

МІНІСТЕРСТВО ОСВІТИ І НАУКИ
ЦЕНТРАЛЬНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

ФАКУЛЬТЕТ БУДІВНИЦТВА І ТРАНСПОРТУ
КАФЕДРА ВИЩОЇ МАТЕМАТИКИ ТА ФІЗИКИ

МЕТОДИ ПРИКЛАДНОЇ МАТЕМАТИКИ В ТРАНСПОРТНИХ ТЕХНОЛОГІЯХ.
КОМБІНАТОРИКА, ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

Методичні рекомендації для студентів спеціальності «Транспортні технології»

КРОПИВНИЦЬКИЙ – 2021

МІНІСТЕРСТВО ОСВІТИ І НАУКИ
ЦЕНТРАЛЬНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

ФАКУЛЬТЕТ БУДІВНИЦТВА І ТРАНСПОРТУ
КАФЕДРА ВИЩОЇ МАТЕМАТИКИ ТА ФІЗИКИ

МЕТОДИ ПРИКЛАДНОЇ МАТЕМАТИКИ В ТРАНСПОРТНИХ ТЕХНОЛОГІЯХ.
КОМБІНАТОРИКА, ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА
Методичні рекомендації для студентів спеціальності «Транспортні технології»

Затверджено
На засіданні кафедри
Вищої математики та фізики
Протокол №2 від 31.08.2021

КРОПИВНИЦЬКИЙ – 2021

Методи прикладної математики в транспортних технологіях. Комбінаторика, теорія ймовірностей та математична статистика. Методичні рекомендації для студентів спеціальності «Транспортні технології» / Філімоніхіна І.І., Гуцул В.І. – Кропивницький: ЦНТУ, 2021 р. – 133 с.

Посібник містить курс лекцій з основних розділів теорії ймовірностей та математичної статистики, питання для самоконтролю, індивідуальні завдання з відповідних розділів та довідково-інформаційні таблиці для розв'язання задач.

Орієнтовано на студентів технічних спеціальностей.

© І.І. Філімоніхіна

© В.І. Гуцул

© ЦНТУ 2021

1. Основні поняття теорії ймовірностей та комбінаторики	7
1.1. Предмет теорії ймовірностей та математичної статистики	7
1.2. Поняття стохастичного експерименту	7
1.3. Простір елементарних подій	8
1.4. Сумісні та несумісні події	8
1.5. Операції над подіями	9
1.6. Основні поняття комбінаторики	11
2. Означення ймовірності події. Аксиоматична побудова теорії ймовірностей .	12
2.1. Класичне означення ймовірності події. Випадки рівно ймовірних наслідків	12
2.2. Статистичне означення ймовірності події. Випадки нерівно ймовірних наслідків	14
2.3. Геометричні ймовірності	15
2.4. Аксиоматична побудова теорії ймовірностей	16
3. Незалежні випадкові події. Умовна ймовірність	17
3.1. Повна група подій	17
3.2. Умовна ймовірність	17
3.3. Формула множення ймовірностей	17
3.4. Формула додавання ймовірностей	18
3.5. Незалежність подій	19
3.6. Формула повної ймовірності	20
3.7. Формула Байеса	20
4. Схема Бернуллі. Граничні теореми	22
4.1. Схема Бернуллі	22
4.2. Випадок непостійної ймовірності появи події в випробуваннях	22
4.3. Наймовірніше число настання подій в схемі Бернуллі	23
4.5. Поняття потоку подій	25
4.6. Локальна теорема Муавра-Лапласа	25
4.7. Інтегральна (глобальна) теорема Муавра-Лапласа	26
5. Випадкова величина і її функція розподілу. Властивості функції розподілу ...	28
5.1. Випадкова величина. Види випадкових величин	28
5.2. Закон розподілу дискретної випадкової величини	28
5.3. Функція розподілу випадкової величини і її властивості	29
5.4. Щільність розподілу ймовірності та її властивості	30
6. Математичне сподівання випадкової величини і його властивості. Дисперсія випадкової величини, її властивості. Середнє квадратичне відхилення	32
6.1. Математичне сподівання випадкової величини та його властивості	32
6.2. Дисперсія випадкової величини і її властивості	33
6.3. Середнє квадратичне відхилення	35
7. Основні розподіли випадкової величини, їх математичні сподівання і дисперсії	36
7.1. Біноміальний розподіл, його математичне сподівання і дисперсія	36
7.2. Розподіл Пуассона	37
7.3. Геометричний розподіл	38

7.4. Гіпергеометричний розподіл (урнова схема)	39
7.5. Рівномірний розподіл	40
7.6. Показниковий розподіл	41
8. Нормальний розподіл	43
8.1. Означення нормального розподілу	43
8.2. Властивості функції Гауса	46
8.3. Ймовірність попадання нормальної випадкової величини в заданий інтервал.....	46
8.4. Функція Лапласа і її властивості	47
9. Нерівність Чебишева. Теорема Чебишева. Центральна гранична теорема	50
9.1. Нерівність Чебишева	50
9.2. Теорема Чебишева	50
9.3. Центральна гранична теорема	53
10. Закон розподілу функції випадкових величин.....	54
10.1. Закон розподілу функції випадкової величини	55
10.2. Закон розподілу лінійної функції	58
10.3. Математичне сподівання і дисперсія функції випадкової величини	59
10.4. Математичне сподівання і дисперсія лінійної функції.....	61
10.5. Математичне сподівання і дисперсія мінімальної із двох величин: випадкової X і не випадкової α	61
10.6. Математичне сподівання і дисперсія максимальної із двох величин: випадкової X і не випадкової α ...	62
10.7. Математичне сподівання і дисперсія модуля функції випадкової величини	63
11. Багатовимірні випадкові величини. Спільна функція розподілу двох випадкових величин. Незалежні випадкові величини	63
11.1. Закон розподілу ймовірності двовимірної випадкової величини	64
11.2. Спільна функція розподілу двох випадкових величин	65
11.3. Щільність спільного розподілу ймовірності неперервної двовимірної випадкової величини	66
11.4. Незалежні випадкові величини.....	67
12. Умовні закони розподілу складових системи дискретних і	67
неперервних випадкових величин. Числові характеристики системи двох випадкових величин.....	67
12.1. Умовний закон розподілу складових системи випадкових величин.....	67
12.2. Кореляційний момент.....	69
12.3. Коефіцієнт кореляції.....	70
13. Вибірковий метод аналізу властивостей генеральної сукупності	72
13.1. Предмет та основні поняття математичної статистики	72
13.2. Вибірковий метод і його основні поняття. Випадкова вибірка, об'єм вибірки	72
13.3. Способи відбору.....	73
13.4. Варіаційний ряд для дискретних і неперервних випадкових величин.....	74
13.5. Полігон і гістограма.....	75
14. Поняття про статистичні оцінки випадкових величин	76
14.1. Емпірична функція розподілу	76
14.2. Найважливіші властивості статистичних оцінок	77
14.3. Вибіркові середнє і дисперсія	78

14.4. Надійність і довірчий інтервал	79
15. Довірчі інтервали для математичного сподівання і дисперсії	81
15.1. Довірчий інтервал для математичного сподівання нормального розподілу при відомій дисперсії.....	81
15.2. Довірчий інтервал для математичного сподівання нормального розподілу при невідомій дисперсії	82
15.3. Довірчий інтервал для оцінки середнього квадратичного відхилення σ нормального розподілу	83
16. Перевірка статистичних гіпотез. Критерій згоди Пірсона про вид розподілу.	85
16.1. Перевірка статистичних гіпотез	85
16.2. Статистичний критерій	87
16.3. Критерій згоди Пірсона про вид розподілу.	89
17.Поняття про кореляційно-регресійний аналіз.....	92
17.1. Поняття про кореляційний аналіз.....	93
17.2. Поняття про регресійний аналіз	95
17.3. Лінійна регресія	97
17.4. Нелінійна регресія	99
ІНДИВІДУАЛЬНІ ЗАВДАННЯ	101
ДОДАТКИ	119
Література.....	133

1. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА КОМБІНАТОРИКИ

1.1. Предмет теорії ймовірностей та математичної статистики

Теорія ймовірностей - спеціальний розділ курсу вищої математики, що вивчає математичні закономірності масових однорідних випадкових подій. Слід особливо підкреслити, що методи теорії ймовірностей за своєю суттю не дають можливості передбачити результат *окремої* випадкової події, але дають можливість *передбачити середній сумарний результат* маси однорідних випадкових подій.

Методи теорії ймовірностей широко використовуються в економіці, в теорії надійності, теорії інформації, теорії масового обслуговування, теорії прийняття рішень, в фізиці, астрономії та інших дисциплінах. Теорія ймовірностей лежить в основі **математичної статистики**, яка, в свою чергу, використовується при плануванні та організації виробництва, при аналізі технологічних процесів, контролі якості продукції та ін.

Математична статистика – наука про математичні методи систематизації та використання статистичних даних для здійснення науково обґрунтованих прогнозів та практичних рекомендацій.

Взагалі кажучи, всі спостережні події (явища) оточуючого нас світу можна поділити на наступні три види: **достовірні**, **неможливі** та **випадкові**.

Достовірною називають подію, яка обов'язково відбудеться, якщо буде виконана певна сукупність умов. *Приклад* – лід плавиться при температурі вище нуля.

Неможливою називають подію, яка точно не відбудеться при виконанні певної сукупності умов. *Приклад* – лід не може існувати при 100 градусах Цельсія, Земля не може без впливу зовні припинити своє обертання.

Випадковою називають подію, яка при виконанні сукупності умов може відбутися або не відбутися. *Приклад* – випадання певного числа очок при киданні грального кубика, попадання снаряду в ціль, вихід з ладу технічного пристрою, отримання певного прибутку фірмою і т.д.

1.2. Поняття стохастичного експерименту

Експериментом називається реалізація наміченої дії, що приводить до деякого результату. Експерименти поділяються на детерміновані та стохастичні (випадкові).

Експеримент називається **детермінованим**, якщо, виходячи з умов, що описують експеримент, його результат передбачуваний.

Приклади: 1) камінь, підкинутий вгору, обов'язково впаде вниз; 2) підвищення життєвого рівня викликає зростання споживання товарів; 3) поломка системного блоку виводить з ладу комп'ютер.

Експеримент вважається **стохастичним (випадковим)**, якщо він може закінчитися будь-яким з деякої сукупності відомих результатів, але до здійснення експерименту не можна сказати, яким саме.

Теорія ймовірностей досліджує саме стохастичні експерименти, вірніше моделі **експериментів з випадковими наслідками**. При цьому розглядаються тільки такі експерименти, які можна повторювати (відтворювати) при незмінному комплексі умов довільну кількість разів (принаймні теоретично).

Будемо розглядати **подію** як результат експерименту.

Приклади: 1) стрілець стріляє по мішені, яка поділена на декілька частин. Постріл – це експеримент, попадання в певну частину мішені – подія. 2) діставання кулі з урни – експеримент, поява кулі певного кольору – подія. 3) здача екзамену – експеримент, отримання оцінки – подія.

1.3. Простір елементарних подій

Нехай в результаті експерименту настає одна і тільки одна з подій $\omega_i (i = 1, 2, \dots, n)$

Події ω_i називають **елементарними подіями**.

Простором елементарних подій Ω називають множину всіх елементарних подій, які можуть з'явитися в експерименті.

Точками простору Ω називаються самі елементарні події.

Простір елементарних подій зазвичай вважається заданим, якщо вказані всі його елементи.

Приклад. Для експерименту з підкиданням грального кубика простір елементарних подій утворює сукупність елементарних подій $\{1, 2, 3, 4, 5, 6\}$; при підкиданні монети $\{\text{Ц}, \text{Г}\}$.

З елементарних подій можна скласти більш складну подію. Іншими словами, кожна випадкова подія A визначається як підмножина в множині елементарних подій Ω .

При цьому ті елементарні наслідки з Ω , при яких подія A настає (тобто належить підмножині A), називають **сприяючими** події A . Кажуть, що подія A відбулася, якщо результатом експерименту стала елементарна подія ω , що належить A ($\omega \in A$).

Приклад. При підкиданні грального кубика події “випадання парної кількості очок” сприяють елементарні події $\{2, 4, 6\}$. Здачі екзамену сприяє отримання 3, 4 чи 5 балів.

1.4. Сумісні та несумісні події

Дві події називаються **сумісними** в даному експерименті, якщо поява одного з них не виключає появу іншої.

Приклади: 1) попадання в ціль двома різними стрільцями, 2) випадання однакового числа очок на двох кубиках.

Дві події називаються **несумісними** в даному експерименті, якщо вони не можуть відбутися разом при одному і тому ж експерименті.

Декілька подій називаються **несумісними**, якщо вони попарно несумісні.

Приклади: 1) попадання і промах при одному пострілі; 2) з ящика з деталями навмання дістали деталь – події “дістали стандартну деталь” та “дістали нестандартну деталь” 3) розорення фірми та отримання нею прибутку.

Іншими словами, події A і B сумісні, якщо відповідні множини A і B мають спільні елементи і несумісні, якщо відповідні множини A і B не мають спільних елементів.

При визначенні ймовірностей подій часто використовується поняття рівноможливих подій.

Декілька подій в даному експерименті називаються **рівновоможливими**, якщо за умовами симетрії є підстава вважати, що жодна з них об'єктивно не є більш можливою, ніж інші.

Приклади: 1) випадання герба і цифри, 2) поява карти будь-якої масті, 3) вибір кулі з урни тощо.

1.5. Операції над подіями

З кожним експериментом пов'язаний ряд подій, які, взагалі кажучи, можуть з'явитися одночасно.

Приклад. При киданні грального кубика подія A - випадання двійки, а подія B - випадання парного числа очок. Очевидно, що ці події не виключають одна одну.

Нехай всі можливі результати випробування здійснюються в ряді єдино можливих частинних випадків, що взаємно виключають один одного. Тоді:

- кожний наслідок випробування представляється однією і тільки однією елементарною подією;
- всяка подія A , пов'язана з цим випробуванням, є множина скінченного чи нескінченного числа елементарних подій;
- подія A відбувається тоді і тільки тоді, коли реалізується одна з елементарних подій, які належать цій множині.

Довільний, але фіксований простір елементарних подій Ω , можна представити у вигляді деякої області на площині. При цьому елементарні події $\omega_1, \omega_2, \dots, \omega_n, \dots$ - це точки площини, що лежать всередині Ω . Так як подія ототожнюється з множиною, то над подіями можна здійснювати всі операції, що здійснюються над множинами.

За аналогією з теорією множин будується **алгебра подій**. При цьому можуть бути визначені наступні операції та співвідношення між подіями:

- 1) $A \subset B$ (відношення включення множин: множина A є підмножиною множини B) - з події A випливає подія B . Іншими словами, подія B відбувається щоразу, коли відбувається подія A .

Приклад - з випадіння двійки випливає випадіння парного числа очок.

- 2) $A = B$ (відношення еквівалентності множин)

Подія A **тотожня** або **еквівалентна** події B , якщо $A \subset B$ і одночасно $B \subset A$, тобто кожна з них відбувається щоразу, коли відбувається інша.

Приклад. Подія A - поломка приладу, подія B - поломка хоча б одного з блоків (деталей) приладу.

- 3) $A + B$ ($A \cup B$) - **сума подій**.

Сума подій - це подія, що полягає в тому, що відбулася хоча б одна з двох подій A або B (логічне "або").

В загальному випадку, під сумою декількох подій розуміють подію, що полягає в появі хоча б однієї з цих подій.

Приклад. Ціль вражена першим стрільцем, другим або обома одночасно.

- 4) AB ($A \cap B$) - **добуток подій**.

Добуток подій - це подія, що полягає в сумісному здійсненні подій A і B (логічне "і").

В загальному випадку, під добутком декількох подій розуміють подію, що полягає в одночасному здійсненні всіх цих подій.

Таким чином, події A і B несумісні, якщо добуток їх є подія неможлива, тобто $AB = \emptyset$.

Приклад. Подія A – виймання з колоди карти бубнової масті, подія B – виймання туза, тоді AB – поява бубнового туза.

5) $A - B$ – *різниця подій*.

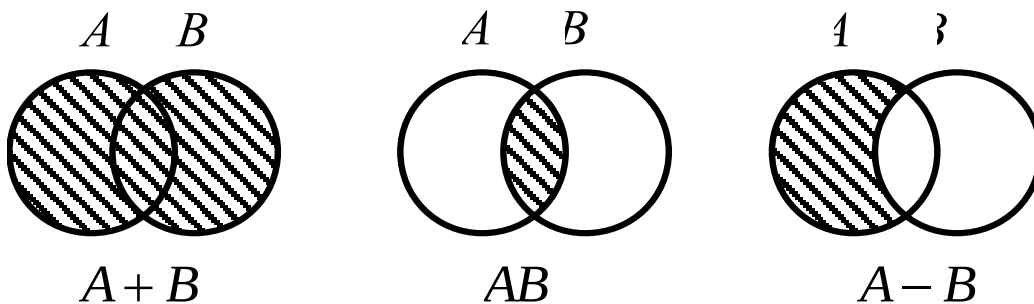
Різниця подій – це подія, що складається з наслідків, які входять в A , але не входять в B . Вона полягає в тому, що відбувається подія A , але при цьому не відбувається подія B .

Приклад. A – здача екзаменаційної сесії, B – отримання стипендії, тоді $A - B$ – здача сесії з недостатньо високим для отримання стипендії результатом.

Протилежною подією (доповняючою) для події A (позначається $\bar{A}$) називається подія, що складається з усіх тих наслідків, які не входять в A .

Настання події $\bar{A}$ означає просто, що подія A не настала.

Часто буває корисною геометрична інтерпретація операцій над подіями. Графічна ілюстрація операцій називається діаграмами Венна.



Властивості операцій над подіями.

Деякі властивості операцій над подіями постулюються, інші легко можна отримати за допомогою діаграм Венна. Наведемо без доведення основні з цих властивостей.

1. $A + B = B + A$
2. $AB = BA$
3. $A + \bar{A} = \Omega$
4. $A\Omega = A$
5. $AB \subset A$
6. $A\bar{A} = \emptyset$
7. $\bar{\bar{A}} = A$
8. $A - B = A\bar{B}$
9. $(A + B)C = AC + BC$

10. $\overline{A+B} = \overline{A} \cdot \overline{B}$
11. $\overline{AB} = \overline{A} + \overline{B}$
12. $(A-B)C = AC - BC$

1.6. Основні поняття комбінаторики

При розв'язанні ряду теоретичних та практичних задач потрібно з скінченної множини елементів за заданими правилами складати різні сполуки та підраховувати кількість всіх можливих таких сполук.

Комбінаторика – це розділ математики, що вивчає розташування об'єктів у відповідності зі спеціальними правилами і методи підрахунку кількості всіх можливих способів, якими ці розташування можна здійснити.

Правила суми та добутку

Правило суми – якщо елемент a може бути вибраний n способами, а елемент b – m способами, то один з цих елементів можна вибрати $n+m$ способами.

Правило добутку – якщо елемент a може бути вибраний n способами та після кожного такого вибору елемент b можна вибрати m способами, то пару (ab) з цих елементів у вказаному порядку можна вибрати nm способами.

Розміщеннями з n елементів по k називаються впорядковані набори, що складаються з k різних елементів, вибраних з n даних елементів, .

Розміщення можуть відрізнятися як елементами, так і порядком.

Теорема. Число всіх розміщень з n елементів по k обчислюється за формулою:

$$A_n^k = n(n-1)(n-2)\dots(n-(k-1)) = \frac{n!}{(n-k)!}$$

Доведення. Дійсно, перший елемент розміщення може бути вибраний n способами. Для кожного з цих варіантів є $n-1$ способів розташування одного з елементів, що залишились, на другому місці. Отже, за правилом добутку, маємо $n \cdot (n-1)$ різних способів вибору елементів на перших двох місцях. Продовжуючи це суждення, отримуємо потрібну формулу.

Приклад: Різними розміщеннями множини з трьох елементів $\{1,2,3\}$ по два будуть набори $(1,2)$, $(2,1)$, $(1,3)$, $(3,1)$, $(2,3)$, $(3,2)$

Перестановками P_n . називаються розміщення при $k=n$.

Так як кожна перестановка містить всі n елементів множини, то різні перестановки відрізняються одна від одної лише порядком елементів і

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

Приклад: Різними перестановками множини елементів $\{1,2,3\}$ будуть $(1,2,3)$, $(1,3,2)$, $(2,3,1)$, $(2,1,3)$, $(3,2,1)$, $(3,1,2)$

Сполученнями (комбінаціями) з n елементів по k називаються неупорядковані набори з k елементів, взятих з даних n елементів,

Теорема. Число сполучень з n елементів по k обчислюється за формулою

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

Доведення можна отримати, враховуючи, що сполучення відрізняються від розміщень тим, що в них не важливий порядок розташування заданих k елементів. Тому при рівних n і k число сполучень менше числа розміщень в $k!$ разів.

Питання для самоперевірки

1. Що вивчає теорія ймовірності?
2. Яка подія називається достовірною?
3. Яка подія називається неможливою?
4. Яка подія називається випадковою?
5. Що таке стохастичний експеримент?
6. Що називають простором елементарних подій?
7. Які події називаються сумісними?
8. Які події називаються несумісними?
9. Що таке рівноможливі події?
10. Перерахувати операції над подіями.
11. Записати властивості операцій над подіями.
12. Що вивчає комбінаторика?
13. Сформулювати правила суми та добутку.
14. Записати формулу для розміщень.
15. Записати формулу для перестановок.
16. Записати формулу для сполучень.

2. ОЗНАЧЕННЯ ЙМОВІРНОСТІ ПОДІЇ. АКСІОМАТИЧНА ПОБУДОВА ТЕОРІЇ ЙМОВІРНОСТЕЙ

Ймовірність є кількісною мірою можливості появи події. Найбільш широке застосування отримали два означення ймовірності події: класичне та статистичне.

2.1. Класичне означення ймовірності події. Випадки рівно ймовірних наслідків

Класичне означення ймовірності пов'язане з означенням сприятливого наслідку.

Наслідок називається **сприятливим** даній події, якщо з його появи впливає настання цієї події.

Ймовірність події дорівнює відношенню числа всіх рівно можливих елементарних наслідків випробування, сприяючих даній події, до всіх рівно можливих елементарних наслідків:

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

де m – число сприятливих події A наслідків;

n – загальна кількість можливих наслідків.

Приклади.

- 1) Яка ймовірність випадіння двійки при киданні кубика?

Розв'язання. Всього наслідків випробування 6, число сприятливих наслідків 1, тому

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

2) Яка ймовірність того, що в довільному двозначному числі дві цифри однакові.

Розв'язання. Всього наслідків випробування $A_{10}^2 = 9 \cdot 10 = 90$, число сприятливих наслідків 9, тому $P(A) = \frac{9}{90}$.

3) З букв слова “диференціал” вибирається одна буква. Яка ймовірність того, що це а) голосна, б) буква “ф”.

Розв'язання. а) всього наслідків випробування 11, число сприятливих наслідків 5, тому $P(A) = \frac{5}{11}$.

б) всього наслідків випробування 11, число сприятливих наслідків 1, тому $P(A) = \frac{1}{11}$.

З означення ймовірності події A випливає, що $0 \leq m \leq n$, тому завжди виконуються нерівності $0 \leq P(A) \leq 1$, тобто ймовірність будь-якої події є невід'ємне число, що не перевищує одиницю.

- Якщо $P(A) = 0$, то подія A неможлива.
- Якщо $P(A) = 1$, то подія A достовірна.
- Рівноможливі елементарні події є **рівноймовірними**, тобто мають однакову ймовірність.

Теорема. Еквівалентні події мають однакові ймовірності, тобто якщо $A = B$, то $P(A) = P(B)$.

Доведення. Дійсно, кожний елементарний наслідок події A є таким ж елементарним наслідком для події B і навпаки. В силу формули $P(A) = \frac{m}{n}$ справедлива рівність $P(A) = P(B)$

Якщо подія B відбувається щоразу після того, як відбулася подія A , то кажуть, що з події A випливає подія B ($A \Rightarrow B$). Наприклад, для будь-яких двох подій A і B справедливо $AB \Rightarrow A$ і $AB \Rightarrow B$.

Теорема. Якщо $A \Rightarrow B$, то $P(A) \leq P(B)$.

Доведення. Нехай m і m' – число сприяючих елементарних наслідків відповідно для подій A і B , а n – загальне число елементарних наслідків.

Так як кожний елементарний наслідок для події A є також елементарним наслідком для події B

$$\text{то } m \leq m' \text{ і, отже, } P(A) = \frac{m}{n} \leq \frac{m'}{n} = P(B).$$

Приклад. Випадіння парного числа очок більш ймовірне, ніж випадіння двійки.

Теорема. Ймовірність події $\bar{A}$, протилежної до події A дорівнює

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

Доведення. Нехай повна система рівноможливих елементарних наслідків містить n подій, з яких m ($m \leq n$) сприяють події A . Тоді $n - m$ наслідків несприятливі події A , тобто сприяють події $\bar{A}$. Таким чином,

$$P(\bar{A}) = \frac{n - m}{n} = 1 - \frac{m}{n} = 1 - P(A).$$

Класичне означення ймовірності припускає що:

- число елементарних наслідків скінченне;
- ці наслідки рівноможливі.

Однак на практиці зустрічаються випробування з нескінченним числом можливих наслідків. Крім того, немає загальних методів, які дозволяють результат випробування, навіть з скінченним числом наслідків, представити у вигляді суми рівноможливих елементарних наслідків. Тому застосування класичного означення ймовірностей обмежене.

Приклад. Кубик зі зміщеним центром ваги.

2.2. Статистичне означення ймовірності події. Випадки нерівно ймовірних наслідків

Класичне означення ймовірності має обмежене застосування. Так, воно неприйнятне, якщо результати випробування не рівноможливі.

В багатьох випадках більш зручним є **статистичне означення ймовірності**, яке пов'язане з поняттям відносної частоти появи події A в випробуваннях.

Відносна частота появи події A – це відношення числа m появи події A в серії з n випробувань до числа випробувань:

$$P^*(A) = \frac{m}{n}.$$

Досвід показує, що при проведенні порівняно невеликого числа випробувань відносна частота $P^*(A)$ приймає значення, які можуть дуже відрізнятися одне від одного. При однотипових **масових** випробуваннях в багатьох випадках спостерігається стійкість відносної частоти події, тобто зі збільшенням числа випробувань відносна частота коливається навколо деякого постійного $P(A)$, причому ці відхилення тим менші, чим більше проведено випробувань.

Ймовірністю події A в статистичному значенні називається число $P(A)$, відносно якого стабілізується (встановлюється) відносна частота $P^*(A)$ при необмеженому збільшенні числа випробувань.

Тому, на практиці за ймовірність події A приймається відносна частота $P^*(A)$ при достатотньо великому числі випробувань.

Властивості ймовірності, що випливають з класичного означення ймовірності, зберігаються і при статистичному означенні ймовірності.

Якщо ймовірність деякої події близька до нуля, то, в відповідності зі сказаним, впливає, що при одиничному випробуванні в більшості випадків така подія не настане. Виникає питання:

наскільки малою повинна бути ймовірність, щоб можна було знехтувати ймовірністю настання деякої події в одиничному випробуванні (наприклад, землетрус в Кіровограді)?

Рівнем значущості називають достатньо малу ймовірність, при якій настання події можна вважати практично неможливим.

На практиці рівень значущості зазвичай приймають рівним 0,05 (пятивідсотковий рівень) або 0,01 (одновідсотковий рівень).

2.3. Геометричні ймовірності

Щоб подолати недолік класичного означення ймовірності, пов'язаний з його незастосовністю до випробувань з нескінченним числом наслідків, вводять поняття **геометричної ймовірності**.

Геометричною ймовірністю називають ймовірність попадання точки в деяку область (відрізок, частину площини тощо).

В подібних випадках простір елементарних наслідків може бути представлений областю G а під подією A можна розуміти наслідки, що входять в деяку область g , яка належить області G . Нехай на область G навмання кидається “точка”. Яка ймовірність того, що ця точка попаде в область g , що є частиною області G ?

1. Нехай відрізок g довжини l_g , складає частину відрізка G , довжина якого l_G . На відрізок G навмання поставлена точка.

Припускається, що

- поставлена точка може опинитися в будь-якій точці відрізка G ;
- ймовірність попадання точки на відрізок g пропорційна довжині цього відрізка та не залежить від його розташування відносно відрізка G .

Тоді ймовірність попадання точки на відрізок g визначається рівністю

$$P = \frac{l_g}{l_G}.$$

2. Нехай плоска фігура g з площею S_g складає частину плоскої фігури G , площа якої S_G . На фігуру G навмання кинута точка.

Припускається, що:

- Кинута точка може опинитися в будь-якій точці фігури G ;
- ймовірність попадання кинutoї точки на фігуру g пропорційна площі цієї фігури і не залежить ні від її розташування відносно фігури G , ні від форми g .

При цих припущеннях ймовірність попадання точки на фігуру g визначається рівністю

$$P = \frac{S_g}{S_G}.$$

3. Аналогічно вводиться поняття геометричної ймовірності при киданні точки в просторову область об'єму V_G , що містить область g об'єму V_g :

$$P = \frac{V_g}{V_G}$$

В загальному випадку поняття геометричної ймовірності вводиться наступним чином. Позначимо міру області g (довжину, площу, об'єм тощо) через $mes(g)$, а міру області G – через $mes(G)$. Тоді ймовірність попадання в область g точки, кинутої в область G , визначається формулою:

$$P(A) = \frac{mes(g)}{mes(G)}.$$

Приклад: протягом доби до причалу можуть підійти 2 пароходи. Час прибуття обох пароходів незалежний і рівноможливий протягом доби. Визначити ймовірність того, що одному з пароходів доведеться чекати звільнення причалу, якщо час розвантаження одного з них дорівнює 1 годині, а іншого – 2 години.

Розв'язання. Застосуємо геометричну ймовірність. Розглянемо прямокутну декартову систему координат xOy , в якій x та y будемо відраховувати в годинах від 0 до 24. Нехай x – час прибуття першого пароплаву, y – час прибуття другого пароплаву. Тоді всі можливі комбінації прибуття пароплавів до причалу зобразяться точками квадрата, для якого $0 \leq x \leq 24$. Очевидно, що положення точок (x, y) в області цього квадрата рівноможливі.

З'ясуємо, які точки (x, y) сприяють події A (один з пароплавів чекає звільнення причалу). Подія A може відбутися лише в тому випадку, якщо момент y прибуття другого пароплаву не більше, ніж на дві години раніше моменту x прибуття першого пароплаву і не більше, ніж на одну годину пізніше прибуття першого пароплаву:

$$x-2 \leq y \leq x+1.$$

Таким чином, область квадрата, сприяюча події A , складається з точок, координати (x, y) яких задовольняють нерівностям, тобто з точок, що лежать між прямими $y=x-2$ та $y=x+1$. Площа квадрата дорівнює $S_G = 24^2 = 576$. Площа меншої області дорівнює $S_g = 576 - \frac{1}{2} \cdot 23^2 - \frac{1}{2} \cdot 22^2 = 69,5$.

$$\text{Звідси } P(A) = \frac{69,5}{576} \approx 0,121.$$

2.4. Аксиоматична побудова теорії ймовірностей

Побудова логічно повноціної теорії ймовірностей основана на аксіоматичному визначенні випадкової події та її ймовірності. В системі аксіом, запропонованій А.Н. Колмогоровим, елементарна подія і ймовірність є поняттями, що не визначаються. Наведемо аксіоми системи Колмогорова

1. Кожній події A поставлено у відповідність невід'ємне дійсне число $P(A)$. Це число називається ймовірністю події A ; $0 \leq P(A_i) \leq 1$
2. Ймовірність достовірної події дорівнює одиниці $P(\Omega) = 1$, де Ω - достовірна подія.
3. Ймовірність настання хоча б однієї з попарно несумісних подій дорівнює сумі ймовірностей цих подій;

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad (\text{скінченний простір елементарних подій})$$

$$P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad (\text{нескінченний простір елементарних подій})$$

Виходячи з цих аксіом, властивості ймовірностей та залежності між ними виводять в якості теорем.

Питання для самоперевірки

1. Що таке сприятливий наслідок для даної події?
2. Сформулювати означення класичної ймовірності.
3. Записати основні властивості ймовірності.
4. Що називають рівноймовірними подіями?
5. Чому дорівнює ймовірність протилежної події?
6. Що таке відносна частота події?
7. Сформулювати статистичне означення ймовірності.
8. Що таке рівень значущості?
9. Коли застосовують геометричну ймовірність?
10. Записати загальну формулу для геометричної ймовірності.
11. Записати аксіоми Колмогорова.

3. НЕЗАЛЕЖНІ ВИПАДКОВІ ПОДІЇ. УМОВНА ЙМОВІРНІСТЬ

3.1. Повна група подій

Повною групою подій називається множина попарно несумісних подій, для якої при будь-якому наслідку випадкового випробування обов'язково настає одна з подій, що входить в цю множину.

Іншими словами, для повної групи подій виконані наступні умови :

- поява однієї з подій даної множини в результаті випробування є достовірною подією, тобто подія $A_1 + A_2 + \dots + A_n = \Omega$;
- події A_i та A_j ($i \neq j$) попарно несумісні і $A_i \cdot A_j$ - подія неможлива для будь-яких $i \neq j$, тобто $A_i \cdot A_j = \emptyset$

Найпростішим прикладом повної групи подій є пара протилежних подій A та $\bar{A}$.

Теорема. Сума ймовірностей повної групи подій $A_1, A_2, \dots, A_n$ дорівнює одиниці :

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

3.2. Умовна ймовірність

У багатьох випадках ймовірності появи одних подій залежать від того, відбулася інша подія чи ні.

Умовною ймовірністю події A називається ймовірність події A , обчислена за умови, що подія B вже відбулася, і позначається $P(A/B) = P_B(A)$.

У тих випадках, коли ймовірність події A розглядається за умови, що мали місце дві інші події B і C , використовується умовна ймовірність відносно добутку подій B і C : $P(A/BC)$.

3.3. Формула множення ймовірностей

Теорема. Ймовірність добутку двох подій дорівнює добутку ймовірностей однієї з них на умовну ймовірність іншої, обчислену за умови, що перша подія відбулася:

$$P(AB) = P(A)P(B/A) = P(B)P(A/B).$$

Теорему множення ймовірності легко узагальнити на будь-яке скінченне число подій.

Теорема. Ймовірність добутку скінченного числа подій дорівнює добутку їх умовної ймовірності відносно добутку попередніх подій :

$$P(ABC...LM) = P(A)P(B/A)P(C/AB)...P(M/AB...L).$$

Для доведення цієї теореми можна використовувати метод математичної індукції.

3.4. Формула додавання ймовірностей

Теорема. Ймовірність суми скінченного числа несумісних подій дорівнює сумі ймовірностей цих подій :

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

Доведення. Доведемо цю теорему для випадку суми двох несумісних подій A_1 і A_2 .

Нехай події A_1 сприяють m_1 елементарних наслідків, а події A_2 - відповідно m_2 наслідків. Оскільки події A_1 і A_2 за умовою теореми несумісні, то події $A_1 + A_2$ сприяють $m_1 + m_2$ елементарних наслідків із загального числа наслідків. Отже,

$$P(A_1 + A_2) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A_1) + P(A_2),$$

де $P(A_1)$ - ймовірність події A_1 ;

$P(A_2)$ - ймовірність події A_2 .

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

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

Доведення. Подія $A + B$, відбудеться, якщо відбудеться одна з несумісних подій $\overline{AB}$, $\overline{AB}$, AB .

За теоремою про додавання ймовірностей несумісних подій

$$P(A + B) = P(\overline{AB}) + P(\overline{AB}) + P(AB)$$

Подія A відбудеться, якщо відбудеться одна з двох несумісних подій $\overline{AB}$, AB .

Знову застосовуючи теорему додавання ймовірностей несумісних подій, отримуємо:

$$P(A) = P(\overline{AB}) + P(AB).$$

Отже, $P(\overline{AB}) = P(A) - P(AB)$.

Аналогічно для події B отримуємо $P(B) = P(\overline{AB}) + P(AB)$

Звідки $P(\overline{AB}) = P(B) - P(AB)$.

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

3.5. Незалежність подій

Події A і B називаються незалежними, якщо при настанні події A ймовірність події B не змінюється.

Теорема. Ймовірність спільної появи двох незалежних подій A і B (добутку A і B) дорівнює добутку ймовірностей цих подій.

Доведення. Дійсно, оскільки події A і B незалежні, то $P(B/A) = P(B)$. В цьому випадку формула ймовірності добутку подій A і B набирає вигляду

$$P(AB) = P(A) \cdot P(B).$$

Події $A_1, A_2, \dots, A_n$ називаються попарно незалежними, якщо незалежні будь-які дві з них.

Події $A_1, A_2, \dots, A_n$ називаються незалежними в сукупності, якщо незалежні кожні дві з них і незалежні кожна подія і усі можливі добутки інших.

Теорема. Ймовірність добутку скінченного числа незалежних в сукупності подій дорівнює добутку ймовірностей цих подій.

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

Проілюструємо відмінність в застосуванні формул ймовірності добутку подій для залежних і незалежних подій на прикладах.

Приклад 1. Ймовірність попадання в ціль першим стрільцем дорівнює 0,85, другим 0,8. Стрільці зробили по одному пострілу. Яка ймовірність того, що в ціль потрапив хоча б один стрілець?

Розв'язання. $P(A+B) = P(A) + P(B) - P(AB)$. Оскільки постріли незалежні, то

$$P(A+B) = P(A) + P(B) - P(A) \times P(B) = 0,97$$

Приклад 2. У урні знаходиться 2 червоних і 4 чорних кулі. З неї виймають підряд 2 кулі. Яка ймовірність того, що обидві кулі червоні.

Розв'язання. 1 випадок. Подія A - поява червоної кулі при першому вийманні, подія B - при другому. Подія C - поява двох червоних куль.

$$P(C) = P(A) \times P(B/A) = (2/6) \times (1/5) = 1/15$$

2 випадок. Перша вийнята куля повертається в урну.

$$P(C) = P(A) \times P(B) = (2/6) \times (2/6) = 1/9$$

3.6. Формула повної ймовірності

Нехай подія A може статися тільки з однією з несумісних подій $H_1, H_2, \dots, H_n$, що утворюють повну групу. Наприклад, в магазин поступає одна і та ж продукція від трьох підприємств і в різній кількості. Ймовірність випуску неякісної продукції на цих підприємствах різна. Випадковим чином відбирається один з виробів. Необхідно визначити ймовірність того, що це виріб неякісний (подія A). Тут події H_1, H_2, H_3 - це вибір виробу з продукції відповідного підприємства.

В цьому випадку ймовірність події A можна розглядати як суму добутків подій

$$A = \sum_{i=1}^n AH_i.$$

За теоремою додавання ймовірностей несумісних подій отримуємо $P(A) = \sum_{i=1}^n P(AH_i)$

Використовуючи теорему множення ймовірностей, знаходимо

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

Отримана формула називається формулою повної ймовірності.

3.7. Формула Байеса

Нехай подія A відбувається одночасно з однією з несумісних подій $H_1, H_2, \dots, H_n$, ймовірність яких $P(H_i)$ ($i = 1, 2, \dots, n$) відома до випробування (ймовірність апіорі). Проводиться випробування, в результаті якого зареєстрована поява події, причому відомо, що ця подія мала певну умовну ймовірність $P(A/H_i)$ ($i = 1, 2, \dots, n$). Необхідно знайти ймовірність подій H_i , якщо відомо, що подія A відбулася (ймовірність апостеріорі).

Завдання полягає в тому, що, маючи нову інформацію (подія A відбулася), треба переоцінити ймовірність подій $H_1, H_2, \dots, H_n$.

На підставі теореми про ймовірність добутку двох подій

$$P(H_i A) = P(A)P(H_i / A) = P(H_i)P(A / H_i),$$

звідки

$$P(H_i / A) = \frac{P(H_i)P(A / H_i)}{P(A)}$$

чи

$$P(H_i / A) = \frac{P(H_i)P(A / H_i)}{\sum_{i=1}^n P(H_i)P(A / H_i)}.$$

Отримана формула носить назву формули Байеса.

Приклад. Для підготовки до змагань були відібрані 4 курсанти першого взводу, 9 курсантів другого взводу та 7 курсантів третього взводу. Ймовірність виграти змагання для

курсанту першого взводу дорівнює 0,7; для другого - 0,8; для третього – 0,9. Виявилося, що навмання вибраний курсант став переможцем змагань. З якого взводу ймовірніше всього був цей курсант?

Розв'язування. Введемо позначення подій:

A - курсант переміг в змаганнях; H_1 – курсант першого взводу; H_2 – курсант другого взводу; H_3 – курсант третього взводу. Тоді

$$P(H_1) = \frac{4}{4+9+7} = \frac{4}{20} = 0,2$$

$$P(H_2) = \frac{9}{4+9+7} = \frac{9}{20} = 0,45$$

$$P(H_3) = \frac{7}{4+9+7} = \frac{7}{20} = 0,35$$

Якщо A – подія, що полягає в тому, що навмання вибраний курсант переміг в змаганнях, то $P(A)$

знаходимо за формулою повної ймовірності: $P(A) = \sum_{i=1}^n P(H_i)P(A/H_i)$

В нашому випадку

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + P(H_3)P(A/H_3)$$

За умовою

$$P(A/H_1) = 0,7$$

$$P(A/H_2) = 0,8$$

$$P(A/H_3) = 0,9$$

Отже,

$$P(A) = 0,7 \cdot 0,2 + 0,8 \cdot 0,45 + 0,9 \cdot 0,35 = 0,815$$

Вибраний навмання курсант переміг в змаганнях. Необхідно визначити ймовірність того, що він з першого, другого або третього взводу.

а) ймовірність того, що курсант з першого взводу. Для цього скористаємося формулою Байеса

$$P(H_i/A) = \frac{P(H_i)P(A/H_i)}{P(A)}$$

В нашому випадку

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{0,2 \cdot 0,7}{0,2 \cdot 0,7 + 0,8 \cdot 0,45 + 0,9 \cdot 0,35} = \frac{0,14}{0,815} = 0,1718$$

б) ймовірність того, що курсант з другого взводу.

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{0,45 \cdot 0,8}{0,2 \cdot 0,7 + 0,8 \cdot 0,45 + 0,9 \cdot 0,35} = 0,4417$$

в) ймовірність того, що курсант з третього взводу.

$$P(H_3/A) = \frac{P(H_3)P(A/H_3)}{P(A)} = \frac{0,35 \cdot 0,9}{0,2 \cdot 0,7 + 0,8 \cdot 0,45 + 0,9 \cdot 0,35} = 0,3865$$

Отже, ймовірніше всього курсант був з другого взводу, так як ця ймовірність більша двох інших.

Питання для самоперевірки

1. Що таке повна група подій?
2. Чому дорівнює сума ймовірностей для повної групи подій?
3. Що називають умовною ймовірністю події A ?
4. Записати формулу множення ймовірностей.
5. Записати формулу додавання ймовірностей для несумісних подій.
6. Записати формулу додавання ймовірностей для сумісних подій.
7. Записати формулу множення ймовірностей для незалежних подій.
8. Записати формулу повної ймовірності.
9. Записати формулу Байєса.

4. СХЕМА БЕРНУЛЛІ. ГРАНИЧНІ ТЕОРЕМИ

4.1. Схема Бернуллі

Схемою Бернуллі називається серія повторних незалежних випробувань, в кожному з яких ця подія A має одну і ту ж ймовірність $P(A) = p$, не залежну від номера випробування.

Таким чином, в схемі Бернуллі для кожного випробування є тільки два результати: подія A (успіх), ймовірність якої $P(A) = p$ і подія $\bar{A}$ (невдача), ймовірність якої $P(\bar{A}) = q = 1 - p$.

Ймовірність, що подія A наступить в m випробуваннях, визначається за формулою Бернуллі

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

Приклад. Ймовірність враження цілі стрільцем при одному пострілі дорівнює 0,8. Знайти ймовірність, що при 10 пострілах ціль вражається 8 разів.

Розв'язання.
$$P_{10}(8) = \frac{10!}{2! 8!} \cdot (0,8)^8 \cdot (0,2)^2 = 0,302.$$

4.2. Випадок непостійної ймовірності появи події в випробуваннях

Ми припускали, що ймовірність настання події в кожному з випробувань постійна. На практиці часто доводиться зустрічатися із складнішим випадком, коли випробування проводяться в неоднакових умовах, і ймовірність події від випробування до випробування змінюється. Наприклад, проводиться серія пострілів при дальності, що змінюється.

Спосіб обчислення ймовірності заданого числа появ подій в таких умовах дає загальна теорема про повторення випробувань.

Нехай проводиться n незалежних випробувань, в кожному з яких може з'явитися або не з'явитися деяка подія A , причому ймовірність появи цієї події в i -му випробуванні дорівнює p_i , а ймовірність її неяви відповідно $q_i = 1 - p_i$ ($i=1, \dots, n$). Потрібно знайти ймовірність $P_{m,n}$ того, що в результаті n випробувань подія A з'явиться рівно m разів.

Розв'язання цієї задачі проводиться за допомогою так званої твірної функції, що має вигляд:

$$\varphi_n(z) = \prod_{i=1}^n (q_i + p_i z).$$

Для того, щоб знайти ймовірність появи події A рівно m разів в серії n випробувань, досить перемножити співмножники у твірній функції. Коефіцієнт при члені z^m і дасть шукану ймовірність.

Приклад. Проводиться 4 незалежні постріли по одній і тій же цілі з різних відстаней. Ймовірність попадання при цих пострілах дорівнює відповідно $p_1=0,1$ $p_2=0,2$ $p_3=0,3$ $p_4=0,4$. Знайти ймовірність трьох попадань.

Розв'язання: Складемо твірну функцію

$$\begin{aligned}\varphi_n(z) &= \prod_{i=1}^n (q_i + p_i z) = (0,9 + 0,1z) \cdot (0,8 + 0,2z) \cdot (0,7 + 0,3z) \cdot (0,6 + 0,4z) = \\ &= 0,302 + 0,440z + 0,215z^2 + 0,040z^3 + 0,002z^4.\end{aligned}$$

Звідси ймовірність трьох попадань дорівнює 0,040. Легко знайти і ймовірність жодного, одного, двох і чотирьох попадань, виписуючи коефіцієнти при z^0 , z^1 , z^2 і z^4 .

4.3. Наймовірніше число настання подій в схемі Бернуллі

Число настання події A називається **наймовірнішим**, якщо воно має найбільшу ймовірність в порівнянні з ймовірністю настання A будь-якої іншої кількості разів.

Теорема. Наймовірніше число настання події A в n незалежних випробуваннях знаходиться між числами $np - q$ і $np + p$.

Слід зазначити, що наймовірніших чисел може бути два або одне залежно від того, є $np + p$ цілим числом або ні. Якщо це число неціле, то наймовірніше число $m_0 = [np + p]$ (ціла частина), інакше є два значення $m'_0 = np - q$ і $m''_0 = np + p$

Приклад. Скільки разів потрібно кинути гральний кубик, щоб наймовірніше число появ трійки дорівнювало 55?

Розв'язання. За умовою задачі $m_0=55$, ймовірність появи трійки дорівнює $p = \frac{1}{6}$, отже, ймовірність не появи трійки дорівнює $q = \frac{5}{6}$. Складемо нерівність

$$n \frac{1}{6} - \frac{5}{6} \leq 55 \leq n \frac{1}{6} + \frac{1}{6}$$

Отримали лінійну систему нерівностей:

$$\begin{cases} n \frac{1}{6} - \frac{5}{6} \leq 55, \\ n \frac{1}{6} + \frac{5}{6} \geq 55 \end{cases} \quad \begin{cases} n - 5 \leq 330, \\ n + 1 \geq 330 \end{cases} \quad \begin{cases} n \leq 335, \\ n \geq 329 \end{cases}$$

Отже, кубик необхідно кинути від 329 до 335 разів.

4.4. Теорема Пуассона

Якщо число випробувань велике, формулу Бернуллі застосовувати незручно. В цьому випадку можна застосовувати наближені формули, точність яких збільшується із зростанням n .

Теорема: Якщо ймовірність p настання події A в кожному випробуванні постійна і мала, а число незалежних випробувань n досить велике, то ймовірність того, що подія A наступить m разів, наближено дорівнює

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

де $\lambda = np$.

Доведення. Введемо позначення $\lambda = np$, виразимо звідси $p = \frac{\lambda}{n}$ і підставимо цей вираз у формулу Бернуллі :

$$\begin{aligned} P_n(m) &= C_n^m \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} = \\ &= \frac{\lambda^m}{n^m} \frac{n(n-1)(n-2)\dots(n-m+1)}{m!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} = \\ &= \frac{\lambda^m}{m!} \frac{n-1}{n} \frac{n-2}{n} \dots \frac{n-m+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} = \\ &= \frac{\lambda^m}{m!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m}. \end{aligned}$$

При $m \ll n$ усі вирази в дужках, за винятком передостаннього, можна прийняти рівними одиниці, тобто

$$P_n(m) = \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n.$$

При $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda},$$

тому

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

що і вимагалось довести.

Приклад. Ткаля обслуговує 1000 веретен. Ймовірність обриву нитки на одному веретені протягом однієї хвилини дорівнює 0,005. Знайти ймовірність того, що протягом однієї хвилини обрив нитки відбудеться на 7 веретенах.

Розв'язання. Так як кількість випробувань велика ($n = 1000$), а ймовірність окремого випробування дуже мала ($p = 0,005$), то для обчислення шуканої ймовірності скористаємося формулою Пуассона. Параметр розподілу $\lambda = 1000 \times 0,005 = 5$, тоді шукана ймовірність дорівнює:

$$P_{1000}(7) = \frac{5^7}{7!} e^{-5} = \frac{78125}{5040} \cdot 0,006738 = 0,1044.$$

4.5. Поняття потоку подій

Формула Пуассона знаходить застосування в теорії масового обслуговування. Вона може розглядатися як математична модель простого потоку подій з інтенсивністю $\lambda = np$. Параметр $\lambda = np$ є при цьому середнє число успіхів.

Потоком подій називають послідовність подій, які настають у випадкові моменти часу.

Інтенсивністю потоку λ називають середнє число подій за одиницю часу

Простим (пуассонівським) називають потік подій, який має властивості стаціонарності, відсутності післядії і ординарності.

Властивість стаціонарності характеризується тим, що ймовірність $P_t(m)$ появи m подій на будь-якому проміжку часу залежить тільки від числа m і від тривалості проміжку часу t і не залежить від початку його відліку.

Властивість відсутності післядії характеризується тим, що ймовірність появи m подій на будь-якому проміжку часу не залежить від того, з'явилися або не з'явилися події в моменти часу, передуючі початку даного проміжку, тобто передісторія потоку не позначається на ймовірності появи подій.

Властивість ординарності характеризується тим, що поява двох і більше подій за малий проміжок часу малоймовірна в порівнянні з ймовірністю появи тільки однієї події.

Якщо інтенсивність простого потоку λ відома, то ймовірність появи m подій за час t визначається формулою

$$P_t(m) = \frac{(\lambda t)^m \cdot e^{-\lambda t}}{m!}.$$

Приклад. Середнє число замовлень таксі, що поступають на диспетчерський пункт за одну хвилину, дорівнює трьом. Знайти ймовірність того, що за 2 хвилини поступить 4 виклики.

Розв'язання. Підставляючи в наведену вище формулу $\lambda = 3$, $t = 2$, $m = 4$, отримаємо:

$$P_2(4) = \frac{(3 \cdot 2) \cdot e^{-3 \cdot 2}}{4!} = 0,135.$$

4.6. Локальна теорема Муавра-Лапласа

Лапласом була отримана важлива наближена формула для ймовірності $P_n(m)$ появи події A рівно m разів, за умови, що n досить велике. На відміну від формули Пуассона тут немає

обмеження на мализну величини p в окремому випробуванні, тобто область застосовності формули Лапласа ширша.

Теорема. Ймовірність того, що в умовах схеми Бернуллі подія A при n випробуваннях з'явиться рівно m разів, виражається наближеною формулою Лапласа

$$P_n(m) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{t^2}{2}}, \text{ де } t = \frac{m - np}{\sqrt{npq}}.$$

Формулу Лапласа іноді називають асимптотичною формулою, оскільки доведено, що відносна помилка формули Лапласа прямує до нуля при $n \rightarrow \infty$.

4.7. Інтегральна (глобальна) теорема Муавра-Лапласа

Інтегральна теорема Муавра-Лапласа містить наближену формулу для ймовірності $P_n(m_1, m_2)$ того, що подія A з'явиться не менше m_1 разів і не більше m_2 разів.

Теорема. Ймовірність того, що подія A з'явиться в n випробуваннях від m_1 до m_2 разів, приблизно дорівнює визначеному інтегралу

$$P_n(m_1, m_2) \approx \int_{t_{m_1}}^{t_{m_2}} \varphi_0(t) dt = \frac{1}{\sqrt{2\pi}} \int_{t_{m_1}}^{t_{m_2}} e^{-\frac{t^2}{2}} dt,$$

де: $t_{m_1} = \frac{m_1 - np}{\sqrt{npq}}; t_{m_2} = \frac{m_2 - np}{\sqrt{npq}}.$

Доведення. На підставі теореми додавання ймовірностей несумісних подій

$$P_n(m_1, m_2) = \sum_{m=m_1}^{m_2} P_n(m).$$

Або, використовуючи локальну теорему Лапласа

$$P_n(m_1, m_2) \approx \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{npq}} \varphi_0(t_m).$$

Введемо позначення $\Delta t_m = t_{m+1} - t_m = \frac{(m+1) - np}{\sqrt{npq}} - \frac{m - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}.$

Запишемо P_n у вигляді $P_n(m_1, m_2) \approx \sum_{m=m_1}^{m_2} \varphi_0(t_m) \Delta t_m.$

Очевидно, при $n \rightarrow \infty$ величина $\Delta t_m \rightarrow 0$, і остання сума прямує до визначеного інтегралу:

$$P_n(m_1, m_2) \approx \int_{t_{m_1}}^{t_{m_2}} \varphi_0(t) dt = \frac{1}{\sqrt{2\pi}} \int_{t_{m_1}}^{t_{m_2}} e^{-\frac{t^2}{2}} dt,$$

що і вимагалось довести.

Введемо стандартний інтеграл Лапласа (функцію Лапласа):

$$\Phi_0(x) = \int_0^x \varphi_0(t) dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt,$$

який, очевидно, є первісною функції Гауса

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тоді на підставі формули Ньютона - Лейбніца можна записати

$$P_m(m_1, m_2) \approx \Phi_0(t_{m_2}) - \Phi_0(t_{m_1}).$$

Значення функцій $\Phi_0(x)$ і $\varphi_0(x)$ зазвичай знаходяться з таблиць, причому таблиці зазвичай дані лише для невід'ємних значень x , оскільки $\varphi_0(x)$ - парна функція, а $\Phi_0(x)$ - непарна.

Приклад. Ймовірність присутності студента на лекції дорівнює 0,8. Знайти ймовірність того, що з 100 студентів на лекції будуть присутні не менше 75 і не більше 90.

Розв'язання. Так як кількість випробувань велика ($n = 100$), то для знаходження ймовірності, що подія A зявиться від 75 до 90 разів, скористаємося інтегральною теоремою Лапласа. Визначимо аргументи інтегральної функції Лапласа x_1 і x_2 :

$$x_1 = \frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = -1,25, \quad x_2 = \frac{90 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = 2,5.$$

Враховуючи, що функція $\Phi(x)$ є непарною, тобто $\Phi(-x) = -\Phi(x)$ за таблицею значень інтегральної функції Лапласа знаходимо:

$$\Phi(-1,25) = -\Phi(1,25) = -0,39435 \text{ і } \Phi(2,5) = 0,49379,$$

тоді

$$P_{100}(75 \leq k \leq 90) = \Phi(2,5) - \Phi(-1,25) = 0,49379 + 0,39435 = 0,888.$$

Питання для самоперевірки

1. Що називають схемою Бернуллі?
2. Записати формулу Бернуллі.
3. Що таке твірна функція?
4. Коли застосовується твірна функція?
5. В яких межах лежить наймовірніше число появи подій в незалежних випробуваннях?
6. Записати формулу Пуассона.
7. Коли застосовують формулу Пуассона?
8. Що називають потоком подій?

9. Що таке інтенсивність потоку?
10. Що називають простим потоком подій?
11. Сформулювати властивості простого потоку подій.
12. Записати локальну теорему Лапласа.
13. Записати інтегральну теорему Лапласа.
14. Записати функції Гаусса та Лапласа.

5. ВИПАДКОВА ВЕЛИЧИНА І ЇЇ ФУНКЦІЯ РОЗПОДІЛУ. ВЛАСТИВОСТІ ФУНКЦІЇ РОЗПОДІЛУ

5.1. Випадкова величина. Види випадкових величин

Одним з найважливіших понять теорії ймовірності є поняття *випадкової величини*.

Випадковою називають величину, яка в результаті випробування прийме одне і тільки одне можливе значення, наперед невідоме і залежне від випадкових причин, які заздалегідь не можуть бути враховані.

Приклад 1. Число хлопчиків, що народилися, серед ста новонароджених - випадкова величина, що має наступні можливі значення: 0, 1, ..., 100.

Приклад 2. Відстань, яку пролетить снаряд при пострілі з гармати, є випадковою величиною, яка залежить не лише від установки прицілу, але і від сили і напрямку вітру, температури, вологості і так далі. Можливі значення цієї випадкової величини належать деякому проміжку (a, b) .

Випадкова величина зазвичай позначається прописною латинською буквою $(X, Y, Z, \dots)$, її конкретні значення - рядковими буквами $(x, y, z, \dots)$.

Більш строгіше формально-математичне визначення випадкової величини: випадковою величиною називається функція $X = X(\omega)$, визначена на множині елементарних подій Ω , $\omega \in \Omega$

Випадкові величини діляться на дискретні і неперервні.

Дискретною (переривчастою) називають випадкову величину, яка набуває окремих, ізольованих можливих значень з певною вірогідністю. Число можливих значень дискретної випадкової величини може бути скінченним або нескінченним.

Неперервною називають випадкову величину, яка може набувати усіх значень з деякого скінченного або нескінченного проміжку. Очевидно, число можливих значень неперервної випадкової величини нескінченне.

5.2. Закон розподілу дискретної випадкової величини

Законом розподілу випадкової величини називається відповідність між усіма можливими значеннями дискретної випадкової величини і їх ймовірностями, тобто сукупність пар чисел (x_i, p_i) .

Закон розподілу можна задавати таблично, аналітично (у вигляді формули) і графічно. При табличному задаванні закону розподілу перший рядок таблиці містить можливі значення, а другий - їх ймовірності:

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

Оскільки в одному випробуванні випадкова величина приймає одне і тільки одне можливе значення, то події $X = x_1, X = x_2, \dots, X = x_n$ утворюють повну групу, у зв'язку з чим сума ймовірностей цих подій дорівнює одиниці:

$$\sum_{i=1}^n p_i = 1$$

Приклад. Нехай схожість насіння деякої рослини визначається ймовірністю 0,6. Знайти закон розподілу X - числа рослин, що з'явилися, з 5 посаджених насінин.

Розв'язання: випадкова величина X може набувати значень 0,1,2,... 5. Завдання описується схемою випробувань Бернуллі з $p = 0.6$. Таким чином $P_n(k) = C_n^k p^k q^{n-k}$ і ми отримаємо

x_k	0	1	2	3	4	5
p_k	0,01	0,077	0,230	0,345	0,259	0,0778

5.3. Функція розподілу випадкової величини і її властивості

Як вже відзначалося, дискретна випадкова величина може бути задана переліком усіх її можливих значень і їх ймовірностей. Такий спосіб непридатний для неперервних випадкових величин, оскільки неможливо скласти перелік усіх можливих значень, що заповнюють інтервал (a, b) . У зв'язку з цим вводиться поняття функції розподілу ймовірності випадкової величини, придатне як для дискретної, так і для неперервної випадкової величини.

Нехай x - дійсне число. Ймовірність події, що полягає в тому, що X набуде значення, менше x , тобто ймовірність події $x < X$, позначимо через $F(x)$. Зрозуміло, якщо x змінюється, то, взагалі кажучи, змінюється і $F(x)$, тобто $F(x)$ є функцією x .

Функцією розподілу називають функцію $F(x)$, що визначає ймовірність того, що випадкова величина X в результаті випробування набуде значення, менше x :

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

Геометрично цю рівність можна представити так: $F(x)$ є ймовірність того, що випадкова величина набуде значення, яке зображається на числовій осі точкою, що лежить лівіше x .

Розглянемо окремо випадки дискретної і неперервної випадкової величин.

1. Дискретна випадкова величина. Розглянемо функцію розподілу $F_X(x)$ дискретної випадкової величини X , що набуває значень $x_1, x_2, \dots, x_n$.

✓ Якщо $x \leq x_1$, то $F(x) = P(X < x) = 0$, оскільки в цьому випадку подія $(X < x)$ є неможливою.

- ✓ Якщо $x_1 < x \leq x_2$, то подія $(X < x)$ наступить тоді і тільки тоді, коли наступить подія $(X < x_1)$, тому $F(x) = P(X < x) = P(X = x_1) = p_1$.
- ✓ Якщо $x_2 < x \leq x_3$, то подія $(X < x)$ дорівнює сумі подій $X = x_1$ і $X = x_2$ і

$$F(x) = P(X < x) = P(X = x_1) + P(X = x_2) = p_1 + p_2.$$
- ✓ Аналогічно, якщо $x_i < x \leq x_{i+1}$, то $F(x) = p_1 + p_2 + \dots + p_i$.

Таким чином, функція розподілу випадкової дискретної величини дорівнює

$$F_X(x) = \sum_{x_i < x} p_i, \text{ де } p_i = P(X = x_i), \text{ і підсумовування ведеться по тих } i, \text{ для яких } x_i < x.$$

Таким чином, в точках $x_1, x_2, \dots, x_n$ функція розподілу має стрибки.

2. Неперервна випадкова величина. На відміну від випадку дискретної випадкової величини в даному випадку X пробігає всю неперервну множину значень, а сама функція $F(x)$ зростає монотонно.

Якщо ймовірність події $X < x_1$ дорівнює $F(x_1)$, а ймовірність події $X < x_2$ дорівнює $F(x_2)$, то ймовірність того, що випадкова величина X знаходиться між x_1 і x_2 , дорівнює різниці відповідних значень функції розподілу: $P(x_1 < X < x_2) = F(x_2) - F(x_1)$.

Ймовірність того, що неперервна випадкова величина X набуде одного певного значення, дорівнює нулю. Має сенс розглядати лише ймовірність попадання її в деякий інтервал, нехай навіть і скільки завгодно малий.

Графік функції розподілу для дискретної випадкової величини є ступінчастою розривною функцією, а неперервної - монотонно зростаючою неперервною функцією.

Приведемо ряд властивостей функції розподілу, що безпосередньо впливають з її означення.

1. Функція розподілу набуває значень з проміжку $[0,1]$:

$$0 \leq F(x) \leq 1.$$
2. Функція розподілу - неспадна функція, тобто $F(x_2) > F(x_1)$ при $x_2 > x_1$.
3. Ймовірність того, що випадкова величина набуде значення з напівінтервалу $[x_1, x_2)$, дорівнює різниці $F(x_2) - F(x_1)$

$$P(x_1 \leq x < x_2) = F(x_2) - F(x_1)$$
4. $P\{X \geq x\} = 1 - F(x)$.
5. Якщо $x \rightarrow \infty$, то $F(x) \rightarrow 1$
6. Якщо $x \rightarrow -\infty$, то $F(x) \rightarrow 0$.

5.4. Щільність розподілу ймовірності та її властивості

Для неперервних випадкових величин, окрім функції розподілу вводиться також поняття щільності розподілу ймовірності, або щільність ймовірності.

Щільністю розподілу ймовірності неперервної випадкової величини називається похідна від її функції розподілу

$$f(x) = F'(x).$$

Знаючи щільність розподілу ймовірності, можна знайти функцію розподілу, інтегруючи щільність ймовірності в загальному випадку від $-\infty$ до даного значення, тобто

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

Властивості щільності розподілу

$$1. \quad f(x) \geq 0$$

Дійсно, оскільки функція розподілу неспадна функція, то її похідна - функція невід'ємна

2. Невласний інтеграл від щільності розподілу ймовірності в межах від $-\infty$ до $+\infty$ дорівнює одиниці:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Дійсно, цей невластний інтеграл виражає ймовірність події, що полягає в тому, що випадкова величина набуде значення, що належить інтервалу $(-\infty, +\infty)$. Оскільки така подія достовірна, то її ймовірність дорівнює одиниці.

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

Імовірнісний зміст щільності розподілу ймовірності : ймовірність того, що випадкова величина набуде значення, що належить інтервалу $(x, x+dx)$ приблизно дорівнює добутку щільності ймовірності в точці x на ширину інтервалу dx .

Приклад. Нехай дальність польоту снаряду при певній установці прицілу описується щільністю розподілу виду

$$f(x) = \begin{cases} -(x-15)^2 + 5, & \text{якщо } 14,9 < x < 15,1 \\ 0 & \text{в інших випадках} \end{cases}$$

Яка ймовірність того, що при одному пострілі буде отриманий переліт в межах від 10 до 20 метрів?

$$\text{Відповідь: } \int_{15,01}^{15,02} f(x)dx = 0,1$$

Питання для самоперевірки

1. Що таке випадкова величина?
2. Які є види випадкових величин?
3. Дати означення закону розподілу випадкової величини.
4. Що таке функція розподілу випадкової величини?
5. Сформулювати властивості функції розподілу.
6. Для яких випадкових величин вводиться поняття щільності розподілу?

7. Дати означення щільності розподілу.
8. Сформулювати властивості щільності розподілу.

6. МАТЕМАТИЧНЕ СПОДІВАННЯ ВИПАДКОВОЇ ВЕЛИЧИНИ І ЙОГО ВЛАСТИВОСТІ. ДИСПЕРСІЯ ВИПАДКОВОЇ ВЕЛИЧИНИ, ЇЇ ВЛАСТИВОСТІ. СЕРЕДНЄ КВАДРАТИЧНЕ ВІДХИЛЕННЯ

У багатьох практичних випадках інформація про випадкову величину, яку дають закон розподілу, функція розподілу або щільність ймовірності, є надлишковою. Часто простіше і зручніше користуватися числами, які описують випадкову величину сумарно. До найбільш важливих з таких числових характеристик випадкових величин належать математичне сподівання, дисперсія і середнє квадратичне відхилення.

6.1. Математичне сподівання випадкової величини та його властивості

Математичне сподівання характеризує середнє очікуване значення випадкової величини, тобто приблизно дорівнює її середньому значенню (імовірнісний зміст математичного сподівання). Іноді знання цієї характеристики вистачає для розв'язання завдання. Наприклад, при оцінці купівельної спроможності населення цілком може вистачити знання середнього доходу, при аналізі ймовірності двох видів діяльності можна обмежитися порівнянням їх середніх прибутковостей. Знання того, що випускники цього університету заробляють в середньому більше випускників іншого, може послужити основою для ухвалення рішення про вступ до цього ВНЗу і тому подібне

Математичне сподівання дискретної випадкової величини визначається співвідношенням:

$$M(X) = M_X = x_1 p_1 + x_2 p_2 + \dots + x_n p_n = \sum_{i=1}^n x_i p_i, \text{ де } \sum_{i=1}^n p_i = 1.$$

Математичне сподівання неперервної випадкової величини дорівнює

$$M(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx, \text{ де } f(x) - \text{щільність ймовірності.}$$

Властивості математичного сподівання

Перш ніж формулювати властивості математичного сподівання, необхідно пояснити зміст арифметичних операцій $X + Y$, $X - Y$, $X \cdot Y$ і тому подібне, де X і Y - дискретні випадкові величини.

Наприклад, під сумою $X + Y$ розуміється випадкова величина Z , значеннями якої є усі допустимі суми $z_{ij} = x_i + y_j$, де x_i і y_j - усі можливі значення відповідно випадкових величин X і Y .

Під добутком $X \cdot Y$ розуміється випадкова величина Z , значеннями якої є усі допустимі добутки $z_{ij} = x_i \cdot y_j$, де x_i і y_j - усі можливі значення відповідно випадкових величин X і Y .

1. Математичне сподівання постійної величини C дорівнює цій величині.

$$M(C) = C \cdot 1 = C.$$

2. Математичне сподівання суми (різниці) двох або декількох випадкових величин X і Y дорівнює сумі (різниці) їх математичних сподівань :

$$M(X \pm Y) = M(X) \pm M(Y).$$

Наслідок. Якщо C - постійна величина, то

$$M(X + C) = M(X) + C$$

3. Математичне сподівання добутку двох незалежних випадкових величин X і Y дорівнює добутку їх математичних сподівань :

$$M(X \cdot Y) = M(X) \cdot M(Y).$$

Наслідок 1. Математичне сподівання добутку декількох взаємно незалежних випадкових величин дорівнює добутку математичних сподівань цих величин.

Наслідок 2. Постійний множник можна виносити за знак математичного сподівання, тобто. $M(CX) = C \cdot M(X)$

6.2. Дисперсія випадкової величини і її властивості

На практиці часто вимагається оцінити розсіювання випадкової величини навколо її середнього значення. Наприклад, акції двох компаній можуть приносити в середньому однакові дивіденди, проте вкладення грошей в одну з них може бути набагато ризикованішою операцією, ніж в іншу. Тому виникає необхідність в числовій характеристиці, що оцінює розкид можливих значень випадкової величини відносно її середнього значення (математичного сподівання). Такою характеристикою є дисперсія.

Дисперсією (розсіюванням) випадкової величини називають математичне сподівання квадрата відхилення цієї величини від її математичного сподівання

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

Легко показати, що наведений вище вираз може бути записаний у вигляді

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

Дійсно, використовуючи основні теореми про математичне сподівання, отримаємо

$$\begin{aligned} D(X) &= M((X - M(X))^2) = M(X^2 - 2XM(X) + M(X)^2) = \\ &= M(X^2) - 2M(X)M(X) + (M(X))^2 = M(X^2) - (M(X))^2 \end{aligned}$$

У разі дискретної випадкової величини, що має закон розподілу (x_i, p_i) ($i = 1, 2, \dots, n$)

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

Для неперервної випадкової величини формула для розрахунку дисперсії має вигляд

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

Властивості дисперсії

1. Дисперсія постійної величини дорівнює нулю.

$$D(C) = 0$$

2. Постійний множник можна виносити за знак дисперсії, підводячи його до квадрату:

$$D(CX) = C^2 D(X).$$

3. Дисперсія суми (різниці) двох незалежних випадкових величин дорівнює сумі дисперсій цих величин :

$$D(X \pm Y) = D(X) + D(Y).$$

Наслідок 1. Дисперсія суми декількох взаємно незалежних випадкових величин дорівнює сумі дисперсій цих величин.

Наслідок 2. Якщо C - постійна величина, то $D(X + C) = D(X)$.

Математичне сподівання і дисперсія випадкової величини є її основними числовими характеристиками.

Приклад 1. Нехай закон розподілу дискретної випадкової величини має вигляд

x_i	1	2	3	4	5
p_i	0,07	0,21	0,55	0,16	0,01

Знайти математичне сподівання і дисперсію випадкової величини X .

Розв'язування: Розрахуємо спочатку математичне сподівання

$$M(X) = 1 \cdot 0,07 + 2 \cdot 0,21 + 3 \cdot 0,55 + 4 \cdot 0,16 + 5 \cdot 0,01 = 2,83$$

Дисперсія дорівнює

$$D(X) = (1 - 2,83)^2 \cdot 0,07 + (2 - 2,83)^2 \cdot 0,21 + (3 - 2,83)^2 \cdot 0,55 + (4 - 2,83)^2 \cdot 0,16 + (5 - 2,83)^2 \cdot 0,01 = 0,661$$

Приклад 2. Щільність ймовірності неперервної випадкової величини дорівнює

$$f(x) = \exp(-x), \quad \text{де } x > 0$$

Знайти її математичне сподівання і дисперсію.

Розв'язування: Знайдемо математичне сподівання:

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x e^{-x} dx = \left[\begin{array}{ll} U = x & dU = dx \\ dV = e^{-x} dx & V = -e^{-x} \end{array} \right] = [-e^{-x} \cdot x] + \int_0^{\infty} e^{-x} dx = 1$$

Далі

$$M(X^2) = \int_0^{\infty} x^2 \cdot e^{-x} dx = \left[\begin{array}{ll} U = x^2 & dU = 2x dx \\ dV = e^{-x} & V = -e^{-x} \end{array} \right] = [-e^{-x} \cdot x^2] + 2 \int_0^{\infty} e^{-x} dx = 2$$

Знайдемо дисперсію, використовуючи формулу

$$D(X) = M(X^2) - (M(X))^2 = 2 - 1 = 1.$$

6.3. Середнє квадратичне відхилення

Для оцінки розсіювання можливих значень випадкової величини навколо її середнього значення окрім дисперсії служать і деякі інші характеристики. До їх числа відноситься середнє квадратичне відхилення.

Середнім квадратичним відхиленням σ (чи стандартом) випадкової величини X називається корінь квадратний з дисперсії цієї величини :

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

Легко показати, що дисперсія має розмірність, рівну квадрату розмірності випадкової величини. Тому розмірність $\sigma(X)$ співпадає з розмірністю X . В тих випадках, коли бажано, щоб оцінка розсіювання мала розмірність випадкової величини, обчислюють середнє квадратичне відхилення, а не дисперсію.

Поняття дисперсії і середнього квадратичного відхилення широко використовується практично в усіх областях людської діяльності, пов'язаних з процесами вимірів. Так, наприклад, в техніці, вони характеризують точність вимірювальної апаратури (чим вище середньоквадратичне відхилення (розкид) при вимірах, тим гірша якість приладу).

Прикладами використання цих параметрів в економіці можуть служити вивчення ризику різних дій з випадковим результатом, зокрема, при аналізі ризику інвестування в ту або іншу галузь, при оцінюванні різних активів в портфелі цінних паперів і так далі.

Приклад. Нехай є два варіанти інвестування з наступними характеристиками

Ймовірності можливого чистого доходу

Порівняння варіантів рішень

Чистий дохід, млн. гр. од.	-3	-	-1	0	1	2	3	4
Ймовірності:								
Інвестиція 1	0	0	0,1	0	0,3	0,2	0,2	0
Інвестиція 2	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2

Очікуваний чистий прибуток інвестування визначається математичним сподіванням і складає:

Інвестиція 1: $M(X) = \sum x_i \cdot p_i = 1,2$ млн.

Інвестиція 2: $M(X) = \sum x_i \cdot p_i = 1,1$ млн.

За очікуваним прибутком прийнятніший 1-й варіант. Проте ми не врахували ризик, пов'язаний з інвестиціями. Цей ризик може бути визначений за допомогою дисперсії і (чи) середнього квадратичного відхилення. Використовуючи результати таблиці, отримаємо

Інвестиція 1: $\sigma(X) = \sqrt{D(X)} = 1,25$

Інвестиція 2: $\sigma(X) = \sqrt{D(X)} = 2,385$

Тобто ризик за варіантом для інвестиції 1 менший. Вибір - за підприємцем.

Питання для самоперевірки

1. Що таке математичне сподівання випадкової величини?
2. Сформулювати властивості математичного сподівання.
3. Що таке дисперсія випадкової величини?
4. Записати формули для обчислення дисперсії дискретної та неперервної величин.
5. Сформулювати властивості дисперсії.
6. Що таке середнє квадратичне відхилення випадкової величини?

7. ОСНОВНІ РОЗПОДІЛИ ВИПАДКОВОЇ ВЕЛИЧИНИ, ЇХ МАТЕМАТИЧНІ СПОДІВАННЯ І ДИСПЕРСІЇ

Випадкову величину повністю задає закон її розподілу (у дискретному випадку), а також функція розподілу або щільність ймовірності (для неперервної випадкової величини).

Найбільш важливими законами розподілу дискретної випадкової величини є біноміальний закон, закон розподілу Пуассона, геометричний і гіпергеометричний розподіли, а неперервної - нормальний, рівномірний і показниковий розподіли.

7.1. Біноміальний розподіл, його математичне сподівання і дисперсія

Закон розподілу випадкової величини X числа появ події A в схемі Бернуллі має вигляд

$$P_m(n) = P(X = m) = C_n^m p^m q^{n-m}$$

$$\text{де } m = 0, 1, 2, \dots, n, \quad q = P(\bar{A}) = 1 - p.$$

Ця формула ще називається біноміальною, оскільки її права частина є $(m + 1)$ -й член бінома

$$\text{Ньютона: } (p + q)^n = \sum_{m=0}^n C_n^m p^m q^{n-m}.$$

Очевидно, що для закону біноміального розподілу ймовірності виконується умова

$$\text{нормування, тобто сума всіх ймовірностей дорівнює одиниці: } \sum_{m=0}^n P_m(n) = (p + q)^n = 1.$$

Біноміальний розподіл для $n=10$ і деяких значень p наведено нижче



Математичне сподівання числа появ події A в n незалежних випробуваннях для біноміального розподілу дорівнює добутку числа випробувань на ймовірність появи події A в кожному випробуванні (тобто середньому числу появи події в цій серії випробувань). $M(X) = np$.

Дисперсія і середнє квадратичне відхилення дорівнюють відповідно: $D(X) = npq$; $\sigma(X) = \sqrt{D(X)} = \sqrt{npq}$.

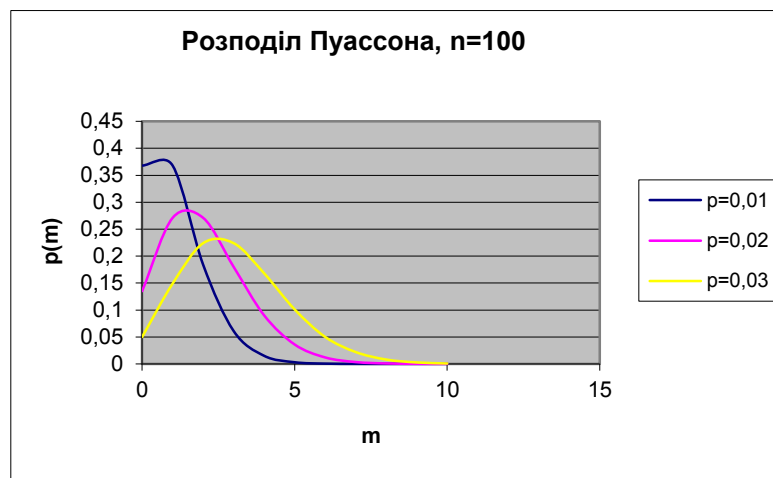
7.2. Розподіл Пуассона

Раніше відзначалося, що якщо при збільшенні числа випробувань добуток $np = \lambda$ залишається постійним, то біноміальний розподіл при великих значеннях n збігається до розподілу Пуассона.

Випадкова величина X називається розподіленою за законом Пуассона, якщо вона може набувати значень $0, 1, 2, \dots, n$, відповідна ймовірність яких визначається за формулою Пуассона:

$$P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad (n \rightarrow \infty).$$

Розподіл Пуассона для $n=100$, $\lambda=1, 2, 3$ наведено нижче.



Для розподілу Пуассона математичне сподівання і дисперсія дорівнюють відповідно:

$$M(X) = \lambda; \quad D(X) = \lambda.$$

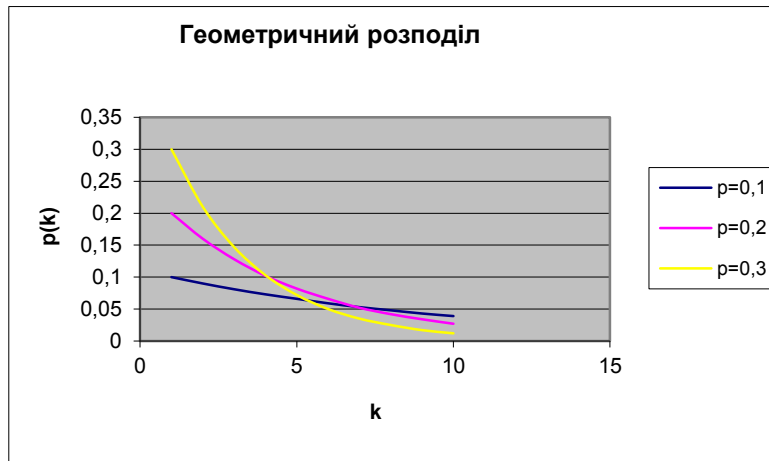
Рівність значень математичного очікування і дисперсії є унікальною властивістю розподілу Пуассона. Ця властивість часто застосовується на практиці для вирішення питання, чи правдоподібна гіпотеза про те, що випадкова величина X розподілена за законом Пуассона. Для цього визначають з експерименту статистичні характеристики випадкової величини - математичне сподівання і дисперсію. Якщо їх значення близькі, то це може служити аргументом на користь гіпотези про пуассонівський розподіл.

7.3. Геометричний розподіл

Дискретна випадкова величина X має геометричний розподіл, якщо вона набуває значень $k=1, 2, 3, \dots$ (зліченна кількість значень) з ймовірностями

$$p_k = P(X = k) = q^{k-1} \cdot p, \quad 0 < p < 1, \quad q = 1 - p.$$

Випадкова величина, що має геометричний розподіл, є числом випробувань в схемі Бернуллі до першого успіху. Геометричний розподіл для деяких конкретних значень p наведений нижче



Можна показати, що математичне сподівання і дисперсія для геометричного розподілу

дорівнюють відповідно: $M(X) = \frac{1}{p}$ $D(X) = \frac{q}{p^2}$

Приклад. У великій партії виробів ймовірність браку рівна p . Контроль якості проводиться до першої появи бракованого виробу. В результаті серії перевірок виявилось, що бракований виріб уперше з'явився в середньому при десятому випробуванні. Оцінити чисельне значення p .

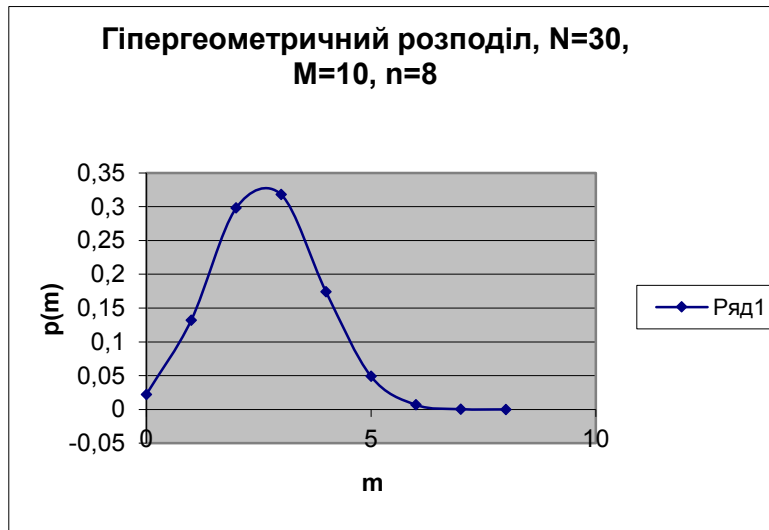
Розв'язання. Нехай X - число випробувань до першої появи бракованого виробу. Ця випадкова величина має геометричний розподіл. За умовою її середнє значення рівне $M(X) = 10$. Таким чином $p = 1/M(X) = 0,1$

7.4. Гіпергеометричний розподіл (урнова схема)

Дискретна випадкова величина X має гіпергеометричний розподіл, якщо вона набуває значення m з ймовірностями

$$p_m = P(X = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$$

p_m представляє ймовірність вибору m об'єктів, що мають задану властивість, з безлічі об'єктів, випадково витягнутих (без повернення) з сукупності N об'єктів, серед яких M об'єктів мають задану властивість. Нижче наведений приклад графіка гіпергеометричного розподілу.



Математичне сподівання і дисперсія випадкової величини, що має гіпергеометричний розподіл з параметрами рівні :

$$M(X) = n \cdot \frac{M}{N} ; \quad D(X) = n \cdot \frac{M}{N-1} \cdot \left(1 - \frac{M}{N}\right) \cdot \left(1 - \frac{n}{N}\right)$$

Приклад. Є 5 фірм, у трьох з яких звітність оформлена неправильно. 2 ревізори перевіряють 2 довільно вибрані фірми. Яка ймовірність того, що при перевірці буде виявлена неправильна звітність: а) ні в одній, б) в одній, в) в двох фірмах?

Розв'язання. Ця задача може бути розв'язана за допомогою гіпергеометричного розподілу. За умовою задачі загальне число об'єктів (фірм) рівне $N = 10$, число фірм з неправильною звітністю $M=3$. Перевіряється всього дві фірми ($n=2$). Число фірм з неправильною звітністю серед двох вибраних - величина змінна ($m=0, 1, 2$). Таким чином, маємо

$$\text{а) } p_0 = P(X = 0) = \frac{C_3^0 \cdot C_2^2}{C_5^2} \quad (\text{жодної неправильної звітності})$$

$$\text{б) } p_1 = P(X = 1) = \frac{C_3^1 \cdot C_2^1}{C_5^2} \quad (\text{одна неправильна звітність})$$

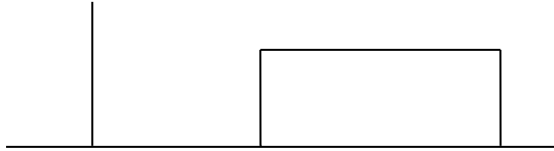
$$\text{в) } p_2 = P(X = 2) = \frac{C_3^2 \cdot C_2^0}{C_5^2} \quad (\text{дві неправильні звітності}).$$

7.5. Рівномірний розподіл

Неперервна випадкова величина вважається рівномірно розподіленою на відрізок (a, b), якщо її щільність ймовірності має вигляд:

$$f(x) = \begin{cases} 0, & x \leq a \\ \frac{1}{b-a}, & a < x \leq b \\ 0, & x > b \end{cases}$$

Графік щільності ймовірності для рівномірного розподілу



Математичне сподівання і дисперсія неперервної випадкової величини, що має рівномірний розподіл, дорівнюють відповідно:

$$M(X) = \frac{a+b}{2} \qquad D(X) = \frac{(b-a)^2}{12}$$

Приклад. Інтервал руху автобуса дорівнює 15 хв. Яка ймовірність того, що пасажир на зупинці чекатиме автобус не більше 5 хвилин?

Розв'язання. Нехай випадкова величина X - час очікування автобуса. Вона має рівномірний розподіл на відрізку $[0,15]$. Маємо

$$P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x) dx.$$

У даному випадку

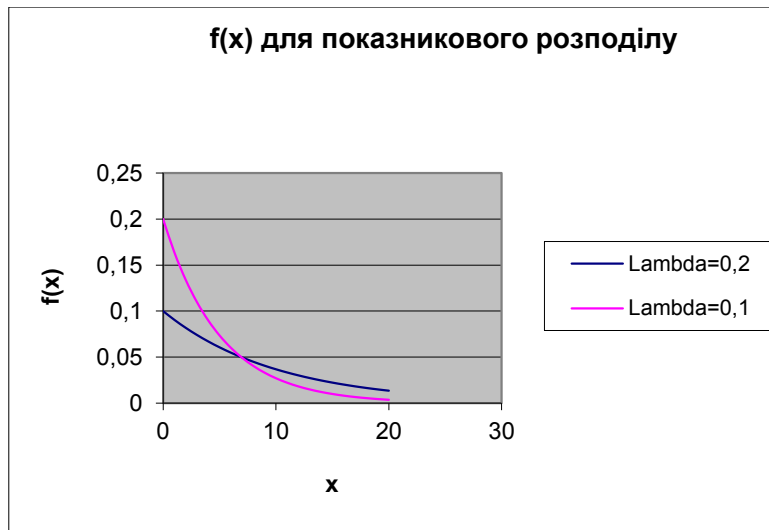
$$P(0 \leq X \leq 5) = \int_0^5 \frac{1}{b-a} dx = \int_0^5 \frac{1}{15-0} dx = \frac{5-0}{15-0} = \frac{1}{3}$$

7.6. Показниковий розподіл

Показниковим (експоненціальним) розподілом неперервної випадкової величини X називається розподіл, що має щільність ймовірності виду :

$$f(x) = \begin{cases} 0, & (x < 0) \\ \lambda e^{-\lambda x}, & (x \geq 0) \end{cases}$$

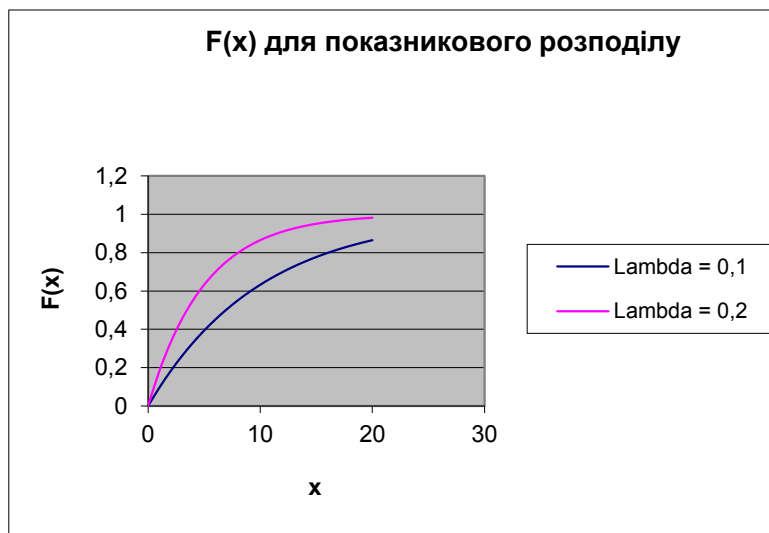
де λ - постійна додатна величина. Щільність ймовірності для показникового розподілу для $\lambda = 0,1; 0,2$ наведена нижче



Функція розподілу ймовірності для показникового розподілу має вигляд:

$$F(x) = \begin{cases} 0, & (x < 0) \\ 1 - e^{-\lambda x}, & (x \geq 0) \end{cases}$$

Функція розподілу для $\lambda = 0,1; 0,2$ наведена нижче



Можна показати, що математичне сподівання і дисперсія випадкової величини, що має експоненціальний розподіл, дорівнюють:

$$M(X) = \frac{1}{\lambda} \quad D(X) = \frac{1}{\lambda^2}$$

Приклад. Встановлено, що час горіння електричної лампочки (T) є випадковою величиною, розподіленою за показниковим законом. Вважаючи, що середнє значення цієї величини дорівнює 6 місяцям, знайти ймовірність того, що лампочка буде справна більше року.

Розв'язання. Оскільки $M(T)=1/\lambda=6$, то $\lambda=1/6$ і функція розподілу випадкової величини T має вигляд

$$F(x) = \begin{cases} 1 - \exp(-x/6), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Тому

$$P(T > 12) = P(12 < T < +\infty) = F(\infty) - F(12) = 1 - (1 - \exp(-12/6)) = \exp(-2) \approx 0.135$$

Питання для самоперевірки

1. Що таке біноміальний розподіл?
2. Записати характеристики біноміального розподілу.
3. Дати означення розподілу Пуассона.
4. Сформулювати основну властивість розподілу Пуассона.
5. Що таке геометричний розподіл?
6. Записати характеристики геометричного розподілу.
7. Коли випадкова величина має гіпергеометричний розподіл?
8. Записати характеристики гіпергеометричного розподілу.
9. Який розподіл називається рівномірним?
10. Записати характеристики рівномірного розподілу.
11. Який розподіл називається показниковим?
12. Як пов'язані між собою характеристики показникового розподілу?

8. НОРМАЛЬНИЙ РОЗПОДІЛ

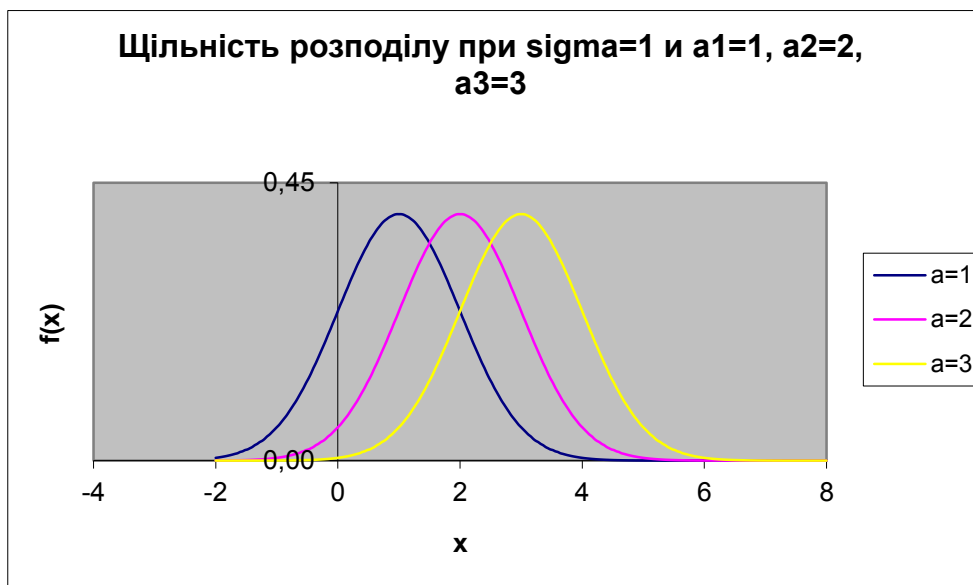
8.1. Означення нормального розподілу

Неперервна випадкова величина має нормальний закон розподілу з параметрами μ і σ , якщо її щільність ймовірності має вигляд функції Гауса

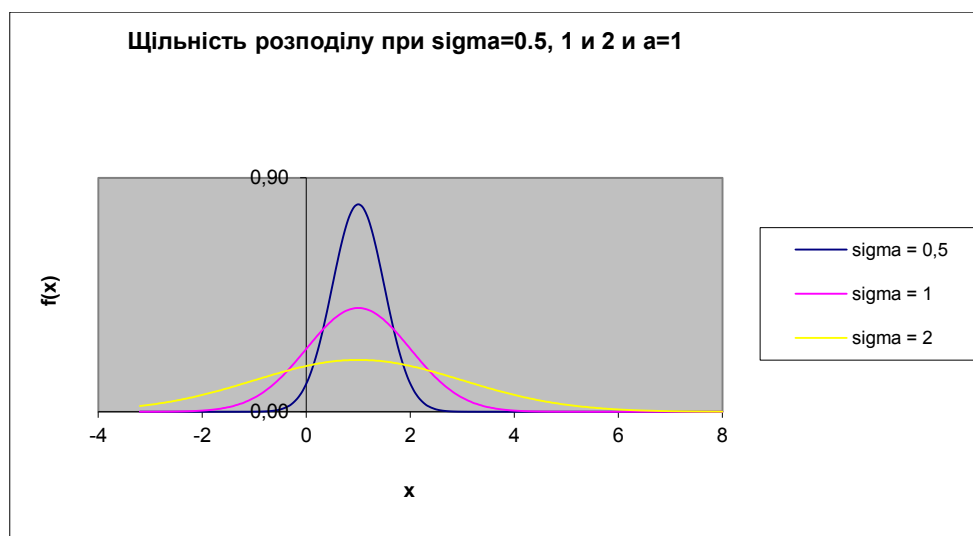
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Де $\sigma > 0$. За допомогою безпосереднього обчислення математичного сподівання і дисперсії нормального розподілу легко з'ясувати імовірнісний зміст його параметрів: μ - є математичне сподівання, а σ - середнє квадратичне відхилення нормального розподілу. При $\mu=0$, $\sigma=1$ розподіл називається **стандартним нормальним розподілом**.

Графіки для ряду конкретних значень математичного сподівання і середнього квадратичного відхилення приведені нижче.



Мал. 1. Зміна вигляду функції $f(x)$ при зміні математичного сподівання

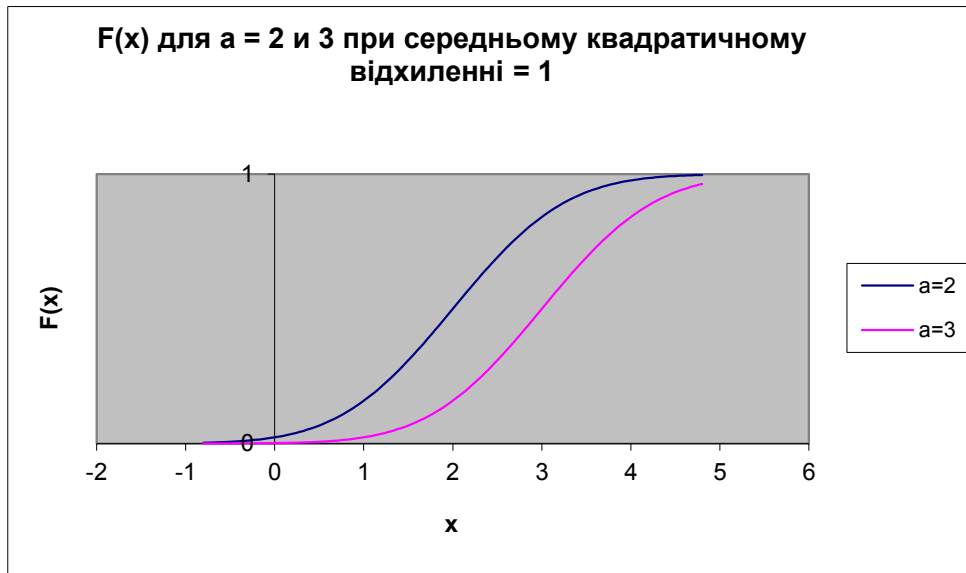


Мал. 2. Зміна вигляду функції $f(x)$ при зміні середнього квадратичного відхилення

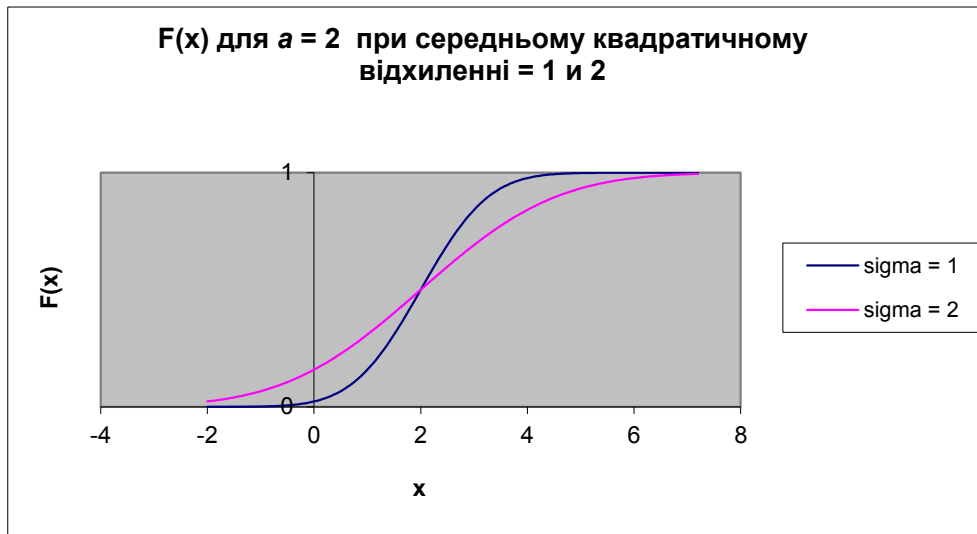
Функція розподілу у випадку нормального розподілу, очевидно, дорівнює

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

Графіки функції для ряду значень математичного сподівання і середнього квадратичного відхилення зображені на малюнках, що наводяться нижче



Мал. 3. Залежність функції розподілу від величини a



Мал. 4. Залежність функції розподілу від величини σ

Нормальний розподіл має виключно важливе значення для практичних застосувань, оскільки багато неперервних випадкових величин описуються саме цим розподілом. Виявляється, що підсумовування великого числа випадкових величин з різними законами розподілу призводить до нормального розподілу результуючої суми. Ця властивість підтверджується центральною граничною теоремою (*теорема Ляпунова*). Зміст цієї теореми полягає в наступному. Якщо випадкова величина X є сумою дуже великого числа взаємно незалежних випадкових величин, вплив кожної з яких на усю суму дуже малий, то X має розподіл, близький до нормального.

Слід мати на увазі, що при посиленні впливу окремих факторів можуть з'являтися відхилення від нормального розподілу результуючого параметра. Тому велике значення на

практиці приділяється експериментальній перевірці висунених гіпотез, у тому числі і гіпотези про нормальний розподіл.

8.2. Властивості функції Гауса

Графік щільності нормального розподілу називають *нормальною кривою Гауса*.

Досліджуємо поведінку функції щільності ймовірності $f(x)$.

1. Очевидно, що функція визначена на усій осі X .
2. Функція набуває лише додатних значень, тобто нормальна крива розташована над віссю Ox .
3. Вісь Ox служить горизонтальною асимптотою графіка. Інших асимптот у графіка немає.
4. При $x = a$ функція має максимум, рівний $\varphi(x)_{\max} = \frac{1}{\sigma\sqrt{2\pi}}$.
5. Функція парна : її графік симетричний відносно прямої $x = a$
6. При $x = a \pm \sigma_x$ графік функції має точки перегину.

При будь-яких значеннях параметрів a і σ , площа, обмежена нормальною кривою і віссю X , дорівнює одиниці.

8.3. Ймовірність попадання нормальної випадкової величини в заданий інтервал

Часто вимагається визначити ймовірність попадання випадкової величини в заданий інтервал. Ця ймовірність може бути виражена у вигляді різниці функції розподілу ймовірності в

граничних точках цього інтервалу : $P(x_1 < X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$.

У разі нормального розподілу

$$P(x_1 < X < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

Використовуючи заміну змінної : $t = \frac{x-a}{\sigma}$, $x = t\sigma + a$, $dx = \sigma dt$, отримаємо

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-\frac{t^2}{2}} dt,$$

де $t_1 = \frac{x_1 - a}{\sigma}$, $t_2 = \frac{x_2 - a}{\sigma}$.

Розіб'ємо отриманий інтеграл на два:

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}} \left(\int_{t_1}^0 e^{-\frac{t^2}{2}} dt + \int_0^{t_2} e^{-\frac{t^2}{2}} dt \right) = \frac{1}{\sqrt{2\pi}} \left(\int_0^{t_2} e^{-\frac{t^2}{2}} dt - \int_0^{t_1} e^{-\frac{t^2}{2}} dt \right)$$

Тоді шукана ймовірність може бути виражена у вигляді:

$$P(x_1 < X < x_2) = \Phi_0(t_2) - \Phi_0(t_1)$$

$$\text{де } \Phi_0(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt \quad - \text{ функція Лапласа.}$$

Функція Лапласа протабульована, що істотно спрощує розрахунок попадання нормально розподіленої випадкової величини в будь-який заданий інтервал.

8.4. Функція Лапласа і її властивості

Функція Лапласа не виражається через елементарні функції.

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

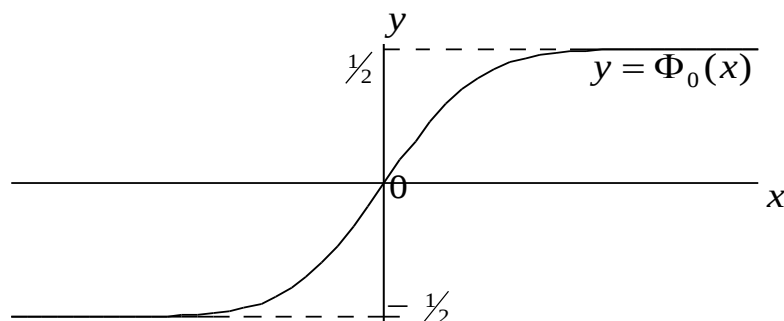
Для її обчислення використовуються спеціальні таблиці або методи наближеного обчислення.

Функція $\Phi_0(x)$ має наступні властивості:

1. $\Phi_0(0) = 0$;
2. $\Phi_0(+\infty) = \frac{1}{2}$;
3. Функція $\Phi_0(x)$ - непарна, тобто $\Phi_0(-x) = -\Phi_0(x)$, тому в таблицях зазвичай приводяться значення $\Phi_0(x)$ тільки для додатних;
4. Функція $\Phi_0(x)$ - монотонно зростаюча функція (це випливає з того, що $\Phi_0'(x) = \varphi_0(x) > 0$).

Відхилення нормальної випадкової величини від її математичного сподівання. Правило "трьох сигм"

Часто вимагається обчислити ймовірність того, що відхилення нормально розподіленої випадкової величини за абсолютною величиною від математичного сподівання менше заданого



додатного числа δ , тобто вимагається знайти ймовірність того, що виконується нерівність .

$$|X - M_X| < \delta$$

Замінімо цю нерівність рівносильною їй подвійною нерівністю .

$$M_X - \delta < X < M_X + \delta$$

Скористаємося формулою: $P(x_1 < X < x_2) = \Phi_0(t_2) - \Phi_0(t_1)$

Отримаємо: $P(M_X - \delta < X < M_X + \delta) =$

$$\begin{aligned} &= \Phi_0\left(\frac{(M_X + \delta) - M_X}{\sigma}\right) - \Phi_0\left(\frac{(M_X - \delta) - M_X}{\sigma}\right) \\ &= \Phi_0\left(\frac{\delta}{\sigma}\right) - \Phi_0\left(-\frac{\delta}{\sigma}\right) = 2\Phi_0\left(\frac{\delta}{\sigma}\right) \end{aligned}$$

Якщо в якості δ узяти потрібне значення середнього квадратичного відхилення, то отримаємо:

$$P(|X - M_X| < 3\sigma) = 2\Phi_0(3) = 0,9973,$$

Таким чином, ймовірність того, що абсолютна величина відхилення перевищить потрібне середнє квадратичне відхилення, дуже мала (0,0027 або 0,27%). Такі події можна вважати практично неможливими.

Іншими словами, якщо випадкова величина розподілена нормально, то абсолютна величина її відхилення від математичного сподівання не перевищує потрібного середнього квадратичного відхилення. У цьому і полягає *суть правила "трьох сигм"*.

На практиці правило "трьох сигм" застосовують так: якщо розподіл випадкової величини, що вивчається, невідомий, але правило "трьох сигм" виконується, то є підстава вважати, що величина, що вивчається, розподілена нормально, і навпаки.

Приклад 1. Поточна ціна цінного паперу є нормально розподіленою випадковою величиною з середнім 100 у.о. і дисперсією 9. Знайти ймовірність того, що ціна активу знаходитиметься в межах від 91 до 109 у.о.

Розв'язання. Оскільки $a=100$, $\sigma=\sqrt{9}=3$, то

$$P(91 < X < 109) = \Phi_0\left(\frac{109-100}{3}\right) - \Phi_0\left(\frac{91-100}{3}\right) = \Phi_0(3) - \Phi_0(-3) = 2\Phi_0(3) = 0.9973$$

Приклад 2. Магазин продає чоловічі костюми. За даними статистики відомо, що розподіл за розмірами є нормальним з математичним очікуванням і середнім квадратичним відхиленням, рівними 48 і 2 відповідно.

Визначити відсоток попиту на 50-й розмір, за умови розкиду значень цієї величини в інтервалі (49,51).

Розв'язання. За умовою, $a=48$, $\sigma = 2$, $\alpha = 49$, $\beta = 51$. Використовуючи формулу

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx$$

отримуємо, що ймовірність попиту на 50-й розмір в заданому інтервалі дорівнює

$$\begin{aligned} P(49 < X < 51) &= \Phi_0\left(\frac{51-48}{2}\right) - \Phi_0\left(\frac{49-48}{2}\right) = \\ &= \Phi_0(1,5) - \Phi_0(0,5) = 0,4332 - 0,1915 = 0,2417 \end{aligned}$$

Висновок: попит на 50-й розмір складе приблизно 24% і магазину треба передбачити це в загальному об'ємі закупівлі.

Питання для самоперевірки

1. Що таке нормальний розподіл?
2. Записати функцію розподілу для нормальної величини.
3. Записати властивості функції Гауса.
4. Як обчислюється ймовірність попадання нормальної величини в заданий інтервал?
5. Записати функцію Лапласа та її властивості.
6. За якою формулою обчислюється ймовірність відхилення нормальної випадкової величини від її математичного сподівання?
7. Сформулювати правило «трьох сигм».

9. НЕРІВНІСТЬ ЧЕБИШЕВА. ТЕОРЕМА ЧЕБИШЕВА. ЦЕНТРАЛЬНА ГРАНИЧНА ТЕОРЕМА

Не дивлячись на те, що заздалегідь не можна передбачити, яке з можливих значень прийме випадкова величина в результаті досліду, за деяких умов сумарна поведінка досить великого числа випадкових величин стає закономірною. Іншими словами, при дуже великому числі випадкових явищ їх середній результат практично перестає бути випадковим і може бути передбачений з великою мірою визначеності.

Для практики дуже важливе знання умов, при виконанні яких це може відбуватися. Ці умови вказуються в теоремах, що носять загальну назву закону великих чисел, найважливішою з яких є теорема Чебишева. Для доведення теореми Чебишева використовується нерівність Чебишева, яку ми зараз розглянемо.

9.1. Нерівність Чебишева

Ймовірність того, що відхилення випадкової величини X від її математичного сподівання за абсолютною величиною менше додатного числа ε , не менша, ніж $1 - D(X) / \varepsilon^2$, тобто

$$P(|X - M(X)| < \varepsilon) \geq 1 - D(X) / \varepsilon^2$$

Приклад. Номінальне значення діаметру втулки дорівнює 5 мм, а дисперсія, через похибку виготовлення, не перевершує 0,01. Оцінити ймовірність того, що розмір втулки відрізнятиметься від номіналу не більше ніж на 0,5 мм.

Розв'язання: За нерівністю Чебишева

$$P(|x - a| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2} = 1 - \frac{0.01}{0.5^2} = 0.96$$

Нерівність Чебишева дає тільки *верхню межу* ймовірності цього відхилення. Вище за цю межу ймовірність не може бути *ні при* якому законі розподілу. Наприклад, якщо ми захочемо з'ясувати, яка ймовірність того, що випадкова величина X відхилиться від свого математичного сподівання не менше, ніж на 3 середньоквадратичні відхилення, то нерівність Чебишева дасть нам верхню межу цього значення $1/9 \approx 0,111$. В той же час, наприклад для нормального розподілу ймовірність такого відхилення набагато менша - 0,0027 (правило трьох сигм).

9.2. Теорема Чебишева

Якщо $X_1, X_2, \dots, X_n$ - попарно незалежні випадкові величини, причому їх дисперсії обмежені (не перевищують постійного числа C), то, яким би малим не було додатне число ε , ймовірність виконання нерівності

$$\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon$$

буде як завгодно близька до одиниці при досить великому числі n . Інакше кажучи

$$\lim_{n \rightarrow \infty} \left(P \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| < \varepsilon \right) = 1$$

Таким чином, теорема Чебишева стверджує, що для досить великого числа незалежних випадкових величин, що мають обмежені дисперсії, майже достовірною можна вважати подію, що полягає в тому, що відхилення середнього арифметичного випадкових величин від середнього арифметичного їх математичних сподівань буде за абсолютною величиною скільки завгодно малим.

Доведення. Розглянемо нову випадкову величину - середнє арифметичне випадкових величин

$$\bar{X} = (X_1 + X_2 + \dots + X_n) / n$$

Знайдемо математичне очікування $\bar{X}$. Користуючись властивостями математичного сподівання, отримаємо

$$M\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}$$

Застосовуючи до величини $\bar{X}$ нерівність Чебишева, маємо

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - M\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)\right| < \varepsilon\right) \geq 1 - \varepsilon^{-2} D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

Або

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}\right| < \varepsilon\right) \geq 1 - \varepsilon^{-2} D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) \quad (1)$$

Користуючись властивостями дисперсії (постійний множник можна винести за знак дисперсії, підвівши його до квадрату; дисперсія суми незалежних випадкових величин дорівнює сумі дисперсій доданків), отримаємо

$$D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{D(X_1) + D(X_2) + \dots + D(X_n)}{n^2}$$

Оскільки за умовою дисперсії усіх випадкових величин обмежені постійним числом C , то

$$\frac{D(X_1) + D(X_2) + \dots + D(X_n)}{n^2} \leq \frac{(C + C + \dots + C)}{n^2} = \frac{C}{n}$$

Таким чином

$$D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) \leq \frac{C}{n}$$

Підставляючи праву частину останньої нерівності в (1) (ця нерівність може бути тільки посилена), отримаємо

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}\right| < \varepsilon\right) \geq 1 - \frac{C}{n\varepsilon^2}$$

Звідси, переходячи до межі при $n \rightarrow \infty$ і враховуючи, що ймовірність не може перевищувати одиницю, отримаємо доведення:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n}\right| < \varepsilon\right) = 1.$$

У важливому окремому випадку, коли випадкові величини мають одне і те ж математичне сподівання (позначимо його a) формула, що виражає теорему Чебишева, набирає вигляду

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| < \varepsilon\right) = 1$$

Суть теореми Чебишева така: хоча окремі незалежні випадкові величини можуть набувати значень, далеких від своїх математичних сподівань, середнє арифметичне досить великого числа випадкових величин з великою ймовірністю набуває значень, близьких до визначеного постійного числа

$$(M(X_1) + (M(X_2) + \dots + (M(X_n))) / n$$

чи - в окремому випадку, до числа a . Іншими словами, *окремі* випадкові величини можуть мати значний розкид, а їх *середнє арифметичне* розсіяне мало. Пояснюється це тим, що відхилення кожної з величин від своїх математичних сподівань можуть бути як додатними, так і від'ємними, а в середньому арифметичному вони взаємно *гасяться*.

Нехай проводиться процес вимірювання деякої величини. Розглядатимемо результати кожного вимірювання як випадкові величини $X_1, X_2, \dots, X_n$. Якщо результат кожного вимірювання не залежить від результатів інших (тобто величини попарно незалежні), а випадкові величини мають однакове математичне сподівання і їх дисперсії обмежені, то, застосовуючи теорему Чебишева, отримаємо, що при досить великому n середнє арифметичне результатів вимірювань як завгодно мало відрізняється від істинного значення вимірюваної величини (математичного сподівання a).

На теоремі Чебишева заснований широко вживаний в статистиці вибірковий метод, суть якого полягає в тому, що за порівняно невеликою випадковою вибіркою судять про усю сукупність (генеральну сукупність) досліджуваних об'єктів.

9.3. Центральна гранична теорема

Причину надзвичайно широкої поширеності випадкових величин, що описуються нормальним розподілом, пояснює центральна гранична теорема, доведена А.М. Ляпуновим.

Центральна гранична теорема: Якщо випадкова величина X є сумою дуже великого числа взаємно незалежних випадкових величин, вплив кожної з яких на усю суму нескінченно малий, то X має розподіл, близький до нормального розподілу.

Нехай $X_1, X_2, \dots, X_n$ - послідовність незалежних випадкових величин, кожна з яких має скінченні математичне сподівання і дисперсію

$$M(X_k) = a_k \quad D(X_k) = b_k^2.$$

Введемо позначення для суми випадкових величин, суми їх математичних сподівань і суми їх дисперсій

$$S_n = X_1 + X_2 + \dots + X_n \quad A_n = \sum_{k=1}^n a_k \quad B_n^2 = \sum_{k=1}^n b_k^2.$$

Розглянемо функцію $\frac{S_n - A_n}{B_n}$, яка, як легко показати, має математичне сподівання і дисперсію, рівні нулю і одиниці відповідно (нормована сума).

Дійсно

$$M\left(\frac{S_n - A_n}{B_n}\right) = \frac{1}{B_n} M[S_n - A_n] = \frac{1}{B_n} [M(S_n) - M(A_n)] = \frac{1}{B_n} [A_n - A_n] = 0,$$

$$D\left(\frac{S_n - A_n}{B_n}\right) = \frac{1}{B_n^2} D[S_n - A_n] = \frac{1}{B_n^2} [D(S_n) - D(A_n)] = \frac{1}{B_n^2} [D(S_n) - 0] = \frac{1}{B_n^2} D(X_1 + X_2 + \dots + X_n) =$$

$$\frac{1}{B_n^2} [D(X_1) + \dots + D(X_n)] = \frac{B_n^2}{B_n^2} = 1$$

Позначимо функцію розподілу нормованої суми

$$F_n(x) = P\left(\frac{S_n - A_n}{B_n} < x\right).$$

Говорять, що до послідовності $X_1, X_2, \dots, X_n$ застосована центральна гранична теорема, якщо при будь-кому x функція розподілу нормованої суми при $n \rightarrow \infty$ прямує до нормальної функції розподілу :

$$\lim_{n \rightarrow \infty} P\left[\frac{S_n - A_n}{B_n} < x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz$$

Зокрема, якщо усі випадкові величини $X_1, X_2, \dots, X_n$ однаково розподілені, то до цієї послідовності застосована центральна гранична теорема, за умови, що дисперсії усіх величин скінченні і відмінні від нуля. У окремому випадку, коли математичні сподівання і дисперсії усіх $X_1, X_2, \dots, X_n$ однакові (a, σ^2), в останній рівності треба покласти $A_n = na, B_n = \sigma\sqrt{n}$.

Центральна гранична теорема знаходить надзвичайно широке застосування в математичній статистиці, зокрема, при обґрунтуванні вибору закону розподілу генеральної сукупності.

На закінчення відмітимо, що використання теореми Чебишева і центральної граничної теореми дозволяє не лише здійснювати наукові прогнози в області випадкових явищ, але і оцінювати точність цих прогнозів.

Питання для самоперевірки

1. Дати тлумачення поняття закон великих чисел.
2. Записати нерівність Чебишева
3. Сформулювати теорему Чебишева.
4. Сформулювати центральну граничну теорему.
5. Де і як застосовується закон великих чисел?

10. ЗАКОН РОЗПОДІЛУ ФУНКЦІЇ ВИПАДКОВИХ ВЕЛИЧИН

Раніше ми розглядали задачі знаходження числових характеристик функцій випадкових величин без знаходження законів їх розподілу. Проте в багатьох застосуваннях, зокрема при визначенні ймовірності попадання цих функцій в певні області їх можливих значень, потрібно знати закони розподілу функцій. Тому при розв'язуванні задач такого типу необхідно знати закони розподілу випадкових величин, що фігурують в постановці задачі. Звичайно при цьому закон розподілу випадкового аргументу або системи випадкових аргументів є відомим, так само як і відома функціональна залежність.

Отже, виникає така задача: задана система випадкових аргументів $(X_1, X_2, \dots, X_n)$, закон розподілу якої відомий. Відома випадкова величина Y

$$Y = \varphi(X_1, X_2, X_n) \quad (1)$$

як функція випадкових аргументів $(X_1, \dots, X_n)$. Потрібно визначити закон розподілу випадкової величини Y .

10.1. Закон розподілу функції випадкової величини

Спочатку розглянемо простішу задачу: про закон розподілу функції одного випадкового аргументу

$$Y = \varphi(X) \quad (2)$$

Нехай: – X – дискретна випадкова величина, задана рядом розподілу

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

Тоді $Y = \varphi(X)$ – теж дискретна випадкова величина, яка приймає можливі значення

$$y_1 = \varphi(x_1), y_2 = \varphi(x_2) \dots, y_n = \varphi(x_n).$$

Тут розрізняють два випадки: а) коли всі значення y_i різні (функція $y = \varphi(x)$ монотонна),

б) коли серед y_i є значення, які співпадають (функція $y = \varphi(x)$ немонотонна).

а) Якщо всі значення y_i ($i = \overline{1, n}$) різні, то для кожного $i=1, 2, \dots, n$ події

$$(X = x_i) \text{ і } (Y = y_i = \varphi(x_i)) \text{ тотожні.}$$

Отже, $P(Y = y_i) = P(X = x_i) = p_i$ і шуканий закон розподілу функції має вигляд

Y	$y_1 = \varphi(x_1)$	$y_2 = \varphi(x_2)$	$\dots$	$y_n = \varphi(x_n)$
P	p_1	p_2	$\dots$	p_n

де $y_i < y_{i+1}$, $(i = \overline{1, n})$; $\sum_{i=1}^n p_i = 1$.

Приклад 1. Задано закон розподілу випадкового аргументу X

X	1	2	3
P	0,25	0,40	0,35

Знайти закон розподілу функції $Y = X^2$.

Розв'язання. Для заданої множини значень X функція $y = x^2$ приймає значення:
 $y_1 = 1^2 = 1$, $y_2 = 2^2 = 4$, $y_3 = 3^2 = 9$

Отже, закон розподілу функції має вигляд

Y	1	4	9
P	0,25	0,40	0,35

б) Якщо серед чисел $y_i = \varphi(x_i)$ є однакові, то кожній групі однакових чисел y_i відводимо в таблиці один стовпчик, а відповідні ймовірності додаємо.

Приклад 2. Задано закон розподілу випадкового аргументу X

X	-2	-1	0	1	2
P	0,15	0,15	0,4	0,12	0,18

Знайти закон розподілу функції $Y = X^2$.

Розв'язання. Можливі значення функції $y_1 = 4$, $y_2 = 1$, $y_3 = 0$, $y_4 = 1$, $y_5 = 4$.

Маємо три різні числові значення $y=0$, $y=1$, $y=4$, причому два останні повторюються, тому додаємо їх відповідні ймовірності.

Отже, закон розподілу функції $Y = X^2$

Y	0	1	2
P	0,4	0,27	0,33

Нехай X – неперервна випадкова величина, щільність розподілу якої відома: $f(x)$. Знайдемо щільність розподілу $g(y)$ випадкової величини $Y = \varphi(X)$.

Припустимо, що функція $\varphi(x)$ монотонно зростає, неперервна і диференційовна на (a, b) .

Функція розподілу $G(y)$ випадкової величини Y визначається за формулою

$$G(y) = P(Y < y) \quad (3)$$

Якщо функція $y = \varphi(x)$ монотонно зростає, то подія $(Y < y)$ еквівалентна події $(X < \varphi^{-1}(y))$, де $x = \varphi^{-1}(y)$ - функція, обернена функції $y = \varphi(x)$, яка теж монотонно зростаюча, неперервна і диференційовна.

Отже,

$$G(y) = P(Y < y) = P(X < \varphi^{-1}(y)) = \int_a^{\varphi^{-1}(y)} f(x) dx \quad (4)$$

Диференціюючи цей вираз по y , отримаємо щільність розподілу випадкової величини Y

$$g(y) = G'(y) = f[\varphi^{-1}(y)][\varphi^{-1}(y)]' \quad (5)$$

Якщо функція $\varphi(x)$ на (a, b) монотонно спадає, то подія $(Y < y)$ еквівалентна події $(X > \varphi^{-1}(y))$.

Отже,
$$G(y) = \int_{\varphi^{-1}(y)}^b f(x) dx \quad (6)$$

і
$$g(y) = -f[\varphi^{-1}(y)][\varphi^{-1}(y)]'. \quad (7)$$

Оскільки щільність розподілу не може бути від'ємною, то формули (5) і (7) можна об'єднати в одну

$$g(y) = f[\varphi^{-1}(y)][\varphi^{-1}(y)]' \quad (8)$$

Отже, щільність розподілу функції (2) визначається формулою (8).

Зауваження. Якщо функція $y = \varphi(x)$ немонотонна, тобто обернена функція $x = \varphi^{-1}(y)$ неоднозначна, то весь інтервал зміни значень функції розбиваємо на інтервали монотонності і для знаходження $g(y)$ підсумовуємо за формулою (8) по всіх інтервалах монотонності.

Приклад 3. Випадкова величина X розподілена нормально з параметрами $m_x = 0$, $\sigma_x = 1$

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

Знайти закон розподілу функції $Y = X^3$.

Розв'язання. Функція $y = x^3$ монотонна на $(-\infty, +\infty)$. Обернена до неї функція $\varphi^{-1}(y) = \sqrt[3]{y}$.

Знаходимо похідну $(\varphi^{-1}(y))' = \frac{1}{3y^{2/3}}$.

Отже, за формулою (8) маємо

$$g(y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{\sqrt[3]{y^2}}{2}} \cdot \frac{1}{3y^{2/3}} = \frac{1}{3\sqrt{2\pi} \cdot y^{2/3}} \cdot e^{-\frac{y^{2/3}}{2}}.$$

Приклад 4. Випадкова величина X розподілена нормально з параметрами $m_x = 0$, $\sigma_x = 1$.

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

Знайти закон розподілу функції $Y = X^2$.

Розв'язання. Обернена функція $x = \varphi^{-1}(y)$ неоднозначна:

$$x_1 = -\sqrt{y} \text{ для } x < 0 \quad \text{і} \quad x_2 = \sqrt{y} \text{ для } x \geq 0.$$

$$\text{Отже, } g(y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-y/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} \cdot e^{-y/2} \left| \frac{1}{2\sqrt{y}} \right| = \frac{1}{\sqrt{2\pi y}} e^{-y/2}, \quad y > 0.$$

10.2. Закон розподілу лінійної функції

Нехай $Y = \alpha X + \beta$, де α, β – не випадкові величини. Оскільки $y = \alpha x + \beta$ монотонна функція, то обернена функція $x = \frac{y - \beta}{\alpha}$ теж монотонна. Маємо $x'_y = \frac{1}{\alpha}$ і за формулою (8)

$$g(y) = f\left(\frac{y - \beta}{\alpha}\right) \cdot \frac{1}{|\alpha|}. \quad (9)$$

Вираз (9) показує, що лінійне перетворення випадкової величини X тотожне зміні масштабу зображення кривої $f(x)$ і переносу початку координат в нову точку. Вигляд кривої $f(x)$ при такому перетворенні не змінюється.

Покажемо, що лінійна функція $Y = \alpha X + \beta$ розподілена нормально, якщо аргумент X – нормально розподілена випадкова величина: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$. Припустивши $\alpha > 0$, знайдемо похідну $x'_y = \frac{1}{\alpha}$. За формулою (9) запишемо щільність розподілу функції

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\frac{y-\beta}{\alpha}-a)^2}{2\sigma^2}} \cdot \frac{1}{\alpha} \quad \text{або} \quad g(y) = \frac{1}{(\sigma\alpha)\sqrt{2\pi}} e^{-\frac{(y-(\beta+\alpha a))^2}{2(\alpha\sigma)^2}}. \quad (10)$$

Таким чином, лінійна функція теж розподілена нормально з параметрами $\sigma_y = \sigma\alpha$ і $m_y = \beta + \alpha a$.

10.3. Математичне сподівання і дисперсія функції випадкової величини

В задачах, пов'язаних з оцінкою точності роботи автоматичних систем, тощо, доводиться розглядати функції однієї або декількох випадкових величин. В найпростішому випадку задача ставиться таким чином: на вхід деякого технічного пристрою поступає випадковий сигнал X , і технічний пристрій, виконуючи над X деяке функціональне перетворення φ , дає на виході випадкову величину Y , яка є функцією від X

$$Y = \varphi(X) \quad (11)$$

Розглянемо таку задачу: за відомим законом розподілу випадкового аргумента X знайти числові характеристики функції $Y = \varphi(X)$, не знаходячи закону розподілу Y .

Нехай X дискретна випадкова величина, задана рядом розподілу

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

де $x_i < x_{i+1}$, $(i = \overline{1, n})$; $\sum_{i=1}^n p_i = 1$, тоді і функція $Y = \varphi(X)$ теж дискретна випадкова величина.

Складемо таблицю значень величини Y і ймовірностей цих значень

$$P(Y = y_i = \varphi(x_i)) = P(X = x_i) = p_i$$

Y	$y_1 = \varphi(x_1)$	$y_2 = \varphi(x_2)$	$\dots$	$y_n = \varphi(x_n)$
P	p_1	p_2	$\dots$	p_n

Ця таблиця не є рядом розподілу, оскільки деякі значення y_i можуть повторюватися. Проте математичне сподівання можна визначати за формулою

$$M[Y] = M[\varphi(x)] = \sum_{i=1}^n \varphi(x_i) p_i \quad (12)$$

Дійсно, величина (2) не може змінитися від того, що під знаком суми деякі члени будуть наперед об'єднані, а порядок членів змінений.

Міркуючи аналогічно, отримаємо формулу для обчислення дисперсії

$$D[Y] = D[\varphi(x)] = \sum_{i=1}^n (\varphi(x_i) - M[Y])^2 p_i$$

або робочу формулу

$$D[Y] = M[Y^2] - (M[Y])^2 = \sum_{i=1}^n \varphi^2(x_i) p_i - \left(\sum_{i=1}^n \varphi(x_i) p_i \right)^2 \quad (13)$$

Якщо аргумент X - неперервна випадкова величина, то і функція $Y = \varphi(X)$ теж неперервна випадкова величина, математичне сподівання якої визначається за формулою

$$M[Y] = \int_{-\infty}^{+\infty} \varphi(x) f(x) dx \quad (14)$$

якщо інтеграл (4) збігається,

$$\text{а дисперсія } D[Y] = \int_{-\infty}^{+\infty} (\varphi(x) - M[Y])^2 f(x) dx \quad \text{або} \quad D[Y] = M[Y^2] - (M[Y])^2 \quad (15)$$

де початковий момент другого порядку $M[Y^2] = \int_{-\infty}^{+\infty} \varphi^2(x) f(x) dx$.

Таким чином, для знаходження числових характеристик функції $Y = \varphi(X)$ досить знати закон розподілу її аргумента.

Зауваження. Надалі будемо записувати тільки вираз для початкового моменту другого порядку, оскільки дисперсія обчислюється за робочою формулою $D[Y] = M[Y^2] - (M[Y])^2$

Приклад 1. Задано закон розподілу випадкового аргументу X

X	-2	-1	0	1	2
P	0,15	0,15	0,4	0,12	0,18

Знайти математичне сподівання і дисперсію функції $Y = X^2$, не знаходячи її закону розподілу.

Розв'язання. Можливі значення функції $y_1 = 4$, $y_2 = 1$, $y_3 = 0$, $y_4 = 1$, $y_5 = 4$.

Складаємо таблицю можливих значень функції та ймовірностей цих значень

$Y = X^2$	4	1	0	1	4
P	0,15	0,15	0,4	0,12	0,18

Отже, $M[Y] = 4 \cdot 0,15 + 1 \cdot 0,15 + 1 \cdot 0,12 + 4 \cdot 0,18 = 1,59$.

$$M[Y^2] = 16 \cdot 0,15 + 1 \cdot 0,15 + 1 \cdot 0,12 + 16 \cdot 0,18 = 4,95.$$

$$D[Y] = 4,95 - (1,59)^2 \approx 2,42.$$

Приклад 2. Задана щільність розподілу аргументу $f(x) = \begin{cases} \sin x, & x \in (0, \pi/2) \\ 0, & x \notin (0, \pi/2) \end{cases}$.

Знайти математичне сподівання і дисперсію функції $Y = X^2$, не знаходячи її закону розподілу.

Розв'язання. $y = \varphi(x) = x^2$. Отже, $M[Y] = \int_0^{\pi/2} x^2 \sin x dx = \pi - 2$.

$$M[Y^2] = \int_0^{\pi/2} x^4 \sin x \, dx = \frac{\pi^3}{2} - 12\pi + 24. \quad D[Y] = \frac{\pi^3}{2} - \pi^2 - 8\pi + 20.$$

В деяких випадках для знаходження числових характеристик функції $Y = \varphi(X)$ не потрібно навіть знати закону розподілу аргумента, а тільки його числові характеристики.

10.4. Математичне сподівання і дисперсія лінійної функції

Нехай випадкові величини X та Y зв'язані між собою лінійно:

$$Y = \alpha X + \beta, \quad (16)$$

де α, β – не випадкові величини, причому відомі $M[X]$ і $D[X]$.

Враховуючи властивості математичного сподівання і дисперсії, отримаємо

$$M[Y] = \alpha M[X] + \beta, \quad (17)$$

тобто математичне сподівання лінійної функції є лінійною функцією математичного сподівання її аргументу, а дисперсія

$$D[Y] = \alpha^2 D[X] \quad (18)$$

10.5. Математичне сподівання і дисперсія мінімальної із двох величин: випадкової X і не випадкової α

Випадкова величина Y як мінімальна із двох величин зв'язана з X залежністю

$$Y = \min\{X, \alpha\} = \begin{cases} X, & X < \alpha \\ \alpha, & X \geq \alpha \end{cases}. \quad (19)$$

Знайдемо її математичне сподівання і дисперсію.

Нехай X – неперервна випадкова величина, щільність розподілу якої $f(x)$.

За формулою (14) знайдемо математичне сподівання

$$M[Y] = \int_{-\infty}^{+\infty} \min\{x, \alpha\} f(x) dx = \int_{-\infty}^{\alpha} x f(x) dx + \alpha \int_{\alpha}^{+\infty} f(x) dx = \int_{-\infty}^{\alpha} x f(x) dx + \alpha \cdot [1 - F(\alpha)] \quad (20)$$

де $F(x)$ – функція розподілу випадкової величини X .

Початковий момент другого порядку

$$M[Y^2] = \int_{-\infty}^{+\infty} [\min\{x, \alpha\}]^2 f(x) dx = \int_{-\infty}^{\alpha} x^2 f(x) dx + \alpha^2 \int_{\alpha}^{+\infty} f(x) dx = \int_{-\infty}^{\alpha} x^2 f(x) dx + \alpha^2 \cdot [1 - F(\alpha)] \quad (21)$$

Нехай X – дискретна випадкова величина, яка приймає значення $X = x_i$, ($i = \overline{1, n}$) з відповідними ймовірностями $p_i = P(X = x_i)$.

За формулою (12) знайдемо математичне сподівання

$$M[Y] = \sum_{i=1}^n \min\{x_i, \alpha\} \cdot p_i = \sum_{i=1}^{(\alpha)} x_i \cdot p_i + \alpha \sum_{i=(\alpha)+1}^n p_i \quad (22)$$

де (α) - номер максимального з можливих значень випадкової величини X , яке не більше α :

$$x_{(\alpha)} \leq \alpha.$$

Початковий момент другого порядку

$$M[Y^2] = \sum_{i=1}^n [\min\{x_i, \alpha\}]^2 \cdot p_i = \sum_{i=1}^{(\alpha)} x_i^2 \cdot p_i + \alpha^2 \sum_{i=(\alpha)+1}^n p_i \quad (23)$$

Приклад 1. Напруга X , яка подається на вхід обмежувача, розподілена за нормальним законом з параметрами a і σ . Обмежувач працює за принципом $Y = \min\{X, \alpha\}$. Знайти математичне сподівання і дисперсію напруги Y на виході обмежувача.

Ввести змінну $\tau = \frac{\alpha - a}{\sigma}$

Приклад 2. В обчислювальний центр за зміну надходить випадкове число X інформаційних документів, яке розподілене за законом Пуассона з параметром $\lambda = M[X]$. Число інформаційних документів, що обробляються в ОЦ за зміну, не може перевищувати α (ціле число): $Y = \min\{X, \alpha\}$. Знайти математичне сподівання і дисперсію випадкової величини Y .

10.6. Математичне сподівання і дисперсія максимальної із двох величин: випадкової X і невідповідної α

Випадкова величина Y як максимальна із двох величин зв'язана з X залежністю

$$Y = \max\{X, \alpha\} = \begin{cases} \alpha, & X < \alpha \\ X, & X \geq \alpha \end{cases} \quad (24)$$

Знайдемо її математичне сподівання і дисперсію.

Нехай X - неперервна випадкова величина, щільність розподілу якої $f(x)$.

За формулою (14) знайдемо математичне сподівання

$$M[Y] = \int_{-\infty}^{+\infty} \max\{x, \alpha\} f(x) dx = \alpha \int_{-\infty}^{\alpha} f(x) dx + \int_{\alpha}^{+\infty} x f(x) dx = \alpha F(\alpha) + \int_{\alpha}^{+\infty} x f(x) dx. \quad (25)$$

Початковий момент другого порядку

$$M[Y^2] = \int_{-\infty}^{+\infty} [\max\{x, \alpha\}]^2 f(x) dx = \alpha^2 \int_{-\infty}^{\alpha} f(x) dx + \int_{\alpha}^{+\infty} x^2 f(x) dx. \quad (26)$$

10.7. Математичне сподівання і дисперсія модуля функції випадкової величини

$$\text{Нехай } Y = |X - \alpha| = \begin{cases} \alpha - X, & X < \alpha \\ X - \alpha, & X \geq \alpha \end{cases} \quad (27)$$

де X - неперервна випадкова величина, щільність якої $f(x)$.

Знайдемо математичне сподівання і дисперсію випадкової величини Y .

За формулою (14) математичне сподівання

$$\begin{aligned} M[Y] &= \int_{-\infty}^{+\infty} |x - \alpha| f(x) dx = \int_{-\infty}^{\alpha} (\alpha - x) f(x) dx + \int_{\alpha}^{+\infty} (x - \alpha) f(x) dx = \\ &= \alpha[2F(\alpha) - 1] + 2 \int_{\alpha}^{+\infty} x f(x) dx - M[X] \end{aligned} \quad (28)$$

Початковий момент другого порядку

$$M[Y^2] = \int_{-\infty}^{+\infty} |x - \alpha|^2 f(x) dx = \int_{-\infty}^{+\infty} (x - \alpha)^2 f(x) dx = M[X^2] - 2\alpha M[X] + \alpha^2 \quad (29)$$

Питання для самоперевірки

1. Дати означення функції випадкового аргументу.
2. Як скласти закон розподілу функції дискретного випадкового аргументу?
3. Як знайти щільність розподілу функції неперервного випадкового аргументу?
4. Записати закон розподілу лінійної функції.
5. Записати формули для математичного сподівання і дисперсії функції випадкової дискретної величини.
6. Записати формули для математичного сподівання і дисперсії функції випадкової неперервної величини.
7. Записати формули для математичного сподівання і дисперсії лінійної функції.

11. БАГАТОВИМІРНІ ВИПАДКОВІ ВЕЛИЧИНИ. СПІЛЬНА ФУНКЦІЯ РОЗПОДІЛУ ДВОХ ВИПАДКОВИХ ВЕЛИЧИН. НЕЗАЛЕЖНІ ВИПАДКОВІ ВЕЛИЧИНИ

Досі ми розглядали випадкові величини, можливі значення яких визначалися одним числом (одновимірні випадкові величини). Наприклад, число очок, яке може випасти при киданні грального кубика (дискретна одновимірна випадкова величина) або відстань від снаряда до місця падіння снаряда (неперервна одновимірна випадкова величина).

Часто доводиться мати справу з величинами, можливі значення яких визначаються двома або більш числами. Такі величини називаються n -мірними випадковими величинами; n -мірну випадкову величину можна розглядати як систему n випадкових величин. У даному контексті використовується також термін *багатовимірний випадковий вектор* $\vec{x} = (x_1, x_2, \dots, x_n)$, де кожна з величин $x_1, x_2, \dots, x_n$ називається складовою (компонентою). Аналогічно

одновимірним випадковим величинам розрізняють **дискретні** багатовимірні випадкові величини (їх складові дискретні) і **неперервні** багатовимірні випадкові величини, складові яких неперервні.

Приклад. Верстат штампує сталеві плити. Якщо контрольованими розмірами є довжина X , ширина Y і висота Z плити, то ми маємо тривимірну випадкову величину (X, Y, Z) .

Зупинимося детальніше на двовимірних випадкових величинах.

11.1. Закон розподілу ймовірності двовимірної випадкової величини

Законом розподілу дискретної двовимірної випадкової величини називають перелік можливих значень цієї величини, тобто пар чисел (x_i, y_j) , де x_i і y_j

$(i = 1, 2, \dots, n ; j = 1, 2, \dots, m)$ - можливі значення величин X і Y , відповідно, і ймовірностей

$p_{ij} = p(x_i, y_j)$ їх спільної появи $P(X = x_i, Y = y_j)$.

Двовимірна дискретна випадкова величина $\xi = (X, Y)$ задається у вигляді **таблиці розподілу** виду :

$\begin{matrix} \cdot & \cdot & Y \\ X & \cdot & \cdot \end{matrix}$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

де перший рядок таблиці вказує можливі значення складової Y , а перший стовпець - усі можливі значення складової X .

Оскільки події $(X = x_i, Y = y_j)$ ($i = 1, 2, \dots, n, j = 1, 2, \dots, m$) утворюють повну

групу, то $\sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1$.

Знаючи закон розподілу двовимірної дискретної випадкової величини, можна знайти закони розподілу кожної з її складових. Так, наприклад, ймовірність того, що X набуде

значення x_k , дорівнює
$$P(X = x_k) = \sum_{j=1}^m p_{kj}$$

11.2. Спільна функція розподілу двох випадкових величин

Функція $F(x, y)$, що визначає для кожної пари чисел x, y ймовірність того, що X набуде значення меншого x , і при цьому Y набуде значення меншого y , називається **спільною функцією розподілу** двох випадкових величин $F(x, y) = P(X < x, Y < y)$.

Геометрично цю рівність можна представити так: $F(x, y)$ - це ймовірність того, що випадкова точка (X, Y) потрапить в нескінченний квадрант з вершиною (x, y) , розташований лівіше і нижче за цю вершину.

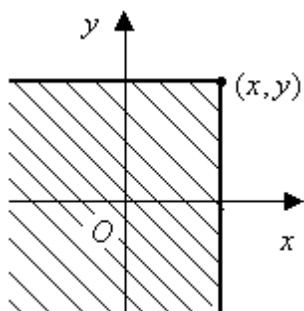


Рисунок 1 – Геометричне тлумачення для функції розподілу $F(x, y)$

Властивості спільної функції розподілу двох випадкових величин

1. Значення спільної функції розподілу задовольняють нерівності:

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

2. $F(x, y)$ - неспадна функція по кожному аргументу, тобто

$$F(x_2, y) \geq F(x_1, y), \text{ якщо } x_2 \geq x_1;$$

$$F(x, y_2) \geq F(x, y_1), \text{ якщо } y_2 \geq y_1.$$

Спільна функція розподілу має наступні граничні значення:

$$F(-\infty, y) = 0; \quad F(x, -\infty) = 0;$$

$$F(-\infty, -\infty) = 0; \quad F(\infty, \infty) = 1.$$

3. При $x = \infty$ або $y = \infty$ спільна функція розподілу системи стає функцією розподілу однієї із

складових :
$$F(x, \infty) = F_1(x); F(\infty, y) = F_2(y)$$

11.3. Щільність спільного розподілу ймовірності неперервної двовимірної випадкової величини

Неперервну двовимірну випадкову величину можна задати за допомогою щільності розподілу. **Щільність спільного розподілу ймовірності** $f(x, y)$ двовимірної неперервної випадкової величини (X, Y) - це друга змішана частинна похідна від функції розподілу $F(x, y)$:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

Знаючи щільність спільного розподілу $f(x, y)$, можна знайти спільну функцію

розподілу $F(x, y)$ за формулою
$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(x, y) dx dy$$
, виходячи з означення

щільності розподілу двовимірної неперервної випадкової величини (X, Y) .

Зміст щільності спільного розподілу ймовірності : ймовірність попадання випадкової точки в прямокутник (з вершиною в точці (x, y) і сторонами Δx і Δy дорівнює добутку $f(x, y)\Delta x\Delta y$, коли сторони цього прямокутника прямують до нуля.

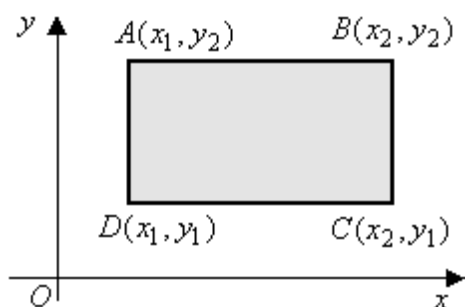


Рисунок 2 – Ймовірність попадання до квадрата ABCD

У зв'язку з цим, ймовірність попадання випадкової точки в **довільну область D** дорівнює подвійному інтегралу по області D від функції $f(x, y)$:

$$P((X, Y) \in D) = \iint_D f(x, y) dx dy$$

Властивості двовимірної щільності ймовірності

1. Двовимірна щільність ймовірності невід'ємна: $f(x, y) \geq 0$.
2. Подвійний невластний інтеграл з нескінченними межами від двовимірної щільності ймовірності дорівнює одиниці:
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$$

11.4. Незалежні випадкові величини

Дві випадкові величини називаються *незалежними*, якщо закон розподілу однієї з них не залежить від того, яких можливих значень набула інша величина.

Теорема. Для того, щоб випадкові величини X і Y були незалежними, необхідно і достатньо, щоб функція розподілу системи (X, Y) дорівнювала добутку функцій розподілу складових: $F(X, Y) = F_1(X)F_2(Y)$.

Наслідок. Для того, щоб випадкові величини X і Y були незалежними, необхідно і достатньо, щоб щільність спільного розподілу системи (X, Y) дорівнювала добутку щільності розподілу складових: $f(X, Y) = f_1(X)f_2(Y)$.

Для незалежних випадкових величин справедливі співвідношення

$$M(XY) = M(X)M(Y)$$

$$D(X \pm Y) = D(X) + D(Y).$$

Питання для самоперевірки

1. Дати означення багатовимірної випадкової величини.
2. Як поділяються багатовимірні випадкові величини?
3. Що таке закон розподілу двовимірної випадкової величини?
4. Що називається спільною функцією розподілу двох випадкових величин?
5. Записати властивості спільної функції розподілу двох випадкових величин.
6. Дати означення щільності спільного розподілу ймовірності неперервної двовимірної випадкової величини.
7. Записати властивості щільності спільного розподілу ймовірності неперервної двовимірної випадкової величини
8. Які випадкові величини називаються незалежними?

12. УМОВНІ ЗАКОНИ РОЗПОДІЛУ СКЛАДОВИХ СИСТЕМИ ДИСКРЕТНИХ І НЕПЕРЕРВНИХ ВИПАДКОВИХ ВЕЛИЧИН. ЧИСЛОВІ ХАРАКТЕРИСТИКИ СИСТЕМИ ДВОХ ВИПАДКОВИХ ВЕЛИЧИН

12.1. Умовний закон розподілу складових системи випадкових величин

Раніше ми розглядали ймовірності добутку залежних і незалежних подій.

Випадкові величини також можуть бути залежними або незалежними одна від одної. Для залежних подій

$$P(AB) = P(A)P(B/A),$$

звідки, умовна ймовірність події B , за умови, що подія A вже відбулася, обчислюється так:

$$P(B/A) = P(AB)/P(A) \quad (*)$$

Для незалежних подій $P(AB) = P(A) \cdot P(B)$.

Щоб охарактеризувати залежність між складовими двовимірної випадкової величини, введено поняття умовного розподілу.

Нехай задана двовимірна *дискретна випадкова величина* $\xi = (X, Y)$ у вигляді **таблиці розподілу** вигляду :

$\begin{matrix} \cdot & \cdot & Y \\ X & \cdot & \cdot \end{matrix}$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

де перший рядок таблиці вказує можливі значення складової Y , а перший стовпець - усі можливі значення складової X .

Якщо, наприклад, величина Y набула значення $Y = y_1$, то величина

X при цьому набуває одне зі своїх можливих значень $x_1, x_2, \dots, x_n$. Позначимо умовну ймовірність того, що X набуде значення x_i за умови, що $Y = y_1$ через $p(x_i|y_1)$. Ця ймовірність, в загальному випадку, не дорівнює безумовній ймовірності $p(x_i)$.

Взагалі, умовні ймовірності позначаються $p(x_i|y_j)$.

Умовним розподілом складової X , при $Y = y_j$, називають сукупність умовних ймовірностей $p(x_1|y_j), p(x_2|y_j), \dots, p(x_n|y_j)$, обчислених у припущенні, що подія $Y = y_j$ вже наступила. Аналогічно визначається умовний розподіл складової Y .

Умовний закон розподілу X , в припущенні, що подія $Y = y_j$ вже відбулася, може бути знайдений за формулою, що аналогічна формулі (*):

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, m \quad (**)$$

Сума ймовірностей умовного розподілу дорівнює одиниці, так як і безумовного.

Умовний закон розподілу *неперервної випадкової величини* можна задавати як функцією розподілу, так і щільністю.

Для системи двох випадкових величин (X, Y) справедлива теорема множення законів розподілу, яка аналогічна теоремі множення ймовірностей подій.

Теорема . Щільність розподілу системи двох випадкових величин дорівнює щільності розподілу однієї з величин, помножену на умовну щільність розподілу іншої величини, обчисленої за умови, що перша величина набула заданого значення:

$$f(x, y) = f_1(x) \cdot f(y|x) \\ \text{або } f(x, y) = f_2(y) \cdot f(x|y)$$

З приведених формул теореми виразимо умовні закони розподілу через безумовні закони розподілу:

$$f(x|y) = \frac{f(x, y)}{f_2(y)}, \quad f(y|x) = \frac{f(x, y)}{f_1(x)}.$$

Числові характеристики системи двох випадкових величин

Для системи двох випадкових величин, окрім математичних сподівань і дисперсій складових використовують і інші характеристики, такі як кореляційний момент і коефіцієнт кореляції.

12.2. Кореляційний момент

Характеристикою залежності між випадковими величинами X і Y служить математичне сподівання добутку відхилень X і Y від їх математичних сподівань. Це так званий **кореляційний момент** або **коваріація** :

$$\mu_{xy} = \text{cov}(X; Y) = M[(X - M_X)(Y - M_Y)] = M(X \cdot Y) - M_X M_Y$$

Для обчислення кореляційного моменту дискретних величин використовують формулу:

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

а для неперервних величин - формулу:
$$\mu_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ((x - M(X))(y - M(Y))) f(x, y) dx dy$$

Якщо кореляційний момент випадкових величин X і Y відмінний від нуля, то ці величини є залежними.

Теорема. Кореляційний момент двох незалежних випадкових величин X і Y дорівнює нулю.

Доведення. Оскільки X і Y - незалежні випадкові величини, то їх відхилення $X - M(X)$ і $Y - M(Y)$ також незалежні. Користуючись властивостями математичного сподівання (математичне сподівання добутку незалежних випадкових величин дорівнює добутку математичних сподівань співмножників) і відхилення (математичне сподівання відхилення дорівнює нулю), отримаємо:

$$\mu_{xy} = M((X - M_X) \cdot (Y - M_Y)) = M(X - M_X) \cdot M(Y - M_Y) = 0.$$

З визначення кореляційного моменту виходить, що він має розмірність, рівну добутку розмірності величин X і Y . Іншими словами, величина кореляційного моменту залежить від одиниць вимірювання випадкових величин і для одних і тих же двох величин величина кореляційного моменту має різні значення залежно від того, в яких одиницях були виміряні величини.

Для усунення цього недоліку вводять нову числову характеристику - коефіцієнт кореляції.

12.3. Коефіцієнт кореляції

Коефіцієнтом кореляції r_{xy} випадкових величин X і Y називають

відношення кореляційного моменту до добутку середніх квадратичних відхилень цих величин:

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

Для *незалежних* X і Y коефіцієнт кореляції дорівнює нулю.

Властивості коефіцієнта кореляції

1. $-1 \leq r_{xy} \leq 1$
2. Якщо $r_{xy} = 1$, то $Y = kX + b$, де k і b - константи, $k > 0$.
3. Якщо $r_{xy} = -1$, то $Y = kX + b$, де $k < 0$.

Коефіцієнт кореляції r_{xy} досягає своїх граничних значень - 1 і 1 в тому і тільки у тому випадку, якщо між X і Y є лінійна залежність.

При $|r_{xy}| < 1$ лінійна залежність відсутня, хоча по мірі наближення до одиниці спільний розподіл X , Y має тенденцію концентруватися поблизу деякої прямої лінії і величину $|r_{xy}|$ можна вважати мірою близькості до повної лінійної залежності між X і Y .

Введемо поняття **кореляційної залежності** між X і Y . Дві випадкові величини називають **корельованими**, якщо їх коваріація або коефіцієнт кореляції відмінні від нуля, і **некорельованими** інакше.

Говорять, що між X і Y існує пряма кореляційна залежність, якщо із зростанням X випадкова величина Y має тенденцію зростати (при великих X з більшою ймовірністю зустрічаються великі значення Y). Якщо із зростанням X випадкова величина Y має тенденцію спадати, говорять, що між X і Y існує обернена кореляційна залежність.

Чим ближче $|r_{XY}|$ до одиниці, тим тісніше глибина кореляційної залежності.

Приклад: Знайти коефіцієнт кореляції між величинами X і Y , спільний закон розподілу яких заданий наступною таблицею

$\begin{matrix} \cdot & \cdot & Y \\ X & \cdot & \cdot \end{matrix}$	1	2	3	4	Σ
10	0,2	0,02	0,01	0	0,23
20	0,03	0,3	0,02	0	0,35
30	0,02	0,1	0,2	0,1	0,42
Σ	0,25	0,42	0,23	0,1	1

Знаходимо:

$$M_X = \sum_i x_i \cdot p_i = 10 \cdot 0,23 + 20 \cdot 0,35 + 30 \cdot 0,42 = 21,9$$

$$M_Y = \sum_i y_i \cdot q_i = 1 \cdot 0,25 + 2 \cdot 0,42 + 3 \cdot 0,23 = 2,18$$

$$M(X \cdot Y) = \sum_{i,j} x_i \cdot y_j \cdot p_{ij} = 10 \cdot 0,2 + 20 \cdot 0,02 + 30 \cdot 0,01 + 20 \cdot 0,03 + 40 \cdot 0,3 + 60 \cdot 0,02 + 30 \cdot 0,02 + 60 \cdot 0,1 + 90 \cdot 0,2 + 120 \cdot 0,1 = 53,1$$

Аналогічно, знайдемо $D(X), D(Y)$ і за ними $\sigma(X) = 7,835$, $\sigma(Y) = 0,92$. Остаточного отримаємо

$$r_{xy} = \frac{M(X \cdot Y) - M(X) \cdot M(Y)}{\sigma(X) \cdot \sigma(Y)} = \frac{53,1 - 21,9 \cdot 2,18}{7,835 \cdot 0,92} = 0,742$$

Питання для самоперевірки

1. Що називається умовним розподілом однієї складової двовимірної дискретної випадкової величини?
2. Сформулюйте теорему множення законів розподілу.
3. Як обчислюються безумовні і умовні закони розподілу складових неперервної двовимірної випадкової величини?
4. Дати означення кореляційного моменту.
5. Сформулювати властивості кореляційного моменту.
6. Дати означення коефіцієнта кореляції.
7. Сформулювати властивості коефіцієнта кореляції.

8. Які величини називаються корельованими?
9. Дати означення прямої і оберненої кореляційної залежності.

13. ВИБІРКОВИЙ МЕТОД АНАЛІЗУ ВЛАСТИВОСТЕЙ ГЕНЕРАЛЬНОЇ СУКУПНОСТІ

13.1. Предмет та основні поняття математичної статистики

Предметом математичної статистики є вивчення випадкових подій і випадкових величин за результатами спостережень. Сукупність предметів або явищ, об'єднаних якою-небудь загальною ознакою, називається **статистичною сукупністю**. Результатом спостережень над статистичною сукупністю є **статистичні дані** - відомості про те, яких значень набула у результаті спостережень ознака (випадкова величина X), що нас цікавить.

Обробка статистичних даних методами математичної статистики призводить до встановлення певних закономірностей, властивих масовим явищам. При цьому **точність** статистичних висновків підвищується із зростанням числа спостережень.

Статистичні дані, як правило, є рядом значень $\{x_1, x_2, \dots, x_n\}$ деякої випадкової величини. Обробка цього ряду значень є першим етапом дослідження випадкової величини.

Перша задача математичної статистики - **вказати способи збору і групування статистичних даних**, отриманих в результаті спостережень або в результаті спеціально поставлених експериментів.

Другою задачею математичної статистики є розробка **методів аналізу** статистичних даних залежно від цілей дослідження. До цієї задачі відносяться:

- Оцінка невідомої **ймовірності події**; оцінка невідомої **функції розподілу**; оцінка **параметрів розподілу**, вид якого відомий; оцінка залежності випадкової величини від однієї або декількох випадкових величин і тому подібне
- Перевірка статистичних гіпотез про вид **невідомого** розподілу або про **величину параметрів розподілу**, вид якого відомий.

У сучасній математичній статистиці є багато спільного з **наукою про ухвалення рішень в умовах невизначеності**, оскільки вона розробляє способи визначення числа необхідних випробувань до початку дослідження (планування експерименту), в процесі дослідження (послідовний аналіз) і вирішує багато інших аналогічних завдань.

13.2. Вибірковий метод і його основні поняття. Випадкова вибірка, об'єм вибірки

Нехай вимагається вивчити сукупність **однорідних** об'єктів відносно деякої якісної або кількісної ознаки, що характеризує ці об'єкти. Наприклад, для партії деталей якісною ознакою може служити стандартність деталі, а кількісною - контрольований розмір деталі.

В принципі, можливе проведення суцільного обстеження, тобто обстеження усіх об'єктів. На практиці таке обстеження застосовується рідко, наприклад,

- ✓ через велике числа об'єктів
- ✓ через дорожнечу проведення операції контролю,
- ✓ через те, що контроль часто пов'язаний з руйнуванням об'єкту (перевірка електролампи на довговічність її роботи), і так далі

У таких випадках випадково відбирається і вивчається **обмежене** число об'єктів з сукупності.

Вибірковою сукупністю або **випадковою вибіркою** називають сукупність випадково відібраних об'єктів.

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

Об'ємом сукупності (вибіркової або генеральної) називають число об'єктів цієї сукупності. Наприклад, якщо з 1000 деталей відбирається для обстеження 100, то об'єм генеральної сукупності $N=1000$, а об'єм вибірки $n = 100$.

Приклад: Число одиниць товару N , виробленого деяким підприємством впродовж року, є генеральна сукупність. Для дослідження якості продукції на практиці розглядається вибірка, що складається з n одиниць товару. Ознакою, або випадковою величиною, може бути число одиниць товару, що задовольняють сертифікаційним вимогам.

При складанні вибірки можна поступати двома способами: після того, як об'єкт відібраний і досліджений, його можна повернути або не повертати в генеральну сукупність. У зв'язку з цим вибірки підрозділяються на **повторні** і **безповторні**.

Повторною називають вибірку, при якій відібраний об'єкт (перед відбором наступного) повертається в генеральну сукупність. При **безповторній** вибірці відібраний об'єкт в генеральну сукупність не повертається.

Для того, щоб за даними вибірки можна було досить упевнено судити про ознаку генеральної сукупності, що нас цікавить, необхідно, щоб об'єкти вибірки правильно його представляли. Вибірка повинна правильно представляти пропорції генеральної сукупності, тобто вибірка має бути репрезентативною (представницькою). Приклад - вивчення громадської думки.

Через закон великих чисел можна стверджувати, що вибірка буде репрезентативною, якщо її здійснити випадково: кожен об'єкт вибірки відібраний випадково з генеральної сукупності, якщо усі об'єкти мають однакову ймовірність потрапити у вибірку.

Якщо об'єм вибірки досить великий, а вибірка складає лише незначну частину сукупності, то відмінність між повторною і безповторною вибіркою стирається.

13.3. Способи відбору

На практиці застосовуються різні способи відбору, які можна підрозділити на два види:

- Відбір, що не вимагає розчленовування генеральної сукупності на частини. Сюди відносяться а) простий випадковий безповторний відбір і б) простий випадковий повторний відбір.
- Відбір, при якому генеральна сукупність розбивається на частини. Сюди відносяться а) типовий відбір, б) механічний відбір і в) серійний відбір.

Простим випадковим називають відбір, при якому об'єкти витягаються по одному з генеральної сукупності. Здійснити такий відбір для генеральної сукупності з N об'єктів можна, наприклад, за допомогою запису на картках номерів від 1 до N , наступному перемішуванню карток і вийманню їх навмання. При цьому обстеженню підлягають об'єкти, що мають номери, співпадаючі з номерами карток. Якщо картки повертаються в пачку, то маємо просту випадкову повторну вибірку, інакше - просту безповторну. При великому об'ємі генеральної сукупності раціональнішим є використання таблиць випадкових чисел. Наприклад, щоб вибрати 50 об'єктів з пронумерованої генеральної сукупності, відкривають будь-яку сторінку таблиці випадкових чисел і виписують 50 чисел підряд; у вибірку потрапляють ті об'єкти, номери яких співпадають з виписаними випадковими числами. Якщо випадкове число таблиці перевершує число N , таке число пропускають. При проведенні безповторної вибірки пропускають також випадкові числа, що вже зустрічалися раніше.

Типовим називають відбір, при якому об'єкти відбираються не з усієї генеральної сукупності, а з кожної її "типової" частини. Наприклад, якщо деталі виготовлені на декількох верстатах, то відбір виробляють з продукції кожного верстата окремо.

Механічним називають відбір, при якому генеральна сукупність механічно ділиться на стільки груп, скільки об'єктів повинно увійти до вибірки, а з кожної групи вибирається один об'єкт. Наприклад, якщо треба відібрати 20% виготовлених верстатом деталей, то відбирають кожну п'яту деталь.

Серійним називають відбір, при якому об'єкти відбирають з генеральної сукупності не по одному, а "серіями", які піддаються суцільному обстеженню. Наприклад, якщо виробниці виробляються великою групою верстатів-автоматів, то піддають суцільному обстеженню продукцію тільки декількох верстатів. Цим видом відбору користуються тоді, коли обстежувана ознака коливається в різних серіях незначно.

На практиці часто застосовують **комбінований** відбір, при якому поєднуються вказані вище способи. Наприклад, розбивають генеральну сукупність на серії однакового об'єму, потім простим випадковим відбором вибирають декілька серій і, нарешті, з кожної серії простим випадковим відбором витягають окремі об'єкти.

13.4. Варіаційний ряд для дискретних і неперервних випадкових величин

Нехай з генеральної сукупності витягнута вибірка, причому значення досліджуваного параметра x_1 спостерігалось n_1 раз, x_2 - n_2 раз і так далі. При цьому $\sum_i n_i = n$ об'єм

вибірки. Спостережувані значення x_i називають **варіантами**, а послідовність варіант, записаних в зростаючому порядку - **варіаційним рядом**. Числа спостережень називають **частотами**, а їх відношення до об'єму вибірки n_i/n - **відносними частотами**. Варіаційний ряд можна представити таблицею:

X	x_1	x_2	..	x_m
n	n_1	n_2	.	n_m

Статистичним розподілом вибірки називають перелік варіант і відповідних їм відносних частот. Статистичний розподіл можна представити як

X	x_1	x_2	..	x_m
w	w_1	w_2	.	w_m

де відносні частоти $w_k = \frac{n_k}{n}$.

Зауважимо, що в теорії ймовірності під розподілом розуміють відповідність між можливими значеннями випадкової величини і їх **ймовірностями**, а в математичній статистиці - відповідність між спостережуваними **варіантами** і їх **частотами** або **відносними частотами**.

Приведений спосіб представлення статистичних даних застосовують у разі дискретних випадкових величин. Для неперервних випадкових величин зручніше розбити відрізок $[a, b]$ можливих значень випадкової величини на часткові напівінтервали $\Delta_k = [a_{k-1}, a_k)$, ($k = 1, \dots, m$) (Δ_m замкнутий також і справа) за допомогою деякої системи точок $a = a_0 < a_1 < \dots < a_m = b$.

Часто розбиття $[a, b]$ проводять на рівні частини, тоді $\Delta_k = [a + (k-1)h, a + kh)$, $k = 1, \dots, m$,

де $h = \frac{b-a}{m}$

В якості частот n_k тепер потрібно брати кількість спостережуваних значень, що потрапили до кожного з часткових інтервалів Δ_k . Варіаційний ряд має у такому разі вигляд

X	Δ_1	Δ_2	..	Δ_m
n	n_1	n_2	.	n_m

а статистичний розподіл -

X	Δ_1	Δ_2	..	Δ_m
n	w_1	w_2	.	w_m

Число інтервалів k часто вибирають на підставі формули Стерджерса

$$k = 1 + 1,4 \ln n$$

13.5. Полігон і гістограма

Графічно статистичний розподіл представляється зокрема, за допомогою **полігону** і **гістограми**.

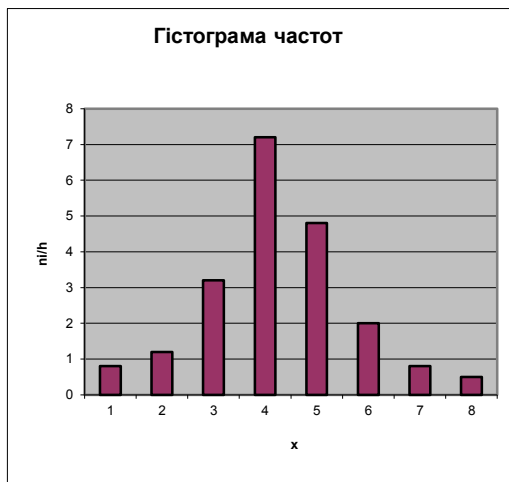
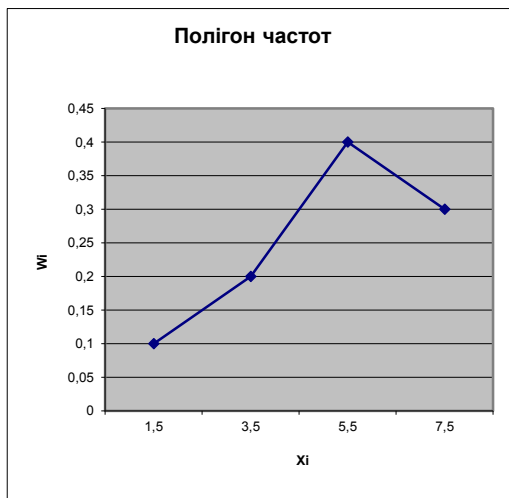
Полігоном частот називають ламану лінію, відрізки якої сполучають точки $(x_1; n_1), (x_2; n_2), \dots, (x_k; n_k)$. Для побудови полігону частот на осі абсцис відкладають варіанти x_i , а на осі ординат - відповідні ним частоти n_i і сполучають точки $(x_i; n_i)$ відрізками прямих.

Полігон відносних частот будується аналогічно, за винятком того, що на осі ординат відкладаються відносні частоти w_i .

У разі неперервної ознаки будується **гістограма**, для чого інтервал, в якому знаходяться усі спостережувані значення ознаки, розбивають на декілька часткових інтервалів завдовжки h і знаходять для кожного часткового інтервалу n_i - суму частот варіант, що потрапили в i -й інтервал.

Гістограмою частот називають ступінчасту фігуру, що складається з прямокутників, основами якої служать часткові інтервали завдовжки h , а висоти дорівнюють відношенню n_i / h . Для побудови гістограми частот на осі абсцис відкладають часткові інтервали, а над ними проводять відрізки, паралельні осі абсцис на відстані (висоті) n_i / h . Площа i -го прямокутника рівна $hn_i / h = n_i$ - сумі частот варіант i -го інтервалу, тому площа гістограми частот дорівнює сумі усіх частот, тобто **об'єму вибірки**.

У разі гістограми відносних частот по осі ординат відкладаються відносні частоти w_i , на осі абсцис - часткові інтервали, над ними проводять відрізки, паралельні осі абсцис на висоті w_i / h . Площа i -го прямокутника дорівнює відносній частоті варіант w_i , що потрапили в i -й інтервал. Тому площа гістограми відносних частот дорівнює сумі усіх відносних частот, тобто одиниці.



Питання для самоперевірки

1. Що вивчає математична статистика?
2. Сформулювати дві основні задачі математичної статистики.
3. Дати означення вибіркової та генеральної сукупності.
4. Як визначається об'єм вибірки?
5. Що таке повторна і безповторна вибірки?
6. Перерахувати види відборів.
7. Дати означення варіанти, частоти, варіаційного ряду.
8. Що таке статистичний розподіл вибірки?
9. Якими способами можна графічно представити вибірку?
10. Як будується полігон частот? Відносних частот?
11. Як будується гістограма частот? Відносних частот?

14. ПОНЯТТЯ ПРО СТАТИСТИЧНІ ОЦІНКИ ВИПАДКОВИХ ВЕЛИЧИН

14.1. Емпірична функція розподілу

Нехай відомий статистичний розподіл частот кількісної ознаки X . Позначимо через n_x число спостережень, при яких спостерігалось значення ознаки, менше x і через n - загальне число спостережень. Очевидно, відносна частота події $X < x$ рівна n_x/n і є функцією x . Оскільки ця функція знаходиться емпіричним (дослідним) шляхом, то її називають емпіричною.

Емпіричною функцією розподілу (функцією розподілу вибірки) називають функцію $F^*(x)$, що визначає для кожного значення x відносну частоту події $X < x$. Таким чином, за визначенням $F^*(x) = n_x/n$, де n_x - число варіант, менших x , n - об'єм вибірки.

На відміну від емпіричної функції розподілу вибірки, функцію розподілу $F(x)$ генеральної сукупності називають **теоретичною функцією розподілу**. Відмінність між цими функціями полягає в тому, що теоретична функція $F(x)$ визначає **ймовірність** події $X < x$, тоді як емпірична - **відносну частоту** цієї ж події.

При зростанні n відносна частота події $X < x$, тобто $F^*(x)$ прямує за ймовірністю до вірогідності $F(x)$ цієї події. Іншими словами $\lim_{n \rightarrow \infty} P[|F(x) - F^*(x)| < \varepsilon] = 1 \quad (\varepsilon > 0)$

Властивості емпіричної функції розподілу :

- 1) Значення емпіричної функції належать відрізку $[0,1]$
- 2) $F^*(x)$ - неспадна функція
- 3) Якщо x_1 - найменша варіанту, то $F^*(x) = 0$ при $x \leq x_1$, якщо x_k - найбільша варіанту, то $F^*(x) = 1$ при $x > x_k$.

Емпірична функція розподілу вибірки служить для оцінки теоретичної функції розподілу генеральної сукупності.

Приклад. Побудуємо емпіричну функцію по розподілу вибірки :

Варіанти	2	6	10
Частоти	12	18	30

Знайдемо об'єм вибірки : $12+18+30=60$. Найменша варіанта рівна 2, тому $F^*(x)=0$ при $x \leq 2$. Значення $x < 6$, тобто $x_1 = 2$, спостерігалось 12 разів, отже, $F^*(x)=12/60=0,2$ при $2 < x \leq 6$. Аналогічно, значення $X < 10$, тобто $x_1 = 2$ і $x_2 = 6$ спостерігалися $12+18=30$ разів, тому $F^*(x) = 30/60 = 0,5$ при $6 < x \leq 10$. Оскільки $x=10$ - найбільша варіанту, то $F^*(x)=1$ при $x > 10$. таким чином, шукана емпірична функція має вигляд:

$$F^*(x) = \begin{cases} 0 & x \leq 2 \\ 0,2 & 2 < x \leq 6 \\ 0,5 & 6 < x \leq 10 \\ 1 & x > 10 \end{cases}$$

14.2. Найважливіші властивості статистичних оцінок

Нехай вимагається вивчити деяку кількісну ознаку генеральної сукупності. Допустимо, що з теоретичних міркувань вдалося встановити, який саме розподіл має ознаку і необхідно оцінити параметри, якими воно визначається. Наприклад, якщо ознака, що вивчається, розподілена в генеральній сукупності нормально, то треба оцінити математичне очікування і середнє квадратичне відхилення; якщо ознака має розподіл Пуассона - то необхідно оцінити параметр λ .

Зазвичай є лише ці вибірки, наприклад значення кількісної ознаки $x_1, x_2, \dots, x_n$, отримані в результаті n незалежних спостережень. Розглядаючи $x_1, x_2, \dots, x_n$ як незалежні випадкові величини $X_1, X_2, \dots, X_n$ можна сказати, що **знайти статистичну оцінку невідомого параметра теоретичного розподілу - означає знайти функцію від спостережуваних випадкових величин, яка дає наближене значення оцінюваного параметра**. Наприклад, для оцінки математичного сподівання нормального розподілу роль функції виконує середнє арифметичне $\bar{X} = (X_1 + X_2 + \dots + X_n) / n$

Для того, щоб статистичні оцінки давали коректні наближення оцінюваних параметрів, вони повинні задовольняти деяким вимогам, серед яких найважливішими є вимоги незміщеності і спроможності оцінки.

Нехай Θ^* - статистична оцінка невідомого параметра Θ теоретичного розподілу. Нехай за вибіркою об'єму n знайдена оцінка Θ_1^* . Повторимо дослід, тобто витягнемо з генеральної

сукупності іншу вибірку того ж об'єму і за її даними отримаємо іншу оцінку Θ_2^* . Повторюючи дослід багаторазово, отримаємо різні числа $\Theta_1^*, \Theta_2^*, \dots, \Theta_k^*$. Оцінку Θ^* можна розглядати як випадкову величину, а числа $\Theta_1^*, \Theta_2^*, \dots, \Theta_k^*$ - як її можливі значення.

Якщо оцінка Θ^* дає наближене значення Θ з **надлишком**, тобто кожне число Θ_i^* ($i=1, \dots, k$) більше істинного значення Θ то, як наслідок, математичне сподівання (середнє значення) випадкової величини Θ^* більше, ніж Θ : $M(\Theta^*) > \Theta$. Аналогічно, якщо Θ^* дає оцінку з **недостачею**, то $M(\Theta^*) < \Theta$.

Таким чином, використання статистичної оцінки, математичне сподівання якої не дорівнює оцінюваному параметру, привело б до систематичних (одного знаку) помилок. Якщо, навпаки, $M(\Theta^*) = \Theta$, то це гарантує від систематичних помилок.

Незміщеною називають статистичну оцінку Θ^* , математичне сподівання якої дорівнює оцінюваному параметру Θ при будь-якому об'ємі вибірки $M(\Theta^*) = \Theta$.

Зміщеною називають оцінку, що не задовольняє цій умові.

Незміщеність оцінки ще не гарантує отримання хорошого наближення для оцінюваного параметра, оскільки можливі значення Θ_i^* можуть бути **сильно розсіяні** навколо свого середнього значення, тобто дисперсія $D(\Theta)$ може бути значною. В цьому випадку знайдена за даними однієї вибірки оцінка, наприклад Θ_1^* , може виявитися значно віддаленою від середнього значення $\bar{\Theta}^*$, а значить, і від самого оцінюваного параметра.

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

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

Спроможною називається статистична оцінка, яка при $n \rightarrow \infty$ прямує за ймовірністю до оцінюваного параметра. Наприклад, якщо дисперсія незміщеної оцінки при $n \rightarrow \infty$ прямує до нуля, то така оцінка виявляється і спроможною.

14.3. Вибіркові середнє і дисперсія

Нехай для вивчення генеральної сукупності відносно кількісної ознаки X витягнута вибірка об'єму n .

Вибірковим середнім $\bar{x}_B$ називають середнє арифметичне значення ознаки вибіркової сукупності. Якщо усі значення $x_1, x_2, \dots, x_n$ ознаки вибірки об'єму n різні, то $\bar{x}_B = (x_1 + x_2 + \dots + x_n) / n$.

Якщо значення ознаки $x_1, x_2, \dots, x_k$ мають частоти $n_1, n_2, \dots, n_k$ відповідно, причому $n_1 + n_2 + \dots + n_k = n$,

то $\bar{x}_B = (n_1 x_1 + n_2 x_2 + \dots + n_k x_k) / n = (1/n) \sum_{i=1}^k n_i x_i$.

Вибіркове середнє, знайдене за даними однієї вибірки, дорівнює певному числу. При витяганні інших вибірок того ж об'єму вибіркове середнє мінятиметься від вибірки до вибірки. Тобто вибіркове середнє можна розглядати як випадкову величину і говорити про її розподіли (теоретичний і емпіричний) і про числові характеристики цього розподілу (наприклад, про математичне сподівання і дисперсію).

Для того, щоб охарактеризувати розсіяння спостережуваних значень кількісної ознаки вибірки навколо середнього значення $\overline{x_B}$ вводиться **вибіркова дисперсія**. **Вибірковою дисперсією** D_B називають середнє арифметичне квадратів відхилення спостережуваних значень ознаки від їх середнього значення $\overline{x_B}$. Якщо усі значення $x_1, x_2, \dots, x_n$ ознаки вибірки об'єму n різні, то

$$D_B = (1/n) \sum_{i=1}^n (x_i - \overline{x_B})^2.$$

Якщо значення ознаки $x_1, x_2, \dots, x_k$ мають частоти $n_1, n_2, \dots, n_k$ відповідно, причому $n_1 + n_2 + \dots + n_k = n$, то

$$D_B = (1/n) \sum_{i=1}^k n_i (x_i - \overline{x_B})^2.$$

Аналогічно вибіровим середньому і дисперсії визначаються **генеральні середнє і дисперсія**, характеризуючі генеральну сукупність в цілому. Для розрахунку цих характеристик досить в наведених вище співвідношеннях замінити об'єм вибірки n на об'єм генеральної сукупності N .

Фундаментальне значення для практики має знаходження середнього і дисперсії ознаки генеральної сукупності за відповідними відомими вибіровими параметрами. Можна показати, що вибіркове середнє є незміщеною спроможною оцінкою генерального середнього. В той же час, незміщеною спроможною оцінкою генеральної дисперсії виявляється не вибіркова дисперсія D_B ,

а так звана "**виправлена**" **вибіркова дисперсія**, що дорівнює $s^2 = \frac{n}{n-1} D_B$.

Таким чином, в якості оцінок генерального середнього і дисперсії в математичній статистиці приймають вибіркове середнє і виправлену вибірову дисперсію.

14.4. Надійність і довірчий інтервал

Досі ми розглядали точкові оцінки, тобто такі оцінки, які визначаються одним числом. При вибірці малого об'єму точкова оцінка може значно відрізнятись від оцінюваного параметра, що призводить до грубих помилок. У зв'язку з цим при невеликому об'ємі вибірки користуються інтервальними оцінками.

Інтервальною називають оцінку, що визначається двома числами - кінцями інтервалу. Нехай знайдена за даними вибірки статистична характеристика Θ^* служить оцінкою невідомого параметра Θ . Очевидно, Θ^* тим точніше визначає параметр Θ , чим менше абсолютна величина різниці $|\Theta - \Theta^*|$. Іншими словами, якщо $\delta > 0$ і $|\Theta - \Theta^*| < \delta$, то чим менше δ , тим точніша оцінка. Таким чином, додатне число δ характеризує **точність оцінки**.

Статистичні методи *не дозволяють* стверджувати, що оцінка Θ^* задовольняє нерівності $|\Theta - \Theta^*| < \delta$; можна говорити лише про ймовірність, з якою ця нерівність здійснюється.

Надійністю (довірчою вірогідністю) оцінки Θ за Θ^* називають ймовірність γ , з якою здійснюється нерівність $|\Theta - \Theta^*| < \delta$. Звичайно надійність оцінки задається заздалегідь, причому в якості γ беруть число, близьке до одиниці, - як правило 0,95; 0,99 або 0,999.

Нехай ймовірність того, що $|\Theta - \Theta^*| < \delta$ дорівнює γ :

$$P[|\Theta - \Theta^*| < \delta] = \gamma.$$

Замінімо нерівність $|\Theta - \Theta^*| < \delta$ рівносильною йому подвійною нерівністю

$$P[\Theta^* - \delta < \Theta < \Theta^* + \delta] = \gamma.$$

Це співвідношення слід розуміти так: ймовірність того, що інтервал $(\Theta^* - \delta, \Theta^* + \delta)$ містить в собі (покриває) невідомий параметр Θ , дорівнює γ .

Таким чином, **довірчим** називають інтервал $(\Theta^* - \delta, \Theta^* + \delta)$, який покриває невідомий параметр із заданою надійністю γ .

Величину $1 - \gamma = \alpha$ називають **рівнем значущості** або **ймовірністю помилки**.

Для побудови інтервальної оцінки параметра необхідно знати закон його розподілу як випадкової величини.

Питання для самоперевірки

1. Дати означення емпіричної функції розподілу.
2. Записати властивості емпіричної функції розподілу.
3. Що означає знайти статистичну оцінку невідомого параметра теоретичного розподілу?
4. Яким вимогам повині задовольняти статистичні оцінки?
5. Що таке систематичні помилки?
6. Яка оцінка називається незміщеною? Зміщеною?
7. Яка оцінка називається спроможною?
8. Яка оцінка називається ефективною?
9. Дати означення вибіркового середнього і вибіркової дисперсії.
10. Дати означення генерального середнього і генеральної дисперсії.
11. Записати формулу для виправленої вибіркової дисперсії.
12. Що таке інтервальна оцінка?
13. Що таке точність оцінки?
14. Що таке надійність оцінки?
15. Дати означення довірчого інтервалу.
16. Що таке рівень значущості?

15. ДОВІРЧІ ІНТЕРВАЛИ ДЛЯ МАТЕМАТИЧНОГО СПОДІВАННЯ І ДИСПЕРСІЇ

15.1. Довірчий інтервал для математичного сподівання нормального розподілу при відомій дисперсії

Нехай кількісна ознака X генеральної сукупності розподілена нормально, причому середнє квадратичне відхилення σ цього розподілу **відоме**. Вимагається оцінити невідоме математичне сподівання a за вибіркоvim середнім $\bar{X}$. Знайдемо довірчі інтервали, що покривають параметр a з надійністю γ .

Розглядатимемо вибіркoве середнє $\bar{X}$ як випадкову величину $\bar{X}$ (оскільки $\bar{X}$ змінюється від вибірки до вибірки) і вибіркoві значення $X_1, X_2, \dots, X_n$ - як однаково розподілені незалежні випадкові величини $X_1, X_2, \dots, X_n$ (ці числа також змінюються від вибірки до вибірки). Іншими словами, математичне сподівання кожної з цих величин дорівнює a і середнє квадратичне відхилення - σ . Оскільки випадкова величина X розподілена нормально, то і вибіркoве середнє $\bar{X}$ також розподілене нормально. Параметри розподілу $\bar{X}$ дорівнюють $M(\bar{X}) = a$, $\sigma(\bar{X}) = \sigma / \sqrt{n}$.

Будемо вимагати, щоб виконувалося співвідношення $P(|\bar{X} - a| < \delta) = \gamma$

де γ - задана надійність. Використаємо формулу

$$P(|X - a| < \delta) = 2 \Phi(\delta / \sigma).$$

Замінюючи X на $\bar{X}$ і σ на $\sigma(\bar{X}) = \sigma / \sqrt{n}$, отримаємо

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

де $t = \delta \sqrt{n} / \sigma$. Виразивши з останньої рівності δ , отримаємо $P(|\bar{X} - a| < t\sigma / \sqrt{n}) = 2 \Phi(t)$

Оскільки ймовірність P задана і дорівнює γ , остаточно маємо

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

Таким чином, з надійністю γ можна стверджувати, що довірчий інтервал $(\bar{x} - t\sigma / \sqrt{n} < a < \bar{x} + t\sigma / \sqrt{n})$ покриває невідомий параметр a , причому точність оцінки дорівнює $t\sigma / \sqrt{n}$.

Число t визначається з рівності $\Phi(t) = \gamma / 2$; за таблицею функції Лапласа знаходять аргумент t , якому відповідає значення функції Лапласа, рівне $\gamma / 2$.

Відмітимо два моменти: 1) при зростанні об'єму вибірки n число δ спадає і, отже, точність оцінки збільшується, 2) збільшення **надійності** оцінки $\gamma / 2 = \Phi(t)$ призводить до збільшення t (оскільки функція Лапласа зростаюча функція) і, отже, до зростання δ , тобто **збільшення надійності** оцінки спричиняє за собою **зменшення її точності**.

Якщо вимагається оцінити математичне сподівання з наперед заданою точністю δ і надійністю γ , то мінімальний об'єм вибірки, який забезпечить цю точність, знаходять за формулою

$$n = t^2 \sigma^2 / \delta^2,$$

що випливає з рівності $\delta = t \sigma / \sqrt{n}$.

15.2. Довірчий інтервал для математичного сподівання нормального розподілу при невідомій дисперсії

Нехай кількісна ознака X генеральної сукупності розподілена нормально, причому середнє квадратичне відхилення σ цього розподілу *невідоме*. Вимагається оцінити невідоме математичне сподівання за допомогою довірчих інтервалів.

Виявляється, що за даними вибірки можна побудувати випадкову величину $T = \frac{\bar{X} - a}{S / \sqrt{n}}$

яка має розподіл Стюдента з $k = n - 1$ ступенями свободи. У останньому виразі - $\bar{X}$ - вибіркове середнє, S - виправлене середнє квадратичне відхилення, n - об'єм вибірки; можливі значення випадкової величини T ми позначатимемо через t . Щільність розподілу Стюдента має вигляд

$S(t, n) = B_n \left[1 + \frac{t^2}{n-1} \right]^{-n/2}$, де B_n деяка стала, така, що виражається через гамма - функції.

Декілька слів про розподіл Стюдента. Нехай $\xi_0, \xi_1, \dots, \xi_n$ - незалежні стандартні

нормальні величини. Тоді випадкова величина $t_n = \frac{\xi_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n \xi_i^2}}$

має розподіл Стюдента з n ступенями вільності. При зростанні числа ступенів вільності розподіл Стюдента прямує до нормального розподілу і вже при $n \geq 30$ використання нормального розподілу дає гарні результати.

Як видно, розподіл Стюдента визначається параметром n - об'ємом вибірки (чи, що те ж саме - числом ступенів вільності $k = n - 1$) і не залежить від невідомих параметрів a, σ . Оскільки

$S(t, n)$ - парна функція від t , то ймовірність виконання нерівності $\left| \frac{\bar{X} - a}{S / \sqrt{n}} \right| < \gamma$

визначається таким чином: $P\left(\left|\frac{\bar{X} - a}{S / \sqrt{n}}\right| < t_\gamma\right) = 2 \int_0^{t_\gamma} S(t, n) dt = \gamma$.

Замінивши нерівність в круглих дужках подвійною нерівністю, отримаємо вираз для шуканого

довірчого інтервалу $P\left(\bar{X} - \frac{t_\gamma S}{\sqrt{n}} < a < \bar{X} + \frac{t_\gamma S}{\sqrt{n}}\right) = \gamma$

Отже, за допомогою розподілу Стюдента знайдений довірчий інтервал $\bar{x} - \frac{t_\gamma s}{\sqrt{n}} < a < \bar{x} + \frac{t_\gamma s}{\sqrt{n}}$, що покриває невідомий параметр a з надійністю γ . За таблицею розподілу Стюдента і заданим n і γ можна знайти t_γ і використовуючи знайдені за вибіркою $\bar{X}$ і S , можна визначити довірчий інтервал.

Приклад. Кількісна ознака X генеральної сукупності розподілена нормально. За вибіркою об'єму $n = 16$ знайдені генеральне середнє $\bar{x} = 20,2$ і виправлене середнє квадратичне відхилення $s = 0,8$. Вимагається оцінити невідоме математичне сподівання за допомогою довірчого інтервалу з надійністю 0,95.

Розв'язання. Знайдемо t_γ за таблицею розподілу Стюдента, використовуючи значення $\gamma = 0,95$; $n = 16$. Цей параметр виявляється рівним 2,13. Знайдемо межі довірчого інтервалу :

$$\bar{x} - t_\gamma s / \sqrt{n} = 20,2 - 2,13 \cdot 0,8 / \sqrt{16} = 19,77$$

$$\bar{x} + t_\gamma s / \sqrt{n} = 20,2 + 2,13 \cdot 0,8 / \sqrt{16} = 20,63$$

Тобто з надійністю 0,95 невідомий параметр a знаходиться в довірчому інтервалі $19,77 < a < 20,63$

Можна показати, що при зростанні об'єму вибірки n розподіл Стюдента прямує до нормального. Тому практично при $n > 30$ можна замість нього користуватися нормальним розподілом. При малих n це призводить до значних помилок.

15.3. Довірчий інтервал для оцінки середнього квадратичного відхилення σ нормального розподілу

Нехай кількісна ознака X генеральної сукупності розподілена нормально і вимагається оцінити невідоме генеральне середнє квадратичне відхилення σ за виправленим вибірковою середнім квадратичним відхиленням s . Знайдемо довірчі інтервали, що покривають параметр σ із заданою надійністю γ .

Будемо вимагати, щоб виконувалося співвідношення

$$P(|\sigma - s| < \delta) = \gamma \quad \text{або} \quad P(s - \delta < \sigma < s + \delta) = \gamma$$

Перетворимо подвійну нерівність $s - \delta < \sigma < s + \delta$ в рівносильну нерівність $s(1 - \delta / s) < \sigma < s(1 + \delta / s)$ і позначимо $\delta/s = q$. Маємо $s(1 - q) < \sigma < s(1 + q)$
(A)

і необхідно знайти q . Для цього введемо в розглядання випадкову величину $\chi = \frac{S}{\sigma} \sqrt{n-1}$

Виявляється, величина $\frac{S^2}{\sigma^2} (n-1)$ розподілена згідно з χ^2 законом з $n - 1$ ступенями вільності.

Декілька слів про розподіл χ^2 . Якщо $\xi_i, i = 1, 2, \dots, n$ - незалежні стандартні нормальні

величини, то говорять, що випадкова величина $\chi_n^2 = \sum_{i=1}^n \xi_i^2$

має розподіл **χ^2 -квадрат** з n ступенями вільності.

Щільність розподілу χ має вигляд

$$R(\chi, n) = \frac{\chi^{n-1} e^{-\chi^2/2}}{2^{(n-3)/2} \Gamma\left(\frac{n-1}{2}\right)}$$

Цей розподіл не залежить від оцінюваного параметра σ , а залежить тільки від об'єму вибірки n .

Перетворимо нерівність (А) так, щоб вона набрала вигляду $\chi_1 < \chi < \chi_2$. Ймовірність

цієї нерівності дорівнює заданій ймовірності γ , тобто $\int_{\chi_1}^{\chi_2} R(\chi, n) d\chi = \gamma$.

Припускаючи, що $q < 1$, перепишемо (А) у вигляді

$$\frac{1}{S(1+q)} < \frac{1}{\sigma} < \frac{1}{S(1-q)},$$

далі, помножимо усі члени нерівності на $S\sqrt{n-1}$:

$$\frac{\sqrt{n-1}}{(1+q)} < \frac{S\sqrt{n-1}}{\sigma} < \frac{\sqrt{n-1}}{(1-q)} \quad \text{або} \quad \frac{\sqrt{n-1}}{(1+q)} < \chi < \frac{\sqrt{n-1}}{(1-q)}.$$

Ймовірність того, що ця нерівність, а також рівносильна їй нерівність (А) буде справедлива, дорівнює

$$\int_{\sqrt{n-1}/(1+q)}^{\sqrt{n-1}/(1-q)} R(\chi, n) d\chi = \gamma.$$

З цього рівняння можна за заданими n та γ знайти q , використовуючи наявні розрахункові таблиці. Обчисливши за вибіркою S і знайшовши за таблицею q , отримаємо шуканий інтервал (А1), що покриває σ із заданою надійністю γ .

Приклад. Кількісна ознака X генеральної сукупності розподілена нормально. За вибіркою об'єму $n = 25$ знайдене виправлене середнє квадратичне відхилення $s = 0,8$. Знайти довірчий інтервал, що покриває генеральне середнє квадратичне відхилення σ з надійністю 0,95.

Розв'язання. За заданими γ і σ за таблицею знаходимо значення $q = 0,32$. Шуканий довірчий інтервал є

$$0,8(1 - 0,32) < \sigma < 0,8(1 + 0,32).$$

Ми припускали, що $q < 1$. Якщо це не так, то ми прийдемо до співвідношень

$$\sqrt{n-1}/(1+q) < \chi < \infty,$$

і значення $q > 1$ може бути знайдене з рівняння

$$\int_{\sqrt{n-1}/(1+q)}^{\infty} R(\chi, n) d\chi = \gamma$$

Питання для самоперевірки

1. Записати довірчий інтервал для математичного сподівання нормального розподілу при відомій дисперсії.
2. Як пов'язані між собою надійність і точність оцінки математичного сподівання нормального розподілу при відомій дисперсії? Довірчий інтервал для оцінки середнього квадратичного відхилення σ нормального розподілу. Записати характеристики біноміального розподілу.
3. Як знаходять мінімальний об'єм вибірки, який забезпечить точність оцінки математичного сподівання з наперед заданою точністю δ і надійністю γ для нормального розподілу з відомою дисперсією?
4. Записати довірчий інтервал для математичного сподівання нормального розподілу при невідомій дисперсії.
5. Записати довірчий інтервал для оцінки середнього квадратичного відхилення σ нормального розподілу.

16. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ. КРИТЕРІЙ ЗГОДИ ПІРСОНА ПРО ВИД РОЗПОДІЛУ.

16.1. Перевірка статистичних гіпотез

Раніше розглядалося завдання побудови довірчих інтервалів для невідомих параметрів генеральної сукупності. Розглянемо питання *перевірки статистичних гіпотез*.

Перевірка статистичних гіпотез є найважливішим етапом процесу ухвалення рішення в управлінській діяльності, дозволяючи проводити підготовчий етап майбутніх дій з урахуванням реальних характеристик процесу виробництва, контролю якості продукції, комерційної діяльності, і тому подібне.

Як відомо, **закон розподілу** визначає кількісні характеристики генеральної сукупності.

Якщо закон розподілу **невідомий**, але є підстави припустити, що він має певний вигляд (наприклад, А), то висувають гіпотезу: генеральна сукупність розподілена згідно із законом А. В цій гіпотезі йдеться про вид передбачуваного розподілу.

Часто закон розподілу відомий, але невідомі його *параметри*. Якщо є підстави припустити, що невідомий параметр Θ дорівнює певному значенню Θ_0 , то може висуватися гіпотеза $\Theta = \Theta_0$. В цій гіпотезі йдеться про *передбачувану величину параметра* відомого розподілу.

Можливі і інші гіпотези: про рівність параметрів двох або декількох розподілів, про незалежність вибірок і т. д.

Приведемо декілька завдань, які можуть бути вирішені за допомогою перевірки статистичних гіпотез.

1. Використовується два методи вимірювання однієї і тієї ж величини. Перший метод дає оцінки $X_1, X_2, \dots, X_m$ цієї величини, другий - $Y_1, Y_2, \dots, Y_n$. Вимагається визначити, чи забезпечують обидва методи *однакову точність вимірювання*.

2. Контроль точності роботи деякої виробничої системи. Отримувані характеристики продукції, що випускається, характеризуються деяким розкидом (дисперсією). Зазвичай величина цього розкиду не повинна перевищувати деякого заздалегідь заданого рівня. Вимагається визначити, чи забезпечує система (наприклад, лінія зборки або окремий верстат) *задану точність*.

Отже, *статистичною* називають гіпотезу про вид невідомого розподілу або про параметри відомих розподілів. Приклади статистичних гіпотез : генеральна сукупність розподілена за законом Пуассона; дисперсії двох нормальних розподілів рівні між собою.

Разом з висуненою гіпотезою завжди розглядають і гіпотезу, що суперечить їй. Якщо висунена гіпотеза буде відхилена, то приймається гіпотеза, що суперечить їй.

Нульовою (основною) називають висунуту гіпотезу H_0 .

Альтернативною (конкуруючою) називають гіпотезу H_1 , яка суперечить нульовій. Наприклад, якщо нульова гіпотеза полягає в припущенні, що математичне сподівання нормального розподілу дорівнює 5, то альтернативна гіпотеза, наприклад, може полягати в припущенні, що $a \neq 5$. Коротке це записують так: $H_0: a=5; H_1: a \neq 5$.

Простою називають гіпотезу, що містить тільки одне припущення. Наприклад, якщо λ - параметр показового розподілу, то гіпотеза $H_0: \lambda=3$ - проста. **Складною** називають гіпотезу, що складається з скінченного або нескінченного числа простих гіпотез. Наприклад, складна гіпотеза $H: \lambda > 3$ складається з нескінченної безлічі простих гіпотез виду $H_i: \lambda = b_i$, де b_i - будь-яке число, більше 3.

Висунута гіпотеза може бути правильною або неправильною, тому виникає необхідність її перевірки. Оскільки перевірку проводять статистичними методами, то її називають статистичною. У результаті статистичної перевірки гіпотези в двох випадках може бути прийняте неправильне рішення, тобто можуть бути допущені помилки двох родів.

Помилка першого роду полягає в тому, що буде **відхилена правильна** гіпотеза. **Помилка другого роду** полягає в тому, що буде **прийнята неправильна** гіпотеза. Слід зазначити, що наслідки помилок можуть виявитися різними. Якщо відхилено правильне рішення "продовжувати будівництво житлового будинку", то ця помилка першого роду спричинить матеріальний збиток; якщо ж прийнято неправильне рішення "продовжувати будівництво", незважаючи на небезпеку обвалу будинку, то ця помилка другого роду може привести до численних жертв. Іноді, навпаки, помилка першого роду спричиняє тяжчі наслідки.

Природно, правильне рішення може бути прийняте також в двох випадках, коли **приймається правильна** гіпотеза або **відхиляється невірна** гіпотеза.

Ймовірність здійснення помилки **першого роду** називають **рівнем значущості** і позначають α . Найчастіше рівень значущості приймають рівним 0,05 або 0,01. Якщо, наприклад, прийнятий рівень значущості 0,05, то це означає, що в п'яти випадках із ста є ризик допустити помилку першого роду (відкинути правильну гіпотезу).

16.2. Статистичний критерій

Для перевірки нульової гіпотези використовують спеціально підібрану випадкову величину, точний або наближений розподіл якої відомий.

Статистичним критерієм (чи просто критерієм) називають випадкову величину (K), яка служить для перевірки нульової гіпотези. Наприклад, якщо перевіряють гіпотезу про рівність дисперсій двох нормальних генеральних сукупностей, то як критерій K приймають відношення виправлених вибірових дисперсій
$$F = s_1^2 / s_2^2.$$

Очевидно, що ця величина випадкова, оскільки в різних дослідах виправлені дисперсії набувають різних, заздалегідь невідомих значень.

Спостережуваним значенням критерію $K_{\text{спост}}$ називають значення критерію, обчислене за вибірками. Наприклад, якщо в наведеному вище випадку $s_1^2 = 20$, $s_2^2 = 5$, то $K_{\text{спост}} = 20/5 = 4$.

Критична область. Область прийняття гіпотези. Критичні точки.

Після вибору певного критерію безліч усіх його можливих значень розбивають на дві підмножини, що не перетинаються, одна з яких містить значення критерію, при яких нульова гіпотеза **відкидається**, а інша - при яких вона **приймається**.

Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу **відкидають**.

Відповідно, **областю прийняття гіпотези** (областю допустимих значень) називають сукупність значень критерію, при яких гіпотезу **приймають**.

Основний принцип перевірки статистичних гіпотез можна сформулювати так: якщо спостережуване значення критерію належить критичній області - гіпотезу відкидають, якщо області прийняття гіпотези - гіпотезу приймають.

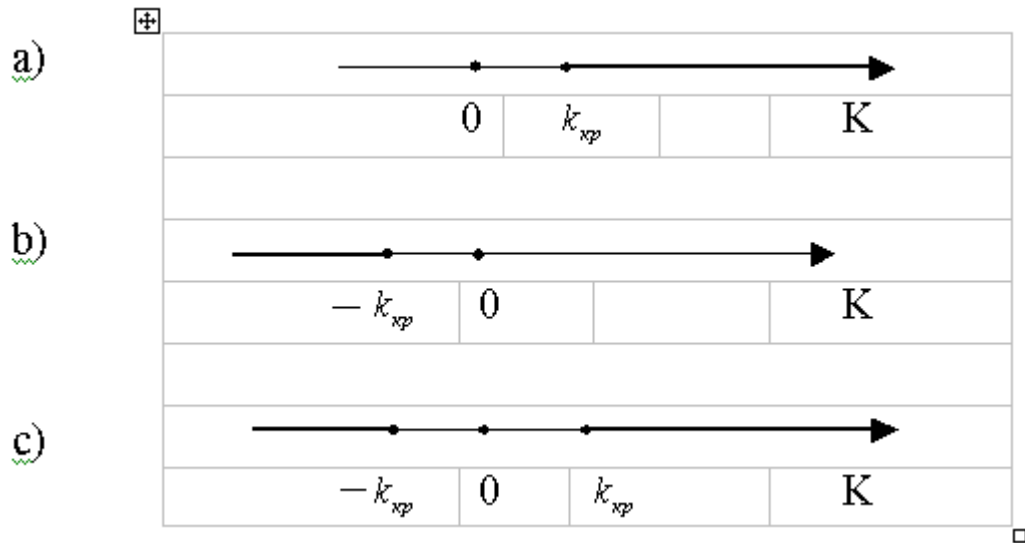
Оскільки критерій K - одновимірною випадкова величина, то усі її можливі значення належать деякому інтервалу і, відповідно, повинні існувати точки, що розділяють критичну область і область прийняття гіпотези. Такі точки називаються **критичними точками**.

Розрізняють **односторонню (правосторонню і лівосторонню)** і **двосторонню** критичні області.

Правосторонньою називають критичну область, що визначається нерівністю $K > k_{кр}$, де $k_{кр}$ - додатне число.

Лівосторонньою називають критичну область, що визначається нерівністю $K < k_{кр}$, де $k_{кр}$ - від'ємне число.

Двосторонньою називають критичну область, що визначається нерівностями $K < k_1$, $K > k_2$, де $k_2 > k_1$. Зокрема, якщо критичні точки симетричні відносно нуля, двостороння критична область визначається нерівностями $K < -k_{кр}$, $K > k_{кр}$ або рівносильною нерівністю $|K| > k_{кр}$. Відмінності між варіантами критичних областей ілюструє наступний малюнок.



Мал. 1. Різні варіанти критичних областей а) правостороння, б) лівостороння, в) двостороння

Ми підійшли до питання про *етапи перевірки статистичних гіпотез*.

Виділимо наступні *етапи*:

- Формулюється нульова гіпотеза H_0
- Визначається критерій K , за значеннями якого можна буде прийняти або відкинути H_0 і вибирається рівень значущості α
- За рівнем значущості визначається критична область
- За вибіркою обчислюється спостережуване значення критерію K , визначається, чи належить воно критичній області і на підставі цього приймається гіпотеза H_0 або альтернативна гіпотеза H_1 .

16.3. Критерій згоди Пірсона про вид розподілу.

Якщо закон розподілу невідомий, але є підстави припускати, що він має певний вигляд A , то перевіряють нульову гіпотезу: генеральна сукупність розподілена згідно з законом A . Перевірка цієї гіпотези проводиться за допомогою спеціально підібраної випадкової величини - *критерія згоди*.

Таким чином, *критерієм згоди* називають критерій перевірки гіпотези про *передбачуваний закон невідомого розподілу*.

Є декілька критеріїв згоди, причому найчастіше використовуваним є критерій згоди K Пірсона ("хі квадрат").

Нехай за вибіркою об'єму n отриманий емпіричний розподіл

Варіанти..... X_i	x_1 x_2 x_s
Емпіричні частоти.. n_i	n_1 n_2 n_s

Для визначеності розглянемо спочатку випадок перевірки статистичної гіпотези про нормальний розподіл генеральної сукупності.

Припустимо, що в припущенні нормального розподілу генеральної сукупності вчислені теоретичні частоти n_i^0 . При рівні значущості α вимагається перевірити нульову гіпотезу: генеральна сукупність розподілена нормально.

Як критерій перевірки нульової гіпотези приймемо випадкову величину

$$\chi^2 = \sum_i (n_i - n_i^0)^2 / n_i^0 \quad (A)$$

Природно, чим менше відрізняються емпіричні і теоретичні частоти, тим менша величина критерію, і, отже, він характеризує **близькість** емпіричного і теоретичного розподілів.

Доведено, що при $n \rightarrow \infty$ закон розподілу випадкової величини (A) прямує до закону розподілу χ^2 з k ступенями свободи незалежно від того, якому закону розподілу підпорядкована генеральна сукупність. Тому сам критерій називають **критерієм згоди** χ^2 .

Число ступенів свободи визначається з рівності $k = s - r - 1$, де s - число груп (часткових інтервалів) вибірки, r - число параметрів передбачуваного розподілу. Зокрема, якщо передбачуваний розподіл - нормальний, то оцінюють два параметри (математичне сподівання і середнє квадратичне відхилення), тому число ступенів свободи $k = n - 3$.

Побудуємо правосторонню критичну область, виходячи з вимоги, щоб ймовірність попадання критерію в цю область в припущенні справедливості нульової гіпотези дорівнювала прийнятому рівню значущості α : $P[\chi^2 > \chi_{кр}^2(\alpha; k)] = \alpha$.

Таким чином, правостороння критична область визначається нерівністю $\chi^2 > \chi_{кр}^2(\alpha; k)$, а область прийняття нульової гіпотези - відповідно нерівністю $\chi^2 < \chi_{кр}^2(\alpha; k)$. Позначимо значення критерію, обчисленого за даними спостережень, через $\chi_{набл}^2$ і сформулюємо правило перевірки нульової гіпотези :

Для того, щоб при заданому рівні значущості перевірити нульову гіпотезу H_0 : генеральна сукупність розподілена нормально, необхідно спочатку обчислити теоретичні частоти, а потім спостережуване значення критерію $\chi^2_{набл} = \sum_i (n_i - n_i^0)^2 / n_i^0$ і по таблиці критичних точок розподілу χ^2 , за заданим рівнем значущості α і числу ступенів свободи $k = n - 3$ знайти критичну точку $\chi^2_{кр}(\alpha; k)$. Якщо $\chi^2_{набл} < \chi^2_{кр}$ - немає підстав відхилити нульову гіпотезу. Інакше нульову гіпотезу відхиляють, вважаючи, що генеральна сукупність не розподілена за нормальним законом.

Відмітимо дві обставини.

- Об'єм вибірки має бути досить великим (не менше 50). Кожна група повинна містити не менше 5-8 варіант, а малочисленні групи слід об'єднувати в одну, додаючи частоти.
- Оскільки можливі помилки першого і другого роду, слід проявляти обережність. Наприклад, можна повторити дослід, збільшити число спостережень, побудувати заздалегідь графік розподілу і тому подібне.

Застосування критерію згоди Пірсона не обмежується випадком нормального розподілу. Наведемо приклади використання критерію Пірсона.

Приклад 1. При рівні значущості 0,05 перевірити гіпотезу про розподіл за законом Пуассона генеральної сукупності, якщо за даними вибірки об'єму $n=60$ отриманий наступний варіаційний ряд:

Варіанти	0	1	2	3	4	5	6	7
Частоти	8	17	16	10	6	2	0	1

Розв'язання: Для розрахунку теоретичних частот використовуємо формулу Пуассона

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

Для оцінки параметра λ використовуємо співвідношення $\lambda = \frac{1}{n} \sum_{k=0}^7 k \cdot n_k = 2$

Обчислимо $P(0)=0,1353$ $P(1)=0,2707$ $P(2)=0,2707$ $P(3)=0,1804$

$P(4)=0,0902$ $P(5)=0,0361$ $P(6)=0,0120$ $P(7)=0,0034$.

Значення χ^2 рівне

$$\chi^2 = \sum_{k=0}^7 \frac{(n_k - n \cdot p_k)^2}{n \cdot p_k} = \frac{(8 - 60 \cdot 0,1353)^2}{60 \cdot 0,1353} + \frac{(17 - 60 \cdot 0,2707)^2}{60 \cdot 0,2707} + \dots = 0,2.$$

Обчислимо число ступенів свободи $k = n - s - 1 = n - 2 = 8 - 2 = 6$. За таблицею критичних точок розподілу хі-квадрат при $n = 6$ і $\alpha = 0,05$ знаходимо $\chi_{кр}^2 = 12,59$. Оскільки спостережене значення менше за критичне, то спостережні значення узгоджуються з розподілом Пуассона і нульова гіпотеза приймається.

Приклад 2. При рівні значущості 0,05 перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні і теоретичні частоти

Емпіричні частоти	6	13	38	74	106	85	30	14
Теоретичні частоти	3	14	42	82	99	76	37	13

Розрахуємо $\chi_{набл}^2 = 7,19$, число ступенів свободи визначимо за співвідношенням $k = s - 3 = 5$ (у нашому випадку $s = 8$). Використовуючи розраховані значення $\chi_{набл}^2$ і k , за таблицею критичних точок розподілу хі-квадрат при рівні значущості $\alpha = 0,05$ знаходимо $\chi_{набл}^2(0,05; 5) = 11,1$. Оскільки $\chi_{набл}^2 < \chi_{кр}^2$, то немає підстав відкидати нульову гіпотезу. Дані спостережень узгоджуються з гіпотезою про нормальний розподіл генеральної сукупності.

Питання для самоперевірки

1. Дати означення статистичної гіпотези.
2. Що таке нульова гіпотеза?
3. Що таке альтернативна гіпотеза?
4. Дати означення простої гіпотези.
5. Дати означення складної гіпотези.
6. Що таке помилка першого роду? Другого роду?
7. Що таке рівень значущості?
8. Дати означення статистичного критерію.
9. Що таке критична область? Область прийняття гіпотези?
10. Які бувають критичні області? Дати їх означення.
11. Перерахувати етапи перевірки статистичних гіпотез.
12. Що таке критерій згоди?

17. ПОНЯТТЯ ПРО КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІЗ

Дві або декілька випадкових величин можуть бути пов'язані або функціональною, або статистичною (стохастичною) залежністю.

У економіці строга функціональна залежність реалізується рідко, оскільки економічні показники схильні до дії випадкових, часто неконтрольованих чинників. Частіше має місце так звана статистична залежність, при якій зміна однієї з величин веде до зміни розподілу іншої. Зокрема, при зміні однієї з величин може змінюватися середнє значення іншої.

Приклад статистичної залежності : урожай зерна Y залежить від кількості внесених добрив X . З однакових по площі ділянок при рівних кількостях внесених добрив знімають різні урожаї. Це пов'язано з впливом випадкових чинників (опаді, температура повітря та ін.). В той же час, середній урожай залежить від кількості добрив, тобто Y пов'язане з X статистичною залежністю.

При вивченні статистичних залежностей розрізняють **кореляцію** і **регресію**. Основним методом дослідження статистичних залежностей виступає **кореляційно - регресійний** аналіз.

Кореляційний аналіз полягає у визначенні **міри зв'язку** між випадковими величинами.

Регресійний аналіз встановлює **форми залежності** між випадковою величиною Y (залежною змінною) і значеннями однієї або декількох змінних величин X (незалежними змінними).

17.1. Поняття про кореляційний аналіз

Одне з найбільш поширених завдань статистичного дослідження полягає у вивченні зв'язку між спостережуваними змінними. Знання взаємозв'язків окремих ознак дає можливість прогнозувати розвиток ситуації при зміні конкретних характеристик об'єкту дослідження. Основний зміст економічної політики, кінець кінцем, може бути зведений до регулювання економічних змінних, здійснюваного на базі виявленої інформації про їх взаємовплив. Тому проблема вивчення взаємозв'язків показників є однією з найважливіших в статистичному аналізі економічних систем.

Кореляція в широкому значенні слова означає зв'язок, співвідношення між об'єктивно існуючими явищами. Якщо випадкові змінні причинно обумовлені, то є кореляція.

Кореляція може бути:

- ✓ додатною або від'ємною;
- ✓ залежно від числа змінних - простою або множинною;
- ✓ залежно від форми зв'язку - лінійною або нелінійною.

Найважливішими завданнями кореляційного аналізу є:

- ✓ вимірювання сили зв'язку двох або більше факторів;

- ✓ відбір факторів, що чинять істотний вплив на результативну ознаку (залежну змінну) на підставі вимірювання тісноти зв'язку між факторами.

У разі лише однієї незалежної змінної X як міра зв'язку між нею і залежною змінною Y служить коефіцієнт кореляції. Він оцінюється за вибіркою об'єму n пов'язаних пар спостережень (x_i, y_i) . У разі декількох змінних необхідно послідовно обчислювати коефіцієнти кореляції по декількох рядах числових даних. Отримані коефіцієнти зводять в таблиці, так звані кореляційні матриці.

Кореляційна матриця є квадратною матрицею, на перетині рядка і стовпця якої знаходиться коефіцієнт кореляції між відповідними змінними.

Якщо в результаті n випробувань система двох випадкових величин (X, Y) набула значень $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, то коефіцієнт кореляції дорівнює

$$\rho_{XY} = \frac{(1/n) \cdot \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \cdot \sigma_y}$$

де $\bar{x}, \bar{y}$ - середні значення, а σ_x, σ_y - середні квадратичні відхилення випадкових величин X, Y відповідно.

Для багатовимірної вибірки (тобто у випадку більше двох факторів) необхідно розрахувати

кореляційну матрицю $\rho(X_i, X_j) = \rho(i, j) = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \dots & \rho_{nn} \end{pmatrix}$, яка є симетричною

відносно головної діагоналі.

Приклад

Є щомісячні дані спостережень за станом погоди і відвідуваністю музею і парку, приведені в таблиці

Число ясних днів (X1)	Кількість відвідувачів музею (X2)	Кількість відвідувачів парку (X3)
8	495	132
14	503	348

20	380	643
25	305	865
20	348	743
15	465	541

Необхідно визначити, чи існує взаємозв'язок між станом погоди і відвідуваністю музеїв і парків. В результаті розрахунку отримаємо кореляційну матрицю

$$\rho(X_i, X_j) = \rho(i, j) = \begin{pmatrix} 1 & -0,921 & 0,975 \\ -0,921 & 1 & -0,919 \\ 0,975 & -0,919 & 1 \end{pmatrix}$$

З кореляційної матриці видно, що кореляція між станом погоди і відвідуваністю музею рівна - 0,921, а між станом погоди і відвідуваністю парку 0,975. Таким чином, виявлена відємна кореляція між відвідуваністю музею і кількістю сонячних днів і практично лінійна додатна кореляція між відвідуваністю парку і станом погоди.

17.2. Поняття про регресійний аналіз

При розгляді взаємозв'язків, як правило, розглядають одну з величин (X) як незалежну (що пояснює), а іншу (Y) як залежну (з'ясовну). При цьому зміна першої з них може служити причиною зміни іншої. Наприклад, зростання доходу веде до збільшення споживання; зростання ціни - до зниження попиту; зниження процентної ставки збільшує інвестиції і так далі. Ця залежність не є однозначною в тому сенсі, що кожному конкретному значенню пояснюючої змінної X може відповідати не одне, а **множина** значень Y . Іншими словами, кожному конкретному значенню незалежної змінної відповідає деякий **імовірнісний розподіл** залежної змінної. Тому аналізують, як пояснююча змінна (чи змінні) впливає (чи впливають) на залежну змінну "в середньому".

Залежність такого типу, що виражається співвідношенням $M(Y|x) = f(x)$ називається **функцією регресії Y на X** . При розгляді залежності двох випадкових величин говорять про **парну регресію**.

Залежність **декількох** змінних, що виражається функцією $M(Y|x_1, x_2, \dots, x_m) = f(x_1, x_2, \dots, x_m)$ називають **множинною регресією**.

Під **регресією** розуміється функціональна залежність між пояснюючими змінними і **умовним математичним сподіванням** (середнім значенням) залежної змінної Y , яка будується з метою передбачення (прогнозування) середнього значення Y при деяких значеннях незалежних змінних.

Встановлення форми залежності і оцінка параметрів функції регресії є завданнями *регресійного аналізу*.

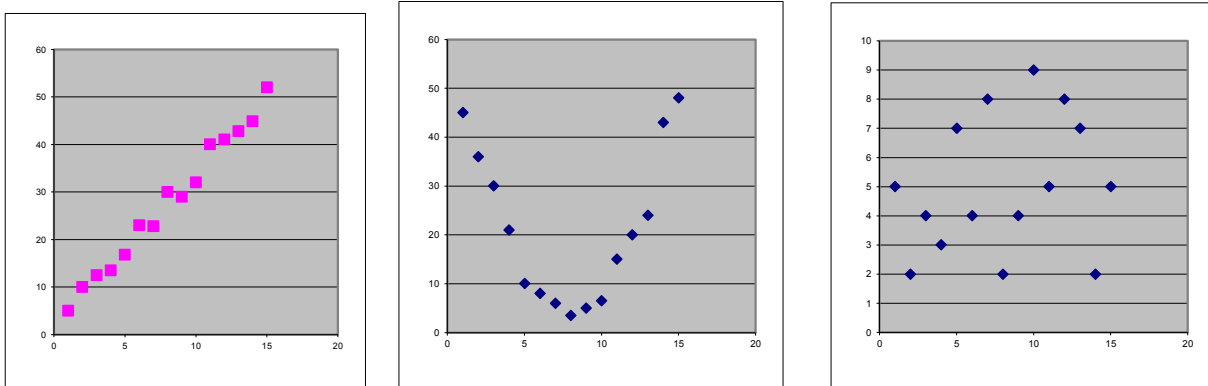
Оскільки реальні значення залежної змінної можуть бути різними при даному X (або $x_1, x_2, \dots, x_m$), залежність має бути доповнена деяким доданком ε , який, по суті, є випадковою величиною. Результируючі співвідношення $Y = f(x) + \varepsilon$, або

$Y = f(x_1, x_2, \dots, x_m) + \varepsilon$ називаються *регресійними рівняннями* (або *моделями*).

Побудова рівняння регресії, що описує емпіричні дані, включає три етапи:

- вибір *формули* рівняння регресії;
- визначення *параметрів* вибраного рівняння;
- аналіз *якості* рівняння і перевірка *адекватності* рівняння емпіричним даним і, при необхідності, *вдосконалення рівняння*.

У випадку парної регресії вибір рівняння зазвичай здійснюється за графічним зображенням реальних статистичних даних - *кореляційному полю*.



Мал. 1 Кореляційні поля. А) - лінійна регресія; Б) - квадратична регресія; В) - відсутність вираженого зв'язку Y і X .

Вибіркові рівняння регресії.

Для визначення значень теоретичних коефіцієнтів, що входять в рівняння регресії, необхідно знати і використовувати *усі* значення змінних генеральної сукупності, що практично неможливо. У зв'язку з цим *за вибіркою обмеженого об'єму* будується так зване *вибіркове (емпіричне) рівняння регресії*. Через *обмеженість вибірки* оцінки коефіцієнтів, що входять у вибіркове рівняння регресії, відрізняються від істинних (теоретичних) значень, що призводить до неспівпадання емпіричною і теоретичною ліній регресії. Різні вибірки з однієї і тієї ж генеральної

сукупності зазвичай призводять до оцінок, що відрізняються одна від одної. Завдання полягає в тому, щоб за конкретною вибіркою (x_i, y_i) , $i = 1, \dots, n$ знайти оцінки невідомих параметрів так, щоб побудована лінія регресії була найкращою серед усіх інших ліній.

17.3. Лінійна регресія

Якщо функція регресії лінійна, то говорять про *лінійну регресію*. Лінійна регресія (лінійне рівняння) є поширеним (і простим) видом залежності між економічними змінними. Для простого випадку *парної лінійної регресії*

$$M(Y|X=x_i) = \beta_0 + \beta_1 x_i \quad \text{або} \quad y_i = M(Y|X=x_i) + \varepsilon_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

де β_0, β_1 - теоретичні параметри регресії; ε_i - випадкове відхилення.

За вибіркою обмеженого об'єму будується *вибіркове* рівняння регресії $\hat{y}_i = b_0 + b_1 x_i$ (1)

де b_0, b_1 - оцінки невідомих параметрів β_0, β_1 , так звані *вибіркові коефіцієнти регресії*, $\hat{y}_i$ - оцінка умовного математичного сподівання $M(Y|X=x_i)$. Для величин y_i справедлива формула

$$y_i = b_0 + b_1 x_i + e_i \quad (2) \text{де } e_i - \text{оцінка теоретичного відхилення } \varepsilon_i.$$

Побудована пряма вибіркової регресії повинна якнайкраще описувати емпіричні дані, тобто коефіцієнти b_0, b_1 мають бути такими, щоб випадкові відхилення e_i були мінімальні. Найбільш поширеним методом знаходження коефіцієнтів рівняння регресії є *метод найменших квадратів (МНК)*.

Якщо за вибіркою (x_i, y_i) , $i = 1, \dots, n$ необхідно визначити оцінки b_0, b_1 вибіркового рівняння регресії (2), то вводиться в розглядання і мінімізується функція

$$Q(b_0, b_1) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2.$$

Необхідною умовою існування мінімуму цієї функції двох змінних є рівність нулю її частинних похідних за невідомими параметрами b_0, b_1 :

$$\frac{\partial Q}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0$$

$$\frac{\partial Q}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i = 0.$$

Звідси
$$n b_0 + b_1 \sum x_i = \sum y_i$$

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i,$$

і виразивши з останніх співвідношень коефіцієнти, отримаємо

$$b_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} \quad b_0 = \bar{y} - b_1 \bar{x}. \quad (3)$$

де введені позначення $\bar{x} = \frac{1}{n} \cdot \sum x_i$, $\bar{y} = \frac{1}{n} \cdot \sum y_i$, $\overline{x^2} = \frac{1}{n} \cdot \sum x_i^2$, $\overline{xy} = \frac{1}{n} \cdot \sum x_i y_i$.

Приклад.

Для аналізу залежності об'єму споживання Y (у.о.) господарства від доходу X (у.о.), що мається, відібрана наступна вибірка об'єму $n=16$

x_i	107	109	110	113	120	122	123	128	136	140	145	150
y_i	102	105	108	110	115	117	119	125	132	130	141	144

Необхідно визначити вид рівняння регресії і за методом найменших квадратів оцінити параметри рівняння регресії, а також спрогнозувати споживання при доході $X=160$.

План розвязку. Будується кореляційне поле. За розташуванням точок на кореляційному полі передбачається, що залежність Y від X - лінійна. За МНК визначаються коефіцієнти $b_0=3,7$; $b_1=0,934$. Таким чином, рівняння парної регресії має вигляд:

$$\hat{Y} = 0,934 X + 3,7$$

Множинна лінійна регресія

На економічний показник найчастіше чинить вплив не один, а декілька факторів. Наприклад, попит на деяке благо визначається не лише ціною цього блага, але і цінами на заміщаючі і доповнюючі блага, доходом споживачів і багатьма іншими факторами. В цьому випадку розглядається множинна регресія $M(Y|x_1, x_2, \dots, x_m) = f(x_1, x_2, \dots, x_m)$.

Теоретичне **лінійне** рівняння регресії має вигляд $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \varepsilon$.

Для оцінки параметрів рівняння множинної лінійної регресії також, як правило, використовується метод найменших квадратів.

17.4. Нелінійна регресія

Багато економічних залежностей не є лінійними. Наприклад, при аналізі еластичності попиту за ціною застосовується так звана логарифмічна модель, при аналізі витрат від обсягу випуску - поліноміальна (кубічна) модель. Часто застосовуються і інші моделі - наприклад, обернена і експоненціальна. Коротко розглянемо деякі з моделей нелінійної регресії.

1. Логарифмічна модель.

Нехай деяка економічна залежність моделюється формулою

$$Y = A X^{\beta}$$

де A, β - параметри моделі. Ця функція може відображати залежність попиту Y на благо від його ціни X (в цьому випадку ($\beta < 0$) або від доходу X ($\beta > 0$ - функція Енгеля). Прологарифмувавши обидві частини останнього співвідношення, отримаємо $\ln Y = \ln A + \beta \ln X$; заміна змінних $Y^* = \ln Y, X^* = \ln X$ дозволяє звести рівняння до лінійного виду

$$Y^* = \beta_0 + \beta X^* + \varepsilon.$$

За МНК можна розрахувати значення параметрів аналогічно випадку лінійної моделі (при цьому замість (x_i, y_i) розглядаються $(\ln x_i, \ln y_i)$).

2. Обернена модель.

Зворотна модель має вигляд

$$Y = \beta_0 + \beta_1 \cdot \frac{1}{X} + \varepsilon.$$

Заміною $X^* = \frac{1}{X}$ ця модель зводиться до лінійної. Обернена модель застосовується, наприклад, для характеристики зв'язку питомих витрат сировини, матеріалів, палива з об'ємом продукції, що випускається.

3. Степенева модель.

Степенева функція вигляду

$$Y = \beta_0 + \beta_1 X + \dots + \beta_m X^m + \varepsilon$$

при $m=3$ (кубічна функція) в мікроекономіці моделює залежність загальних витрат від обсягу випуску; квадратична функція ($m=2$) відображає залежність між обсягом випуску і середніми або

граничними витратами. Модель може бути зведена до *лінійної* моделі множинної регресії за допомогою заміни $X \rightarrow X_1, X^2 \rightarrow X_2, \dots, X^m \rightarrow X_m$. Параметри моделі визначають за допомогою МНК.

4. Показникова модель.

Показникова функція

$$Y = b_0 e^{\beta X} = b_0 m^X$$

може використовуватися при аналізі зміни змінної Y з постійним темпом приросту в часі. Прикладом може служити виробнича функція Кобба - Дугласа з урахуванням науково - технічного прогресу

$$Y = AK^\alpha L^\beta e^{\gamma t},$$

де K - витрати капіталу, L - витрати праці, γ характеризує темпи зростання обсягу виробництва.

Прологарифмувавши, отримуємо співвідношення

$$\ln Y = \ln A + \alpha \ln K + \beta \ln L + \gamma t,$$

яке зводиться до лінійного виду за допомогою замін

$$y \rightarrow \ln Y, k \rightarrow \ln K, l \rightarrow \ln L, a \rightarrow \ln A.$$

На закінчення відмітимо, що побудова і перевірка якості рівняння регресії вимагають застосування методів кореляційного аналізу, що дозволяють проводити відбір істотних для опису регресійної залежності факторів.

Питання для самоперевірки

1. Що таке статистична залежність?
2. Що вивчає кореляційний аналіз?
3. Що вивчає регресійний аналіз?
4. Дати означення кореляції.
5. Перерахувати види кореляції.
6. Що таке функція регресії?
7. Дати означення регресійних рівнянь.
8. Що таке кореляційне поле?
9. Як виглядає рівняння лінійної регресії? Множинної лінійної регресії?
10. Перерахувати основні види нелінійних моделей нелінійної регресії.

ІНДИВІДУАЛЬНІ ЗАВДАННЯ

Завдання 1. Елементи комбінаторики

1. В групі є 25 студентів. Скількома способами можна вибрати серед них старосту та заступника?
2. Шість викладачів мають розподілити між собою 6 розділів книги так, щоб кожен викладач писав один розділ. Скількома способами можна це зробити?
3. На екзамені з вищої математики має бути комісія з 3 викладачів. Скількома способами можна утворити комісію з 12 осіб?
4. У групі налічується 20 студентів. Скількома способами можна обрати: а) студентську раду із трьох студентів; б) голову, заступника та секретаря студентської ради?
5. Скількома способами можна скласти список з 9 учнів?

6. У пасажирському поїзді 14 вагонів. Скількома способами можна розподілити по цих вагонах 14 провідників, якщо за кожним вагоном закріплювати одного провідника?

7. З цифр 0, 1, 2, 3 складені різні чотирицифрові числа так, що в кожному немає однакових цифр. Скільки вийшло чисел?

8. З цифр 1, 2, 3, 4, 5, не повторюючи їх, склали всі можливі п'ятицифрові числа. З'ясуйте, скільки серед цих чисел таких, які: а) починаються цифрою 5; б) не починаються цифрою 5; в) починаються з 54; г) не починаються з 543.

9. Скільки різних акордів можна зіграти на 10 вибраних клавішах рояля, якщо в кожному акорді може бути від трьох до десяти звуків.

10. У підрозділі 80 солдат і 6 офіцерів. Скількома способами можна призначити патруль, у складі якого два офіцери і три солдати?

11. Необхідно виділити для охорони підприємства 6 солдат, 3 сержантів і 2 офіцерів. Скількома способами це можна зробити, якщо в роті є 30 солдат, 6 сержантів і 4 офіцери?

12. Для карнавального походу готують різні прапори. Скількома способами з п'яти сувоїх стрічок різного кольору можна виготовити різні прапори, які мають по три горизонтальні смуги, якщо в одному прапорі жоден з кольорів не повинен повторюватись?

13. Номер автомобільного причепа містить дві букви і чотири цифри. Скільки різних номерів можна скласти з 30 букв і 10 цифр.

14. Із Києва до Чернігова можна добратися чотирма способами — поїздом, автобусом, літаком та теплоходом, а з Чернігова до Новгород–Сіверська двома — теплоходом і автобусом. Скількома способами можна здійснити подорож від Києва до Новгород–Сіверська із зупинкою у Чернігові?

15. У першості України з футболу у вищій лізі приймають участь 16 команд. Скількома способами теоретично між ними могли б бути розподілені золоті, срібні та бронзові медалі?

16. Необхідно вибрати делегацію в складі чотирьох учнів одного класу і одного вчителя. Скількома способами це можна здійснити, коли в школі працює 20 вчителів, а у вибраному класі навчаються 36 учнів?

16. Автомобільний номер складається з п'яти цифр і двох літер, які стоять після цифр. Скільки різних номерів можна скласти, використовуючи 30 літер і 10 цифр?

17. У профспілковий комітет обрано 7 працівників. З них необхідно обрати голову, його заступника і секретаря. Скількома способами можна це зробити?

18. Є шість видів конвертів і 5 видів поштових марок. Скількома способами для відправки листа можна вибрати конверт і марку?

19. В аудиторії 8 світильників. Скільки існує способів освітлення цієї аудиторії?

20. Три хлопчики зірвали 40 яблук. Скількома способами вони можуть розділити ці яблука?

21. Скільки різних чотирицифрових чисел, які діляться на 4 можна записати за допомогою цифр 3, 4, 5, 6, 7, якщо цифри в записі числа можуть повторюватись?

22. У поштовому кіоску продають 12 видів листівок. Скількома способами можна закупити 7 листівок? Скількома способами можна вибрати 8 різних листівок?

23. Скільки пар різних підмножин множини $M=\{1; 2; 3; 4; \dots; 10\}$ можна утворити, якщо в одну підмножину можна включати два, три або чотири різних елементи множини?

24. Скільки існує чотирицифрових чисел, в записах яких усі цифри різні?

25. Скільки існує чотирицифрових чисел, в записах яких цифри можуть повторюватись?

Завдання 2. Класичне означення ймовірності. Геометрична ймовірність

1. Знайти класичну ймовірність того, що випадковим чином вибране двозначне число ділиться: а) на 2 або на 3; б) на 2 і на 3.

2. В податковій адміністрації зареєстровано 6 приватних і 4 державних підприємства. Знайти ймовірність того, що серед навмання вибраних трьох підприємств приватними будуть: а) три; б) два; в) не більше одного.

3. Є п'ять карточок на кожній з яких написано по одній із букв «О», «Р», «С», «Т», «П». Знайти ймовірність того, що випадковим чином складені карточки утворять слово «СПОРТ».

4. У бібліотеці є 9 книг, з них 5 з економіки і 4 з математики. Знайти ймовірність того, що серед навмання вибраних 6 книг є 3 з економіки.

5. Серед 10 працівників є 4 з вищою освітою. Визначити ймовірність того, що із навмання вибраних 6 працівників буде два з вищою освітою.

6. З карток, які складають слово «ціноутворення», навмання вибирають одну. Визначити ймовірність того, що на ній буде написана буква: а) «н»; б) «у».

7. В ящику 40 куль: 28 червоних і 12 синіх. З ящика навмання виймають дві кулі. Визначити ймовірність того, що ці кулі: а) червоні; б) сині.

8. Із 40 доброякісних і 4 бракованих деталей для контролю взято навмання вісім. Визначити ймовірність того, що серед них 5 деталей буде доброякісними.

9. На книжковій полиці в довільному порядку розміщені книги з п'ятитомного зібрання творів. Яка ймовірність того, що всі томи розміщені в порядку зростання номерів?

10. На книжковій полиці в довільному порядку розміщені книги з десяти томного зібрання творів. Яка ймовірність того, що третій, четвертий, п'ятий томи після переставляння стоятимуть поруч в порядку зростання номерів?

11. На книжковій полиці в довільному порядку розміщені книги з восьмитомного зібрання творів. Яка ймовірність того, що п'ятий, шостий, сьомий томи після переставляння стоятимуть поруч в порядку зростання номерів?

12. Серед 4000 перших простих чисел натурального ряду є 551 просте число. Знайдіть частоту появи простого числа.

13. Проведено 100 пострілів у мішень. Виявлено, що відносна частота влучень дорівнює 0,85. Скільки пострілів були влучними?

14. Під час перевірки готової продукції було виявлено 5 бракованих одиниць товару з 200 перевірених. Знайти відносну частоту бракованих одиниць товару.

15. На проміжку $[0; 5]$ випадково вибирають два цілих числа. Яка ймовірність того, що: а) сума чисел менша за 4; б) добуток вибраних чисел більший за 5; в) модуль різниці вибраних чисел менший за 2, а їх добуток більший за 3.

16. На відрізку $[0, 10]$ навмання вибирають два цілих числа. Знайти ймовірність того, що сума цих чисел не перевищує 1.
17. На кожній з 9 карток написано по літері: «м», «с», «к», «н», «е», «о», «і», «т», «о». Знайти ймовірність того, що, навмання викладаючи ці картки одна біля одної, ви отримаєте слово «економіст».
18. П'ять книжок, серед яких 2 підручники з математики, довільно розміщують на полиці. Яка ймовірність того, що ці 2 підручники стоятимуть поряд?
19. В урні знаходиться 8 карток з цифрами 1, 2, 3, 4, 5, 6, 7, 8. Із урни навмання без повернення виймають по одній усі картки і цифри, що з'являються, записують зліва направо. Яка ймовірність того, що: а) записане число закінчується рівно двома непарними цифрами; б) усі парні цифри стоятимуть поруч?
20. Навмання вибирається п'ятицифрове число. Знайти ймовірність того, що це число: а) кратне 5; б) складається з непарних цифр; в) однаково читається як зліва направо, так і справа наліво (як, наприклад, 23132).
21. На шести однакових картках написані цифри від 1 до 6. Навмання вибрано чотири картки і розкладено їх в порядку появи в ряд зліва направо. Знайти ймовірність того, що з'явиться число, яке: а) не містить цифр 1 і 5; б) кратне 3.
22. Маємо 5 відрізків, довжини яких — 1 см, 3 см, 5 см, 7 см, 9 см. Визначити ймовірність того, що з трьох відрізків, навмання взятих з даних п'яти, можна скласти трикутник.
23. 7. На кожній із 5 карточок написана одна з таких літер «А», «І», «З», «К», «Л». Знайти ймовірність утворення слова «ЗАЛІК».
24. Знайти ймовірність того, що навмання складене із цифр 1, 2, 3, 4 без повторення чотирицифрове число починається цифрами: а) 2; б) 34.
25. Обчислити ймовірність того, що для навмання заданого трицифрового числа із цифри 1, 2, 3, 4 без повторення, сума усіх його цифр: а) рівна 8; б) менша 8.

Завдання 3. Теорема додавання і множення ймовірностей.

1. Три студенти складають іспит. Ймовірність того, що перший студент складе іспит, дорівнює 0,95, другий-0,9, третій-0,85. Визначити ймовірність того, що: а) два студенти складуть іспит; б) всі три студенти складуть іспит.
2. В класі 35 учнів: 20 дівчаток та 15 хлопчиків. Вирішено за допомогою жеребу розподілити серед учнів чотири квитки в театр. Яка ймовірність, що квитки отримають: (а) чотири дівчинки; (б) два хлопчики та дві дівчинки?
3. У вазі 5 троянд рожевого, 7 червоного та 3 білого кольору. Навмання дістають 2 троянди. Яка ймовірність того, що вони будуть: а) одного кольору; б) різних кольорів?
4. Із 10 лотерейних білетів книжкової лотереї, що перебувають у продажу, 2 виграшні. Визначити ймовірність того, що серед куплених 5 білетів: а) один виграшний; б) хоча б один виграшний.
5. Перший стрілець влучає у ціль з ймовірністю 0,8, другий — з ймовірністю 0,9, а третій — з ймовірністю 0,85. Яка ймовірність того, що хоча б один стрілець влучить у ціль?

6. Робітник обслуговує одночасно 3 верстати. Ймовірність порушення роботи протягом години для першого дорівнює 0,1, для другого — 0,2, для третього — 0,2. Яка ймовірність того, що: а) усі три верстати працюватимуть протягом години; б) хоча б один із них вийде з ладу?
7. Двоє стрільців виконали по одному пострілу у мішень. Ймовірність влучення для першого стрільця становить 0,7, а для другого — 0,8. Яка ймовірність того, що не відбудеться жодного влучення?
8. Знайти ймовірність того, що навмання вибране двозначне число ділиться на 5 або на 10.
9. Робітник обслуговує одночасно 3 верстати. Ймовірність порушення роботи протягом години для першого дорівнює 0,1, для другого — 0,2, для третього — 0,2. Яка ймовірність того, що: а) усі три верстати працюватимуть протягом години; б) хоча б один із них вийде з ладу?
10. Для виконання завдання замовник звернувся до двох виконавців. Ймовірність того, що перший виконавець надасть згоду виконати замовлення дорівнює 0,8, а другий — 0,9. Знайти ймовірність того, що замовлення буде прийнято принаймні одним з цих виконавців.
11. Три стрільці незалежно один від одного стріляють по мішені. Ймовірність влучення в мішень для першого дорівнює 0,7, для другого — 0,8, для третього — 0,9. Яка ймовірність, що: а) хоча б один з них влучить у мішень; б) тільки двоє влучать у мішень; в) жоден не влучить у мішень?
12. До іспиту з теорії ймовірності студентові запропоновано 100 задач. При підготовці до іспиту студент розв'язав 90 задач. Яка ймовірність того, що в білеті, який містить 3 задачі, всі будуть з тих, які студентом не розв'язані?
13. Мішень складається з трьох областей, які попарно не перетинаються. Ймовірність влучення стрільцем у першу область мішені дорівнює 0,45, у другу область — 0,35, а у третю — 0,15. Яка ймовірність того, що при одному пострілі стрілець влучить: а) у першу або в другу область; б) не влучить у мішень?
14. Робітник обслуговує три верстати. Ймовірність того, що протягом години він буде біля першого верстата, дорівнює 0,7, біля другого - 0,5, третього - 0,75. Визначити ймовірність того, що протягом години уваги робітника зажадають: а) всі верстати; б) який-небудь один верстат; в) хоча б один верстат.
15. Ймовірність того, що студент відповість на перше із трьох питань екзаменаційного квитка, дорівнює 0,95, на друге - 0,9 і на третє - 0,85. Визначити ймовірність того, що студент здасть іспит, якщо для цього йому необхідно відповісти: а) на всі питання; б) хоча б на два питання.
16. Для повідомлення про пожежу в приміщенні встановлено три сигналізатори, які працюють незалежно. Ймовірність того, що при пожежі спрацює перший сигналізатор дорівнює 0,9, для другого та третього сигналізаторів ці ймовірності відповідно дорівнюють 0,75 та 0,8. Яка ймовірність того, що при пожежі спрацює: а) тільки перший сигналізатор; б) принаймні два сигналізатори; в) тільки один сигналізатор?
17. Студент знає 28 з 35 питань програми. Білет містить чотири питання програми. Яка ймовірність того, що студент: а) не відповість на жодне питання навмання взятого білета; б) відповість тільки на два питання з білета; в) відповість хоча б на одне питання з білета?

18. В першому ящику 16 білих і 4 червоних кулі, в другому – 8 білих і 12 червоних. З кожного ящика навмання беруть по одній кулі. Яка ймовірність того, що серед вибраних буде: а) тільки одна червона куля; б) хоча б одна червона куля?

19. Партія деталей містить 6 деталей вищого, 7 деталей першого та 4 деталі другого сорту. З партії навмання взято 4 деталі. Яка ймовірність того, що: а) серед них буде три першого і одна другого сорту; б) серед них хоча б одна буде другого сорту; в) серед них жодної першого сорту?

20. Студент вивчив 8 питань із 12. Знайти ймовірність того, що він дасть відповідь хоч би на два питання із трьох навмання вибраних.

21. Ймовірності появи двох несумісних, незалежних подій дорівнюють 0,6 і 0,3 відповідно. Яка ймовірність того, що в результаті випробування відбудеться хоча б одна з цих подій?

22. Три елементи електричного ланцюга виходять зі строю незалежно один від одного з ймовірностями 0,1, 0,05 та 0,08 відповідно. Знайти ймовірність того, що в ланцюгу настане розрив, якщо елементи з'єднані паралельно?

23. По мішені зроблено залп із двох гармат. Ймовірність вразити ціль першою гарматою дорівнює 0,8, другою – 0,7. Для знищення цілі необхідне одне попадання. Яка ймовірність того, що ціль знищена?

24. В урні 5 синіх і 10 червоних кульок. Навмання витягують по черзі три кульки. Знайти ймовірність того, що перша куля буде синя, а дві наступні – червоні.

25. Для зруйнування крижаного затору на річці досить одного влучення авіаційної бомби. Знайти ймовірність зруйнування затору при трьох бомбометаннях, якщо ймовірність влучення бомб при цьому відповідно рівні 0,4; 0,5; 0,6.

Завдання 4. Формула повної ймовірності. Формула Байєса.

1. На трьох дочок у сім'ї (старшу, середню та молодшу); покладений обов'язок мити посуд. Старша дочка виконує 40% усієї роботи, інші по 30%. Ймовірність розбити при митті посуд для старшої дочки становить 0,02, для середньої та молодшої відповідно 0,03 та 0,04. Невідомо, хто напередодні мив посуд, але одна тарілка виявилася розбитою. Яка ймовірність того, що посуд мила старша дочка?

2. В урні було 4 чорних і 3 білих кульки. З неї випадковим чином вийняли 3 кульки. Яка ймовірність того, що вийняті після цього 2 кульки виявляться чорними?

3. Клієнт має можливість покласти гроші на довготривалий рахунок в один з трьох банків. Ризик (тобто можливість банкрутства банку протягом зазначеного терміну) при цьому становить 2% для першого банку, 3% та 4% відповідно для другого та третього. Яка ймовірність не вигідного вкладання коштів в банк, вибраний навмання?

4. В магазин надійшли телевізори, виготовлені на трьох заводах. Продукція першого заводу містить 18% телевізорів із скритим дефектом, другого – 12%, третього – 5%. Яка ймовірність придбати справний телевізор, якщо в магазин поступило 30% телевізорів з першого заводу, 25% – з другого і 45% – з третього?

5. Для участі в студентських відбіркових спортивних змаганнях виділено з першої групи курсу 4, з другої – 6, з третьої – 5 студентів. Ймовірності того, що студент першої, другої, третьої групи попаде в збірну інституту, відповідно рівні 0,9; 0,7; 0,8. Навмання вибраний студент в результаті змагань попав до збірної. До якої із груп найімовірніше він належав?

6. Чотири верстати виготовляють однакові деталі. За зміну перший верстат виготовив 20 деталей, другий – 30, третій – 40, а четвертий – 10 деталей. Ймовірність появи браку на першому верстаті рівна 0,1; на другому – 0,15; на третьому – 0,2; на четвертому – 0,05. Яка ймовірність того, що навмання взята деталь виготовлена на третьому верстаті, якщо вона виявилася якісною?

7. З 20 стрільців 8 влучають у мішень з імовірністю 0,8; 7 – з імовірністю 0,6 і 5 – з імовірністю 0,5. Навмання вибраний стрілець, зробивши 1 постріл в мішень, не влучив. Знайти ймовірність того, що він належав до першої групи.

8. Для контролю з 3 партій деталей навмання взято 1 деталь. Яка ймовірність того, що деталь бракована, якщо в першій партії 10% бракованих деталей, в другій – 20% бракованих деталей, а в третій партії всі деталі стандартні?

9. Відомо, що 5% всіх чоловіків і 0,25% всіх жінок дальтоніки. Навмання вибрана особа – дальтонік. Яка ймовірність того, що це жінка? (Вважати, що чоловіків і жінок однакова кількість).

10. В електронно-обчислювальній машині 5% з усіх неполадок припадає на арифметичний пристрій, 10% на оперативну пам'ять, 15% на пристрої зовнішньої пам'яті, решта на периферійні пристрої. Ймовірність ліквідації неполадки у цих пристроях з допомогою служби ТО відповідно дорівнює 0,1, 0,2, 0,4 і 0,8. Неполадка, що виникла, була ліквідована службою ТО. Яка ймовірність того, що вона стосувалася оперативної пам'яті?

11. У коробці 10 чорних, 5 білих та 15 синіх кульок. Навмання витягують і відставляють в сторону 1 кульку. Потім витягують ще 1 кульку. Яка ймовірність того, що друга кулька буде синього кольору?

12. Є 2 партії деталей з 150 і 100 деталей. В кожній партії по 5 деталей бракованих. Одну деталь, навмання вибрану з першої партії, переклали в другу партію. Потім з другої партії вибрали навмання 1 деталь. Яка ймовірність того, що остання деталь не бракована?

13. В обчислювальній лабораторії є три партії комп'ютерів індивідуального користування, виготовлених на різних заводах у кількостях 20, 30 і 50 штук відповідно. Ймовірності того, що комп'ютери, поставлені різними заводами, пропрацюють без ремонту заданий термін, відповідно дорівнюють для цих партій 0,7, 0,8, 0,9. Яка ймовірність того, що вибраний навмання комп'ютер із ста даних пропрацює без ремонту заданий час?

14. Є три партії деталей по 30 деталей в кожній. Число стандартних деталей в кожній партії дорівнює відповідно 20, 21, 18. З навмання вибраної партії вийняли деталь, яка виявилась стандартною. Знайти ймовірність того, що деталь була взята з другої партії.

15. В урні міститься 10 чорних і 5 білих кульок. Навмання витягнули 3 кульки і чорні повернули назад, а білі відклали в сторону. Знайти ймовірність того, що витягнуті після цього 2 кульки будуть білі.

16. В урні 3 білих і 7 чорних кульок. Можна або додати в урну 2 білих кульки або навмання витягнути з неї 2 кульки. В якому випадку ймовірність витягнути 1 чорну кульку буде більшою?

17. У першій урні 5 білих і 10 чорних кульок, а в другій – 7 білих і 8 чорних. В одну з урн додають 2 білих кульки. Після чого з навмання вибраної урни витягують 2 кульки. Яка ймовірність того, що ці кульки білі?

18. В урні 3 червоних і 12 білих кульок. З неї витягують навмання 2, а потім ще 2 кульки. Яка ймовірність того, що 2 останні кульки будуть білого кольору?

19. В ящику знаходяться однакові деталі двох заводів у співвідношенні 1:3. Серед деталей першого заводу брак становить 3%, серед деталей другого – 2%. Навмання взята з ящика деталь виявилася бракованою. Яка ймовірність того, що вона виготовлена першим заводом?

20. В кожній з 8 урн міститься по 6 білих і 4 чорних кулі, а в кожній з наступних 5 урн міститься по 3 білих і 7 чорних куль. З навімання взятої урни навімання вийнята куля. Яка ймовірність того, що вона біла?

21. Три заводи поставляють на конвеєр однотипні деталі. Перший завод поставляє 20%, другий – 30%, третій – 50% деталей. Перший завод випускає 3% браку, другий – 2%, а третій – 1%. При перевірці навімання вибрана деталь виявилася бракованою. Яка ймовірність того, що вона виготовлена першим заводом?

22. Є 3 однакових урни. В першій 6 білих і 8 чорних, в другій 4 білих і 6 чорних, а в третій 2 білих і 4 чорних кульки. Знайти ймовірність того, що кулька, витягнута з навімання вибраної урни, виявиться білою.

23. Є три партії деталей по 30 деталей в кожній. Число стандартних деталей в кожній партії дорівнює відповідно 20, 21, 18. З навімання вибраної партії вийняли деталь, яка виявилась стандартною. Знайти ймовірність того, що деталь була взята з другої партії.

24. При масовому виробництві деякого виробу ймовірність того, що він виявиться стандартним, рівна 0,97. Для контролю проводиться деяка спрощена перевірка стандартності виробу, яка дає позитивний результат в 99% для стандартного виробу і в 2% – для нестандартного. Яка ймовірність стандартності виробу, який пройшов спрощену перевірку?

25. В групі з 25 чоловік 5 студентів підготувались до екзамену відмінно, 10 – добре, 7 – посередньо і 3 – погано. На екзамен виноситься 25 питань. Відмінно підготовлений студент може відповісти на всі 25 питань, добре підготовлений – на 20, посередньо підготовлений – на 13, а погано підготовлений – на 6. Викликаний навімання студент відповів на всі 3 поставлені питання. Знайти ймовірність того, що це був посередньо підготовлений студент.

Завдання 5. Повторні незалежні випробування

1. Ймовірність того, що кожний із 100 верстатів працює в даний момент, дорівнює 0,9. Знайти ймовірність того, що в даний момент працює від 50 до 70 верстатів.
2. Серед виробів даного підприємства 20% виробів вищого сорту. Знайти ймовірність того, що із 5 придбаних виробів 3 вищого сорту.
3. Ймовірність влучення в ціль при одному пострілі дорівнює 0,8. Знайти ймовірність 100 влучень при 130 пострілах.
4. Завод відправив на базу 5000 доброякісних виробів. Ймовірність пошкодження кожного виробу в дорозі дорівнює 0,0002. Знайти ймовірність того, що серед 5000 виробів в дорозі буде пошкоджено більше 3 виробів.
5. Ймовірність хоча б одного влучення в ціль при двох пострілах дорівнює 0,99. Знайти ймовірність трьох влучень в ціль при чотирьох пострілах.
6. В квартирі чотири електролампочки. Для кожної лампочки ймовірність того, що вона вийде з ладу на протязі року, дорівнює $\frac{5}{6}$. Яка ймовірність того, що на протязі року доведеться замінити не менше половини лампочок?
7. Ймовірність пошкодження при транспортуванні з заводу до пункту збуту для одного пристрою, незалежно від інших, становить 0,002. Яка ймовірність того, що при транспортуванні 1000 одиниць продукції буде пошкоджено не більше двох одиниць?
8. Гральний кубик підкидають тричі. Знайти ймовірність того, що рівно двічі з'явиться число очок кратне трьом.
9. Ймовірність отримати в лотереї білет без виграшу дорівнює 0,1. Яка ймовірність того, що серед 400 навмання куплених білетів не менше 30 і не більше 40 без виграшу?
10. Ймовірність того, що телевізор вимагатиме ремонту впродовж гарантійного строку дорівнює 0,2. Знайти ймовірність того, що на протязі гарантійного строку з трьох телевізорів не більше, ніж одному буде потрібен гарантійний ремонт.
11. Гральний кубик підкинули 500 разів. Яка ймовірність того, що число очок, кратне трьом, випаде не менше 160 і не більше 200 разів?
12. Відділ технічного контролю перевіряє деталі на стандартність. Ймовірність того, що деталь нестандартна, дорівнює 0,1. Знайти ймовірність того, що з 5 перевірених деталей принаймні 4 деталі будуть стандартними.
13. Ймовірність влучення в ціль при одному пострілі становить 0,8. Яка ймовірність того, що при 100 незалежних пострілах буде зафіксовано не більше 75 влучень?
14. Ймовірність народження хлопчика дорівнює 0,515. У деякій сім'ї шестеро дітей. Знайти ймовірність того, що серед них не більше двох дівчаток.
15. Радіоапаратура складається із 1000 мікроелементів. Ймовірність відмови кожного елемента протягом доби дорівнює 0,001 і не залежить від стану інших елементів. Знайти ймовірність відмови не менше двох елементів за добу.
16. Для забезпечення нормальної роботи банку потрібно, щоб справними були не менше ніж 80% із наявних 50 комп'ютерів. Яка ймовірність нормальної роботи банку, якщо ймовірність вийти з ладу для кожного з комп'ютерів дорівнює 0,1?

17. Яка ймовірність того, що серед 200 чоловік буде не менше чотирьох ліворуких, якщо вони в середньому складають 1% від загальної кількості?
18. Яка ймовірність того, що в стовпчику з 100 навімання відібраних монет, розташованих гербом угору, буде від 45 до 60?
19. Було посаджено 400 дерев. Знайти ймовірність того, що число дерев, які прийнялись, менше 100, якщо ймовірність того, що окреме дерево прийметься, дорівнює 0,8.
20. Підприємство має 5 постачальників, ймовірність виконання договору для кожного з яких дорівнює 0,7. Знайти ймовірність того, що менше 40% постачальників виконають договір.
21. Два рівносильних суперники грають в шахи. Що ймовірніше: виграти не менше 4 партій із 7 чи не менше 5 партій із 9? Нічий до уваги не беруться.
22. Гральний кубик кидають 80 разів. Знайти ймовірність того, що 26 разів з'явиться одиниця.
23. Гральний кубик кидають 125 разів. Знайти ймовірність того, що рівно 32 рази з'явиться шістка.
24. Керівництво застави зібрало дані, які вказують, що 80% машин, які прибувають на прикордонну заставу, – це легкові автомобілі. Якщо до в'їзду прибуло 9 машин, то яка ймовірність того, що від 4 до 7 з них будуть легкові?
25. Ймовірність виготовлення робітником деталі відмінної якості становить 0,8. Яка ймовірність того, що серед 6 виготовлених робітником деталей хоча б дві будуть відмінної якості.

Завдання 6. Випадкові величини. Числові характеристики випадкових величин

1. Монету кинуто 4 рази. Знайти закон розподілу та функцію розподілу випадкової величини X – числа випадань герба і обчислити її основні числові характеристики: математичне сподівання $M(X)$, дисперсію $D(X)$, середнє квадратичне відхилення $\sigma(X)$.
2. В ящику є 6 деталей, 4 з яких стандартні. Навмання відібрано 3 деталі. Скласти закон розподілу дискретної випадкової величини X – числа стандартних деталей серед відібраних. Побудувати функцію розподілу та обчислити середнє квадратичне відхилення даної випадкової величини.
3. Ймовірність виконання договору для кожного з чотирьох заводів дорівнює 0,6. Знайти закон розподілу та функцію розподілу випадкової величини X – числа заводів, які виконають договір, і обчислити її математичне сподівання $M(X)$ та середнє квадратичне відхилення $\sigma(X)$.
4. В ящику знаходиться 3 білих та 5 чорних куль. Навмання виймають три кулі. Знайти закон розподілу випадкової величини X – числа появ білої кулі та обчислити її математичне сподівання $M(X)$.
5. Проведено 3 постріли по мішені. Ймовірність влучення в мішень при першому пострілі дорівнює 0,1, при другому – 0,2, при третьому – 0,3. Знайти закон розподілу та функцію розподілу випадкової величини X – числа влучень в мішень і обчислити її математичне сподівання $M(X)$.
6. З колоди карт (36 шт.) витягають навмання 4 карти. Знайти закон розподілу випадкової величини X – кількості фігурних карт серед витягнутих. Побудувати закон розподілу випадкової

величини X та обчислити її середнє квадратичне відхилення $\sigma(X)$ та математичне сподівання $M(X)$.

7. Гральний кубик кидають двічі. Знайти закон розподілу та функцію розподілу випадкової величини X – числа випадань 5 очок і обчислити її середнє квадратичне відхилення $\sigma(X)$.

8. Стрілець, маючи 4 патрони, стріляє до першого влучення в ціль. Ймовірність влучення в ціль при кожному пострілі дорівнює 0,7. Знайти закон розподілу та функцію розподілу випадкової величини X – числа використаних патронів і обчислити її математичне сподівання $M(X)$.

9. В урні міститься 4 зелених і 3 чорних кулі. З урни витягають по одній кулі до появи чорної. Знайти закон розподілу випадкової величини X – кількості витягнутих куль і обчислити її математичне сподівання $M(X)$.

10. В партії із 7 деталей є 5 стандартних. Навмання відібрано 4 деталі. Знайти закон розподілу та функцію розподілу випадкової величини X – числа стандартних деталей серед відібраних і обчислити її математичне сподівання $M(X)$.

11. Абонент при наборі номера телефону забув останню цифру і набирає її навмання. Знайти закон розподілу випадкової величини X – кількості наборів номера телефону до попадання на потрібного абонента. Обчислити математичне сподівання $M(X)$.

12. В ящику знаходиться 3 білих та 7 чорних куль. Навмання виймають чотири кулі. Знайти закон розподілу та функцію розподілу випадкової величини X – числа появ білої кулі та обчислити її математичне сподівання $M(X)$.

13. Гральний кубик кидають тричі. Знайти закон розподілу випадкової величини X – числа випадань 6 очок і обчислити її математичне сподівання $M(X)$.

14. Два гральних кубики одночасно кидають три рази. Знайти закон розподілу випадкової величини X – числа випадань парного числа очок на двох гральних кубиках. Обчислити математичне сподівання $M(X)$ та середнє квадратичне відхилення $\sigma(X)$ даної випадкової величини.

15. В партії із 8 деталей є 5 стандартних. Навмання відібрано 4 деталі. Знайти закон розподілу та функцію розподілу випадкової величини X – числа стандартних деталей серед відібраних і обчислити її середнє квадратичне відхилення $\sigma(X)$.

16. В ящику знаходиться 4 білих та 8 чорних куль. Навмання виймають три кулі. Знайти закон розподілу та функцію розподілу випадкової величини X – числа появ чорної кулі та обчислити її дисперсію $D(X)$.

17. Технологічний процес є таким, що брак складає 10% усіх виробів. Навмання взято 3 вироби. Знайти закон розподілу та функцію розподілу випадкової величини X – числа бракованих виробів серед вибраних і обчислити її математичне сподівання $M(X)$ та середнє квадратичне відхилення $\sigma(X)$.

18. В партії із 10 деталей є 7 стандартних. Навмання відібрано 3 деталі. Знайти закон розподілу випадкової величини X – числа стандартних деталей серед відібраних і обчислити її середнє квадратичне відхилення $\sigma(X)$ та ймовірність події $X > 1$.

19. Монету кидають 5 разів. Знайти закон розподілу випадкової величини X – числа випадань герба і обчислити її основні числові характеристики: математичне сподівання $M(X)$, дисперсію $D(X)$, середнє квадратичне відхилення $\sigma(X)$.

20. Технічний пристрій складається з трьох елементів, що працюють незалежно. Ймовірність відмови першого елемента дорівнює 0,2, другого – 0,1, третього – 0,3. Знайти закон розподілу випадкової величини X – числа елементів, що відмовили, і обчислити її середнє квадратичне відхилення $\sigma(X)$ та ймовірність події $X < 3$.

21. В партії з 15 деталей є 10 стандартних. Навмання відібрано 3 деталі. Знайти закон розподілу випадкової величини X – числа нестандартних деталей серед відібраних і обчислити її математичне сподівання $M(X)$ та ймовірність події $X < 2$.

22. В урні знаходиться 4 білих та 5 зелених куль. Навмання виймають три кулі. Знайти закон розподілу і функцію розподілу випадкової величини X – числа появ зеленої кулі та обчислити її середнє квадратичне відхилення $\sigma(X)$.

23. П'ять приладів перевіряють на надійність. Кожний наступний прилад підлягає перевірці лише тоді, коли попередній перевірений прилад виявиться ненадійним. Ймовірність того, що прилад витримає перевірку на надійність, однакова для всіх приладів і дорівнює 0,8. Знайти закон розподілу випадкової величини X – числа перевірених приладів і обчислити її математичне сподівання $M(X)$, середнє квадратичне відхилення $\sigma(X)$ та ймовірність події $X \geq 3$.

24. В партії з 12 деталей є 9 стандартних. Навмання відібрано 4 деталі. Знайти закон розподілу випадкової величини X – числа нестандартних деталей серед відібраних і обчислити її математичне сподівання $M(X)$, дисперсію $D(X)$ та ймовірність події $X \geq 2$.

25. Стрілець, маючи 3 патрони, стріляє до першого влучення в ціль. Ймовірність влучення в ціль при кожному пострілі дорівнює 0,6. Знайти закон розподілу та функцію розподілу випадкової величини X – числа використаних патронів і обчислити її математичне сподівання $M(X)$ та середнє квадратичне відхилення $\sigma(X)$.

Завдання 7. Типові закони розподілу випадкових величин

Знайти ймовірність потрапляння випадкової величини X в інтервал (α, β) , якщо вона розподілена: а) рівномірно на відрізку $[a, b]$; б) за нормальним законом і має математичне сподівання a і середнє квадратичне відхилення α ; в) за показниковим законом і має математичне сподівання b .

Варіант	a	b	α	β
1	1	5	2	4
2	3	6	4	6
3	10	13	11	12
4	5	8	6	7
5	3	9	4	8

6	4	10	6	9
7	2	7	3	5
8	7	10	8	10
9	1	6	2	4
10	5	9	6	8
11	10	15	11	14
12	7	13	8	11
13	3	9	4	7
14	6	11	8	10
15	1	9	3	7
16	2	8	4	7
17	5	10	6	9
18	3	8	5	7
19	5	9	4	8
20	6	12	9	11
21	3	10	5	9
22	1	7	3	6
23	7	15	9	14
24	6	16	8	12
25	1	7	3	6

Завдання 8.

Якість продукції контролюється за наявністю в ній дефектів двох видів X та Y. Ці дефекти є випадковими величинами, що мають закон розподілу, вказаний в таблиці. Потрібно знайти: а) закони розподілу компонент X та Y; б) умовний розподіл Y за умови, що X приймає своє найменше значення; в) коваріацію та коефіцієнт кореляції дефектів та з'ясувати залежні вони чи ні.

Варіант 1

X \ Y	-1	0	1	2
1	0,1	0,1	0	0
2	0,1	0,2	0,2	0
3	0	0,1	0,1	0,1

Варіант 2

X \ Y	1	2	3	4
-1	0,1	0,1	0,1	0
0	0	0,2	0,2	0
1	0	0	0,1	0,2

Варіант 3

X \ y	1	3	5	7
-3	0,1	0,1	0,1	0,1
-1	0,2	0,1	0,1	0
1	0	0	0,2	0

Варіант 10

X \ y	1	3	4	5
1	0,1	0,1	0,1	0,05
2	0,1	0,1	0,2	0,05
3	0	0,1	0	0,1

Варіант 4

X \ y	2	5	8	11
-1	0,1	0	0	0
0	0,1	0,1	0,2	0,2
1	0	0,2	0	0,1

Варіант 11

X \ y	2	4	6	7
-3	0,1	0,1	0	0
-2	0,05	0,1	0,2	0
-1	0,05	0,1	0,2	0,1

Варіант 5

X \ y	1	3	5	7
-2	0,1	0	0	0
-1	0,1	0,2	0,3	0
0	0	0,1	0,1	0,1

Варіант 12

X \ y	-4	-2	0	2
6	0,1	0,1	0,1	0,1
7	0	0,15	0,2	0,05
8	0	0	0,1	0,1

Варіант 6

X \ y	2	3	4	5
3	0,1	0,1	0,2	0
5	0	0,3	0,1	0,1
7	0	0	0,1	0

Варіант 13

X \ y	-5	-3	-1	1
2	0,1	0,05	0,05	0
4	0,1	0,2	0,2	0,1
7	0,1	0,1	0	0

Варіант 7

X \ y	4	5	6	7
1	0,1	0,1	0	0
3	0,1	0,2	0,1	0,1
5	0	0	0,2	0,1

Варіант 14

X \ y	2	4	5	7
5	0,1	0	0	0
7	0	0,2	0,2	0
9	0	0,2	0,2	0,1

Варіант 8

X \ y	-3	-2	-1	1
3	0,05	0,1	0,1	0
4	0,1	0,2	0,1	0,05
5	0	0	0,1	0,2

Варіант 15

X \ y	1	2	4	6
-7	0,1	0,1	0,05	0
-6	0,2	0,15	0,15	0
-4	0,1	0,1	0	0,05

Варіант 9

X \ y	-3	-1	1	2
5	0,1	0,05	0,05	0,1
6	0	0,2	0,2	0,15
7	0	0	0,1	0,1

Варіант 16

X \ y	-2	-1	0	1
-3	0,1	0	0	0
-1	0	0,1	0,3	0,1
1	0,1	0,1	0,2	0

Варіант 17

X \ y	-7	-5	-4	-2
4	0,15	0,1	0,1	0,1
7	0	0,2	0,1	0,1
9	0	0	0,1	0,05

Варіант 19

X \ y	2	4	5	7
-2	0,1	0,2	0,1	0,1
-1	0	0,1	0,05	0,05
0	0	0,1	0,1	0,1

Варіант 18

X \ y	-3	-1	1	3
1	0,1	0	0	0
2	0,1	0,2	0,25	0
3	0,1	0,1	0,05	0,1

Варіант 20

X \ y	-2	-1	0	1
4	0,1	0,1	0,1	0,1
6	0	0,15	0,2	0,1
7	0	0,05	0,05	0,05

Варіант 21

X \ y	-2	-1	2	3
2	0	0,1	0	0,1
4	0	0,15	0,15	0,1
	0,1	0	0,1	0,2

Варіант 23

X \ y	5	6	7	8
2	0,1	0	0,1	0
3	0,05	0,2	0,1	0
4	0,1	0,15	0,1	0,1

Варіант 22

X \ y	1	2	4	5
-5	0	0	0	0,1
-4	0,1	0,05	0,2	0,1
-2	0,1	0,1	0,15	0,1

Варіант 24

X \ y	3	4	5	7
-3	0,1	0,1	0	0,1
-1	0	0,1	0,2	0,05
1	0,1	0	0,05	0,1

Варіант 25

X \ y	1	2	4	6
2	0,05	0	0	0
3	0,1	0,2	0,15	0,1
5	0,1	0,1	0,1	0,1

Варіант 26

X \ y	-2	-1	0	1
4	0,1	0,1	0,1	0,1
6	0	0,15	0,2	0,1
7	0	0,05	0,05	0,05

Завдання 9. Критерії згоди

За даним інтервальним розподілом вибірки об'єму n при рівні значущості α за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл генеральної сукупності.

1. $n=100, \alpha = 0,01$

(x_i, x_{i+1})	(2, 3)	(3, 4)	(4, 5)	(5, 6)	(6, 7)	(7, 8)
n_i	9	20	27	23	10	11

2. $n=200, \alpha = 0,025$

(x_i, x_{i+1})	(0, 3)	(3, 6)	(6, 9)	(9, 12)	(12, 15)	(15, 18)	(18, 21)
n_i	10	8	34	79	53	12	4

3. $n=150, \alpha = 0,01$

(x_i, x_{i+1})	(1, 3)	(3, 5)	(5, 7)	(7, 9)	(9, 11)	(11, 13)	(13, 15)	(15, 17)
n_i	8	11	18	26	32	25	18	12

4. $n=120, \alpha = 0,05$

(x_i, x_{i+1})	(1, 4)	(4, 7)	(7, 10)	(10, 13)	(13, 16)	(16, 19)
n_i	11	21	32	28	17	11

5. $n=200, \alpha = 0,025$

(x_i, x_{i+1})	(2, 4)	(4, 6)	(6, 8)	(8, 10)	(10, 12)	(12, 14)	(14, 16)	(16, 18)
n_i	7	31	51	46	19	24	16	6

6. $n=130, \alpha = 0,05$

(x_i, x_{i+1})	(6, 9)	(9, 12)	(12, 15)	(15, 18)	(18, 21)	(21, 24)	(24, 27)
n_i	4	13	36	39	27	7	4

7. $n=140, \alpha = 0,01$

(x_i, x_{i+1})	(0, 5)	(5, 10)	(10, 15)	(15, 20)	(20, 25)	(25, 30)	(30, 35)
n_i	21	9	28	34	13	8	37

8. $n=100, \alpha = 0,01$

(x_i, x_{i+1})	(4, 7)	(7, 10)	(10, 13)	(13, 16)	(16, 19)	(19, 22)
n_i	11	27	45	31	19	7

9. $n=160, \alpha = 0,025$

(x_i, x_{i+1})	(0, 2)	(2, 4)	(4, 6)	(6, 8)	(8, 10)	(10, 12)	(12, 14)	(14, 16)
n_i	7	12	19	29	36	28	18	11

10. $n=120, \alpha = 0,05$

(x_i, x_{i+1})	(9, 11)	(11, 13)	(13, 15)	(15, 17)	(17, 19)	(19, 21)
n_i	12	24	31	26	16	11

11. $n=140, \alpha = 0,01$

(x_i, x_{i+1})	(3, 5)	(5, 7)	(7, 9)	(9, 11)	(11, 13)	(13, 15)	(15, 17)
n_i	7	10	22	38	32	21	10

12. $n=150, \alpha = 0,025$

(x_i, x_{i+1})	(0, 4)	(4, 8)	(8, 12)	(12, 16)	(16, 20)	(20, 24)	(24, 28)	(28, 32)
n_i	6	14	20	26	32	25	18	9

13. $n=130, \alpha = 0,01$

(x_i, x_{i+1})	(1, 7)	(7, 13)	(13, 19)	(19, 25)	(25, 31)	(31, 37)
n_i	13	23	40	25	15	14

14. $n=160, \alpha = 0,05$

(x_i, x_{i+1})	(2, 4)	(4, 6)	(6, 8)	(8, 10)	(10, 12)	(12, 14)	(14, 16)	(16, 18)
n_i	2	16	39	43	24	23	12	1

15. $n=200, \alpha = 0,025$

(x_i, x_{i+1})	(6, 8)	(8, 10)	(10, 12)	(12, 14)	(14, 16)	(16, 18)	(18, 20)
n_i	5	19	51	59	42	15	9

16. $n=100, \alpha = 0,05$

(x_i, x_{i+1})	(0, 5)	(5, 10)	(10, 15)	(15, 20)	(20, 25)	(25, 30)	(30, 35)
n_i	14	19	11	14	23	11	8

17. $n=110, \alpha = 0,01$

(x_i, x_{i+1})	(4, 7)	(7, 10)	(10, 13)	(13, 16)	(16, 19)	(19, 22)
n_i	8	25	32	31	9	5

18. $n=150, \alpha = 0,05$

(x_i, x_{i+1})	(1, 2)	(2, 3)	(3, 4)	(4, 5)	(5, 6)	(6, 7)	(7, 8)	(8, 9)
n_i	6	11	19	26	37	25	17	9

19. $n=120, \alpha = 0,025$

(x_i, x_{i+1})	(2, 5)	(5, 8)	(8, 11)	(11, 14)	(14, 17)	(17, 20)	(20, 23)
n_i	2	5	34	42	29	7	1

20. $n=130, \alpha = 0,05$

(x_i, x_{i+1})	(0, 1)	(1, 2)	(2, 3)	(3, 4)	(4, 5)	(5, 6)	(6, 7)
n_i	18	9	23	24	13	8	35

21. $n=120, \alpha = 0,01$

(x_i, x_{i+1})	(7, 16)	(16, 25)	(25, 34)	(34, 43)	(43, 52)	(52, 61)
n_i	8	24	47	26	11	4

22. $n=200, \alpha = 0,025$

(x_i, x_{i+1})	(1, 3)	(3, 5)	(5, 7)	(7, 9)	(9, 11)	(11, 13)	(13, 15)	(15, 17)
n_i	5	12	46	63	41	24	7	2

23. $n=150, \alpha = 0,025$

(x_i, x_{i+1})	(1, 6)	(6, 11)	(11, 16)	(16, 21)	(21, 26)	(26, 31)
n_i	4	22	72	28	19	5

24. $n=130, \alpha = 0,05$

(x_i, x_{i+1})	(3, 7)	(7, 11)	(11, 15)	(15, 19)	(19, 23)	(23, 27)	(27, 31)
n_i	6	13	20	63	17	9	2

25. $n=120, \alpha = 0,01$

(x_i, x_{i+1})	(-1, 0)	(0, 1)	(1, 2)	(2, 3)	(3, 4)	(4, 5)	(5, 6)	(6, 7)
n_i	3	9	14	57	18	10	7	2

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	0,664	0,81	0,445	3,10	0,045
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
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

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

Таблиця 2.

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3445	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3075	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0503	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

$$\text{Значення функції } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt.$$

Таблиця 3.

x	0	1	2	3	4	3	6	7	8	9
0,0	0,0000	0,00399	0,00798	0,01197	0,01595	0,01994	0,02392	0,02790	0,03188	0,03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14310	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774
1,4	41924	42073	42220	42384	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46582	46638	46712	46784	46856	46926	46995	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48923	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3,0	49865	49869	49873	49877	49882	49886	49889	49893	49897	49900
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976
3,5	49977	49978	49978	49979	49980	49981	49982	49982	49983	49984
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49993
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998

Квантилі U_p нормального розподілу $N(0,1)$.

Таблиця 4.

p	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
U_p	1,282	1,645	1,960	2,326	2,576	3,090	3,291

k/p	0,005	0,010	0,025	0,05	0,10	0,90	0,95	0,975	0,990	0,995	0,999
1	0,04393	0,03157	0,03982	0,02393	0,0158	2,71	3,84	5,02	6,63	7,88	10,8
2	0,0100	0,0201	0,0506	0,103	0,211	4,61	5,99	7,38	9,21	10,6	13,3
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,3	12,8	16,3
4	0,207	0,297	0,484	0,711	1,06	7,78	9,49	11,1	13,3	14,9	18,5
5	0,412	0,554	0,831	1,15	1,61	9,24	11,1	12,8	15,1	16,7	20,5
6	0,676	0,872	1,24	1,64	2,20	10,6	12,6	14,4	16,8	18,5	22,5
7	0,989	1,24	1,69	2,17	2,83	12,0	14,1	16,0	18,5	20,3	24,3
8	1,34	1,65	2,18	2,73	3,49	13,4	15,5	17,5	20,1	22,0	26,1
9	1,73	2,09	2,70	3,33	4,17	14,7	16,9	19,0	21,7	23,6	27,9
10	2,16	2,56	3,25	3,94	4,87	16,0	18,3	20,5	23,2	25,2	29,6
11	2,60	3,05	3,82	4,57	5,58	17,3	19,7	21,9	24,7	26,8	31,3
12	3,07	3,57	4,40	5,23	6,30	18,5	21,0	23,3	26,2	28,3	32,9
13	3,57	4,11	5,01	5,89	7,04	19,8	22,4	24,7	27,7	29,8	34,5
14	4,07	4,66	5,63	6,57	7,79	21,1	23,7	26,1	29,1	31,3	36,1
15	4,60	5,23	6,26	7,26	8,55	22,3	25,0	27,5	30,6	32,8	37,7
16	5,14	5,81	6,91	7,96	9,31	23,5	26,3	28,8	32,0	34,3	39,3
17	5,70	6,41	7,56	8,67	10,1	24,8	27,6	30,2	33,4	35,7	40,8
18	6,26	7,01	8,23	9,39	10,9	26,0	28,9	31,5	34,8	37,2	42,3
19	6,84	7,63	8,91	10,1	11,7	27,2	30,1	32,9	36,2	38,6	43,8
20	7,43	8,26	9,59	10,9	12,4	28,4	31,4	34,2	37,6	40,0	45,3
25	10,5	11,5	13,1	14,6	16,5	34,4	37,7	40,6	44,3	46,9	52,6
30	13,8	15,0	16,8	18,5	20,6	40,3	43,8	47,0	50,9	53,7	59,7
35	17,2	18,5	20,6	22,5	24,8	46,1	49,8	53,2	57,3	60,3	66,6
50	28,0	29,7	32,4	34,8	37,7	63,2	67,5	71,4	76,2	79,5	86,7
100	67,3	70,1	74,2	77,9	82,4	118,5	124,3	129,6	135,6	140,2	149,4

Квантилі розподілу Стьюдента $t_p(k)$.

Таблиця 6.

$k \setminus p$	0,750	0,900	0,950	0,975	0,990	0,995	0,999
1	1,000	3,078	6,314	12,706	31,821	63,657	318
2	0,816	1,886	2,920	4,303	6,965	9,925	22,3
3	0,765	1,638	2,353	3,182	4,541	5,841	10,2
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160

Квантилі розподілу Фішера $F_p(k_1, k_2)$:

$p = 0,9$

Таблиця 7.

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	62,26	63,06
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,46	9,48
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,17	5,14
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,82	3,73
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,17	3,12
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,80	2,74
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,56	2,49
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,38	2,32
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,25	2,18
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,16	2,08
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,08	2,00
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,01	1,93
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14	1,96	1,88
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10	1,91	1,83
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	1,87	1,79
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03	1,84	1,75
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00	1,81	1,72
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1,78	1,69
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96	1,76	1,67
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,74	1,64
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87	1,66	1,56
120	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65	1,41	1,26

$p = 0,95$

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	161,4	199,5	215,7	224,6	230,2	234,0	2,368	238,9	240,5	241,9	250,1	253,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,46	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,62	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,75	5,66
5	6,61	5,79	5,41	5,19	5,05	4,35	4,88	4,82	4,77	4,74	4,50	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,81	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,38	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,08	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	2,86	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,70	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,57	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,30	2,75	2,47	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,38	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,31	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,25	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,19	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,15	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,11	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,07	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,04	1,90
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	1,92	1,77
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,55	1,35

$$p = 0,975$$

Продовження табл. 7

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	1001	1014
2	38,51	39,00	34,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,46	39,49
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,08	13,95
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,46/	8,31
5	10,01	8,43	7,76	7,39,	7,15	6,98	6,35	6,76	6,68	6,62	6,23	6,07
6	8,71	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,07	4,90
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,87	4,76	4,36	4,20
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	3,89	3,73
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,56	3,39
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,31	3,14
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,12	2,94
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	2,96	2,79
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	2,84	2,66
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	2,73	2,55
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,64	2,46
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,57	2,38
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,50	2,32
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,44	2,26
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,39	2,20
20	5,37	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,35	2,16
25	5,67	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,18	1,93
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	1,69	1,43

$$p = 0,990$$

Продовження табл. 7

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6261	6339
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,47	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	26,50	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	13,84	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,38	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,23	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	5,99	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,20	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,65	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,25	4,00
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	3,94	3,69
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	3,70	3,45
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,51	3,25
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,35	3,09
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,21	2,96
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,10	2,84
17	8,40	6,11	5,13	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,00	2,75
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	2,92	2,66
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	2,84	2,58
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	2,78	2,52
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,54	2,27
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	1,86	1,53

$$p = 0,995$$

Продовження табл. 7

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	25044	25359
2	198,5	199,0	199,2	199,2	199,3	199,3	199,4	199,4	199,4	199,4	199,5	199,5
3	55,55	49,80	47,47	46,19	45,39	44,84	44,43	44,13	43,88	43,69	42,47	41,99
4	31,13	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	19,89	19,47
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	12,66	12,27
6	18,63	14,54	12,92	12,03	11,46	11,07	10,79	10,57	10,39	10,25	9,36	9,00
7	16,24	12,40	10,88	10,05	9,52	9,16	8,89	8,68	8,51	8,38	7,53	7,19
8	14,69	11,04	9,60	8,81	8,30	7,95	7,69	7,50	7,34	7,21	6,40	6,06
9	13,61	10,11	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	5,62	5,30
10	12,83	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,07	4,75
11	12,23	8,91	7,60	6,88	6,42	6,10	5,86	5,68	5,54	5,42	4,65	4,34
12	11,75	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,33	4,01
13	11,37	8,19	6,93	6,23	5,79	5,48	5,25	5,08	4,94	4,82	4,07	3,76
14	11,06	7,92	6,68	6,00	5,56	5,26	5,03	4,86	4,72	4,60	3,86	3,55
15	10,80	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	3,69	3,37
16	10,58	7,51	6,30	5,64	5,21	4,91	4,69	4,52	4,38	4,27	3,54	3,22
17	10,38	7,35	6,16	5,50	5,07	4,78	4,56	4,39	4,25	4,14	3,41	3,10
18	10,22	7,21	6,03	5,37	4,96	4,66	4,44	4,28	4,14	4,03	3,30	2,99
19	10,07	7,09	5,92	5,27	4,85	4,56	4,34	4,18	4,04	3,93	3,21	2,89
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,12	2,81
25	9,48	6,60	5,46	4,84	4,43	4,15	3,94	3,78	3,64	3,54	2,82	2,50
120	8,18	5,54	4,50	3,92	3,55	3,28	3,09	2,93	2,81	2,71	1,98	1,61

$$p = 0,999$$

Закінчення табл. 7

$k_2 \setminus k_1$	1	2	3	4	5	6	7	8	9	10	30	120
1	4053*	5000*	5404*	5625*	5764*	5859*	5929*	5981*	6023*	6056*	6261*	6340*
2	998,5	999,0	999,2	999,2	999,3	999,3	999,4	999,4	999,4	999,4	999,5	999,5
3	167,0	148,5	141,1	137,1	134,6	132,8	131,6	130,6	129,9	129,2	125,4	124,0
4	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,47	48,05	45,43	44,40
5	47,18	37,12	33,20	31,09	29,75	28,54	28,16	27,64	27,24	26,92	24,87	24,06
6	35,51	27,00	23,70	21,92	20,81	20,03	19,46	19,03	18,69	18,41	16,67	15,99
7	29,25	21,69	18,77	17,19	16,21	15,52	15,02	14,63	14,33	14,08	12,53	11,91
8	25,42	18,49	15,83	14,39	13,49	12,86	12,40	12,04	11,77	11,54	10,11	9,53
9	22,86	16,39	13,90	12,56	11,71	11,13	10,70	10,37	10,11	9,89	9,55	8,00
10	21,04	14,91	12,55	11,28	10,48	9,92	9,52	9,20	8,96	8,75	7,47	6,94
11	19,69	13,81	11,56	10,35	9,58	9,05	9,66	8,35	8,12	7,92	6,68	6,17
12	18,64	12,97	10,80	9,63	8,89	8,38	8,00	7,71	7,48	7,29	6,09	5,59
13	17,81	12,31	10,21	9,07	8,35	7,86	7,49	7,21	6,98	6,80	5,63	5,14
14	17,14	11,78	9,73	8,62	7,92	7,43	7,08	6,80	6,58	6,40	5,25	4,77
15	16,59	11,34	9,34	8,25	7,57	7,09	6,74	6,47	6,26	6,08	4,95	4,47
16	16,12	10,97	9,00	7,94	7,27	6,81	6,46	6,19	5,98	5,81	4,70	4,23
17	15,72	10,66	8,73	7,68	7,02	6,56	6,22	5,96	5,75	5,58	4,48	4,02
18	15,38	10,39	8,49	7,46	6,81	6,35	6,02	5,76	5,56	5,39	4,30	3,84
19	15,08	10,16	8,28	7,26	6,62	6,18	5,85	5,59	5,39	5,22	4,14	3,68
20	14,82	9,95	8,10	7,10	6,46	6,02	5,69	5,44	5,24	5,08	4,00)	3,54
25	13,88	9,22	7,45	6,49	5,88	5,46	5,15	4,91	4,71	4,56	3,52	3,06
120	11,38	7,32	5,79	4,95	4,42	4,04	3,77	3,55	3,38	3,24	2,56	1,76

4053* дорівнює $4053 \cdot 10^2$

Таблиця 8.

Числові значення контрольного критерію $Q(P, n)$

n	Q		
	$P = 90 \%$	$P = 95 \%$	$P = 99 \%$
3	0,89	0,94	0,99
4	0,68	0,77	0,89
5	0,56	0,64	0,76
6	0,48	0,56	0,70
7	0,43	0,51	0,64
8	0,40	0,48	0,58
9	0,38	0,46	0,55

Таблиця 9.

Критерії значущості для перевірки гіпотез щодо дисперсії
нормально розподіленої генеральної сукупності

Гіпотеза H_0 , яку перевіряють	Припущення відносно m	Статистика Z критерію ^{*)}	Розподіл Z : $f(z/H_0)$	Область прийняття гіпотези H_0 для двостороннього критерію	Альтернативна гіпотеза і область прийняття гіпотези H_0 для правостороннього критерію
$\sigma^2 = \sigma_0^2$	m відоме	$\frac{nS_0^2}{\sigma_0^2}$	$\chi^2(n)$	$\chi_{\alpha/2}^2(n) < \frac{nS_0^2}{\sigma_0^2} < \chi_{1-\alpha/2}^2(n)$	$H_1: \sigma^2 > \sigma_0^2$ $\frac{nS_0^2}{\sigma_0^2} < \chi_{1-\alpha}^2(n)$
$\sigma^2 = \sigma_0^2$	m невідоме $m = \bar{X}$	$\frac{(n-1)S_0^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi_{\alpha/2}^2(n-1) < \frac{(n-1)S_0^2}{\sigma_0^2} < \chi_{1-\alpha/2}^2(n-1)$	$H_1: \sigma^2 > \sigma_0^2$ $\frac{(n-1)S_0^2}{\sigma_0^2} < \chi_{1-\alpha}^2(n-1)$
$\sigma_1^2 = \sigma_2^2$	m_1 і m_2 відомі	S_{01}^2/S_{02}^2 , $S_{01}^2 > S_{02}^2$	$F(n_1, n_2)$	$\frac{S_{01}^2}{S_{02}^2} < F_{1-\alpha/2}(n_1, n_2)$	$H_1: \sigma_1^2 > \sigma_2^2$ $\frac{S_{01}^2}{S_{02}^2} < F_{1-\alpha}(n_1, n_2)$
$\sigma_1^2 = \sigma_2^2$	m_1 і m_2 невідомі $m_1 = \bar{X}_1, m_2 = \bar{X}_2$	S_1^2/S_2^2 , $S_1^2 > S_2^2$	$F(n_1 - 1, n_2 - 1)$	$\frac{S_1^2}{S_2^2} < F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$	$H_1: \sigma_1^2 > \sigma_2^2$ $\frac{S_1^2}{S_2^2} < F_{1-\alpha}(n_1 - 1, n_2 - 1)$

^{*)}Значення статистики Z наведені при умові, що гіпотеза H_0 вірна.

Таблиця 10.

Критерії значущості для перевірки гіпотез щодо математичного сподівання
нормально розподіленої генеральної сукупності

Гіпотеза H_0 , яку перевіряють	Припущення відносно σ^2	Статистика Z критерію ^{*)}	Розподіл Z : $f(z/H_0)$	Область прийняття гіпотези H_0 для двостороннього критерію	Альтернативна гіпотеза і область прийняття гіпотези H_0 для правостороннього критерію
$m = m_0$	σ^2 відома	$\frac{\bar{x} - m_0}{\sigma/\sqrt{n}}$	$N(0,1)$	$\frac{ \bar{x} - m_0 }{\sigma/\sqrt{n}} < u_{1-\alpha/2}$	$H_1 : m > m_0$; $\frac{\bar{x} - m_0}{\sigma/\sqrt{n}} < u_{1-\alpha}$
$m = m_0$	σ^2 невідома	$\frac{\bar{x} - m_0}{S/\sqrt{n}}$	$T(n-1)$	$\frac{ \bar{x} - m_0 }{s/\sqrt{n}} < t_{1-\alpha/2}(n-1)$	$H_1 : m > m_0$; $\frac{\bar{x} - m_0}{s/\sqrt{n}} < t_{1-\alpha}(n-1)$
$m_1 = m_2$	σ_1^2 і σ_2^2 відомі	$\frac{\bar{x}_1 - \bar{x}_2}{\sigma_{12}}$, де $\sigma_{12} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$N(0,1)$	$\frac{ \bar{x}_1 - \bar{x}_2 }{\sigma_{12}} < u_{1-\alpha/2}$	$H_1 : m_1 > m_2$; $\frac{\bar{x}_1 - \bar{x}_2}{\sigma_{12}} < u_{1-\alpha}$

Гіпотеза H_0 , яку перевіряють	Припущення відносно σ^2	Статистика Z критерію ^{*)}	Розподіл Z : $f(z/H_0)$	Область прийняття гіпотези H_0 для двостороннього критерію	Альтернативна гіпотеза і область прийняття гіпотези H_0 для правостороннього критерію
$m_1 = m_2$	σ_1^2 і σ_2^2 відомі та гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ приймається, $\sigma_1^2 = S_1^2, \sigma_2^2 = S_2^2$	$\frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{1/n_1 + 1/n_2}}$, де $S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$	$T(n_1 + n_2 - 2)$	$\frac{ \bar{x}_1 - \bar{x}_2 }{s\sqrt{1/n_1 + 1/n_2}} < t_{1-\alpha/2}(n_1 + n_2 - 2)$	$H_1 : m_1 > m_2;$ $\frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{1/n_1 + 1/n_2}} < t_{1-\alpha}(n_1 + n_2 - 2)$
$m_1 = m_2$	σ_1^2 і σ_2^2 відомі та гіпотеза $H_0 : \sigma_1^2 = \sigma_2^2$ відкидається, $\sigma_1^2 = S_1^2, \sigma_2^2 = S_2^2$	$\frac{\bar{x}_1 - \bar{x}_2}{\sigma_{12}}$, де $S_{12} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$	$T(k)$, де $k = \frac{S_{12}^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$	$\frac{ \bar{x}_1 - \bar{x}_2 }{S_{12}} < t_{1-\alpha/2}(k)$	$H_1 : m_1 > m_2;$ $\frac{\bar{x}_1 - \bar{x}_2}{S_{12}} < t_{1-\alpha}(k)$

^{*)}Значення статистики Z наведені при умові, що гіпотеза H_0 вірна

Література

1. Вентцель Е.С. Теория вероятностей / Е.С. Вентцель. – М.: Высшая школа, 1999. – 576 с.
2. Волощенко А.Б. Теорія ймовірностей та математична статистика: навч.-метод. посібник для самост. вивч. дисц. / А.Б. Волощенко, І.А. Джалладова. – К.: КНЕУ, 2003. – 256 с.
3. Гихман И.И. Теория вероятностей и математическая статистика / И.И. Гихман, А.В. Скороход, М.И. Ядренко. – К.: Вища школа, 1988. – 438 с.
4. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М.: Высш. шк., 1979. – 400 с.
5. Гмурман В.Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М.: Высш. шк., 1977.-479 с.
6. Гнеденко Б.В. Элементарное введение в теорию вероятностей / Б.В. Гнеденко, А.Я. Хинчин. – М.: ГПТИ, 1946. – 128 с.
7. Жлуктенко В.І. Теорія ймовірностей і математична статистика: навч.-метод. посібник. У 2ч. Ч.1. Теорія ймовірностей / В.І.Жлуктенко, С.І. Наконечний. – К.: КНЕУ, 2000.–304с.
8. Колмогоров А.Н. Основные понятия теории вероятностей / А.Н. Колмогоров. – М.: Наука, 1974. – 130 с. 9. Колмогоров А.Н. Введение в теорию вероятностей / А.Н. Колмогоров, И.Г. Журбенко, А.В. Прохоров. – М.: Наука, 1982. – 160 с.
10. Кремер Н.Ш. Теория вероятностей и математическая статистика / Н.Ш. Кремер. – М.: Юнити-Дана, 2007. – 573 с.
11. Гончаров В.В., Гончарова С.Я., Личук М.В. Теорія ймовірностей і математична статистика. – Кіровоград: КНТУ, 2006 (р/н 1167).
12. Булдигін В.В., Буценко Ю.П., Диховичний О.О. Теорія ймовірностей. – К.: ТІМС. –1999.
13. Моцний Ф.В. Курс лекцій з теорії ймовірностей: навч. посіб./ Ф.В. Моцний. – К.: ДП „Інформаційно-аналітичне агенство”, 2010. – 122 с. 169