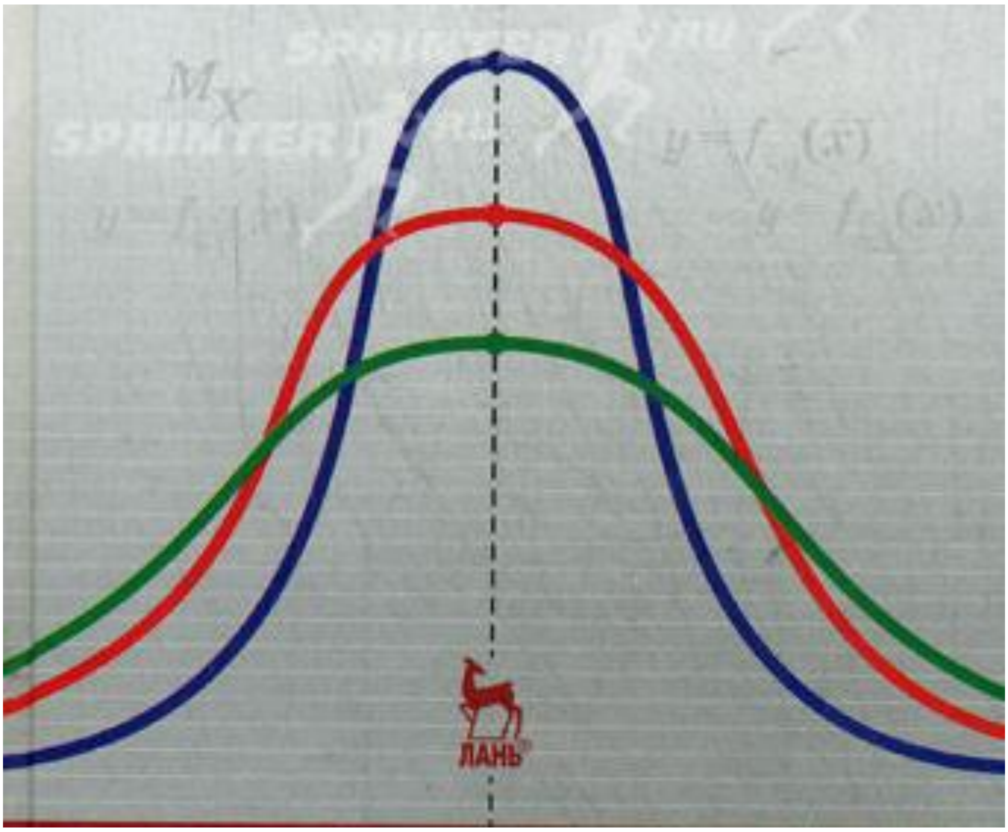
Міністерство освіти і науки, молоді та спорту України

Чернігівський національний педагогічний університет імені Т.Г.Шевченка

КІЛОЧИЦЬКА Т.В.

ЕЛЕМЕНТИ ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ



Чернігів, 2011

Навчально-методичний посібник «Елементи теорії ймовірностей та математичної статистики» написаний з використанням елементів історії математики.

Кожна лекція містить стисло викладений теоретичний матеріал з прикладами, малюнками, зразками розв’язування типових завдань.

Для студентів вищих педагогічних навчальних закладів зі спеціальності 6.010104 „Професійна освіта“, істориків математики, методистів і всіх тих, хто цікавиться історією науки.

Рекомендовано до друку та використання в електронному вигляді Вченою радою ЧНПУ імені Т.Г.Шевченка (Протокол № 8 від 30 березня 2011 року)

ЗМІСТ

ВСТУП

Навчально-методичний посібник «Елементи теорії ймовірностей та математичної статистики» побудований з залученням максимальної кількості літературних джерел з історії науки, зокрема з історії теорії ймовірностей та математичної статистики. В посібнику історичний матеріал органічно переплітається з викладом основного матеріалу навчального курсу, доповнює і не заважає його вивченню.

Використання елементів історії математики при викладанні теорії йморівностей має метою:

- підвищення інтересу студентів до вивчення вищої математики, до наукової творчості;
- посилення міжпредметних зв'язків, гуманізації освіти;
- поглиблення розуміння нового матеріалу, розвиток критичного мислення;
- з'ясування ролі і місця математики в практичній діяльності людей різних історичних епох;
- визначення передумов створення нових математичних теорій, методів доведення теорем.

Для спеціальності «професійна освіта» теорія ймовірності складає всього 24 години (14 лекційних і 10 практичних занять). Це є одним з змістових модулів курсу «Вища математика» і вивчається наприкінці ІІІ семестру. Цей модуль містить такі теми:

Тема 1. (1 лекція, 1 практика)

Елементи комбінаторики

Елементи історії комбінаторики. Поняття комбінаторної задачі. Основні правила комбінаторики. Розміщення, перестановки, сполучення без повторень. Розміщення, перестановки, сполучення з повтореннями.

Тема 2. (1 лекція, 1 практика)

Випадкові події

Стохастичний експеримент. Випадкові події. Дії над подіями. Статичне та класичне означення ймовірності. Властивості ймовірності.

Тема 3. (2 лекції, 1 практика)

Властивості ймовірності

Теорема про додавання ймовірностей для несумісних подій. Умовні ймовірності та теорема множення ймовірностей. Теорема додавання ймовірностей двох довільних подій.

Формула повної ймовірності та формули Байєса.

Тема 4. (2 лекції, 1 практика)

Випадкові величини

Поняття випадкової величини. Закон розподілу дискретної випадкової величини. Основні характеристики дискретної випадкової величини: математичне сподівання, дисперсія.

Інтегральна функція (функція розподілу) неперервної випадкової величини та її властивості. Диференціальна функція (щільність розподілу ймовірностей). Основні характеристики неперервної випадкової величини: математичне сподівання, дисперсія. Два важливі розподіли випадкових величин (біноміальний розподіл, нормальний розподіл).

Тема 5. (1 лекція, 1 практика)

Кореляційний зв'язок між випадковими величинами. Регресія

Встановлення залежності між випадковими величинами. Коваріація і коефіцієнт кореляції. Кореляційне поле, його побудова. Лінія регресії. Знаходження коефіцієнтів лінії регресії.

ЗАРОДЖЕННЯ І РОЗВИТОК ТЕОРІЇ ЙМОВІРНОСТЕЙ

Теорія ймовірностей виникла в середині XVII ст. Перші праці з теорії ймовірностей належать французьким ученим Б. Паскалю і П. Ферма (листування, 1654) і голландському вченому Християну Гюйгенсу (перший трактат з теорії ймовірностей «Про розрахунки при азартних іграх», 1657). Швейцарський математик Я. Бернуллі встановив закон великих чисел для схеми незалежних випробувань з двома результатами (опубліковано в 1713).

Наступний (другий) період історії теорії ймовірностей (XVIII ст. – початок XIX ст.) пов'язаний з іменами вчених А. Муавра (Англія), П. Лапласа (Франція), К. Гауса (Німеччина) і С. Пуассона (Франція). При розрахунку орбіт комет, який доводилося проводити при малій кількості спостережень, Ньютон, Лежандр (Франція, 1805), Гаус (1808, 1815) користувалися методом, який в подальшому розвитку отримав назву метод найменших квадратів. В цей період теорія ймовірностей знаходить ряд дуже актуальних застосувань у природознавстві і техніці (головним чином у теорії помилок спостережень, що розвилася в зв'язку з потребами геодезії та астрономії, і в теорії стрільби). У першій половині XIX ст. український математик М.В. Остроградський займався задачами теорії ймовірностей та математичної статистики, російський математик В.Я. Буняковський – застосуванням теорії ймовірностей до страхової справи, статистики та демографії).

Третій період історії теорії ймовірностей (II половина XIX ст.) пов'язаний з іменами російських математиків П.Л. Чебишева, О. М. Ляпунова та А.А. Маркова (старшого). П.Л. Чебишев та його учні О.М.Ляпунов і А.А.Марков сформулювали і розв’язали низку спільних задач теорії ймовірностей, узагальнюючи теореми Бернуллі і Лапласа. У 1867 р. Чебишев довів закон великих чисел в загальному випадку. Він вперше сформулював (1887) центральну граничну теорему для сум незалежних випадкових величин і вказав один з методів її доведення. Іншим методом О.М.Ляпунов отримав (1901) близьке до сучасного доведення цієї теореми. Марков вперше розглянув (1907) один випадок залежних випробувань, який згодом отримав назву ланцюгів Маркова.

У Західній Європі в ІІ-й половині XIX ст. отримали великий розвиток праці з математичної статистики (у Бельгії - А. Кетле, в Англії - Ф. Гальтон) та статистичної фізики (в Австрії - Л. Больцман), які поряд з основними теоретичними роботами Чебишева, Ляпунова та Маркова створили основу для розширення проблематики теорії ймовірностей у четвертому (сучасному) періоді її розвитку. Цей період історії теорії ймовірностей характеризується розширенням кола її застосувань (в астрономії, фізиці, біології, сільському господарстві, промисловості, медицині тощо), створенням декількох систем математичного обґрунтування теорії ймовірностей, нових потужних методів, що вимагають застосування (крім класичного аналізу) засобів теорії множин, теорії функцій дійсного змінного і функціонального аналізу. У цей період відбувся значний розвиток теорії ймовірностей за кордоном (у Франції - Е. Борель, П. Леві, М. Фреше, у Німеччині - Р. Мізес, в США - Н. Вінер, В. Феллер, Дж. Дуб, у Швеції - Г. Крамер). Радянська наука продовжує займати значне, а в більшості напрямків і провідне становище. Зокрема, С.Н. Бернштейн значно узагальнив класичні граничні теореми Чебишева, Ляпунова та Маркова і вперше в Росії широко застосував теорію ймовірностей до природознавства. У Москві А. Я. Хинчин і А.Н. Колмогоров застосували до задач теорії ймовірностей методи теорії функцій дійсної змінної. Пізніше (у 30-х рр.) вони (і Є. Є. Слуцький) заклали основи теорії випадкових процесів. В. І. Романовський (Ташкент) і М.В. Смирнов (Москва) застосували теорію ймовірностей до математичної статистики. Крім московської групи фахівців з теорії ймовірностей, в даний час розробкою проблем теорії ймовірностнй займаються вчені в Петербурзі і в Києві.

ЛЕКЦІЯ 1. ЕЛЕМЕНТИ КОМБІНАТОРИКИ

План лекції

1. *Елементи історії комбінаторики. Поняття комбінаторної задачі*
2. *Основні правила комбінаторики.*
3. *Основні формули комбінаторики*

1. Елементи історії комбінаторики. Поняття комбінаторної задачі

Комбінаторика – одна з важливих розділів математики. З комбінаторними задачами мають справу економісти, менеджери, соціологи, політологи. Комбінаторні методи покладені в основу розв'язання багатьох задач теорії ймовірностей.

Комбинаторика широко використовується при розв’язанні задач алгебри, геометрії, математичного аналізу, і використовує геометричні, аналітичні та алгебраїчні методи досліджень. Комбінаторні методи використовуються в криптографії - як для розробки шифрів, так і для їх злому.

У практичному житті, серед різних математичних задач часто зустрічаються такі, в яких треба вибирати з деякої множини об'єктів підмножини елементів, які мають ті чи інші властивості, розміщувати їх у певному порядку за певними правилами знаходити число способів, за якими таке розташування можна здійснити. Наприклад, знайти кількість можливих шляхів доставки продукції до певного магазину даним поставщиком і т.д.

Оскільки в таких задачах ідеться про комбінації об'єктів, то їх називають комбінаторними задачами.

Розділ математики, в якому обґрунтовується теорія розв'язування комбінаторних задач, називається комбінаторикою.

Окремі комбінаторні задачі з'явилися дуже давно. У працях стародавніх індійських вчених знайдені основні формули комбінаторики.

Класична задача комбінаторики: «скільки є способів витягти m елементів з n можливих» згадується ще в сутрах стародавньої Індії (починаючи приблизно з IV століття до н. е.). Античні греки також розглядали окремі комбінаторні задачі. Хрісіпп (III століття до н. е.) та Гіппарх (II століття до н. е.) підраховували скільки наслідків можна отримати з 10 аксіом. Аристотель в праці «Перша аналітика» безпомилково перерахував всі можливі типи тричленних силогізмів.

У XII столітті індійський математик Бхаскара у основній праці «Лілаваті» докладно дослідив задачі, пов'язані з перестановками і поєднаннями, включаючи перестановки з повтореннями.

У Західній Європі відкриття в галузі комбінаторики зробили два єврейських дослідника, Авраам ібн Езра (XII століття) і Леві бен Гершем (Герсонід, XIV століття). Герсонід дав формули для обчислення числа розміщень і сполучень.

Кілька комбінаторних задач містить «Книга абака» (Фібоначчі, XIII століття). Наприклад, він поставив завдання знайти найменше число гирь, достатню для зважування будь-якого товару вагою від 1 до 40 фунтів.

Джироламо Кардано написав математичне дослідження гри в кістки, опубліковане посмертно.

В працях Н. Тортальї (XVI ст.) є елементи комбінаторики, але повне дослідження перестановок, розміщень, сполучень він не зробив. У зв’язку розвитком алгебри многочленів подальший розвиток теорія ймовірностей набула в XVII ст. в працях французьких математиків Б.Паскаля (1623-1662) і П.Ферма (1601-1665).

Термін "комбінаторика" введений німецьким вченим Готфридом Вильгельмом Лейбніцем. Він включив в це поняття всю кінцеву математику і навіть логіку (розумів дуже широко). В 1666 р. Лейбніц в праці "Про комбінаторне мистецтво" ввів спеціальні символи, терміни для підмножин і операцій над ними і знайшов всі k-сполучення з n елементів, вивів властивості сполучень. Він побудував таблиці сполучень до n = k = 12, займався застосуваннями комбінаторики до логіки, арифметики, до проблем складання віршів тощо.

В 1713 р. опублікована праця учня В.Лейбніца Я. Бернуллі "Мистецтво передбачень", в якому комбінаториці присвячена друга частина. В ній є формули для числа перестановок з n елементів, для числа сполучень без повторень і з повтореннями, для числа розміщень з повтореннями і без повторень. Ця праця в XVIII ст. була досить популярною. В працях Я. Бернуллі та Лейбніца вивчались властивості сполучень, розміщень, перестановок.

Остаточно комбінаторика як самостійний розділ математики оформилася в працях Ейлера (1707-1783). Крім перестановок і сполучень, Ейлер вивчав розбиття, а також поєднання і розміщення з умовами.

Наприкінці XVIII ст. науковці комбинаторної школи Гінденбурга, намагались побудувати загальну комбінаторну теорію, використовуючи нескінченні ряди.

В XX ст. почався процес алгебраізації комбінаторики в працях Дж.-К. Рота (1964), пізніше Р. Стенлі. Вони вивчали частково впорядковані множини, властивостей функцій Мьобіуса, абстрактних властивостей лінійної залежності та їх використання при розв’язуванні комбінаторних задач сприяли збагаченню комбінаторних методів дослідження та подальшої інтеграції комбінаторики в сучасну математику.

В XX столітті комбінаторика почала широко застосовуватись в обчислювальній техніці, кібернетиці, економіці та інших науках.

В останні десятиріччя інтерес до комбінаторних задач значно пожвавився у зв’язку з розробкою оптимальних планів виробництва, транспортування, розміщення підприємств. Ці задачі вимагають надзвичайно великої кількості варіантів, тому сучасні комбінаторні методи пов'язані з застосуванням швидкодіючих електронних обчислювальних машин, комп'ютерної техніки.

2. Основні правила комбінаторики.

Правило суми. Якщо елемент *a* з множини *A* можна вибрати m способами, а елемент *b* множини *B* можна вибрати *n* способами, причому ніякий вибір *a* не збігається з жодним з виборів *b* , то число способів, якими можна здійснити хоча б один з цих виборів, дорівнює сумі $m+n$.

Приклад 1. Якщо маємо 2 кошика, в одному 5 яблук, в іншому 6 груш, то кількість способів вибору довільного фрукту 5+6=11.

Правило суми легко узагальнюється.

Узагальнене правило суми. Нехай елемент a_1 множини A_1 можна вибрати m_1 способами, елемент a_2 множини A_2 - m_2 способами,..., елемент $a_к$ - множини $A_к$ можна вибрати $m_к$ способами. Тоді число способів, якими можна здійснити хоча б один з цих виборів, дорівнює сумі $m_1 + m_2 + ...+ m_к$.

Правило добутку. Якщо елемент a множини A можна вибрати m способами і при кожному a цих виборів елемент b множини B можна вибрати n способами, то впорядковану пару (a; b) можна вибрати $m \cdot n$ способами.

У справедливості правила добутку можна переконатись з таких міркувань. Нехай $A=\{a_1,a_2,...,a_m\}$ і $B=\{b_1,b_2,...,b_m\}$.

Тоді пари виду (a,b) можна записати у вигляді такої таблиці:

$(a_1; b_1), (a_1; b_2),...,(a_1; b_n)$,

$(a_2; b_1), (a_2; b_2),...,(a_2; b_n)$,

…………………………..

$(a_m; b_1), (a_m;b_2),...,(a_m; b_n)$.

Ця таблиця складається з m рядків, у кожному з яких міститься п елементів. Отже, загальне число пар дорівнює добутку $m \cdot n$.

Правило добутку можна узагальнити.

Узагальнене правило добутку. Нехай елемент a_1 з множини A_1, можна вибрати m_1 способами, елемент a_2 з множини A_2 - m_2 способами,..., елемент a_k з множини A_k можна вибрати m_k способами.

Тоді послідовний вибір елементів $(a_1,a_2,...,a_k;)$ можна здійснити $m_1 \cdot m_2 \cdot ... \cdot m_k$ способами.

Приклад 2. З Києва до Чернігова можна доїхати пароплавом, поїздом, автобусом, літаком, автомобілем. З Чернігова до Новгород-Сіверського - пароплавом і автобусом, автомобілем. Скількома способами можна здійснити подорож за маршрутом Київ - Чернігів - Новгород-Сіверський?

Розв'язання. З Києва до Чернігова можна доїхати 5-ма способами. З Чернігова до Н.-Сіверського – 3-ма способами.

Тоді за правилом добутку число способів вибору упорядкованої пари дорівнює добутку 5·3=15. Отже, з міста Києва до міста Н.-Сіверського через місто Чернігів можна вибрати 15 способів подорожі. Відповідь: 15 способів.

3. Основні формули комбінаторики: формула розміщень, формула перестановок, формула сполучень.

У комбінаторних задачах по різному підходять до поняття "рівні підмножини": в одних задачах підмножини, які відрізняються тільки порядком розташування в них елементів, треба вважати різними, а в інших порядок слідування елементів для розв'язування задачі значення не має.

В комбінаторних задачах розглядаються тільки скінченні множини.

Комбінаторні задачі поділяються на задачі на розміщення, на перестановки і на сполучення як без повторення, так і з повтореннями.

Термін «перестановка» (permutation) вжив у 1713 р. Я. Бернуллі в книзі "Мистецтво передбачень" (хоча епізодично він зустрічався і раніше).

Я.Бернуллі використовував і термін «розміщення» (arrangement).

Термін «сполучення» (combination) вперше зустрічається у Паскаля (1653, опублікований в 1665 році).

Розміщення

Розміщення – комбінації, які складаються з n різних елементів по k і відрізняються складом елементів і порядком їх слідування.

Приклад 3. Дана множина A ={1,2,3,4}. Подмножини B =(1,2,3), C=(1,3,2), D=(2,3,4) – це приклади розміщень з 4-х елементів по 3 елемента.

Число всіх розміщень з *n елементов по k элементов* позначають A_n^k («A з *n* по *k»)* і обчислюють за формулою:

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

де $n! = 1 \cdot 2 \cdot ... \cdot n$. Прийнято $0! = 1$.

Приклад 4. В турнірі з футболу беруть участь 8 команд. Скольки існує варіантів призової трійки?

Так як порядок команд в призовій трійці має значення (від цього залежить цінність медалей), то маємо розміщення. Треба визначити число всіх розміщень з $n = 8$ команд по $k = 3$ команди.

Тоді $A_n^k = A_8^3 = \frac{8!}{(8-3)!} = \frac{8!}{5!} = \frac{5! \cdot 6 \cdot 7 \cdot 8}{5!} = 6 \cdot 7 \cdot 8 = 336$ варіантів.

Приклад 5. Розклад одного дня містить 5 різних пар. Знайти кількість можливих розкладів, якщо вивчається 9 дисциплін.

Розв'язання. Маємо розміщення з 9 елементів по 5 без повторень, їх кількість дорівнює

$$A_9^5 = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 15120$$

Відповідь. 15120 розкладів.

У розміщеннях без повторень не має однакових елементів у вибірці: після вибору першого елемента для вибору другого елемента залишається на одиницю менше можливостей і т.д.

Перестановки

Розміщення з *n* елементів *по n* елементи називаються *перестановками з n елементів.* Число всех перестановок з *n* елементів позначають через P_n і обчислюють за формулою

$$P_n = n! = 1 \cdot 2 \cdot .... \cdot n.$$

Приклад 6. В спортивному випробуванні на 100 м беруть участь 5 спортсменів. Скільки існує варіантів протокола випробування?

В даному випадку маємо перестановки з $n = 5$ елементів. Тоді $P_n = P_5 = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$ варіантів.

Приклад 7. Скількома способами можна поставити на полиці 6 різних книг?

Число таких способів дорівнює числу перестановок з шести елементів, тобто

$P_6 = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720$ (способами).

Сполуки

Нехай маємо множину з n елементів. Довільна його підмножина з k елементів называється сполукою з n елементів по k елементи. Підмножини відрізняються один від одного лише складом елементів, порядок їх слідування значення немає.

Приклад 8. В прикладі 3 підмножини В і С — це приклад однієї сполуки, так як вони відрізняються лише порядком слідування елементів. Підмножини *B* і *D* — це приклад різних сполук.

Число всіх сполук з n елементів по k елементи позначають C_n^k («С з n по k ») і знаходять за формулою:

$$C_n^k = \frac{n!}{k!(n-k)!}, \text{ де } n! = 1 \cdot 2 \cdot ... \cdot n.$$

Приклад 9. В турнірі з футболу беруть участь 8 команд. Скільки всього матчей буде зіграно?

В даному випадку маємо число всіх сполук з $n = 8$ команд по $k = 2$ команд.

Тоді $C_n^k = C_8^2 = \frac{8!}{2!(8-2)!} = \frac{8!}{2 \cdot 6!} = \frac{6! \cdot 7 \cdot 8}{2 \cdot 6!} = \frac{7 \cdot 8}{2} = 28$ матчів.

Приклад 10. Скількома способами можна вибрати з 20 осіб делегацію в складі 4 осіб?

Різними вважатимемо ті делегації, які відрізняються хоча б однією особою. Кількість способів вибрати з 20 осіб делегацію в складі 4 осіб дорівнює числу комбінацій з 20 елементів по 4:

$$C_{20}^4 = \frac{A_{20}^4}{P_4} = \frac{20 \cdot 19 \cdot 18 \cdot 17}{1 \cdot 2 \cdot 3 \cdot 4} = 4845.$$

Приклад 11. Для проведення іспиту треба створити комісію з трьох викладачів. Скільки можна створити таких комісій з 7 викладачів?

Так як ми не можемо в комісію вибрати декілька разів одну й ту саму людину, то маємо комбінації без повторень елементів, порядок вибору людини (чи першою її вибрати, чи другою) значення не має. Треба вибрати з 7 викладачів 3 викладачі. Отже, маємо сполуки без повторень:

$$C_n^k = C_7^3 = \frac{7!}{3!(7-3)!} = \frac{7!}{6 \cdot 4!} = \frac{4! \cdot 7 \cdot 6 \cdot 5}{6 \cdot 4!} = 7 \cdot 5 = 35.$$

Наведемо схему вибору виду комбінації елементів

Чи всі елементи входять в комбінацію?

Так

Ні

Перестановки

$P_n = n!$

Чи має значення порядок?

Да

Ні

Розміщення

Сполуки

4. Розміщення, перестановки, сполучення з повтореннями

Розміщенням з повтореннями з n елементів по m елементів називається будь-який упорядкований *m*-елементний набір виду $(a_1, a_2, ..., a_к)$, де $a_1, a_2, ..., a_к$ - елементи множини М={ $a_1, a_2, ..., a_n$}, в якому хоч би один елемент повторювався.

Число всіх розміщень з повтореннями з n елементів по m елементів позначають символом $\overline{A}_n^m$. На відміну від розміщень без повторень, де $m \le n$, для розміщень з повтореннями може бути і $m>n$. Особливістю розміщень з повторенням є те, що після вибору першого (другого, третього і т.д.) елемента $a_1 \in M$, записавши його на першому місці набору, його повертають у множину *M,* тобто число елементів множини *M* залишається сталим.

Число розміщень з повтореннями з n елементів по m елементів обчислюють за формулою

$$\overline{A}_n^m = n^m \quad , \quad \text{де } m \text{ і } n \text{ - натуральні числа.}$$

<u>Приклад 12.</u> До шестицифрових номерів телефонів входять цифри від 0 до 9. Скільки абонентів може обслуговувати телефонна станція?

Розв'язання. Оскільки цифри в номерах можуть повторюватись, то кількість шестицифрових номерів телефонів дорівнює числу розміщень з 10 елементів по 6 з повторенням, тобто

$$\overline{A}_{10}^6 = 10^6 = 1000000.$$

Кожний упорядкований *n*–елементний набір з елементів множини $M=\{a_1,a_2,...,a_k\}$, - в якому елемент a_1 повторюється n_1 раз, елемент a_2 повторюється n_2 раз, і т.д., елемент a_k повторюється n_k раз, причому $n_1+n_2+\ldots+n_k=n$ називається *перестановкою з повтореннями з n елементів.*

Кількість усіх перестановок з повтореннями з n елементів, які відповідають означенню, позначають символом $P_n(n_1,n_2,...n_k)$, де $1 \le k \le n$.

Число $P_n(n_1,n_2,\ldots, n_k)$, всіх перестановок $n=n_1+n_2+\ldots+n_k$ з повтореннями з елементів $a_1,a_2,\ldots,a_k$, які повторюються відповідно $n_1,n_2,\ldots, n_k$ разів, *обчислюється за формулою*

$$P_n(n_1,n_2,\ldots,n_k) = \frac{n!}{n_1!\cdot n_2!\cdot\ldots n_k!}$$

<u>Приклад 13.</u> Абонент пам'ятає, що потрібний йому шестицифровий номер телефона починається з цифри 7 і містить три п'ятірки і дві дев'ятки.

Проте розташування цифр він не пам'ятає. Скільки спроб повинен зробити абонент, щоб набрати потрібний номер?

Маємо перестановки з повтореннями, всього спроб буде

$$P_5(3,2) = \frac{5!}{3!\,2!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 2\cdot 1} = 10.$$

Сполукою з повтореннями з n елементів по m елементів називається будь-який m-елементний набір виду $\{a_1, a_2, ..., a_m\}$, де кожний з елементів $a_1, a_2, ..., a_m$ належить до одного з n типів.

Позначають $\overline{C}_n^m$ і обчислюють за формулою

$$\overline{C}_n^m = \frac{(n+m-1)!}{(n-1)!\,m!}$$

<u>Приклад 14.</u> У поштовому відділенні продаються листівки 8 різних видів. Скількома способами можна купити в ньому 10 листівок?

Розв'язання. Оскільки порядок покупки листівок не має значення, а купити їх не можна всі різними (10>8), то маємо сполучення з повтореннями з 8 елементів по 10 елементів. Число їх дорівнює

$$\overline{C}_8^{10} = \frac{(8+10-1)!}{10!(8-1)!} = \frac{17!}{10!7!} = 19448.$$

При розв’язуванні комбінаторних задач доцільно користуватися таким *правилом*. Треба з’ясувати:

1. Чи має дана комбінація повторення;
2. Чи всі елементи містяться в комбінації;
3. Чи має значення порядок розміщення елементів.

ЛЕКЦІЯ 2. ВИПАДКОВІ ПОДІЇ.

План лекції

1. *Стохастичний експеримент. Випадкові події. Дії над подіями.*
2. *Статичне та класичне означення ймовірності. Властивості ймовірності.*

1. Стохастичний експеримент. Випадкові події. Дії над подіями

Поліпшення технології виробництва гральних кісток у XVIII столітті стимулювало розвиток теорії ймовірності. Виник дослід, який можна повторювати необмежену кількість разів в однакових умовах (пізніше названий "статистичним експериментом").

На практиці часто доводиться мати справу з тим чи іншим експериментом, точний результат проведення якого передбачити, взагалі кажучи, неможливо. Такий експеримент називають *випадковим*, або *стохастичним*.

Результат будь-якого проведення випадкового експерименту (випробування) називають *елементарною подією*, що відповідає цьому експерименту. Множину всіх елементарних подій, які відповідають даному експерименту, позначають Ω і називають *простором елементарних подій*.

Нехай випадковому експерименту відповідає простір елементарних подій Ω. Тоді випадковою подією (або просто подією) називають певну підмножину A множини Ω.

Кажуть, що подія A відбулась у даному випробуванні (проведенні експерименту), якщо результатом випробування є елементарна подія $E \in A$. При цьому елементарну подію E називають *сприятливою* для події A або кажуть, що E сприяє A.

Оскільки результатом будь-якого випробування є елементарна подія E, що належить до простору Ω елементарних подій, то у будь-якому випробуванні подія $A = \Omega$ обов’язково відбувається, і тому цю подію називають *вірогідною*.

Якщо у будь-якому випробуванні подія $B = \varnothing$ не відбувається, то її називають *неможливою.*

Нехай A і B – події, які є підмножинами простору Ω елементарних подій. Якщо $A \subset B$, тобто подія B відбувається завжди, коли відбувається A, то кажуть, що подія B є наслідком події A.

Якщо A і B такі, що кожна з них є наслідком іншої, тобто $A \subset B$ і $B \subset A$, то такі події називають рівними і позначають $A = B$.

Події А і В *сумісні* при даному випробуванні, якщо поява однієї не виключає можливості появи іншої.

Подія А - поява числа

Подія В - поява непарного числа

Події А і В несумісні, якщо поява однієї виключає можливість появи іншої.

Наприклад, кидання монети.

Подія А - поява герба.

Подія В - поява решки.

Події А і В залежні, якщо подія А залежить від події В.

Події А і В незалежні, якщо результат однієї не впливає на можливість появи іншої події.

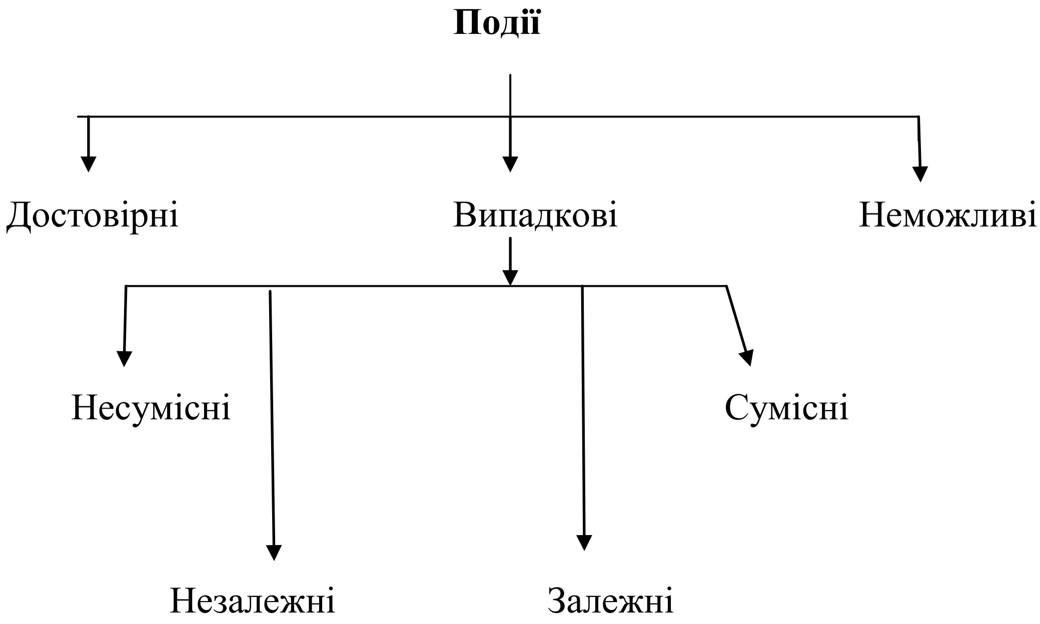
Наприклад, стрілянина двох стрільців.

Подія А - влучення в ціль першим стрільцем.

Подія В - влучення в ціль другого стрільцем.

Події А і В протилежні, якщо вони не сумісні і одна з них відбудеться обов'язково (Ă).

Цю класифікацію подій можна подати у вигляді схеми:



Нехай A і B є підмножинами даного простору елементарних подій Ω.

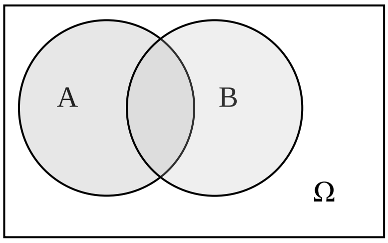


Рис.1. 1

Сумою подій A і B називають таку подію C, яка відбувається тоді й тільки тоді, коли відбувається принаймні одна з подій A або B.

Тоді суму подій A і B позначають $A \cup B$ або $A + B$. Геометрично суму подій A і B зображено на рис.1.1.

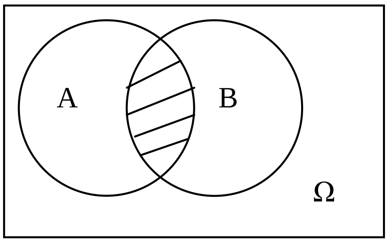


Рис.1. 2

Добутком подій A і B називають подію C, що відбувається тоді й тільки тоді, коли відбуваються обидві події A і B.

Тому добуток подій A і B позначають $A \cap B$ або $A \cdot B$. Геометрично добуток подій A і B зображено на рис.1.2.

Події A і B називають *несумісними*, якщо $A \cap B = \varnothing$, тобто ці події не можуть відбутися одночасно в жодному випробуванні.

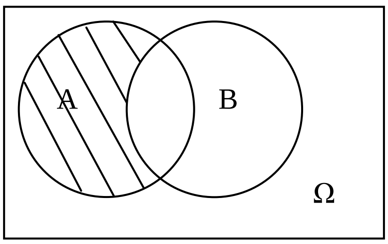


Рис.1. 3

Різницею подій A і B (A мінус B) називають таку подію C, яка відбувається тоді й тільки тоді, коли відбувається подія A і не відбувається подія B. Різницю A мінус B позначають $A \setminus B$ або $A - B$.

Геометрично різницю подій *A* і *B* зображено на рис.1.3.

2. Статичне та класичне означення ймовірності. Властивості ймовірності

Ймовірність є одним із основних понять теорії ймовірностей. Існує 7 означень ймовірностей. Розглянемо класичне і статистичне визначення.

Класичне визначення

Ймовірністю *P(A)* даної події *A* називається відношення числа результатів *m*, які сприяють появі даної події, до загального числа n рівно можливих і єдино можливих результатів випробувань:

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

<u>Приклад 1.</u> В урні десять куль. З них – шість білих і чотири чорних. Яка ймовірність вилучення білої кулі? Яка ймовірність вилучення чорної кулі?

Розв’язання:

Подія А - вилучення білої кулі.

Подія В - витяг чорної кулі.

$$P(A)=\frac{m}{n}=\frac{6}{10}=\frac{3}{5}=0{,}6,$$

$$P(B)=\frac{m}{n}=\frac{4}{10}=\frac{2}{5}=0{,}4.$$

<u>Приклад 2.</u> З 34 екзаменаційних білетів, занумерованих за допомогою цілих чисел від 1 до 34, навмання дістається 1. Яка ймовірність того, що номер витягнутого білета є число, що ділиться на 3.

Розв'язання. Простір елементарних подій складається з 34 елементів: N = 34. Подій, сприятливих до події *A*: $N(A)=[34:3]=11$.

Тоді: $P(A)=\frac{N(A)}{N}=\frac{11}{35}\approx 0.31$.

Приклад 3. Група спортсменів з 15 юнаків і 5 дівчат обирають команду в складі чотирьох чоловік. Яка ймовірність того, що в складі цієї команди не буде жодного юнака.

Розв'язання

A - в складі команди жодного юнака.

Всього		Вибрали
M=20	спортсменів	*m=4*
R=5	дівчат	*r=4*
M-R=15	юнаків	*m-r=0*

$$P(A)=\frac{C_5^4\cdot C_{15}^0}{C_{20}^4}=0.001.$$

Найпростіші властивості ймовірності:

1) Ймовірність вірогідної події дорівнює одиниці.

2) Ймовірність неможливої події дорівнює нулю $P(\varnothing)=0$.

3) $P(\overline{A})=1-P(A)$, або $P(A)+P(\overline{A})=1$;

4) Ймовірність випадкової події є додатнє число, обмежене нулем і одиницею: $0\le P(A)\le 1$;

5) $P(A+B)=P(A)+P(B)-P(AB)$ (теорема додавання).

При розв'язанні задач на підрахунок ймовірності можна користуватись таким *алгоритмом*:

1. З'ясувати за змістом задачі у чому конкретно полягає випробування.
2. З'ясувати, чи є елементарні події рівноможливими.
3. Підрахувати число *N* всіх можливих подій.
4. Підрахувати число *N(A)* всіх подій, сприятливих появі події *A*.
5. Обчислити ймовірність $P(A)=\frac{N(A)}{N}$.

Статистичне визначення ймовірності

Статистична ймовірність *P*(A)* - відносна частота появи події при випробуваннях апостеріорі (після досліду).

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

де *m* – число випробувань, при яких подія *A* відбулася;

n – число всіх випробувань.

Приклад 4. Студент потрапив 15 разів на мішень зі ста. Знайти ймовірність потрапляння в мішень.

Розв’язання.

m = 15, n = 100.

$$P^*(A) = \frac{m}{n} = \frac{15}{100} = 0{,}15.$$

ЛЕКЦІЯ 3. ВЛАСТИВОСТІ ЙМОВІРНОСТІ

План

1. Теорема про додавання ймовірностей для несумісних подій.

2.Умовні ймовірності та теорема множення ймовірностей.

3. Теорема додавання ймовірностей двох довільних подій.

1. Теореми додавання і множення ймовірностей

Теореми додавання ймовірностей безпосередньо випливають з найпростіших властивостей імовірності.

При введенні поняття умовна ймовірність розглянемо наступну задачу. Нехай двічі підкидається гральний кубик. Подія A полягає в тому, що сума очок при двох підкиданнях дорівнює 5. Знайти ймовірність цієї події. Припустимо, що при першому киданні випало «3» (подія B). Знайти ймовірність події A, якщо відбулася подія B.

Розв’язання. Простір елементарних подій Ω складається в цьому експерименті з 36 елементів: $\Omega = \{(m,n): m,n = 1,2,3,4,5,6\}$. Ці можливі результати випробування для зручності запишемо у вигляді таблиці (перше число – кількість очок, що випало при першому киданні, друге – при другому):

(1;1)	(1;2)	(1;3)	(1;4)	(1;5)	(1;6)
(2;1)	(2;2)	(2;3)	(2;4)	(2;5)	(2;6)
(3;1)	(3;2)	(3;3)	(3;4)	(3;5)	(3;6)
(4;1)	(4;2)	(4;3)	(4;4)	(4;5)	(4;6)
(5;1)	(5;2)	(5;3)	(5;4)	(5;5)	(5;6)
(6;1)	(6;2)	(6;3)	(6;4)	(6;5)	(6;6)

З таблиці видно, що $A = \{(4;1),(3;2),(2;3),(1;4)\}$, тому $P(A) = \dfrac{4}{36} = \dfrac{1}{9}$.

Розглянемо другу частину задачі. Нам стало відомо, що при першому киданні випала трійка (подія *B*). Тоді ймовірність події *A* стане меншою: 5 очок в сумі тепер може випасти лише тоді, коли при другому киданні випаде двійка, тобто в одному випадку з шести. Отже, ймовірність події *A* при умові, що настала подія *B*, дорівнює 1/6. Цю ймовірність позначають $P(A/B)$ або $P_B(A)$. Таким чином $P(A) = \dfrac{1}{9}$, $P_B(A) = \dfrac{1}{6}$.

Умовна ймовірність $P(A/B)$ або $P_B(A)$ - це ймовірність того, що відбудеться подія *A* за умови, що в цьому ж стохастичному експерименті відбулася подія *B*. Поняття умовної ймовірності є одним із основних інструментів теорії ймовірностей. Тому наведемо деякі міркування, які визначають формальне означення.

Розглянемо як виглядає ця характеристика в рамках «класичного означення» ймовірності – коли всі N наслідків стохастичного експерименту рівноможливі. Позначимо N_A, N_B, N_{AB} - число елементарних наслідків, які сприяють подіям *A*, *B* і *AB* відповідно. Для того, щоб, виходячи з цих даних, знайти формулу для умовної ймовірності $P(A/B)$, потрібно у відповідності до події *B* ввести новий імовірнісний простір Ω_1 – це ті і тільки ті елементарні наслідки із Ω, які сприяють події *B*. Всі вони рівноможливі і кількість їх N_B. Застосування класичного означення ймовірності дозволяє записати:

$P(A/B)=\frac{N_{AB}}{N_B}$. Якщо поділити чисельник і знаменник останнього дробу на число N, то прийдемо до формули $P(A/B)=\frac{N_{AB}/N}{N_B/N}=\frac{P(AB)}{P(B)}$.

Зрозуміло, наведені міркування мають смисл лише при $\boldsymbol{P}(\boldsymbol{B})>0$.

Повертаючись до загальних рамок аксіоматичного означення ймовірності, введемо умовну ймовірність, як аналог тільки що виведеної формули.

Нехай $\boldsymbol{P}(\boldsymbol{B})>0$. Умовною ймовірністю $P(A/B)$ події A за умови, що відбулася подія B (або просто: за умови B), називається відношення $P(A/B)=\frac{P(AB)}{P(B)}$.

Теорема про ймовірність добутку подій. Якщо $\boldsymbol{P}(\boldsymbol{B})>0$, то $P(AB)=P(A/B)P(B)$.

Легко бачити, що коли події A і B незалежні і $\boldsymbol{P}(\boldsymbol{B})>0$, то $P(A/B)=P(A)$. Отже, можна стверджувати, що призгаданих вище умовах означення незалежності еквівалентні між собою. Саме рівність типу $P(A/B)=P(A)$ на практиці кладеться в основу критерію незалежності подій A і B.

Теорема. Якщо події A і B незалежні, то незалежними будуть події A і $\overline{B}$ а також події $\overline{A}$ і $\overline{B}$.

Доведення.

Маємо $P(A)=P(AB\cup A\overline{B})=P(AB)+P(A\overline{B})$, звідки

$$P(A\overline{B})=P(A)-P(AB)=P(A)-P(A)P(B)=P(A)\cdot(1-P(B))=P(A)P(\overline{B}).$$

Друге твердження теореми є прямим наслідком першого твердження.

Із попарної незалежності подій не випливає їх незалежність в сукупності. Наступний приклад належить видатному вітчизняному математику С.Н. Бернштейну.

Аналіз і розв’язування задач доцільно проводити за *правилом*:

1. Визначити, в чому полягає розглянуте в задачі випробування.
2. Позначити буквами події (з умови задачі).
3. За допомогою введених позначень виразити подію, ймовірність появи якої необхідно знайти.
4. Якщо необхідно знайти ймовірність суми подій, з'ясувати сумісні чи несумісні розглянуті події. Якщо необхідно знайти ймовірність добутку події, з'ясувати залежні чи незалежні розглянуті події.
5. Вибрати відповідну умові задачі формулу і виконати необхідні обчислення.

Ймовірність суми подій

Чи відомо, що події несумісні?

- так:
  - для двох подій: $P(A+B)=P(A)+P(B)$.
  - для трьох подій: $P(A+B+C)=P(A)+P(B)+P(C)$.
  - для *n* подій: $P(A_1+A_2+..+A_n)=P(A_1)+P(A_2)+...+P(A_n)$.
- ні:
  - для двох подій: $P(A+B)=P(A)+P(B)-P(AB)$.
  - для трьох подій:

$$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC).$$

Ймовірність добутку подій

Чи відомо, що події незалежні?

- так:
  - для двох подій: $P(AB)=P(A)\cdot P(B)$.
  - для трьох подій: $P(A_1A_2A_3)=P(A_1)\cdot P(A_2)\cdot P(A_3)$.
  - для *n* подій: $P(A_1A_2...A_n)=P(A_1)\cdot P(A_2)\cdot ...\cdot P(A_n)$.
- ні:
  - для двох подій: $P(AB)=P(A)\cdot P(B/A)=P(B)\cdot P(A/B)$.
  - для трьох подій: $P(A_1A_2A_3)=P(A_1)\cdot P(A_2/A_1)\cdot P(A_3/A_1A_2)$.

- для n подій:

$$P(A_1 A_2 \ldots A_n) = P(A_1) \cdot P(A_2 / A_1) \cdot P(A_3 / A_1 A_2) \cdot \ldots \cdot P(A_n / A_1 A_2 \ldots A_{n-1}).$$

Приклад 1. Ймовірність того, що в жіночому магазині спочатку буде продана пара взуття сорок четвертого розміру дорівнює 0,12; сорок п'ятого 0,04; сорок шостого 0,01. Знайти ймовірність того, що спочатку буде продана пара взуття не менше 44 розміру.

Розв’язання:

Подія A - продана пара взуття сорок четвертого розміру

P (A) = 0,12.

Подія B - продана пара взуття сорок п'ятого розміру

P (B) = 0,04.

Подія C - продана пара взуття сорок шостого розміру

P (C) = 0,01.

Подія D - продана пара взуття не менше сорок четвертого розміру

P (D) = P (A) + P (B) + P (C) = 0,17.

Відповідь: ймовірність того, що спочатку буде продана пара взуття не менше 44 розміру дорівнює 0,17.

Приклад 2. Знайти ймовірність того, що спочатку продадуть пару взуття менш сорок четвертого розміру.

Розв’язання:

Подія - продана пара взуття менш сорок четвертого розміру

P (A) = 1 - P (D) = 1 - 0,17 = 0,83

Відповідь: ймовірність того, що спочатку продадуть пару взуття менш сорок четвертого розміру дорівнює 0,83.

Приклад 3. У цеху працюють 7 чоловіків і 3 жінки. За табельними номерами навмання вибирають трьох осіб. Знайти ймовірність того, що відібрані будуть всі чоловіки.

Розв’язання:

Подія A – відібрані будуть всі чоловіки, а події A_i – i-м відібраний чоловік ($i=1,2,3$). Тоді $A=$ A_1, A_2, A_3 і за формулою множення ймовірностей маємо:

$$P(A)=P(A_1A_2A_3)=P(A_1)\cdot P(A_2/A_1)\cdot P(A_3/A_1A_2)=\frac{7}{10}\cdot\frac{6}{9}\cdot\frac{5}{8}=\frac{7}{24}.$$

Приклад 4. В урні десять куль, з них чотири білих і шість чорних, почали витягувати кулі.

Розв’язання:

Подія A_1 - поява білої кулі при першому вийманні

P(A_1) = $\frac{4}{10}$.

Подія A_2- поява білої кулі при другому вийманні, з урахуванням повернення першої витягнутого кулі в урну.

P(A_2) = $\frac{3}{9}$.

Подія B_1- витягання чорної кулі при першому вийманні.

Подія B_2- витягання чорної кулі при другому вийманні.

$P_{B_1}(A_2)$ = $\frac{4}{9}$ - ймовірність витягання білої кулі з урахуванням раніше витягнутої чорної.

$P_{A_1}(B_2)$ = $\frac{6}{9}$- ймовірність витягання чорної кулі з урахуванням раніше витягнутої білої.

Приклад 5. Ймовірність влучення в мішень для першого стрільця дорівнює 0,7, а для другого дорівнює 0,8. Обидва стрілка роблять по одному пострілу і потім кожен з них стріляє ще раз, якщо при першому він промахнувся. Знайти ймовірність того, що в мішені буде дві пробоїни.

Розв’язання:

Подія A_1 - перший стрілок потрапляє в мішень з першого разу

P(A_1) = 0,7.

P($\overline{A}_1$) = 0,3.

Подія B_1 -другий стрілок потрапляє в мішень з першого разу

Р(B_1) = 0,8.

Р($\overline{B}_1$) = 0,2.

Подія A_2- перший стрілок потрапляє в мішень з другого разу.

Подія B_2 - другий стрілок потрапляє в мішень з другого разу.

Подія С - перший і другий стрілок потрапляють з першого разу.

С = $A_1\,B_1 + \overline{A}_1\,B_1\,A_2 + A_1\,\overline{B}_1\,B_2 + \overline{A}_1\,\overline{B}_1\,A_2\,B_2$.

Р(С) = Р($A_1\,B_1$) + Р($\overline{A}_1\,B_1\,A_2$) + Р($A_1\,\overline{B}_1\,B_2$) + Р($\overline{A}_1\,\overline{B}_1\,A_2\,B_2$) = 0,8736.

Відповідь: ймовірність того, що в мішені буде дві пробоїни дорівнює 0,8736.

Приклад 6. Ймовірність складання іспиту з математики дорівнює 0,5; з економіки дорівнює 0,7; з англійської мови дорівнює 0,6. Знайти ймовірність здачі двох, трьох або одного іспиту.

Розв'язання:

Подія *А* – студент здав три іспити,

подія A_1 – здав іспит з математики,

подія A_2 – з економіки,

подія A_3 – з англійської мови, тоді

$A = A_1 * A_2 * A_3$.

Подія *В* – студент здав два іспити

В = $A_1 * A_2 * \overline{A}_3 + A_1 * \overline{A}_2 * A_3 + \overline{A}_1 * A_2 * A_3$.

Подія С - один іспит

С = $A_1 * \overline{A}_2 * \overline{A}_3 + \overline{A}_1 * A_2 * \overline{A}_3 + \overline{A}_1 * \overline{A}_2 * A_3$.

Подія D – не здав жодного іспиту

D = $\overline{A}_1 * \overline{A}_2 * \overline{A}_3$.

А = 0,5 * 0,7 * 0,6 = 0,21.

В = 0,5 * 0,7 * 0,4 + 0,5 * 0,3 * 0,6 + 0,5 * 0,7 * 0,6 = 0,44.

С = 0,5 * 0,3 * 0,4 + 0,5 * 0,7 * 0,4 + 0,5 * 0,3 * 0,6 = 0,29.

D = 0,5 * 0,3 * 0,4 = 0,06.

Отже, ймовірність того, що студент здасть три іспити дорівнює 0,21; два іспити дорівнює 0,44; один іспит дорівнює 0,29, не здасть жодного іспиту дорівнює 0,06.

ЛЕКЦІЯ 4. ВЛАСТИВОСТІ ЙМОВІРНОСТІ

План

1.Формула повної ймовірності

2.Формула Байєса

1.Формула повної ймовірності

Бейєс або Байєс Томас (1702-1761) – англійський математик, член Лондонського Королівського товариства з 1742р. Народився в м.Лондоні. Був священником. У сучасній теорії ймовірностей і математичній статистиці важливу роль відіграють формули Байєса, які дають можливість емпірично оцінювати ймовірність подій (гіпотез).

Наслідком двох теорем ймовірностей – теореми додавання і теореми множення – є формули повної ймовірності і формула Байєса.

Теорема. Формула повної ймовірності. Якщо подія A настане за умови, що настане одна з подій H_1, H_2, …, H_n, які є попарно несумісними і в сумі дають вірогідну подію Ω ($P(H_1)+P(H_2)+\ldots+P(H_n)=1$), то має місце формула повної ймовірності:

$$P(A)=P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)+\ldots+P(H_n)\cdot P(A/H_n),$$

або записати так:

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

2. Формула Байєса

Нехай подія A настане за умови, що настане одна з подій (гіпотез) H_1, H_2, …, H_n, які є попарно несумісними і в сумі дають вірогідну подію Ω ($P(H_1)+P(H_2)+\ldots+P(H_n)=1$), то ймовірність гіпотез, одержаних після проведення дослідів, обчислюється за формулою:

$$P(H_i/A)=\frac{P(H_i)\cdot P(A/H_i)}{P(A)}, \text{ або } P(H_i/A)=\frac{P(H_i)\cdot P(A/H_i)}{\sum_{i=1}^{n} P(H_i)\cdot P(A/H_i)}. \qquad (*)$$

Доведення. За теоремою множення для двох подій $P(H_i A)=P(H_i)\cdot P(A/H_i)=P(A)\cdot P(H_i/A)$, звідки одержимо (*).

Значення формули Байєса можна тлумачити так, що при настанні події A, тобто по мірі отримання нової інформації, ми можемо перевіряти і корегувати висловлені до випробування гіпотези. Такий підхід називають байєсовським. Він дає можливість корегувати управлінські рішення в економіці, оцінці невідомих параметрів розподілу вивчаємих ознак у статистичному аналізі і т.п.

При обчисленні ймовірності події за формулою повної ймовірності і формулою Байєса можна користуватися таким *алгоритмом*:

1. Позначити подію, ймовірність настання якої необхідно знайти буквою A.
2. Скласти множину попарно-несумісних подій (гіпотез) $H_1, H_2, \ldots, H_n$. Перевірити, що об'єднання гіпотез співпадає з простором елементарних подій, що розглядаються.
3. Обчислити ймовірність кожної з гіпотез (їх сума повинна дорівнювати <u>одиниці</u>) і умовні ймовірності настання події A за кожною з умов, що відбулися події H_i, ($i=1, 2, \ldots, n$) (якщо вони не задані в умові задачі).
4. За формулою $P(A)=P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)+\ldots+P(H_n)\cdot P(A/H_n)$ обчислити ймовірність події A.
5. Якщо з умови задачі відомо, що подія A вже відбулась, то за формулою Байєса $P(H_i/A)=\dfrac{P(H_i)\cdot P(A/H_i)}{P(A)}$ необхідно обчислити ймовірність гіпотез, за умови що подія A відбулась.

Приклад 1. У першій урні сім білих і дев'ять чорних куль. У другій урні – шість білих і чотири чорних. З першої у другу переклали дві кулі, з другої витягли кулю, яка ймовірність того, що ця куля біла?

Розв’язання:

Подія A - вилучення білої кулі з першої урни.

Подія $H_1 - A_1 * A_2$ - дві білі в другу урну,

Подія $H_2 - A_1 * \overline{A}_2$ - білу і чорну,

Подія $H_3 - \overline{A}_1 * A_2$ - чорну і білу,

Подія $H_4 - \overline{A}_1 * \overline{A}_2$ - дві чорні кулі.

$$P(H_1) = P(A_1) * P_{A_1}(A_2) = \frac{7}{16} * \frac{6}{15} = \frac{42}{240}.$$

$$P(H_2) = \frac{7}{16} * \frac{9}{15} = \frac{63}{240}.$$

$$P(H_3) = \frac{9}{16} * \frac{7}{15} = \frac{63}{240}.$$

$$P(H_4) = \frac{9}{16} * \frac{8}{15} = \frac{72}{240}.$$

$$\sum P = 1,0.$$

$$P_{H_1}(A) = \frac{8}{12}.$$

$$P_{H_2}(A) = \frac{7}{12}.$$

$$P_{H_3}(A) = \frac{7}{12}$$

$$P_{H_4}(A) = \frac{6}{12}.$$

$$P(A) = \sum_{i=1}^{4} P(H_i) * P_{H_i}(A) = \frac{42}{240} * \frac{8}{12} + \frac{63}{240} * \frac{7}{12} + \frac{63}{240} * \frac{7}{12} + \frac{72}{240} * \frac{6}{12} = 0,5729 -$$

ймовірність того, що ця куля біла.

Приклад 2. У кожній з трьох груп по 25 студентів. Кількість студентів групи, які склали іспит з математики дорівнює відповідно 22, 20, 18. Яка ймовірність того, що випадково обраний студент склав іспит з математики? Розв’язання:

Подія A - студент склав іспит.

Подія H_1 - студент з першої групи.

Подія H_2 - студент з другої групи.

Подія H_3 - студент з третьої групи.

$P_{H_1}(A) = \frac{22}{25}$.

$P_{H_2}(A) = \frac{20}{25}$.

$P_{H_3}(A) = \frac{18}{25}$.

Р(H_1) = $\frac{1}{3}$.

Р(H_2) = $\frac{1}{3}$.

Р(H_3) = $\frac{1}{3}$.

$P(A) = \sum_{i=1}^{3} P(H_1) * P_{H_1}(A) = \frac{22}{25} * \frac{1}{3} + \frac{20}{25} * \frac{1}{3} + \frac{18}{25} * \frac{1}{3} = \frac{4}{5}$ – ймовірність того, що студент склав іспит.

Приклад 3. У торгівельну фірму надійшли телевізори від трьох постачальників у відношенні 1:4:5. Практика показала, що телевізори, які надійшли від 1-го, 2-го і 3-го постачальників, не потребують ремонту протягом гарантійного терміну відповідно у 98, 88 і 92% випадків. Знайти ймовірність того, що: а) телевізор, який надійшов у торгівельну фірму не потребує ремонту протягом гарантійного терміну; б) телевізор, який було реалізовано, потребував ремонту протягом гарантійного терміну, в) визначити, від якого постачальника ймовірніше за все надійшов телевізор, який потребує ремонту.

Розв'язування.

а) Подія *А* – телевізор не потребує ремонту протягом гарантійного терміну.

Подія H_i – телевізор надійшов від *і*-го постачальника (*і=1, 2, 3*),

подія A/H_i – телевізор не потребує ремонту, якщо він від *і*-го постачальника.

За умовою:

$$\begin{aligned}
&P(H_1)=\frac{1}{1+4+5}=0{,}1 \qquad && P(A/H_1)=0{,}98,\\
&P(H_2)=\frac{4}{1+4+5}=0{,}4 \qquad && P(A/H_2)=0{,}88,\\
&P(H_3)=\frac{5}{1+4+5}=0{,}5 \qquad && P(A/H_3)=0{,}92.\\
&\overline{\qquad\qquad\qquad}=1 &&
\end{aligned}$$

За формулою Байєса (*)

$$P(A)=0{,}1\cdot 0{,}98+0{,}4\cdot 0{,}88+0{,}5\cdot 0{,}92=0{,}91\,.$$

б) Подія $\overline{A}$ – телевізор потребує ремонту протягом гарантійного терміну. Ймовірність даної події можна знайти таким чином:

$$\begin{aligned}
&P(H_1)=0{,}1 \qquad && P\left(\overline{A/H_1}\right)=1-0{,}98=0{,}02.\\
&P(H_2)=0{,}4 \qquad && P\left(\overline{A/H_2}\right)=1-0{,}88=0{,}12.\\
&P(H_3)=0{,}5 \qquad && P\left(\overline{A/H_3}\right)=1-0{,}92=0{,}08.\\
&\overline{\qquad\quad}=1 &&
\end{aligned}$$

$$P\left(\overline{A}\right)=0{,}1\cdot 0{,}02+0{,}4\cdot 0{,}12+0{,}5\cdot 0{,}08=0{,}002+0{,}048+0{,}04=0{,}09.$$

в)

$$P\left(\overline{A}\right)=1-P(A)=1-0{,}91=0{,}09.$$

За формулою Байєса (*)

$$P\left(H_1/\overline{A}\right)=\frac{P(H_1)\cdot P\left(\overline{A/H_1}\right)}{P\left(\overline{A}\right)}=\frac{0{,}1\cdot 0{,}02}{0{,}09}=0{,}022.$$

$$P\left(H_2/\overline{A}\right)=\frac{P(H_2)\cdot P\left(\overline{A/H_2}\right)}{P\left(\overline{A}\right)}=\frac{0{,}4\cdot 0{,}12}{0{,}09}=0{,}533.$$

$$P\left(H_3/\overline{A}\right)=\frac{P(H_3)\cdot P\left(\overline{A/H_3}\right)}{P\left(\overline{A}\right)}=\frac{0{,}5\cdot 0{,}08}{0{,}09}=0{,}444.$$

Отже, ймовірність найбільша у другого постачальника.

ЛЕКЦІЯ 5. ВИПАДКОВІ ВЕЛИЧИНИ

План

1.Поняття випадкової величини.

2. Закон розподілу дискретної випадкової величини.

3. Основні характеристики дискретної випадкової величини: математичне сподівання, дисперсія.

1.Поняття випадкової величини

У другій половині XIX століття в роботах Максвелла, Больцмана і Гіббса набула бурхливого розвитку статистична механіка, яка описувала стан розряджених систем, що містять величезне число частинок (порядку числа Авогадро). Якщо раніше поняття розподілу випадкової величини було переважно пов'язане з розподілом помилок вимірювання, то тепер розподіленими виявилися самі різні величини - швидкості, енергії, довжини вільного пробігу.

Випадковою величиною називають числову функцію $X = X(E)$, визначену на просторі елементарних подій Ω, який відповідає певному випробуванню. Надалі випадкові величини будемо позначати великими латинськими буквами – X, Y, Z, …. , а їх можливі значення – відповідними малими буквами – x, y, z, … .

Наприклад:

Випробування полягає в киданні грального кубика. Простір елементарних подій є $\Omega=\{E_1,...,E_6\}$ – випадання грані з i очками). Нехай $X(E_i)=i$, тобто кожному елементарному результату одного кидання кубика поставимо у відповідність число очок, що випало при цьому киданні. Можливі значення цієї випадкової величини – числа 1, 2, 3, 4, 5, 6.

Наприклад:

Випробування полягає у тому, що стрілець робить один постріл у круглу мішень. Простір елементарних подій – це множина всіх точок площини мішені. Нехай X(E) – відстань від центра мішені до точки E. Можливі значення цієї випадкової величини – всі невід’ємні числа, що не перевищують r, де r – радіус мішені. Очевидно вони складають нескінченну незчисленну множину.

Випадкові величини бувають дискретними і неперервними.

Випадкова величина називається *дискретною*, якщо вона приймає окремі, ізольовані одне від одного значення.

Множина значень дискретної випадкової величини може бути як скінченною так і нескінченною. В останньому випадку множина має бути такою, щоб її значення можна було пронумерувати натуральними числами 1, 2, 3, …. (зчисленна множина). Якщо ця множина скінчена: $X(\Omega)=\{x_1, x_2, ..., x_m\}$, $m \in N$, де числа x_k попарно різні, то *випадкову величину* називають *простою*.

Наприклад:

1) Кількість влучень в мішень при десяти незалежних пострілах.
2) Кількість нестандартних виробів, що опинились у партії з 100 виробів.
3) Кількість очок, що випали на верхній грані при підкиданні грального кубика.
4) Довжина слова в технічному (художньому) тексті.
5) Сума очок на випали гранях двох кубиків.
6) Кількість годин, пропущених студентами за семестр.

Дискретна випадкова величина чи неперервна залежить від того, розривна чи неперервна функція $F(x)$ розподілу ймовірностей випадкової величини *X*. Функцією розподілу ймовірностей випадкової величини X називається функція $F(x)$, яка для довільного числа x задає ймовірність того, що випадкова величина X набуває значення, менше x, тобто:

$$F(x)=P_x((-\infty;x))=P(\{E : X(E)<x\})=F\left(X^{-1}(-\infty;x)\right).$$

Щільністю розподілу ймовірностей на множині значень випадкової величини X називається похідна (за умови її існування) від функції розподілу ймовірностей на тій же множині:

$$f_x(x)=F'_x(x).$$

Математичним сподіванням простої випадкової величини X називається сума добутків всіх її значень на відповідні їм ймовірності:

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

$$\bar{x}=\frac{x_1 m_1 + x_2 m_2 + \ldots + x_k m_k}{m_1 + m_2 + \ldots + m_k}, \text{ або } \bar{x}=\sum_{i=1}^{m} x_i \frac{m_i}{n}.$$

Позначимо дріб $\frac{m_i}{n}=p_i$ і запишемо тепер середнє арифметичне таким чином:

$$\bar{x}=\sum_{i=1}^{m} x_i p_i .$$

Тобто математичне сподівання дорівнює середньому арифметичному спостережних значень випадкової величини.

Математичне сподівання не може в достатній мірі характеризувати дискретну випадкову величину. Дискретні значення ймовірностей можуть бути сильно відмінними від значення математичного сподівання. Тому розглядають *відхилення (розсіювання) значень ймовірностей відносно статистичного математичного сподівання*. У якості такої характеристики розглядають *дисперсію*. Слово дисперсія означає „розсіювання". Дисперсійний аналіз набув інтенсивного розвитку на початку XX століття в Англії при розв'язанні задач сільського господарства.

Для характеристики розсіювання використовують квадрат відхилення значень випадкової величини від її математичного сподівання (оскільки за властивостями математичного сподівання саме відхилення дорівнює нулю).

Дисперсія випадкової величини дорівнює різниці між статистичним математичним сподіванням квадрату простої випадкової величини та квадратом статистичного математичного сподівання, тобто $D_n^*(X)=M_n^*(X^2)-\left[M_n^*(X)\right]^2$,

де $M_n^*(X)=x_{ц.м.}=\dfrac{x_1 p_1 + x_2 p_2 + \ldots + x_m p_m}{p_1 + p_2 + \ldots + p_m}$.

Розглянемо будь-яку операцію, прибуток від якої буде простою випадковою величиною *X*, тоді середній очікуваний прибуток - це статистичне математичне сподівання. А статистичне середнє квадратичне відхилення $\sigma_n^*(X)$ - це міра розсіювання можливих значень прибутку, який вважають мірою ризику.

Приклад 1. Дискретна випадкова величина X – температура повітря протягом березня задана таблицею розподілу (x_i – значення температури, p_i – відповідні ймовірності):

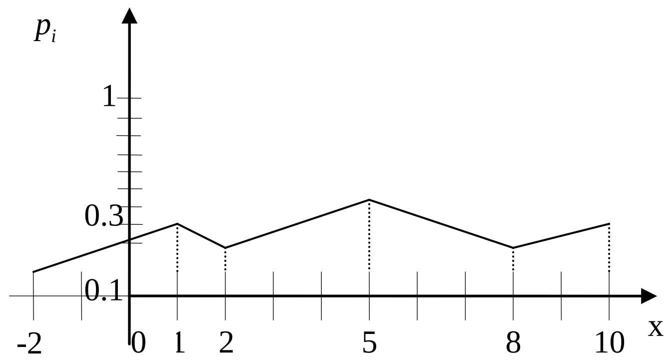
$X=x_i$	-2	1	2	5	8	10
p_i	0.1	0.2	0.1	0.3	a	0.2

Знайти *a,* побудувати полігон розподілу, функцію розподілу $\boldsymbol{F_n^*(x)}$ і побудувати її графік. Обчислити ймовірність події: а) $P(x \le 1)$; б) $P(1 \le x \le 5)$. Знайти $\boldsymbol{M_n^*(X)}$, $\boldsymbol{D_n^*(X)}$, $\boldsymbol{\sigma_n^*(X)}$ і Mo.

Розв'язання.

$a = 1 - (0.1 + 0.2 + 0.1 + 0.3 + 0.2) = 0.1$.

На рис. зображено полігон розподілу.



Для дискретної випадкової величини функція розподілу $\boldsymbol{F(x) = P(X < x) = \sum_{x_i \prec x} P(X = x_i)}$.

Якщо $x \le -2$, то $F(x) = 0$;

Якщо $-2 < \boldsymbol{x} \le 1$,то $F(x) = 0.1$;

Якщо $1 < \boldsymbol{x} \le 2$, то $F(x) = 0.3$;

Якщо $2 < \boldsymbol{x} \le 5$, то $F(x) = 0.4$;

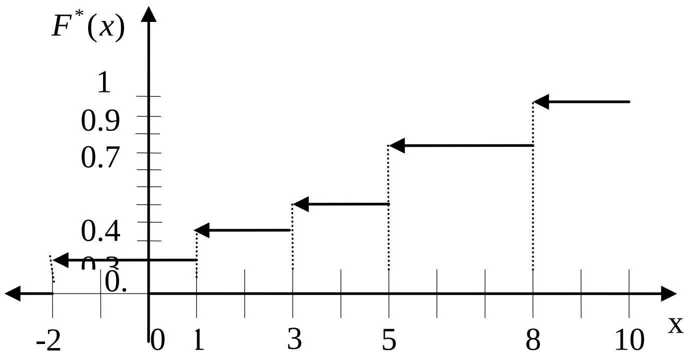
Якщо $5 < \boldsymbol{x} \le 8$, то $F(x) = 0.7$;

Якщо $8 < \boldsymbol{x} \le 10$, то $F(x) = 0.8$;

Якщо $x > 10$, то $F(x) = 1$.

Тоді, $F(x) = \begin{cases} 0, & x \le -2, \\ 0.1, & -2 < x \le 1, \\ 0.3, & 1 < x \le 2, \\ 0.4, & 2 < x \le 5, \\ 0.7, & 5 < x \le 8, \\ 0.8, & 8 < x \le 10, \\ 1, & x > 10. \end{cases}$

Графік функції має такій вигляд



Ймовірність події $x \le 1$ дорівнює $P(x \le 1) = 0.1$ (по функції розподілу ймовірностей).

Ймовірність події $1 \le x \le 5$ дорівнює $P(1 \le x \le 5) = F(5) - F(1) = 0.4 - 0.1 = 0.3$

$Mo = 5$ – найімовірніша температура (мода розподілу дискретної випадкової величини).

$M(X) = -2 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot 0.1 + 5 \cdot 0.3 + 8 \cdot 0.1 + 10 \cdot 0.2 = 4.5$ – середня температура у березні (математичне сподівання).

$M(X^2) = 4 \cdot 0.1 + 1 \cdot 0.2 + 4 \cdot 0.1 + 25 \cdot 0.3 + 64 \cdot 0.1 + 100 \cdot 0.2 = 34.9$.

Дисперсія $D(X) = 34.9 - 20.25 = 14.65$.

Статистичне середнє квадратичне відхилення $\sigma(X) = \sqrt{D(X)} = 3.83$.

ЛЕКЦІЯ 6. НЕПЕРЕРВНА ВИПАДКОВА ВЕЛИЧИНА

План

1.Інтегральна функція (функція розподілу) неперервної випадкової величини та її властивості.

2. Диференціальна функція (щільність розподілу ймовірностей)

3. Основні характеристики неперервної випадкової величини: математичне сподівання, дисперсія.

4. Два важливі розподіли випадкових величин (біноміальний розподіл, нормальний розподіл).

1. Інтегральна функція (функція розподілу) неперервної випадкової величини та її властивості

Неперервною називають *випадкову величину*, що може набувати всіх значень з деякого скінченого або нескінченного проміжку.

Приклади неперервних випадкових величин:

1. температура повітря,
2. тиск,
3. сила струму,
4. зріст людини,
5. вік людини,
6. довжина стопи людини,
7. вага немовлят,
8. швидкість руху автотранспорту,
9. інтервал часу між проїжджаючими вантажними автомобілями.

Неперервні випадкові величини задати таблицею неможливо, бо їх значення заповнюють цілий проміжок. Тому треба задавати не ймовірності окремих значень, а ймовірності того, що значення випадкової величини потрапить в певний інтервал.

Інтегральною функцією розподілу $F(x)$ ймовірностей випадкової величини називається ймовірність того, що випадкова величина X у результаті випробування набуде значення меншого за *x,* де *x* – довільне дійсне число:

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

Звідси: випадкова величина *X* називається неперервною**,** якщо її функція розподілу (інтегральна функція) *F(x)* є неперервною кусково - диференційовною.

Основні властивості інтегральної функції

1. Інтегральна функції задовільняє умови $0 \le F(x) \le 1$, що випливає з означення *F(x)* як ймовірності.

2. Інтегральна функція $F=F(x)$ - неспадна, тобто $F(x_2) \ge F(x_1)$ при $x_2 \ge x_1$.

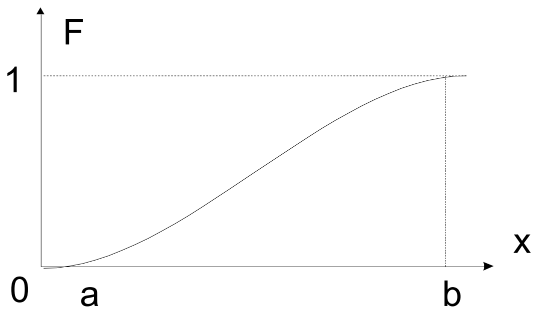
Наслідок 2.1. Імовірність того, що випадкова величина *X* набуде значення x *є* $(a;\beta)$ дорівнює приросту функції $F=F(x)$ на цьому інтервалі, тобто

$$P(a \le x \le b)=F(b)–F(a)$$

Наслідок 2.2. Ймовірність того, що неперервна величина *X* набуде одного значення дорівнює нулю $P(X=x_1)=0.$

Наслідок 2.3. Якщо всі можливі значення випадкової величини *X* належать інтервалу *(a,b),* то $F(x)=0$, при $x \le a$, $F(x)=1$, при $x \ge b$.

Графік інтегральної функції розподілу



2. Щільність розподілу ймовірностей (диференціальна функція) та її властивості

Диференціальною функцією розподілу *y=f(x)* (або щільністю розподілу ймовірностей) неперервної випадкової величини називається перша похідна від її інтегральної функції *F(x):*

f(x)=F'(x).

Отже, інтегральна функція *F(x)* є первісною для диференціальної функції *y=f(x)*.

Теорема 1. Ймовірність того, що неперервна випадкова величина Х набуде значення з інтервалу (α; β*)* дорівнює визначеному інтегралу від диференціальної функції (щільності розподілу) в межах від *a* до *b*:

$$P(a\langle X\langle b)=\int_a^b f(x)dx,$$

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

Доведення. Оскільки визначений інтеграл з геометричної точки зору дорівнює площі криволінійної трапеції, обмеженої неперервною кривою *y=f(x),* віссю *ox* і прямими *x=a* і *x=b,* а *f(x)=F'(x),* то маємо

$\int_a^b f(x)dx=\int_a^b F'(x)dx=F(x)\Big|_a^b=F(b)-F(a)=P(a\langle X\langle b)$. Теорема доведена.

Інтегральна функція *F(x)* визначаться з рівності $F(x)=\int_{-\infty}^{+\infty} f(x)dx$.

Для дискретної випадкової величини поняття щільності розподілу не розглядається.

Властивості щільності розподілу (диференціальної функції)

1. Диференціальна функція невід'ємна f(x) ≥ 0, бо інтегральна функція *F(x)* неспадна.

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

3. Якщо $x\in(a;b), \quad то \quad P(a<X<b)=\int_a^b f(x)dx=1$

4 . Якщо x ∈ *(a, b)*, то *f(x) = 0* при *x<a* (*при x>b*).

5 . $f(x)=\lim\limits_{\Delta x\to 0}\frac{P(x\le X\le x+\Delta x)}{\Delta x}$.

3. Основні характеристики неперервної випадкової величини: математичне сподівання, дисперсія.

Математичним сподіванням неперервної випадкової величини X, можливі значення якої x ∈ *(a;b),* називають число, яке дорівнює визначеному інтегралу від добутку *x* на диференціальну функцію *f(x)*, тобто

$$M(X)=\int_a^b xf(x)dx.$$

Дисперсією неперервної випадкової величини X, можливі значення якої x ∈ *(a;b),* називають математичне сподівання квадрату її відхилення від математичного сподівання, тобто

$$D(X)=\int_a^b [x-M(X)]^2 f(x)dx=\int_a^b x^2 f(x)dx-M^2(X).$$

Зауваження. Якщо $X \in$ R, то

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

Середнім квадратичним відхиленням неперервної випадкової величини називають корінь квадратний з дисперсії, тобто $\sigma(X)=\sqrt{D(X)}$.

Число M_e називається *медіаною*, якщо виконується рівність

$$\int_{-\infty}^{+\infty} f(x)dx=\int_{M_1}^{+\infty} f(x)dx=\frac{1}{2} \quad \text{або} \quad P(X\langle M_e)=P(X\rangle M_e)=\frac{1}{2}.$$

<u>Приклад 1.</u> Випадкова величина *X* задана законом розподілу:

$$F(x) = \begin{cases} \mathrm{a}\cdot\sin \mathrm{x}, \text{ при } 0<\mathrm{x}<\pi, \\ 0, \text{ при } \mathrm{x}<0, \\ 0, \text{ при } \mathrm{x}>\pi. \end{cases}$$

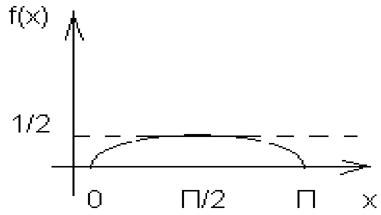
Знайти коефіцієнт *a*, побудувати графік цієї функції.

Розв’язання: $\int_{-\infty}^{\infty} a \sin xdx =1.$

$$\int_0^{\Pi} a \sin x dx = a(-\cos x)\Big|_0^{\pi} = a[1+1]=2a=1.$$

$a=\frac{1}{2}$.

Графік диференціальної функції розподілу



Приклад 2. **З**найти *f(x)* і ймовірність влучення випадкової величини *X* в інтервал $1<X<2.5$. Побудувати графік *f(x)*.

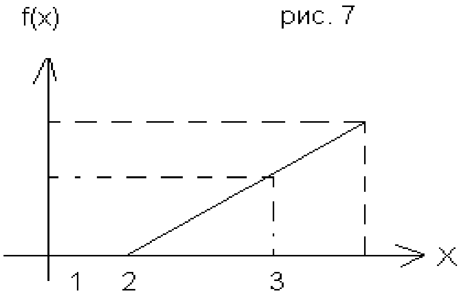
$$F(x) = \begin{cases} 0, \text{ при } x \le 2 \\ (x-2)^2, \text{ при } 2 < x \le 3 \\ 1, \text{ при } x > 3 \end{cases}$$

Розв'язання:

$$f(x)= F^{/}(x) = \begin{cases} 0, \text{ при } x \le 2 \\ 2(x-2), \text{ при } 2 < x \le 3 \\ 0, \text{ при } x > 3 \end{cases}$$

$$P(1<X<2.5) = F(b) - F(a) = F(2.5) - F(1) = (0.5)^2 - 0 = 0.25.$$

Графік диференціальної функції розподілу



4. Два важливі розподіли випадкових величин (біноміальний розподіл, нормальний розподіл).

Біноміальний закон розподілу

Закон розподілу випадкової величини *X*:

$$P_n(X = k) = C_n^k p^k q^{n-k}, \text{ де } q = 1 - p,\ k = 0,1,2,...,n.$$

Запишемо у вигляді таблиці:

X	0	1	...	k	...	n
p	q^n	npq^{n-1}	...	$C_n^k p^k q^{n-k}$	...	p^n

Закон визначається двома параметрами: *n*, *p*. Використовується у теорії і практиці статистичного контролю якості продукції, теорії стрільби, демографії, у фізичному експерименті при реєстрації нейтронів з повною енергією.

Розглянемо два способи знаходження числових характеристик для біноміального розподілу.

1 спосіб. Розглядають у якості випадкової величини *X* число появ події *A* в *n* незалежних випробуваннях. Зрозуміло, загальне число *X* появ подій в окремих дослідженнях. Тому якщо X_1 – число появ події у першому випробуванні, X_2 – у другому, ..., X_n – в *n*-му, то загальне число появ подій $X = X_1 + X_2 + ... + X_n$. Величини X_1, X_2, ..., X_n взаємно незалежні, оскільки

результат кожного випробування не залежить від результатів інших. За властивістю математичного сподівання

$$M(X)=M(X_1)+M(X_2)+...+M(X_n).$$

Кожен з доданків правої частини рівності є математичне сподівання числа появ подій в одному випробуванні: $M(X_1)$ - у першому, $M(X_2)$ – у другому тощо. Оскільки математичне сподівання числа появ подій в одному випробуванні дорівнює ймовірності події, то $M(X_1)=M(X_2)=M(X_n)=p$. Підставляємо у праву частину рівності, замість кожного доданку, одержимо $M(X)=np$.

Аналогічно обчислимо дисперсію, використовуючи її властивості:

$$D(X)=D(X_1)+D(X_2)+...+D(X_n).$$

Оскільки, X_1 – число появ події A у першому випробуванні, то $M(X_1)=p$. Знайдемо математичне сподівання величини X_1^2, яке може приймати тільки два значення, а саме: 1^2 з імовірністю p і 0^2 з імовірністю q: $M(X_1^2)=1^2\cdot p+0^2\cdot q=p$. Підставляючи знайдені результати у співвідношення $D(X_1)=M(X_1^2)-[M(X_1)]^2$, одержимо $D(X_1)=p-p^2=p(1-p)=pq$. Зрозуміло, дисперсія кожної з інших випадкових величин теж дорівнює pq. Замінивши кожен доданок правої частини через pq, отримаємо $D(X)=npq$.

<u>2 спосіб.</u> За означенням математичного сподівання для дискретної випадкової величини маємо:

$$M(X)=\sum_{i=0}^{n}x_i p_i=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}$$

Запишемо біном Ньютона:

$$(p+q)^n=\sum_{k=0}^{n}C_n^k p^k q^{n-k}$$

і диференціюємо його по p, у результаті одержимо:

$$n(p+q)^{n-1}=\sum_{k=0}^{n}kC_n^k p^{k-1}q^{n-k}.$$

Помножимо одержану рівність на p, маємо:

$$np(p+q)^{n-1}=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}. \qquad (*****)$$

Порівняємо останню рівність з початковою, звідки одержимо, що $np(p+q)^{n-1} = M(X)$, звідки, оскільки $p+q=1$, випливає

$$M(X) = np.$$

Диференціюємо рівність (*****) двічі по p, знаходимо

$$n(n-1)(p+q)^{n-2} = \sum_{k=0}^{n} k(k-1)C_n^k p^{k-2} q^{n-k}.$$

Помножимо останню частину на p^2, а потім перетворимо праву частину. Маємо $n(n-1)p^2(p+q)^{n-2} = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} - \sum_{k=0}^{n} kC_n^k p^k q^{n-k}$, звідки при $p+q=1$ випливає, що $n^2p^2 - np^2 = M(X^2) - M(X) = M(X^2) - np$. Тоді $M(X^2) = n^2p^2 - np^2 + np$, а значить $D(X) = M(X^2) - [M(X)]^2 = n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p) = npq$.

Нормальний закон розподілу

Неперервна випадкова величина x розподілена за нормальним законом (розподіл Гауса), якщо її щільність ймовірностей

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in R.$$

Функція розподілу ймовірностей $F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$

Обчислюючи невласні інтеграли $\int_{-\infty}^{+\infty} xf(x)dx$ і $\int_{-\infty}^{+\infty} (x-a)^2 f(x)dx$ матимемо

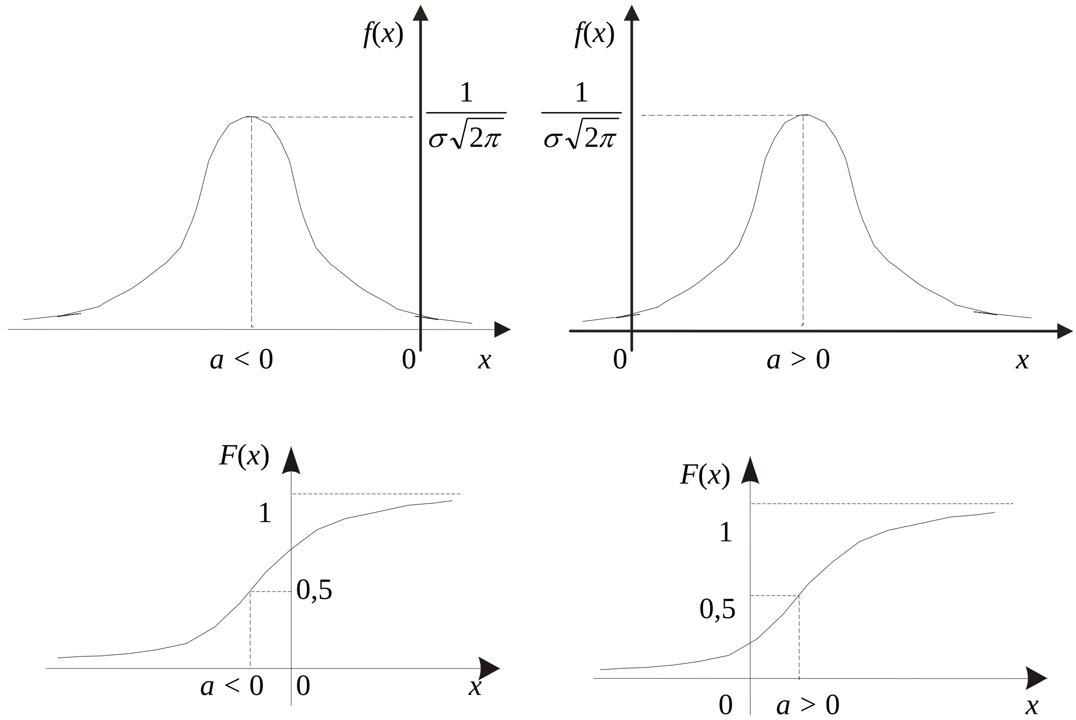
$$M(x) = a, \qquad D(x) = \sigma^2.$$

Нормальний закон визначається двома параметрами:

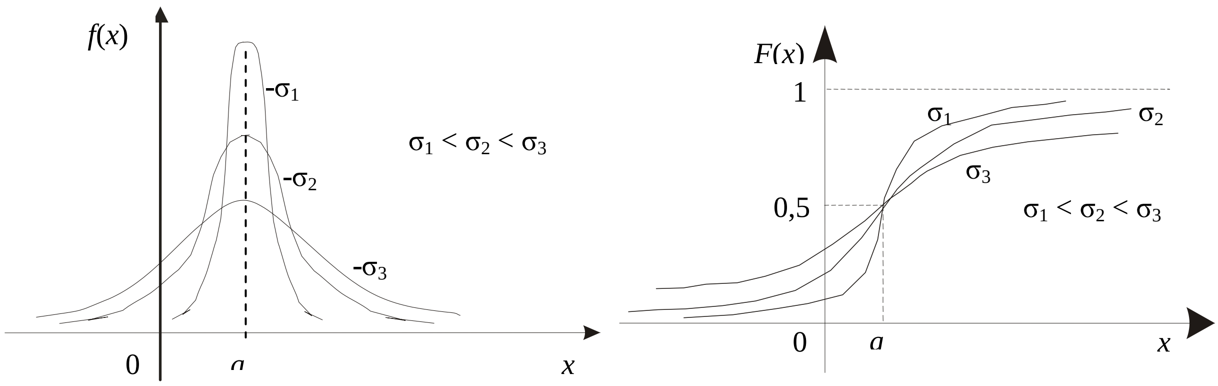
a – математичним сподіванням;

σ – середнім квадратичним відхиленням.

Вплив параметра a на графіки функцій $f(x)$, $F(x)$:



Вплив параметра σ на графіки функцій $f(x)$, $F(x)$:



Якщо $a = 0$, $\sigma = 1$, то нормальний закон розподілу ймовірностей називається нормованим.

Щільність ймовірностей і функція розподілу для нормованого нормального розподілу

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in R$$

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

Формули для обчислення ймовірностей випадкових подій для нормального закону

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

де $\Phi(x)$ – функція Лапласа, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{t} e^{-\frac{t^2}{2}} dt$.

Формули для обчислення ймовірностей випадкових подій для нормованого закону:

$$P(\alpha < x < \beta) = \Phi(\beta) - \Phi(\alpha),$$
$$P(|x| < \delta) = 2\Phi(\delta).$$

Правило трьох сигм.

$$P(|x - a| < 3\sigma) \approx 2\Phi(3) = 0{,}9973.$$

Отже, якщо випадкова величина X має нормальний розподіл, то ймовірність абсолютної величини її відхилення від математичною сподівання, яка менше трьох середніх квадратичних відхилень, близька до одиниці. Тобто, практично вірогідно, що розподілена за нормальним законом випадкова величина x набуде значень з інтервалу (a - 3σ ; a + 3σ).

$$\xrightarrow[\alpha - 3\sigma \qquad a \qquad a + 3\sigma]{} \alpha$$

Приклад 3. Випадкова величина x розподілена за нормальним законом розподілу з числовими характеристиками $M(x) = 3$, $D(x) = 4$. Знайти $f(x)$.

Розв’язання:

$M(x)$ = a=3, $\sigma=\sqrt{D(x)}=2$. Підставимо знайдені значення a і σ в диференціальну функцію нормального розподілу $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$, $x\in R$.

Маємо: $f(x)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-3)^2}{8}}$, $x\in R$.

Приклад 4. Термін роботи приладу є випадкова величина, що відповідає нормальному закону розподілу з гарантією на 15 років і середнім квадратичним відхиленням, що дорівнює 3 роки. Визначити ймовірність того, що прилад працюватиме від 10 до 20 років.

Розв’язання:

$a=15,\ \sigma=3,\ \alpha=10,\ \beta=20.$ Використаємо формулу

$$P(\alpha<x<\beta)=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right),$$

$$P(10<x<20)=\Phi\left(\frac{20-15}{3}\right)-\Phi\left(\frac{10-15}{3}\right)=2\Phi\left(\frac{5}{3}\right)\approx 2\Phi(1{,}67)=2\cdot 0{,}4525=0{,}905.$$

Значення функції Лапласа знайдемо по таблиці (див. додаток), маємо:

Ф(1,67) = 0,4525 – ймовірність того, що прилад працюватиме від 10 до 20 років.

При розв’язанні задач, в яких необхідно скласти таблицю розподілу ймовірностей випадкової величини, рекомендується використовувати наступний алгоритм:

1. Встановити, що є випадковою величиною у розглянутій задачі.

2. Перерахувати всі можливі значення випадкової величини.

3. Із умови задачі встановити закон розподілу ймовірностей випадкової величини.

4. Використовуючи відповідну формулу, знайти ймовірності появи можливої випадкової величини.

5. Скласти таблицю розподілу ймовірностей випадкової величини і перевірити, що $\sum_{i=1}^{n} p_i = 1$.

При розв'язанні задач можна користуватися такою додатковою таблицею:

	Закон розподілу	Ймовірність можливих значень	Числові характеристики
	Біномний розподіл	$P(X=k)=C_n^k p^k q^{n-k}$, де $p+q=1$, $k=0,1,2,...,n$.	$M(X)=np$, $D(X)=npq$.
	Нормальний розподіл	$P(\alpha<x<\beta)=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right)$, $P(\|x-a\|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right)$.	$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}}dx.$ $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$, $x\in R.$ a – математичне сподівання; σ – середнє квадратичне відхилення.

Ці розподіли мають значні практичні застосування. Зокрема, завдяки використанню закону розподілу в територіальному управлінні спортивних лотерей, була розкрита група злочинців, які використовуючи службове положення, здійснювали великі розкрадання з лотереї „Спортлото”. Статистичний аналіз виграшів показав значне відхилення від теоретичного розподілу. За даними теоретико-ймовірнісної експертизи у відповідному місті було встановлено спостереження і злочин розкрито.

ЛЕКЦІЯ 7. КОРЕЛЯЦІЙНИЙ ЗВ'ЯЗОК МІЖ ВИПАДКОВИМИ ВЕЛИЧИНАМИ. РЕГРЕСІЯ.

План

1.Встановлення залежності між випадковими величинами. Коваріація і коефіцієнт кореляції.

2. Кореляційне поле, його побудова. Лінія регресії. Знаходження коефіцієнтів лінії регресії.

1.Встановлення залежності між випадковими величинами. Коваріація і коефіцієнт кореляції

У 1870-1900 рр. бельгієць Кетле і англійці Френсіс Гальтон і Карл Пірсон заснували новий науковий напрям – біометрію, в якій вперше стала систематично і кількісно вивчатися мінливість живих організмів і успадкування кількісних ознак. У науковий обіг були введені нові поняття теорії ймовірностей – регресії і кореляції.

Поступово введемо ці поняття. Для цього спочатку розглянемо, які величини називаються незалежними.

За означенням, дві випадкові величини називаються *незалежними,* якщо закон розподілу однієї з них не залежить від того, яких можливих значень набуде інша випадкова величина.

Теорема І. Для того, щоб випадкові величини *X* і У були незалежними, необхідно і достатньо, щоб інтегральна функція розподілу $F(x;\ y)$ дорівнювала добутку функцій розподілу компонент, тобто $F(x;y)=F_1(x)*F_2(y)$.

<u>Наслідок.</u> Для того, щоб випадкові величини X і *У* були незалежними, необхідно і достатньо, щоб диференціальна функція розподілу $f(x;y)$ дорівнювала добутку диференціальних функцій компонент, тобто:

$$f(x;y)=f_1(x)\cdot f_2(y).$$

<u>Зауваження 1</u>**.** Рівності $F(x;y)=F_1(x)*F_2(y)$ і $f(x;y)=f_1(x)\cdot f_2(y)$ можна вважати означеннями незалежності випадкових величин *X* і *Y*.

Крім уже відомих числових характеристик *M(X), M(У), D(X), D(У*) для випадкових величин X і Y вводиться поняття *кореляційного моменту (коваріації)* μ_{xy} *і коефіцієнта кореляції* r_{xy}.

Кореляційним моментом (коваріацією) випадкових величин *X* і *Y* називають число

$$\mu_{xy} = M(XY) - M(X)M(Y).$$

Із означення кореляційного моменту випливає: а) Якщо *X* і *Y* незалежні, то $\mu_{xy}=0$. б) Якщо $\mu_{xy} = 0$, то не завжди *X* і *Y* незалежні.

Коваріація характеризує не лише залежність випадкових величин *X* і *У*, а і міру розсіювання пари *(X,У)* навколо центра розсіювання *(M(X), M(У))*.

Коефіцієнтом кореляції випадкових величин *X* і *Y* називають число

$$r_{xy} = \frac{\mu_{xy}}{\sigma(X)\sigma(Y)},$$

де $\sigma(X) = \sqrt{D(X)}, \qquad \sigma(Y) = \sqrt{D(Y)}.$

Якщо $r_{xy}=0$, то випадкові величини *X і Y* називають *корельованими,* а якщо $r_{xy}=0$, то *некорельованими***.**

Лінійна кореляційна залежність між випадковими величинами *X* і *Y* має вигляд:

$$Y - M(Y) = \frac{\sigma_y}{\sigma_x} r_{xy}(X - M(X)).$$

Ця пряма називається *прямою регресії*, а коефіцієнт $\beta = r_{xy}(\sigma_y / \sigma_x)$ називається *коефіцієнтом регресії***.**

У теорії ймовірностей вивчаються різні поняття, пов'язані з випадковими подіями і випадковими величинами, найважливіші з них - це поняття ймовірності, функції розподілу, математичного сподівання тощо. Але в більшості практичних задач точне значення ймовірності або точний вираз функції розподілу нам невідомі. Тому постає питання про їхнє експериментальне визначення. Математична статистика вивчає методи, які дають змогу за результатами випробувань робити певні ймовірносні висновки. Розглянемо деякі найпростіші задачі, що вивчаються в математичній статистиці.

2. Кореляційне поле, його побудова. Лінія регресії. Знаходження коефіцієнтів лінії регресії.

Розглянемо деякі загальні поняття статистики. Нехай *X* - випадкова величина, функція розподілу якої *F(x)* нам невідома. Досліджуючи цю величину, ми здійснюємо *n* раз той самий експеримент, в результаті чого дістанемо *n* значень величини *X:* x_1 , x_2 …, x_n. Експеримент полягає в тому, що ми вибираємо навмання один з елементів множини *A,* реєструємо деяку його характеристику *X* (наприклад, довжину) і повертаємо цей елемент до множини. Множину *A* називають *генеральною сукупністю***,** а групу елементів, які спостерігали при *n* повтореннях експерименту, - *випадковою вибіркою.*

У більшості випадків нас цікавитимуть не самі елементи множини *A,* а лише згадана їх характеристика *X* і закон її розподілу. Тому у випадку довільної випадкової величини *X* множину всіх її можливих значень також називають *генеральною сукупністю***,** а утворені в результаті *n* експериментів числа $x_1,x_2,..,x_n$ – *вибіркою* з цієї сукупності.

Числа $x_1,x_2,..,x_n$ утворені в результаті *n* експериментів називають *варіантами,* а послідовність варіант, записаних у порядку зростання – *варіаційним рядом.*

Суму $n_1 + n_2 + ... + n_k = n$ називають *об'ємом вибірки***.**

Число $\omega_i = \frac{n_i}{n} (\sum_{n=1}^{k} \omega_n = 1)$ називають *відносною частотою вибірки***.**

Оскільки умови експерименту вважаємо незмінними, а результати експериментів– незалежними один від одного, то результати *n* експериментів x_1, x_2, ..., x_n є незалежними в сукупності випадковими величинами, що мають ту саму функцію *F(x).*

На практиці вибір з генеральної сукупності часто здійснюється так, що після кожного експерименту досліджуваний елемент не повертається до генеральної сукупності перед наступним вибором (так, наприклад, здійснюють вибірковий контроль якості продукції). При цьому не можна вважати, що всі експерименти проводяться в однакових умовах і їх результати незалежні один від одного, оскільки після кожного експерименту склад генеральної сукупності змінюється. Такий вибір називають вибором без повернення на відміну від простого випадкового вибору, або вибору з поверненням. Але якщо генеральна сукупність достатньо велика, а вибірка по своєму об'єму становить незначну її частину, то відміна між цими двома типами вибору незначна, і нею можна нехтувати.

Емпіричною функцією розподілу (функцією розподілу вибірки) називається функція $F_n(x)$, яка визначає для кожного значення $\boldsymbol{x}$ відносну частоту події $X < x$ тобто $F_n(x) = \frac{n_x}{n}$ де n_x - число варіант, менших $\boldsymbol{x}$; n - об'єм вибірки.

Емпірична функція має такі властивості:

1) $0 \le F_n(x) \le 1$,
2) $F_n(x)$ - неспадна;
3) якщо $\boldsymbol{x_1}$ найменша варіанта, а $\boldsymbol{x_k}$ – найбільша, то $F_n(x)=0$ при $x<x_1$, $F(x)=1$ при $x>x_k$.

Полігоном частот називають ломану лінію, відрізки якої з'єднують точки M_1 $(x_1;n_1)$, $M_2(x_2;\ n_2)$, ..., $M_n(x_n;\ y_n)$ – варіанти вибірки і n_i – відповідні їм частоти.

Полігоном відносних частот називається ломана лінія, відрізки якої з'єднують точки $N_1\ (x_1;\ w_i)$, $N_2\ (x_2;\ w_2)$, ..., $N_n\ (x_n;\ w_n)$, де x_i - варіанти вибірки і w_i,- відповідні їм відносні частоти.

Гістограмою частот називається східцева фігура, яка складається з прямокутників, основами яких є частинні інтервали довжиною h і висотою рівною щільності частоти n_i/h.

Площа гістограми відносної частоти дорівнює об’єму вибірки n.

Гістограмою відносних частот називається східцева фігура, яка складається з прямокутників, основами яких є частинні інтервали довжиною *h* і висотою рівною щільності частоти w_i/h_i.

Площа гістограми відносної частоти дорівнює сумі всіх відносних частот, тобто одиниці.

<u>Задача № 1</u>

Складання дискретного варіаційного ряду, обчислення характеристик вариаційного ряду та побудова полігону відносних частот.

Приклад 1.

На кінцевій зупинці була вивчена завантаженість маршрутних таксі. Кількість пасажирів в кажному таксі наведено в табл. 1.1.

Завантаженість таксі на кінцевій зупинці

Табл. 1.1

1	2	6	5	7	5	6	7	6	5
5	4	5	6	8	7	4	8	4	8
6	5	2	4	6	4	5	3	5	4
3	8	6	7	8	2	4	8	7	6
4	6	5	5	6	5	8	5	3	5
7	3	4	1	7	6	2	4	5	7
6	2	3	7	3	2	6	6	4	3
4	6	5	4	8	7	1	2	5	4
6	8	4	5	2	8	7	5	3	6
7	4	5	3	5	5	3	4	5	7

Для статистичного аналізу завантаженості таксі треба:

1. Скласти дискретний варіаціоний ряд величини X – кількості пасажирів в кажному таксі;

2. Обчислити вибіркове середнє $\overline{x}$ (середня кількість пасажирів в кажному таксі), вибіркову дисперсію $D_x = S_x^2$ і вибіркове середне квадратичне відхилення σ_x ;

3. Побудувати полігон відносних частот величини $\boldsymbol{X}$.

4. Зробити висновки щодо завантаженості таксі.

Розв’язання:

1. Складання дискретного варіаційного ряду

Складання дискретного варіаційного ряду полягає в заповненні таблиці (см. табл. 1.2).

У першому рядку записуємо значення дискретної випадкової величини X (кількість пасажирів у кожному таксі), в другій – частоти (n_i), тобто число спостережень, що потрапили в даний інтервал швидкості руху, в третій - відносні частоти $W_i = \frac{n_i}{n}$. При цьому повинні дотримуватися умови $\sum_{i=1}^{k} n_i = n$, $\sum_{i=1}^{k} W_i = 1{,}00$, де k - кількість значень дискретної випадкової величини.

Таблиця 1.2

Дискретний варіаційний ряд

x_i	1	2	3	4	5	6	7	8	Σ
n_i	3	8	10	17	22	17	13	10	100
$W_i = \frac{n_i}{n}$	0,03	0,08	0,1	0,17	0,22	0,17	0,13	0,1	1,00

2. Обчислення вибіркових характеристик дискретного варіаційного ряду

Якщо вибірка представлена у вигляді дискретного варіаційного ряду (табл. 1.1), то вибіркове середнє $\overline{x}$, вибіркову дисперсію S_x^2 і вибіркове середнє квадратичне відхилення σ_x обчислюють за такими формулами:

$$\overline{x} = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \sum_{i=1}^{k} x_i W_i ;$$

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

$$\sigma_x = \sqrt{S_x^2} \ ,$$

де x_i – значеня дискретної випадкової величини X ;

n_i – відповідні частотыи дискретного варіаційного ряду.

Використовуючи дані табл.1.2 і формули (1.1), знаходимо числові характеристики дискретного варіаційного ряду:

$$\bar{x} = 1 \cdot 0{,}03 + 2 \cdot 0{,}08 + 3 \cdot 0{,}1 + 4 \cdot 0{,}17 + 5 \cdot 0{,}22 + 6 \cdot 0{,}17 + 7 \cdot 0{,}13 + 8 \cdot 0{,}1;$$

$$\bar{x} = 5;$$

$$S_x^2 = (1-5)^2 \cdot 0{,}03 + (2-5)^2 \cdot 0{,}08 + (3-5)^2 \cdot 0{,}1 + (4-5)^2 \cdot 0{,}17 +$$
$$+ (0-0)^2 \cdot 0{,}22 + (6-5)^2 \cdot 0{,}17 + (7-5)^2 \cdot 0{,}13 + (8-5)^2 \cdot 0{,}1 = 3{,}36;$$

$$\sigma_x = \sqrt{3{,}36} = 1{,}833.$$

3. Побудова полігону відносних частот

Геометричною ілюстрацією дискретного варіаційного ряду служить полігон відносних частот.

Для побудови полігона частот (рис. 1.1) використовують дані дискретного варіаційного ряду (табл. 1.2). При цьому по осі Ox відкладають значення дискретної величини X, а по осі Oy – відповідні відносні частоти W_i.

4. Висновки

1) Середня завантаженість маршрутних таксі на кінцевій зупинці складаює $\bar{x}$ =5 пасажирів.

2) Середнє квадратичне відхилення завантаженості таксі становить 1,833.

Рис. 1.1. Полігон відносних частот

Задача № 2. Складання неперервного варіаційного ряду, обчислення характеристик вариаційного ряду та побудова гістограми відносних частот.

Приклад 2. При вивченні швидкісного режиму на автомобільній дорозі Київ-Чернігів зареєстрована наступна швидкість руху автотранспорту (див. табл.1.3).

71	112	95	50	98	102	135	142	91	128
98	73	100	125	140	115	108	95	85	135
145	126	75	136	120	95	88	105	125	140
125	118	98	77	100	125	145	122	105	152
155	125	120	105	79	60	142	125	155	100
85	120	106	160	105	80	165	95	75	100
165	100	124	85	140	115	82	105	160	135
75	105	136	120	175	110	107	84	125	185
110	95	80	135	125	105	100	68	87	136
145	96	120	100	98	160	140	115	95	89

Для статистичного аналізу інформації про швидкісний режим руху автотранспорту на даній ділянці дороги потрібно:

1. Скласти інтервальний варіаційний ряд;

2. Обчислити вибіркове середнє $\overline{x}$ (середню швидкість руху), вибіркову дисперсію $D_x = S_x^2$ і вибіркове середне квадратичне відхилення $б_x$;

3. Побудувати гістограму швидкості руху автотранспорту;

4. Визначити середню швидкість автомобілів, що перевищують дозволену швидкість руху 110 км / год;

5. Зробити висновки про швидкісний режим.

Розв'язання

1.Складання інтервального варіаційного ряду

При складанні інтервального варіаційного ряду об'єднують в один інтервал близькі значення спостережуваної випадкової величини x_i. Оптимальну ширину кожного з таких інтервалів (крок) знаходимо за формулою Стерджеса

$$h = \frac{x_{\max} - x_{\min}}{1 + 3{,}2\lg n}$$

де $x_{\max}$, $x_{\min}$ - відповідно найбільше та найменше значення випадкової величини X;

n - обсяг вибірки (кількість вимірів);

$1 + 3{,}2\lg n$ – число інтервалів варіаційного ряду, визначається за обсягом вибірки.

Для нашої задачі з табл. 1.3 знаходимо найбільшу і найменшу швидкості руху: $x_{\max}$ = 185 км/год, $x_{\min}$ = 50 км/год, n = 100 і за формулою знаходимо ширину кожного з інтервалів швидкості руху

$$h = \frac{185 - 50}{1 + 3{,}2\lg 100} = \frac{135}{7{,}4} \approx 18{,}24 \approx 20 \text{ км/год.}$$

Для зручності обчислень ширину інтервалу швидкості руху округляємо до 20 км /год і приймаємо цю ширину за оптимальну. Отже, при побудові інтервального варіаційного ряду будемо об'єднувати в один інтервал значення швидкості, які відрізнялася один від одного не більше ніж на 20 км/год.

Складання інтервального варіаційного ряду полягає в заповненні таблиці (див. табл. 1.4).

Таблиця 1.4

Інтервальний варіаційний ряд

i - номер інтервалу	**1**	**2**	**3**	**4**	**5**	**6**	**7**	$\sum$
x_{i-1} - x_i- межі інтервалу	50-70	70-90	90-110	110-130	130-150	150-170	170-190	
m_i - частота	3	17	30	24	16	8	2	100
$P_i = \frac{m_i}{n}$	0,03	0,17	0,30	0,24	0,16	0,08	0,02	1,00
F_i	0,03	0,20	0,50	0,74	0,90	0,98	1,00	

У першому рядку таблиці записуємо номери інтервалів швидкості руху (i), у другому - межі інтервалів швидкості руху $(x_{i-1} - x_i)$, в третьому - частоти (m_i), тобто кількість спостережень, що потрапили в даний інтервал швидкості руху, у четвертому – відносні частоти (p_i = $\frac{m_i}{n}$). При цьому повинні дотримуватися умови $\sum_{i=1}^{7} m_i = n$, $\sum_{i=1}^{7} p_i = 1{,}00$, які дозволяють перевірити правильність обчислень. У п'ятому рядку записується накопичена відносна частота F_i , яка визначається як сума відносних частот ***і***-го і всіх попередніх інтервалів.

2. Обчислення вибіркових характеристик

Якщо вибірка представлена у вигляді індивідуальних даних (табл. 1.3), то вибіркове середнє значення $\overline{x}$, вибіркову дисперсію S_x^2 і вибіркове середнє квадратичне відхилення $б_x$ обчислюють за формулами:

$$\overline{x} = \frac{\sum x_i}{n}, \qquad S_x^2 = \frac{\sum (x_i - \overline{x})^2}{n}, \quad б_x = \sqrt{S_x^2}.$$

**

Використовуючи ці формули і дані табл. 1.3 знаходимо числові характеристики вибірки для несгруппірованних даних:

$$\overline{x} = \frac{71 + 112 + 95 + ... + 95 + 89}{100} = 111{,}9 \approx 112,$$

$$S_x^2 = \frac{(71 - 111{,}9)^2 + (112 - 111{,}9)^2 + ... + (89 - 111{,}9)^2}{100} = 727{,}1,$$

$$б_x = \sqrt{727{,}1} = 26{,}96 \approx 27.$$

Якщо вибірка представлена у вигляді інтервального варіаційного ряду (табл. 1.4), то вибіркове середнє $\overline{x}$, вибіркову дисперсію S_x^2 і вибіркове середнє квадратичне відхилення $б_x$ можна обчислити за такими формулами:

$$\overline{x} = \frac{\sum c_i m_i}{n} = \sum c_i p_i$$

$$S_x^2 = \frac{\sum (c_i - \overline{x})^2 m_i}{n} = \sum (c_i - \overline{x})^2 p_i;$$

$$б_x = \sqrt{S_x^2},$$

де $c_i = \dfrac{x_{i-1} + x_i}{2}$ – середина інтервалу $(x_{i-1} - x_i)$;

m_i – частоти варіаційного ряду.

Для зручності обчислення вибіркових характеристик інтервального варіаційного ряду складаємо розрахункову таблицю 1.5.

Таблиця 1.5

Вихідні дані для обчислення вибіркових характеристик

$x_{i-1} - x_i$ - межі інтервалу	50-70	70-90	90-110	110-130	130-150	150-170	170-190	$\sum$
c_i - середина інтервалу	**60**	**80**	**100**	**120**	**140**	**160**	**180**	

$\boldsymbol{m_i}$ - частота	**3**	**17**	**30**	**24**	**16**	**8**	**2**	**100**
$\boldsymbol{c_i m_i}$	**180**	**1360**	**3000**	**2880**	**2240**	**1280**	**360**	**11300**
$\boldsymbol{c_i} - \bar{x}$	**-53**	**-33**	**-13**	**7**	**27**	**47**	**67**	
$(\boldsymbol{c_i} - \bar{x})^2$	**2809**	**1089**	**169**	**49**	**729**	**2209**	**4489**	
$(\boldsymbol{c_i} - \bar{x})^2 \boldsymbol{m_i}$	**8427**	**18513**	**5070**	**1176**	**11664**	**17672**	**8978**	**71500**

Використовуючи вихідні дані табл.1.5 і формули, знайдемо числові характеристики інтервального варіаційного ряду:

$$\bar{x} = \frac{180+1360+3000+2880+2240+1280+360}{100} = 113 \text{ км/год},$$

$$S_x^2 = \frac{2809 \bullet 3 + 1089 \bullet 17 + 169 \bullet 30 + 49 \bullet 24 + 729 \bullet 16 + 2209 \bullet 8 + 4489 \bullet 2}{100} = 715 (\text{км/год})^2,$$

$$б_x = \sqrt{715} = 26{,}74 \approx 27 \text{ км/год}.$$

Для порівняння числових характеристик вибірки, знайдених за індивідуальними даними і за згрупованими даними, випливає, що ці швидкості і дисперсії різні. При цьому слід знати, що розрахунок середньої швидкості і дисперсії за згрупованими даними менш точний, але його переваги в меншій трудомісткості.

3. Побудова гістограми відносних частот

Гістограма є геометричною ілюстрацією статистичних даних. Для побудови гістограми швидкості руху автотранспорту використовуємо дані інтервального варіаційного ряду (табл.1.4). При цьому (див. рис.1.2) по осі ОХ відкладаємо межі інтервалів випадкової величини Х (межі інтервалів швидкості руху автотранспорту $\boldsymbol{x_{i-1}}$ - $\boldsymbol{x_i}$), а по осі ОУ - відповідні відносні частоти p_i.

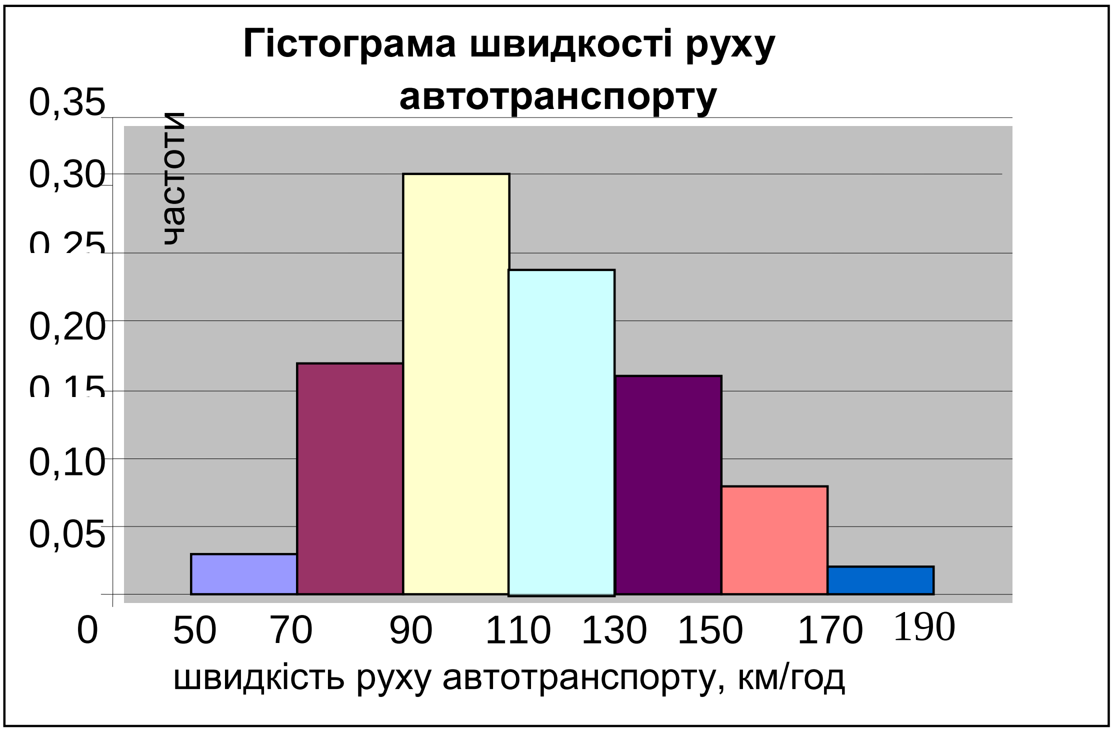


Рис. 1.2. Гістограма швидкості руху автотранспорту

4. Середню швидкість автомобілів, що перевищують дозволену швидкість 110 км /год, обчислюємо за формулою (1.3):

$$\bar{x} = \frac{112 + 135 + ... + 115}{48} = \frac{6534}{48} = 136{,}125 \approx 136.$$

5. Висновки

1) Середня швидкість руху автотранспорту на 60 км траси Київ - Чернігів = 112 км / год.

2) Середнє квадратичне відхилення швидкості руху автотранспорту $\boldsymbol{б_x} \approx 27$ км / год.

3) 3% автотранспорту на даній ділянці дороги рухається зі швидкістю менше 70 км / год.

4) 2% автотранспорту рухається зі швидкістю понад 170 км / год.

5) 30% автомобілів рухається зі швидкістю 90-110 км / год.

6) 50% автомобілів перевищують дозволену швидкість руху-ня (110 км / год).

7) Середня швидкість автомобілів, що перевищують дозволену швидкість, дорівнює 136 км / год.

2. Кореляційне поле, його побудова. Лінія регресії. Знаходження коефіцієнтів лінії регресії.

Коефіцієнт кореляції між випадковими величинами *X,Y* дорівнює + 1 або - 1 тоді і тільки тоді, коли **X** і Y зв'язані лінійною залежністю *Y= aX + C*. Ця пряма називається *прямою регресії* (або *прямою середньої квадратичної регресії*) Y на *X*, якщо коефіцієнти *a* і *C* вибрано так, щоб середнє квадратичне відхилення *aX+C* від *Y* було мінімальним $M(У\text{-}aX\text{-}C)^2=min$. Аналогічно визначається пряма регресії *X* на *Y*.

2. Лінійна кореляція

Якщо обидві *лінії регресії* (залежності між *X і Y) Y* на *X* (*Y* -функція, *X* - аргумент) і *X* на *Y* (*X* - функція, *Y* - аргумент) – ***прямі***, то кореляція називається ***лінійною***.

Вибіркове рівняння прямої лінії регресії (залежності між *X і Y*) Y на *X* (Y – функція, *X* – аргумент) має вигляд

$$\overline{y_x} - \overline{y} = r_R \frac{\sigma_x}{\sigma_y}(x - \overline{x}),$$

де y_x - умовна середня, *x і y* - вибіркові середні ознак *X* і *У*, $о_x$; $о_y$ - вибіркові середні квадратичні відхилення, r_B- вибірковий коефіцієнт кореляції

$$r_B = \left[\sum n_{xy} - n\overline{xy}\right] / n\sigma_x\sigma_y .$$

Вибіркове рівняння прямої лінії регресії (залежності між *X* і *У*) *X* на *У* (*X* - функція, *У*- аргумент) має такий вигляд

$$\overline{x_y} - \overline{x} = r_B \frac{\sigma_x}{\sigma_y}(y - \overline{y}).$$

Задача № 3.

Дослідження лінійного зв'язку між випадковими величинами та знаходження рівняння лінії регресії.

Приклад 2. Органами МВС України в 2009 році були зафіксовані правопорушення та злочини в 15 областях. Їх кількість на 10 тис. осіб наведена в табл. 2.1.

Таблиця 2.1

Кількість зареєстрованих правопорушень і злочинів

Область	X – кількість правопорушень на 10 тис. чол.	Y – кількість злочинів на 10 тис. чол.
1	1800	79
2	231	21
3	1799	115
4	1667	97
5	487	66
6	746	35
7	1004	37
8	131	44
9	1358	81
10	1402	53
11	1603	90
12	1898	127
13	1974	132
14	1327	48
15	1408	63

Для аналізу криміногенної ситуації і прийняття правильних управлінських рішень потрібно:

1. Дослідити наявність зв'язку між кількістю правопорушень і кількістю злочинів;

2. Знайти рівняння лінійної залежності (рівняння лінії регресії) між кількістю правопорушень і кількістю злочинів;

3. Побудувати кореляційне поле і нанести на нього лінію регресії, середнє число правопорушень і середнє число злочинів;

4. Зробити висновки про криміногенну ситуацію.

Розв’язання

1. Дослідження взаємозв'язку між правопорушеннями і злочинами

Для дослідження взаємозв'язку між кількістю правопорушень і кількістю злочинів використовуємо елементи кореляційного аналізу. Оцінку тісноти лінійного зв'язку між випадковими величинами (ознаками) зробимо, обчисливши вибірковий коефіцієнт кореляції *r* за формулою

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y}, \qquad ***$$

де $\overline{x}, \overline{y}, \overline{xy}$ – середне значення величин X і Y і їх попарних добутків XY;

σ_x, σ_y – вибіркові середні квадратичні відхилення величин X і Y, які знаходять за формулами:

$$S_x^2 = \overline{(x^2)} - (\overline{x})^2, \qquad S_y^2 = \overline{(y^2)} - (\overline{y})^2,$$

$$\sigma_x = \sqrt{S_x^2} \qquad \sigma_y = \sqrt{S_y^2}$$

Коефіцієнт кореляції r вимірює тісноту лінійного зв'язку між випадковими величинами X і Y і має наступні властивості:

1. Якщо X і Y незалежні, то $r = 0$;

2. Для будь-яких X і Y має місце нерівність $-1 \le r \le 1$;

3. $r = \pm 1$ тоді і тільки тоді, коли між випадковими величинами X і Y існує функціональна лінійна залежність виду $y = kx + b$.

Якщо процедура обчислення коефіцієнта кореляції і параметрів рівняння лінії регресії відбувається на калькуляторі, то вихідні дані для обчислення оформляються у вигляді допоміжної таблиці (див. табл. 2.2).

Таблиця 2.2

Вихідні дані для кореляційного та регресійного аналізів

№	*X*	*У*	*XУ*	X^2	$У^2$
1	1800	79	142200	3240000	6241
2	231	21	4851	53361	441
3	1799	115	206885	3236401	13225
4	1667	97	161699	2778889	9409
5	487	66	32142	237169	4356
6	746	35	26110	556516	1225
7	1004	37	37148	1008016	1369
8	131	44	5764	17161	1936
9	1358	81	109998	1844164	6561
10	1402	53	74306	1965604	2809
11	1603	90	144270	2569609	8100
12	1898	127	241046	3602404	16129
13	1974	132	260568	3896676	17424
14	1327	48	63696	1760929	2304
15	1408	63	88704	1982464	3969
$\sum$	18835	1088	1599387	28749363	95498
Середне значення	1255,(6)	72,53	106625,8	1916624,2	6366,53

Використовуючи дані табл. 2.2 і формулу (**) знаходимо середні значення випадкових величин X і Y, а середні квадратичні відхилення знаходимо за формулами:

$$\sigma_x = \sqrt{\overline{(x^2)} - (\bar{x})^2} = \sqrt{1916624{,}2 - (1255{,}6)^2} = \sqrt{340092{,}84} = 584{,}175;$$

$$\sigma_y = \sqrt{\overline{(y^2)} - (\bar{y})^2} = \sqrt{6366{,}53 - (72{,}53)^2} = \sqrt{1105{,}929} = 33{,}26.$$

За формулою (***) знаходимо вибірковий коефіцієнт кореляції

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{б_x \cdot б_y} = \frac{106625{,}8 - 1255{,}6 \cdot 72{,}53}{583{,}175 \cdot 33{,}26} = \frac{15557{,}132}{583{,}175 \cdot 33{,}26} = 0{,}802.$$

Високий вибірковий коефіцієнт кореляції $r = 0{,}802$ вказує на тісний лінійний зв'язок між X (числом виявлених правопорушень) і Y (числом злочинів у областях).

Для встановлення надійності вибіркового коефіцієнта кореляції r обчислюємо його середнє квадратичне відхилення σ_r і показник Ляпунова μ за формулами:

$$\sigma_r = \frac{1 - r^2}{\sqrt{n}}, \qquad \mu = \frac{|r|}{\sigma_r} \qquad (****)$$

Встановлено, якщо показник Ляпунова μ > 2,6, то між величинами X і Y існує тісний лінній1ний зв'язок. В нашому випадку

$$\sigma_r = \frac{1 - (0{,}8)^2}{\sqrt{15}} = \frac{0{,}36}{3{,}87} = 0{,}093, \quad \text{а} \quad \mu = \frac{0{,}8}{0{,}093} = 8{,}6\,.$$

Так як показник Ляпунова μ = *8,6 >2,6* (тобто перевищує критичне значеня), то вважаємо, що між числом виявлених правопорушень і числом злочинів існує тісний лінійний зв'язок.

2. Проведення регресійного аналізу

Регресійний аналіз дозволяє визначити вид залежності між випадковими величинами. Для знаходження рівняння лінії регресії, що встановлює залежність між випадковими величинами у вигляді $y = kx + b$, необхідні експериментальні дані. Обчислюємо коефіцієнти (параметри) рівняння k і b лінії регресії за формулами:

$$\kappa = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x^{\ 2}}; \qquad b = \overline{y} - \kappa \overline{x}.$$

Використовуючи дані табл. 2.2 за цими формулами знаходимо коефіцієнти рівняння лінії регресії

$$\kappa = \frac{106625{,}8 - 12555{,}6 \cdot 72{,}53}{340092{,}84} = 0{,}04574;$$

$$b = 72{,}53 - 0{,}046 \cdot 1255,(6) = 14{,}7696.$$

Отже, рівняння лінії регресії

$$y = 0{,}046 \cdot x + 14{,}77$$

,

де x – кількість правопорушень на 10 тис. чоловік,

y – кількість злочинів на 10 тис. чоловік.

3. Побудова кореляційного поля

При побудові кореляційного поля розглядаємо кожний рядок табл.2.1 як точку $\boldsymbol{M}_i\,(x_i\,,\,y_i\,)$ на координатній площині Oxy з координатами $(x_i\,,\,y_i\,)$. За характером розташування точок можна скласти уявлення про форму залежності випадкових величин. Таким чином, кореляційне поле в нашому прикладі містить 15 точок (областей). Для нанесення лінії регресії на кореляційне поле знаходимо координати двох точок A і B, що лежать на прямій: *A(200; 23,97)* і *B(2000; 106,77)*. Графік лінії регресії та кореляційне поле наведені на рис. 2.1.

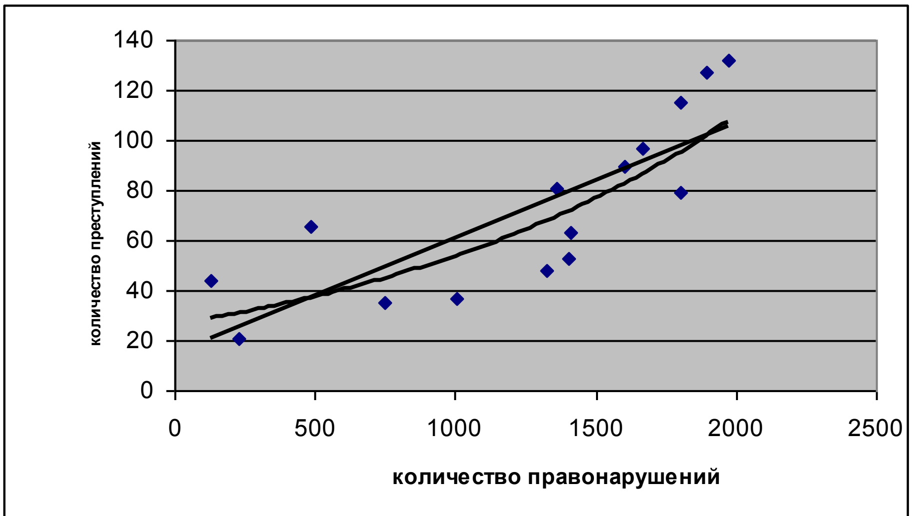


Рис.2.1. Залежність між кількістю злочинів і кількістю правопорушень

Точки не лежать на одній лінії $y = a + bx$. Але це не дивно. Адже крім x на поведінку y впливають і інші фактори. Подальший аналіз отриманого рівняння дозволяє сказати, наскільки сильний вплив неврахованих факторів, чи дійсно модель лінійна тощо. На змінні x, y накладаються умови. Для опису природи зв'язку використовується термін «регресія». Коефіцієнт b називається показником нахилу лінії лінійної регресії.

Зауваження. Замість обчислень коефіцієнтів a і b за формулами можна скористатися відповідно статистичними функціями ОТРЕЗОК (изв_знач_у; изв_знач_х) і НАКЛОН (изв_знач_у; изв_знач_х) майстра функцій f_x пакету Excel. Тут изв_знач_у і изв_знач_х - це посилання на клітинки, які містять значення змінних y і x відповідно.

4. Висновки

1. Середнє число правопорушень в 2009 році на 10 тисяч осіб склало *1255*, а середня кількість злочинів - *72*.

2. Між числом зареєстрованих правопорушень і числом злочинів існує тісний лінійний зв'язок, який характеризується високим коефіцієнтом кореляції $r = 0,802$.

3. Аналізуючи рівняння лінії регресії, бачимо, що збільшення виявлених правопорушень на 1000 одиниць статистично тягло за собою *46* злочинів.

4. З рис. 2.1 видно, що точки кореляційного поля розташовані близько від прямої лінії, що свідчить про наявність лінійного зв'язку. Але ще менше відхилення експериментальних точок від кривої лінії вказує на те, що нелінійна регресія характеризується більш тісним зв'язком. Рівняння лінії

регресії і тісноту нелінійного зв’язку також можна встановити за допомогою регресійного і кореляційного аналізів.

ІНДИВІДУАЛЬНІ ЗАВДАННЯ З ТЕОРІЇ ЙМОВІРНОСТЕЙ

1. Комбінаторика

1.1 Розклад одного дня містить 3 різних пари. Знайти кількість можливих розкладів, якщо вивчається 8 дисциплін.

Відповідь: 336.

1.2 У студентську раду обрані 7 осіб. З них треба обрати голову і заступника. Визначити кількість способів, за допомогою яких це можна зробити.

Відповідь: 42.

1.3 Скільки трьохцифрових чисел можна скласти з цифр 0, 2, 4, 5, 8 так, щоб у кожному числі цифри не повторювалися?

Відповідь: 48.

1.4 Скільки чотирьохцифрових чисел, в яких цифри не повторюються, можна отримати з цифр 0, 1, 3, 5, 4, 8?

Відповідь: 300.

1.5 Студенту треба здати 4 іспити протягом 10 днів. Скільки існує способів це зробити, якщо відомо, що перший іспит він здасть в перший день, а останній - в останній день і в один день він здасть не більше одного іспиту?

Відповідь: 56.

1.6 З 6 співробітників фірми треба обрати голову та заступника. Скільки існує способів це зробити?

Відповідь: 30.

1.7 Фірма має 3 вакантних посади і 9 претендентів на ці посади. Скільки існує способів заповнення цих посад?

Відповідь: 504.

1.8 Визначити кількість способів, за допомогою яких можна розсадити 5 студентів на 12 місцях.

Відповідь: 1320.

1.9 Скільки існує способів присудження золотої, срібної і бронзової медалі на змаганнях, в яких беруть участь 15 чоловік.

Відповідь: 2730.

1.10 Визначити кількість способів, за допомогою яких можна отримати 3-х кольоровий прапорець з горизонтальними смужками однакової ширини, якщо тканина є 6-ти кольорів.

Відповідь: 120.

2. Класичне визначення ймовірності

2.1 З 250 автомобілів, що проходять техогляд, 45 мають великі несправності, а 150 - дрібні. Яка ймовірність того, що випадковим чином вибраний автомобіль: а) має велику несправність, б) немає несправностей.

Відповідь: $P(A)=\frac{9}{50}$, $P(B)=\frac{11}{50}$.

2.2 У азартній грі "рулетка" кулька кидають на обертальний диск. Диск має 38 секторів: 18 червоних, 18 чорних і 2 зелених. Попадання кульки в будь-який із секторів рівноможливими. Знайти ймовірність того, що кулька потрапила в: а) червоний сектор, б) чорний або зелений.

Відповідь: $P(A)=\frac{9}{19}$, $P(B)=\frac{10}{19}$.

2.3 Свідок ДТП забув останню цифру номера автомобіля, який скоїв наїзд на пішохода і зник з місця події. Яка ймовірність, що навмання вибрана цифра: а) виявиться правильною, б) виявиться правильною, якщо він пам'ятає, що цифра ділиться на 3.

Відповідь: $P(A)=\frac{1}{10}$, $P(B)=\frac{1}{4}$.

2.4 З 30 екзаменаційних білетів студент встиг вивчити на іспит лише 25. Яка ймовірність, що він витягне: а) непідготовлений білет; б) підготовлений білет.

Відповідь: $P(A)=\frac{1}{6}$, $P(B)=\frac{5}{6}$

2.5 Кількість злочинів, скоєних особами, які не які досягли повноліття, в Чернігові за літні місяці становить 40, у віці 18-25 років - 37, між 26 і 35 роками - 23, за 35 - 60. Знайти ймовірність того, що злочин буде вчинено: а) неповнолітнім, б) повнолітнім.

Відповідь: $P(A)=\frac{1}{4}$, $P(B)=\frac{3}{4}$.

2.6 З продуктів, доставлених до супермаркету, 400 випущено не ліцензованою фірмою А, 350 - ліцензованою фірмою В, 250 – ліцензованою фірмою С. З метою контролю продукції навмання вибирається один з продуктів. Яка ймовірність, що він випущений: а) не ліцензованою фірмою А; б) ліцензованою фірмою.

Відповідь: $P(A)=\frac{2}{5}$, $P(B)=\frac{3}{5}$.

2.7 На митниці конфісковано 14 комп'ютерів: типу А - 4 шт., типу В – 2 шт., типу С - 8 шт. Яка ймовірність того, що випадковим чином вибраний комп'ютер: а) типу А, б) типу В або С.

Відповідь: $P(A)=\frac{2}{7}$, $P(B)=\frac{5}{7}$.

2.8 У страховій компанії характерна така тенденція страхування майна приватних осіб: у 2003 році застрахували своє майно 600 осіб, у 2004 році - 647, у 2005 - 753. Яка ймовірність того, що довільно обрана людина з цих списків застрахувала своє майно: а) у 2005 році; б) у 2003 або 2004 роках.

Відповідь: $P(A)=\frac{3}{10}$, $P(B)=\frac{7}{10}$.

2.9 З 12 викрадених автомобілів 3 викрадено правослухняними особами, 2 автомобілі - особами, що мають судимість, 7 автомобілів-особами, що стоять на обліку в ОВС. Яка ймовірність того, що довільно вибраний автомобіль а) викрадений особою, яка має судимість, б) викрадений правослухняну особою або особою, що стоять на обліку в ОВС.

Відповідь: $P(A)=\frac{1}{6}$, $P(B)=\frac{5}{6}$.

2.10 У групі 25 студентів, серед яких 7 відмінників. Для змагань за списком навмання призначено 10 студентів. Яка ймовірність того, що серед них буде 5 відмінників.

Відповідь: 0,11.

3. Основні теореми теорії ймовірностей

3.1 З трьох знарядь намагаються влучити в ціль. Ймовірність влучення в ціль при одному пострілі з першого знаряддя дорівнює 0,9; для другого і третього знаряддя ймовірності відповідно рівні 0,8 і 0,9. Знайти ймовірність того, що: а) тільки один влучить у ціль, б) тільки два потраплять в ціль, в) всі три потраплять у ціль; г) хоча б один влучить у ціль.

3.2 Експедиція видавництва відправила газети в два поштових відділення. Імовірність своєчасної доставки газет у кожне відділення пошти дорівнює 0,85. Знайти ймовірність того, що а) обидва поштових відділення одержать газети вчасно; б) обидва поштових відділення одержать газети з запізненням, в) тільки одне поштове відділення отримає газети вчасно; г) хоча б одне поштове відділення отримає газети вчасно.

3.3 Працівник обслуговує три верстата, які працюють незалежно один від одного. Імовірність того, що протягом години перший верстат працюватиме без нагляду, дорівнює 0,82. Для другого і третього верстатів ці ймовірності рівні відповідно 0,9 і 0,87. Яка ймовірність того, що протягом години: а) три верстата працюватимуть без нагляду; б) всі три верстата будуть потребувати нагляду; в) хоча б один верстат потребуватиме нагляду працівника.

3.4 Ймовірність своєчасного виконання студентом контрольної роботи по кожній з трьох дисциплін відповідно дорівнює 0,7, 0,6 і 0,9. Знайти ймовірність того, що студент вчасно виконає контрольну роботу: а) з двох дисциплін, б) хоча б з двох дисциплін; в) хоча б з однієї дисципліни.

3.5 Відділ технічного контролю перевіряє вироби на стандартність. Імовірність того, що виріб нестандартний, дорівнює 0,2. Знайти ймовірність того, що: а) з трьох перевірених виробів тільки один виявиться нестандартним, б) нестандартним виявиться тільки четвертий за порядком перевірений виріб; в) хоча б один виріб виявиться нестандартним.

3.6 Ймовірність того, що потрібна деталь міститься в першому, другому, третьому і четвертому ящиках відповідно рівні 0,7, 0,9, 0,6 і 0,8. Знайти ймовірність того, що деталь міститься: а) не більше ніж у трьох ящиках, б) не менш ніж у двох ящиках, в) хоча б в одному ящику.

3.7 Два клієнти зайшли до магазину. Імовірність того, що перший клієнт захоче зробити покупку дорівнює 0,3, другий - 0,8. Знайти ймовірність того, що захочуть зробити покупку: а) обидва клієнти; б) тільки один клієнт, в) хоча б один клієнт; г) жоден з клієнтів не захоче зробити покупку.

3.8 Два клієнти звернулися в туристичну фірму. Імовірність того, що перший клієнт захоче скористатися послугами фірми дорівнює 0,6, другий - 0,7. Знайти ймовірність того, що захочуть скористатися послугами туристичної фірми: а) обидва клієнти; б) тільки один клієнт, в) хоча б один клієнт; г) жоден з клієнтів.

3.9 Два стрільці провели по одному пострілу по мішені. Можливість поразки мішені кожним із стрільців дорівнює 0,65. Знайти ймовірність того, що: а) обидва стрілка вразять мішень, б) обидва стрілка промахнуться; в) тільки один стрілець вразить мішень; г) хоча б один стрілець вразить мішень.

3.10 Три стрілка влучають в ціль. Ймовірність поразки першого стрільця дорівнює 0,8; для другого і третього стрільців ці ймовірності відповідно рівні 0,7 і 0,9. Знайти ймовірність того, що: а) тільки один із стрільців попаде в ціль, б) тільки два стрілка попадуть в ціль; в) всі три стрілка попадуть в ціль; г) хоча б один стрілець попаде в ціль.

Варіант 1

1.Задано закон розподілу дискретної випадкової величини. Знайти: p_2; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-2	-1	3	5	8
p	0.3	p_2	0.2	0.1	0.3

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X набуде значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0,\ при\ x\le -2, \\ \dfrac{(x+2)^2}{3},\ при -2<x\le 2, \\ 1,\ при\ x>2. \end{cases} \qquad \alpha=0; \beta=1$$

Варіант 2

1.Задано закон розподілу дискретної випадкової величини . Знайти: p_1; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-3	-2	2	4	5
p	p_1	0.1	0.3	0.3	0.2

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X прийме значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0\,, при\ x\le 1, \\ \dfrac{(x-1)^2}{5}\,, при\ 1<x\le 6, \\ 1,\ при\ x>6. \end{cases} \qquad \alpha=3; \beta=4$$

Варіант 3

1.Задано закон розподілу дискретної випадкової величини. Знайти: p_2; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-3	-1	2	4	8
p	0.3	p_2	0.1	0.2	0.1

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X прийме значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0\,, при\ x\le 1, \\ \dfrac{(x+1)^2}{3}\,, при\ 1<x\le 7, \\ 1, при\ x>7. \end{cases} \qquad \alpha=2; \beta=3$$

Варіант 4

1.Задано закон розподілу дискретної випадкової величини. Знайти: p_3; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-6	-4	2	3	7
p	0.3	0,4	p_3	0.1	0.2

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X прийме значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0, при\ x\le -4, \\ \dfrac{(x+4)^2}{5}, при\ -4<x\le 3, \\ 1\,, при\ x>3. \end{cases} \qquad \alpha=0; \beta=1$$

Варіант 5

1.Задано закон розподілу дискретної випадкової величини. Знайти: p_2; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-3	0	4	6	9
p	0.3	p_2	0.3	0.1	0.1

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X прийме значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0 \; при \; x\le 5, \\ \dfrac{(x-5)^2}{4} \; при \; 5<x\le 7, \\ 1 \; при \; x>7. \end{cases} \qquad \alpha=5; \beta=6$$

Варіант 6

1.Задано закон розподілу дискретної випадкової величини. Знайти: p_5; Mo ; $M(X)$; $Д(X)$; $\sigma(X)$.

x	-7	-2	5	8	10
p	0.3	0,1	0.2	0.1	p_5

2.Випадкова величина X задана інтегральною функцією розподілу ймовірності F(x). Потрібно :

1) знайти диференціальну функцію (щільність ймовірності) f(x);

2) знайти ймовірність того, що випадкова величина X прийме значення які належать інтервалу (α ; β);

3) побудувати графіки функцій F(x) і f(x);

4) знайти математичне сподівання M(x), дисперсію D(x), середнє квадратичне відхилення σ(x) випадкової величини X;

5) знайти моду M_0; 6) знайти медіану M_e.

$$F(x)=\begin{cases} 0 \, , при \; x\le -4, \\ \dfrac{(x+4)^2}{7} \, , при \; -4<x\le -1, \\ 1, \; при \; x>0. \end{cases} \qquad \alpha=-3; \beta=-2$$

ДОДАТКИ

Таблиця значень функції Гауса $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	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	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046

3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
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	0003	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	0002	000

Таблиця значень функції Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt$

x	$\Phi(x)$	***x***	$\Phi(x)$	***x***	$\Phi(x)$	***x***	$\Phi(x)$	***x***	$\Phi(x)$	***x***	$\Phi(x)$
0,00	0,0000	0,43	0,1664	0,86	0,3051	1,29	0,4015	1,72	0,4573	2,30	0,4893
0,01	0,0040	0,44	0,1700	0,87	0,3078	1,30	0,4032	1,73	0,4582	2,32	0,4898
0,02	0,0080	0,45	0,1736	0,88	0,3106	1,31	0,4049	1,74	0,4591	2,34	0,4904
0,03	0,0120	0,46	0,1772	0,89	0,3133	1,32	0,4066	1,75	0,4599	2,36	0,4909
0,04	0,0160	0,47	0,1808	0,90	0,3159	1,33	0,4082	1,76	0,4608	2,38	0,4913
0,05	0,0199	0,48	0,1844	0,91	0,3186	1,34	0,4099	1,77	0,4616	2,40	0,4918
0,06	0,0239	0,49	0,1879	0,92	0,3212	1,35	0,4115	1,78	0,4625	2,42	0,4922
0,07	0,0279	0,50	0,1915	0,93	0,3238	1,36	0,4131	1,79	0,4633	2,44	0,4927
0,08	0,0319	0,51	0,1950	0,94	0,3264	1,37	0,4147	1,80	0,4641	2,46	0,4931
0,09	0,0359	0,52	0,1985	0,95	0,3289	1,38	0,4162	1,81	0,4649	2,48	0,4934
0,10	0,0398	0,53	0,2019	0,96	0,3315	1,39	0,4177	1,82	0,4656	2,50	0,4938
0,11	0,0438	0,54	0,2054	0,97	0,3340	1,40	0,4192	1,83	0,4664	2,52	0,4941
0,12	0,0478	0,55	0,2088	0,98	0,3365	1,41	0,4207	1,84	0,4671	2,54	0,4945
0,13	0,0517	0,56	0,2123	0,99	0,3389	1,42	0,4222	1,85	0,4678	2,56	0,4948
0,14	0,0557	0,57	0,2157	1,00	0,3413	1,43	0,4236	1,86	0,4686	2,58	0,4951
0,15	0,0596	0,58	0,2190	1,01	0,3438	1,44	0,4251	1,87	0,4693	2,60	0,4953
0,16	0,0636	0,59	0,2224	1,02	0,3461	1,45	0,4265	1,88	0,4699	2,62	0,4956
0,17	0,0675	0,60	0,2257	1,03	0,3485	1,46	0,4279	1,89	0,4706	2,64	0,4959
0,18	0,0714	0,61	0,2291	1,04	0,3508	1,47	0,4292	1,90	0,4713	2,66	0,4961
0,19	0,0753	0,62	0,2324	1,05	0,3531	1,48	0,4306	1,91	0,4719	2,68	0,4963
0,20	0,0793	0,63	0,2357	1,06	0,3554	1,49	0,4319	1,92	0,4726	2,70	0,4965
0,21	0,0832	0,64	0,2389	1,07	0,3577	1,50	0,4332	1,93	0,4732	2,72	0,4967
0,22	0,0871	0,65	0,2422	1,08	0,3599	1,51	0,4345	1,94	0,4738	2,74	0,4969
0,23	0,0910	0,66	0,2454	1,09	0,3621	1,52	0,4357	1,95	0,4744	2,76	0,4971
0,24	0,0948	0,67	0,2486	1,10	0,3643	1,53	0,4370	1,96	0,4750	2,78	0,4973
0,25	0,0987	0,68	0,2517	1,11	0,3665	1,54	0,4382	1,97	0,4756	2,80	0,4974
0,26	0,1026	0,69	0,2549	1,12	0,3686	1,55	0,4394	1,98	0,4761	2,82	0,4976
0,27	0,1064	0,70	0,2580	1,13	0,3708	1,56	0,4406	1,99	0,4767	2,84	0,4977
0,28	0,1103	0,71	0,2611	1,14	0,3729	1,57	0,4418	2,00	0,4772	2,86	0,4979
0,29	0,1141	0,72	0,2642	1,15	0,3749	1,58	0,4429	2,02	0,4783	2,88	0,4980
0,30	0,1179	0,73	0,2673	1,16	0,3770	1,59	0,4441	2,04	0,4793	2,90	0,4981
0,31	0,1217	0,74	0,2703	1,17	0,3790	1,60	0,4452	2,06	0,4803	2,92	0,4982
0,32	0,1255	0,75	0,2734	1,18	0,3810	1,61	0,4463	2,08	0,4812	2,94	0,4984
0,33	0,1293	0,76	0,2764	1,19	0,3830	1,62	0,4474	2,10	0,4821	2,96	0,4985
0,34	0,1331	0,77	0,2794	1,20	0,3849	1,63	0,4484	2,12	0,4830	2,98	0,4986
0,35	0,1368	0,78	0,2823	1,21	0,3869	1,64	0,4495	2,14	0,4838	3,00	0,49865
0,36	0,1406	0,79	0,2852	1,22	0,3888	1,65	0,4505	2,16	0,4846	3,20	0,49931
0,37	0,1443	0,80	0,2881	1,23	0,3907	1,66	0,4515	2,18	0,4854	3,40	0,49966
0,38	0,1480	0,81	0,2910	1,24	0,3925	1,67	0,4525	2,20	0,4861	3,60	0,499841
0,39	0,1517	0,82	0,2939	1,25	0,3944	1,68	0,4535	2,22	0,4868	3,80	0,499928
0,40	0,1554	0,83	0,2967	1,26	0,3962	1,69	0,4545	2,24	0,4875	4,00	0,499968
0,41	0,1591	0,84	0,2995	1,27	0,3980	1,70	0,4554	2,26	0,4881	4,50	0,499997
0,42	0,1628	0,85	0,3023	1,28	0,3997	1,71	0,4564	2,28	0,4887	5,00	0,4999997

Таблиця значень $P(k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$

k \ λ	0,1	0,2	0,3	0,4	0,5	0,6
0	0,904837	818731	740818	670320	606531	548812
1	090484	163746	222245	268128	303265	329287
2	004524	016375	033337	053626	075816	098786
3	000151	001091	003334	007150	012636	019757
4	000004	000055	000250	000715	001580	000964
5		000002	000015	000057	000158	000356
6			000001	000004	000013	000035
7					000001	000003

k \ λ	0,7	0,8	0,9	1,0	2,0	3,0
0	0,496585	449329	406570	367879	135335	049787
1	347610	359463	365913	367879	270671	149361
2	121663	143785	164661	183940	270671	224042
3	028388	03834	049398	061313	180447	224042
4	004968	007669	011115	015328	090224	168031
5	000695	001227	002001	003066	036089	100819
6	000081	000164	000300	000511	012030	050409
7	000008	000019	000039	000073	003437	021604
8		000002	000004	000009	000859	008101
9				000001	000191	002701
10					000038	000810
11					000007	000221
12					000001	000055
13						000013
14						000003
15						000001

k \ λ	4,0	5,0	6,0	7,0	8,0	9,0
0	0,018316	006738	002479	000912	000335	000123
1	073263	033990	004873	006383	002684	001111
2	146525	084224	044618	022341	010735	004998
3	195367	140374	089235	052129	028626	014994
4	195367	175467	133853	091226	057252	033737
5	196293	175467	160623	127717	091604	060727
6	104194	146223	160623	149003	122138	091090
7	059540	104445	137677	149003	139587	117116
8	029770	065278	103258	130377	139587	131756
9	013231	036266	068838	101405	124077	131756
10	005292	018133	041303	070983	099262	118580
11	001925	008242	022539	045171	072190	097020
12	000642	003434	011262	026350	048127	072765
13	000197	001321	005199	014188	029616	050376
14	000056	000472	002228	007094	016924	032384
15	000015	000157	000891	003311	009026	019431

16	000004	000049	000334	001449	004513	010930
17	000001	000014	000118	000596	002124	005786
18		000004	000039	000232	000944	002893
19		000001	000012	000085	000397	001370
20			000004	000030	000159	000617
21			000001	000010	000061	000264
22				000003	000022	000108
23				000001	000008	000042
24					000003	000016
25					000001	000006
26						000002
27						000001

Література

1. Вентцель Е.С., Овчаров Л.А. Теория вероятностей. –М.: Наука, 1973. – 368с.
2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. – 5-е издание, М.: Высшая школа, 2002. – 405с.
3. Гмурман В.Е. Теория вероятностей и математическая статистика. – 8-е издание, М.: Высшая школа, 2002. – 479с.
4. Жлуктенко В.І., Наконечний С.І. Теорія ймовірностей і математична статистика. У 2ч. –ДО.: КНЕУ , 2000. – 304с.
5. Шефтель З.Г. Теорія ймовірностей. К.: Вища школа, 1994. – 193с.
6. Гмурман В.Е. Теория вероятностей и математическая статистика, М., Высшая школа, 2004. – 470с.
7. Данко П.Е. Высшая математика в упражнениях и задачах. В 2 ч.Ч.2: Учеб. Пособие для вузов. – М.: 2005.
8. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – 2-е изд., перераб. и доп. – М.: 2003.
9. Сборник задач по высшей математике для экономистов: Учебное пособие/ Под ред. В.И. Ермакова. – М.: ИНФРА-М, 2004. – 575 с. – (Высшее образование).
10. Турецкий В.Я. Математика и информатика. -М.: ИНФРА-М, 2004.-560с.