = P(A)$ – імовірність появи події A в кожному випробуванні, $q = 1 - p$, X – кількість випробувань до появи події A в серії незалежних повторних випробувань.

Ряд імовірностей цього розподілу буде нескінченно спадною геометричною прогресією із знаменником q , сума якої дорівнює одиниці.

Геометричний розподіл застосовують у різноманітних задачах статистичного контролю якості виробів, в теорії надійності та у страхових розрахунках.

4. Гіпергеометричний розподіл. Цей розподіл має вигляд

$$P(X = m) = \frac{C_k^m \cdot C_{N-k}^{n-m}}{C_N^n}, \quad m = 0, 1, 2, \dots, n, \quad k \geq n.$$

Він вказує імовірність появи m елементів з певною властивістю серед n елементів, взятих із сукупності N елементів, яка містить k елементів саме такої властивості.

Цей розподіл використовують у багатьох задачах статистичного контролю якості.

Зауваження 3. Якщо об'єм вибірки n малий у порівнянні з об'ємом

N сукупності, тобто $\frac{n}{N} \leq 0,1$; $\frac{n}{k} \leq 0,1$, то імовірності у гіпергеометричному розподілі будуть близькими до відповідних імовірностей біноміального розподілу з $p = \frac{k}{N}$.

У статистиці це означає, що розрахунки імовірностей для безповторної вибірки будуть мало відрізнятися від розрахунків імовірностей для повторної вибірки.

5. Поліноміальний розподіл. Цей розподіл має вигляд

$$P_n = (X_1 = m_1; X_2 = m_2; \dots; X_s = m_s) = \frac{n!}{m_1! m_2! \dots m_s!} \cdot p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_s^{m_s}$$

Він застосовується тоді, коли внаслідок кожного із здійснених повторних незалежних випробувань може з'явитися s різних подій

A_i з імовірністю p_i , причому $\sum_{i=1}^s p_i = 1$.

4.2.2. Числові характеристики

Закони розподілу ДВВ повністю характеризують випадкові величини і дозволяють розв'язувати усі пов'язані з ними задачі.

Але в практичній діяльності не завжди вдається одержати закон розподілу, або закон надто складний для практичних розрахунків. Тому з'явилася потреба характеризувати ДВВ за допомогою числових характеристик, які характеризують особливості випадкових величин достатньо.

Найчастіше використовують три числових характеристики: **математичне сподівання**, **дисперсію** та **середнє квадратичне відхилення від математичного сподівання**.

Ознайомимось із цими числовими характеристиками та їх властивостями.

Математичне сподівання та його основні властивості

Означення 1. *Математичним сподіванням дискретної випадкової величини X називають число, яке дорівнює сумі добутків усіх можливих значень X на відповідні їм імовірності.*

Математичне сподівання ДВВ X позначають $M(X)$ або m_x , тобто

$$M(X) = \sum_{k=1}^n x_k \cdot p_k \quad (1)$$

Якщо X приймає нескінчену кількість значень, то

$$M(X) = \sum_{k=1}^{\infty} x_k \cdot p_k.$$

Математичне сподівання ДВВ X характеризує середнє значення випадкової величини X із врахуванням імовірностей його можливих значень. У практичній діяльності під математичним сподіванням розуміють центр розподілу випадкової величини.

Основні властивості математичного сподівання

1. Математичне сподівання постійної величини дорівнює самій постійній

$$M(C) = C.$$

2. Постійний множник можна виносити за знак математичного сподівання

$$M(CX) = C \cdot M(X).$$

Властивості 1 і 2 впливають безпосередньо з Означення 1.

3. Математичне сподівання добутку декількох взаємно незалежних дискретних випадкових величин дорівнює добутку їх математичних сподівань, тобто

$$M(X_1 \cdot X_2 \cdot \dots \cdot X_n) = M(X_1) \cdot M(X_2) \cdot \dots \cdot M(X_n).$$

Доведення. Якщо дві величини X та Y розподілені за законами

X	x_1	x_2
P	p_1	p_2

Y	y_1	y_2
G	g_1	g_2

(для спрощення викладок взято лише по 2 можливих значення), тоді закон розподілу добутку $X \cdot Y$ буде

$X \cdot Y$	$x_1 y_1$	$x_2 y_1$	$x_1 y_2$	$x_2 y_2$
$P \cdot G$	$p_1 g_1$	$p_2 g_1$	$p_1 g_2$	$p_2 g_2$

За формулою (1) одержимо математичне сподівання

$$M(X \cdot Y) = y_1 g_1(x_1 p_1 + x_2 p_2) + y_2 g_2(x_1 p_1 + x_2 p_2) = \\ = (x_1 p_1 + x_2 p_2) \cdot (y_1 g_1 + y_2 g_2) = M(X) \cdot M(Y).$$

У випадку трьох випадкових величин маємо

$$M(X \cdot Y \cdot Z) = M((X \cdot Y) \cdot Z) = M(X \cdot Y) \cdot M(Z) = M(X) \cdot M(Y) \cdot M(Z).$$

Методом математичної індукції тепер не важко завершити доведення.

Аналогічно, але дещо складніше, можна довести наступну властивість.

4) Математичне сподівання суми випадкових величин дорівнює сумі їх математичних сподівань, тобто

$$M(X_1 + X_2 + \dots + X_n) = M(X_1) + M(X_2) + \dots + M(X_n).$$

Приклад 2. Незалежні випадкові величини X та Y розподілені так

X	5	2	4
P	0,6	0,1	0,3

Y	8	10
P	0,8	0,2

Знайти математичне сподівання випадкової величини $X \cdot Y$.

Розв'язання. Спочатку знайдемо математичні сподівання кожної з цих величин. За формулою (1) маємо.

$$M(X) = 5 \cdot 0,6 + 2 \cdot 0,1 + 4 \cdot 0,3 = 4,4, \quad M(Y) = 8 \cdot 0,8 + 10 \cdot 0,2 = 8,4.$$

Випадкові величини X і Y незалежні, тому згідно властивості 3 математичного сподівання одержимо

$$M(X \cdot Y) = M(X) \cdot M(Y) = 4,4 \cdot 8,4 = 36,96.$$

Приклад 3. Знайти математичне сподівання суми числа очок, які можуть з'явитися при киданні двох гральних кубиків.

Розв'язання. Позначимо кількість очок, які можуть з'явитись на першому кубіку X , а на другому – Y . Можливі значення цих величин 1, 2, 3, 4, 5, 6 однакові, імовірність кожного з цих значень дорівнює $\frac{1}{6}$. Тому

нює $\frac{1}{6}$. Тому

$$M(X) = M(Y) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

Згідно властивості 4 математичного сподівання, одержимо

$$M(X + Y) = M(X) + M(Y) = \frac{7}{2} + \frac{7}{2} = 7.$$

Отже, математичне сподівання суми числа очок, що можуть з'явитись при киданні двох гральних кубиків, дорівнює 7.

Дисперсія та її властивості

Математичне сподівання характеризує *центр розподілу* дискретної випадкової величини. Але цієї характеристики недостатньо, бо можливе значне відхилення можливих значень від центру розподілу. Для характеристики розсіювання можливих значень X відносно центру розподілу введемо нову числову характеристику.

Означення 2. *Дисперсією дискретної випадкової величини X називають число, яке дорівнює математичному сподіванню квадрата відхилення ДВВ X від її математичного сподівання.*

Дисперсію величини X позначають $D(X)$ або D_X . Це означення математично виглядає так

$$D(X) = M\left(\left(X - M(X)\right)^2\right). \quad (2)$$

Основні властивості $D(X)$

1. Дисперсія будь-якої ДВВ X невід'ємна

$$D(X) \geq 0.$$

Дійсно, $\left(X - M(X)\right)^2$ невід'ємна, тому згідно означення математичного сподівання та властивостей імовірностей p_k , $k = 1, 2, \dots, n$, $D(X)$ також невід'ємна.

2. Дисперсія постійної величини C дорівнює нулеві

$$D(C) = 0.$$

Дійсно, якщо $X = C$, то $M(C) = C$, тому $C - M(C) = 0$.

3. Постійний множник C можна виносити за знак дисперсії, при цьому постійний множник треба піднести у квадрат

$$D(CX) = C^2 D(X). \quad (3)$$

Дійсно,

$$CX - M(CX) = C(X - M(X)),$$

тому

$$(CX - M(CX))^2 = C^2 (X - M(X))^2.$$

Постійний множник C^2 можна виносити за знак математичного сподівання, тому з формули (2) випливає потрібна рівність (3).

4. Дисперсія ДВВ X дорівнює різниці між математичним сподіванням квадрата випадкової величини X та квадрата її математичного сподівання

$$D(X) = M(X^2) - (M(X))^2. \quad (4)$$

Дійсно,

$$\begin{aligned} D(X) &= M\left(\left(X - M(X)\right)^2\right) = M\left(X^2 - 2XM(X) + M^2(X)\right) = \\ &= M(X^2) - 2M^2(X) + M^2(X) = M(X^2) - M^2(X). \end{aligned}$$

Зауваження 4. Формула (2) визначає дисперсію випадкової величини X , а за формулою (4) її доцільно знаходити.

5. Дисперсія алгебраїчної суми ДВВ X та Y дорівнює сумі їх дисперсій

$$D(X \pm Y) = D(X) + D(Y).$$

Доведення. Спочатку доведемо цю властивість для $X + Y$. Згідно з формулою (4) маємо

$$\begin{aligned} D(X + Y) &= M\left(\left(X + Y\right)^2\right) - M^2(X + Y) = M\left(X^2 + 2XY + Y^2\right) - \\ &- \left(M(X) + M(Y)\right)^2 = M(X^2) + 2M(X)M(Y) + M(Y^2) - \end{aligned}$$

$$-M^2(X) - 2M(X)M(Y) - M^2(Y) = (M(X^2) - M^2(X)) + (M(Y^2) - M^2(Y)) = D(X) + D(Y).$$

Тепер розглянемо дисперсію різниці X та Y

$$D(X - Y) = D(X) + (-1)^2 \cdot D(Y) = D(X) + D(Y).$$

Зауваження 5. П'ята властивість дисперсії має місце для алгебраїчної суми не лише двох, але й скінченного числа дискретних випадкових величин.

Приклад 4. Знайти дисперсію випадкової величини X , що задана законом

X	-5	0	4	5
P	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

Розв'язання. Будемо шукати $D(X)$ з використанням формули (4). Математичним сподіванням X згідно з формулою (1) буде

$$M(X) = (-5) \cdot \frac{1}{8} + 0 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 5 \cdot \frac{1}{8} = 1 \Rightarrow M^2(X) = 1.$$

Щоб знайти математичне сподівання X^2 , тобто $M(X^2)$, запишемо закон розподілу X^2 у вигляді таблиці

X^2	25	0	16	25
P	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

Відмітимо, що усі значення X^2 отримані шляхом піднесення до квадрату відповідних значень X . Елементи другого рядка – імовірності цих значень – не змінюються.

За формулою (1) знаходимо

$$M(X^2) = 25 \cdot \frac{1}{8} + 0 \cdot \frac{1}{2} + 16 \cdot \frac{1}{4} + 25 \cdot \frac{1}{8} = \frac{82}{8}.$$

Згідно з формулою (4) тепер одержуємо

$$D(X) = \frac{82}{8} - 1 = \frac{74}{8} = 9,25.$$

Середнє квадратичне відхилення дискретної випадкової величини

У більшості випадків випадкова величина X має розмірність, наприклад, метр, міліметр, грам, тому її дисперсія $D(X)$ буде вимірюватись у квадратних одиницях цієї розмірності.

У практичній діяльності доцільно знати величину розсіювання випадкової величини в розмірності цієї величини. Для цього використовують середнє квадратичне відхилення, яке дорівнює квадратному кореню з дисперсії і позначається

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

Поняття моментів розподілу

Означення 3. Початковим моментом порядку k випадкової величини X називають математичне сподівання величини X^k і позначають

$$\nu_k = M(X^k), \quad k = 1, 2, \dots, n.$$

Центральним моментом порядку k випадкової величини X називають математичне сподівання величини $(X - M(X))^k$ і позначають

$$\mu_k = M\left(\left(X - M(X)\right)^k\right), \quad k = 1, 2, \dots, n.$$

Відмітимо, що

$$\nu_1 = M(X), \quad \nu_2 = M(X^2),$$

тому

$$D(X) = \nu_2 - \nu_1^2; \quad \mu_1 = M(X - M(x)) = 0;$$

$$\mu_2 = M\left(\left(X - M(X)\right)^2\right) = D(X).$$

Початкові та центральні моменти порядку $k \geq 2$ дозволяють краще враховувати вплив на математичне сподівання (центр розподілу випадкової величини X) тих можливих значень X , які великі та мають малу імовірність.

Приклад 5. Дискретна випадкова величина задана законом

X	1	2	5	100
p	0,6	0,2	0,19	0,01

Математичним сподіванням X буде

$$M(X) = 1 \cdot 0,6 + 2 \cdot 0,2 + 5 \cdot 0,19 + 100 \cdot 0,01 = 2,95.$$

Законом розподілу X^2 буде

X^2	1	4	25	10000
p	0,6	0,2	0,19	0,01

Тому

$$M(X^2) = 1 \cdot 0,6 + 4 \cdot 0,2 + 25 \cdot 0,19 + 10000 \cdot 0,01 = 106,15.$$

Отже, $M(X^2)$ значно більше $M(X)$, а це означає, що роль значення $X = 100$ зростає.

Зауваження 6. Доцільно знати числові характеристики основних законів розподілу дискретних випадкових величин, які можна привести у вигляді наступної таблиці.

Розділ 4. Випадкові величини

Закон розподілу X та його математичний запис	Математичне сподівання $M(X)$	Дисперсія $D(X)$	Середнє квадратичне відхилення $\sigma(X)$
1. Біноміальний $P(X = m) = C_n^m p^m (1-p)^{n-m}$, $m = 0, 1, 2, \dots, n$.	np	npq	\sqrt{npq}
2. Пуассона $P(X = m) = \frac{a^m}{m!} e^{-a}$, $a > 0$.	a	a	\sqrt{a}
3. Геометричний $P(X = m) = pq^{m-1}$, $m = 1, 2, \dots$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{\sqrt{q}}{p}$
4. Гіпергеометричний $P(X = m) = \frac{C_k^m \cdot C_{N-k}^{n-m}}{C_N^n}$, $m = 0, 1, 2, \dots, n, k \geq n$.	$\frac{kn}{N}$	$\frac{nk(N-k)}{N^2} \times$ $\times \frac{(N-n)}{(N-1)}$	
5. Поліноміальний $P_n(X_1 = m_1, \dots, X_s = m_s) = \frac{n!}{m_1! m_2! \dots m_s!} \times p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_s^{m_s}$.	$M(X_i) = np_i$ $i = 1, \dots, s$	$D(X_i) = np_i q_i$ $q_i = 1 - p_i$ $i = 1, \dots, s$	$\sigma(X_i) = \sqrt{np_i q_i}$ $q_i = 1 - p_i$ $i = 1, \dots, s$

4.3. Числові характеристики законів розподілу неперервних випадкових величин

4.3.1. Числові характеристики неперервних випадкових величин

У випадку неперервних випадкових величин (НВВ) математичне сподівання, дисперсія та середнє квадратичне відхилення мають такий же смисл та властивості, як і для дискретних випадкових величин, але обчислюють їх за іншими формулами.

Нехай можливі значення неперервної випадкової величини X заповнюють відрізок $[a, b]$. Поділимо $[a, b]$ на n частин довжиною

$$\Delta x = \frac{b-a}{n},$$

У кожній частині візьмемо точку ξ_k , $k = 1, 2, \dots, n$.

Тоді щільність імовірності $f(x)$ в точці ξ_k буде $f(\xi_k)$ – імовірність того, що X прийме значення Одержимо розподіл НВВ X вигляду

X	ξ_1	ξ_2	\dots	ξ_n
P	$f(\xi_1)$	$f(\xi_2)$	\dots	$f(\xi_n)$

Сума

$$\sum_{k=1}^n \xi_k f(\xi_k)$$

характеризує математичне сподівання X тим точніше, чим менше буде Δx . Ця сума буде дорівнювати математичному сподіванню $M(X)$ неперервної величини X , якщо перейти до границі при $\Delta x \rightarrow 0$. Згідно з означенням визначеного інтеграла маємо

$$M(X) = \lim_{\Delta x \rightarrow 0} \sum_{k=1}^n \xi_k f(\xi_k) = \int_a^b x f(x) dx.$$

Таким чином, доведена

Теорема 2. *Якщо неперервна випадкова величина приймає значення на відрізку $[a, b]$ та має щільність імовірності $f(x)$, то її математичне сподівання знаходиться за формулою*

$$M(X) = \int_a^b xf(x)dx. \quad (1)$$

Аналогічно доводиться

Теорема 3. *Якщо $f(x)$ є щільність імовірності X , неперервна випадкова величина Y є функцією випадкової величини X , тобто $Y = \varphi(X)$, тоді математичне сподівання Y знаходиться за формулою*

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

Зауваження 1. *Якщо можливі значення X належать відрізку $[a, b]$, то центр розподілу $M(X)$ величини X знаходиться на цьому проміжку тому, що із нерівностей*

$$\int_a^b af(x)dx < \int_a^b xf(x)dx < \int_a^b bf(x)dx$$

та умови нормування $\int_a^b f(x)dx = 1$ випливають співвідношення

$$a < M(X) = \int_a^b xf(x)dx < b.$$

Якщо щільність імовірності $f(x)$ – парна функція, тобто $f(-x) = f(x)$, то центр розподілу X співпадає з початком $M(X) = 0$. Якщо графік функції $f(x)$ симетричний відносно прямої $x = a$, то $M(X) = a$.

Як і у випадку дискретних випадкових величин, дисперсію неперервних випадкових величин X визначають так

$$D(X) = M\left(\left(X - M(X)\right)^2\right) = \int_{-\infty}^{\infty} (x - M(X))^2 \cdot f(x) dx, \quad (2)$$

а обчислюють за формулою

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - M^2(X). \quad (3)$$

Якщо можливі значення X належать лише скінченному проміжку (a, b) , то рівності (2) та (3) приймають вигляд

$$D(X) = \int_s^b (x - M(X))^2 f(x) dx, \quad (4)$$

$$D(X) = \int_a^b x^2 f(x) dx - M^2(X). \quad (5)$$

Середнє квадратичне відхилення неперервної випадкової величини визначають та обчислюють так

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

Приклад 1. Знайти числові характеристики випадкової величини X , яка задана функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x^2}{25}, & 0 < x < 5 \\ 1, & x \geq 5. \end{cases}$$

Розв'язання. Спочатку знайдемо диференціальну функцію розподілу, тобто щільність імовірності $f(x) = F'(x)$

$$f(x) = \begin{cases} \frac{2}{25}x, & 0 \leq x \leq 5; \\ 0, & x \notin [0, 5]. \end{cases}$$

Тепер за формулою (1) знайдемо математичне сподівання

$$M(X) = \int_0^5 x \cdot \frac{2x}{25} dx = \frac{2}{25} \cdot \frac{x^3}{3} \Big|_0^5 = \frac{10}{3}.$$

Дисперсію знайдемо за формулою (4)

$$D(X) = \int_0^5 x^2 \cdot \frac{2x}{25} dx - \left(\frac{10}{3}\right)^2 = \frac{2}{25} \cdot \frac{x^4}{4} \Big|_0^5 - \frac{100}{9} = \frac{25}{18}.$$

Середнє квадратичне відхилення одержимо за формулою (5)

$$\sigma(X) = \sqrt{\frac{25}{18}} = \frac{5}{3\sqrt{2}} = \frac{5\sqrt{2}}{6} \approx 1,17.$$

4.3.2. Закони розподілу НВВ та їх числові характеристики

Основні закони розподілу неперервних випадкових величин розділяють за виглядом їх диференціальних функцій розподілу (щільності імовірностей) $f(x)$.

Найчастіше використовують наступні закони розподілу.

Рівномірний розподіл

Означення 1. Величина X розподілена **рівномірно** на проміжку (a, b) , якщо усі її можливі значення належать цьому проміжку і щільність її імовірностей на цьому проміжку постійна, тобто

$$f(x) = \begin{cases} C = \frac{1}{b-a}, & \text{при } x \in (a, b), \\ 0, & \text{при } x \notin (a, b). \end{cases}$$

Величина постійної $C = \frac{1}{b-a}$ визначається **умовою нормування**

$$P(a < X < b) = C(b-a) = 1.$$

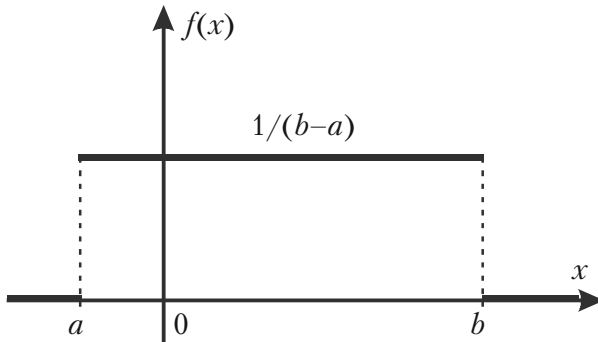
Якщо X рівномірно розподілена на проміжку (a, b) , то імовірність належності X будь-якому інтервалу $(x_1, x_2) \in (a, b)$ пропорційна довжині цього інтервалу

$$P(x_1 < X < x_2) = C(x_2 - x_1) = \frac{x_2 - x_1}{b - a}.$$

Іншими словами, імовірність влучення X в інтервал (x_1, x_2) дорівнює відношенню довжини цього інтервалу до довжини усього проміжку (a, b) .

Цей розподіл задовольняють, наприклад, похибки округлення різноманітних розрахунків.

Графік рівномірного розподілу НВВ X зображено на мал. 12.



Мал. 12.

Числовими характеристиками НВВ X , що розподілена за рівномірним законом, будуть

$$\begin{aligned} M(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_{-\infty}^a x \cdot f(x) dx + \int_a^b x \cdot f(x) dx + \int_b^{\infty} x \cdot f(x) dx = \\ &= \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}; \end{aligned}$$

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

$$= \int_a^b \frac{(x - M(X))^2}{b-a} dx = \frac{\left(x - \frac{b+a}{2}\right)^3}{3(b-a)} \Big|_a^b = \frac{(b-a)^2}{12};$$

$$\sigma(X) = \frac{(b-a)\sqrt{3}}{6}.$$

Показниковий розподіл

Означення 2. Випадкову величину X називають **розподіленою за показниковим законом**, якщо щільність її імовірностей має вигляд

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{при } x \geq 0; \\ 0, & \text{при } x < 0, \end{cases}$$

де $\lambda > 0$ – параметр.

Показниковому розподілу задовольняють: час телефонної розмови, час ремонту техніки, час безвідмовної роботи комп'ютера.

Числовими характеристиками показникового розподілу будуть

$$M(X) = \frac{1}{\lambda}; \quad D(X) = \frac{1}{\lambda^2}; \quad \sigma(X) = \frac{1}{\lambda}. \quad (6)$$

Отже, якщо НВВ X розподілена за показниковим законом, то вона має математичне сподівання та середнє квадратичне відхилення рівні.

Приклад 2. Знайти числові характеристики випадкової величини, розподіленої за законом

$$f(x) = \begin{cases} 4e^{-4x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Розв'язання. У даному випадку випадкова величина X розподілена за показниковим законом із параметром $\lambda = 4$. Згідно з формулами (6) маємо

$$M(X) = \sigma(X) = \frac{1}{4} = 0,25; \quad D(X) = \frac{1}{4^2} = \frac{1}{16} = 0,0625.$$

Зауваження 2. Якщо випадкова величина X розподілена за показниковим законом, то її функція розподілу (інтегральна функція розподілу) має вигляд $F(x) = 1 - e^{-\lambda x}$. Тому основна формула теорії імовірностей набуде вигляду

$$P(a < X < b) = e^{-a\lambda} - e^{-b\lambda}. \quad (7)$$

Приклад 3. Величина X розподілена за законом

$$f(x) = \begin{cases} 3e^{-3x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Знайти імовірність того, що X потрапить в інтервал $(0,4; 1)$.

Розв'язання. Випадкова величина X розподілена за показниковим законом із параметром $\lambda = 3$. Використовуючи формулу (7), отримаємо

$$P(0,4 < X < 1) = e^{-0,4 \cdot 3} - e^{-1 \cdot 3} = e^{-1,2} - e^{-3}.$$

Відповідні значення e^x можна взяти з таблиці значень цієї функції.

Нормальний розподіл

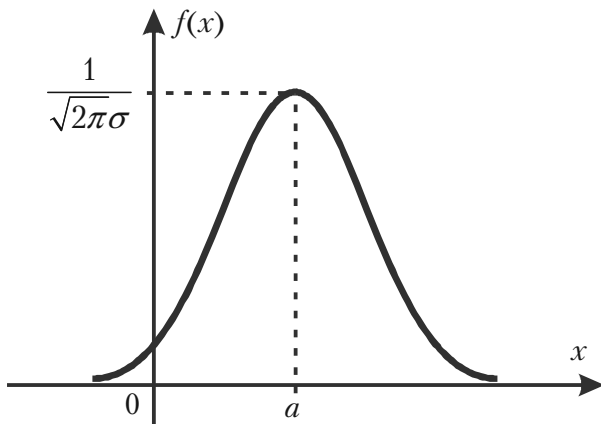
Означення 3. Випадкову величину X називають **розподіленою нормально**, якщо щільність її імовірностей має вигляд

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\alpha)^2}{2\sigma^2}},$$

де α та σ – параметри розподілу.

Графік цієї функції $f(x)$ називають **нормальною кривою** або **кривою Гаусса**.

Повне дослідження цієї функції методами диференціального числення дозволяє побудувати графік нормальної кривої, який зображено на мал. 13.



Мал. 13.

При $a = 0$ та $\sigma = 1$ нормальну криву називають **нормованою**, її рівняння буде

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тобто це є функція Лапласа, яка табульована.

Заміна змінної $Z = \frac{X - a}{\sigma}$, використання інтеграла Пуассона

$$\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}$$

та формул (1), (2) та (5) дозволяють одержати числові характеристики нормально розподіленої НВВ X у вигляді

$$M(X) = a; \quad D(X) = \sigma^2; \quad \sigma(X) = \sigma.$$

Отже, математичне сподівання нормального розподілу дорівнює параметру a цього розподілу, а середнє квадратичне відхилення дорівнює параметру σ .

Зауваження 3. Якщо випадкова величина X розподілена за нормальним законом з параметрами a та σ , то випадкова величина

$Z = \frac{X-a}{\sigma}$ буде розподілена за нормованим нормальним законом і

$$M(Z) = 0, \quad \sigma(Z) = 1.$$

Інтегральною функцією нормального закону розподілу буде

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(z-a)^2}{2\sigma^2}} dz,$$

а для нормованого нормального закону

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz. \quad (8)$$

Імовірність влучення в інтервал (c, d) нормально розподіленої випадкової величини X знаходять за формулою

$$P(c < X < d) = \Phi\left(\frac{d-a}{\sigma}\right) - \Phi\left(\frac{c-a}{\sigma}\right), \quad (9)$$

де функція Лапласа $\Phi(x)$ має вигляд (8).

Приклад 4. Випадкова величина X розподілена за нормальним законом, її математичне сподівання дорівнює 30, а серед квадратичне відхилення — 10. Знайти імовірність того, що X матиме значення з інтервалу $(10, 50)$.

Розв'язання. Згідно умови $a = 30$, $\sigma = 10$, тому за формулою (9) одержимо

$$\begin{aligned} P(10 < X < 50) &= \Phi\left(\frac{50-30}{10}\right) - \Phi\left(\frac{10-30}{10}\right) = 2\Phi(2) = \\ &= 2 \cdot 0,4772 = 0,9544. \end{aligned}$$

Тут використано властивості непарності інтегральної функції Лапласа $\Phi(-x) = -\Phi(x)$ та значення $\Phi(2)$ з таблиці значень цієї функції.

Приклад 5. Зріст студентів розподілено за нормальним законом. Математичне сподівання зросту студентів дорівнює 175 см., а се-

Розділ 4. Випадкові величини

редне квадратичне відхилення — 6 см. Визначити імовірність того, що хоча б один із п'яти викликаних навмання студентів буде мати зріст від 170 до 180 см.

Розв'язання. Зріст студента X — випадкова величина, яка за умовою задачі розподілена нормально з математичним сподіванням $M(X) = 175$ см. та середнім квадратичним відхиленням $\sigma = 6$ см.

Позначимо події: A — із 5 викликаних студентів зріст хоча б одного належить проміжку (170, 180); \bar{A} — зріст усіх 5 викликаних студентів не належать проміжку (170, 180).

Величина X розподілена нормально, тому за формулою (9) знайдемо імовірність того, що зріст одного викликаного студента належить проміжку (170, 180)

$$\begin{aligned} P(170 < X < 180) &= \Phi\left(\frac{180-175}{6}\right) - \Phi\left(\frac{170-175}{6}\right) = \Phi\left(\frac{5}{6}\right) - \\ &- \Phi\left(-\frac{5}{6}\right) = 2\Phi\left(\frac{5}{6}\right) = 2\Phi(0,833). \end{aligned}$$

Використовуючи табличне значення інтегральної функції Лапласа, отримаємо

$$P(170 < X < 180) = 2 \cdot 0,2967 = 0,5934.$$

Імовірність того, що зріст одного викликаного студента не належить проміжку (170, 180) буде

$$p = 1 - P(170 < X < 180) = 1 - 0,5934 = 0,4066.$$

Застосовуючи теорему множення імовірностей незалежних подій, знайдемо імовірність події \bar{A}

$$P(\bar{A}) = (0,4066)^5 = 0,0111.$$

Отже, імовірність шуканої події A буде

$$P(A) = 1 - P(\bar{A}) = 1 - 0,0111 = 0,9889 \approx 0,99.$$

Правило трьох сигм

Якщо випадкова величина X розподілена нормально, то

$$P(|X - a| > 3\sigma) \rightarrow 0,$$

тобто імовірність того, що абсолютна величина відхилення X від її математичного сподівання більше 3σ прямує до 0, а це означає, що $|X - a| < 3\sigma$ – практично достовірна подія.

У практиці це правило використовують так:

Якщо закон розподілу випадкової величини X невідомий, але $|X - a| < 3\sigma$, тоді можна припустити, що X розподілена нормально.

Розподіл χ^2 («хі-квадрат»)

Нехай X_i ($i = 1, 2, \dots, n$) – нормальні, нормовані незалежні величини, тобто їх математичне сподівання дорівнює нулю, середнє квадратичне відхилення дорівнює одиниці і кожна з них розподілена за нормальним законом. Тоді сума квадратів цих величин

$$\chi^2 = \sum_{i=1}^n X_i^2$$

розподілена по закону χ^2 з $k = n$ степенями вільності.

Якщо величини X_i зв'язані одним лінійним співвідношенням, наприклад, $\sum_{i=1}^n X_i = n\bar{X}$, то число степеней вільності буде $k = n - 1$.

Диференціальна функція розподілу χ^2 має вигляд

$$f(x) = \begin{cases} 0, & \text{при } x \leq 0; \\ \frac{1}{2^{\frac{k}{2}} \cdot \Gamma\left(\frac{k}{2}\right)} \cdot e^{-\frac{x}{2}} \cdot x^{\frac{k}{2}-1}, & \text{при } x > 0, \end{cases}$$

де $\Gamma(x) \int_0^{\infty} t^{x-1} e^{-t} dt$ – гама-функція, $\Gamma(n+1) = n!$.

Відмітимо, що розподіл χ^2 визначається параметром – числом степеней вільності k . Коли k зростає, розподіл χ^2 прямує до нормального розподілу дуже повільно.

Розподіл Стьюдента

Нехай X – нормальна нормована випадкова величина, а Y – незалежна від X величина, яка розподілена за законом хі-квадрат з k степенями вільності. Тоді величина

$$T = \frac{X}{\sqrt{\frac{Y}{k}}}$$

має розподіл, який називають ***t-розподілом*** або ***розподілом Стьюдента*** (це є псевдонім англійського статистика Вільяма Госсета) з k степенями вільності.

При зростанні k розподіл Стьюдента швидко наближається до нормального розподілу.

4.4. Закон великих чисел та центральна гранична теорема

4.4.1. Загальні поняття

Граничні теореми теорії імовірностей встановлюють відповідність між теоретичними та дослідними характеристиками випадкових величин або випадкових подій при великій кількості випробувань. Граничні теореми описують також граничні закони розподілу. У підрозділах 3.2 та 3.4 ми ознайомились з деякими граничними теоремами схеми Бернуллі.

Граничні теореми, які встановлюють відповідність між теоретичними та дослідними характеристиками випадкових подій, об'єднують загальною назвою – ***закону великих чисел***.

Закон великих чисел грає важливу роль у різних процесах, пов'язаних з масовим виробництвом.

Граничні теореми, що встановлюють граничні закони розподілу випадкових величин, об'єднують загальною назвою – **центральна гранична теорема**.

Необхідність граничних теорем обумовлена потребою розв'язання, наприклад, таких задач:

1. Коли сума багатьох випадкових величин мало відрізняється від постійної величини, тобто майже перестає бути випадковою величиною і тому її поведінка може прогнозуватись із значною імовірністю?

2. При яких умовах можна із значною імовірністю прогнозувати число появ деякої випадкової події при великій кількості незалежних випробувань?

3. При яких обмеженнях сума багатьох випадкових величин буде розподілена за нормальним законом?

4.4.2. Нерівність Чебишова

При доведенні різних граничних теорем, а також при розв'язанні різних задач важливу роль грає нерівність Чебишова, яка має дві форми.

Перша форма нерівності Чебишова

Для довільної випадкової величини X , яка приймає невід'ємні значення та має скінчене математичне сподівання

$$P(X \geq 1) \leq M(X).$$

Якщо X – дискретна випадкова величина, то

$$P(X \geq 1) = \sum_i p(x_i) \leq \sum_i x_i p(x_i) = M(X).$$

Якщо X – неперервна випадкова величина, $f(x)$ – щільність її імовірностей, то

$$P(X \geq 1) = \int_1^{\infty} f(x) dx \leq \int_1^{\infty} x f(x) dx \leq \int_0^{\infty} x f(x) dx = M(X).$$

Наслідок. Якщо X приймає лише невід'ємні значення, $M(X) < \infty$, $\alpha > 0$, то

$$P(X \geq \alpha) \leq \frac{M(X)}{\alpha}. \quad (1)$$

Дійсно,

$$P(X \geq \alpha) = P\left(\frac{X}{\alpha} \geq 1\right) \leq M\left(\frac{X}{\alpha}\right) = \frac{M(X)}{\alpha}.$$

Нерівність (1) іноді називають **нерівністю Маркова**.

Друга форма нерівності Чебишова

Якщо випадкова величина X має скінчені математичне сподівання та дисперсію, то для довільного $\varepsilon > 0$ має місце нерівність

$$P(|X - M(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}. \quad (2)$$

Доведення. Спочатку розглянемо протилежну подію $|X - M(X)| \geq \varepsilon$. Легко бачити, що ця подія еквівалентна події $(X - M(X))^2 \geq \varepsilon^2$, тому до неї можна застосувати першу форму нерівності Чебишова

$$\begin{aligned} P(|X - M(X)| \geq \varepsilon) &= P\left(\frac{1}{\varepsilon^2} |X - M(X)|^2 \geq 1\right) \leq \\ &\leq \frac{M\left((X - M(X))^2\right)}{\varepsilon^2} = \frac{D(X)}{\varepsilon^2}. \end{aligned} \quad (3)$$

Тепер імовірність протилежної події $|X - M(X)| < \varepsilon$ задовольняє нерівність (2), що і треба було довести.

Приклад 1. Дисперсія випадкової величини X дорівнює 0,001. Яка імовірність того, що випадкова величина X відрізняється від її математичного сподівання $M(X)$ більше ніж на 0,1?

Розв'язання. За нерівністю Чебишова (3) маємо

$$P\left(|X - M(X)| > 0,1\right) \leq \frac{D(X)}{(0,1)^2} = \frac{0,001}{0,01} = 0,1.$$

Нерівності Чебишова дозволяють довести граничну теорему Бернуллі (див. підрозділ 3.4) та інші важливі граничні теореми про стійкість середніх.

4.4.3. Важливі граничні теореми

Теорема Бернуллі. *Нехай імовірність появи події A в кожному із n незалежних повторних випробувань дорівнює p , m – число появ події A (частота події) в n випробуваннях. Тоді*

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1, \quad \varepsilon > 0. \quad (4)$$

Доведення. Частота $\frac{m}{n}$ можна розглядати як невід'ємну випадкову величину X . Знайдемо її математичне сподівання

$$M(X) = M\left(\frac{m}{n}\right) = \frac{1}{n} M(m) = \frac{1}{n} \cdot np = p.$$

Отже, необхідно оцінити імовірність відхилення випадкової величини X від її математичного сподівання. Для цього знайдемо дисперсію цієї випадкової величини

$$D(X) = D\left(\frac{m}{n}\right) = \frac{1}{n^2} D(m) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}.$$

За нерівністю Чебишова (2) одержимо

$$P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) = 1 - \frac{p(1-p)}{n \cdot \varepsilon^2}.$$

Розділ 4. Випадкові величини

Звідси граничним переходом при $n \rightarrow \infty$ одержуємо (4), що й треба було довести.

Теорема Чебишова. *Нехай X_1, X_2, \dots, X_n – послідовність попарно незалежних випадкових величин, які задовольняють умовам*

$$1) M(X_i) = a_i, \quad 2) D(X_i) \leq c$$

для усіх $i = 1, 2, \dots, n$.

Тоді

$$\lim_{x \rightarrow \infty} P \left(\left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n a_i}{n} \right| < \varepsilon \right) = 1. \quad (5)$$

Доведення. Знайдемо математичне сподівання та дисперсію середньої випадкових величин, тобто

$$\frac{X_1 + X_2 + \dots + X_n}{n}$$

$$M \left(\frac{\sum_{i=1}^n X_i}{n} \right) = \frac{1}{n} M \left(\sum_{i=1}^n X_i \right) = \frac{1}{n} \sum_{i=1}^n M(X_i) = \frac{1}{n} \sum_{i=1}^n a_i ;$$

$$D \left(\frac{\sum_{i=1}^n X_i}{n} \right) = \frac{1}{n^2} D \left(\sum_{i=1}^n X_i \right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{c \cdot n}{n^2} = \frac{c}{n}.$$

Застосуємо для випадкової величини $\frac{1}{n} \sum_{i=1}^n X_i$, нерівність Чебишова (2)

$$P \left(\left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n a_i}{n} \right| < \varepsilon \right) \geq 1 - \frac{c}{n \cdot \varepsilon^2}. \quad (6)$$

Границя цієї імовірності при $n \rightarrow \infty$ дорівнює одиниці, тобто рівність (5) доведено.

Центральна гранична теорема. *Нехай задана послідовність незалежних однаково розподілених випадкових величин X_1, X_2, \dots, X_n , $M(X_i) = 0$, $D(X_i) = b$, $i = 1, 2, \dots$*

Розглянемо випадкову величину $Y_n = \sum_{i=1}^n X_i$. Тоді

$$M(Y_n) = \sum_{i=1}^n M(X_i) = 0; \quad D(Y_n) = \sum_{i=1}^n D(x_i) = n \cdot b.$$

При $n \rightarrow \infty$ функція розподілу

$$F_{Y_n}(x) = P(Y_n < x) \rightarrow \frac{1}{\sqrt{2\pi nb}} \int_{-\infty}^x e^{-\frac{z^2}{2nb}} dz,$$

тобто сума Y_n буде розподілена за нормальним законом з математичним сподіванням 0 та дисперсією \sqrt{nb} .

Для доведення цієї теореми треба знайти границю характеристичної функції, побудованої для нормованої випадкової величини

$$Z_n = \frac{Y_n}{\sqrt{nb}}.$$

Наслідок. При $n \geq 30$ розподіл суми однаково розподілених випадкових величин мало відрізняється від нормального розподілу.

Теорема Ляпунова. *Нехай задана послідовність незалежних випадкових величин $X_1, X_2, \dots, X_n, \dots$ таких, що $M(X_i) = 0$,*

$$D(X_i) = b_i^2, \quad i = 1, 2, \dots, n, \dots$$

Побудуємо суму випадкових величин $Y_n = \sum_{i=1}^n X_i$. Позначимо

$B_n^2 = \sum_{i=1}^n b_i^2$. Якщо виконується умова рівномірної малості величин, що утворюють суму

$$\frac{1}{B_n^3} \cdot \sum_{i=1}^n M(X_i)^3 \rightarrow 0 \quad \text{при } n \rightarrow \infty,$$

то сума Y_n буде розподіленою нормально з математичним сподіванням $M(Y_n) = 0$ та дисперсією $D(Y_n) = B_n^2$.

Доведення цієї теореми досить складне, але відмітимо, що у випадку, коли $M(X_i) = a_i \neq 0$, можна розглядати випадкові величини $X'_i = X_i - a_i$. Величини X'_i будуть задовольняти умову теореми Ляпунова.

Приклад 2. Скільки додатків треба взяти у теоремі Чебишова, щоб з надійністю 96% і точністю до 0,01 виконувалась наближена рівність

$$\frac{1}{n} \sum_{i=1}^n X_i \approx \frac{1}{n} \sum_{i=1}^n M(X_i).$$

Розв'язання. В цьому прикладі $\varepsilon = 0,01$. Щоб одержати надійність 96% згідно формули (6) достатньо підібрати таке n , яке задовольняє нерівність

$$\frac{c}{\varepsilon^2 n} \leq 0,04 \Rightarrow n \geq \frac{c}{0,04 \cdot 0,0001} = 250000c.$$

Зауваження 1. Приклад 2 показує, що навіть у випадку не дуже великих точності та надійності, треба брати значну кількість додатків (n – досить велике число). Це означає, що оцінки, одержані з використанням нерівності (6), – завищені. Більш точні оцінки можна одержати за допомогою теореми Ляпунова.

4.5. Закони розподілу та числові характеристики двовимірних випадкових величин

4.5.1. Загальні поняття

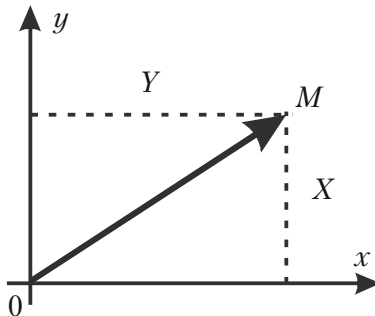
Вище розглянуті випадкові величини X , які при кожному випробуванні визначались одним можливим числовим значенням. Тому таку випадкову величину X називають **одновимірною**.

Якщо можливі значення випадкової величини визначаються у кожному випробуванні 2, 3, ..., n числами, то такі величини називають відповідно **дво, трьох, ..., n -вимірними** відповідно.

Двовимірну випадкову величину будемо позначати (X, Y) , X та Y при цьому будуть компонентами. Величини X та Y , що розглядаються одночасно, утворюють систему двох випадкових величин. Аналогічно можна розглядати систему n випадкових величин.

Означення 1. Сукупність n одночасно розглядаємих випадкових величин (X_1, X_2, \dots, X_n) називають **системою випадкових величин**.

Систему n випадкових величин (X_1, X_2, \dots, X_n) можна розглядати як випадкову точку в n -вимірному просторі з координатами (X_1, X_2, \dots, X_n) або як випадковий вектор, напрямлений з початку системи координат у точку $M(X_1, X_2, \dots, X_n)$.



Мал. 14.

При $n = 2$ маємо систему двох випадкових величин (X, Y) , яку можна зобразити як випадкову точку $M(X, Y)$ на площині xOy або як випадковий вектор \overline{OM} (див. мал. 14).

Багатовимірні випадкові величини бувають **дискретними** та **неперервними** (компоненти цих величин відповідно будуть дискретними та неперервними).

4.5.2. Закон розподілу імовірностей дискретної двовимірної випадкової величини

Означення 2. *Законом розподілу дискретної двовимірної випадкової величини називають перелік можливих значень цієї величини (x_i, y_k) та їх імовірностей $p(x_i, y_k)$, $i = 1, 2, \dots, n$; $k = 1, 2, \dots, m$.*

Найчастіше закон розподілу двовимірної дискретної випадкової величини задають у вигляді таблиці з двома входами

У першому рядку таблиці записують усі можливі значення компоненти X . У першому стовпчику таблиці записують усі можливі значення компоненти Y . На перетині k -того рядка та i -того стовпчика записують імовірність $p(x_i, y_k)$ того, що двовимірна випадкова величина (X, Y) прийме значення (x_i, y_k) , $i = 1, 2, \dots, n$; $k = 1, 2, \dots, m$.

Y	X					
	x_1	x_2	...	x_i	...	x_n
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$...	$p(x_i, y_1)$...	$p(x_n, y_1)$
y_2	$p(x_1, y_2)$	$p(x_2, y_2)$...	$p(x_i, y_2)$...	$p(x_n, y_2)$
...
y_k	$p(x_1, y_k)$	$p(x_2, y_k)$...	$p(x_i, y_k)$...	$p(x_n, y_k)$
...
y_m	$p(x_1, y_m)$	$p(x_2, y_m)$...	$p(x_i, y_m)$...	$p(x_n, y_m)$

Події $(X = x, Y = y_k)$, $i = 1, 2, \dots, n$; $k = 1, 2, \dots, m$, утворюють повну групу, тому сума імовірностей таблиці дорівнює одиниці, тобто

$$\sum_{i=1}^n \sum_{k=1}^m p(x_i, y_k) = 1.$$

Закон розподілу двовимірної випадкової величини дозволяє отримати закони розподілу кожної компоненти.

Дійсно, події (x_i, y_1) , (x_i, y_2) , ..., (x_i, y_m) несумісні, тому імовірність $P(x_i)$ того, що X прийме значення x_i за теоремою додавання імовірностей буде

$$P(x_i) = P(x_i, y_1) + P(x_i, y_2) + \dots + P(x_i, y_m),$$

тобто дорівнює сумі імовірностей, що розташовані в i -тому стовпчику таблиці розподілу.

Аналогічно, додаванням імовірностей k -того рядка таблиці, одержимо імовірність

$$P(y_k) = P(x_1, y_k) + P(x_2, y_k) + \dots + P(x_n, y_k)$$

Приклад 1. Знайти закони розподілу компонент двовимірної випадкової величини, закон розподілу якої заданий таблицею

Y	X		
	x_1	x_2	x_3
y_1	0,1	0,3	0,2
y_2	0,06	0,18	0,16

Розв'язання. Закони розподілу X та Y будуть мати вигляд

X	x_1	x_2	x_3
P	0,16	0,48	0,36

Y	y_1	y_2
P	0,6	0,4

Імовірності відповідних значень X та Y знаходимо так

$$p(x_1) = 0,1 + 0,06 = 0,16;$$

$$p(x_2) = 0,3 + 0,18 = 0,48; \quad p(x_3) = 0,2 + 0,16 = 0,36;$$

$$p(y_1) = 0,1 + 0,3 + 0,2 = 0,6; \quad p(y_2) = 0,06 + 0,18 + 0,16 = 0,4.$$

$$\text{Контроль: } 0,16 + 0,48 + 0,36 = 1; \quad 0,6 + 0,4 = 1.$$

Означення 3. *Інтегральною функцією розподілу (функцією розподілу) двовимірної випадкової величини (X, Y) називають функцію двох змінних $F(x, y)$, яка визначає для кожної пари чисел (X, Y) імовірність виконання нерівностей $X < x; Y < y$, тобто*

$$F(x, y) = P(X < x; Y < y).$$

Аналогічно визначають функцію розподілу системи n випадкових величин

$$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n).$$

Властивості імовірності P та Означення 3 функції розподілу дозволяють довести такі властивості функції розподілу, які у випадку двовимірної випадкової величини виглядають

1) $0 \leq F(x, y) \leq 1$;

2) $F(x, y)$ не спадна функція за кожним аргументом, тобто

$$F(x_2, y) \geq F(x_1, y), \text{ якщо } x_2 > x_1;$$

$$F(x, y_2) > F(x, y_1), \text{ якщо } y_2 > y_1.$$

3) Мають місце граничні співвідношення

$$F(-\infty, y) = 0; \quad F(x, -\infty) = 0; \quad F(-\infty, -\infty) = 0; \quad F(\infty, \infty) = 1.$$

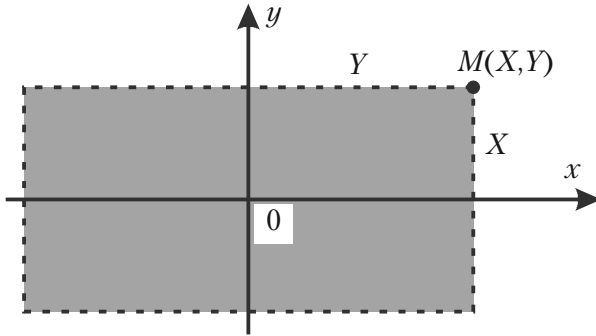
4) Імовірність влучення випадкової точки до прямокутника

$$\{x_1 \leq X \leq x_2; y_1 \leq Y \leq y_2\}$$

можна знайти за формулою

$$\begin{aligned} &P(x_1 < X < x_2; y_1 < Y < y_2) = \\ &= \{F(x_2, y_2) - F(x_1, y_2)\} - \{F(x_2, y_1) - F(x_1, y_1)\}. \end{aligned} \quad (1)$$

Геометричний зміст функції розподілу $F(x, y)$ – це імовірність того, що випадкова точка $M(X, Y)$ попаде у нескінчений прямокутник з вершиною в точці (X, Y) і розміщений нижче та лівіше цієї вершини (див. мал. 15).



Мал. 15.

Приклад 2. Знайти імовірність влучення випадкової точки (X, Y) у прямокутник, обмежений лініями

$$x = \frac{\pi}{6}; \quad x = \frac{\pi}{2}; \quad y = \frac{\pi}{3}; \quad y = \frac{\pi}{4},$$

якщо задана функція розподілу вигляду

$$F(x, y) = \sin x \cdot \sin y; \quad 0 \leq x \leq \frac{\pi}{2}; \quad 0 \leq y \leq \frac{\pi}{2}.$$

Розв'язання. У заданому випадку

$$x_1 = \frac{\pi}{6}; \quad x_2 = \frac{\pi}{2}; \quad y_1 = \frac{\pi}{4}; \quad y_2 = \frac{\pi}{3}.$$

Згідно з формулою (1) одержуємо

$$\begin{aligned} P\left(\frac{\pi}{6} < X < \frac{\pi}{2}; \frac{\pi}{4} < Y < \frac{\pi}{3}\right) &= \left[\sin \frac{\pi}{2} \cdot \sin \frac{\pi}{3} - \sin \frac{\pi}{6} \cdot \sin \frac{\pi}{3} \right] - \\ &- \left[\sin \frac{\pi}{2} \cdot \sin \frac{\pi}{4} - \sin \frac{\pi}{6} \cdot \sin \frac{\pi}{4} \right] = \left[\frac{\sqrt{3}}{2} - \frac{1}{2} \cdot \frac{\sqrt{3}}{2} \right] - \left[\frac{\sqrt{2}}{2} - \frac{1}{2} \cdot \frac{\sqrt{2}}{2} \right] = \\ &= \frac{\sqrt{3} - \sqrt{2}}{4} \approx 0,08. \end{aligned}$$

4.5.3. Неперервна двовимірна випадкова величина

Двовимірну випадкову величину можна задавати функцією розподілу $F(x, y) = P(X < x; Y < y)$ або диференціальною функцією розподілу.

Означення 4. *Диференціальною функцією розподілу (двовимірною щільністю імовірностей) $f(x, y)$ двовимірної випадкової величини (X, Y) називають мішану частинну похідну другого порядку від інтегральної функції розподілу*

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

Аналогічно визначають щільність імовірностей n -вимірної випадкової величини, тобто

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \cdot \partial x_2 \cdot \dots \cdot \partial x_n}.$$

Таким чином, якщо функція розподілу $F(x, y)$ двовимірної випадкової величини відома, то за формулою (2) можна знайти диференціальну функцію розподілу $f(x, y)$ цієї випадкової величини.

Якщо відома щільність імовірностей $f(x, y)$ двовимірної випадкової величини, то її функцію розподілу знаходять за формулою

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy,$$

тобто з використанням невластного двократного інтегралу.

Імовірність влучення випадкової точки (X, Y) в довільну область D знаходять за формулою

$$P((X, Y) \in D) = \iint_D f(x, y) dx dy.$$

Диференціальна функція розподілу $f(x, y)$ задовольняє властивостям:

1) $f(x, y) \geq 0$, тобто вона не від'ємна;

$$2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$$

4.5.4. Залежні та незалежні випадкові величини

Дві випадкові величини **незалежні**, якщо закон розподілу однієї з них не залежить від того, які можливі значення прийняла друга величина.

Випадкові величини **залежні**, якщо закон розподілу однієї величини залежить від того, які значення прийняла друга величина.

У теорії імовірностей доведено таке твердження.

Теорема. *Щоб випадкові величини X та Y були незалежні, необхідно і достатньо, щоб інтегральна функція системи (X, Y) дорівнювала добутку інтегральних функцій кожної з них*

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

Наслідок. Щоб неперервні випадкові величини X та Y були незалежними, необхідно і достатньо, щоб диференціальна функція системи (X, Y) дорівнювала добутку диференціальних функцій складових

$$f(x, y) = f_1(x) \cdot f_2(y).$$

4.5.5. Числові характеристики двовимірної випадкової величини

Математичне сподівання двовимірної випадкової величини (X, Y) характеризує координати **центру розподілу** випадкової величини. Ці координати у випадку неперервних величин знаходять за формулами

$$m_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy; \quad m_y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy. \quad (3)$$

Дисперсії D_x та D_y характеризують розсіювання випадкової точки (X, Y) вздовж координатних осей Ox та Oy , відповідно. Їх знаходять за формулами

$$Dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 \cdot f(x, y) dx dy - m_x^2, \quad (4)$$

$$D_Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 \cdot f(x, y) dx dy - m_Y^2. \quad (5)$$

Для опису двовимірної випадкової величини крім математичного сподівання, дисперсії та середніх квадратичних відхилень $\sigma_X = \sqrt{D_X}$, $\sigma_Y = \sqrt{D_Y}$ використовують також інші характеристики, а саме – **кореляційний момент** (або **коваріація**)

$$\text{cov}(X, Y) = K_{XY} = M((X - m_X)(Y - m_Y)).$$

Для неперервних величин X та Y

$$K_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y) f(x, y) dx dy.$$

Коефіцієнт кореляції

$$r_{XY} = \frac{K_{XY}}{\sigma_X \sigma_Y}.$$

Коефіцієнт кореляції є кількісна характеристика залежності випадкових величин X та Y і часто використовується в статистиці.

Якщо випадкові величини X та Y дискретні, то в формулах (3)–(6) знаки інтегралів замінюють знаками суми по усім можливим значенням випадкових величин.

Означення 5. *Випадкові величини X та Y зуть **некорельованими**, якщо їх кореляційний момент або коефіцієнт кореляції дорівнює нулеві.*

Властивості коефіцієнта кореляції

- 1) $|r_{XY}| \leq 1$;
- 2) якщо X та Y незалежні, то $r_{XY} = 0$;
- 3) якщо між X та Y є лінійна залежність $Y = aX + b$, де a та b – постійні, то $|r_{XY}| = 1$.

Зауваження 1. *Якщо момент кореляції або коефіцієнт кореляції не дорівнює нулеві, тоді випадкові величини X та Y – корельовані. Дві*

корельовані величини обов'язково залежні. Але дві залежні випадкові величини можуть бути корельованими або некорельованими, тобто їх коефіцієнт кореляції може дорівнювати нулеві, а може і не дорівнювати нулеві.

Із незалежності двох величин випливає їх некорельованість, але із некорельованості ще не випливає незалежність цих величин. У випадку нормального розподілу величин із некорельованості випадкових величин випливає їх незалежність.

4.6. Функції випадкової величини та їх характеристики

4.6.1. Поняття функції

У багатьох випадках треба розглядати дві випадкові величини X та Y . Так, наприклад, при аналізі діяльності підприємства треба враховувати кількість усіх працюючих X та кількість зроблених виробів Y . З різних причин кількість працюючих та зроблених виробів кожного дня можуть бути різними, тобто X та Y будуть випадковими величинами.

Означення 2. Якщо вказано закон, за яким кожному можливому значенню випадкової величини X відповідає певне значення випадкової величини Y , то Y звать **функцією X** і позначають $Y = \varphi(X)$.

Відзначимо, що іноді різним можливим значенням випадкової величини X відповідають однакові значення Y . Наприклад, якщо $Y = X^2$, то значенням -3 та 3 випадкової величини X відповідає одне значення випадкової величини $Y = 9$.

Однією із можливих задач теорії імовірностей є визначення законів розподілу та числових характеристик функцій випадкового аргументу, закон розподілу якого відомий. Вкажемо основні формули для розв'язування цієї задачі.

4.6.2. Закон розподілу та числові характеристики функції дискретного випадкового аргументу

Нехай $Y = \varphi(X)$, аргумент X – дискретна випадкова величина з можливими значеннями x_1, x_2, \dots, x_n , імовірності яких дорівнюють p_1, p_2, \dots, p_n відповідно, тобто X задана законом

X	x_1	x_2	\dots	x_n
$P(X)$	p_1	p_2	\dots	p_n

У цьому випадку Y також дискретна випадкова величина з можливими значеннями $y_1 = \varphi(x_1), y_2 = \varphi(x_2), \dots, y_n = \varphi(x_n)$.

Із події «величина X прийняла значення x_k » випливає подія «величина Y прийняла значення $\varphi(x_k)$ », тому імовірності можливих значень Y також дорівнюють p_1, p_2, \dots, p_n . Це означає, що закон розподілу Y буде мати вигляд

Y	$\varphi(x_1)$	$\varphi(x_2)$	\dots	$\varphi(x_n)$
$P(y)$	p_1	p_2	\dots	p_n

Математичне сподівання, дисперсію та середнє квадратичне відхилення функції Y обчислюють за формулами

$$M(Y) = \sum_{k=1}^n \varphi(x_k) p_k, \tag{1}$$

$$D(Y) = M(Y^2) - M^2(Y) = \sum_{k=1}^n [\varphi(x_k)]^2 p_k - M^2(Y),$$

$$\sigma(Y) = \sqrt{D(Y)}.$$

Початкові та центральні моменти розподілу знаходять за формулами

$$\nu_k = \sum_{i=1}^n (\varphi(x_i))^k p_i,$$

$$\mu_k = \sum_{i=1}^n (\varphi(x_i) - M(Y))^k p_i.$$

Приклад 1. Дискретна випадкова величина задана законом розподілу

X	1	3	5
P	0,2	0,5	0,3

Знайти математичне сподівання функції $Y = X^2 + 1$.

Розв'язання. Можливими значеннями Y будуть

$$y_1 = 1^2 + 1 = 2; \quad y_2 = 3^2 + 1 = 10; \quad y_3 = 5^2 + 1 = 26.$$

За формулою (1) знаходимо математичне сподівання Y

$$M(Y) = M(X^2 + 1) = 2 \cdot 0,2 + 10 \cdot 0,5 + 26 \cdot 0,3 = 13,2.$$

4.6.3. Закон розподілу та числові характеристики функції неперервного випадкового аргументу

Нехай Y — неперервна випадкова величина, закон розподілу якої заданий диференціальною функцією розподілу (щільність ймовірностей) $f(x)$; випадкова величина $Y = \varphi(X)$.

Якщо φ — диференційовна функція, монотонна на усьому проміжку можливих значень X , то щільність розподілу функції $Y = \varphi(X)$ визначають так

$$g(y) = f(\psi(y)) \cdot |\psi'(y)|, \quad (2)$$

де ψ — функція, обернена по відношенню до функції φ ,

ψ' — похідна першого порядку.

Якщо φ — не монотонна функція в області визначення аргументу X , то обернена функція неоднозначна і щільність розподілу

Розділ 4. Випадкові величини

$g(y)$ визначається як сума додатків, кількість яких дорівнює кількості значень оберненої функції, тобто

$$g(y) = \sum_{i=1}^k f(\psi_i(y)) |\psi'_i(y)|, \quad (3)$$

де $\psi_i(y)$ – обернені функції при заданому y .

Приклад 2. Випадкова величина X розподілена за нормальним законом з математичним сподіванням, що дорівнює нулеві. Знайти закон розподілу функції $Y = X^3$.

Розв'язання. Згідно означенню нормального розподілу неперервної випадкової величини X та умови прикладу диференціальна функція розподілу X має вигляд

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

Функція $Y = X^3$ диференційовна, $Y' = 3X^2 > 0$ тому вона зростає для усіх $x \in (-\infty, \infty)$. Отже, можна застосувати формулу (2) для знаходження диференціальної функції розподілу $g(y)$ випадкової величини Y .

У даному випадку з рівності

$$Y = X^3 \Rightarrow X = \sqrt[3]{Y},$$

тобто

$$\psi(y) = \sqrt[3]{y} = y^{\frac{1}{3}}.$$

Тому формула (2) прийме вигляд

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{y^{\frac{2}{3}}}{2\sigma^2}} \cdot \left| \left(y^{\frac{1}{3}} \right)' \right| = \frac{e^{-\frac{y^{\frac{2}{3}}}{2\sigma^2}}}{3\sigma\sqrt{2\pi} \cdot y^{\frac{2}{3}}}.$$

Для знаходження математичного сподівання від $Y = \varphi(X)$ можна спочатку знайти $g(y)$ – диференціальну функцію розподілу величини Y за формулою (2) або (3), а потім використати формулу

$$M(Y) = \int_{-\infty}^{\infty} y \cdot g(y) dy.$$

Але більш доцільно знаходити математичне сподівання функції неперервного випадкового аргументу $\varphi(X)$ безпосередньо за формулою

$$M(\varphi(X)) = \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \quad (4)$$

де $f(x)$ – щільність імовірностей величини X .

Якщо величина X може приймати значення лише в проміжку $[a, b]$, то формула (4) спрощується

$$M(\varphi(X)) = \int_a^b \varphi(x) f(x) dx. \quad (5)$$

Приклад 3. Неперервна випадкова величина X задана диференціальною функцією розподілу

$$f(x) = \begin{cases} \sin x, & \text{при } x \in \left(0, \frac{\pi}{2}\right); \\ 0, & \text{при } x \notin \left(0, \frac{\pi}{2}\right). \end{cases}$$

Знайти математичне сподівання функції $Y = X^2$.

Розв'язання. У даному випадку $\varphi(X) = X^2$, $x \in \left(0, \frac{\pi}{2}\right)$, тому за формулою (5) одержимо

$$M(Y) = M(X^2) = \int_0^{\frac{\pi}{2}} x^2 \sin x dx.$$

Інтегруючи частинами два рази, одержимо потрібне математичне сподівання

$$\begin{aligned} \int_0^{\frac{\pi}{2}} x^2 \sin x dx &= -x^2 \cos x \Big|_0^{\frac{\pi}{2}} + 2 \int_0^{\frac{\pi}{2}} x \cos x dx = \\ &= 2 \left(x \sin x \Big|_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} \sin x dx \right) = \pi + 2 \cos x \Big|_0^{\frac{\pi}{2}} = \pi - 2. \end{aligned}$$

Отже, одержали

$$M(Y) = M(X^2) = \pi - 2.$$

Дисперсію функції Y неперервного випадкового аргументу X визначають звичайним чином $D(Y) = M(Y^2) - M^2(Y)$, а обчислюють за формулою

$$D(Y) = D(\varphi(X)) = \int_{-\infty}^{\infty} \varphi^2(x) f(x) dx - \left(\int_{-\infty}^{\infty} \varphi(x) f(x) dx \right)^2. \quad (6)$$

У випадку, коли X змінюється лише в проміжку $[a, b]$, дисперсію функції $Y = \varphi(X)$ знаходять за формулою

$$D(\varphi(X)) = \int_a^b \varphi^2(x) f(x) dx - \left(\int_a^b \varphi(x) f(x) dx \right)^2. \quad (7)$$

У формулах (6) та (7) $f(x)$ – це щільність імовірностей неперервної випадкової величини X (диференціальна функція розподілу X).

Питання для самоперевірки та вправи

1. Як визначають випадкові величини, дискретні та неперервні випадкові величини?
2. Якими способами можна визначити дискретну випадкову величину?
3. Вказати основні закони розподілу дискретної випадкової величини та умови їх використання.
4. Як визначаються і що характеризують числові характеристики дискретних випадкових величин?
5. За якими формулами обчислюють числові характеристики дискретних випадкових величин?
6. Як визначають функцію розподілу та щільності імовірностей неперервних випадкових величин? Які властивості мають ці функції?
7. Який існує зв'язок між інтегральною та диференціальною функціями розподілу імовірностей?
8. За якими формулами можна обчислити імовірність влучення випадкової величини в проміжок (a, b) , використовуючи інтегральну або диференціальну функції розподілу?
9. Які числові характеристики існують для неперервних випадкових величин та що характеризує кожна з них?
10. За якими формулами обчислюють числові характеристики неперервних випадкових величин?
11. Вказати основні властивості математичного сподівання та дисперсії.
12. Вказати основні закони розподілу неперервних випадкових величин та їх вигляд.
13. Чому дорівнюють числові характеристики основних законів розподілу дискретних та неперервних випадкових величин?
14. За якими формулами треба знаходити імовірність влучення випадкової величини X в проміжок (a, b) , якщо X розподілено за рівномірним, показниковим або нормальним законами?
15. Як знайти функції розподілу функції $Y = \varphi(X)$, якщо X – дискретна або неперервна випадкова величина?
16. За якими формулами обчислюють числові характеристики функції дискретного та неперервного випадкового аргументу?

17. Як визначають початкові та центральні моменти, коефіцієнт кореляції та як пов'язані поняття кореляції, залежності та незалежності випадкових величин?

18. Правило 3σ та його використання.

19. Центральна гранична теорема Ляпунова.

Розв'язати задачі

1. По мішені проведено 3 постріли. Імовірність влучення у мішень першого пострілу становить 0,1, другого — 0,2, а третього — 0,3. Знайти ряд розподілу кількості влучень при трьох пострілах. Побудувати графік функції розподілу.

Відповідь:

X	0	1	2	3
P	0,504	0,398	0,092	0,006

Графік відповідний.

2. Серед 10 виробів є один бракований. Щоб його знайти, беруть навмання один виріб за іншим і кожен взятий виріб перевіряють. Побудувати ряд розподілу кількості перевірених виробів.

Відповідь:

$$X\{1, 2, \dots, 10\} \quad P(X = k) = p \cdot q^{k-1}, \quad k = 1, 2, \dots, 10,$$

де p — імовірність браку.

3. Виконується 4 незалежних пострілів по мішені. Імовірність влучення при одному пострілі дорівнює 0,25. Знайти закон розподілу кількості влучень X , найімовірніше число влучень та його імовірність.

Відповідь:

X	0	1	2	3	4
P	$\frac{81}{256}$	$\frac{108}{256}$	$\frac{54}{256}$	$\frac{12}{256}$	$\frac{1}{256}$

$$m_0 = 1, \quad P_4(1) = \frac{108}{256}.$$

4. Монету кидають 6 разів. Скласти ряд розподілу і побудувати функцію розподілу відношення частоти появ герба до числа появ надпису.

Відповідь:

X	0	$\frac{1}{5}$	$\frac{1}{2}$	1	2	5	∞
P	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

5. На шляху руху автомобіля 6 світлофорів, кожен з яких дозволяє або забороняє рух автомобіля з імовірністю 0,5. Скласти ряд розподілу і побудувати функцію розподілу кількості світлофорів, які автомобіль минув без зупинки. Чому дорівнює математичне сподівання та дисперсія цієї випадкової величини?

Відповідь:

X	0	1	2	3	4	5	6
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$

$$M(X) = \frac{63}{64} \approx 0,984; D(X) \approx 1,797.$$

6. Знайти закон розподілу дискретної випадкової величини X , яка може приймати лише два значення: x_1 з імовірністю 0,6 та x_2 ($x_1 < x_2$), якщо $M(X) = 1,4$; $D(X) = 0,24$.

Відповідь:

X	1	2
P	0,6	0,4

7. Знайти закон розподілу дискретної випадкової величини X , яка може приймати лише два значення: x_1 з імовірністю 0,4 та x_2 ($x_1 < x_2$), якщо $M(X) = 3,2$; $D(X) = 0,96$.

Розділ 4. Випадкові величини

Відповідь:

X	2	4
P	0,4	0,6

8. Нехай X – кількість появ числа 5 при двох киданнях грального кубика. Знайти закон розподілу величини X та її математичне сподівання.

Відповідь:

X	0	1	2
P	$\frac{25}{36}$	$\frac{5}{18}$	$\frac{1}{36}$

$$M(X) = \frac{1}{3}.$$

9. Технологічний процес є таким, що брак складає 5% усіх виробів. Навмання взято 20 виробів. Знайти математичне сподівання та дисперсію кількості бракованих виробів.

Відповідь: $M(X) = 1$; $D(X) = 0,95$.

10. Знайти дисперсію дискретної випадкової величини X – числа появ події A в двох незалежних випробуваннях, якщо імовірність появ події у цих випробуваннях однакові і відомо, що $M(X) = 1,2$.

Відповідь: 0,56.

11. Проекція X радіус-вектора випадкової точки кола радіуса a на діаметр має щільність імовірності

$$f(x) = \begin{cases} \frac{1}{\pi\sqrt{a^2 - x^2}}, & \text{при } |x| < a; \\ 0, & \text{при } |x| \geq a. \end{cases}$$

Визначити математичне сподівання та дисперсію випадкової величини X .

Відповідь: $M(X) = 0$; $D(X) = \frac{a^2}{2}$.

12. Випадкова величина X задана інтегральною функцією

$$F(x) = \begin{cases} 0, & x \leq -1; \\ \frac{3}{4}x + \frac{3}{4}, & -1 < x \leq \frac{1}{3}; \\ 1, & x > \frac{1}{3}. \end{cases}$$

Знайти імовірність того, що X прийме значення з проміжку $\left(0, \frac{1}{3}\right)$.

Відповідь: $P\left(0 < X < \frac{1}{3}\right) = \frac{1}{4}$.

13. Дана функція розподілу неперервної випадкової величини X

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \sin x, & 0 < x \leq \frac{\pi}{2}; \\ 1, & x > \frac{\pi}{2}. \end{cases}$$

Знайти диференціальну функцію розподілу $f(x)$ (щільність імовірності).

Відповідь:

$$f(x) = \begin{cases} \cos x, & x \in \left(0, \frac{\pi}{2}\right]; \\ 0, & x \notin \left(0, \frac{\pi}{2}\right]. \end{cases}$$

14. Задана щільність імовірності величини X

$$f(x) = \begin{cases} \frac{2}{3} \sin 3x, & x \in \left(0, \frac{\pi}{3}\right]; \\ 0, & x \notin \left(0, \frac{\pi}{3}\right]. \end{cases}$$

Знайти імовірність того, що X прийме значення із $\left(\frac{\pi}{6}, \frac{\pi}{4}\right)$.

Відповідь: $\frac{\sqrt{2}}{9}$.

15. Задана диференціальна функція розподілу неперервної випадкової величини X

$$f(x) = \begin{cases} x - \frac{1}{2}, & 1 < x \leq 2; \\ 0, & x \notin (1, 2). \end{cases}$$

Знайти інтегральну функцію розподілу X .

Відповідь:

$$F(x) = \begin{cases} 0, & x \leq 1; \\ \frac{x^2}{2} - \frac{x}{2}, & x \in (1, 2]; \\ 1, & x > 2. \end{cases}$$

16. Задана щільність імовірності випадкової величини X

$$f(x) = \frac{2C}{1+x^2}.$$

Знайти постійний параметр C .

Відповідь: $\frac{1}{2\pi}$.

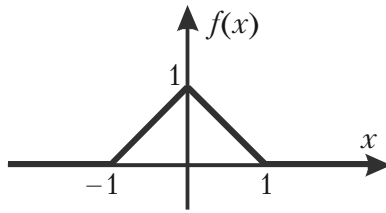
17*. Щільність імовірності випадкової величини X задана виразом

$$f(x) = \begin{cases} a \cos x, & 0 \leq x \leq \frac{\pi}{2}; \\ 0, & x < 0, x > \frac{\pi}{2}. \end{cases}$$

Знайти коефіцієнт a , математичне сподівання та дисперсію.

Відповідь: $a = 1$; $M(X) = \frac{\pi}{2} - 1$; $D(X) = \pi - 3$.

18. Графік диференціальної функції розподілу випадкової величини X (закон Сімпсона) зображено на малюнку.



Мал. 16.

Знайти щільність імовірності X , її математичне сподівання, дисперсію та середнє квадратичне відхилення.

Відповідь:

$$f(x) = \begin{cases} x+1, & -1 < x \leq 0; \\ -x+1, & 0 < x \leq 1; \\ 0, & x \leq -1, x > 1; \end{cases}$$

$$M(X) = 0; \quad D(X) = \frac{1}{6}; \quad \sigma(X) = \frac{\sqrt{6}}{6}.$$

19. Задана функція

$$f(x) = \begin{cases} 0, & x < 0; \\ \frac{1}{2} \sin x, & 0 \leq x \leq \pi; \\ 0, & x > \pi. \end{cases}$$

Розділ 4. Випадкові величини

Показати, що вона буде диференціальною функцією розподілу деякої випадкової величини X . Знайти $M(X)$, $D(X)$, $\sigma(X)$.

Відповідь:

$$f(x) \geq 0 \quad \text{та} \quad \int_{-\infty}^{\infty} f(x) dx = 1; \quad M(X) = \frac{\pi}{2}; \quad D(X) = \frac{\pi^2}{4} - 2;$$

$$\sigma(X) = \sqrt{\frac{\pi^2}{4} - 2} \approx 0,69.$$

20*. Щільність імовірності випадкових амплітуд бокової хита-виці човна має вигляд (закон Релея)

$$f(x) = \frac{x}{a^2} e^{-\frac{x^2}{2a^2}}, \quad x \geq 0.$$

Визначити: $M(X)$, $D(X)$ та $\sigma(X)$.

Відповідь:

$$M(X) = a\sqrt{\frac{\pi}{2}}; \quad D(X) = a^2\left(2 - \frac{\pi}{2}\right); \quad \sigma(X) = a\sqrt{2 - \frac{\pi}{2}}.$$

21. Функція розподілу випадкової величини X задана у вигляді

$$F(x) = \begin{cases} 0, & x \leq 0; \\ ax^3, & 0 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Знайти коефіцієнт a та щільність розподілу величини X .

Відповідь: $a = 1$;

$$f(x) = \begin{cases} 0, & x \leq 0; \\ 3x^2, & 0 < x \leq 1; \\ 0, & x > 1. \end{cases}$$

22. Задано щільність розподілу кількості прибутку X

$$f(x) = ae^{-|x|}.$$

Знайти коефіцієнт a та імовірність одержання величини прибутку X із відрізка $[0,5, 1]$ мл. гривень.

$$\text{Відповідь: } a = \frac{1}{2}; P(0,5 < X < 1) = \frac{\sqrt{e} - 1}{2e}.$$

23. Випадкова величина доходу підприємства X має диференціальну функцію розподілу

$$f(x) = \begin{cases} 0,2, & -2 \leq x \leq 3; \\ 0, & x < -2 \text{ TA } x > 3. \end{cases}$$

Знайти математичне сподівання, дисперсію та імовірність одержання доходу $x \in (1, 5]$.

$$\text{Відповідь: } M(X) = \frac{1}{2}; D(X) = \frac{6,25}{3} \approx 2,1; P(1 < X < 5) = 0,4.$$

24. Диференціальна функція розподілу прибутку X підприємства відома

$$f(x) = \begin{cases} \frac{2}{25}x, & x \in (0, 5); \\ 0, & x \notin (0, 5). \end{cases}$$

Знайти математичне сподівання прибутку, дисперсію та імовірність одержання прибутку $x \in [3, 4]$ мл. гривень.

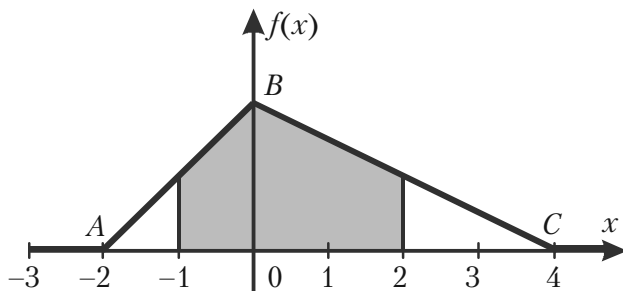
$$\text{Відповідь: } M(X) = 1; D(X) = \frac{25}{18}; P(3 < X < 4) = \frac{7}{25}.$$

25. Знайти дисперсію та середнє квадратичне відхилення випадкової величини X , заданою функцією розподілу

$$F(x) \begin{cases} 0, & x \leq -2; \\ \frac{x}{4} + \frac{1}{2}, & -2 < x \leq 2; \\ 1, & x > 2. \end{cases}$$

Відповідь: $D(X) = \frac{4}{3}$; $\sigma(X) = \frac{2\sqrt{3}}{3}$.

26*. Графік щільності розподілу імовірності задано мал. 17.



Мал. 17.

Знайти аналітичний вираз щільності розподілу, функцію розподілу, математичне сподівання, дисперсію та імовірність

$$P(-1 \leq X \leq 2).$$

Рекомендація: з властивості щільності розподілу $\int_{-\infty}^{\infty} f(x) dx = 1$,

геометричного змісту інтегралу та рівності $S_{\Delta ABC} = \frac{6h}{2}$ випливає, що

точка B має координати $(0, \frac{1}{3})$.

Відповідь:

$$f(x) = \begin{cases} 0, & x \leq -2 \\ \frac{1}{6}(2+x), & -2 < x \leq 0; \\ \frac{1}{12}(4-x), & 0 < x \leq 4; \\ 0, & x > 4, \end{cases}$$

$$M(X) = \frac{2}{3}; \quad D(X) = \frac{14}{9}; \quad P(-1 \leq X \leq 2) = \frac{3}{4};$$

$$F(x) = \begin{cases} 0, & x \leq -2; \\ \frac{1}{12}(2+x)^2, & -2 < x \leq 0; \\ 1 - \frac{1}{24}(4-x)^2, & 0 < x \leq 4; \\ 1, & x > 4. \end{cases}$$

27. Випадкова величина X розподілена рівномірно на $[a, b]$. Знайти її функції щільності та розподілу ймовірності, побудувати їх графіки, знайти математичне сподівання та дисперсію, якщо задано відрізок

а) $[-3; 7]$; б) $[1; 5]$; в) $[-10; -4]$; г) $[2,3; 2,5]$.

28. Знайти математичне сподівання та дисперсію випадкової величини, розподіленої за біноміальним законом, якщо

а) $n = 10, p = 0,4$; б) $n = 30, p = 0,1$; в) $n = 100, q = 0,2$;
г) $n = 50, q = 0,8$.

29. Нормально розподілена випадкова величина має математичне сподівання $M(X)$ та дисперсію $D(X)$. Записати диференціальну функцію її розподілу, якщо

а) $M(X) = 3, D(X) = 4$; б) $M(X) = -1, D(X) = 9$;
в) $M(X) = 0, D(X) = 1$; г) $M(X) = -2, D(X) = \sqrt{3}$.

30. Випадкова величина X розподілена за законом Пуассона з параметром a , випадкова величина Y розподілена за тим же законом з параметром b . Ці випадкові величини незалежні. Знайти закон розподілу, математичне сподівання та дисперсію їх суми, якщо

1) $a = 0,3, b = 1,2$; 2) $a = 1,7, b = 2,1$; 3) $a = 5, b = 2$;
4) $a = 0,2, b = 1,5$;

31. Для нормально розподіленої випадкової величини X знайти

а) $P(12 < X < 14)$, якщо $M(X) = 10, D(X) = 4$;
б) $P(15 < X < 25)$, якщо $M(X) = 20, \sigma(X) = 5$;
в) $P(5 < X < 15)$, якщо $M(X) = 20, D(X) = 25$;

г) $P(1 < X < 5)$, якщо $M(X) = 3$, $\sqrt{D(X)} = 4$.

32. За яким законом розподілена випадкова величина X та чому дорівнюють її математичне сподівання та дисперсія, якщо щільність імовірності цієї величини має вигляд

$$\text{а) } f(x) = \begin{cases} 2e^{-2x}, & x \geq 0; \\ 0, & x < 0; \end{cases} \quad \text{б) } f(x) = \begin{cases} 0,1e^{-0,1x}, & x \geq 0; \\ 0, & x < 0; \end{cases}$$

$$\text{в) } f(x) = \begin{cases} 16e^{-16x}, & x \geq 0; \\ 0, & x < 0; \end{cases} \quad \text{г) } f(x) = \begin{cases} 25e^{-25x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

33. Неперервна випадкова величина X розподілена за законом, заданим диференціальною функцією

$$f(x) = \begin{cases} 3e^{-3x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Знайти імовірність того, що X потрапить в інтервал

$$\text{а) } (0,13; 0,7); \quad \text{б) } (0,8; 1,2); \quad \text{в) } \left(\frac{1}{3}; \frac{2}{3}\right); \quad \text{г) } (0,2; 1).$$

Відповідь:

$$\text{а) } P(0,13 < X < 0,7) = e^{-0,39} - e^{-2,1} = 0,677 - 0,122 = 0,555.$$

34. Знайти щільність імовірності функції $Y = aX + b$ від аргументу X , розподіленого нормально.

Відповідь:

Y розподілена нормально з параметрами $m_Y = am_X + b$, $\sigma_Y = |a|\sigma_X$, де m_X та σ_X – параметри нормального закону величини X .

35. Випадкова величина X рівномірно розподілена на інтервалі $(-1; 1)$. Знайти щільність імовірності випадкової величини $Y = X^2$.

Відповідь:

$$q(y) = \begin{cases} \frac{1}{4\sqrt{y}}, & 0 < y < 1; \\ 0, & \text{при } y \leq 0 \text{ та } y \geq 1. \end{cases}$$

36*. Закон розподілу похибок при вимірюванні радіуса r кола є нормальним з математичним сподіванням 50 см та дисперсією 0,25 см². Знайти закон розподілу похибок при обчисленні площі S кола радіуса r .

Відповідь:

$$g(s) = \begin{cases} \frac{1}{\pi\sqrt{2s}} \left[e^{-\frac{(\sqrt{s}-50\sqrt{\pi})^2}{0,5\pi}} + e^{-\frac{(\sqrt{s}+50\sqrt{\pi})^2}{0,5\pi}} \right], & s > 0; \\ 0, & s \leq 0. \end{cases}$$

37. Випадкова величина X розподілена рівномірно на $(0, 1)$. Знайти щільність імовірності випадкової величини $Y = X^3$.

Відповідь:

$$g(y) = \begin{cases} \left(3y^{\frac{2}{3}}\right)^{-1}, & 0 < y < 1; \\ 0, & \text{при } y \leq 0 \text{ та } y \geq 1. \end{cases}$$

38. Випадкова величина X розподілена нормально з параметрами $a = 2$ та $\sigma = 3$. Знайти щільність імовірності $Y = X^5$.

Відповідь:

$$g(y) = \frac{1}{15 \cdot \sqrt[10]{32\pi^5 y^8}} \cdot e^{-\frac{(\sqrt[5]{y}-2)^2}{18}}.$$

39*. Задані розподіли незалежних випадкових величин X та Y

X	-1	0	1
$P(X)$	0,3	0,5	0,2

Y	0	1
$P(Y)$	0,4	0,6

Розділ 4. Випадкові величини

Знайти математичне сподівання та дисперсію випадкової величини $Z = X^2 \cdot Y^3$, використовуючи властивість математичного сподівання $M(Z) = M(X^2 \cdot Y^3) = M(X^2)M(Y^3)$.

Відповідь: $M(Z) = 0,3$; $D(Z) = 0,21$.

40. Задана щільність імовірності випадкової величини X

$$f(x) = \begin{cases} \frac{1}{2} \cos x, & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}; \\ 0, & x \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \end{cases}$$

Знайти математичне сподівання та дисперсію випадкової величини $Y = \sin 2X$.

Відповідь: $M(Y) = 0$; $D(Y) = \frac{8}{15}$.

41. Випадкова величина X розподілена рівномірно на $(0, 2)$. Знайти дисперсію випадкової величини $Y = 3 - 2X$.

Відповідь: $\frac{3}{4}$.

42. Випадкова величина X розподілена нормально з параметрами $a = 1$ та $\sigma = 2$. Знайти щільність розподілу величини $Y = X^3$.

Відповідь:

$$g(y) = \begin{cases} \frac{1}{2\sqrt{2\pi}} \cdot e^{-\frac{(\sqrt[3]{y^2}-1)^2}{8}} \cdot \frac{1}{3y^{\frac{2}{3}}}, & y > 0; \\ 0, & y \leq 0. \end{cases}$$

43. Задані розподіли випадкових величин X та Y

X	3	5	7
$P(X)$	0,3	0,5	0,2

Y	2	6
$P(X)$	0,6	0,4

Скласти розподіл величини $Z = X + Y$ та знайти математичне сподівання Z .

Відповідь:

Z	5	7	9	11	13
$P(Z)$	0,18	0,3	0,24	0,2	0,08

$$M(Z)=8,4.$$

44. Випадкова величина X розподілена рівномірно на інтервалі $(1, 2)$. Знайти $M(3X^2 - 2X + 1)$ та $D(X^2 + 1)$.

Відповідь:

$$M(3X^2 + 2X + 1) = 5; \quad D(X^2 + 1) = \frac{34}{45}.$$

45. Кількість елементів комп'ютера, що псуються за деякий час, розподілено за законом Пуассона з параметром a . Час ремонту комп'ютера залежить від кількості m зіпсованих елементів і визначається формулою

$$t_m = T \cdot (1 - e^{-am}).$$

Визначити математичне сподівання часу ремонту та збитків від псування комп'ютера, якщо збитки пропорційні до квадрату часу ремонту, тобто $S = kt_m^2$.

Відповідь:

$$M(t_m) = T(1 - e^{-a(1-e^a)});$$

$$M(S) = kT^2 \left[1 - 2e^{-a(1-e^a)} + e^{-a(1-e^{2a})} \right]^2.$$

46. Похибка спостереження X при обчислюванні довжини розподілена нормально з параметрами $a = 5$ мм і $\sigma = 4$ мм. Знайти імовірність того, що значення, що вимірюється, відхиляється від дійсного більше ніж на 10 мм.

Відповідь:

$$P(|X| \geq 10) = 1 - P(|X| < 10) = 1 - \Phi(1,25) - \Phi(3,75) = 0,1057.$$

47. Відхилення розміру деталі від стандарту є випадкова величина, яка розподілена нормально з $M(X) = 0$ та $D(X) = 36$. Скільки потрібно виготовити деталей, щоб з імовірністю не менше 0,95 можна було стверджувати, що серед них буде хоч би одна стандартна, якщо дозволено відхилення розміру деталі від стандарту в інтервалі $(-2, 4)$.

Відповідь: $n \geq 7$.

48. Довести, що при центруванні та нормуванні нормально розподіленої випадкової величини X одержимо нормально розподілену випадкову величину з математичним сподіванням 0 та дисперсією 1.

Рекомендація: центрування та нормування випадкової величини еквівалентне функції $Y = \frac{X - a}{\sigma}$.

Відповідь: $f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$.

Частина II

**МАТЕМАТИЧНА
СТАТИСТИКА**

*Вимірйой все,
щой піддіється вимірйованню
та зроби таким усе,
щой не піддається вимірйованню.*

Галілео Галілей

Розділ 5

ОСНОВНІ ПОНЯТТЯ ТА СТАТИСТИЧНИЙ РОЗПОДІЛ

*Бажання та увага
завжди допомагають
засвоїти початок
будь-якої справи.*

5.1. Предмет математичної статистики та коротка історична довідка

Мета кожного наукового дослідження – виявлення закономірностей явищ, які спостерігають, та використання цих закономірностей у повсякденній практичній діяльності. Для встановлення цих закономірностей проводять спеціальні дослідження та спостерігають одиничні явища. Далі роблять узагальнений висновок у вигляді закону.

У тих випадках, коли явище знаходиться під дією багатьох факторів і неможливо виявити вплив усіх цих факторів, застосовують інший метод вивчення – **статистичний**, тобто систематизація та обробка статистичних даних однорідних дослідів.

Звичайним є використання статистичних методів в економіці, соціології, політології.

Нехай, наприклад, темп приросту промислового виробництва за певний період часу дорівнює 5%. Це означає, що в середньому для усієї сукупності підприємств показник 5% є статистичною закономірністю зростання промислового виробництва. Цей середній показник не виключає, а, навпаки, припускає, що на окремих підприємствах темп приросту може бути більше або менше 5%.

Предмет математичної статистики полягає в розробці методів збору та обробки статистичних даних для одержання наукових та практичних висновків.

Вкажемо основні задачі, які розв'язує математична статистика:

- 1) вказати способи збору та групування (якщо даних дуже багато) статистичних відомостей;
- 2) визначити закон розподілу випадкової величини або системи випадкових величин за статистичними даними;
- 3) визначити невідомі параметри розподілу;
- 4) перевірити правдоподібність припущень про закон розподілу випадкової величини, про форму зв'язку між випадковими величинами або про значення параметра, який оцінюють.

Можна сказати, що основна задача математичної статистики – розробка методів аналізу статистичних даних в залежності від мети дослідження.

Методи математичної статистики ефективно використовують при розв'язанні багатьох задач науки, організації технологічного процесу, планування, управління та ціноутворення.

Математична статистика виникла (XVII ст.) та почала розвиватись паралельно з теорією імовірностей. Подальшим розвитком (кінець XIX – початок XX ст.) математична статистика зобов'язана П.Л.Чебишову, А.А.Маркову, О.М.Ляпунову, а також К.Гауссу, Ф.Гальгону, К.Пірсону та іншим.

У XX ст. найбільший вклад у математичну статистику зробили В.І.Романовський, Е.Е.Слущкий, А.Н.Колмогоров, Ст'юдент (псевдонім У.Госсета), Е.Пірсон, Ю.Нейман, А.Вальд, А.В.Скороход, В.С.Корольок та інші вчені.

5.2. Генеральна та вибіркова сукупності

Нехай потрібно вивчити сукупність об'єктів відносно деякої якісної або кількісної ознаки, які характеризують ці об'єкти. Кожен об'єкт, який спостерігають, має декілька ознак. Розглядаючи лише одну ознаку кожного об'єкта, ми припускаємо, що інші ознаки рівноправні, або що множина **об'єктів однорідна**.

Такі множини однорідних об'єктів називають **статистичною сукупністю**.

Наприклад, якщо досліджують партію деталей, то якісною ознакою може бути стандартність або нестандартність кожної деталі, а

Розділ 5. Основні поняття та статистичний розподіл

кількісною ознакою – розмір деталі. Кількісні ознаки бувають **неперервними** та **дискретними**.

Перевірку сукупності деталей можна провести двома способами:

- 1) провести перевірку (контроль) усіх деталей;
- 2) перевірити лише певну частину деталей.

Якщо деталей дуже багато або перевірка пов'язана з руйнуванням деталі (наприклад, випробування деталі на міцність), тоді перший спосіб перевірки не доцільний. Якщо дослідити усі деталі неможливо, тоді відбирають із усієї сукупності обмежене число деталей і перевіряють лише їх.

Вибірковою сукупністю (вибіркою) називають сукупність випадково взятих об'єктів.

Генеральною називають сукупність об'єктів, з яких зроблено вибірку.

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

Наприклад, якщо з 5000 виробів для дослідження взято 50, тоді об'єм генеральної сукупності $N = 5000$, а об'єм вибірки $n = 50$.

Наведемо приклади вибірок.

Первинним результатом статичного спостереження є перелік членів сукупності та відповідних їм значень ознаки.

Приклад 1.

Спостерігають величину врожаю пшениці на 10 дослідницьких ділянках. Результати спостережень представлені у Таблиці 1 (ознака N – номер ділянки, ознака X – врожай у центнерах з га)

Таблиця 1

N	1	2	3	4	5	6	7	8	9	10
X	15,2	19,1	17,1	20,8	18,4	16,4	22	20,4	17,6	18,2

Таке зведення називають **рядом варіант** або **простим статистичним рядом**.

Вибірки бувають **повторні** та **безповторні**. **Повторною** називають вибірку, при якій відібраний об'єкт повертається до генеральної сукупності перед відбором іншого об'єкта. Вибірку називають

безповторною, якщо взятий об'єкт до генеральної сукупності не повертається. Найчастіше використовують безповторні вибірки.

Альтернативою вибірки є перепис. **Переписами** називають обстеження, що мають своєю метою дослідження кожного елементу сукупності (генеральної сукупності), що вивчається.

Зразками перепису є перепис населення у країні, звіт про виробничі показники усіх підприємств у галузі (наприклад, шахт вугільної промисловості).

Перевага вивчення вибірки порівняно з переписом: малі затрати коштів, обладнання та часу.

Вибірку можна ефективно використовувати для вивчення відповідної ознаки генеральної сукупності лише тоді, коли дані вибірки вірно відображають цю ознаку. Коротко ця умова формулюється таким чином – вибірка повинна бути **репрезентативною**, тобто **представницькою**.

Згідно із законом великих чисел теорії імовірностей можна стверджувати, що вибірка буде репрезентативною лише тоді, коли її здійснюють випадково.

В більшості випадків для математичної статистики найбільш підходящим засобом здійснення випадкового відбору є **простий випадковий**.

Означення 1. *Простим випадковим є такий відбір з генеральної сукупності, при якому кожний об'єкт, що витягається, має однакову імовірність потрапити до вибірки.*

Вибірка, що здійснена за допомогою **простого випадкового** відбору, називається **простою випадковою**. Способи здійснення простого випадкового відбору розглянуто в параграфі 5.5.

Варто відмітити, що альтернативою для простої випадкової вибірки у статистиці є **розширована випадкова вибірка**.

Способи здійснення відбору розглянуто в параграфі 5.4.

5.3. Джерела даних у статистиці

Дослідники і менеджери отримують дані, необхідні для прийняття рішення, в основному з трьох джерел:

Розділ 5. Основні поняття та статистичний розподіл

Вибіркові обстеження, спеціально поставлені експерименти і дані, що є результатом повсякденної (рутинної) роботи у бізнесі.

Розглянемо приклади застосування вищезгаданих джерел.

При вибірковому обстеженні засобами збирання даних вибірки можуть бути індивідуальні опитування (інтерв'ю), опитування по пошті, телефонні інтерв'ю, і таке інше. Способи організації вибірки описані у параграфі 5.4. Наведемо приклад вибірки.

Приклад 1. Видавництво газети вибирає 1000 потенційних виборців для опитування з метою вивчення рейтингу певного кандидата на виборах.

При спеціально спланованому експерименті у дослідника є можливість, у певних рамках, управління процесом. Наведемо приклад планування експерименту і використання його як джерела даних у статистиці.

Приклад 2. У одній з японських фірм розроблений наступний бланк бальної оцінки здібностей керівника [11] (див. Таблицю 2).

Зібравши дані про кожного керівного співробітника у вигляді Таблиці 2 згідно тестуванню, керівник фірми може використати ці дані для об'єктивної оцінки роботи керівного складу, для оцінки впливу реформи системи управління фірмою, що впроваджується, на прибуток і так далі.

Найчастіше джерелом даних є дані, що збираються у повсякденній, рутинній роботі у бізнесі. Наведемо приклади.

Приклад 3. Керівник крамниці, аналізуючи дані рівня продажу «Вид товару – сезон року», може більш оптимально планувати свою роботу для отримання найбільшого рівня прибутку за рахунок збільшення об'єму продажу ходових видів товару, зменшення витрат на створення надлишків запасів товару на складах крамниці і так далі.

Другим прикладом джерела такого роду даних є різноманітні офіційні джерела статистичних даних.

Приклад 4.

Книга: Народне господарство України у 1994 році. Статистичний щорічник України: Відповідальний за випуск В.В.Самченко – К.: Техніка, 1994, 494 ст.

Таблиця 2

Оцінка здібностей керівника
(для чергового підвищення заробітної плати)

1. Потенціал (можливості удосконалення здібностей і результатів праці)	Гнучкість мислення, активність, наявність потенціалу внутрішнього росту. Постійне прагнення до вдосконалення, незадоволення досягнутим. Бажання прийняти на себе більш високу відповідальність.	3–15
2. Лідерство; мотивація підлеглих	Хороший контакт з підлеглими	5–25
3. Результативність роботи: прогноз на майбутній рік	Відношення до планів продажу, прибутку	4–20
4. Вміння вести переговори і взаємодіяти з партнерами	Вміння добре говорити і (або) слухати. Прагнення зрозуміти інших. Спокійна манера мови, витриманість. Здібність до взаємодії з партнерами.	3–15
5. Креативність (спроможність до творчості у майбутньому)	Постійне прагнення до розв'язання складних проблем. Творчий підхід до розв'язання проблем.	2–10
6. Знання і поінформованість	Забезпечення необхідною інформацією. Вміння працювати за планом, вчасно реагуючи на зміни. Вміння оцінювати та прогнозувати ситуацію	3–15
	Всього	20–100

Джерела даних бувають *первинними і вторинними*.

Первинні дані збираються спеціально для статистичного дослідження. Для цих даних є відомості про методи збирання, точність даних і так далі.

Вторинними даними є дані, що використовуються у статистиці, але спочатку збиралися для інших цілей.

Очевидно, що рутинні записи про діяльність фірм, офіційні статистичні звіти є вторинними даними.

Безумовно, більш цінними даними у статистиці є первинні дані, але їх не завжди можливо отримати, тому часто використовуються вторинні дані.

5.4. Способи відбору

У практичній діяльності використовують різноманітні способи відбору об'єктів із генеральної сукупності. Усі способи відбору можна поділити на два види:

1. Вибір, який не потребує розділення генеральної сукупності на частини. До цього виду вибору відносять:

- простий випадковий неповторний відбір;
- простий випадковий повторний відбір.

2. Вибір, при якому генеральна сукупність розділяється на частини (розшарований випадковий відбір). До цього виду вибору відносять:

- типовий відбір;
- механічний відбір;
- серійний відбір.

Типовим називають відбір, при якому об'єкти відбирають не з усієї генеральної сукупності, а лише з її типових частин. Наприклад, якщо вироби виготовлені на різних станках, то відбір проводять лише з виробів кожного станка окремо.

Типовий відбір доцільно використовувати тоді, коли однакові вироби виготовляють на станках, серед яких є більш та менш досконалі, або у випадку виготовлення однакових виробів різними підприємствами.

Механічним називають відбір, при якому генеральна сукупність механічно поділяється на стільки частин, скільки має бути об'єктів у вибірці. З кожної частини випадковим чином відбирають один об'єкт. Наприклад, якщо потрібно перевірити 25% усіх виготовлених станком-автоматом виробів, то відбирають кожен четвертий виріб. Щоб механічний відбір був репрезентативним, треба враховувати специфіку технологічного процесу.

Серійним називають відбір, при якому об'єкти із генеральної сукупності відбирають не по одному, а серіями, які і досліджують. Серійний відбір використовують тоді, коли ознака, яку досліджують, мало змінюється в різних серіях.

В економічних дослідженнях іноді використовують **комбінований відбір**. Наприклад, спочатку поділяють генеральну сукупність на серії однакового об'єму, випадковим чином відбирають декілька серій і, нарешті, з кожної серії випадковим чином беруть окремі об'єкти.

5.5. Проста випадкова вибірка

У цьому розділі детально розглянуто питання побудови простої випадкової вибірки за допомогою таблиці випадкових чисел. Розв'язання цієї же задачі за допомогою електронної таблиці Excel 97 розглянуто у параграфі 8.2 розділу 8.

5.5.1. Умови здійснення простої випадкової вибірки

Для здійснення простої випадкової вибірки необхідна наявність **основи вибірки**, тобто такого представлення генеральної сукупності, при якому її елементи були хоча б перераховані. Наведемо приклади основ вибірки.

Приклад 1.

а) Генеральна сукупність – всі покупці крамниці. Основою вибірки можуть бути робочі списки покупців, що веде крамниця.

б) Генеральна сукупність – всі жителі міста, що мають телефон. Основою вибірки може бути довідкова телефонна книга.

Як правило, дані для утворення випадкової вибірки представляються у вигляді деякої, заздалегідь складеної таблиці і тому основою вибірки є нумерація елементів цієї таблиці.

Основа вибірки повинна повністю відбивати признак генеральної сукупності, що вивчається. Порушення цієї вимоги може зробити вибірку не репрезентативною. Пояснимо це на прикладі.

Приклад 2. Вимагається обстежити всі молоді сім'ї невеликого міста на предмет кількості дітей дошкільного віку у сім'ях.

З цією метою співробітник відділу, що займається цим питанням, випадковим чином із допомогою телефонного довідника обдзвонює сім'ї з 18.00 до 21.00 кожний день.

Чи буде вибірка даних, отримана таким чином, репрезентативною?

Основою вибірки у цьому прикладі є телефонний довідник. Проста випадкова вибірка, що складена за цією основою, не буде репрезентативною по цілому ряду причин: не всі сім'ї міста, що обстежуються, мають телефон, деякі сім'ї у цей період не будуть знаходитися вдома, або члени цих сімей не зможуть підійти до телефону, деякі сім'ї користуються телефонами, номери яких не записані в довідник і так далі.

Наведемо приклади, де може бути використана проста випадкова вибірка.

Приклад 3.

а) Телефонна компанія перевіряє рахунки 10% всіх міжнародних телефонних переговорів з метою визначення їхньої середньої величини.

б) Аудиторська перевірка накладних 20% фірм регіону з метою контролю правильності сплати податків.

5.5.2. Випадкові вибірккові числа

Загальновідомо, що найкращим способом здійснення простої вибірки є використання випадкових вибірккових чисел.

Ці числа складаються з цифр від 0 до 9, генеруються випадковим чином (як правило, за допомогою комп'ютера) і записуються до спеціальної таблиці.

Вираз «генеруються випадковим чином» відбиває той факт, що шанс появи будь-якої цифри у будь-якому місці таблиці не більше й не менше шансу появи будь-якої іншої цифри з десяти названих цифр.

У Додатку 4 наведена типова таблиця випадкових чисел (Таблиця 8), що згруповані; для зручності при читанні, у блоки по п'ять цифр.

Використання таблиць випадкових чисел гарантує, що не буде зроблена систематична помилка (тобто помилка, що робить дані не репрезентативними).

У наступному підрозділі розглянуто докладно питання про застосування таблиці випадкових чисел для здійснення простої випадкової вибірки.

5.5.3. Здійснення простої випадкової вибірки з використанням випадкових чисел

Опишемо процедуру одержання простої випадкової вибірки на наступному прикладі.

Приклад 4. У Таблиці А наведено результати тестування 180 співробітників фірми по певній методиці з метою прийняття рішення про чергове підвищення зарплати. У ній перераховані 180 двозначних цифр – число балів, що дістав кожний, хто обстежувався. Необхідно виробити просту випадкову вибірку по Таблиці А.

Побудова простої випадкової вибірки здійснюється в такій послідовності.

1. Генеральна сукупність – дані Таблиці А. Основа вибірки – нумерація елементів таблиці. Об'єм генеральної сукупності $N = 180$.

2. Задаємося числом n елементів простої випадкової вибірки, наприклад, прийmemo $n = 10$.

3. Для забезпечення випадковості відбору скористуємось Таблицею 8 додатку 4. Виберемо в ній будь-який рядок або стовпець, наприклад, почнемо з шостого ряду. Представимо тут його частину

14704|55808|05863|72452|68374|02760|79464|36170|19001... (1)

Таблиця А

Число балів, набране при тестуванні 180 співробітників фірми

001.	30	031.	48	061.	34	091.	46	121.	47	151.	36
002.	38	032.	35	062.	25	092.	30	122.	64	152.	25
003.	32	033.	26	063.	61	093.	64	123.	55	153.	28
004.	49	034.	62	064.	42	094.	28	124.	50	154.	56
005.	33	035.	51	065.	48	095.	64	125.	65	155.	51
006.	42	036.	67	066.	57	096.	46	126.	32	156.	53
007.	60	037.	30	067.	26	097.	59	127.	53	157.	40
008.	30	038.	57	068.	55	098.	60	128.	44	158.	33
009.	34	039.	50	069.	36	099.	46	129.	38	159.	26
010.	61	040.	62	070.	33	100.	27	130.	36	160.	42
011.	49	041.	59	071.	63	101.	59	131.	53	161.	62
012.	64	042.	56	072.	48	102.	43	132.	44	162.	29
013.	63	043.	65	073.	37	103.	50	133.	27	163.	37
014.	37	044.	50	074.	49	104.	51	134.	40	164.	27
015.	25	045.	54	075.	46	105.	39	135.	43	165.	36
016.	38	046.	61	076.	31	106.	59	136.	45	166.	43
017.	65	047.	57	077.	26	107.	33	137.	33	167.	30
018.	56	048.	55	078.	28	108.	60	138.	60	168.	59
019.	55	049.	26	079.	63	109.	26	139.	62	169.	41
020.	43	050.	41	080.	37	110.	72	140.	30	170.	63
021.	68	051.	64	081.	65	111.	25	141.	51	171.	55
022.	38	052.	25	082.	42	112.	44	142.	49	172.	32
023.	71	053.	28	083.	73	113.	58	143.	31	173.	32
024.	47	054.	49	084.	44	114.	49	144.	29	174.	35
025.	65	055.	53	085.	54	115.	35	145.	36	175.	33
026.	54	056.	25	086.	67	116.	57	146.	50	176.	58
027.	74	057.	33	087.	49	117.	37	147.	54	177.	58
028.	36	058.	28	088.	38	118.	66	148.	38	178.	31
029.	62	059.	25	089.	59	119.	55	149.	60	179.	38
030.	31	060.	60	090.	42	120.	66	150.	65	180.	29

Ми будемо вибирати з цього ряду числа, що утворені трьома цифрами, оскільки нумерація елементів Таблиці А не перевищує числа 180.

З цієї причини зручніше ряд (1) переписати по три цифри в групі

$$\underline{147}|045|580|805|863|724|526|837|402|760|794|643|617|019|001\dots \quad (2)$$

4. Вибираємо з ряду (2) лише ті числа, що входять в нашу основну вибірку, тобто всі числа більші 180 ігноруються. (В ряді (2) існує чотири таких числа – вони підкреслені).

Відібравши таким чином 10 випадкових чисел, ми здійснимо просту випадкову вибірку, наведену в Таблиці 3.

Таблиця 3

Проста випадкова вибірка об'єму $n=10$,
що отримана з Таблиці А результатів тестування

Випадкове число	Номер числа у вибірці	Число балів
147	1	54
045	2	54
019	3	55
001	4	30
102	5	43
003	6	32
111	7	25
011	8	49
159	9	26
013	10	63

В цій таблиці кожному випадковому числу (перший стовпець) ставиться у відповідність число балів (третій стовпець) згідно Таблиці А. Для вигоди користування всі елементи простої випадкової вибірки пронумеровано (другий стовпець).

5.6. Організація даних: статистичний розподіл вибірки

5.6.1. Впорядкування даних

Дані у статистиці, отримані за допомогою спеціальних досліджень або із звичайних робочих (рутинних) записів у бізнесі, надходять до дослідника чи менеджера у вигляді неорганізованої маси, незалежно від того, чи є вони вибірковими даними, чи даними з генеральної сукупності.

У математичній статистиці замість слова «дані» вживається термін «варіанти». Числову характеристику варіанти при цьому називають **ознакою**.

Нехай із генеральної сукупності взята вибірка об'єктів $\{x_1, x_2, \dots, x_n\}$ об'єму n , для вивчення ознаки X . Тобто, значення $x_1, x_2, \dots, x_n \in$ *варіанти ознаки X* .

Першим кроком обробки є впорядкування варіант. Розглянемо цей процес на прикладі.

Приклад 1. У Таблиці В наведена вибірка середньомісячної платні 100 співробітників фірми N . Треба впорядкувати вибірку.

Таблиця В

Вибірка середньомісячної платні 100 співробітників фірми N

338	348	304	314	326	314	324	304	342	308
336	304	302	338	314	304	320	321	322	321
312	323	336	324	312	312	364	356	362	302
322	310	334	292	362	381	304	366	298	304
381	368	304	298	368	290	340	328	316	322
302	314	292	342	321	322	290	332	298	296
296	298	324	338	352	326	318	304	332	322
360	312	331	331	304	316	332	282	342	338
342	322	324	325	302	328	354	330	316	324
334	350	334	324	332	340	324	314	326	323

У нашому прикладі ознака є число, що виражає середньомісячну зарплату співробітників фірми *N*. Отже, у Таблиці В наведено 100 значень варіант.

Розмістимо дані Таблиці Б у порядку зростання (див. Таблицю В.1).

Таблиця В.1

Впорядкована вибірка середньомісячної платні
100 співробітників фірми (у порядку зростання)

282	298	304	314	321	323	326	332	340	356
290	302	304	314	321	324	326	334	340	360
290	302	304	314	321	324	328	334	342	362
292	302	304	314	322	324	328	334	342	362
292	302	308	314	322	324	330	336	342	364
296	304	310	316	322	324	331	336	342	366
296	304	312	316	322	324	331	338	348	368
298	304	312	316	322	324	332	338	350	368
298	304	312	318	322	325	332	338	352	381
298	304	312	320	323	326	332	338	354	381

Варіанти, що записані до таблиці у зростаючому (спадаючому) порядку, називають **варіаційним рядом**. Тобто, Таблиця В.1 є варіаційним рядом, що налічує 100 варіант.

Після упорядкування можна отримати більше інформації, наприклад, про межі зміни середньомісячної платні.

5.6.2. Розподіл частот

Нехай у нашій вибірці із n варіант x_1, x_2, \dots, x_m ознака X прийняла значення $x_1 - n_1$ раз, значення $x_2 - n_2$ раз, ..., значення $x_m - n_m$ раз.

Додатне число, що вказує, скільки раз та чи інша варіанта зустрічається в таблиці даних, називається **частотою**.

Ряд

$$n_1, n_2, \dots, n_m$$

Розділ 5. Основні поняття та статистичний розподіл

називається **рядом частот**. Відмітимо, що сума усіх частот повинна дорівнювати об'єму вибірки

$$\sum_{i=1}^m n_i = n. \quad (1)$$

Статистичний розподіл вибірки встановлює зв'язок між рядом варіант, що зростає або спадає, і відповідними частотами. Він може бути представлений таблицею

x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

де n – об'єм вибірки, $n = n_1 + n_2 + \dots + n_m$.

Статистичний розподіл вибірки, заданий цією таблицею, також називають **простим** чи **незгрупованим статистичним розподілом** або **розподілом частоти варіанти x_i (рядом розподілу частоти варіанти x_i)**.

Приклад 2. Для вивчення потреб у певних розмірах взуття власник магазину спостерігає розміри взуття, проданого на протязі дня: 40, 35, 37, 39, 40, 41, 36, 42, 40, 39, 36, 43, 43, 41, 38, 37, 36, 42, 40, 38.

Статистичний розподіл цієї вибірки (розподіл частоти розміру взуття) буде мати такий вигляд

Розмір взуття x_i	35	36	37	38	39	40	41	42	43
Кількість n_i	1	3	2	2	2	4	2	2	2

Контроль: $1 + 3 + 2 + 2 + 2 + 4 + 2 + 2 + 2 = 20$, $n = 20$.

Зауваження. Властивість елементів ряду частот (див. формулу (1)) використовують для контролю одержаного статистичного розподілу вибірки.

Наведемо подальші удосконалення вибірки, що фігурує у Прикладі 1 (Таблиця В.1), перетворивши її в розподіл частот середньомісячної платні співробітників фірми N (Таблиця В.2).

Таблиця В.2

Розподіл частоти середньомісячної платні
співробітників фірми N

x_i	n_i	x_i	n_i	x_i	n_i	x_i	n_i
282	1	314	5	328	2	350	1
290	2	316	3	330	1	352	1
292	2	318	1	331	2	354	1
296	2	320	1	332	4	356	1
298	4	321	3	334	3	360	1
302	4	322	6	336	2	362	2
304	9	323	2	338	4	364	1
308	1	324	7	340	2	366	1
310	1	325	1	342	4	368	2
312	4	326	3	348	1	381	2
30		32		25		13	
Разом:							100

Подальший крок в обробці даних, що призводить до суттєвого спрощення дослідження, є їх *згрупування*.

Як видно з Таблиці В.2 максимальне та мінімальне значення варіанти будуть

$$x_{\max} = 381, \quad x_{\min} = 282.$$

Різниця цих чисел

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

називається **розмахом варіант**. У нашому випадку $R = 99$.

Введемо для варіанти інтервали зміни платні.

$$\begin{array}{cccc}
 280-290 & 310-320 & 340-350 & 370-380 \\
 290-300 & 320-330 & 350-360 & 380-390 \\
 300-310 & 330-340 & 360-370 &
 \end{array} \quad (1)$$

Кожний інтервал називається **класом інтервалів** або **класом**. Всього маємо $k = 11$ класів платні.

Використовуючи дані Таблиці В.2, просумуємо частоти для кожного класу інтервалів (1), причому значення x_i , що знаходяться на границі класів, заносимо до того класу, що є слідуючим до

Розділ 5. Основні поняття та статистичний розподіл

класу, де це число зустрілось вперше. Результат перепишемо її у вигляді Таблиці В.3.

Таблиця В.3

Згрупований розподіл частоти середньомісячної платні співробітників фірми N

Інтервали платні	Частоти n_i	
280–290	1	(=1)
290–300	10	(=2+2+2+4)
300–310	14	(=4+9+1)
310–320	14	(=1+4+5+3+1)
320–330	25	(=1+3+6+2+7+1+3+2)
330–340	16	(=1+2+4+3+2+4)
340–350	7	(=2+4+1)
350–360	4	(=1+1+1+1)
360–370	7	(=1+2+1+1+2)
370–380	0	(=0)
380–390	2	(=2)
Разом:	100	

Таблиця вигляду В.3, яка встановлює зв'язок між згрупованим рядом варіант, що зростає або спадає, та сумами їхніх частот по класах, називається **згрупованим розподілом частоти варіанти** x_i .

Для кожного класу маємо верхню та нижню границі, наприклад, для першого і другого класу інтервалів маємо

$$x_{1\min} = 280, \quad x_{1\max} = 290$$

$$x_{2\min} = 290, \quad x_{2\max} = 300.$$

Шириною класу h_i називається число одиниць виміру у цьому класі, тобто різниця

$$h_i = x_{i\max} - x_{i\min}.$$

У нашому випадку ширина класів однакова і дорівнює $h = 10$.

Введені величини R – розмах варіант, k – число класів, h – ширина класу пов'язані співвідношенням

$$h = \frac{R}{k}$$

Наведемо інший приклад згрупованого розподілу частот вибірки. Так, наприклад, згрупування даних Таблиці 1 (див. стор. 155) по інтервалах врожайності дозволяє одержати таблицю

Інтервали (ц/га)	15–17	17–19	19–21	21–23	Усього
Кількість ділянок	2	4	3	1	10

5.6.3. Згрупований розподіл накопиченої частоти

Часто поряд з розподілом частоти варіанти необхідно мати розподіл **накопиченої (кумулятивної) частоти**. Розподіл накопиченої частоти одержується послідовним додаванням частот чергового інтервалу, починаючи з першого і кінчаючи останнім (див. Таблицю В.4).

Розподіл накопиченої частоти (позначається F_i) дозволяє відповісти на питання: «Скільки існує варіант, що менше ніж, наприклад, 350?». З Таблиці В.4 знаходимо: таких варіант 87. Це можна записати так

$$F_{350} = 87.$$

Таблиця В.4

Згрупований розподіл накопиченої частоти середньомісячної платні співробітників фірми N

Інтервали платні	Частоти n_i	Платня		Накопичені частоти F_i	
280–290	1	менше ніж	290	1	(=1)
290–300	10	менше ніж	300	11	(=1+10)
300–310	14	менше ніж	310	25	(=11+14)
310–320	14	менше ніж	320	39	(=25+14)
320–330	25	менше ніж	330	64	(=39+25)
330–340	16	менше ніж	340	80	(=64+16)
340–350	7	менше ніж	350	87	(=80+7)
350–360	4	менше ніж	360	91	(=87+4)
360–370	7	менше ніж	370	98	(=91+7)
370–380	0	менше ніж	380	98	(=98+0)
380–390	2	менше ніж	390	100	(=98+2)
Разом:	100				

5.6.4. Розподіл відносної частоти (частоти) вибірки

Нерідко замість значень частот використовуються відносні частоти. Нехай існує m частот n_1, \dots, n_m ($n_1 + \dots + n_m = n$). Відношення частоти n_i варіанти x_i до об'єму вибірки n

$$W_i = \frac{n_i}{n}$$

називається **відносною частотою** або **частотою**, причому, сума усіх відносних частот

$$\sum_{i=1}^m W_i = 1. \quad (2)$$

Залежність між упорядкованим рядом варіант і відповідними їм відносними частотами або частотями також називається **статистичним розподілом вибірки**, тобто маємо табличне представлення розподілу

x_i	x_1	x_2	...	x_m
W_i	$\frac{n_1}{n}$	$\frac{n_2}{n}$...	$\frac{n_m}{n}$

де n – об'єм вибірки і $x_1 < x_2 < \dots < x_m$.

Приклад 3. Заданий розподіл частот вибірки

x_i	4	6	9
n_i	10	3	7

Знайти розподіл частостей.

Розв'язання. Об'єм вибірки $n = 10 + 3 + 7 = 20$. Частостями будуть

$$W_1 = \frac{10}{20} = 0,5; \quad W_2 = \frac{3}{20} = 0,15; \quad W_3 = \frac{7}{20} = 0,35.$$

Тому розподіл частостей цієї вибірки буде

x_i	4	6	9
W_i	0,5	0,15	0,35

Зауваження. Властивість елементів ряду частостей (дивись формулу (2)) використовують для контролю одержаного статистичного розподілу вибірки.

Проведемо подальшу організацію даних вибірки, що задана Таблицею В.

Розглянемо Таблицю В4. Якщо поділити частоти (другий стовпчик) та накопичені частоти (четвертий стовпчик) на об'єм вибірки $n = 100$, то дістанемо відповідно **розподіл відносної частоти** W_i

та **накопиченої відносної частоти** $\frac{F_i}{n}$ вибірки (див. Таблицю В.5).

Таблиця В.5

Згруповані розподіли відносної та накопиченої відносної частот середньомісячної платні співробітників фірми N

Інтервали платні	Відносні частоти $W_i = n_i / n$	Платня	Накопичені відносні частоти F_i/n
280–290	0.01	менше ніж 290	0.01
290–300	0.10	менше ніж 300	0.11
300–310	0.14	менше ніж 310	0.25
310–320	0.14	менше ніж 320	0.39
320–330	0.25	менше ніж 330	0.64
330–340	0.16	менше ніж 340	0.80
340–350	0.07	менше ніж 350	0.87
350–360	0.04	менше ніж 360	0.91
360–370	0.07	менше ніж 370	0.98
370–380	0.00	менше ніж 380	0.98
380–390	0.02	менше ніж 390	1
Разом:	100		

Розділ 5. Основні поняття та статистичний розподіл

Розподіл накопиченої відносної частоти одержується послідовним додаванням відносних частот чергового інтервалу, починаючи з першого і кінчаючи останнім.

Розподіл накопиченої відносної частоти дозволяє відповісти на питання: «Яка пропорція варіант, що менше ніж, наприклад, 350?».

З Таблиці В.5 знаходимо: пропорція цих варіант становить 0,88, тобто це є доля середньомісячної платні, що менше 350 грн.

5.6.5. Ряди розподілу частоти із змінною шириною класів інтервалів

Інколи неможливо, або небажано вибирати ширину класів інтервалів однаковою. Нерівна ширина класів бажана, наприклад, коли значення частоти одного чи декількох класів набагато більше (менше) значень частот інших інтервалів. Як правило, ширина інтервалів зростає (або спадає) і може містити інтервали відкритого типу «більше ніж...», «менше ніж...».

Приклад розподілу частот із зростаючою шириною інтервалів наведений у Таблиці 4.

Таблиця 4

Розподіл частоти віку 30 співробітників фірми *М*

Вік співробітників (інтервали у роках)	Число співробітників
17–18	2
18–20	3
20–25	15
25–35	7
35 та старше	3
Разом:	30

5.6.6. Згрупований розподіл щільності частоти і щільності відносної частоти (частоти)

Якщо поділити всі частоти (другий стовпчик) Таблиці В.3 на ширину інтервалу $h = 10$, то отримуємо **розподіл щільності частоти вибірки**

$$\frac{n_i}{h}.$$

Якщо поділити всі відносні частоти (другий стовпчик) Таблиці В.5 на ширину інтервалу $h = 10$, то отримуємо **розподіл щільності відносної частоти (частоти)** вибірки

$$\frac{W_i}{h}.$$

Щоб підсумувати результати, які одержані для Прикладу 1, зведемо разом в одну Таблицю В.6 розглянуті раніше (див. Таблиці В.4 та В.5) і знову введені розподіли.

Таблиця В.6

Згрупований розподіл частоти (n_i), відносної частоти (W_i), щільності частоти (n_i/h), щільності відносної частоти (W_i/h), накопиченої частоти (F_i) і накопиченої відносної частоти (F_i/n) вибірки середньомісячної платні 100 співробітників фірми N

x	n_i	W_i	$\frac{n_i}{h}$	$\frac{W_i}{h}$	x	F_i	$\frac{F_i}{n}$
280–290	1	0,01	0,1	0,001	менше ніж 290	1	0,01
290–300	10	0,10	1,0	0,010	менше ніж 300	11	0,11
300–310	14	0,14	1,4	0,014	менше ніж 310	25	0,25
310–320	14	0,14	1,4	0,014	менше ніж 320	39	0,39
320–330	25	0,25	2,5	0,025	менше ніж 330	64	0,64
330–340	16	0,16	1,6	0,016	менше ніж 340	80	0,80
340–350	7	0,07	0,7	0,007	менше ніж 350	87	0,87
350–360	4	0,04	0,4	0,004	менше ніж 360	91	0,91
360–370	7	0,07	0,7	0,007	менше ніж 370	98	0,98
370–380	0	0,00	0,0	0,000	менше ніж 380	98	0,98
380–390	2	0,02	0,2	0,002	менше ніж 390	100	1,00
Разом:	100		1,00				

Ця таблиця знайомить з усіма важливими статистичними розподілами вибірки.

Розподіли частоти та відносної частоти будуть використані у побудові полігонів частот та гістограм (див. параграф 5.8).

Розподіли накопиченої частоти та накопиченої відносної частоти будуть використані у побудові полігону накопичених частот та емпіричної функції розподілу (див. параграф 5.8).

Введені поняття розподілу щільності частоти та розподілу щільності відносної частоти (частоти) вибірки мають глибокий імовірнісний зміст і будуть також використані у графічному представленні розподілів (див. пункт 5.8.3 параграфу 5.8).

5.6.7. Загальна схема побудови згрупованого розподілу частот

Викладений в підпункті 5.6.2 алгоритм згрупування даних вибірки можна представити у вигляді наступної загальної схеми послідовності дій:

1. Визначити найбільше x_{\max} та найменше x_{\min} значення варіанти x_i і визначити розмах варіант $R = x_{\max} - x_{\min}$.

2. Задатися певним числом класів k . Число класів k слід приймати непарним, і при загальному числі замірів $n \geq 100$ доцільно брати $9 \leq k \leq 15$, а при $n \leq 100$ можна вважати $k = 7$.

3. Визначити ширину класу $h = R/k$. Для спрощення розрахунків, отримане значення h округлити у будь-яку сторону.

4. Встановити границі класів і підрахувати кількість варіант у кожному класі. При підрахунку числа варіант значення x_i , що знаходиться на границі класів, слід відносити завжди до одного й того ж класу, наприклад, до того класу, що є слідуючим до класу, де це число зустрілось вперше. Воно, таким чином, стає нижньою границею класу.

5. Визначити частоту для кожного класу і записати ряд розподілу.

5.7. Емпірична функція розподілу та її властивості

Нехай є статистичний розподіл частот деякої ознаки X . Позначимо через n загальну кількість спостережень, тобто об'єм вибірки; n_x – кількість спостережень, при яких спостерігались ознаки X менше x .

Тоді відносна частота (або частість) події $X < x$ дорівнює $\frac{n_x}{n}$.

Якщо x змінюється, то може змінюватись відносна частота, тобто $\frac{n_x}{n}$ є функція від x . Ця функція знаходиться емпіричним (дослідним) шляхом, тому її називають **емпіричною**.

Означення 1. *Емпіричною функцією розподілу (або функцією розподілу вибірки) називають функцію $F^*(x)$, яка визначає для кожного значення x частість події $X < x$.*

Математично це означення має вигляд

$$F^*(x) = \frac{n_x}{n},$$

де n_x – кількість варіант, які менше від x , n – об'єм вибірки.

Таким чином, щоб знайти, наприклад, $F^*(x_3)$, треба кількість варіант, що менше x_3 , поділити на об'єм вибірки, тобто

$$F^*(x_3) = \frac{n_1 + n_2}{n}.$$

Зауваження. *Інтегральну функцію розподілу $F(x)$ генеральної сукупності у математичній статистиці називають теоретичною функцією розподілу. Вона відрізняється від емпіричної функції розподілу $F^*(x)$ тим, що визначає імовірність події $X < x$, а не частість цієї події.*

З теореми Бернуллі випливає, що частість

$$F^*(x) = \frac{n_x}{n} \quad \text{події } X < x$$

Розділ 5. Основні поняття та статистичний розподіл

прямує до імовірності

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

цієї події. Тому $F(x)$ та $F^*(x)$ мало відрізняються одна від одної.

Доцільно використовувати $F^*(x)$ для наближеного представлення функції розподілу $F(x)$ генеральної сукупності.

Емпірична функція розподілу $F^*(x)$ має такі властивості:

1. $0 \leq F^*(x) \leq 1$.
2. $F^*(x)$ – зростаюча функція.

$$3. F^*(x) = \begin{cases} 0, & x \leq x_1; \\ 1, & x > x_m, \end{cases}$$

де x_1 – найменша варіанта, x_m – найбільша варіанта.

Приклад 1. Знайти емпіричну функцію розподілу за статистичним розподілом вибірки

Таблиця 5

x_i	2	6	10
n_i	12	18	30

та побудувати її графік.

Розв'язаним. Об'єм цієї вибірки буде $n = 12 + 18 + 30 = 60$. Найменша варіанта дорівнює 2, тому $F^*(x) = 0$ для $x \leq 2$. Найбільша варіанта дорівнює 10, тому $F^*(x) = 1$ для $x > 10$. Значення $x < 6$,

тобто $X = \{x_1 = 2\}$, спостерігалось 12 разів, тому $F^*(x) = \frac{12}{60} = 0,2$

при $2 < x \leq 6$.

Значення $X < 10$, тобто $X = \{x_1 = 2\}$ та $X = \{x_2 = 6\}$, спостерігались $12 + 18 = 30$ разів, тому $F^*(x) = \frac{30}{60} = 0,5$ при $6 < x \leq 10$.

Тобто, простий статистичний розподіл частоти, що заданий Таблицею 5, замінюється згрупованим розподілом частоти (див. Таблицю 6).

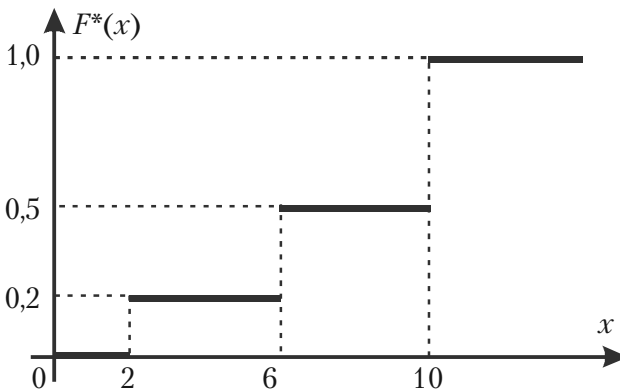
Таблиця 6

Варіанта x_i	Частота n_i	Варіанта x_i	Накопичені частоти F_i
$x \leq 2$	0	менше ніж 2	0
$2 < x \leq 6$	12	менше ніж 6	12
$6 < x \leq 10$	18	менше ніж 10	30
$10 < x$	30	більше ніж 10	60
Разом:	60		

Тут же побудований розподіл накопиченої частоти. Таким чином, одержали емпіричну функцію розподілу вигляду

$$F^*(x) = \begin{cases} 0, & x \leq 2; \\ 0,2, & 2 < x \leq 6; \\ 0,5, & 6 < x \leq 10; \\ 1, & x > 10. \end{cases}$$

Графік цієї функції зображено на мал. 18.



Мал. 18.

Цей графік можна розуміти як наближений графік теоретичної функції розподілу $F(x)$.

Встановимо зв'язок між емпіричною функцією розподілу $F^*(x)$ і функцією накопичених частот F_i . Як слідує з Означення 1, Прикладу 1 і Таблиці 6 емпірична функція розподілу $F^*(x)$ визначається як кусково-постійна функція, рівна значенню накопиченої відносної частоти

$$F^*(x) = \frac{n_x}{n} = \frac{F_i}{n},$$

на кожному класі інтервалів

$$J_i = (x_{i\min}, x_{i\max}) \quad i = 1, 2, \dots, k.$$

5.8. Графічне зображення статистичних розподілів

Усі статистичні розподіли, що вивчались у параграфі 5.6, можуть бути представлені графічно. Завдяки цьому ми можемо побачити характерні зміни ряду розподілу, не користуючись аналізом цифрових даних.

Графічне зображення статистичних розподілів потрібно розглядати окремо для незгрупованих та згрупованих даних.

5.8.1. Незгруповані дані: полігони частот та частостей, гістограма

Якщо в результаті вибірки ми одержали статистичний розподіл ознаки X , яку треба дослідити, то будемо мати перелік варіант ознаки

$$x_1, x_2, \dots, x_m$$

та відповідних їм частот

$$n_1, n_2, \dots, n_m$$

або частостей

$$W_1, W_2, \dots, W_m.$$

Значення варіант та частот або частостей можна розглядати як координати точок

$$M_1(x_1, n_1), M_2(x_2, n_2), M_m(x_m, n_m)$$

або

$$M_1(x_1, W_1), M_2(x_2, W_2), M_m(x_m, W_m).$$

Означення 1. *Полігоном частот називають ламану, відрізки якої з'єднують точки*

$$(x_1, n_1), (x_2, n_2), \dots, (x_m, n_m).$$

Полігоном відносних частот (частостей) називають ламану, відрізки якої проходять через точки

$$(x_1, W_1), (x_2, W_2), \dots, (x_m, W_m).$$

Полігони частот та частостей є аналогами щільності імовірностей.

Для побудови полігону частот на осі абсцис відкладають варіанти x_k ознаки X , а на осі ординат - відповідні їм частоти. Точки (x_k, n_k) ($k = 1, 2, \dots, m$) з'єднують відрізками прямих і одержують полігон частот.

Для побудови полігону відносних частот (частостей) на осі ординат відкладають частості W_k , ($k = 1, 2, \dots, m$), а потім точки (x_k, W_k) з'єднують відрізками прямих.

Приклад 1. У результаті вибірки одержали такі значення ознаки X

$$-3, 2, -1, -3, 5, -3, 2.$$

Побудувати полігон частот цієї вибірки.

Розв'язання. У цьому випадку варіантами будуть

$$x_1 = -3, x_2 = -1, x_3 = 2, x_4 = 5.$$

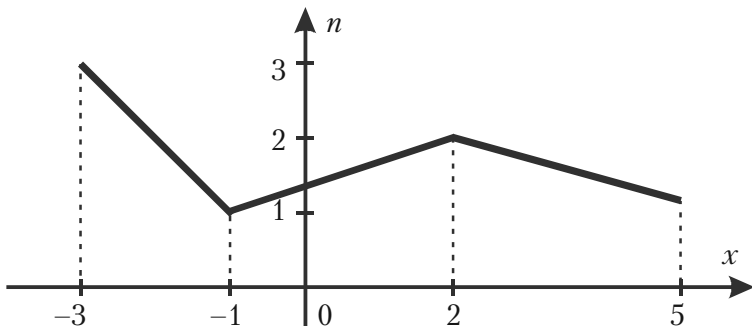
Відповідні їм частоти

$$n_1 = 3, n_2 = 1, n_3 = 2, n_4 = 1.$$

Відклавши у системі координат (xOy) точки

$$M_1(-3, 3), M_2(-1, 1), M_3(2, 2), M_4(5, 1)$$

та з'єднавши їх відрізками прямих, одержимо полігон частот цієї вибірки (див. мал. 19).



Мал. 19.

Означення 2. *Гістограмою частот* називають ступінчасту фігуру, яка складається з прямокутників, основами яких є часткові інтервали варіант довжиною $h = x_k - x_{k-1}$, а висоти дорівнюють $\frac{n_k}{h}$ (щільність частоти).

Гістограмою відносних частот (частостей) називають ступінчасту фігуру, яка складається з прямокутників, основами яких є часткові інтервали варіант, а висоти дорівнюють відношенню $\frac{W_k}{h}$ (щільність частості).

Площа гістограми частот дорівнює об'єму вибірки, а площа гістограми частостей – одиниці.

Для побудови гістограми частот (частостей) проміжок варіант $[x_{\min}, x_{\max}]$, тобто від найменшого значення x_{\min} , що спостерігали, до найбільшого значення x_{\max} , розбивають на декілька відрізків рівної довжини h . Потім підраховують суму частот (частостей) тих

значень варіант ознаки X , які належать кожному із одержаних відрізків.

Якщо на k -ому відрізку ($k = 1, 2, \dots, m$) кількість варіант, що спостерігали, з врахуванням їх частот, дорівнює n_k , то будують прямокутник Π_k , основою якого буде k -тий відрізок довжиною h , а висотою буде $\frac{n_k}{h}$ (для частостей висота – $\frac{W_k}{h}$).

Площа такого прямокутника дорівнює $h \cdot \frac{n_k}{h} = n_k$ (у випадку частостей $h \cdot \frac{W_k}{h} = W_k$). Тому площа усіх прямокутників буде дорівнювати сумі $n_k(W_k)$, тобто $S = \sum_{k=1}^m n_k = n$ – об'єм вибірки.

У випадку гістограми частостей площа прямокутників буде дорівнювати сумі усіх частостей

$$S = \sum_{k=1}^m W_k = \sum_{k=1}^m \frac{n_k}{n} = \frac{1}{n} \sum_{k=1}^m n_k = \frac{1}{n} \cdot n = 1.$$

Приклад 2. У результаті спостереження одержали розподіл ознаки X вибірки у вигляді

Варіанти x_k	-2	0	1	2	3	5	7
Частоти n_k	4	5	7	8	6	2	1

Побудувати гістограму частот цього розподілу.

Розв'язання. У даному випадку найменше значення варіанти $x_{\min} = -2$, а найбільше значення – $x_{\max} = 7$, тому довжина проміжка $[x_{\min}, x_{\max}]$ дорівнює 9. Розіб'ємо цей відрізок на 4 рівних

частини довжиною $h = \frac{9}{4} = 2,25$.

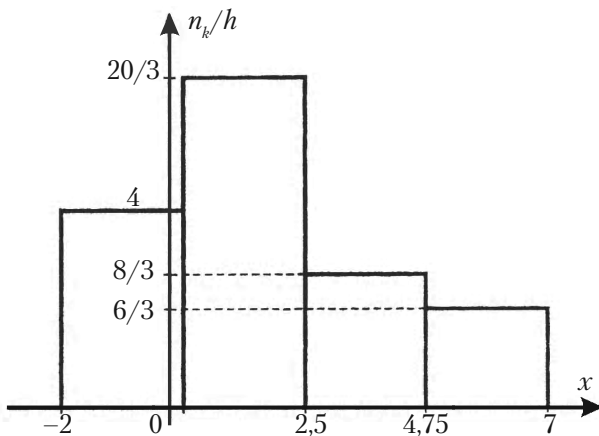
Розділ 5. Основні поняття та статистичний розподіл

Для побудови гістограми доцільно скласти Таблицю 7: у перший рядок таблиці записують одержані відрізки, у другий рядок таблиці записують суму частот варіант, що належать відповідному відрізку, а у третій рядок записують висоти відповідних прямокутників. У даному випадку ця таблиця буде мати вигляд

Таблиця 7

Відрізки довжини $h = 2,25$	$[-2; 0,25]$	$[0,25; 2,5]$	$[2,5; 4,75]$	$[4,75; 7]$
Кількість значень n_k	9	15	6	3
Щільність частоти $\frac{n_k}{h}$	4	$\frac{20}{3}$	$\frac{8}{3}$	$\frac{4}{3}$

Відмітимо що
відрізку $[-2; 0,25]$ належать 9 значень варіант (-2 чотири рази та 0 п'ять разів);
відрізку $[0,25; 2,5]$ належать 15 значень варіант (1 – сім разів та 2 – вісім разів);
відрізку $[2,5; 4,75]$ належать 6 значень варіант (3 – шість разів);
відрізку $[4,75; 7]$ належать 3 значення (5 – 2 рази та 7 – один раз).
За даними Таблиці 7 будемо відповідну гістограму частот (див. мал. 20).



Мал. 20.

Зауваження 1. Іноді для побудови гістограми перші два рядки розподілу вигляду Таблиці 7 вже задані. У цьому випадку треба підрахувати елементи третього рядка таблиці і побудувати відповідну гістограму.

Приклад 3. Побудувати гістограму частостей заданого розподілу

Інтервали $h = 5$	Частоти варіант
5–10	4
10–15	6
15–20	16
20–25	36
25–30	24
30–35	10
35–40	4

Розв'язання. У даному випадку об'єм вибірки

$$n = 4 + 6 + 16 + 36 + 24 + 10 + 4 = 100,$$

довжина часткового інтервалу варіант $h = 5$.

За формулою $W_k = \frac{n_k}{n}$ знаходимо частоти кожного інтервалу

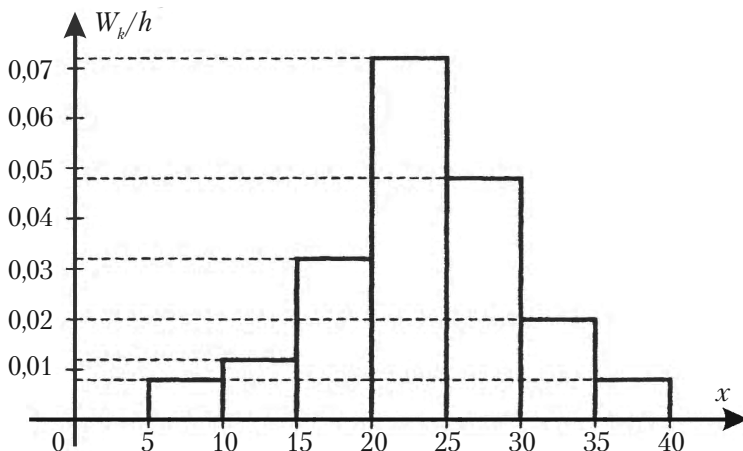
$$\begin{aligned} W_1 &= 0,04, & W_2 &= 0,06, & W_3 &= 0,16, & W_4 &= 0,36, \\ W_5 &= 0,24, & W_6 &= 0,1, & W_7 &= 0,04. \end{aligned}$$

Для побудови гістограми частостей за Означенням 2, знайдемо щільності частоти

$$\frac{W_1}{h} = 0,008, \quad \frac{W_2}{h} = 0,012, \quad \frac{W_3}{h} = 0,032, \quad \frac{W_4}{h} = 0,072,$$

$$\frac{W_5}{h} = 0,048, \quad \frac{W_6}{h} = 0,02, \quad \frac{W_7}{h} = 0,008.$$

Отже, потрібна гістограма частостей прийме вигляд, що зображений на мал. 21.



Мал. 21.

Зауваження 2. При дослідженні практичних проблем часто вибірка має значну кількість варіант. Розглянемо у наступному прикладі спосіб їх обробки.

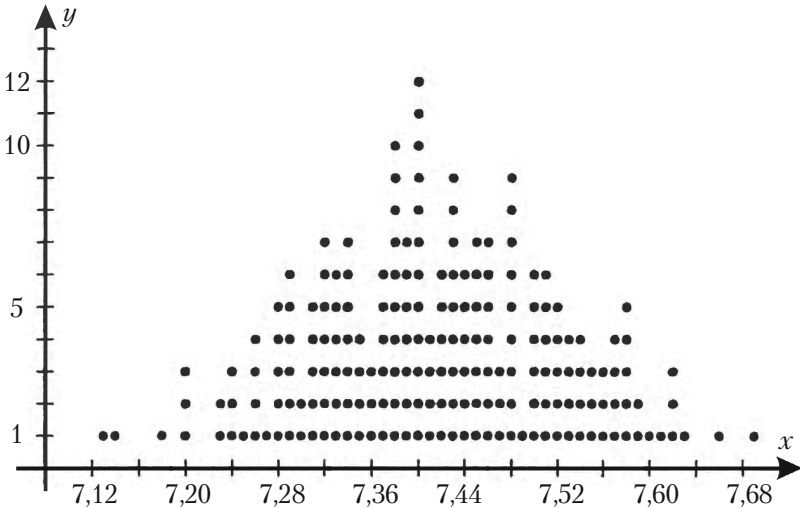
Приклад 4. Контрольні виміри радіуса 200 циліндрів дали наступні результати

7,39;	7,43;	7,54;	7,64;	7,40;	7,55;	7,40;	7,26;	7,42;	7,50;
7,32;	7,31;	7,28;	7,52;	7,46;	7,63;	7,38;	7,44;	7,52;	7,53;
7,37;	7,33;	7,24;	7,13;	7,53;	7,53;	7,39;	7,57;	7,51;	7,34;
7,39;	7,47;	7,51;	7,48;	7,62;	7,58;	7,57;	7,33;	7,51;	7,40;
7,30;	7,48;	7,40;	7,57;	7,51;	7,40;	7,52;	7,56;	7,40;	7,34;
7,23;	7,37;	7,48;	7,48;	7,62;	7,35;	7,36;	7,40;	7,45;	7,29;
7,48;	7,58;	7,44;	7,56;	7,28;	7,59;	7,47;	7,62;	7,54;	7,20;
7,38;	7,43;	7,35;	7,56;	7,51;	7,47;	7,40;	7,29;	7,20;	7,46;
7,42;	7,44;	7,41;	7,29;	7,48;	7,39;	7,50;	7,38;	7,45;	7,50;
7,45;	7,42;	7,29;	7,53;	7,34;	7,55;	7,33;	7,32;	7,69;	7,46;
7,32;	7,46;	7,32;	7,48;	7,38;	7,43;	7,51;	7,43;	7,60;	7,44;
7,25;	7,29;	7,31;	7,45;	7,43;	7,44;	7,31;	7,58;	7,28;	7,24;
7,34;	7,49;	7,50;	7,38;	7,48;	7,43;	7,37;	7,29;	7,54;	7,33;
7,36;	7,46;	7,23;	7,44;	7,38;	7,27;	7,52;	7,40;	7,26;	7,66;
7,59;	7,48;	7,46;	7,40;	7,24;	7,41;	7,34;	7,43;	7,38;	7,50;
7,26;	7,43;	7,37;	7,55;	7,42;	7,41;	7,38;	7,14;	7,42;	7,52;
7,46;	7,39;	7,35;	7,32;	7,18;	7,30;	7,54;	7,38;	7,37;	7,34;
7,50;	7,61;	7,42;	7,32;	7,35;	7,40;	7,57;	7,31;	7,40;	7,36;
7,28;	7,58;	7,38;	7,58;	7,26;	7,37;	7,28;	7,39;	7,32;	7,20;
7,43;	7,34;	7,45;	7,33;	7,41;	7,33;	7,45;	7,31;	7,45;	7,39.

Побудувати гістограму розподілу частот вибірки.

Розв'язання. У заданому вигляді варіанти не придатні для висновку контролю. Ці варіанти треба якось упорядкувати. Найпростіший спосіб упорядкування – представити варіанти графічно: на осі абсцис відкласти результати вимірів, а над кожним значенням варіанти відкласти стільки точок, скільки разів зустрічається ця варіанта.

У результаті ми одержимо точкову діаграму, яка до-зволяє скласти деяке представлення про вибірку і може бути покладена в основу подальшої обробки результатів контролю (див. мал. 22).



Мал. 22.

Щоб спростити побудову гістограми, в таблиці розподілу вказують не відрізки, в яких згруповані результати вимірів, а їх середини. Ці середини визначають відрізки і віддалені одна від одної на відстань h , яка дорівнює довжині відрізків.

Дійсно, якщо довжина відрізка дорівнює h , а його серединою є число a , то відрізок буде

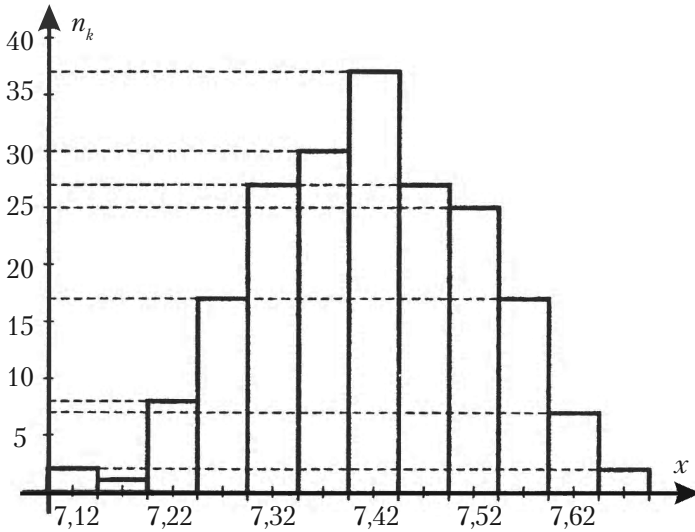
$$\left[a - \frac{h}{2}, a + \frac{h}{2} \right].$$

Для побудови гістограми частот вибірки із мал. 22 при $h = 0,05$ одержимо таблицю вигляду

Розділ 5. Основні поняття та статистичний розподіл

x_k	7,12	7,17	7,22	7,27	7,32	7,37
n_k	2	1	8	17	27	30
x_k	7,42	7,47	7,52	7,57	7,62	7,67
n_k	37	27	25	17	7	2

За даними таблиці будемо гістограму частот вибірки.

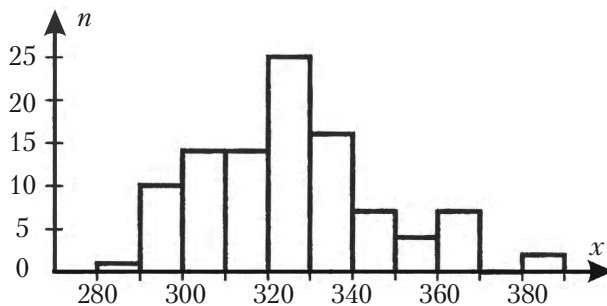


Мал. 23.

5.8.2. Згруповані дані: гістограма і полігон частот

Для згрупованого розподілу частоти, відносної частоти, щільності частоти і щільності відносної частоти можуть бути побудовані *спеціальні діаграми*, складені з прямокутників ступінчасті фігури, що називаються **гістограмами**.

Для побудови гістограми на горизонтальну вісь наносяться класи інтервалів. На кожному класі будується прямокутник, висота якого рівна значенню частоти (або відносної частоти, або щільності частоти, або щільності відносної частоти) на цьому інтервалі.



Мал. 24. Гістограма частот середньомісячної зарплати 100 співробітників фірми N

Таблиця 8

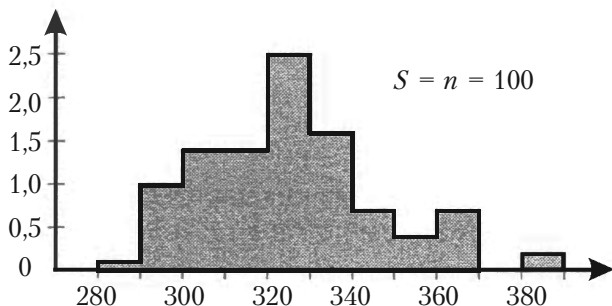
Згрупований розподіл частоти, відносної частоти, щільності частоти та щільності відносної частоти

x_i	n_i	W_i	$\frac{n_i}{h}$	$\frac{W_i}{h}$
280–290	1	0,01	0,01	0,001
290–300	10	0,10	1,0	0,010
300–310	14	0,14	1,4	0,014
310–320	14	0,14	1,4	0,014
320–330	25	0,25	2,5	0,025
330–340	16	0,16	1,6	0,016
340–350	7	0,07	0,7	0,007
350–360	4	0,04	0,4	0,004
360–370	7	0,07	0,7	0,007
370–380	0	0,00	0,0	0,000
380–390	2	0,02	0,2	0,002
Разом:	100	1,00		

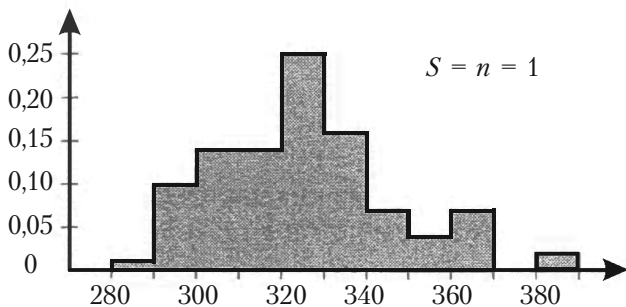
Розділ 5. Основні поняття та статистичний розподіл

На мал. 24 зображена гістограма частот згідно даних Таблиці 8, що побудована по Таблиці В.6 параграфу 5.6.

Далі зобразимо заштриховані контури гістограм для щільності частоти і відносної щільності частоти згідно даних Таблиці 8 (див. малюнки 25 та 26).



Мал. 25. Контур гістограми для щільності частоти n/h . Площа гістограми 5 рівна об'єму вибірки $S = n = 100$

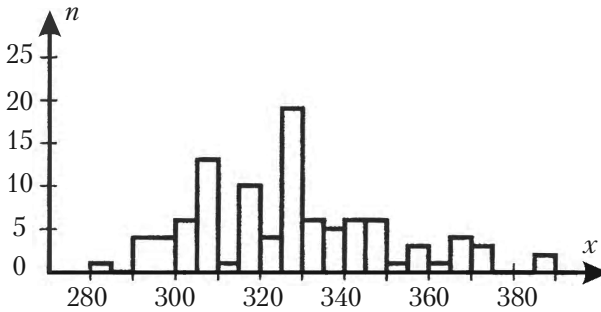


Мал. 26. Контур гістограми для щільності відносної частоти W/h . Площа гістограми рівна 1.

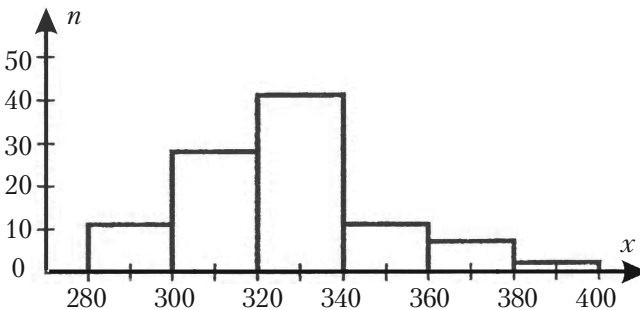
Порівнюючи мал. 24, 25, 26 легко бачити, що всі три гістограми геометрично подібні і відрізняються лише масштабом вертикальної осі.

Від способу обрання ширини класу інтервалів h залежить виразність гістограми. Якщо h досить мале, то гістограма містить багато випадкового. Якщо h досить велике, то в гістограмі зникають індивідуальні особливості вибірки.

На мал. 27 та 28 зображені гістограми для Прикладу 1 при значеннях ширини класів інтервалів $h = 5$ та $h = 20$ згідно відповідних розподілів частоти, що наведено у Таблиці 9.



Мал. 27. Гістограма частот середньомісячної зарплати 100 співробітників фірми N (ширина класів інтервалів $h = 5$)



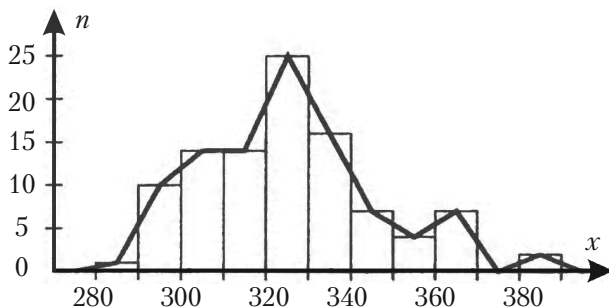
Мал. 28. Гістограма частот середньомісячної зарплати 100 співробітників фірми N (ширина класів інтервалів $h = 20$)

Таблиця 9

Згрупований розподіл частоти середньомісячної платні співробітників фірми N для двох значень ширини класів інтервалів $h = 5$ та $h = 20$

Ширина класів $h = 5$				Ширина класів $h = 20$	
Інтервали платні	Частоти	Інтервали платні	Частоти	Інтервали платні	Частоти
280–285	1	335–340	6	280–300	11
285–290	0	340–345	6	300–320	28
290–295	4	345–350	1	320–340	41
295–300	6	350–355	3	340–360	11
300–305	13	355–360	1	360–380	7
305–310	1	360–365	4	380–400	2
310–315	10	365–370	3	100	
315–320	4	370–375	0		
320–325	19	375–380	0		
325–330	6	380–385	2		
330–335	10	385–390	0		
Разом:	74		26		

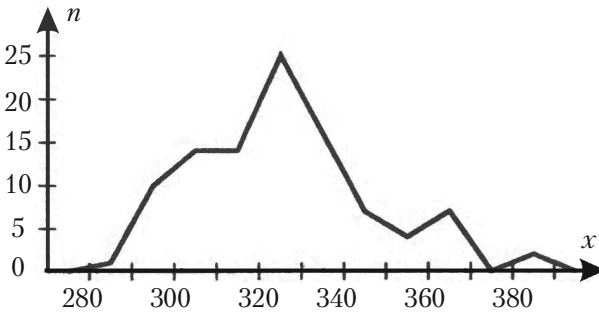
Альтернативою гістограми для розподілу частот є **полігон частот**. Для побудови цього графіка над серединою кожного інтервалу варіант відкладається крапка на висоті, відповідній частоті цього інтервалу. Після цього ці крапки сполучаються відрізками прямих (див. мал. 29).



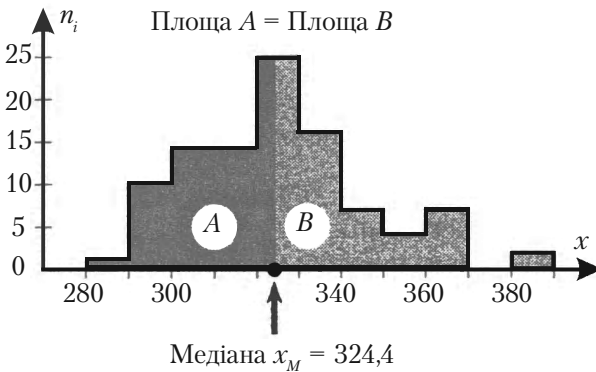
Мал. 29. Полігон частот середньомісячної платні 100 співробітників фірми N

На графіку, де розміщена гістограма частот (мал. 29), полігон частот зображений так, що кінцеві точки графіку торкаються горизонтальної осі у точках 275 та 395.

Суміщення двох типів графіків на одному малюнку зроблено з метою підкреслити різний геометричний зміст цих двох графічних зображень розподілу частот. Звичайно ж полігон частот зображується на окремому кресленні, як це показано на мал. 30.



Мал. 30. Полігон частот середньомісячної платні 100 співробітників фірми N



Мал. 31. Медіана згрупованого розподілу частот, заданого Таблицею 8

З допомогою гістограми зручно визначити наступне важливе поняття.

Означення 3. *Медіаною згрупованого розподілу частот (при постійній ширині класів інтервалів) називається значення або точка на горизонтальній осі гістограми розподілу частот така, в якій перпендикулярна лінія, що проходить через неї, поділяє площу гістограми на дві рівні частини (див. мал. 31).*

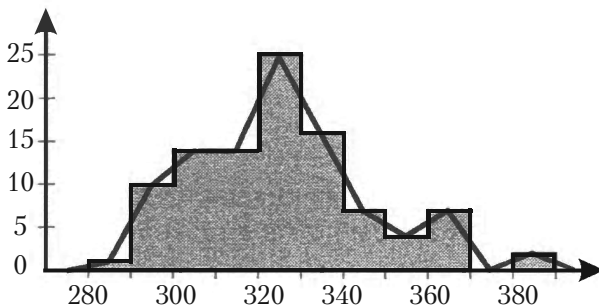
Алгоритм знаходження медіани буде розглянуто в кінці цього розділу.

5.8.3. Імовірнісний зміст гістограми та полігону частот

Щоб виявити, який імовірнісний зміст гістограми і полігона частот, зобразимо на одному малюнку (див. мал. 32) контур гістограми для щільності відносної частоти W/h та полігон щільності відносних частот.

Площа фігури, обмежена полігоном щільності відносних частот приблизно дорівнює площі контура діаграми, що в свою чергу дорівнює 1. Отже, полігон щільності відносних частот є приблизним зображенням функції щільності імовірності генеральної сукупності.

Якщо збільшувати об'єм вибірки n , то полігон щільності відносних частот буде більш точно зображати функцію щільності імовірності генеральної сукупності.



Мал. 32. Контур гістограми для щільності відносної частоти W/h та полігон щільності відносних частот. Площа гістограми дорівнює 1.0

5.8.4. Полігони накопичених частот та частостей. Знаходження їх медіани

Для накопиченої частоти і накопиченої відносної частоти можуть бути побудовані графіки схожі на полігон частот. Ці графіки називаються *полігоном накопиченої частоти* або *полігоном накопиченої відносної частоти*. У статистиці також їх називають *огівою* або *кумулятивною кривою*. Полігон накопиченої частоти зручно використовувати у цілому ряді задач статистики.

Розглянемо побудову полігона накопиченої відносної частоти на прикладі розподілу накопиченої відносної частоти вибірки середньомісячної платні 100 співробітників фірми N (див. Таблицю 10, яку було побудовано за даними Таблиці В.6 стор.189).

Таблиця 10

Згруповані розподіли накопиченої частоти і накопиченої відносної частоти вибірки середньомісячної платні 100 співробітників фірми N

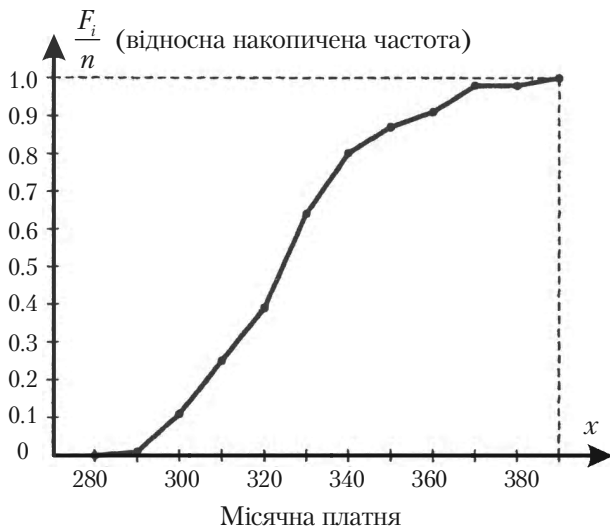
x_i	F_i	$\frac{F_i}{n}$
менше ніж 290	1	0,01
менше ніж 300	11	0,11
менше ніж 310	25	0,25
менше ніж 320	39	0,39
менше ніж 330	64	0,64
менше ніж 340	80	0,80
менше ніж 350	87	0,87
менше ніж 360	91	0,91
менше ніж 370	98	0,98
менше ніж 380	98	0,98
менше ніж 390	100	1,00

Для побудови полігону ми ставимо точку над нижньою границею кожного класу інтервалів на горизонтальній осі на висоті, що відповідає величині накопиченої відносної частоти на попередніх інтервалах.

Розділ 5. Основні поняття та статистичний розподіл

Так на верхній границі $x_{1\max} = 280$ інтервалу «270–280» відкладена точка на висоті 0, бо накопичена відносна частота на попередніх інтервалах рівна 0 (див. 3-й стовпчик Таблиці 10).

На верхній границі $x_{2\max} = 290$ інтервалу «280–290» відкладена точка на висоті 0,01, бо накопичена відносна частота на попередніх інтервалах рівна 0,01 (див. 3-й стовпчик Таблиці 10) і так далі. Остаточний вигляд побудованого полігону накопиченої відносної частоти наведено на мал. 33.



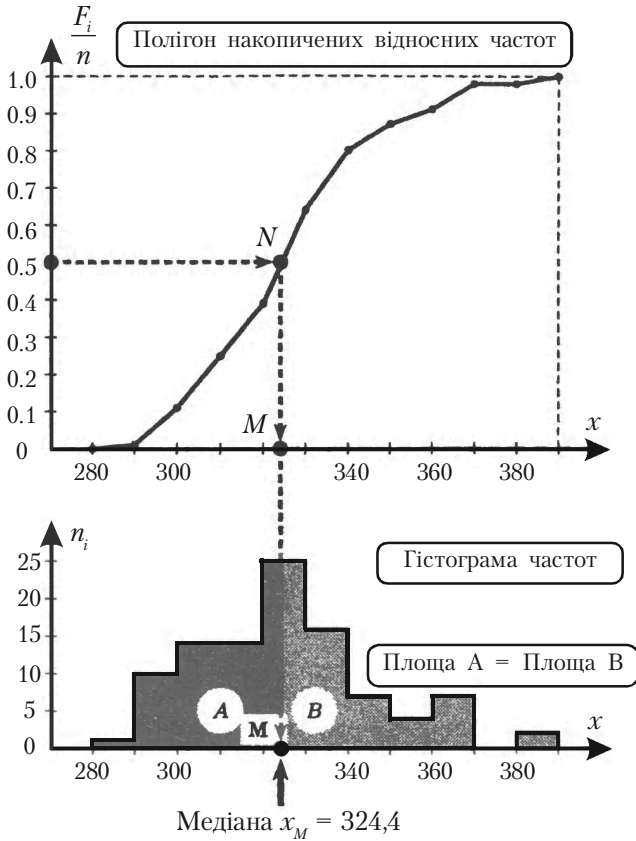
Мал. 33. Полігон накопиченої відносної частоти

В заключення параграфу розглянемо, як буде утворюватися медіана згрупованого розподілу частоти вибірки. Існують два засоби: **формульний** і **графічний**.

Розглянемо графічний засіб побудови медіани згрупованого розподілу частоти вибірки.

Звернемось до мал. 34. У верхній частині малюнку зображений полігон накопичених відносних частот, а у нижній частині – гістограма частот.

З точки із ординатою 0,5 на осі $\frac{F_i}{n}$ проводимо горизонтальну лінію до перетину з ламаною прямою полігону накопичених відносних частот у точці N .



Мал. 34. Графічна побудова медіани групованого розподілу частот за допомогою полігона накопичених відносних частот

Далі, з точки N проводимо вертикальну лінію до перетинання осі Ox графіка гістограми частот в точці M , що і є медіаною. По графіку на мал. 34 можна лише наближено визначити значення координати точки M : $x_M \approx 324$.

Розглянемо формульний засіб побудови медіани згрупованого розподілу частоти вибірки.

Точне значення x_M можна знайти по наступній інтерполяційній формулі. Нехай медіана належить інтервалу $J_m = (L_m, L_{m+1})$, тоді

$$x_M = L_m + \left[\frac{0.5 - F_{m-1}}{W_m} \right] h, \quad (1)$$

де L_m – нижня границя інтервалу J_m , F_{m-1} – значення функції накопичених відносних частот на попередньому інтервалі $m - 1$, W_m – відносна частота варіанти в інтервалі J_m , h – ширина інтервалу.

Інтервал J_m це той клас інтервалів, що містить число $n/2$, де n – об'єм вибірки. В нашому випадку $n/2 = 50$. Звернемося до Таблиці В.4 (стор. 185) і скористуємось функцією накопиченої частоти.

З таблиці видно, що числу 50 відповідає інтервал $J_{320} = (L_{320}, L_{330})$.

Отже, маємо

$$L_m = 320, \quad L_{m+1} = 330, \quad F_{m-1} = 0,39, \quad W_m = 0,25, \quad h = 10.$$

Після підстановки цих значень в формулу (1) одержуємо

$$x_M = 320 + \frac{0,5 - 0,39}{0,25} 10 = 320 + \frac{0,11}{0,25} 10 = 324,4.$$

Порівнюючи два засоби побудови медіани згуртованого розподілу частоти вибірки формульний і графічний, приходимо до наступного висновку.

Графічне знаходження медіани згрупованого розподілу частот більш наочне, але дає лише наближене значення медіани і вказує той клас інтервалів, де знаходиться медіана. Цією обставиною можна скористуватись для знаходження її точного значення за інтерполяційною формулою (1), що наведена вище.

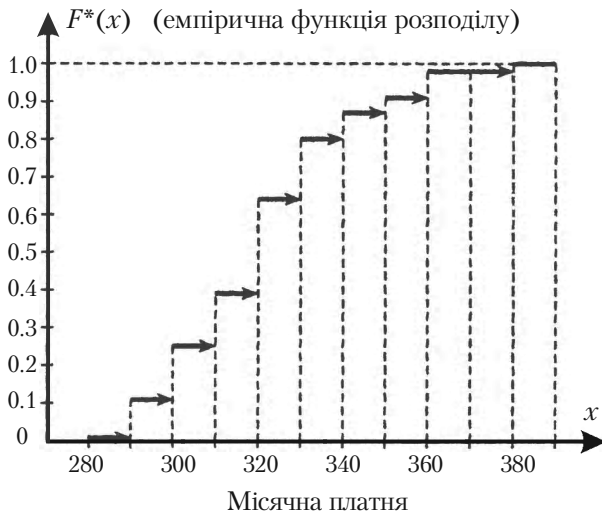
5.8.5. Графік емпіричної функції розподілу

Користуючись даними Таблиці 9, запишемо емпіричну функцію розподілу $F^*(x)$ (стор. 198) для нашого прикладу (теорію див. параграф 5.7).

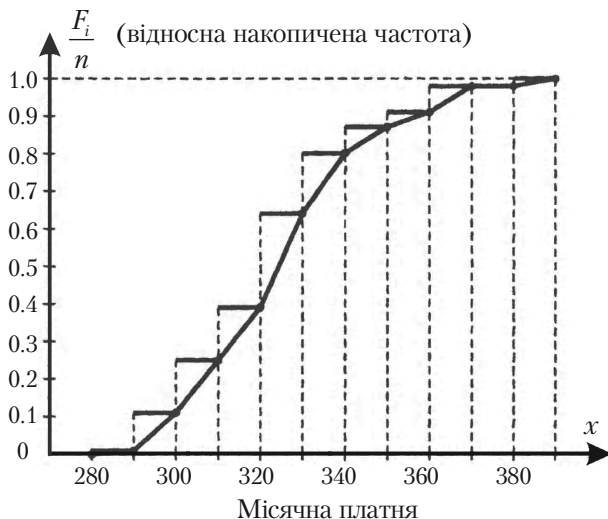
Очевидно, що функція $F^*(x)$ визначає відносну частоту події «Значення варіанти менше x », тобто $X < x$.

Графік емпіричної функції розподілу $F^*(x)$ (див. мал. 35) є близьким зображенням графіка теоретичної функції розподілу $F(x)$ випадкової величини ξ , з елементів якої укладена вибірка.

$$F^*(x) = \begin{cases} 0,00, & \text{якщо } x < 280 \\ 0,01, & \text{якщо } 280 \leq x < 290 \\ 0,11, & \text{якщо } 290 \leq x < 300 \\ 0,25, & \text{якщо } 300 \leq x < 310 \\ 0,39, & \text{якщо } 310 \leq x < 320 \\ 0,64, & \text{якщо } 320 \leq x < 330 \\ 0,80, & \text{якщо } 330 \leq x < 340 \\ 0,87, & \text{якщо } 340 \leq x < 350 \\ 0,91, & \text{якщо } 350 \leq x < 360 \\ 0,98, & \text{якщо } 360 \leq x < 370 \\ 0,98, & \text{якщо } 370 \leq x < 380 \\ 1,00, & \text{якщо } 380 \leq x < 390. \end{cases}$$



Мал. 35. Емпірична функція розподілу



Мал. 36. Полігон накопиченої відносної частоти (кускова лінія) та емпірична функція розподілу $F^*(x)$ (східчаста лінія)

На мал. 36 суміщені графіки полігона накопиченої відносної частоти (мал. 33) та емпіричної функції розподілу $F^*(x)$ (мал. 35).

З цього графіка видно, що розподіл накопиченої відносної частоти може бути зображений двома різними графіками: ступінчастим графіком або ламаною прямою.

Першому зображенню відповідає емпірична функція розподілу, а другому – полігон накопиченої відносної частоти.

У той же час, цілком ясно, що обидва ці графіки є різним, але близьким, зображенням теоретичної функції розподілу $F(x)$, вказану вище.

Питання для самоперевірки та вправи

1. Що є предметом математичної статистики?
2. Вказати основні задачі математичної статистики.
3. Що називають статистичною, генеральною та вибірковою сукупністю, об'ємом цих сукупностей?
4. Вказати джерела даних у статистиці.
5. Що називають простою випадковою вибіркою? Як здійснюється проста випадкова вибірка за допомогою випадкових чисел?
6. Що називають незгрупованим і згрупованим розподілом частоти вибірки?
7. Що називають згрупованим розподілом накопиченої частоти вибірки?
8. Як визначають гістограми частот або частостей для згрупованих даних вибірки? Як визначають полігон частот для згрупованих даних вибірки?
9. Який імовірнісний зміст мають гістограма та полігон частот для згрупованих даних вибірки?
10. Як визначають та позначають емпіричну функцію розподілу? Які основні властивості цієї функції?
11. Який геометричний зміст медіани і як вона підраховується для згрупованих даних вибірки?
12. Як ширина класу інтервалів згрупованих даних вибірки впливає на якість гістограми? Які рекомендації до вибору числа класів інтервалів?

Розв'язати задачі

1. Для заданої вибірки із генеральної сукупності скласти розподіли частот та частостей.
 - а) 7, 4, 4, 8, 12, 12, 12, 7, 8, 12, 8, 12, 4, 12, 12, 4, 12, 4, 12, 12;
 - б) 6, 9, 5, 3, 6, 6, 9, 3, 5, 6, 9, 5, 6, 6, 9, 6, 9, 6, 6, 6;
 - в) 2, 2, 8, 5, 4, 2, 5, 4, 2, 8, 8, 2, 4, 8, 2, 2, 8, 2, 2, 2;
 - г) 7, 3, 10, 4, 7, 7, 3, 4, 10, 3, 7, 7, 4, 7, 4, 7, 4, 7, 7, 7.

Розділ 5. Основні поняття та статистичний розподіл

Відповідь:

а)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>4</td><td>7</td><td>8</td><td>12</td></tr> <tr><td>n_k</td><td>5</td><td>2</td><td>3</td><td>10</td></tr> <tr><td>$\frac{n_k}{n}$</td><td>$\frac{1}{4}$</td><td>$\frac{1}{10}$</td><td>$\frac{3}{20}$</td><td>$\frac{1}{2}$</td></tr> </table>	x_k	4	7	8	12	n_k	5	2	3	10	$\frac{n_k}{n}$	$\frac{1}{4}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$; б)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>3</td><td>5</td><td>6</td><td>9</td></tr> <tr><td>n_k</td><td>2</td><td>3</td><td>10</td><td>5</td></tr> <tr><td>$\frac{n_k}{n}$</td><td>$\frac{1}{10}$</td><td>$\frac{3}{20}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{4}$</td></tr> </table>	x_k	3	5	6	9	n_k	2	3	10	5	$\frac{n_k}{n}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$	$\frac{1}{4}$
x_k	4	7	8	12																													
n_k	5	2	3	10																													
$\frac{n_k}{n}$	$\frac{1}{4}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$																													
x_k	3	5	6	9																													
n_k	2	3	10	5																													
$\frac{n_k}{n}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$	$\frac{1}{4}$																													
в)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>2</td><td>4</td><td>5</td><td>8</td></tr> <tr><td>n_k</td><td>10</td><td>3</td><td>2</td><td>5</td></tr> <tr><td>$\frac{n_k}{n}$</td><td>$\frac{1}{2}$</td><td>$\frac{3}{20}$</td><td>$\frac{1}{10}$</td><td>$\frac{1}{4}$</td></tr> </table>	x_k	2	4	5	8	n_k	10	3	2	5	$\frac{n_k}{n}$	$\frac{1}{2}$	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{1}{4}$; г)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>3</td><td>4</td><td>7</td><td>10</td></tr> <tr><td>n_k</td><td>3</td><td>5</td><td>10</td><td>2</td></tr> <tr><td>$\frac{n_k}{n}$</td><td>$\frac{3}{20}$</td><td>$\frac{1}{4}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{10}$</td></tr> </table>	x_k	3	4	7	10	n_k	3	5	10	2	$\frac{n_k}{n}$	$\frac{3}{20}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{10}$
x_k	2	4	5	8																													
n_k	10	3	2	5																													
$\frac{n_k}{n}$	$\frac{1}{2}$	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{1}{4}$																													
x_k	3	4	7	10																													
n_k	3	5	10	2																													
$\frac{n_k}{n}$	$\frac{3}{20}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{10}$																													

2. Знайти емпіричну функцію заданого розподілу вибірки та побудувати її графік.

а)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>2</td><td>5</td><td>7</td><td>8</td></tr> <tr><td>n_k</td><td>1</td><td>3</td><td>2</td><td>4</td></tr> </table>	x_k	2	5	7	8	n_k	1	3	2	4	; б)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>4</td><td>7</td><td>8</td></tr> <tr><td>n_k</td><td>5</td><td>2</td><td>3</td></tr> </table>	x_k	4	7	8	n_k	5	2	3
x_k	2	5	7	8																	
n_k	1	3	2	4																	
x_k	4	7	8																		
n_k	5	2	3																		
в)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>1</td><td>4</td><td>6</td></tr> <tr><td>n_k</td><td>10</td><td>15</td><td>25</td></tr> </table>	x_k	1	4	6	n_k	10	15	25	; г)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>3</td><td>5</td><td>8</td></tr> <tr><td>n_k</td><td>15</td><td>10</td><td>25</td></tr> </table>	x_k	3	5	8	n_k	15	10	25		
x_k	1	4	6																		
n_k	10	15	25																		
x_k	3	5	8																		
n_k	15	10	25																		

3. Побудувати полігон частот та частостей заданого розподілу вибірки.

а)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>2</td><td>3</td><td>5</td><td>6</td></tr> <tr><td>n_k</td><td>10</td><td>15</td><td>5</td><td>20</td></tr> </table>	x_k	2	3	5	6	n_k	10	15	5	20	; б)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>15</td><td>20</td><td>25</td><td>30</td></tr> <tr><td>n_k</td><td>10</td><td>15</td><td>30</td><td>20</td></tr> </table>	x_k	15	20	25	30	n_k	10	15	30	20
x_k	2	3	5	6																			
n_k	10	15	5	20																			
x_k	15	20	25	30																			
n_k	10	15	30	20																			
в)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>3</td><td>4</td><td>5</td><td>7</td></tr> <tr><td>n_k</td><td>20</td><td>5</td><td>10</td><td>15</td></tr> </table>	x_k	3	4	5	7	n_k	20	5	10	15	; г)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>x_k</td><td>10</td><td>15</td><td>18</td><td>20</td></tr> <tr><td>n_k</td><td>6</td><td>4</td><td>5</td><td>5</td></tr> </table>	x_k	10	15	18	20	n_k	6	4	5	5
x_k	3	4	5	7																			
n_k	20	5	10	15																			
x_k	10	15	18	20																			
n_k	6	4	5	5																			

4. Побудувати гістограму частот заданого розподілу вибірки.

а)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>Інтервали</td><td>(1, 5)</td><td>(5, 9)</td><td>(9, 13)</td><td>(13, 17)</td><td>(17, 21)</td></tr> <tr><td>Частоти</td><td>10</td><td>20</td><td>50</td><td>12</td><td>8</td></tr> </table>	Інтервали	(1, 5)	(5, 9)	(9, 13)	(13, 17)	(17, 21)	Частоти	10	20	50	12	8	; б)	<table style="border-collapse: collapse; display: inline-table;"> <tr><td>Інтервали</td><td>(3, 5)</td><td>(5, 7)</td><td>(7, 9)</td><td>(9, 11)</td><td>(11, 13)</td><td>(13, 15)</td><td>(15, 17)</td></tr> <tr><td>Частоти</td><td>4</td><td>6</td><td>20</td><td>40</td><td>20</td><td>4</td><td>6</td></tr> </table>	Інтервали	(3, 5)	(5, 7)	(7, 9)	(9, 11)	(11, 13)	(13, 15)	(15, 17)	Частоти	4	6	20	40	20	4	6
Інтервали	(1, 5)	(5, 9)	(9, 13)	(13, 17)	(17, 21)																										
Частоти	10	20	50	12	8																										
Інтервали	(3, 5)	(5, 7)	(7, 9)	(9, 11)	(11, 13)	(13, 15)	(15, 17)																								
Частоти	4	6	20	40	20	4	6																								

в)

Інтервали	(2, 7)	(7, 12)	(12, 17)	(17, 22)	(22, 27)
Частоти	5	10	25	6	4

;

г)

Інтервали	(10, 15)	(15, 20)	(20, 25)	(25, 30)	(30, 35)
Частоти	2	4	8	4	2

.

5. В наступній таблиці наведено згрупований розподіл частоти кількості студентів університету в залежності від часу, що вони затратили на виконання контрольної роботи у хвиликах.

Кількість хвилин	Число студентів (частоти)	Накопичені частоти
85–90	8	8
90–95	17	25
95–100	29	54
100–105	93	147
105–110	43	190
110–115	8	198
115–120	2	200
Разом	200	

Побудувати гістограму частот, полігон частот, полігон накопичених частот.

6. Використовуючи точкову діаграму, одержати розподіл частот заданої вибірки, побудувати графік емпіричної функції розподілу та полігон частот.

а) 3, 6, 6, 5, 6, 2, 3, 6, 6, 5, 2, 3, 3, 6, 2, 3, 6, 6, 6, 5, 2, 6, 2, 3, 2, 5, 5, 6, 6, 3, 2, 3, 6, 2, 6, 2, 2, 3, 3, 6, 3, 6, 6, 3, 3, 6, 3, 6, 6, 3.

б) 6, 6, 9, 5, 3, 6, 9, 3, 5, 6, 9, 5, 6, 9, 6, 9, 6, 9, 6, 9, 6, 6, 3.

в) 2, 2, 4, 5, 8, 8, 5, 2, 4, 2, 8, 4, 2, 8, 8, 2, 2, 5, 2, 5, 4, 2, 4, 2, 5.

г) 3, 7, 3, 3, 4, 3, 7, 5, 5, 4, 7, 3, 7, 4, 3, 3, 7, 4, 5, 3, 5, 3, 7, 4, 5, 3, 3, 7, 3, 5, 3, 7, 5, 5, 3, 7, 5, 3, 7, 5, 7, 7, 3, 3, 7, 3, 7, 3, 7, 3.

Розділ 6

СТАТИСТИЧНІ ОЦІНКИ ПАРАМЕТРІВ РОЗПОДІЛУ

Застосування математики до обробки вибірки дозволяє нам знайти числові характеристики.

6.1. Основні вимоги до статистичних оцінок

У багатьох випадках треба дослідити кількісну ознаку X генеральної сукупності, використовуючи результати вибірки. Часто для цього достатньо знати наближені значення математичного сподівання $M(X)$, дисперсію $D(X)$, середньоквадратичне відхилення $\sigma(X)$, початкові або центральні моменти випадкової величини X .

Іноді з деяких міркувань вдається встановити закон розподілу X . Тоді треба вміти оцінювати параметри цього закону розподілу.

Наприклад, відомо, що випадкова величина X розподілена рівномірно; треба по даних вибірки наближено знайти відрізок, в якому знаходяться значення випадкової величини X .

Якщо X розподілена у генеральній сукупності за нормальним законом, то її щільність імовірностей має вигляд

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

Необхідно оцінити (знайти наближені значення) параметра a , який дорівнює $M(X)$, та σ , який дорівнює $\sigma(X)$. Ці параметри повністю визначають нормальний розподіл X .

Якщо X розподілена за законом Пуассона, то необхідно оцінювати лише один параметр a , яким цей розподіл визначається.

Дослідник має у своєму розпорядженні лише дані вибірки, одержані в результаті спостережень. Саме через ці дані і треба виразити потрібний параметр випадкової величини X генеральної сукупності.

Означення 1. *Статистичною оцінкою невідомого параметра випадкової величини X генеральної сукупності (теоретичного розподілу X) називають функцію від випадкових величин (результатів вибірки), що спостерігаються.*

Щоб статистичні оцінки давали найкращі наближення параметрів, вони повинні задовольняти певним вимогам. Розглянемо ці вимоги.

Нехай θ^* є статистична оцінка невідомого параметра θ теоретичного розподілу.

Припустимо, що за вибіркою об'єму n знайдена оцінка θ_1^* . При інших вибірках того ж об'єму одержимо деякі інші оцінки θ_2^* , θ_3^* , ..., θ_m^* . Саме оцінку θ^* можна розглядати як випадкову величину, а числа θ_1^* , θ_2^* , ..., θ_m^* як її можливі значення.

Якщо числа θ_k^* ($k = 1, 2, \dots, m$) будуть більші значення θ , тоді оцінка θ^* дає наближене значення θ з надлишком. У цьому випадку математичне сподівання випадкової величини θ^* буде більше θ , $M(\theta^*) > \theta$.

Якщо θ^* дає оцінку θ з недостачею, тоді $M(\theta^*) < \theta$.

Таким чином, використання статистичної оцінки, математичне сподівання якої не дорівнює параметру θ , приводить до систематичних (одного знака) похибок.

Вимога $M(\theta^*) = \theta$ застерігає від систематичних похибок.

Означення 2. *Статистичну оцінку θ^* параметра θ називають незсунутою, якщо $M(\theta^*) = \theta$.*

Оцінку θ^* називають **зсунутою**, якщо ця рівність не виконується.

Розділ 6. Статистичні оцінки параметрів розподілу

Вимога про незсунутість оцінки θ^* є недостатньою тому, що можливі значення θ^* можуть бути сильно розсіяні від свого середнього значення, дисперсія $D(\theta^*)$ може бути великою. Тоді знайдена за даними однієї вибірки оцінка, наприклад, θ_k^* може набагато відрізнятись від середнього значення θ^* , отже і від параметра θ .

Якщо $D(\theta^*)$ буде малою, тоді можливість допустити велику помилку буде виключена. Тому до статистичної оцінки виникає вимога про її ефективність.

Означення 3. Ефективною називають таку статистичну оцінку θ^* , яка при заданому об'єму n має найменшу можливу дисперсію.

При розгляданні вибірки великого об'єму ($n \rightarrow \infty$) до статистичних оцінок пред'являють вимогу їх обґрунтованості.

Означення 4. Обґрунтованою називають статистичну оцінку, яка при $n \rightarrow \infty$ прямує за імовірністю до оцінюваного параметра.

Наприклад, якщо дисперсія незсунутої оцінки при $n \rightarrow \infty$ прямує до нуля, то оцінка буде і обґрунтованою.

6.2. Числові характеристики вибіркової сукупності

6.2.1. Вибіркові характеристики

У доповнення до табличних та графічних методів представлення даних наступним найважливішим засобом обробки даних є обчислення їх числових характеристик. Найважливіші з них: **середнє значення, дисперсія, середнє квадратичне відхилення (стандартне відхилення)**.

Ці характеристики можуть бути обчислені за даними, що входять у вибірці або за даними, що входять у кінцеву генеральну сукупність.

Числові характеристики, обчислені по вибірці або ті, що використовуються для опису даних вибірки, називають **статистиками**.

Числові характеристики, обчислені по генеральній сукупності або ті, що використовуються для опису даних генеральної сукупності, називають **параметрами**.

По аналогії з математичним сподіванням, дисперсією та середнім квадратичним відхиленням дискретної випадкової величини визначають вибіркові характеристики, замінюючи при цьому імовірності p_i частотами вибірки $\frac{n_i}{n}$.

Але у статистиці застосовують ще інші числові характеристики.

Означення 1. *Простою середньоарифметичною вибірки називають суму варіант вибірки, поділену на об'єм вибірки. Її позначають*

$$\bar{x}_a = \frac{1}{n} \sum_{i=1}^m x_i,$$

де x_i ($i = 1, 2, \dots, m$) – варіанти вибірки, n – об'єм вибірки.

Означення 2. *Вибірковою середньою або зваженою середньоарифметичною називають середню арифметичну варіант вибірки з врахуванням їх частостей і позначають*

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^m n_i x_i,$$

де n – об'єм вибірки, m – число різних варіант,

n_1, \dots, n_m – частоти варіант ($n = n_1 + \dots + n_m$),

x_i – значення i -тої варіанти.

Вибіркова середня є аналогом математичного сподівання і використовується дуже часто. Вона може приймати різні числові значення при різних вибірках однакового об'єму.

Тому можна розглядати розподіли (теоретичний та емпіричний) вибіркової середньої та числові характеристики цього розподілу (цей розподіл називають **вибірковим**).

Основні властивості вибіркової середньої

1. При множенні усіх варіант вибірки на однаковий множник вибіркова середня також множитья на цей множник

$$\frac{1}{n} \sum_{i=1}^m n_i (cx_i) = \frac{c}{n} \sum_{i=1}^m n_i x_i = c\bar{x}_B.$$

2. Якщо додати (відняти) до всіх варіант вибірки однакове число, то вибіркова середня зростає (зменшується) на це число

$$\frac{1}{n} \sum_{i=1}^m n_i (x_i + c) = \frac{1}{n} \sum_{i=1}^m n_i x_i + \frac{1}{n} c \sum_{i=1}^m n_i = \bar{x}_B + c.$$

Ці властивості можна поєднати в одну формулу, яку називають **формулою моментів**

$$\mu = \frac{k}{n} \sum_{i=1}^m n_i \frac{x_i - c}{k} \tag{2}$$

і використовують у статистиці.

Зауваження 1. Якщо ввести умовну варіанту

$$u_i = \frac{x_i - c}{k},$$

то формула моментів (2) приймає простий вигляд

$$\bar{x}_B = \frac{k}{n} \sum_{i=1}^m n_i u_i + c.$$

Означення 3. *Степеневою середньою вибірки називають таку середню, яку знаходять за формулою*

$$\bar{x}_C = \left(\sum_{i=1}^m \frac{n_i}{n} x_i^\alpha \right)^{\frac{1}{\alpha}}.$$

При $\alpha = 1$ одержимо формулу (1), тобто вибірккову середню.

При $\alpha = 2$ одержимо **середньоквадратичну вибірки**

$$\bar{x}_2 = \sqrt{\sum_{i=1}^m \frac{n_i}{n} x_i^2}.$$

При $\alpha = -1$ одержуємо **середню гармонічну**

$$\bar{x}_{-1} = \bar{x}_{\text{гар.}} = \frac{n}{\sum_{i=1}^m \frac{n_i}{x_i}}.$$

Середню гармонічну застосовують у тому випадку, коли шуканий показник є величина, що обернена середньому значенню ознаки.

При $\alpha = 0$ вираз (3) буде невизначеним. Застосовуючи логарифмування та правило Лопітала розкриття невизначеності, одержимо середню геометричну

$$\bar{x}_G = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_m^{n_m}}.$$

Ця середня обчислюється лише при умові, що усі варіанти є додатні $x_i > 0$, $i = 1, 2, \dots, m$.

Середня геометрична застосовується у статистиці для визначення темпу зростання при дослідженні змін ознаки з часом.

Зауваження 2. *Обрання тієї чи іншої середньої для характеристики розподілу пов'язано з якісним аналізом цього розподілу.*

Зауваження 3. *Крім вказаних степеневих середніх, у статистиці застосовуються ще **структурні середні**, які не залежать від зна-*

чень варіант, що розташовані на краях розподілу, а пов'язані з рядом частот.

До структурних середніх відносять **моду** та **медіану**. Нагадаємо, що модою називають значення варіанти, яка має найбільшу частоту. Означення медіани та способи її побудови див. параграф 5.8.

Означення 4. *Вибірковою дисперсією D_B називають середню квадратів відхилення варіант від вибіркової середньої з врахуванням відповідних частостей*

$$D_B = \frac{1}{n} \sum_{i=1}^m n_i (x_i - \bar{x}_B)^2. \quad (4)$$

Зауваження 4. *Обчислення вибіркової дисперсії спрощується, якщо її знаходити за формулою*

$$D_B = \frac{1}{n} \sum_{i=1}^m n_i x_i^2 - \left(\frac{1}{n} \sum_{i=1}^m n_i x_i \right)^2 = \frac{1}{n} \sum_{i=1}^m n_i x_i^2 - (\bar{x}_B)^2. \quad (5)$$

Означення 5. *Вибірковим середньоквадратичним відхиленням (стандартом) називають квадратний корінь із вибіркової дисперсії*

$$\sigma_B = \sqrt{D_B}. \quad (6)$$

Зауваження 5. *Вибіркова дисперсія дає занижені значення для дисперсії $D(X)$ генеральної сукупності, вона буде зсунутою оцінкою $D(X)$. Але математичне сподівання D_B буде*

$$M(D_B) = \frac{n-1}{n} D(X).$$

Тому вибіркoву дисперсію доцільно виправити таким чином, щоб вона стала незсунутою оцінкою. Для цього достатньо D_B помножити на дріб $\frac{n}{n-1}$.

Виправлену вибіркoву дисперсію позначають

$$s^2 = \frac{n}{n-1} D_B = \frac{1}{n-1} \sum_{i=1}^m n_i (x_i - \bar{x}_B)^2. \quad (7)$$

Тоді виправленим середньоквадратичним відхиленням вибірки буде $s = \sqrt{s^2}$.

Із формул (4) та (7) випливає, що при досить великих n (об'єм вибірки) вибіркова дисперсія D_B та виправлена вибіркова дисперсія s^2 різняться дуже мало. Тому в практичних задачах виправлену дисперсію s^2 та виправлене середньоквадратичне відхилення вибірки s використовують лише при об'ємі вибірки $n < 30$.

Приклад 1. Вибіркова сукупність задана таблицею

x_i	1	2	3	4
n_i	20	15	10	5

Знайти вибіркові характеристики.

Розв'язання. У даному випадку об'єм вибірки дорівнює

$$n = 20 + 15 + 10 + 5 = 50.$$

За формулою (1) знаходимо вибіркву середню

$$\bar{x}_B = \frac{1}{50}(1 \cdot 20 + 2 \cdot 15 + 3 \cdot 10 + 4 \cdot 5) = \frac{100}{50} = 2.$$

За формулою (4) знаходимо вибіркву дисперсію

$$D_B = \frac{1}{50}(20 \cdot (1-2)^2 + 15 \cdot (2-2)^2 + 10 \cdot (3-2)^2 + 5 \cdot (4-2)^2) = \frac{50}{50} = 1.$$

За формулою (6) знаходимо вибіркве середньоквадратичне відхилення (стандарт)

$$\sigma_B = \sqrt{1} = 1.$$

6.2.2. Обчислення вибіркових характеристик методом добутоків

Як правило, обчислення \bar{x}_B та D_B за формулами (1) та (4) або (5) проводиться з використанням комп'ютерної техніки. Часто розрахунки можна спростити, використовуючи **метод добутоків**, в основі якого лежать рівновіддалені варіанти та наступна розрахункова таблиця

1	2	3	4	5	6
x_k	n_k	u_k	$n_k \cdot u_k$	$n_k \cdot u_k^2$	$n_k (u_k + 1)^2$

Дамо необхідні пояснення до цього методу

Алгоритм методу добутоків

1. У перший стовпчик таблиці записують рівновіддалені варіанти x_k вибірки, розміщуючи їх у зростаючому порядку.

2. У другий стовпчик таблиці записують відповідні частоти n_k варіант. Суму усіх елементів цього стовпчика (об'єм вибірки n) записують у останню клітинку цього стовпчика.

3. Третій стовпчик містить умовні варіанти u_k вибірки. Для знаходження умовних варіант вибірки треба:

а) значення варіанти вибірки з найбільшою частотою C обрати за умовний нуль. Це значення варіанти називається **модюю**;

б) знайти різницю h між будь-якими двома сусідніми варіантами;

в) обчислити умовні варіанти вибірки за формулою

$$u_k = \frac{x_k - C}{h}. \quad (8)$$

Відмітимо, що умовні варіанти завжди будуть цілими числами.

4. У четвертий стовпчик записують добутки частот та відповідних умовних варіант $n_k \cdot u_k$. Суму елементів стовпчика записують в останню клітинку цього стовпчика.

5. Знаходять добутки частот та квадратів умовних варіант $n_k \cdot u_k^2$ і записують їх у п'ятий стовпчик. Суму елементів стовпчика

$\sum_{k=1}^m n_k u_k^2$ записують в останню клітинку цього стовпчика.

6. Знаходять добутки частот та квадратів умовних варіант, збільшених на одиницю, $n_k (u_k + 1)^2$ і записують їх у шостий конт-

рольний стовпчик. Суму елементів стовпчика $\sum_{k=1}^m n_k (u_k + 1)^2$ записують в останню клітинку цього стовпчика.

7. Перевіряють обчислення так: сума елементів шостого стовпчика повинна задовольняти тотожність

$$\sum_{k=1}^m n_k (u_k + 1)^2 = \sum_{k=1}^m n_k \cdot u_k^2 + 2 \sum_{k=1}^m n_k u_k + n. \quad (9)$$

8. Обчислюють умовні моменти за формулами

$$M_1^* = \frac{1}{n} \cdot \sum_{k=1}^m n_k u_k; \quad M_2^* = \frac{1}{n} \cdot \sum_{k=1}^m n_k u_k^2. \quad (10)$$

9. Обчислюють вибіркові середню та дисперсію за формулами

$$\bar{x}_B = M_1^* \cdot h + C; \quad D_B = \left[M_2^* - (M_1^*)^2 \right] \cdot h^2. \quad (11)$$

Приклад 2. Знайти методом добутоків вибіркові середню \bar{x}_B та дисперсію D_B заданої вибірки

x_k	18,6	19	19,4	19,8	20,2	20,6
n_k	4	6	30	40	18	2

Розв'язання. Будемо використовувати розрахункову таблицю методу добутоків. У даному випадку: варіанти вибірки рівновіддалені $h = x_2 - x_1 = 19 - 18,6 = 0,4$; найбільша частота 40 у варіанті 19,8. Тому умовним нулем буде $C = 19,8$.

У перші два стовпчики розрахункової таблиці записуємо варіанти та частоти заданої вибірки, а елементи третього стовпчика обчислюємо за формулою (8) при вказаних $C = 19,8$ та $h = 0,4$.

1	2	3	4	5	6
x_k	n_k	u_k	$n_k \cdot u_k$	$n_k \cdot u_k^2$	$n_k (u_k + 1)^2$
18,6	4	-3	-12	36	16
19	6	-2	-12	24	6
19,4	30	-1	-30	30	0
19,8	40	0	0	0	40
20,2	18	1	18	18	72
20,6	2	2	4	8	18
	$n = 100$		$\sum = -32$	$\sum = 116$	$\sum = 152$

Для контролю перевіряємо умову (9)

$$116 + 2 \cdot (-32) + 100 = 152; \quad 152 \equiv 152.$$

Знайдемо умовні моменти за формулами (10)

$$M_1^* = \frac{1}{100} \cdot (-32) = -0,32; \quad M_2^* = \frac{1}{100} \cdot 116 = 1,16.$$

За формулами (11) знаходимо шукані вибірку середню та вибірку дисперсію

$$\bar{x}_B = (-0,32) \cdot 0,4 + 19,8 = -0,128 + 19,8 = 19,672;$$

$$D_B = [1,16 - (-0,32)^2] \cdot (0,4)^2 = [1,16 - 0,1024] \cdot 0,16 = 0,169216.$$

6.2.3. Статистичні моменти розподілу

Визначимо аналогічно початковому та центральному моментам розподілу із теорії імовірностей деякі числові характеристики вибірки.

Означення 6. Моментом порядку k називають середнє значення k -го степеня різниці $x_i - C$.

При $C = 0$ одержимо початковий момент порядку k вибірки

$$v_k^* = \frac{1}{n} \sum_{i=1}^m x_i^k.$$

При $C = \bar{x}_B$ одержимо центральний момент порядку k вибірки

$$\mu_k^* = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x}_B)^k.$$

Моменти порядку k та умовні моменти M_1^* та M_2^* , які обчислюють за формулами (10), часто використовують у статистиці.

6.2.4. Приклади знаходження статистик вибірки

У випадку згрупованої вибірки припускається, що всяке значення варіанти, що потрапило в даний клас інтервалів, дорівнює середньому значенню варіанти в цьому класі.

1. **Вибіркове середнє** обчислюється по формулі

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k n_i \tilde{x}_i, \quad (12)$$

де n – об'єм вибірки, \tilde{x}_i – середнє значення варіанти на класі, n_i – частота i -того класу інтервалів, k – кількість класів.

2. Для дисперсії маємо формулу

$$D_B = \frac{1}{n} \sum_{i=1}^k n_i (\tilde{x}_i - \bar{x}_B)^2 \quad (13)$$

і формулу для обчислень

$$D_B = \frac{1}{n} \sum_{i=1}^k n_i \tilde{x}_i^2 - (\bar{x}_B)^2. \quad (14)$$

Для **виправленої дисперсії** маємо формулу

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

і формулу для обчислень

$$s^2 = \frac{n \sum_{i=1}^k n_i \tilde{x}_i^2 - \left(\sum_{i=1}^k n_i \tilde{x}_i \right)^2}{n(n-1)}$$

або

$$s^2 = \frac{\sum_{i=1}^k n_i x_i^2 - n \cdot \tilde{x}_B^2}{n-1}.$$

Приклад 3. Обчислити числові характеристики вибірки згрупованого розподілу частот середньомісячної платні співробітників фірми N (див. Таблицю 11).

Таблиця 11

Платня (класи інтервалів)	Частота n_i
280–290	1
290–300	10
300–310	14
310–320	14
320–330	25
330–340	16
340–350	7
350–360	4
360–370	7
370–380	0
380–390	2
Разом:	100

Таблиця 12

Платня (середнє в класі), \tilde{x}_i	Частота n_i
285	1
295	10
305	14
315	14
325	25
335	16
345	7
355	4
365	7
375	0
385	2
Разом:	100

Перейдемо від Таблиці 11 до Таблиці 12, замінивши класи інтервалів на середні значення варіант в класі. Користуючись даними з Таблиці 12, за формулою (12) отримуємо

$$\bar{x}_B = \frac{1}{100} (285 \cdot 1 + 295 \cdot 10 + \dots + 385 \cdot 2) = 325,6 \text{ грн.}$$

Порівнюючи отримане значення вибіркової середньої $\bar{x}_B = 325,6$ із точним значенням $\bar{x}_B = 324,3$, бачимо що вони відрізняються незначним чином.

Помилка виникла за рахунок округлення всіх варіант у класі до середнього значення. Проте, як показує практика, помилка, що вноситься при цьому, незначна.

Далі, за формулами (13), (14) отримаємо

$$D_B = 439,64 \text{ грн}^2.$$

Для середньоквадратичного відхилення маємо

$$\sigma_B = \sqrt{D_B} = \sqrt{439,64} = 20,97 \text{ грн.}$$

Згідно із запровадженої вище термінології \bar{x}_B , D_B , σ_B , s^2 є статистиками.

Як видно з формул (13) і (14) дисперсія є мірою розсіювання варіант у вибірці навколо їхнього середнього значення. Пояснимо це на наступному прикладі.

Приклад 4. Обстежені по 65 випадків виплати страхових сум двома страховими компаніями N і R за деякий період часу. За одиницю виплати прийнята деяка стандартна сума. Виплата може приймати будь-які значення від 0 до 4. Знак «мінус» перед числом означає, що виплату робить страхова компанія, а знак «плюс» – що компанія отримує страховий внесок. Треба підрахувати числові характеристики цих вибірок.

Розподіл частот виплат страхових сум обох компаній наведено у Таблицях 13 та 14.

Підрахуємо вибіркове середнє $\bar{x}_B(N)$ і $\bar{x}_B(R)$ для обох компаній

$$\bar{x}_B(N) = \frac{1}{65}(-4 \cdot 1 - 3 \cdot 4 - 2 \cdot 9 - 1 \cdot 16 + 0,25 + 1 \cdot 16 + 2 \cdot 9 + 3 \cdot 4 + 4 \cdot 1) = 0,$$

$$\bar{x}_B(R) = \frac{1}{65}(-4 \cdot 1 - 3 \cdot 5 - 2 \cdot 8 - 1 \cdot 17 + 0 \cdot 23 + 1 \cdot 17 + 2 \cdot 8 + 3 \cdot 4 + 4 \cdot 1) = 0.$$

Підрахуємо вибіркові дисперсії $D_B(N)$ і $D_B(R)$

$$D_B(N) = \frac{1}{65}(1 \cdot 4^2 + 4 \cdot 3^2 + 9 \cdot 2^2 + 16 \cdot 1^2) \cdot 2 = \frac{208}{65} = 3,2,$$

$$D_B(R) = \frac{1}{65}(1 \cdot 4^2 + 5 \cdot 3^2 + 4 \cdot 8 + 1 \cdot 17) \cdot 2 = \frac{55}{65} = 0,85.$$

Таблиця 13

Розподіл частот виплат страхових сум компанією N

Страховий внесок	Частота
-4	1
-3	4
-2	9
-1	16
0	25
1	16
2	9
3	4
4	1
Разом:	65

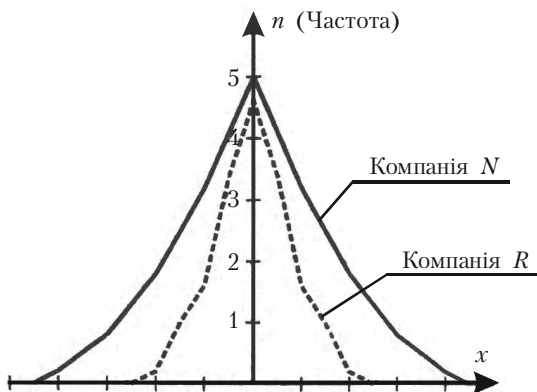
Таблиця 14

Розподіл частот виплат страхових сум компанією R

Страховий внесок	Частота
-2	1
-1,5	5
-1	8
-0,5	17
0	23
0,5	17
1	8
1,5	5
2	1
Разом:	65

Полігон частот обох розподілів зображень на мал. 37.

З мал. 37 видно, що чим менша дисперсія, тим у більш вузькому інтервалі дані вибірки групуються навколо свого середнього значення (у нашому випадку $\bar{x}_B = 0$).



Мал. 37

6.3. Точкові та інтервальні оцінки

6.3.1. Загальні поняття

Означення 1. *Точковими оцінками параметрів розподілу генеральної сукупності називають такі оцінки, які визначаються одним числом.*

Наприклад, вибіркова середня \bar{x}_B , вибіркова дисперсія та вибіркове середньоквадратичне σ_B – точкові оцінки відповідних числових характеристик генеральної сукупності.

Точкові оцінки параметрів розподілу є випадковими величинами, їх можна вважати первинними результатами обробки вибірки тому, що невідомо, з якою точністю кожна з них оцінює відповідну числову характеристику генеральної сукупності.

Якщо об'єм вибірки досить великий, то точкові оцінки задовольняють практичні потреби точності.

Якщо об'єм вибірки малий, то точкові оцінки можуть давати значні похибки, тому питання точності оцінок у цьому випадку дуже важливе і використовують інтервальні оцінки.

Означення 2. *Інтервальною називають оцінку, яка визначається двома числами – кінцями інтервалу.*

Інтервальні оцінки дозволяють встановити точність та надійність оцінок. Ознайомимось з цими поняттями.

Нехай знайдена за даними вибірки статистична оцінка θ^* буде оцінкою невідомого параметра θ .

Ясно, що θ^* тим точніше визначає θ , чим менше абсолютна величина різниці $\theta - \theta^*$.

Іншими словами, якщо $\delta > 0$ і $|\theta - \theta^*| < \delta$, тоді меншому δ відповідає більш точна оцінка. Тому число δ характеризує точність оцінки.

Але статистичні методи не дозволяють категорично стверджувати, що оцінка θ^* задовольняє нерівність $|\theta - \theta^*| < \delta$.

Таке твердження можна зробити лише з імовірністю γ .

Означення 3. *Надійністю (довірчою імовірністю) оцінки параметра θ за θ^* називають імовірність*

$$\gamma = P(|\theta - \theta^*| < \delta), \quad (1)$$

з якою виконується нерівність $|\theta - \theta^*| < \delta$.

Найчастіше число γ задається наперед і, залежно від обставин, воно дорівнює 0,95 або 0,99 або 0,999.

Формулу (1) можна записати у вигляді

$$P(\theta^* - \delta < \theta < \theta^* + \delta) = \gamma. \quad (2)$$

З цієї рівності випливає, що інтервал $(\theta^* - \delta, \theta^* + \delta)$ містить невідомий параметр θ генеральної сукупності.

Означення 4. *Інтервал $(\theta^* - \delta, \theta^* + \delta)$ називають довірчим, якщо він покриває невідомий параметр θ із заданою надійністю γ .*

Зауваження 1. *Кінці довірчого інтервалу є випадковими величинами.*

6.3.2. Довірчий інтервал для оцінки математичного сподівання нормального розподілу

Нехай кількісна ознака X генеральної сукупності розподілена за нормальним законом, середньоквадратичне відхилення σ відомо. Треба знайти довірчий інтервал, що покриває математичне сподівання a генеральної сукупності із заданою надійністю γ .

Згідно із властивістю нормально розподіленої випадкової величини X маємо

$$\begin{aligned} P(|X - a| < \delta) &= P(a - \delta < X < a + \delta) = \\ &= \Phi\left(\frac{a + \delta - a}{\sigma}\right) - \Phi\left(\frac{a - \delta - a}{\sigma}\right) = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(-\frac{\delta}{\sigma}\right). \end{aligned}$$

Оскільки інтегральна функція Лапласа Φ є непарна, то одержимо

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

Але \bar{x}_B випадкова величина, $M(\bar{x}_B) = a$, $\sigma(\bar{x}_B) = \frac{\sigma}{\sqrt{n}}$, тому

при заміні X на \bar{x}_B , σ на $\frac{\sigma}{\sqrt{n}}$ одержимо

$$P(|\bar{x}_B - a| < \delta) = 2\Phi\left(\frac{\delta \cdot \sqrt{n}}{\sigma}\right) = 2\Phi(t),$$

де

$$t = \frac{\delta \cdot \sqrt{n}}{\sigma} \Rightarrow \delta = \frac{t \cdot \sigma}{\sqrt{n}}.$$

Використовуючи формули (3) та (2), одержимо

$$P\left(\bar{x}_B - \frac{t\sigma}{\sqrt{n}} < a < \bar{x}_B + \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma, \quad (4)$$

тобто з надійністю γ довірчий інтервал

$$\left(\bar{x}_B - \frac{t\sigma}{\sqrt{n}}, \bar{x}_B + \frac{t\sigma}{\sqrt{n}}\right)$$

покриває невідомий параметр a . Точність оцінки буде

$$\delta = \frac{t\sigma}{\sqrt{n}}. \quad (5)$$

Число t визначається із рівності

$$2\Phi(t) = \gamma \Rightarrow \Phi(t) = \frac{\gamma}{2} \quad (6)$$

з використанням таблиці значень інтегральної функції Лапласа.

Розділ 6. Статистичні оцінки параметрів розподілу

Зауваження 2. З формули (5) випливає, що при зростанні об'єму вибірки n число δ зменшується, а це означає, що точність оцінки збільшується. Коли надійність γ збільшується, функція $\Phi(t)$ зростає і згідно з її властивістю зростає t і, як наслідок, зростає δ . Отже, збільшення надійності оцінки зменшує її точність.

Приклад 1. Випадкова величина розподілена за нормальним законом з параметром $\sigma = 2$. Зроблена вибірка об'єму $n = 25$. З надійністю $\gamma = 0,95$ знайти довірчий інтервал невідомого параметра a цього розподілу.

Розв'язання. Із рівності

$$\Phi(t) = \frac{1}{2}\gamma \Rightarrow \Phi(t) = 0,475.$$

З таблиці інтегральної функції Лапласа Φ знайдемо число $t = 1,96$. Тоді за формулою (5) точність оцінки буде

$$\delta = \frac{t \cdot \sigma}{\sqrt{n}} = \frac{2 \cdot 1,96}{\sqrt{25}} = 0,784.$$

Отже, довірчий інтервал буде $(\bar{x}_B - 0,784; \bar{x}_B + 0,784)$. Якщо $\bar{x}_B = 2,8$, то з надійністю 95% інтервал (2,016; 3,584) покриває параметр $a = 2,8$ з точністю до 0,8.

Зауваження 3. Знаходження об'єму вибірки. Нехай ознака X генеральної сукупності розподілена за нормальним законом з параметром σ і треба знайти об'єм вибірки n , який із заданою точністю δ та надійністю γ дозволить знайти оцінку параметра a . Із формули (5) одержуємо рівність

$$t = \frac{\delta\sqrt{n}}{\sigma},$$

з якої випливає

$$n = \frac{t^2\sigma^2}{\delta^2}. \quad (7)$$

Для надійності γ , використовуючи (6) та таблицю значень інтегральної функції Лапласа, знайдемо відповідне число t .

Тепер t , δ та σ відомі, тому за формулою (7) можна знайти потрібний об'єм вибірки.

Приклад 2. Випадкова величина розподілена за нормальним законом з параметром σ . Знайти мінімальний об'єм n вибірки, щоб з надійністю γ та точністю δ виконувалась рівність $\bar{x}_B = a$, якщо $\sigma = 0,5$; $\gamma = 0,95$; $\delta = 0,1$.

Розв'язання. Для $\gamma = 0,95$ згідно з формулою (6) маємо

$$\Phi(t) = 0,475 \Rightarrow t = 1,96.$$

Використовуючи формулу (7), знайдене t та задані σ , δ , одержуємо

$$n = \left(\frac{1,96 \cdot 0,5}{0,1} \right)^2 = (9,8)^2 = 96,04.$$

Отже, мінімальний об'єм вибірки $n = 96$.

Зауваження 4. Якщо невідоме середньоквадратичне відхилення σ ознаки X генеральної сукупності, то використовують розподіл Стюдента (див. параграф 16 роботи [4]), можна також в формулах (3)–(4), (7) замість σ використати σ_B .

6.4. Обробка вибірки методом найменших квадратів

6.4.1. Основні поняття

Припустимо, що нам відома функціональна залежність між випадковими величинами Y та X вигляду

$$Y = f(X, a_1, a_2, \dots, a_m)$$

з невідомими параметрами a_1, a_2, \dots, a_m .

Наприклад, можна розглядати залежність між собівартістю продукції (ознака Y) та об'ємом продукції (ознака X) деякої кількості однотипних підприємств.

Звичайно при зростанні об'єму продукції X собівартість Y повинна спадати. Але ця залежність не є однозначною. Внаслідок різних причин при випуску однакового об'єму продукції собівартість її на різних підприємствах буде неоднаковою.

Нехай внаслідок n незалежних випробувань одержані варіанти ознак Y та X , які оформлені у статистичній таблиці вигляду

Номери випробувань	1	2	...	k	...	n
X	x_1	x_2	...	x_k	...	x_n
Y	y_1	y_2	...	y_k	...	y_n

Для знаходження оцінок параметрів функціональної залежності a_1, a_2, \dots, a_m за даними вибірки застосуємо метод найменших квадратів. Цей метод базується на тому, що найімовірніші значення параметрів a_1, a_2, \dots, a_m повинні давати мінімум функції

$$S = \sum_{k=1}^n [y_k - f(x_k, a_1, a_2, \dots, a_m)]^2. \quad (1)$$

Якщо функція $f(x_k, a_1, a_2, \dots, a_m)$ має неперервні частинні похідні відносно невідомих параметрів a_1, a_2, \dots, a_m , то необхідною

умовою існування мінімуму функції S буде система m рівнянь з m невідомими

$$\frac{\partial S}{\partial a_k} = 0, \quad k = 1, 2, \dots, m$$

Знаходження функціональної залежності між випадковими величинами X та Y з використанням даних випробувань (або вибірки) називають **вирівнюванням емпіричних даних вздовж кривої**

$$y = f(x_k, a_1, a_2, \dots, a_m).$$

Нижче розглянемо детальніше оцінки параметрів лінійної та параболічної функціональної залежностей, які використовуються найчастіше.

6.4.2. Оцінка параметрів лінійної функції

Нехай між випадковими величинами X та Y існує лінійна функціональна залежність

$$Y = aX + b, \tag{2}$$

параметри a та b якої невідомі.

Згідно з формулою (1) маємо

$$S = \sum_{k=1}^n (y_k - (ax_k + b))^2.$$

Ця функція S неперервно диференційовна, тому згідно з необхідними умовами існування мінімуму S повинні виконуватись

$$\text{рівності } \frac{\partial S}{\partial a} = 0 \text{ та } \frac{\partial S}{\partial b} = 0.$$

У нашому випадку ці рівності мають вигляд

$$\begin{cases} \sum_{k=1}^n (y_k - ax_k - b)x_k = 0; \\ \sum_{k=1}^n (y_k - ax_k - b) = 0; \end{cases} \Rightarrow \begin{cases} a \sum_{k=1}^n x_k + nb = \sum_{k=1}^n y_k, \\ a \sum_{k=1}^n x_k^2 + b \sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k. \end{cases}$$

Розділ 6. Статистичні оцінки параметрів розподілу

Виписана система є неоднорідною лінійною системою двох рівнянь відносно двох невідомих a та b . За правилом Крамера можна знайти єдиний розв'язок цієї системи у вигляді

$$a = \frac{\left(\sum_{k=1}^n x_k\right) \cdot \left(\sum_{k=1}^n y_k\right) - n \cdot \sum_{k=1}^n x_k \cdot y_k}{\left(\sum_{k=1}^n x_k\right)^2 - n \cdot \sum_{k=1}^n x_k^2}; \quad (3)$$

$$b = \frac{\left(\sum_{k=1}^n x_k\right) \cdot \left(\sum_{k=1}^n x_k \cdot y_k\right) - \left(\sum_{k=1}^n x_k^2\right) \cdot \left(\sum_{k=1}^n y_k\right)}{\left(\sum_{k=1}^n x_k\right)^2 - n \cdot \sum_{k=1}^n x_k^2}. \quad (4)$$

Якщо кількість значень x_k та y_k велика, то обчислення параметрів a та b за формулами (3), (4) ускладнюється. Для спрощення обчислень початок розрахунків величин x_k переносять в середнє значення усіх, тобто в точку

$$\bar{x}_a = \frac{1}{n} \sum_{k=1}^n x_k.$$

Тоді після деяких проміжних викладок одержують

$$a = \frac{\sum_{k=1}^n (x_k - \bar{x}_a) \cdot y_k}{\sum_{k=1}^n (x_k - \bar{x}_a)^2}; \quad b = \frac{1}{n} \cdot \sum_{k=1}^n y_k - a\bar{x}_a. \quad (5)$$

Формули (5) дозволяють визначити параметри a та b лінійної функціональної залежності (2) шляхом обчислення простішим, ніж за формулами (3), (4).

6.4.3. Оцінка параметрів параболічної функціональної залежності

Нехай між випадковими величинами X та Y існує функціональна залежність вигляду

$$Y = aX^2 + bX + c. \quad (6)$$

Методом найменших квадратів на основі даних випробувань знайдемо значення невідомих параметрів a , b , c . Тепер формула (1) буде мати вигляд

$$S = \sum_{k=1}^n (y_k - ax_k^2 - bx_k - c)^2.$$

Необхідні умови існування мінімуму функції S є рівності нулю частиних похідних першого порядку

$$\begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{k=1}^n (y_k - ax_k^2 - bx_k - c) x_k^2 = 0, \\ \frac{\partial S}{\partial b} = -2 \sum_{k=1}^n (y_k - ax_k^2 - bx_k - c) x_k = 0, \\ \frac{\partial S}{\partial c} = -2 \sum_{k=1}^n (y_k - ax_k^2 - bx_k - c) = 0. \end{cases}$$

Цю систему можна записати у вигляді

$$\begin{cases} a \sum_{k=1}^n x_k^4 + b \sum_{k=1}^n x_k^3 + c \sum_{k=1}^n x_k^2 = \sum_{k=1}^n x_k^2 y_k, \\ a \sum_{k=1}^n x_k^3 + b \sum_{k=1}^n x_k^2 + c \sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k, \\ a \sum_{k=1}^n x_k^2 + b \sum_{k=1}^n x_k + cn = \sum_{k=1}^n y_k. \end{cases} \quad (7)$$

Система (7) є неоднорідною лінійною системою трьох рівнянь з невідомими a , b , c . Розв'язок цієї системи можна знайти різними методами (матричним, за правилом Крамера, методом Гаусса-Жор-

Розділ 6. Статистичні оцінки параметрів розподілу

дана), а його вигляд буде громіздкий при досить великій кількості випробувань n .

Система (7) та її розв'язок значно спрощуються, коли значення x_k рівновіддалені ($x_2 - x_1 = x_3 - x_2 = \dots = x_n - x_{n-1}$) та виконується умова

$$\sum_{k=1}^n x_k = 0,$$

яку можна одержати за допомогою нового аргументу

$$x'_k = x_k - \bar{x}_a, \quad \text{де } \bar{x}_a = \frac{1}{n} \sum_{k=1}^n x_k.$$

Припустимо, що вказані умови виконуються, тоді замість системи (7) одержимо систему

$$\begin{cases} a \sum_{k=1}^n x_k^4 + c \sum_{k=1}^n x_k^2 = \sum_{k=1}^n x_k^2 y_k, \\ b \sum_{k=1}^n x_k^2 = \sum_{k=1}^n x_k y_k, \\ a \sum_{k=1}^n x_k^2 + cn = \sum_{k=1}^n y_k. \end{cases}$$

Розв'язок цієї системи можна знайти за формулами

$$a = \frac{n \sum_{k=1}^n x_k^2 y_k - \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{k=1}^n y_k \right)}{n \sum_{k=1}^n x_k^4 - \left(\sum_{k=1}^n x_k^2 \right)^2}, \quad b = \frac{\sum_{k=1}^n x_k y_k}{\sum_{k=1}^n x_k^2},$$

$$c = \frac{\left(\sum_{k=1}^n x_k^4 \right) \left(\sum_{k=1}^n y_k \right) - \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{k=1}^n x_k^2 y_k \right)}{n \sum_{k=1}^n x_k^4 - \left(\sum_{k=1}^n x_k^2 \right)^2}.$$

Приклад 1. Застосовуючи метод найменших квадратів, скласти рівняння параболи (6), яка проходить найближче до точок

x_k	-4	-3	-2	-1	0	1	2	3	4
y_k	6	3	1	0,3	-0,1	-0,2	0	0,2	1

Розв'язання. У даному випадку значення x_k рівновіддалені,

$$n = 9, \quad \sum_{k=1}^9 x_k = -4 - 3 - 2 - 1 + 0 + 1 + 2 + 3 + 4 = 0.$$

і виконуються умови, які дозволяють знаходити параметри a , b , c за спрощеними формулами. Тому, підставивши значення x_k та y_k з таблиці у ці формули, одержимо

$$a = \frac{9 \cdot 144,81 - 60 \cdot 11,2}{9 \cdot 708 - 3600} = \frac{1303 - 672}{6312 - 3600} = \frac{631,29}{2712} = 0,20003;$$

$$b = \frac{-24 - 9 - 2 - 0,3 - 0,2 + 0,6 + 4}{60} = \frac{-30,9}{60} = -0,515;$$

$$c = \frac{708 \cdot 11,2 - 60 \cdot 144,81}{2712} = \frac{-759}{2712} = -0,27987,$$

тому що

$$\sum_{k=1}^9 x_k^2 y_k = 144,81; \quad \sum_{k=1}^9 x_k^4 = 708; \quad \sum_{k=1}^9 x_k^2 = 60;$$

$$\sum_{k=1}^9 y_k = 11,2; \quad \sum_{k=1}^9 x_k y_k = -30,9.$$

Таким чином, рівнянням шуканої параболи буде

$$Y = 0,20003 \cdot X^2 - 0,515 \cdot X - 0,27987.$$

Питання для самоперевірки та вправи

1. Як визначають статистичні оцінки числових характеристик та умови їх незсунутості, ефективності, обґрунтованості?

2. Вказати числові характеристики вибірки та формули, за якими їх обчислюють.

3. Які властивості має вибіркова середня?

4. Як визначають степеневу середню вибірки, середню квадратичну, середню гармонічну та середню геометричну вибірки?

5. Що називають вибірковим середньо-квадратичним відхиленням (стандартом)?

6. В яких випадках використовують виправлену вибіркову дисперсію і як вона пов'язана із вибірковою дисперсією?

7. В яких випадках обчислюють характеристики вибірки методом добутків? Який порядок дій при використанні цього методу?

8. Які статистичні оцінки називають точковими, інтервальними?

В яких випадках використовують інтервальні оцінки та що вони дозволяють встановити?

9. Який порядок дій знаходження довірчого інтервалу для оцінки математичного сподівання нормального розподілу при відомому та невідомому σ ?

10. Як знаходять об'єм вибірки, який із заданими точністю та надійністю дозволить знайти оцінку математичного сподівання нормально розподіленої випадкової величини?

11. Маємо криву $y = f(x, a_1, a_2, \dots, a_m)$. Що називають вирівнюванням емпіричних даних вздовж цієї кривої?

12. Яка суть методу найменших квадратів знаходження невідомих параметрів функціональної залежності випадкових величин?

Які дії та в якій послідовності треба виконати при застосуванні цього методу?

Розв'язати задачі

1. Вимірювальним приладом, який не має систематичних похибок, були зроблені незалежні вимірювання деякої величини. Знайти незсунуті оцінки математичного сподівання та дисперсії цієї величини за даними результатами вимірювань.

а) 8, 9, 11, 12; б) 4, 6, 10, 20; в) 5, 15, 10, 20; г) 2, 3, 5, 10; д) 2,3,5;
 є) 2504, 2486, 2526, 2495, 2515, 2528, 2492, 2494.

Відповідь:

а) $\overline{x_B} = 10$, $s^2 = 0$; б) $\overline{x_B} = 10$, $s^2 = 0$; в) $\overline{x_B} = 12,5$, $s^2 = 0$;

г) $\overline{x_B} = 5$, $s^2 = 0$; д) $\overline{x_B} = 2505$, $s^2 = 260,3$.

2. Знайти методом добутоків вибірккову середню та вибірккову дисперсію заданої вибірки.

а)

x_k	12	14	16	18	20	22
n_k	5	15	50	16	10	4

б)

x_k	65	70	75	80	85
n_k	2	5	25	15	3

в) 2.30, 2.28, 2.29, 2.28, 2.30, 2.28, 2.32, 2.29, 2.31, 2.32, 2.31, 2.32, 2.30, 2.31, 2.30, 2.28, 2.29, 2.28, 2.30, 2.28, 2.32, 2.29, 2.32, 2.31, 2.30, 2.32, 2.30, 2.31;

г)

x_k	7.14	7.21	7.28	7.35	7.42	7.49	7.56	7.63	7.70
n_k	2	9	24	43	51	37	25	8	1

Відповідь:

а) $\overline{x_B} = 16,46$, $D_B = 4,87$; б) $\overline{x_B} = 76,2$, $D_B = 18,56$;

в) $\overline{x_B} = 2,3008$, $D_B = 0,00013$; г) $\overline{x_B} = 7,416$, $D_B = 0,116$.

3. Знайти з надійністю 0,95 довірчий інтервал оцінки математичного сподівання a нормально розподіленої ознаки X генеральної сукупності, якщо відомі вибірккова середня $\overline{x_B}$, об'єм вибірки n та середнє квадратичне відхилення σ генеральної сукупності.

Розділ 6. Статистичні оцінки параметрів розподілу

а) $\bar{x}_B = 14, n = 25, \sigma = 5$; б) $\bar{x}_B = 10,2, n = 16, \sigma = 4$;

в) $\bar{x}_B = 16,8, n = 25, \sigma = 5$; г) $\bar{x}_B = 2000, n = 1600, \sigma = 40$.

Відповідь: а) (12,04; 15,96); б) (7,63; 12,77); в) (14,23; 19,37); г) (1998,04; 2001,96).

4. Знайти мінімальний об'єм вибірки, при якому з надійністю γ точність оцінки математичного сподівання нормально розподіленої генеральної сукупності буде дорівнювати 0,2.

Відомо середнє квадратичне відхилення генеральної сукупності $\sigma = 1,5$.

а) $\gamma = 0,925$; б) $\gamma = 0,95$; в) $\gamma = 0,978$; г) $\gamma = 0,99$.

Відповідь: а) $n = 179$; б) $n = 216$; в) $n = 298$; г) $n = 380$.

5. За дослідними даними методом найменших квадратів знайти параметри лінійної залежності $Y = aX + b$.

а)

x_k	0,5	1,5	2,5	3,5
y_k	1,32	0,81	0,18	-0,46

б)

x_k	1	2	3	4
y_k	1,1	0,48	-0,1	-0,8

в)

x_k	2	4	6	8
y_k	0,81	2,3	3,4	4,54

г)

x_k	1	3	5	7
y_k	0,48	1,26	2,85	3,96

д)

x_k	0	0,5	1	1,5	2	2,5	3	3,5	4	4,5
y_k	1,67	1,32	1,1	0,81	0,48	0,18	-0,1	-0,46	-0,8	-1,15

Відповідь:

а) $Y = -0,6X + 1,6625$; б) $Y = -0,628X + 1,74$;

в) $Y = -0,6145X - 0,31$; г) $Y = 0,60125X - 0,26825$;

д) $Y = -0,6188X + 1,6972$.

6. Використовуючи метод найменших квадратів, знайти коефіцієнти залежності $y = ax^2 + bx + c$ за даними вибірки

а)

x_k	0	1	2	3	4	5	6	7	8
y_k	0,1	0,48	0,81	1,26	2,3	2,85	3,4	3,96	4,54

б)

x_k	-4	-3	-2	-1	0	1	2	3	4
y_k	6	3	1	0,3	-0,1	-0,2	0	0,2	1

Відповідь:

а) $y = 0,06835x^2 + 0,5576x + 0,66$;

б) $y = 0,288x^2 - 0,515 - 0,275$.

Розділ 7

СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

*Гіпотези корисні
у багатьох випадках.
Вони бувають різні.
Як їх перевіряти?*

7.1. Статистичні гіпотези та їх різновиди

Часто необхідно знати закон розподілу генеральної сукупності. Якщо закон розподілу невідомий, але є міркування для припущення його певного вигляду A , наприклад, розподіл рівномірний, показниковий або нормальний, тоді висувають гіпотезу:

генеральна сукупність розподілена за законом A .

У цій гіпотезі йде мова про вигляд невідомого розподілу.

Іноді закон розподілу генеральної сукупності відомий, але його параметри (числові характеристики) невідомі. Якщо є міркування припустити, що невідомий параметр θ дорівнює певному значенню θ_0 то висувають гіпотезу: $\theta = \theta_0$. Ця гіпотеза вказує припущену величину параметра відомого розподілу.

Можливі також інші гіпотези: про рівність параметрів двох різних розподілів, про незалежність вибірок, про те, що у листопаді 2000 року буде кінець світу, та багато інших.

Означення 1. *Статистичними називають гіпотези про вигляд розподілу генеральної сукупності або про параметри відомих розподілів.*

Наприклад, статистичними будуть гіпотези:

- генеральна сукупність розподілена за нормальним законом;
- дисперсії двох сукупностей, розподілених за законом Пуассона, рівні між собою.

Відомо, що на творчі можливості людей впливають не тільки гени та умови життя, але й космос. Розглянемо гіпотези:

- значна частина народжених у першому півріччі має краще розвинену ліву частину мозку, яка здійснює логічне мислення;
- значна частина людей, народжених у другому півріччі, має краще розвинену праву частину мозку, яка здійснює образне мислення.

Ці гіпотези не статистичні, бо в них не йде мова ні про вигляд, ні про параметри розподілу. Але для вказаної ситуації можна сформулювати декілька статистичних гіпотез.

Разом з припущеною гіпотезою завжди можна розглядати протилежну їй гіпотезу. Якщо припущена гіпотеза була відхилена, тоді має місце протилежна гіпотеза. Отже, ці гіпотези доцільно відрізняти.

Означення 2. *Основною (нульовою) називають припущену гіпотезу і позначають H_0 .*

Означення 3. *Альтернативною (конкурентною) називають гіпотезу, що суперечить основній, її позначають H_1 .*

Наприклад, якщо $H_0 : M(X) = 6$, то $H_1 : M(X) \neq 6$.

Гіпотези можуть містити тільки одне або більше одного припущення.

Означення 4. *Гіпотезу звать простою, якщо вона містить лише одне припущення.*

Наприклад, якщо λ – параметр показникового розподілу, то гіпотеза $H_0 : \lambda = 5$ буде проста.

Означення 5. *Гіпотезу називають складною, якщо вона складається із скінченної або нескінченної кількості простих гіпотез.*

Наприклад, гіпотеза H_0 : математичне сподівання нормального розподілу дорівнює 2 – складна гіпотеза тому, що середнє квадратичне відхилення σ невідоме і може приймати будь-яке значення.

Гіпотеза H : показниковий розподіл має параметр $\lambda > 2$ складається із нескінченної множини гіпотез $H_k : \lambda = c_k$, де $c_k > 2$, $k = 1, 2, \dots$

7.2. Похибки перевірки гіпотез

Статистична гіпотеза, яка висунута, може бути правильною або неправильною, тому виникає необхідність її перевірки.

Перевірка гіпотези здійснюється за даними вибірки, тобто статистичними методами. Тому перевірку гіпотези за даними вибірки називають **статистичною**.

При перевірці статистичної гіпотези за даними **випадкової вибірки** можна зробити хибний висновок. При цьому можуть бути похибки першого та другого роду.

Означення 1. *Якщо за висновком буде відкинута правильна гіпотеза, то кажуть, що це **похибка першого роду**.*

Означення 2. *Якщо за висновком буде прийнята неправильна гіпотеза, то кажуть, що це **похибка другого роду**.*

Відмітимо, що наслідки цих похибок можуть бути різними. Наприклад, якщо відкинути правильну гіпотезу «продовжити будівництво м'ясокомбінату», то ця похибка першого роду буде сприяти матеріальним витратам.

Якщо прийняти неправильну гіпотезу «продовжити будівництво, не враховуючи можливості обвалу об'єкта будівлі», то внаслідок цієї похибки другого роду можуть загинути люди.

Означення 3. *Імовірність здійснити похибку першого роду позначають α і називають **рівнем значущості**.*

Найчастіше рівень значущості приймають рівним 0,05 або 0,01. Якщо прийнято рівень значущості рівним 0,05, то це означає, що в п'яти випадках із 100 ми ризикуємо одержати похибку першого роду (відкинути правильну гіпотезу).

Зауваження. *При контролі якості продукції імовірність признати нестандартними стандартні вироби називають **ризиком виробника**, а імовірність признати придатними браковані вироби називають **ризиком споживача**.*

7.3. Критерії узгодження для перевірки гіпотез

7.3.1. Статистичний критерій перевірки основної гіпотези

Перевірку статистичної гіпотези можна здійснити лише з використанням даних вибірки. Для цього слід обрати, деяку випадкову статистичну характеристику (вибіркову функцію), точний або наближений розподіл якої відомий, і за допомогою цієї характеристики перевірити основну гіпотезу.

Означення 1. *Статистичним критерієм узгодження перевірки гіпотези (або просто критерієм) називають випадкову величину K , розподіл якої (точний або наближений) відомий і яка застосовується для перевірки основної гіпотези.*

Зауваження 1. *В Означенні 1 не враховується вид розподілу статистичної характеристики.*

Якщо статистична характеристика розподілена нормально, то критерій позначають не літерою K , а літерами U або Z . Якщо статистична характеристика розподілена за законом Фішера-Снедекора, то її позначають F .

У випадку розподілу статистичної характеристики за законом Стьюдента її позначають T , а у випадку закону «хі-квадрат» – χ^2 .

Наприклад, для перевірки гіпотез про рівність дисперсії двох нормальних генеральних сукупностей за статистичну характеристику K вибирають відношення виправлених вибірових дисперсій

$$F = \frac{S_1^2}{S_2^2}.$$

В різних дослідях дисперсія буде приймати різні, наперед невідомі значення, тому ця величина випадкова. Вона розподілена за законом Фішера-Снедекора.

Означення 2. *Спостереженим значенням критерію узгодження називають значення відповідного критерію, обчислене за даними вибірки.*

Наприклад, якщо за даними вибірок із двох нормальних генеральних сукупностей знайдено виправлені вибіркові дисперсії $S_1^2 = 18$ та $S_2^2 = 6$, тоді спостереженим значенням критерію узгодження буде

$$F_{cn} = \frac{18}{6} = 3.$$

Існує багато критеріїв узгодження. Наприклад, найбільш точний (асимптотично) критерій Неймана-Пірсона використовує нерівності або відношення функцій правдоподібності.

7.3.2. Критична область

Після обрання певного критерію узгодження, множину усіх його можливих значень поділяють на дві підмножини, що не перетинаються: одна з них містить значення критерію, при яких основна гіпотеза відхиляється, а друга – при яких вона приймається.

Означення 3. *Критичною областю називають сукупність значень критерію, при яких основна гіпотеза відхиляється.*

Означення 4. *Областю прийняття гіпотези (областю допустимих значень) називають множину значень критерію, при яких гіпотезу приймають.*

Критерій узгодження K – одновимірна випадкова величина, усі її можливі значення належать деякому інтервалу. Тому критична область та область прийняття гіпотези також будуть інтервалами, а це означає, що існують точки, які ці інтервали відокремлюють.

Означення 5. *Критичними точками (межами) критерію K називають точки k_{kp} , які відокремлюють критичну область від області прийняття гіпотези.*

Розрізняють одnobічну (правобічну та лівобічну) та двобічну критичні області.

Означення 6. *Правобічною називають критичну область, що визначається нерівністю $K > k_{kp}$ де – додатне число.*

Означення 7. *Лівобічною називають критичну область, що визначається нерівністю де – від'ємне число.*

7.3.3. Знаходження критичних областей

Щоб знайти однобічну критичну область, треба знайти критичну точку k_{kp} . Для цього задають достатньо малу імовірність – **рівень значущості** α , а потім шукають критичну точку з врахуванням вимоги

$$P(K > k_{kp}) = \alpha$$

у випадку правобічної критичної області або

$$P(K < k_{kp}) = \alpha$$

у випадку лівобічної критичної області.

У випадку двобічної критичної області повинно виконуватись тотожність

$$P(K < k_1) + P(K > k_2) = \alpha.$$

Для кожного критерію узгодження є відповідні таблиці (дивись, наприклад, таблиці додатка), які дозволяють знайти таку точку k_{kp} яка задовольняє потрібну умову.

При знаходженні критичної області доцільно враховувати потужність критерію.

Означення 8. *Потужністю критерію називають імовірність належності критерію критичній області при умові, що правильна альтернативна гіпотеза.*

Іншими словами, потужність критерію є імовірність того, що основна гіпотеза буде відхилена, якщо альтернативна гіпотеза правильна.

Якщо рівень значущості α вже обрано, то критичну область доцільно будувати так, щоб потужність критерію була максимальною. Виконання цієї вимоги забезпечує мінімальну імовірність помилки другого роду.

Зауваження 2. *Єдиний спосіб одночасного зменшення імовірностей помилок першого та другого роду це збільшення об'єму вибірки.*

7.3.4. Порядок дій при перевірці статистичних гіпотез

Для перевірки правильності основної статистичної гіпотези H_0 необхідно:

- 1) визначити гіпотезу H_1 , альтернативну до гіпотези H_0 ;
- 2) обрати статистичну характеристику перевірки;
- 3) визначити допустиму імовірність похибки першого роду, тобто рівень значущості α ;
- 4) знайти за відповідною таблицею критичну область (критичну точку) для обраної статистичної характеристики.

До критичної області належать такі значення статистичної характеристики, при яких гіпотеза H_0 відхиляється на користь альтернативної гіпотези H_1 .

Підкреслимо, що між рівнем значущості α та критичною областю існує такий зв'язок: *якщо гіпотеза H_0 правильна, то з імовірністю значення вибіркової функції будуть належати критичній області.*

Так, при перевірці гіпотези про рівність дисперсій двох нормальних сукупностей при альтернативній

$$H_1 : D(X) > D(Y)$$

треба знайти спостережене значення критерія Фішера-Снедекора, тобто

$$F_{cn} = \frac{S_1^2}{S_2^2},$$

а потім з таблиці критичних точок цього розподілу по заданому рівню значущості α та степенях вільності $k_1 = n_1 - 1$ та $k_2 = n_2 - 1$ знайти $F_{kp}(\alpha; k_1; k_2)$.

Якщо $F_{cn} < F_{kp}$, то гіпотеза H_0 приймається.

Якщо $F_{cn} > F_{kp}$, то H_0 відхиляють.

7.4. Деякі критерії перевірки статистичних гіпотез

7.4.1. Перевірка гіпотези про рівність математичних сподівань нормальних генеральних сукупностей

Нехай дві нормально розподілені генеральні сукупності мають рівні дисперсії, а математичні сподівання можуть бути різними.

Із сукупностей зробили вибірку об'єму n_1 та n_2 і знайшли вибіркові середні \bar{x}_1 та \bar{x}_2 , а також виправлені дисперсії S_1 та S_2 , відповідно.

Потрібно перевірити гіпотезу H_0 : *різниця математичних сподівань цих сукупностей дорівнює числу c_0*

$$H_0 : a_1 - a_2 = c_0 .$$

Альтернативна гіпотеза буде

$$H_0 : a_1 - a_2 \neq c_0 .$$

Для перевірки гіпотези в якості статистичної характеристики (вибіркової функції) візьмемо функцію

$$v_{cn} = \frac{\bar{x}_1 - \bar{x}_2 - c_0}{\sqrt{\frac{n_1 \cdot S_1^2 + n_2 \cdot S_2^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} , \quad (1)$$

яка розподілена за законом Стюдента з степенями вільності, що дорівнюють $n_1 + n_2 - 2$.

Для заданого рівня значущості α можна знайти критичну область для статистичної характеристики v з врахуванням альтернативної гіпотези H_1 .

Приклад 1. Підприємство виготовляє однакові деталі двома способами. Першим способом виготовлено 10 деталей, витрати сировини були такими

$$1.4, 1.6, 1.2, 1.5, 1.4, 1.6, 1.5, 1.8, 1.1, 1.4.$$

Розділ 7. Статистична перевірка гіпотез

Другим способом виготовлено 6 деталей, витрати сировини були такими

$$1.8, 1.7, 1.9, 1.3, 1.6, 1.5.$$

Припускаючи, що дисперсія витрат сировини однакова, при рівні значущості $\alpha = 0,02$ перевірити гіпотезу

$$H_0 : a_1 - a_2 = 0$$

при альтернативній гіпотезі

$$H_1 : a_1 - a_2 \neq 0.$$

Розв'язаним. Треба перевірити гіпотезу про рівність математичних сподівань двох нормальних генеральних сукупностей. Згідно з гіпотезою H_1 критична область буде двобічною і визначається умовою

$$P\left(v > v_{\frac{\alpha}{2}}\right) = P\left(v < -v_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2},$$

де статистична характеристика v визначена формулою (1).

Степінь вільності дорівнює

$$n_1 + n_2 - 2 = 10 + 6 - 2 = 14.$$

З таблиці критичних значень $v_{\frac{\alpha}{2}}$ для 14 степеней вільності одержимо

$$\left(v_{\frac{\alpha}{2}}\right)_{kp} = 2,6.$$

За даними вибірки можна знайти

$$\bar{x}_1 = 1,45; \quad S_1^2 = 0,04; \quad \bar{x}_2 = 1,63; \quad S_2^2 = 0,05.$$

Тепер за формулою (1) одержимо

$$v_{cn} = \frac{1,45 - 1,63}{\sqrt{\frac{10 \cdot 0,04 + 6 \cdot 0,05}{10 + 6 - 2} \cdot \left(\frac{1}{10} + \frac{1}{6}\right)}} = \frac{-0,18}{0,115} = -1,58.$$

Отже, v_{cn} не належить до критичної області, тому гіпотеза H_0 може бути прийнята.

7.4.2. Критерій дисперсійного аналізу

Нехай є N нормально розподілених генеральних сукупностей з рівними дисперсіями та, можливо, з різними математичними сподіваннями.

Із кожної сукупності робимо вибірку об'єму

$$\{n_i\}, \quad i = 1, 2, \dots, N,$$

тоді $\sum_{i=1}^n n_i = n$ – об'єм усієї вибірки.

Позначимо j варіанту випадкової величини X з i -тої сукупності x_{ij} . Тоді середня арифметична вибірки із i -тої сукупності буде

$$x_i = \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} x_{ij}, \quad (2)$$

а середня усієї вибірки буде

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^N x_i \cdot n_i. \quad (3)$$

При рівні значущості α треба перевірити основну гіпотезу про рівність математичних сподівань розглядаємих сукупностей

$$H_0 : a_1 = a_2 = \dots = a_N.$$

При рівності дисперсій статистична характеристика буде мати розподіл Фішера з $N - 1$ та $n - N$ степенів вільності. Тому в якості статистичної характеристики для перевірки цієї гіпотези візьмемо функцію

$$F = \frac{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 n_i}{\frac{1}{n-N} \sum_{i=1}^N \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2}. \quad (4)$$

Критичну область у цьому випадку знаходять з урахуванням умови

Розділ 7. Статистична перевірка гіпотез

$$P(F > f_a) = a,$$

де f_a – критичне (і табульоване) значення розподілу Фішера.

Приклад 2. Є дані про вартість (в тис. гривень) проданих трьох видів виробів певним магазином в окремі дні тижня

вівторок	середа	четвер	п'ятниця	субота
10,2	10,8	10,7	13,0	12,0
11,5	9,8	11,5	13,2	11,5
12,0	12,1	12,0	11,5	11,8

Припускаючи нормальний закон розподілу одержаної суми кожного дня та рівність дисперсій, перевірити гіпотезу H_0 :

$a_1 = a_2 = a_3 = a_4 = a_5$ при рівні значущості $\alpha = 0,05$.

Розв'язання. Умови прикладу дозволяють застосувати до розв'язання задачі критерій дисперсійного аналізу.

У цьому випадку маємо:

$$N = 5; n_i = 3, i = 1, 2, \dots, 5; n = 15.$$

За формулами (2) та (3) знаходимо

$$x_1 = 11,2, x_2 = 10,8, x_3 = 11,4, x_4 = 12,6, x_5 = 11,8, \bar{x} = 11,6.$$

Зробимо обчислення сум, що входять до формули (4)

$$\sum_{i=1}^5 (x_i - \bar{x})^2 \cdot n_i = 0,48 + 1,92 + 0,12 + 3 + 0,12 = 5,64,$$

$$\sum_{i=1}^5 \sum_{j=1}^3 (x_{ij} - \bar{x})^2 = 1,92 + 0,01 + 0,16 + 1 + 3,24 + 0,25 + 0,81 + 0,01 +$$

$$+ 0,16 + 1,96 + 2,56 + 0,01 + 0,15 + 0,01 + 0,04 = 12,34.$$

Тепер за формулою (4) знайдемо значення статистичної характеристики

$$F_{cn} = \frac{\frac{1}{4} \cdot 5,64}{\frac{1}{10} \cdot 12,34} = 1,14.$$

Із таблиці критичних значень розподілу Фішера із степенями вільності $N - 1 = 5 - 1 = 4$ та $n - N = 15 - 5 = 10$ і рівнем значущості $\alpha = 0,05$ знаходимо

$$F_{kp} = f_{0,05} = 3,48.$$

Одержали, що $F_{cn} = 1,14 < F_{kp} = 3,48$, тому гіпотеза H_0 може бути прийнята.

7.4.3. Критерій узгодження Пірсона (χ^2)

Критерій узгодження Пірсона (χ^2 -квадрат) ефективно використовують для перевірки гіпотези про розподіл генеральної сукупності, тобто що розподіл випадкової величини має певний функціональний вираз.

Обмежимося застосуванням цього критерію для перевірки гіпотези про нормальний розподіл генеральної сукупності.

Нехай вибірка має такий розподіл об'єму n

варіанти x_k	x_1	x_2	...	x_m
частоти n_k	n_1	n_2	...	n_m

або

інтервали I_k	(x_1, x_2)	(x_2, x_3)	...	(x_m, x_{m+1})
частоти n_k	n_1	n_2	...	n_m

Потрібно з рівнем значущості α перевірити основну гіпотезу H_0 : **генеральна сукупність розподілена нормально.**

Критерієм перевірки цієї гіпотези беруть випадкову величину χ^2 , яка у різних випробуваннях приймає різні, наперед невідомі значення.

Критичне значення цієї випадкової величини залежить від рівня значущості α та степенів вільності її розподілу k

$$\chi_{kp}^2 = \chi^2(\alpha, k).$$

Ці критичні значення табульовані (дивись Таблицю 5 у додатку) для різних α та k .

Для розподілу генеральної сукупності за нормальним законом степінь вільності буде

$$K = m - 3,$$

де m – кількість варіант вибірки або часткових інтервалів варіант.

Правило Пірсона. *Щоб при заданому рівні значущості α перевірити основну гіпотезу H_0 : генеральна сукупність розподілена нормально, треба*

- 1) обчислити теоретичні частоти n'_k для варіант вибірки;
- 2) обчислити спостережене значення критерія χ^2 за формулою

$$\chi_{cn}^2 = \sum_{k=1}^m \frac{(n_k - n'_k)^2}{n'_k}; \quad (6)$$

- 3) знайти степінь вільності χ^2 за формулою (5);
- 4) знайти з таблиці критичну точку χ_{kp}^2 , яка відповідає заданому рівню значущості α та степені вільності k ;
- 5) порівняти χ_{cn}^2 та χ_{kp}^2 зробити висновок:

якщо $\chi_{cn}^2 < \chi_{kp}^2$, то гіпотезу H_0 треба прийняти;

якщо $\chi_{cn}^2 > \chi_{kp}^2$ то гіпотезу треба відхилити.

Приклад 3. При рівні значущості перевірити $\alpha = 0,05$ гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні та теоретичні частоти

n_k	6	13	38	74	106	85	30	14
n'_k	3	14	42	82	99	76	37	13

Розв'язання. У даному випадку теоретичні частоти n'_k задані, кількість варіант вибірки $m = 8$. Тому за формулою (5) знаходимо степінь вільності $k = 8 - 3 = 5$.

З таблиці критичних точок розподілу $\chi^2(\alpha, \bar{k})$ для $\alpha = 0,05$ та $k = 5$ знаходимо $\chi_{kp}^2 = 11,1$.

Для обчислення χ_{cn}^2 за формулою (6) використаємо розрахункову таблицю

n_k	n'_k	$(n_k - n'_k)^2$	$(n_k - n'_k)^2$	$\frac{(n_k - n'_k)^2}{n'_k}$
6	3	3	9	3
13	14	-1	1	0,07
38	42	-4	16	0,38
74	82	-8	64	0,78
106	99	7	49	0,49
85	76	9	81	1,07
30	37	-7	49	1,32
14	13	1	1	0,08
				$\chi_{cn}^2 = 7,19$

Таким чином, $\chi_{cn}^2 = 7,19 < 11,1 = \chi_{kp}^2$ тому за правилом Пірсона гіпотезу H_0 слід прийняти. Отже, дані вибірки узгоджуються з гіпотезою H_0 , тому що розбіжність емпіричних та теоретичних частот незначна.

Знаходження теоретичних частот нормального розподілу

Згідно з класичним означенням імовірності

$$p_k = \frac{n'_k}{n} \Rightarrow n'_k = p_k \cdot n, \quad k = 1, 2, \dots, m.$$

Отже, для знаходження теоретичних частот n'_k треба знайти імовірність

$$p_k = P(X = x_k) \quad \text{або} \quad p_k = P(x_k < X < x_{k+1}),$$

відповідно.

Імовірність $p_k = P(X = x_k)$ можна знайти, використовуючи локальну функцію Лапласа $\varphi(x)$ та дані вибірки за формулою

$$p_k = P(X = x_k) = \frac{h}{\sigma_B} \varphi(u_k),$$

де

$$h = x_{k+1} - x_k; \quad u_k = \frac{x_k - \bar{x}_B}{\sigma_B}; \quad \varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}},$$

варіанти x_k рівновіддалені.

Імовірність $p_k = P(x_{k-1} < X < x_k)$ можна знайти, використовуючи інтегральну функцію Лапласа $\Phi(x)$ за формулою

$$p_k = P(x_k < X < x_{k+1}) = \Phi\left(\frac{x_{k+1} - \bar{x}_B}{\sigma_B}\right) - \Phi\left(\frac{x_k - \bar{x}_B}{\sigma_B}\right).$$

Питання для самоперевірки та вправи

1. Які гіпотези називають статистичними, основною та альтернативною, простою та складною?
2. Що таке похибки першого та другого роду перевірки статистичної гіпотези?
3. Який смисл рівня значущості α ?
4. Що називають статистичним критерієм, критичною областю та критичною точкою перевірки гіпотези?
5. Як перевіряють гіпотезу про рівність дисперсій двох нормальних сукупностей?
6. Який смисл потужності критерія перевірки гіпотези?
7. Вказати порядок дій при перевірці гіпотез.
8. Як здійснюють перевірку гіпотези про рівність математичних сподівань?
9. За яким критерієм здійснюють перевірку гіпотези про рівність математичних сподівань N нормально розподілених сукупностей?
10. Коли застосовують критерій узгодження Персона (χ^2 -квдрат)?
11. Як формулюють правило Пірсона?
12. Як знаходять теоретичні частоти нормального розподілу для перевірки гіпотези за правилом Пірсона?

Розв'язати задачі

1. По даних двох незалежних вибірок об'єму $n_1 = 11$ та $n_2 = 14$ із нормальних сукупностей X та Y знайдені виправлені вибіркові дисперсії $s_1^2 = 0,76$ та $s_2^2 = 0,38$. При рівні значущості $\alpha = 0,05$ перевірити гіпотезу $H_0 : D(X) = D(Y)$ при альтернативній $H_1 : D(X) > D(Y)$.

Відповідь: $F_{cn} = 2$, $F_{kp}(0,5; 10; 13) = 2,67$, H_0 приймається.

Розділ 7. Статистична перевірка гіпотез

2. По даних двох незалежних вибірок об'єму $n_1 = 9$ та $n_2 = 16$ із нормальних сукупностей X та Y знайдені виправлені вибіркові дисперсії $s_1^2 = 34,02$ та $s_2^2 = 12,15$. При рівні значущості $\alpha = 0,01$ перевірити гіпотезу $H_0: D(X) = D(Y)$ при альтернативній $H_1: D(X) > D(Y)$.

Відповідь: $F_{cn} = 2,8$, $F_{kp}(0,01; 8; 15) = 2,64$, відхиляється.

3. Використовуючи критерій Пірсона (χ^2 -квадрат) з рівнем значущості α , перевірити гіпотезу про нормальний розподіл генеральної сукупності X , якщо відомі емпіричні n_k та теоретичні n'_k частоти.

а) $\alpha = 0,05$;

n_k	15	26	25	30	26	21	24	20	13
n'_k	9,1	16,5	25,3	32	33,9	29,8	22	13,5	7

Відповідь: $\chi_{kp}^2 = 16,9$; $\chi_{cn}^2 = 177,13$, H_0 відхиляємо.

б) $\alpha = 0,01$;

n_k	15	26	25	30	26	21	24	20	13
n'_k	9,1	16,5	25,3	32	33,9	29,8	22	13,5	7

Відповідь: $\chi_{kp}^2 = 21,7$; H_0 відхиляємо.

в) $\alpha = 0,05$;

n_k	8	16	40	72	36	18	10
n'_k	6	18	36	76	39	18	7

Відповідь: $\chi_{kp}^2 = 9,5$; $\chi_{cn}^2 = 3,06$, H_0 приймаємо.

г) $\alpha = 0,05$;

n_k	8	16	40	72	36	18	10
n'_k	6	18	36	76	39	18	7

Відповідь: $\chi_{kp}^2 = 13,3$; H_0 приймаємо.

д) $\alpha = 0,05$;

n_k	5	10	20	8	7
n'_k	6	14	18	7	5

Відповідь: $\chi_{kp}^2 = 2,47 < \chi_{cn}^2 = 6$, H_0 приймаємо.

е) $\alpha = 0,05$;

n_k	5	10	20	8	7
n'_k	6	14	18	7	5

Відповідь: $\chi_{kp}^2 = 9,2$; H_0 приймаємо.

4. Використовуючи критерій χ^2 перевірити гіпотезу H_0 : наведений у таблиці розподіл зросту 1000 дорослих чоловіків є вибіркою з нормальної генеральної сукупності з $a = \bar{x}_B$ та $\sigma = \sqrt{\sigma_B^2}$.

Зріст см.	Кількість чоловіків	Зріст см.	Кількість чоловіків
144–147	1	168–171	172
147–150	3	171–174	120
150–153	7	174–177	63
153–156	26	177–180	27
156–159	66	180–183	9
159–162	114	183–186	3
162–165	186	186–189	1
165–168	200	189–192	2

Відповідь: Вибірка узгоджується з гіпотезою,

$$n'_k = 1000 p_k; p_k = P(x_k < X < x_{k+1}) = \Phi\left(\frac{x_{k+1} - a}{\sigma}\right) - \Phi\left(\frac{x_k - a}{\sigma}\right).$$

Частина III

**КОМП'ЮТЕРНИЙ
АНАЛІЗ**

*Щоб у новому столітті
бути добрим фахівцем,
застосовувай комп'ютер
та можли вості Excel.*

Розділ 8

РОЗВ'ЯЗАННЯ ЗАДАЧ ЗАСОБАМИ EXCEL

8.1. Комп'ютерне розв'язання задач теорії імовірностей та математичної статистики у бізнесі

8.1.1. Огляд основних задач, що вирішені за допомогою Excel

Надзвичайне розповсюдження персональних ЕОМ в останній час має своїм наслідком величезний вплив комп'ютерного аналізу у питаннях управління, різного роду дослідженнях та, зокрема, у вирішенні завдань теорії імовірностей та статистики у бізнесі.

Комп'ютери звільняють менеджера чи бізнесмена від трудомістких ручних обчислень. Завдяки цьому дослідник у галузі бізнесу може більше приділити уваги якісній підготовці статистичних даних і їх інтерпретації.

Поряд із широким розповсюдженням персональних комп'ютерів з'явилося багато програмних пакетів, орієнтованих на задачі теорії імовірностей та статистики.

У цьому підручнику всі обчислення по статистиці орієнтовані на електронну таблицю MS Excel 2003 (скорочено Excel), що входить у пакет програмного забезпечення Microsoft Office 2003, Windows XP (русіфікована версія). В програмному забезпеченні Excel передбачена можливість вирішення багатьох важливих задач теорії імовірностей та математичної статистики. При цьому забезпечується висока точність обчислень, можливість роботи з дуже великим об'ємом статистичних даних.

У підручнику засобами Excel повністю вирішені наступні важливі задачі теорії імовірності та математичної статистики:

- побудова простої випадкової вибірки заданого об'єму по заданій скінченій вибірці;
- впорядкування і організація даних у вигляді розподілів частот, накопичених частот і т.д.;
- подання отриманих розподілів у графічному вигляді: гістограм, полігонів частот і т.д.;
- обчислення числових характеристик вибірки: середнього значення, дисперсії і середнє квадратичного відхилення;
- обчислення, пов'язані з комбінаторикою;
- обчислення, пов'язані з біноміальним розподілом, розподілом Пуассона, нормальним розподілом;
- побудова довірчого інтервалу для оцінки математичного сподівання нормального розподілу.

Окремо розглянуто алгоритми Excel, які корисні в статистичних обчисленнях: введення даних, редагування графіків за допомогою майстра діаграм, оформлення результатів та підготовка до друку.

Умовно відзначимо три рівня використання Excel у задачах теорії імовірностей та математичній статистиці.

1. Створення (простих) власних алгоритмів за допомогою Excel.
2. Використання вмонтованих в Excel спеціальних функцій по статистиці.
3. Використання вмонтованого в Excel пакету «Статистический анализ».

Наголосимо, що для свідомого використання можливостей усіх трьох рівнів Excel необхідні початкові знання теорії імовірностей та математичної статистики.

При написанні Частина III автори використали всі три рівня. Перевага віддавалася роботі зі спеціальними статистичними функціями 2-го і 3-го рівнів. Ці функції орієнтовані на розв'язання складних задач, швидко приводять до поставленої мети і, разом з тим, вимагають для свого застосування мінімальних зусиль з боку користувача. Створення власних програм за допомогою Excel надає в лаві випадків більш широкі можливості, однак вимагає навиків у роботі з електронною таблицею.

Excel багато функціональна складна система. Один і той же алгоритм може бути здійснений по різному. Серед них може бути і найкращий (або оптимальний).

Автори при відборі алгоритмів для розв'язання задач теорії імовірностей та математичної статистики в Excel дотримувались наступного «критерію оптимальності»:

алгоритми мають бути найбільше близькими до матеріалу, що викладається у відповідних розділах тексту підручника.

Після засвоєння основних теоретичних положень і формул, студент може створити свої власні алгоритми, оптимальні в будь-якому іншому сенсі (простоти, швидкодії, вигоди і так далі). Надалі, при самостійному поглибленому вивченні матеріалу, треба також більш широко використовувати спеціальні функції статистики і пакет «Статистичний аналіз».

На заключення дамо деякі практичні рекомендації.

- Постійно зберігайте дані на диску, а після завершення роботи – на дискеті.

Оптимально встановити в Excel опцію «Автозберігання», що забезпечить автоматичне, через певний період часу, зберігання даних.

Ці прості засоби дозволять не втратити дані і затрачену працю при випадкових збоях у роботі комп'ютера.

- Кожний окремий етап в алгоритмі корисно розпочинати з нового листа.

Якщо ви зробили на цьому етапі якісь неправильні і безповоротні дії (наприклад, щось стерли, невірно набрали дані і так далі), то простіше весь цей етап розпочати спочатку.

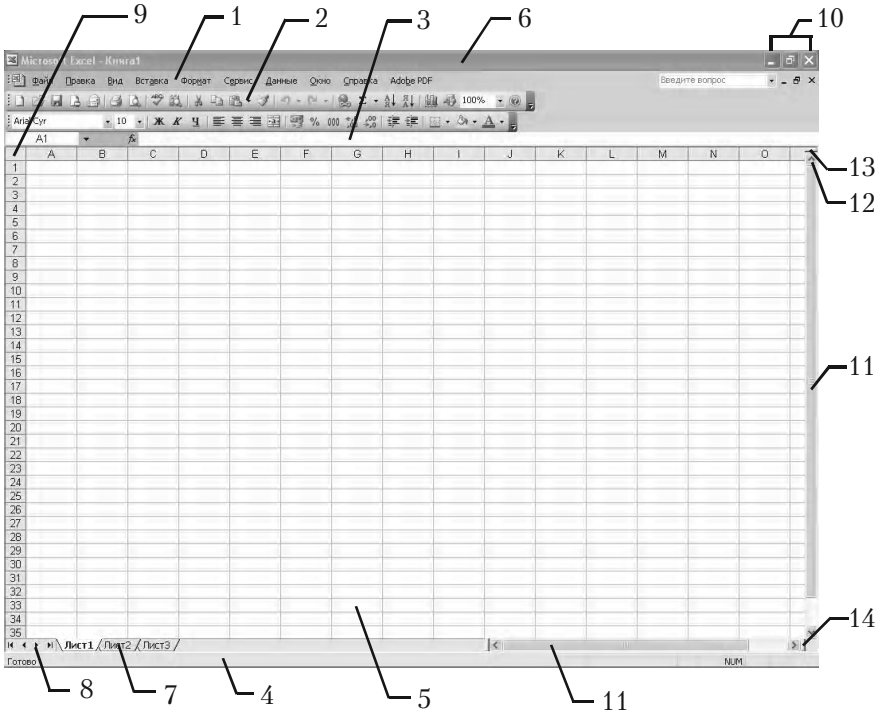
Для цього все витирається з даного листа, і необхідні вихідні дані переносяться з попереднього листа.

Такий простий засіб зекономить багато зусиль та нервів.

8.1.2. Електронна таблиця Excel: елементи вікна та умовні скорочення

При викликові програми Excel з'явиться вікно, зображене на мал. 38.

Розглянемо основні компоненти вікна, з якими ми будемо працювати при розв'язанні задач теорії імовірності та математичної статистики.



Мал. 38. Головне меню, панель інструментів і робочий лист Excel

1 – головне меню;

2 – панель інструментів;

3 – рядок формул;

4 – рядок становища;

5 – робоча книга: основний документ електронної таблиці Excel.

Вона є файлом, що зберігається, із розширенням xls;

6 – заголовок вікна робочої книги. Служить для переміщення вікна за допомогою миші;

7 – ярлик листя;

Лист є «сторінкою» книги. Ярлички листів: Лист 1, Лист 2 і т.д. є їх іменами. Ці стандартні назви можуть бути змінені за бажанням на інші власні імена.

Розділ 8. Розв'язання задач засобами Excel

Кожний лист – це самостійна електронна таблиця. Вона містить 16394 рядки, що нумеруються по порядку, і 256 стовпчиків, що позначаються літерами

A, B, ..., Z, AA, BB, ..., IV.

- 8 – кнопки прокрутки листів;
- 9 – елемент виділення всіх клітин вікна;
- 10 – три кнопки управління окном Excel: «згорнути» (у піктограму), «розгорнути» (на повний екран), «закрити».
- 11 – смуги прокрутки поля робочого листа;
- 12 – стрілки прокрутки поля робочого листа;
- 13 – горизонтальний дільник вікна;
- 14 – вертикальний дільник вікна;

В Excel передбачена можливість введення однієї і тієї самої команди різними засобами. Віддаємо перевагу роботі з мишею.

Будемо застосовувати позначення, що прийняті в [1]. Позначення дій з мишею зведені в наступну таблицю.

ДІЯ	КЛАВІША	
	ліва	права
Один щиглик		
Подвійний щиглик		–
Натиск та переміщення		–

Наприклад, запис A1:B5 відповідає виконанню наступних дій:

1. Розташувати курсор у вічку A1.
2. Натиснути та утримувати ліву кнопку миші.
3. Пересунути курсор у B5.
4. Відпустити ліву кнопку миші.

Далі запис, наприклад, **Правка**, **Спеціальная вставка** відповідає виконанню наступних дій:

1. Виконати команду «Правка».
2. В меню, що з'явилося, виконати команду «Спеціальная вставка».

Далі записи, наприклад, < **Enter** >, < **Del** > відповідають натискуванню на клавіші відповідних назв.

Запис < **Shift** > + < **Enter** > відповідає одночасному натискуванню на клавіші відповідних назв.

Запис **ОК** відповідає натискуванню кнопки **ОК** у діалоговому вікні за допомогою миші.

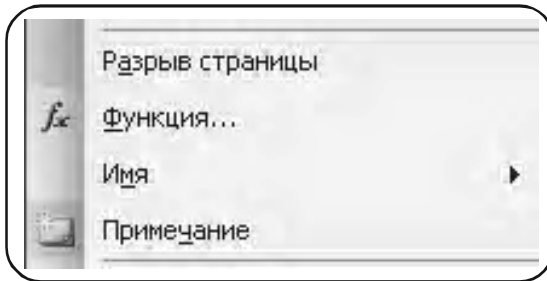
8.1.3. Статистичні функції Excel

У цьому розділі ознайомимось із спеціальними можливостями Excel в задачах теорії імовірності та математичної статистики у бізнесі. Тут можна виділити два рівня.

1. Використання вмонтованих спеціальних функцій по статистиці.
2. Використання вмонтованого пакету «Статистический анализ».

У Excel існує 78 **функцій для статистичних розрахунків**.

Щоб познайомитися з набором цих функцій, слід у вікні Excel вибрати в меню опцію «Вставка», в меню, що з'явилося, вибрати опцію «Функція» (на мал. 39 зображена частина меню).



Мал. 39

У вікні меню, що з'явилося, вибрати опцію «Статистические» (див. мал. 40).

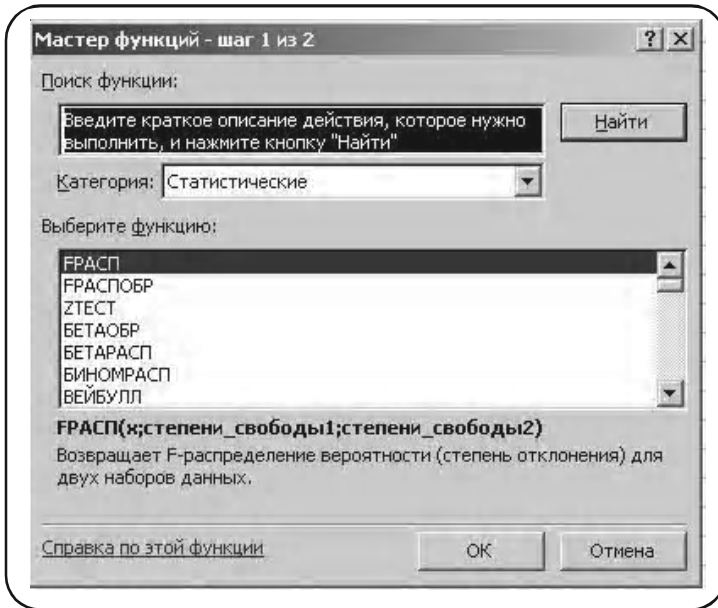
За допомогою клавіш прокрутки можна вибрати будь-яку з наведених функцій. Наприклад, на малюнку 40 вибрана функція ДИСП, що обчислює дисперсію по вибірці. Опис функції можна отримати при натиску клавіші зі знаком «?» (Help).

У склад Excel входить набір засобів аналізу даних, що називається **пакетом аналізу**, призначений для вирішення складних статистичних завдань.

Для проведення аналізу даних з допомогою цих інструментів, слід вказати вхідні дані і вибрати параметри; аналіз буде проведений за

Розділ 8. Розв'язання задач засобами Excel

допомогою підхожої статистичної макрофункції, а результат буде поміщений у вихідний діапазон. Майстер діаграм дозволить представити результати аналізу у графічному вигляді.



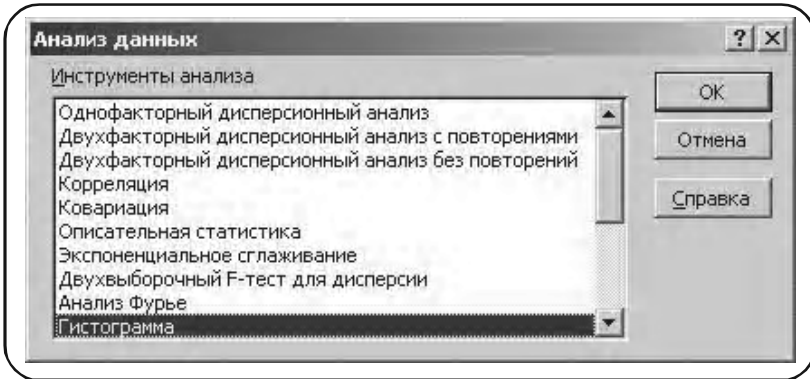
Мал. 40

Щоб ознайомитися з цим пакетом, слід у меню вікна Excel вибрати опцію «Сервис» і в меню, що з'явилося, – опцію «Анализ данных» (на мал. 41 зображена частина меню).



Мал. 41

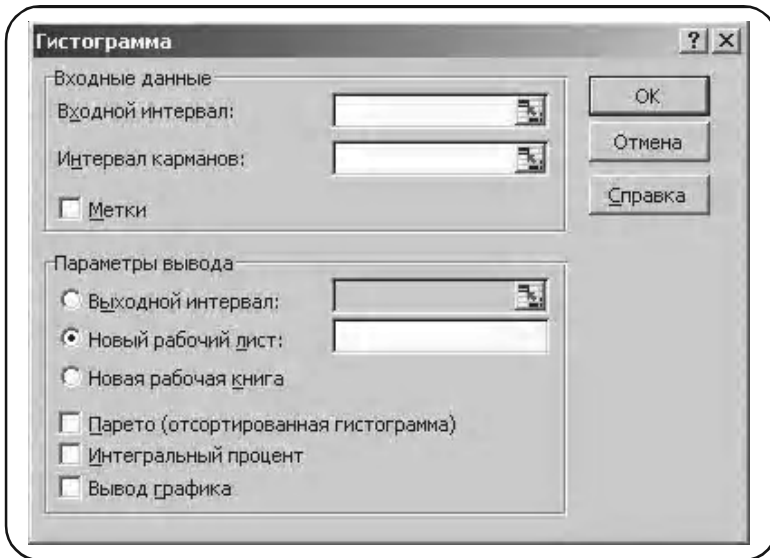
За допомогою клавiші прокрутки можна вибрати будь-яку з наведених функцій аналізу. Наприклад, при виборі опції «Гистограма» (див. мал. 42).



Мал. 42

На екрані з'явиться нове меню «Гистограмма» (див. мал. 43).

За допомогою цієї функції аналізу створюється гiстограма по вибірковим даним. Опис функції можна отримати при натиску клавiші «Справка».



Мал. 43

8.2. Здійснення простої випадкової вибірки за допомогою функції «Виборка»

Теоретичні положення: розділ 5, пункт 5.5.3, стор. 162.

Функція «Виборка» із пакету «Анализ данных» використовується для побудови простої випадкової вибірки заданого об'єму по генеральній сукупності.

Приклад 1. За даними таблиці А середньомісячної платні співробітників фірми *N* (див. стор. 176) побудувати просту випадкову вибірку, наприклад, об'єму $n = 10$. (При вирішенні цієї задачі обмежись малим об'ємом вибірки для спрощення викладення матеріалу).

Розв'язання. Нехай вихідна Таблиця А (стор. 163) вже набрана в Excel і представлена на мал. 44.

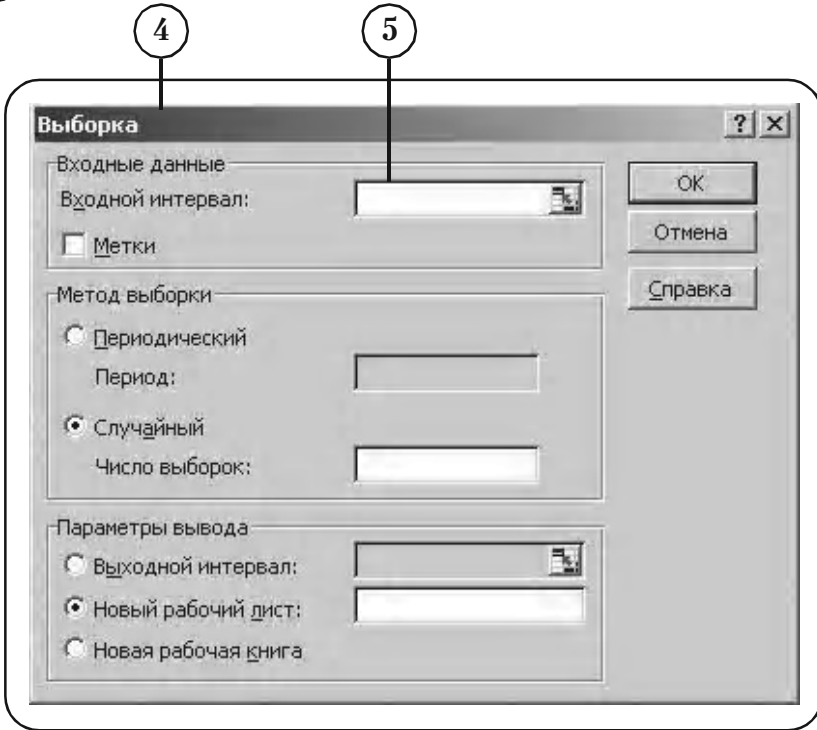
Таблиця А у вікні Excel, що розділено на дві частини: ми бачимо водночас початок і кінець таблиці

	A	B	C	D	E	F	G	H
1	Номер	Бал	Вибірка					
2	1	30						
3	2	38						
4	3	32						
5	4	49						
6	5	33						
7	6	42						
8	7	60						
9	8	30						
10	9	34						
11	10	61						
180	179	38						
181	180	29						

Мал. 44

① Массив даних – це вічка B2:B181.

- 2 В якості вихідного інтервалу прийнемо вічка C2:C11. Вибрано-му інтервалу вічок дамо назву «Вибірка».
- 3 Сервис → Анализ данных → Выборка.
- 4 У вікні з'явилось меню функції «Выборка» (див. мал. 45).



Мал. 45

Заповнюємо вікна меню «Выборка» значеннями, які відповідають нашій задачі.

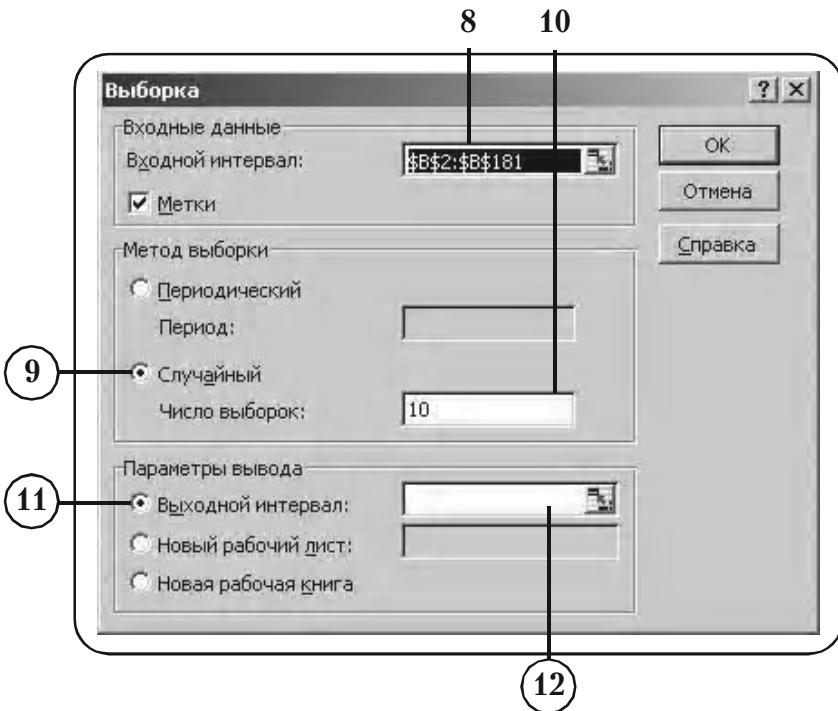
Здійснимо автоматичне внесення до меню «Выборка» адрес, звідки будуть взяті вхідні дані.

- 5 Активізуємо вікно «Входной интервал» (встановлюємо стрілку миші у вікні та натискаємо ліву клавішу).
- 6 ↵ 1 B2 (Активізуємо вічко B2, див. мал. 46).

	А	В	С
1	Номер	Бал	Вибірка
2	1	30	
3	2	38	
4	3	32	
5	4	49	
6	5	33	
7	6	42	
8	7	60	
9	8	30	
10	9	34	
11	10	61	
180	179	38	
181	180	29	

Мал. 46

- 7 < Shift > + \leftarrow 1 В181 (Активізуємо вічко В181 при натиснутій клавіші < Shift >).
- 8 Вічка, звідкіля будуть взяті входні дані, автоматично заносяться у вікно «Входной інтервал» (див. мал. 47).
Призначимо засіб вибірки і кількість елементів у вибірці.
- 9 У меню «Метод выборки» вибираємо опцію «Случайный».
- 10 У вікні «Число выборок» набираємо число 10.
Здійсимо автоматичне внесення адрес, куди будуть поміщені дані виробленої вибірки, до меню «Выборка».
- 11 Вибираємо опцію «Выходной интервал».
- 12 Активізуємо вікно «Выходной интервал» (встановлюємо стрілку миші в вікні і натискаємо ліву клавішу).
- 13 \leftarrow 1 С2 (Активізуємо вічко С2, див. мал. 48).




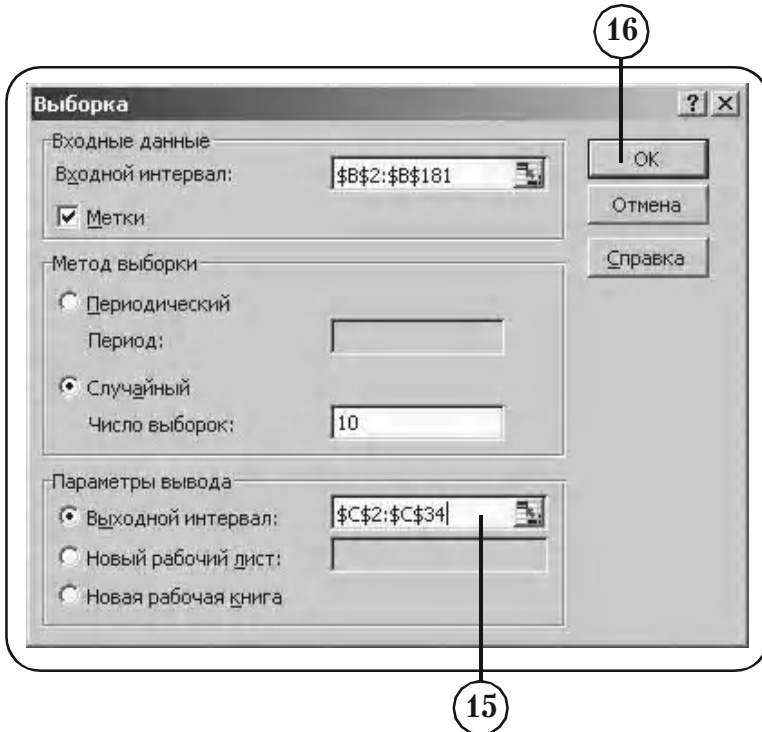
Мал. 47

	А	В	С
1	Номер	Бал	Вибірка
2	1	30	
3	2	38	
4	3	32	
5	4	49	
6	5	33	
7	6	42	
8	7	60	
9	8	30	
10	9	34	
11	10	61	
180	179	38	
181	180	29	

Мал. 48

Розділ 8. Розв'язання задач засобами Excel

- 14 < Shift > +  1 C11 (Активізуємо вічко C11 при натиснутій клавіші < Shift >).
- 15 Вічка, куди будуть поміщені дані виробленої вибірки, автоматично заносяться в вікно «Выходной интервал» (див. мал. 49).



Мал. 49

- 16 ОК.
- 17 У вічках C2:C11 записаний остаточний результат – проста випадкова вибірка об'єму $n = 10$ (див. мал. 50).

Остаточний результат автоматичної побудови простої випадкової вибірки об'єму $n = 10$ за даними вихідної таблиці

	А	В	С
1	Номер	Бал	Вибірка
2	1	30	33
3	2	38	28
4	3	32	42
5	4	49	30
6	5	33	72
7	6	42	62
8	7	60	26
9	8	30	57
10	9	34	30
11	10	61	32
180	179	38	
181	180	29	

Вправи для самостійної роботи

1. Проведіть просту випадкову вибірку об'єму $n = 20$ по даних Таблиці А числа балів, що набрані при тестуванні 180 співробітників фірми на сторінці 176.

2. Проведіть просту випадкову вибірку об'єму $n = 15$ по даних Таблиці В середньомісячної платні 100 співробітників фірми N на сторінці 179.

8.3. Організація даних вибіркової сукупності

Теоретичні положення: параграф 5.6 розділ 5.

8.3.1. Побудова варіаційного ряду

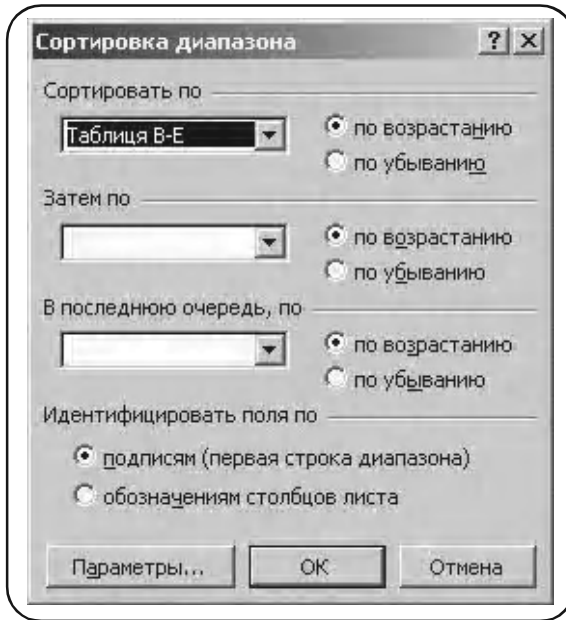
Запис вхідних даних у вигляді таблиці у зростаючому порядку (тобто у вигляді варіаційного ряду) дає багато корисної інформації: максимальне та мінімальне значення варіант, межі їхньої зміни. Ця інформація використовується при підрахунку потрібного числа класів інтервалів при побудові згуртованого розподілу частот. Нижче розглянута задача побудови варіаційного ряду вибірки.

- 1 Представимо дані Таблиці В (стор. 165) у вигляді Таблиці В-Е (див. мал. 51) з одним стовпцем.
- 2 ↗ A3:A102 (виділення даних).
- 3 **Данные → Сортировка.**
У вікні з'являється меню «Сортировка даних» (див. мал. 52).

Вибірка з середньомісячної платні співробітників фірми N (Таблиця В) у вікні Excel набрано в один стовпчик

	А	В	С	Д	Е	F	Г
1	Таблиця В-Е						
2	Платня						
3	338						
4	336						
5	312						
6	322						
7	381						
8	302						
99	322						
100	338						
101	324						
102	323						

Мал. 51



Мал. 52

Результат упорядкування даних Таблиці В-Е по зростанню

↓

	А	В	С	Д	Е	Ф
1	Таблиця В-1-Е					
2	Платня					
3	282					
4	290					
5	290					
6	292					
7	292					
8	296					
99	368					
100	381					
101	381					
102	386					

Мал. 53

Вибираємо опцію «по возрастанию».

4 ОК.

Результат сортування наведений на мал. 53 у вигляді Таблиці В 1-Е.

8.3.2. Побудова розподілів частот за допомогою функції ЧАСТОТА

Функція ЧАСТОТА дає розподіл частот по заданим інтервалам у вигляді вертикального масиву. Вона має формат

ЧАСТОТА {массив-данных; массив_карманов},

«массив_данных» – це масив або посилання на множину даних, для яких обчислюються частоти;

«массив_карманов» – це масив чи посилання на множину інтервалів, в які групуються значення аргументу «массив-данных».

Приклад 1. Для вибірки середньомісячної платні 100 співробітників фірми N, що наведено у Таблиці В (див. мал. 54, а також приклад на стор. 165),

а) побудувати згрупований розподіл частот по інтервалах

280 – 290 310 – 320 340 – 350 370 – 380
 290 – 300 320 – 330 350 – 360 380 – 390
 300 – 310 330 – 340 360 – 370

б) побудувати простий (незгрупований) розподіл частот.

	A	B	C	D	E	F	G	H	I	J	K
1											
2		Таблиця									
3		Вибірка середньомісячної платні 100 співробітників фірми N									
4	338	348	304	314	326	314	324	304	342	308	
5	336	304	302	338	314	304	320	321	322	321	
6	312	323	336	324	312	312	364	366	362	302	
7	322	310	334	292	362	381	304	366	298	304	
8	381	368	304	298	386	290	340	328	316	322	
9	302	314	292	342	321	322	290	332	298	296	
10	296	298	324	338	352	326	318	304	332	322	
11	360	312	331	331	304	316	332	282	342	338	
12	342	322	324	325	302	328	354	330	316	324	
13	334	350	334	324	332	340	324	314	326	323	

Мал. 54

Роз'язання.

Задача а)

У нашому випадку (див. мал. 54, 55).

	A	B	C	D	E	F	G	H	I	J
15	Інтервали платні				Частоти		Масив кишень			
16	280-289						289			
17	290-299						299			
18	300-309						309			
19	310-319						319			
20	320-329						329			
21	330-339						339			
22	340-349						349			
23	350-359						359			
24	360-369						369			
25	370-379						379			
26	380-389						389			

Мал. 55

1 «массив_данных» – це вічка A3: J12.

2 «массив_карманов» – це вічка G16:G26.

Занесені у цей масив дані значать, що частоти варіант вибірки будуть підраховуватись для інтервалів

$$\begin{array}{ll}
 x < 289 & 340 < x < 349 \\
 290 < x < 299 & 250 < x < 359 \\
 300 < x < 309 & 260 < x < 369 \\
 310 < x < 319 & 270 < x < 379 \\
 320 < x < 329 & 280 < x < 389 \\
 330 < x < 339 & x > 390
 \end{array}$$

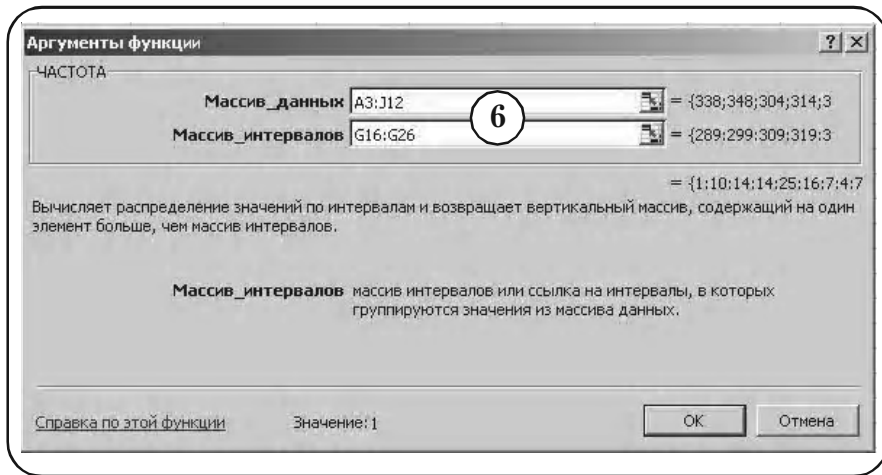
3 «Інтервали платні» – це вічка A16:A26.

4 Виділимо вічка E16:E27 (на одне більше, чим масив кишень), куди будуть виведені частоти.

5 Вставка → Функція → Статистическая → ЧАСТОТА.

6 У меню функції ЧАСТОТА, що з'явилося у вікні, вводимо вічка, де розміщений масив даних і масив кишень (див. мал. 56).

Розділ 8. Розв'язання задач засобами Excel



Мал. 56

7 < Ctrl > + < Shift > + < Enter >.

8 Результат – частоти варіант по виділеним інтервалам, виводиться у вічка E16:E26 (див. мал. 57).

	A	B	C	D	E	F	G	H	I	J
15	Інтервали платні				Частоти		Масив_кишень			
16	280-289				1		289			
17	290-299				10		299			
18	300-309				14		309			
19	310-319				14		319			
20	320-329				25		329			
21	330-339				16		339			
22	340-349				7		349			
23	350-359				4		359			
24	360-369				7		369			
25	370-379				0		379			
26	380-389				2		389			

Мал. 57

Порівняйте результат з отриманим у таблиці ВЗ на стор. 169.

Зауваження 1. *Стовпці «Частоти» і «Масив_кишень» зв'язані між собою і це призводить до певних незручностей, наприклад, не*

можна змінювати вміст масиву. Щоб позбавитися від цих незручностей, проведемо наступні прості операції.

- 9 Виділимо вічка A15:I27 (див. мал. 58).

	A	B	C	D	E	F	G	H	I
15	Інтервали платні			Частоти		Масив кишень			
16	280-289				1		289		
17	290-299				10		299		
18	300-309				14		309		
19	310-319				14		319		
20	320-329				25		329		
21	330-339				16		339		
22	340-349				7		349		
23	350-359				4		359		
24	360-369				7		369		
25	370-379				0		379		
26	380-389				2		389		

Мал. 58

- 10 Правка → Спеціальна вставка → Значення.

- 11 Після нескладних перетворень придамо їй вигляд, що зображено на мал. 59.

	A	B	C	D	E	F	G
1				Таблиця В-3-Е	Масив кишень		
2				Платня	Частоти		
3				280-289	1		
4				290-299	10		
5				300-309	14		
6				310-319	14		
7				320-329	25		
8				330-339	16		
9				340-349	7		
10				350-359	4		
11				360-369	7		
12				370-379	0		
13				380-389	2		
14					100		

Мал. 59

Розділ 8. Розв'язання задач засобами Excel

Задача б)

- 1 Складаємо таблицю показану на мал. 60.

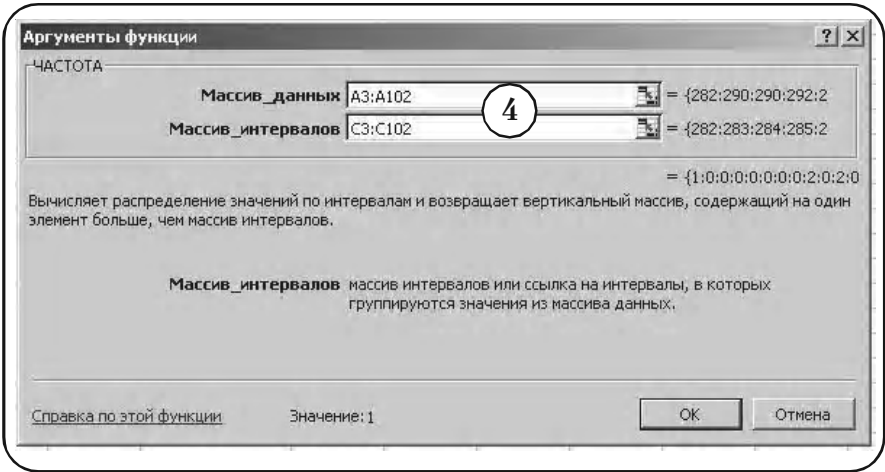
	A	B	C	D	E	F	G	H
1	Таблиця В.1-Е							
2	Платня		Масив_кишень	Частота				
3	282		282					
4	290		283					
5	290		284					
6	292		285					
7	292		286					
98	366		377					
99	368		378					
100	368		379					
101	381		380					
102	381		381					
103								

Мал. 60

В стовпці «Платня» внесена упорядкована по зростанню вибірка. Стовпець «Масив_кишень» заповнений числами від 282 до 381. Стовпець «Частота» передвизначений для виводу частот варіант.

- 2 Виділимо вічка D3:D103 (на одне більше, чим масив кишень), куди будуть виведені частоти.
- 3 Вставка → Функція → Статистическая → ЧАСТОТА.
- 4 У меню функції ЧАСТОТА, що з'явилося у вікні, вводимо адреси вічок, де розміщений масив даних і масив кишень (див. мал. 61).
- 5 < Ctrl > + < Shift > + < Enter >.
- 6 Результат – частоти варіант, виводиться у вічка D3:D103 (див. мал. 62).

Для одержання розподілу частот потрібно з чисел, що входять в «Масив_кишень» витягти ті, що входять у вхідний масив «Платня», тобто усунути рядки, що містять нульові частоти. Це можна зробити за допомогою фільтрації даних.



Мал. 61

	A	B	C	D	E	F	G	H
1	Таблица В.1-Е			Частота				
2	Платня		Массив_кишень	0				
3	282		282	1				
4	290		283	0				
5	290		284	0				
6	292		285	0				
7	292		286	0				
98	366		377	0				
99	368		378	0				
100	368		379	0				
101	381		380	0				
102	381		381	2				

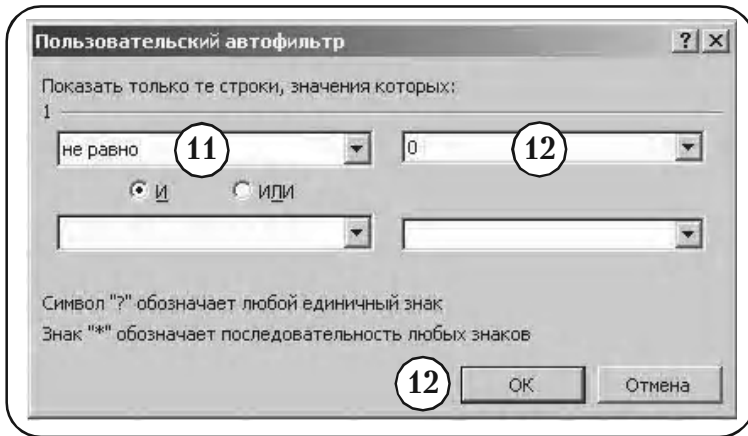
Мал. 62

- 7) Виділимо масив даних у вічках C2:D102.
- 8) Данные > Фильтр > Автофильтр.
- 9) Натискуємо кнопку меню, що з'явилось (див. мал. 63).
- 10) Вибираємо опцію «Условия...».

	A	B	C	D	E	F	G	H
1	Таблица В.1-Е			Частота				
2	Платня		Массив_кишень	Частота				
3	282			282				
4	290			(Все)				
5	290			(Первые 10...)				
6	292			(Условие...)				
7	292			0				
				1				
				2				
				3				
98	366			4				
99	368			5				
100	368			6				
101	381			7				
102	381			9				

Мал. 63

- 11 З'являється меню «Пользовательский автофильтр». Заповнюємо його, як показано на мал. 64.



Мал. 64

- 12 ОК.
- 13 В стовпцях C і D одержуємо шуканий розподіл частоти (див. мал. 65).
- 14 Зважаючи на Зауваження 1, перепишемо отриманий розподіл частоти на окремому листі (див. мал. 66).

Остаточний результат: простий розподіл частоти вибірки

	A	B	C	D	E	F	G	H
1	Таблиця В.1-Е							
2	Платня		Массив_кишень	Частота				
3	282		282					
11	298		290	2				
13	298		292	2				
17	302		296	2				
19	304		298	4				
81	338		360	1				
83	340		362	2				
85	342		364	1				
87	342		366	1				
89	348		368	2				
102	381		381	2				

Мал. 65

	A	B	C	D	E	F	G
1	Платня	Частота					
2	282	1					
3	298	2					
4	298	2					
5	302	2					
6	304	4					
7	304	4					
37	340	2					
38	342	1					
39	342	1					
40	348	2					
41	381	2					
42		100					

Мал. 66

- 15 Скопіюємо отриманий розподіл частоти в Таблицю В.2-Е (див. мал. 67). Таким чином, дістаємо остаточний результат: згрупований розподіл частот і простий (незгрупований) розподіл частот для вибірки середньомісячної платні 100 співробітників фірми N.

Остаточний результат: згрупований розподіл частот і простий (незгрупований) розподіл частот для вибірки

	A	B	C	15 D	E	F	G
1	Таблиця В.2-Е			Таблиця В.3-Е			
2	Платня	Частота		Платня	Частоти		
3	282	1		280-289	1		
4	298	2		290-299	10		
5	298	2		300-309	14		
6	302	2		310-319	14		
7	304	4		320-329	25		
8	304	4		330-339	16		
9	304	9		340-349	7		
10	312	1		350-359	4		
11	312	1		360-369	7		
12	314	4		370-379	0		
13	314	5		380-389	2		
41	348	2					
42	381	2					

Мал. 67

8.3.3. Побудова розподілів частостей та накопичених частот і частостей

Розглянемо побудову розподілів вибірки засобами Excel, виходячи з відповідних означень.

Приклад 2. Для простого і згрупованого розподілів частот, наведених на мал. 67, побудувати:

- а) простий розподіл частостей (відносних частот);
- б) розподіл накопичених частот;
- в) розподіл накопичених частостей.

Розв'язання. Перепишемо Таблицю, зображену на мал. 67, на новий лист і введемо стовпчики «Частоті», «Накопичена частота» та «Накопичена частість» (див. мал. 68).

Вихідна таблиця для побудови розподілів частотей та накопичених частот і частостей

	A	B	C	D	E	F	G
1	Таблиця В.2.1-Е			Таблиця В.3.1-Е			
2	Платня	Частота	Частоти	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1		менше 290	1		
4	298	2		менше 300	10		
5	298	2		менше 310	14		
6	302	2		менше 320	14		
7	304	4		менше 330	25		
8	304	4		менше 340	16		
9	304	9		менше 350	7		
10	312	1		менше 360	4		
11	312	1		менше 370	7		
12	314	4		менше 380	0		
13	314	5		менше 390	2		
14	314	3					
42	381	2					

Мал. 68

Задача а)

Виконуємо наступні послідовності дій.

- 1 У вічку С3 набираємо функцію (див. мал. 69)

$$= B3/100.$$
- 2 Ставимо курсор у правий нижній кут вічка С3; курсор змінює форму із стрілки на хрестик. Натискуємо ліву клавішу миші і, не відпускаючи її, протягуємо до кінця таблиці. В результаті у стовпці С отримуємо ряд відносних частот (див. мал. 70).

Задача б)

- 1 Для отримання ряду накопичених частот необхідно (див. мал. 68) у вічко F3 помістити вміст вічка E3; у вічко F4 помістити суму вмісту вічок E4 і F3 і так далі (алгоритм побудови дивись пункт 5.6.3 параграфу 5.6, розділ 5). Процес складання можна автоматизувати.

Розділ 8. Розв'язання задач засобами Excel

Введення функції для одержання розподілу відносної частоти

	A	B	C	D	E	F	G
1	Таблиця В.2.1-Е			Таблиця В.3.1-Е			
2	Платня	Частота	Частоті	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1	=В3:В3100		1		
4	298	2		менше 300	10		
5	298	2		менше 310	14		
6	302	2		менше 320	14		
7	304	4		менше 330	25		
8	304	4		менше 340	16		
9	304	9		менше 350	7		
10	312	1		менше 360	4		
11	312	1		менше 370	7		
12	314	4		менше 380	0		
13	314	5		менше 390	2		
14	314	3					
42	381	2					

Мал. 69

Результат побудови простого розподілу відносної частоти середньомісячної платні співробітників фірми N

32	331	1	0,01				
33	332	1	0,01				
34	332	1	0,01				
35	334	1	0,01				
36	336	1	0,01				
37	338	1	0,01				
38	340	2	0,02				
39	342	1	0,01				
40	342	1	0,01				
41	348	2	0,02				
42	381	2	0,02				
43		100	1				

Мал. 70

- 2) Активізуємо вічко F3 і внесемо до нього

$$= E3$$

Знак « \Rightarrow » означає, що у вічко вводиться формула.

- 3) < Enter >.

У вічку F3 з'явилось число 1 (див. мал. 71).

Введення функції для одержання розподілу накопиченої частоти

	A	B	C	D	E	F	G	H
1	Таблиця В.2.1-Е			Таблиця В.3.1-Е				
2	Платня	Частота	Частоті	Платня	Частоти	Накопичена частота	Накопичена частість	
3	282	1	0,01	менше 290	1	1		
4	298	2	0,02	менше 300	10	=E4+F3		
5	298	2	0,02	менше 310	14			

Мал. 71

- 4) У вічку F4 наберемо

$$= E4 + E3$$

- 5) < Enter >.

У вічку F4 з'явилось число «11» (див. мал. 72).

	A	B	C	D	E	F	G
1	Таблиця В.2.1-Е			Таблиця В.3.1-Е			
2	Платня	Частота	Частоті	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1	0,01	менше 290	1	1	
4	298	2	0,02	менше 300	10	11	
5	298	2	0,02	менше 310	14		

Мал. 72

- 6) Активізуємо вічко F4 і ставимо курсор миші на квадратик у правому нижньому куті вічка. Курсор змінить форму зі стрілки на хрестик. Не відпускаючи лівої кнопки миші, протягнемо її до кінця нашої таблиці.

Розділ 8. Розв'язання задач засобами Excel

7 Відпускаємо ліву кнопку миші.

На екрані з'явиться зображення розподілу накопиченої частоти платні співробітників фірми N, що і треба було побудувати (див. мал. 73).

Остаточний результат: згрупований розподіл накопиченої частоти середньомісячної платні співробітників фірми N

	А	В	С	Д	Е	Ф	Г
1	Таблиця В.2.1-Е			Таблиця В.3.1-Е			
2	Платня	Частота	Частоти	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1	0,01	менше 290	1	1	
4	298	2	0,02	менше 300	10	11	
5	298	2	0,02	менше 310	14	25	
6	302	2	0,02	менше 320	14	39	
7	304	4	0,04	менше 330	25	64	
8	304	4	0,04	менше 340	16	80	
9	304	9	0,09	менше 350	7	87	
10	312	1	0,01	менше 360	4	91	
11	312	1	0,01	менше 370	7	98	
12	314	4	0,04	менше 380	0	98	
13	314	5	0,05	менше 390	2	100	

Мал. 73

Задача в)

Побудуємо в стовпці G розподіл відносної накопиченої частоти (див. мал. 73).

1 У вічку G3 набираємо функцію (див. мал. 74)
 $= F3/100$

2 Ставимо курсор у правий нижній кут вічка G3; курсор змінює форму із стрілки на хрестик. Натискаємо ліву клавішу миші і, не відпускаючи її, протягуємо до кінця таблиці. В результаті у стовпці G отримуємо ряд відносних частот (див. мал. 75).

Введення функції для одержання розподілу накопиченої відносної частоти

	A	B	C	D	E	F	G
1	Таблиця В.2.1-Е		Таблиця В.3.1-Е				
2	Платня	Частота	Частоті	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1	0,01	менше 290	1	1	=F3/100
4	298	2	0,02	менше 300	10	11	

Мал. 74

Результат побудови згрупованого розподілу частоти середньомісячної платні співробітників фірми N

	A	B	C	D	E	F	G
1	Таблиця В.2.1-Е		Таблиця В.3.1-Е				
2	Платня	Частота	Частоті	Платня	Частоти	Накопичена частота	Накопичена частість
3	282	1	0,01	менше 290	1	1	0,01
4	298	2	0,02	менше 300	10	11	0,11
5	298	2	0,02	менше 310	14	25	0,25
6	302	2	0,02	менше 320	14	39	0,39
7	304	4	0,04	менше 330	25	64	0,64
8	304	4	0,04	менше 340	16	80	0,8
9	304	9	0,09	менше 350	7	87	0,87
10	312	1	0,01	менше 360	4	91	0,91
11	312	1	0,01	менше 370	7	98	0,98
12	314	4	0,04	менше 380	0	98	0,98
13	314	5	0,05	менше 390	2	100	1

Мал. 75

Вправи для самостійної роботи

Побудувати

- упорядковану вибірку;
- незгрупований розподіл частот;

Розділ 8. Розв'язання задач засобами Excel

- в) згрупований розподіл частот;
- г) згрупований розподіл накопиченої частоти.

1. У наступній таблиці наведена вибірка стажу роботи (у роках) 144 працівників фірми *N* (див. Таблицю 15).

Таблиця 15

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

2. Для прийняття на роботу фірма робить перевірку кандидатів на здатність. У число перевірок входить тестування по певній 100 бальній системі. Нижче наведені результати тестування 100 кандидатів (див. Таблицю 16).

Таблиця 16

71	77	47	54	62	54	61	47	73	59
69	47	46	71	54	47	58	59	60	46
53	60	69	61	53	53	88	83	87	47
60	51	68	39	87	98	47	90	43	60
95	91	47	43	91	38	72	64	55	42
46	54	39	73	59	60	38	66	43	60
42	43	61	71	80	62	57	47	66	71
86	53	66	66	47	55	66	32	73	61
73	60	61	62	46	64	81	65	55	60
68	79	68	61	66	72	61	54	62	50

3. Перевірені депозитні рахунки 100 банків. На рахунках виявилися наступні суми у млн. гривень (див. Таблицю 17).

Таблиця 17

3,7	14,8	20,1	14,2	20,6	23,6	6,1	23	24,1	13,1
8,4	23,6	8,4	6,6	18,9	0,8	0,8	14,2	10,7	3,7
4,9	23	27,7	1,4	24,1	2,6	21,8	7,8	28,8	23,6
14,8	7,8	13,6	22,4	15,4	14,8	10,7	14,8	11,9	2,6
5,5	0,8	24,1	16	17,7	17,1	14,2	13,1	17,7	23,6
10,7	8,4	17,7	25,3	21,8	0,8	19,5	4,3	25,3	13,1
21,2	24,1	29,4	3,7	19,5	5,5	1,4	1,4	14,8	20,6
3,7	18,9	7,2	19,5	18,3	2,6	18,3	2,6	8,4	21,2
6,1	18,3	22,4	15,4	1,4	0,8	7,2	23	20,6	13,1
21,8	11,3	4,3	22,4	10,1	21,2	5,5	7,8	10,7	2,0

4. Нижче наведено ціну у гривнях одного квадратного метра жилої площі в залежності від якості та міста розташування житла по даних фірми нерухомості *N* (див. Таблицю 18).

Таблиця 18

599	1400	1293	759	652	634	688	563
1168	563	955	1061	1026	777	688	616
883	901	1257	901	990	599	670	1115
1008	1150	1097	599	670	759	741	1026
1026	990	1008	830	634	883	705	1061
812	741	1275	883	759	652	1150	830
1168	1133	688	919	1008	848	1150	705
705	77	1061	705	1079	1168	670	581
1186	1293	901	1186	794	1239	794	866
581	1097	794	1222	1186	1097	634	1275

8.4. Графічне представлення розподілів вибіркової сукупності

Теоретичні положення: параграф 5.8 розділ 5.

8.4.1. Графічне представлення розподілів за допомогою функції «Гистограмма»

Функція «Гистограмма» із пакета «Анализ данных» використовується для обчислення вибірових і інтегральних частот попадання даних у вказані інтервали значень. Використання функції «Гистограмма» припускає існування трьох масивів:

«Входной_интервал» – це вічка, де розміщені дані вибірки;

«Интервал_карманов» – це вічка, що вказують як вибираються інтервали згуртовування частот вибірки;

«Выходной_интервал» – це вічка, куди буде виведений розподіл згрупованих частот.

Задання масиву кишеней не обов'язково. Однак, якісні діаграми можна отримати лише коли масив кишень заданий. Тому, перед графічними побудовами необхідно вибрати класи інтервалів згрупованих даних. Для цього потрібно побудувати варіаційний ряд (тобто записати вибірку в порядку зростання варіант (див. пункт 8.3.1 параграфу 8.3)) і провести всі необхідні обчислення (див. пункт 5.6.7 параграфу 5.6, розділ 5).

Приклад 1. Для вибірки середньомісячної платні 100 співробітників фірми *N*, що наведено у Таблиці В (див. мал. 76), побудувати а) гістограму частот, б) полігон частот, в) полігон накопичених відносних частот.

Розв'язання. У нашому випадку

«Входной_интервал» – це вічка A3: J12;

«Интервал-карманов» – це вічка L2:L12, які заповнені рядом чисел

289, 299, 309, 319, 329, 339, 349, 359, 369, 379, 389.

Занесені дані значать, що частоти варіант вибірки будуть підраховуватись для інтервалів

Вихідна таблиця даних

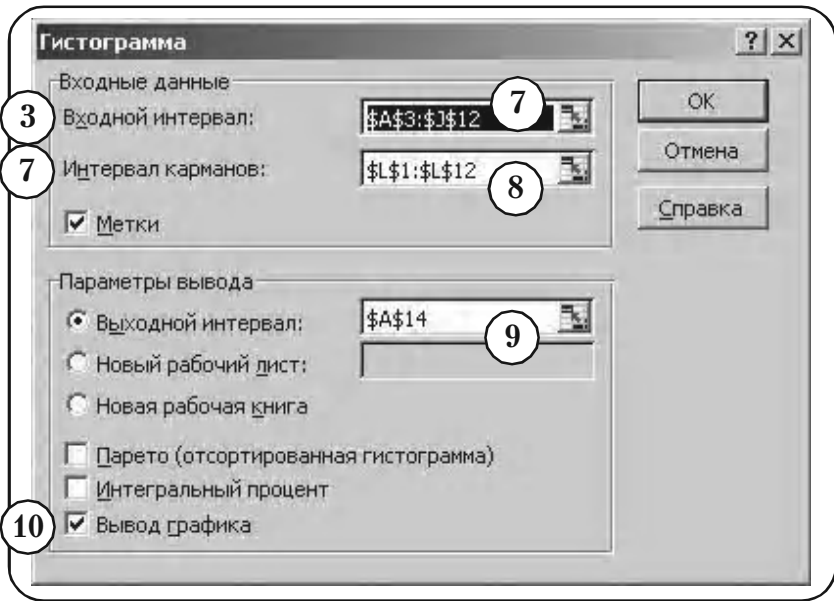
	A	B	1	D	E	F	G	H	I	J	K	L
1	Таблиця В											кишені
2	Вибірка середньомісячної платні 100 співробітників фірми											289
3	338	348	304	314	326	314	324	304	342	308		299
4	336	304	302	338	314	304	320	321	322	321		309
5	312	323	336	324	312	312	364	356	362	302		319
6	322	310	334	292	362	381	304	366	298	304		329
7	381	368	304	298	368	290	340	328	316	322		339
8	302	314	292	342	321	322	290	332	298	296		349
9	296	298	324	338	352	326	318	304	332	322		359
10	360	312	331	331	304	316	332	282	342	338		369
11	342	322	324	325	302	328	354	330	316	324		379
12	334	350	334	324	332	340	324	314	326	323		389

Мал. 76

$x \leq 289$	$340 \leq x \leq 349$
$290 \leq x \leq 299$	$250 \leq x \leq 359$
$300 \leq x \leq 309$	$260 \leq x \leq 369$
$310 \leq x \leq 319$	$270 \leq x \leq 379$
$320 \leq x \leq 329$	$280 \leq x \leq 389$
$330 \leq x \leq 339$	$x > 390.$

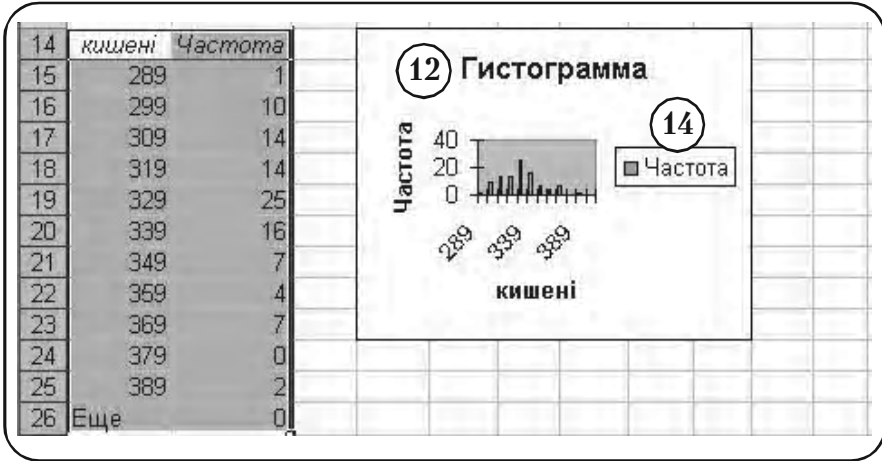
Задача а)

- 1 Сервіс > Аналіз даних > Гістограма.
- 2 У вікні з'явилося меню функції Гістограма (див. мал. 77).
- 3 Активізуємо вікно «Входной интервал».
- 4 ⚡ 1 А3 (Активізуємо вічко А3) (див. мал. 76).
- 5 < SHIFT > + ⚡ 1 J12 (Активізуємо вічко J12).



Мал. 77

- 6) Интервал вічок A3:J12 автоматично заноситься в вікно «Входной интервал».
- 7) Активізуємо вікно «Интервал карманов».
- 8) Аналогічно крокам 3—9 заповнюємо вікно «Интервал карманов».
- 9) В якості вихідного інтервалу вибираємо вічка A14:A24.
- 10) Ставимо прапорець у вікні «Вывод графики».
- 11) ОК.
- 12) У полі екрану з'являється «Гистограмма» (див. мал. 78).
Отриману гистограму необхідно підредагувати.
- 13) Вилучити надпис «Гистограмма».
- 14) Вилучити «легенду» – надпис «Частота».



Мал. 78

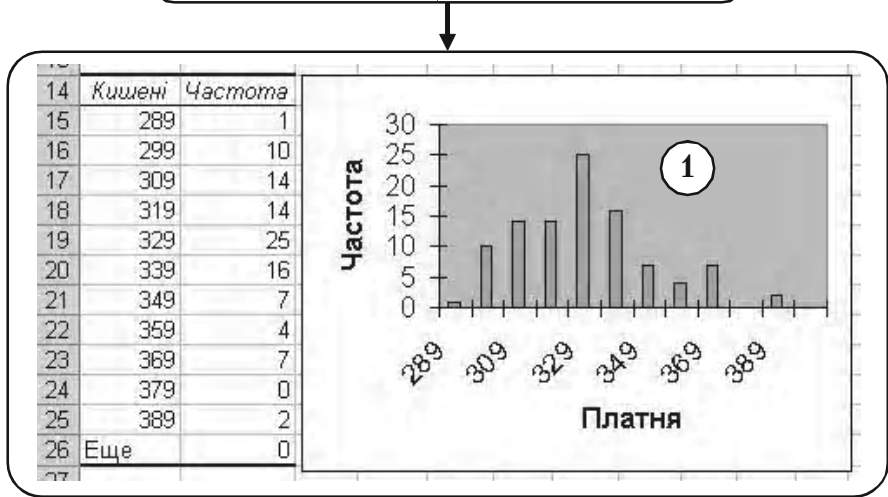
- 15 Змінити надпис «Карман» на надпис «Платня».
- 16 Вибрати відповідні масштаби на осі x і осі y .
- 17 Змінити розміри гістограми.

Операції 13–17 детально описані в пункті 8.9.2, параграф 8, тому приведемо тут кінцевий результат (див. мал. 79).

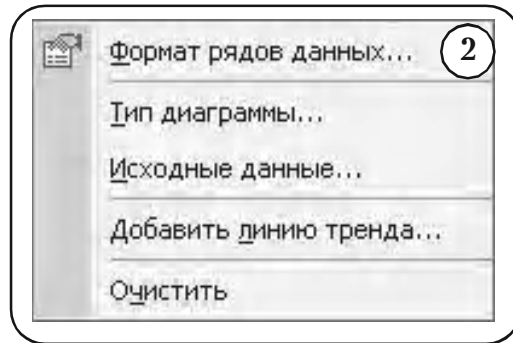
Нарешті, опишемо алгоритм знищення зазорів між прямокутниками на гістограмі. Здійснюється це за допомогою меню «Формат рядов данных».

- 1 Встановлюємо курсор лінії на поле одного з прямокутників діаграми і натискаємо праву клавішу (див. мал. 79).
- 2 У меню, що з'явилося, вибираємо опцію «Формат рядов данных» (див. мал. 80).
- 3 У меню «Формат рядов данных» вибираємо опцію «Параметры» (див. мал. 81).
- 4 У вікні «Ширина зазора» встановлюємо значення 0.
- 5 У вікні меню з'являється зображення гістограми.

Гістограма частот після редагування



Мал. 79

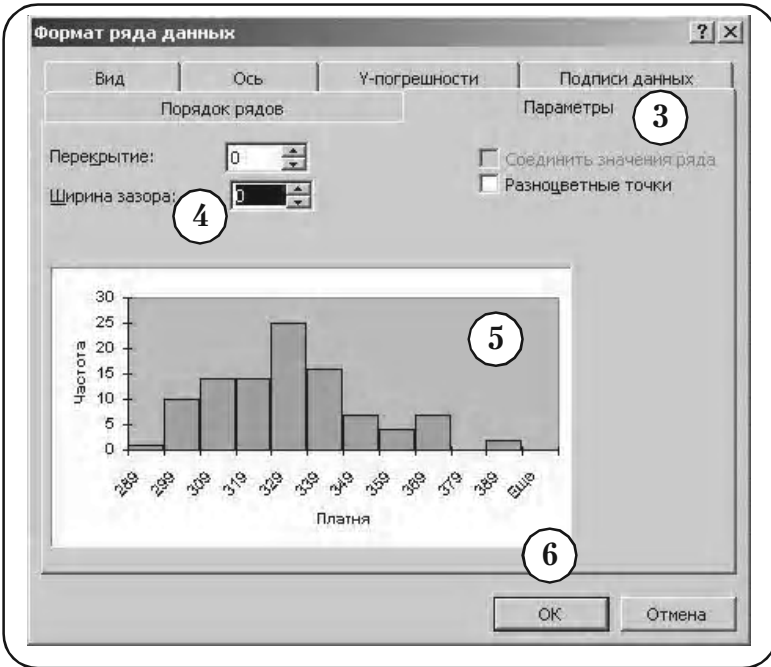


Мал. 80

6 ОК.

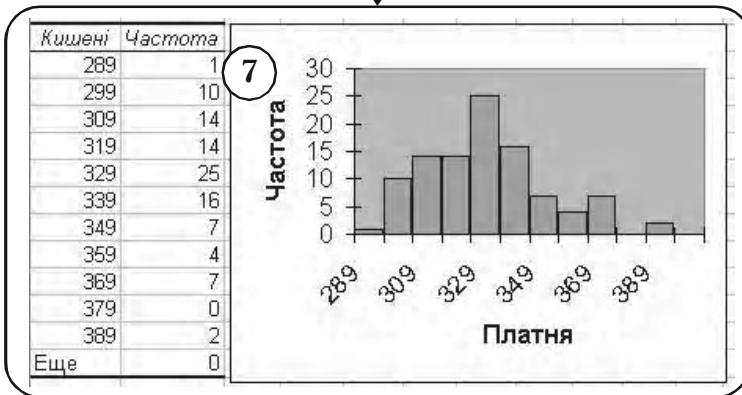
7 Кінцевий вигляд побудованої діаграми показаний на мал. 82.

Подальші можливі перетворення гістограми надаємо на розгляд читачам.



Мал. 81

Остаточний результат: гістограма частот вибірки середньомісячної платні

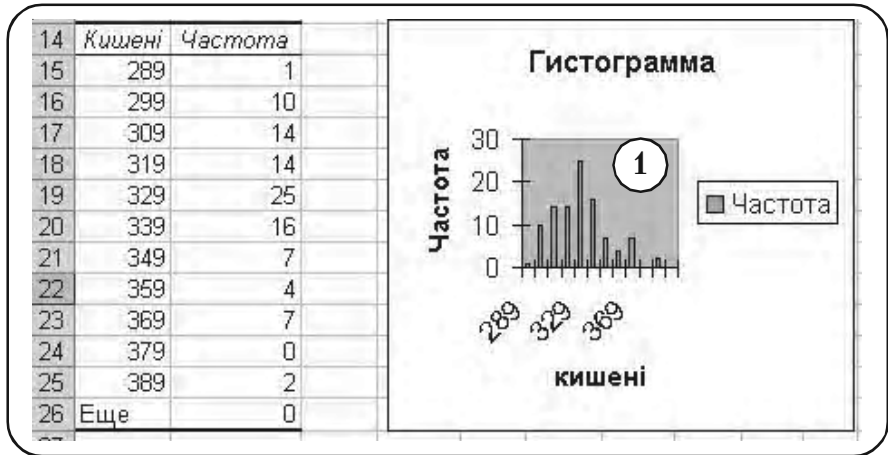


Мал. 82

Розділ 8. Розв'язання задач засобами Excel

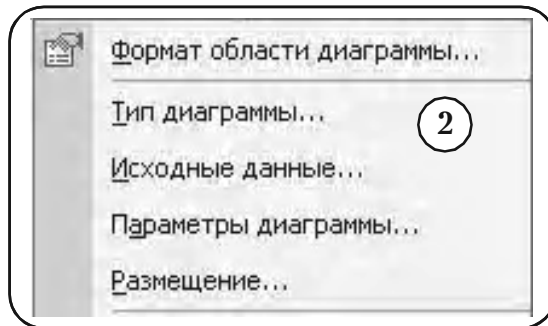
Задача б)

Повторимо всі кроки алгоритму Задачі а) аж до кроку (12) (див. мал. 83).



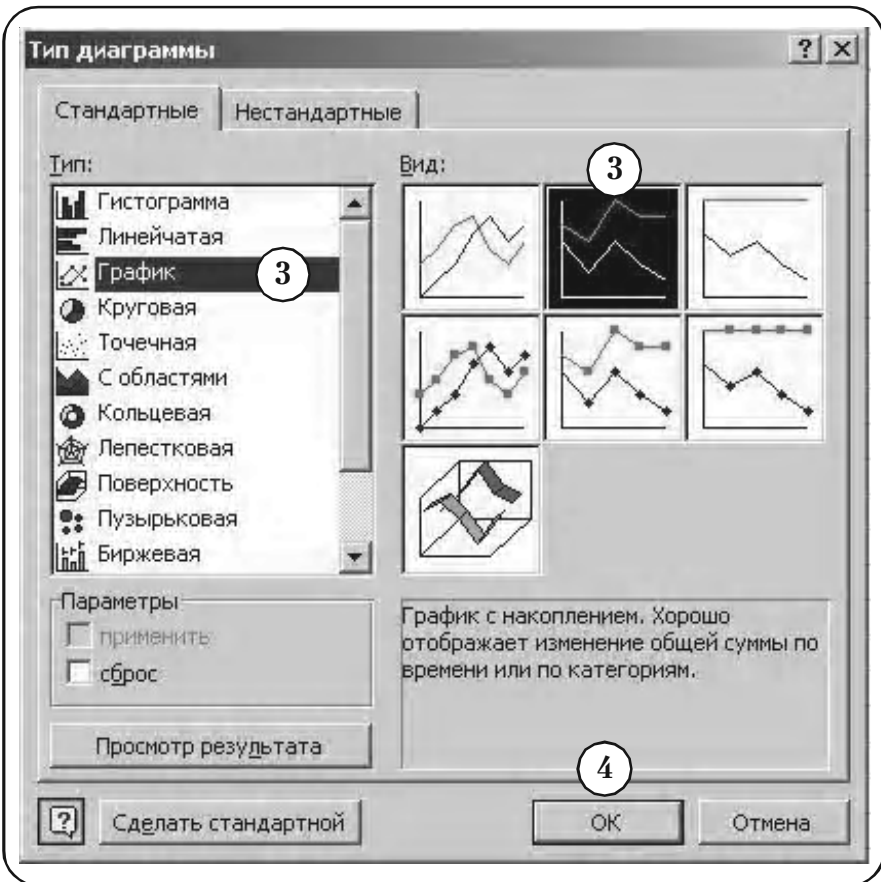
Мал. 83

- 1 Ставимо мишу в «Область диаграммы» і натискуємо праву клавішу миші.
- 2 У меню, що з'явилося, вибираємо опцію «Тип диаграммы» (див. мал. 84). Тут зображено лише частину меню.



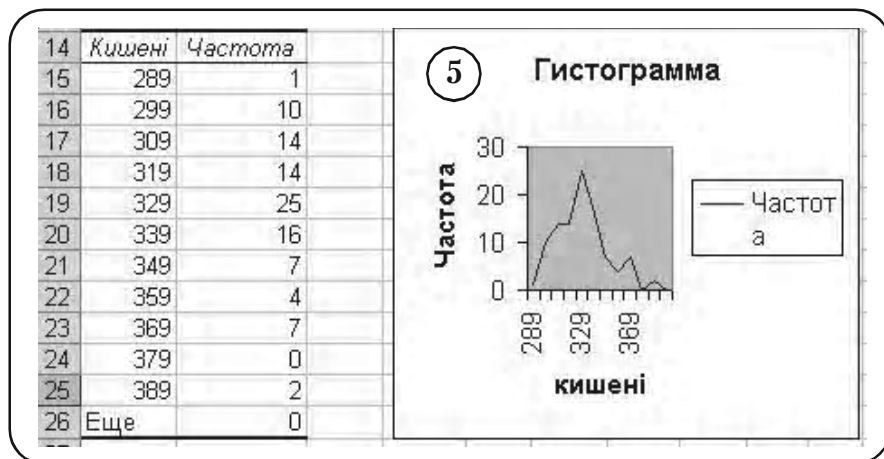
Мал. 84

- 3 У меню «Тип діаграмми», що з'явилося, вибираємо опції «Стандартные» та «График», як показано на мал. 85.



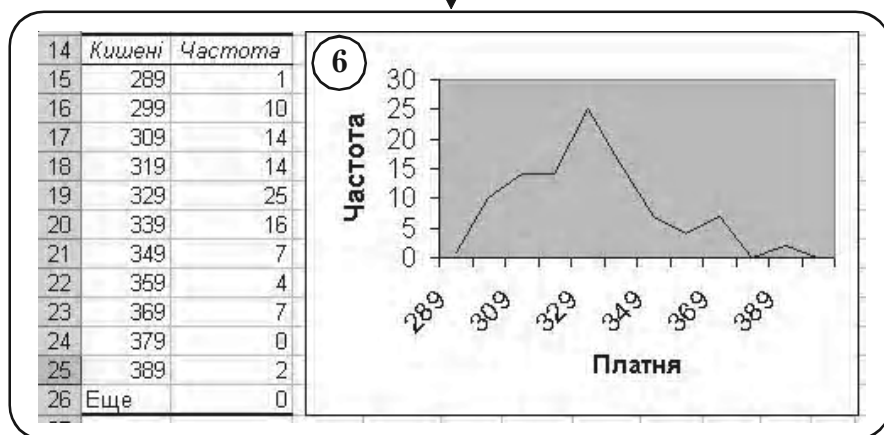
Мал. 85

- 4 ОК.
- 5 На екрані з'являється зображення полігона частот (див. мал. 86)
- 6 Після редагування графіка (пункт 8.9.2, параграф 8) дістаємо шуканий полігон частот (див. мал. 87).



Мал. 86

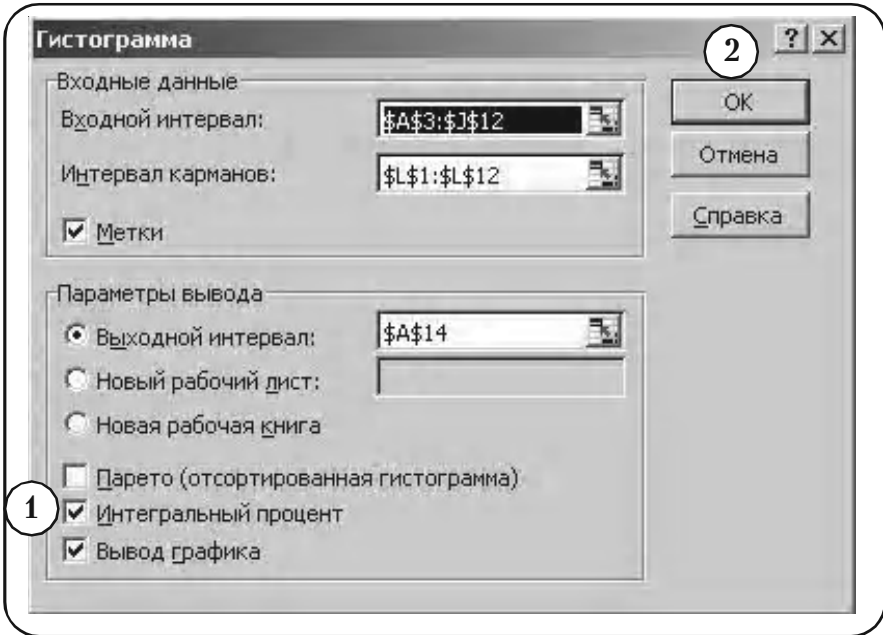
Остаточний результат: полігон частот вибірки середньомісячної платні



Мал. 87

Задача в)

Повторимо всі кроки алгоритму Задачі а) аж до кроку (10) (див. мал. 88).

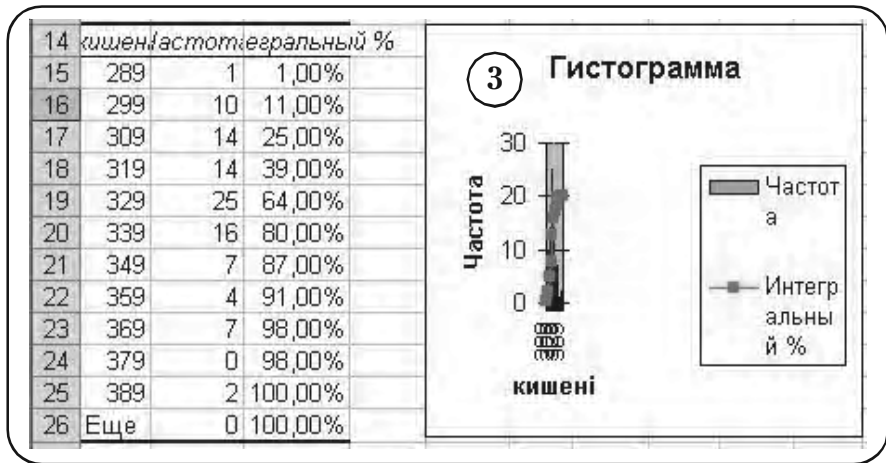


Мал. 88

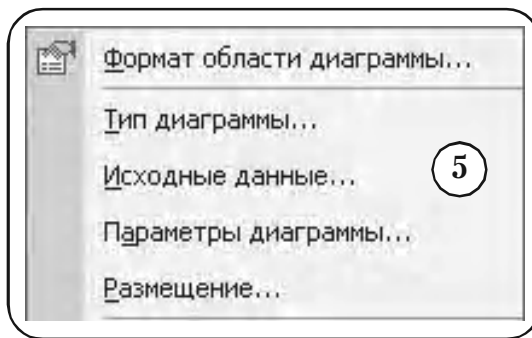
- 1 Ставимо прапорець у вікні «Интегральный процент».
- 2 ОК.
- 3 На екрані з'являється зображення графіка (див. мал. 89).

На графіку водночас зображені гистограма частот і полігон накопичених частот. Залишимо лише зображення полігона накопичених частот.

- 4 Ставимо мишу в «Область диаграммы» і натискаємо праву клавішу миші.
- 5 У меню, що з'явилося, вибираємо опцію «Исходные данные» (див. мал. 90). Тут зображено лише частину меню.

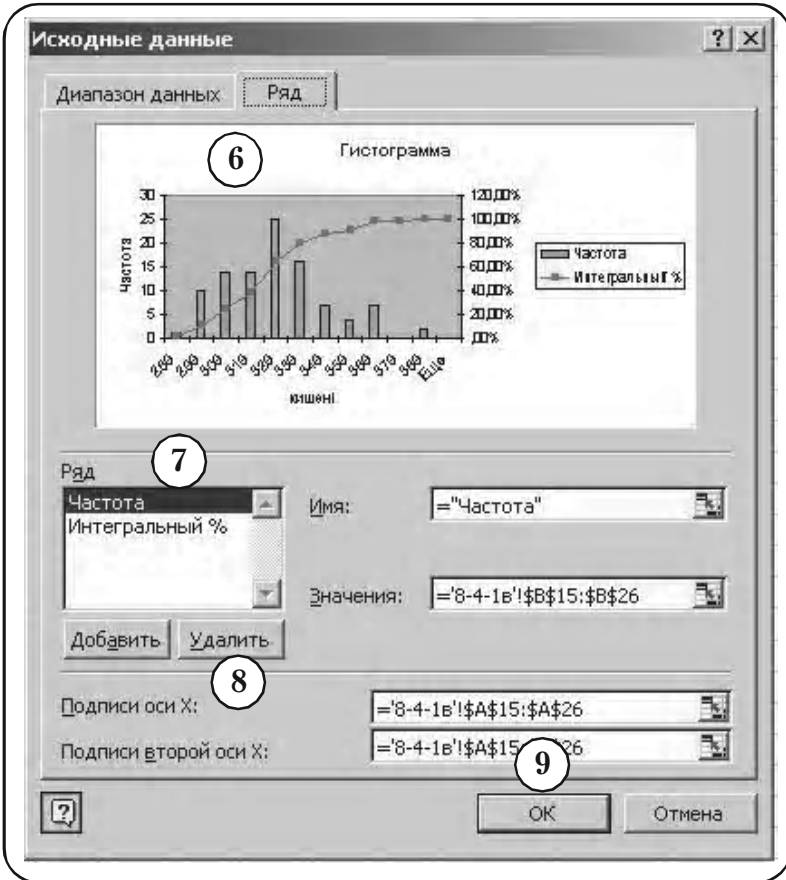


Мал. 89



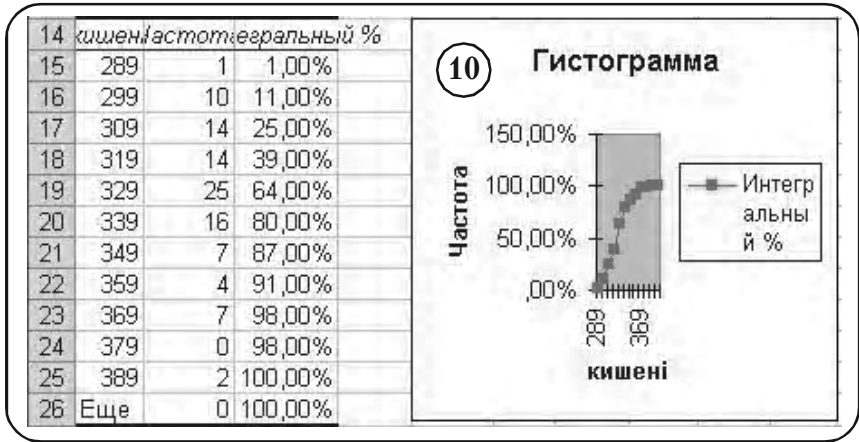
Мал. 90

- 6 У меню «Исходные данные», що з'явилося, вибираємо опцію «Ряд», як показано на мал. 91.
- 7 У вікні «Ряд» вибираємо опцію «Частота».
- 8 < Удалить >.
- 9 ОК.
- 10 На графіку залишаться лише зображення полігона накопичених частот (див. мал. 92).



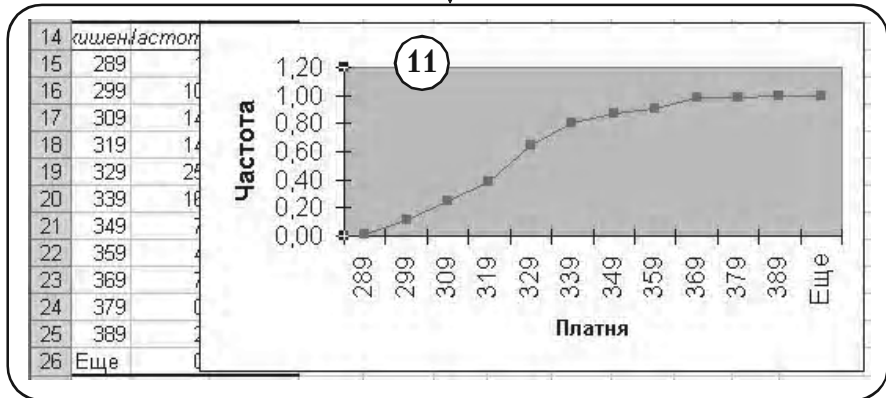
Мал. 91

- 11 Після редагування графіка (пункт 8.9.2, параграф 8) дістаємо шуканий полігон накопичених частот (див. мал. 93).



Мал. 92

Остаточный результат: полигон накопленных частот выборки средньюмесячной платни



Мал. 93

8.4.2. Графічне представлення розподілів за допомогою «Мастера діаграмм»

Якщо розподіл вибірки вже існує, то ефективним засобом його графічного представлення є спеціальна функція Excel під назвою «Мастер діаграмм».

Організуючи дані розподілу вибірки певним чином, з допомогою «Мастера діаграмм» можна побудувати різноманітні графіки.

Нижче розглянуті три види графіків:

- гістограма частот;
- контурна гістограма;
- емпірична функція розподілу.

1. Побудова гістограми частоти

Побудова гістограми частоти вже розглянута в пункті 8.4.1. Для порівняння методик побудови, ця ж гістограма будується тут за допомогою «Мастера діаграмм». За даними Таблиці В.3-Е (стор. 272) створюємо Таблицю В.3-Е.3.1 структури, показаної на мал. 94. Для кращої наочності результати розміщені у чотири стовпчики. Щоб користуватись при графічних побудовах «Майстром діаграм», що входить в Excel, запишемо цю таблицю у вигляді Таблиці В.3-Е.3 з двох стовпчиків (див. мал. 95). Щоб розібратися у структурі таблиці, побудуємо частину діаграми для даних, що розташовані у вічках А3:В8. Для цього необхідно виконати наступні дії.

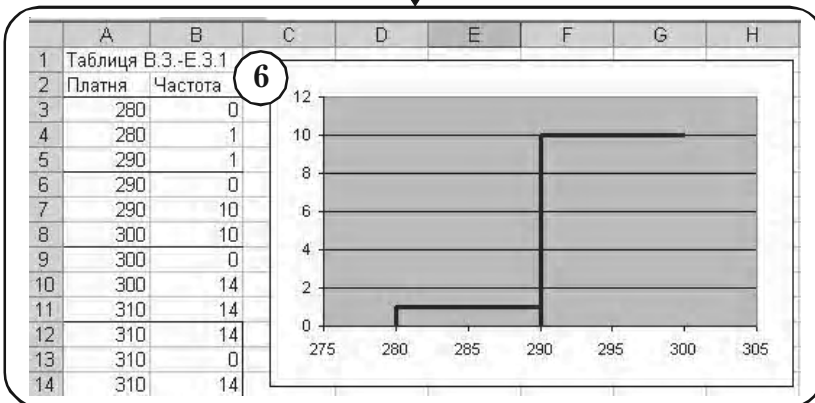
- 1 Виділимо вічка А3:А8 в Таблиці В.3-Е.3, для яких буде побудована частина гістограми.
- 2 **Вставка → Діаграмма.**
- 3 Вменю «Мастера діаграмм», що з'явилося (див. мал. 96), вибираємо опції «Стандартная», «Точечная».
- 4 **<Готово>.**
- 5 З'явиться графічне зображення частини гістограми (див. мал. 97).
- 6 Після редагування (див. пункт 8.9.2, параграф 8) одержимо остаточний результат – дві сходинки гістограми (див. мал. 95).

Таблиця В.5-Е.3.1 вихідних даних для побудови гістограми, яку сформовано за даними Таблиці В.3-Е

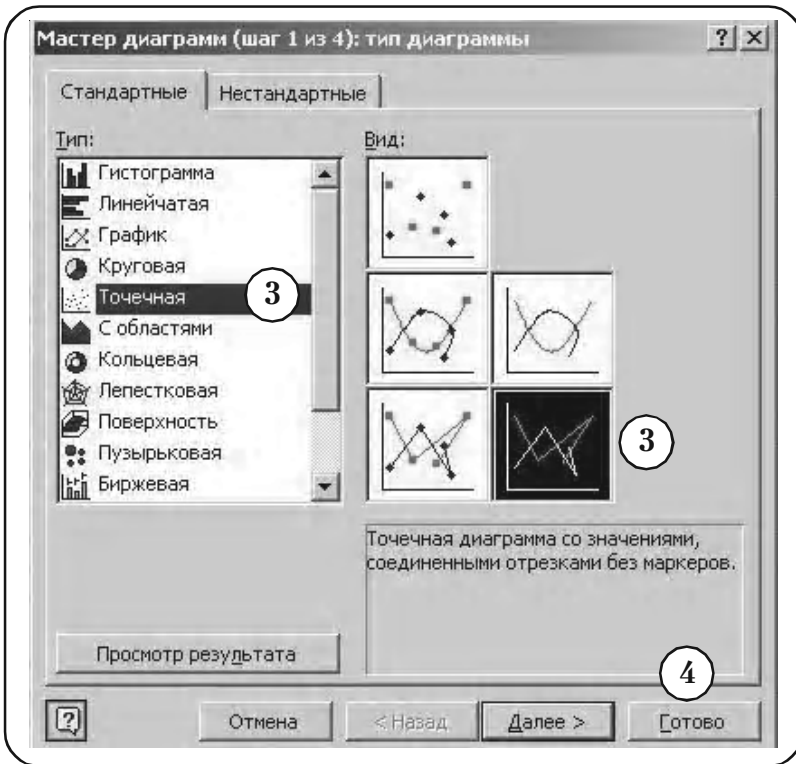
	А	В	С	Д
1	Таблиця В.3.-Е.3.1			
2	Платня	Частота	Платня	Частота
3	280	0	330	18
4	280	1	340	18
5	290	1	340	0
6	290	0	340	7
7	290	10	350	7
8	300	10	350	0
9	300	0	350	4
10	300	14	360	4
11	310	14	360	0
12	310	14	360	7
13	310	0	370	7
14	310	14	370	0
15	320	14	380	0
16	320	0	380	2
17	320	25	390	2
18	330	25	390	0
19	330	0		

Мал. 94

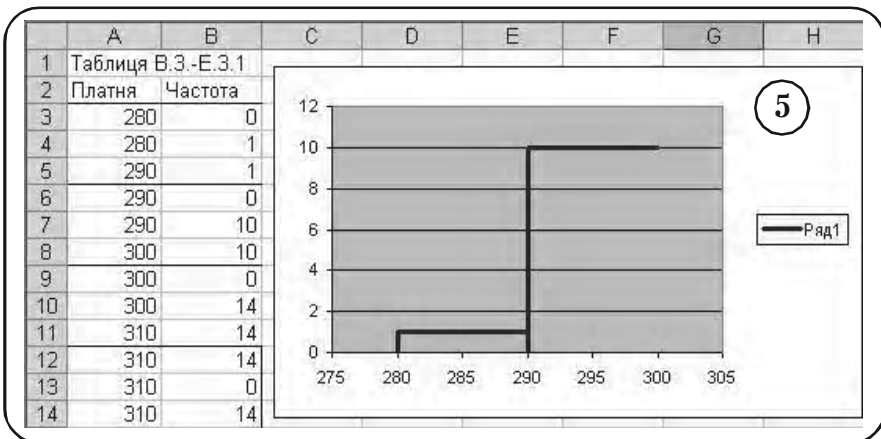
Побудова перших двох сходинок гістограми за даними рядків 3–8 Таблиці В.3-Е3



Мал. 95



Мал. 96



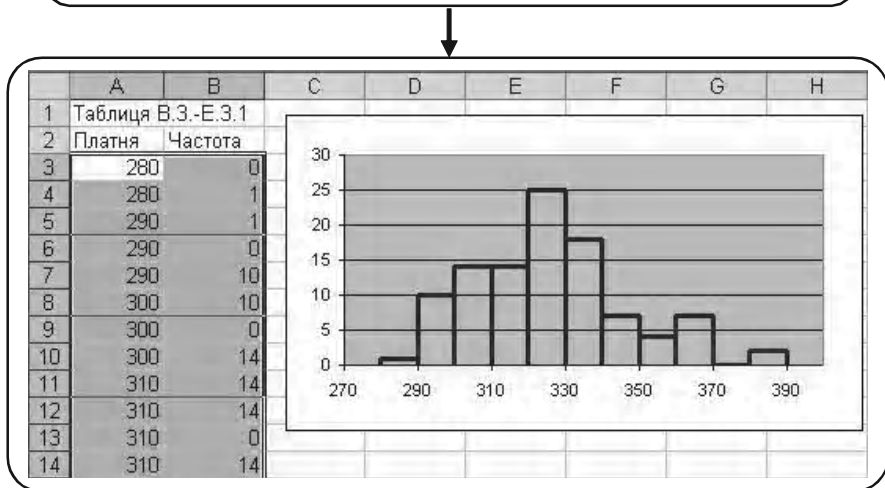
Мал. 97

Розділ 8. Розв'язання задач засобами Excel

Як видно з мал. 95, рядки 3, 4, 5 відповідають за формування першої сходинки, а рядки 6, 7, 8 – за формування другої сходинки. Зокрема, вертикальна лінія першої сходинки формується рядками 3, 4, а другої сходинки формується рядками 6, 7.

Після того, як з'ясована структура Таблиці В.З-Б.3 виділимо усі дані таблиці і побудуємо для неї гістограму. На мал. 98 наведено остаточний результат – гістограма частоти середньомісячної платні співробітників фірми N.

Остаточний вигляд гістограми частот середньомісячної платні співробітників фірми N (Таблиця В.З-Е)



Мал. 98

Подальші можливі вдосконалення у гістограмі надаємо читачу ввести самостійно.

2. Побудова контурної гістограми

① Створюємо Таблицю В.З-Е.2.1 структури, показану на мал. 99.

При її створенні використано Таблицю В.З-Е згрупованого розподілу частоти середньомісячної платні співробітників фірми N (див. стор. 292). Для кращої наочності результати розміщені у чотири стовпчики. Щоб користуватись при графічних побудовах «Майстром

Таблиця В.5-Е.2.1 вхідних даних для побудови контурної гістограми, яку сформовано за даними Таблиці В.3-Е

	A	B	C	D	E	F	G
1	Таблиця В.3.-Е.2.1						
2	Платня	Частота	Платня	Частота			
3	280	0	340	16			
4	280	1	340	7			
5	290	1	350	7			
6	290	10	350	4			
7	300	10	360	4			
8	300	14	360	7			
9	310	14	370	7			
10	310	14	370	0			
11	320	14	380	0			
12	320	25	380	2			
13	330	25	390	2			
14	330	16	390	0			

Мал. 99

Таблиця В.5-Е.2, у вигляді двох стовпчиків є зображенням Таблиці В.5-Е.2.1

	A	B	C	D	E	F	G
1	Таблиця В.3.-Е.2.1						
2	Платня	Частота					
3	280	0					
4	280	1					
5	290	1					
6	290	10					
7	300	10					
8	300	14					
9	310	14					
10	310	14					
11	320	14					
25	390	2					
26	390	0					

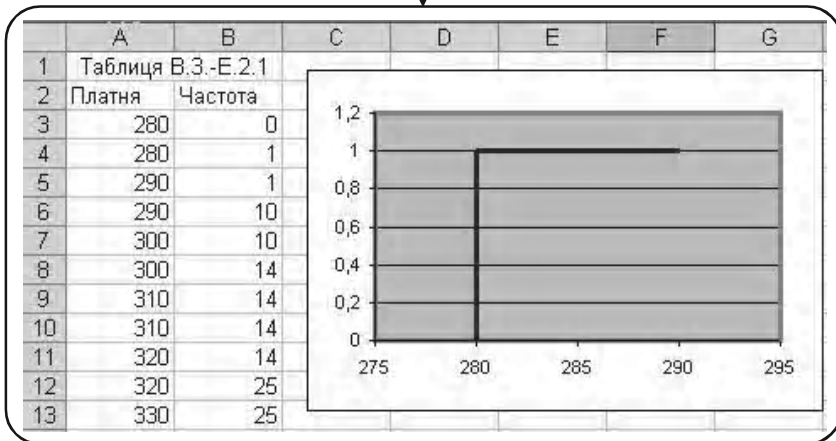
Мал. 100

Розділ 8. Розв'язання задач засобами Excel

діаграм», що входить в Excel, треба записати цю таблицю у вигляді Таблиці В.3-Е.2 з двох стовпчиків (див. мал. 100).

Щоб зрозуміти структуру Таблиці В.3-Е.2, виділимо у ній рядки 3–5 і побудуємо для них діаграму, застосовуючи майстер діаграм (див. мал. 101).

Побудова першої сходинки контурної гістограми за даними рядків 3–5 Таблиці В.3–Е.2



Мал. 101

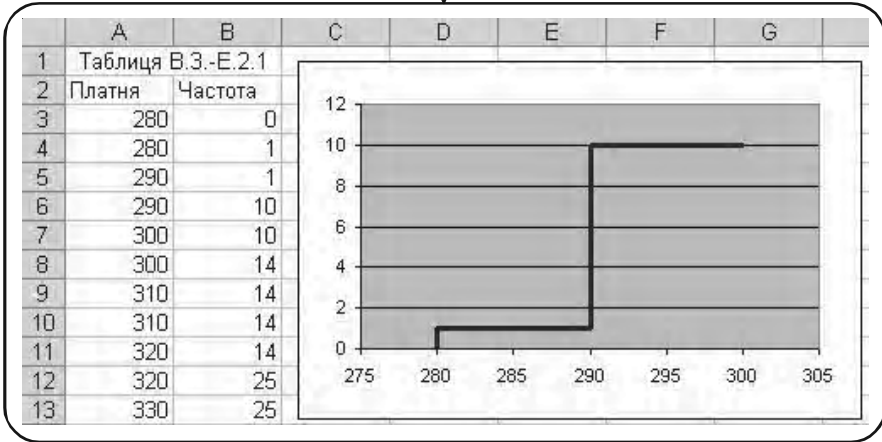
З мал. 101 видно, що 3 та 4 рядки формують вертикальну лінію на діаграмі, а 4 і 5 рядки – горизонтальну лінію.

Зробимо наступний крок у побудові гістограми. Виділимо вічка А3:В7 у Таблиці В.3-Е.2 і побудуємо для них діаграму (див. мал. 102).

Як видно з мал. 102, рядки 6 і 7 з різними значеннями аргументу (Платня), але однаковими значеннями функції (Частота), формують на діаграмі горизонтальні лінії графіка.

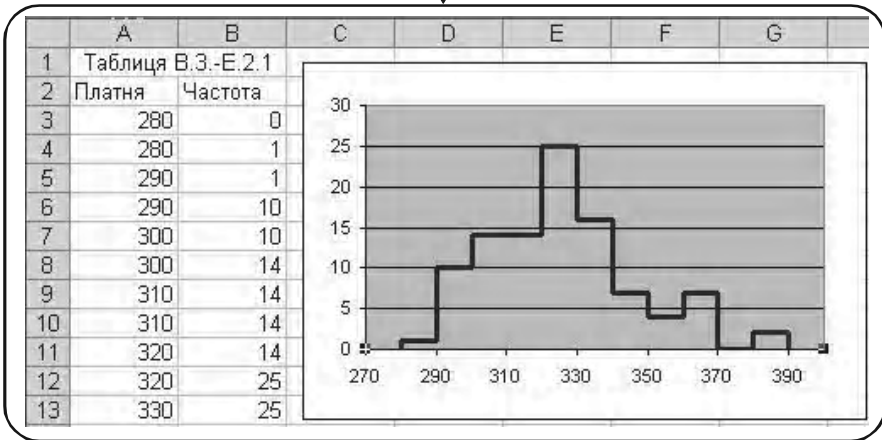
Рядки 5 та 6 з однаковими значеннями аргументу (Платня), але різними значеннями функцій (Частота), формують на діаграмі вертикальну ділянку графіка.

Побудова перших двох сходинок контурної гістограми за даними рядків 3–7 Таблиці В.З-Е.2



Мал. 102

Остаточний результат: контурна гістограма частот середньомісячної платні співробітників фірми N (Таблиця В.З-Е)



Мал. 103

- 2 Після з'ясування структури таблиці, виділимо всі числові дані Таблиці В.3-Е.2 і побудуємо всю гістограму (див. мал. 103).

Подальше можливе оформлення гістограми надаємо читачу.

3. Побудова емпіричної функції розподілу

- 1 Побудова діаграми у вигляді прямокутників.

Таблиця даних для побудови емпіричної функції розподілу наведена на мал. 104 (Таблиці В.5-Е.3).

Для більшої наочності дані згруповані у шість стовпчиків. Дані взято з Таблиці В.3.1-Е (мал. 75, стор. 287). Таблиця даних формується так, як якщо б ми будували гістограму у вигляді прямокутників (див. алгоритм «1. Побудова гістограми частоти»).

Таблиця В.5-Е.3 вхідних даних для побудови емпіричної функції розподілу, яку сформовано за даними Таблиці В.5-Е

	A	B	C	D	E	F
1	Таблиця В.5-Е.3					
		Накопичена відносна частота		Накопичена відносна частота		Накопичена відносна частота
2	Зарплата		Зарплата		Зарплата	
3	270	0	320	0	360	0
4	280	0	320	0,64	360	0,98
5	280	0	330	0,64	370	0,98
6	280	0,01	330	0	370	0
7	290	0,01	330	0,8	370	0,98
8	290	0	340	0,8	380	0,98
9	290	0,11	340	0	380	0
10	300	0,11	340	0,87	380	1
11	300	0	350	0,87	390	1
12	300	0,25	350	0	390	0
13	310	0,25	350	0,91	390	1
14	310	0	360	0,91	400	
15	310	0,39				
16	320	0,39				
17						

Мал. 104

Запишемо Таблицю В.5-Е.3 у вигляді Таблиці В.5-Е.4 з двома стовпчиками (див. мал. 105) і побудуємо діаграму у вигляді прямокутників (див. мал. 106).

Таблиця В.5-Е.4 у вигляді двох стовпчиків є зображенням Таблиці В.5-Е.3

	A	B	C	D	E	F	G
1	Таблиця В.5.-Е.4						
2	Платня	Накопичена відносна частота					
3	270	0					
4	280	0					
5	280	0					
6	280	0,01					
7	290	0,01					
36	380	0					
37	380	1					
38	390	1					
39	390	0					
40	400						

Мал. 105

2) Перетворення діаграми у вигляді прямокутників в емпіричну функцію розподілу

Суть цієї побудови полягає в перетворенні гістограми на мал. 106 в емпіричну функцію розподілу: усі горизонтальні лінії зробимо жирними (основні лінії), а вертикальні лінії (допоміжні) – пунктирними.

Стаavimo курсор миші на довільну лінію гістограми і натискаємо ліву кнопку миші. З'явиться надпис

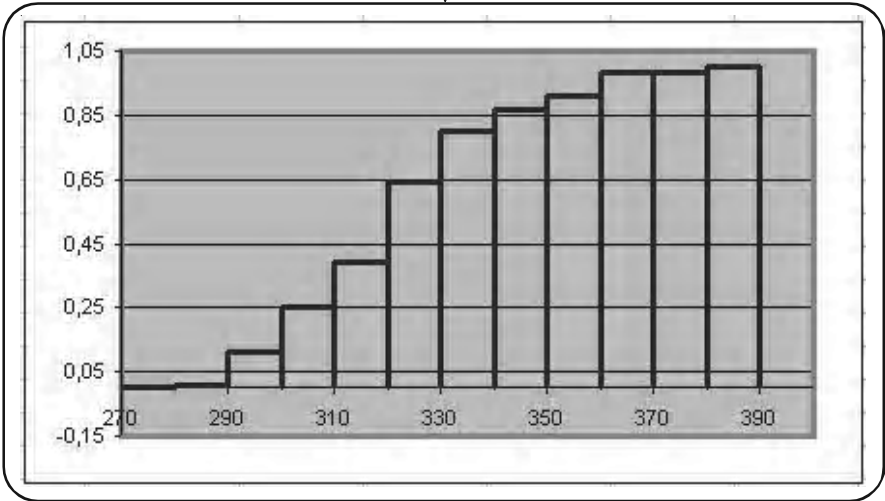
Ряд 1 точка «310»

та маркери (координата точки в лапках може бути будь-якою, див. мал. 107).

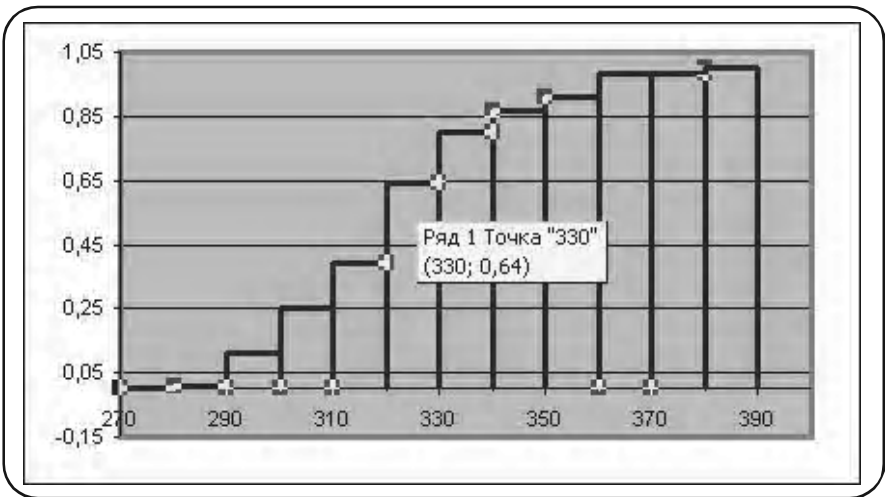
Натискуємо праву кнопку миші і в меню, що випало, вибираємо меню «Формат рядов данных» (див. мал. 108).

Розділ 8. Розв'язання задач засобами Excel

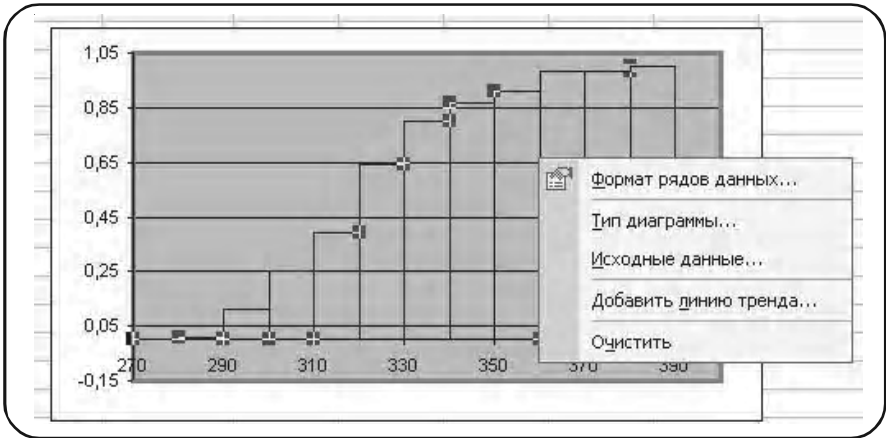
Гістограма побудована за даними Таблиці В.5-Е.4 і буде поступово перетворена в емпіричну функцію розподілу



Мал. 106



Мал. 107

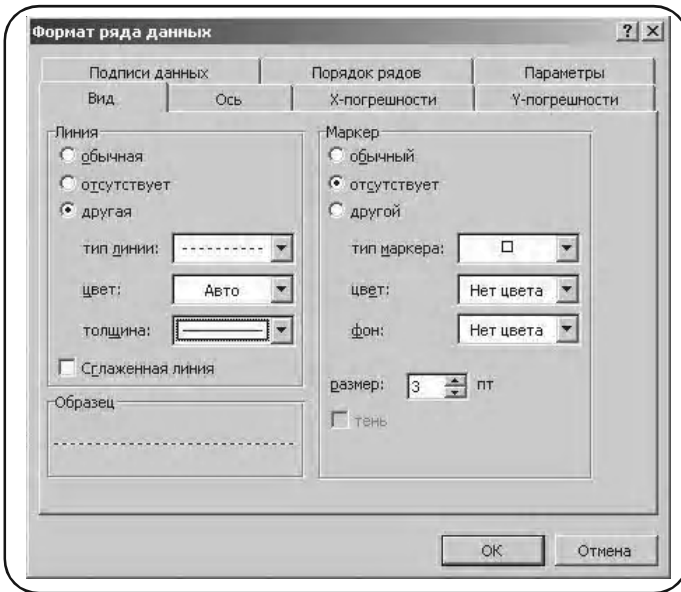


Мал. 108

У меню «Формат рядов данных», що з'явилось, вибираємо тип лінії «пунктирная», як це показано на мал. 109.

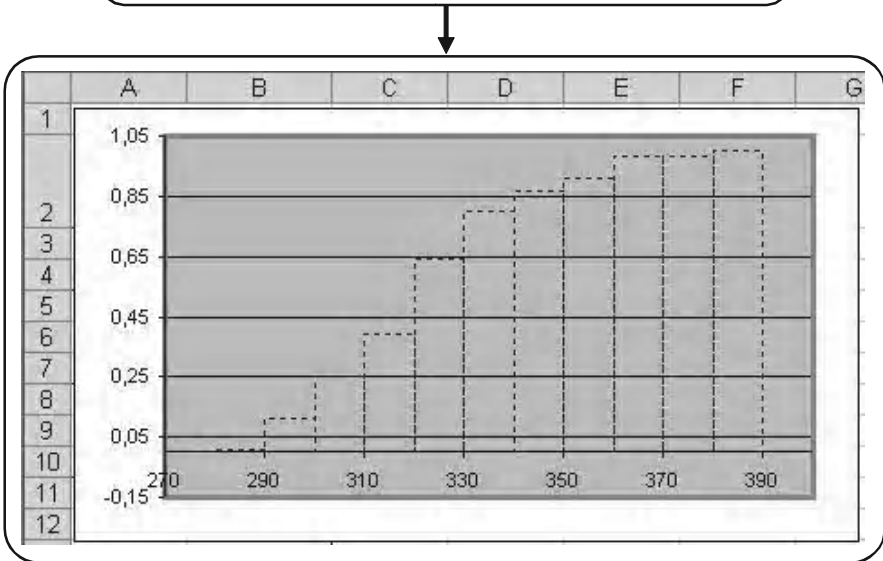
ОК,

на екрані з'являється зображення (див. мал. 110).



Мал. 109

Проміжний результат перетворення вихідної гистограми в емпіричну функцію розподілу



Мал. 110

Тепер намалюємо горизонтальні лінії жирно. Покажемо на прикладі першої горизонтальної лінії як це робиться.

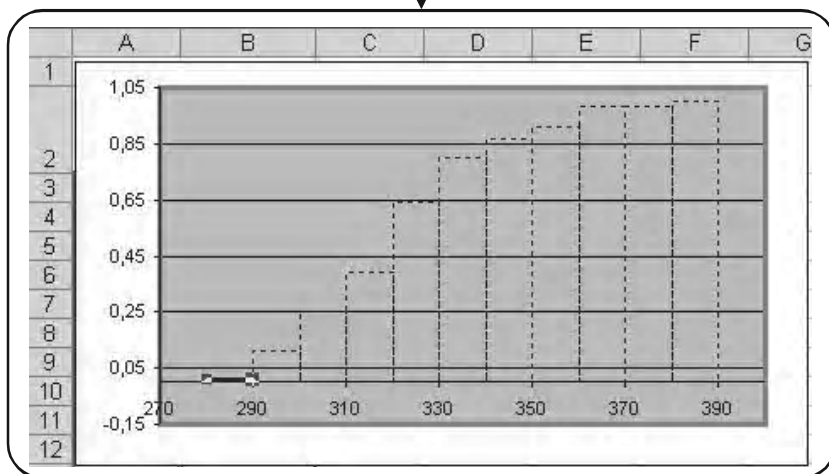
Ставимо курсор миші на першу горизонтальну лінію і двічі клацаємо лівою кнопкою миші. У результаті лінія виділяється двома маркерами.

Подальший крок полягає в натиску правої клавіші миші, виборі меню «Формат рядов даних» і виборі у ньому товстої лінії. У результаті отримуємо (див. мал. 111).

Аналогічно наводяться жирною лінією всі горизонтальні ділянки графіка, що залишилися. Остаточний результат — емпіричну функцію розподілу середньомісячної платні співробітників фірми N (Таблиця В.3-Е) наведено на мал. 112.

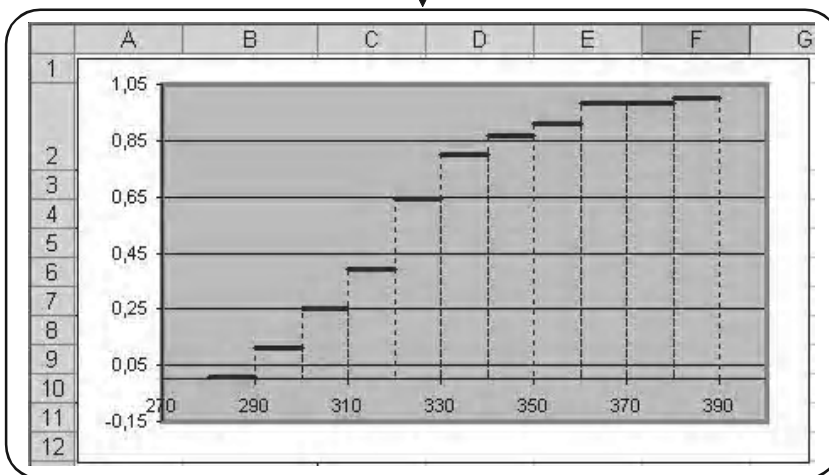
Як завжди, подальше вдосконалення діаграми надаємо читачу.

Креслення жирно першої горизонтальної лінії вихідної гістограми



Мал. 111

Остаточний вигляд емпіричної функції розподілу середньомісячної платні співробітників фірми N (Таблиця В.3-Е)



Мал. 112

Вправи для самостійної роботи

Побудувати

а) побудувати гістограму; б) полігон частот; в) полігон накопичених частот; г) емпіричну функцію розподілу.

1. Наступний розподіл є результатом фіксації числа звернення студентів факультету до студентської лікарні за певний період часу (див. Таблицю 19).

Таблиця 19

Кількість звернень	1	2	3	4	5	6	7	8	9
Кількість студентів	25	20	17	12	8	7	3	0	2

2. Наступний розподіл згрупованих частот є результатом аналізу числа вкладів у банк 1000 вкладникам у гривнях (див. Таблицю 20).

3. У наступному згрупованому розділі наведено число замовлень, прийняте хімчисткою за 120 робочих днів (див. Таблицю 21).

4. У наступній таблиці наведений сумарний бал, набраний 50 студентами по 10 дисциплінам (по 10 бальній системі, див. Таблицю 22).

Таблиця 20

Вклад	Кількість вкладників
100–300	115
130–190	90
160–190	180
190–220	146
220–250	155
250–280	190
280–310	124
	1000

Таблиця 21

Кількість замовлень	Кількість днів
10–20	7
20–30	13
30–40	35
40–50	2
50–60	19
60–70	16
70–80	21
80–90	7
90–100	0
	120

5. У таблиці наведена кількість автомобілів, що паркуються на платній вуличній денній стоянці (див. Таблицю 23).

Таблиця 22

Сумарний бал	Кількість студентів
60–65	1
65–70	2
70–75	4
75–80	13
80–85	14
85–90	6
90–95	7
95–100	3
	50

Таблиця 23

Час парковки (години)	Кількість автомобілів
менше 1 години	23
1–2	38
2–3	32
3–4	48
5–6	23
6–7	12
7–8	8
8–9	4
9–10	2
	190

8.5. Знаходження числових характеристик вибірових даних

Теоретичні положення: параграф 6.2 розділ 6.

8.5.1. Знаходження числових характеристик за допомогою статистичних функцій

Найважливішими числовими характеристиками вибірки є

- вибірове середнє \bar{x}_B ;
- дисперсія D_B ;
- середньоквадратичне відхилення σ_B ,

а також

- виправлена вибірова дисперсія

$$s^2 = \frac{n}{n-1} D_B,$$

- виправлене середньоквадратичне відхилення s .

Для обчислення вибірового середнього \bar{x}_B використовується функція

СРЗНАЧ.

Для обчислення виправленої вибірової дисперсії s^2 і виправленого середньоквадратичного відхилення s використовуються функції

ДИСП, СТАНДОТКЛ

відповідно.

Для обчислення вибірової дисперсії D_B і середньоквадратичного відхилення використовуються функції

ДИСПР, СТАНДОТКЛП

відповідно.

При об'ємі вибірки $n > 30$ функції **ДИСП** і **ДИСПР**, **СТАНДОТКЛ** і **СТАНДОТКЛП** практично дають однакові результати.

Зауваження 1. Перечислені функції застосовуються безпосередньо до даних вибірки генеральної сукупності і не потребують ніякої їх попередньої підготовки. Дивись також функції:

СРЗНАЧ;
ДИСПА, ДИСПРА ;
СТАНДОТКЛА, СТАНДОТКЛПА.

Приклад 1. Для вибірки середньомісячної платні 100 співробітників фірми N, що наведено у Таблиці В (див. стор. 165), підрахувати

- а) вибіркоче середнє \bar{x}_B ;
 - б) вибіркочу дисперсію (виправлену) s^2 ;
 - в) вибіркоче середньоквадратичне відхилення s (виправлене).
- Розв'язання.*

- 1 Представимо початкові дані у вигляді таблиці на мал. 113.
- 2 У лівому нижньому куті таблиці вводимо назви:
 «Вибіркове середнє» (вічко A13),
 «Вибіркова дисперсія» (виправлена, вічко D13),
 «Середньоквадратичне відхилення» (виправлене, вічко G13).

	A	B	C	D	E	F	G	H	I	J
1	Таблиця В									
2	Вибіркове середньомісячне платні 100 співробітників фірми N									
3	338	348	304	314	326	314	324	304	342	308
4	336	304	302	338	314	304	320	321	322	321
5	312	323	336	324	312	312	364	356	362	302
6	322	310	334	292	362	381	304	366	298	304
7	381	368	304	298	368	290	340	328	316	322
8	302	314	292	342	321	322	290	332	298	296
9	296	298	324	338	352	326	318	304	332	322
10	360	312	331	331	304	316	332	282	342	338
11	342	322	324	325	302	328	354	330	316	324
12	334	350	334	324	332	340	324	314	326	323
13	Вибіркове середнє		Вибіркова дисперсія		Вибіркове середнє значення					

Мал. 113

Розділ 8. Розв'язання задач засобами Excel

Задача а)

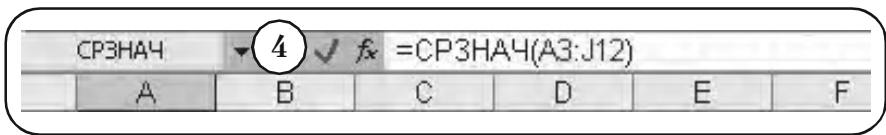
Обчислення вибіркового середнього \bar{x}_B проводимо за допомогою функції

СРЗНАЧ (число 1; число 2; ...),

«число 1», «число 2», ... — це від 1 до 30 числових аргументів, що відповідають вибірці із генеральної сукупності.

- 3 Активізуємо вічко С13.
- 4 Вводимо функцію (див. мал. 114).

= **СРЗНАЧ** (А3 : J12).



Мал. 114

- 5 < Enter >.
- 6 Результат $\bar{x}_B = 324.3$ заноситься у вічко С13 (див. мал. 115).

The image shows a portion of the Excel spreadsheet. Cell C13 contains the value 324,3, and cell D13 contains the value 13. The text 'Вибіркове середнє' is in C13, 'Вибіркова дисперсія' is in D13, and 'Вибіркове середнє значення' is in E13. A circled number 6 is over C13 and a circled number 7 is over E13.

	Вибіркове	Вибіркова	Вибіркове	
13	середнє	324,3	дисперсія	середнє значення

Мал. 115

Задача б)

Обчислення виправленої вибіркової дисперсії s^2 проводимо за допомогою функції

ДИСП (число 1; число 2;...),

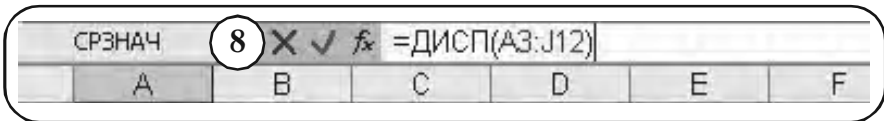
«число 1», «число 2», ... — це від 1 до 30 числових аргументів, що відповідають вибірці із генеральної сукупності.

ДИСП використовує формулу

$$s^2 = \frac{n \sum_{i=1}^n x^2 - \left(\sum_{i=1}^n x \right)^2}{n(n-1)}.$$

7) Активізуємо вічко F13.

8) Вводимо формулу (див. мал. 116)
= ДИСП (A3 : J12).



Мал. 116

9) < Enter >.

10) Результат $s^2 = 436.8$ заноситься у вічко F13 (див. мал. 117).

Вибіркове		Вибіркова	10	Вибіркове		
13 середнє	324,3	дисперсія	436,8	середнє		
			значення			

Мал. 117

Задача в)

Обчислення виправленого середньоквадратичного відхилення s проводимо за допомогою функції

СТАНДОТКЛ (число 1; число 2;...),

«число 1», «число 2»,... — це від 1 до 30 числових аргументів, що відповідають вибірці із генеральної сукупності.

СТАНДОТКЛ використовує формулу

$$s = \sqrt{\frac{n \sum_{i=1}^n x^2 - \left(\sum_{i=1}^n x \right)^2}{n(n-1)}}.$$

$$\frac{|s - \sigma_B|}{s} = \frac{|20,9 - 20,8|}{20,9} = \frac{0,1}{20,9} \approx 0,5\%.$$

Остаточний результат — числові характеристики вибіркової сукупності наведено на мал. 119.

8.5.2. Знаходження числових характеристик вибірки у вигляді згрупованого розподілу частоти

В деяких випадках вимагається знайти числові характеристики для вибірки, що записана у вигляді згрупованого розподілу частоти. У цьому випадку не можна застосувати статистичні функції, вивчені в пункті 8.5.1.

Проведемо їх обчислення шляхом реалізації формул, по яким знаходяться необхідні числові характеристики, засобами Excel.

Приклад 2. Для вибірки середньомісячної платні співробітників фірми N , що задана у вигляді згрупованого розподілу частоти (дивись таблицю В.3 на стор. 169), підрахувати

- а) вибіркоче середнє \bar{x}_B ,
- б) вибіркочу дисперсію D_B ,
- в) вибіркоче середньоквадратичне відхилення σ_B .

Розв'язання.

Задача а)

Використаємо для обчислень формулу (12) пункту 6.2.4 параграфу 6.2 (стор. 214)

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k n_i \tilde{x}_i.$$

Вводимо позначення

\tilde{x}_i — середнє значення (варіанти);

n_i — частота;

$n_i \tilde{x}_i$ — частота_помн_ (середнє_значення);

Розділ 8. Розв'язання задач засобами Excel

$$\frac{n_i \tilde{x}_i}{n} \text{ — (частота_помн_ (середнє_значення))_поділ_об'єм.}$$

- ① Нехай таблиця В.3 вже набрана в Excel у вигляді Таблиці В.3–Е (див. мал. 120).

	А	В
1	Таблиця В.3-Е	
2	Платня	Частоти
3	280-289	1
4	290-299	10
5	300-309	14
6	310-319	14
7	320-329	25
8	330-339	16
9	340-349	7
10	350-359	4
11	360-369	7
12	370-379	0
13	380-389	2
14		100

Мал. 120

Перепишемо Таблицю В.3-Е на новий лист у вигляді нової Таблиці В.3-ЕЕ.1 (див. мал. 121), де введено:

- ② Перші два стовпчика вихідної таблиці;
③ Середнє значення варіанти на кожному класі;
④ Стовпчики для запису значень

$$n_i \tilde{x}_i, \frac{n_i \tilde{x}_i}{n}.$$

- ⑤ Внесемо наступні функції у вічка (див. мал. 122).

	А	В	С	Д	Е
1	Таблиця В.3-ЕЕ.1				
			Середньє_	Частота_помн_	(Частота_)помн_се
2	Платня	Частота	значення	(середньє_зна	редньє_значення)
3	280-289	1		чення)	поділ_об'єм
4	290-299	10			
5	300-309	14			
6	310-319	14			
7	320-329	25			
8	330-339	16			
9	340-349	7			
10	350-359	4			
11	360-369	7			
12	370-379	0			
13	380-389	2			

Мал. 121

	А	В	С	Д	Е
1	Таблиця В.3-ЕЕ.1				
			Середньє_	Частота_помн_	(Частота_)помн_се
2	Платня	Частота	значення	(середньє_зна	редньє_значення)
3	280-289	1	285	285	=D3/100

Мал. 122

Вічко D3 (частота_помн_ (середнє_значення))
 $=B3 * C3$.

Вічко E3 (((частота_помн_ (середнє_значення)) _поділ_об'єм).
 $=D3 / 100$.

Обчислення у стовпчиках D і E проводимо згідно наступної схеми.

- 6 Виділимо вічка D3:E3 (див. мал. 123).
- 7 Ставимо мишу на квадратик у правому нижньому куті виділеної рамки.

Розділ 8. Розв'язання задач засобами Excel

	А	В	С	Д	Е
1	Таблиця В.3-ЕЕ.1				
			Середнє_значення	Частота_помн_(середнє_значення)	(Частота_помн_середнє_значення) поділ об'єм
2	Платня	Частота			
3	280-289	1	285	285	2,85

Мал. 123

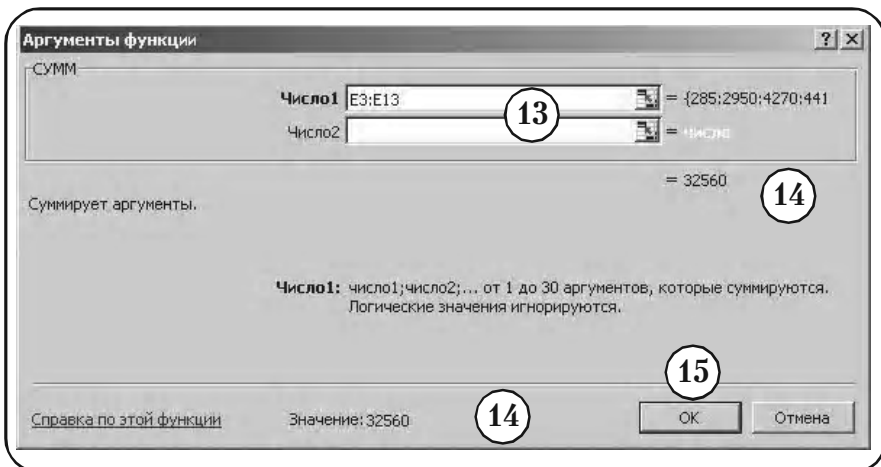
- 8 Курсор у вигляді стрілки змінюється на курсор у вигляді хрестика.
- 9 Не відпускаючи лівої клавіші, розтягуємо вікно, що утворилося, до останнього числового вічка у таблиці.
- 10 Відпускаємо ліву кнопку миші. На мал. 124 показаний результат обчислень у двох стовпчиках таблиці.

	А	В	С	Д	Е
1	Таблиця В.3-ЕЕ.1				
			Середнє_значення	Частота_помн_(середнє_значення)	(Частота_помн_середнє_значення) поділ об'єм
2	Платня	Частота			
3	280-289	1	285	285	2,85
4	290-299	10	295	2950	29,5
5	300-309	14	305	4270	42,7
6	310-319	14	315	4410	44,1
7	320-329	25	325	8125	81,25
8	330-339	16	335	5360	53,6
9	340-349	7	345	2415	24,15
10	350-359	4	355	1420	14,2
11	360-369	7	365	2555	25,55
12	370-379	0	375	0	0
13	380-389	2	385	770	7,7
14				Вибіркове середнє:	

Мал. 124

Переходимо безпосередньо до обчислення вибіркового середнього.

- 11 Активізуємо вічко E14.
- 12 Вставить → Функция → Математическая → СУММ.
- 13 В меню функції СУМ, що з'явилося, вводимо адреси вічок E3:E13, дані в яких будуть просумовані (див. мал. 125).



Мал. 125

- 14 Результат сумування – вибіркоче середнє $\tilde{x}_B = 325,6$, можна прочитати в меню функції СУММ.
- 15 □ ОК.
- 16 Результат сумування – вибіркоче середнє $\tilde{x}_B = 325,6$, заноситься у вічко E14 (див. мал. 126).

У вічку E14 представлено остаточний результат: вибіркоче середнє $\tilde{x}_B = 325,6$ середньомісячної платні співробітників фірми N (Таблиця В.3-Е).

Розділ 8. Розв'язання задач засобами Excel

Остаточний результат: вибіркове середнє середньомісячної платні співробітників фірми N (Таблиця В.3-Е)

	А	В	С	Д	Е
1	Таблиця В.3-ЕЕ.1				
2	Платня	Частота	Середнє значення	Частота_помн_ (середнє_значення)	(Частота_помн_реднє_значення)_поділ_об'єм
3	280-289	1	285	285	2,85
4	290-299	10	295	2950	29,5
5	300-309	14	305	4270	42,7
6	310-319	14	315	4410	44,1
7	320-329	25	325	8125	81,25
8	330-339	16	335	5360	53,6
9	340-349	7	345	2415	24,15
10	350-359	4	355	1420	14,2
11	360-369	7	365	2555	25,55
12	370-379	0	375	0	0
13	380-389	2	385	770	7,7
14				Вибіркове середнє:	325,6

Мал. 126

Як видно, отриманий результат $\tilde{x}_B = 325,6$ незначним чином відрізняється від точного результату $\tilde{x}_B = 324,3$. Помилка виникла за рахунок округлення всіх варіант у класі до середнього значення. Проте, як показує практика, помилка, що вноситься при цьому, незначна. Цей факт також підтверджують проведені вище обчислення.

Задача б)

Використаємо для обчислення вибіркової дисперсії формулу (14) пункту 6.2.4 параграфу 6.2 (стор. 214) для згрупованих даних

$$D_B = \frac{1}{n} \sum_{i=1}^k n_i \tilde{x}_i^2 - (\tilde{x}_B)^2 .$$

Вибіркове середнє $\tilde{x}_B = 325,6$.

Введемо позначення.

\tilde{x}_i — середнє значення (варіанти);

n_i — частота;

\tilde{x}_i^2 — квадрат_(середнє_значення);

$n_i \tilde{x}_i^2$ — частота_помн_(квадрат_(середнє_значення));

$\frac{n_i \tilde{x}_i^2}{n}$ — (частота_помн_(квадрат_(середнє_значення))_поділ_об'єм).

- 1) Перепишемо Таблицю В.3-Е на новий лист у вигляді нової Таблицю В.3-ЕЕ.3 (див. мал. 127), де введено стовпчики для запису значень

$$\tilde{x}_i, \tilde{x}_i^2, n_i \tilde{x}_i^2, \frac{n_i \tilde{x}_i^2}{n}.$$

- 2) Внесемо наступні функції у вічка (див. мал. 128).

	A	B	C	D	E	F
1				Таблиця В.3-ЕЕ.3		
	1		Середнє	квадрат_(середнє_значення)	Частота_помн_(квадрат_(середнє_значення))	(Частота_помн_(квадрат_(середнє_значення)))_поділ_об'єм
2	Платня	Частота	значення			
3	280-289	1	285			
4	290-299	10	295			
5	300-309	14	305			
6	310-319	14	315			
7	320-329	25	325			
8	330-339	16	335			
9	340-349	7	345			
10	350-359	4	355			
11	360-369	7	365			
12	370-379	0	375			
13	380-389	2	385			

Мал. 127

Розділ 8. Розв'язання задач засобами Excel

	A	B	C	D	E	F
1	Таблиця В.3-ЕЕ.3					
			Середньє	квадрат_(се редньє_зна чення)	Частота_помн_(квадрат_(серед ньє_значення))	(Частота_помн_(квадрат_(середньє _значення)))_поділ об'єм
2	Платня	Частота	значення			
3	280-289	1	285	81225	81225	=E3/100

Мал. 128

Вічко D3 (квадрат_(середньє_значення))

$$=C3*C3.$$

Вічко E3 (частота_помн_(квадрат_(середньє_значення)))

$$=B3*D3.$$

Вічко F3 ((частота_помн_(квадрат_(середньє_значення)))_поділ_об'єм)

$$=E3/100.$$

3 Виділимо вічка D3:F3 (див. мал. 129).

	A	B	C	D	E	F
1	Таблиця В.3-ЕЕ.3					
			Середньє	квадрат_(се редньє_зна чення)	Частота_помн_(квадрат_(серед ньє_значення))	(Частота_помн_(квадрат_(середньє _значення)))_поділ об'єм
2	Платня	Частота	значення			
3	280-289	1	285	81225	81225	81225

Мал. 129

4 Ставимо мишу на квадратик у правому нижньому куті виділеної рамки.

5 Курсор у вигляді стрілки змінюється на курсор у вигляді хрестика.

6 Не відпускаючи лівої клавіші, розтягуємо вікно, що утворилося, до останнього числового вічка у таблиці.

- 7 Відпускаємо ліву кнопку миші. На мал. 130 показаний результат обчислень у трьох стовпчиках.

	A	B	C	D	E	F
1	Таблиця В.3-ЕЕ.3					
2	Платня	Частота	Середнє значення	квадрат_(середнє_значення)	Частота_помн_(квадрат_(середнє_значення))	(Частота_помн_(квадрат_(середнє_значення)))/поділ об'єм
3	280-289	1	285	81225	81225	812,25
4	290-299	10	295	87025	870250	8702,5
5	300-309	14	305	93025	1302350	13023,5
6	310-319	14	315	99225	1389150	13891,5
7	320-329	25	325	105625	2640625	26406,25
8	330-339	16	335	112225	1795600	17956
9	340-349	7	345	119025	833175	8331,75
10	350-359	4	355	126025	504100	5041
11	360-369	7	365	133225	932575	9325,75
12	370-379	0	375	140625	0	0
13	380-389	2	385	148225	296450	2964,5

Мал. 130

Переходимо безпосередньо до обчислення вибіркової дисперсії.

- 8 У вічко F14 вносимо функцію
 $=\text{СУММ}(F3:F13)-325,6 * 325,6$.

9 < Enter > .

- 10 На мал. 131 показано остаточний результат: вибірка дисперсія середньомісячної платні співробітників фірми N (Таблиця В.3-Е)

$$D_B = 439,64 \text{ грн}^2.$$

Задача в)

Вибіркове середнє квадратичне відхилення знаходиться за формулою $\sigma_B = \sqrt{D_B}$.

- 1 У вічко E15 вводимо текст
 «Вибіркове_середньоквадратичне_відхилення».

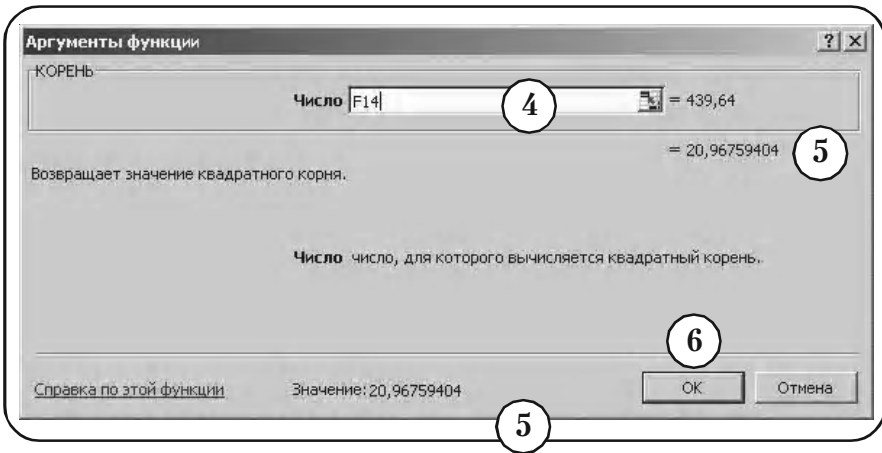
Розділ 8. Розв'язання задач засобами Excel

Остаточний результат: вибіркова дисперсія середньомісячної платні співробітників фірми N (Таблиця В.3-Е)

	A	B	C	D	E	F
1	Таблиця В.3-ЕЕ.3					
2	Платня	Частота	Середньє значення	квадрат_(середньє значення)	Частота_помн_(квадрат_(середньє значення))	(Частота_помн_(квадрат_(середньє значення)))/об'єм
3	280-289	1	285	81225	81225	812,25
4	290-299	10	295	87025	870250	8702,5
5	300-309	14	305	93025	1302350	13023,5
6	310-319	14	315	99225	1389150	13891,5
7	320-329	25	325	105625	2640625	26406,25
8	330-339	16	335	112225	1795600	17956
11	360-369	7	365	133225	932575	9325,75
12	370-379	0	375	140625	0	0
13	380-389	2	385	148225	296450	2964,5
14		100			Вибіркова дисперсія	439,64

Мал. 131

- 2 Активізуємо вічко F15.
- 3 Вставка → Функція → Математические → КОРЕНЬ.
- 4 У меню, що з'явилося (див. мал. 132), набираємо число 439,64, з якого буде взятий корінь.
- 5 Результат обчислення – вибіркоче середнє квадратичнє відхилення $\sigma_B = 20,97$, можна прочитати в меню функції **КОРЕНЬ**.
- 6 ОК.
- 7 Результат обчислення – вибіркоче середнє квадратичнє відхилення $\sigma_B = 20,97$, заноситься у вічко F15 (див. мал. 133).



Мал. 132

Остаточний результат: вибіркове середнє квадратичне відхилення середньомісячної платні співробітників фірми N (Таблиця В.3-Е)

	A	B	C	D	E	F
1	Таблиця В.3-ЕЕ.3					
2	Платня	Частота	Середнє значення	квадрат_(середнє_значення)	Частота_помн_(квадрат_(середнє_значення))	(Частота_помн_(квадрат_(середнє_значення)))/поділ_об'єм
3	280-289	1	285	81225	81225	812,25
4	290-299	10	295	87025	870250	8702,5
5	300-309	14	305	93025	1302350	13023,5
6	310-319	14	315	99225	1389150	13891,5
7	320-329	25	325	105625	2640625	26406,25
8	330-339	16	335	112225	1795600	17956
14		100			Вибіркова дисперсія	439,64
15					Вибіркове середнє квадратичне відхилення	20,97

Мал. 133

На мал. 133 представлено остаточний результат: вибіркове середнє квадратичне відхилення середньомісячної платні співробітників фірми N (Таблиця В.3-Е) $\sigma_B = 20,97$.

Вправи для самостійної роботи

Обчислити

- а) вибіркове середнє;
- б) вибіркєву дисперсію;
- в) середнє квадратичне відхилення вибірки.

1. Обстеження, проведене у науково-дослідному інституті N , вказує середнє число телефонних розмов, що відбуваються за робочий день у вибірці з 200 співробітників (див. Таблицю 24).

Таблиця 24

Кількість телефонних розмов	Кількість співробітників
0	10
1	50
2	100
3	28
4	10
5	2
	200

2. Нижче наведений середній час, що витрачається студентами на дорогу від будинку до університету N (див. Таблицю 25).

3. Нижче наведені відсотки виконання учбової програми до зимової сесії у вибірці з 200 студентів університету N (див. Таблицю 26).

4. Обстеження 180 власників автомобілів виявило наступний середній вік автомобілів у власників (див. Таблицю 27).

5. У Таблиці 28 наведено віковий склад службовців однієї з фірм.

Таблиця 25

Вклад	Кількість вкладників
100–130	115
130–160	90
160–190	180
190–220	146
220–250	155
250–280	190
280–310	124
	1000

Таблиця 26

Відсотки виконання програми	Кількість студентів
менше 10	2
10–20	3
20–30	10
30–40	15
40–50	25
50–60	55
70–80	90
80–90	150
90–100	50
	200

Таблиця 27

Вік автомобіля	Кількість власників
1–3	19
3–6	15
6–9	20
9–12	48
12–15	40
15–18	20
18–21	18
	180

Таблиця 28

Вік	Кількість службовців
18–23	8
23–28	15
28–33	36
33–38	48
38–43	40
43–48	38
48–53	20
53–58	15
58–63	10
	230

8.6. Комбінаторика

8.6.1. Обчислення факторіалу числа за допомогою функції ФАКТР

Функція **ФАКТР** повертає факторіал числа

$$n! = 1 \cdot 2 \cdot \dots \cdot n.$$

Вона має представлення:

ФАКТР(число),

де «число» — це невід'ємне число, факторіал якого обчислюється, якщо число не ціле, то здійснюється відсікання.

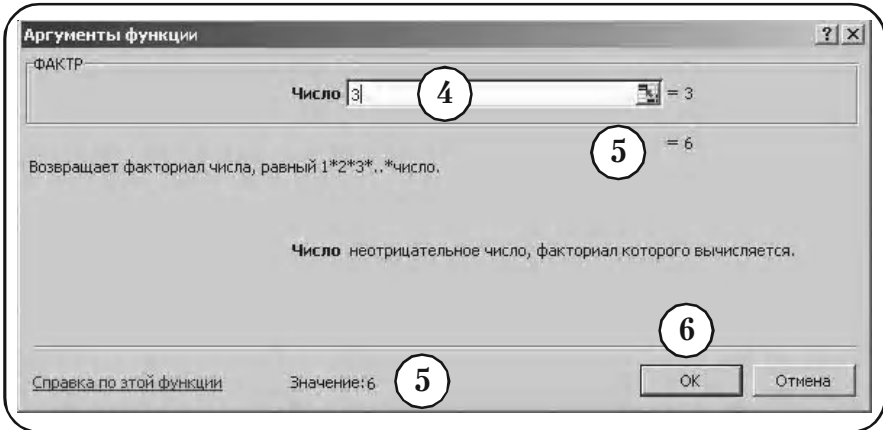
Приклад 1. Обчислити факторіал 3!, 9!, 0!, 3.7!.

- ① Вводимо назву вічка (див. мал. 134).

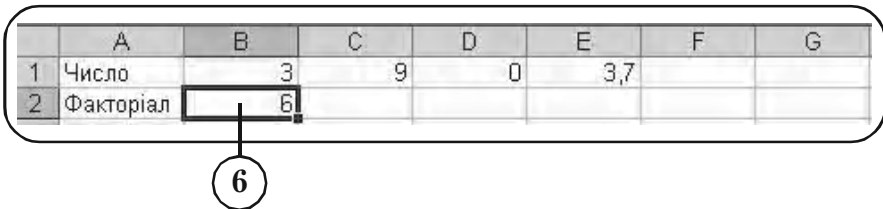
	A	B	C	D	E	F	G
1	Число	3	9	0	3,7		
2	Факторіал						

Мал. 134

- ② Активізуємо вічко B2.
- ③ **Вставка** → **Функція** → **Математичні** → **ФАКТР**.
- ④ В меню функції ФАКТР у вікні «Число» вводимо 3 (див. мал. 135).
- ⑤ Результат обчислений 3! = 6 можна прочитати в меню функції.
- ⑥ **ОК**.
- ⑦ Результат 3! = 6 заноситься у вічко (див. мал. 136).
- ⑧ Аналогічно обчислюється факторіал чисел 9, 0, 3.7 (див. мал. 137).

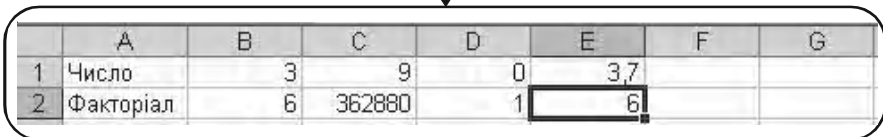


Мал. 135



Мал. 136

Остаточный результат обчислень



Мал. 137

8.6.2. Обчислення сполучень за допомогою функції ЧИСЛОКОМБ

Функція **ЧИСЛОКОМБ** використовується для визначення числа всіляких сполучень об'єктів у групі і реалізує формулу

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Вона має представлення:

ЧИСЛОКОМБ (число; число_вибраних),

де «число» — це число об'єктів (n);

«число_выбраних» — це число об'єктів у кожній комбінації (m).

Приклад 2. Потрібно вибрати команду із 3 студентів для участі в конкурсі. Є 8 кандидатів. Підрахуйте можливе число команд.

Розв'язання.

В умовах задачі

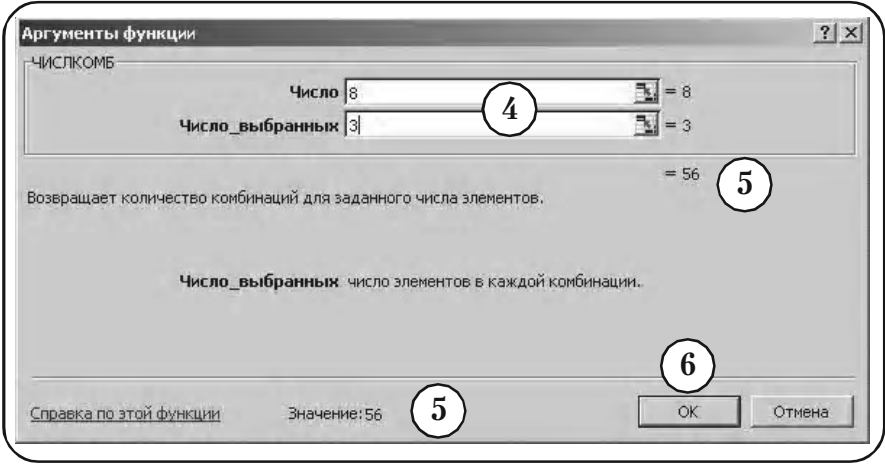
число = 8; число_выбраних = 3.

- ① Складемо таблицю (див. мал. 138).

	A	B	C	D	E	F
1	Кількість команд					

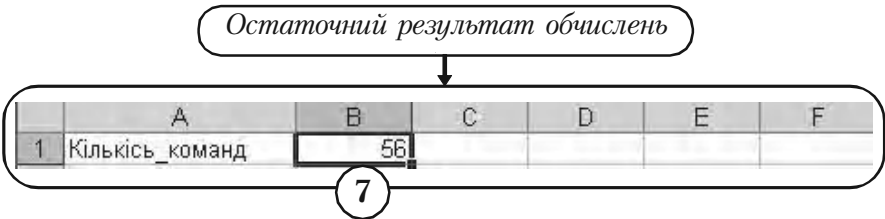
Мал. 138

- ② Активізуємо вічко B1.
- ③ Вставка → Функція → Математичні → ЧИСЛОКОМБ.
- ④ В меню функції **ЧИСЛОКОМБ**, що з'явилося у вікні, вводимо параметри задачі (див. мал. 139).
- ⑤ Результат $C_8^3 = 56$ можна прочитати в меню функції.
- ⑥ □ ОК.



Мал. 139

7 Результат $C_8^3 = 56$ заноситься у вічко В1 (див. мал. 140).



Мал. 140

8.6.3. Обчислення розміщень за допомогою функції ПЕРЕСТ

Функція ПЕРЕСТ повертає кількість перестановок для заданого числа об'єктів, які вибираються із загального числа об'єктів. У перестановці суттєвий внутрішній порядок об'єктів. Цим перестановки відрізняються від сполучень, для яких внутрішній порядок не суттєвий. Функція ПЕРЕСТ реалізує формулу

$$A_n^m = \frac{n!}{(n-m)!}$$

Вона має представлення:

ПЕРЕСТ (число; число_выбранных),

де «число» — це ціле число, що задає кількість об'єктів (n);

«число_выбранных» — це ціле число, що задає кількість об'єктів у кожній перестановці (m).

Приклад 3. В яхт-клубі N із 15 членів потрібно вибрати правління із президента, скарбника і секретаря. Скільки існує варіантів вибрати правління?

Розв'язання. В умовах задачі

число = 15; число_выбранных = 3.

① Складемо таблицю (див. мал. 141).

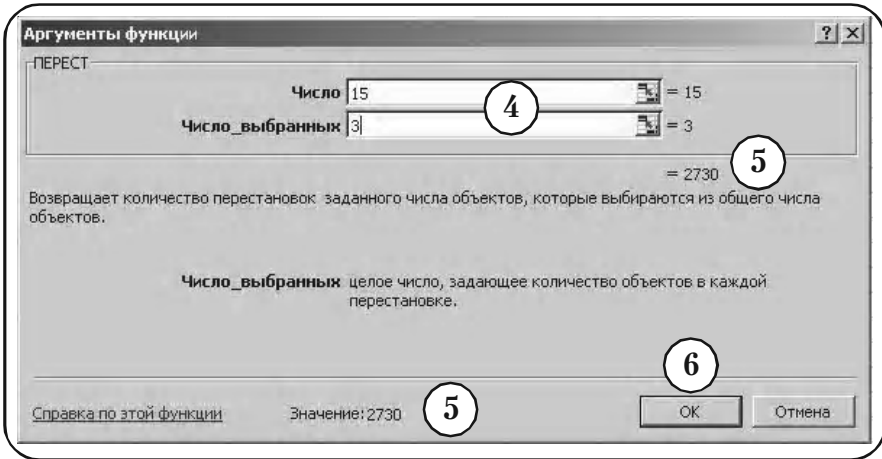
	A	B	C	D	E	F
1	Кількість правління					

Мал. 141

② Активізуємо вічко В1.

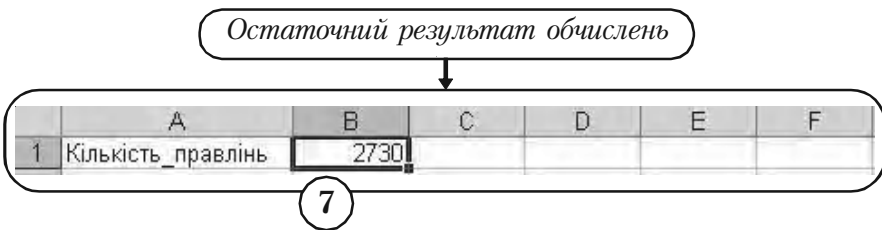
③ Вставка → Функція → Статистичні → ПЕРЕСТ.

④ В меню функції ПЕРЕСТ, що з'явилося у вікні, вводимо параметри задачі (див. мал. 142).



Мал. 142

- 5 Результат $A_{15}^3 = 2730$ можна прочитати в меню функції.
- 6 ОК.
- 7 Результат $A_{15}^3 = 2730$ заноситься у вічко B1 (див. мал. 143).



Мал. 143

8.7. Обчислення, пов'язані з законами розподілу випадкових величин

8.7.1. Біноміальний розподіл

Теоретичні положення: пункт 4.2.1 параграфу 4.2, розділ 4.

Нижче використаємо функцію

БИНОМРАСПР (число_успехов; число_испытаний; вероятность_успеха; интегральная),

де «число_успехов» — це кількість успішних випробувань (m);

«число_испытаний» — це число незалежних випробувань (n);

«вероятность_успеха» — це імовірність успіху кожного випробування (p);

«интегральная» — це логічне значення, що визначає форму функції.

Якщо «интегральная» = **ЛОЖЬ**, то функція **БИНОМРАСПР** використовує формулу Бернуллі

$$P_n(m) = C_n^m p^m (1-p)^{n-m} \quad (1)$$

для визначення імовірності події, яка заключається в тому, що настануть рівно m успіхів в n іспитах.

Якщо «интегральная» = **ИСТИНА**, то функція **БИНОМРАСПР** використовує формулу інтегрального біноміального розподілу

$$P_n(k \leq m) = \sum_{y=0}^m P_n(y) \quad (2)$$

для визначення імовірності події, що настануть не більше m успіхів в n іспитах.

Зауваження 1. Замість слова «**ЛОЖЬ**» можна набрати число «0», а замість слова «**ИСТИНА**» — число «1».

Приклад 1. Після багаторічних спостережень власник магазину встановив, що ймовірність здійснення покупки покупцем, що випадково зайшов до магазину, складає 0,4. У припущенні, що в певний день випадковим чином відібрано 25 покупців, визначити ймовірності

наступних подій (вважаємо, що виконуються умови повторних випробувань схеми Бернуллі)

- а) покупку зроблять рівно 10 покупців,
- б) не більше 10 покупців,
- в) менше 10 покупців,
- г) не менше 10 покупців,
- д) від 10 до 15 покупців.

Розв'язання. За умовою задачі випадкова величина — кількість покупок серед 25 покупців, має біноміальний розподіл, де

$$\text{«число_испытаний»} = 25, \text{ «вероятность успеха»} = 0,4.$$

- 1 Складемо таблицю (див. мал. 144).

	А	В	С	Д	Е	Ф	Г
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність						

Мал. 144

Задача а)

Для підрахування імовірності $P_{25}(10)$ події, що кількість успіхів в 25 випробуваннях дорівнює рівно 10, використаємо формулу (1).

- 2 Активізуємо вічко В3.
- 3 Вставка → Функція → Статистическая → БИНОМРАСП
- 4 У вікні меню функції **БИНОМРАСП**, що з'явилося, набираємо параметри задачі (див. мал. 145, 146).
- 5 Результат $P_{25}(10) = 0,161$ можна прочитати в меню функції.
- 6 ОК.
- 7 Результат $P_{25}(10) = 0,161$ заноситься у вічко В3 (див. мал. 147).

4

	A	B	C	D	E	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	РАСП()					

Мал. 145

Аргументы функции

-БИНОМРАСП-

Число_успехов: 10 = 10

Число_испытаний: 25 = 25

Вероятность_успеха: 0,4 = 0,4

Интегральная: 0 = ЛОЖЬ

= 0,161157939

Возвращает отдельное значение биномиального распределения.

Интегральная: логическое значение, определяющее вид функции: интегральная функция распределения (ИСТИНА) или весовая функция распределения (ЛОЖЬ).

Справка по этой функции Значение: 0,161157939

OK Отмена

Мал. 146

Імовірність події, що кількість успіхів в 25 випробуваннях дорівнює 10

7

	A	B	C	D	E	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	0,161					

Мал. 147

Задача б)

Для підрахування імовірності $P_{25}(k \leq 10)$ події, що кількість успіхів в 25 випробуваннях не більше 10, використаємо формулу (2).

8) Активізуємо вічко С3.

9) Набираємо функцію (див. мал. 148)
= **БИНОМРАСПР** (10; 25; 0,4; **ИСТИНА**).

Імовірність події, що кількість успіхів в 25 випробуваннях не більше 10

	А	В	С	Д	Е	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	0,161	0,586				

11

Мал. 148

10) < Enter >.

11) Результат $P_{25}(k \leq 10) = 0.586$ заноситься у вічко С3.

Задача в)

Для підрахування імовірності $P_{25}(k < 10)$ події, що кількість успіхів в 25 випробуваннях менше 10, використаємо співвідношення $P_{25}(k < 10) = P_{25}(k \leq 9)$.

12) Активізуємо вічко D3.

13) Набираємо функцію (див. мал. 149)
= **БИНОМРАСПР** (9; 25; 0,4; **ИСТИНА**).

14) < Enter >.

Розділ 8. Розв'язання задач засобами Excel

Імовірність події, що кількість успіхів в 25 випробуваннях менше 10

13

D3	fx = БИНОМРАСП(9;25;0,4;1)						
	A	B	C	D	E	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число_покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	0,161	0,586	0,425			

15

Мал. 149

15) Результат $P_{25}(k < 10) = P_{25}(k \leq 9) = 0,425$ заноситься у вічко D3.

Задача г)

Для підрахування імовірності $P_{25}(k \geq 10)$ події, що кількість успіхів в 25 випробуваннях не менше 10, використаємо співвідношення $P_{25}(k \geq 10) = 1 - P_{25}(k < 10)$.

16) Активізуємо вічко E3 і заносимо функцію (див. мал. 150) $= 1 - D3$.

Імовірність події, що кількість успіхів в 25 випробуваннях не менше 10

E3

E3	fx = 1-D3						
	A	B	C	D	E	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число_покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	0,161	0,586	0,425	0,575		

17

Мал. 150

17 < Enter >.

18 У вічку E3 читаємо результат

$$P_{25}(k \geq 10) = 1 - P_{25}(k < 10) = 1 - 0,425 = 0,575.$$

Задача д)

Для підрахування імовірності $P_{25}(10 \leq k \leq 15)$ події, що кількість успіхів в 25 випробуваннях знаходиться в межах від 10 до 15, використаємо співвідношення

$$P_{25}(10 \leq k \leq 15) = P_{25}(k \leq 15) - P_{25}(k < 10).$$

19 Активізуємо вічко F3.

20 Набираємо функцію (див. мал. 151)

= **БИНОМРАСП**(15; 25; 0,4; **ИСТИНА**) – D3.

Імовірність події, що кількість успіхів в 25 випробуваннях знаходиться в межах від 10 до 15

F3		fx =БИНОМРАСП(15;25;0,4;1)-D3					
	A	B	C	D	E	F	G
1		задача а)	задача б)	задача в)	задача г)	задача д)	
2	число покупок	рівно 10	не більше 10	менше 10	не менше 10	від 10 до 15	
3	імовірність	0,161	0,586	0,425	0,575	0,562	

Мал. 151

21 < Enter >.

22 У вічку F3 читаємо кінцевий результат

$$P_{25}(10 \leq k \leq 15) = P_{25}(k \leq 15) - P_{25}(k < 10) = 0,987 - 0,425 = 0,562.$$

8.7.2. Розподіл Пуассона

Теоретичні положення: пункт 4.2.1 параграфу 4.2, розділ 4.

Нижче використаємо функцію

ПУАССОН (m ; среднее; интегральное),

де m — це кількість подій;

«среднее» — це очікуване число подій за інтервал (a);

«интегральная» — це логічне значення, що визначає форму функції.

Якщо аргумент «интегральная» = **ЛОЖЬ**, то функція використовує формулу

$$P(X = m) = \frac{e^{-a} a^m}{m!} \quad (3)$$

для визначення імовірності подій, що настануть рівно m успіхів за розглядуваний інтервал. Якщо аргумент «интегральная» = **ИСТИНА**, то функція використовує формулу інтегрального пуассонівського розподілу

$$P(X \leq m) = \sum_{k=0}^m \frac{e^{-a} a^k}{k!} \quad (4)$$

для визначення імовірності події, що настануть не більше m успіхів за розглядуваний інтервал.

Приклад 2. Покупець заходить в магазин N випадковим чином і середнє сподіване значення складає 3,4 чол./хвилину. За припущенням, що випадкова величина — число покупців, що заходять до магазину, підпорядкована закону Пуассона, визначити, що на протязі двох хвилин зайде

- а) рівно 5 покупців;
- б) не більше 5 покупців;
- в) не менше 6 покупців.

Розв'язання. За умовою задачі

$$\text{среднее} = 3,4 \times 2 = 6,8 \text{ чол./хвилину.}$$

① Складемо таблицю (див. мал. 152).

	A	B	C	D	E	F
1		задача а)	задача б)	задача в)		
2	Число покупців	рівно 5	не більше 5	не менше 6		
3	Імовірність					

Мал. 152

Задача а)

Для підрахування імовірності $P(X = 5)$ події, що кількість успіхів на протязі двох хвилин дорівнює 5, використовуємо формулу (3).

- 2) Активізуємо вічко В3.
- 3) Вставка → Функція → Статистическая → ПУАССОН.
- 4) У вікні меню функції ПУАССОН, що з'явилося, вводимо параметри задачі (див. мал. 153, 154).

	A	B	C	D	E	F
1		задача а)	задача б)	задача в)		
2	Число покупців	рівно 5	не більше 5	не менше 6		
3	Імовірність	=ПУАССОН(5;6,8;0)				

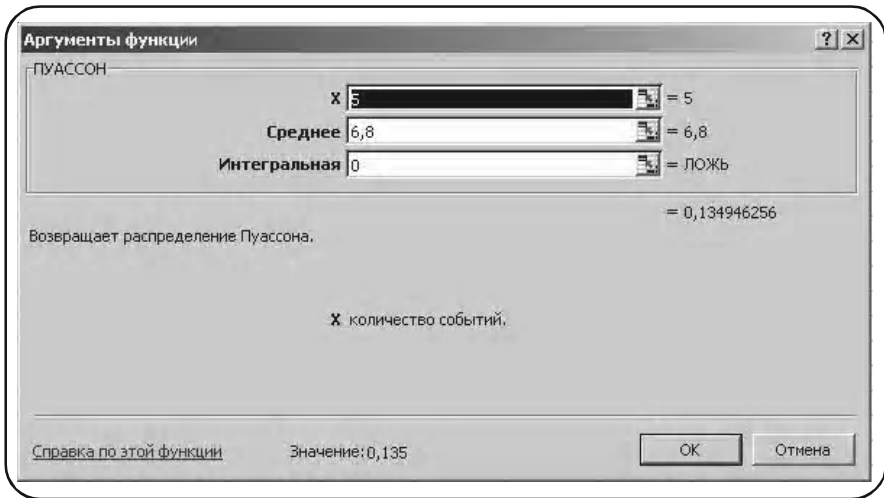
Мал. 153

- 5) Результат $P(X = 5) = 0,135$ можна прочитати в меню функції.
- 6) ОК.
- 7) Результат $P(X = 5) = 0.135$ заноситься у вічко В3 (див. мал. 155).

Задача б)

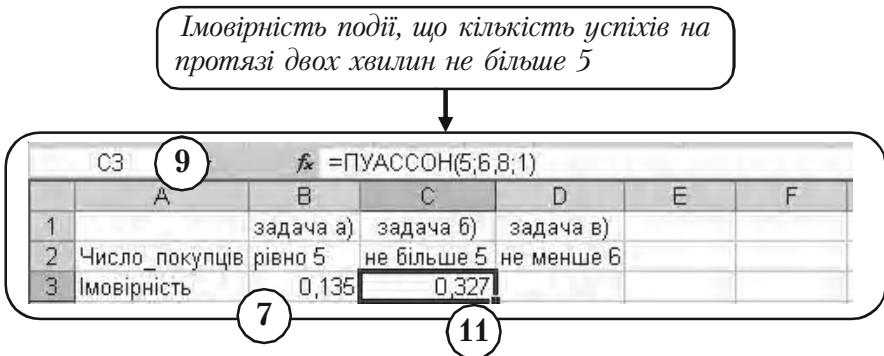
Для підрахування імовірності $P(X < 5)$ події, що кількість успіхів на протязі двох хвилин не більше 5, використаємо формулу (4).

- 8) Активізуємо вічко С3.



Мал. 154

- 9 Набираємо функцію (див. мал. 155)
=ПУАССОН(5; 6,8; ИСТИНА).
- 10 < Enter >.
- 11 Результат $P(X \leq 5) = 0,327$ заноситься у вічко C3.



Мал. 155

Задача в)

Для підрахування імовірності $P(X \geq 6)$ події, що кількість успіхів на протязі двох хвилин не менше 6, використаємо співвідношення

$$P(X \geq 6) = 1 - P(X < 6) = 1 - P(X \leq 5).$$

12) Активізуємо вічко D3.

13) Набираємо функцію (див. мал. 156)
 $= 1 - C(3)$.

Імовірність події, що кількість успіхів на протязі двох хвилин не менше 6

	A	B	C	D	E	F
1		задача а)	задача б)	задача в)		
2	Число покупок равно 5	рівно 5	не більше 5	не менше 6		
3	Імовірність	0,135	0,327	0,673		

Мал. 156

12) < Enter >.

13) Результат

$$P(X \geq 6) = 1 - P(X < 6) = 1 - P(X \leq 5) = 1 - 0,327 = 0,673$$

заноситься у вічко D3.

8.7.3. Нормальний розподіл

Теоретичні положення: пункт 4.3.2 параграфу 4.3, розділ 4.

Нижче використаємо функцію

НОРМРАСП(x ; среднее; стандартное _откл; интегральная),

Розділ 8. Розв'язання задач засобами Excel

x — це значення, для якого будується розподіл;
«среднее» — це середнє арифметичне розподілу (математичне сподівання);

«стандартное_откл» — це стандартне відхилення розподілу;

«интегральная» — це логічне значення, що визначає форму функції.

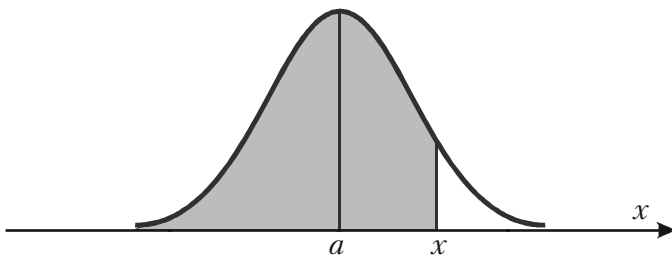
Якщо «интегральная» = **ЛОЖЬ**, то функція **НОРМРАСПР** використовує функцію щільності імовірності нормального розподілу

$$f(x, a, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \quad (5)$$

Якщо «интегральная» = **ИСТИНА**, то функція **НОРМРАСПР** використовує інтегральну функцію розподілу

$$F(x, a, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^x e^{-\frac{(y-a)^2}{2\sigma^2}} dy. \quad (6)$$

і значення функції дає площу криволінійної фігури, обмеженої кривою $y = f(x, a, \sigma)$ і вертикальною лінією, що проходить через точку x (див. мал. 157).



Мал. 157

Зауваження 1. Якщо покласти

среднее = 0, стандартное_откл = 1,

то отримуємо стандартний нормальний розподіл.

Зауваження 2. Дивись також функції

НОРМАЛИЗАЦИЯ, НОРМОБР, НОРМСТОБР, НОРМСТРАСП.

Приклад 3. Картопля розфасована в пакети по 5 кг. Вага пакетів підпорядкована нормальному закону із середнім значенням 5 кг і стандартним відхиленням 0,1. Знайти ймовірність, що навмання взятий пакет

- а) важить менше 4,8 кг;
- б) важить більше 5,1 кг;
- в) вага пакета лежить у межах від 4,8 кг до 5,1 кг.

Розв'язання. За умовами нашої задачі маємо

$$\text{середнее} = 5 \text{ кг, стандартное_откл} = 0.1 \text{ кг.}$$

- 1) Утворюємо таблицю (див. мал. 158).

	А	В	С	1	Д	Е
1		задача а)	задача б)	задача в)		
2	Вага пакунку	менше 4,8	більше 5,1	у межах від 4,8 до 5,1		
3	Імовірність					

Мал. 158

Задача а)

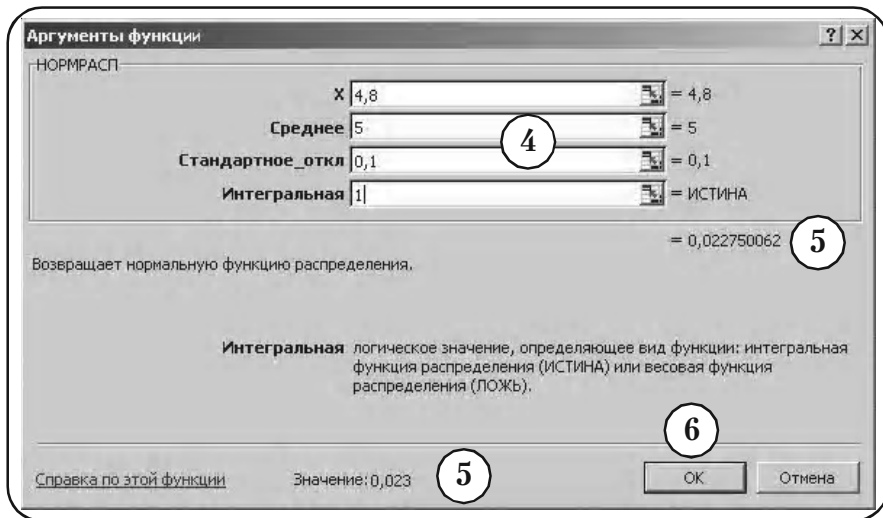
Для підрахування ймовірності $P(X < 4,8)$ події, що навмання взятий пакет важить менше 4,8 кг, використаємо формулу (6).

- 2) Активізуємо вічко ВЗ.
- 3) Вставка → Функція → Статистическая → НОРМРАСП.
- 4) У вікні меню функції **НОРМРАСП**, що з'явилося, набираємо параметри задачі (див. мал. 159, 160).

3>НОРМРАСП 4 X ✓ ✎ =НОРМРАСП(4,8;5;0,1;1)						
	А	В	С	Д	Е	
1		задача а)	задача б)	задача в)		
2	Вага пакунку	менше 4,8	більше 5,1	у межах від 4,8 до 5,1		
3	Імовірність	,8;5;0,1;1)				

Мал. 159

Розділ 8. Розв'язання задач засобами Excel

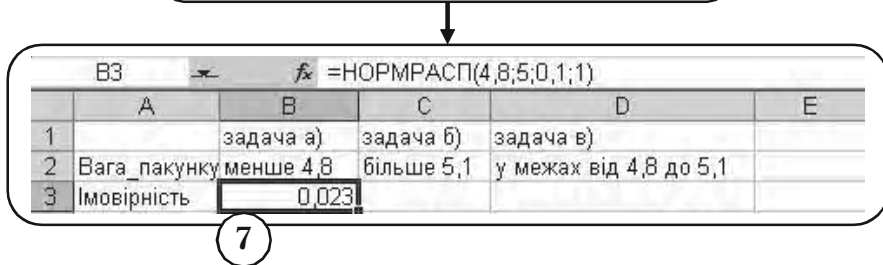


Мал. 160

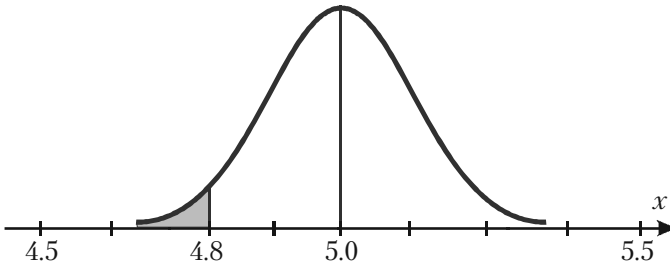
- 5 Результат $P(X < 4.8) = 0,0228$ можна прочитати в меню функції.
- 6 ОК.
- 7 Результат $P(X < 4.8) = 0,0228$ заноситься у вічко В3 (див. мал. 161).

Геометрично ця імовірність дорівнює площі, що вказана на мал. 162.

Імовірність події, що навмання взятий пакет важить менше 4,8 кг



Мал. 161



Мал. 162

Задача б)

Для підрахування імовірності $P(X > 5,1)$ події, що навання взятий пакет важить більше 5,1 кг використаємо співвідношення

$$P(X > 5,1) = 1 - P(X < 5,1).$$

8) Активізуємо вічко С3.

9) Набираємо функцію (див. мал. 163)

$$=1 - \text{НОРМРАСП}(5,1; 5; 0,1; \text{ИСТИНА}).$$

Імовірність події, що навання взятий пакет важить більше 5.1 кг

↓

С3		=1-НОРМРАСП(5,1;5;0,1;1)			
	A	B	C	D	E
1		задача а)	задача б)	задача в)	
2	Вага пакунку	менше 4,8	більше 5,1	у межах від 4,8 до 5,1	
3	Імовірність	0,023	0,159		

Мал. 163

10) < Enter >.

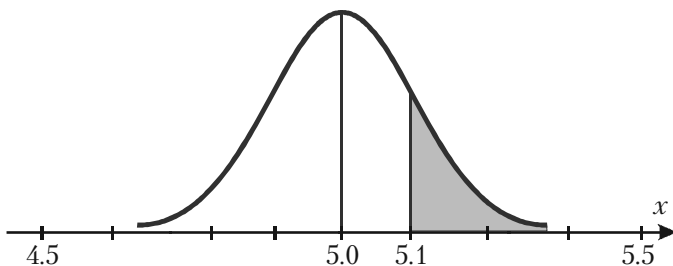
11) Результат

$$P(X > 5,1) = 1 - P(X < 5,1) = 1 - 0,8413 = 0,1587$$

заноситься у вічко С3.

Розділ 7. Статистична перевірка гіпотез

Геометрично ця імовірність дорівнює площі, що вказана на мал. 164.



Мал. 164

Задача в)

Для підрахування імовірності $P(4,8 < X < 5,1)$ події, що вага наважання взятого пакету лежить у межах від 4,8 кг до 5,1 кг, використаємо співвідношення

$$P(4,8 < X < 5,1) = P(X < 5,1) - P(X < 4,8).$$

(12) Активізуємо вічко D3.

(13) Набираємо функцію (див. мал. 165)

$$= \text{НОРМРАСПР}(5,1; 5; 0,1; \text{ИСТИНА}) - \text{НОРМРАСПР}(4,8; 5; 0,1; \text{ИСТИНА})$$

Імовірність події, що вага пакету лежить в межах від 4,8 кг до 5,1 кг

D3 (13) $=\text{НОРМРАСПР}(5,1;5;0,1;1)-\text{НОРМРАСПР}(4,8;5;0,1;1)$					
	A	B	C	D	E
1		задача а)	задача б)	задача в)	
2	Вага пакунку менше 4,8	більше 5,1	у межах від 4,8 до 5,1		
3	Імовірність	0,023	0,159	(15) 0,819	

Мал. 165

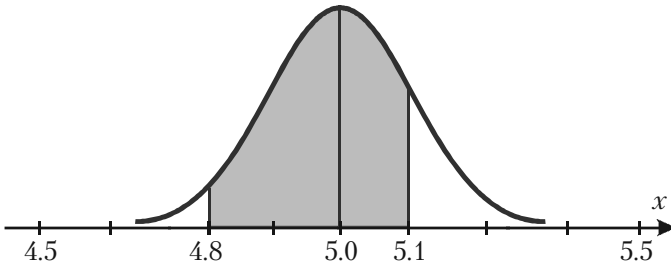
14 < Enter >.

15 Результат

$$P(4,8 < X < 5,1) = P(X < 5,1) - P(X < 4,8) = 0,8414 - 0,0228 = 0,8186$$

заноситься у вічко D3.

Геометрично ця імовірність дорівнює площі, що вказана на мал. 166.



Мал. 166

Приклад 4. У деякій фірмі N 40% працюючих складають жінки. Випадковим чином відібрано 200 службовців фірми (вважаємо, що виконуються умови повторних випробувань схеми Бернуллі). Визначити ймовірність того, що серед вибраних співробітників

- буде рівно 90 жінок;
- число жінок буде знаходитись в межах від 70 до 90.

Розв'язання.

За умовою задачі випадкова величина – кількість жінок серед 200 вибраних співробітників, має біноміальний розподіл, де

«число_испытаний» = 200, «вероятность_успеха» = 0,4. При $n > 100$ і $npq > 20$ біноміальний розподіл наближено замінюється нормальним з такими параметрами

$$\text{среднее} = a = np = 80,$$

$$\text{стандартное_отклонение} = \sigma = \sqrt{npq} = \sqrt{40} \approx 6,93.$$

Задача а)

Для обчислення числа $m = 90$ успіхів у $n = 200$ випробуваннях використовуємо локальну теорему Муавра-Лапласа

Розділ 8. Розв'язання задач засобами Excel

$$P_{200}(90) \approx \frac{1}{\sigma} \varphi(z_m), \tag{7}$$

де $z_m = \frac{m-a}{\sigma}$, $\varphi(z_m)$ – локальна функція Лапласа.

Легко перевірити, що права частина формули (7) співпадає з функцією щільності розподілу ймовірності нормального розподілу (5). Тому

$$P_{200}(90) \approx f(x, a, \sigma) = f(90, 80, 6,93) =$$

=НОРМРАСП(90,80,6,93,ЛОЖЬ).

- 1) Складемо таблицю (див. мал. 167).

	А	В	С	Д	Е
1		задача а)	задача б)		
2	Кількість жінок	рівно 90	у межах від 70 до 80		
3	Імовірність				

Мал. 167

- 2) Активізуємо вічко В3.
 3) Набираємо функцію (див. мал. 168)
=НОРМРАСП(90;80;6,93;ЛОЖЬ).

- 4) < Enter >.

Імовірність події, що з'явиться 90 успіхів у 200 випробуваннях

	А	В	С	Д	Е
1		задача а)	задача б)		
2	Кількість жінок	рівно 90	у межах від 70 до 80		
3	Імовірність	0,020			

Мал. 168

5) Результат $P_{200}(90) = 0,0203$ заноситься у вічко В3.

Задача б)

Згідно інтегральної теореми Муавра-Лапласа імовірність $P_{200}(70 \leq m \leq 90)$ події, що число успіхів знаходиться у межах від 70 до 90, знаходимо з допомогою інтегральної функції нормального розподілу (6)

$$P_{200}(70 \leq m \leq 90) = f(90, a, \sigma) - F(70, a, \sigma) =$$

$$= \text{НОРМРАСПР}(90; 80; 6,93; \text{ИСТИНА}) -$$

$$- \text{НОРМРАСПР}(70; 80; 6,93; \text{ИСТИНА}).$$

6) Активізуємо вічко С3.

7) Набираємо функцію (див. мал. 169)

$$= \text{НОРМРАСПР}(90; 80; 6,93; \text{ИСТИНА}) -$$

$$- \text{НОРМРАСПР}(70; 80; 6,93; \text{ИСТИНА}).$$

Імовірність події, що число успіхів знаходиться у межах від 70 до 90

	А	В	С	Д	Е	F
1		задача а) задача б)				
2	Кількість жінок	рівно 90	у межах від 70 до 80			
3	Імовірність	0,020	0,851			

Мал. 169

8) < Enter >.

9) Результат

$$P_{200}(70 \leq m \leq 90) = 0,8510$$

заноситься у вічко С3.

8.8. Побудова довірчого інтервалу для оцінки математичного сподівання

Теоретичні положення: пункт 6.3.2 параграфу 6.3, розділ 6.

Нехай маємо вибірку з генеральної сукупності, що розподілена за нормальним законом і середньоквадратичне відхилення якої відоме.

Довірчий інтервал для невідомого середнього μ (математичного сподівання) генеральної сукупності — це інтервал з обох сторін від вибіркової середньої \bar{x}_B

$$\bar{x}_B \pm \delta, \quad (1)$$

δ — точність оцінки, яка вибирається таким чином, щоб гарантувати попадання μ в інтервал (1) з заданою імовірністю (надійністю) γ .

Нижче точність оцінки δ знаходиться за допомогою функції **ДОВЕРИТ**. Ця функція має наступний формат:

ДОВЕРИТ (альфа; станд_откл; размер),

«альфа» — рівень значущості;

«альфа» = $1 - \gamma$;

«станд_откл» — це стандартне відхилення генеральної сукупності, яке вважаємо відомим;

«размер» — об'єм вибірки.

Приклад 5. Є вибірка із 100 рахунків компанії *N*. Випадкова величина — розмір суми рахунку, розподілена за нормальним законом. Середнє значення величини суми рахунку складає 45,5 грн, стандартне відхилення генеральної сукупності дорівнює 3,24 грн.

Визначити довірчий інтервал для середнього значення μ всіх рахунків з заданою надійністю $\gamma = 0,95$.

Розв'язання. В умовах нашої задачі маємо

«альфа» = $1 - \gamma = 0,05$,

«станд_откл» = 3,24,

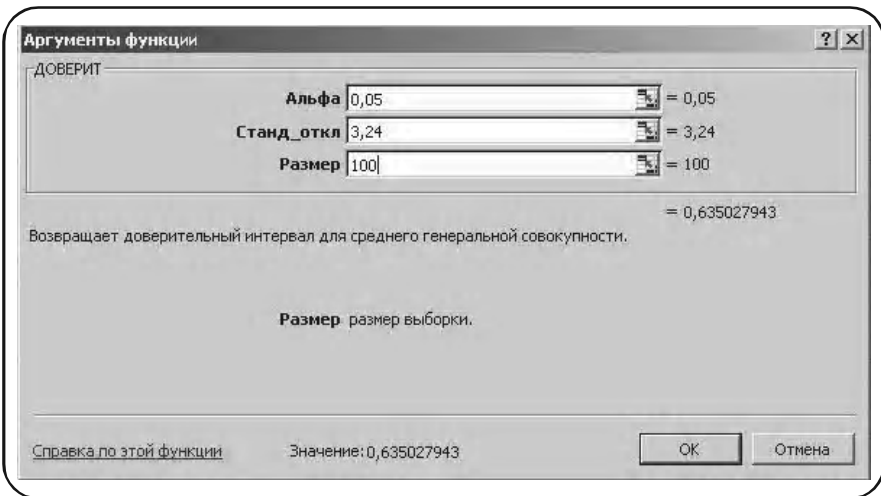
«размер» = 100.

- 1 Складемо таблицю з заголовком «Точність оцінки» (див. мал. 170).

	A	B
1	Точність оцінки	
2		

Мал. 170

- 2 Активізуємо вічко A2.
- 3 Вставка → Функція → Статистическая → ДОВЕРИТ.
- 4 У вікні меню функції **ДОВЕРИТ**, що з'явилося на екрані, набираємо числові значення аргументів функції, як показано на мал. 171.



Мал. 171

- 5 Результат – точність оцінки $\delta = 0,6350$, читаємо у вікні меню функції **ДОВЕРИТ**.
- 6 ОК.

Розділ 8. Розв'язання задач засобами Excel

- 7 Результат – точність оцінки $\delta = 0,6350$, заноситься у вічко A2 (див. мал. 172).

	A	B	C	D	E
1	Точність оцінки				
2	0,6350				

Мал. 172

Приходимо до висновку: з надійністю 95% можна стверджувати, що середня сума рахунків компанії N знаходиться в межах

$$45,50 \pm 0,635 \text{ грн,}$$

або

від 44,865 до 46,135 грн.

8.9. Алгоритми Excel корисні в статистичних обчисленнях


8.9.1. Введення даних у таблиці Excel

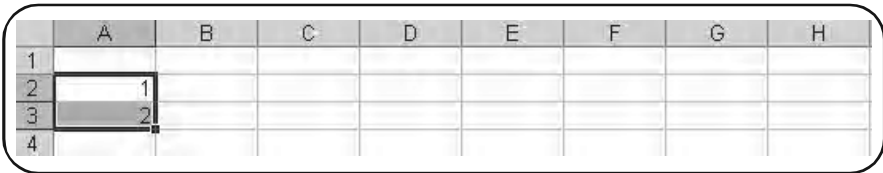
Першим кроком у роботі з електронною таблицею Excel є введення даних.

Приклад 1. Записати дані Таблиці А (стор. 176) числа балів, набраних при тестуванні 180 співробітників фірми *N*, у вигляді таблиці в Excel.

Розв'язання. Всі дані введемо у вигляді двох стовпців.



1. Формування першого стовпця

1.  A2 (встановлюємо мишу на клітину A2 та активізуємо її) і набираємо цифру 1 (див. мал. 173).



	A	B	C	D	E	F	G	H
1								
2	1							
3	2							
4								

Мал. 173

2.  **Стрілка додолу** > (переходимо у вічко A3) та набираємо цифру 2.
3.  A2:A3 (виділення чисел A2 і A3).
В результаті у вікні з'явиться прямокутник із квадратиком у правому нижньому куту.
4. Ставимо курсор на цей квадрат. Курсор у вигляді стрілки змінюється на курсор у вигляді хрестика.
Натискаємо ліву клавішу миші і, утримуючи її в нажатому положенні, переміщуємо її вниз до тих пір, доки не досягнемо 180 рядка. Номер рядка автоматично висвічується праворуч при переміщенні курсору (мал. 174).

	A	B	C	D	E	F	G	H
1	Номер							
2	1							
3	2							
4								
5								
6								
7								

Мал. 174

Після цього потрібно відпустити ліву клавішу миші. Перший стовпчик заповниться послідовністю цілих чисел, починаючи з 1 і кінчаючи 180 (мал. 177).

Результат формування першого стовпця

	A	B	C	D	E	F	G	H
1								
2	1							
3	2							
178	177							
179	178							
180	179							
181	180							

Мал. 175

2. Формування другого стовпця

- 1) B2 (Виділення і активізація клітини B2). Набираємо у вічку B2 перше число 30 Таблиці А (див. мал. 176).
- 2) < Стрілка додолу >.
- 3) Набираємо у вічку B3 друге число 38 Таблиці А.

На мал. 178 показаний результат набору перших п'яти чисел.

Після занесення 175 чисел, що залишалися, отримуємо Таблицю А у вікні електронної таблиці Excel (див. мал. 177).

Результат формування другого стовпця

	A	B	C	D	E	F	G	H
1								
2	1	30						
3	2	38						
4	3	32						
5	4	49						
6	5	33						

Мал. 176

Таблиця А у вікні Excel

	A	B	C	D	E	F	G	H
1	Номер	Бали						
2	1	30						
3	2	38						
4	3	32						
5	4	49						
6	5	33						
7	6	42						
8	7	60						
9	8	30						
180	179	38						
181	180	29						

Мал. 177

Озаглаavimo колонки таблиці.

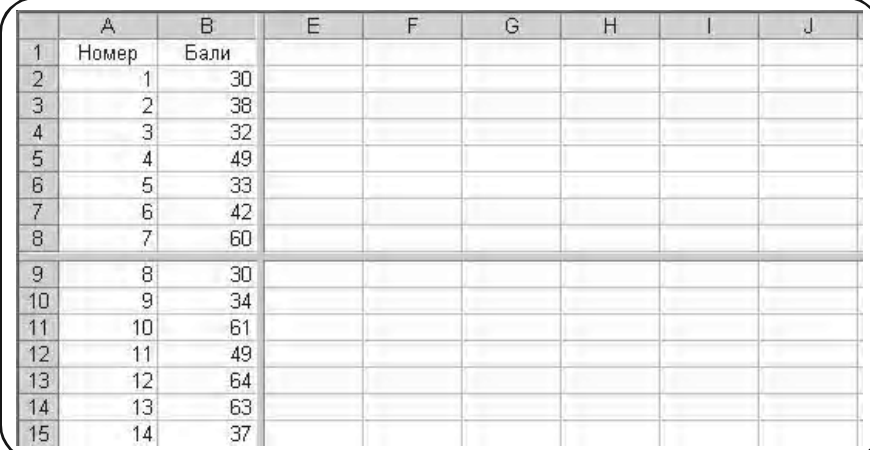
- 6 ↻ 1 A1. Набираємо текст «Номер».
- 7 ↻ 1 B2. Набираємо текст «Бали».

3. Розподіл екрану на два вікна

Складена таблиця містить 180 рядків і незручна для огляду. З метою поліпшення огляду поділимо поле таблиці Excel на два вікна.

1 Окно → Разделить.

Робоче поле електронної таблиці поділиться на чотири частини, як показано на мал 178.



	A	B	E	F	G	H	I	J
1	Номер	Бали						
2	1	30						
3	2	38						
4	3	32						
5	4	49						
6	5	33						
7	6	42						
8	7	60						
9	8	30						
10	9	34						
11	10	61						
12	11	49						
13	12	64						
14	13	63						
15	14	37						

Мал. 178

2 Ставимо курсор миші на вертикальну лінію. Курсор у вигляді стрілки змінює свою форму.

3 ↗ , пересуваємо мишу до правого краю вікна. При цьому пересувається і вертикальна розділова лінія (мал. 179).

Коли миша досягне правого краю вікна, ця розділова лінія зникне. Користуючись клавішею вертикальної прокрутки, встановимо в другому вікні нижню частину нашого ряду (мал. 180).

4. Зняття розподілу вікон

1 Окно → Снять разделение.

	A	B	C	D	E	F	G	H
1	Номер	Бали						
2	1	30						
3	2	38						
4	3	32						
5	4	49						
6	5	33						
7	6	42						
8	7	60						
9	8	30						
10	9	34						
11	10	61						
12	11	49						
13	12	64						
14	13	63						
15	14	37						

Мал. 179

Таблиця А у вікні Excel, що розділено на дві частини: ми бачимо водночас початок і кінець таблиці

	A	B	C	D	E	F	G
1	Номер	Бали					
2	1	30					
3	2	38					
4	3	32					
5	4	49					
6	5	33					
7	6	42					
8	7	60					
176	175	33					
177	176	58					
178	177	58					
179	178	31					
180	179	38					
181	180	29					

Мал. 180

8.9.2. Редагування графіків за допомогою майстра діаграм

Розглянемо докладно, як редагується графік за допомогою функції «Мастер діаграм».

Приклад 2. Побудувати полігон частот для згрупованого розподілу вибірки платні, наведеної в Таблиці В.3-Е.0 на мал. 181.

	А	В
1	Таблиця В-3-Е.0	
2	Платня	Частоти
3	280-289	1
4	290-299	10
5	300-309	14
6	310-319	14
7	320-329	25
8	330-339	16
9	340-349	7
10	350-359	4
11	360-369	7
12	370-379	0
13	380-389	2
14		100

Мал. 181

Розв'язання.

Побудову графіка робимо за допомогою «Майстра діаграм» Excel. Для цього переробимо Таблицю В.3-Е.0 наступним чином. У відповідності з алгоритмом побудови полігону частот, значення частоти відкладається у точці, що відповідає середині інтервалу, на висоті, що відповідає значенню частоти (див. пункт 5.8.2 параграфу 5.8, розділ 5).

1. Формування таблиці даних для «Майстра діаграм»

① Створюємо Таблицю В.3-Е.1, показану на мал. 182.

У перший стовпчик занесемо значення середини інтервалів, у другий – частоти. Введемо нульові значення частоти на кінцях інтервалів.

	А	В
1	Таблиця В-3-Е.1	
2	Платня	Частоти
3	275	1
4	285	10
5	295	14
6	305	14
7	315	25
8	325	16
9	335	7
10	345	4
11	355	7
12	365	0
13	375	2

Мал. 182

При побудові Таблиці В.3-Е. 1 заголовок та стовпчик частот були скопійовані. В перший стовпчик занесені серединні значення інтервалів.

2. Побудова полігона частот

Виділимо числові дані Таблиці В.3-Е.1 (див. мал. 182).

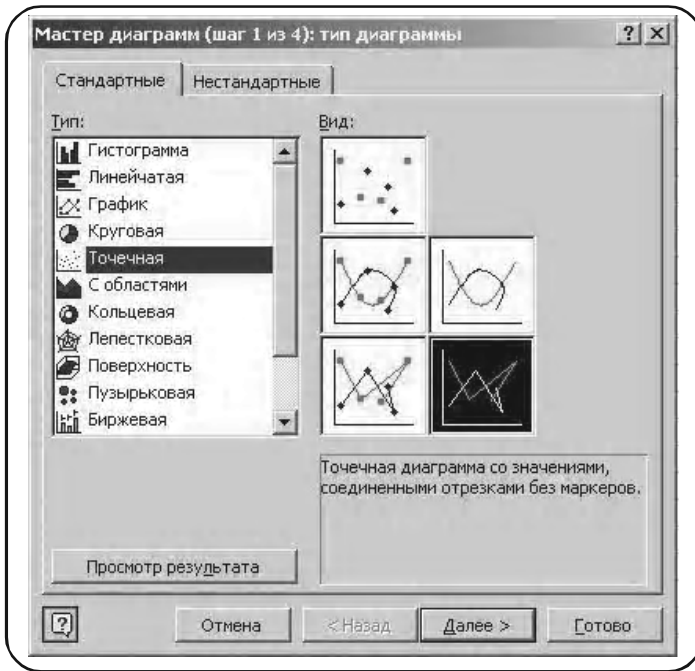
① ↗ А3:В15

② Вставка → Диаграмма.

У вікні з'явиться меню майстра діаграм (див. мал. 183). Для наших цілей у меню «Тип» слід вибрати опцію «Точечная», а серед усіх діаграм цього типу – кускову лінійну (вона виділена чорним кольором у вікні «Вид»).

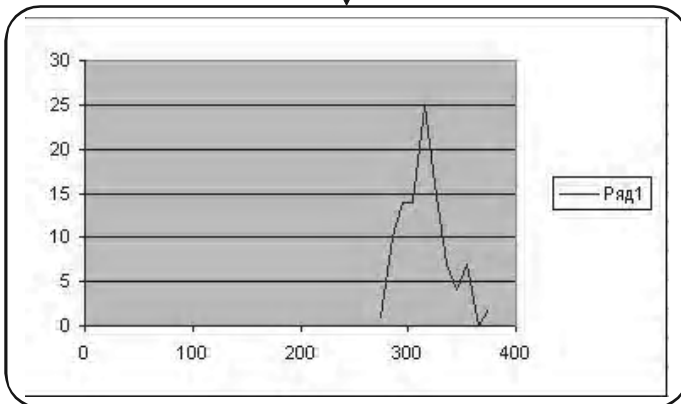
③ Натискуємо кнопку «Готово» на меню майстра діаграм.

У вікні з'явився полігон частот, відповідний Таблиці В.3-Е.1 (див. мал. 184).



Мал. 183

Первинний вигляд полігону частот, що потребує редагування



Мал. 184

3. Редагування графіка

Наведений на мал. 186 полігон частот сформований в Excel у відповідності з тими установками, що існували в ньому на час виведення графіка. Тому він потребує редагування.

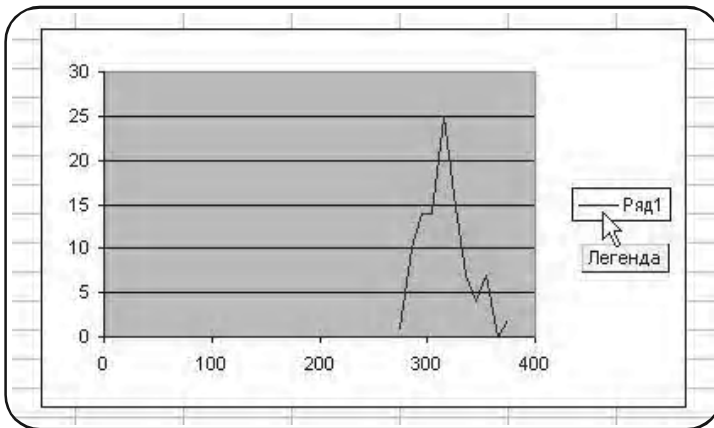
Майстер діаграм представляє дуже широкі можливості для редагування. При мінімальних зусиллях і фантазії, користувач може отримати графік високої технічної та естетичної якості.

Тут зупинимося на мінімальних, що вимагаються від нас, прийомах перетворення графіка.


1. Вилучення легенди

Так називається надпис, «Ряд 1», розміщений праворуч від діаграми.

Ставимо мишку на надпис «Ряд 1». З'являється рамка з надписом «Легенда» (див. мал. 185).

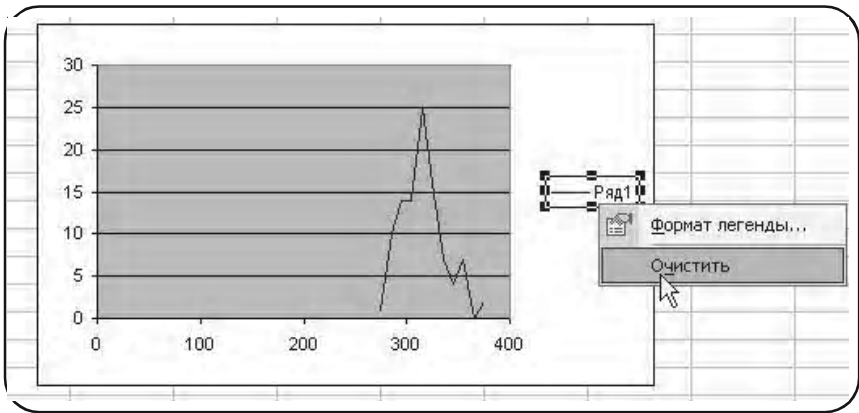


Мал. 185

Натискуємо праву клавішу мишки. На меню, що з'явилося (див. мал. 186), вибираємо опцію «Очистить»,  1 .

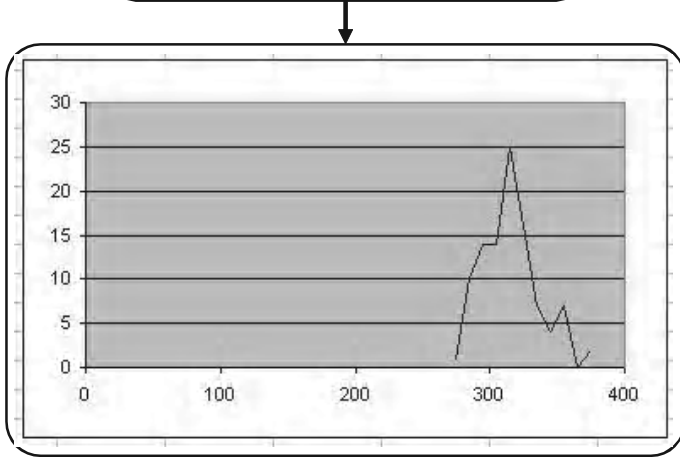
В результаті діаграма набуває вигляду (див. мал. 187).

З мал. 187 видно, що вимагається редагування шкали осі x -ів та осі y -ів.



Мал. 186

Графік полігону частот після
усунення легенди

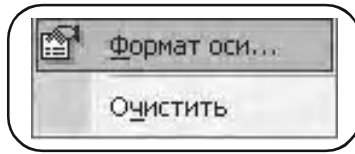


Мал. 187

2. Редагування надписів по осі x та осі y

Ставимо мишу на надписи у нижній шкалі та натискаємо праву клавішу, (**↵** 1).

З'явиться вікно меню редагування осі x (див. мал. 188).



Мал. 188

Ставимо курсор миші на «Формат осі», \rightarrow 1 .

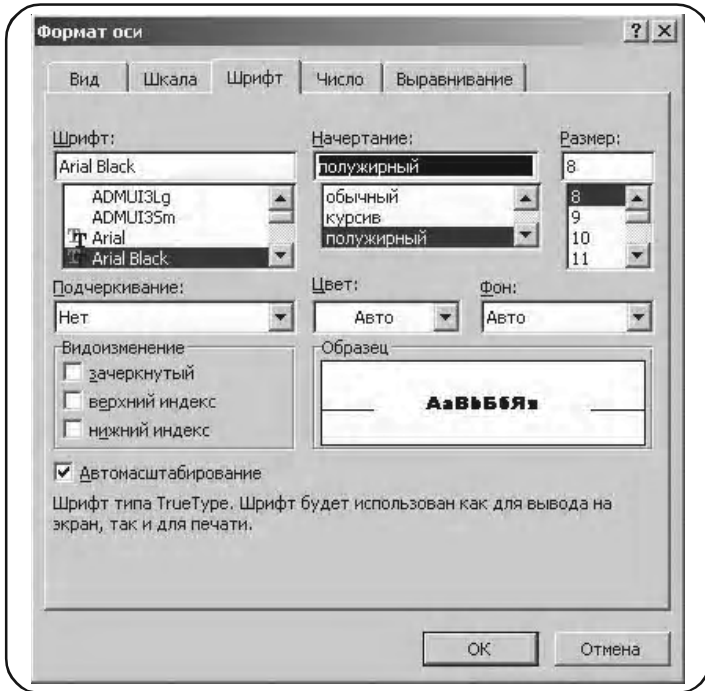
На екрані з'являється меню «Формат осі» (див. мал. 189).



Мал. 189

Вибираємо опцію «Шкала» і вносимо відповідні значення цифрових даних з Таблиці В.3-Е.0 (мал. 181), як показано на мал. 189.

Ставимо курсор мишки на опцію «Шрифт» і вносимо відповідні параметри, як показано на мал. 190.

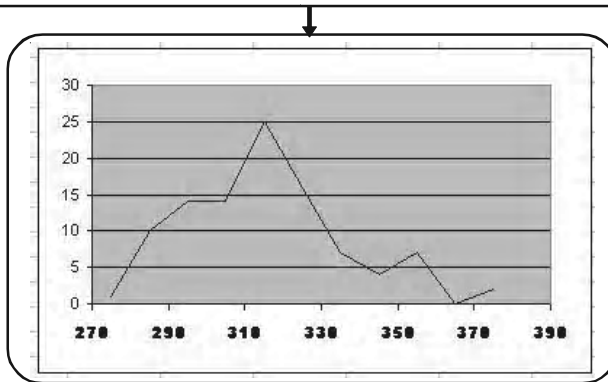


Мал. 190

□ OK

Діаграма набуває вигляду (див. мал. 191).

Графік полігону частот після редагування осі x-ів

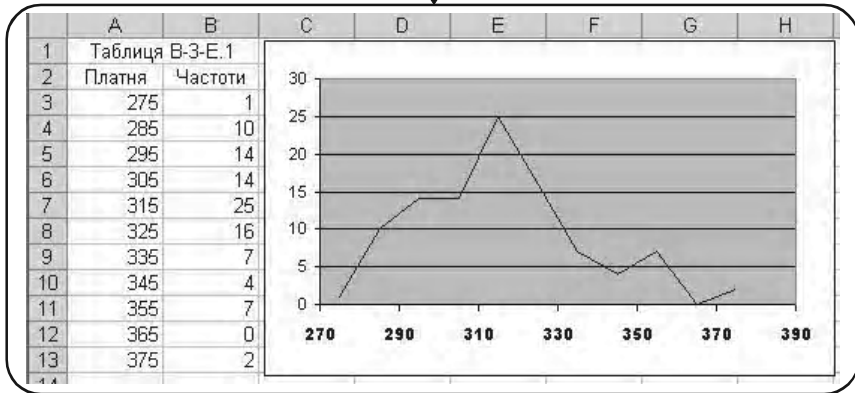


Мал. 191

Аналогічно редагуємо вісь y -ків.

Остаточний результат редагування осей x та y показано на мал. 192.

Графік полігону частот після редагування осей x -ів та осей y -ів



Мал. 192

3. Зміна розмірів діаграми

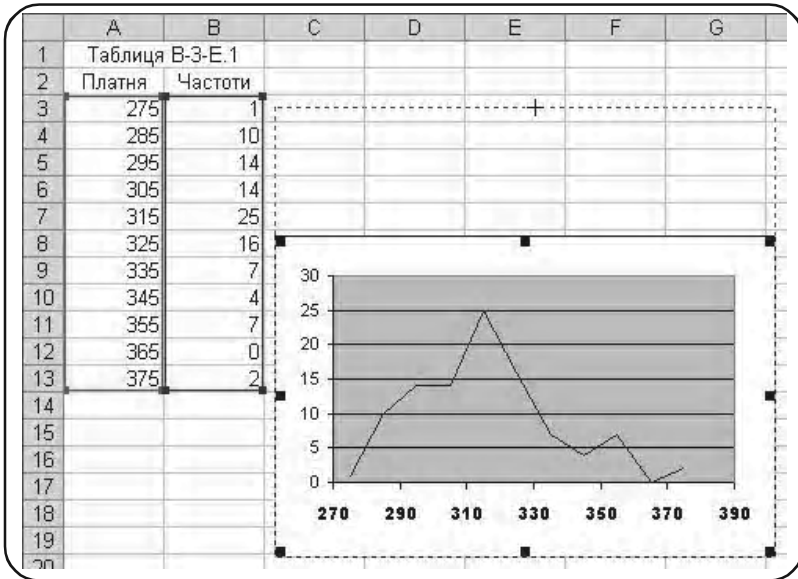
Збільшимо масштаб діаграми у напрямку осей y . Для цього ставимо курсор на середньому квадраті верхньої кромки малюнка. Курсор змінює форму на подвійну стрілку. Натискаємо ліву клавішу мишки і, не відпускаючи її, розтягуємо малюнок до потрібної величини у вертикальному напрямку.

Контур майбутньої діаграми позначений пунктирною лінією (див. мал. 193).

Відпускаємо ліву кнопку мишки. Діаграма набуває вигляду (див. мал. 194).

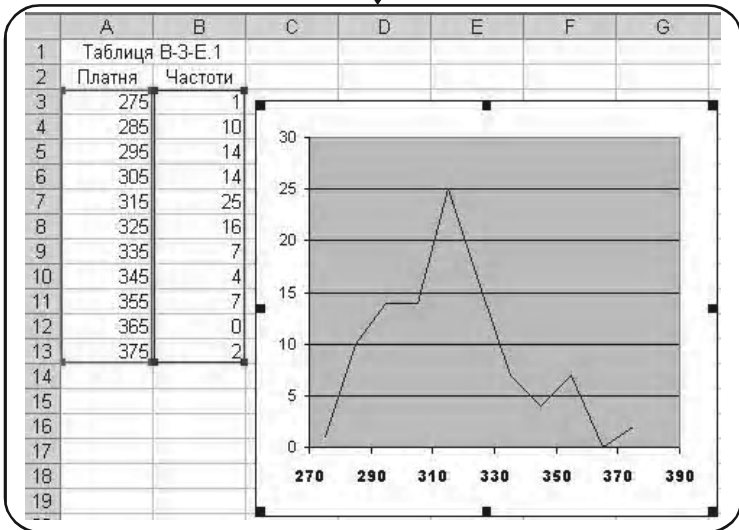
Аналогічно змінюється горизонтальний розмір діаграми. При цьому використовуються середні квадратики на бокових сторонах обмежуючої рамки.

Для зміни діаграми по діагоналі використовуються кутові квадратики обмежуючої рамки.



Мал. 193

Зміна розмірів графіку полігону частот у напрямку осі y-ів



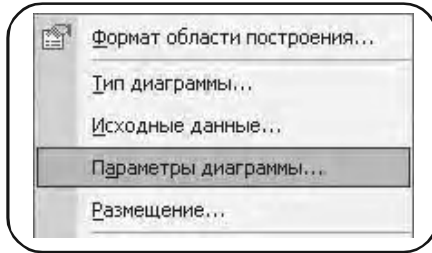
Мал. 194

4. Введення заголовків осей

Розмістимо курсор між зовнішньою обмежуючою рамкою і полем, на якому зображено графік.

З'явиться надпис у рамці «Область діаграми».

Натискуємо праву клавішу мишки. З'являється меню (мал. 195).

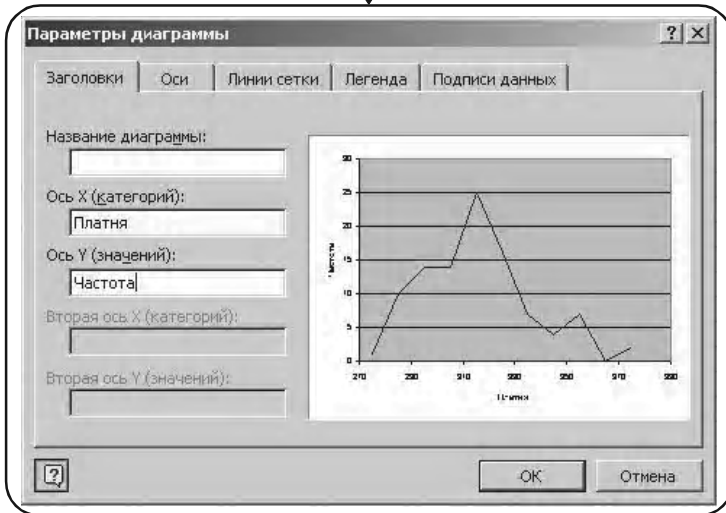


Мал. 195

Вибираємо опцію «Параметри діаграми» (на малюнку показана верхня частина меню).

Після натиску лівої кнопки мишки з'являється меню «Параметри діаграми» (див. мал. 196).

Введення надписів для осі x-ів та осі y-ів



Мал. 196

Розділ 8. Розв'язання задач засобами Excel

Вибираємо опцію «Заголовки» і вносимо назви осей, як показано на малюнку.

ОК.

Отримуємо полігон частот (див. мал. 197).



Мал. 197

5. Зміна товщини лінії графіка

Встановлюємо курсор миші на лінію графіка у будь-якій точці. З'явиться надпис (див. мал. 198).

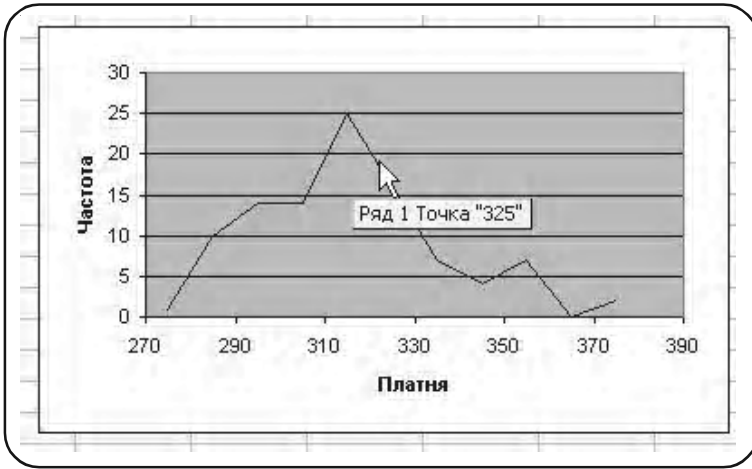
«Ряд 1».

Натискуємо ліву клавішу миші, і на графіку з'являться маркери (див. мал. 199). Натискаємо праву клавішу миші, і з'явиться меню, у якому вибираємо опцію «Формат рядов даних».

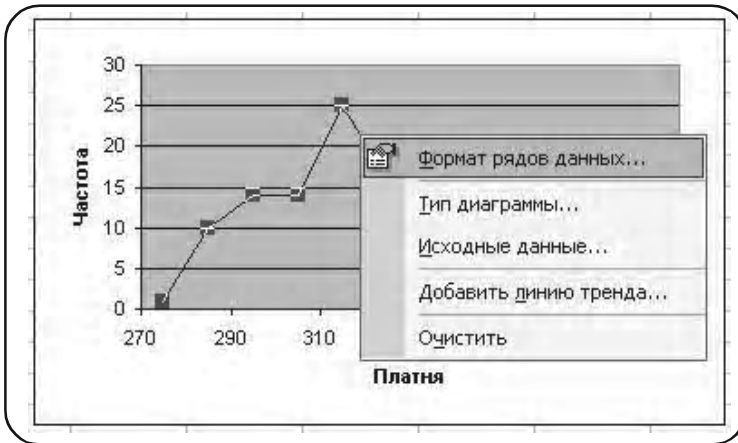
➤ 1: на екрані з'являється вікно (див. мал. 200).

Товщину лінії вибираємо, як показано на малюнку,

ОК.



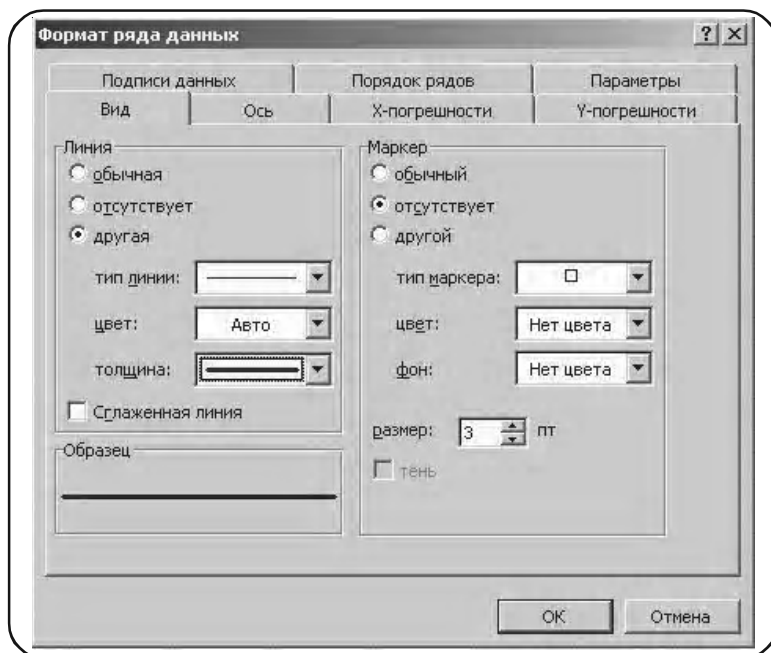
Мал. 198



Мал. 199

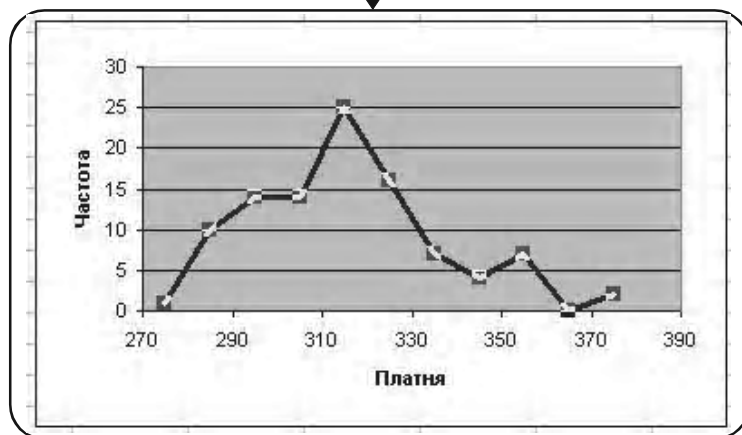
На екрані з'явиться зображення (див. мал. 201).

На ньому зображені присутні маркери у вигляді білих прямокутників.



Мал. 200

Результат зміни товщини лінії графіку полігону частот

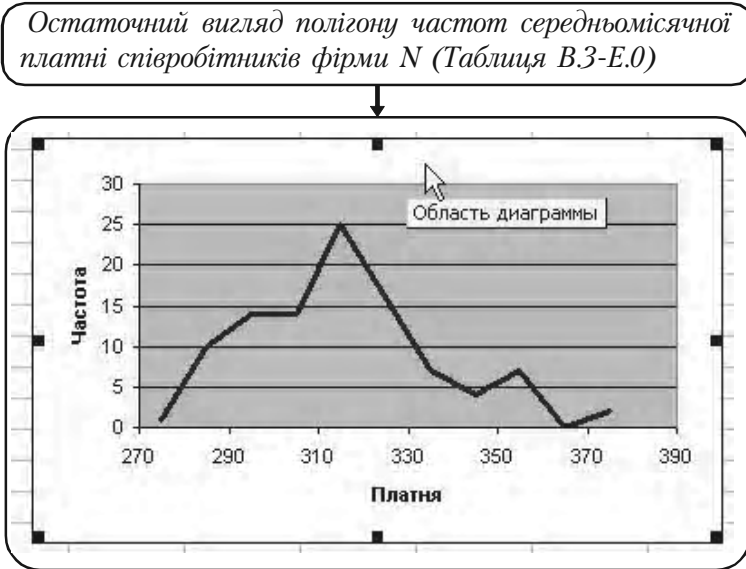


Мал. 201

6. Вилучення маркерів

Для знищення маркерів необхідно встановити курсор між зовнішньою рамкою малюнка і полем, на якому зображений графік. З'явиться надпис у рамці «Область діаграми». Натиснути ліву клавішу мишки.

Остаточний результат: полігон частот середньомісячної платні співробітників фірми N (Таблиця В.3-Е.0) показано на мал. 202.



Мал. 202

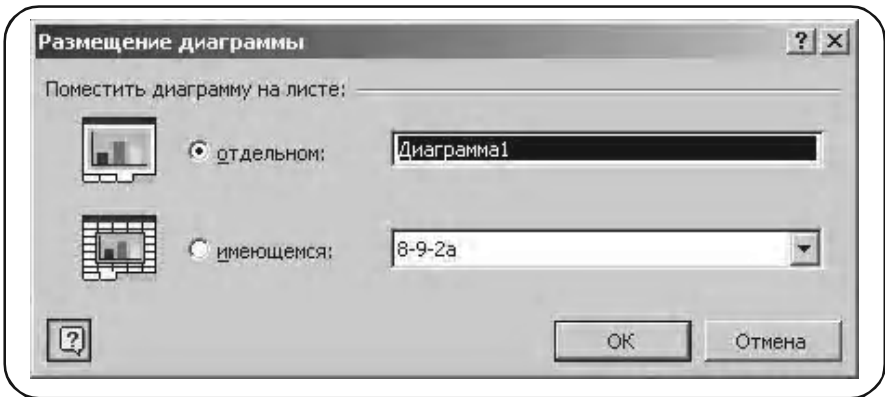
4. Перенос діаграми на окремий лист

Для подальшого роздрукування діаграми нам зручніше винести діаграму на окремий лист.

- 1 Ставимо курсор в область діаграми, натискаємо праву мишу і в меню, що з'явилося, вибираємо опцію «Размещение...».

З'явиться меню «Размещение диаграммы» (див. мал. 203).

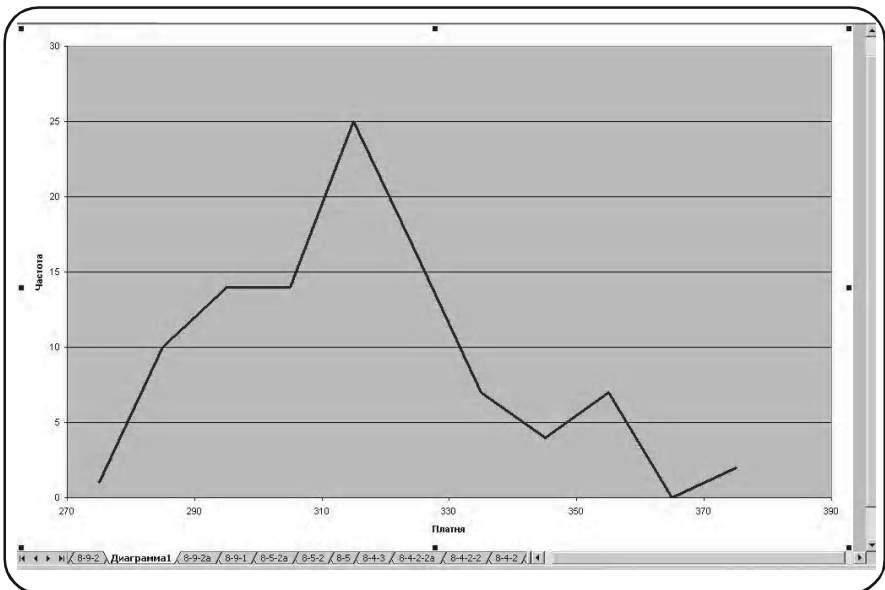
- 2 У цьому меню вибираємо опцію «отдельном», як показано на малюнку.



Мал. 203

ОК.

На екрані з'явиться зображення полігону частот на окремому листі під назвою «Диаграмма 1» (див. мал. 204).

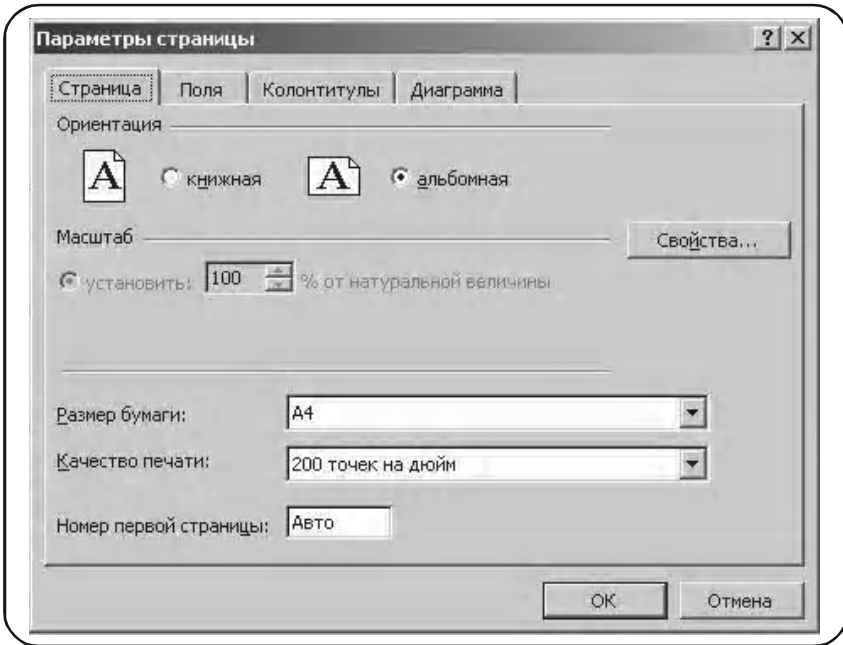


Мал. 204

5. Виведення діаграми до друку

1 Файл → Попередній перегляд.

З'являється меню «Параметри сторінки» (див. мал. 205).



Мал. 205

2 Вибираємо опцію «Сторінка».

У цьому меню вибирається формат паперу, розташування діаграми відносно листа та інші параметри для роздрукування.

3 ОК.

4 < Print >.

8.9.3. Виведення цифрових даних із заданою точністю

Часто вимагається вивести цифрові дані результатів підрахунків у вічко з заданим числом десяткових знаків. Нехай, наприклад, вимагається знайти вибіркове середньо-квадратичне відхилення вибірки, що складається з чисел 2, 4, 8, 6 із двома цифрами після коми (див. Таблицю 1 на мал. 206).

	А	В
1	Таблиця 1	
2		2
3		4
4		8
5		6
6		
7	Вибіркове _середньо _квадрати чне_відхіл ення	1

Мал. 206

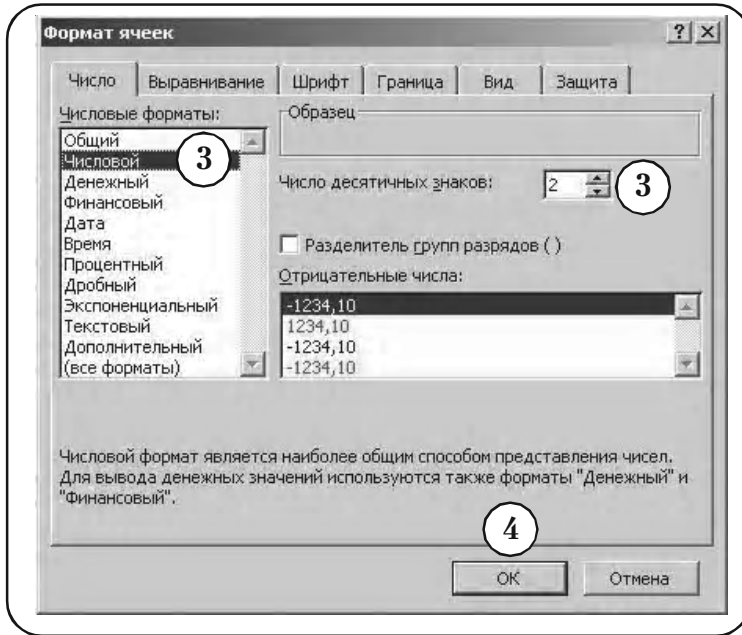
Перші кроки алгоритму забезпечують необхідний формат чисел.

- ① Активізуємо вічко B7.
- ② **Формат** → **Ячейки**.
- ③ В меню «Формат ячеек», що з'явилося, необхідно вибрати опцію «Числової». У вікні «Число десятичних знаків» поставити 2 (див. мал. 207).
- ④ **ОК**.

Подальші кроки алгоритму забезпечують необхідні обчислення.

- ⑤ Набираємо функцію

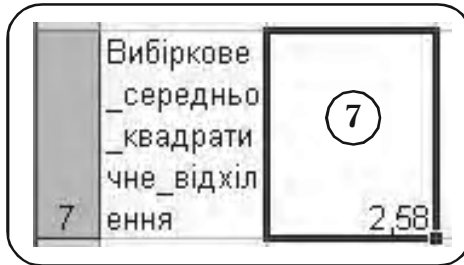
$$=СТАНДОТКЛН(В1:Е1).$$



Мал. 207

6 < Enter >.

7 Остаточний результат – вибіркоче середньоквадратичне відхилення $\sigma_B = 2,58$, заноситься у вічко B7.



Мал. 208

Пропонуємо читачу провести для порівняння ті ж обчислення, але без попереднього формування вічка B7.

8.10. Висновки

1. Побудована проста випадкова вибірка із скінченної сукупності даних.

Використання випадкових чисел при утворенні вибірки дозволяє уникнути систематичних помилок і робить вибірку репрезентативною.

Алгоритм утворення простої випадкової вибірки гарантує, що кожний елемент вихідної сукупності даних, з яких утворюється вибірка, має рівний шанс бути включеним у вибірку.

2. Зроблено впорядкування і організацію даних.

Це сприяє кращому розумінню залежностей, що описуються цими даними.

Впорядкування даних дозволяє краще зрозуміти характер їхнього зростання і межі зміни.

Наочну інформацію про характер зміни даних дають простий (незгрупований) розподіл частот вибірки і згрупований розподіл частот вибірки.

Згрупований розподіл частот дає більш наочну інформацію про вибірку, але він не відбиває індивідуальних значень даних.

Накопичений розподіл частот несе в собі інформацію про те, яка кількість даних лежить нижче або вище деякої довільно встановленої границі.

3. Наступний етап в обробці даних полягає в їхньому графічному представленні. Графічне представлення більш наочно і не вимагає аналізу даних у вигляді таблиць.

Гістограма і полігон частот є основними формами графічного представлення згуртованого розподілу частот вибірки. Полігон частот і частостей дають близьке уявлення про форму невідомої функції щільності імовірності даних з генеральної сукупності.

Полігон накопиченої частоти (відносної частоти) і емпірична функції розподілу дають близьке уявлення про форму невідомої теоретичної функції розподілу даних з генеральної сукупності.

Полігон накопиченої частоти (відносної частоти) використовується для графічного знаходження медіани згрупованого розподілу частот вибірки.

4. Числові характеристики розподілу частот вибірки (згруповані або ні) — вибіркове середнє, дисперсія і середнє квадратичне відхилення, що називаються статистиками, використовуються як наближені значення невідомих числових характеристик генеральної сукупності — математичного сподівання, дисперсії і середнє-квадратичного відхилення, що називаються параметрами. Такі оцінки називаються точковими.

Названі числові характеристики розподілу частот вибірки також можуть бути використані в інтервальних оцінках невідомих параметрів як наближені значення.

5. Реалізовані формули підрахунку переставлень, сполучень і розміщень. Ці формули широко застосовуються для знаходження теоретичної імовірності.

6. Розглянутий біноміальний розподіл дискретної випадкової величини. Підрахунок імовірності $P(X)$ появи випадкової величини X можна здійснити для окремих значень, або для інтервалу значень від 0 до X .

Розподіл Пуассона є апроксимацією біноміального розподілу. Найважливішим у теорії імовірностей і математичній статистиці є нормальний розподіл неперервної випадкової величини.

Площа під кривою щільності розподілу імовірності, обмеженою двома вертикальними лініями, що проходять через точки A та B , дорівнює імовірності попадання випадкової величини в інтервал (A, B) .

Розглянуті функції

БИНОМРАСПР, ПУАССОН, НОРМРАСПР

можна використати в якості електронних аналогів відповідних статистичних таблиць випадкових величин.

7. Побудований довірчий інтервал для оцінки математичного сподівання генеральної сукупності, розподіленої по нормальному закону з відомим стандартним відхиленням σ .

Побудований довірчий інтервал є прикладом статистичних оцінок параметрів генеральної сукупності.

8. Проведені в даному розділі обчислення і графічні побудови засобами Excel показують широкі можливості застосування електронної таблиці Excel у задачах теорії імовірності та математичної статистики в бізнесі завдяки високому ступеню автоматизації обчислень і графічних побудов, точності і можливості обробки великого числа даних.

ДОДАТКИ

*Тепер вельми важливо
перевірити себе.*

*Чи засвоєно системно
в цій науці головне?*

Д.1. Завдання для контролю навичок розв'язування задач з теорії імовірностей

1.1. Контрольна робота 1

Варіант 1

Завдання 1. Розв'язати задачі.

а) Проведено 30 пострілів та зареєстровано 22 влучення у ціль. Знайти частоту влучення у ціль.

б) У точці C , положення якої на телефонній лінії AB довжиною 1 км рівноможливі, трапився розрив. Знайти імовірність того, що точка C віддалена від точки A на відстань

- не менше 400 метрів;
- не менше 300 метрів.

в) Монету кидають три рази. Яка імовірність того, що випаде 2 рази герб та один раз надпис?

Завдання 2. Для зменшення загальної кількості ігор $2N$ футбольні команди розподілені на дві підгрупи з однаковою кількістю команд. Скільки є способів розподілу команд, щоб дві найсильніші команди були:

- а) в різних підгрупах;
- б) в одній підгрупі?

Завдання 3. Серед студентів другого курсу вищого навчального закладу є 45% тих, що палять, 40% тих, що інколи вживають алкоголь, 18% тих, що палять та іноді вживають алкоголь. Знайти імовірність того, що будь-який студент цього курсу:

- а) не палить, але іноді вживає алкоголь;
- б) палить, але ніколи не вживає алкоголь;
- в) не палить та ніколи не вживає алкоголь;
- г) або лише палить або лише іноді вживає алкоголь.

Завдання 4. Два станки виготовляють деталі, які поступають на загальний конвеєр. Імовірність виготовлення нестандартної деталі першим станком дорівнює 0,075, а другим — 0,09. Другий станок виготовив деталей в два рази більше, ніж перший. Знайти імовірність того, що навмання взята з конвеєра деталь буде нестандартною.

Варіант 2

Завдання 1. Розв'язати задачі.

а) Відділом технічного контролю виявлено 8 бракованих виробів серед 160 перевірених. Знайти частоту появи бракованого виробу.

б) У прямокутник $\{0 \leq x \leq 2; 0 \leq y \leq 1\}$ навмання ставлять крапку. Знайти імовірність того, що ця крапка лежить вище кривої

$$y = \frac{x^2}{4}.$$

в) Кидають два гральних кубика. Знайти імовірність того, що сума очок на гранях цих кубиків дорівнює 7.

Завдання 2. Правління підприємства складається з 8 осіб. Скільки можна скласти варіантів обрання з них трьох керівників — президента, директора та комерційного директора?

Завдання 3. Серед студентів другого курсу вищого навчального закладу є 50% тих, що одержують на екзаменах оцінки «добре», 20% тих, що одержують «відмінно», 10% тих, що одержують «добре» та «відмінно». Знайти імовірність того, що будь-який студент цього курсу:

- а) не одержує 5, а одержує 4;
- б) одержує 5 і не одержує 4;
- в) не одержує 4 та не одержує 5;
- г) одержує або лише 4 або лише 5.

Д.1. Завдання для контролю навичок розв'язування задач з теорії імовірності

Завдання 4. В окремих цехах взуттєвої фабрики виготовляють підметки, підбори та верхи чобітків. Дефектними виявляються 0,5% підборів, 0,2% підметок та 4% верхів.

Виготовлені підбори, підметки та верхи випадковим чином комбінують у пошивочному цеху. Знайти імовірність того, що виготовлена пара:

- не має дефектів;
- має дефекти.

Варіант 3

Завдання 1. Розв'язати задачу.

а) В урні 15 кульок, серед яких 10 червоних. Навмання беруть три кульки. Знайти імовірність того, що усі вони будуть червоні.

б) Частота появи стандартних виробів дорівнює 0,8. Знайти кількість стандартних виробів серед перевірених 500.

в) На відрізок L довжиною 20 см намалювали відрізок l довжиною 6 см. Знайти імовірність того, що крапка, яка поставлена навмання на L , попаде також і на l .

Завдання 2. В навчальному плані семестру є 9 дисциплін. Кожен день можна включати до розкладу 3 різних дисципліни. Скількома способами можна розподілити учбові дисципліни на один день?

Завдання 3. Розповсюджено 1000 лотерейних білетів. Серед них два виграшу по 50 гривень, п'ять виграшів по 20 гривень, десять виграшів по 10 гривень та 25 виграшів по 5 гривень. Студент має один лотерейний білет. Яка імовірність

- виграшу не менше 20 гривень;
- будь-якого виграшу?

Завдання 4. Група спортсменів складається з 20 лижників, 6 велосипедистів та 4 бігунів. Імовірність виконання норми майстра спорту для лижника дорівнює 0,9, для велосипедиста — 0,8, для бігуна — 0,75.

Знайти імовірність того, що навмання обраний спортсмен виконає норму майстра спорту.

1.2. Контрольна робота 2

Варіант 1

Завдання 1. Технологічний процес підприємства дозволяє одержувати 90% виготовлених виробів найвищої якості. Знайти найімовірніше число виробів вищої якості серед 250 виробів, які виготовлені підприємством.

Завдання 2. Середня кількість замовлень, що надходять до комбінату побутового обслуговування за одну годину, дорівнює 5. Знайти імовірність того, що за три години надійде

- 2 замовлення;
- менше 2 замовлень;
- не менше 2 замовлень.

Завдання 3. Дискретна випадкова величина X може приймати лише два значення x_1 та x_2 , причому $x_1 < x_2$. Відомі імовірність $p_1 = 0,2$ появи можливого значення x_1 , математичне сподівання $M(X) = 3,8$ та дисперсія $D(X) = 0,16$. Знайти закон розподілу цієї випадкової величини.

Завдання 4. Випадкову величину X задано функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x}{4}, & 0 < x \leq 4; \\ 1, & x > 4. \end{cases}$$

Знайти щільність розподілу імовірностей, математичне сподівання та дисперсію цієї випадкової величини. Побудувати графіки $F(X)$ та щільності $f(x)$.

Завдання 5. Випадкова величина X розподілена за нормальним законом з математичним сподіванням $a = 8$ та середньоквадратичним відхиленням $\sigma = 1$. Знайти імовірність влучення цієї величини в інтервал $(2, 13)$.

Варіант 2

Завдання 1. При випуску деякої продукції третьій сорт складає в середньому 20%. Знайти імовірність того, що з п'яти навмання взятих виробів цієї продукції не менше, ніж 3 будуть третього сорту.

Завдання 2. Проростання пшеничного насіння становить 85%. Знайти імовірність того, що з 2000 посіяних насінин проросте від 1880 до 1920.

Завдання 3. В урні 4 золотих та 6 срібних кульок однакового розміру. Навмання беруть 3 кульки. Нехай X кількість взятих золотих кульок. Побудувати ряд розподілу величини X та знайти її математичне сподівання.

Завдання 4. Випадкову величину X задано функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 0; \\ x^3, & 0 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Знайти щільність розподілу імовірностей, математичне сподівання та дисперсію цієї випадкової величини. Побудувати графік щільності $f(x)$.

Завдання 5. Випадкова величина X розподілена рівномірно в інтервалі $(3, 8)$. Знайти її диференціальну та інтегральну функції розподілу, математичне сподівання, дисперсію та імовірність того, що $X \in (1, 4)$.

Варіант 3

Завдання 1. На складі є вироби двох сортів, причому виробів другого сорту в 1,5 рази більше, ніж виробів першого сорту. Знайти імовірність того, що серед трьох навмання взятих виробів хоча б один буде першого сорту.

Завдання 2. З кожних 100 виробів, виготовлених верстатом-автоматом, 5 є бракованими. Взято навмання 400 виробів. Знайти імовірність того, що серед них без браку буде від 360 до 375 виробів.

Завдання 3. Знайти дисперсію дискретної випадкової величини X — кількості появ події A в двох незалежних випробуваннях, якщо імовірність появи події в цих випробуваннях однакові і відомо, що $M(X) = 1,5$.

Завдання 4. Випадкову величину X задано функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 0; \\ x^2, & 0 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Знайти щільність розподілу імовірностей, математичне сподівання та дисперсію цієї випадкової величини. Побудувати графік $F(x)$.

Завдання 5. Випадкова величина X має показниковий розподіл з параметром $\lambda = 4$. Знайти її математичне сподівання, дисперсію та імовірність того, що $X \in (0,25; 2)$.

Д.2. Завдання для індивідуальної семестрової роботи (N – номер відповідного варіанту)

Завдання 1. Скількома способами можна скласти букет з $(12 + N)$ червоних та $(9 + N)$ білих троянд таким чином, щоб у ньому було 3 червоних та 2 білих троянди?

Завдання 2. До зборки надходять: $(40 + N)\%$ деталей з підприємства A ; $(35 - N)\%$ з підприємства B ; 25% – з підприємства C . Нестандартних деталей виготовляють: підприємство A – $0,2\%$; підприємство B – $0,3\%$; підприємство C – $0,5\%$. Знайти імовірність того, що взята навмання деталь для зборки буде нестандартною.

Завдання 3. Методом добутоків за даними вибірки знайти статистичні характеристики: вибірккову середню \bar{x}_B , вибірккову дисперсію D_B та середньоквадратичне відхилення (стандарт) σ_B .

x_k	$[N, 328]$	$[(N+3), 569]$	$[(N+6), 8]$	$[(N+10), 51]$
n_k	10	8	7	10

$[(N+13), 292]$	$[(N+16), 533]$	$[(N+19), 774]$	$[(N+23), 15]$
22	18	14	11

Завдання 4. Скласти таблицю основних законів розподілу неперервних випадкових величин та їх числових характеристик. За якою формулою знаходять $P(\alpha < X < \beta)$ для кожного з цих законів?

Завдання 5. Випадкова величина X генеральної сукупності в послідовності з 20 незалежних випробувань прийняла значення $N, 5; N, 7; (N + 1), 5; (N + 1), 6; (N + 1), 7; (N + 1), 9; N, 5; (N + 1), 6; N, 5; (N + 1), 6; (N + 1), 9; N, 7; (N + 1), 6; (N + 1), 9; N, 5; (N + 1), 9; N, 7; (N + 1), 7; (N + 1), 6; (N + 1), 5$.

Треба:

- скласти статистичний розподіл X ;
- обчислити емпіричне середнє значення \bar{x}_B ;
- знайти виправлене середньоквадратичне відхилення s .

Завдання 6.

Ознака X генеральної сукупності розподілена нормально з відомим середньоквадратичним відхиленням $\sigma = 2,4$ та невідомим математичним сподіванням a . За результатами вибірки об'єму $n = 100$ знайдена вибіркова середня $\bar{x}_B = N$. Визначити довірчий інтервал для оцінки параметра a з надійністю $\gamma = 0,9 + 0,01N$.

Завдання 7. За даними вибірки об'єму $n = 200 + 2N$ та відомими теоретичними частотами перевірити гіпотезу про нормальний розподіл ознаки X генеральної сукупності, використовуючи критерій Пірсона (χ^2) з рівнем значущості

$$\alpha = 0,05 - 0,001N.$$

Д.3. Тест для перевірки знань та навичок

Д.3.1. Методичні рекомендації до тесту

Контроль за тестом можна провадити з частин *Теорія імовірностей* та *Математична статистика* разом або окремо.

Для виконання тестових завдань з двох частин доцільно надати 3 академічні години.

Вибрати тестові завдання студент може самостійно або за вказівкою викладача, але завдання обов'язково вибирають із різних розділів.

Тестові завдання розподілено на розділи трьох рівнів:

- А. Репродукція або відтворення теоретичних знань;
- Б. Навички застосовувати знання;
- В. Застосування знань в дещо ускладнених ситуаціях (творче завдання).

Після виконання відповідних завдань та перевірки правильності розв'язання (відповіді до розділів Б та В подаються на стор. 270), студент зможе перевірити свої знання.

Спробуйте вибрати завдання, розв'язати їх та підрахувати набрані бали.

Завдання рівня репродукції (розділ А), що вимагають відтворення в пам'яті теоретичних знань, оцінюються 5 балами, рівня навичок застосовування знань (розділ Б) — 12 балами, творчі завдання (розділ В) — 20 балами. Неправильні розв'язання завдань оцінюються 0 балів.

Якщо у завданні подається декілька підзавдань (питань), і студент відповідає правильно лише на декілька з них, то бали до розрахунків треба перерахувати за такою схемою. Наприклад, розв'язується завдання першої групи (на репродукцію). За повне його розв'язання надається 5 балів. Студент з 3 завдань вирішив лише 2. Три вирішені завдання варті 5 балам. Рішення кожного завдання

дорівнює $\frac{5}{3}$ бала. Тоді вірна відповідь на 2 завдання дорівнює

$$\frac{5}{3} \cdot 2 \approx 3,3 \text{ бали.}$$

Вам треба виконати, наприклад, з першої групи А (на репродукцію) — 4 завдання, з другої групи Б (основні навички) — 5 завдань, з третьої групи (творчі завдання) — 1 завдання.

Якщо кількість завдань відповідної групи перемножити на відповідну кількість балів, якими оцінюються правильне виконання завдань, то можна визначити кількість набраних балів. Максимальна сума дорівнюватиме 100 балам $[4 \cdot 5 + 5 \cdot 12 + 1 \cdot 20]$.

Кількість N набраних балів відповідає таким оцінкам

«відмінно»: $90 \leq N \leq 100$,

«добре»: $76 \leq N \leq 89$,

«задовільнено»: $60 \leq N \leq 75$,

«незадовільнено»: $N < 60$.

У процесі виконання завдань фіксуйте тривалість у хвилинах.

Виконайте (у рамках запропонованого часу) декілька варіантів вибору запропонованих тестів завдань і виведіть середню оцінку.

Щоб уникнути прогалин у Ваших знаннях, повторіть ті розділи чи окремі теми, з яких запропоновані завдання Ви виконали неправильно. Потім виберіть тестові завдання з цих розділів чи тем і повторіть випробування.

Д3.2. Завдання тесту

Розділ А

1. Які події називають достовірними, неможливими, випадковими?
2. Як визначають та позначають суму, добуток випадкових подій, протилежну подію, повну групу подій?
3. Які події називають сумісними, несумісними, рівноможливими?
4. Які комбінації називають переставленням, розміщенням, сполученням? Як позначають та обчислюють кількість цих сполук?
5. Як визначають та в яких випадках використовують класичне та геометричне означення імовірності?
6. Які основні властивості імовірності та частоті?
7. Як формулюють і якими формулами записують теореми додавання імовірностей сумісних та несумісних подій?
8. Як формулюють і якими формулами записують теореми множення імовірностей залежних та незалежних випадкових подій?
9. За якою формулою можна обчислити імовірність появи хоча б однієї з n сумісних подій?
10. Яким умовам повинна задовольняти подія, щоб її імовірність можна було знаходити за формулою повної імовірності? Який вигляд має ця формула?
11. Яка послідовність випробувань утворює схему Бернуллі?
12. Яку формулу звать формулою Бернуллі і що вона дозволяє обчислювати?
13. Коли доцільно застосовувати формули Пуассона, локальну або інтегральну формули Муавра-Лапласа? Який вони мають вигляд?
14. Як формулюється теорема Бернуллі і який вона має наслідок?
15. Який вигляд має твірна функція і коли вона використовується?
16. Як визначають випадкові величини, дискретні та неперервні випадкові величини?
17. За якими формулами обчислюють числові характеристики дискретних випадкових величин? Який їх зміст?
18. Який існує зв'язок між інтегральною та диференціальною функціями розподілу імовірностей? Що характеризують ці функції?
19. За якими формулами обчислюють числові характеристики неперервних випадкових величин? Що вони характеризують?
20. Вказати основні задачі математичної статистики.

21. Що називають статистичною, генеральною та вибірковою сукупністю, об'ємом цих сукупностей?
22. Що називають статистичним розподілом вибірки?
23. Як визначають та позначають емпіричну функцію розподілу? Які основні властивості цієї функції?
24. Як визначають полігони частот та частостей?
25. Як визначають статистичні оцінки числових характеристик та умови їх незсунутості, ефективності, обґрунтованості?
26. Вказати числові характеристики вибірки та формули, за якими їх обчислюють.
27. В яких випадках використовують виправлену вибірку дисперсію і як вона пов'язана із вибірковою дисперсією?
28. В яких випадках обчислюють характеристики вибірки методом добутоків? Який порядок дій при використанні цього методу?
29. Які статистичні оцінки називають точковими, інтервальними?
30. Як визначають надійність статистичних оцінок та їх довірчий інтервал?
31. Вказати послідовність дій методу добутоків знаходження вибірових характеристик.
32. Які гіпотези називають статистичними, основною та альтернативною, простою та складною?
33. Що таке похибки першого та другого роду перевірки статистичної гіпотези?
34. Який смисл рівня значущості α та потужності критерію перевірки статистичної гіпотези?
35. Що називають статистичним критерієм, критичною областю та критичною точкою перевірки гіпотези?
36. Коли застосовують критерій узгодження Персона (χ^2 -кватрат)?
37. Як формулюють правило Пірсона?
38. Як знаходять теоретичні частоти нормального розподілу для перевірки гіпотези за правилом Пірсона?

Розділ Б

1. У розигранні першості країни з футболу бере участь 17 команд. Скількима способами можуть бути розподілені золота, срібна та бронзова медалі?

Д.3. Тест для перевірки знань та навичок

2. Проведено три постріли у мішень. Розглядають такі елементарні події

A_i — влучення у мішень при i -тому пострілі;

\bar{A}_i — промах при i -тому пострілі, ($i = 1, 2, 3$).

Виразити через A_i та \bar{A}_i такі події

A — усі три влучення; B — рівно два влучення;

C — усі три промахи; D — хоча б одне влучення;

E — більш одного влучення; F — не більш одного влучення.

3. Гральний кубик підкидають двічі. Описати простір елементарних подій. Описати події: A — сума очок, яка з'явилась, дорівнює 8; B — принаймні один раз з'явиться 6.

4. Монету кинуто двічі. Знайти імовірність того, що хоча б один раз вона впаде гербом вгору.

5. Із 1000 лотерейних квитків є один з виграшем 50 гривень, п'ять — з виграшем по 20 гривень, 20 — по 10 гривень, 50 — по 5 гривень. Знайти імовірність виграшу на один квиток а) не менш 10 гривень; б) будь-якого виграшу.

6. У колі радіуса R лежить коло радіуса r . Знайти імовірність того, що точка, навмання поставлена у велике коло, попаде також і до області малого кола.

7. Серед 25 студентів групи, в якій є 10 дівчат, розігрують 5 квитків на концерт. Визначити імовірність того, що білети виграють дві дівчини.

8. Стрілок влучить у «десятку» з імовірністю 0,05, в «дев'ятку» — з імовірністю 0,20, у «вісімку» — з імовірністю 0,6.

Зроблено один постріл. Яка імовірність того, що а) вибито не менше 8 очок; б) вибито більше 8 очок?

9. Транспортування вантажу для замовника виконують два автопідприємства, кожне з яких повинно виділити для цього по одній вантажівці. Імовірність виходу на лінію вантажівки з першого автопарку дорівнює 0,7, а з другого — 0,6. Знайти імовірність того, що замовник одержить потрібні вантажівки.

10. В урні 10 білих, 15 чорних, 20 блакитних та 25 червоних куль однакового розміру. Навмання взято одну кулю. Знайти імовірність того, що ця куля буде:

а) біла або чорна; б) блакитна або червона; в) біла, чорна або блакитна.

11. У білому ящику 12 червоних та 6 блакитних куль. У чорному ящику — 15 червоних та 10 блакитних куль. Кидають гральний кубик. Якщо випаде кількість очок, кратна 3, то навмання беруть кулю з білого ящика. Якщо випаде будь-яка інша кількість очок, то беруть кулю з чорного ящика. Якою буде імовірність узяти червону кулю?

12. У трьох ящиках маємо однакові деталі з різних заводів: у першому — 20 стандартних та 5 нестандартних деталей; у другому — 15 стандартних та 3 нестандартних; у третьому — 14 стандартних та 2 нестандартних.

Із навмання взятого ящика навмання взята деталь, яка виявилася стандартною. Знайти імовірність того, що цю деталь взято:

а) з першого ящика; б) із другого ящика; в) із третього ящика.

13. Імовірність виграшу однієї облігації за весь час займу дорівнює 0,6. Знайти імовірність того, що у власника 5 облігацій:

а) виграш випаде на дві облігації; б) виграш випаде більш ніж на дві облігації;

14. Імовірність появи події A хоча б один раз в 5 незалежних випробуваннях дорівнює 0,9. Яка імовірність появи події A в одному випробуванні, якщо при кожному випробуванні ця імовірність однакова?

15. З кожних 40 виробів, виготовлених верстатом-автоматом, 4 є бракованими. Взято навмання 400 виробів автомату. Знайти імовірність того, що серед них 350 без дефекту.

16. Імовірність присутності студента на лекції дорівнює 0,8. Знайти імовірність того, що із 100 студентів на лекції будуть а) не менше 75 та не більше 90;

17. Середня кількість замовлень таксі, що надходять до диспетчерського пункту щохвилини, дорівнює 3, знайти імовірність того, що за дві хвилини надійде а) 4 замовлення; б) менше за 4 замовлення; в) не менше 4 замовлень.

18. Нехай X — кількість появ числа 5 при двох киданнях грального кубика. Знайти закон розподілу величини X та її математичне сподівання.

19. Випадкова величина X задана інтегральною функцією

$$F(x) = \begin{cases} 0, & x \leq -1; \\ \frac{3}{4}x + \frac{3}{4}, & -1 < x \leq \frac{1}{3}; \\ 1, & x > \frac{1}{3}. \end{cases}$$

Знайти імовірність того, що X прийме значення з проміжку $\left(0, \frac{1}{3}\right)$.

20. Випадкова величина X задана функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 2; \\ 0,5x - 1, & 2 < x < 4; \\ 1, & x > 4. \end{cases}$$

Знай їй імовірність того, що X прийме значення менше 0,2.

21. Функція розподілу випадкової величини X має вигляд

$$F(x) = \begin{cases} 0, & x \leq 0; \\ x^3, & 0 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Знайти математичне сподівання та дисперсію X .

22. Неперервна випадкова величина X розподілена за законом, заданим диференціальною функцією

$$f(x) = \begin{cases} 3e^{-3x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Знайти імовірність того, що X потрапить в інтервал $(0,13; 0,7)$.

23. Для заданої вибірки із генеральної сукупності скласти розподіли частот та частостей

7, 4, 4, 8, 12, 12, 12, 7, 8, 12, 8, 12, 4, 12, 12, 4, 12, 4, 12, 12;

24. Для заданої вибірки із генеральної сукупності скласти розподіли частот та частостей

6, 9, 5, 3, 6, 6, 9, 3, 5, 6, 9, 5, 6, 6, 9, 6, 9, 6, 6, 6;

25. Знайти емпіричну функцію заданого розподілу вибірки та побудувати її графік.

x_k	2	5	7	8
n_k	1	3	2	4

26. Побудувати полігон частот та частостей заданого розподілу вибірки.

x_k	2	3	5	6
n_k	10	15	5	20

27. Побудувати гістограму частот заданого розподілу вибірки.

Інтервали	(1; 5)	(5; 9)	(9; 13)	(13; 17)	(17; 21)
Частоти	10	20	50	12	8

28. Вимірювальним приладом, який не має систематичних похибок, були зроблені незалежні вимірювання деякої величини. Знайти незсунуті оцінки математичного сподівання та дисперсії цієї величини за даними результатами вимірювань 8, 9, 11, 12.

29. Вимірювальним приладом, який не має систематичних похибок, були зроблені незалежні вимірювання деякої величини. Знайти незсунуті оцінки математичного сподівання та дисперсії цієї величини за даними результатами вимірювань 4, 6, 10, 20.

30. Знайти методом добутоків вибірккову середню та вибірккову дисперсію заданої вибірки.

n_k	12	14	16	18	20	22
n_k	5	15	50	16	10	4

31. Знайти методом добутоків вибірккову середню та вибірккову дисперсію заданої вибірки.

n_k	65	70	75	80	85
n_k	2	5	25	15	3

Д.3. Тест для перевірки знань та навичок

32. Знайти методом добутків вибірку середню та вибірку дисперсію заданої вибірки.

n_k	7,14	7,21	7,28	7,35	7,42	7,49	7,56	7,63	7,70
n_k	2	9	24	43	51	37	25	8	1

33. По даних двох незалежних вибірок об'єму $n_1 = 11$ та $n_2 = 14$ із нормальних сукупностей X та Y знайдені виправлені вибірккові дисперсії $s_1^2 = 0,76$ та $s_2^2 = 0,38$. При рівні значущості $\alpha = 0,05$ перевірити гіпотезу $H_0 : D(X) = D(Y)$ при альтернативній $H_1 : D(X) > D(Y)$.

34. По даних двох незалежних вибірок об'єму $n_1 = 9$ та $n_2 = 16$ із нормальних сукупностей X та Y знайдені виправлені вибірккові дисперсії $s_1^2 = 34,02$ та $s_2^2 = 12,15$. При рівні значущості $\alpha = 0,01$ перевірити гіпотезу $H_0 : D(X) = D(Y)$ при альтернативній $H_1 : D(X) > D(Y)$.

35. Використовуючи критерій Пірсона (χ^2 -квадрат) з рівнем значущості α , перевірити гіпотезу про нормальний розподіл генеральної сукупності X , якщо відомі емпіричні n_k та теоретичні n'_k частоти.

$$\alpha = 0,05;$$

n_k	15	26	25	30	26	21	24	20	13
n_k	9.1	16.5	25.3	32	33.9	29.8	22	13.5	7

Розділ В

1. Товариство з n чоловік сідає за круглий стіл. Знайти ймовірність того, що певні дві особи будуть поряд.

2. Із колоди карт (52) вибирають 6. Знайти ймовірність того, що серед цих 6 карт будуть представники 4 мастей.

3. У урні 15 червоних, 9 синіх та 6 зелених куль однакового розміру. Навмання беруть 6 куль. Яка імовірність того, що будуть узяті 1 зелена, 2 синіх та 3 червоних кулі?

4. У першій урні 10 куль, з них 8 білих. У другій урні 20 куль, з них 4 білих. Із кожної урни навмання взято по одній кулі, а потім із двох обраних навмання взято одну. Знайти ймовірність того, що остання куля буде білою.

5. На відрізку $[0, 2]$ навмання вибрані числа x, y . Знайти ймовірність того, що $y - x < 1, x + y > 1$.

6. У лотереї n — білетів, серед них m таких, що виграють. Обчислити ймовірність виграшу для того, хто має r білетів.

7. Визначити ймовірність того, що 100 лампочок, взятих навмання із 1000 виявляться незіпсованими, якщо відомо, що число зіпсованих лампочок на 1000 штук рівноможливо від 0 до 5.

8. З урни в якій n та m чорних куль загублено 1 кулю. Для того, щоб визначити склад куль з урни виймають дві кулі, які виявилися білими. Визначити ймовірність того, що загублена біла куля.

9. Є n урн, в кожній з яких m білих та k чорних куль. З першої урни навмання виймається одна куля і перекладається в другу. Потім з другої навмання виймається одна куля і перекладається в третю урну і т.д. Визначити ймовірність вилучення після такого перекладання білої кулі з останньої урни.

10. Троє бізнесменів на полюванні одночасно зробили постріли по ведмедю, який був убитий однією кулею. Визначити ймовірність того, що медвідь був убитий першим, другим або третім бізнесменом, якщо ймовірності попадання для них дорівнюють відповідно 0,2, 0,4, 0,6.

11. Подія A може з'явитись у випробуванні з ймовірністю p . Провели n випробувань при однакових умовах незалежним чином. Чому дорівнює ймовірність того, що будуть виконані нерівності

$$np - 2\sqrt{pq} \leq m \leq np + 2\sqrt{pq},$$

де m — число випробувань, в яких подія A з'явилась.

12. Випадкова величина X рівномірно розподілена в інтервалі $(-1, 1)$. Знайти щільність ймовірності випадкової величини $Y = X^2$.

13. (Задача де-Мере). Скільки раз треба підкинути два гральних кубика, щоб ймовірність принаймні однієї появи шестірок була

більше за $\frac{1}{2}$.

14. Задана функція розподілу прибутку за рік особам, які обкладені податком

Д.3. Тест для перевірки знань та навичок

$$F(x) \begin{cases} 1 - \left(\frac{x_0}{x}\right)^a, & \text{при } x \geq x_0, \quad a \geq 0; \\ 0, & \text{при } x \leq x_0. \end{cases}$$

Визначити розмір прибутку за рік, який для випадково вибраного платіжника податку може бути перевищеним з імовірністю 0,5.

15. Використовуючи точкову діаграму, одержати розподіл частот заданої вибірки, побудувати графік емпіричної функції розподілу та полігон частот

3, 6, 6, 5, 6, 2, 3, 6, 6, 5, 2, 3, 3, 6, 2, 3, 6, 6, 6, 5, 2, 6, 2, 3, 2, 5, 5, 6, 6, 3, 2, 3, 6, 2, 6, 2, 2, 3, 3, 6, 3, 6, 6, 3, 3, 6, 3, 6, 6, 3.

16. Знайти з надійністю 0,95 довірчий інтервал оцінки математичного сподівання a нормально розподіленої ознаки X генеральної сукупності, якщо відомі вибіркова середня \bar{x}_B , об'єм вибірки n та середнє квадратичне відхилення σ генеральної сукупності.

$$\bar{x}_B = 14, \quad n = 25, \quad \sigma = 5.$$

17. Використовуючи критерій χ^2 перевірити гіпотезу H_0 : наведений у таблиці розподіл зросту 1000 дорослих чоловіків є вибіркою з нормальної генеральної сукупності з

$$a = \bar{x}_B \quad \text{та} \quad \sigma = \sqrt{\sigma_B^2}.$$

Зріст см.	Кількість чоловіків	Зріст см.	Кількість чоловіків
144–147	1	168–171	172
147–150	3	171–174	120
150–153	7	174–177	63
153–156	26	177–180	27
156–159	66	180–183	9
159–162	114	183–186	3
162–165	186	186–189	1
165–168	200	189–192	2

Відповіді до розділу Б

1. $17 \cdot 16 \cdot 15$.

2. $A = A_1 \cdot A_2 \cdot A_3$; $B = A_1 \cdot A_2 \cdot \bar{A}_3 \cup A_1 \cdot \bar{A}_2 \cdot A_3 \cup \bar{A}_1 \cdot A_2 \cdot A_3$.

3. $\Omega = \{\Gamma\Gamma, \Gamma\text{P}, \text{P}\Gamma, \text{P}\text{P}\}$, $A = \{\Gamma\Gamma, \Gamma\text{P}, \text{P}\Gamma\}$, $B = \{\Gamma\Gamma, \text{P}\Gamma\}$.

4. $\frac{3}{4}$.

5. а) 0,026; б) 0,076.

6. $p = \frac{r^2}{R^2}$.

7. $\frac{195}{253} \approx 0,771$.

8. а) 0,85; б) 0,25.

9. 0,88;

10. а) $\frac{5}{14}$; б) $\frac{9}{14}$; в) $\frac{9}{14}$.

11. $\frac{28}{45}$.

12. а) 0,319; б) 0,332; в) 0,349.

13. а) $P_5(2) = 0,2304$; б) 0,3072.

14. 0,37.

15. $P_{400}(350) = \frac{1}{6}\varphi(-1,67) = 0,0165$.

16. а) $P_{100}(75 \leq m \leq 90) = 0,4938 + 0,3944 = 0,8882$;

17. а) 0,135; б) 0,1525; в) 0,8475.

18.

X	0	1	2
p	$\frac{25}{36}$	$\frac{5}{18}$	$\frac{1}{36}$

; $M(X) = \frac{1}{3}$.

Д.3. Тест для перевірки знань та навичок

19. $P\left(0 < X < \frac{1}{3} = \frac{1}{4}\right)$.

20. 0.

21. $M(X) = \frac{3}{4}$; $D(X) = \frac{3}{80}$.

22. а) $P(0,13 < X < 0,7) = e^{-0,39} - e^{-2,1} = 0,677 - 0,122 = 0,555$.

23.

x_k	4	7	8	12
n_k	5	2	3	10
$\frac{n_k}{n}$	$\frac{1}{4}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$

24.

x_k	3	5	6	9
n_k	2	3	10	50
$\frac{n_k}{n}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{2}$	$\frac{1}{4}$

28. $\overline{x_B} = 10$, $S^2 = 0$.

29. $\overline{x_B} = 10$, $S^2 = 0$.

30. $\overline{x_B} = 16,46$, $D_B = 4,87$.

31. $\overline{x_B} = 76,2$, $D_B = 18,56$.

32. $\overline{x_B} = 7,416$, $D_B = 0,116$.

33. $F_{cn} = 2$, $F_{kp}(0,5; 10; 13) = 2,67$, H_0 приймається.

34. $F_{cn} = 2,8$, $F_{kp}(0,01; 8; 15) = 2,64$, H_0 відхиляється.

35. $\chi_{kp}^2 = 16,9$; $\chi_{kp}^2 = 177,13$, H_0 відхиляємо.

Відповіді до розділу В

$$1. p = \frac{2}{n-1}.$$

$$2. p = 1 - \frac{4 \cdot C_{39}^6}{C_{52}^6} - \frac{6 \cdot C_{26}^6}{C_{52}^6} - \frac{4 \cdot C_{13}^6}{C_{52}^6}.$$

$$3. \frac{24}{145} \approx 0,17.$$

$$4. 0,5.$$

$$5. p(A) = \frac{3}{4}.$$

$$6. p(A) = \sum_{s=1}^r \frac{C_m^s \cdot C_{n-m}^{r-s}}{C_n^r}.$$

$$7. p \approx 0,78.$$

$$8. \frac{n-2}{n+m-2}.$$

$$9. p = \frac{m}{m+k}.$$

$$10. 0,103; 0,277; 0,620.$$

$$11. 2\Phi(2)=0,9545.$$

$$12. q(y) = \begin{cases} \frac{1}{4\sqrt{y}}, & 0 < y < 1; \\ 0, & \text{при } y \leq 0 \text{ та } y \geq 1. \end{cases}$$

$$13. 1 - \left(\frac{35}{36}\right)^n \geq \frac{1}{2} \text{ при } n \geq 25.$$

$$14. x \geq 2^{1/\alpha} x_0.$$

Д.3. Тест для перевірки знань та навичок

16. (12,04; 15,96).

17.

$$n'_k = 1000 p_k;$$

$$p_k = P(x_k < X < x_{k+1}) = \Phi\left(\frac{x_{k+1} - a}{\sigma}\right) - \Phi\left(\frac{x_k - a}{\sigma}\right).$$

Вибірка узгоджується з гіпотезою.

Д.4. Таблиці

Таблиця 1

Значення функції e^{-x}

x	e^{-x}	x	e^{-x}	x	e^{-x}	x	e^{-x}
0,00	1,000	0,40	0,670	0,80	0,449	3,00	0,050
0,01	0,990	0,41	0,644	0,81	0,445	3,10	0,040
0,02	0,980	0,42	0,657	0,82	0,440	3,20	0,041
0,03	0,970	0,43	0,650	0,83	0,436	3,30	0,037
0,04	0,961	0,44	0,644	0,84	0,432	3,40	0,033
0,05	0,951	0,45	0,638	0,85	0,427	3,50	0,030
0,06	0,942	0,46	0,631	0,86	0,423	3,60	0,027
0,07	0,932	0,47	0,625	0,87	0,419	3,70	0,025
0,08	0,923	0,48	0,619	0,88	0,415	3,80	0,022
0,09	0,914	0,49	0,613	0,89	0,411	3,90	0,020
0,10	0,905	0,50	0,606	0,90	0,407	4,00	0,0183
0,11	0,896	0,51	0,600	0,91	0,403	4,10	0,0166
0,12	0,887	0,52	0,595	0,92	0,399	4,20	0,0150
0,13	0,878	0,53	0,589	0,93	0,395	4,30	0,0136
0,14	0,869	0,54	0,583	0,94	0,391	4,40	0,0123
0,15	0,861	0,55	0,577	0,95	0,387	4,50	0,0111
0,16	0,852	0,56	0,571	0,96	0,383	4,60	0,0101
0,17	0,844	0,57	0,565	0,97	0,379	4,70	0,0091
0,18	0,835	0,58	0,560	0,98	0,375	4,80	0,0082
0,19	0,827	0,59	0,554	0,99	0,372	4,90	0,0074
0,20	0,819	0,60	0,549	1,00	0,368	5,00	0,0067
0,21	0,811	0,61	0,543	1,10	0,333	5,10	0,0061
0,22	0,803	0,62	0,538	1,20	0,302	5,20	0,0055
0,23	0,795	0,63	0,533	1,30	0,273	5,30	0,0050
0,24	0,787	0,64	0,527	1,40	0,247	5,40	0,0045
0,25	0,779	0,65	0,522	1,50	0,223	5,50	0,0041
0,26	0,771	0,66	0,517	1,60	0,202	5,60	0,0037
0,27	0,763	0,67	0,512	1,70	0,183	5,70	0,0033
0,28	0,756	0,68	0,507	1,80	0,165	5,80	0,0030
0,29	0,748	0,69	0,502	1,90	0,150	5,90	0,0027
0,30	0,741	0,70	0,497	2,00	0,135	6,00	0,0025
0,31	0,733	0,71	0,492	2,10	0,122	6,10	0,0022

Таблиця 1а (продовження)

Значення функції e^{-x}

x	e^{-x}	x	e^{-x}	x	e^{-x}	x	e^{-x}
0,32	0,726	0,72	0,487	2,20	0,111	6,20	0,0020
0,33	0,719	0,73	0,482	2,30	0,100	6,30	0,0018
0,34	0,712	0,74	0,477	2,40	0,091	6,40	0,0017
0,35	0,705	0,75	0,472	2,50	0,082	6,50	0,0015
0,36	0,698	0,76	0,468	2,60	0,074	6,60	0,0014
0,37	0,691	0,77	0,463	2,70	0,067	6,70	0,0012
0,38	0,684	0,78	0,458	2,80	0,061	6,80	0,0011
0,39	0,677	0,79	0,454	2,90	0,055	6,90	0,0010

Таблиця 2

Значення локальної функції Лапласа

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

x	0	1	2	3	4
0,0	0,3989	0,3989	0,3989	0,3988	0,3986
0,1	0,3970	0,3965	0,3961	0,3956	0,3951
0,2	0,3910	0,3902	0,3894	0,3885	0,3876
0,3	0,3814	0,3802	0,3790	0,3778	0,3765
0,4	0,3683	0,3668	0,3653	0,3637	0,3621
0,5	0,3521	0,3503	0,3485	0,3467	0,3448
0,6	0,3332	0,3312	0,3292	0,3271	0,3251
0,7	0,3123	0,3101	0,3079	0,3056	0,3034
0,8	0,2897	0,2874	0,2850	0,2827	0,2903
0,9	0,2661	0,2637	0,2613	0,2589	0,2565
1,0	0,2420	0,2396	0,2371	0,2347	0,2323
1,1	0,2179	0,2155	0,2131	0,2107	0,2083
1,2	0,1942	0,1919	0,1876	0,1872	0,1849
1,3	0,1714	0,1691	0,1669	0,1647	0,1626
1,4	0,1497	0,1476	0,1456	0,1435	0,1415

Таблиця 2а (продовження)
Значення локальної функції Лапласа

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

<i>x</i>	0	1	2	3	4
1,5	0,1295	0,1276	0,1257	0,1238	0,1219
1,6	0,1109	0,1092	0,1074	0,1057	0,1040
1,7	0,0940	0,0925	0,0909	0,0893	0,0878
1,8	0,0790	0,0775	0,0761	0,0748	0,0734
1,9	0,0656	0,0644	0,0632	0,0620	0,0608
2,0	0,0540	0,0519	0,0519	0,0508	0,0498
2,1	0,0440	0,0431	0,0422	0,0413	0,0404
2,2	0,0355	0,0347	0,0339	0,0332	0,0325
2,3	0,0283	0,0277	0,0270	0,0264	0,0258
2,4	0,0224	0,0219	0,0213	0,0208	0,0203
2,5	0,0175	0,0171	0,0167	0,0163	0,0158
2,6	0,0136	0,0132	0,0129	0,0126	0,0122
2,7	0,0104	0,0101	0,0099	0,0096	0,0093
2,8	0,0079	0,0077	0,0075	0,0073	0,0071
2,9	0,0060	0,0058	0,0056	0,0055	0,0053
3,0	0,0044	0,0043	0,0042	0,0040	0,0039
3,1	0,0033	0,0032	0,0031	0,0030	0,0029
3,2	0,0024	0,0023	0,0022	0,0022	0,0021
3,3	0,0017	0,0017	0,0016	0,0016	0,0015
3,4	0,0012	0,0012	0,0012	0,0011	0,0011
3,5	0,0009	0,0008	0,0008	0,0008	0,0007
3,6	0,0006	0,0006	0,0006	0,0005	0,0005
3,7	0,0004	0,0004	0,0004	0,0004	0,0004
3,8	0,0003	0,0003	0,0003	0,0003	0,0003
3,9	0,0002	0,0002	0,0002	0,0002	0,0002

Таблиця 2б (продовження)

Значення локальної функції Лапласа

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

<i>x</i>	5	6	7	8	9
0,0	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3752	0,3739	0,3726	0,3712	0,3697
0,4	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3230	0,3209	0,3387	0,3166	0,3144
0,7	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2780	0,2757	0,2732	0,2709	0,2685
0,9	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1023	0,1006	0,0983	0,0973	0,0957
1,7	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0396	0,0388	0,0379	0,0371	0,0363
2,2	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0198	0,0194	0,0189	0,0184	0,0180

Таблиця 2в (продовження)
Значення локальної функції Лапласа

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

<i>x</i>	5	6	7	8	9
2,5	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0015	0,0015	0,0014	0,0013	0,0013
3,4	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0001	0,0001

Таблиця 3

Значення інтегральної функції Лапласа

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

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,0	0,000	0,55	0,2088	1,3	0,4032	2,5	0,4938
0,01	0,0040	0,06	0,2257	1,4	0,4192	2,6	0,4953
0,05	0,0199	0,65	0,2422	1,5	0,4332	2,7	0,4965
0,1	0,0398	0,07	0,2580	1,6	0,4452	2,8	0,4974
0,15	0,0596	0,075	0,2734	1,7	0,4554	2,9	0,4981
0,2	0,0798	0,08	0,2881	1,8	0,4641	3,0	0,4986
0,25	0,0987	0,85	0,3023	1,9	0,4713	3,2	0,4993
0,3	0,1179	0,9	0,3159	2,0	0,4772	3,4	0,4997
0,35	0,1368	0,95	0,3289	2,1	0,4821	3,6	0,49984
0,4	0,1554	1,00	0,3413	2,2	0,4861	3,8	0,49992
0,45	0,1736	1,1	0,3643	2,3	0,4893	4,0	0,49996
0,5	0,1915	1,2	0,3849	2,4	0,4918	4,5	0,49997
—	—	—	—	—	—	5,0	0,49998

Таблиця 4

Значення $P_k = \frac{a^k}{k!} e^{-a}$ (розподіл Пуассона)

k	a				
	1	2	3	4	5
0	0,9048	0,8187	0,7408	0,6703	0,6065
1	0,0905	0,1638	0,2222	0,2681	0,3033
2	0,0045	0,0164	0,0333	0,0536	0,0758
3	0,0002	0,0019	0,0033	0,0072	0,0126
4		0,0001	0,0002	0,0007	0,0016
5				0,0001	0,0002
6					

Таблиця 4а (продовження)

Значення $P_k = \frac{a^k}{k!} e^{-a}$ (розподіл Пуассона)

k	a			
	0,6	0,7	0,8	0,9
0	0,5488	0,4966	0,4493	0,4066
1	0,3293	0,3476	0,3595	0,3659
2	0,0988	0,1217	0,1438	0,1647
3	0,0198	0,0284	0,0383	0,0494
4	0,0030	0,0050	0,0077	0,0111
5	0,0004	0,0007	0,0012	0,0020
6		0,0001	0,0002	0,0003

Таблиця 4б (продовження)

Значення $P_k = \frac{a^k}{k!} e^{-a}$ (розподіл Пуассона)

k	a				
	1	2	3	4	5
0	0,3679	0,1353	0,0498	0,0183	0,0067
1	0,3679	0,2707	0,1494	0,0733	0,0337
2	0,1839	0,2707	0,2240	0,1465	0,0842
3	0,0613	0,1804	0,2240	0,1954	0,1404
4	0,0153	0,0902	0,1680	0,1954	0,1755
5	0,0031	0,0361	0,1008	0,1563	0,1755
6	0,0005	0,0120	0,0504	0,1042	0,1462
7	0,0001	0,0037	0,0216	0,0595	0,1044
8		0,0009	0,0081	0,0298	0,0653
9		0,0002	0,0027	0,0132	0,0363
10			0,0008	0,0053	0,0181
11			0,0002	0,0019	0,00825
12			0,0001	0,0006	0,0034
13				0,0002	0,0013
14				0,0001	0,0005
15					0,0002

Значення $P_k = \frac{a^k}{k!} e^{-a}$ (розподіл Пуассона)

k	a				
	6	7	8	9	10
0	0,0025	0,0009	0,0003	0,0001	0,0000
1	0,0149	0,0064	0,0027	0,0011	0,0005
2	0,0446	0,0223	0,0107	0,0050	0,0023
3	0,0892	0,0521	0,0286	0,0150	0,0076
4	0,1339	0,0912	0,0572	0,0337	0,0189
5	0,1606	0,1277	0,0916	0,0607	0,0378
6	0,1606	0,1490	0,1221	0,0911	0,0631
7	0,1377	0,1490	0,1396	0,1171	0,0901
8	0,1033	0,1304	0,1396	0,1318	0,1126
9	0,0688	0,1014	0,1241	0,1318	0,1251
10	0,0413	0,0710	0,0933	0,1186	0,1251
11	0,0225	0,0452	0,0722	0,0970	0,1377
12	0,0126	0,0263	0,0481	0,0728	0,0948
13	0,0052	0,0142	0,0296	0,0504	0,0729
14	0,0022	0,0071	0,0169	0,0324	0,0521
15	0,0009	0,0033	0,0090	0,0194	0,0347
16	0,0003	0,0014	0,0045	0,0109	0,0217
17	0,0001	0,0006	0,0021	0,0058	0,0128
18		0,0002	0,0009	0,0029	0,0071
19		0,0001	0,0004	0,0014	0,0037
20			0,0002	0,0006	0,0019
21			0,0001	0,0003	0,0009
22				0,0001	0,0004
23					0,0002
24					0,0001

Таблиця 5

Значення критерію Пірсона $P(\chi_{cn}^2 > \chi_{кр}^2) = \alpha$
 (ступінь вільності $k = m - s - 1$)

k	α						
	0,99	0,98	0,95	0,9	0,8	0,7	0,5
1	0,000	0,001	0,004	0,016	0,064	0,148	0,455
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386
3	0,115	0,185	0,352	0,584	1,005	1,424	2,370
4	0,297	0,429	0,711	1,064	1,649	2,200	3,360
5	0,554	0,752	1,145	1,610	2,340	3,000	4,350
6	0,872	1,134	1,635	2,200	3,070	3,830	5,350
7	1,239	1,564	2,170	2,830	3,820	4,670	6,350
8	1,646	2,030	2,730	3,490	4,590	5,530	7,340
9	2,090	2,530	3,320	4,170	5,380	6,390	8,340
10	2,560	3,060	3,940	4,860	6,180	7,270	9,340

Таблиця 5а (продовження)

Значення критерію Пірсона $P(\chi_{cn}^2 > \chi_{кр}^2) = \alpha$
 (ступінь вільності $k = m - s - 1$)

k	α						
	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	1,074	1,642	2,710	3,840	5,410	6,640	10,830
2	2,410	3,220	4,600	5,990	7,820	9,210	13,820
3	3,660	4,640	6,250	7,820	9,840	11,340	16,270
4	4,880	5,990	7,780	9,490	11,670	13,280	18,460
5	6,060	7,290	9,240	11,070	13,390	15,090	20,500
6	7,230	8,560	10,640	12,590	15,030	16,810	22,500
7	8,380	9,800	12,020	14,070	16,620	18,480	24,300
8	9,520	11,030	13,360	15,510	18,170	20,100	27,100
9	10,660	12,240	14,680	16,920	19,680	21,700	27,900
10	11,780	13,440	15,990	18,310	21,200	23,200	29,600

Критичні точки розподілу Стюдента

Число степенів вільності k	Рівень значущості α (двохстороння критична область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
	0,05	0,025	0,01	0,005	0,001	0,0005
	Рівень значущості α (одностороння критична область)					

Таблиця 6 (продовження)
Критичні точки розподілу Стюдента

Число степенів вільності k	Рівень значущості α (двохстороння критична область)					
	0,10	0,05	0,02	0,01	0,002	0,001
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	Рівень значущості α (одностороння критична область)					

Критичні точки розподілу F Фішера-Снедекора
 (k_1 – число степенів вільності більшої дисперсії)
 (k_2 – число степенів вільності меншої дисперсії)
 Рівень значущості $\alpha = 0,01$

k_2	k_1					
	1	2	3	4	5	6
1	4052	4999	5403	5625	5764	5889
2	98,49	99,01	99,17	99,25	99,33	99,30
3	34,12	30,81	29,46	28,71	28,24	27,91
4	21,20	18,00	16,69	15,98	15,52	15,21
5	16,26	13,27	12,06	11,39	10,97	10,67
6	13,74	10,92	9,78	9,15	8,75	8,47
7	12,25	9,55	8,45	7,85	7,46	7,19
8	11,26	8,65	7,59	7,01	6,63	6,37
9	10,56	8,02	6,99	6,42	6,06	5,80
10	10,04	7,56	6,55	5,99	5,64	5,39
11	9,85	7,20	6,22	5,67	5,32	5,07
12	9,33	6,93	5,95	5,41	5,06	4,82
13	9,07	6,70	5,74	5,20	4,86	4,62
14	8,86	6,51	5,56	5,03	4,69	4,46
15	8,68	6,36	5,42	4,89	4,56	4,32
16	8,53	6,23	5,29	4,77	4,44	4,20
17	8,40	6,11	5,18	4,67	4,34	4,10

Таблиця 7а (продовження)

Критичні точки розподілу F Фішера-Снедекора $(k_1$ — число степенів вільності більшої дисперсії) $(k_2$ — число степенів вільності меншої дисперсії)Рівень значущості $\alpha = 0,01$

k_2	k_1					
	7	8	9	10	11	12
1	5928	5981	6022	6056	6082	6106
2	99,34	99,36	99,36	99,40	99,41	99,42
3	27,67	27,49	27,34	27,23	27,13	27,05
4	14,98	14,80	14,66	14,54	14,45	14,37
5	10,45	10,27	10,15	10,05	9,96	9,89
6	8,26	8,10	7,98	7,87	7,79	7,72
7	7,00	6,84	6,71	6,62	6,54	6,47
8	6,19	6,03	5,91	5,82	5,74	5,67
9	5,62	5,47	5,35	5,26	5,18	5,11
10	5,21	5,06	4,95	4,85	4,78	4,71
11	4,88	4,74	4,63	4,54	4,46	4,40
12	4,65	4,50	4,39	4,30	4,22	4,16
13	4,44	4,30	4,19	4,10	4,02	3,96
14	4,28	4,14	4,03	3,94	3,86	3,80
15	4,14	4,00	3,89	3,80	3,73	3,67
16	4,03	3,89	3,78	3,69	3,61	3,55
17	3,93	3,79	3,68	3,59	3,52	3,45

2220 випадкових чисел

01.	49487	25252	02431	69414	77285	52852	98740	85022	17778	81833	63789	61840
02.	52802	97738	42193	89353	35179	11444	98054	58736	03840	93449	54958	81740
03.	28667	23901	96960	70724	92042	71868	30195	12138	21636	57781	33167	60986
04.	62058	11106	19620	67893	67518	34534	09891	35146	56269	94621	10909	12498
05.	87822	86864	29188	23218	56773	69124	18453	62085	08149	90998	40343	71546
06.	14704	55808	05863	72452	68374	02760	79464	36170	19001	37561	81023	42249
07.	18519	22557	92900	03095	71115	06406	01156	25433	67367	59688	61590	13812
08.	17889	23214	06836	68333	98166	95234	95522	80787	13138	93299	44474	59902
09.	45869	15021	13433	13751	43352	87995	06884	96496	02400	27726	39810	27864
10.	14454	54268	21709	37260	06414	78560	55073	40579	89515	82167	10305	21809
11.	42243	45236	40338	54040	49158	80958	07636	37227	99460	60248	95076	20944
12.	10153	09129	42477	71253	20908	03808	04876	80750	45915	75845	79089	97852
13.	20891	53031	78804	88789	44859	83655	61063	08261	45637	37296	87380	26586
14.	90883	12260	36272	98203	29089	18415	57571	97048	41353	33783	28982	32796
15.	15782	01278	72053	54999	76130	96563	69434	60438	35335	42393	97750	51513
16.	98167	14404	07958	96564	51442	43582	14965	75053	69087	28185	82221	47475
17.	86837	40969	67158	00789	34453	82207	20911	05939	57536	31880	35584	48621
18.	99166	33419	60979	68879	98590	53322	73162	34414	68418	00241	27444	20067
19.	92143	14188	79891	47134	37353	30419	33576	16685	10247	31642	85793	88975
20.	82441	69557	92409	83941	61137	64435	52839	32103	93253	37526	69755	39506
21.	30458	38905	96545	21944	36910	48745	32519	75757	07911	89887	30543	68442
22.	49207	91282	15638	28328	71928	47626	91993	12965	97756	03328	99488	55201
23.	62358	79309	90114	00692	63327	28856	59374	29285	89561	76911	75363	33946
24.	41532	49022	93730	89164	00980	28382	83994	11481	27464	93168	94187	42495
25.	30057	17405	13741	96025	32154	60639	59873	31744	25133	56236	32885	28384
26.	53017	18830	70177	01383	46006	51370	51217	41754	50026	39056	23887	89889
27.	10375	09136	49175	50252	62289	70091	62806	24428	16436	67905	10872	50278
28.	97204	07629	42113	67044	28079	58261	20028	81819	75846	94933	22793	91985
29.	98675	01785	21600	70596	03076	70135	26545	02354	83718	05456	26232	58185
30.	77634	78317	69625	58266	15619	88259	16820	37895	08533	52347	87356	19124
31.	22403	70701	69804	65806	43902	49145	47363	26244	72875	09065	68256	38744
32.	56698	36907	96122	22398	53070	71587	36295	87033	39496	16283	51225	81018
33.	88524	51242	42342	19470	54319	14273	62126	90247	06385	61398	92645	41909
34.	13692	52083	28467	63653	19347	62440	42358	79131	48458	08288	77747	70458
35.	55012	43126	79037	27055	59506	15770	20322	38773	30545	00708	33104	72459
36.	25343	90379	13218	02606	75440	03281	82000	67687	74383	21816	81206	66136
37.	76391	60380	63510	43347	90826	58124	52830	45541	22814	39615	00112	97266

*На закінчення бажаємо
придбати знання
при можливості завжди
застосовувать в житті,
регулярно поновляти,
творчість в праці виявляти.*

Список літератури

До I та II частин

[1] *Белинский В.А., Калихман И.Л., Майстров Л.Е., Митькин А.М.* Высшая математика с основами математической статистики — М.: Высшая школа, 1965.

[2] *Володин Б.Г., Ганин М.П., Динер И.Я., Комаров Л.Б., Свешников А.А., Старобин К.Б.* Сборник задач по теории вероятностей, математической статистике и теории случайных функций — М.: Наука, 1965.

[3] *Волощенко А.Б.* Теория вероятностей и математическая статистика инженеров-экономистов — К.: КИНХ, 1972.

[4] *Гмурман В.Е.* Теория вероятностей и математическая статистика — М.: Высшая школа, 1980.

[5] *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике, — М.: Высшая школа, 1975.

[6] *Гнеденко Б.В.* Курс теории вероятностей — М.: Наука, 1988.

[7] *Гурский Е.И., Скобля Т.В., Юшкевич В.Э.* Методическое пособие по теории вероятностей и математической статистике — Минск: Высшая школа, 1973.

[8] *Жлуктенко В.І., Наконечний С.І.* Теорія ймовірностей і математична статистика, — К.: ІЗМН, 1997. — 408 с.

[9] *Леоненко М.М., Мішура Ю.С., Пархоменко В.М., Ядренко М.І.* Теоретико-імовірнісні та статистичні методи в економетриці та фінансовій математиці — К.: Інформтехніка, 1995.

[10] *Справочник по теории вероятностей и математической статистике (под редакцией акад. Королюка В.С.)* — К.: Наукова думка, 1978.

[11] *Бондаренко А.* Управление персоналом по-японски. Персонал, 1996, № 4, стр. 16–28.

До III частини

[1] *Кривцов В.В., Корнев В. П. Саватьев В.А.* Компьютерная техника и программирование II, Операционная среда Microsoft Windows (основные составляющие) – К: НАУ, 1997.

[2] *Кривцов В.В., Корнев В.П. Саватьев В.А., Николайчук В.Й., Саватьева О.В.* Комп'ютерна техніка та програмування, Електронні таблиці Excel, Навчально-методичні матеріали для курсу практичних занять по вивченню Microsoft Excel – К: НАУ, 1998.

[3] *Гончаров А.* Microsoft Excel 97 в примерах – Санкт-Петербург: Питер, 1997. – 336 с.

[4] *Курицкий Б.* Поиск оптимальных решений средствами Excel 7.0 в примерах – Дюссельдорф–Киев–Москва–Санкт-Петербург: ВНУ-Санкт-Петербург, 1997. – 384 с.

НАВЧАЛЬНЕ ВИДАННЯ

В. В. БАРКОВСЬКИЙ

Н. В. БАРКОВСЬКА

О. К. ЛОПАТИН

ТЕОРІЯ ІМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

5-те видання

Керівник видавничих проектів – Б. А. Сладкевич

Друкується в авторській редакції

Дизайн обкладинки – Б. В. Борисов

Підписано до друку 05.09.2009. Формат 60x84 1/16.

Друк офсетний. Гарнітура PetersburgС.

Умовн. друк. арк. 24,65.

Наклад 600 прим.

Видавництво “Центр учбової літератури”

вул. Електриків, 23

м. Київ, 04176

тел./факс 425-01-34, тел. 451-65-95, 425-04-47, 425-20-63

8-800-501-68-00 (безкоштовно в межах України)

e-mail: office@uabook.com

сайт: WWW.CUL.COM.UA

Свідоцтво ДК №2458 від 30.03.2006